

SPASQR applies smoothed quantile regression to detect heterogeneous genetic associations

Wenjian Bi

wenjianb@pku.edu.cn

Peking University <https://orcid.org/0000-0002-5108-9311>

Qiang Heng

Southeast University

Lin Miao

School of Basic Medical Sciences, Peking University

He Xu

Peking University

Dajiang Liu

Pennsylvania State University College of Medicine

Seunggeun Lee

Seoul National University <https://orcid.org/0000-0002-8097-3878>

Xiang Zhan

Southeast University <https://orcid.org/0000-0001-9650-143X>

Article

Keywords:

Posted Date: September 8th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10408022/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

**SPA_{SQR} applies smoothed quantile regression to detect
heterogeneous genetic associations**

Qiang Heng^{1,#}, Lin Miao^{2,#}, He Xu², Dajiang Liu^{3,4}, Seunggeun Lee⁵,
Wenjian Bi^{2,6,7,8,9*}, Xiang Zhan^{1,*}

¹ School of Statistics and Data Science, Southeast University, Nanjing, China.

² Department of Medical Genetics, School of Basic Medical Sciences, Peking University, Beijing, China.

³ Department of Public Health Sciences, Penn State College of Medicine, Hershey, United States.

⁴ Department of Molecular and Precision Medicine, Penn State College of Medicine, Hershey, United States.

⁵ Graduate School of Data Science, Seoul National University, Seoul, South Korea.

⁶ Center for Medical Genetics, School of Basic Medical Sciences, Peking University, Beijing, China.

⁷ Medicine Innovation Center for Fundamental Research on Major Immunology-related Diseases, Peking University,
Beijing, China.

⁸ State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing, China.

⁹ Lead Contact

#These authors contributed equally

*Correspondence: Xiang Zhan (zhanx@seu.edu.cn); Wenjian Bi (wenjianb@pku.edu.cn)

20 Genome-wide association studies (GWAS) of quantitative traits are mostly based on
21 linear regression, which associates genetic variants with the mean of the phenotype but
22 has limited power to detect dynamic and context-dependent genetic effects. Quantile
23 regression (QR) offers a powerful alternative to address this limitation, however,
24 existing QR-based GWAS methods remain underdeveloped. Here, we introduce
25 SPA_{SQR}, a scalable and robust QR-based framework designed for biobank-scale GWAS.
26 SPA_{SQR} includes three key innovations: (i) convolution-smoothed quantile regression
27 to enhance statistical power; (ii) score tests based on a retrospective framework to
28 effectively control for relatedness; and (iii) saddlepoint approximation to accurately
29 calibrate p-values for rare variants. In extensive simulations, SPA_{SQR} outperformed
30 existing QR methods in both type I error control and statistical powers. Applied to 66
31 UK Biobank quantitative traits, SPA_{SQR} detected 1,618 loci with heterogeneous effect
32 patterns that were missed by conventional linear approaches. Notably, SPA_{SQR} achieves
33 over 100-fold speedup relative to existing QR-based methods, with computational
34 efficiency comparable to popular GWAS tools such as REGENIE and LDAK-KVIK.

Introduction

For quantitative traits, the majority of genome-wide association studies (GWAS) are based on linear regression models designed to test for genetic effects on mean value of the trait. Through decades of methodology development, the linear models have been extensively optimized to analyze large-scale biobank data with hundreds of thousands of participants while adjusting for cryptic sample relatedness¹⁻⁶. Although effective to detect mean-based genetic associations, linear models may oversimplify the complex nature of biology, in which genetic effects are often highly heterogeneous and context-dependent, varying with sex, age, and lifestyle factors⁷. This oversimplification could result in significant power loss, such that important genetic signals are missed.

Quantile regression (QR) is a regression method that models the quantiles of the trait distribution rather than the mean value, allowing for the detection of more complex and dynamic association patterns⁸. For example, QR at low or high quantiles can more effectively identify association localized to tails of the trait distribution⁹. Recently, there have been a growing number of QR applications in GWAS⁹⁻¹⁵. By integrating genome-wide results across a wide range of quantiles, QR-based GWAS can identify novel genetic variants with heterogeneous genetic effects that traditional linear models miss^{9,15}. Despite important advances, QR-based GWAS methodology remains significantly less developed than conventional linear approaches.

The computational burden in large-scale biobank data analyses poses a major bottleneck. While Wang et al.⁹ have introduced a framework based on QR rank score inference to improve efficiency, it still runs considerably slower than well-established tools such as REGENIE⁴. To address sample relatedness and improve statistical power, leave-one-chromosome-out (LOCO) polygenic scores (PGS) have been incorporated into null quantile regression models¹⁵ as offsets. However, inflated type I error rates persist in cases of strong sample relatedness. Additionally, existing QR-based GWAS mostly analyze common variants, with limited performance evaluation on low-

frequency and rare variants. Our simulations suggest that existing QR-based GWAS methods using rank score inference results in inflated type-I error for rare variants. These limitations collectively hinder a wider application of QR-based GWAS in large-scale biobank data analyses.

In this paper, we introduce SPASQR, a scalable, robust, and powerful method for large-scale QR-based GWAS. SPASQR employs a fast score test based on a retrospective framework¹⁶⁻²⁰, treating genotypes as a multivariate random variable. By incorporating a sparse genetic relationship matrix (GRM)²¹ to adjust the variance of score statistics, SPASQR can effectively control for sample relatedness while maintaining low computational burden. Being a saddlepoint approximation (SPA) method, SPASQR can produce well-calibrated p-values across a wide range of minor allele frequencies (MAF). Motivated by the insight that smoothed QR estimators can be more statistically efficient than regular non-smoothed QR estimators, we leverage convolution-smoothed quantile regression^{22,23} (SQR) instead of classic nonsmooth quantile regression to fit null QR models, which leads to considerable improvement in the overall power of QR-based GWAS. In addition to methodological contributions, we provide a highly computationally efficient C++ implementation of the proposed SPASQR method in our GRAB package, which substantially outperforms existing QR-based GWAS tools in computation efficiency and rivals widely used tools such as REGENIE⁴ and LDAK-KVIK⁶.

Results

An Overview of SPASQR Method

SPASQR method is designed to identify genetic effects across a wide range of phenotype quantiles. Building on a retrospective framework¹⁸, SPASQR is scalable to analyze large cohorts with hundreds and thousands of participants, can effectively account for sample relatedness, and is accurate to analyze low-frequency and rare variants. A comparison

of key methodological features between SPA_{SQR} and existing QR-based GWAS methods can be found in Table 1. In general, the SPA_{SQR} consists of two main steps.

In Step 1, under the null hypothesis of no genetic effects, we fit a smoothed QR model for each quantile τ to adjust for the effects of covariates. In addition to standard covariates such as age, sex, and genetically-derived principal components²⁴, we also incorporate leave-one-chromosome-out (LOCO) polygenetic scores (PGS) from LDAK-KVIK⁶ as offsets. SPA_{SQR} follows a retrospective framework as in SPA_{GRM}¹⁸ and thus it is not required to fit a mixed model to control for sample relatedness, which is technically challenging for quantile regressions. The retrospective framework circumvents fitting mixed models by taking relatedness into account when computing the variance of the score statistic. It also enables an easy implementation of saddlepoint approximation since the genotypes are considered as $Binomial(2, \mu)$ random variables, whose moment generating function (MGF) has a closed form expression.

In Step 2, for each genetic variant and quantile τ , we construct score statistics $S = \mathbf{R}(\tau)^T \mathbf{G}$, where $\mathbf{G} = (G_1, G_2, \dots, G_n)^T$ is the genotype vector, n is the number of participants, and $\mathbf{R}(\tau)$ is the residual vector obtained from the null model fitting at quantile τ . More details on the score statistics and residuals can be found in Methods and Supplementary Note. Following the retrospective framework in SPA_{GRM}, SPA_{SQR} treats the genotype vector $\mathbf{G}$ as a multivariate random variable and calibrates p values by incorporating a sparse genetic relationship matrix²¹ (GRM) and saddlepoint approximation (SPA). The original SPA_{GRM} method¹⁸ uses a relatively involved procedure to model relatedness, requiring the computation of Identical-By-Descent sharing probabilities and Chow-Liu trees. SPA_{SQR} greatly simplifies the SPA_{GRM} framework by capitalizing on a score statistic scaling technique, which significantly improved its accessibility and computational efficiency. Using the Cauchy combination (CCT) method^{25,26}, we aggregate the quantile-specific p values to calculate an overall

p value. To simplify notation, when no confusion arises, we use $\mathbf{R} = (R_1, R_2, \dots, R_n)^T$ to denote $\mathbf{R}(\boldsymbol{\tau})$.

To remain scalable for large-scale GWAS, SPA_{SQR} employs a suite of strategies to enhance computational efficiency. The LOCO PGS and sparse GRM can be reused across different quantiles. An efficient and stable quadratic majorization minimization algorithm^{27,28} is employed to fit null smoothed QR models. Similar to prior studies^{2,18,29-33}, SPA_{SQR} adopts a hybrid approach incorporating both regular normal distribution approximation and SPA to calibrate p-values. We also adapt the partial normal distribution approximation idea in the fastSPA algorithm³³ to the context of quantile regression to further accelerate our SPA algorithm.

Simulation Studies

We conducted simulation studies to evaluate type I error rates and statistical powers of SPA_{SQR}. For comparison, we applied REGENIE.QRS¹⁵, an existing QR-based GWAS algorithm, and two methods based on linear regression, REGENIE and LDAK-KVIK. We simulated genotype data based on real genotypes of White British participants from UK Biobank by performing gene-dropping simulations³⁴. Two cohorts were generated: one includes 50,000 unrelated individuals, and another one includes 25,000 unrelated individuals and 25,000 related individuals. To mimic sample relatedness, we generated genotypes for 2,500 families of size 10, following the family structure as in SPA_{GRM}¹⁸. More details on genotype generation can be found in Supplementary Figures 2 and 3.

We simulated 100,000 common variants (minor allele frequency, MAF > 0.05) and 100,000 low-frequency and rare variants (MAF < 0.05 and minor allele count > 10)³⁴, using genotype calls (field ID: 22418) and sequencing data (field ID: 23155), respectively. Following previous studies¹⁵, we simulated quantitative traits covering four scenarios (see Figure 1). The homogeneous errors scenario aligns with a linear model, whereas the other three scenarios, including dominance effects, heteroskedastic

effects, and heterogeneous effects, deviate from the linear model and genetic effects on the phenotypic distribution are quantile-dependent. Causal SNPs were randomly selected from odd-numbered chromosomes. More details can be found in Methods.

In addition to the default SPA_{SQR} method, we evaluated three modified versions including Normal_{QR}, Normals_{SQR}, and SPA_{QR}, for a comprehensive comparison. SPA_{QR} employs a standard non-smooth QR, as in REGENIE.QRS, rather than the convolution-smoothed QR used in SPA_{SQR}. See Figure 2 for an illustration of quantile regression residuals produced by non-smooth QR and smoothed QR. Normal_{QR} and Normals_{SQR} apply a normal distribution approximation for all genetic variants without using SPA to calibrate p-values. For all QR-based GWAS methods, we first calculated quantile-specific p-values at 9 quantiles $\tau = 0.1, 0.2, \dots, 0.9$ and then combined them via Cauchy combination to calculate an overall p value. More details can be found in Methods. We used PLINK2³⁵ to compute a sparse GRM using a threshold of 0.05. In the main text, we used LDAK-KVIK LOCO PGS as offsets for REGENIE.QRS and SPA_{SQR}. We also applied REGENIE LOCO PGS as offsets and the results can be found in Supplementary Notes (Supplementary Tables 2–3 and Supplementary Figures 6–7).

Type I Error Rates

We analyzed genetic variants on even-numbered chromosomes to evaluate type I error rates. For common and low-frequency/rare variants, we conducted more than 2×10^7 tests and evaluated type I error rates at significance levels 5×10^{-7} and 5×10^{-5} (see Table 2, Supplementary Tables 1 and 6–7, Supplementary Figures 4–5).

