KnockoffTrio: A knockoff framework for the identification of putative causal variants in genome-wide association studies with trio design

Yi Yang[1], Chen Wang[1], Linxi Liu[2], Joseph Buxbaum[3], Zihuai He[4,5], Iuliana Ionita-Laza[1,#]

[1] Department of Biostatistics, Columbia University, New York, NY 10032
[2] Department of Statistics, University of Pittsburgh, Pittsburgh, PA, 15260, USA
[3] Departments of Psychiatry, Neuroscience, and Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA
[4] Quantitative Sciences Unit, Department of Medicine, Stanford University, Stanford, CA, 94305, USA
[5] Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA 94305, USA
[#] Correspondance: ii2135@columbia.edu

Abstract

Family-based designs can eliminate confounding due to population substructure and can distinguish direct from indirect genetic effects but they are underpowered due to limited sample sizes. Here we propose KnockoffTrio, a novel statistical method to identify putative causal genetic variants for father-mother-child trio design built upon a recently developed knockoff framework in statistics. KnockoffTrio controls the false discovery rate (FDR) in the presence of arbitrary correlations among tests, and is less conservative and thus more powerful than the conventional methods that control the family-wise error rate via Bonferroni correction. Furthermore, KnockoffTrio is not restricted to family-based association tests and can be used in conjunction with more powerful, potentially nonlinear models to improve power of standard family-based tests. We show using empirical simulations that KnockoffTrio can prioritize causal variants over associations due to linkage disequilibrium. In applications to 14,200 trios from three study cohorts for autism spectrum disorders (ASD), including AGP, SPARK, and SSC we show that KnockoffTrio can identify multiple significant associations that are missed by conventional tests applied to the same data. In particular, we replicate known ASD association signals with variants in several genes such as *MACROD2*, *NRXN1*, *PRKAR1B*, *CADM2*, *PCDH9*, and *DOCK4* and identify additional associations with variants in other genes including *ARHGEF10*, *SLC28A1*, *ZNF589*, and *HINT1* at FDR 10%.

Introduction

The father-mother-child trio design is a popular family-based design, especially for early-onset diseases. One important example is autism spectrum disorders (ASD) where several prominent studies have successfully employed such a design[1,2,3]. The main advantages of the family-based design are that it is robust to external confounders such as population structure[4,5] and can help distinguish between direct and indirect effects[6]. Although popular methods have been proposed to account for confounding effects of population structure in the context of population-based designs[7,8,9], a more reliable approach to eliminating such confounders is to use randomized experiments, and family-based designs provide an analogy to such experiments because of the randomness in transmission of genetic material from parents to offspring[10]. However, a main limitation of GWAS studies with family-based designs is the modest sample sizes, which ultimately lead to reduced power.

Most of the existing studies have focused on controlling the family-wise error rate (FWER) to account for multiple testing in genome-wide association studies. Given the polygenic nature of many complex traits, with a large number of small effect loci accounting for most of the trait heritability, a more meaningful and powerful strategy is to control the false-discovery rate (FDR) that quantifies the expected proportion of false discoveries. Control of FDR has been previously suggested in genome-wide association studies [11,12], and has been successfully employed in genetic association studies of ASD [13,14]. Valid control of FDR is however difficult to achieve using the standard Benjamini-Hochberg procedure due to possible complex correlations among genetic variants. The knockoff-based framework we employ here allows valid FDR control under arbitrary correlations.

The idea of the knockoff-based inference is to construct knockoff copies of the original features (genotypes) that preserve the correlation structure and are independent of the trait conditional on the original features [15]. These knockoff features serve as negative controls and when compared with the original features help identify the truly causal ones. The knockoff-based inference provides rigorous control of FDR under arbitrary correlation structure and is thus more versatile than the Benjamini-Hochberg (BH) procedure that requires independence or positive dependence [16] for the FDR control. Several knockoff procedures have been proposed with applications to population-based designs, including KnockoffZoom [17] for genome-wide association studies based on hidden Markov models and KnockoffScreen [18] for whole-genome sequencing data based on the sequential conditional independent tuples (SCIT) algorithm. These methods, however, were designed for independent individuals in population-based studies, making them unsuitable to family-based studies as considered in this article. A related approach to construct synthetic offspring has been proposed before in order to perform causal inference with trio designs [10]. Specifically, Bates et al. proposed a digital twin test based on the conditional randomization test [19], a method related to the knockoff but which can produce valid empirical p-values. Computational cost is a concern for this test especially in high-dimensional genome-wide settings where a large number of random drawings are needed to get small empirical p-values.

In this paper we propose KnockoffTrio, a knockoff-based framework for the analysis of trio data in genome-wide association studies. Conventional association tests for family-based designs include the family-based association test (FBAT) [5], a generalization of the transmission disequilibrium test (TDT) [20] to handle various practical complexities such as missing parental data, covariate adjustment, and different types of phenotypes. Methods based on kernel machine regression under a generalized linear mixed model framework have also been proposed for family-based designs [21,22] and for population-based designs adjusting for population structure and relatedness [8]. Compared to these conventional testing strategies, KnockoffTrio enjoys several advantages of the general knockoff-based inference such as higher statistical power, prioritization of causal variants over associations due to linkage disequilibrium, and robustness in controlling false positives in the presence of linkage disequilibrium between causal and non-causal variants [18,17]. Furthermore, KnockoffTrio can leverage more general machine learning models while increasing power and maintaining proper FDR control regardless of the validity of the assumed model.

Methods

Knockoff generation for trio design

We assume a study with n trios and p genetic variants. We denote the matrix of trio genotypes by $\boldsymbol{G} \in \{0,1,2\}^{3n\times p}$. Our goal is to test the conditional null hypothesis

$$H_0 : \boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{G}_g | \boldsymbol{G}_{-g}$$

where $\boldsymbol{Y}$ are the phenotypes and $g \subset [p]$ is a continuous block. That is, variant(s) in group g (e.g., a gene or a region) are null if $\boldsymbol{Y}$ is independent of $\boldsymbol{G}_g$ given variants outside g.

We describe a knockoff generation method for the trio design to capture sample relatedness and test the above hypothesis. Our method assumes knowledge of haplotype phase; most phasing algorithms are able to provide highly accurate estimates of haplotypes when applied to trio data sets [23]. We first generate knockoff haplotypes for the parents, and then conditional on them we generate the knockoff haplotypes for the offspring. We describe the algorithm as follows:

Algorithm 1 Generation of knockoff trios

1. Sample one haplotype from each father into a group; assign the remaining haplotypes to the second group.
2. Repeat Step 1 for mothers and obtain two additional groups of haplotypes.
3. Apply the SCIP algorithm[18] to each group of haplotypes and obtain the corresponding knockoffs (see below).
4. Generate knockoff offspring haplotypes conditional on the knockoff parental haplotypes (see below).

Note that in Steps 1 and 2, we assign an individual's two haplotypes to two separate groups when generating their knockoffs so that the permutation-based SCIP algorithm below does not use the residual from one haplotype to generate the other haplotype's knockoff. This is done to increase the contrast between the original and knockoff genotypes in an individual, and hence to improve power.

SCIP algorithm to generate knockoff parental haplotypes. We adopt the residual permutation method proposed in KnockoffScreen[18] to generate knockoff haplotypes for the parents. The residual permutation method is based on the general sequential conditional independent pairs (SCIP) algorithm[19], defined as follows:

Algorithm 2 SCIP algorithm for knockoff haplotype generation

j=1
while $j \leq p$ **do**
  Sample $\tilde{H}_j$ independently from $L(H_j|H_{k\in B_j}, \tilde{H}_{1\leq k\leq j-1,k\in B_j})$
  $j = j + 1$
end while

where H_j and $\tilde{H}_j$ denote the original and knockoff parental haplotypes for the jth variant, respectively, and B_j denotes the subset of variants in a neighborhood of the jth variant (+/-100kb from the variant). Algorithm 2 has been shown to generate knockoffs that preserve the exchangeability conditions between the original and the knockoff genotypes necessary for controlling the FDR[18]. In the context of genetic data, the exchangeability describes the invariance in the LD structure when one swaps a subset S of genetic variants with their knockoffs, i.e., $(H, \tilde{H})_{\text{swap}(S)} \overset{D}{=} (H, \tilde{H})$, in which $(H, \tilde{H})_{\text{swap}(S)}$ is obtained from $(H, \tilde{H})$ by swapping H_j and $\tilde{H}_j$, $\forall j \in S$.

As in He et al.[18] we consider a semiparametric model for $L(H_j|H_{k\in B_j}, \tilde{H}_{1\leq k\leq j-1,k\in B_j})$ in KnockoffTrio:

$$H_j = \beta_0 + \sum_{k\neq j, k\in B_j} \beta_k H_k + \sum_{k\leq j-1, k\in B_j} \gamma_k \tilde{H}_k + \epsilon_j,$$

where ϵ_j is a random error term with a mean of zero. We obtain $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$, fitted values $\hat{\boldsymbol{H}}_{\boldsymbol{j}}$, and residuals $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{j}} = \boldsymbol{H}_{\boldsymbol{j}} - \hat{\boldsymbol{H}}_{\boldsymbol{j}}$ by minimizing the mean squared loss. We then obtain permuted residuals $\hat{\boldsymbol{\epsilon}}^*_{\boldsymbol{j}}$ and define the parental knockoffs $\tilde{\boldsymbol{H}}_{\boldsymbol{j}} = \hat{\boldsymbol{H}}_{\boldsymbol{j}} + \hat{\boldsymbol{\epsilon}}^*_{\boldsymbol{j}}$.

Generating knockoff offspring haplotypes. Conditional on the knockoff parental haplotypes generated as above, we then proceed to generate the knockoff offspring haplotypes. Given the phased haplotypes of the original trio for a region, we first infer which parental haplotypes were transmitted to the offspring by matching parental haplotypes with offspring haplotypes. We assume that no recombination occurs in the transmission of haplotypes from parents to offspring in any small region. We then use the knockoff haplotypes that correspond to the transmitted haplotypes in the original trio as the offspring's knockoff haplotypes.

Exchangeability property. As with independent samples, we need certain exchangeability properties to hold for the trio design in order for the FDR control to hold[15]. We formally prove the exchangeability property and FDR control for the trio design in Appendix C.

Multiple knockoffs to improve power and stability. The knockoff generation algorithm described above generates one single knockoff haplotype for each original haplotype. However, the inference based on a single knockoff often has limited power due to the detection threshold of $\frac{1}{q}$, i.e. the number of independent signals required for making any discoveries at the target FDR q. In particular, there is no power at the target FDR q if there are fewer than $\frac{1}{q}$ discoveries to be made, which is not uncommon when q is low and the signal is sparse. Moreover, the randomness in the sampling of a single knockoff makes the results unstable particularly for weak causal effects. Therefore, to further improve the power and stability, we extend the above single knockoff algorithm to generating multiple knockoffs. For M knockoffs, the detection threshold decreases from $\frac{1}{q}$ to $\frac{1}{Mq}$, making it more powerful to detect sparse signals even when the target FDR level q is low. Furthermore, multiple knockoffs help improve the stability and reproducibility of the results.

Algorithm 3 SCIT algorithm for multiple knockoffs

j=1
while $j \leq p$ **do**
  Sample $\tilde{H}_j^1, ..., \tilde{H}_j^M$ independently from $L(H_j|H_{k\in B_j}, \tilde{H}^1_{1\leq k\leq j-1,k\in B_j}, ..., \tilde{H}^M_{1\leq k\leq j-1,k\in B_j})$
  $j = j + 1$
end while

The semiparametric model for $L(H_j|H_{k\in B_j}, \tilde{H}^1_{1\leq k\leq j-1,k\in B_j}, ..., \tilde{H}^M_{1\leq k\leq j-1,k\in B_j})$ in the multiple-knockoff setting is:

$$H_j = \beta_0 + \sum_{k\neq j,k\in B_j} \beta_k H_k + \sum_{1\leq m\leq M} \sum_{k\leq j-1,k\in B_j} \gamma_k^m \tilde{H}_k^m + \epsilon_j,$$

where ϵ_j is a random error term with a mean of zero. We obtain $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\gamma}}$, fitted values $\hat{\boldsymbol{H}}_{\boldsymbol{j}}$, and the residuals $\hat{\boldsymbol{\epsilon}}_{\boldsymbol{j}}$ and their permutations $\hat{\boldsymbol{\epsilon}}^*_{\boldsymbol{j}}$. We then define the mth knockoff $\tilde{\boldsymbol{H}}^{\boldsymbol{m}}_{\boldsymbol{j}}=\hat{\boldsymbol{H}}_{\boldsymbol{j}}+\hat{\boldsymbol{\epsilon}}^{*\boldsymbol{m}}_{\boldsymbol{j}}$.

KnockoffTrio: A Knockoff Framework for Trio Design

We describe here a knockoff-based test using a family-based association test (FBAT) to compute the importance scores.

KnockoffTrio-FBAT. Once the knockoff generation for the father-mother-child trio data is completed, KnockoffTrio-FBAT performs a genome-wide scanning procedure with a window ϕ_{kl} in both the original and the knockoff data. We consider several candidate window sizes (e.g. in our applications 1bp, 1kb, 5kb, 10kb, 20kb, and 50kb), for ϕ_{kl}, with half of each window overlapping with neighboring windows of the same size. We employ the weighted burden FBAT[24], which is a generalization of the SNP-based FBAT for a set of variants. Let n denote the number of trios and p denote the number of variants in a window. When $p = 1$, the weighted burden FBAT is equivalent to the SNP-based FBAT. The weighted burden FBAT statistic W_w for trio design is computed as:

$$W_w = \sum_{j=1}^{p} w_j U_j,$$

$$U_j = \sum_{i=1}^{n} (Y_i - u) U_{ij},$$

$$U_{ij} = X_{ij} - E(X_{ij}|P^1_{ij}, P^2_{ij}),$$

in which w_j is a weight associated with the jth variant, Y_i is a dichotomous or quantitative trait for the offspring in the ith trio, u is an offset parameter, X_{ij} is the offspring genotype, P^1_{ij} and P^2_{ij} are the parental genotypes, and $E(X_{ij}|P^1_{ij}, P^2_{ij})$ is the expected value of the offspring genotype conditional on parental genotypes. Typically, $u = 0$ for dichotomous traits and $u = \bar{\boldsymbol{Y}}$ for quantitative traits. The choice of w_j is flexible and can reflect any prior functional information on the variant; in this study we consider $w_j = (\sqrt{np_j(1-p_j)})^{-1}$,

in which n is the number of trios and p_j is the minor allele frequency (MAF) for the jth variant. We can further obtain the variance of W_w as

$$Var(W_w) = \sum_{i=1}^{n}(Y_i - \mu_i)^2\Big[\sum_{j=1}^{p} w_j^2 Var(X_{ij}|P_{ij}^1, P_{ij}^2) + \sum_{j\neq k} w_j w_k Cov(X_{ij}, X_{ik}|P_{ij}^1, P_{ij}^2, P_{ik}^1, P_{ik}^2)\Big].$$

Therefore, the standardized test statistic $Z = W_w/\sqrt{Var(W_w)}$ approximately follows a standard normal distribution in large samples under the null hypothesis of no association between any of the p variants and the trait.

Aggregated Cauchy association test to compute importance scores. For a given window ϕ_{kl} we compute an importance score as follows:

- For a 1bp window, KnockoffTrio-FBAT implements SNP-based FBAT for variants with a MAF ≥ 0.01 and obtain $p_{\phi_{kl}}$ and $p_{\phi_{kl}}^m$ (for the mth knockoff).
- For a 1kb, 5kb, 10kb, 20kb, or 50kb window, KnockoffTrio-FBAT implements
  1. Weighted burden FBAT for variants with MAF$\geq$0.01.
  2. SNP-based FBAT for variants with MAF$\geq$0.01.
  3. The aggregated Cauchy association test (ACAT)[25] to combine the p-values in Steps 1 and 2 and obtain $p_{\phi_{kl}}$ and $p_{\phi_{kl}}^m$.

KnockoffTrio-X. The application of KnockoffTrio is not restricted to the FBAT test. Alternatively, p-values can be obtained from different, more sophisticated methods that can help increase power in complex scenarios, e.g., the error terms for quantitative traits are not normally distributed. As a proof of concept, we investigate in simulations KnockoffTrio-iQRAT, in which we replace FBAT with iQRAT, a novel gene-level association test that integrates quantile rank score process to accommodate more complex, non-linear associations[26]. iQRAT considers a quantile model for quantitative trait Y:

$$Q_{Y_i}(\tau) = \alpha_0(\tau) + \boldsymbol{\beta}(\tau)^\top \boldsymbol{X}_i^{\text{adj}},$$

where $\tau \in (0,1)$ is the quantile level, $\boldsymbol{\beta}(\tau)^\top = (\beta_1(\tau), \beta_2(\tau), ..., \beta_p(\tau))$ is the quantile coefficient functions, $\alpha_0(\tau)$ is the intercept function, and $\boldsymbol{X}_i^{\text{adj}} = \boldsymbol{X}_i - E(\boldsymbol{X}_i|\boldsymbol{P}_i^1, \boldsymbol{P}_i^2)$ is the adjusted offspring genotype where we subtract the conditional expectation (conditional on parental genotypes) so that it corresponds to FBAT formulation. iQRAT tests the null hypothesis $\boldsymbol{\beta}(\tau) = \boldsymbol{0}, \forall \tau \in (0,1)$. The iQRAT statistics that generalize the sequence kernel association tests (S) and Burden tests (B) are respectively computed as:

$$Q_S^\varphi = \boldsymbol{S}^{\varphi\top}\boldsymbol{W}^2\boldsymbol{S}^\varphi,$$

$$Q_B^\varphi = \boldsymbol{S}^{\varphi\top}\boldsymbol{W}1_p 1_p^\top \boldsymbol{W}\boldsymbol{S}^\varphi,$$

where $\boldsymbol{S}^\varphi = n^{-1/2}\sum_{i=1}^n \boldsymbol{X}_i^{\text{adj}\top}\hat{\phi}_i^\varphi$, $\hat{\phi}_i^\varphi = \int_0^1 \hat{a}_i(\tau)d\varphi(\tau)$, $\hat{a}_i(\tau) = \mathbf{1}\{Y_i < \hat{\alpha}_0(\tau)\} - \tau$, $\varphi(\tau)$ is the weight function, $\hat{\alpha}_0(\tau)$ is the estimated intercept via quantile regression under the null, and $\boldsymbol{W} = \text{diag}(w_1, ..., w_p)$ is the weight matrix. iQRAT considers four different weight functions and combines the results using ACAT. We use Q_B^φ, the burden version of iQRAT, in KnockoffTrio-iQRAT so that it is comparable to the burden FBAT in KnockoffTrio-FBAT.

