nature genetics

Article

https://doi.org/10.1038/s41588-025-02141-1

Rare loss-of-function variants in *HECTD2* and *AKAP11* confer risk of bipolar disorder

Received: 18 January 2024

Accepted: 25 February 2025

Published online: 25 March 2025

Check for updates

Thorgeir E. Thorgeirsson[1] ✉, Vinicius Tragante[1], Gardar Sveinbjornsson[1], Gudrun A. Jonsdottir[1], G. Bragi Walters[1], Erna V. Ivarsdottir[1], Gudny A. Arnadottir[1], Arni Sturluson[1], Brynjar O. Jensson[1], Run Fridriksdottir[1], Astros Th. Skuladottir[1], Gudmundur Einarsson[1], Gyda Bjornsdottir[1], Arni F. Gunnarsson[1], Rosa S. Gisladottir[1,2], Asgeir Sigurdsson[1], Asmundur Oddsson[1], Hakon Jonsson[1], Olafur Th. Magnusson[1], Hannes Helgason[1], Gudmundur Norddahl[1], Gudmar Thorleifsson[1], Magnus Haraldsson[3,4], Engilbert Sigurdsson[3,4], Hilma Holm[1], Gisli Masson[1], Daniel F. Gudbjartsson[1,5], Hreinn Stefansson[1], Patrick Sulem[1] & Kari Stefansson[1,4] ✉

Bipolar disorder is a highly heritable psychiatric disorder; genome-wide association studies of bipolar disorder have yielded over 60 risk loci harboring common variants. To harness the information contained in rare loss-of-function (LOF) variants, holding promise for informing on the underlying biology, we performed a variant burden analysis for bipolar disorder using gene-based aggregation of LOF variants in whole-genome sequencing data from Iceland (4,197 cases, more than 200,000 controls) and the UK Biobank (1,881 cases, 426,622 controls). We found that *HECTD2* was associated with bipolar disorder and confirmed it using the Bipolar Exome dataset. Meta-analysis with Bipolar Exome also revealed that LOF variants in *AKAP11* were associated with bipolar disorder. Both associations with bipolar disorder are new, but *AKAP11* has previously been associated with psychosis and schizophrenia. The products of *AKAP11* and *HECTD2* interact with GSK3β, a protein inhibited by lithium, the most effective mood stabilizer available to treat bipolar disorder.

Bipolar disorder is characterized by bouts of mania (or hypomania), usually accompanied by episodes of depression; it has a global prevalence of approximately 2% with some differences between populations[1]. Bipolar disorder is a serious psychiatric condition and has a large impact on individuals and society[1,2]; if untreated, it comes with a high suicide rate[3]. Several mood-stabilizing drugs are available for treating bipolar disorder. Lithium[4] is the best known, but anticonvulsants and antipsychotics are also used, as well as joint administration of antidepressants and antipsychotics[1]. The pharmacological treatments of bipolar disorder can all be accompanied by serious side effects, and better treatments are urgently needed. While the mood-stabilizing effects of lithium have been known for a long time, the underlying biology is not well understood[4,5]; this may have hampered necessary drug development. Among suggested targets of lithium is glycogen synthase kinase-3 beta (GSK3β), a serine threonine kinase with over 100 known substrates[6].

Bipolar disorder has high heritability, between 60% and 85% according to twin and family studies[7–9]; several genome-wide association studies (GWASs) have identified common variants associated with bipolar disorder, with a 2021 study reporting 64 associated genomic

[1]deCODE genetics/Amgen, Reykjavik, Iceland. [2]School of Humanities, University of Iceland, Reykjavik, Iceland. [3]Department of Psychiatry, Landspitali University Hospital, Reykjavik, Iceland. [4]Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. [5]Faculty of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. ✉e-mail: thorgeir@decode.is; kstefans@decode.is

https://doi.org/10.1038/s41588-025-02141-1

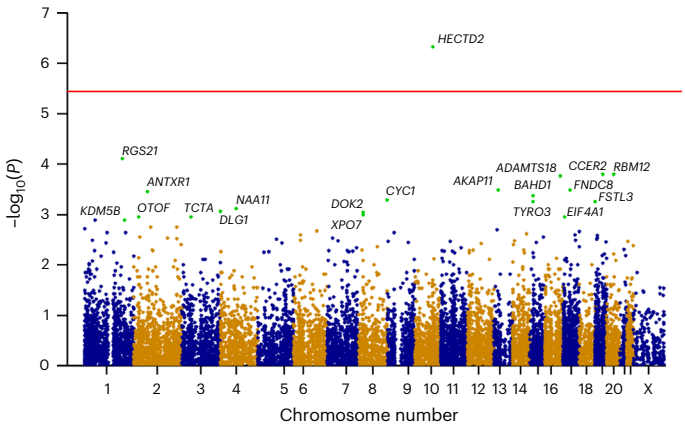

Fig. 1 | Gene-based Manhattan plot for the ICE-UKB meta-analysis. The plot shows $-\log_{10}P$ values against the chromosomal position of each of the genes studied. The top 20 findings are indicated by the green dots; the corresponding gene name is represented by the most proximal gene symbol. The results for the top 20 genes, for the meta-analysis and for each study (ICE and UKB), are provided in Supplementary Table 1. Logistic regression was used to test for association in ICE and UKB, obtaining the *P* values from a likelihood ratio test corrected for cryptic relatedness and stratification. The meta-analysis used the fixed-effect inverse variance-weighted method. The *P* values presented were not corrected for multiple testing, but the red line represents a Bonferroni threshold of $P \leq 3.6 \times 10^{-6}$.

loci[7]. Based on two recent GWASs, the common single-nucleotide polymorphism (SNP)-based heritability is around 20% (refs. 7,10). With increased efforts in whole-genome sequencing (WGS) and exome sequencing, it has become possible to study the role of rare high-impact variants. Often these variants are studied in aggregation at the gene level; genome-wide rare variant burden analyses for schizophrenia and bipolar disorder were recently performed in two separate studies[11,12]. The Schizophrenia Exome Sequencing Meta-analysis (SCHEMA)[11] study focused on the exomes of up to 24,248 schizophrenia cases and 97,322 controls, as well as de novo mutations from 3,402 parent–proband trios. The Bipolar Exomes (BipEx)[12] study analyzed the exomes of up to 14,210 cases with bipolar disorder and 14,422 controls. The results are available through online browsers (see 'Data availability'). While the SCHEMA study yielded several genes with high-penetrance risk variants for schizophrenia, the BipEx results did not yield any genes with significant association in a burden test; however, the top gene in the loss-of-function (LOF) burden analysis for bipolar disorder ($P = 1.15 \times 10^{-5}$), *AKAP11*, had a significant association in a joint analysis of bipolar disorder and schizophrenia[12] and was subsequently found to associate with schizophrenia[13].

