Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Global and local genetic correlations between hypertrophic cardiomyopathy and atrial fibrillation: comparison of two linkage disequilibrium reference panels

Donglin Xu
Qilu Hospital of Shandong University
Zhen Yu
Qilu Hospital of Shandong University
Ang Li
Qilu Hospital of Shandong University
Zengkun Liu
Qilu Hospital of Shandong University
Qiduo Chen
Qilu Hospital of Shandong University
Jing Liu
Qilu Hospital of Shandong University
Hourong Sun
sunhourong2016@126.com
Qilu Hospital of Shandong University

Research Article

Keywords: hypertrophic cardiomyopathy, atrial fibrillation, genetic correlation, linkage disequilibrium, reference panels

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11546852/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Hypertrophic cardiomyopathy and atrial fibrillation are clinically associated, but their genetic associations may differ in direction across the genome. We examined their overall and local genetic correlations and whether local findings persisted after changing the linkage disequilibrium reference panel.

Methods

We reanalyzed public genome-wide association summary statistics, retaining the original variant set and 695 previously eligible regions. Each region used the same final variants with the 1000 Genomes European and UK Biobank European reference panels. Genome-wide uncertainty was calculated from the complete sampling covariance matrix. Local analyses used separate Benjamini–Hochberg and Benjamini–Yekutieli corrections, each with a fixed family of 695 tests.

Results

The genome-wide genetic correlation was 0.0964 (standard error 0.0442; 95% confidence interval 0.0098–0.1830; P = 0.0291). Local correlations were estimable in 542 regions with the 503-person 1000 Genomes panel and 192 with the 100000-person UK Biobank panel. Benjamini–Hochberg correction identified 38 and 10 significant regions, respectively. Among 187 regions estimable with both panels, 150 had the same sign and 37 had opposite signs. Eight regions were significant with both panels, including four positive and four negative correlations; four also passed Benjamini–Yekutieli correction with both. All eight belonged to the 45 regions significant in the original analysis.

Conclusions

A small positive overall correlation coexisted with local correlations in both directions. However, local estimability and significance depended on the reference panel, even with matched variants. The comparison reused the same disease association data and is not independent replication. Mixed ancestry in the cardiomyopathy study and incomplete variant-level cohort information limit interpretation of these regional findings.

Clinical trial registration

Clinical trial number: not applicable.

Background

Atrial fibrillation (AF) is a clinically important complication of hypertrophic cardiomyopathy (HCM). In people with HCM, AF is associated with symptoms and thromboembolic events, and clinical guidelines address its detection and management [1–3]. These clinical observations do not establish how common genetic variants contribute to the two diagnoses. We examined whether their genetic associations tend to point in the same direction throughout the genome or differ between regions.

Common variants contribute to HCM susceptibility in addition to rare disease-causing variants [4, 5]. Large AF studies have also identified associations across many genomic regions [6]. An AF polygenic score has been associated with AF development in people with HCM [7]. Our analysis addresses a different question: how genetic associations with HCM and AF compare in population-level association studies. It does not estimate the probability that a person with HCM will develop AF.

Linkage disequilibrium score regression (LDSC) estimates genetic correlation across the genome [8, 9]. Local genetic correlation analysis estimates the relationship within individual regions [10]. Local analyses require a reference panel to describe linkage disequilibrium (LD), the correlation between nearby genetic variants. Changing that panel can change whether a region can be analyzed, the direction of its estimated correlation, or its statistical significance. These outcomes need to be reported separately.

We estimated the genome-wide HCM–AF correlation and compared local estimates using 1000 Genomes and UK Biobank reference panels. To avoid comparing different variants, we matched the variants used by the two panels within each region. We reassessed the 695 regions eligible in the original analysis and tracked the 45 that were originally significant. The comparison retained eight regions with significant correlations under both references, while showing substantial differences in which other regions could be estimated.

Methods

Study design and data sources

We analyzed public genome-wide association study (GWAS) summary statistics and existing LD reference resources. We did not recruit participants or analyze patient-level HCM or AF clinical records. Ethics approval and consent for the source studies are described in their publications. This secondary analysis makes no claim of a new institutional approval or waiver.

The overall HCM GWAS, GCST90435254, included 5,900 cases and 68,359 controls [5]. It combined several ancestries, although most participants were European. The AF GWAS, GCST90624413, included 192,851 European cases and 1,239,541 controls and excluded UK Biobank [6]. We kept study-wide participant counts separate from the sample size reported for each variant. Additional file 1 describes the input data, and Additional file 2 records the source-file identities.

Variants and regions included in the comparison

The original analysis used variants available in four files: overall HCM, two HCM subtypes (GCST90435255 and GCST90435256), and AF. We retained this common set for the present reanalysis. The subtype files determined variant availability and remained in the four-trait LDSC analysis, but we report only the overall HCM–AF result. This choice preserves the original input; it can exclude variants that would be available in an analysis of HCM and AF alone.

The common set contained 1,162,323 autosomal single-nucleotide polymorphisms (SNPs) from HapMap3 [11]. We retained nonpalindromic single-nucleotide variants with distinct alleles, finite effect estimates, positive standard errors and sample sizes, valid P values, and effect-allele frequencies of 0.01–0.99. We removed duplicate rsIDs, chromosome conflicts, and position conflicts where coordinates could be compared directly. We excluded chromosome 6:25–34 Mb. HCM positions were reported on GRCh37 and AF positions on GRCh38; we matched variants by rsID and alleles and assigned regional positions from the GRCh37 reference. Additional file 1 gives the full filtering rules.