When testing for common variants, REGENIE shows inflated type I error rates in the presence of sample relatedness, consistent with previous reports^{5,6}. Meanwhile, due to using LDAK-KVIK LOGO PGS as offsets, LDAK-KVIK and REGENIE.QRS exhibit a mild level of deflation, which was also noted in the LDAK-KVIK paper⁶. The

proposed default SPA_{SQR} method, together with SPA_{QR}, Normal_{QR}, and Normal_{SQR}, can accurately control type I error rates.

When testing for low-frequency and rare variants, SPA_{QR} and SPA_{SQR} can still well control type I error rates across all scenarios. Meanwhile, without incorporating SPA, REGENIE.QRS, Normal_{QR}, and Normal_{SQR}, fail in controlling type I error rates when analyzing rare variants, especially at a significance level of 5×10^{-7} . These results are consistent with previous studies, highlighting the necessity of SPA when analyzing rare variants in QR-based GWAS. We note that the type I error inflation of QR methods based on the normal distribution approximation is more severe at extreme quantiles (Supplementary Tables 6-7) such as $\tau = 0.1$ and $\tau = 0.9$. Thus, the type I error inflation in the combined p value is mostly driven by test miscalibration at the low and high quantiles.

Statistical Power

We evaluated statistical powers for common and rare variants. For each scenario, we generated 100 sets of phenotypes and evaluated powers at a genome-wide significance level 5×10^{-8} . Figure 3 illustrates the powers across different scenarios when testing for common variants. Under the homogeneous model scenario, LDAK-KVIK is more powerful than all QR-based methods, which is expected as the scenario aligns with the assumption of linear models. Using regular QR, Normal_{QR} and SPA_{QR} exhibit similar power to REGENIE.QRS, demonstrating that the retrospective framework performs comparably to conventional methods. Meanwhile, using smoothed QR, Normal_{SQR} and SPA_{SQR} substantially boost statistical power, highlighting the advantages of smoothed QR over standard non-smooth QR methods. REGENIE exhibits the lowest power among the methods tested, as it is the only approach that incorporates REGENIE LOCO PGS instead of LDAK-KVIK LOCO PGS, which explains less phenotypic variance. In Supplementary Note, we demonstrate that using REGENIE LOCO PGS also reduces powers for REGENIE.QRS and SPA_{SQR} (Supplementary Figure 8).

Under the other three scenarios in which genetic effects differ across quantiles, QR-based methods always outperform LDAK-KVIK and REGENIE. Using smoothed QR, $\text{Normal}_{\text{SQR}}$ and SPA_{SQR} are more powerful than regular-QR-based methods including REGENIE.QRS, $\text{Normal}_{\text{QR}}$, and SPA_{QR} . Power results when testing for rare variants can be found in Supplementary Figures 9-10. Since methods without incorporating SPA cannot accurately control type I error rates, we primarily compare REGENIE, LDAK-KVIK, SPA_{SQR} and SPA_{QR} and the results are similar to those for common variants, in which QR-based methods are more powerful than linear methods when the genetic effect is heterogeneous and SPA_{SQR} is uniformly more powerful than SPA_{QR} .

UK Biobank Analysis

We applied REGENIE, LDAK-KVIK, REGENIE.QRS, and SPA_{SQR} to analyze 66 quantitative traits in UK Biobank³⁶, comprising $\sim 408,000$ white British participants as a discovery cohort and $\sim 50,000$ self-identified non-British European participants as a replication cohort. For each trait, we excluded participants with missing phenotype data and applied rank-based inverse normal transformation prior to analysis. After quality control, we retained $\sim 630,000$ hard-called variants for LOCO PGS computation and ~ 16.9 million imputed variants for association testing. In Step 1, we incorporated LDAK-KVIK LOCO PGS as an offset for LDAK-KVIK, REGENIE.QRS, and SPA_{SQR} .

For REGENIE.QRS and SPA_{SQR} , we applied Cauchy combination to combine 9 quantile-specific p values into an overall p value as in simulation studies. Of the genetic variants whose discovery p values less than 5×10^{-8} , the variants with validation p values less than 0.05 and consistent effect directions were defined as nominally replicated variants, or simply replicated variants for brevity. The replicated variants were clumped into 1Mb loci, which served as a primary unit for comparison of different methods. Due to the fact that the replication cohort has a relatively small sample size,

following REGENIE.QRS¹⁵ and LDAK-KVIK⁶, we use 0.05 as the replication threshold instead of a Bonferroni-corrected significance level.

SPASQR is more Powerful and Robust than REGENIE.QRS

We grouped the 66 traits into 14 categories and aggregated the number of replicable loci for each category (see Supplementary Table 8). In total, SPASQR and REGENIE.QRS identified 22,559 and 18,793 nominally replicable loci, respectively. Meanwhile, the linear GWAS methods LDAK-KVIK and REGENIE identified 22,836 and 21,450 nominally replicable loci, respectively (see Figure 3a).

SPASQR identified a considerably greater number of loci than REGENIE.QRS, demonstrating the enhanced overall power of the smoothed QR approach. These results are consistent with simulation studies, in which applying smoothed QR can boost powers to detect both homogeneous and heterogeneous genetic effects. As QR-based approaches, both SPASQR and REGENIE.QRS identify genetic loci that are missed by LDAK-KVIK and REGENIE (see Figure 4b). A replicable locus discovered by SPASQR is considered missed by linear methods if all replicated variants of LDAK-KVIK and REGENIE are at least 200kb away from the lead variant. Across the 66 traits analyzed, SPASQR identified 1,618 such loci, while REGENIE.QRS only identified ~1,000 (~62%), demonstrating that SPASQR can substantially enhance the ability to detect genetic variants with heterogeneous effects.

Supplementary Table 9 reports the total number of genome-wide significant variants, stratified by MAF. REGENIE.QRS and Normal_{SQR} produced an abnormally high number of significant associations for rare variants, especially in the validation cohort. These results suggest that many rare variant associations found by REGENIE.QRS and Normal_{SQR} are likely false discoveries, highlighting that SPA is necessary to accurately control type I error rates for rare variants in QR-based GWAS.

Quantile Regression Associations Replicate as well as Linear Regression

We calculate replication rates for all genome-wide significant variants in the discovery cohort across 66 traits. Consistent to previous studies^{9,15}, GWAS signals discovered by QR typically replicate at least as well as linear regression.

Supplementary Figure 14 shows the replication rates for each trait at thresholds of 0.05 and 0.0005. Across the 66 traits, both linear model and QR-based GWAS methods show consistent replication rates. At a threshold of 0.05, the median replication rates were between 0.6 and 0.65 for all methods evaluated. At a more stringent threshold of 0.0005, the median replication rates all reduced to around 0.22. Among all genome-wide significant variants in the discovery cohort, the effect directions of 96.5% were consistent to those in the replication cohort for REGENIE and LDAK-KVIK, 97.3% for REGENIE.QRS, and 96.4% for SPASQR.

Examples of Heterogeneous Associations

To illustrate the performance of SPASQR on real data, we present the GWAS results of glucose, apolipoprotein B, body fat percentage, sitting height, cystatin C, and vitamin D. For each trait, we demonstrate GWAS results for the four methods and highlight up to 10 loci identified by SPASQR but missed by baseline methods. Loci missed by linear regression methods are highlighted in blue, while loci missed by all three baseline methods (including REGENIE.QRS) are highlighted in orange. Results for glucose, apolipoprotein B, and body fat percentage are shown in Figures 5-7. Results for the remaining three traits can be found in Supplementary Figures 15-17. Further details on the lead SNPs of the highlighted loci are given in Supplementary Tables 10–15.

In addition to the Manhattan plots, we provide regional zoom plots for two loci per trait, selected as examples of genetic effects that are heterogeneous across quantiles. For these loci, we visualize the effect sizes $\hat{\gamma}(\tau)$ across quantiles (see Section 3 of

Supplementary Note on effect size estimation). Among the 12 loci featured in regional zoom plots, rs1849554 has been previously reported in a large-scale meta-analysis study for fasting blood glucose³⁷. The remaining 11 loci represent potentially novel findings, as they have not been reported for the same trait in the GWAS catalog³⁸.

Variant rs1849554 (IGF1R, MAF = 0.36) associated with glucose reached genome-wide significance in the discovery cohort only under SPA_{SQR} ($p = 3.6 \times 10^{-8}$), whereas the other three methods all fell short. IGF1R encodes the insulin-like growth factor 1 receptor, a close homolog of the insulin receptor³⁹. Its negative effect is concentrated in the lower quantiles of trait distribution (Figure 5b). The second featured locus of glucose, rs71026083 (OAT, MAF = 0.18, INFO=0.99), is an imputed indel variant that is outside of known glycemic loci. It is GWS under QR methods but not linear methods, with a positive effect also concentrated in the lower quantiles.

For apolipoprotein B, rs4334340 (B4GALNT2, MAF = 0.35) reached genome-wide significance only under SPA_{SQR} ($p = 7.8 \times 10^{-9}$). This signal is in tight LD ($r^2 = 0.796$) with rs2233368 in the neighboring ABI3 ($p = 1.2 \times 10^{-8}$). Figure 6b shows that its negative effect is concentrated in the upper quantiles of the trait distribution. We also highlight a novel rare variant rs751518494 (near OPN1SW, MAF = 2.8×10^{-4}). It is found to have a very large effect size and is GWS only under the QR methods (SPA_{SQR} $p = 3.3 \times 10^{-8}$, REGENIE.QRS $p = 3.2 \times 10^{-8}$) in the discovery cohort. It is worth emphasizing that this rare variant is nominally replicated by all four methods in the replication cohort (SPA_{SQR} $p = 0.011$, REGENIE.QRS $p = 0.024$, REGENIE $p = 8.5 \times 10^{-3}$, LDAK-KVIK $p = 4.7 \times 10^{-3}$).

For body-fat percentage, rs4759276 (LRP1, MAF = 0.39) is GWS under all four methods in the discovery cohort. However, it is only nominally replicated by the QR methods in the replication cohort (SPA_{SQR} $p = 0.036$, REGENIE.QRS $p = 0.034$, REGENIE $p = 0.078$, LDAK-KVIK $p = 0.062$). It demonstrates an effect increasing across quantiles (Figure 7b). LRP1 encodes an endocytic adipocyte receptor

whose adipocyte-specific deletion in mice reduces fat mass and confers resistance to diet-induced obesity⁴⁰. The second featured locus, rs34440469 (near FOXP1, MAF = 0.29), reached genome-wide significance only under SPA_{SQR} ($p = 6.3 \times 10^{-9}$), and likewise showed an effect strengthening toward the upper quantiles (Figure 7b). FOXP1 encodes a transcription factor that controls adipocyte differentiation and thermogenesis⁴¹ and regulates adipose energy metabolism⁴², providing a plausible link to adiposity. Its regional zoom suggests that it is a regulatory hub with numerous spikes in the 1Mb window (Figure 7b).

Computational Efficiency

We benchmarked the computational efficiency of SPA_{SQR} against LDAK-KVIK, REGENIE, and REGENIE.QRS on a machine equipped with an AMD EPYC 9965 processor (192 physical cores, 384 logical cores) and 2 TB of RAM. Using PLINK2, we simulated two PLINK bed files, each containing 10 million variants: one with a sample size of 400,000 and the other with 50,000. Of the 10 million variants, 500,000 were used for LOCO PGS construction. We simulated a quantitative trait with heritability set at 40%, with all causal variants residing within the 500,000-variant set used for LOCO PGS construction.

