BTS: scalable Bayesian Tissue Score for prioritizing GWAS variants and their functional contexts across omics data

Pavel P. Kuksa[1,†], Matei Ionita[1,†], Luke Carter[1], Jeffrey Cifello[1], Kaylyn Clark[1], Otto Valladares[1], Yuk Yee Leung[1,*], Li-San Wang[1,*]

[1] Penn Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

† The authors wish it to be known that, in their opinions, the first two authors should be regarded as Joint First Authors

* To whom correspondence should be addressed

Email: yyee@pennmedicine.upenn.edu
Email: lswang@pennmedicine.upenn.edu

Abstract

Motivation: Summary statistics from genome-wide association studies (GWAS) are widely used in fine-mapping and colocalization analyses to identify causal variants and their enrichment in functional contexts, such as affected cell types and genomic features. With the expansion of functional genomic (FG) datasets, which now include hundreds of thousands of tracks across various cell and tissue types, it is critical to establish scalable algorithms integrating thousands of diverse FG annotations with GWAS results.

Results: We propose BTS (Bayesian Tissue Score), a novel, highly efficient algorithm uniquely designed for 1) identifying affected cell types and functional elements (context-mapping) and 2) fine-mapping potentially causal variants in a context-specific manner using large collections of cell type-specific FG annotation tracks. BTS leverages GWAS summary statistics and annotation-specific Bayesian models to analyze genome-wide annotation tracks, including enhancers, open chromatin, and histone marks. We evaluated BTS on GWAS summary statistics for immune and cardiovascular traits, such as Inflammatory Bowel Disease (IBD), Rheumatoid Arthritis (RA), Systemic Lupus Erythematosus (SLE), and Coronary Artery Disease (CAD). Our results demonstrate that BTS is over *100x* more efficient in estimating functional annotation effects and context-specific variant fine-mapping compared to existing methods. Importantly, this large-scale Bayesian approach prioritizes both known and novel annotations, cell types, genomic regions, and variants and provides valuable biological insights into the functional contexts of these diseases.

Availability and implementation: Docker image is available at https://hub.docker.com/r/wanglab/bts with pre-installed BTS R package (https://bitbucket.org/wanglab-upenn/BTS-R) and BTS GWAS summary statistics analysis pipeline (https://bitbucket.org/wanglab-upenn/bts-pipeline).

Introduction

Typical single-marker-based genome-wide association studies (GWASs) do not consider correlation between variants (linkage disequilibrium, LD) and the potential cellular or epigenetic context of the genetic variants. Interpretation and prioritization of GWAS results and variants thus require downstream analyses using various types of functional genomic (FG) data.

Many existing methods can perform fine-mapping of the GWAS signals within a region, based on the summary statistics alone (approximate Bayes factor), or summary statistics with LD information. Examples of these methods include conditional analysis (Galarneau et al., 2010; Knight et al., 2012; Uffelmann et al., 2021; Yang et al., 2012), CAVIAR/CAVIARBF (Chen et al., 2015; Chen et al., 2016; Hormozdiari et al., 2014), DAP-G (Wen et al., 2016), FINEMAP (Benner et al., 2016), and SuSiE (Wang et al., 2020; Zou et al., 2022). Other methods prioritize variants based on their functional annotations, such as coding, promoter or enhancer regions, e.g., RegulomeDB (Boyle et al., 2012; Dong et al., 2023). Colocalization methods such as coloc (Giambartolomei et al., 2014; Wallace, 2021), HyPrColoc (Foley et al., 2021), eCaviar (Hormozdiari et al., 2016), ENLOC (Wen et al., 2017) and their variants jointly analyze summary statistics from GWAS and molecular traits such as expression quantitative trail loci (eQTL). There are also tools for integrative analysis (e.g., INFERNO (Amlie-Wolf et al., 2018), SparkINFERNO (Kuksa et al., 2020), FUMA (Watanabe et al., 2017)) which combine data from multiple sources: for example, prioritizing variants based on both colocalization posterior probabilities and overlap with functional annotations. Additionally, some methods (e.g., PAINTOR (Kichaev et al., 2014), fGWAS (Pickrell, 2014), BFGWAS_QUANT (Chen et al., 2022), PolyFun (Weissbrod et al., 2020), CARMA (Yang et al., 2023)) further formalize overlap with functional annotations and incorporate functional annotations as priors for inferring causal variant status.

The recent increase in availability of large functional annotation databases such as ENCODE (Consortium, 2012; Consortium et al., 2020), ROADMAP (Consortium et al., 2020; Roadmap Epigenomics et al., 2015), and FILER (Kuksa et al., 2022) provides an opportunity to carry out unbiased functional analyses of GWAS results across thousands of cell types and genome-wide annotations. However, such large-scale, systematic analyses can come at a significant computational cost, as they need to build and evaluate different statistical models for each annotation or combination of annotations and cell types of interest. In addition, each model evaluation often requires running time exponential in the number of potentially causal variants (Asimit et al., 2019; Kichaev et al., 2014).

Moreover, most methods that model LD are susceptible to errors due to a mismatch between GWAS summary statistics and LD. Such mismatches can occur when the GWAS and the reference genotype panel used to compute LD differ demographically. Even more importantly, mismatch is all but guaranteed in the case of GWAS meta-analyses, as reviewed recently in the SLALOM publication (Kanai et al., 2022). Briefly, if two variants belong to the same haplotype, most models expect them to have similar summary statistics and behave erratically if this expectation is violated. But different studies in a meta-analysis can incorporate one variant and

not the other, which can cause large differences in the amount of evidence supporting each variant. To accommodate this common situation, fine-mapping methods must be robust to LD mismatch (Chen et al., 2021; Kanai et al., 2022; Yang et al., 2023).

To address these large-scale analysis issues, we present BTS, a Bayesian Tissue Score model and an efficient implementation of this model. BTS performs joint fine-mapping of variants and their context-mapping based on FG annotations and provides easy-to-interpret summaries of the results. The main features of the proposed BTS framework are:

- BTS can perform joint context-mapping (inference of cell types, genomic features) and context-specific fine-mapping of variants (**Figs. 1**,**3**,**5**; **Methods**)
- End-to-end GWAS summary statistics analysis pipeline (**Fig. 1**; **Section "**BTS GWAS summary statistics analysis workflow**"; Supplementary Methods**): users can provide their own functional annotations for running with BTS, or use the FILER FG database (Kuksa et al., 2022) to obtain cell type-specific annotations from data sources such as ENCODE (Consortium, 2012; Consortium et al., 2020), EPIMAP (Boix et al., 2021), and GTEx (Consortium, 2020). By default, the user only needs to provide GWAS summary statistics as input.
- Scalability: BTS can conduct a systematic and exhaustive search through thousands of genome-wide annotation tracks and has running times two orders-of-magnitude faster per genome-wide track (**Section** "Running time improvement"; **Fig. 6**) using a novel, more efficient factored Bayesian model (**Methods**; **Section** "BTS statistical model") for FG annotations, LD and GWAS summary statistics.
- Robustness to mismatch between GWAS summary statistics and LD estimates (**Fig. 2**). The BTS model introduces a single model parameter (the prior on variance of the true effect sizes) to control most of the sensitivity to LD mismatch. We also provide guidelines on how to choose this parameter. (**Methods; Section** "Mitigation of LD mismatch").

We applied BTS to GWAS datasets from four different diseases: Coronary Artery Disease (CAD) (van der Harst & Verweij, 2018), Inflammatory Bowel Disease (IBD) (Liu et al., 2015), Rheumatoid Arthritis (RA) (Stahl et al., 2010) and Systemic Lupus Erythematosus (SLE) (Bentham et al., 2015). In each case, BTS took under one hour (**Fig. 6**) to prioritize cell types, tissues, genomic regions and variants relevant to the disease across a variety of FG annotation tracks (>900; see **Supplementary Table S2** for details on annotation tracks used). This shows BTS can serve as a tool for systematic and unbiased functional context mapping and context-specific variant mapping.

The BTS GWAS summary statistics analysis pipeline, including scripts for obtaining relevant LD and functional annotations starting from the GWAS summary statistics, is freely available at https://bitbucket.org/wanglab-upenn/BTS-pipeline. BTS model estimation is implemented as an R package and is freely available at https://bitbucket.org/wanglab-upenn/BTS-R.
BTS Docker including pre-installed GWAS summary statistics pipeline is also available at https://hub.docker.com/r/wanglab/bts.

Results

BTS overview

To jointly estimate variant posterior probabilities and enrichment of annotations in causal variants, we rely on a well-known Bayesian model (PAINTOR (Kichaev et al., 2014), CAVIARBF (Chen et al., 2015), see also **Methods; Section "**BTS statistical model**"; Supplementary Methods**). Briefly, for each possible configuration of causal variants within a locus, its Bayes factor only depends on the GWAS summary statistics and the correlation between variants (LD). On the other hand, causal variant configuration prior depends only on the functional annotations of the variants, as well as the annotation importance (e.g., enrichments of the annotations in causal variants). We use an Expectation-Maximization (EM) algorithm to maximize overall data likelihood by iteratively updating the annotation enrichments and evaluating configuration posterior probabilities, until a convergence criterion is reached (**Fig. 1; Methods; Supplementary Methods**).

Our key observation is that the posterior probability decomposes into a factor which only involves the GWAS data, and one which only involves annotations (**Section "**BTS statistical model**"**; **Eq. 9**). Furthermore, the factor involving GWAS data is the same for all iterations of the EM algorithm. Because of this, BTS can compute Bayes factors only once and then re-use them hundreds or thousands of times across various annotation evaluations (**Figs. 1, 6**).

BTS algorithm further improves running time by introducing two main computational improvements 1) by using the matrix inversion lemma to compute Bayes factors (**Section "**BTS statistical model**"; Supplementary Methods; Lemma 2**, **Eq. 4**) and 2) by more efficiently computing variant configuration probabilities (**Section "**BTS statistical model**"; Eq. 8**).

These observations, alongside computational improvements detailed in the **Methods, and Supplementary Methods**, are the main reason for the improved runtime of BTS (**Fig. 6**).