Of the original 2,495 genomic regions, 695 had passed the 1000 Genomes-based eligibility screen. We reanalyzed these 695 regions and did not retest the other 1,800. The comparison therefore describes a previously selected set, rather than all regions that either reference might allow us to analyze. We also retained the original 45 significant regions as a fixed subset, including those that later became non-estimable.

Genome-wide genetic correlation

We reran the original four-trait LDSC analysis in GenomicSEM using the original source files and variant set, European LD scores, separate regression weights, all 22 autosomes, and 200 jackknife blocks [8, 9, 12]. Sample and population prevalences were unspecified, so no population-liability conversion was applied. The rerun saved the genetic covariance matrix S, its sampling covariance matrix V, the intercept matrix, and the block-level estimates. These are newly reconstructed results from the original inputs, not the recovered historical fit object.

We calculated the HCM–AF genetic correlation as $S(HCM,AF)/\sqrt{[S(HCM,HCM) \times S(AF,AF)]}$. Its standard error was calculated by the delta method using the full V matrix. This includes uncertainty in both heritabilities and their covariance, together with the sampling covariances between these quantities. We used this standard error for the normal 95% confidence interval and two-sided P value throughout the main report. Additional file 1 compares it with the earlier calculation that treated the heritabilities as fixed. We checked V against the saved 200-block records without generating new bootstrap samples.

The LDSC intercept matrix was passed to LAVA for its sample-overlap adjustment. Excluding UK Biobank from the AF study does not exclude overlap through other cohorts. The adjustment also cannot recover missing information about each variant’s contributing cohorts or effective sample size.

Local correlations using matched variants

We used LAVA 0.1.5 to compare the 503-person 1000 Genomes European panel (1KG EUR) with the UKB100K European reference, version 1.1, based on 100,000 unrelated participants [10, 13–15]. Both analyses used the same HCM and AF association statistics and the original region boundaries. Both references were European and therefore did not fully match the ancestry composition of the HCM study.

Within each region, we matched variants by rsID, chromosome, GRCh37 position, and strand-normalized alleles. We aligned effect statistics to the reference dosage alleles and checked that the effects and per-variant sample sizes agreed between inputs. We then restricted both panels to variants retained after their native quality checks and verified the final sets after local processing. The final variant sets matched in all 695 regions.

We retained LAVA's native principal-component processing and did not replace or manually repair the LD matrices. Matching variants does not force the panels to have identical allele frequencies, LD estimates, or numbers of retained components. The comparison therefore removes differences in final variant membership, but it does not isolate the effect of LD estimates alone. Reference checks and regional component counts are provided in Additional file 2.

A region qualified for bivariate analysis only if it had at least 100 variants and 50 retained principal components, passed LAVA's native variance checks, and had a univariate P value below 0.05/2,495 for each trait. We retained the original bivariate settings, AF as the target, and the region-specific random seeds. We recorded computational failures, failed eligibility criteria, and non-significant estimates separately. Missing estimates were not assigned a correlation of zero or an artificial P value.

Multiple testing and result reporting

We applied Benjamini–Hochberg (BH) correction separately for each reference, keeping the number of planned tests fixed at 695 [16]. We also applied Benjamini–Yekutieli (BY) correction, which allows for dependence between tests [17]. Missing tests did not reduce this number. Corrections used the native LAVA P values, reported to six significant digits; higher-precision diagnostic quantities were not substituted. We used $q < 0.05$ for corrected statistical significance.

We reported three separate outcomes: whether a local correlation could be estimated, whether the two estimates had the same sign, and whether both passed correction. The set significant with both panels is the overlap of two separately corrected result sets, not a new joint test. The original 45-region subset used the same 695-test correction. Region IDs refer to genomic intervals, not causal genes, and adjacent regions were not assumed to be independent signals.

The analyses used R 4.4.1 [18]. Additional file 2 contains the numerical results and plotting inputs. Additional file 3 provides the analysis scripts for inspection, numerical checks, and code to reproduce the figures from the supplied results. Preparing this manuscript did not add new association, fine-

mapping, colocalization, or functional analyses. We followed the relevant STREGA reporting principles [19].

Use of artificial intelligence tools

ChatGPT and Codex assisted with workflow and code drafting, language editing and manuscript organization. All submitted statistical figures were rendered programmatically from the supplied numerical results; no generative image was used as a data figure. The authors are responsible for the analyses, interpretations and submitted content.

Results

Genome-wide correlation was small; local estimability differed between panels

The genome-wide HCM–AF genetic correlation was 0.0964 (standard error 0.0442; 95% confidence interval 0.0098–0.1830; P = 0.0291; Fig. 1A). The earlier calculation that treated the heritabilities as fixed gave the same point estimate and a standard error of 0.0439 (Additional file 1). Both calculations indicate a small positive correlation, not a strong overall genetic relationship.

Figure 1. Genome-wide correlation and regions eligible for local analysis. (A) HCM–AF genetic correlation and its normal 95% confidence interval, calculated using the full sampling covariance matrix. (B) Estimable and non-estimable regions in the fixed set of 695 previously eligible regions. Grey segments indicate regions without an eligible bivariate estimate, not regions with zero correlation. 1KG EUR, 1000 Genomes European reference; UKB100K EUR, UK Biobank European reference based on 100,000 participants.

Observed-scale SNP heritability was 0.08548 for HCM (standard error 0.01027) and 0.04669 for AF (standard error 0.00334). The corresponding LDSC intercepts were 1.04971 and 1.45100; the cross-trait intercept was 0.01390. The elevated AF intercept remains relevant when interpreting the results.

