Title: The genetic architecture of human cerebellar morphology supports a key role for the cerebellum in human evolution and psychopathology

Short title: Genetic architecture of cerebellar morphology

Authors: Torgeir Moberget*[1,2], Dennis van der Meer†[1,3], Shahram Bahrami†[1], Daniel Roelfs[1], Oleksandr Frei[1], Tobias Kaufmann[1,4], Sara Fernandez-Cabello[1], Milin Kim[1], Thomas Wolfers[1,4], Joern Diedrichsen[5], Olav B. Smeland[1], Alexey Shadrin[1], Anders Dale[6], Ole A. Andreassen[1] & Lars T. Westlye[1,7].

*Corresponding author

†These two authors contributed equally

Affiliations:

1: Norwegian Centre for Mental Disorders Research (NORMENT), Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway

2: Department of Behavioral Science, School of Health Sciences, Oslo Metropolitan University - OsloMet, Oslo, Norway.

3: School of Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands

4: Department of Psychiatry and Psychotherapy, Tübingen Center for Mental Health, University of Tübingen, Germany

5: Brain and Mind Institute, University of Western Ontario, London, Canada

6: Center for Multimodal Imaging and Genetics, University of California at San Diego, La Jolla, CA, USA

7: Department of Psychology, Faculty of Social Sciences, University of Oslo, Norway

Abstract:
The functional domain of the cerebellum has expanded beyond motor control to also include cognitive and affective functions. In line with this notion, cerebellar volume has increased over recent primate evolution, and cerebellar alterations have been linked to heritable mental disorders. To map the genetic architecture of human cerebellar morphology, we here studied a large imaging genetics sample from the UK Biobank (n discovery = 27,302; n replication: 11,264) with state-of-the art neuroimaging and biostatistics tools. Multivariate GWAS on regional cerebellar MRI features yielded 351 significant genetic loci (228 novel, 94% replicated). Lead SNPs showed positive enrichment for relatively recent genetic mutations over the last 20-40k years (i.e., overlapping the Upper Paleolithic, a period characterized by rapid cultural evolution), while gene level analyses revealed enrichment for human-specific evolution over the last ~6-8 million years. Finally, we observed genetic overlap with major mental disorders, supporting cerebellar involvement in psychopathology.

Teaser: Genome-wide analysis of cerebellar morphology reveals links to recent human evolution and psychopathology

Introduction

The cerebellum contains ~80% of all neurons in the human brain[1] and has rapidly expanded in volume over recent primate evolution[2]. Indeed, the surface area of the cerebellar cortex extends to almost 80% of the surface area of the cerebral cortex[3]. Comparative genetic analyses suggest that protein coding genes with known roles in cerebellar development have been subject to a similar, or even greater, rate of hominid evolution as compared to cerebro-cortical developmental genes[4]. Thus, the evolution of the cerebellum may have played a key role in the emergence of human cognition, including language[5].

A growing number of neuroimaging and clinical studies in humans also link cerebellar structure and function to a wide range of cognitive and affective functions[6-8], as well as to a number of heritable developmental[9] and psychiatric[10] disorders where these abilities either fail to develop properly or are compromised later in life. However, compared to supra-tentorial brain structures such as the cerebral cortex[11] and the hippocampus[12], few studies have mapped the genetic architecture of the cerebellum.

Of note, the few existing cerebellar genome-wide association studies (GWAS) have mostly been restricted to total cerebellar volume[13,14], thus largely ignoring regional variation in cerebellar morphology. Importantly, such variation in the relative volumes of cerebellar subregions (i.e., variation in cerebellar shape independent of total cerebellar volume) has been associated with variation in behavioral repertoires in several species[15,16], including domain-general cognition in primates[16].

We here performed a multivariate GWAS of MRI-derived regional cerebellar morphological features in a large population-based sample from the UK biobank (n discovery = 27,302; n replication = 11,264), functionally characterized the genetic signal, tested for enrichment of SNPs and genes linked to human evolution, and assessed genetic overlap with major mental disorders.

Results

Data-driven decomposition of cerebellar grey matter maps reveals highly reproducible structural covariance patterns.

Since traditional atlases of the cerebellar cortex based on gross anatomical landmarks (i.e., lobules) only partially overlap with more recent functional parcellations of the cerebellum[17-25], we first used a data-driven approach (non-negative matrix factorization, NNMF[26-28]) to parcellate voxel-based morphometry (VBM) based maps of cerebellar grey matter volume (containing 147,121 1*1*1 mm voxels) from 28,212 participants into a smaller number of structural covariance patterns (SCPs), i.e., cerebellar sub-regions where voxel-volumes co-vary consistently across individuals (see Online Methods for details about the study sample, MRI processing and quality control procedures). Like other dimensionality reduction methods, such as principal component analysis (PCA) or independent component analysis (ICA), NNMF decomposes the input matrix (here 147k voxels * 28k participants)

into two lower rank matrices (voxel weights: W, and participant weights: H), so that the product of W and H approximates the original input data. The defining feature of NNMF is that it requires these lower rank matrices to be non-negative, which has previously been shown to result in sparse, reproducible and easily interpretable parcellations of high-dimensional brain imaging data[26-30]. In essence, when applied to voxel-wise indices of grey matter volume, NNMF yields distinct maps of voxels that show similar patterns of volume-variation across individuals, commonly referred to as "structural covariance patterns"[26,27,29]. For each structural covariance pattern (common across all participants), NNMF also provides individual subject weighs expressing the degree to which these patterns are expressed, i.e., reflecting individual variation in regional cerebellar volumes.

In the current study NNMF yielded highly reproducible cerebellar structural covariance patterns (across split-half datasets) for model orders (i.e., number of components/patterns specified) ranging from 2 to 100 (see Supplementary Figures 2 for summary maps of all tested model orders). After observing that the improved fit to the original data seen with higher model orders tended to level off between 15 and 30 components (indicating that the intrinsic dimensionality of the data might have been reached (see Supplementary Figure 3), we decided on a model order of 23 based on its good split-half reproducibility (see Fig. 1B, and Online Online Methods for details regarding model order selection). Importantly, when subject weights were summed across the 23 structural covariance maps, the resulting values corelated tightly (r = 0.9995) with estimates of total cerebellar grey matter volume (see also Supplementary Figure 4), demonstrating that our data-driven decomposition preserved inter-individual variation in cerebellar volume.

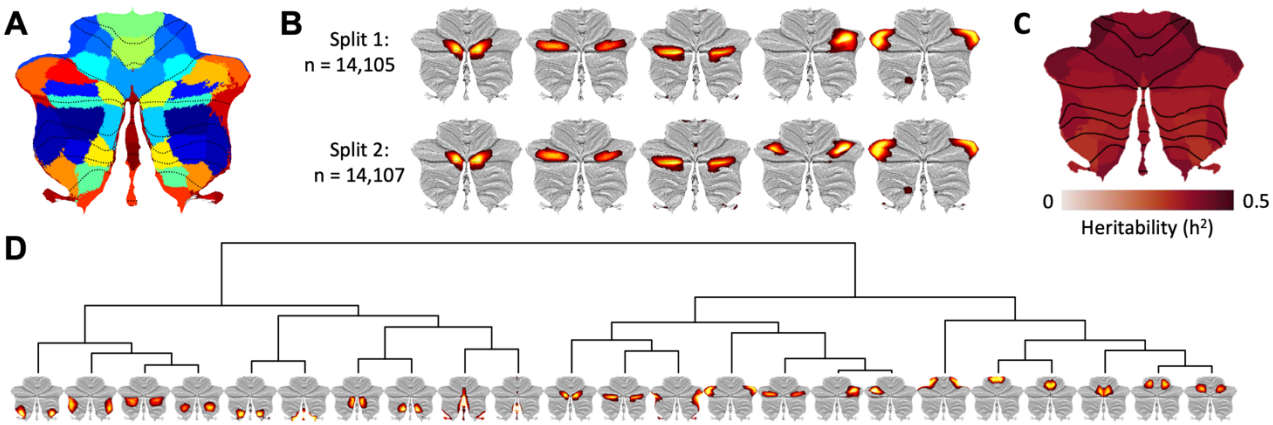


Figure 1: Data driven decomposition of cerebellar grey matter maps yields highly reproducible and moderately heritable structural covariance patterns (SCPs). A*: Binarized winner-takes-all map for the 23-component solution based on data-driven decomposition of cerebellar grey matter maps from 28,212 participants. Note that empirically derived boundaries between cerebellar regions only partially follow traditional lobular borders (marked with dotted black lines);* ***B:*** *Five distinct components overlapping cerebellar Crus I derived from the split-half reliability analyses. While one of these SCPs emerged as bilateral in Split 2, the remaining four SCPs were almost identical, despite being derived from two independent samples;* ***C:*** *Narrow sense (SNP-based) heritability of the 23 SCPs (see Supplementary Data 5 for numerical values).* ***D:*** *Hierarchical clustering of the 23 SCPs (derived from the full sample decomposition) based on their pairwise*

genetic correlations revealed a primary division between the anterior and posterior cerebellum, with additional separations between medial and lateral regions. The full genetic correlation matrix can be found in Supplementary Data 6.