To run the BTS algorithm, the user has the option to supply three types of information for each genomic region to be analyzed: per-locus GWAS summary statistics, LD matrix, and functional annotations. Alternatively, the user can provide full GWAS summary statistics only and use BTS end-to-end GWAS summary statistics analysis pipeline (**Fig. 1**) which (1) identifies genomic regions of interest, (2) computes per-locus variant LD and functional annotation matrices, then (3) runs BTS model estimation on the assembled input data (Steps 3,4 in **Supplementary Fig. S1**), and (4) aggregates across estimated models to prioritize annotations, genomic regions, and variants. This GWAS summary statistics analysis pipeline leverages the FILER FG data repository (Kuksa et al., 2022) for tissue and cell type-level functional annotations, and the 1000 Genomes project genotype reference panel (Byrska-Bishop et al., 2022; Genomes Project et al., 2015) to estimate LD between variants. The pipeline outputs cell type and functional context-specific posterior probabilities for individual variants within each of the identified genomic regions of interest, annotation and functional context importance (enrichment scores, causal

variant prior odds), as well as cell type, region, and variant mapping summary plots (e.g., **Figs. 3, 4**)

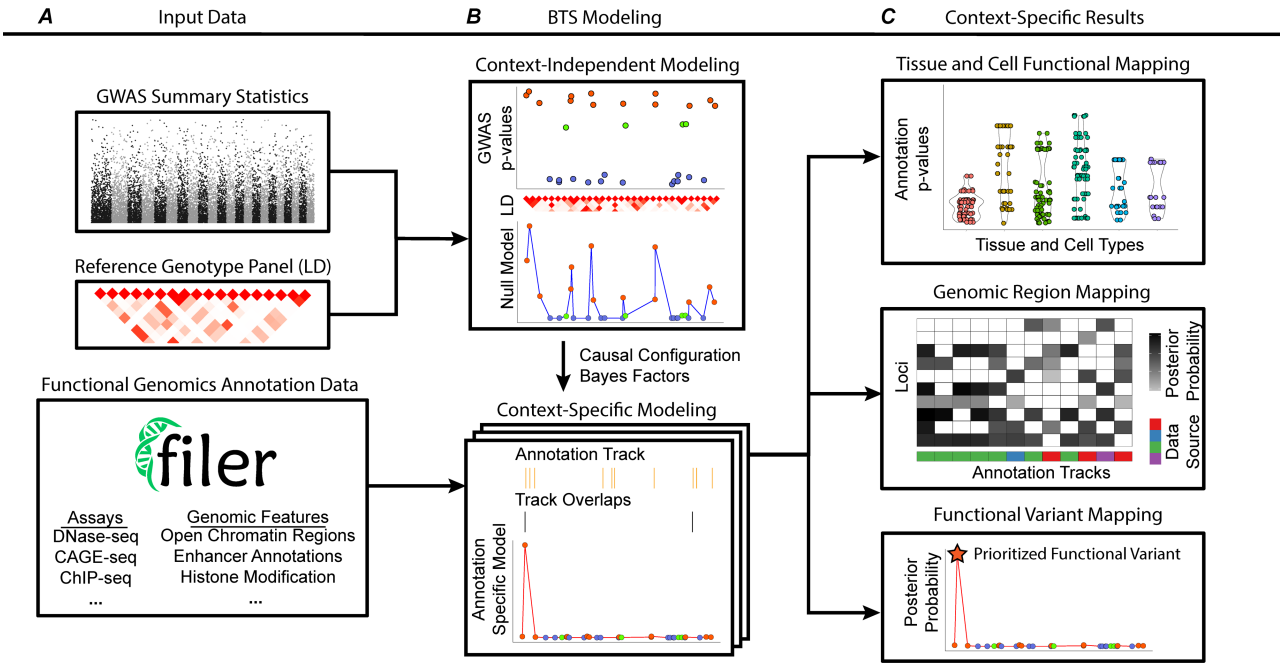

Figure 1. Overview of the BTS framework. **A)** Input GWAS summary statistics and LD data (top sub-panel) are analyzed across cell type and tissue-level FG annotations (bottom sub-panel) to jointly perform context-mapping and identify context-specific set of genomic regions and potentially causal variants for each context and region. **B)** The two-factor structure of BTS model (**Section** "BTS statistical model") allows it to be evaluated efficiently across thousands of genome-wide annotations (**Fig. 6; Section** "Running time improvement"). BTS estimates context-independent GWAS+LD null model (top sub-panel) and many annotation-specific models (bottom sub-panel). **C)** BTS outputs annotation relevance (enrichment *p*-values) (top sub-panel), functional genomic regions within each of the prioritized contexts (middle sub-panel), and annotation-specific causal variant posterior probabilities (bottom sub-panel) (**Fig. 3**; **Sections "**Prioritizing regions, variants and their contexts with BTS**", "**Cross-trait BTS evaluation**"**). **Supplementary Fig. S1** and **Section** "BTS GWAS summary statistics analysis workflow" detail BTS analysis workflow and its main steps.

As illustrated in the example in **Fig. 2**, given the same summary statistics, BTS discovers two independent signals if the variants are uncorrelated (**Fig. 2**, first column, LD=0), and one independent signal shared between two variants, if these are correlated (**Fig. 2**, second column, LD=1). Crucially, in the baseline model with no functional annotations, the posterior probability mass is shared equally between the two variants, but in the model that uses an annotation the probability is assigned to the variant which overlaps trait-relevant annotation.

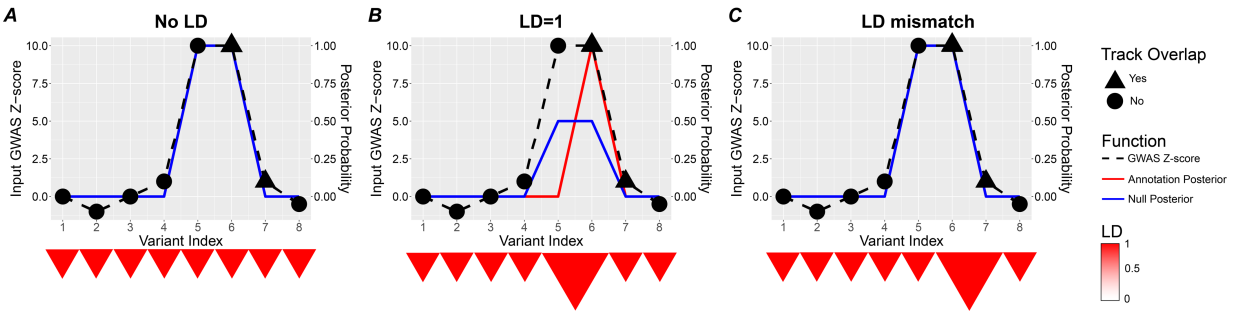

Figure 2. Variant prioritization with BTS using summary statistics, LD, and annotation information. **A**) In case of two independent signals (numbered 5 and 6 on the x-axis), both variants are prioritized by BTS (posterior probability for both variants is 1). **B**) If the two variants (5 and 6) are in LD, variant 6 is prioritized (red line posterior probability) based on its overlap with relevant annotation track (black triangle). Note that the null model (blue line) will not be able to prioritize between these variants (posterior=0.5) since both variants have similar Z-scores and are in LD. **C**) In case of LD and GWAS mismatch, two variants (6 and 7) are in LD but have very different GWAS summary statistics (Z-scores). As shown here, BTS is robust with respect to the LD mismatch and prioritizes variant 6 with a more significant association statistic (Z-score) and avoids prioritizing false-positive variant 7 by default model.

Finally, BTS is robust to mismatch between LD and summary statistics (**Fig. 2**, third column, LD mismatch): when two variants with LD close to 1 have wildly different summary statistics, the one with low phenotype association is not prioritized. We emphasize that this common-sense result was difficult to achieve: the same statistical model with different parameter settings has a propensity for prioritizing variants with low phenotype association in such cases of LD mismatch. One of our contributions in this article is to identify parameter values that make the model robust to LD mismatch (**Supplementary Methods; Methods** **"**Mitigation of LD mismatch**"**).

Prioritizing regions, variants and their contexts with BTS

To illustrate the performance of BTS, we started with summary statistics from a Coronary Artery Disease (CAD) GWAS (van der Harst & Verweij, 2018). Our end-to-end pipeline first performed LD pruning of the 4,298 genome-wide significant GWAS variants (*p*-value<*5e-8)* to identify 389 tag variants (pairwise-independent, r^2<0.7). We further filtered out 2 loci belonging to the HLA region on chromosome 6, where complicated linkage patterns make fine-mapping extremely difficult. LD expansion of the remaining tag variants yielded 6,513 candidate variants. Using the LD-expanded candidate set, we then defined the LD block for each tag variant using the left-most and right-most variants in LD with the tag variant 1) within a 1 megabase window and 2) with no more than 1000 variants separating the boundary variants and the tag variant. Any of the overlapping LD blocks were then merged to obtain 167 non-overlapping regions of interest. By extracting variants within these regions from the initial GWAS summary statistics, we obtained a total of 24,727 variants for further analysis. We note that each of the regions will not only contain variants that are highly associated with the CAD phenotype or are in LD with them

but also any unrelated or unassociated variants in between. For each of the identified regions of interest, the BTS pipeline used PLINK (Chang et al., 2015; Purcell et al., 2007) to compute LD matrices based on the 1000 Genomes dataset (Byrska-Bishop et al., 2022; Genomes Project et al., 2015), using genotypes from samples belonging to the European super-population as reference.

BTS used FILER (Kuksa et al., 2022) to obtain a discovery set of 943 various regulatory annotations for testing including FANTOM5 enhancers (Andersson et al., 2014), Roadmap Epigenomics ChromHMM enhancers (Roadmap Epigenomics et al., 2015), ENCODE (Consortium, 2012; Consortium et al., 2020) DNase hypersensitivity sites, EpiMap enhancers (Boix et al., 2021) and all available tissues and cell types therein (see **Supplementary Tables S2, S4**). After excluding annotations which did not overlap any of the regions of interest, 532 genome-wide annotation tracks remained, across all data sources and tissues. Then BTS was run on the 167 genomic regions, with the default setting of d=2, i.e. looking for at most two distinct causal signals in each region (we note that even though d=2, the credible variant set may include more than d variants).