With matched variants, local correlations could be estimated in 542 of 695 regions using 1KG EUR and 192 using UKB100K EUR (Fig. 1B; Table 1). There were 153 eligibility-screen failures with 1KG EUR. With UKB100K EUR, 501 regions failed that screen and two failed native variance checks. In total, 187 regions were estimable with both panels, 355 only with 1KG EUR, five only with UKB100K EUR, and 148 with neither.

Table 1
Local analysis of the same 695 regions with matched variants.

Outcome	1KG EUR	UKB100K EUR
Fixed regions	695	695
Estimable regions	542	192
Not estimable	153	503
Eligibility screen failed	153	501
Native variance check failed	0	2
BH q < 0.05	38	10
Positive BH results	23	4
Negative BH results	15	6
BY q < 0.05	12	5
Estimable under both panels	187	187

Each reference used a correction family of 695 tests. The eligibility screen includes the univariate, variant-count, and component-count criteria. A non-estimable region has no eligible bivariate result; this is not evidence of a null association. Counts describe the matched-variant reanalysis, not the original analysis.

Significant local correlations included both positive and negative estimates

After BH correction, 38 regions were significant with 1KG EUR: 23 positive and 15 negative. UKB100K EUR yielded 10 significant regions: four positive and six negative (Fig. 2). BY correction retained 12 and five regions, respectively. These counts use the same 695-test family for each reference and replace neither the original 45-region result nor its analysis definition.

Figure 2. Local genetic correlations across the 695 selected regions. Points show correlations at GRCh37 region midpoints for (A) 1KG EUR and (B) UKB100K EUR. Orange and blue points indicate positive and negative BH-significant estimates; grey points indicate estimable results with BH q ≥ 0.05. Black rings identify the eight regions significant with both panels. Short ticks on a separate lower track mark non-estimable regions; their height is not a correlation value. BH correction used 695 tests for each panel. Alternating bands separate chromosomes; horizontal spans follow the represented intervals, not complete chromosome lengths.

Both reference panels therefore identified local correlations in both directions. However, the number of significant regions differed even though the final variant sets were matched. The plots retain non-

significant and non-estimable regions so that this difference is not reduced to a comparison of significant findings alone.

Most jointly estimable regions agreed in sign, but 37 did not

Among the 187 regions with estimates from both panels, 150 (80.2%) had the same sign and 37 had opposite signs (Fig. 3). Seventy-eight were positive with both panels and 72 were negative with both. Twenty were positive with 1KG EUR and negative with UKB100K EUR; 17 showed the reverse pattern. These counts compare point estimates. They do not test whether the two correlations differ statistically.

Figure 3. Agreement in local correlation estimates between reference panels. Each point represents one of the 187 regions estimable with both panels, using the same final variants. The dashed diagonal marks equal estimates; horizontal and vertical lines mark zero. Orange and blue points identify the eight regions with BH $q < 0.05$ in both analyses; grey points show the remaining regions. Quadrant labels give the sign combinations and their counts. The two axes use the same disease GWAS, so agreement is a reference comparison rather than independent replication.

Of the 45 regions significant in the original analysis, 40 remained estimable with 1KG EUR and 22 with UKB100K EUR. Those 22 were estimable with both panels, and all 22 agreed in sign. Only eight of the original 45 remained BH-significant with both panels. The other 23 could not be estimated with both references and were not classified as null findings.

Eight regions remained significant with both reference panels

The eight regions were 972, 1046, 1047, 1207, 1267, 1555, 1843, and 2469 (Fig. 4). Four had positive correlations and four had negative correlations. All eight belonged to the original 45-region subset. Regions 972, 1046, 1047, and 1555 also passed BY correction with both panels.

Figure 4. Paired estimates for the eight regions significant with both panels. Circles and diamonds show the 1KG EUR and UKB100K EUR estimates. Short connecting lines join the two point estimates; they are not confidence intervals. The columns give separate BH-adjusted q values with 695 tests per panel. Asterisks mark regions passing BY correction with both panels. Intervals are labeled in GRCh37 megabases; exact boundaries appear in Table 2 and Additional file 2. Adjacent regions are not assumed to represent independent signals.

The exact estimates and coordinates are listed in Table 2. Adjacent chromosome 6 regions 1046 and 1047 had UKB100K EUR correlations of 0.7609 and 0.8645, respectively. Region 972 had the most negative UKB100K EUR estimate among these eight regions (− 0.7411). Chromosome 10 region 1555 was also negative with both panels (− 0.4823 and − 0.4348). These results identify intervals for further study, not the variants or genes responsible for the associations.

Table 2
Local correlations significant with both reference panels.

ID	GRCh37 interval	ρ 1KG	ρ UKB	BH q 1KG	BH q UKB	Both BY
972	chr6:36346354–37570051	-0.6611	-0.7411	1.71e-08	2.01e-07	Yes
1046	chr6:117674077–118529071	0.8275	0.7609	5.52e-09	1.91e-05	Yes
1047	chr6:118529072–119451294	0.8262	0.8645	3.68e-13	1.88e-12	Yes
1207	chr7:127903739–128779173	-0.6357	-0.5420	1.59e-04	0.0244	No
1267	chr8:21650454–22895018	-0.3793	-0.4327	0.0089	0.0345	No
1555	chr10:73514215–75694023	-0.4823	-0.4348	4.60e-08	2.23e-05	Yes
1843	chr12:114685671–115439451	0.3911	0.3830	4.14e-04	0.0086	No
2469	chr22:23950733–25282436	0.4671	0.5948	0.0074	0.0011	No

Each panel was corrected separately using 695 tests. “Both BY” indicates BY q < 0.05 in both analyses. This overlap is not a joint test or independent replication. Local confidence limits were absent from the supplied numerical export and were not inferred from P values.