Of note, our data-driven decomposition differed markedly from the standard cerebellar atlas based on gross anatomical features, shown as dotted lines in Fig 1A (see Supplementary Figure 3 for all 23 components and Supplementary Data 1 for quantification of overlap between NNMF-derived structural covariance patterns (SCPs) and standard cerebellar anatomical regions, i.e. lobules). For instance, five distinct SCPs overlapped Crus I of cerebellar lobule VII (shown in Fig 1B), an anatomical region which already started to split into separate components at a model order of three. We further observed only partial overlap with task-based functional parcellations of the cerebellar cortex (Supplementary Data 2). While some SCPs clearly overlapped cerebellar regions previously associated with hand movements, eye-movements/saccades or autobiographical recall, other data-driven SCPs overlapped multiple functionally defined cerebellar regions (Supplementary Data 2).

Cerebellar structural covariance patterns are heritable and reveals a distinct anterior-posterior pattern based on their bivariate genetic correlations.

After removal of one of each genetically related pair of individuals (n = 910), 27,302 participants remained for the genetic analyses. In addition to the 23 regional cerebellar structural covariance patterns of primary interest, we also included total cerebellar volume, estimated total intracranial volume and 9 cerebral brain phenotypes to serve as covariates and/or comparison phenotypes. Prior to all genetic analyses, morphological features were adjusted for effects of scanner site, sex, age, estimated total intracranial volume, 40 genetic population components, genetic analysis batch and a quantitative structural MRI quality index (the Euler number[31]) using general additive models, before finally being rank-order normalized (see Online Methods for details).

To validate our analysis approach, we computed genetic correlations (using LD-score regression, LDSC[32]) between univariate GWAS results on the comparison features (see Manhattan- and QQ-plots in Supplementary Figure 5) and previously published neuroimaging GWAS studies on these brain phenotypes. Results showed a mean r_g of .90 (range: .80-.99, see Supplementary Data 3). For the 23 cerebellar morphological features, we used the univariate GWAS summary statistics (see Supplementary Figures 6-7 for Manhattan- and QQ-plots) to compute genetic correlations between discovery (n = 27,302) and replication (n = 11,264) samples using LDSC[32]. These genetic correlations were high (mean rG: .92; range: .83-1), indicating reliable genetic signals (see Supplementary Data 4).

Genetic complex trait analysis (GCTA[33]) revealed SNP-based heritability estimates (h^2) ranging from .33 to .44 (Figure 1C and Supplementary Data 5). Analyses of total cerebellar volume (h^2 = .35), estimated total intracranial volume (h^2 = .41) and the 9 cerebral comparison phenotypes (h^2 range: .26 to .45) gave a similar range of heritability estimates (Supplementary Data 5).

181 Hierarchical clustering of cerebellar features based on their bivariate genetic correlation
182 matrix (GCTA-based rG ranging from 0.35 to 0.98, Supplementary Data 8) revealed a primary
183 anterior-posterior division running along the horizontal fissure separating Crus I and Crus II, with
184 secondary divisions grouping features into more lateral or medial, as well as more anterior and
185 posterior features within the major anterior and posterior regions (Fig 1D). Of note, the primary
186 division along the horizontal fissure was also evident from the (genetically naïve) two-component
187 NNMF decomposition, while medial-to-lateral divisions already began to emerge with a model order
188 of three (see Supplementary Figure 1).
189 In order to examine whether this phenotypic and genetic correlation structure was also
190 reflected in regional cerebellar gene expression patterns, we used the abagen toolbox[34] to extract
191 Allen Human Brain Atlas[35] gene expression profiles for 22 of the 23 morphological features and
192 computed their bivariate Pearson correlations (across 15,631 genes; Supplementary Data 7) and
193 hierarchical clustering solution (Supplementary Figure 8). The anterior-posterior boundary across
194 the horizontal fissure was also evident in the gene expression data, which in addition highlighted
195 distinct gene expression patterns for the posterior midline (grouped together with the horizontal
196 fissure), as well as for the most lateral regions of the cerebellar cortex.
197
198 **Multivariate GWAS reveals 351 genetic loci associated with cerebellar morphology.**
199 Figure 2A shows the main results for the multivariate GWAS across the 23 cerebellar structural
200 covariance patterns applying MOSTest[36]. We observed 35,098 genome-wide significant (GWS)
201 SNPs, which were mapped by FUMA[37] to a total of 51,803 candidate SNPs by adding reference
202 panel SNPs in high LD (r >.6) with GWS SNPs. The 51,803 candidate SNPs (see Supplementary
203 Data 8) were represented by 1,936 independent significant SNPs and 560 lead SNPs in 351 genomic
204 loci (see Supplementary Data 10). The LDSC intercept of the MOSTest summary statistics was
205 1.0352 (i.e., indicative of no or minimal genetic inflation), while the QQ-plot of results based on
206 permuted data (under the null hypothesis) confirmed the validity of the MOSTest analytical approach
207 (Supplementary Figure 9).

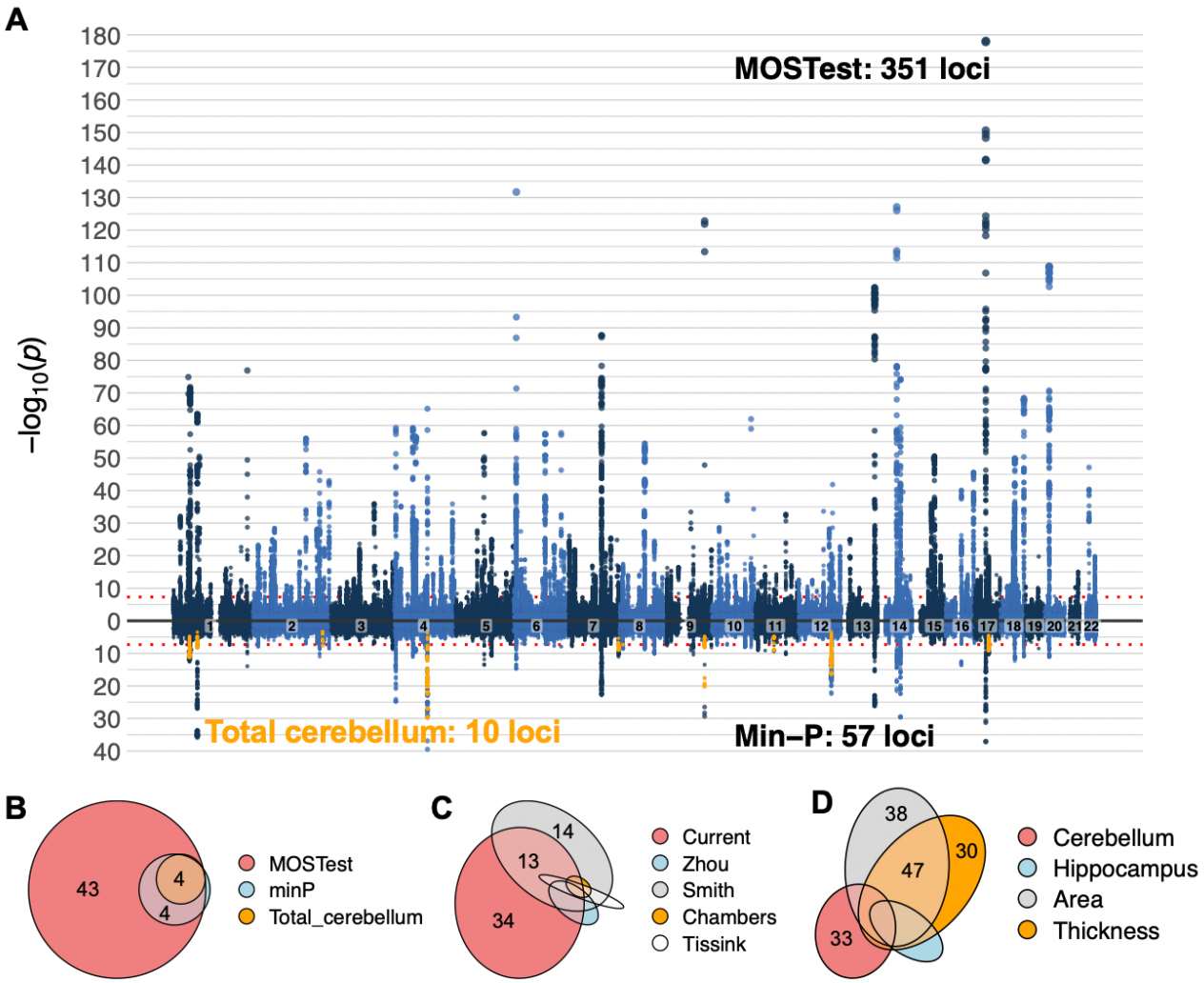


Figure 2: Multivariate GWAS analysis of the 23 cerebellar structural covariance patterns revealed 351 independent genome-wide significant (GWS) loci. A: *The upper half of the Miami plot shows the main results from the multivariate analysis. The lower half displays results from a series of 23 univariate analyses (corrected for multiple comparisons using the standard min-P approach), as well as results from a univariate analysis of total cerebellar grey matter (marked in orange).* ***B-D:*** *Euler diagrams showing the relative numbers of - and overlaps between - candidate SNPs (in thousands) mapped by the three analysis approaches employed in the current study* ***(B)****, the current and four recent studies reporting genetic associations with cerebellar morphology* ***(C)****, as well as results from multivariate GWASs on hippocampal and cerebrocortical morphology* ***(D)****. For full results on overlap between all cerebellar candidate SNPs and other brain phenotypes, see Supplementary Data 8.*

Annotation of all candidate SNPs using ANNOVAR[38] as implemented in FUMA[35] revealed that the majority of candidate SNPs were intronic (57.8%) or intergenic (38.3%). While only 0.7% were exonic, about 81% of the candidate SNPs were assigned minimal chromatin states between 1 and 7 (i.e., open chromatin states), implying effects on active transcription[37] (see Supplementary Data 8-9).