Results

To expand the search for genes implicated in bipolar disorder, we performed a genome-wide LOF burden scan of bipolar disorder using data from Iceland (ICE) and the UK Biobank (UKB); we subsequently performed a meta-analysis of our results and those from BipEx. In our study, we combined two large genetic datasets of European ancestry: 428,503 whole-genome-sequenced British or Irish individuals from the UKB and 325,104 Icelanders, of whom 58,449 were whole-genome-sequenced. We tested the association between bipolar disorder and the burden of LOF variants in 13,786 genes. The number of tests performed amounted to a Bonferroni significance threshold of $P \leq 3.6 \times 10^{-6}$. We used the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) diagnoses for bipolar disorder (ICD-10 code F31 and subcodes) to define cases in the UKB from healthcare records, clinical records and general practitioners (GPs). Most of the Icelandic cases are defined based on ICD-10 diagnoses from electronic medical records on centralized databases containing diagnoses from Iceland's main hospitals, psychiatrists and GPs (Methods).

The full results for the ICE-UKB meta-analysis are available online (see 'Data availability'). The quantile–quantile plots for both meta-analyses (Supplementary Fig. 1) show that our tests are well calibrated. The results for the top 20 genes from both the ICE-UKB and the ICE-UKB-BipEx LOF burden meta-analyses (Supplementary Tables 1 and 2) are labeled with gene symbols on the Manhattan plots (Figs. 1 and 2); brief descriptions of the genes involved are provided in the Supplementary Note.

HECTD2 and *AKAP11* are associated with bipolar disorder

We determined that combined rare LOF variants in *HECTD2* were significantly associated with bipolar disorder based on the ICE-UKB analysis (frequency = 0.011%, odds ratio (OR) = 9.1, $P = 4.7 \times 10^{-7}$, 95% confidence interval (CI) = 3.85 to 21.49). We replicated the association of LOF variants in *HECTD2* with bipolar disorder using data from the BipEx study (13,933 cases and 14,422 controls, OR = Inf, $P = 0.0069$); the significance was greater when combining the two studies (overall $P = 4.1 \times 10^{-8}$) (Table 1 and Figs. 1 and 2). The ICE-UKB analysis included 45 LOF variants in *HECTD2* (140 carriers) and 38 in *AKAP11* (76 carriers) (Table 1 and Supplementary Table 3).

Systematic meta-analysis of our study and BipEx revealed one additional association with bipolar disorder: *AKAP11* (OR = 11.8, $P = 7.4 \times 10^{-9}$) (Table 1 and Fig. 2). For both *HECTD2* and *AKAP11*, we observed at least a nominal association in the same direction in all three groups tested; however, gene-burden-wide significance was not reached in any of them individually (ICE, UKB and BipEx) (Table 1).

HECTD2 and *AKAP11* are constrained in LOF variants

Genes can be classified according to how constrained they are with regard to LOF variants, estimated by comparing the number of observed versus expected very rare LOF variants, with genes scored according to the probability of being LOF intolerant (pLI) and the ratio of LOF observed/expected upper bound fraction; these metrics are available from the Genome Aggregation Database (gnomAD) for most genes.

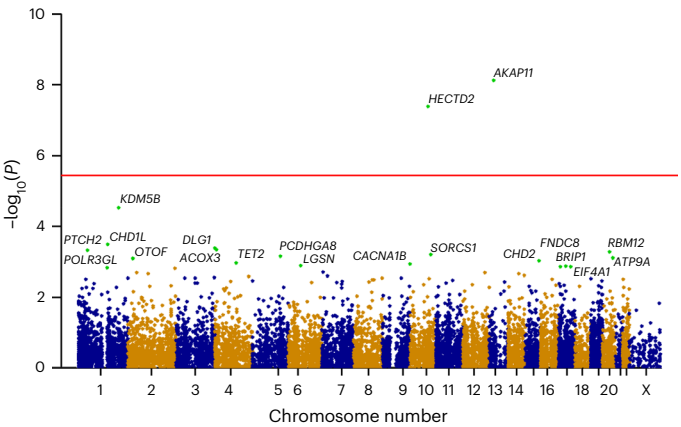

Fig. 2 | Gene-based Manhattan plot for the ICE-UKB-BipEx meta-analysis. The plot shows $-\log_{10}$ two-sided *P* values against the chromosomal position of each of the genes studied. The top 20 findings are indicated by the green dots; the corresponding gene name is represented by the most proximal gene symbol. Logistic regression was used to test for association in ICE and UKB; *P* values were obtained from a likelihood ratio test adjusted for cryptic relatedness and stratification. The results, including ORs, for the top 20 genes for the meta-analysis and for each study (ICE, UKB and BipEx) are provided in Supplementary Table 2. To obtain the overall *P* values, the results of the ICE-UKB fixed-effect meta-analysis (Fig. 1) were combined with the BipEx results on the *z*-score. The *P* values presented were not corrected for multiple testing, but the red line represents a Bonferroni threshold of $P \leq 3.6 \times 10^{-6}$.

https://doi.org/10.1038/s41588-025-02141-1

Table 1 | Results of the LOF burden analysis of bipolar disorder for *HECTD2* and *AKAP11*