Discussion

We found a small positive genome-wide HCM–AF genetic correlation and local correlations in both directions. Eight regions remained significant with both reference panels, including four that passed BY correction with both. The broader results were less consistent: only 187 of the 695 selected regions were estimable with both panels, and 37 of those had opposite signs. The eight-region overlap and the changes elsewhere are both necessary to describe the result.

Matching the variants makes the comparison easier to interpret. If different variants enter the two analyses, an apparent reference effect may partly reflect a change in variant coverage. Here, the final

variant identities and aligned association statistics were the same within each region. Remaining differences can still arise from LD estimates, reference allele frequencies, and retained principal components. This analysis does not separate those effects or show which panel is unbiased for the mixed-ancestry HCM data.

The original 45-region subset illustrates why agreement in sign should not be reported alone. All 22 regions estimable with both panels agreed in sign, but those 22 represented less than half of the original 45. Only eight retained BH significance with both references. Reporting “22 of 22 agreed” without the 45-region denominator would conceal the regions that could no longer be compared. Conversely, calling the remaining 23 negative replications would treat missing estimates as observed null results.

A positive local correlation means that common-variant associations with the two diagnoses tend to align within the fitted regional model. A negative correlation means that they tend to oppose one another. Neither result shows that one disease causes or prevents the other. The analysis also does not identify a shared causal variant or establish that an HCM-related pathway protects against AF. This differs from the patient-level AF prediction question studied using polygenic scores [7]. Our findings describe regional genetic associations, not a clinical risk score.

We do not use the earlier covariance cancellation index to explain the small genome-wide correlation. LAVA’s native covariance parameter is normalized by the number of retained components, and its unweighted sum across regions is not a calibrated genome-wide covariance. That historical calculation is documented in Additional file 1, but it is not an endpoint of the present report. The local results stand on their estimated directions and corrected P values, not on a claim that opposing biological mechanisms cancel one another.

The study has several limitations. The 695 regions were selected by an earlier 1KG-based eligibility screen, so the comparison excludes regions that might be estimable only with UKB100K outside this set. HCM summary statistics combined ancestries, whereas the AF statistics and both LD panels were European. Matching variants does not resolve this mismatch or the earlier concern about the calibration of absolute local-variance sums. Variant-specific ancestry and case/control contributions were also incomplete, so study-wide sample proportions may not represent the information available at each variant.

The elevated AF LDSC intercept and incompletely known sample overlap add uncertainty. The intercept adjustment cannot establish that all confounding or overlap has been removed. In addition, both reference analyses reused the same GWAS, so their agreement is not external replication. Further disease-association data are needed to test the regional findings independently; functional data are needed to assign them to a mechanism. No patient-level prediction, treatment effect, or clinical utility was assessed here.

Conclusions

HCM and AF showed a small positive genome-wide genetic correlation. Within the 695 selected regions, eight had significant local correlations with both LD references, with four positive and four negative estimates. Most original regions did not retain significance with both panels, and many could not be jointly estimated. The results support follow-up of the specified intervals while showing why reference-panel sensitivity must be reported alongside local genetic associations.

Abbreviations

AF
atrial fibrillation
BH
Benjamini–Hochberg
BY
Benjamini–Yekutieli
CI
confidence interval
EUR
European
GRCh37/GRCh38
Genome Reference Consortium human builds 37/38
GWAS
genome-wide association study
HCM
hypertrophic cardiomyopathy
ID
identifier
1KG
1000 Genomes
LAVA
Local Analysis of [co]Variant Association
LD
linkage disequilibrium
LDSC
linkage disequilibrium score regression
SNP
single-nucleotide polymorphism
STREGA
STrengthening the REporting of Genetic Association studies
UKB100K
UK Biobank reference panel of 100000 unrelated European participants.

Declarations

Ethics approval and consent to participate

This study used public, deidentified GWAS summary statistics and existing LD reference resources. No participants were recruited and no patient-level HCM or AF clinical records were analyzed. Ethics approval and informed consent for the source studies are described in their original publications. No new institutional approval or waiver is claimed for this secondary analysis.

Consent for publication

Not applicable. No identifiable individual information is reported.

Competing interests

The authors declare that they have no competing interests.

Additional files

Additional file 1. Additional_file_1.docx. Detailed methods and numerical checks. Describes source data, variant matching, the full-covariance uncertainty calculation, regional eligibility, multiple-testing correction and interpretation limits. (DOCX)

Additional file 2. Additional_file_2.xlsx. Numerical results and source identities. Contains all 695 regions, the original 45-region subset, the eight-region overlap, reference checks, global matrices, saved jackknife records and a data dictionary. Numerical content is unchanged. (XLSX)

Additional file 3. Additional_file_3.zip. Numerical checks and figure-generation code. Reproduces summary counts, the uncertainty calculation and four figures from the included data. Original analysis scripts are retained with the external inputs required for full model reruns. (ZIP)

Funding

This research was funded by the Shandong Provincial Natural Science Foundation, grant number ZR2020MH029, awarded to Hourong Sun. The funder had no role in study design, data acquisition, analysis, interpretation, manuscript preparation or the decision to submit the article.

Author Contribution

DX and HS conceived the study. DX developed the methods and software, performed the formal analysis and visualization, and drafted the manuscript. DX, ZY, AL, ZL, QC and JL contributed to validation. DX, ZY,

AL, ZL and QC curated the data. DX, ZY, AL, ZL, QC, JL and HS reviewed and edited the manuscript. HS supervised the study, administered the project and acquired funding.