We evaluated the robustness of these multivariate results using a multivariate replication procedure established in Loughnan et al[39], which computes a composite score from the mass-univariate *z*-statistics (i.e., applying multivariate weights from the discovery sample to the replication sample input data) and then tests for associations between this composite score and genotypes in the replication sample (for mathematical formulation see Loughnan et al[39]). Results showed that 97% of the 339 loci lead SNPs present in both samples (i.e., 94% of the 351 reported loci lead SNPs from the discovery sample) replicated at a nominal significance threshold of $p < .05$ (Supplementary Figure 10 and Supplementary Data 10), and that 74% remained significant after Bonferroni correction (Supplementary Data 10). Moreover, 99% of loci lead SNPs showed the same effect direction across discovery and replication samples (Supplementary Figure 9 and Supplementary Data 10).

In addition, we assessed the robustness of the multivariate patterns by computing bivariate correlations between feature z-score vectors assigned to the discovery sample lead SNPs in an independent multivariate GWAS (MOSTest) performed on the replication sample. These correlations (restricted to the 339 loci lead SNPs present in both samples) were relatively high (mean r: .70, see Supplementary Data 10 and Supplementary Figure 11). Figure 3 and Supplementary Figure 11 also give some examples of discovery and replication sample multivariate patterns projected back onto the cerebellar cortex.

To compare our main multivariate MOSTest results to univariate approaches, the lower part of the Miami plot in Figure 2A, and Figure 2B, displays results from a set of univariate GWASs on the cerebellar morphological features (which yielded 8370 candidate SNPs and 57 genomic loci, corrected for multiple comparisons using the min-P approach[40,41], see also Supplementary Figures 6-7 for univariate Manhattan and QQ-plots), as well as the 4044 candidate SNPs and 10 significant loci resulting from the univariate GWAS on total cerebellar grey matter volume (marked in orange). 52 of the 57 loci identified in the univariate analyses of regional cerebellar features overlapped 55 of the 351 loci identified using the multivariate method, while 9 out of 10 loci identified in the univariate analysis of total cerebellar volume were also identified in the multivariate analysis (the slightly mismatching numbers are due to loci of unequal lengths, causing some larger loci to overlap with several smaller loci). Thus, our multivariate analysis of regional cerebellar features increased the locus yield ~35-fold relative to analyzing total cerebellar volume and ~6-fold relative to performing a set of univariate analyses on the same regional features.

We next compared the current findings with previously reported genetic loci for cerebellar morphology by extracting summary statistics from two recent GWAS studies using the UKBB sample (n = 19k[42] and 33k[43]) that included regional cerebellar volumes among the full set of analyzed brain imaging-derived phenotypes (101[42] and 3,144[43], respectively), as well as two recent GWASs on total cerebellar volume (n = 33k[13] and 27k[14], respectively). Candidate SNPs and independent GWS loci were identified in FUMA using the same settings as for our primary analyses and employing a liberal *p*-value threshold of 5e-8 (i.e., not correcting for the total number of brain imaging features analyzed).

Results are displayed in Figure 2c and Supplementary Data 8 and 10. In brief, we found that 19,527 of the 51,803 candidate SNP (i.e., 36%) and 123 of the 351 identified genomic loci (i.e., 35%) overlapped with candidate SNPs and loci extracted from these three previous studies. Thus, 228 of the 351 (i.e., 65%) genetic loci reported here are novel to the literature on of cerebellar morphology genetics (see Supplementary Data 10.

Overlap of cerebellar candidate SNPs and genetic loci with results from recent multivariate analyses of hippocampal and cerebrocortical morphology are displayed in Figure 2D and Supplementary Data 8 and 10 (final columns). Of note, we found that 32 and 29 percent of the candidate SNPs discovered here for the cerebellum overlapped with candidate SNPs for vertex-wise cerebrocortical surface area and thickness[44], respectively, while 11.4 percent overlapped with candidate SNPs found for hippopcampal subregions[45]. 95 of the 351 genetic loci overlapped loci linked to the other multivariate brain phenotypes (Supplementary Data 10). Thus, 64% of the candidate SNPs and 73% of genetic loci appeared to be specifically associated with cerebellar morphology.

Significant genetic variants show heterogeneous effects across the cerebellar cortex, influencing both regional and total volumes.

A major advantage of our multivariate analysis approach is its sensitivity to both highly localized and more generally distributed effects of SNPs on cerebellar morphology. This is illustrated in Figure 3, which displays the 351 loci lead SNPs as a function of both the most extreme individual Z-score across all cerebellar features (e.g, analogous to the strongest “local” effect) and of the mean Z-score across these features (i.e. analogous to the main effect on overall cerebellar volume).

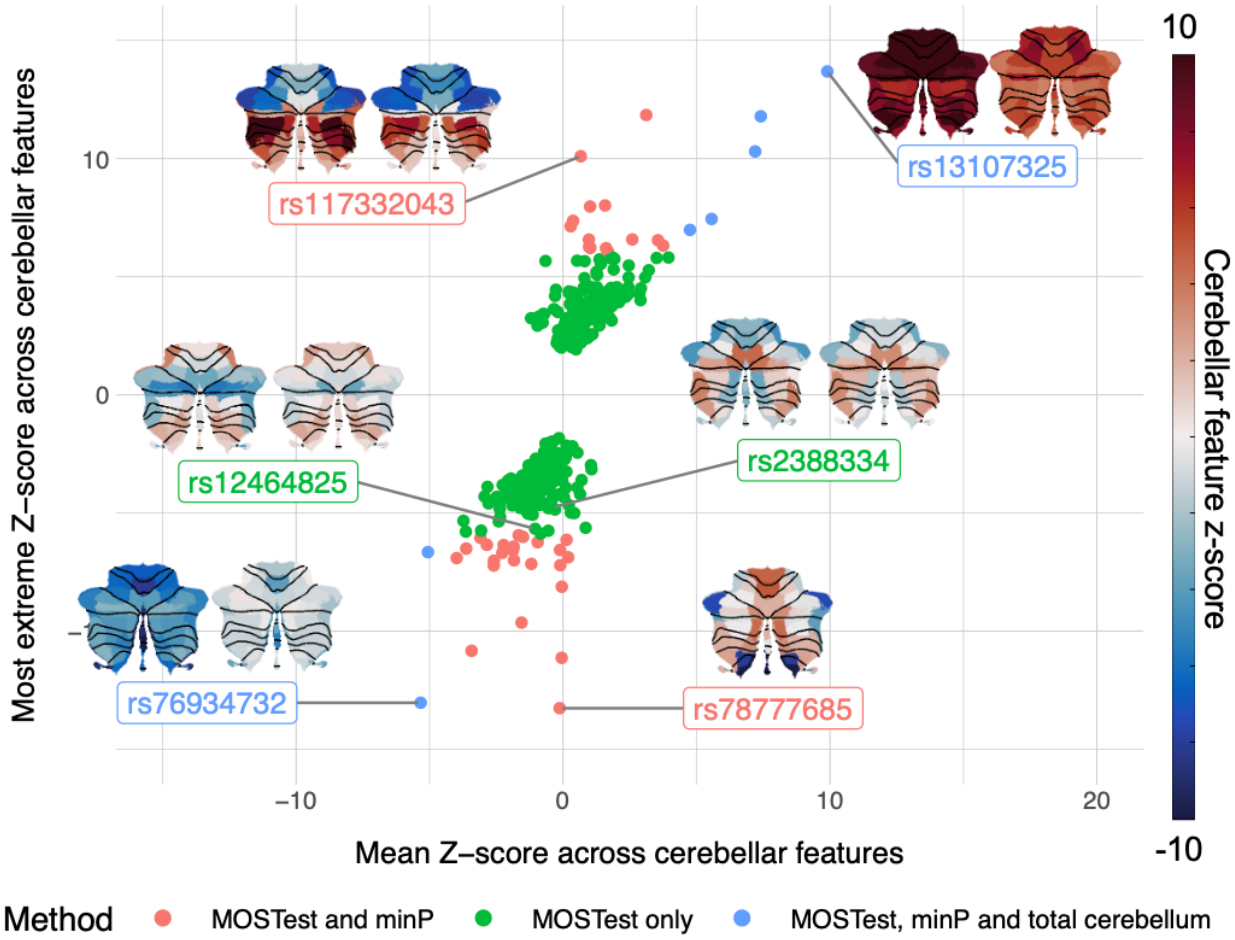
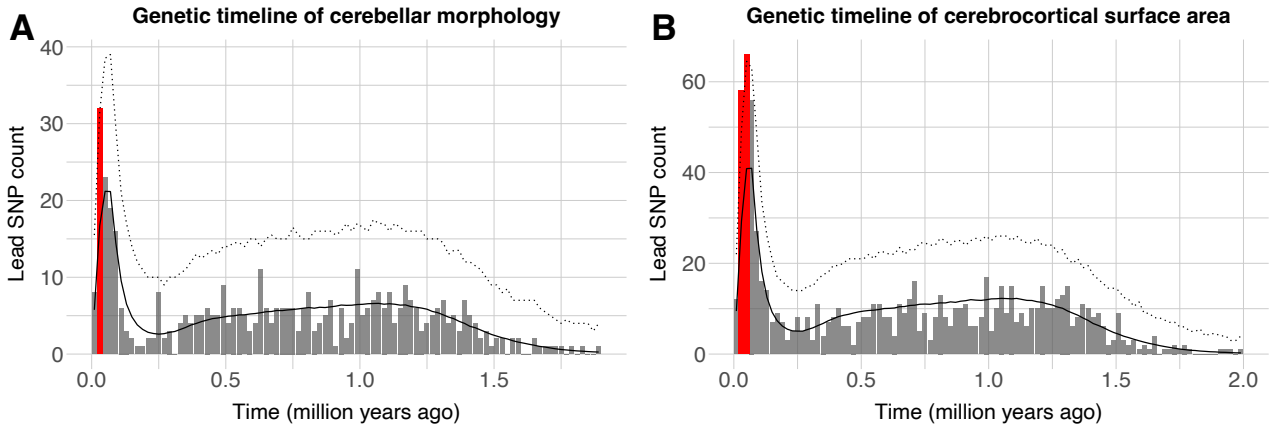