Knockoff filter procedure for FDR control. For each given window ϕ_{kl}, KnockoffTrio calculates a feature statistic, defined as

$$W_{\phi_{kl}} = (T_{\phi_{kl}} - \text{median } T_{\phi_{kl}}^m)I_{T_{\phi_{kl}} \geq \max T_{\phi_{kl}}^m}, \quad (1)$$

in which $T_{\phi_{kl}} = -\log_{10} p_{\phi_{kl}}$ and $T_{\phi_{kl}}^m = -\log_{10} p_{\phi_{kl}}^m$ where $p_{\phi_{kl}}$ and $p_{\phi_{kl}}^m$ are the p-values computed above for the original and the knockoff trios, respectively. KnockoffTrio then calculates a threshold τ and selects

windows with $W_{\phi_{kl}} > \tau$ while controlling the FDR at a target level q. The corresponding value of τ is computed as (see also KnockoffScreen[18]):

$$\tau = \min\Big\{t > 0 : \frac{\frac{1}{M} + \frac{1}{M}\#\{\phi_{kl} : \kappa_{\phi_{kl}} \geq 1, \tau_{\phi_{kl}} \geq t\}}{\#\{\phi_{kl} : \kappa_{\phi_{kl}} = 0, \tau_{\phi_{kl}} \geq t\}} \leq q\Big\}, \tag{2}$$

where $\tau_{\phi_{kl}} = T^{(0)}_{\phi_{kl}} - \text{median}T^{(m)}_{\phi_{kl}}$ is the largest importance score minus the median of the remaining importance scores; $\kappa_{\phi_{kl}} = 0$ when $T_{\phi_{kl}}$ is the largest importance score, and $\kappa_{\phi_{kl}} = m$ when $T^m_{\phi_{kl}}$ for the mth knockoff is the largest importance score.

We show a schematic flowchart for KnockoffTrio in Figure 1.

Calculation of q-values. We also calculate a q-value $q_{\phi_{kl}}$ for ϕ_{kl}, which is the p-value analogue in the FDR setting and unifies $W_{\phi_{kl}}$ and τ for declaring significance. Specifically, the q-value is the minimum FDR when all tests that show evidence against the null hypothesis at least as strong as the current test are declared as significant. Under the knockoff framework, we follow KnockoffScreen and define the q-value for window ϕ as

$$q_\phi = \min_{t \leq \tau_\phi} \frac{\frac{1}{M} + \frac{1}{M}\#\{\phi_{kl} : \kappa_{\phi_{kl}} \geq 1, \tau_{\phi_{kl}} \geq t\}}{\#\{\phi_{kl} : \kappa_{\phi_{kl}} = 0, \tau_{\phi_{kl}} \geq t\}},$$

where $\frac{\frac{1}{M} + \frac{1}{M}\#\{\phi_{kl}:\kappa_{\phi_{kl}} \geq 1, \tau_{\phi_{kl}} \geq t\}}{\#\{\phi_{kl}:\kappa_{\phi_{kl}}=0, \tau_{\phi_{kl}} \geq t\}}$ is the estimated FDR if we declare significant windows with feature statistics $\kappa_{\phi_{kl}} = 0, \tau_{\phi_{kl}} \geq t$. We define $q_\phi = 1$ for windows with $\kappa_{\phi_{kl}} > 0$ so that they will not be selected. By definition, the windows selected by $W_\phi > \tau$ are equivalent to those selected by $q_\phi < q$, where q is the target FDR.

Meta-analysis for KnockoffTrio. For a variant or set of variants, meta-analysis can be performed by integrating summary statistics from individual studies into a combined summary statistic. KnockoffTrio can be naturally extended to the meta-analysis setting because KnockoffTrio's feature statistics are defined based on summary statistics for the original and the knockoff cohorts. Here, we implement the sample-size-based meta-analysis[27] into KnockoffTrio. Specifically, KnockoffTrio's meta-analysis procedure is defined as follows:

1. For the ith study, obtain $Z_{\phi_{kl},i}$ for a window ϕ_{kl} in the original cohort and $Z^m_{\phi_{kl},i}$ for the same window in the mth knockoff cohort; $Z_{\phi_{kl},i}$ and $Z^m_{\phi_{kl},i}$ are the standardized SNP-based FBAT statistics for a single-variant window or the set-based FBAT statistics for a multi-variant window.

2. Calculate $Z_{\phi_{kl},meta} = \frac{\Sigma_i w_i Z_{\phi_{kl},i}}{\sqrt{\Sigma_i w_i^2}}$ for the original cohort and $Z^m_{\phi_{kl},meta} = \frac{\Sigma_i w_i Z^m_{\phi_{kl},i}}{\sqrt{\Sigma_i w_i^2}}$ for the mth knockoff cohort, in which $w_i = \sqrt{N_i}$ is the weight and N_i is the sample size (i.e., the number of trios) for the ith study.

3. Calculate $p_{\phi_{kl},meta} = 2\Phi(-|Z_{\phi_{kl},meta}|)$ for the original cohort and $p^m_{\phi_{kl},meta} = 2\Phi(-|Z^m_{\phi_{kl},meta}|)$ for the mth knockoff cohort.

4. Calculate $W_{\phi_{kl},meta}$ and τ_{meta} using Equations (1) and (2).

Code availability. KnockoffTrio has been implemented in an R package available at: https://cran.r-project.org/web/packages/KnockoffTrio.

Results

Simulation Studies

We simulate genetic data based on the Autism Genome Project (AGP) cohort. The AGP cohort consists of 798,961 common (MAF$\geq$0.05) and low-frequency (0.01$\leq$MAF$<$0.05) variants for 1,266 trio families of European ancestry. For a simulation replicate, we simulate 10,000 trios with common and low-frequency variants sampled from a 1Mb region (chr20:15,981,843-16,981,842; 495 variants with MAF$\geq$0.01) near the

MACROD2 gene. In line with previous studies[28,18], we applied hierarchical clustering such that variants from different clusters have correlation no greater than 0.7 and then randomly selected one representative variant from each cluster to be included in the replicate. For a trio, we sampled four haplotypes from the phased AGP data for the parents and simulated the genotypes for the offspring using two of the four haplotypes, each randomly selected from a parent.

KnockoffTrio preserves exchangeability in trio studies. The rationale of the proposed algorithm is to augment the original trios with synthetic trios. The knockoff construction proposed here ensures the exchangeability property between the original and synthetic genotypes, i.e. if we swap any subset of variants with their synthetic counterparts, the joint haplotype distribution for the trio remains the same (see formal proof in Appendix C). This exchangeability property is a necessary condition for the FDR control. We verify the exchangeability for the offspring haplotypes using simulations. We generated a replicate of 10,000 trios with variants sampled from a 1Mb region as described above. To validate the exchangeability, we generated the offspring knockoff haplotypes using the proposed algorithm in KnockoffTrio and evaluated whether the covariance between each pair of variants is exchangeable for the common variants in the region. As shown in Figure 2, the exchangeability property holds in simulations.

Empirical power and FDR in single-locus simulations. We performed simulations to evaluate the power and empirical FDR of KnockoffTrio. We simulated 500 replicates as described above. We generated the dichotomous trait for the offspring using a logit model:

$$\text{logit}(Y_i) = \beta_0 + \beta_1 X_{i1} + ... + \beta_p X_{ip},$$

and the quantitative trait using a linear model:

$$Y_i = \beta_1 X_{i1} + ... + \beta_p X_{ip} + \epsilon_i,$$

where $\epsilon_i \sim N(0,1)$, β_0 was set such that the disease prevalence is 1% and $\text{logit}(x) = \log\frac{x}{1-x}$. We randomly selected three variants within a 1kb signal window to be causal with the causal effect $\beta_j = 0.2|\log_{10}\text{MAF}_j|$. For dichotomous traits, we include a trio only when $Y_i = 1$ to mimic the usual ascertainment in real trio design studies with dichotomous traits.

For each replicate, we generated multiple knockoffs ($M = 1, 4, 6, 8,$ and 10) and used several window sizes to scan the region (1bp, 1kb, 5kb, 10kb, 20kb, and 50kb). We evaluated the performance of KnockoffTrio in terms of different numbers of knockoffs for both dichotomous and quantitative traits. For each replicate, the power is the proportion of detected causal windows (i.e. windows that contain at least one causal variant) among all causal windows and the FDR is the proportion of non-causal windows among all detected windows. The power and FDR were averaged over the 500 replicates. As shown in Figure 3, KnockoffTrio controls the FDR at the target level in all scenarios considered. The power of KnockoffTrio increases when the number of knockoffs increases, especially at low target FDR levels as expected due to the detection threshold issue mentioned in the Methods section.

KnockoffTrio prioritizes causal variants over false positive associations due to linkage disequilibrium. Based on the single-locus simulations, we further compared KnockoffTrio with the conventional association test that controls the FWER in terms of (1) the proportion of selected windows that overlap with the 1kb signal window, and (2) the median distance of selected windows to the 1kb signal window. For the conventional association test, we used the same aggregated Cauchy association test implemented in KnockoffTrio for each window and controlled the FWER using the Bonferroni correction. As shown in Figures 4A and 4B, the windows selected by KnockoffTrio have a substantially higher chance of overlapping with the signal window and a shorter distance to the signal window than the conventional method. We also randomly selected 200 false positives identified by the conventional association test with Bonferroni correction from all simulated replicates and showed the relationship between their significance and the maximum correlation with any causal variants in the left panel of Figure 4C. As the correlation increases, the conventional association test yields more significant p-values for the false positives. On the other hand, for these same 200 variants, KnockoffTrio has a much higher chance of correctly identifying these non-causal variants as true negatives as shown in the right panel of Figure 4C, and thus is substantially more robust in controlling false positives in the presence of linkage disequilibrium between causal and non-causal variants.

Empirical power and FDR in multi-locus simulations in the presence of noise loci. We additionally conducted multi-locus simulations to compare KnockoffTrio with conventional FDR and FWER control methods in the presence of multiple causal and non-causal (noise) loci. We adopted the same simulation method in single-locus simulations to randomly generate 100 1Mb causal loci and 2,000 200kb non-causal loci. A causal locus contains a 1kb signal window, in which three variants were randomly selected to be causal.

We compared KnockoffTrio with M=10 to the Bonferroni correction that controls the FWER and the BH procedure that controls the FDR. Both the Bonferroni correction and the BH procedure were applied to the ACAT-combined p-values used to compute importance scores in KnockoffTrio. We also applied the Bonferroni correction to the weighted burden FBAT, a commonly used test in family-based studies. A method's power is the proportion of detected causal windows (i.e. windows that contain at least one causal variant) among all causal windows. We evaluated power at a target FDR of 0.1 for FDR-control methods or a target FWER of 0.05 for FWER-control methods. The empirical FDR is defined as the proportion of non-causal windows at least 50/25/0kb away from the nearest signal windows among all detected windows. As shown in Figure 5, KnockoffTrio was more powerful than the Bonferroni correction, as expected given the more liberal FDR control, while preserving the FDR at the target level of 0.1. The BH procedure failed to control the FDR at the target level due to the complex correlations among genetic variants. We also note that the FDR for each method decreased as the distance to the signal windows increased. This is expected because the non-causal windows closer to the signal windows are more likely to be false positives due to stronger linkage disequilibrium with variants in the signal windows. Such decrease in FDR is particularly evident for the BH procedure, which is more affected by the correlation among tests.

KnockoffTrio-iQRAT improves power in detecting complex associations. We performed simulations to compare the power of KnockoffTrio-iQRAT with KnockoffTrio-FBAT in complex scenarios where the normality of quantitative traits is violated. Specifically, we generated quantitative trait values using a location model:

$$Y_i = \beta_1 X_{i1} + ... + \beta_p X_{ip} + \epsilon_i,$$

where $\epsilon_i \sim$ Cauchy(μ =0, γ =1), μ is the location parameter and γ is the scale parameter for the Cauchy distribution. We generated 500 replicates, each of which consists of 1,000 trios and 500 variants near the *MACROD2* gene using the AGP cohort as above. We randomly selected three variants within a 1kb window to be causal with the causal effect $\beta_j = 1.2|\log_{10}\text{MAF}_j|$. We applied quantile and rank normalization to Y_i's before analysis. For KnockoffTrio-iQRAT, to make fair comparisons with FBAT, we only analyzed the offspring data, and adjusted the offspring genotypes by subtracting the conditional expectation (conditional on parental genotypes), i.e. $\boldsymbol{X}_i - E(\boldsymbol{X}_i|\boldsymbol{P}_i^1, \boldsymbol{P}_i^2)$. As shown in Figure S4 in the Appendix, KnockoffTrio-iQRAT is more powerful than KnockoffTrio-FBAT in the scenario with non-Gaussian errors as expected.

Applications to trio data on Autism Spectrum Disorders

We applied KnockoffTrio with multiple knockoffs (M=10) to several cohorts on autism spectrum disorders (ASD), including the family trio data from the Autism Genome Project (AGP) (dbGaP accession: phs000267.v5.p2)[29] and two cohorts collected by the Simons Foundation Autism Research Initiative (SFARI) to study the risk genetic variants for ASD, including the Simons Foundation Powering Autism Research (SPARK)[30] and the Simons Simplex Collection (SSC)[31]. The details of the individual cohorts are described below.

Data descriptions

AGP. Our AGP analysis included 798,961 common (MAF$\geq$0.05) and low-frequency (0.01$\leq$MAF$<$0.05) variants for 1,266 trio families of European ancestry, each of which consists of two parents and their offspring diagnosed with strict ASD, i.e., met the criteria for autism on both the ADI-R[32] and the ADOS[33].

SPARK. Our SPARK analysis included 10,540 trio families from the first three releases of the SPARK cohort. The probands in the two SFARI cohorts received a professional diagnosis of ASD from a physician, psychologist, or therapist. We have focused on 381,063 common and low-frequency variants.

SSC. Our SSC analysis included 2,394 trio families from the pilot and phases 1, 2, 3-1, and 3-2 studies of the SSC cohort, with whole-genome sequencing data available. We have focused on 5,772,421 common and low-frequency variants.

KnockoffTrio analyses

We adopted a quality control procedure that excluded variants with MAFs $< 1\%$, missing call rates $> 5\%$, Mendelian error rates $> 0.1\%$, and Hardy-Weinberg equilibrium p-values $< 10^{-7}$ for all cohorts. Genotype data were phased using SHAPEIT2[34]. The genomic coordinates in the AGP data were converted from hg18 to hg38 using the NCBI Genome Remapping Service. We adjusted for gender of offspring in all analyses. We present results from individual cohorts at a target FDR of 0.1 and 0.2 and compared them to the conventional association test with the Bonferroni correction and with the usual BH procedure for FDR control (Figures 6, 7, 8, and Table 1). We also present results from meta-analyses of the AGP and SPARK cohorts (Figure 9 and Table 2).

For the AGP cohort, the conventional association test (Bonferroni and BH) did not identify any significant association, whereas KnockoffTrio identified five significant regions including Neurexin 1 (*NRXN1*), Rho Guanine Nucleotide Exchange Factor 10 (*ARHGEF10*), Lamin Tail Domain Containing 1 (*LMNTD1*) - Ras Association Domain Family Member 8 (*RASSF8*), Alpha Kinase 3 (*ALPK3*) - Solute Carrier Family 28 Member 1 (*SLC28A1*), and Mono-ADP Ribosylhydrolase 2 (*MACROD2*) at FDR=0.1 (Figure 6). Among them, *MACROD2* and *NRXN1* have been reported in previous studies as risk genes associated with ASD[35,36,37,38]. *ARHGEF10* has been associated with impaired social interaction in mice[39], one of the main features of ASD. *SLC28A1* has a brain-biased expression and shows an excess of introgressed segments in EAS and EUR[40]. *SLC28A1* also belongs to the SLC (Solute Carrier) family, several members of which have previously been associated to behavioral traits (depression, mood disorders, and smoking behavior), and autism susceptibility and attention-deficit/hyperactivity disorder[40]. Furthermore, rs4842996, 8 kb upstream of *SLC28A1*, has been associated with ASD in a meta-analysis of GWAS findings from literature[41].

For the SPARK cohort, KnockoffTrio identified nine significant loci including Zinc Finger Protein 589 (*ZNF589*), Cell Adhesion Molecule 2 (*CADM2*), Chondroitin Sulfate Synthase 3 (*CHSY3*) - Histidine Triad Nucleotide Binding Protein 1 (*HINT1*), Platelet Derived Growth Factor Subunit A (*PDGFA*) - Protein Kinase CAMP-Dependent Type I Regulatory Subunit Beta (*PRKAR1B*), Dedicator Of Cytokinesis 4 (*DOCK4*), MT-RNR2 Like 6 (*MTRNR2L6*) - Serine Protease 1 (*PRSS1*), La Ribonucleoprotein 4B (*LARP4B*) - GTP Binding Protein 4 (*GTPBP4*), Isopentenyl-Diphosphate Delta Isomerase 2 (*IDI2*), and Protocadherin 20 (*PCDH20*) - Protocadherin 9 (*PCDH9*) at FDR=0.1, and additionally, Serine Peptidase Inhibitor Kazal Type 8 (*SPINK8*), Solute Carrier Family 22 Member 23 (*SLC22A23*)/Proteasome Assembly Chaperone 4 (*PSMG4*), BAG Cochaperone 4 (*BAG4*), and Cyclin B1 Interacting Protein 1 (*CCNB1IP1*) - Poly (ADP-Ribose) Polymerase 2 (*PARP2*) at FDR=0.2 (Figure 7). *PRKAR1B* has been implicated in several neurodevelopmental disorders including ASD[42,43,44,45]. Similarly, *CADM2* has been associated with ASD in multiple studies[46,47,48,49]. *PCDH9* has been implicated as a genetic risk factor for multiple psychiatric disorders, including major depression[50] and ASD[51]. It is a cell adhesion molecule involved in neuronal migration, synaptic plasticity, and circuit formation. Previous studies have shown that homozygous knockout *PCDH9*-deficient mice have deficits in specific long-term social and object recognition[52]. *DOCK4* has been associated with ASD[53,54]. Furthermore, *DOCK4* knockout mice displayed a series of ASD-like behaviors, including impaired social novelty preference, abnormal isolation-induced pup vocalizations, elevated anxiety, and perturbed object and spatial learning[55]. *BAG4* resides at a locus that has been genome-wide significant in a combined ASD-schizophrenia GWAS[56]. A deleterious variant c.956T>A, p.(Leu319His) in *ZNF589* segregated with the phenotype (intellectual disability) and was identified as homozygous in two affected siblings in a consanguineous family from Northern Pakistan[57]; the variant was absent from 200 ethnically matched control individuals. *HINT1* regulates the function of PKC (protein kinase C) which is a prime gene to regulate regression in autism[58,59]. *SPINK8* resides at a GWAS significant locus associated with multiple psychiatric disorders[60]. In comparison, the conventional association test (BH) identified five loci *PDGFA-PRKAR1B*, *DOCK4*, *LARP4B-GTPBP4*, *IDI2*, and *PCDH20-PCDH9*, at FDR=0.1, and three loci *ZNF589*, *CHSY3-HINT1*, and *MTRNR2L6-PRSS1* at FDR=0.2, all of which have been identified by KnockoffTrio as well.