Acknowledgements

We thank the investigators and participants of the source GWAS and reference datasets.

Data Availability

The numerical datasets supporting the conclusions of this article are included in Additional file 2 and the CSV exports in Additional file 3. The source GWAS are identified by the GWAS Catalog records GCST90435254, GCST90435255, GCST90435256 and GCST90624413; exact source-file identities are recorded in Additional file 2. Linkage disequilibrium reference resources are described in the LAVA reference documentation. Additional file 3 contains numerical checks, figure-generation code and the original analysis scripts for inspection. The supplied files reproduce the figures and numerical summaries. Rerunning LDSC and LAVA additionally requires the source GWAS, linkage disequilibrium resources and historical project inputs listed in the code README; a self-contained full model rerun is not claimed.

References

1. Olivotto I, Cecchi F, Casey SA, Dolara A, Traverse JH, Maron BJ. Impact of atrial fibrillation on the clinical course of hypertrophic cardiomyopathy. Circulation. 2001;104(21):2517–24. 10.1161/hc4601.097997.
2. Rowin EJ, Hausvater A, Link MS, Abt P, Gionfriddo W, Wang W, et al. Clinical Profile and Consequences of Atrial Fibrillation in Hypertrophic Cardiomyopathy. Circulation. 2017;136(25):2420–36. 10.1161/circulationaha.117.029267.
3. Ommen SR, Ho CY, Asif IM, Balaji S, Burke MA, Day SM, et al. 2024 AHA/ACC/AMSSM/HRS/PACES/SCMR Guideline for the Management of Hypertrophic Cardiomyopathy: A Report of the American Heart Association/American College of Cardiology Joint Committee on Clinical Practice Guidelines. Circulation. 2024;149(23):e1239–311. 10.1161/cir.0000000000001250.
4. Harper AR, Goel A, Grace C, Thomson KL, Petersen SE, Xu X, et al. Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. Nat Genet. 2021;53(2):135–42. 10.1038/s41588-020-00764-0.
5. Tadros R, Zheng SL, Grace C, Jordà P, Francis C, West DM, et al. Large-scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. Nat Genet. 2025;57(3):530–8. 10.1038/s41588-025-02087-4.

6. Yuan S, Chen J, Ruan X, Li Y, Abramowitz SA, Wang L, et al. Cross-population GWAS and proteomics improve risk prediction and reveal mechanisms in atrial fibrillation. Nat Commun. 2025;16(1):6426. 10.1038/s41467-025-61720-2.
7. Lee YS, Yang PS, Jang E, Kim D, Yu HT, Kim TH, et al. Clinical and Genetic Risk Factors Predict Atrial Fibrillation on the Basis of Hypertrophic Cardiomyopathy. J Am Heart Assoc. 2025;14(6):e038074. 10.1161/jaha.124.038074.
8. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291–5. 10.1038/ng.3211.
9. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236–41. 10.1038/ng.3406.
10. Werme J, van der Sluis S, Posthuma D, de Leeuw CA. An integrated framework for local genetic correlation analysis. Nat Genet. 2022;54(3):274–82. 10.1038/s41588-022-01017-y.
11. International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. Nature. 2010;467(7311):52–8. 10.1038/nature09298.
12. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav. 2019;3(5):513–25. 10.1038/s41562-019-0566-x.
13. 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015;526(7571):68–74. 10.1038/nature15393.
14. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–9. 10.1038/s41586-018-0579-z.
15. Werme J, de Leeuw CA. LAVA reference data documentation: UK Biobank reference, version 1.1. https://github.com/josefin-werme/LAVA/blob/main/REFERENCE.md. Accessed 27 Oct 2026.
16. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B (Methodological). 1995;57:289–300. 10.1111/j.2517-6161.1995.tb02031.x.
17. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat. 2001;29(4):1165–88. 10.1214/aos/1013699998.
18. R Core Team. R: A language and environment for statistical computing. Version 4.4.1. Vienna: R Foundation for Statistical Computing; 2024. https://www.R-project.org/.
19. Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, et al. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. PLoS Med. 2009;6(2):e22. 10.1371/journal.pmed.1000022.
20. NHGRI-EBI GWAS Catalog. Study GCST90435254. https://www.ebi.ac.uk/gwas/studies/GCST90435254. Accessed 27 Oct 2026.

21. NHGRI-EBI GWAS Catalog. Study GCST90435255. https://www.ebi.ac.uk/gwas/studies/GCST90435255. Accessed 27 Oct 2026.
22. NHGRI-EBI GWAS Catalog. Study GCST90435256. https://www.ebi.ac.uk/gwas/studies/GCST90435256. Accessed 27 Oct 2026.
23. NHGRI-EBI GWAS Catalog. Study GCST90624413. https://www.ebi.ac.uk/gwas/studies/GCST90624413. Accessed 27 Oct 2026.

