Supplementary methods and sensitivity analyses

CD3 on CD39-Positive Resting CD4 Regulatory T Cells and ALS Risk: Genetic Localization to the CD247 Region

Molecular Neurobiology

Rongrong Du[1], Hongfen Wang[1], Fei Yang[1], Qiang Shi[1], Xusheng Huang[1], Mao Li[1]

[1] Department of Neurology, The First Medical Center, Chinese PLA General Hospital, Beijing, China

Correspondence: Mao Li, Department of Neurology, The First Medical Center, Chinese PLA General Hospital, 28 Fuxing Road, Haidian District, Beijing 100853, China. Email: limao301shennei@163.com

Phenotype selection and analysis flow

A case-insensitive selection expression combined Treg, T reg, T-reg, regulatory T, T regulatory, FOXP3, CD25.*CD127, CD127.*CD25, CD4.*CD25, CD25hi, CD39, and CTLA.?4, selecting 74 of 731 traits. All selected traits had harmonized data. One trait, CD39+ activated CD4 regulatory T cell %activated CD4 regulatory T cell, retained a single SNP and was excluded from the multi-SNP IVW screen. The testing family therefore comprised 73 traits: 63 regulatory and 10 non-regulatory gates, classified from their flow-cytometry labels. These labels do not establish the cellular origin of a causal effect. Fig. S4 shows the analysis flow.

Variance conventions and follow-up selection

The original screen used multiplicative random-effects IVW with residual variance allowed below one. The main analysis applies a lower bound of one to this variance. Point estimates are identical under the two conventions; standard errors differ when Q is below its degrees of freedom. For the focal CD3 phenotype, the standard error and unadjusted P value are also unchanged. Its FDR changes from 0.042 to 0.084 because the other P values and their ranking change. Benjamini–Hochberg correction uses the same 73 traits under each convention (Fig. S1). Follow-up analyses retain the three phenotypes selected in the original screen. Strict-threshold, region-deletion, related-gate, eQTL sensitivity and prior-sensitivity analyses were conducted post hoc; their nominal P values are exploratory.

IVW estimation and uncertainty

The original estimator, TwoSampleMR::mr_ivw_mre, permits residual underdispersion to reduce standard errors. A common variance floor was applied to the main screen and follow-up IVW estimates. The original estimates are provided for comparison using the same data (Online Resource 2). Let X denote SNP–exposure estimates, Y denote SNP–outcome estimates, w = 1/SE(Y)², and K denote the number of instruments. The IVW slope, residual heterogeneity and standard error were calculated as follows:

$$\beta = \frac{\Sigma\, wXY}{\Sigma\, wX^2}$$

$$Q = \Sigma\, w(Y - \beta X)^2$$

$$SE(\beta) = \sqrt{\frac{max[1, Q/(K - 1)]}{\Sigma\, wX^2}}$$

Two-sided normal P values and 95% confidence intervals were calculated consistently. A single instrument used the first-order Wald ratio without a residual-variance estimate. Exposure-association uncertainty remains a limitation of these first-order estimators.

Instrument and diagnostic traceability

The archived harmonized data contain effect and other alleles, SNP identifiers, positions, effect estimates, standard errors, allele frequencies, exposure sample sizes, F statistics and retention flags. The primary screen uses rows with mr_keep = TRUE and at least two instruments per exposure. Online Resource 2 also includes the original secondary MR-estimator, heterogeneity, Egger-intercept and MR-PRESSO outputs, identified by filename. These diagnostic analyses were not all rerun from genome-wide source data. The main IVW, leave-one-out, reverse, locus

and colocalization calculations were reconstructed from the archived inputs. Non-significant diagnostic tests do not exclude pleiotropy.

Historical and FinnGen outcome sensitivity

The GCST005647 comparison uses both the available harmonized instruments and the subset shared with the primary outcome. For CD3, these sets contain nine and eight SNPs, respectively. Historical IVW FDR values use the three-phenotype follow-up family and are separate from the primary 73-trait correction. The strict-threshold historical estimate uses rs1723016 alone. Because GCST005647 overlaps the 2021 ALS dataset, this comparison was not treated as independent replication or combined in a meta-analysis. OpenGWAS metadata identify finn-b-G6_ALS as Motor neuron disease, 2021, with 219 cases and 111,621 controls, consistent with the FinnGen R5 Risteys endpoint count. IVW estimates use SNP effect estimates and standard errors, rather than these metadata sample counts. The nine-SNP CD3 estimate is 0.775 (95% CI 0.549–1.092; P = 0.145); the common eight-SNP estimate is 0.718 (95% CI 0.458–1.126; P = 0.149). No external three-trait analysis passes FDR correction. The FinnGen endpoint is broader than strictly adjudicated ALS, and participant nonoverlap could not be verified.

