nature portfolio

Corresponding author(s): Maria Koromina and Niamh Mullins

Last updated by author(s): 04/07/2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were generated at many sites using standard genotype calling softwares from commercial sources (Affymetrix and Illumina).
Data analysis	Analysis scripts are available online at [Github: https://github.com/mkoromina/SAFFARI]. Additional scripts to recreate the visuals/graphs are available online at [Github: https://github.com/Mullins-Lab/Post-finemap_processing/]. Other software used include: DENTIST [Github: https://github.com/Yves-CHEN/DENTIST], PolyPred [Github: https://github.com/omerwe/polyfun/wiki/6.-Trans-ethnic-polygenic-risk-prediction-with-PolyPred], PRS-CS [Github: https://github.com/getian107/PRScs], r2redux [Github: https://github.com/mommy003/r2redux] and RICOPILI [Github: https://github.com/Ripkelab/ricopili]. All software used is publicly available at the URLs or references cited.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

GWAS data were retrieved from (Mullins et al., 2021) from the following Figshare link: https://figshare.com/articles/dataset/PGC3_bipolar_disorder_GWAS_summary_statistics/14102594. The PGC's policy is to make genome-wide summary results public. All results are made available through the Figshare open access repository at the following DOI links: (https://doi.org/10.6084/m9.figshare.27871677.v2, https://doi.org/10.6084/m9.figshare.27880524.v1, https://doi.org/10.6084/m9.figshare.27886110.v1). Data provided include MHC fine-mapping analyses of the PGC3 BIP study, as well as aggregated fine-mapping results using various methods (PolyFun+SuSiE, PolyFun+FINEMAP, SuSiE, FINEMAP) across four LD reference panels (UKB, HRC, LD, noLD) and GWS locus windows, provided in both .txt.gz and .merged.csv formats. Additional files include genome-wide fine-mapping results from SuSiE and PRS-CS protocols, and a detailed Excel file on credible sets for 12 fine-mapping analyses, specifying the SNPs and loci involved (https://doi.org/10.6084/m9.figshare.28027706.v1).
Individual-level genetic data are accessible via Secondary Analysis Proposals to the Bipolar Disorder Working Group of the PGC (https://www.med.unc.edu/pgc/shared-methods/how-to/). This study included some publicly available datasets accessed through dbGaP - PGC bundle phs001254.v1.p1.
Additional annotations were retrieved from the following databases: gnomAD database v4.0.0 (https://gnomad.broadinstitute.org), CADD (https://cadd.gs.washington.edu/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	While sex and gender-based analyses are important, they are beyond the scope of the present study, which addresses a broader analytical framework for fine-mapping.
Reporting on race, ethnicity, or other socially relevant groupings	Only genetic ancestry information is used in this study.
Population characteristics	We did not have extensive access to detailed covariate-relevant population characteristics of the human research participants. As such, information such as age, socioeconomic status, and other relevant demographics was unavailable.
Recruitment	This is fully described in the Supplementary Note.
Ethics oversight	All local IRBs approved of this study. This is fully described in the Supplementary Note.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We used summary statistics and cohort data from the latest published bipolar disorder GWAS. Sample sizes in that GWAS were larger than previous analyses, which revealed some significant findings and showed that additional true positives remained to be discovered. Most GWAS fine-mapping methods require a sample size of at least 20,000 samples (Weissbrod et al., 2020) and the current one had an effective sample size of 101,962. Therefore, there is efficient sample size to perform robust fine-mapping.
Data exclusions	Predetermined phenotypic data exclusions, for both cases and controls, have been described in detail in the initial GWAS. In brief, genotype data exclusions were also predetermined and were performed for quality control; these included high missing call rate, high or low heterozygosity, inconsistent genotype versus clinical data sex, and ancestry outlier status based on visual inspection of genotype principal component analysis results.
Replication	All available cohorts of bipolar disorder cases and controls were included in this study. External cohorts were also used, where possible, to replicate findings in the polygenic risk scoring analyses. We also used different statistical genetics methods and integrated different publicly

	available biological datasets to further validate our findings. We also wish to note that all fine-mapping analyses have been performed 5 times confirming the robustness of the results.
Randomization	Randomization processes have been described in detail in the initial GWAS study. This information is not relevant to our study. In this study, we also performed association analyses in each dataset and then meta-analyzed across datasets. Ancestry covariates derived from genotype principal components analysis were included in association tests, which were logistic regression.
Blinding	Standard quality control and analysis pipelines were run such that blinding is not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	N/A
Study protocol	N/A
Data collection	N/A
Outcomes	N/A

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A