Figure 3: Loci lead SNPs show spatially heterogeneous and replicable effects across the cerebellar cortex. *The 351 loci lead SNPs identified by MOSTest are plotted as a function of main overall effect across all cerebellar features (x-axis: mean Z-score) and most extreme effect for any single cerebellar feature (y-axis: most extreme Z-score across features), and colour coded by SNP discovery method. The cerebellar flat-maps show discovery (left) and replication (right) sample regional distributions of Z-scores (color-scale range from -10 to 10) for a few selected lead SNPs (see Supplementary Data 10 for individual feature Z-scores for all 351 discovery sample loci lead SNPs). SNP rs7877685 was only present in the discovery sample.*

As can be seen, some loci lead SNPs (e.g. rs13107325; rs76934732) show pronounced positive or negative mean z-scores, indicating a relatively consistent direction of effect across cerebellar features. See also inset figures displaying feature Z-scores projected back onto the cerebellar cortex. Many of these SNPs also emerged in the univariate analysis of total cerebellar volume and were recently reported in GWASs on total cerebellar volume[13,14].

Many other loci lead SNPs, however, show strong “local” signals with opposite effect directions across features, yielding very weak global signals (e.g. rs117332043; rs78777685). Thus, while several of the most significant SNPs in this category have previously been reported in GWASs including local cerebellar morphological features[42,43], they did not emerge from analyses of total cerebellar volume, neither in the current nor in previous studies.

306 Finally, our multivariate MOSTest approach is also sensitive to SNPs displaying weaker
307 effects distributed across several cerebellar features, often with opposite effect directions (e.g.
308 rs12464825; rs2388334). For this category of SNPs, neither of the univariate methods have sufficient
309 power at the current sample size. In contrast, these two example SNPs robustly emerged from the
310 multivariate analysis (discovery sample p-values < 1e-56; replication sample p-values <1e-15).
311
312 **Genetic variants associated with cerebellar morphology are enriched for evolutionary recent**
313 **mutations in the human genome.**
314 We next mapped the evolutionary age of cerebellar lead SNPs (thresholded for linkage
315 disequilibrium at r2 <0.1) by merging them with a recently published dataset on dated mutations in
316 the human genome[46]. Following the analysis procedure established by Libedinsky et al.[47], we plotted
317 the histogram of dated lead SNPs over the last 2 million years in bins of 20.000 years (Figure 4A
318 and Supplementary Data 11), and tested for positive or negative enrichment by comparing them to
319 empirical null distributions derived from 10,000 randomly drawn and equally sized sets of all SNPs
320 in the full human genome dating dataset (after matching them to cerebellum-associated lead SNPs
321 in terms of minor allele frequency; see Online Methods for details).
322



323
324 ***Figure 4: Lead SNPs associated with cerebellar morphology are enriched for evolutionary recent***
325 ***mutations in the human genome. A:*** *Histogram of estimated SNP age (ranging from 0 to 2 million years, in*
326 *bins on 20,000 years) for 548 independent lead SNPS associated with cerebellar morphology. The solid black*
327 *and dotted lines denote the mean and upper 95th confidence interval (Bonferroni corrected across 100 time-*
328 *bins) derived from a null model constructed from 10,000 of equally sized sets of SNPs randomly drawn from*
329 *the Human Genome Dating Atlas of Variant Age (after matching these to cerebellar lead SNPs in terms of*
330 *minor allele frequencies). Red bars denote time-bins of significant positive enrichment.* ***B:*** *Histogram showing*
331 *comparative evolutionary enrichment effects for lead SNPs identified in a multivariate GWASs of*
332 *cerebrocortical area (for cerebrocortical thickness, see Supplementary Figure 12), See Supplementary Data*
333 *11-12 for full numerical results.*
334
335 Results revealed positive enrichment for cerebellar lead SNPs in the time bin ranging from 20-40,000
336 years ago, i.e., overlapping the Upper Paleolithic, a period characterized by rapid cultural evolution

337 and the first evidence of several uniquely human behaviors (often referred to as behavioral
338 modernity), such as the recording of information onto objects[48].
339 For comparison, Figure 4B also shows results from an analysis of lead SNPs identified in a
340 previous multivariate GWAS study of regional cerebrocortical surface area [44], while Supplementary
341 Figure 12 shows the corresponding histogram for regional cerebrocortical thickness. As can be seen,
342 both these cerebrocortical phenotypes also showed significant enrichment in overlapping time bins
343 (i.e., 20-60,000 years ago). For full numerical results in Supplementary Data 12. Given the relatively
344 low number of independent lead SNPs (range: 548-862 SNPs across phenotypes) , we also ran
345 validation analyses using all independent significant SNPs (LD-thesholded at r2 < 0.6, range: 1574-
346 2883 SNPs), which yielded very similar results (see Supplementary Figure 13 and Supplementary
347 Table 13).
348
349 **Genes associated with cerebellar morphology show selective expression in cerebellar and**
350 **prenatal brain tissue, as well as enrichment for genes linked to human accelerated regions.**
351 To functionally characterize the multivariate GWAS signal, we mapped the full set of GWAS p-values
352 to 19,329 protein coding genes using MAGMA[49] and used the resulting gene-level p-values to test
353 for 1) GWS genes, 2) gene expression in brain tissue; and 3) enrichment for genes linked to human
354 accelerated regions (HARs), i.e. sections of DNA that have remained relatively conserved
355 throughout mammalian evolution, before being subject to a burst of changes in humans since
356 divergence of humans from chimpanzees[50,51]. These analyses yielded a total of 534 GWS genes
357 (i.e., 2.78% of all protein coding genes, see Figure 5A and Supplementary Data 14). Using the full
358 set of 19,329 gene-level p-values in MAGMA gene property analyses revealed significant and
359 specific gene expression in cerebellar and prenatal brain tissue (Figure 5B-C and Supplementary
360 Data 15-16), with the selective cerebellar expression seen in two independent datasets (Allen
361 Human Brain Atlas[35] and The Genotype-Tissue Expression (GTEx) Project).
362 The MAGMA gene set analysis of HAR-linked genes revealed significant enrichment (p =
363 7.09e-08) for genes associated with cerebellar morphology (Figure 5D, Supplementary Data 17). Of
364 note, running this same HAR gene set analysis on summary statistics from recent multivariate GWAS
365 studies on cerebrocortical[44] or hippocampal[45] morphological features yielded similar (for cortical
366 features) or significantly weaker (hippocampal features) enrichment effects (Figure 5d, see Online
367 Methods for processing pipeline).
368