Binary-outcome Steiger calculation

ALS liability correlations were calculated using the TwoSampleMR get_r_from_lor logit formula, observed outcome allele frequencies and 27,205/110,881 case/control counts. This replaces the original automatic calculation, which treated both traits as quantitative. The Python implementation was checked against the archived official R source. Prevalence of 5 per 100,000 was the reference scenario, informed by the epidemiological estimate near 5.05; sensitivity scenarios used 1, 10 and 50 per 100,000. Results are provided for total N and the TwoSampleMR effective-N expression 2/(1/Ncase + 1/Ncontrol). Exposure correlations use beta, standard error and phenotype-specific N. Measurement-error sensitivity used reliability values of 0.1, 0.25, 0.5, 0.75 and 1.0 for each trait, with observed R^2 divided by assumed reliability. Of 300 combinations across three phenotypes, four prevalences and 25 reliability pairs, 40 implied $R^2 \geq 1$ and were excluded. All 260 admissible combinations retained forward ordering. The reliability values are illustrative rather than empirically estimated; this analysis does not resolve winner's curse or establish the exclusion restriction.

Joint linkage disequilibrium

The 30-SNP MVMR union was evaluated in 503 participants from the 1000 Genomes Phase 3 EUR panel: CEU 99, FIN 99, GBR 91, IBS 107 and TSI 107. Genotype calls were complete for all variants. PLINK v1.9.0-b.7.14 (4 September 2026) used the minimum exposure P value to prioritize index SNPs, clump-p1 = clump-p2 = 1 for the selected union, $r^2 = 0.001$ and a 10-Mb window. All 30 SNPs were retained. The 31 within-window pairs had maximum haplotype $r^2 = 0.000917627$, measured using the R2 field in the PLINK output. A threshold check at 0.0005 produced 25 clumps. Genotype-dosage correlations exceeded 0.001 for nine local pairs, with maximum $r^2 = 0.0027112241$; P-prioritized greedy dosage pruning retained 26 SNPs. These thresholds apply to local pairs within 10 Mb. The selected-SNP matrix does not provide the dense regional LD reference required for coloc.susie.

Conditional strength and multivariable inference

The source covariance estimates were obtained from 15,866,292 aligned genome-wide variants using jointly null absolute-z thresholds of 0.4, 0.5 and 0.6. The 0.5 setting is the reference specification. SNP-specific covariance matrices use the corresponding exposure standard errors. Conditional F values were recalculated with the official MVMR phenocov_mvmr and strength_mvmr functions; independent Python and R results agreed to approximately 4.3×10^{-14}. The 30-SNP reference model has conditional F of 11.34/24.06/6.81 and the 26-SNP dosage model 13.07/23.68/7.85 for CD3/CD39/CD25. Coefficient covariance is inverse(X′WX) multiplied by max[1, Q/(K − 3)]. Both normal P values with normal confidence intervals and matched Student-t P values with t confidence intervals are supplied. The 21-SNP positional-pruning model is retained as a separate post hoc sensitivity analysis.

Colocalization assumptions and credible sets

Regions are chr1:166933056–167933056 for the CD3 phenotype, chr10:97241071–98241071 for CD39 percentage, and chr1:198330942–199330942 for CD25 (GRCh37). Variants were aligned by alleles, outcome effects flipped as required, ambiguous high-frequency palindromes removed, and duplicate identifiers resolved using the original

quality-control rules. The quantitative and case-control approximate Bayes factors use prior effect standard deviations 0.15 and 0.20, respectively. Default p1/p2 = 10^{-4} and p12 = 10^{-5} are accompanied by p12 = 10^{-6}, 5×10^{-6} and 5×10^{-5}. Further p1/p2 combinations of 5×10^{-5} and 2×10^{-4} are tabulated, including asymmetric settings. Across this wider grid, the principal H4 probability ranges from 0.632 to 0.999, compared with 0.856 to 0.997 when only p12 varies at the default p1/p2. The seven-variant credible set represents uncertainty in variant identity conditional on a shared signal. Low H4 alone does not establish distinct causal variants. Quantitative sample sizes were 2,649, 3,437 and 3,435 for CD3, CD39 percentage and CD25, respectively, with sdY = 1. ALS used case fraction 27,205/138,086; trait-specific allele frequencies were retained.