For LDAK-KVIK, REGENIE, and SPA_{SQR}, we evaluated performance using 1, 4, and 10 threads. Note that REGENIE.QRS does not support multi-threading. Table 3 reports the total wall-clock time and peak memory footprint. The computation time of REGENIE.QRS and SPA_{SQR} in step 1 is not recorded as they are very fast. For step 2, LDAK-KVIK is the fastest when run single-threaded, but REGENIE and SPA_{SQR} are much more effective using multiple threads, whose wall-clock time drops substantially as thread count increases. Supplementary Note Section 4 gives more details about the multi-threading strategy in SPA_{SQR}.

The original R implementation of REGENIE.QRS is not scalable in large-scale data analyses. Based on progress made in 24 hours, the projected wall-clock time to complete the genome-wide analysis at $N = 400,000$ exceeded one week. To make computation feasible, we reimplemented REGENIE.QRS in Python, reading genotypes through the Python library `bed_reader`, which substantially improved its efficiency. Even so, REGENIE.QRS remained the slowest. SPASQR is competitive in both memory footprint and computational efficiency. In particular, with 10 threads it completes Step 2 in 1.15 hours at $N = 400,000$, roughly 16 times as fast as our Python implementation of REGENIE.QRS and over 100 times faster than the original R implementation. SPASQR is also efficient in analyzing multiple traits simultaneously. Supplementary Table 16 demonstrates results when analyzing 10 and 20 traits simultaneously. The computational efficiency of SPASQR can be attributed to several factors, including its advanced C++ implementation parallelizing computation across multiple threads, the adoption of SIMD vectorization, as well as algorithmic innovations such as the partial normal distribution approximation used in the saddle point approximation algorithm (detailed in the Supplementary Note).

Discussion

In this paper, we propose SPASQR, a powerful, robust, and scalable QR-based approach for GWAS of quantitative traits. Instead of fitting standard non-smoothed QR models, SPASQR employs convolution-smoothed QR, exhibiting substantial power gains in both simulation studies and real data analyses. Built on a retrospective framework, SPASQR can effectively address sample relatedness while maintaining accurate type I error control for low-frequency and rare variants. The software implementation is user-friendly and highly computationally efficient, with runtime comparable to widely-used GWAS tools such as REGENIE and LDAK-KVIK, potentially facilitating broader applications of QR-based GWAS.

As a QR-based approach, SPASQR can detect genetic variants with heterogeneous effects across different quantiles of quantitative traits. Such complex effect patterns may arise from gene-environment interactions, epistasis^{43,44}, or the underlying biology of the trait itself. Across the 66 quantitative trait analyses in UK Biobank, SPASQR identified 1,618 loci missed by conventional methods, suggesting that the quantile-dependent genetic effects are widespread yet have largely escaped detection. Even at well-established loci, QR offers additional insights by profiling how the effect size varies across quantiles. We emphasize that QR-based methods are not intended to replace conventional methods, but rather to complement existing GWAS findings.

There are several limitations in SPASQR. As a score test approach within a retrospective framework, SPASQR implicitly assumes Hardy-Weinberg Equilibrium (HWE), that is, genotypes marginally follow a binomial distribution. This assumption is appropriate for relatively homogeneous populations, such as non-British European in real data analyses, but ill-suited for admixed populations or when participants are from multiple ancestries. To extend SPASQR to admixed populations or multi-ancestry data, we plan to employ a strategy similar to SPAmix¹⁶. More specifically, we may estimate individual-specific allele frequencies using genetically derived principal components, instead of using a fixed allele frequency for all individuals.

As a single-variant approach, SPASQR has limited power for rare variants. To boost power for rare-variant associations, gene-level or region-based association analyses offer a promising direction. Widely used approaches include burden tests⁴⁵⁻⁴⁷, SKAT⁴⁸, SKAT-O⁴⁹, ACAT⁵⁰, STAAR⁵¹. An interesting direction is to develop a meta-analysis framework that enables gene-based testing directly from single-variant summary statistics^{49,52-56}, which will also allow for integration of results from multiple cohorts. The Z-scores produced by SPASQR may also be used as the input to the GhostKnockoff^{57,58} framework, which facilitates genome-wide causal variant or loci identification through controlling the false discovery rate (FDR).

Beyond biomarker discovery, another application of QR GWAS is to construct quantile-specific polygenic scores. Conventional PGS for quantitative traits focus on predicting the conditional mean of the phenotype. Quantile-specific association signals, by contrast, could be used to build PGS for a specific quantile of the phenotype. Quantile-specific PGS at extreme quantiles may capture genetic predisposition toward the tails of the trait distribution. For important endophenotypes such as LDL and glucose, such quantile-specific PGS may be of clinical relevance. When combined with conformal prediction⁵⁹, quantile-specific PGS may also be used to construct individualized prediction intervals⁶⁰ with correct nominal coverage.

QR-based GWAS refines the genetic architecture of complex traits by modelling quantiles rather than the mean value of quantitative traits. Despite its promise as a standard tool in genetic analysis, applying QR methods to large-scale GWAS remains technically challenging, particularly when addressing genetic confounders such as sample relatedness. Recently, retrospective frameworks¹⁶⁻²⁰ have demonstrated robustness and generality in applying statistically powerful methods to genetic studies. In this paper, we validate the unique advantage of the framework in terms of statistical power, robustness to confounders, and scalability to large-scale biobank data analyses. Looking ahead, we expect that the retrospective framework will continue serving as an important bridge between statistics and genetics.

Methods

Regular Quantile Regression and Rank-score Inference

We first briefly review the rank score inference of Wang et al⁹, in which QR assumes that τ quantile of response Y_i , i.e. $Q_{Y_i|X_i,G_i}(\tau)$, depends linearly on the covariates X_i and the genotype G_i as below,

$$Q_{Y_i|X_i,G_i}(\tau) = X_i^\top \boldsymbol{\beta}(\tau) + G_i \gamma(\tau).$$

399 Regular quantile regression utilizes a non-smoothed pinball loss $\rho_\tau(r) = \frac{1}{2}|r| + (\tau -$
400 $\frac{1}{2})r$. To construct a score statistic at quantile level τ , the following loss function
401 computes the null model:

$$402 \quad f_\tau(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^n \rho_\tau(Y_i - X_i^\top \boldsymbol{\beta}).$$

403 If we use $\tilde{\boldsymbol{\beta}}(\tau)$ to denote the optimal solution to the above non-smooth quantile
404 regression problem, a score statistic for quantile level τ is then:

$$405 \quad S_\tau = \sum_{i=1}^n \tilde{G}_i \left(\tau - I \left(e_i \left(\tilde{\boldsymbol{\beta}}(\tau) \right) < 0 \right) \right),$$

406 where $e_i \left(\tilde{\boldsymbol{\beta}}(\tau) \right) = Y_i - X_i^\top \tilde{\boldsymbol{\beta}}(\tau)$ and $\tilde{\mathbf{G}} = (\mathbf{I} - \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top) \mathbf{G}$ is the genotype
407 vector projected onto the null space of the covariates $\mathbf{X}$. We call $R_i = \tau -$
408 $I \left(e_i \left(\tilde{\boldsymbol{\beta}}(\tau) \right) < 0 \right)$ non-smooth quantile regression residuals.

409 Note that R_i takes only two possible values: τ and $\tau - 1$. The values τ and
410 $\tau - 1$ occur in proportions $1 - \tau$ and τ (see Figure 2), respectively. The asymptotic
411 variance of S_τ is given by

$$412 \quad V_\tau = \tau(1 - \tau) \tilde{\mathbf{G}}^\top \tilde{\mathbf{G}}.$$

413 Under the null hypothesis of no genetic effects, $\frac{S_\tau}{\sqrt{V_\tau}}$ is asymptotically standard
414 normal⁶¹, so that we may obtain p values for the hypothesis $\gamma(\tau) = 0$ using the
415 cumulative distribution function (CDF) of a normal distribution.

416 To aggregate the statistical evidence across different quantiles, the common
417 practice is to combine the p values using Cauchy combination¹¹:

$$p_{cct} = \frac{1}{2} - \frac{1}{\pi} \arctan \left(\frac{1}{d} \sum_{i=1}^d \tan \{ (0.5 - p_i) \pi \} \right),$$

where p_i are the p values at different quantiles. The combined p values for each genetic variant are the primary genome-wide association results.

Smoothed Quantile Regression

Recent statistical literature proposes a convolution-smoothed alternative¹²⁻¹³ to the traditional non-smooth quantile regression model. When fitting the null model, convolution-smoothed quantile regression minimizes the following loss function:

$$f_{h,\tau}(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^n \ell_{h,\tau}(Y_i - X_i^\top \boldsymbol{\beta}),$$

where $\ell_{h,\tau}(u) = \int_{-\infty}^{+\infty} \rho_\tau(v) K_h(v - u) dv$. K_h is a kernel density function that integrates to 1. h is a bandwidth parameter that controls the proximity of the convolution smoothed loss to the original non-smooth pinball loss. As h goes to 0, the convolution-smoothed pinball loss converges to the original non-smooth pinball loss. See supplementary Figure 1 for a visualization of the smoothed QR loss function. In this work we focus on the Gaussian kernel $K_h(u) = \frac{1}{\sqrt{2\pi}h} e^{-\frac{u^2}{2h^2}}$.

If we use $\tilde{\boldsymbol{\beta}}_h(\tau)$ to denote the optimal solution to the above smooth quantile regression problem, then a score statistic for the hypothesis $\gamma(\tau) = 0$ is

$$S = \sum_{i=1}^n G_i \left[\tau - \mathcal{K} \left(-\frac{e_i(\tilde{\boldsymbol{\beta}}_h(\tau))}{h} \right) \right],$$

where $\mathcal{K}$ is the cumulative distribution function of a standard normal distribution. For convenience, we also denote $\left[\tau - \mathcal{K} \left(-\frac{e_i(\tilde{\boldsymbol{\beta}}_h(\tau))}{h} \right) \right]$ as R_i and refer to them as smoothed quantile regression residuals. Unlike non-smooth quantile regression

residuals which take two discrete values, smoothed quantile regression residuals range continuously in $[\tau - 1, \tau]$, with most values still cluster near the two extremes (See Figure 2 for an illustration). The reason why smoothed quantile regression improves power over non-smooth quantile regression can be understood as an example of bias–variance trade-off. Smoothing introduces a small bias—the smoothed loss no longer targets the quantile τ exactly—but in return it substantially reduces the variance of the score statistic. When the bandwidth is chosen appropriately, this variance reduction outweighs the induced bias, concentrating the test statistic and thereby increasing power. Another intuitive explanation is that under smoothed QR, the quantile regression residuals retain more fine-grained information. The residuals are no longer a binary assignment between τ and $\tau - 1$, and the magnitude of the residuals can be interpreted as the model’s confidence in choosing between τ and $\tau - 1$ for a particular participant.

We note that the estimation bias in β does not alter the null center of the score statistic, which still remains at zero, so that the bandwidth choice bears no tangible impact on type I error control (Supplementary Tables 4-5). However, the bandwidth h does influence the power of the score tests for detecting true signals to some extent (Supplementary Figures 11-12).

Score Tests using Normal Distribution Approximation

In a retrospective framework, assuming unrelated individuals, we model the distribution of G_i as independently distributed $\text{Binom}(2, \mu)$, where μ is the alternative allele frequency. Under the null hypothesis, we assume that the distribution of G_i conditional on the quantile regression residuals $\mathbf{R}$ remains unchanged, i.e. $G_i | \mathbf{R} \sim i.i.d. \text{Binom}(2, \mu)$. The null expectation and variance of the score statistic S are then given by

463

$$E(S) = 2 \sum_{i=1}^n R_i \cdot \mu,$$

464

$$Var(S) = \left(\sum_{i=1}^n R_i^2 \right) \cdot 2\mu(1 - \mu).$$

465

Notably, provided the null model includes an intercept term, we have $\sum_{i=1}^n R_i = 0$. The

466

term $2\mu(1 - \mu)$ reflects our assumption that G_i follows a $Binom(2, \mu)$ distribution.