Gene	Dataset	*n* variants	MAF ICE, UKB (%)	LOF carriers n_{tot}^{a} (n_{aff})[b]	OR	*P*	P_{het}^{c}	BipEx LOF carriers n_{tot} (n_{aff})	BipEx OR	BipEx *P*	ICE, UKB and BipEx *P*
HECTD2	Meta-analysis	45	–	145 (11)	9.1	4.8×10^{-7}	0.72	7 (7)	Infinity[d]	0.0069	4.1×10^{-8}
	ICE	3	0.012	46	10.1	1.4×10^{-5}					
	UKB	42	0.011	99	7.25	0.010					
AKAP11	Meta-analysis	38	–	80 (5)	11.8	3.3×10^{-4}	0.73	16 (16)	Infinity	1.2×10^{-5}	7.4×10^{-9}
	ICE	3	0.0024	7	17.4	0.031					
	UKB	35	0.0092	73	10.2	3.8×10^{-3}					

The results from ICE, UKB and the ICE-UKB meta-analysis are provided along with information on the frequency and number of variants and carriers. Logistic regression was used to test for association in ICE and UKB; the two-sided *P* values were obtained using a likelihood test. The last four columns provide the results from the BipEx study and the overall *P* value from the ICE-UKB-BipEx meta-analysis. The ORs provided for the meta-analyses are based on the ICE-UKB fixed-effect meta-analysis; the overall *P* value in the last column is based on combining the ICE-UKB results and BipEx on *z*-scores. The 95% CIs for the ORs from the ICE-UKB study are 3.8–21.5 for *HECTD2* and 3.1–45.4 for *AKAP11*. The fifth column shows the total number of carriers, along with the number of affected carriers in parentheses. The *P* values provided are all two-sided and have not been corrected for multiple testing. [a]Total number of carriers. [b]Number of carriers affected with bipolar disorder. [c]Heterogeneity *P* value. [d]No carriers among the controls in the BipEx study. To protect privacy, carrier numbers are not broken down according to case-control status in the ICE and UKB studies. MAF, minor allele frequency.

Enrichment of ultra-rare LOF variants in pLI genes (pLI ≥ 0.9) has been observed for bipolar disorder[12], although this pattern is far weaker than that observed for schizophrenia[11,13]. Both *AKAP11* and *HECTD2* are pLI genes (Supplementary Table 4).

Effects of LOF variants on cognition in *HECTD2* and *AKAP11*

A general intelligence factor (*g*-factor) was derived from the results of cognitive tests available for parts of both the ICE and UKB samples, as outlined previously[14]. When analyzed jointly for ICE and UKB, individuals with bipolar disorder, compared to controls, exhibited significantly lower cognitive performance (effect = −0.40, $P = 1.0\times10^{-18}$, 95% CI = −0.49 to −0.31). We next looked at carriers of LOF variants in *HECTD2* and *AKAP11*, focusing on carriers not diagnosed with bipolar disorder because cognitive measures were available for only one LOF carrier diagnosed with bipolar disorder (*AKAP11* in the UKB sample). In the combined analysis, there was nominal evidence for a lower *g*-factor among carriers of *HECTD2* LOF variants who had not been diagnosed with bipolar disorder (effect = −0.38, $P = 0.030$, 95% CI = −0.72 to −0.04). Cognitive tests were not available for carriers of *AKAP11* LOF variants in ICE, but there was no evidence for lowered *g*-factor among carriers of LOF variants in *AKAP11* (not diagnosed with bipolar disorder) in the UKB sample (effect = −0.09, $P = 0.76$, $n_{aff} = 12$, $n_{ctrl} = 74{,}558$, 95% CI = −0.67 to 0.49) (Supplementary Note).

Burden analysis combining LOF and damaging missense variants

We also performed a burden analysis jointly analyzing LOF variants combined with predicted deleterious missense variants or in-frame insertions (for the list of markers, see Supplementary Table 3b). Briefly, adding this group of variants to the analysis removed all association with *AKAP11* (OR = 0.92, $P = 0.42$) and did not improve that of *HECTD2* (OR = 6.35, $P = 3.4\times10^{-6}$). The results for the top 20 genes from this analysis revealed no new significant signals but some suggestive associations that may be of interest, although these require further study in larger samples (Supplementary Table 5).

Bipolar disorder is not especially enriched for variants in pLI genes

Enrichment of damaging variants in pLI genes has been reported for bipolar disorder[12], and both *HECTD2* and *AKAP11* are pLI genes. Hence, we explored whether there was overall enrichment of LOF or LOF + missense variants in bipolar disorder for this group of genes in our data, but did not find unequivocal evidence for particular enrichment of the whole group (Supplementary Note).

Differences in methodology

Burden studies involve rare variants and can thus be sensitive with regard to variant inclusion, experimental error or some underlying confounders, raising questions regarding using ICE, UKB and BipEx in a meta-analysis. BipEx itself is a meta-analysis of samples from several different studies and is thus subject to batch effects; however, the study implemented methods to minimize them[12]. We note that the UKB sample and around 40% of the ICE sample were whole-genome-sequenced with the same technology, with a minimum depth of at least 20× coverage, while the BipEx samples were whole-exome-sequenced with a somewhat higher coverage (~55×)[12]. Almost all variants identified using whole-exome sequencing were found using WGS[15], so these differences in methodology are not likely to affect the study. One weakness in our approach is that we may have overlooked a considerable fraction of de novo mutations in Iceland because the Icelandic sample was not fully whole-genome-sequenced. Another difference is that the summary statistics available from BipEx are from analyses confined to ultra-rare variants (minor allele count (MAC) ≤ 5)[12], and this type of filtering may lead to differences for some of the genes studied.