For the SSC cohort, KnockoffTrio identified Potassium Channel Regulator (*KCNRG*) - Deleted In Lymphocytic Leukemia 7 (*DLEU7*) at FDR=0.1, and additionally, Potassium Voltage-Gated Channel Interacting Protein 4 (*KCNIP4*) at FDR=0.2 (Figure 8). The finding of *KCNRG*, a gene in the KCTD (potassium channel tetramerization domain) family, provides further evidence for the role of KCTD family in neurodevelopmental and neuropsychiatric disorders[61]. *KCNIP4* is a gene with the largest number of differential RNA editing sites that have been suggested for aberrant synaptic formation in ASD[62]; *KCNIP4* has also been associated with nonverbal communication and social skills in ASD twins[63]. In comparison, the conventional association tests identified no significant loci.

Meta-analyses. We conducted meta-analyses of the AGP and SPARK cohorts with and without the SSC cohort because the signals in the SSC cohort are particularly weak. When the SSC cohort was excluded, KnockoffTrio identified five significant loci *CHSY3-HINT1, PDGFA-PRKAR1B, DOCK4, LARP4B-GTPBP4*, and *IDI2* at FDR=0.1, and additionally, *ARHGEF10* and Coiled-Coil Domain Containing 89 (*CCDC89*) / Synaptotagmin Like 2 (*SYTL2*) at FDR=0.25 (Figure 9). Among them, *CCDC89/SYTL2* has not been identified in analyses of individual cohorts. In comparison, the conventional association test (BH) identified no significant association at FDR=0.1 and two significant loci *PDGFA-PRKAR1B* and *DOCK4* at FDR=0.25. We show the results for meta-analysis of all three cohorts in Figure S3 in the Appendix.

Replicability of analyses. Given the random nature of the knockoff procedure, we have attempted to assess the replicability of the results by re-analyzing the individual cohorts and the meta-analyses with different random seeds for the knockoff generation. As shown in Figures S5, S6, S7, and S8 in the Appendix, the replications produced results that are in good concordance with the original results. For the AGP cohort, the replication analysis identified *NRXN1*, *ARHGEF10*, *LMNTD1-RASSF8*, and *MACROD2*, all of which have been identified in the original analysis. For the SPARK cohort, the replication analysis identified *ZNF589*, *SPINK8*, *CHSY3-HINT1*, *PDGFA-PRKAR1B*, *DOCK4*, *MTRNR2L6-PRSS1*, *LARP4B-GTPBP4*, *IDI2*, *PCDH20-PCDH9*, all of which have been identified in the original analysis. For the SSC cohort, the replication analysis identified RANBP2 Like And GRIP Domain Containing 2 (*RGPD2*), *KCNIP4*, and *KCNRG-DLEU7*, the latter two of which have been identified in the original analysis. For the meta-analysis of the AGP and SPARK cohorts, the replication analysis identified *CHSY3-HINT1, PDGFA-PRKAR1B, DOCK4, LARP4B-GTPBP4*, and *IDI2*, all of which have been identified in the original analysis. This shows the replicability of results from KnockoffTrio despite the randomness in knockoff generation.

Discussion

We propose KnockoffTrio, a novel association test with trio design for GWAS data built upon the knockoff framework. As an FDR-controlling procedure that accounts for arbitrary correlation structure, KnockoffTrio has been shown in both simulations and real-data analysis to be more powerful than the conventional FWER-controlling methods while possessing better FDR control than the conventional FDR-controlling methods such as Benjamini-Hochberg. Because it is built conditional on parental genotypes, KnockoffTrio is also by construction robust against bias induced by population substructure. Furthermore, an important advantage of KnockoffTrio is that it can leverage more sophisticated machine learning models to model the association between genotypes and phenotypes while maintaining valid FDR control and with potential increases in power. These properties make KnockoffTrio an appealing and promising strategy for the analysis of trio designs for which conventional methods are known to be underpowered.

Although we have focused the current manuscript on the trio design, the method can be easily extended to handle larger pedigrees by breaking each pedigree into all possible trios, and applying KnockoffTrio on the individual trios. The method can also be extended to combine trios and population-based designs. For example, we can obtain the estimated coefficient $\hat{\beta}_j$ for variant j from the external population-based GWAS and use it as weight w_j in the weighted FBAT when constructing the importance scores. Alternatively we can perform knockoff analysis for population-based data as in He et al.[18] and use a meta-analysis approach as discussed in the Methods section to combine the trio and population-based results. Note that this alternative approach is no longer robust to confounding due to population structure. Transfer learning methods that leverage information from such external population-based data could also be of interest[64].

KnockoffTrio has been implemented in a computationally efficient R package. The runtime for completing the analyses of the AGP, SPARK, and SSC cohorts with 10 knockoffs is 8 minutes, 46 minutes, and 173 minutes, respectively, with 1,000 parallel jobs performed in a high-performance computing cluster environment of Intel(R) Xeon(R) CPU E5-2630 0 @ 2.30GHz. This demonstrates that KnockoffTrio is a highly scalable method and can be effectively used for any large-scale datasets in whole-genome sequencing studies.

KnockoffTrio reduces the randomness in the knockoff generation by using a multiple-knockoff generation procedure. As shown in the simulations and real-data applications, KnockoffTrio with 10 knockoffs is more powerful than using a single knockoff and has good replicability in terms of identifying significant loci. Although the marginal gain in power appears to diminish as the number of knockoffs increases, given the computational efficiency of KnockoffTrio, we suggest researchers to use at least 10 knockoffs for better power and reproducibility.

Although KnockoffTrio assumes no recombination events given a 200kb region to simplify the inference about the transmission pattern, KnockoffTrio can be extended to handle recombination events at the cost of more complex construction of offspring knockoffs, which may potentially help further improve the power in real-data applications. In addition to the haplotype-based knockoff generation algorithm that KnockoffTrio adopts, another possible approach is to use summary statistics and apply knockoff-based methods for summary statistics directly instead of generating knockoffs for individual trio data. We leave these potential extensions to future studies.

GWAS with family-based designs are appealing due to built-in robustness to population substructure, but are underpowered due to limited sample sizes, much smaller than for GWAS studies with unrelated individuals. KnockoffTrio provides a more powerful alternative to classical family-based association tests in this setting. Furthermore, by design KnockoffTrio reduces the confounding effect of LD and prioritizes causal loci over associations due to LD. KnockoffTrio has been implemented in a computationally efficient R package.

Figure 1: KnockoffTrio workflow. Knockoff generation based on original trios, calculation of importance scores using sliding windows, and examples of hypothesis testing using conventional association testing and KnockoffTrio.

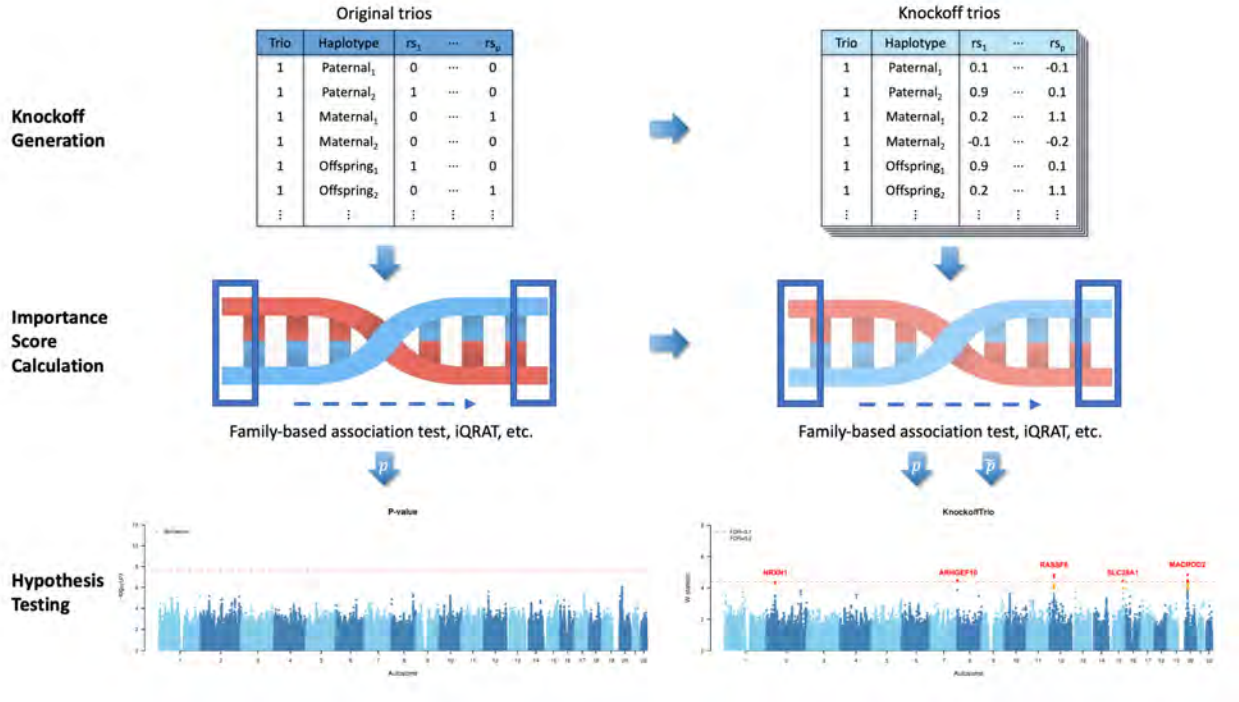


Figure 2: Empirical validation for exchangeability property in KnockoffTrio. To validate the exchangeability, we generated offspring knockoff genotypes (Xk) using the proposed algorithm and evaluated whether the second order (covariance between each pair of genetic variants) is exchangeable for common variants in the region. "Cov.X_X" is the covariance between each pair of original variants, "cov.Xk_Xk" is the covariance between each pair of knockoff variants, and "cov.X_Xk" is the covariance between each pair of original and knockoff variants.

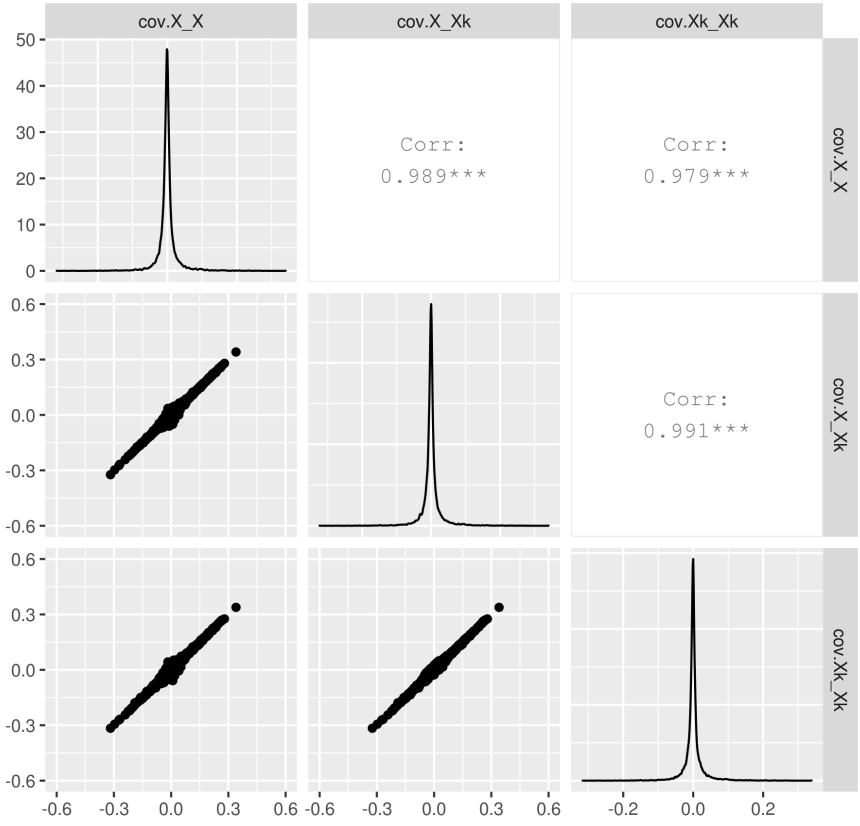


Figure 3: KnockoffTrio's power and FDR in single-locus simulations. The two panels show the power and FDR for dichotomous and quantitative traits. We evaluate KnockoffTrio's power and FDR with a target FDR ranging from 0 to 0.2 and with different numbers of knockoffs. The solid lines indicate KnockoffTrio's power and the dotted lines indicate KnockoffTrio's observed FDR. The different colors indicate different numbers of knockoffs. The grey dashed line indicates the expected FDR.

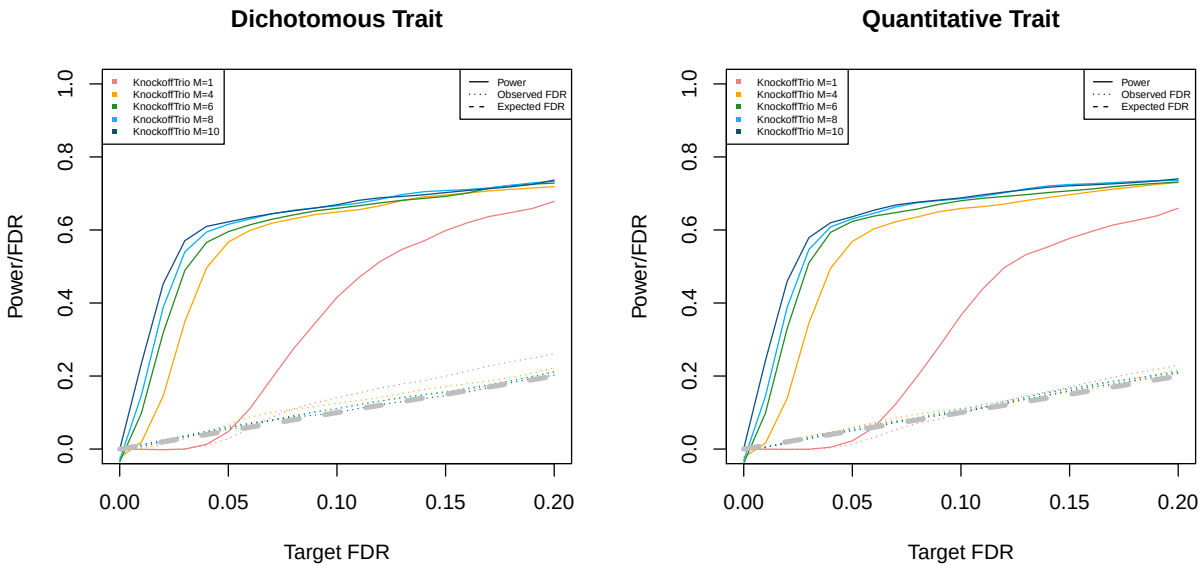


Figure 4: KnockoffTrio prioritizes causal variants over associations due to linkage disequilibrium. Comparisons between KnockoffTrio and the conventional method in terms of **A**. the proportion of selected windows that overlap with the signal window, **B**. the median distance of selected windows to the signal window, and **C**. the robustness in controlling false positives in the presence of linkage disequilibrium between causal and non-causal variants. The conventional method is the same aggregated Cauchy association test implemented in KnockoffTrio and controls the FWER using the Bonferroni correction. For KnockoffTrio, the target FDR is 0.1 and the number of multiple knockoffs is 10. The $|r|$ in panel **C** is the maximum absolute correlation between the false positive and any causal variants. The variants in the right figure in panel **C** correspond to the variants in the left figure.

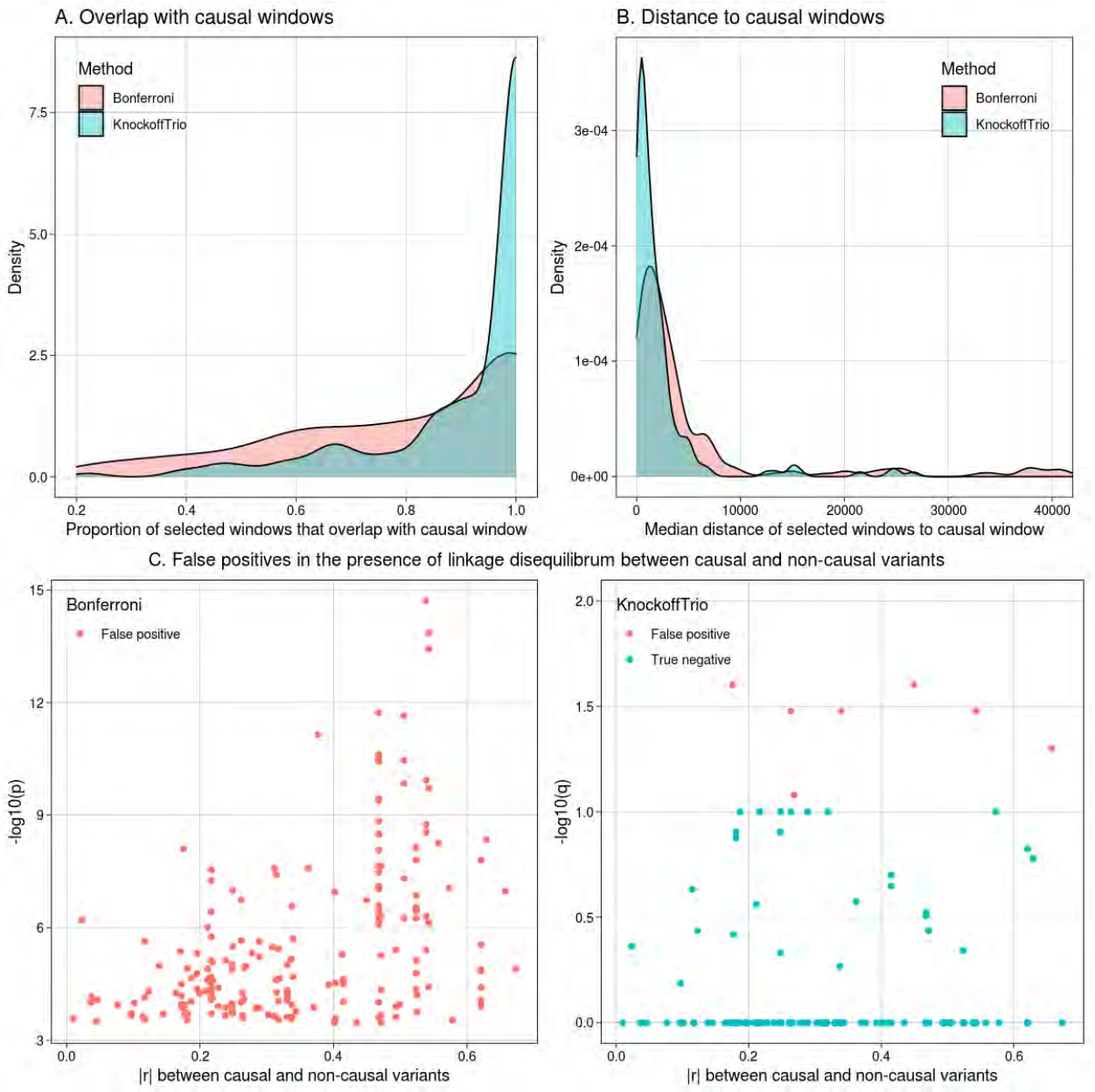


Figure 5: Genome-wide power and FDR in the presence of noise loci. The left panel presents each method's power (target FDR 0.1), defined as the proportion of detected causal windows among all causal windows. A causal window is a window that contains any causal variants. The right panel presents each method's false discovery rate (target FDR 0.1) at different resolutions, defined as the proportion of non-causal windows at least 50/25/0kb away from the nearest signal windows among all detected windows.