Figures

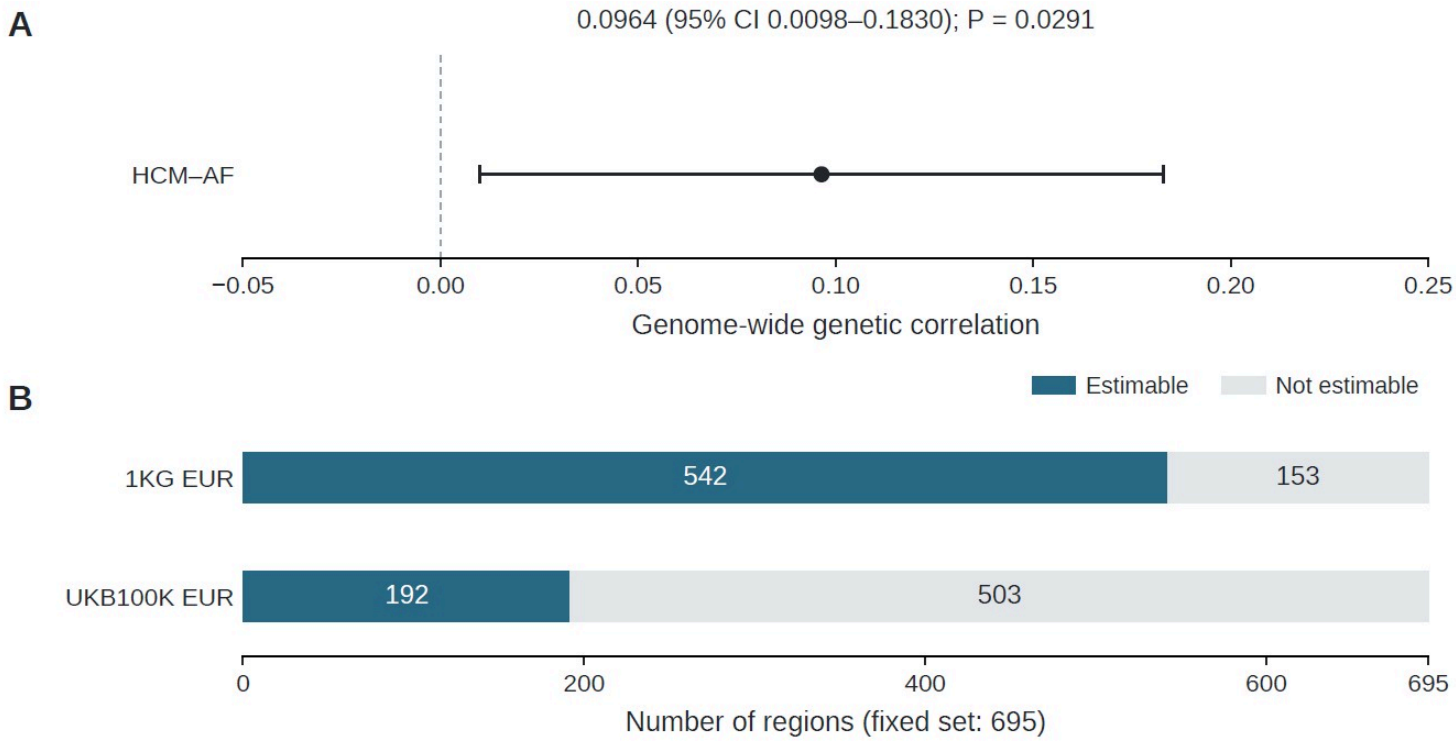

Figure 1

Genome-wide correlation and regions eligible for local analysis. (A) HCM–AF genetic correlation and its normal 95% confidence interval, calculated using the full sampling covariance matrix. (B) Estimable and non-estimable regions in the fixed set of 695 previously eligible regions. Grey segments indicate regions without an eligible bivariate estimate, not regions with zero correlation. 1KG EUR, 1000 Genomes European reference; UKB100K EUR, UK Biobank European reference based on 100,000 participants.

Observed-scale SNP heritability was 0.08548 for HCM (standard error 0.01027) and 0.04669 for AF (standard error 0.00334). The corresponding LDSC intercepts were 1.04971 and 1.45100; the cross-trait intercept was 0.01390. The elevated AF intercept remains relevant when interpreting the results.