467

However, this variance estimate may be inaccurate if the variant significantly deviate

468

from the Hardy-Weinberg Equilibrium or if we impute a substantial proportion of

469

missing variants using the column mean. Thus, we substitute $Var(S)$ by

470

$$Var(S) = \left(\sum_{i=1}^n R_i^2 \right) \cdot \frac{1}{n} \sum_{i=1}^n (G_i - \bar{G})^2,$$

471

where $\bar{G}$ is the mean of the variant. Invoking a normal approximation, the association

472

p value can be computed as

473

$$p_{norm} = 2\mathcal{K}\left(-\frac{|s|}{\sqrt{Var(S)}}\right),$$

474

where s is the observed score statistic. We refer to the above normal-approximation-

475

based score test using smoothed quantile regression residuals as $Normal_{\text{SQR}}$. When the

476

non-smooth rank scores $R_i = \tau - I\left(e_i\left(\tilde{\beta}(\tau)\right) < 0\right)$ are used instead to compute the

477

score statistic and the null variance, we call the resulting test $Normal_{\text{QR}}$. Notice that

478

$Normal_{\text{QR}}$ is in fact the same test as the rank score inference if the covariates only

479

include an intercept term, since $\mathbf{G} - \bar{\mathbf{G}}$ is simply $\mathbf{G}$ projected onto the null space of

480

$\mathbf{1}_n$ and for non-smooth quantile regression, $\sum_{i=1}^n R_i^2 \approx n\tau(1 - \tau)$.

481

Moment Generating Function of Score Statistics

It is well known that the null distribution of a score statistic is not always well approximated by a normal approximation, especially when testing for rare variants. To address this issue, we use saddlepoint approximation⁶² (SPA) to more accurately calibrate p values. Saddlepoint approximation relies upon the moment generating function (MGF) to compute p-values. In what follows, the residuals $\mathbf{R}$ can be produced by a smooth quantile regression null model or a classic non-smooth quantile regression model.

Recall that the score statistic is $S = \sum_{i=1}^n R_i G_i$. Treating G_i as i.i.d. $\text{Binom}(2, \mu)$ random variables and R_i as fixed coefficients, the MGF of S can be computed as

$$M_S(t) = \prod_{i=1}^n [1 - \mu + \mu e^{R_i t}]^2,$$

so that the cumulant generating function (CGF) can be expressed as

$$K_S(t) = \log M_S(t) = \sum_{i=1}^n 2 \log(1 - \mu + \mu e^{R_i t}).$$

The above CGF can be numerically expensive to evaluate due to the exp and log operations. To enable efficient computation, we turn to the partial normal approximation idea of fastSPA³³. See section 1.1 of the Supplementary Note for more information.

Saddlepoint approximation to calibrate p-values

After obtaining the CGF, we use saddle point approximation to calibrate p values. The CGF $K_S(t)$ encodes variance $\text{Var}_{CGF}(S) = K_S''(0) = (\sum_{i=1}^n R_i^2) \cdot 2\mu(1 - \mu)$. To ensure consistency with our normal-approximation-based tests, we scale the observed score statistic as $s_{adj} = s \cdot \sqrt{\text{Var}_{CGF}(S)} / \sqrt{\text{Var}(S)}$, where $\text{Var}(S) = \frac{1}{n} \sum_{i=1}^n (G_i -$

504 $\bar{G})^2 \cdot R^\top R$. We then use Newton's algorithm to find root ζ such that $K'_S(\zeta) = s_{adj}$.
 505 Next, we compute auxiliary statistics $\omega = \text{sgn}(\zeta)\sqrt{2(\zeta s_{adj} - K_S(\zeta))}$ and $\nu =$
 506 $\zeta\sqrt{K''_S(\zeta)}$. According to the Barndorff-Nielsen method⁶³, the cumulative distribution
 507 function of S at s is approximated by

$$508 \quad P(S < s) \approx \mathcal{K}\left(\omega + \frac{1}{\omega} \log\left(\frac{\nu}{\omega}\right)\right),$$

509 and a two-sided p-value for the score test is given by $P(S < -|s|) + 1 - P(S < |s|)$.
 510 To further accelerate computation, we adopt a hybrid strategy: when $|s| < r\sqrt{\text{Var}(S)}$,
 511 we use the normal approximation to compute p values; otherwise we employ SPA as
 512 described above. In this work, we set $r = 2$.

513 Supplementary Note section 1.2 gives the complete workflow of SPA_{SQR}.
 514 Depending on the choice of quantile regression residuals, we obtain two variants of the
 515 saddlepoint approximation test. When smoothed quantile regression residuals are used
 516 for $\mathbf{R}$, we refer to the method as SPA_{SQR}. When the non-smooth quantile regression
 517 residuals are used instead, we refer to the method as SPA_{QR}.

518 Using LOCO PGS as an Offset

519 The tests described above rely on the key assumption that, independent to the quantile
 520 regression residuals $\mathbf{R}$, the distributions of G_i are independently distributed
 521 $\text{Binom}(2, \mu)$. This assumption may be severely violated in the presence of relatedness,
 522 since G_i themselves are no longer independent. However, notice that the null variance
 523 of S can be decomposed as

$$524 \quad \text{Var}(S) = \sum_{i=1}^n R_i^2 \text{Var}(G_i) + \sum_{i \neq j} R_i R_j \text{Cov}(G_i, G_j).$$

If $R_i R_j$ have randomly flipped signs among related pairs, then $\sum_{i \neq j} R_i R_j \text{Cov}(G_i, G_j) \approx 0$ through cancellation. In other words, to restore the validity of the test, it suffices to ensure that the covariance structure of $\mathbf{R}$ is independent of that of $\mathbf{G}$. This is commonly achieved by modeling relatedness as a random effect with covariance structure proportional to the genetic relationship matrix Φ in linear mixed-models^{64,65}. However, mixed-model extensions^{66,67} of quantile regression typically involve relatively complex computation so that they may not be scalable enough for modern biobank applications.

This motivates the use of leave-one-chromosome-out (LOCO) polygenic scores as an offset, which is a standard technique in modern GWAS algorithms^{1,4-6,18,65,68}. For example, the popular GWAS software REGENIE employs stacked ridge regression to build a predictive model for the trait using all chromosomes except the chromosome being tested. The resulting predictions, referred to as LOCO predictions, are then used as an offset in step 2 of REGENIE to enhance association power and reduce inflation due to relatedness. LDAK-KVIV constructs LOCO PGS under an elastic-net prior, fitted via a novel variational Bayes solver. LOCO PGS reduces confounding from relatedness by removing a substantial portion of the polygenic background. It also improves power because the LOCO predictions capture a portion of trait heritability, thereby improving the signal-to-noise ratio for variants on the chromosome under testing. REGENIE.QRS adopted a similar idea, using LOCO PGS from REGENIE as an offset to improve power and reduce inflation in QR association studies. Here, we also adopt this established practice. We use LOCO PGS from LDAK-KVIV as an offset in the null model, since LDAK-KVIV provides the most accurate LOCO PGS to date. More concretely, we subtract the LOCO PGS for chromosome c from the trait values before testing variants on chromosome c . We recommend applying rank-based inverse normal transformation (INT) prior to PGS construction and prior to association testing.

In our UKB analysis, association power is typically greater under INT (Supplementary Figure 13).

Using sparse GRM to adjust for sample relatedness

While the LOCO scheme can serve as an approximation to the full mixed-model, in practice we find that the accuracy of LOCO PGS substantially affects test calibration in the presence of strong relatedness. On the related data set constructed for our simulations, using LOCO PGS from REGENIE as an offset does not fully control for relatedness and results in type I error inflation (Supplementary Figure 6). On the other hand, using LOCO PGS from LDAK-KVIK as an offset induces noticeable deflation in type-I error (Supplementary Table 1). To protect against either inflation or deflation, SPA_{SQR} optionally incorporates a sparse genetic relationship matrix (GRM) into the null variance, following our prior work SPA_{GRM}¹⁸. More specifically, assuming that the marginal distribution of G_i are identical, the variance of the score statistic is

$$Var(S) = \sum_{i=1}^n R_i^2 \sigma_g^2 + \sum_{i \neq j} R_i R_j \sigma_g^2 \rho_{ij} = \sigma_g^2 \mathbf{R}^\top \mathbf{\Phi} \mathbf{R},$$

where ρ_{ij} is (i, j) -th entry of the genetic correlation matrix $\mathbf{\Phi}$. We may estimate σ_g^2 as $\frac{1}{n} \sum_{i=1}^n (G_i - \bar{G})^2$ and use PLINK2 to compute a sparse GRM in which only genetic correlation coefficients above a chosen threshold (such as 0.05) are retained. The resulting variance estimate $Var(S) = \frac{1}{n} \sum_{i=1}^n (G_i - \bar{G})^2 \mathbf{R}^\top \mathbf{\Phi} \mathbf{R}$ is then used in the test statistic scaling step $s_{adj} = s \cdot \sqrt{Var_{CGF}(S)} / \sqrt{Var(S)}$ during saddlepoint approximation. We note that for cohorts with relatively low levels of relatedness, such as the UK Biobank, the two versions of $Var(S)$ yield comparable GWAS results; the GRM-aware variance is therefore recommended primarily for cohorts with substantial family structure.

Smoothing Bandwidth h

Smoothed quantile regression requires the specification of a bandwidth parameter, h . When h is very small, smoothed quantile regression behaves overwhelmingly similar to non-smooth quantile regression. To fully realize the benefits of variance reduction in quantile smoothing, it is generally advisable to choose a non-negligible value for h . On the other hand, if h is very large, then our GWAS results will mostly resemble linear GWAS results. In this sense, SPASQR interpolates between GWAS based on linear regression and GWAS based on non-smooth quantile regression.

In our experiments, we find that while the choice of h has no tangible impact on controlling Type I error, it affects the power of our score tests to some extent (Supplementary Figures 11-12). After extensive experiments, we find that $h = IQR/3$ performs robustly across a wide range of scenarios, where IQR is the interquartile range of $\tilde{\mathbf{Y}} - \hat{\mathbf{Y}}_{-c}$, $\tilde{\mathbf{Y}}$ is the trait after INT, $\hat{\mathbf{Y}}_{-c}$ is the LOCO PGS for chromosome c . This is the default bandwidth for SPASQR. A larger bandwidth, $h = IQR/2$ yields power comparable to that of $h = IQR/3$, and performed similarly in our real data analysis. $h = IQR/5$ is another reasonable choice, but it produced significantly fewer associations than $h = IQR/3$ or $h = IQR/2$ in our UKB analysis. We therefore recommend using $h = IQR/3$ or $h = IQR/2$ in practice.

Data Simulation

When generating the phenotype, we consider the following four simulation settings, which largely mirrors the phenotype generation protocol of REGENIE.QRS¹⁵. These settings are designed to mimic genetic effects that go beyond a simple shift in the mean. In all settings, G_{il} denotes standardized genotypes of 1000 causal variants with additive effects, and X_i is a standard normal covariate with coefficient β .

For the homogeneous model, we randomly selected 1000 causal SNPs from the odd-numbered chromosomes, and statistical power was defined as the proportion of these causal SNPs that were successfully identified. On the other hand, for the dominance, heteroskedastic, and heterogeneous models, we selected 100 special SNPs out of the 1000 causal SNPs to also exhibit dominance, heteroskedastic, or heterogeneous effects, respectively, and focus on power in detecting those model-specific causal SNPs with complex effect patterns.

Homogeneous effects Phenotypes are generated as

$$Y_i = X_i\beta + \sum_{l=1}^{1000} G_{il} \gamma_l + e_i,$$

where $\gamma_l \sim \mathcal{N}(0, 0.4/1000)$, $\beta \sim \mathcal{N}(0, 0.3)$, and $e_i \sim \mathcal{N}(0, 0.3)$.