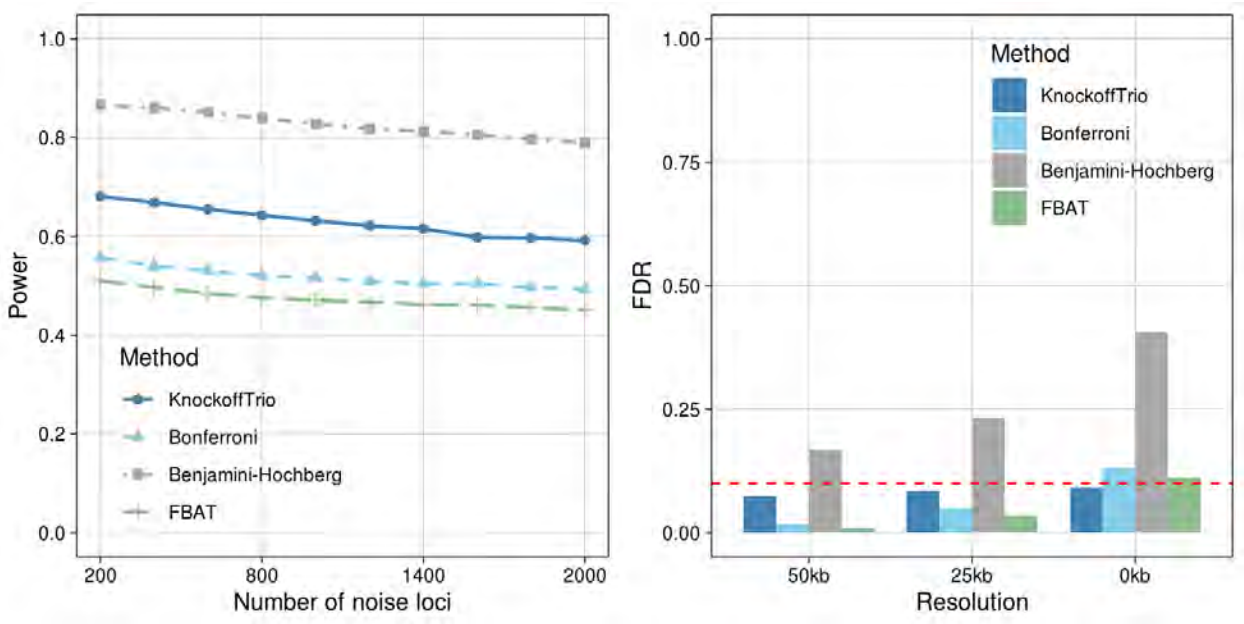


Figure 6: Manhattan plots from KnockoffTrio analysis for the Autism Genome Project (AGP). The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.2. Each locus is annotated with the closest gene name.

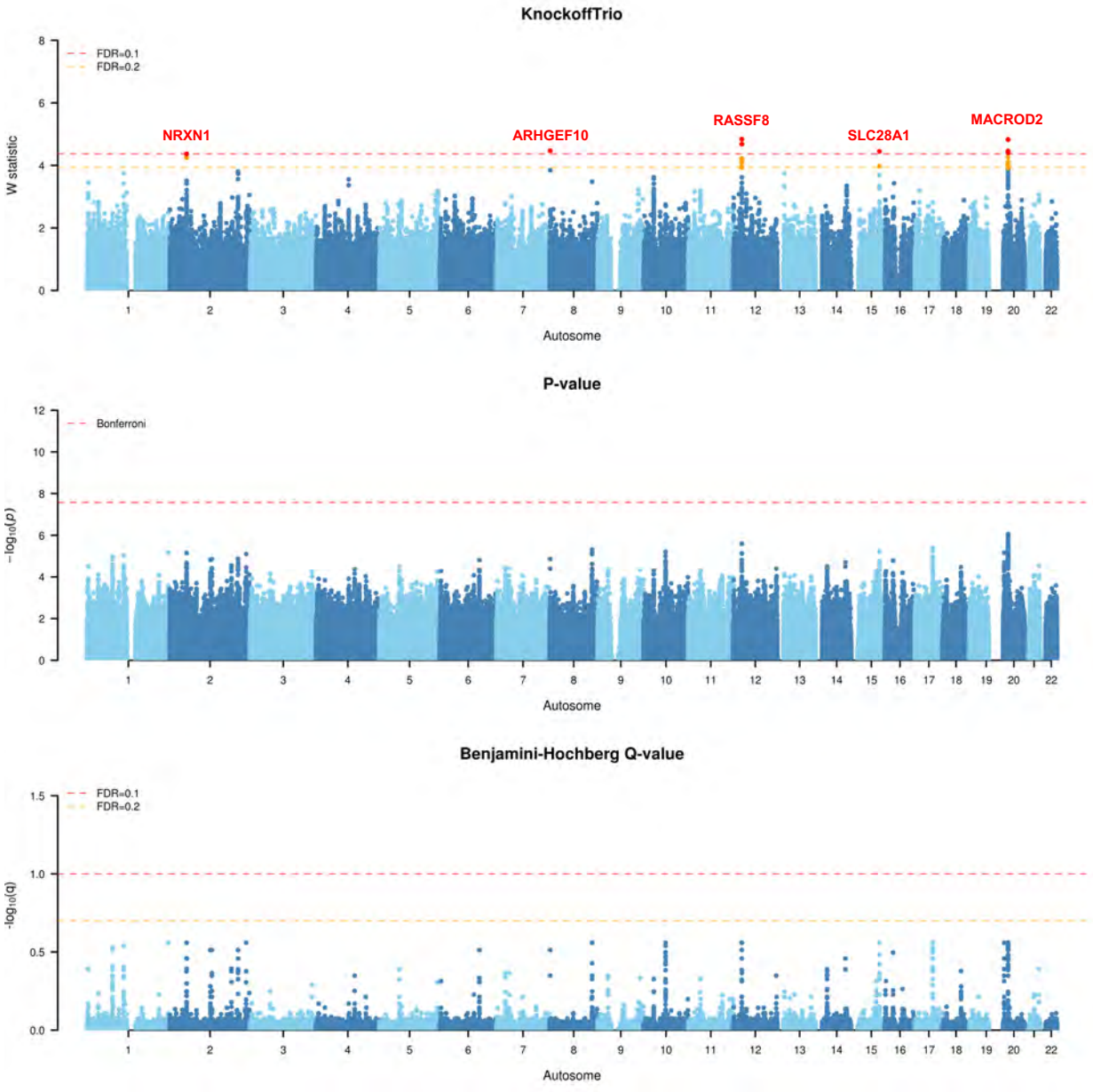


Figure 7: Manhattan plots from KnockoffTrio analysis for the Simons Foundation Powering Autism Research (SPARK). The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.2. Each locus is annotated with the closest gene name.

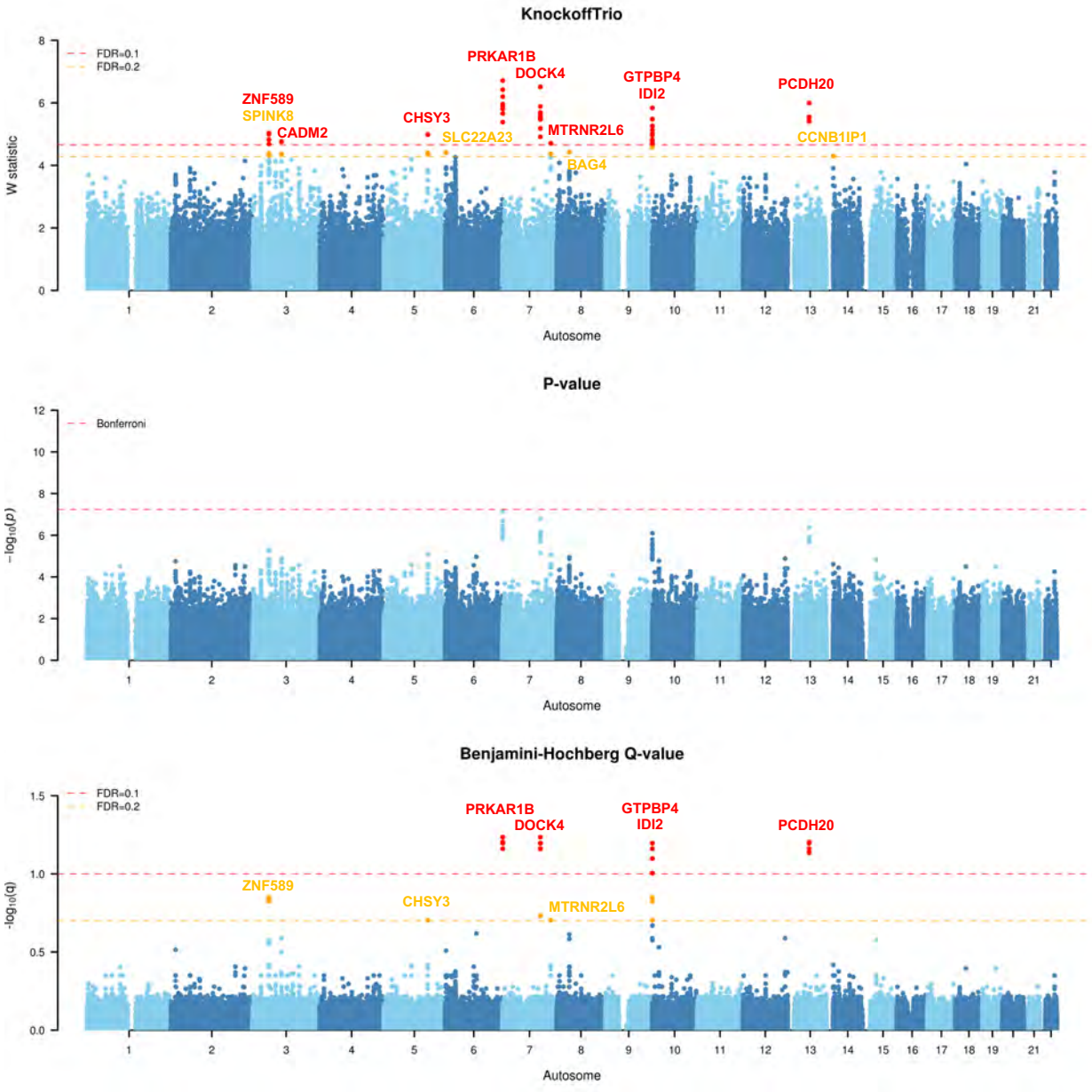


Figure 8: Manhattan plots from KnockoffTrio analysis for the Simons Simplex Collection (SSC). The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.2. Each locus is annotated with the closest gene name.

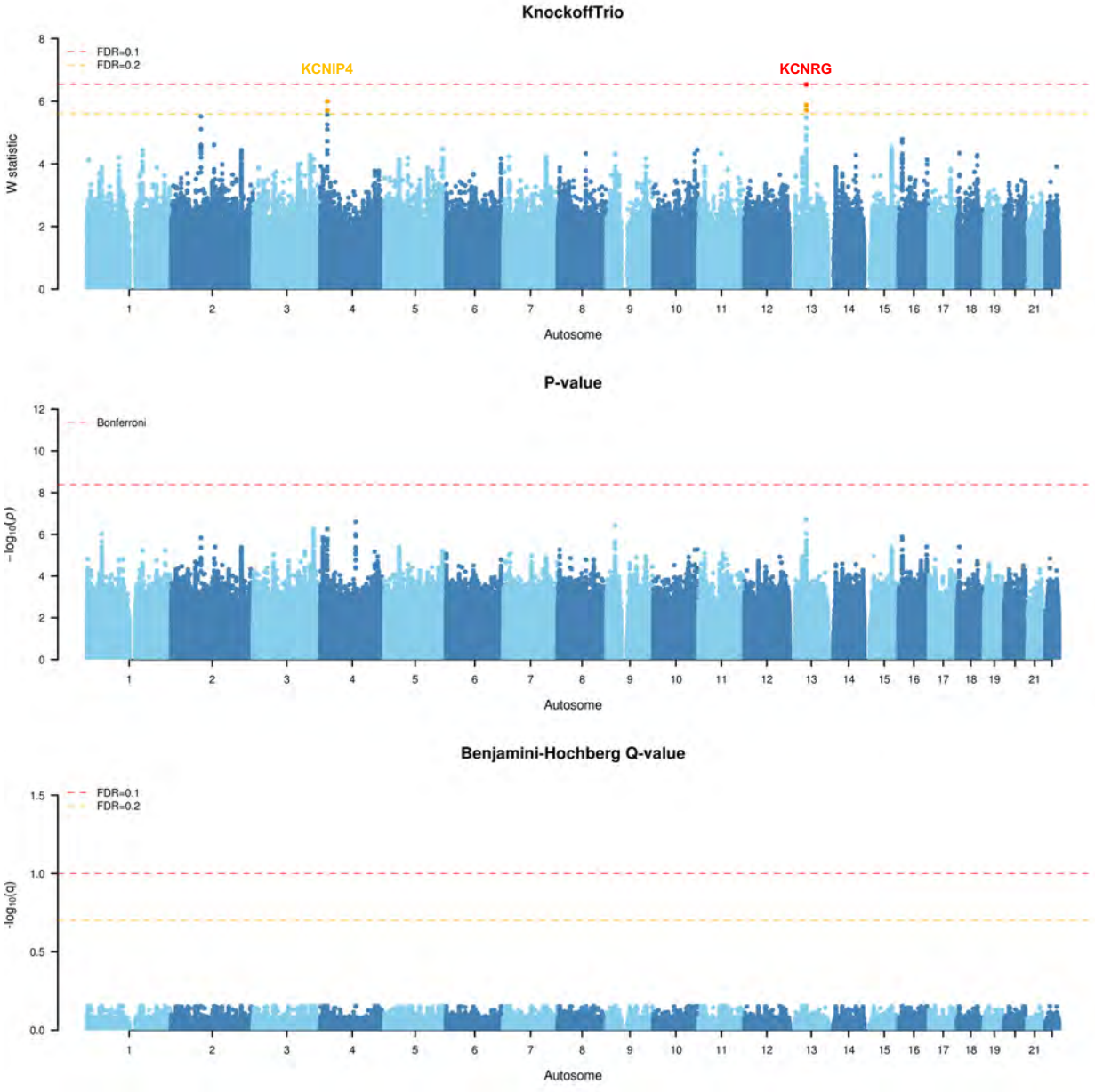


Figure 9: Manhattan plots from KnockoffTrio analysis for the meta-analysis of the AGP and SPARK cohorts. The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.25. Each locus is annotated with the closest gene name.

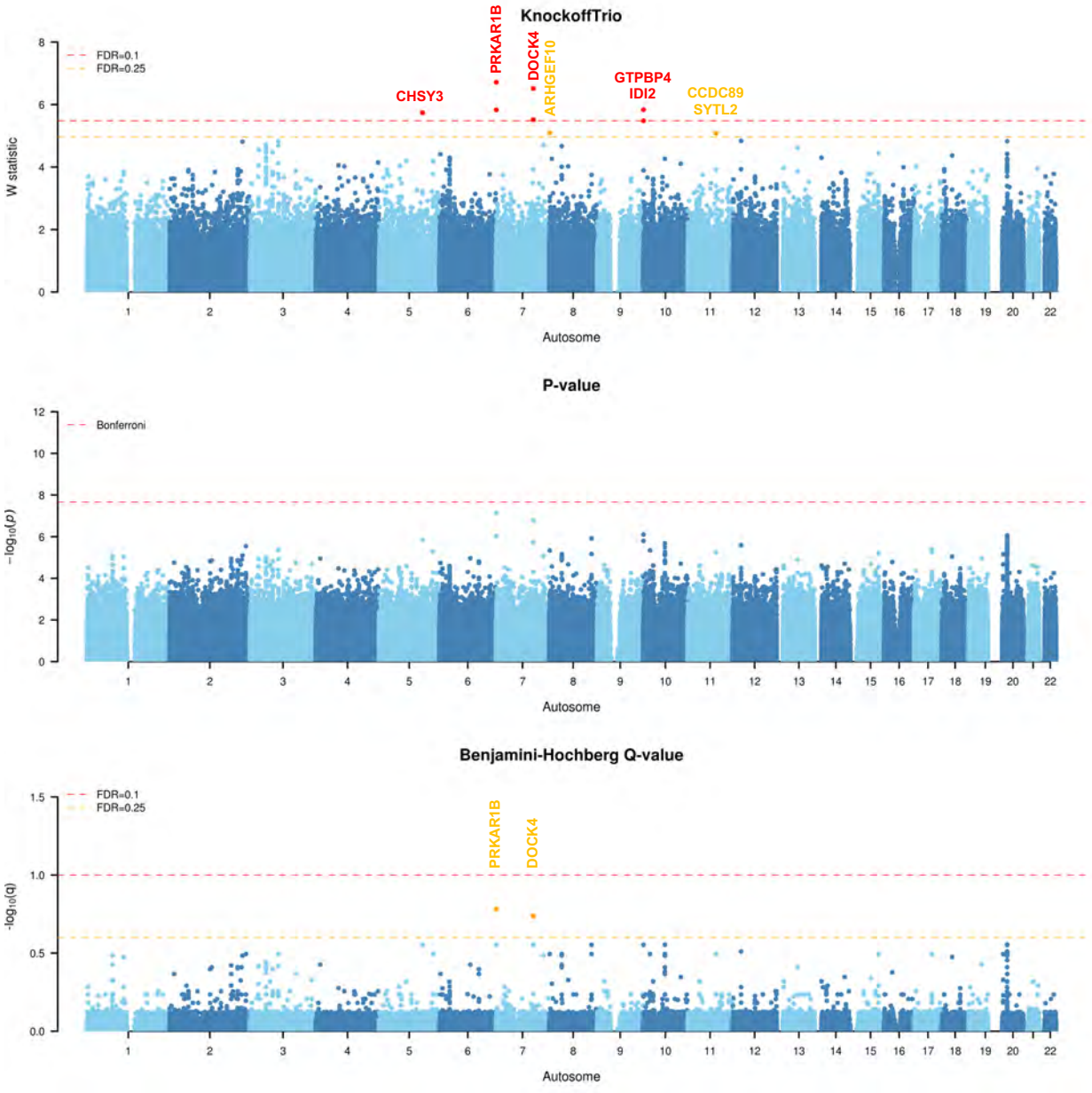


Table 1: Genome-wide significant loci from KnockoffTrio analysis.