Expression sample reconciliation

GSE288704 contains 69 processed-matrix columns but 66 deposited sample records. The unmatched columns COUNTS_C11, COUNTS_PLMN and COUNTS_PUMN were excluded because no corresponding sample records were available. The dataset narrative describes 48 cases, while the matched deposited sample records support 47 cases and 19 controls. GSE308462 contains 260 samples from 159 subject IDs, up to six samples per person, and 28 cerebrospinal-fluid samples. The post hoc BA4 comparison contains 34 unique donors and adjusts only age and sex. It does not correct unavailable batch, RNA integrity, postmortem interval or cellular composition. The two blood comparisons use Welch tests on the indicated transformed scales. Twelve genes in each contrast produce 36 tests; both within-study and all-36 Benjamini–Hochberg values are supplied. The analyses are descriptive and are not transcriptome-wide differential-expression models.

Files and reproducibility

Online Resource 2 contains CSV result tables, an index and explanatory notes. Online Resource 3 provides analysis-ready inputs, scripts, reference outputs, selected public genotypes and software requirements. Its LD/MVMR module supports offline recomputation. Reproduction starts from the supplied harmonized and transformed inputs; original GWAS acquisition, genome-wide harmonization and initial exposure clumping are outside this scope. Source accessions, input manifests and software-version records document provenance.

Public resource links

OpenGWAS: https://opengwas.io/ ; official dataset metadata: https://opengwas.io/data/gwasinfo/gwasinfo.json ; FinnGen R5 endpoint: https://r5.risteys.finngen.fi/phenocode/G6_ALS ; GEO: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE234297 , https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE288704 , https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE308462 ; Ensembl genotype queries: https://grch37.rest.ensembl.org/variation/human/rs1723016?genotypes=1 ; Phase 3 population panel: https://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/integrated_call_samples_v3.20130502.ALL.panel