Dominance effects A subset of 100 SNPs has dominance effects in addition to mean effects:

$$Y_i = X_i\beta + \sum_{l=1}^{1000} G_{il} \gamma_l + \sum_{m=1}^{100} \tilde{G}_{im} \gamma_m + e_i,$$

where $\tilde{G}_{im}$ denotes standardized genotypes of the 100 special SNPs under dominance encoding $(-\mu/(1-\mu), 1, -(1-\mu)/\mu)$. Here $\beta \sim \mathcal{N}(0, 0.2)$, $\gamma_l \sim \mathcal{N}(0, 0.3/1000)$, $\gamma_m \sim \mathcal{N}(0, 0.2/100)$, and $e_i \sim \mathcal{N}(0, 0.2)$.

Heteroskedastic effects A subset of 100 SNPs not only affects phenotype mean, but also phenotypic variability:

$$Y_i = X_i\beta + \sum_{l=1}^{1000} G_{il} \gamma_l + e_i \cdot \left(1 + \sum_{m=1}^{100} \tilde{G}_{im} \gamma_m\right),$$

617 where $\tilde{G}_{im}$ denotes standardized genotypes for the 100 SNPs that induce
 618 heteroskedasticity. $\beta \sim \mathcal{N}(0,0.1)$, $\gamma_l \sim \mathcal{N}(0,0.2/500)$, $\gamma_m \sim \mathcal{N}(0,0.5/50)$, and
 619 $e_i \sim \mathcal{N}(0,0.3)$.

620 **Heterogeneous effects** A subset of 100 SNPs has effects are active only at upper
 621 quantiles:

$$622 \quad Y_i = X_i\beta + \sum_{l=1}^{1000} G_{il} \gamma_l + \mathbf{1}(u_i > 0.7) \cdot u_i \cdot \sum_{m=1}^{100} \tilde{G}_{im} \gamma_m + e_i,$$

623 where $u_i \sim \text{Uniform}(0,1)$ represents the conditional quantile of individual i , and
 624 $\tilde{G}_{im}$ denotes standardized genotypes for 100 SNPs whose effects act only when $u_i >$
 625 0.7 . Here $\beta \sim \mathcal{N}(0,0.1)$, $\gamma_l \sim \mathcal{N}(0,0.3/500)$, $\gamma_m \sim \mathcal{N}(0,0.4/50)$, and the error
 626 term is $e_i = Q^{-1}(u_i)$ with Q^{-1} being the inverse CDF of $\mathcal{N}(0,0.2)$.

627 **GWAS of 66 Quantitative Traits in the UK Biobank**

628 Directly genotyped array variants were used to compute leave-one-chromosome-out
 629 (LOCO) polygenic scores (PGS) for REGENIE and LDAK-KVIK, whereas association
 630 testing was performed on imputed variants. Using PLINK2, we excluded variants with
 631 a genotype missingness rate greater than 10%, a minor allele frequency (MAF) below
 632 10^{-4} , or a Hardy–Weinberg equilibrium p-value below 10^{-6} ; imputed variants were
 633 additionally required to have an INFO score of at least 0.8. After quality control,
 634 629,374 directly genotyped autosomal variants remained for LOCO PGS construction
 635 and 16,852,077 imputed variants for association testing, with the number of individuals
 636 carrying a non-missing phenotype ranging from 352,829 to 407,800 across traits. On
 637 the validation cohort (~ 50000 unrelated individuals), we applied an additional $\text{MAF} \geq$
 638 10^{-4} filter (recomputed within the validation data), leaving 16,355,878 imputed variants
 639 for association testing; non-missing phenotype sample sizes ranged from 43,162 to
 640 50,067 across traits.

For the QR-based GWAS methods (REGENIE.QRS and SPA_{SQR}), we considered nine quantile levels, $\tau = 0.1, 0.2, \dots, 0.9$. For SPA_{SQR}, the smoothing bandwidth was set to $h = \text{IQR}(\tilde{Y} - \hat{Y}_{-c})/3$, where $\tilde{Y}$ is the trait after rank-based inverse normal transformation and $\hat{Y}_{-c}$ is the LOCO PGS for chromosome c . All association analyses adjusted for an intercept, age, sex, age², age×sex, age²×sex, and the top 10 principal components (PCs). A sparse GRM computed by PLINK2 using variants with MAF at least 0.05 and genetic correlation cut-off 0.05 is supplied to GRAB for the discovery data set. No GRM is computed for the replication data set.

A variant is considered nominally replicated if its p-value is less than 5×10^{-8} in the discovery cohort, less than 0.05 in the validation cohort, and had consistent effect directions between discovery and validation. For QR-based methods, we focused on the quantile with the smallest discovery p value to check for direction consistency. We clumped (nominally) replicated variants of each method into 1Mb loci using PLINK2 (`--clump-p1  $5 \times 10^{-8}$ , --clump-kb 500, --clump-r2 0`). We define that SPA_{SQR} have identified a locus missed by another method if all replicated variants of that method is at least 200kb away from the lead variant of the SPA_{SQR}. A locus is considered missed by linear methods if all LDAK-KVIK and REGENIE replicated variants are at least 200kb away.

Data Availability

The individual-level genotype and phenotype data are available through formal application to the UK Biobank (<http://www.ukbiobank.ac.uk>). GWAS summary statistics for 66 quantitative traits in both the discovery and the replication cohort are available on Zenodo at <https://doi.org/10.5281/zenodo.20944191>⁶⁹.

Code Availability

Software mentioned in this paper are available at the following sites

- 666 SPA_{SQR}: <https://qhengncsu.github.io/SPAsqr.github.io/>
- 667 REGENIE.QRS: <https://github.com/FanWang0216/Regenie.QRS>
- 668 LDAK-KVIK: <https://www.ldak-kvik.com>
- 669 REGENIE: <https://rgcgithub.github.io/regenie/>
- 670 PLINK2: <https://www.cog-genomics.org/plink/2.0/>

Table 1: A comparison between SPASQR and existing QR GWAS methods. Key methodological features across three QR-based GWAS approaches.

Feature	QRS ⁹	REGENIE.QRS ¹⁵	SPASQR
Control for sample relatedness	✗	✓	✓
Accurate for analyzing rare variants	✗	✗	✓
Use LOCO PGS to improve power	✗	✓	✓
Use smoothed QR to improve power	✗	✗	✓
Highly efficient at biobank scale	✗	✗	✓

Table 2: Type I error inflation ratio at significance level 5×10^{-7} for REGENIE, LDAK-KVIV, REGENIE.QRS, Normal_{QR}, SPA_{QR}, Normal_{SQR}, and SPA_{SQR} using more than 2×10^7 null tests. LDAK-KVIK LOCO PGS are used as an offset for QR GWAS methods. Type I error inflation ratio is defined as type I error divided by nominal significance level. Normal_{SQR} and SPA_{SQR} uses the bandwidth $h = \text{IQR}(\tilde{\mathbf{Y}} - \hat{\mathbf{Y}}_{-c})/3$. We consider two cohorts. The unrelated cohort consists of 50000 unrelated individuals, while the related cohort consists of 25000 unrelated individuals and 25000 related individuals. Common variants have an MAF in the range (0.05, 0.5), while rare variants have an MAF in the range (1e-4, 0.05). Bold numbers represent uncontrolled type I error rates (inflation ratio greater than 2).

	Variant	Method	Homogeneous	Dominance	Heteroskedastic	Heterogeneous
Unrelated	Common	REGENIE	1.21	0.96	1.13	1.45
		LDAK-KVIK	1.77	1.13	1.21	1.29
		REGENIE.QRS	1.13	0.80	0.72	1.29
		Normal _{QR}	1.13	0.88	1.45	0.88
		SPA _{QR}	1.13	0.88	1.45	0.88
		Normal _{SQR}	1.45	1.05	0.96	1.21
		SPA _{SQR}	1.37	1.05	0.96	1.21
	Rare	REGENIE	1.42	2.08	0.94	1.04
		LDAK-KVIK	1.13	1.79	1.89	1.79
		REGENIE.QRS	19.64	19.35	21.62	19.07
		Normal _{QR}	23.19	22.43	24.43	20.62
		SPA _{QR}	0.94	0.94	1.51	1.60
		Normal _{SQR}	24.92	23.98	24.83	19.92
		SPA _{SQR}	0.76	0.38	0.19	0.47
Related	Common	REGENIE	9.24	6.18	8.11	4.90
		LDAK-KVIK	0.72	0.72	0.40	0.96
		REGENIE.QRS	0.88	0.64	1.04	1.20
		Normal _{QR}	1.12	0.80	0.72	0.88
		SPA _{QR}	1.04	0.80	0.72	0.88
		Normal _{SQR}	0.48	1.12	0.80	1.04
		SPA _{SQR}	0.48	1.12	0.80	1.04
	Rare	REGENIE	10.73	6.50	8.47	8.19
		LDAK-KVIK	0.56	0.75	0.75	0.66
		REGENIE.QRS	13.09	17.32	17.13	22.59
		Normal _{QR}	13.75	17.06	15.27	18.11
		SPA _{QR}	0.47	0.94	0.47	0.56
		Normal _{SQR}	16.57	21.94	22.41	20.24
		SPA _{SQR}	0.38	0.85	0.38	0.47

Table 3. Computational efficiency and memory footprint of four GWAS methods in analyzing a single simulated trait using different numbers of threads. We used the linux command `command /usr/bin/time -v` to record the total wall clock time and peak memory footprint. Experiments were run on a machine equipped with an AMD EPYC 9965 processor (192 physical cores, 384 logical cores) and 2 TB of RAM.

Method	Step	Threads	N	Wall (h)	Peak memory (GB)
REGENIE	Step 1	1	400,000	7.76	14.4
			50,000	1.42	2.0
		4	400,000	3.53	14.3
			50,000	0.91	2.0
		10	400,000	1.96	14.3
			50,000	0.61	2.0
	Step 2	1	400,000	17.49	5.3
			50,000	1.98	3.6
		4	400,000	5.15	5.4
			50,000	0.41	3.6
		10	400,000	2.75	5.4
			50,000	0.29	3.6
LDAK-KVIK	Step 1	1	400,000	2.58	4.4
			50,000	0.56	1.6
		4	400,000	2.20	4.4
			50,000	0.51	1.6
		10	400,000	2.22	4.4
			50,000	0.39	1.5
	Step 2	1	400,000	2.48	6.6
			50,000	0.50	3.8
		4	400,000	1.42	6.7
			50,000	0.40	3.8
		10	400,000	1.43	6.7
			50,000	0.38	3.8
REGENIE.QRS (R)	Step 2	1	400,000	>168	6.24
			50,000	>24	0.84
REGENIE.QRS (Python)	Step 2	1	400,000	18.78	3.9
			50,000	2.24	1.2
SPA_{SQR}	Step 2	1	400,000	9.23	3.1
			50,000	1.13	2.8
		4	400,000	2.51	3.5
			50,000	0.30	2.8
		10	400,000	1.15	4.2
			50,000	0.17	2.9

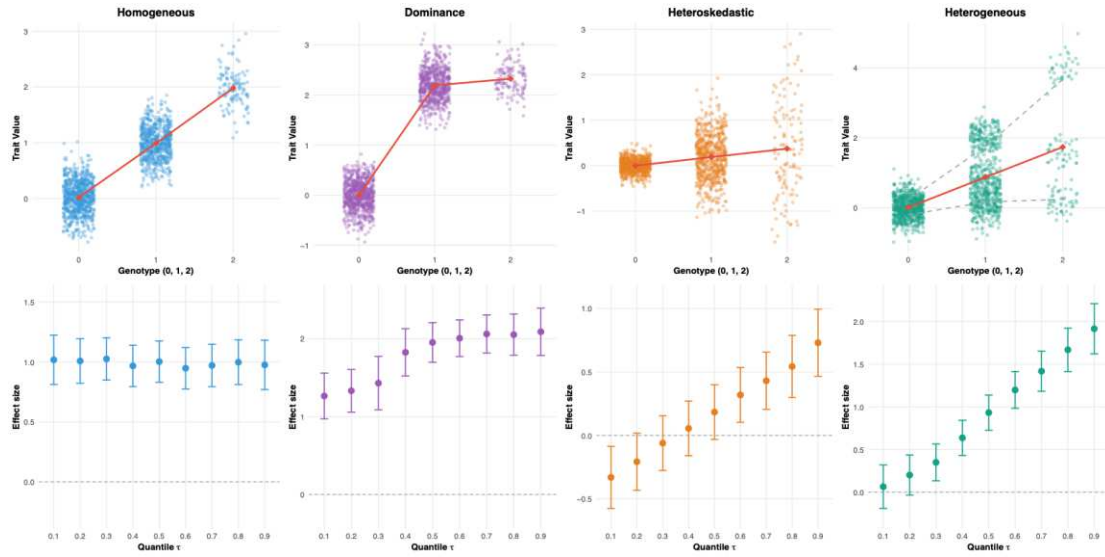


Fig. 1: Complex genetic effects that deviate from a simple linear model. The top row displays trait values across three genotype groups (0, 1, 2 minor alleles); red lines connect genotype means. **a. Homogeneous:** classical additive effects with constant within-group variance, the ideal setting for mean-based GWAS. **b. Dominance:** heterozygotes resemble minor-allele homozygotes rather than lying halfway, violating the additive assumption. **c. Heteroskedastic:** the mean effect is small, but trait variance grows with minor allele count. **d. Heterogeneous:** distinct subgroups (dashed lines) show divergent genotype effects, yielding a modest overall mean (solid line) that masks strong subgroup-specific effect. The bottom row illustrates how these genetic effects translate into varying effect sizes across different quantiles of the phenotype distribution.