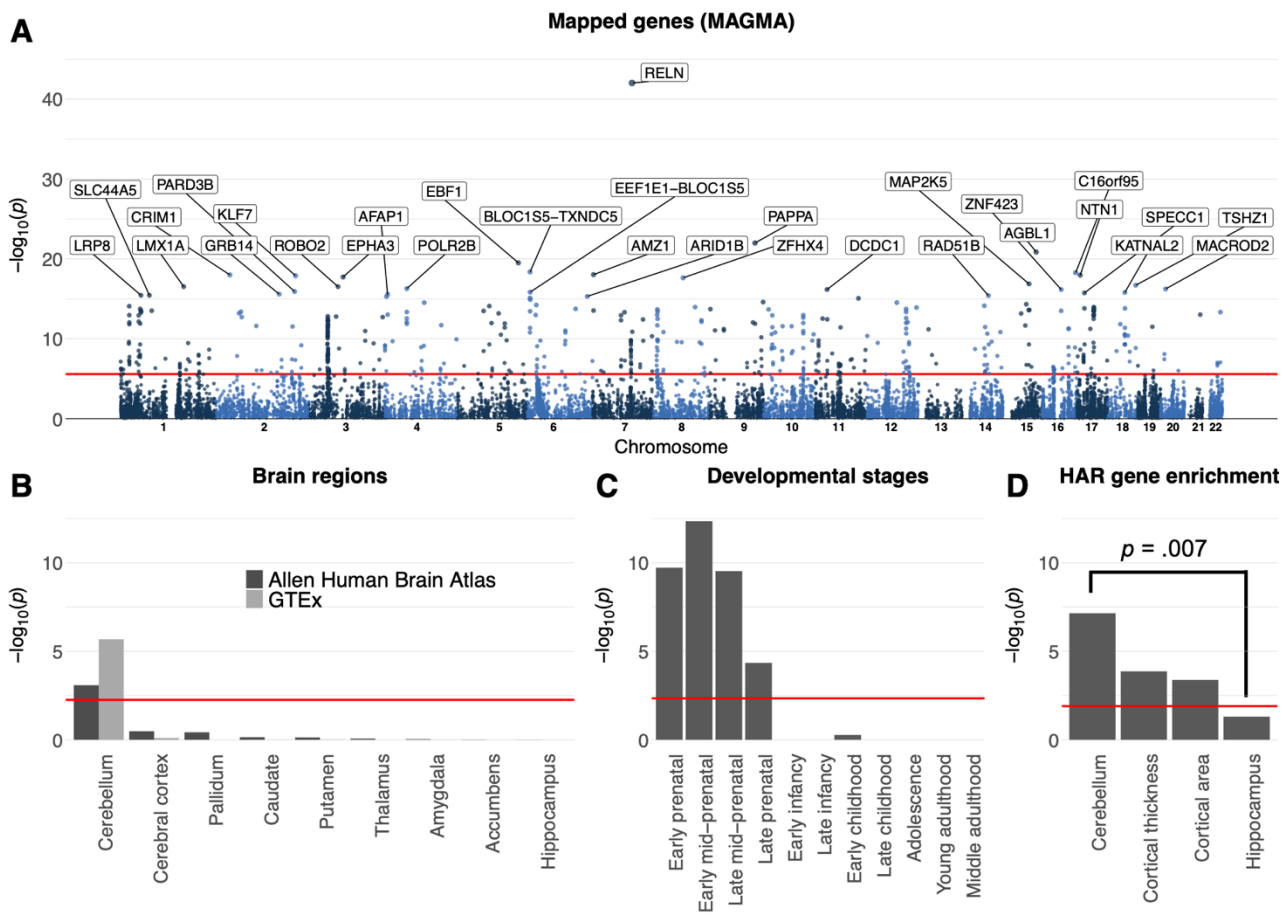


Figure 5: Gene mapping reveals selective enrichment for cerebellar and prenatal brain tissue, as well as for genes linked to human accelerated regions. A: *MAGMA mapped the full range of SNPs from the multivariate GWAS to 19.329 protein-coding genes, 534 of which were genome-wide significant (Bonferroni corrected threshold marked with red line), with the 30 most significant genes labeled in the gene-based Manhattan plot.* ***B-C:*** *MAGMA gene-property analyses revealed selective brain expression in the cerebellum and during prenatal developmental stages.* ***D:*** *MAGMA gene-set analysis revealed significant enrichment for sets of genes previously linked to human accelerated regions (HARs). Figure 4d also shows results from comparative analyses using summary statistics from other multivariate GWASs of MRI-based brain, as well as significant results from statistical tests comparing beta-weights for HAR-linked genes across multivariate brain features. Horizontal red lines mark the Bonferroni-corrected significance threshold for each subplot. See Supplementary Data 14-17 for full results.*

Genes linked to human cerebellar morphology show enrichment for gene sets linked to selective cerebellar gene expression, altered cerebellar morphology in mouse models, human clinical/anthropometric traits, as well as specific biological processes and pathways. In line with the continuous brain tissue gene expression results described above, we also observed significant and relatively selective enrichment for smaller curated sets of genes previously found to be highly and selectively expressed in mouse (Figure 6A, Supplementary Data 18) and human (Figure 6C, Supplementary Data 20) cerebellar brain tissue, as well as for sets of genes previously shown to affect cerebellar morphology in mouse gene perturbation experiments (Figure 6B,

Supplementary Data 19) and various clinical conditions and anthropometric traits in humans (Figure 6D, Supplementary Data 21). The most significant gene ontology and curated gene sets from the MSigDB[52,53] database were related to brain development (e.g., neurogenesis, axon guidance and neuron differentiation, Figure 6E, Supplementary Data 22), and highlighted the reelin signaling pathway (Figure 6F, Supplementary Data 23).

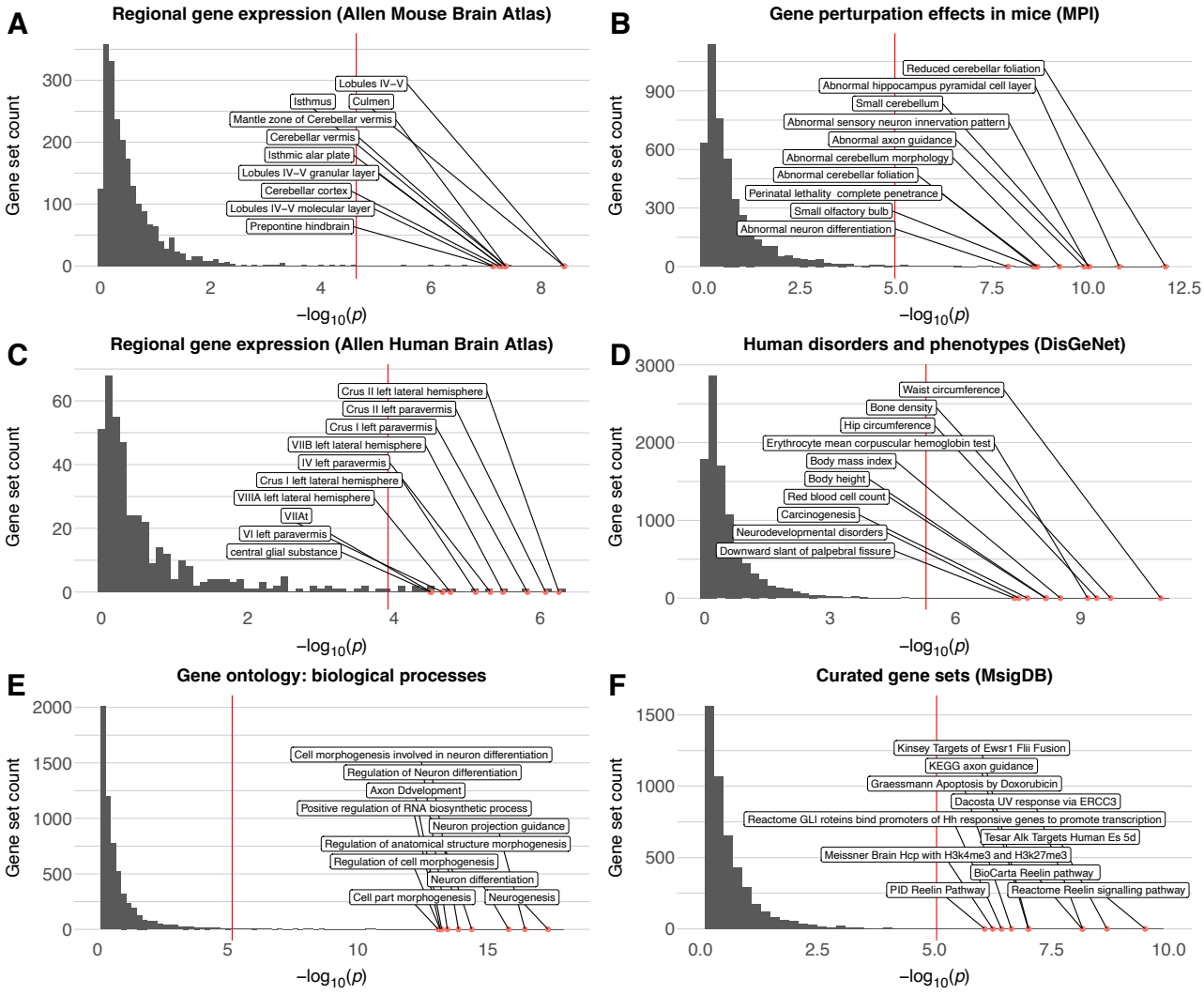


Figure 6: Gene mapping reveals selective enrichment across brain tissues and curated gene sets. *MAGMA gene-set analyses revealed significant enrichment for sets of genes previously linked to preferential expression in human* ***(A)*** *and mouse* ***(C)*** *cerebellar tissue, as well as effects on cerebellar morphology (and other brain phenotypes) in mouse gene perturbation experiments* ***(B)*** *and human clinical disorders and phenotypes* ***(D).*** *Significant gene ontology terms were related to neural development* ***(E)****, while curated gene sets highlighted the Reelin signaling pathway* ***(F)****. Across all subplots the x-axis shows the -log10 p-value, the y-axis marks the number of gene sets, and the top 10 most significant gene-sets are labelled. Red lines mark the Bonferroni-corrected significance threshold for each subplot. See Supplementary Data 18-23 for full results.*

As can be seen in Figure 5A (and Supplementary Data 14), *RELN* (encoding the protein Reelin) was also the most significant gene mapped by MAGMA. See also Supplementary Figure 14 for a regional locus plot showing the 12 lead SNPs mapped to *RELN* and their associated z-score maps.