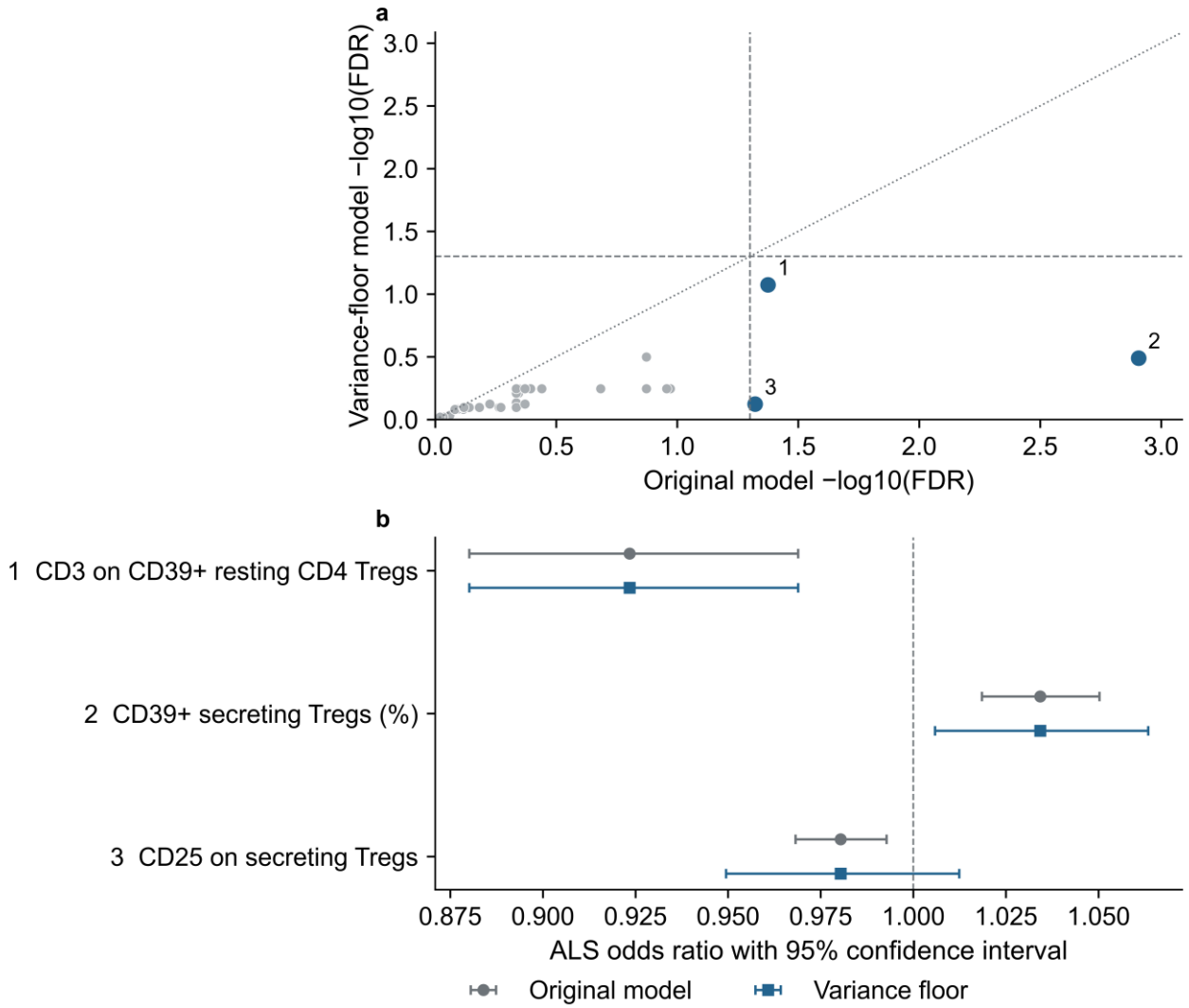


Fig. S1 Comparison of IVW variance conventions in the original 73-trait screen. a Original-model and variance-floor FDR values on negative log10 scales. Dashed lines mark FDR = 0.05 and the dotted line marks equality. Labels 1–3 identify the phenotypes in panel b. b The same point estimates with original-model circles and variance-floor squares and their 95% confidence intervals. The focal CD3 estimate, standard error and unadjusted P value are identical; its FDR changes from 0.042 to 0.084 because the other P values and the full-screen ranking change. Both conventions use the same data. The percentage phenotype denotes CD39-positive secreting Tregs as a proportion of secreting Tregs. No phenotype passes the revised 73-trait FDR threshold

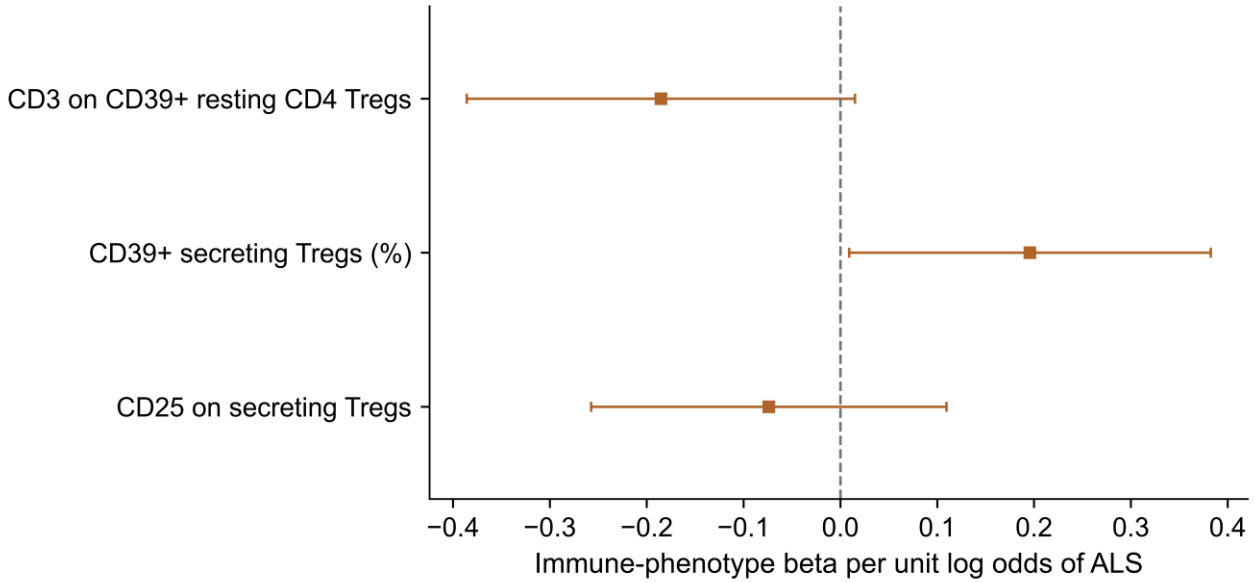


Fig. S2 Reverse MR from ALS liability to the three originally selected immune phenotypes. Points and bars show continuous immune-phenotype beta coefficients and 95% confidence intervals per unit higher genetically predicted log odds of ALS, using 11 ALS instruments and the uniform variance floor. The dashed line marks zero. The focal CD3 estimate is −0.185 (95% CI −0.386 to 0.015; P = 0.0699; FDR across three tests = 0.105). No reverse estimate passes FDR correction