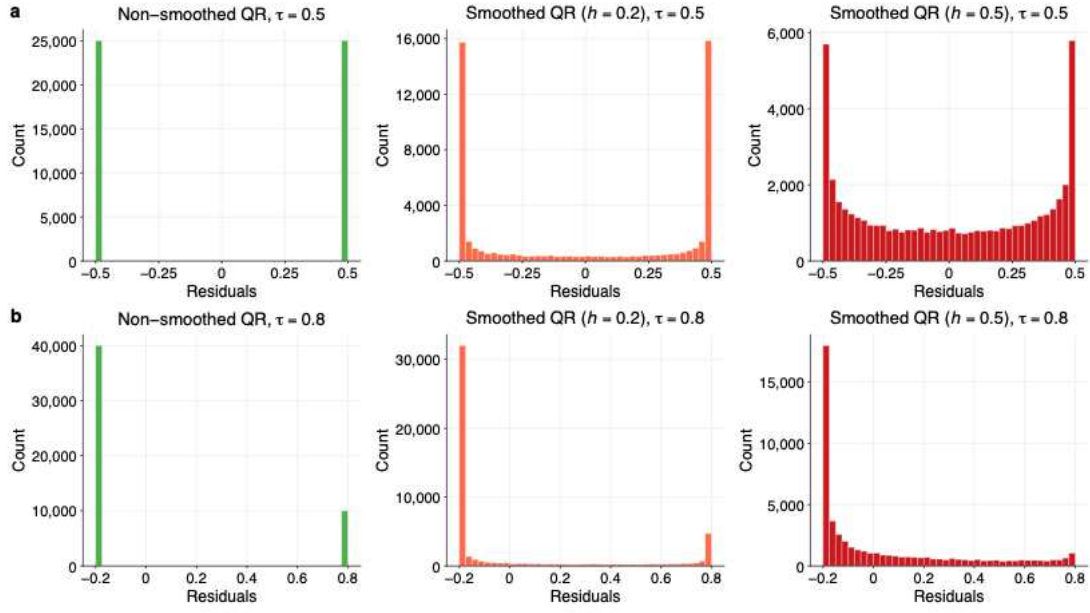


Fig. 2: Histograms of the residual R_i for a simulated trait ($n = 50,000$), for both non-smooth and smoothed quantile regression using bandwidths $h = 0.2, 0.5$. For non-smooth quantile regression, R_i takes two discrete values, $\tau - 1$ and τ ; for smooth quantile regression, R_i ranges continuously in $[\tau - 1, \tau]$. **a.** histograms of quantile regression residuals at $\tau = 0.5$. **b.** histograms of quantile regression residuals at $\tau = 0.8$.

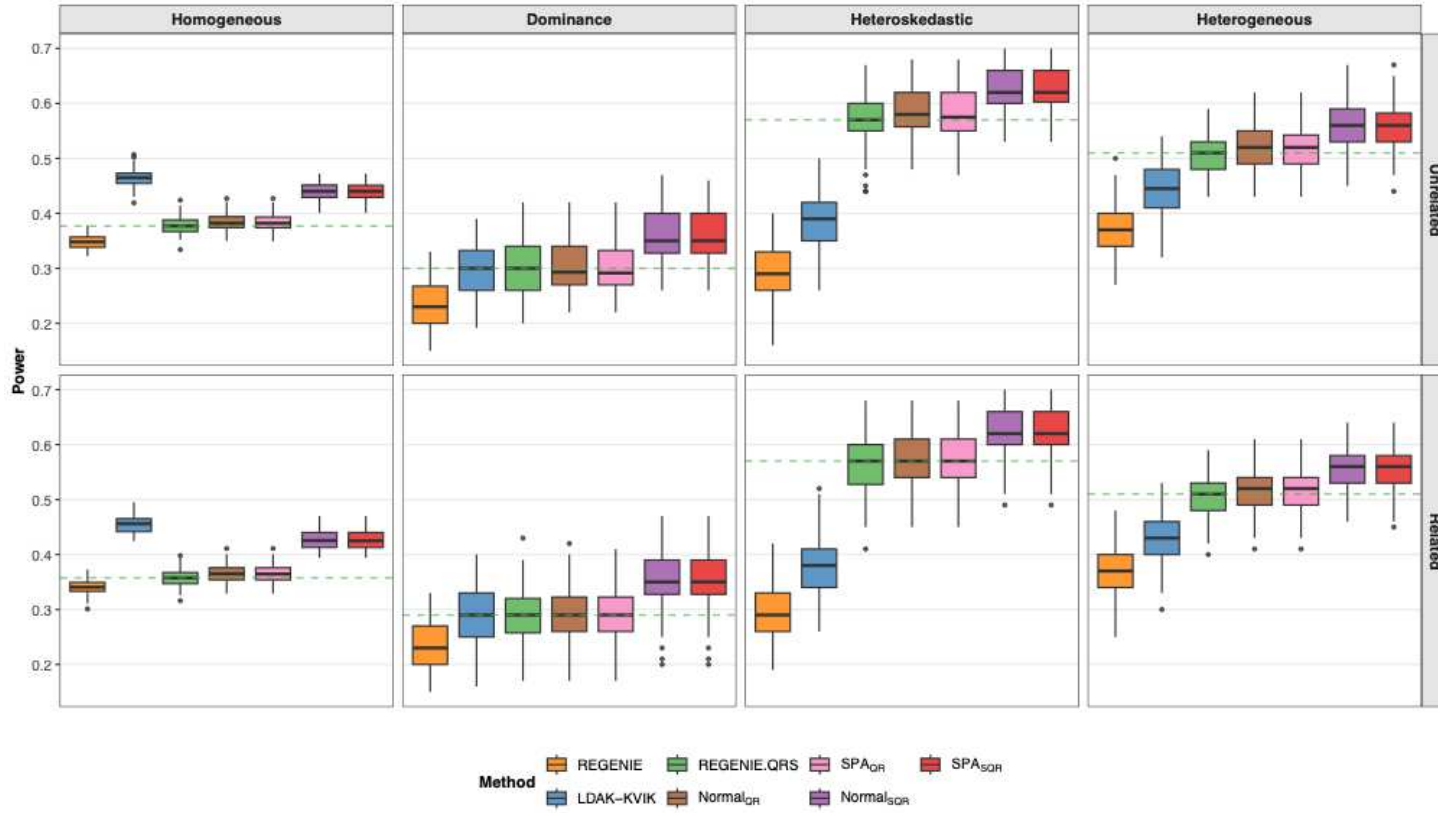


Fig. 3: Power box plots (common variants, $MAF > 0.05$) at significance level 5×10^{-8} for REGENIE, LDAK-KVIK, REGENIE.QRS, Normal_{QR}, SPA_{QR}, Normal_{SQR}, and SPA_{SQR} across 100 simulations. Per simulation, 1,000 causal common variants were tested under Homogeneous and 100 under the other three scenarios. Normal_{SQR} and SPA_{SQR} uses bandwidth $h = \text{IQR}(\tilde{\mathbf{Y}} - \hat{\mathbf{Y}}_c)/3$. The dashed green line marks the median power of REGENIE.QRS.

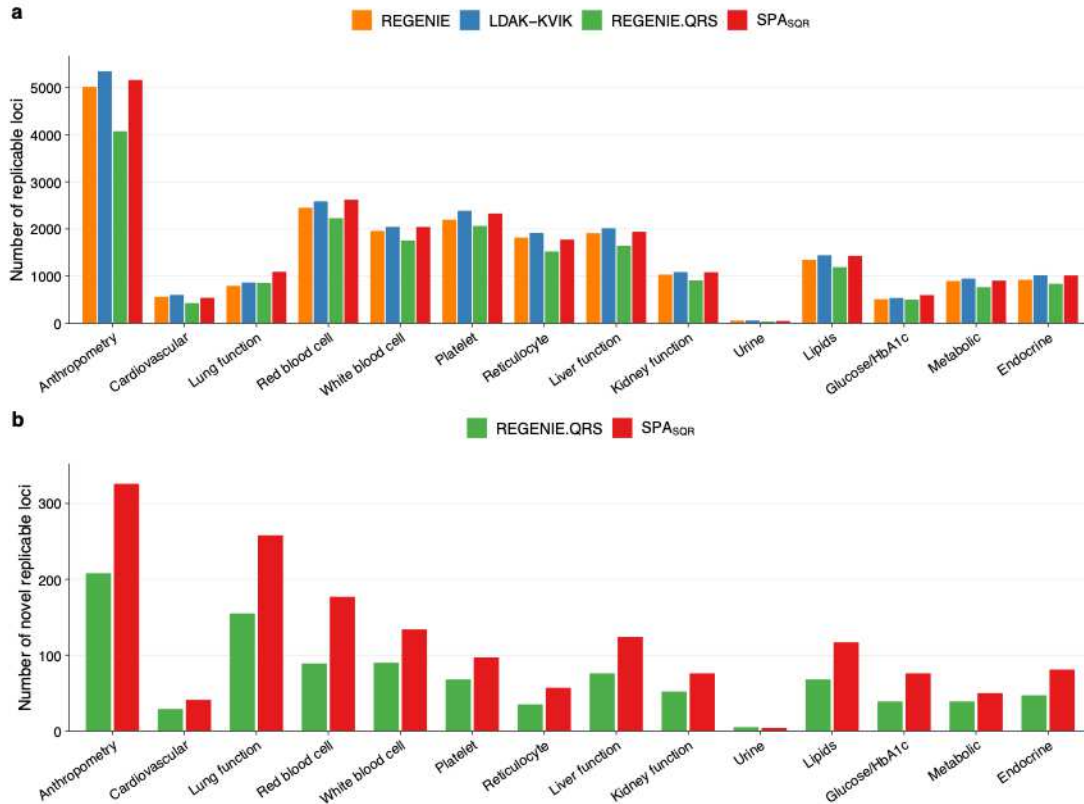


Fig. 4: Number of replicable loci discovered by different methods across 14 trait categories across 66 UK Biobank quantitative traits. a. Total number of replicable loci aggregated over each trait category for REGENIE, LDAK-KVIK, REGENIE.QRS and SPA_{SQR}. **b.** Number of loci missed by linear methods per trait category for REGENIE.QRS and SPA_{SQR}. A replicated locus is considered missed by linear methods if its lead variant lies at least 200 kb from every LDAK-KVIK and REGENIE replicated variant.