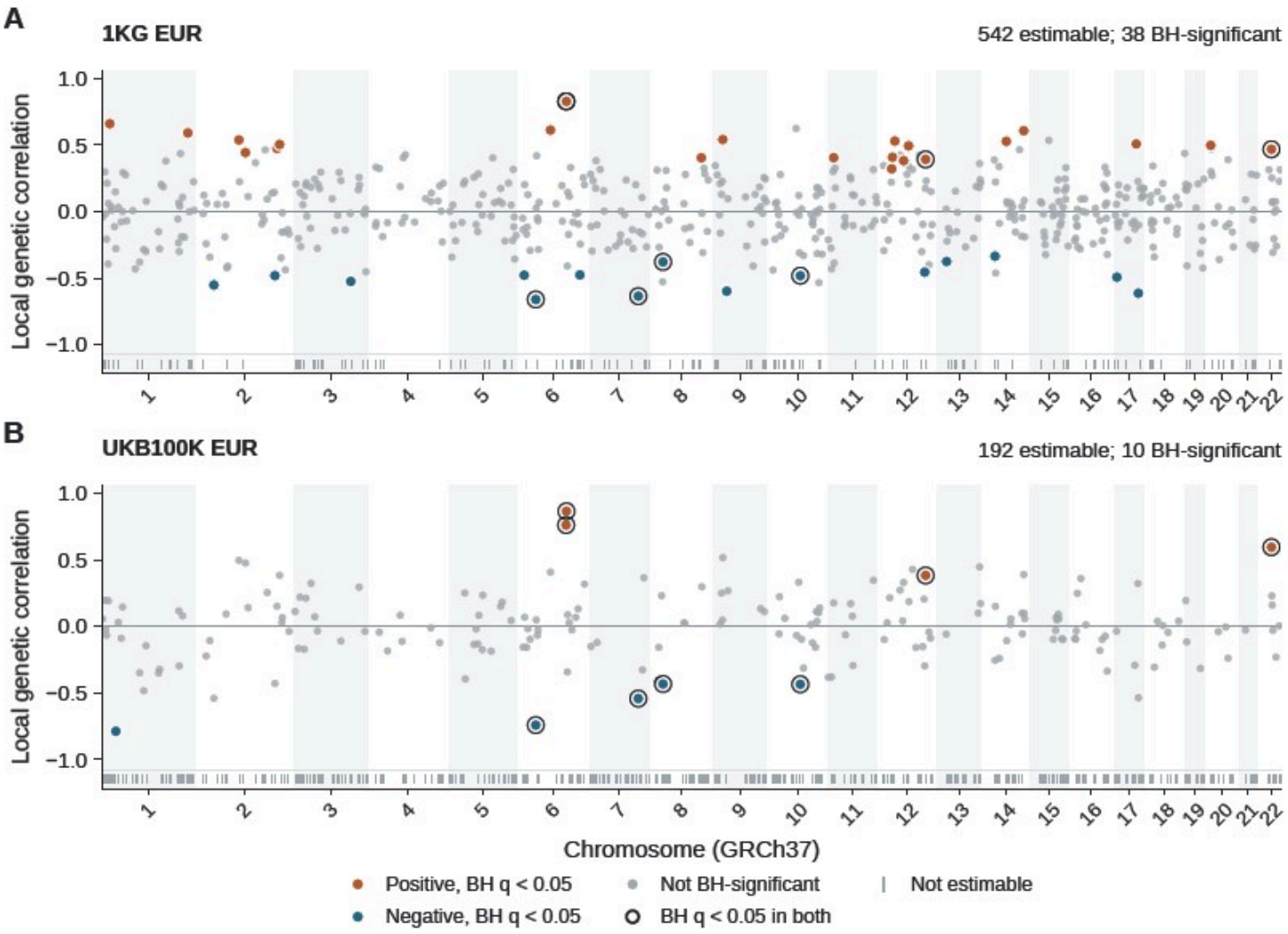

Figure 2

Local genetic correlations across the 695 selected regions. Points show correlations at GRCh37 region midpoints for (A) 1KG EUR and (B) UKB100K EUR. Orange and blue points indicate positive and negative BH-significant estimates; grey points indicate estimable results with BH q ≥ 0.05. Black rings identify the eight regions significant with both panels. Short ticks on a separate lower track mark non-estimable regions; their height is not a correlation value. BH correction used 695 tests for each panel. Alternating bands separate chromosomes; horizontal spans follow the represented intervals, not complete chromosome lengths.