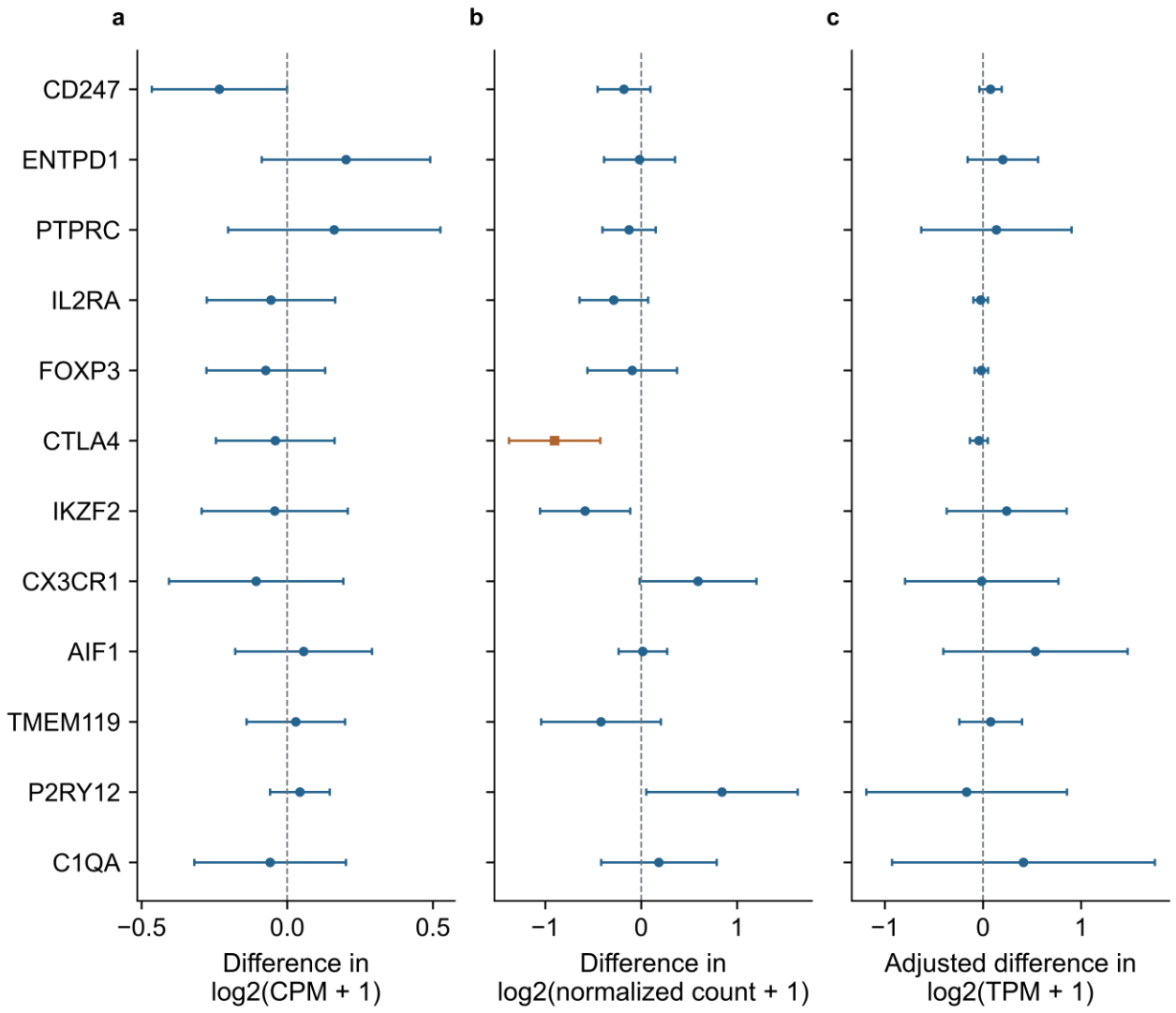


Fig. S3 Exploratory expression comparisons. a GSE234297 blood, 96 ALS cases and 48 controls. b Metadata-matched GSE288704 PBMCs, 47 cases and 19 controls. c GSE308462 motor-cortex BA4, 29 ALS donors and five controls. Points show ALS-minus-control differences with 95% confidence intervals. Blood comparisons use Welch inference; BA4 uses age- and sex-adjusted OLS with HC3 standard errors and t inference. Squares indicate within-contrast Benjamini–Hochberg $P < 0.05$ across 12 genes; circles do not meet that threshold. Axis scales represent different transformations and cannot be read as directly comparable fold changes. CD247 does not pass correction in any panel