BTS prioritized open chromatin annotations for blood vessel, monocyte and B cells and other relevant cell types and tissues (**Fig. 3**) consistent with the tissues and cell types involved in the disease (Ghattas et al., 2013; Libby & Theroux, 2005; Srikakulapu & McNamara, 2017). Blood vessel, blood, and immune-related annotations accounted for 23 out of 103 (22%) of the significant annotations (annotation statistical significance assessed with a likelihood ratio test, $p<0.01$).

Importantly, BTS provides context-specific variant fine-mapping as shown in **Fig. 4**, where BTS prioritized two variants (rs9295128 and rs9457927) within the same SLC22A2 region, with one variant (rs9295128) prioritized in coronary artery context, and another variant (rs9457927) prioritized in the cardiac muscle cell.

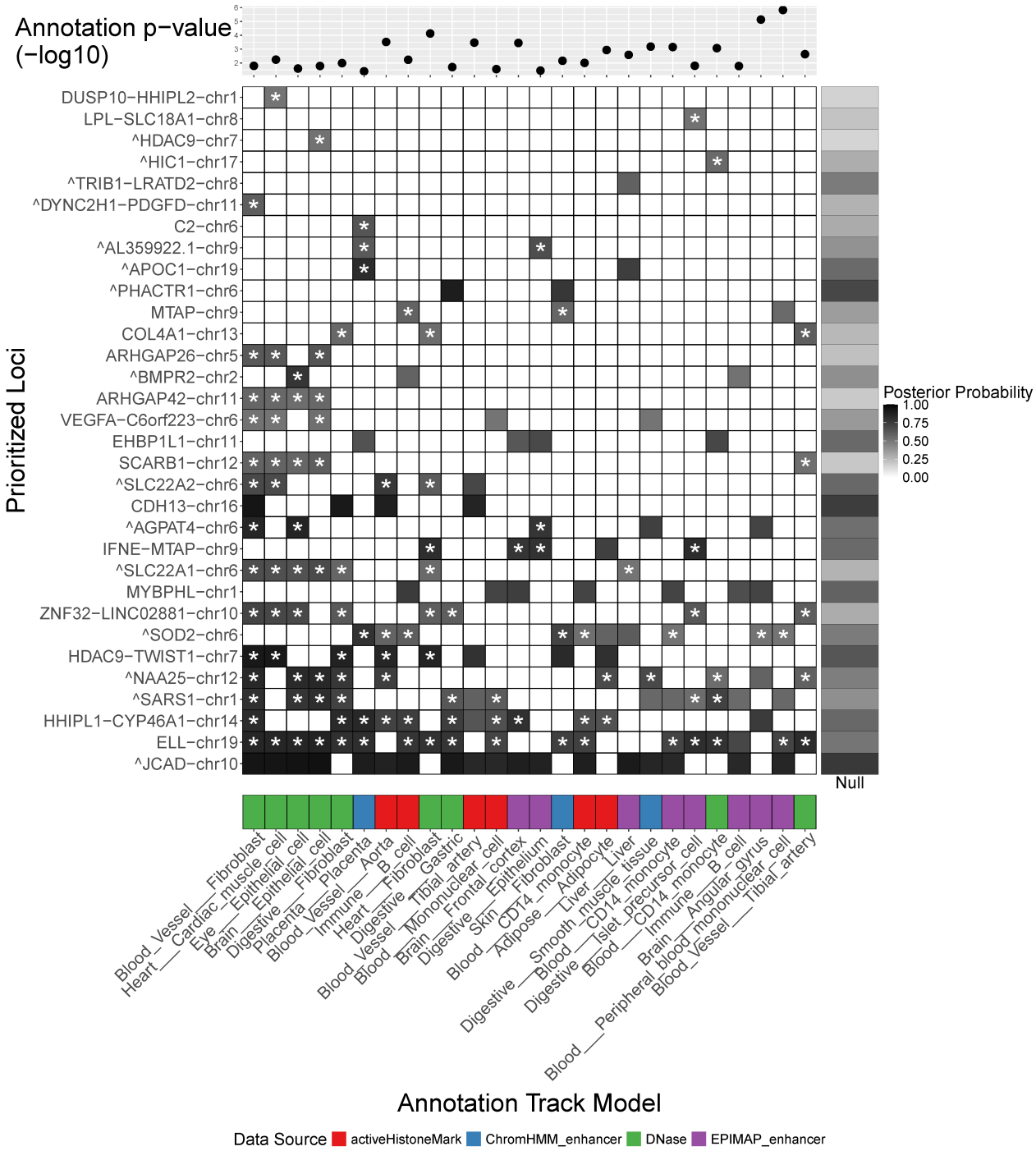

Figure 3. BTS prioritization for functional contexts (X axis) and genomic regions (Y axis) in CAD GWAS. **A**) Effect size and enrichment of the functional annotation genome-wide across all identified regions of interest and using all variants located within these regions. Shown are the annotation significance (*p*-values) as given by the likelihood ratio test of the BTS model with annotation and the model without annotation (see **Section** "BTS algorithm"; **Supplementary Methods** for details on estimating annotation significance*)*. Out of 103 significant annotations ($p < 0.01$), shown are 25 annotations overlapping with at least five prioritized genomic regions

(variant posterior > 0.5). In general, annotations with greater enrichment in top GWAS variants have lower, more significant *p*-values. **B**) BTS-prioritized regions of interest (Y axis) and their potential functional contexts (X axis). Shown are regions with annotation overlaps and top variant posterior > 0.5. Asterisks (*) mark regions with increased causal variant posterior compared to the null model (GWAS+LD only, without annotation) in one or more functional contexts. Darker colors correspond to regions and annotations with greater top variant posteriors. Merged regions (i.e. regions containing more than one overlapping LD blocks) are labeled with a caret (^) before the region name. For CAD, BTS prioritizes blood vessel, monocyte, cardiac muscle, and immune cell types across active histone marks, open chromatin, and enhancer genomic features.

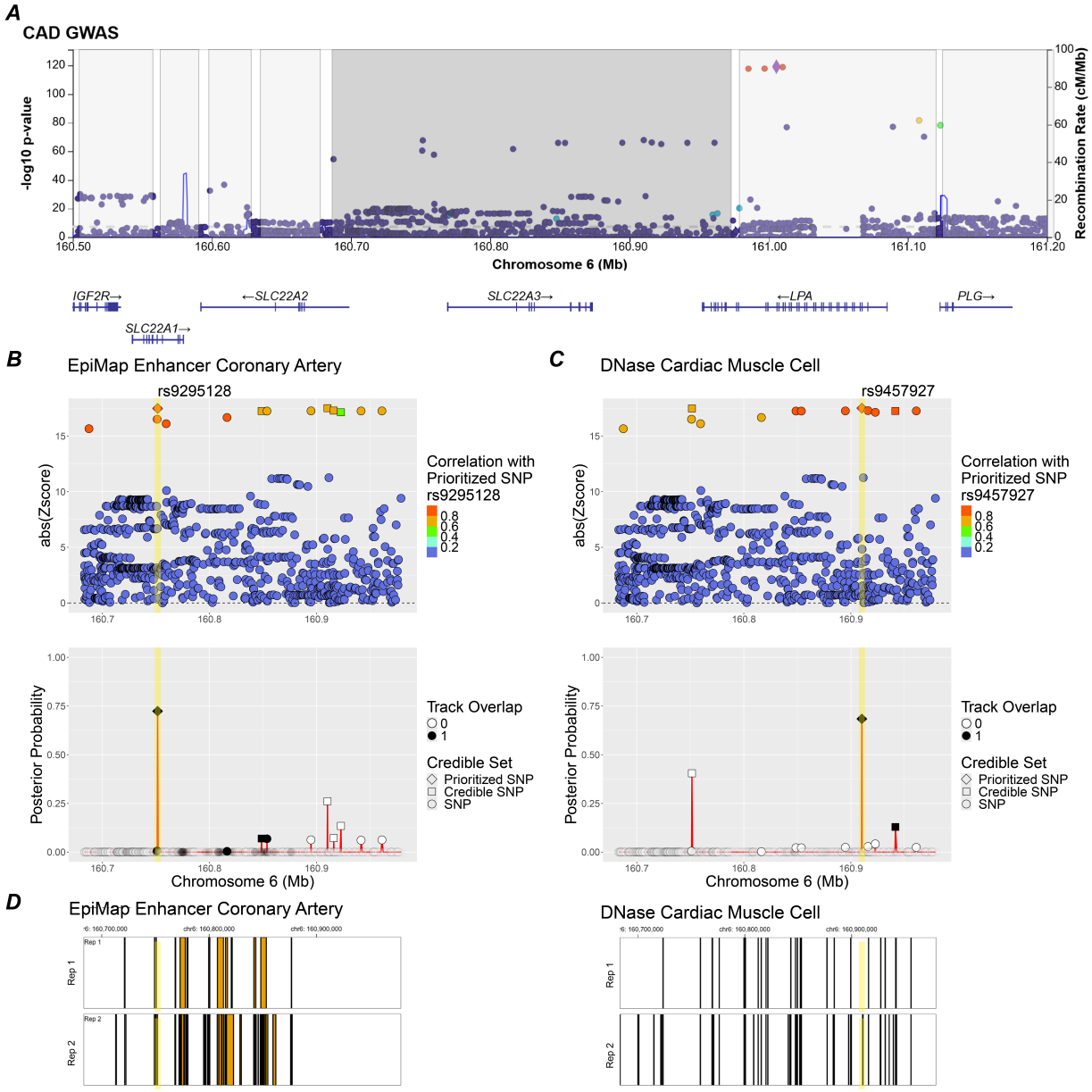

Figure 4. BTS within-region context-specific variant fine-mapping. Shown is an example of differential variant prioritization for SLC22A2 region (GRCh37/hg19 chr6:160683145-160978997) in two different contexts, coronary cell enhancers and open chromatin regions in

cardiac muscle cells. **A)** GWAS results for genomic region encompassing SLC22A, with discrete LD blocks shaded separately. Subfigures **B**, **C**, **D** show BTS analysis results for the middle (dark grey) LD block. **B)** BTS prioritizes variant rs9295128 in coronary artery context as it overlaps with annotated enhancer from EpiMAP (BTS coronary artery enhancer annotation *p*-value=1.7e-3; prior odds=2.85). **C)** In a different context (heart), BTS prioritizes variant rs9457927 in the cardiac muscle cell based on its location within an annotated open chromatin region defined by DNase assay (BTS annotation *p*-value=7.1e-4; prior odds=5.51). **D)** Annotated enhancer regions for coronary artery (left) and open chromatin regions for cardiac muscle cell (right) across two biological replicates (Rep1, Rep2) each.