Verification of imputed carriers

Our methodology for sequencing and imputations in ICE has been described previously[16,17]. For this study, we used a high cutoff of genotype imputation probability (0.9); the genotypes with this quality of imputation are correct more than 90% of the time. Nevertheless, we tested all carriers from imputations of the rare LOF variants in *HECTD2* and *AKAP11* in ICE using Sanger sequencing (Methods). There were no imputed carriers for *AKAP11* and 14 for *HECTD2*. Sequencing did not work for one study participant, but the other 13 imputed carriers were all confirmed Sanger sequencing. In our Sanger sequencing analysis, we also included several individuals predicted as unlikely carriers ($P < 0.7$, $n = 1$, and $P < 0.1$, $n = 15$) and found no new, unexpected carriers.

Sensitivity to variant inclusion

In our analysis of LOF variants, we included start-loss and stop-loss variants in addition to stop-gain, essential splice and frameshift variants. We performed the LOF of the burden analyses in ICE and UKB excluding start-loss and stop-loss variants. The results indicated that variants in this category neither drive the signals nor significantly attenuate them. For *AKAP11*, we obtained an OR = 12.7 (95% CI = 3.8 to 42.7); for *HECTD2*, we obtained an OR = 9.1 (95% CI = 3. 9 to 21.3).

Sensitivity to psychiatric comorbidities

Comorbidity among individuals with bipolar disorder, particularly with schizophrenia and schizoaffective disorders, and the fact that

our phenotype definition is based on at least one diagnosis of bipolar disorder, raises the concern that the results might be driven by study participants who have also been diagnosed with other psychiatric disorders. To evaluate the contribution of schizophrenia or schizoaffective diagnoses on the results, we conducted burden analysis removing individuals with at least one diagnosis of either schizophrenia or schizoaffective disorder from both cases and controls. We found that the observed ORs were robust to this removal (Supplementary Note). We also performed an analysis of schizophrenia in UKB and ICE, removing all individuals with at least one bipolar disorder diagnosis from cases and controls. For *HECTD2*, no risk was observed for this subgroup as carriers diagnosed with schizophrenia all had dual diagnoses of bipolar disorder; for *AKAP11*, we obtained an OR = 4.3 (P = 0.25), suggestive of a residual risk consistent with LOF variants in *AKAP11* conferring a risk of schizophrenia[13].

Thus, for both *AKAP11* and *HECTD2*, the results presented in Table 1 represent estimates that are based on verified carriers of LOF variants and are driven by bipolar disorder without substantial contribution from carriers with dual diagnoses of bipolar disorder and schizophrenia or schizoaffective disorder. For *AKAP11* and *HECTD2*, we found no evidence for differences in the prevalence of anxiety or depression among carriers unaffected by bipolar disorder (Supplementary Note and Supplementary Table 6); however, further studies using larger samples are required for strong inference on questions relating to risk of other psychiatric disorders.

Discussion

By conducting gene-based LOF burden analyses using the ICE and UKB (British/Irish) samples, we discovered an association of LOF variants in *HECTD2* with bipolar disorder; we confirmed this finding using a meta-analysis with data from BipEx. We also found an additional association with LOF variants in *AKAP11* in the three groups combined. The largest published sequencing-based effort is the BipEx study[12]. The ICE-UKB study, despite a lower number of bipolar disorder cases, has an effective sample size (n_{eff} = 24,034) similar to that of the BipEx (n_{eff} = 28,347). Larger samples are needed for further discovery of variants; the results from both the ICE-UKB study and BipEx will be useful resources for those conducting analyses of bipolar disorder in the future.

Bipolar disorder and schizophrenia partially share their genetics, with a genetic correlation of ~0.6 (ref. 18). LOF variants in *AKAP11* have previously been implicated in a joint analysis of bipolar disorder and schizophrenia[12], and in an analysis of only schizophrenia patients[13]. Our study demonstrated an association with bipolar disorder. While the association of *AKAP11* with schizophrenia has been confirmed[13], there is currently no evidence for an association between schizophrenia and *HECTD2* (SCHEMA OR = 1.61, P = 0.23, 95% CI = 0.74 to 3.50). The results of the sensitivity analyses showed that our results for both *AKAP11* and *HECTD2* are driven by bipolar disorder. We found no evidence for a different prevalence of other common psychiatric disorders among carriers not diagnosed with bipolar disorder. As sample sizes increase, it will be possible to further delineate the relationship between LOF burden in *AKAP11* and *HECTD2* and the risks of other psychiatric disorders through studies focused on, or excluding, individuals diagnosed with other psychiatric conditions.

The evaluation of cognitive function (*g*-factor) in carriers of the LOF variants in *HECTD2* and *AKAP11* did not reveal large differences between carriers and noncarriers; however, the number of carriers studied was small and based on a combined analysis in ICE and UKB. *HECTD2* carriers had at least a subtle impairment of cognition independent of disease state. Both AKAP11 and HECTD2 have, through functional studies, been tied to GSK3β, a suggested target of lithium. AKAP11 is a scaffolding protein acting to localize protein kinase A to facilitate phosphorylation of proteins at particular cellular locations; it also binds GSK3β[19]. Protein kinase A-mediated inhibition of AKAP11-bound GSK3β is stronger than for unbound GSK3β[19]. Thus, AKAP11 both localizes and moderates GSK3β activity. HECTD2 is an understudied E3 ubiquitin ligase; its substrates are largely unknown. However, HECTD2 ubiquinates protein inhibitor of activated STAT1 (PIAS1), thereby marking it for proteasomal degradation[20]; GSK3β-mediated phosphorylation is required to initiate the ubiquitination of PIAS1 by HECTD2 (ref. 20). PIAS1 is a transcriptional coregulator of many cellular pathways, including STAT1, p53, steroid hormone signaling and nuclear factor kappaB pathways[21]. PIAS1 and other PIAS proteins are E3 ligases that mediate posttranslational modification of several target proteins with small ubiquitin-like modifiers (SUMOs)[21]. Although PIAS1 has been studied extensively, not all substrates are known. PIAS1 has been shown to perform SUMOylation of CREB1 (ref. 22), FOXP2 (ref. 23) and group III metabotropic glutamate receptors[24]. Furthermore, products of the splice variants of *PIAS* genes (PIASxα) SUMOylate the serotonin 1A receptor[25].