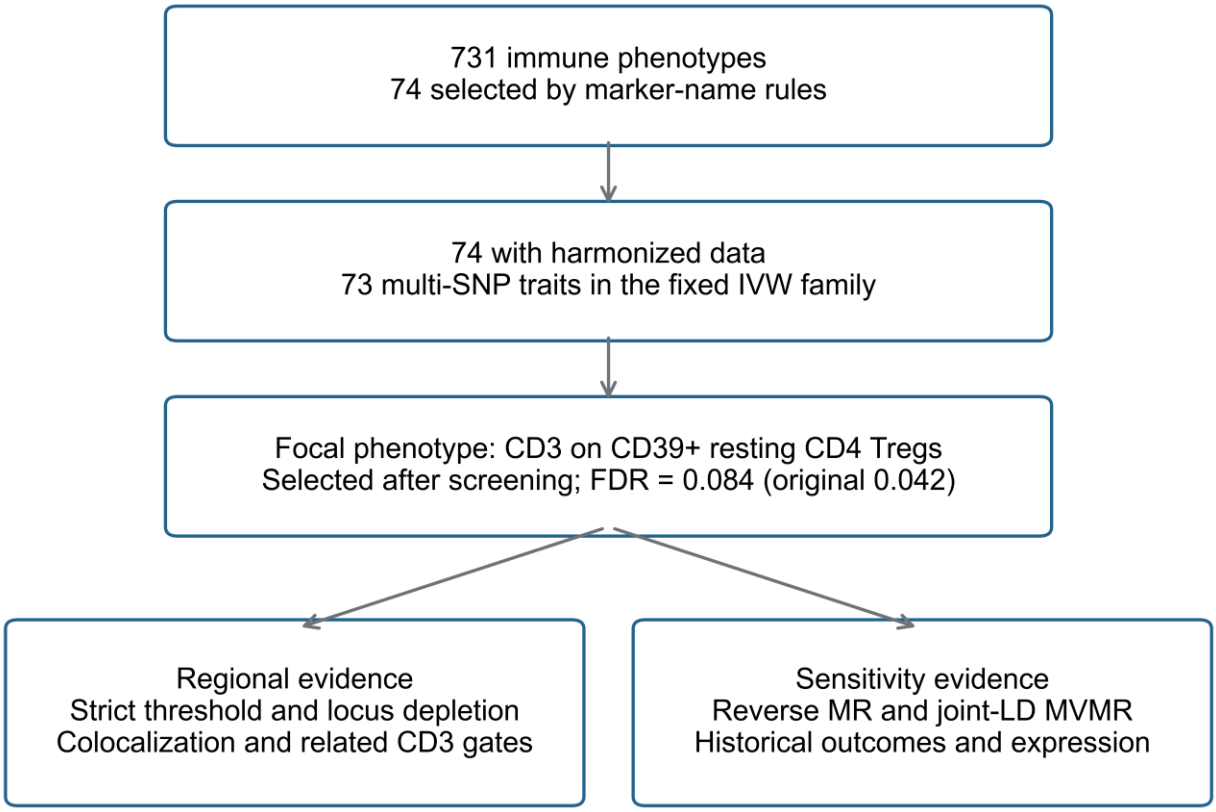


Fig. S4 Study flow and evidence roles. Marker-name rules selected 74 immune traits with harmonized data; 73 multi-instrument traits formed the original IVW screening family. The focal CD3 phenotype was followed up after screening. One single-SNP trait remained outside that family. Regional and sensitivity analyses were exploratory, with the original testing family retained. The three originally selected traits remain in the supplementary tables. The displayed FDR values refer to the focal CD3 phenotype under the revised and original variance conventions