Gene mapping reveals sets of plausible causal genes.

In addition to the gene-based mapping strategy using all SNPs described above (MAGMA), we also mapped candidate SNPs to genes using two complementary gene mapping strategies: 1) positional mapping of deleterious SNPs (defined as having a CADD-score[54] > 12.37); and 2) mapping of SNPs previously shown to affect gene expression in cerebellar tissues (i.e., eQTL mapping). Across all three strategies, we mapped a total of 674 unique genes; 531 using MAGMA, 310 using positional and 197 using eQTL mapping (see Figure 7A and Supplementary Data 24). 298 genes were identified by at least two strategies, while 65 genes were mapped by all three strategies. Out of these 674 genes, 61 have previously been associated with cerebellar pathology in humans and/or altered cerebellar morphology in mouse gene perturbation experiments, while 121 have been linked to human accelerated regions (Supplementary Data 24). As can be seen in Figure 7B and Supplementary Data 24, the 674 genes mapped to cerebellar morphology showed some overlap with genes mapped to hippocampal[45] and cerebrocortical[44] morphology using the same mapping strategies, but 264 genes (39%) appeared relatively specific to the cerebellum.

Results from gene set analyses on the 674 mapped genes (Figure 7, Supplementary Data 25-29) largely mirrored results from the MAGMA analyses described above, but in addition revealed that this set of 674 mapped genes was also enriched for gene sets associated with several complex clinical phenotypes and anthropometric traits in humans, including cognitive ability, neuroticism and schizophrenia (Figure 7F, Supplementary Data 29).

Of note, while the full set of mapped genes showed significant enrichment for sets of genes known to alter cerebellar morphology in mouse mutation or knock-down experiments (Supplementary Data 27), we also note that 613 of the 674 mapped genes have not to our knowledge previously been linked to cerebellar development, anatomy or pathology in mice or humans (Supplementary Data 24), and thus constitute potential targets for future gene perturbation experiments in animal models.

Restricting the above analyses to the 298 genes mapped across at least two strategies did not markedly affect the results (Supplementary Data 25-29, final columns).

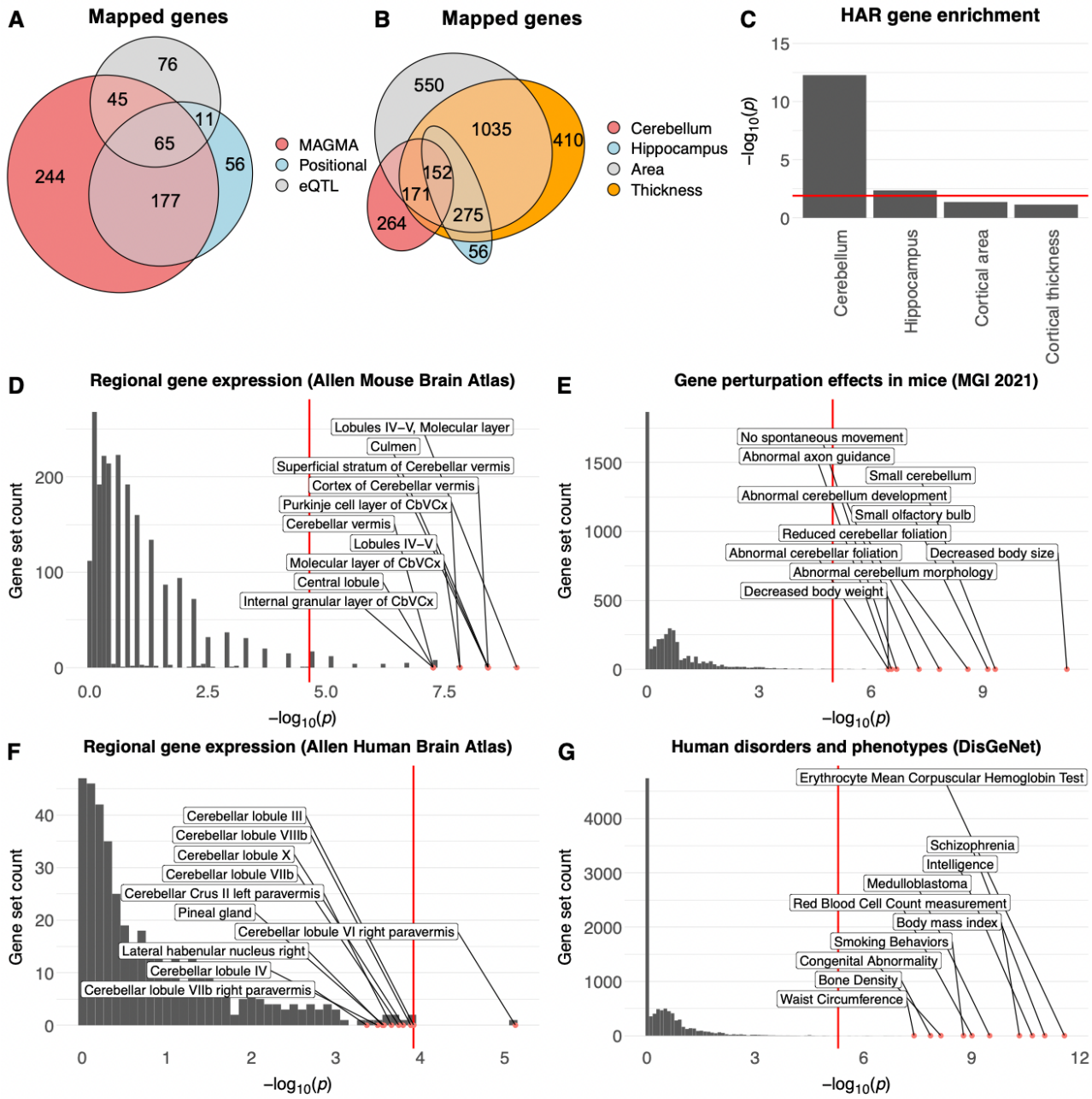


Figure 7: Mapping and functional characterization of plausible causal genes. *The 674 plausible causal genes mapped using three complementary strategies* ***(A)*** *show partial overlap with genes mapped to hippocampal*[45] *and cerebrocortical*[44] *morphology* ***(B)****; significant enrichment for sets of genes linked to human accelerated regions (HAR)* ***(C)****; selective expression in the cerebellum in mice* ***(D)*** *and humans* ***(E)****, as well as effects on cerebellar morphology in mouse gene perturbation studies* ***(F)*** *and human disorders and anthropometric phenotypes* ***(G)****. Across all subplots the x-axis shows the -log10 p-value, the y-axis marks the number of gene sets, and the top 10 most significant gene-sets are labelled. Red lines mark the Bonferroni-corrected significance threshold for each subplot. See Supplementary Data 24-29 for full results.*

Cerebellar morphology shows significant genetic overlap with psychiatric disorders.

We finally tested for overlap between the multivariate genetic profile for cerebellar morphology and genetic profiles for five major developmental/psychiatric disorders (attention deficit hyperactivity disorder: ADHD; autism spectrum disorder: ASD; bipolar disorder: BIP; major depressive disorder:

MDD; and schizophrenia: SCZ) using conditional/conjunctional FDR analysis[55]. As can be seen in the conditional QQ-plots in Figure 8, these analyses revealed clear patterns of enriched association with the clinical phenotypes when selecting subsets of SNPs with increasingly stronger association with cerebellar morphology (Fig. 8A and figure insets in (C-F).