Further functional studies, identifying the substrates of HECTD2 in particular, are required to understand the rich biology involved, and how GSK3β, AKAP11 and HECTD2 act to influence the regulation of proteins and pathways relevant to the underlying biology of bipolar disorder (Supplementary Fig. 2). The association of rare variants in two genes that are functionally linked to GSK3β with bipolar disorder is interesting in light of the hypothesis that GSK3β may be the target of lithium in the treatment of the disorder. Furthermore, one of the limiting aspects of lithium in long-term treatment of bipolar disorder is nephrotoxicity, which is common and somewhat dose-dependent[26]. This toxicity may prevent lithium from providing the full benefit that could be obtained by inhibiting specific targets. Hence, GSK3β and the products of *HECTD2* and *AKAP11* have to be considered as promising targets in the search for new treatments for bipolar disorder.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-025-02141-1.

References

1. Nierenberg, A. A. et al. Diagnosis and treatment of bipolar disorder: a review. *JAMA* **330**, 1370–1380 (2023).
2. Crump, C., Sundquist, K., Winkleby, M. A. & Sundquist, J. Comorbidities and mortality in bipolar disorder: a Swedish national cohort study. *JAMA Psychiatry* **70**, 931–939 (2013).
3. Jamison, K. R. Suicide and bipolar disorder. *J. Clin. Psychiatry* **61**, 47–51 (2000).
4. Beaulieu, J.-M. & Caron, M. G. Looking at lithium: molecular moods and complex behaviour. *Mol. Interv.* **8**, 230–241 (2008).
5. Freland, L. & Beaulieu, J.-M. Inhibition of GSK3 by lithium, from single molecules to signaling networks. *Front. Mol. Neurosci.* **5**, 14 (2012).
6. Beurel, E., Grieco, S. F. & Jope, R. S. Glycogen synthase kinase-3 (GSK3): regulation, actions, and diseases. *Pharmacol. Ther.* **148**, 114–131 (2015).
7. Mullins, N. et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).
8. Bienvenu, O. J., Davydow, D. S. & Kendler, K. S. Psychiatric 'diseases' versus behavioral disorders and degree of genetic influence. *Psychol. Med.* **41**, 33–40 (2011).
9. Song, J. et al. Bipolar disorder and its relation to major psychiatric disorders: a family-based study in the Swedish population. *Bipolar Disord.* **17**, 184–193 (2015).
10. Stahl, E. A. et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).

11. Singh, T. et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* **604**, 509–516 (2022).
12. Palmer, D. S. et al. Exome sequencing in bipolar disorder identifies *AKAP11* as a risk gene shared with schizophrenia. *Nat. Genet.* **54**, 541–547 (2022).
13. Liu, D. et al. Schizophrenia risk conferred by rare protein-truncating variants is conserved across diverse human populations. *Nat. Genet.* **55**, 369–376 (2023).
14. Jonsdottir, G. A. et al. Genetic propensities for verbal and spatial ability have opposite effects on body mass index and risk of schizophrenia. *Intelligence* **88**, 101565 (2021).
15. Halldorsson, B. V. et al. The sequences of 150,119 genomes in the UK Biobank. *Nature* **607**, 732–740 (2022).
16. Gudbjartsson, D. F. et al. Large-scale whole-genome sequencing of the Icelandic population. *Nat. Genet.* **47**, 435–444 (2015).
17. Gudbjartsson, D. F. et al. Sequence variants from whole genome sequencing a large group of Icelanders. *Sci. Data* **2**, 150011 (2015).
18. Cardno, A. G. & Owen, M. J. Genetic relationships between schizophrenia, bipolar disorder, and schizoaffective disorder. *Schizophr. Bull.* **40**, 504–515 (2014).
19. Tanji, C. et al. A-kinase anchoring protein AKAP220 binds to glycogen synthase kinase-3β (GSK-3β) and mediates protein kinase A-dependent inhibition of GSK-3β. *J. Biol. Chem.* **277**, 36955–36961 (2002).
20. Coon, T. A. et al. The proinflammatory role of HECTD2 in innate immunity and experimental lung injury. *Sci. Transl. Med.* **7**, 295ra109 (2015).
21. Rytinki, M. M., Kaikkonen, S., Pehkonen, P., Jääskeläinen, T. & Palvimo, J. J. PIAS proteins: pleiotropic interactors associated with SUMO. *Cell. Mol. Life Sci.* **66**, 3029–3041 (2009).
22. Chen, Y.-C., Hsu, W.-L., Ma, Y.-L., Tai, D. J. C. & Lee, E. H. Y. CREB SUMOylation by the E3 ligase PIAS1 enhances spatial memory. *J. Neurosci.* **34**, 9574–9589 (2014).
23. Estruch, S. B., Graham, S. A., Deriziotis, P. & Fisher, S. E. The language-related transcription factor FOXP2 is post-translationally modified with small ubiquitin-like modifiers. *Sci. Rep.* **6**, 20911 (2016).
24. Tang, Z., El Far, O., Betz, H. & Scheschonka, A. Pias1 interaction and sumoylation of metabotropic glutamate receptor 8. *J. Biol. Chem.* **280**, 38153–38159 (2005).
25. Gupta, S., Wang, M., Azuma, Y. & Muma, N. A. Regulation of serotonin 1A receptor SUMOylation by SENP2 and PIASxα. *Int. J. Mol. Sci.* **22**, 13176 (2021).
26. Castro, V. M. et al. Stratifying risk for renal insufficiency among lithium-treated patients: an electronic health record study. *Neuropsychopharmacology* **41**, 1138–1143 (2016).