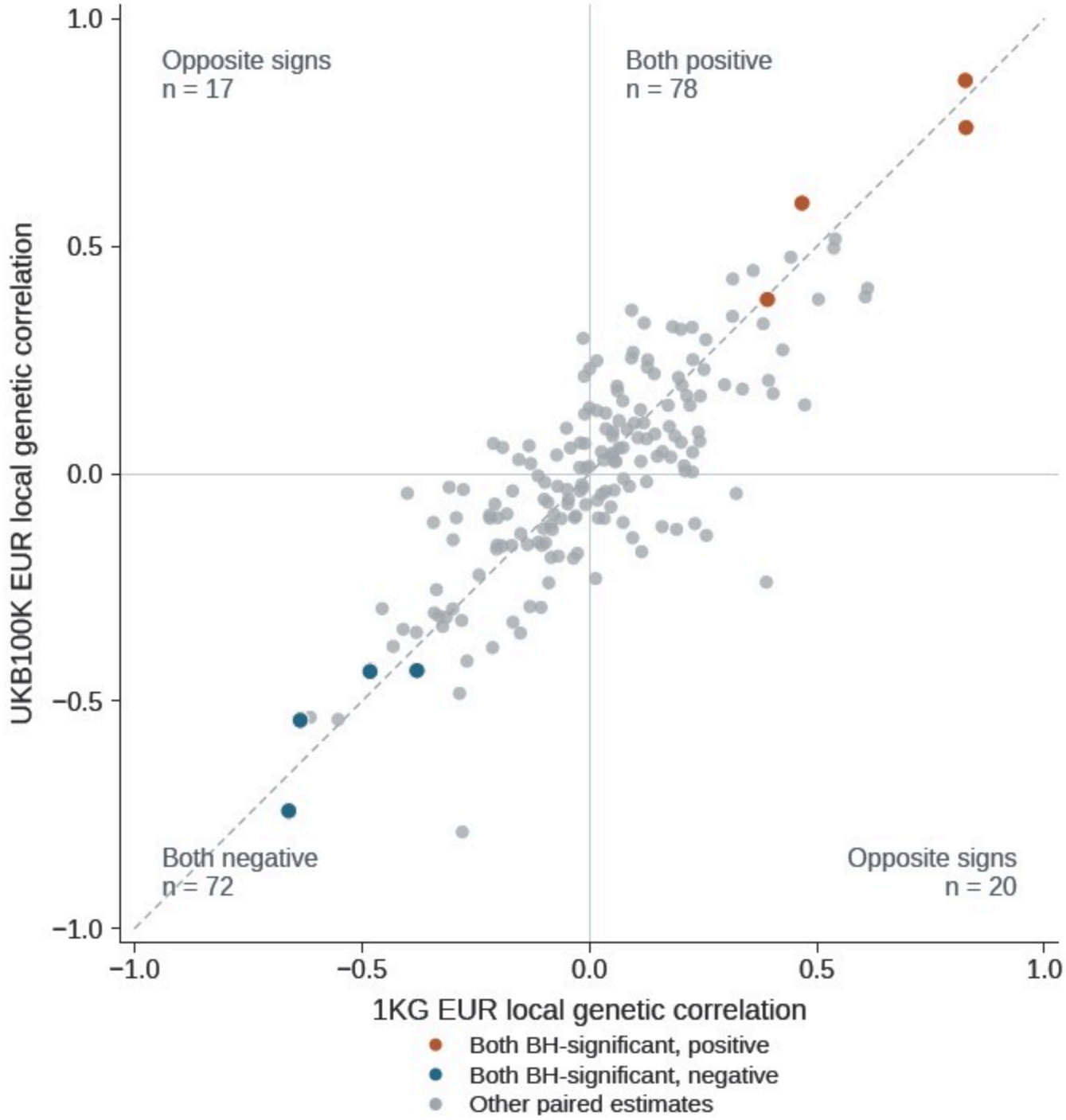

Figure 3

Agreement in local correlation estimates between reference panels. Each point represents one of the 187 regions estimable with both panels, using the same final variants. The dashed diagonal marks equal estimates; horizontal and vertical lines mark zero. Orange and blue points identify the eight regions with BH q < 0.05 in both analyses; grey points show the remaining regions. Quadrant labels give the sign combinations and their counts. The two axes use the same disease GWAS, so agreement is a reference comparison rather than independent replication.

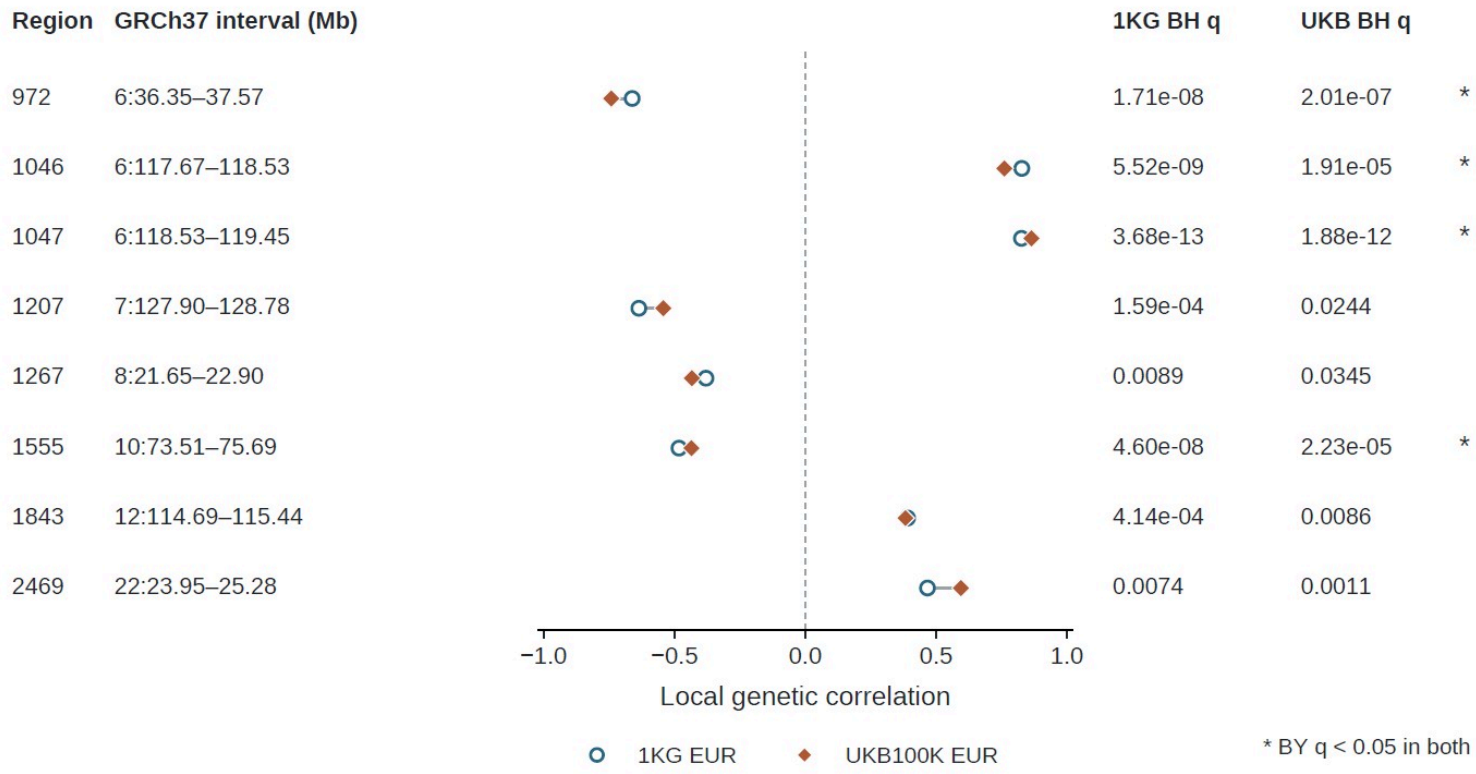

Figure 4

Paired estimates for the eight regions significant with both panels. Circles and diamonds show the 1KG EUR and UKB100K EUR estimates. Short connecting lines join the two point estimates; they are not confidence intervals. The columns give separate BH-adjusted q values with 695 tests per panel. Asterisks mark regions passing BY correction with both panels. Intervals are labeled in GRCh37 megabases; exact boundaries appear in Table 2 and Additional file 2. Adjacent regions are not assumed to represent independent signals.

Supplementary Files

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

- Additionalfile1.docx
- Additionalfile3.zip
- Additionalfile2.xlsx