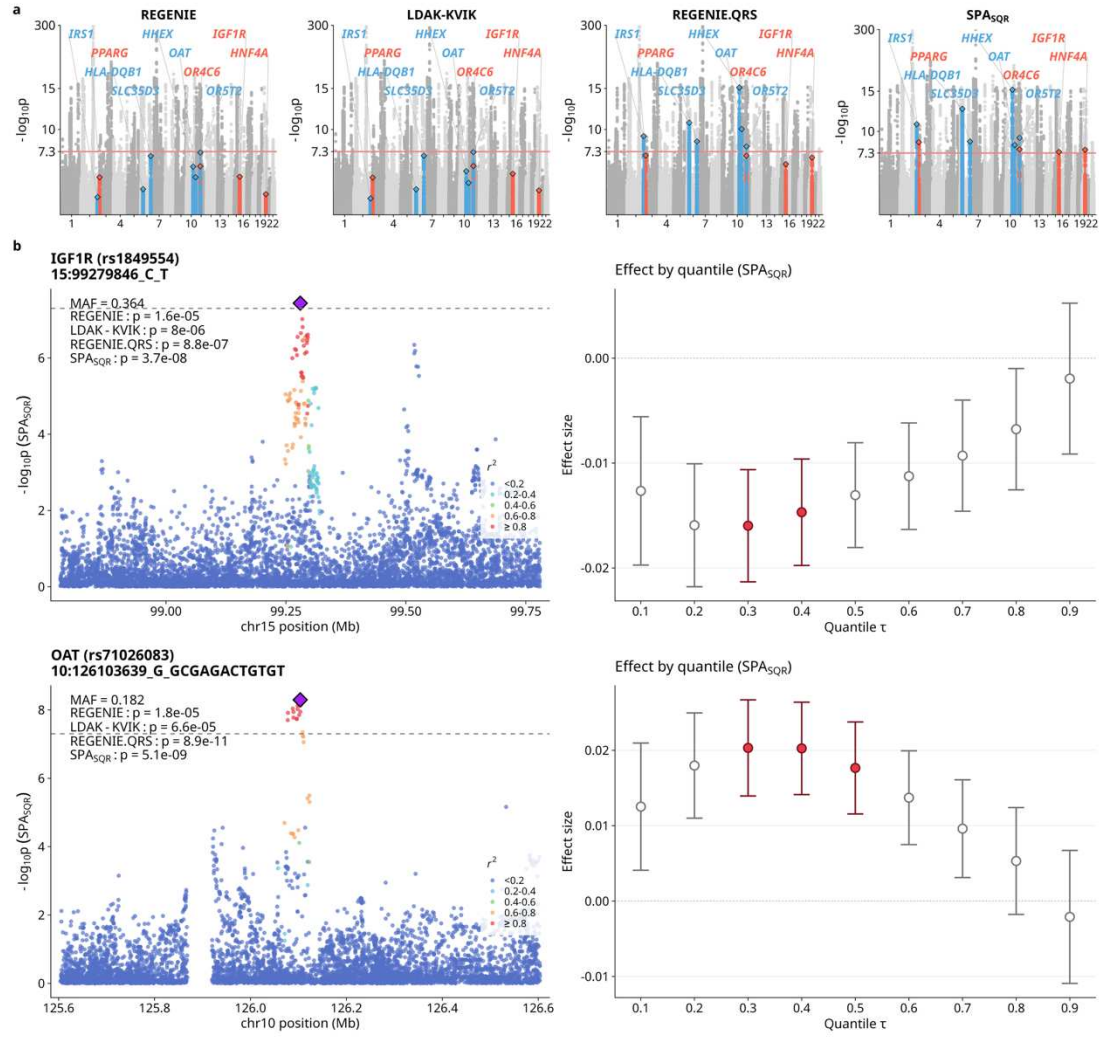


Fig. 5: GWAS results for glucose (UKB field 30740) using UK Biobank white British individuals. **a.** Genome-wide Manhattan plots for REGENIE, LDAK-KVIK, REGENIE.QRS, and SPASQR. The red horizontal line marks the genome-wide significance threshold ($p = 5 \times 10^{-8}$). In each panel, we highlight up to 10 loci that were identified by SPASQR but missed by the baseline methods. Loci discovered by SPASQR but missed by both linear GWAS methods are highlighted in blue, while those missed by all three competing methods are highlighted in orange. **b.** Two GWAS loci with heterogeneous effect patterns. For each locus, the left panel shows the an 1Mb regional zoom plot, with each variant colored by its LD r^2 relative to the lead variant. The upper left corner annotates its minor allele frequency and p-values from all four GWAS methods in the discovery cohort. The right panel displays the quantile-specific effect size $\hat{\gamma}(\tau)$ estimated at $\tau = 0.1, 0.2, \dots, 0.9$, with error bars indicating 95% confidence intervals. Quantiles with genome-wide significant Wald test p values are highlighted as filled red circles.

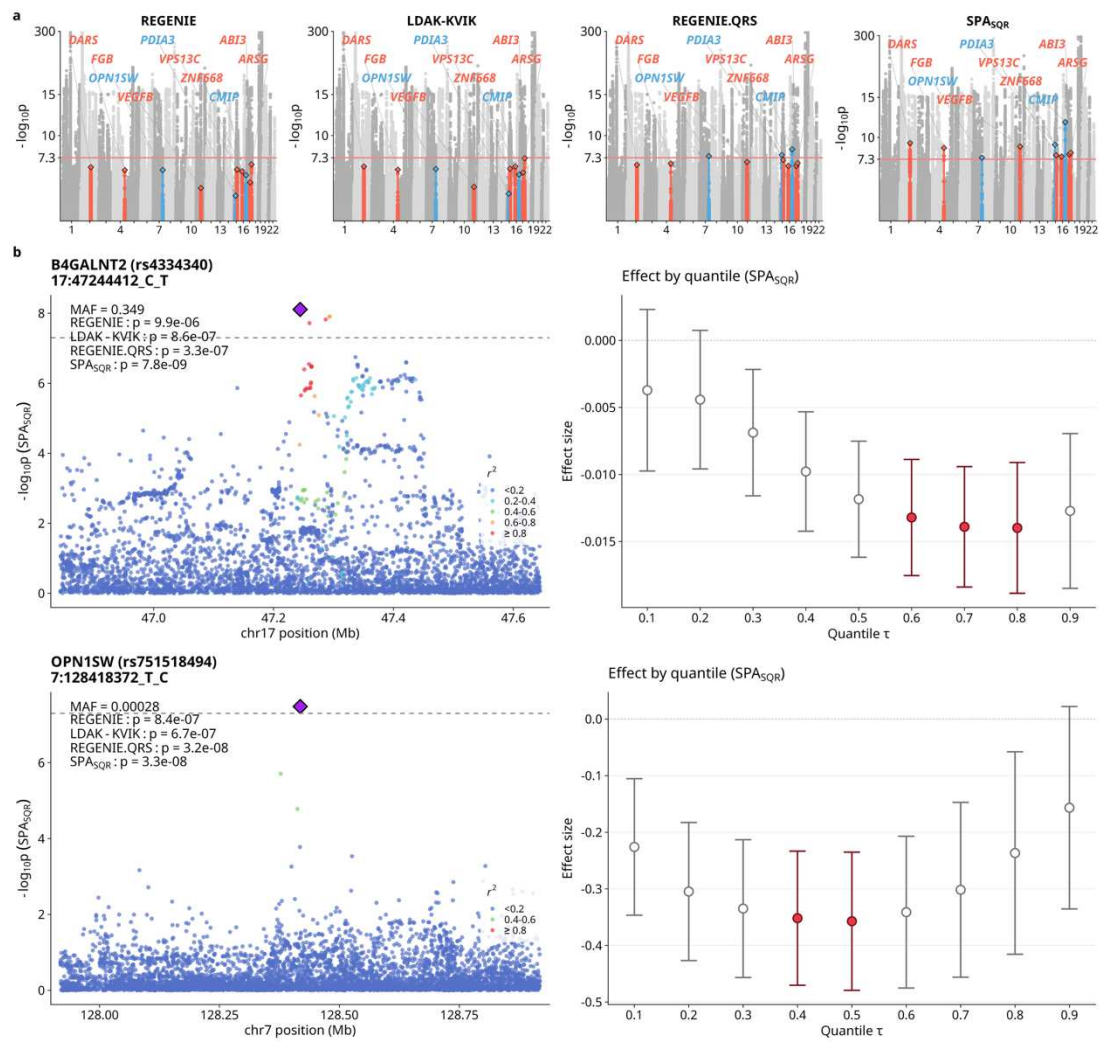


Fig. 6: GWAS results for apolipoprotein B (UKB field 30640) using white British individuals.

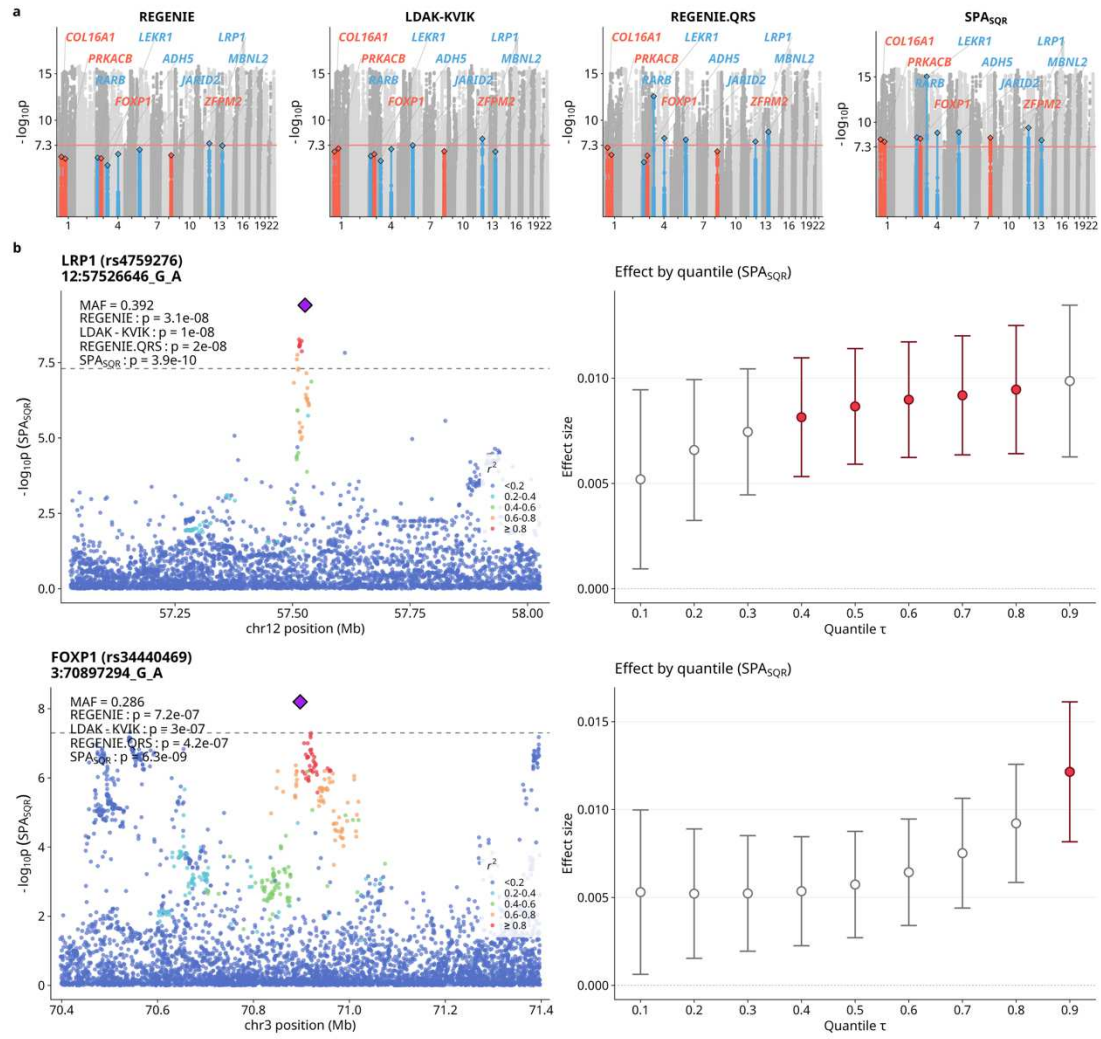


Fig. 7: GWAS results for body fat percentage (UKB field 23099) using white British individuals.