Methods

Ethical considerations

The use of Icelandic data was approved by the National Bioethics Committee (NBC) (nos. VSN-16-067 and VSN-21-24). All genotyped participants signed a written informed consent allowing the use of their samples and data in projects at deCODE genetics approved by the NBC. Data were anonymized and encrypted using a third-party system, approved and monitored by the Icelandic Data Protection Authority[27].

The UKB data were obtained under application no. 42256. All participants provided written informed consent for the use of genotype data and link to electronic health records. The North West Research Ethics Committee reviewed and approved the UKB protocol (ref. 06/MRE08/65).

Sample descriptions

The genomes for ICE were characterized using WGS of 58,449 Icelanders, using the Illumina standard TruSeq methodology to a mean depth of 35× (s.d. = 8×) with subsequent long-range phasing[28] and imputing the information into 155,241 individuals chip-genotyped using multiple Illumina platforms[16], and using familial imputation in up to ~170,000 Icelandic individuals. Phenotypic data come from: (1) Landspitali-the National University Hospital of Iceland in Reykjavik; (2) primary healthcare clinics of the Capital area; (3) specialists in private practice (psychiatry); and (4) recruitment for the deCODE studies. International Statistical Classification of Diseases and Related Health Problems, Ninth Revision (ICD-9) and ICD-10 codes were available from electronic records between 1987 and 2022. Older records were accessed by the Department of Psychiatry at the National University Hospital of Iceland as part of ongoing genetic studies. In total, 4,197 study participants had at least one diagnosis of bipolar disorder (1,739 males and 2548 females), with year of birth ranging from 1890 to 2005. Most of the data (n = 4,056) came from electronic records (ICD-10 F31 and subcodes and ICD-9 codes 296.0, 296.2, 296.3, 296.4 and 296.5) from psychiatrists and hospital records (n = 2,969); and from primary healthcare clinics (n = 2,903), with 1,816 in both datasets. The remaining information came through genetics research projects aimed at depression and anxiety (n = 128), and schizophrenia and bipolar disorder (n = 380), which used a combination of patient records and various diagnostic interviews using a number of diagnostic systems. Most of the individuals diagnosed through these studies also had ICD diagnoses from the sources detailed above (n = 367). The controls used were population controls and included individuals diagnosed with other psychiatric conditions than bipolar disorder.

The UKB dataset consists of 428,503 WGS White British or Irish individuals (as identified using principal component analysis)[15,29] who enrolled in the study between 2006 and 2010 throughout the UK and were aged 40–65 years at recruitment. Phenotypic data provided by the UKB were primarily from hospital records and increasingly from GPs. The analysis included 1,881 individuals with at least one instance of the ICD-10 F31 diagnosis code or any of its subcodes. This included 811 males and 1,070 females, 1,704 from the hospital inpatient diagnosis data (field 41234) and 619 from primary care clinical event records (field 42040), with 442 in both datasets. The controls used were population controls and included individuals diagnosed with other psychiatric conditions than bipolar disorder.

Variant quality control and annotation

Variant calling was performed using GraphTyper v.2.6 (ref. 30), and chip data were phased using SHAPEIT4 (ref. 31). Only high-quality sequence variants were considered for selection. To estimate the quality of the sequence variants, we regressed the alternative allele counts (AD) on the depth (DP) conditioned on the genotypes (GT) reported by GraphTyper. We filtered variants with a slope of less than 0.5. Additionally, we only included variants with a GraphTyper AAscore greater than 0.8 and minor allele frequency smaller than 0.02.

We used Variant Effect Predictor[32] to attribute predicted consequences to the variants sequenced in each dataset. We classified as high-impact variants those predicted as start-loss, stop-gain, stop-loss, splice donor, splice acceptor or frameshift, collectively called LOF variants.

Association analyses

For the case-control analyses, we used logistic regression under an additive model to test for association between gene-level LOF variant burden and bipolar disorder. Disease status was the dependent variable and genotype counts was the independent variable. A likelihood ratio test was used to compute two-sided P values. For sequenced individuals, all identified LOF variants passing the quality control criteria were used; for imputed individuals, LOF variants were used if the imputation information was greater than 0.7. Sequenced or imputed individuals were coded with genotype count 1 if they carried any such LOF variants in the autosomal gene being tested and their expected allele count was greater than 0.9 (0 otherwise). For ICE, the expected genotypes of first-degree and second-degree relatives of participants donating blood was incorporated in the analysis by integrating over possible genotypes. Family imputation occurred in silico, that is, it was performed without the genotypes being kept in storage. For ICE, sex, age, sequencing status and county of origin (equivalent to principal components) were used as covariates. For the UKB, 20 principal components were used to adjust for population substructure; age and sex were included as covariates in the logistic regression model. For the analyses, we used software developed at deCODE genetics[16]. We used linkage disequilibrium score regression intercepts[33] to adjust the chi-squared statistics and avoid inflation due to cryptic relatedness and stratification, using a set of 1.1 million variants. P values were calculated from the adjusted chi-squared results.

The meta-analysis was performed on the summary results from ICE and UKB, when available, using a fixed-effect inverse variance-weighted method[34], in which the datasets were allowed to have different population frequencies for alleles and genotypes but were assumed to have a common OR and weighted with the inverse of the variance of the effect estimate derived from the logistic regression.

Sanger sequencing

Primers were designed using Primer 3 (https://www.broadinstitute.org). PCR and cycle sequencing reactions were performed using MJ Research PTC-255 Thermal Cyclers and the Big Dye Terminator Cycle Sequencing Kit v.3.1 (Thermo Fisher Scientific) and Agencourt AMPure XP and CleanSeq kits (Beckman Coulter) to clean up the PCR products. The sequencing products were loaded onto the 3730 XL Genetic Analyzer (Applied Biosystems) and analyzed with the Sequencher software v.5.4.6 (GeneCodes Corporation).