Gene	Chr	Position	Variant	Allele	MAF	P	Z	W	Q	BH Q
AGP (FDR=0.1)										
NRXN1	2	50805721	rs9284756	A	0.03	7.10E-6	4.49	4.37	0.10	0.28
ARHGEF10	8	1920247-1920676	rs17756915-rs11136442	-	0.41	1.38E-5	-	4.47	0.10	0.31
LMNTD1-RASSF8	12	25946268	rs4963941	A	0.10	2.56E-6	4.70	4.84	0.10	0.28
ALPK3-SLC28A1	15	84881866	rs12917429	T	0.21	6.19E-6	-4.52	4.45	0.10	0.28
MACROD2	20	14781064	rs6074798	A	0.49	1.02E-6	4.89	4.83	0.10	0.28
SFARI: SPARK (FDR=0.1)										
ZNF589	3	48262179	rs11709691	G	0.28	4.87E-6	-4.57	5.03	0.06	0.14
CADM2	3	85395534-85410981	rs75005531-rs1549979	-	0.22	1.30E-5	-	4.76	0.09	0.26
CHSY3-HINT1	5	130661503	rs17714209	C	0.28	8.25E-6	4.46	4.99	0.06	0.20
PDGFA-PRKAR1B	7	536383	rs62431385	C	0.10	7.20E-8	-5.39	6.71	0.02	0.06
DOCK4	7	111986531	rs73210911	A	0.12	1.59E-7	-5.24	6.51	0.02	0.06
MTRNR2L6-PRSS1	7	142688332	rs13223009	C	0.02	8.42E-6	-4.45	4.71	0.09	0.20
LARP4B-GTPBP4	10	975370	rs117732138	A	0.02	1.60E-6	4.80	5.48	0.02	0.07
IDI2	10	1020654	rs77782977	C	0.02	7.95E-7	4.94	5.84	0.02	0.06
PCDH20-PCDH9	13	63204555	rs12184522	T	0.23	4.21E-7	5.06	6.00	0.02	0.06
SFARI: SPARK (FDR=0.2)										
SPINK8	3	48316110-48329279	rs74735576-rs13090538	-	0.17	1.58E-5	-	4.39	0.17	0.28
SLC22A23/PSMG4	6	3285062	rs41301847	G	0.02	1.85E-5	4.28	4.41	0.17	0.31
BAG4	8	38205717	rs7836805	A	0.24	2.83E-5	-4.19	4.43	0.17	0.40
CCNB1IP1-PARP2	14	20334133	rs72671266	T	0.02	2.45E-5	-4.22	4.30	0.19	0.38
SFARI: SSC (FDR=0.1)										
KCNRG-DLEU7	13	50197099	rs2703087	A	0.04	1.88E-7	5.21	6.54	0.10	0.70
SFARI: SSC (FDR=0.2)										
KCNIP4	4	20917151	rs185413018	T	0.02	5.59E-7	5.00	6.00	0.13	0.70

Only the top signal is shown if multiple signals were identified for a locus. **Gene**: A single gene name indicates the signal is within or overlaps with the gene. "Gene1/Gene2" indicates the signal overlaps with two genes. "Gene1-Gene2" indicates the signal is between two genes. **MAF**: minor allele frequency of a variant, or average minor allele frequency if a signal contains multiple variants. **P**: KnockoffTrio's ACAT-combined p-values. For single variants, ACAT-combined p-values are equivalent to FBAT p-values. **Z**: FBAT Z-scores for single variants. **W**: KnockoffTrio's feature statistics. **Q**: KnockoffTrio's Q-values. **BH Q**: Benjamini-Hochberg Q-values.

Table 2: Genome-wide significant loci from KnockoffTrio meta-analysis (AGP and SPARK).

					Meta-Analysis				AGP				SPARK			
Gene	Chr	Position	Allele	MAF	P	Z	W	Q	P	Z	W	Q	P	Z	W	Q
FDR=0.1																
CHSY3-HINT1	5	130661503	C	0.33	1.45E-6	4.82	5.74	0.10	6.48E-2	1.85	0	1	8.25E-6	4.46	4.99	0.06
PDGFA-PRKAR1B	7	536383	C	0.10	7.20E-8	-5.39	6.71	0.10	-	-	-	-	7.20E-8	-5.39	6.71	0.02
DOCK4	7	111986531	A	0.12	1.59E-7	-5.24	6.51	0.10	-	-	-	-	1.59E-7	-5.24	6.51	0.02
LARP4B-GTPBP4	10	975370	A	0.02	1.60E-6	4.80	5.48	0.10	-	-	-	-	1.60E-6	4.80	5.48	0.02
IDI2	10	1020654	C	0.02	7.95E-7	4.94	5.84	0.10	-	-	-	-	7.95E-7	4.94	5.84	0.02
FDR=0.25																
ARHGEF10	8	1920247-1920676	-	0.41	4.62E-6	-4.58	5.10	0.21	1.38E-5	-	4.47	0.10	-	-	-	-
CCDC89/SYTL2	11	85684675-85727667	-	0.11	5.73E-6	4.54	5.08	0.21	1.69E-1	-	0.70	1	2.61E-4	-	3.47	0.49

Only the top signal is shown if multiple signals were identified for a locus. **Gene**: A single gene name indicates the signal is within or overlaps with the gene. "Gene1/Gene2" indicates the signal overlaps with two genes. "Gene1-Gene2" indicates the signal is between two genes. **MAF**: weighted minor allele frequency of a variant, or weighted average minor allele frequency if a signal contains multiple variants. A cohort's weight is the number of trios in the cohort divided by the total number of trios in the meta-analysis. **P**: KnockoffTrio's ACAT-combined p-values. For single variants, ACAT-combined p-values are equivalent to FBAT p-values. **Z**: FBAT z-scores for single variants. **W**: KnockoffTrio's feature statistics. **Q**: KnockoffTrio's Q-values.

References

[1] B. Al-Mubarak, M. Abouelhoda, A. Omar, H. AlDhalaan, M. Aldosari, M. Nester, H. A. Alshamrani, M. El-Kalioby, E. Goljan, R. Albar, et al., "Whole exome sequencing reveals inherited and de novo variants in autism spectrum disorder: a trio study from saudi families," Scientific reports, vol. 7, no. 1, pp. 1–14, 2017.

[2] T. H. Wassink, J. Piven, V. J. Vieland, J. Huang, R. E. Swiderski, J. Pietila, T. Braun, G. Beck, S. E. Folstein, J. L. Haines, et al., "Evidence supporting wnt2 as an autism susceptibility gene," American journal of medical genetics, vol. 105, no. 5, pp. 406–413, 2001.

[3] B. J. O'Roak, P. Deriziotis, C. Lee, L. Vives, J. J. Schwartz, S. Girirajan, E. Karakoc, A. P. MacKenzie, S. B. Ng, C. Baker, et al., "Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations," Nature genetics, vol. 43, no. 6, p. 585, 2011.

[4] N. M. Laird and C. Lange, "The role of family-based designs in genome-wide association studies," Statistical Science, vol. 24, no. 4, pp. 388–397, 2009.

[5] N. M. Laird and C. Lange, "Family-based designs in the age of large-scale gene-association studies," Nat Rev Genet, vol. 7, pp. 385–94, May 2006.

[6] A. Kong, G. Thorleifsson, M. L. Frigge, B. J. Vilhjalmsson, A. I. Young, T. E. Thorgeirsson, S. Benonisdottir, A. Oddsson, B. V. Halldorsson, G. Masson, D. F. Gudbjartsson, A. Helgason, G. Bjornsdottir, U. Thorsteinsdottir, and K. Stefansson, "The nature of nurture: Effects of parental genotypes," Science, vol. 359, pp. 424–428, Jan 2018.

[7] A. L. Price, N. A. Zaitlen, D. Reich, and N. Patterson, "New approaches to population stratification in genome-wide association studies," Nat Rev Genet, vol. 11, pp. 459–63, Jul 2010.

[8] H. Chen, J. E. Huffman, J. A. Brody, C. Wang, S. Lee, Z. Li, S. M. Gogarten, T. Sofer, L. F. Bielak, J. C. Bis, et al., "Efficient variant set mixed model association tests for continuous and binary traits in large-scale whole-genome sequencing studies," The American Journal of Human Genetics, vol. 104, no. 2, pp. 260–274, 2019.

[9] W. Zhou, Z. Zhao, J. B. Nielsen, L. G. Fritsche, J. LeFaive, S. A. G. Taliun, W. Bi, M. E. Gabrielsen, M. J. Daly, B. M. Neale, et al., "Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts," Nature genetics, vol. 52, no. 6, pp. 634–639, 2020.

[10] S. Bates, M. Sesia, C. Sabatti, and E. Candès, "Causal inference in genetic trio studies," Proc Natl Acad Sci U S A, vol. 117, pp. 24117–24126, 09 2020.

[11] C. P. Nelson, A. Goel, A. S. Butterworth, S. Kanoni, T. R. Webb, E. Marouli, L. Zeng, I. Ntalla, F. Y. Lai, J. C. Hopewell, et al., "Association analyses based on false discovery rate implicate new loci for coronary artery disease," Nature genetics, vol. 49, no. 9, p. 1385, 2017.

[12] M. Sesia, S. Bates, E. Candès, J. Marchini, and C. Sabatti, "Controlling the false discovery rate in gwas with population structure," bioRxiv, 2020.

[13] F. K. Satterstrom, J. A. Kosmicki, J. Wang, M. S. Breen, S. De Rubeis, J.-Y. An, M. Peng, R. Collins, J. Grove, L. Klei, C. Stevens, J. Reichert, M. S. Mulhern, M. Artomov, S. Gerges, B. Sheppard, X. Xu, A. Bhaduri, U. Norman, H. Brand, G. Schwartz, R. Nguyen, E. E. Guerrero, C. Dias, Autism Sequencing Consortium, iPSYCH-Broad Consortium, C. Betancur, E. H. Cook, L. Gallagher, M. Gill, J. S. Sutcliffe, A. Thurm, M. E. Zwick, A. D. Børglum, M. W. State, A. E. Cicek, M. E. Talkowski, D. J. Cutler, B. Devlin, S. J. Sanders, K. Roeder, M. J. Daly, and J. D. Buxbaum, "Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism," Cell, vol. 180, pp. 568–584.e23, 02 2020.

[14] S. De Rubeis, X. He, A. P. Goldberg, C. S. Poultney, K. Samocha, A. E. Cicek, Y. Kou, L. Liu, M. Fromer, S. Walker, T. Singh, L. Klei, J. Kosmicki, F. Shih-Chen, B. Aleksic, M. Biscaldi, P. F. Bolton, J. M. Brownfeld, J. Cai, N. G. Campbell, A. Carracedo, M. H. Chahrour, A. G. Chiocchetti, H. Coon, E. L. Crawford, S. R. Curran, G. Dawson, E. Duketis, B. A. Fernandez, L. Gallagher, E. Geller, S. J. Guter, R. S. Hill, J. Ionita-Laza, P. Jimenz Gonzalez, H. Kilpinen, S. M. Klauck, A. Kolevzon, I. Lee, I. Lei, J. Lei, T. Lehtimäki, C.-F. Lin, A. Ma'ayan, C. R. Marshall, A. L. McInnes, B. Neale, M. J. Owen, N. Ozaki, M. Parellada, J. R. Parr, S. Purcell, K. Puura, D. Rajagopalan, K. Rehnström, A. Reichenberg, A. Sabo, M. Sachse, S. J. Sanders, C. Schafer, M. Schulte-Rüther, D. Skuse, C. Stevens, P. Szatmari, K. Tammimies, O. Valladares, A. Voran, W. Li-San, L. A. Weiss, A. J. Willsey, T. W. Yu, R. K. C. Yuen, DDD Study, Homozygosity Mapping Collaborative for Autism, UK10K Consortium, E. H. Cook, C. M. Freitag, M. Gill, C. M. Hultman, T. Lehner, A. Palotie, G. D. Schellenberg, P. Sklar, M. W. State, J. S. Sutcliffe, C. A. Walsh, S. W. Scherer, M. E. Zwick, J. C. Barett, D. J. Cutler, K. Roeder, B. Devlin, M. J. Daly, and J. D. Buxbaum, "Synaptic, transcriptional and chromatin genes disrupted in autism," Nature, vol. 515, pp. 209–15, Nov 2014.

[15] E. Candes, Y. Fan, L. Janson, and J. Lv, "Panning for gold:'model-x'knockoffs for high dimensional controlled variable selection," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 80, no. 3, pp. 551–577, 2018.

[16] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," Annals of statistics, pp. 1165–1188, 2001.

[17] M. Sesia, E. Katsevich, S. Bates, E. Candès, and C. Sabatti, "Multi-resolution localization of causal variants across the genome," Nat Commun, vol. 11, p. 1093, 02 2020.

[18] Z. He, L. Liu, C. Wang, Y. Le Guen, J. Lee, S. Gogarten, F. Lu, S. Montgomery, H. Tang, E. K. Silverman, et al., "Identification of putative causal loci in whole-genome sequencing data via knockoff statistics," Nature communications, vol. 12, no. 1, pp. 1–18, 2021.

[19] E. Candes, Y. Fan, L. Janson, J. Lv, et al., "Panning for gold: Model-x knockoffs for high-dimensional controlled variable selection," Journal of the Royal Statistical Society Series B, vol. 80, no. 3, pp. 551–577, 2018.

[20] R. S. Spielman, R. E. McGinnis, and W. J. Ewens, "Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (iddm)," Am J Hum Genet, vol. 52, pp. 506–16, Mar 1993.

[21] H. Chen, J. B. Meigs, and J. Dupuis, "Sequence kernel association test for quantitative traits in family samples," Genetic epidemiology, vol. 37, no. 2, pp. 196–204, 2013.

[22] Q. Yan, H. K. Tiwari, N. Yi, G. Gao, K. Zhang, W.-Y. Lin, X.-Y. Lou, X. Cui, and N. Liu, "A sequence kernel association test for dichotomous traits in family samples under a generalized linear mixed model," Human heredity, vol. 79, no. 2, pp. 60–68, 2015.

[23] J. Marchini, D. Cutler, N. Patterson, M. Stephens, E. Eskin, E. Halperin, S. Lin, Z. S. Qin, H. M. Munro, G. R. Abecasis, P. Donnelly, and International HapMap Consortium, "A comparison of phasing algorithms for trios and unrelated individuals," Am J Hum Genet, vol. 78, pp. 437–50, Mar 2006.

[24] G. De, W.-K. Yip, I. Ionita-Laza, and N. Laird, "Rare variant analysis for family-based design," PloS one, vol. 8, no. 1, p. e48495, 2013.

[25] Y. Liu, S. Chen, Z. Li, A. C. Morrison, E. Boerwinkle, and X. Lin, "Acat: A fast and powerful p value combination method for rare-variant analysis in sequencing studies," Am J Hum Genet, vol. 104, pp. 410–421, 03 2019.

[26] T. Wang, I. Ionita-Laza, and Y. Wei, "Integrated quantile rank test (iqrat) for gene-level associations," arXiv preprint arXiv:1910.10102, 2019.

[27] C. J. Willer, Y. Li, and G. R. Abecasis, "Metal: fast and efficient meta-analysis of genomewide association scans," Bioinformatics, vol. 26, pp. 2190–1, Sep 2010.

[28] M. Sesia, C. Sabatti, and E. J. Candès, "Gene hunting with hidden markov model knockoffs," Biometrika, vol. 106, pp. 1–18, Mar 2019.

[29] Autism Genome Project Consortium, P. Szatmari, A. D. Paterson, L. Zwaigenbaum, W. Roberts, J. Brian, X.-Q. Liu, J. B. Vincent, J. L. Skaug, A. P. Thompson, L. Senman, L. Feuk, C. Qian, S. E. Bryson, M. B. Jones, C. R. Marshall, S. W. Scherer, V. J. Vieland, C. Bartlett, L. V. Mangin, R. Goedken, A. Segre, M. A. Pericak-Vance, M. L. Cuccaro, J. R. Gilbert, H. H. Wright, R. K. Abramson, C. Betancur, T. Bourgeron, C. Gillberg, M. Leboyer, J. D. Buxbaum, K. L. Davis, E. Hollander, J. M. Silverman, J. Hallmayer, L. Lotspeich, J. S. Sutcliffe, J. L. Haines, S. E. Folstein, J. Piven, T. H. Wassink, V. Sheffield, D. H. Geschwind, M. Bucan, W. T. Brown, R. M. Cantor, J. N. Constantino, T. C. Gilliam, M. Herbert, C. Lajonchere, D. H. Ledbetter, C. Lese-Martin, J. Miller, S. Nelson, C. A. Samango-Sprouse, S. Spence, M. State, R. E. Tanzi, H. Coon, G. Dawson, B. Devlin, A. Estes, P. Flodman, L. Klei, W. M. McMahon, N. Minshew, J. Munson, E. Korvatska, P. M. Rodier, G. D. Schellenberg, M. Smith, M. A. Spence, C. Stodgell, P. G. Tepper, E. M. Wijsman, C.-E. Yu, B. Rogé, C. Mantoulan, K. Wittemeyer, A. Poustka, B. Felder, S. M. Klauck, C. Schuster, F. Poustka, S. Bölte, S. Feineis-Matthews, E. Herbrecht, G. Schmötzer, J. Tsiantis, K. Papanikolaou, E. Maestrini, E. Bacchelli, F. Blasi, S. Carone, C. Toma, H. Van Engeland, M. de Jonge, C. Kemner, F. Koop, F. Koop, M. Langemeijer, M. Langemeijer, C. Hijmans, C. Hijimans, W. G. Staal, G. Baird, P. F. Bolton, M. L. Rutter, E. Weisblatt, J. Green, C. Aldred, J.-A. Wilkinson, A. Pickles, A. Le Couteur, T. Berney, H. McConachie, A. J. Bailey, K. Francis, G. Honeyman, A. Hutchinson, J. R. Parr, S. Wallace, A. P. Monaco, G. Barnby, K. Kobayashi, J. A. Lamb, I. Sousa, N. Sykes, E. H. Cook, S. J. Guter, B. L. Leventhal, J. Salt, C. Lord, C. Corsello, V. Hus, D. E. Weeks, F. Volkmar, M. Tauber, E. Fombonne, A. Shih, and K. J. Meyer, "Mapping autism risk loci using genetic linkage and chromosomal rearrangements," Nat Genet, vol. 39, pp. 319–28, Mar 2007.