References

1. Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature Genetics* **47**, 284–290 (2015).
2. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature genetics* **50**, 1335–1341 (2018).
3. Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association tool for biobank-scale data. *Nature Genetics* **53**, 1616–1621 (2021).
4. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nature Genetics* **53**, 1097–1103 (2021).
5. Loya, H., Kalantzis, G., Cooper, F. & Palamara, P.F. A scalable variational inference approach for increased mixed-model association power. *Nature Genetics* **57**, 461–468 (2025).
6. Hof, J.P. & Speed, D. LDAK-KVIK performs fast and powerful mixed-model association analysis of quantitative and binary phenotypes. *Nature Genetics* **57**, 2116–2123 (2025).
7. Namba, S. *et al.* A cross-population compendium of gene–environment interactions. *Nature*, 1–10 (2026).
8. Koenker, R. *Quantile regression*, (Cambridge university press, 2005).
9. Wang, C. *et al.* Genome-wide discovery for biomarkers using quantile regression at biobank scale. *Nature Communications* **15**, 6460 (2024).
10. Mefford, J. *et al.* Beyond predictive R²: Quantile regression and non-equivalence tests reveal complex relationships of traits and polygenic scores. *The American Journal of Human Genetics* **112**, 1363–1375 (2025).
11. Miao, J. *et al.* A quantile integral linear model to quantify genetic effects on phenotypic variability. *Proceedings of the National Academy of Sciences* **119**, e2212959119 (2022).
12. Nascimento, M. *et al.* Quantile regression for genome-wide association study of flowering time-related traits in common bean. *PLoS One* **13**, e0190303 (2018).
13. Oliveira, G.F. *et al.* Population size in QTL detection using quantile regression in genome-wide association studies. *Scientific Reports* **13**, 9585 (2023).
14. Pozarickij, A., Williams, C., Hysi, P.G. & Guggenheim, J.A. Quantile regression analysis reveals widespread evidence for gene-environment or gene-gene interactions in myopia development. *Communications Biology* **2**, 167 (2019).
15. Wang, F. *et al.* Computationally efficient whole-genome quantile regression at biobank scale. in *Proc Natl Acad Sci U S A*, Vol. 122 e2513007122 (United States, 2025).
16. Ma, Y. *et al.* SPAmix: a scalable, accurate, and universal analysis framework for large-scale genetic association studies in admixed populations. *Genome biology* **26**, 356 (2025).

17. Ma, Y., Zhao, Y., Zhang, J.-F. & Bi, W. Efficient and accurate framework for genome-wide gene-environment interaction analysis in large-scale biobanks. *Nature Communications* **16**, 3064 (2025).
18. Xu, H. *et al.* SPAGRM: effectively controlling for sample relatedness in large-scale genome-wide association studies of longitudinal traits. *Nature Communications* **16**, 1413 (2025).
19. Li, Y. *et al.* Applying weighted Cox regression to genome-wide association studies of time-to-event phenotypes. *Nature Computational Science* **5**, 1064–1079 (2025).
20. Xu, H. *et al.* Leveraging longitudinal data to boost statistical power for gene–environment interaction analysis. *Nature Computational Science*, 1–13 (2026).
21. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics* **88**, 76–82 (2011).
22. Fernandes, M., Guerre, E. & Horta, E. Smoothing quantile regressions. *Journal of Business & Economic Statistics* **39**, 338–357 (2021).
23. He, X., Pan, X., Tan, K.M. & Zhou, W.-X. Smoothed quantile regression with large-scale inference. *Journal of Econometrics* **232**, 367–388 (2023).
24. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38**, 904–909 (2006).
25. Liu, Y. & Xie, J. Cauchy Combination Test: A Powerful Test With Analytic p-Value Calculation Under Arbitrary Dependency Structures. *Journal of the American Statistical Association* **115**, 393–402 (2020).
26. Long, M., Li, Z., Zhang, W. & Li, Q. The Cauchy combination test under arbitrary dependence structures. *The American Statistician* **77**, 134–142 (2023).
27. Heng, Q. & Wang, C. Inertial Quadratic Majorization Minimization with Application to Kernel Regularized Learning. *arXiv preprint arXiv:2507.04247* (2025).
28. Heng, Q., Zhou, H. & Lange, K. Tactics for improving least squares estimation. *Statistical Science* **To appear**(2026).
29. Bi, W., Fritsche, L.G., Mukherjee, B., Kim, S. & Lee, S. A fast and accurate method for genome-wide time-to-event data analysis and its application to UK Biobank. *The American Journal of Human Genetics* **107**, 222–233 (2020).
30. Bi, W. *et al.* Efficient mixed model approach for large-scale genome-wide association studies of ordinal categorical phenotypes. *The American Journal of Human Genetics* **108**, 825–839 (2021).
31. Ko, S. *et al.* GWAS of longitudinal trajectories at biobank scale. *The American Journal of Human Genetics* **109**, 433–445 (2022).
32. Dey, R. *et al.* Efficient and accurate frailty model approach for genome-wide survival association analysis in large-scale biobanks. *Nature communications* **13**, 5437 (2022).

33. Dey, R., Schmidt, E.M., Abecasis, G.R. & Lee, S. A fast and accurate algorithm to test for binary phenotypes and its application to PheWAS. *The American Journal of Human Genetics* **101**, 37–49 (2017).
34. Abecasis, G.R., Cherny, S.S., Cookson, W.O. & Cardon, L.R. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genetics* **30**, 97–101 (2002).
35. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742–015–0047–8 (2015).
36. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
37. Jee, Y.H. *et al.* Genome-wide association studies in a large Korean cohort identify quantitative trait loci for 36 traits and illuminate their genetic architectures. *Nature Communications* **16**, 4935 (2025).
38. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic acids research* **51**, D977–D985 (2023).
39. Siddle, K. Signalling by insulin and IGF receptors: supporting acts and new players. *Journal of molecular endocrinology* **47**, R1–R10 (2011).
40. Singh, R. *et al.* Autophagy regulates adipose mass and differentiation in mice. *The Journal of clinical investigation* **119**, 3329–3339 (2009).
41. Liu, P. *et al.* Foxp1 controls brown/beige adipocyte differentiation and thermogenesis through regulating β 3-AR desensitization. *Nature communications* **10**, 5070 (2019).
42. Ma, H. *et al.* The transcription factor Foxp1 regulates aerobic glycolysis in adipocytes and myocytes. *Journal of Biological Chemistry* **299**(2023).
43. Zuk, O., Hechter, E., Sunyaev, S.R. & Lander, E.S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences* **109**, 1193–1198 (2012).
44. Fu, B. *et al.* A biobank-scale test of marginal epistasis reveals genome-wide signals of polygenic interaction effects. *Nature Genetics*, 1–10 (2025).
45. Li, B. & Leal, S.M. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *The American Journal of Human Genetics* **83**, 311–321 (2008).
46. Morgenthaler, S. & Thilly, W.G. A strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a cohort allelic sums test (CAST). *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **615**, 28–56 (2007).
47. Han, F. & Pan, W. A data-adaptive sum test for disease association with multiple common or rare variants. *Human Heredity* **70**, 42–54 (2010).
48. Wu, M.C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel association test. *The American Journal of Human Genetics* **89**, 82–93 (2011).
49. Lee, S., Teslovich, T.M., Boehnke, M. & Lin, X. General framework for meta-analysis of rare variants in sequencing association studies. *The American Journal of Human Genetics* **93**, 42–53 (2013).

50. Liu, Y. *et al.* ACAT: a fast and powerful p value combination method for rare-variant analysis in sequencing studies. *The american journal of human genetics* **104**, 410–421 (2019).
51. Li, X. *et al.* Dynamic incorporation of multiple in silico functional annotations empowers rare variant association analysis of large whole-genome sequencing studies at scale. *Nature genetics* **52**, 969–983 (2020).
52. Joseph, T.A. *et al.* Computationally efficient meta-analysis of gene-based tests using summary statistics in large-scale genetic studies. *Nature Genetics*, 1–8 (2025).
53. Park, E. *et al.* Scalable and accurate rare variant meta-analysis with Meta-SAIGE. *Nature Genetics*, 1–8 (2025).
54. Liu, D.J. *et al.* Meta-analysis of gene-level tests for rare variant association. *Nature genetics* **46**, 200–204 (2014).
55. Li, X. *et al.* Powerful, scalable and resource-efficient meta-analysis of rare variant associations in large whole genome sequencing studies. *Nature genetics* **55**, 154–164 (2023).
56. Feng, S., Liu, D., Zhan, X., Wing, M.K. & Abecasis, G.R. RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics* **30**, 2828–2829 (2014).
57. He, Z. *et al.* GhostKnockoff inference empowers identification of putative causal variants in genome-wide association studies. *Nature Communications* **13**, 7209 (2022).
58. He, Z. *et al.* Beyond guilty by association at scale: searching for causal variants on the basis of genome-wide summary statistics. *BioRxiv*, 2024.02.28.582621 (2025).
59. Romano, Y., Patterson, E. & Candes, E. Conformalized quantile regression. *Advances in neural information processing systems* **32**(2019).
60. Wang, C. *et al.* Individualized Uncertainty Quantification in Polygenic Risk Scores Using Conformalized Quantile Regression. *bioRxiv*, 2025.10.13.682102 (2025).
61. Koenker, R. & Machado, J.A. Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association* **94**, 1296–1310 (1999).
62. Daniels, H.E. Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, 631–650 (1954).
63. Barndorff-Nielsen, O.E. Approximate interval probabilities. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **52**, 485–496 (1990).
64. Kang, H.M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* **42**, 348–354 (2010).
65. Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M. & Price, A.L. Advantages and pitfalls in the application of mixed-model association methods. *Nature Genetics* **46**, 100–106 (2014).

66. Geraci, M. & Bottai, M. Linear quantile mixed models. *Statistics and Computing* **24**, 461–479 (2014).
67. Galarza, C.E., Lachos, V.H. & Bandyopadhyay, D. Quantile regression in linear mixed models: a stochastic approximation EM approach. *Statistics and its Interface* **10**, 471 (2017).
68. Jurgens, S.J. *et al.* Adjusting for common variant polygenic scores improves yield in rare variant association analyses. *Nature Genetics* **55**, 544–548 (2023).
69. Heng, Q., Miao, L. & Bi, W. Summary statistics of SPASQR applied to quantitative traits of UK Biobank white British and European participants. in *Zenodo* (2026).

Acknowledgements

This research was supported by National Natural Science Foundation of China (62411540239, 82441004, 32541007, and 62273010, W.B.) and National Research Foundation of Korea (2024K2A9A2A0601368311, S.L.). UK Biobank data was accessed under the accession number 78795. We thank Professor Qizhai Li for reading manuscript and offering helpful comments. This research is supported by the high-performance computing platform of Peking University.

Author Contributions

Qiang Heng, Xiang Zhan, and Wenjian Bi conceptualized the project. Qiang Heng, Lin Miao, Xiang Zhan, and Wenjian Bi designed methods. Lin Miao and Qiang Heng developed the software. Qiang Heng performed simulation studies. Lin Miao and Wenjian Bi analyzed UK Biobank data. He Xu, Seunggeun Lee, and Dajiang Liu consulted on genetics and gave useful suggestions to improve the paper. Wenjian Bi and Xiang Zhan supervised the project. Qiang Heng and Wenjian Bi wrote the paper, with input from all authors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [spasqrsupplement.pdf](#)