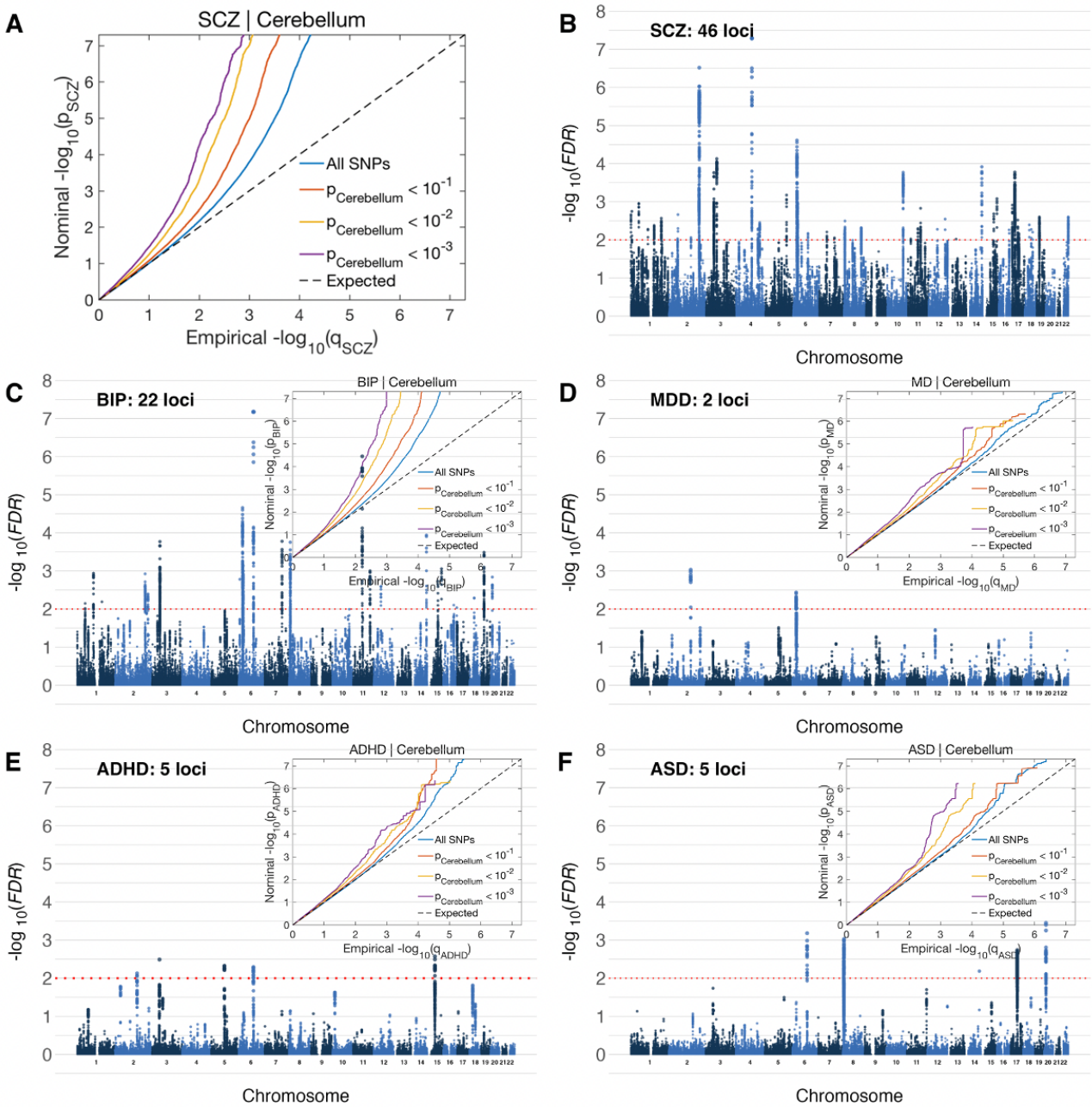
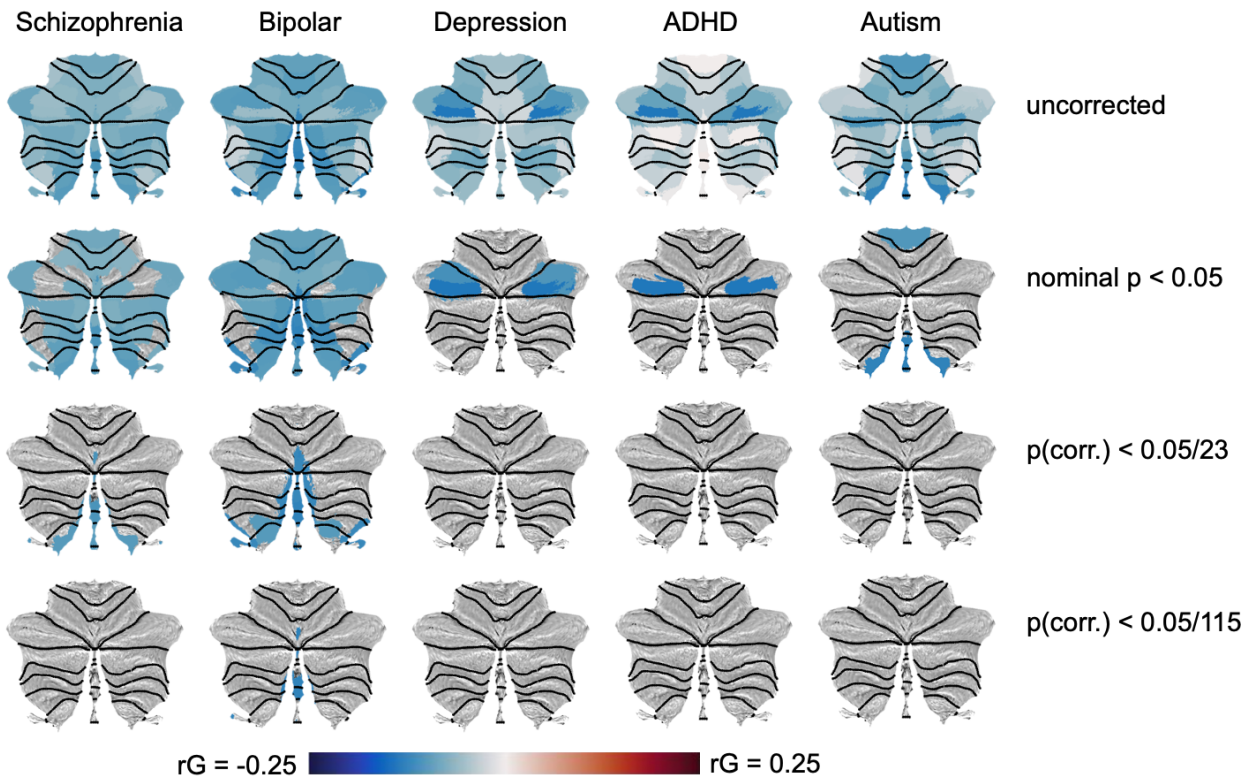


Figure 8: Genetic variants influencing cerebellar morphology overlap with variants associated with five major mental disorders. *Conditional QQ plots (A and figure insets in C-F) show an incremental incidence of association with five mental disorders (leftward deflection) as a function of the significance of association with cerebellar morphology. Manhattan plots show SNPs with significant association with both traits, thresholded at a conjunctional FDR threshold of p > .01 (red dotted line). SCZ: Schizophrenia; BIP: Bipolar Disorder; MDD: Major Depressive Disorder; ADHD: Attention Deficit Hyperactivity Disorder; ASD: Autism Spectrum Disorders.*

467 See also Supplementary Figure 15 for QQ-plots depicting the reverse association, i.e., enriched
468 association with cerebellar morphology when conditioning on the association with
469 developmental/psychiatric disorders. Specific genetic variants jointly influencing the two phenotypes
470 were identified using conjunctional FDR analyses at a conservative statistical threshold of p<.01.
471 Results revealed shared genetic loci with all disorders; namely 48 with SCZ, 22 with BIP, 2 with
472 MDD, 5 with ADHD and 5 with ASD (Figure 6; Supplementary Data 30-34). We mapped lead and
473 candidate SNPs for each of these loci to genes using positional and eQTL mapping and checked for
474 gene overlap across disorders (Supplementary Data 35). Of note, the *LRP8* gene (a HAR-linked
475 gene[50,51] encoding a Reelin receptor) emerged from the conjunctional FDR analyses of both BIP and
476 SCZ, thus again highlighting the Reelin signaling pathway.
477 To complement these multivariate analyses, we also conducted a set of univariate analyses,
478 using LD-score regression[32] to calculate the genetic correlations between each individual cerebellar
479 feature and the five developmental/psychiatric disorders.
480



481
482 ***Figure 9: Univariate genetic correlations between cerebellar morphological features and five***
483 ***major mental disorders are negative.*** *The top row display unthresholded genetic correlations,*
484 *while these are filtered at increasingly strict statistical thresholds in the following rows, i.e. nominal*
485 *p-value < .05 (second row); Bonferroni correction for the 23 cerebellar features tested (third row);*
486 *and Bonferroni correction for both 23 features and 5 clinical conditions (bottom row). Black dotted*
487 *lines denote lobular boundaries.*
488

As can be seen in the first row of Figure 9 (and Supplementary Data 36), genetic correlations with cerebellar morphological features were predominantly negative across diagnoses, indicating that genetic variants associated with a clinical diagnosis tended to also be associated with reduced cerebellar volumes (with 42 out of 115 tested associations showing nominally significant negative correlations; Fig 9, second row). However, all univariate genetic correlations were relatively weak, and only a few negative genetic correlations with BIP and SCZ survived Bonferroni correction for the 23 features tested. When also correcting for the five clinical conditions, only the negative correlation between BIP and cerebellar feature 23 (primarily overlapping vermal lobules VIIIa and VIIIb) remained significant.