Enrichment of variants in pLI genes in bipolar disorder

We tested for an association between a bipolar disorder diagnosis and the number of conserved genes (pLI > 0.8) in which each individual had putative LOF(s) using logistic regression. An additive model was used, where homozygotes were counted 2× compared to heterozygotes. We adjusted for the number of genes where an individual had putative LOF(s), that is, not restricting to conserved ones, and for age, sex and county of origin in ICE, and principal components in the UKB.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Summary data of the gene burden results for the ICE-UKB meta-analysis of bipolar disorder are available on the deCODE genetics website (https://www.decode.com/summarydata/). For those using this

resource for discovery, deCODE will provide the necessary support regarding details on the individual markers involved while ensuring the protection of privacy, and will be available for more extensive collaboration, such as follow-up studies on correlations with other phenotypes than bipolar disorder and potential recruitment of carriers for evaluation. Our previously described Icelandic population WGS data have been deposited at the European Variation Archive under accession no. PRJEB15197. The UKB data were downloaded under application no. 42256. WGS, genotype and phenotypic data for UKB participants can be accessed by approved researchers via the UKB Research Analysis Platform (https://ukbiobank.dnanexus.com/landing). Guidance on access can be found at the UKB website (application for access required; ukbiobank.ac.uk). Other data generated and analyzed in this study are included in the article and the Supplementary Information. Other information and data used are as follows: SCHEMA (https://schema.broadinstitute.org/); BipEx (https://bipex.broadinstitute.org/); gnomAD (https://gnomad.broadinstitute.org/); GWAS Catalog (https://www.ebi.ac.uk/gwas/home; 26 October 2021 release) for reported GWAS associations; and NCBI (gene) (https://www.ncbi.nlm.nih.gov/gene/).

Code availability

We used publicly available software, which is available upon request from the following sources: graphtyper (v.2.0-beta; https://github.com/DecodeGenetics/graphtyper); Eagle (v.2.4.1; http://www.hsph.harvard.edu/alkes-price/software/); SHAPEIT4 (https://odelaneau.github.io/shapeit4/); ADMIXTURE (v.1.23; https://dalexander.github.io/admixture/); BOLT-LMM (v.2.1; http://www.hsph.harvard.edu/alkes-price/software/); R (v.3.6.3; https://www.r-project.org/); ggplot for data visualization (v.3.3.3; https://ggplot2.tidyverse.org/); Ensembl (v.87; https://www.ensembl.org/index.html); dbSNP (v.140; https://www.ncbi.nlm.nih.gov/snp/); kallisto (v.0.46; https://github.com/pachterlab/kallisto); ensembl-vep (release 100; https://github.com/Ensembl/ensembl-vep); BioRender (https://www.biorender.com/) accessed on 30 January 2025; and NCBI Build 38 (https://www.ncbi.nlm.nih.gov/). No custom code was written for this study.

References

27. Gulcher, J. R., Kristjánsson, K., Gudbjartsson, H. & Stefánsson, K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet.* **8**, 739–742 (2000).
28. Kong, A. et al. Detection of sharing by descent, long-range phasing and haplotype imputation. *Nat. Genet.* **40**, 1068–1075 (2008).
29. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
30. Eggertsson, H. P. et al. GraphTyper2 enables population-scale genotyping of structural variation using pangenome graphs. *Nat. Commun.* **10**, 5402 (2019).
31. Delaneau, O., Zagury, J.-F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T. Accurate, scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019).
32. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
33. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
34. Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl Cancer Inst.* **22**, 719–748 (1959).

Acknowledgements

We thank all participants who contributed the data and samples used in this study. Their contributions are essential for research such as that reported in this article. We thank all investigators and colleagues who collaborated on the many aspects of this study, including data collection, sample handling, phenotypic characterization of clinical samples, genotyping and analysis of the whole-genome association data. We thank participants and investigators of the UKB study. The UKB Resource is a major biomedical database (application no. 42256; https://www.ukbiobank.ac.uk/). This work uses linked data from NHS England and Public Health Scotland, provided by patients and collected by the NHS as part of their care and support, and data assets made available by National Safe Haven as part of the Data and Connectivity National Core Study led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (grant nos. MC_PC_20029 and MC_PC_20058).

Author contributions

T.E.T., V.T., G.S., H.J., D.F.G., H.S., P.S. and K.S. designed the study and interpreted the results. T.E.T., G.B., G.B.W., R.S.G., D.F.G., P.S., H.S. and K.S. collected and managed the Icelandic phenotypes and samples for the study. T.E.T., V.T., G.S., G.A.J., G.B.W., E.V.I., G.A.A., A. Sturluson, B.O.J., R.F., A.Th.S., G.B., A.F.G., A. Sigurdsson, H.J., H. Helgason, A.O., G.E., O.Th.M., G.N., G.T., D.F.G., H.S. and P.S. analyzed the data. M.H. and E.S. conducted the clinical work and recruited the study participants. H. Holm and G.A.J. oversaw recruitment and cognitive phenotyping of carriers and noncarriers. T.E.T., V.T., H.S., P.S. and K.S. drafted the manuscript with input from the other authors, who reviewed and contributed to the final version of the manuscript. H. Holm, G.M., D.F.G., H.S., P.S. and K.S. provided oversight and direction for the project.

Competing interests

Authors affiliated with deCODE genetics/Amgen declare competing financial interests as employees: T.E.T., V.T., G.S., G.A.J., G.B.W., E.V.I., G.A.A., A. Sturluson, B.O.J., R.F., A.Th.S., G.E., G.B., A.F.G., A. Sigurdsson, A.O., H.J., O.Th.M., H. Helgason, G.N., G.T., H. Holm, G.M., D.F.G., H.S., P.S. and K.S. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-025-02141-1.

Correspondence and requests for materials should be addressed to Thorgeir E. Thorgeirsson or Kari Stefansson.