Using cell type and tissue-specific annotations with BTS helped identify a more precise set of potentially causal variants across genomic regions for CAD. Compared to the GWAS+LD null model without annotation, we observed a significant reduction in average credible set sizes for BTS models that use annotations. For example, across 46 genomic regions prioritized by BTS and 200 annotations we observed the average credible set size of 2.44±2.7, while the null model had a significantly larger credible set size of 6.5±7.77 across the same genomic regions. Importantly, using annotations allowed BTS to identify 44 additional potentially causal variants that were otherwise not prioritized by the null model. We also note 38 out of 72 variants identified by GWAS+LD-only model were de-prioritized as they were not found to be located within any of the prioritized annotations or functional elements. Overall, using annotations with BTS allowed to identify a total of 78 potentially causal CAD variants across 46 loci (see **Supplementary Table 5a**).

Cross-trait BTS evaluation

We further tested performance of BTS on several other GWAS datasets for immune-related traits (IBD, RA, SLE) using the same set of 943 annotation tracks as input. As shown in **Fig. 5**, BTS prioritizes disease and cell type associations consistent with previous studies, including myeloid dendritic cells for IBD (Baumgart et al., 2005), T cells for RA (Weyand et al., 2000), and monocytes for SLE (Hirose et al., 2019) (ChromHMM enhancers) and CAD (Ghattas et al., 2013) (active histone marks, enhancer and DNase-hypersensitive regions). See **Supplementary Tables 5a-d** for detailed list of all prioritized annotations, genomic regions, and variants for each tested GWAS dataset.

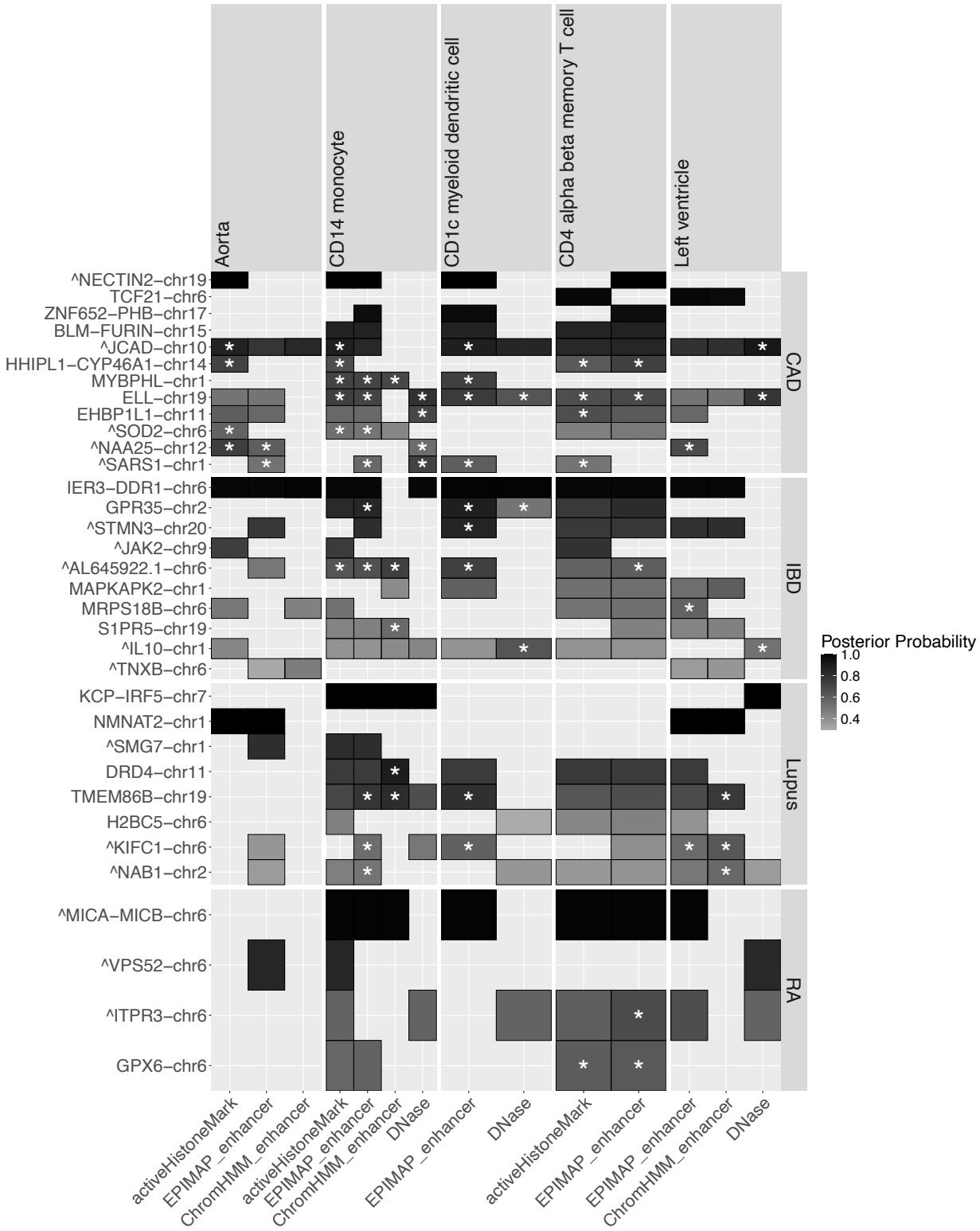

Figure 5. BTS results across four immune-related and cardiovascular GWAS traits. Shown are identified functional contexts (X axis) and genomic regions (Y axis) at the cell-type level for each of the traits (CAD, IBD, SLE, and RA panels on the right). For each region and functional context, the top variant posterior is shown (shades of gray) with a star (*) indicating posterior increase of at least 0.2 in that context compared to the null model without annotation. BTS

identifies specific cell types and genomic feature types (open chromatin, histone marks, enhancers on the X axis) for each of the traits (e.g., CD4 T cell for RA; myeloid dendritic cells for IBD; aorta blood vessel and CD4 alpha-beta T cells for CAD; heart left ventricle for SLE).

Moreover, BTS prioritized regions which have been implicated in IBD (**Supplementary Fig. S2**). Among these are an intergenic region near the gene coding for the transcription factor CEBPA (Zhou et al., 2019), or genes coding for interleukins and their receptors, such as IL10 (Ip et al., 2017) and ILR1 (Dosh et al., 2019). Importantly, while recent functional work on IBD (Stankey et al., 2024) identified intergenic region directing macrophage inflammation through ETS2, the ETS2-PSMG1 region was prioritized by BTS in granulocytes (**Supplementary Fig. S2**), belonging to the same myeloid family with macrophage.

Running time improvement

BTS only took minutes to process genome-wide GWAS data, which involved tens of thousands of variants and hundreds of annotation tracks (**Fig. 6**). Compared to reference implementation fastPaintor/Paintor v3.0 (Kichaev et al., 2017), BTS is faster by two orders of magnitude (**Fig. 6c; Section** “BTS statistical model”), being able to compute an annotation-specific model in under one second on average (**Fig. 6a,b**). This comes at the cost of a pre-processing step that takes a few seconds but is only carried out once for the entire analysis. Moreover, BTS scales linearly (**Fig. 6a,b**) with both the number of genomic regions and the number of annotation tracks, which guarantees reasonable running times for even larger datasets.

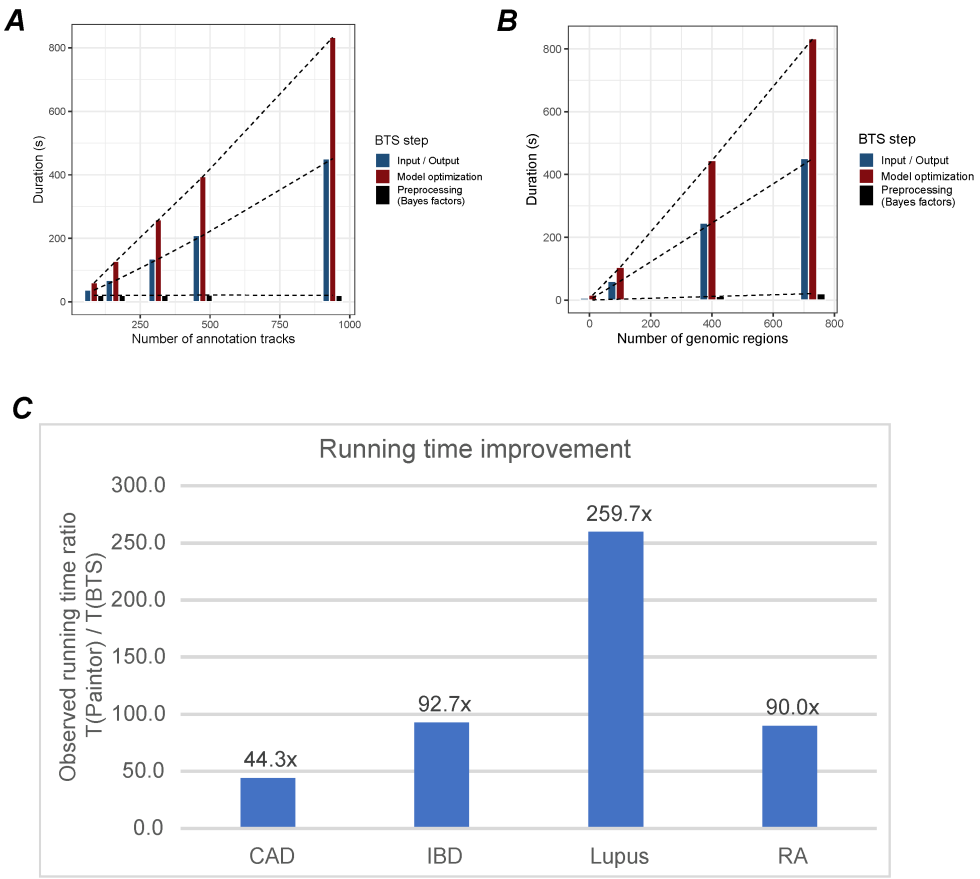

Figure 6. BTS running time. **A.** BTS running time as a function of the number of annotation tracks used for functional evaluation. BTS scales linearly with the number of input annotation tracks. Note the constant running time for pre-computing annotation-independent Bayes factors. **B.** BTS running time as a function of the number of genomic regions. BTS scales linearly with the number of genomic regions of interest. **C.** Comparison of the per-model BTS running time and the running time using the reference implementation (fastPaintor 3.0 (Kichaev et al., 2017)). BTS improves running time by a factor of 44-259x (average 120x) across the four GWAS datasets analyzed.