[30] SPARK Consortium. Electronic address: pfeliciano@simonsfoundation.org and SPARK Consortium, "Spark: A us cohort of 50,000 families to accelerate autism research," Neuron, vol. 97, pp. 488–493, 02 2018.

[31] G. D. Fischbach and C. Lord, "The simons simplex collection: a resource for identification of autism genetic risk factors," Neuron, vol. 68, pp. 192–5, Oct 2010.

[32] C. Lord, M. Rutter, and A. Le Couteur, "Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders," J Autism Dev Disord, vol. 24, pp. 659–85, Oct 1994.

[33] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, Jr, B. L. Leventhal, P. C. DiLavore, A. Pickles, and M. Rutter, "The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism," J Autism Dev Disord, vol. 30, pp. 205–23, Jun 2000.

[34] O. Delaneau, J. Marchini, and J.-F. Zagury, "A linear complexity phasing method for thousands of genomes," Nat Methods, vol. 9, pp. 179–81, Dec 2011.

[35] R. Anney, L. Klei, D. Pinto, R. Regan, J. Conroy, T. R. Magalhaes, C. Correia, B. S. Abrahams, N. Sykes, A. T. Pagnamenta, J. Almeida, E. Bacchelli, A. J. Bailey, G. Baird, A. Battaglia, T. Berney, N. Bolshakova, S. Bölte, P. F. Bolton, T. Bourgeron, S. Brennan, J. Brian, A. R. Carson, G. Casallo, J. Casey, S. H. Chu, L. Cochrane, C. Corsello, E. L. Crawford, A. Crossett, G. Dawson, M. de Jonge, R. Delorme, I. Drmic, E. Duketis, F. Duque, A. Estes, P. Farrar, B. A. Fernandez, S. E. Folstein, E. Fombonne, C. M. Freitag, J. Gilbert, C. Gillberg, J. T. Glessner, J. Goldberg, J. Green, S. J. Guter, H. Hakonarson, E. A. Heron, M. Hill, R. Holt, J. L. Howe, G. Hughes, V. Hus, R. Igliozzi, C. Kim, S. M. Klauck, A. Kolevzon, O. Korvatska, V. Kustanovich, C. M. Lajonchere, J. A. Lamb, M. Laskawiec, M. Leboyer, A. Le Couteur, B. L. Leventhal, A. C. Lionel, X.-Q. Liu, C. Lord, L. Lotspeich, S. C. Lund, E. Maestrini, W. Mahoney, C. Mantoulan, C. R. Marshall, H. McConachie, C. J. McDougle,

J. McGrath, W. M. McMahon, N. M. Melhem, A. Merikangas, O. Migita, N. J. Minshew, G. K. Mirza, J. Munson, S. F. Nelson, C. Noakes, A. Noor, G. Nygren, G. Oliveira, K. Papanikolaou, J. R. Parr, B. Parrini, T. Paton, A. Pickles, J. Piven, D. J. Posey, A. Poustka, F. Poustka, A. Prasad, J. Ragoussis, K. Renshaw, J. Rickaby, W. Roberts, K. Roeder, B. Roge, M. L. Rutter, L. J. Bierut, J. P. Rice, J. Salt, K. Sansom, D. Sato, R. Segurado, L. Senman, N. Shah, V. C. Sheffield, L. Soorya, I. Sousa, V. Stoppioni, C. Strawbridge, R. Tancredi, K. Tansey, B. Thiruvahindrapduram, A. P. Thompson, S. Thomson, A. Tryfon, J. Tsiantis, H. Van Engeland, J. B. Vincent, F. Volkmar, S. Wallace, K. Wang, Z. Wang, T. H. Wassink, K. Wing, K. Wittemeyer, S. Wood, B. L. Yaspan, D. Zurawiecki, L. Zwaigenbaum, C. Betancur, J. D. Buxbaum, R. M. Cantor, E. H. Cook, H. Coon, M. L. Cuccaro, L. Gallagher, D. H. Geschwind, M. Gill, J. L. Haines, J. Miller, A. P. Monaco, J. I. Nurnberger, Jr, A. D. Paterson, M. A. Pericak-Vance, G. D. Schellenberg, S. W. Scherer, J. S. Sutcliffe, P. Szatmari, A. M. Vicente, V. J. Vieland, E. M. Wijsman, B. Devlin, S. Ennis, and J. Hallmayer, "A genome-wide scan for common alleles affecting risk for autism," Hum Mol Genet, vol. 19, pp. 4072–82, Oct 2010.

[36] J. Grove, S. Ripke, T. D. Als, M. Mattheisen, R. K. Walters, H. Won, J. Pallesen, E. Agerbo, O. A. Andreassen, R. Anney, S. Awashti, R. Belliveau, F. Bettella, J. D. Buxbaum, J. Bybjerg-Grauholm, M. Bækvad-Hansen, F. Cerrato, K. Chambert, J. H. Christensen, C. Churchhouse, K. Dellenvall, D. Demontis, S. De Rubeis, B. Devlin, S. Djurovic, A. L. Dumont, J. I. Goldstein, C. S. Hansen, M. E. Hauberg, M. V. Hollegaard, S. Hope, D. P. Howrigan, H. Huang, C. M. Hultman, L. Klei, J. Maller, J. Martin, A. R. Martin, J. L. Moran, M. Nyegaard, T. Nærland, D. S. Palmer, A. Palotie, C. B. Pedersen, M. G. Pedersen, T. dPoterba, J. B. Poulsen, B. S. Pourcain, P. Qvist, K. Rehnström, A. Reichenberg, J. Reichert, E. B. Robinson, K. Roeder, P. Roussos, E. Saemundsen, S. Sandin, F. K. Satterstrom, G. Davey Smith, H. Stefansson, S. Steinberg, C. R. Stevens, P. F. Sullivan, P. Turley, G. B. Walters, X. Xu, Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium, BUPGEN, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, 23andMe Research Team, K. Stefansson, D. H. Geschwind, M. Nordentoft, D. M. Hougaard, T. Werge, O. Mors, P. B. Mortensen, B. M. Neale, M. J. Daly, and A. D. Børglum, "Identification of common genetic risk variants for autism spectrum disorder," Nat Genet, vol. 51, pp. 431–444, 03 2019.

[37] J. Gauthier, T. J. Siddiqui, P. Huashan, D. Yokomaku, F. F. Hamdan, N. Champagne, M. Lapointe, D. Spiegelman, A. Noreau, R. G. Lafrenière, F. Fathalli, R. Joober, M.-O. Krebs, L. E. DeLisi, L. Mottron, E. Fombonne, J. L. Michaud, P. Drapeau, S. Carbonetto, A. M. Craig, and G. A. Rouleau, "Truncating mutations in nrxn2 and nrxn1 in autism spectrum disorders and schizophrenia," Hum Genet, vol. 130, pp. 563–73, Oct 2011.

[38] H.-G. Kim, S. Kishikawa, A. W. Higgins, I.-S. Seong, D. J. Donovan, Y. Shen, E. Lally, L. A. Weiss, J. Najm, K. Kutsche, M. Descartes, L. Holt, S. Braddock, R. Troxell, L. Kaplan, F. Volkmar, A. Klin, K. Tsatsanis, D. J. Harris, I. Noens, D. L. Pauls, M. J. Daly, M. E. MacDonald, C. C. Morton, B. J. Quade, and J. F. Gusella, "Disruption of neurexin 1 associated with autism spectrum disorder," Am J Hum Genet, vol. 82, pp. 199–207, Jan 2008.

[39] D.-H. Lu, H.-M. Liao, C.-H. Chen, H.-J. Tu, H.-C. Liou, S. S.-F. Gau, and W.-M. Fu, "Impairment of social behaviors in arhgef10 knockout mice," Mol Autism, vol. 9, p. 11, 2018.

[40] A. Gouy and L. Excoffier, "Polygenic patterns of adaptive introgression in modern humans are mainly shaped by response to pathogens," Mol Biol Evol, vol. 37, pp. 1420–1433, 05 2020.

[41] J. Lee, M. J. Son, C. Y. Son, G. H. Jeong, K. H. Lee, K. S. Lee, Y. Ko, J. Y. Kim, J. Y. Lee, J. Radua, M. Eisenhut, F. Gressier, A. Koyanagi, B. Stubbs, M. Solmi, T. B. Rais, A. Kronbichler, E. Dragioti, D. F. P. Vasconcelos, F. R. P. d. Silva, K. Tizaoui, A. R. Brunoni, A. F. Carvalho, S. Cargnin, S. Terrazzino, A. Stickley, L. Smith, T. Thompson, J. I. Shin, and P. Fusar-Poli, "Genetic variation and autism: A field synopsis and systematic meta-analysis," Brain Sci, vol. 10, Sep 2020.

[42] F. Marbach, G. Stoyanov, F. Erger, C. A. Stratakis, N. Settas, E. London, J. A. Rosenfeld, E. Torti, C. Haldeman-Englert, E. Sklirou, et al., "Variants in prkar1b cause a neurodevelopmental disorder with autism spectrum disorder, apraxia, and insensitivity to pain," Genetics in medicine, vol. 23, no. 8, pp. 1465–1473, 2021.

[43] E. K. Ruzzo, L. Pérez-Cano, J.-Y. Jung, L.-K. Wang, D. Kashef-Haghighi, C. Hartl, C. Singh, J. Xu, J. N. Hoekstra, O. Leventhal, V. M. Leppä, M. J. Gandal, K. Paskov, N. Stockham, D. Polioudakis, J. K. Lowe, D. A. Prober, D. H. Geschwind, and D. P. Wall, "Inherited and de novo genetic risk for autism impacts shared networks," Cell, vol. 178, pp. 850–866.e26, 08 2019.

[44] T. N. Turner, F. Hormozdiari, M. H. Duyzend, S. A. McClymont, P. W. Hook, I. Iossifov, A. Raja, C. Baker, K. Hoekzema, H. A. Stessman, M. C. Zody, B. J. Nelson, J. Huddleston, R. Sandstrom, J. D. Smith, D. Hanna, J. M. Swanson, E. M. Faustman, M. J. Bamshad, J. Stamatoyannopoulos, D. A. Nickerson, A. S. McCallion, R. Darnell, and E. E. Eichler, "Genome sequencing of autism-affected families reveals disruption of putative noncoding regulatory dna," Am J Hum Genet, vol. 98, pp. 58–74, Jan 2016.

[45] S. Chen, X. Zhou, E. Byington, S. L. Bruce, H. Zhang, and Y. Shen, "Dissecting autism genetic risk using single-cell rna-seq data," bioRxiv, 2020.

[46] J. P. Casey, T. Magalhaes, J. M. Conroy, R. Regan, N. Shah, R. Anney, D. C. Shields, B. S. Abrahams, J. Almeida, E. Bacchelli, A. J. Bailey, G. Baird, A. Battaglia, T. Berney, N. Bolshakova, P. F. Bolton, T. Bourgeron, S. Brennan, P. Cali, C. Correia, C. Corsello, M. Coutanche, G. Dawson, M. de Jonge, R. Delorme, E. Duketis, F. Duque, A. Estes, P. Farrar, B. A. Fernandez, S. E. Folstein, S. Foley, E. Fombonne, C. M. Freitag, J. Gilbert, C. Gillberg, J. T. Glessner, J. Green, S. J. Guter, H. Hakonarson, R. Holt, G. Hughes, V. Hus, R. Igliozzi, C. Kim, S. M. Klauck, A. Kolevzon, J. A. Lamb, M. Leboyer, A. Le Couteur, B. L. Leventhal, C. Lord, S. C. Lund, E. Maestrini, C. Mantoulan, C. R. Marshall, H. McConachie, C. J. McDougle, J. McGrath, W. M. McMahon, A. Merikangas, J. Miller, F. Minopoli, G. K. Mirza, J. Munson, S. F. Nelson, G. Nygren, G. Oliveira, A. T. Pagnamenta, K. Papanikolaou, J. R. Parr, B. Parrini, A. Pickles, D. Pinto, J. Piven, D. J. Posey, A. Poustka, F. Poustka, J. Ragoussis, B. Roge, M. L. Rutter, A. F. Sequeira, L. Soorya, I. Sousa, N. Sykes, V. Stoppioni, R. Tancredi, M. Tauber, A. P. Thompson, S. Thomson, J. Tsiantis, H. Van Engeland, J. B. Vincent, F. Volkmar, J. A. S. Vorstman, S. Wallace, K. Wang, T. H. Wassink, K. White, K. Wing, K. Wittemeyer, B. L. Yaspan, L. Zwaigenbaum, C. Betancur, J. D. Buxbaum, R. M. Cantor, E. H. Cook, H. Coon, M. L. Cuccaro, D. H. Geschwind, J. L. Haines, J. Hallmayer, A. P. Monaco, J. I. Nurnberger, Jr, M. A. Pericak-Vance, G. D. Schellenberg, S. W. Scherer, J. S. Sutcliffe, P. Szatmari, V. J. Vieland, E. M. Wijsman, A. Green, M. Gill, L. Gallagher, A. Vicente, and S. Ennis, "A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder," Hum Genet, vol. 131, pp. 565–79, Apr 2012.

[47] B. Namjou, K. Marsolo, R. J. Caroll, J. C. Denny, M. D. Ritchie, S. S. Verma, T. Lingren, A. Porollo, B. L. Cobb, C. Perry, L. C. Kottyan, M. E. Rothenberg, S. D. Thompson, I. A. Holm, I. S. Kohane, and J. B. Harley, "Phenome-wide association study (phewas) in emr-linked pediatric cohorts, genetically links plcl1 to speech language development and il5-il13 to eosinophilic esophagitis," Front Genet, vol. 5, p. 401, 2014.

[48] E. D. Gamsiz, E. W. Viscidi, A. M. Frederick, S. Nagpal, S. J. Sanders, M. T. Murtha, M. Schmidt, Simons Simplex Collection Genetics Consortium, E. W. Triche, D. H. Geschwind, M. W. State, S. Istrail, E. H. Cook, Jr, B. Devlin, and E. M. Morrow, "Intellectual disability is associated with increased runs of homozygosity in simplex autism," Am J Hum Genet, vol. 93, pp. 103–9, Jul 2013.

[49] S. Calderoni, I. Ricca, G. Balboni, R. Cagiano, D. Cassandrini, S. Doccini, A. Cosenza, D. Tolomeo, R. Tancredi, F. M. Santorelli, et al., "Evaluation of chromosome microarray analysis in a large cohort of females with autism spectrum disorders: a single center italian study," Journal of personalized medicine, vol. 10, no. 4, p. 160, 2020.

[50] X. Xiao, F. Zheng, H. Chang, Y. Ma, Y.-G. Yao, X.-J. Luo, and M. Li, "The gene encoding protocadherin 9 (pcdh9), a novel risk factor for major depressive disorder," Neuropsychopharmacology, vol. 43, pp. 1128–1137, 04 2018.

[51] C. R. Marshall, A. Noor, J. B. Vincent, A. C. Lionel, L. Feuk, J. Skaug, M. Shago, R. Moessner, D. Pinto, Y. Ren, et al., "Structural variation of chromosomes in autism spectrum disorder," The American Journal of Human Genetics, vol. 82, no. 2, pp. 477–488, 2008.

[52] H. Bruining, A. Matsui, A. Oguro-Ando, R. S. Kahn, H. M. Van't Spijker, G. Akkermans, O. Stiedl, H. van Engeland, B. Koopmans, H. A. van Lith, H. Oppelaar, L. Tieland, L. J. Nonkes, T. Yagi, R. Kaneko, J. P. H. Burbach, N. Yamamoto, and M. J. Kas, "Genetic mapping in mice reveals the involvement of pcdh9 in long-term social and object recognition and sensorimotor development," Biol Psychiatry, vol. 78, pp. 485–95, Oct 2015.

[53] E. Maestrini, A. T. Pagnamenta, J. A. Lamb, E. Bacchelli, N. H. Sykes, I. Sousa, C. Toma, G. Barnby, H. Butler, L. Winchester, T. S. Scerri, F. Minopoli, J. Reichert, G. Cai, J. D. Buxbaum, O. Korvatska, G. D. Schellenberg, G. Dawson, A. de Bildt, R. B. Minderaa, E. J. Mulder, A. P. Morris, A. J. Bailey, A. P. Monaco, and IMGSAC, "High-density snp association study and copy number variation analysis of the auts1 and auts5 loci implicate the immp2l-dock4 gene region in autism susceptibility," Mol Psychiatry, vol. 15, pp. 954–68, Sep 2010.

[54] A. T. Pagnamenta, E. Bacchelli, M. V. de Jonge, G. Mirza, T. S. Scerri, F. Minopoli, A. Chiocchetti, K. U. Ludwig, P. Hoffmann, S. Paracchini, E. Lowy, D. H. Harold, J. A. Chapman, S. M. Klauck, F. Poustka, R. H. Houben, W. G. Staal, R. A. Ophoff, M. C. O'Donovan, J. Williams, M. M. Nöthen, G. Schulte-Körne, P. Deloukas, J. Ragoussis, A. J. Bailey, E. Maestrini, A. P. Monaco, and International Molecular Genetic Study Of Autism Consortium, "Characterization of a family with rare deletions in cntnap5 and dock4 suggests novel risk loci for autism and dyslexia," Biol Psychiatry, vol. 68, pp. 320–8, Aug 2010.

[55] D. Guo, Y. Peng, L. Wang, X. Sun, X. Wang, C. Liang, X. Yang, S. Li, J. Xu, W.-C. Ye, et al., "Autism-like social deficit generated by dock4 deficiency is rescued by restoration of rac1 activity and nmda receptor function," Molecular psychiatry, vol. 26, no. 5, pp. 1505–1519, 2021.

[56] The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, "Meta-analysis of gwas of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24. 32 and a significant overlap with schizophrenia," Molecular autism, vol. 8, pp. 1–17, 2017.

[57] Z. Agha, Z. Iqbal, M. Azam, H. Ayub, L. E. L. M. Vissers, C. Gilissen, S. H. B. Ali, M. Riaz, J. A. Veltman, R. Pfundt, H. van Bokhoven, and R. Qamar, "Exome sequencing identifies three novel candidate genes implicated in intellectual disability," PLoS One, vol. 9, no. 11, p. e112687, 2014.