Peer review information *Nature Genetics* thanks Tadafumi Kato and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature portfolio

Corresponding author(s): Thorgeir E. Thorgeirsson

Last updated by author(s): Feb 13, 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	No software was used for data collection
Data analysis	In combination with methods developed at deCODE Genetics as described in the methods section, we used publicly available software that is available on request under the following URLs: GraphTyper (v2.0-beta, GNU GPLv3 license), https://github.com/DecodeGenetics/graphtyper; Eagle (version 2.4.1), http://www.hsph.harvard.edu/alkes-price/software/; SHAPEIT4, https://odelaneau.github.io/shapeit4/; ADMIXTURE (v1.23), https://dalexander.github.io/admixture/; BOLT-LMM (v.2.1), http://www.hsph.harvard.edu/alkes-price/software/; R (version 3.6.3), https://www.r-project.org/; R package ggplot for visualization (version 3.3.3), https://ggplot2.tidyverse.org/; Ensembl v.87, https://www.ensembl.org/index.html; dbSNP v.140, http://www.ncbi.nlm.nih.gov/SNP/; kallisto v.0.46, https://github.com/pachterlab/kallisto; VEP (release100), https://github.com/Ensembl/ensembl-vep; Biorender, https://www.biorender.com/ (Publication license no MW27UR4KHP, accessed on Jan 30, 2025); NCBI Build 38, https://www.ncbi.nlm.nih.gov/. No custom code was written for this project.

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

Previously described Icelandic WGS data have been deposited at the European Variant Archive under accession PRJEB15197 (https://www.ebi.ac.uk/ena/browser/view/PRJEB15197). The summary statistics for the LOF burden analysis of the Icelandic and UK datasets will be available at https://www.decode.com/summarydata/. The UKB data are available at the UK biobank, and were downloaded under application no 42256. Other data generated or analyzed are included in the article and its supplements.

The full Data Availability statement is as follows:
Data availability
Summary data of the gene-burden meta-analysis results for the ICE-UKB meta-analysis of bipolar disorder will be available at the time of publication on the deCODE genetics website (https://www.decode.com/summarydata/). For those utilizing this resource for discovery, deCODE will provide necessary support regarding details on individual markers involved while ensuring the protection of privacy, and will be available for more extensive collaboration, such as follow-up studies on correlations with other phenotypes than bipolar disorder, and potential recruitment of carriers for evaluation.
Our previously described Icelandic population whole-genome sequence data have been deposited at the European Variant Archive under accession PRJEB15197. The UKB data were downloaded under application 42256. WGS, genotype and phenotypic data for UKB participants can be accessed by approved researchers via the UKB research analysis platform: https://ukbiobank.dnanexus.com/landing. Guidance on access can be found here: apply for access (ukbiobank.ac.uk). Other data generated / analyzed in this study are included in the article and the Supplementary Information. URLs for other information /data used are as follows:
SCHEMA: https://schema.broadinstitute.org/
BipEx: https://bipex.broadinstitute.org/
gnomAD: https://gnomad.broadinstitute.org/
GWAS catalog : (https://www.ebi.ac.uk/gwas/home 26/10/2021 release) for reported GWAS associations.
NCBI (gene):https://www.ncbi.nlm.nih.gov/gene/

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	No sex-specific results are reported.
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	The study is based on two populations of European descent from Iceland (N = up to ~325,000) and the British / Irish part of the UK biobank (N = 428,503 WGS subjects) (defined by principal component analysis). Genetic ancestry filtering and principal components determining European ancestry are described in the methods. The Icelandic sample had 58,449 WGS subjects, with imputed genotypes available for 155,241 chip genotyped subjects, and up to 170,000 subjects through familial imputation. In the UK biobank cases were defined by ICD10: F31 and subcodes, and included 811 males and 1070 females, 1704 from the hospital inpatient diagnosis data (field 41234) and 619 from primary care clinical event records (field 42040), with 442 in both datasets. In Iceland cases (N = 4,197) were defined using phenotype data from (i) Landspitali - The National University Hospital of Iceland (LUH) in Reykjavik; (ii) Primary Health Care Clinics of the Capital area, (iii) specialists in private practice (psychiatry), and (iv) recruitment for deCODE studies. Most of the cases (N = 4,056) were defined based on electronic records (ICD10-F31 and subcodes and/or ICD9 codes 296.0, 296.2, 296.3, 296.4 and 296.5) from psychiatrists and hospital records (N = 2,969) and Primary Health Care Clinics (2,903) with 1,816 in both datasets. The Icelandic cases were comprised of 1739 males and 2548 females with year of birth ranging from 1890 to 2005.
Recruitment	Recruitment and phenotype assessment is described in the methods section, and above under Population characteristics. Briefly, UK subjects enrolled in the UK biobank study between 2006 and 2010 throughout the UK and were aged 40-65 years at recruitment, and the study has extensive phenotype information available on about 500k participants. Icelandic subjects providing DNA samples (N ~ 155k) were recruited through various research projects ongoing since 1996, with over half of the adult population participating.
Ethics oversight	Use of Icelandic data was approved by the National Bioethics Committee (VSN-16-067 and VSN-21-24). All genotyped participants signed a written informed consent allowing the use of their samples and data in projects at deCODE genetics approved by the NBC. Data were anonymized and encrypted by a third-party system, approved and monitored by the Icelandic Data Protection Authority. The UKB data were obtained under application number 42256. All participants provided an informed consent for the use of genotype data and link to electronic health records (EHR). The North West Research Ethics Committee reviewed and approved the UKB protocol (ref. 06/MRE08/65).

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	Sample sizes corresponding to all available data are reported in the manuscript.
Data exclusions	Data from participants of non-European descent were excluded from the study as described in the manuscript, but no available data from the remaining subjects were excluded.
Replication	We used the BipEx study (https://bipex.broadinstitute.org/results) to replicate a signal from the joint analysis of Icelandic and UK biobank data. Systematic meta-analysis of the joint study and BipEx was then conducted, identifying an additional signal. No further replication studies were attempted, as the additional signal was highly significant with support from all three samples.
Randomization	Not applicable. The study is a gene burden association study, not a randomized trial.
Blinding	Not relevant to the study, as there was no grouping performed prior to analysis. The DNA sequencing, allele calling, and imputations were blind to phenotype, and hence no need for additional blinding.

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*