Methods

BTS GWAS summary statistics analysis workflow

We provide an end-to-end functional variant fine-mapping and context-mapping pipeline for analysis of GWAS summary statistics based on user-supplied full GWAS summary statistics as input (**Fig. 1; Supplementary Fig. S1; Supplementary Methods**). The BTS GWAS summary statistics pipeline aims to automate genome-wide post-GWAS functional analysis and provide a systematic and more complete report of all potentially causal variants, genomic loci, and likely

functional contexts including cell type and tissue-specific regulatory mechanisms underlying observed GWAS signals.

To do this, input GWAS summary statistics are first preprocessed to resolve reference and non-reference (alternative) alleles and normalize effect sizes to consistently reflect effects of the alternative alleles. Using the normalized GWAS summary statistics, a set of genomic regions for further downstream analyses and all potential candidate variants are then identified. This is achieved through identification of pairwise-independent GWAS signals (*tag* variants) by performing linkage disequilibrium-based pruning (LD r^2>0.7) of all genome-wide significant (p < 5e-8) GWAS variants. Based on the identified tag variants, the candidate set of potentially causal variants is then formed as a set of variants in LD with the tagging variants (LD r^2>0.7), including the tag variants themselves. LD-based genomic regions for analysis are constructed by defining the LD region for each of the tag variants as a genomic region with the left and right boundaries corresponding to the leftmost and rightmost variants linked with the tag variant. The leftmost and rightmost variants are restricted to be within 1Mbp from the tag variant and have no more than 1000 variants between them and the tag variant. The final set of non-overlapping genomic regions for downstream analyses is obtained by merging any overlapping LD-based regions into larger regions.

For each such identified genomic region, the pipeline will then generate all the information required to fit the BTS model (**Section** "BTS statistical model") including the pairwise LD-based variant correlation matrix L (n x n, where n is the number of variants in the region), functional annotation matrix A (n x N_A), and a vector of GWAS summary statistics Z (Z-scores) (n x 1).

Pairwise LD calculation for all variants located in the locus is conducted based on the reference genotype panel (Byrska-Bishop et al., 2022; Genomes Project et al., 2015). Functional annotation matrix **A** is obtained by querying FILER FG database (Kuksa et al., 2022) for each of the N_A genomic annotations and FG data tracks of interest and noting annotation overlaps for each of the variants ($\mathbf{A}_{i,j}$ will be set to 1 if variant i overlaps annotation j). The summary of included annotations and detailed list of annotation tracks used are provided in **Supplementary Tables S2**,**S4**. Summary statistics (Z-scores) are extracted from the input GWAS summary statistics after reference and alternative allele resolution and effect normalization to consistently reflect the effect of the non-reference (alternative) allele.

BTS algorithm (**Fig. 1**; **Methods; Section "**BTS algorithm**"**) will then be applied to fit the model and estimate variant and locus posteriors and functional annotation enrichment for each of the target annotations (**Fig. 3**). To find potentially causal variants within each locus, BTS uses variant LD matrix and Z-scores from GWAS summary statistics to pre-compute and store Bayes factors for each possible causal variant configuration in every locus. BTS then uses an EM-based algorithm to iteratively estimate annotation enrichment coefficients and compute annotation-specific causal priors for each of the analyzed variants. These functional annotation-specific priors are then combined with pre-computed configuration Bayes factors to obtain context-specific causal variant posteriors.

Evaluation results reported in **Sections** “Prioritizing regions, variants and their contexts with BTS”, “Cross-trait BTS evaluation” were generated by applying this pipeline to CAD (van der Harst & Verweij, 2018), IBD (Liu et al., 2015), RA (Bentham et al., 2015), and SLE (Stahl et al., 2010) GWAS summary statistics.

BTS statistical model

To accommodate for correlation between variants (linkage disequilibrium, LD) as well as the tissue and cell-type-specific functions of variants and genomic regions, we adopt the Bayesian probabilistic framework first proposed in (Kichaev et al., 2014). We consider the following information for all *n* variants in a genomic region of interest: 1) a vector **Z** of GWAS Z-scores (standardized regression coefficients; observed), 2) a vector ***A*** of variant tissue and cell type-specific annotations (observed), 3) LD (linkage disequilibrium) correlation matrix **Σ,** and a vector **C** of variant causal status (unobserved; latent binary variable: set to 1 for causal variants, and 0 otherwise), a vector **Λ** of unknown true effect sizes for each variant. To model each genomic region, we use a Bayesian model in which the likelihood of observing **Z** is a multivariate normal parametrized by causal configuration **C**, true effect sizes **Λ**, and LD correlation matrix **Σ**:

$$P(Z|C,\Lambda,\Sigma) = N(Z;\Sigma(\Lambda \circ C),\Sigma). \tag{1}$$

where $\Lambda \circ C$ is an element-wise vector product.
The true effect size **Λ** is also modeled as normal, with mean 0 and diagonal variance, using the scalar W as a model parameter:

$$P(\Lambda|C) = N(\Lambda;0,WI_C) \tag{2}$$

where WI_C is a scaled diagonal matrix, with diagonal elements of I_C set to 0 and 1 according to the causal configuration C.

Integrating out **Λ** gives the following formula for the full likelihood as proved in (Kichaev et al., 2014):

$$P(Z|C,\Sigma) = N(Z;0,\Sigma + W\Sigma I_C\Sigma). \tag{3}$$

Bayes factor (BF) for a causal variant configuration C in any particular genomic region:

$$\begin{aligned} BF_C &= \frac{P(Z|C)}{P(Z|C=0)} \\ &= \frac{N(Z;0,\Sigma + W\Sigma \mathrm{I}_C\Sigma)}{N(Z;0,\Sigma)} \\ [\boldsymbol{Lemma\ 1}] &= \frac{N(Z_1;0,\Sigma_{11} + W\Sigma_{11}^2)}{N(Z_1;0,\Sigma_{11})} \\ [\boldsymbol{Lemma\ 2}] &= det(I + W\Sigma_{11})^{-1/2} exp\left(\frac{W}{2} Z_1^T (I + W\Sigma_{11})^{-1} Z_1\right), \end{aligned} \tag{4}$$

where the first simplification (**Lemma 1)** reduces BF computation from full vectors and matrices to the much smaller $Z_1 = Z_{i:C_i=1}$, and $\Sigma_{11} = \Sigma_{i:C_i=1}$ corresponding to the Z-scores and correlation between the causal variants ($C_i = 1$), and the second simplification (**Lemma 2**) further reduces

BF computation to a single matrix inversion of a positive semi-definite matrix (see **Supplementary Methods**).

The prior probability of causality for each variant *i* is modeled as a logistic function of its annotation *A*:

$$Prior(C_i) = P(C_i = 1|A, E) = 1/\big(1 + exp(-EA)\big) \tag{5}$$

where the annotation effect size coefficient *E* for annotation *A* is estimated genome-wide and is shared by all variants in all regions. Note that for variants overlapping annotation A with positive effect E, their prior probability of causality will be greater than prior probabilities for variants located outside of annotation A. Prior for causal configuration C_j for region *j* is then a product of all *n* individual variant priors $P\big(C_{i,j}\big)$

$$Prior\big(C_j\big) = \prod_i P\big(C_{ij}\big|A, E\big) = \prod_i P\big(C_{ij} = 1\big|A, E\big)^{C_{ij}} P\big(C_{ij} = 0\big|A, E\big)^{1-C_{ij}} \tag{6}$$

The full data likelihood across all genomic regions *j* is a product of individual region data likelihoods

$$L(Z; E, A) = \prod_j \sum_{C_j} P\big(Z_j\big|C_j\big)\, P\big(C_j\big|E, A\big), \tag{7}$$

The computational complexity for each region *j* is then consists of prior computation and data likelihood computation (Eq. 3) for every possible causal variant configuration *C*, $O(|C| \times (L + d^3))$.

To improve computational efficiency, we first note that $P(C_i = 0|A, E) = 1 - P(C_i = 1|A, E) = 1/(1 + exp(EA))$ and a full variant configuration probability can be computed in *O(d)* time (where *d* is the maximum number of independent causal variants) as an update to the null configuration probability:

$$P(C|E, A) = P(C_0|E) \prod_{i:C_i=1} P(C_i = 1|E, A)/P(C_i = 0|E), \tag{8}$$

where C_0 is a null configuration (all variants are non-causal) and the $P(C_0)$ term is only computed once per locus.

Marginalizing over C, and restricting to configurations that have at most *d* causal variants, we obtain the posterior probability that a variant *i* is causal in a particular genomic region:

$$\begin{aligned} P(C_i = 1|Z, \Sigma, A, E) &= \sum_{C:C_i=1} P(C|Z, \Sigma, A, E) \\ &= \frac{\sum_{C:C_i=1} P(Z|C, \Sigma) P(C|E, A)}{\sum_{all\ C} P(Z|C, \Sigma) P(C|E, A)} \\ &= \frac{\sum_{C:C_i=1} BF_C\, P(C|E, A)}{\sum_{all\ C} BF_C\, P(C|E, A)} \end{aligned} \tag{9}$$

expressed in terms of the Bayes factors $BF_C = P(Z|C)/P(Z|C_0)$ (Eq. 4) and configuration priors P(C) (Eq. 6).