[58] M. A. Bemben, Q.-A. Nguyen, T. Wang, Y. Li, R. A. Nicoll, and K. W. Roche, "Autism-associated mutation inhibits protein kinase c-mediated neuroligin-4x enhancement of excitatory synapses," Proc Natl Acad Sci U S A, vol. 112, pp. 2551–6, Feb 2015.

[59] L. Ji, A. Chauhan, and V. Chauhan, "Reduced activity of protein kinase c in the frontal cortex of subjects with regressive autism: relationship with developmental abnormalities," Int J Biol Sci, vol. 8, no. 7, pp. 1075–84, 2012.

[60] A. J. Schork, H. Won, V. Appadurai, R. Nudel, M. Gandal, O. Delaneau, M. Revsbech Christiansen, D. M. Hougaard, M. Bækved-Hansen, J. Bybjerg-Grauholm, M. Giørtz Pedersen, E. Agerbo, C. Bøcker Pedersen, B. M. Neale, M. J. Daly, N. R. Wray, M. Nordentoft, O. Mors, A. D. Børglum, P. Bo Mortensen, A. Buil, W. K. Thompson, D. H. Geschwind, and T. Werge, "A genome-wide association study of shared risk across psychiatric disorders implicates gene regulation during fetal neurodevelopment," Nat Neurosci, vol. 22, pp. 353–361, 03 2019.

[61] X. Teng, A. Aouacheria, L. Lionnard, K. A. Metz, L. Soane, A. Kamiya, and J. M. Hardwick, "Kctd: A new gene family involved in neurodevelopmental and neuropsychiatric disorders," CNS Neurosci Ther, vol. 25, pp. 887–902, 07 2019.

[62] S. S. Tran, H.-I. Jun, J. H. Bahn, A. Azghadi, G. Ramaswami, E. L. Van Nostrand, T. B. Nguyen, Y.-H. E. Hsiao, C. Lee, G. A. Pratt, et al., "Widespread rna editing dysregulation in brains from autistic individuals," Nature neuroscience, vol. 22, no. 1, pp. 25–36, 2019.

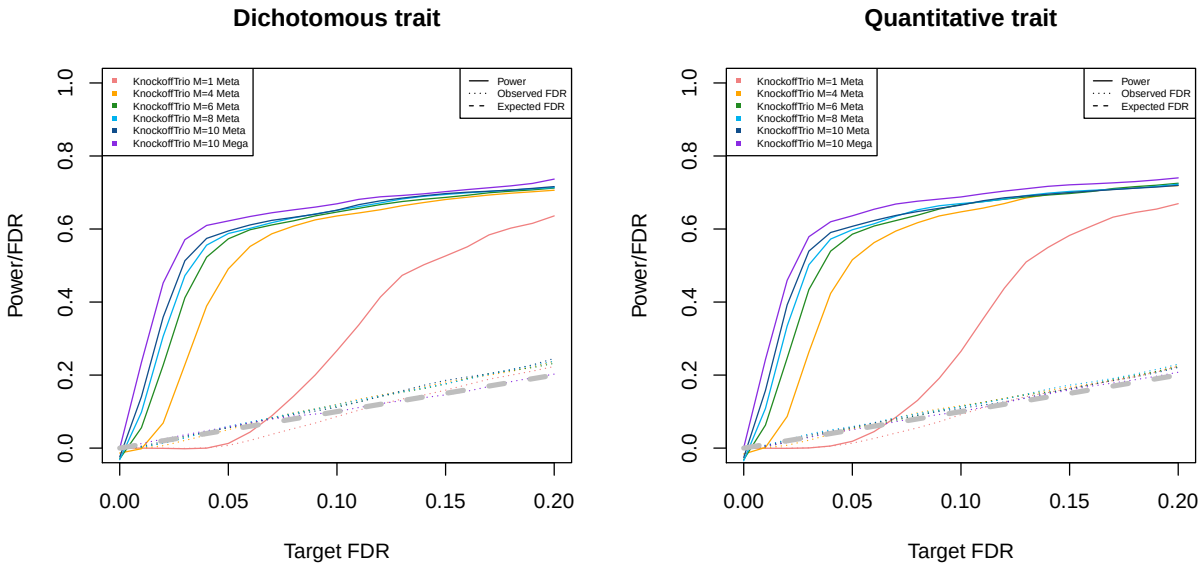
[63] V. W. Hu, C. A. Devlin, and J. J. Debski, "Asd phenotype-genotype associations in concordant and discordant monozygotic and dizygotic twins stratified by severity of autistic traits," Int J Mol Sci, vol. 20, Aug 2019.

[64] S. Li, Z. Ren, C. Sabatti, and M. Sesia, "Transfer learning in genome-wide association studies with knockoffs," arXiv preprint arXiv:2108.08813, 2021.

A Empirical power and FDR in meta-analysis simulations

We adopted the same simulations in the "Empirical power and FDR in single-locus simulations" Section except that in each replicate we partitioned the trios into two subcohorts of 5,000 trios each. We then applied KnockoffTrio to the two subcohorts respectively and used KnockoffTrio's meta-analysis procedure to combine the results from the two subcohorts. In Figure S1, we show the empirical power and FDR for KnockoffTrio's meta-analysis of the two subcohorts, compared to the corresponding mega-analysis of the combined cohort. KnockoffTrio's meta-analysis has comparable power to the mega-analysis (when $M = 10$) while preserving the FDR in all scenarios.

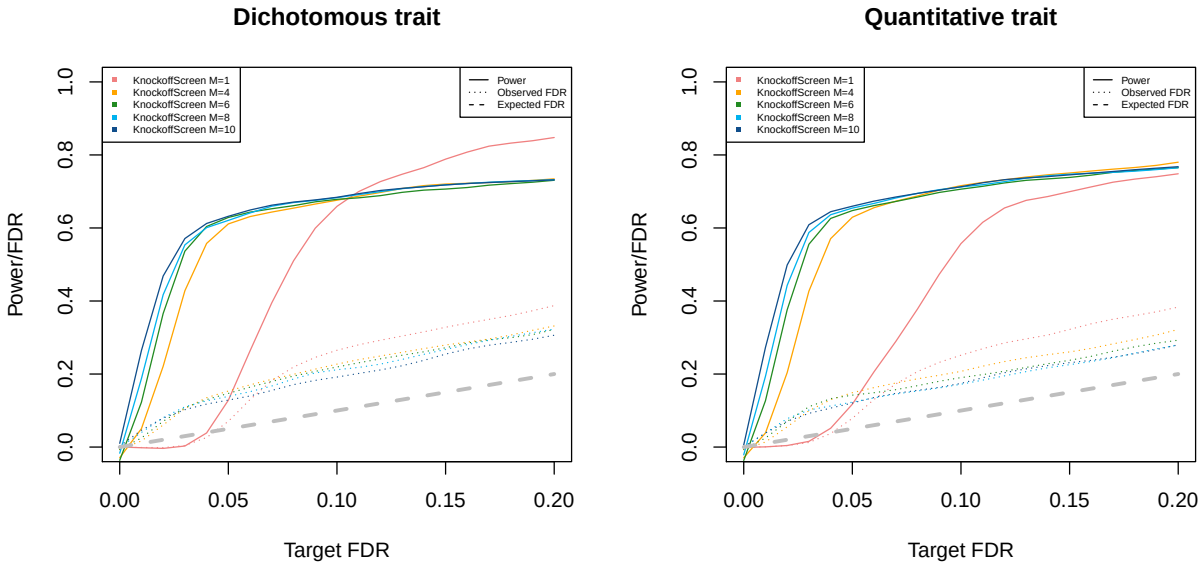
Figure S1: KnockoffTrio's power and FDR in meta-analysis. The two panels show the power and FDR for dichotomous and quantitative traits. We evaluate KnockoffTrio's power and FDR with a target FDR ranging from 0 to 0.2 and with different numbers of knockoffs. The solid lines indicate KnockoffTrio's power and the dotted lines indicate KnockoffTrio's observed FDR. The different colors indicate different numbers of knockoffs. The grey dashed line indicates the expected FDR.



B KnockoffScreen in trio studies

KnockoffScreen was designed for independent individuals in population-based studies. We investigated the power and FDR of KnockoffScreen in trio studies. Specifically, we used KnockoffScreen to generate knockoffs of trio data disregarding the family structure and treating family members as unrelated individuals. We adopted the same simulations in the "Empirical power and FDR in single-locus simulations" Section. As shown in Figure S2, KnockoffScreen has inflated FDR when applied to trio data.

Figure S2: KnockoffScreen's power and FDR in single-locus simulations. The two panels show the power and FDR for dichotomous and quantitative traits. We evaluate KnockoffScreen's power and FDR with a target FDR ranging from 0 to 0.2 and with different numbers of knockoffs. The solid lines indicate KnockoffTrio's power and the dotted lines indicate KnockoffScreen's observed FDR. The different colors indicate different numbers of knockoffs. The grey dashed line indicates the expected FDR.



C Exchangeability and FDR control in trio studies

We now formally define the exchangeability and FDR control in trio studies.

C.1 Notations

Let $\mathbf{P}_i$, a $4 \times p$ matrix, denote the parental haplotypes for the i-th trio:

$$\mathbf{P}_i = \begin{bmatrix} H_i^{f,1} \\ H_i^{f,2} \\ H_i^{m,1} \\ H_i^{m,2} \end{bmatrix}.$$

We assume that the genome has been divided into K contiguous regions of equal sizes, e.g., 200kb, and that no recombination occurs within each region. The indices of genetic variants in the k-th region are denoted as $\mathcal{T}_k$, and $t_k = |\mathcal{T}_k|$. We have $\cup_{k=1}^{K} \mathcal{T}_k = [p]$. Let $\mathbf{P}_{i,k}$, a $4 \times t_k$ matrix, denote the parental haplotypes in the k-th region.

Given the parental haplotypes, the offspring haplotypes within a region are a function of the parental haplotypes, either viewed as deterministic (observed) or random. As we have assumed no recombination events, the form of the function remains the same for all genetic variables in $\mathcal{T}_k$. A further observation is that this function is a linear form of $\mathbf{P}_{i,k}$. Therefore, it can be represented by $\mathbf{\Gamma}_{i,k}$, a 2×4 matrix. Now let $\mathbf{\Gamma}_i$ be a $2K \times 4K$ block diagonal matrix of $\mathbf{\Gamma}_{i,k}$'s

$$\mathbf{\Gamma}_i = \begin{bmatrix} \mathbf{\Gamma}_{i,1} & & \\ & \ddots & \\ & & \mathbf{\Gamma}_{i,K} \end{bmatrix},$$

and $\mathbf{P}_i^{\text{block}}$ be a $4K \times p$ block diagonal matrix of $\mathbf{P}_{i,k}$'s

$$\mathbf{P}_i^{\text{block}} = \begin{bmatrix} \mathbf{P}_{i,1} & & \\ & \ddots & \\ & & \mathbf{P}_{i,K} \end{bmatrix}.$$

Then, the offspring haplotypes can be represented as

$$\begin{bmatrix} X_i^f \\ X_i^m \end{bmatrix} = \mathbf{A}\mathbf{\Gamma}_i\mathbf{P}_i^{\text{block}},$$

where A is a $2 \times 2K$ matrix of 0's and 1's, with 1's at odd positions in the first row and at even positions in the second row.

C.2 Exchangeability for trios

We first construct the knockoff variables for parental haplotypes using the SCIT algorithms as described in the Methods section. Let $\tilde{\mathbf{P}}_i$, a $4 \times p$ matrix, denote the knockoff parental haplotypes for the i-th trio

$$\tilde{\mathbf{P}}_i = \begin{bmatrix} \tilde{H}_i^{f,1} \\ \tilde{H}_i^{f,2} \\ \tilde{H}_i^{m,1} \\ \tilde{H}_i^{m,2} \end{bmatrix}.$$

We also define $\tilde{\mathbf{P}}_i^{\text{block}}$ similarly. The knockoff offspring haplotypes are obtained by setting

$$\begin{bmatrix} \tilde{X}_i^f \\ \tilde{X}_i^m \end{bmatrix} = \mathbf{A}\mathbf{\Gamma}_i\tilde{\mathbf{P}}_i^{\text{block}},$$

where we assume the transmission patterns are the same for both the original and synthetic trios.

The exchangeability is defined at the matrix level. We say that $[\mathbf{P}_i, \tilde{\mathbf{P}}_i]$ satisfies the exchangeability condition if for any $\mathcal{S} \subset [p]$

$$[\mathbf{P}_i, \tilde{\mathbf{P}}_i] \overset{D}{=} [\mathbf{P}_i, \tilde{\mathbf{P}}_i]_{\text{swap}(\mathcal{S})}$$

where $[\mathbf{P}_i, \tilde{\mathbf{P}}_i]_{\text{swap}(\mathcal{S})}$ is obtained by swapping the j-th column of $\mathbf{P}_i$ and $\tilde{\mathbf{P}}_i$ for $j \in \mathcal{S}$. Note that if the exchangeability condition holds for $[\mathbf{P}_i, \tilde{\mathbf{P}}_i]$, then it also holds for $[\mathbf{P}_i^{\text{block}}, \tilde{\mathbf{P}}_i^{\text{block}}]$. Therefore,

$$\begin{bmatrix} X_i^f & \tilde{X}_i^f \\ X_i^m & \tilde{X}_i^m \end{bmatrix}_{\text{swap}(\mathcal{S})} = \mathbf{A}\mathbf{\Gamma}_i[\mathbf{P}_i^{\text{block}}, \tilde{\mathbf{P}}_i^{\text{block}}]_{\text{swap}(\mathcal{S})}.$$

This implies that, if we consider all haplotypes in a trio, the exchangeability holds in the sense that

$$\begin{bmatrix} \mathbf{P}_i & \tilde{\mathbf{P}}_i \\ X_i^f & \tilde{X}_i^f \\ X_i^m & \tilde{X}_i^m \end{bmatrix} \overset{D}{=} \begin{bmatrix} \mathbf{P}_i & \tilde{\mathbf{P}}_i \\ X_i^f & \tilde{X}_i^f \\ X_i^m & \tilde{X}_i^m \end{bmatrix}_{\text{swap}(\mathcal{S})}. \tag{3}$$

Here, we treat the transmission pattern $\mathbf{\Gamma}_i$ as given. In the case that we also consider the randomness of $\mathbf{\Gamma}_i$, the exchangeability can be understood as (3) holds conditional on $\mathbf{\Gamma}_i$.

C.3 FDR control for trios

Our goal is to test the conditional null hypothesis

$$H_0 : \boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{G}_g | \boldsymbol{G}_{-g}, g \subset [p].$$

For trios, the null hypothesis is essentially

$$H_{0,l} : \boldsymbol{Y} \perp\!\!\!\perp (\boldsymbol{X}_{gl}, \boldsymbol{P}_{gl}) | (\boldsymbol{X}_{-gl}, \boldsymbol{P}_{-gl}), l = 1, \ldots, L$$

where $\boldsymbol{X} \in \{0,1\}^{2n \times p}$ are the offspring haplotypes, $\boldsymbol{P} \in \{0,1\}^{4n \times p}$ are the parental haplotypes, and $g_1, \ldots, g_L \subset [p]$ are a collection of subsets of indices of p genetic variables. Let KnockoffTrio's feature importance statistic for a window be

$$W_l = w_l\left(\begin{bmatrix} \boldsymbol{P} & \tilde{\boldsymbol{P}} \\ \boldsymbol{X} & \tilde{\boldsymbol{X}} \end{bmatrix}, \boldsymbol{y} \right)$$

for some function w_l. Because the p-values for calculating W_l's are obtained from marginal tests for each genetic variable in a window, we can see that for any $S \subset 1, \ldots, L$

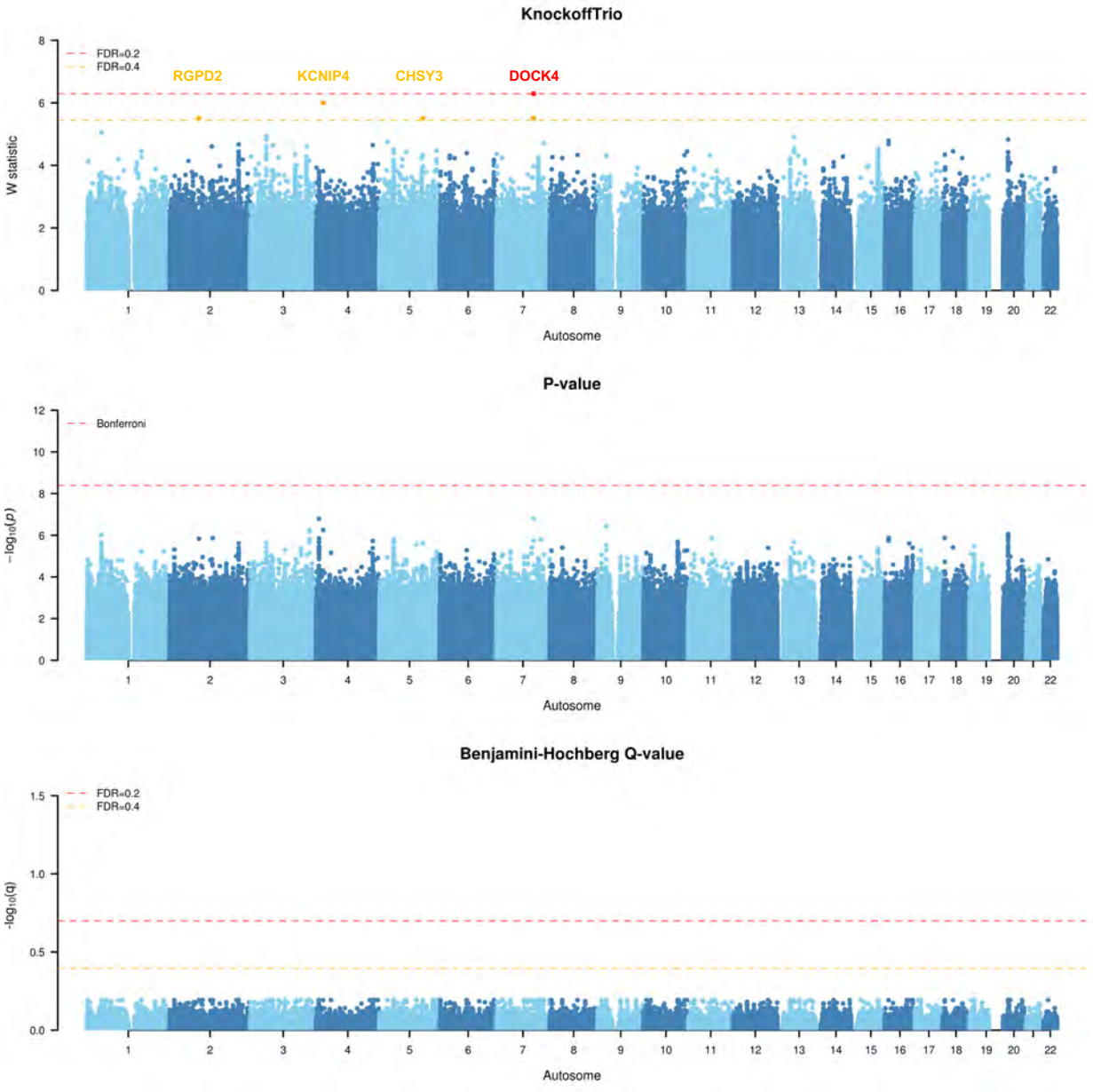
$$w_l\left(\begin{bmatrix} \boldsymbol{P} & \tilde{\boldsymbol{P}} \\ \boldsymbol{X} & \tilde{\boldsymbol{X}} \end{bmatrix}_{\text{swap}(S)}, \boldsymbol{y} \right) = \begin{cases} w_l\left(\begin{bmatrix} \boldsymbol{P} & \tilde{\boldsymbol{P}} \\ \boldsymbol{X} & \tilde{\boldsymbol{X}} \end{bmatrix}, \boldsymbol{y} \right), & l \notin S, \\ -w_l\left(\begin{bmatrix} \boldsymbol{P} & \tilde{\boldsymbol{P}} \\ \boldsymbol{X} & \tilde{\boldsymbol{X}} \end{bmatrix}, \boldsymbol{y} \right), & l \in S \end{cases} \tag{4}$$

where $\begin{bmatrix} \boldsymbol{P} & \tilde{\boldsymbol{P}} \\ \boldsymbol{X} & \tilde{\boldsymbol{X}} \end{bmatrix}_{\text{swap}(S)}$ is defined by swapping original genetic variables in all windows $g_l, l \in S$ with their knockoffs. The flip-sign property (4) in combination with the exchangeability (3) leads to valid FDR control for trios. The proof is for single knockoff construction and can be easily generalized to the case of multiple knockoffs with similar arguments.