More conceptually, the variant posterior in (Eq. 9) is a dot product between a vector of annotation-independent Bayes factors BF_c and a vector of annotation-dependent variant priors $P(C)$, where each vector is indexed by causal variant configurations C with up to *d* causal variants (i.e. each vector is of |C|=$\sum_{l=0}^{l=d}\binom{n}{l}$ dimensionality, where *n* is the number of variants in the locus). These vectors can be computed independently from each other, as Bayes factors (BF) only depend on GWAS summary statistics and LD matrix (Eq. 4), while the variant priors only depend on annotations and their enrichment coefficients (Eq. 5,6).

Given access to the precomputed Bayes factors for each possible configuration C, the overall complexity of computing variant posteriors (using Eq. 8,9) is then linear $O(|C| \times d)$ for any given genome-wide annotation *A*, which is a $O(L + d^2)$ improvement compared to non-factorized model (Eq. 7) $O(|C| \times (L + d^3))$ with the on-the-fly prior and Bayes factor computation.

The model outputs the variant causal posterior probabilities (Eq. 9) for each of analyzed variants and the estimated annotation effect size coefficients E_A for each tested annotation *A*.

BTS algorithm

We use an expectation-maximization (EM) algorithm (Kichaev et al., 2014) to fit the statistical model in **Eq. 9**. Intuitively, this is an iterative algorithm which optimizes overall likelihood and takes turns updating the posterior probabilities and enrichment coefficients, until a convergence criterion is reached.

Given multiple annotations, with a possibly complex correlation structure, it is standard practice (Kichaev et al., 2014; Pickrell, 2014) to perform feature selection, by first fitting a separate model for each annotation, and then selecting a few high-ranking annotations for a final model. Therefore, our systematic approach to annotations requires that thousands of models be fitted for all annotation, tissue, and cell types.

Our key observation is that, in (**Eq. 9**), the posterior probability decomposes into a factor which only involves the GWAS data, and one which only involves annotations. Furthermore, the factor involving GWAS data is the same for all iterations of the EM algorithm. Because of this, BTS computes it only once, and distributes it to all models and EM iterations as necessary. Since likelihood computation is the most time-intensive part of fitting the model, the choice to only perform it once is responsible for the bulk of our computational improvements.

This algorithm design choice is a trade-off between speed and memory: to compute the likelihood only once, BTS must store it until all models have been processed. The amount of

time and memory spent on likelihood computations is proportional to the number of allowed causal configurations:

- If there are N variants, and any subset of them could be the causal set, then there are 2^N causal sets to be enumerated. Due to the nature of exponential growth, this is impractical even for moderately large regions (N>30) and impossible for large ones (N>80).
- BTS implements a common solution to this problem, which is to only consider configurations of size smaller or equal to *d* (Asimit et al., 2019; Kichaev et al., 2014), a user-provided parameter, with default value 2. Then storing the likelihood of such configurations requires $O(N^d)$ memory for a region with N variants.
- If *d*=2, the necessary memory is equal to that of storing the LD matrix, so BTS gains two orders of magnitude in speed, at the cost of only doubling its memory use.
- For *d*>2, we mitigate the memory use by only storing those likelihoods which are at most t orders of magnitude smaller than the largest, where t is a user-provided parameter, with default value 12. In our experiments with *d*=3,4,5, this optimization reduces memory use by 100-fold, and does not change the final results within the first 5 significant digits. Our experiments suggest that BTS can accommodate values of *d* up to 5 without significant memory issues, while for *d*>5 runtime increases severely.

We obtain further computational improvements by using the matrix inversion lemma (**Lemma 2**) to compute Bayes factors (**Supplementary Methods**) and more efficiently computing variant configuration probabilities (**Eq. 8**). When computing likelihood ratios, BTS needs to evaluate the ratio of multivariate normal densities with different variance matrices. Naively, this involves inverting each variance matrix. The matrix inversion lemma provides an equivalent expression in which a single matrix needs to be inverted and has the following benefits:

- Decreased computation time, since matrix inversion is the most time-consuming part of likelihood computation.
- The covariance matrices are singular in the case of variants in perfect LD. In our formulation, the matrix to be inverted is strictly positive definite, which improves numerical stability and removes the need for regularization.

Figure 1 summarizes the BTS algorithm:

- A module for computing Bayesian factors and likelihoods.
- A loop which distributes annotations and likelihoods to each model to be fit.
- Aggregation of results, and prioritization of annotations, loci and variants.

BTS GWAS summary statistics pipeline using core BTS algorithm is outlined in **Supplementary Figure S1**.

Mitigation of LD mismatch

We investigated the effects of LD mismatch, which can occur whenever in-sample LD for the GWAS cohort is unavailable, and a reference genotype panel is used instead. There are existing methods to flag regions with high suspicion of LD mismatch (DENTIST (Chen et al.,

2021), SLALOM (Kanai et al., 2022)), but they do not address the problem after identifying it. Moreover, SLALOM only flags a region if the mismatch involves the variant with highest association, which is not completely general.

Our approach is to quantify the spurious effect of LD mismatch on likelihood computations and provide guidance in choosing algorithm parameters so that this effect is minimized. In the supplementary material (see **Supplementary Methods**), we show that, for two variants in perfect LD, with Z-scores a, b, the likelihood of the configuration where both are causal is:

$$(1+2W)^{-1/2}\ exp(\frac{W}{2(1+2W)}[\,(a^2\ +b^2)\ +\ W(a-b)^2\,]\,), \tag{10}$$

where W is the prior variance from (Eq. 2). Since the variants are perfectly correlated, it should be the case that a=b in the absence of LD mismatch. In practice, we often observe LD=1 but a≠b; one Z-score could be large while the other is close to zero. In this case, the term $W(a-b)^2$ in the exponent is the spurious effect which should be minimized. If W is much larger than 1, then the spurious second term can end up dominating the first one. However, if W is much smaller than 1, then the null configuration can end up dominating all others, leading to posterior probabilities close to 0. To balance these requirements, BTS uses W=1.

Discussion

In this paper we report BTS, a new algorithm that performs joint fine-mapping of variants and context-mapping using genome-wide functional annotations. The algorithm has multiple statistical and algorithmic innovations to allow one to characterize genetic association signals across the genome against thousands of functional assay experiments across different tissue and cell types. The algorithm provides easily interpretable results that highlight important cellular and tissue context of genetic trait associations with genome-wide enrichment and statistical confidence. The BTS algorithm implementation is highly scalable and can allow multiple (up to *d*=5) independent association signals in each locus.

We applied BTS to summary statistics from four GWASs (immune-related and cardiovascular) and compared them against different publicly available functional genomic annotations across >200 tissue and cell types. BTS successfully prioritized relevant tissue and cellular types known to be associated with the disease biology, validating the underlying statistical model and showing the value of our approach to translate genetic findings to biological mechanisms of disease.

BTS has certain limitations. First, currently available functional annotations are still limited in their specificity in cell and tissue types, and this remains to be addressed in the future as the research community continues to generate highly specific functional experiments. Additionally, users can provide their own specialized functional annotations when running BTS. Second, the number of independent causal variants (parameter *d*) per locus needs to be chosen in advance and the memory use is still exponential, which prohibits analysis of high-order genetic interactions. In practice this might not be a serious limitation as the algorithm is still scalable up to *d*=5.

There are several useful usage scenarios for BTS by biologists, geneticists and other researchers. As complex diseases involve multiple cell types and their interactions, the choice of input tracks and genomic features can reflect multiple hypotheses involving suspected or potentially causal tissues, cell types, and disease mechanisms. BTS can then be used to test these hypotheses and select and prioritize tissues, cell types, genomic mechanisms, functional genomic regions, and variants. We note that compared to enrichment-based frameworks such as LDSC (Bulik-Sullivan et al., 2015) and S-LDSC (Finucane et al., 2015) that are commonly used to gain insight into potentially relevant cell types, tissues, BTS not only identifies relevant functional context(s), but also allows to identify and prioritize functional genomic regions for each of these potential contexts and provides fine-mapping/prioritization for variants within each of the analyzed genomic regions and for each of analyzed contexts. While multiple overlapping tracks are typically used to show or prioritize functional regions and individual variants (e.g. in the genome browser), BTS can also output per-annotation or per-locus variant posteriors which can be used to select variants with high causal posteriors across annotations. Overall, BTS provides data-driven discovery and prioritization by combining GWAS signals and FG and annotation data.

There are many directions for future development of this work. First, although BTS ranks region and annotation pairs in its output, it does not rank regions by themselves, nor does the algorithm provide affected target genes (e.g., as in Activity-by-contact (Fulco et al., 2019), Effector Index (Forgetta et al., 2022) models) regulated by the causal variants. Incorporation of expression QTL data or chromatin interaction data could address the second problem, and this idea is explored in many other papers including the INFERNO algorithm we reported previously (INFERNO (Amlie-Wolf et al., 2018), FUMA (Watanabe et al., 2017), SparkINFERNO (Kuksa et al., 2020)). BTS may also be expanded to accommodate for other types of annotations such as experimental functional assays (MPRAs) at gene or transcript level, and predicted functional activity or pathogenicity of variants (e.g., GENO-NET (He et al., 2018), CADD (Rentzsch et al., 2021; Rentzsch et al., 2019); JARVIS (Vitsios et al., 2021); PoPS (Weeks et al., 2023)), but new statistical models need to be developed to integrate these types of data.

An advantage of BTS is its scalability which allows us to explore thousands of genome-wide functional annotations and systematically explore the biological mechanism and cellular and tissue context without bias. One may extend our algorithm to analyze multiple traits and tissues jointly or carry out combinatorial analysis of multiple cell type and tissue functional surveys across many genomic features. Applications of such approach to the growing body of GWAS and WGS-based studies, along with further experimental validations, could lead to significant insights into disease-underlying variants, loci, and molecular mechanisms.

Data availability

GWAS summary statistics used in this study were obtained from EMBL-EBI GWAS catalog for IBD (https://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST003001-GCST004000/GCST003043/IBD_trans_ethnic_association_summ_stats_b37.txt.gz),

RA (http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST000001-GCST001000/GCST000679/stahl_2010_20453842_ra_efo0000685_1_gwas.sumstats.tsv.gz), and SLE (http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST003001-GCST004000/GCST003156/bentham_2015_26502338_sle_efo0002690_1_gwas.sumstats.tsv.gz), and from MRC IEU OpenGWAS database for CAD (https://gwas.mrcieu.ac.uk/files/ebi-a-GCST005195/ebi-a-GCST005195.vcf.gz).
ENCODE (https://encodeproject.org), Roadmap Epigenomics (http://www.roadmapepigenomics.org/), and EpiMap (https://compbio.mit.edu/epimap/) functional annotations used in BTS evaluations and 1000 Genomes (https://www.internationalgenome.org/) reference genotype panel data for LD computation were obtained from the FILER database (https://lisanwanglab.org/FILER). BTS results on tested GWASs are available from https://doi.org/10.5281/zenodo.14521100. BTS runs on HPC clusters, single servers or cloud-based instances and is available at https://hub.docker.com/r/wanglab/bts (Docker container) and https://bitbucket.org/wanglab-upenn/bts-pipeline (BTS pipeline source code). BTS core R package is also available at https://bitbucket.com/wanglab-upenn/BTS-R.

Acknowledgments

We thank Wang lab members, Mingyao Li, and Struan Grant for their valuable feedback and comments.

Funding

This work was supported by the National Institute on Aging [U24-AG041689, U54-AG052427, U01-AG032984].

References

Amlie-Wolf, A., Tang, M., Mlynarski, E. E., Kuksa, P. P., Valladares, O., Katanic, Z., Tsuang, D., Brown, C. D., Schellenberg, G. D., & Wang, L. S. (2018). INFERNO: inferring the molecular mechanisms of noncoding genetic variants. *Nucleic Acids Res*, *46*(17), 8740-8753. https://doi.org/10.1093/nar/gky686

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., Ntini, E., Arner, E., Valen, E., Li, K., Schwarzfischer, L., Glatz, D., Raithel, J., Lilje, B., Rapin, N., . . . Sandelin, A. (2014). An atlas of active enhancers across human cell types and tissues. *Nature*, *507*(7493), 455-461. https://doi.org/10.1038/nature12787

Asimit, J. L., Rainbow, D. B., Fortune, M. D., Grinberg, N. F., Wicker, L. S., & Wallace, C. (2019). Stochastic search and joint fine-mapping increases accuracy and identifies previously unreported associations in immune-mediated diseases. *Nat Commun*, *10*(1), 3216. https://doi.org/10.1038/s41467-019-11271-0

Baumgart, D. C., Metzke, D., Schmitz, J., Scheffold, A., Sturm, A., Wiedenmann, B., & Dignass, A. U. (2005). Patients with active inflammatory bowel disease lack immature peripheral blood plasmacytoid and myeloid dendritic cells. *Gut*, *54*(2), 228-236. https://doi.org/10.1136/gut.2004.040360

Benner, C., Spencer, C. C., Havulinna, A. S., Salomaa, V., Ripatti, S., & Pirinen, M. (2016). FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics*, *32*(10), 1493-1501. https://doi.org/10.1093/bioinformatics/btw018

Bentham, J., Morris, D. L., Graham, D. S. C., Pinder, C. L., Tombleson, P., Behrens, T. W., Martin, J., Fairfax, B. P., Knight, J. C., Chen, L., Replogle, J., Syvanen, A. C., Ronnblom, L., Graham, R. R., Wither, J. E., Rioux, J. D., Alarcon-Riquelme, M. E., & Vyse, T. J. (2015). Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet*, *47*(12), 1457-1464. https://doi.org/10.1038/ng.3434

Boix, C. A., James, B. T., Park, Y. P., Meuleman, W., & Kellis, M. (2021). Regulatory genomic circuitry of human disease loci by integrative epigenomics. *Nature*, *590*(7845), 300-307. https://doi.org/10.1038/s41586-020-03145-z

Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M., Karczewski, K. J., Park, J., Hitz, B. C., Weng, S., Cherry, J. M., & Snyder, M. (2012). Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*, *22*(9), 1790-1797. https://doi.org/10.1101/gr.137323.112

Bulik-Sullivan, B. K., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson, N., Daly, M. J., Price, A. L., & Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*, *47*(3), 291-295. https://doi.org/10.1038/ng.3211

Byrska-Bishop, M., Evani, U. S., Zhao, X., Basile, A. O., Abel, H. J., Regier, A. A., Corvelo, A., Clarke, W. E., Musunuri, R., Nagulapalli, K., Fairley, S., Runnels, A., Winterkorn, L., Lowy, E., Human Genome Structural Variation, C., Paul, F., Germer, S., Brand, H., Hall, I. M., . . . Zody, M. C. (2022). High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios. *Cell*, *185*(18), 3426-3440 e3419. https://doi.org/10.1016/j.cell.2022.08.004

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, *4*, 7. https://doi.org/10.1186/s13742-015-0047-8

Chen, J., Wang, L., De Jager, P. L., Bennett, D. A., Buchman, A. S., & Yang, J. (2022). A scalable Bayesian functional GWAS method accounting for multivariate quantitative functional annotations with applications for studying Alzheimer disease. *HGG Adv*, *3*(4), 100143. https://doi.org/10.1016/j.xhgg.2022.100143

Chen, W., Larrabee, B. R., Ovsyannikova, I. G., Kennedy, R. B., Haralambieva, I. H., Poland, G. A., & Schaid, D. J. (2015). Fine Mapping Causal Variants with an Approximate Bayesian Method Using Marginal Test Statistics. *Genetics*, *200*(3), 719-736. https://doi.org/10.1534/genetics.115.176107

Chen, W., McDonnell, S. K., Thibodeau, S. N., Tillmans, L. S., & Schaid, D. J. (2016). Incorporating Functional Annotations for Fine-Mapping Causal Variants in a Bayesian Framework Using Summary Statistics. *Genetics*, *204*(3), 933-958. https://doi.org/10.1534/genetics.116.188953

Chen, W., Wu, Y., Zheng, Z., Qi, T., Visscher, P. M., Zhu, Z., & Yang, J. (2021). Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. *Nat Commun*, *12*(1), 7117. https://doi.org/10.1038/s41467-021-27438-7

Consortium, E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57-74. https://doi.org/10.1038/nature11247

Consortium, E. P., Moore, J. E., Purcaro, M. J., Pratt, H. E., Epstein, C. B., Shoresh, N., Adrian, J., Kawli, T., Davis, C. A., Dobin, A., Kaul, R., Halow, J., Van Nostrand, E. L., Freese, P., Gorkin, D. U., Shen, Y., He, Y., Mackiewicz, M., Pauli-Behn, F., . . . Weng, Z. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, *583*(7818), 699-710. https://doi.org/10.1038/s41586-020-2493-4

Consortium, G. T. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, *369*(6509), 1318-1330. https://doi.org/10.1126/science.aaz1776

Dong, S., Zhao, N., Spragins, E., Kagda, M. S., Li, M., Assis, P., Jolanki, O., Luo, Y., Cherry, J. M., Boyle, A. P., & Hitz, B. C. (2023). Annotating and prioritizing human non-coding variants with RegulomeDB v.2. *Nat Genet*, *55*(5), 724-726. https://doi.org/10.1038/s41588-023-01365-3

Dosh, R. H., Jordan-Mahy, N., Sammon, C., & Le Maitre, C. (2019). Interleukin 1 is a key driver of inflammatory bowel disease-demonstration in a murine IL-1Ra knockout model. *Oncotarget*, *10*(37), 3559-3575. https://doi.org/10.18632/oncotarget.26894

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P. R., Anttila, V., Xu, H., Zang, C., Farh, K., Ripke, S., Day, F. R., ReproGen, C., Schizophrenia Working Group of the Psychiatric Genomics, C., Consortium, R., Purcell, S., Stahl, E., Lindstrom, S., Perry, J. R., . . . Price, A. L. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet*, *47*(11), 1228-1235. https://doi.org/10.1038/ng.3404

Foley, C. N., Staley, J. R., Breen, P. G., Sun, B. B., Kirk, P. D. W., Burgess, S., & Howson, J. M. M. (2021). A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat Commun*, *12*(1), 764. https://doi.org/10.1038/s41467-020-20885-8

Forgetta, V., Jiang, L., Vulpescu, N. A., Hogan, M. S., Chen, S., Morris, J. A., Grinek, S., Benner, C., Jang, D. K., Hoang, Q., Burtt, N., Flannick, J. A., McCarthy, M. I., Fauman, E., Greenwood, C. M. T., Maurano, M. T., & Richards, J. B. (2022). An effector index to predict target genes at GWAS loci. *Hum Genet*, *141*(8), 1431-1447. https://doi.org/10.1007/s00439-022-02434-z

Fulco, C. P., Nasser, J., Jones, T. R., Munson, G., Bergman, D. T., Subramanian, V., Grossman, S. R., Anyoha, R., Doughty, B. R., Patwardhan, T. A., Nguyen, T. H., Kane, M., Perez, E. M., Durand, N. C., Lareau, C. A., Stamenova, E. K., Aiden, E. L., Lander, E. S., & Engreitz, J. M. (2019). Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat Genet*, *51*(12), 1664-1669. https://doi.org/10.1038/s41588-019-0538-0

Galarneau, G., Palmer, C. D., Sankaran, V. G., Orkin, S. H., Hirschhorn, J. N., & Lettre, G. (2010). Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. *Nat Genet*, *42*(12), 1049-1051. https://doi.org/10.1038/ng.707

Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68-74. https://doi.org/10.1038/nature15393

Ghattas, A., Griffiths, H. R., Devitt, A., Lip, G. Y., & Shantsila, E. (2013). Monocytes in coronary artery disease and atherosclerosis: where are we now? *J Am Coll Cardiol*, *62*(17), 1541-1551. https://doi.org/10.1016/j.jacc.2013.07.043

Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., & Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association

studies using summary statistics. *PLoS Genet*, *10*(5), e1004383. https://doi.org/10.1371/journal.pgen.1004383

He, Z., Liu, L., Wang, K., & Ionita-Laza, I. (2018). A semi-supervised approach for predicting cell-type specific functional consequences of non-coding variation using MPRAs. *Nat Commun*, *9*(1), 5199. https://doi.org/10.1038/s41467-018-07349-w