D KnockoffTrio meta-analysis including the SSC cohort

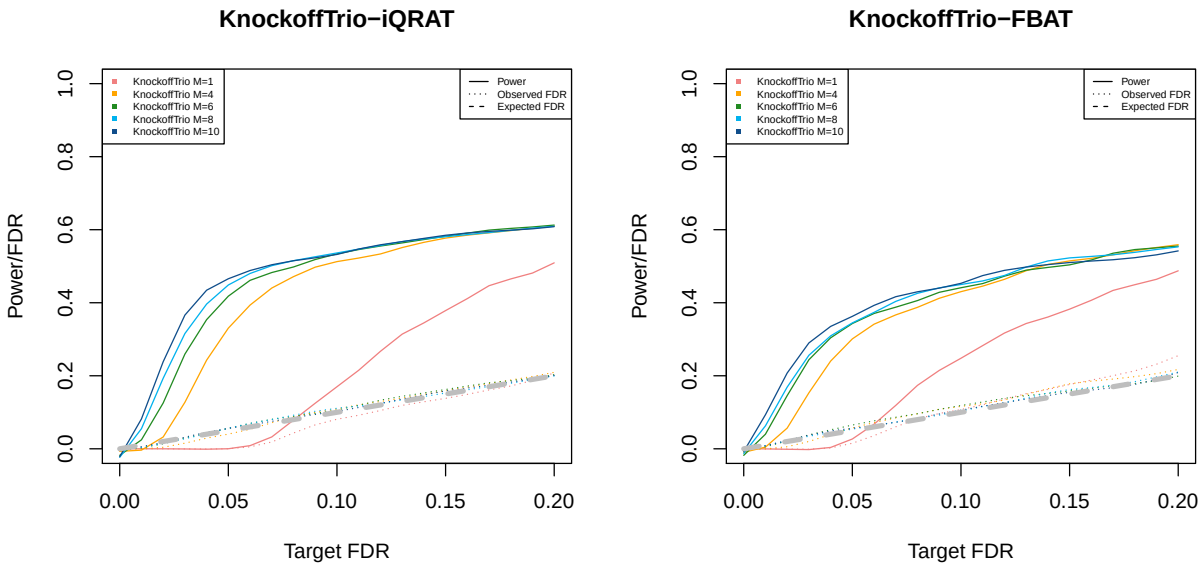
Here we show the results for the meta-analysis of all three cohorts. We note that the number of identified loci is lower than that in the meta-analysis excluding the SSC cohort due to the overall weak signals from the SSC cohort.

Figure S3: Manhattan plots from KnockoffTrio analysis for the meta-analysis of the AGP, SPARK, and SSC cohorts. The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. The FDR for KnockoffTrio and the BH procedure is 0.2 and 0.4. Each locus is annotated with the closest gene name.



E KnockoffTrio-iQRAT

Figure S4: KnockoffTrio-iQRAT improves power in detecting complex associations. The two panels show the power and FDR for quantitative traits with Cauchy error terms using KnockoffTrio-iQRAT and KnockoffTrio-FBAT, respectively. We evaluate KnockoffTrio's power and FDR with a target FDR ranging from 0 to 0.2 and with different numbers of knockoffs. The solid lines indicate KnockoffTrio's power and the dotted lines indicate KnockoffTrio's observed FDR. The different colors indicate different numbers of knockoffs. The grey dashed line indicates the expected FDR.



F Replicability of KnockoffTrio ASD Analyses

Here we assess the replicability of the KnockoffTrio results for the AGP, SPARK, and SSC cohorts and the meta-analysis of the AGP and SPARK cohorts using the same, original haplotype data but different random seeds in knockoff generation.

Figure S5: Replication of Autism Genome Project (AGP) KnockoffTrio results. Manhattan plots from KnockoffTrio analysis for the Autism Genome Project (AGP) with different random seeds in knockoff generation. The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. Different random seeds were used to generate knockoffs than in the main manuscript. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.25. Each locus is annotated with the closest gene name.

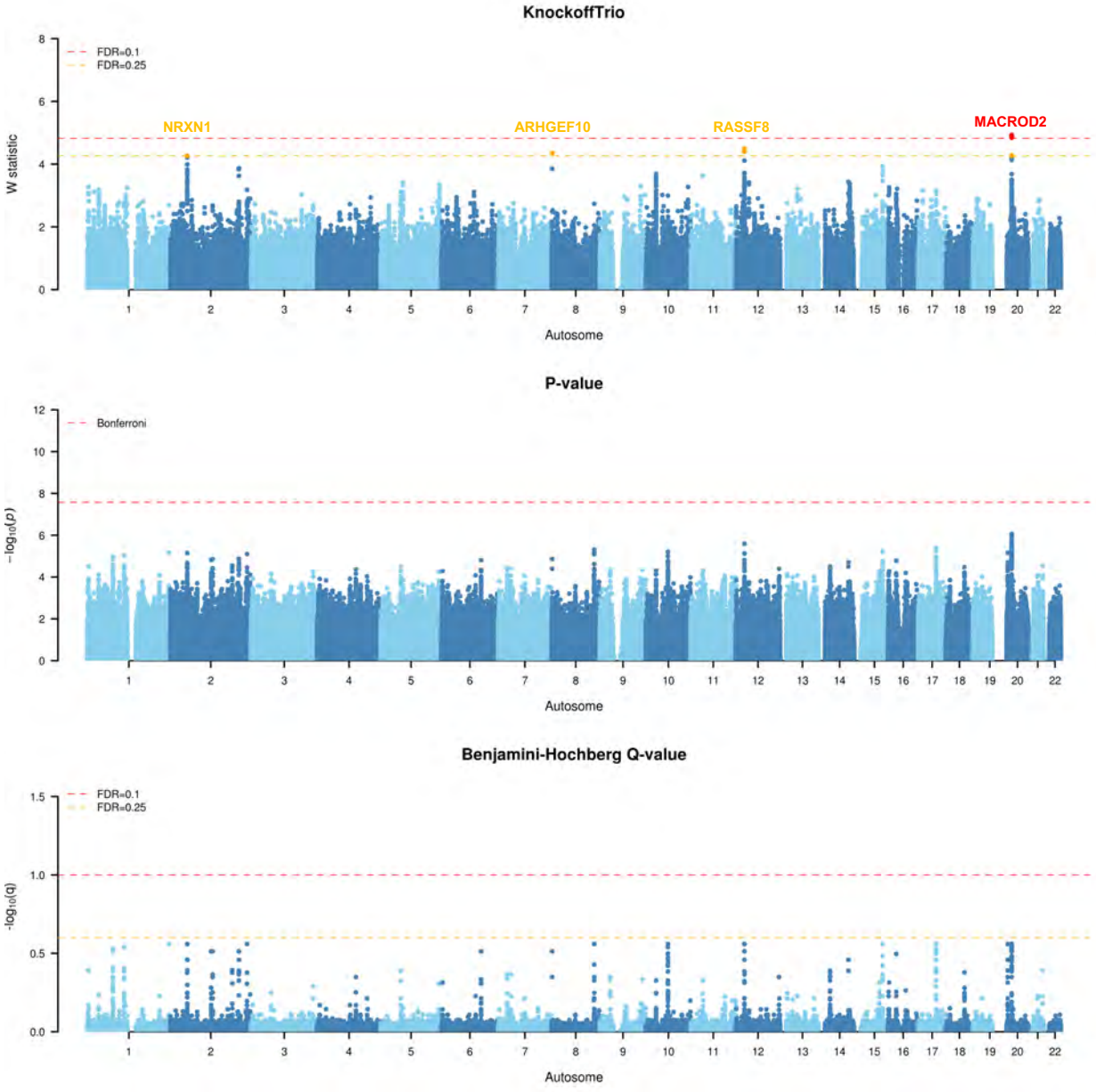


Figure S6: Replication of the Simons Foundation Powering Autism Research (SPARK) KnockoffTrio results. Manhattan plots from KnockoffTrio analysis for the Simons Foundation Powering Autism Research (SPARK) with different random seeds in knockoff generation. The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. Different random seeds were used to generate knockoffs than in the main manuscript. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.2. Each locus is annotated with the closest gene name.

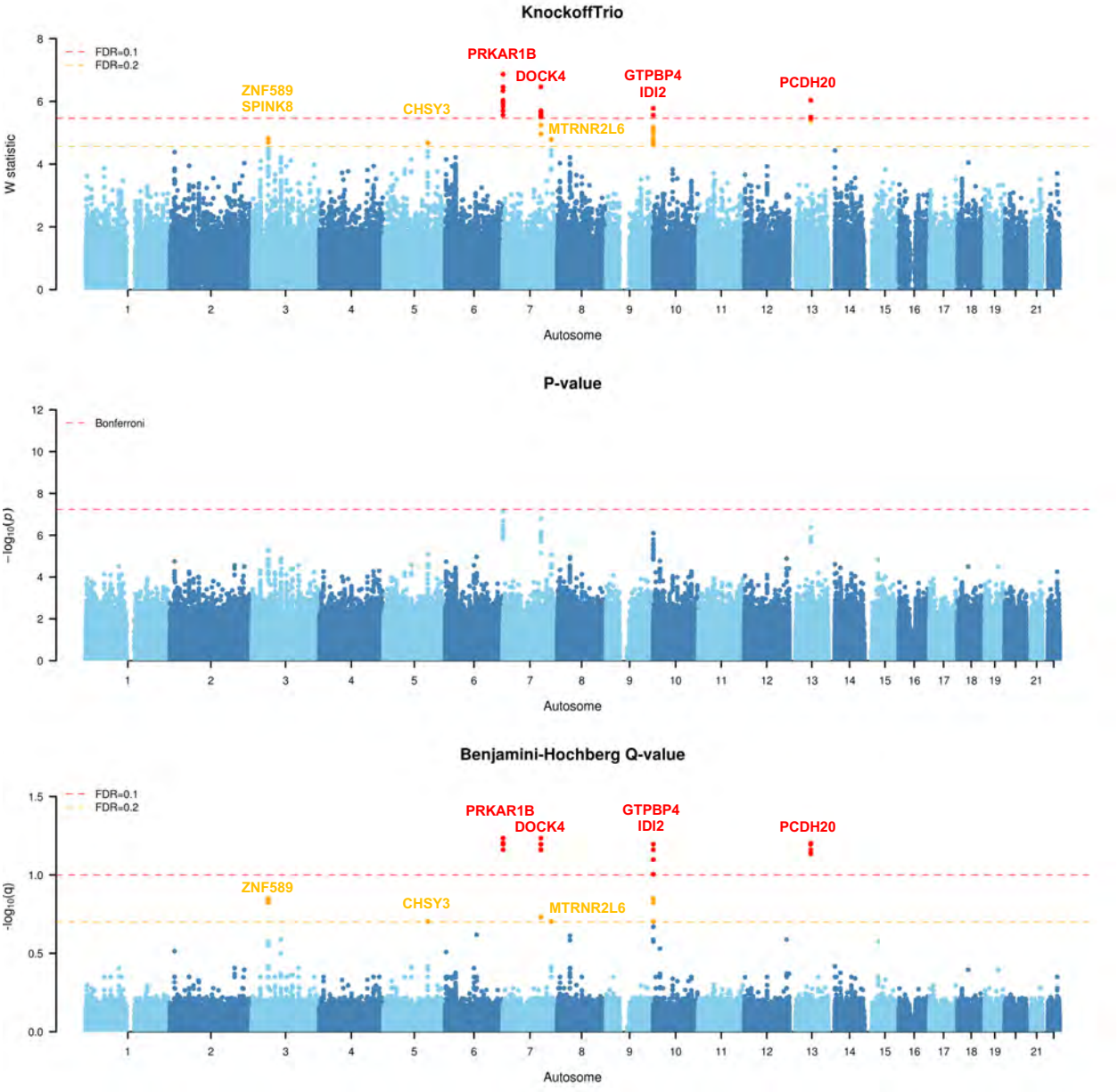


Figure S7: Replication of the Simons Simplex Collection (SSC) KnockoffTrio results. Manhattan plots from KnockoffTrio analysis for the Simons Simplex Collection (SSC) with different random seeds in knockoff generation. The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. Different random seeds were used to generate knockoffs than in the main manuscript. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.2. Each locus is annotated with the closest gene name.

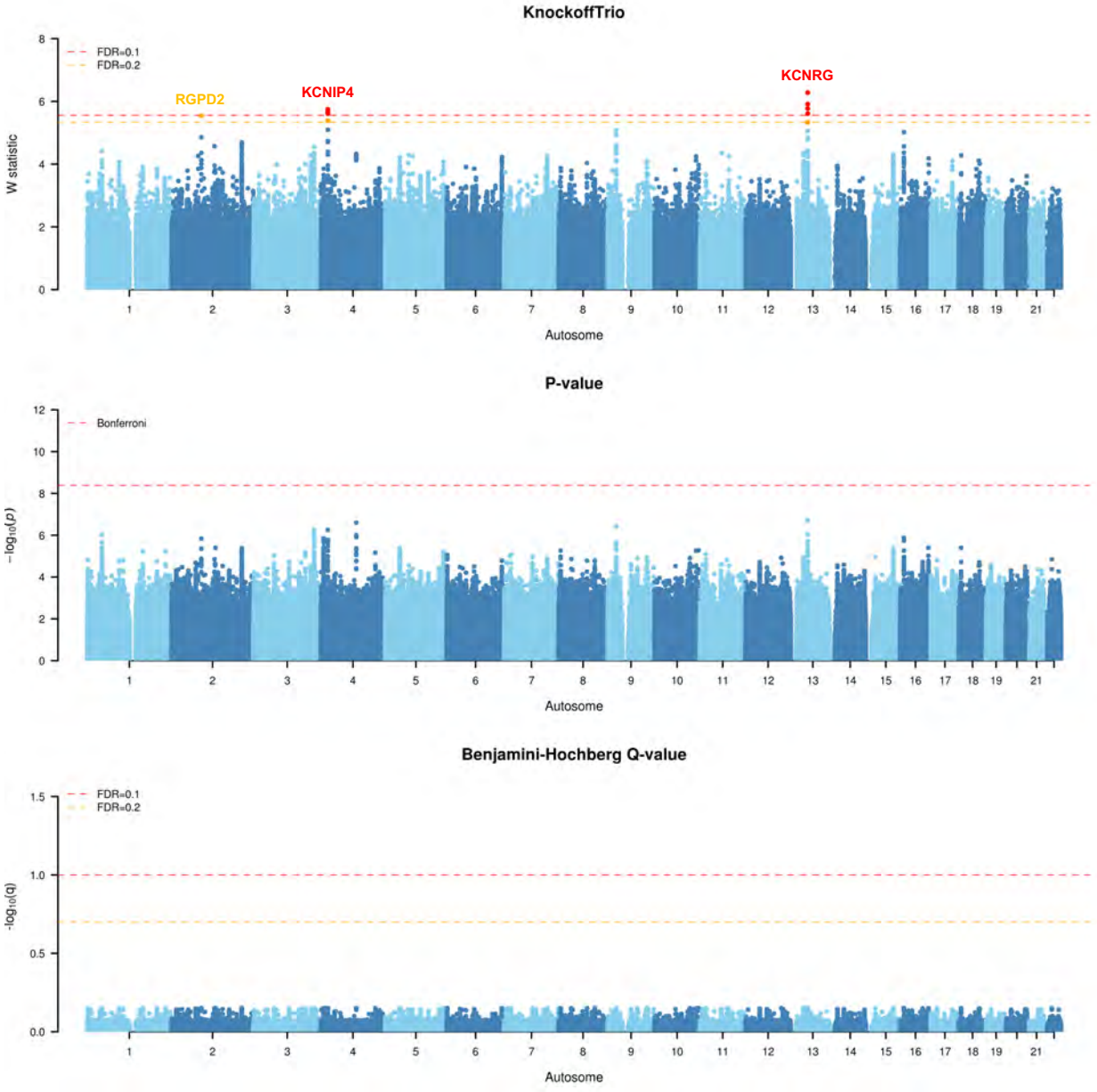


Figure S8: Replication of the AGP and SPARK cohorts KnockoffTrio results. Manhattan plots from KnockoffTrio analysis for the meta-analysis of the AGP and SPARK cohorts with different random seeds in knockoff generation. The Manhattan plots of the W statistics from KnockoffTrio, the p-values from the conventional association tests with the Bonferroni correction for controlling the FWER, and the Q-values from the BH procedure for controlling the FDR. Different random seeds were used to generate knockoffs than in the main manuscript. The FDR target level for KnockoffTrio and the BH procedure is 0.1 or 0.25. Each locus is annotated with the closest gene name.