Hirose, S., Lin, Q., Ohtsuji, M., Nishimura, H., & Verbeek, J. S. (2019). Monocyte subsets involved in the development of systemic lupus erythematosus and rheumatoid arthritis. *Int Immunol*, *31*(11), 687-696. https://doi.org/10.1093/intimm/dxz036

Hormozdiari, F., Kostem, E., Kang, E. Y., Pasaniuc, B., & Eskin, E. (2014). Identifying causal variants at loci with multiple signals of association. *Genetics*, *198*(2), 497-508. https://doi.org/10.1534/genetics.114.167908

Hormozdiari, F., van de Bunt, M., Segre, A. V., Li, X., Joo, J. W. J., Bilow, M., Sul, J. H., Sankararaman, S., Pasaniuc, B., & Eskin, E. (2016). Colocalization of GWAS and eQTL Signals Detects Target Genes. *Am J Hum Genet*, *99*(6), 1245-1260. https://doi.org/10.1016/j.ajhg.2016.10.003

Ip, W. K. E., Hoshi, N., Shouval, D. S., Snapper, S., & Medzhitov, R. (2017). Anti-inflammatory effect of IL-10 mediated by metabolic reprogramming of macrophages. *Science*, *356*(6337), 513-519. https://doi.org/10.1126/science.aal3535

Kanai, M., Elzur, R., Zhou, W., Global Biobank Meta-analysis, I., Daly, M. J., & Finucane, H. K. (2022). Meta-analysis fine-mapping is often miscalibrated at single-variant resolution. *Cell Genom*, *2*(12). https://doi.org/10.1016/j.xgen.2022.100210

Kichaev, G., Roytman, M., Johnson, R., Eskin, E., Lindstrom, S., Kraft, P., & Pasaniuc, B. (2017). Improved methods for multi-trait fine mapping of pleiotropic risk loci. *Bioinformatics*, *33*(2), 248-255. https://doi.org/10.1093/bioinformatics/btw615

Kichaev, G., Yang, W. Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A. L., Kraft, P., & Pasaniuc, B. (2014). Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet*, *10*(10), e1004722. https://doi.org/10.1371/journal.pgen.1004722

Knight, J., Spain, S. L., Capon, F., Hayday, A., Nestle, F. O., Clop, A., Wellcome Trust Case Control, C., Genetic Analysis of Psoriasis, C., Consortium, I. c. f. P., Barker, J. N., Weale, M. E., & Trembath, R. C. (2012). Conditional analysis identifies three novel major histocompatibility complex loci associated with psoriasis. *Hum Mol Genet*, *21*(23), 5185-5192. https://doi.org/10.1093/hmg/dds344

Kuksa, P. P., Lee, C. Y., Amlie-Wolf, A., Gangadharan, P., Mlynarski, E. E., Chou, Y. F., Lin, H. J., Issen, H., Greenfest-Allen, E., Valladares, O., Leung, Y. Y., & Wang, L. S. (2020). SparkINFERNO: a scalable high-throughput pipeline for inferring molecular mechanisms of non-coding genetic variants. *Bioinformatics*, *36*(12), 3879-3881. https://doi.org/10.1093/bioinformatics/btaa246

Kuksa, P. P., Leung, Y. Y., Gangadharan, P., Katanic, Z., Kleidermacher, L., Amlie-Wolf, A., Lee, C. Y., Qu, L., Greenfest-Allen, E., Valladares, O., & Wang, L. S. (2022). FILER: a framework for harmonizing and querying large-scale functional genomics knowledge. *NAR Genom Bioinform*, *4*(1), lqab123. https://doi.org/10.1093/nargab/lqab123

Libby, P., & Theroux, P. (2005). Pathophysiology of coronary artery disease. *Circulation*, *111*(25), 3481-3488. https://doi.org/10.1161/CIRCULATIONAHA.105.537878

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., Ripke, S., Lee, J. C., Jostins, L., Shah, T., Abedian, S., Cheon, J. H., Cho, J., Dayani, N. E., Franke, L., Fuyuno, Y., Hart, A., Juyal, R. C., Juyal, G., . . . Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet*, *47*(9), 979-986. https://doi.org/10.1038/ng.3359

Pickrell, J. K. (2014). Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am J Hum Genet*, *94*(4), 559-573. https://doi.org/10.1016/j.ajhg.2014.03.004

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, *81*(3), 559-575. https://doi.org/10.1086/519795

Rentzsch, P., Schubach, M., Shendure, J., & Kircher, M. (2021). CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med*, *13*(1), 31. https://doi.org/10.1186/s13073-021-00835-9

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*, *47*(D1), D886-D894. https://doi.org/10.1093/nar/gky1016

Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., . . . Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*(7539), 317-330. https://doi.org/10.1038/nature14248

Srikakulapu, P., & McNamara, C. A. (2017). B cells and atherosclerosis. *Am J Physiol Heart Circ Physiol*, *312*(5), H1060-H1067. https://doi.org/10.1152/ajpheart.00859.2016

Stahl, E. A., Raychaudhuri, S., Remmers, E. F., Xie, G., Eyre, S., Thomson, B. P., Li, Y., Kurreeman, F. A., Zhernakova, A., Hinks, A., Guiducci, C., Chen, R., Alfredsson, L., Amos, C. I., Ardlie, K. G., Consortium, B., Barton, A., Bowes, J., Brouwer, E., . . . Plenge, R. M. (2010). Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet*, *42*(6), 508-514. https://doi.org/10.1038/ng.582

Stankey, C. T., Bourges, C., Haag, L. M., Turner-Stokes, T., Piedade, A. P., Palmer-Jones, C., Papa, I., Silva Dos Santos, M., Zhang, Q., Cameron, A. J., Legrini, A., Zhang, T., Wood, C. S., New, F. N., Randzavola, L. O., Speidel, L., Brown, A. C., Hall, A., Saffioti, F., . . . Lee, J. C. (2024). A disease-associated gene desert directs macrophage inflammation through ETS2. *Nature*, *630*(8016), 447-456. https://doi.org/10.1038/s41586-024-07501-1

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, *1*(1), 59. https://doi.org/10.1038/s43586-021-00056-9

van der Harst, P., & Verweij, N. (2018). Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res*, *122*(3), 433-443. https://doi.org/10.1161/CIRCRESAHA.117.312086

Vitsios, D., Dhindsa, R. S., Middleton, L., Gussow, A. B., & Petrovski, S. (2021). Prioritizing non-coding regions based on human genomic constraint and sequence context with deep learning. *Nat Commun*, *12*(1), 1504. https://doi.org/10.1038/s41467-021-21790-4

Wallace, C. (2021). A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet*, *17*(9), e1009440. https://doi.org/10.1371/journal.pgen.1009440

Wang, G., Sarkar, A., Carbonetto, P., & Stephens, M. (2020). 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. https://doi.org/10.1111/rssb.12388

Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*, *8*(1), 1826. https://doi.org/10.1038/s41467-017-01261-5

Weeks, E. M., Ulirsch, J. C., Cheng, N. Y., Trippe, B. L., Fine, R. S., Miao, J., Patwardhan, T. A., Kanai, M., Nasser, J., Fulco, C. P., Tashman, K. C., Aguet, F., Li, T., Ordovas-Montanes, J., Smillie, C. S., Biton, M., Shalek, A. K., Ananthakrishnan, A. N., Xavier, R. J., . . . Finucane, H. K. (2023). Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. *Nat Genet*, *55*(8), 1267-1276. https://doi.org/10.1038/s41588-023-01443-6

Weissbrod, O., Hormozdiari, F., Benner, C., Cui, R., Ulirsch, J., Gazal, S., Schoech, A. P., van de Geijn, B., Reshef, Y., Marquez-Luna, C., O'Connor, L., Pirinen, M., Finucane, H. K., & Price, A. L. (2020). Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat Genet*, *52*(12), 1355-1363. https://doi.org/10.1038/s41588-020-00735-5

Wen, X., Lee, Y., Luca, F., & Pique-Regi, R. (2016). Efficient Integrative Multi-SNP Association Analysis via Deterministic Approximation of Posteriors. *Am J Hum Genet*, *98*(6), 1114-1129. https://doi.org/10.1016/j.ajhg.2016.03.029

Wen, X., Pique-Regi, R., & Luca, F. (2017). Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. *PLoS Genet*, *13*(3), e1006646. https://doi.org/10.1371/journal.pgen.1006646

Weyand, C. M., Bryl, E., & Goronzy, J. J. (2000). The role of T cells in rheumatoid arthritis. *Arch Immunol Ther Exp (Warsz)*, *48*(5), 429-435. https://www.ncbi.nlm.nih.gov/pubmed/11140470

Yang, J., Ferreira, T., Morris, A. P., Medland, S. E., Genetic Investigation of, A. T. C., Replication, D. I. G., Meta-analysis, C., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Weedon, M. N., Loos, R. J., Frayling, T. M., McCarthy, M. I., Hirschhorn, J. N., Goddard, M. E., & Visscher, P. M. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*, *44*(4), 369-375, S361-363. https://doi.org/10.1038/ng.2213

Yang, Z., Wang, C., Liu, L., Khan, A., Lee, A., Vardarajan, B., Mayeux, R., Kiryluk, K., & Ionita-Laza, I. (2023). CARMA is a new Bayesian model for fine-mapping in genome-wide association meta-analyses. *Nat Genet*, *55*(6), 1057-1065. https://doi.org/10.1038/s41588-023-01392-0

Zhou, J., Li, H., Xia, X., Herrera, A., Pollock, N., Reebye, V., Sodergren, M. H., Dorman, S., Littman, B. H., Doogan, D., Huang, K. W., Habib, R., Blakey, D., Habib, N. A., & Rossi, J. J. (2019). Anti-inflammatory Activity of MTL-CEBPA, a Small Activating RNA Drug, in LPS-Stimulated Monocytes and Humanized Mice. *Mol Ther*, *27*(5), 999-1016. https://doi.org/10.1016/j.ymthe.2019.02.018

Zou, Y., Carbonetto, P., Wang, G., & Stephens, M. (2022). Fine-mapping from summary data with the "Sum of Single Effects" model. *PLoS Genet*, *18*(7), e1010299. https://doi.org/10.1371/journal.pgen.1010299