Supplementary Data 37 show genetic correlations between the five disorders and the ten comparison brain phenotypes, as well as total cerebellar volume. In general agreement with the observed pattern for regional cerebellar features, total cerebellar grey matter volume showed nominally significant negative genetic correlations with BIP (rg: -.10; p = 0.0052) and SCZ (rg: -.09; p = 0.0133). Pallidal volume also showed nominally significant negative genetic correlations with these two disorders (BIP: rg: -.10; p = 0.0087; SCZ: rg: -.08; p = 0.0172, while ADHD displayed negative genetic correlations with estimated total intracranial volume (rg: -.15; p = 0.0003) and total cortical surface area (rg:-.14 ; p = 0.0065), as well as a positive genetic correlation with hippocampal volume (rg: .13; p = 0.0164). Finally, MDD showed nominally significant negative genetic correlations with volumes of the hippocampus (rg -.08: p = 0.025) and thalamus (rg: -.09; p = 0.0258). Only the negative genetic correlation between ADHD and estimated total intracranial volume survived Bonferroni correction across the 55 tests performed.

Discussion

The current study identified novel features of the genetic architecture of cerebellar morphology, supported the notion of recent changes over human evolution, implicated specific neurobiological pathways, and demonstrated genetic overlap with major mental disorders.

With respect to the main features of cerebellar morphology, it is worth noting that results from our data-driven decomposition of cerebellar grey matter maps (which was not informed by genetic data), the genetic correlation analyses of the chosen 23-feature solution and gene expression data from the Allen Human Brain Atlas all converge on a similar general pattern. The first boundary to emerge in the data-driven decompositions ran along the horizontal fissure separating Crus I and II of lobule VII, reflecting its centrality in characterizing phenotypic variability in cerebellar morphology at the population level. Of note, this boundary also emerged from clustering of the 23 cerebellar morphological features based on their bivariate genetic correlations or gene expression profiles (based on the Allen Human Brain Atlas). This latter finding essentially mirrors results from a recent report using a different analysis strategy on the same gene expression data[56]. Interestingly, the horizontal fissure has been suggested to mark the border between two separate cerebellar

528 representations of the cerebral cortex, with a possible third representation in lobules IX-X[57]. The
529 current cerebellar results thus complement previous work on the hierarchical genetic organization of
530 the cerebral cortex, which has identified the Rolandic fissure (separating the frontal and parietal
531 lobes) as a main boundary with respect to genetic effects of effects of surface area[58], as well as a
532 superior-inferior gradient for genetic influences on cortical thickness[59]

533 Our multivariate GWAS using MOSTest identified 351 independent GWS loci associated with
534 cerebellar morphology, increasing the yield ~35-fold relative to analyzing total cerebellar volume and
535 ~6-fold relative to performing a set of univariate analyses on the same regional features in our current
536 sample. 329 (94%) loci from the multivariate analyses were replicated in an independent sample,
537 indicating robust results. After applying a liberal threshold to summary statistics from previous well-
538 powered studies, we find that 228 (65%) of our reported loci are novel. Importantly, among the genes
539 mapped to these novel loci we find several that are known to play important roles in cerebellar
540 development in mice (e.g., *RORA*[60], *FGF8*[61] and *BAHRL1*[62], see Supplementary Data 23). While
541 candidate SNPs associated with cerebellar morphology partially overlap with SNPs previously
542 mapped to other multivariate brain phenotypes, we note that a substantial number of SNPs appear
543 to be relatively selectively linked to cerebellar morphology, a finding that is in in line with the distinct
544 gene expression profile found for the cerebellum[35].

545 SNP- and gene-level results from the current study also bolster – and refine – the notion of
546 relatively recent changes in cerebellar morphology over human evolution[2,4,15,16]. Specifically, we
547 found that lead SNPs associated with cerebellar morphology were enriched for SNPs with an
548 estimated age of 20-40 thousand years. This overlaps the Upper Paleolithic (10-50k years ago), a
549 period characterized by rapid cultural evolution, and coinciding with the first evidence of several
550 uniquely human behaviors (often referred to as behavioral modernity), such as the recording of
551 information onto objects[48]. Converging evidence for changes in brain anatomy around this
552 evolutionary period comes from the fossil scull record, where only fossils dated less than 35.000
553 years old fall within the range of shape-variation seen in modern humans[63,64]. Of note, one key
554 feature of modern skulls is their "globular" shape, in part characterized by an enlarged posterior
555 fossa (the cranial compartment housing the cerebellum)[63,64]. A similar pattern of enrichment for
556 recent evolutionary time bins (20-60 thousand years ago) was also found when analyzing lead SNPs
557 derived from previous studies on multivariate cerebro-cortical[44] morphology.

558 Gene level analyses further showed that genes associated with inter-individual variation in
559 cerebellar morphology are enriched for genes linked to human accelerated regions (HARs)[50] of the
560 genome. HARs denote previously conserved regions of the genome that were subject to a burst of
561 changes in humans after the divergence of humans from chimpanzees about 6-8 million years ago[50].
562 Of note, a recent GWAS on total cerebellar volume found no enrichment for HAR-linked genes[14],
563 suggesting that SNPs associated with regional cerebellar variation may be driving this effect. When
564 comparing this finding with results from other multivariate brain phenotypes, we observed that HAR
565 gene enrichment was nominally stronger for cerebellar morphology than for vertex-wise

566 cerebrocortical thickness and area, and significantly stronger than for hippocampal regional volumes
567 (Figure 5D).

568 Together, these SNP- and gene-level results suggest that genetic variants influencing
569 cerebellar morphology in modern humans have been subject to selection over relatively recent
570 human evolution, and that cerebellar changes – in concert with other brain regions – may thus have
571 played a central part in the emergence of uniquely human cognitive abilities.

572 Results from the MAGMA gene property and gene set analyses bolster our confidence in the
573 genetic signal by showing selective gene expression in human cerebellar brain tissue across two
574 independent datasets (Allen Human Brain Atlas and GTEx.8) and significant enrichment for sets of
575 genes which have previously been shown to affect cerebellar morphology in mouse gene
576 perturbation experiments. The MAGMA results further show significant enrichment for sets of genes
577 known to play key roles in neurodevelopment (e.g., neurogenesis & axon guidance) and preferential
578 expression in prenatal brain tissue, thus supporting a primarily developmental origin of genetically
579 determined effects on adult cerebellar morphology. Of note, our MAGMA enrichment analysis of
580 curated gene sets strongly implicated the Reelin signaling pathway. Indeed, both the gene coding
581 for the Reelin protein (*RELN*) and genes coding for its two receptors (*LRP8* & *VLDLR*) were identified
582 across at least two gene mapping strategies, with *RELN* emerging as the single most significant
583 gene by MAGMA. The Reelin pathway is known to play important roles in neurodevelopment (e.g.
584 neuronal migration), and mutations in the *RELN* and VLDLR[65] (and to a lesser extent LRP8; also
585 known as ApoER2[66]) have been associated with cerebellar malformations and/or dysfunction. LRP8
586 is also among the genes linked to human accelerated regions (HARs) of the genome.

587 The sets of genes mapped by the three complementary mapping strategies provide a
588 database for future studies investigating the genetic architecture of cerebellar morphology. For
589 instance, we mapped 616 genes associated with inter-individual variability in human cerebellar
590 morphology that have not yet to our knowledge been examined in mouse gene perturbation
591 experiments and/or associated with cerebellar pathology in humans. Among these, we highlight
592 *MAP2K5* and *GRB14*, two HAR-linked genes mapped across all strategies and associated with lead
593 SNP p-values <1e-50, but whose functions in the brain are largely unknown.

594 The reported results for previously discovered variants, loci and genes add important
595 information regarding regional effects on cerebellar morphology. For instance, while genetic variants
596 linked to the *RELN* gene have previously been associated with volumes of cerebellar vermal lobules
597 VI-X and hemispheric lobule IX[42,43], we here mapped 12 lead SNPs to *RELN* showing heterogeneous
598 effects across the entire cerebellar cortex (but with peak effects overlapping previously described
599 midline and posterior cerebellar regions, see Supplementary Figure 13).

600 The observed genetic overlap between cerebellar morphology and the five mental disorders
601 reinforces the recent notion of the cerebellum as a key brain structure in complex clinical traits and
602 disorders[6-10]. Across the five diagnoses, the strongest evidence for genetic overlap with cerebellar
603 morphology was found for SCZ and BIP, likely at least in part because these disorder GWASs were

relatively well-powered. While an in-depth discussion of genetic loci jointly influencing psychiatric disorders and cerebellar morphology is beyond the scope of this report, we note that the Reelin pathway again emerges in the genetic overlap analyses for SCZ and BIP. Specifically, the current finding of *LRP8* (a reelin receptor, and HAR-linked gene) as one of the genes jointly associated with cerebellar morphology and the aforementioned severe mental disorders points towards a potential molecular pathway involved in the cerebellar abnormalities previously reported in SCZ[10,67,68] and BIP[67,68]. Indeed, in line with its key importance for brain development and function, the Reelin pathway is also increasingly seen as relevant for a wide range of neurodevelopmental, psychiatric and neurodegenerative disorders[69]. Of particular relevance to the current findings, converging evidence supports LRP8 as a key susceptibility gene for psychosis[70].

The main limitations of the current study concern the ancestral homogeneity of the sample, the sample size, the exclusion of very rare genetic variants and the limitations on follow-up analyses placed by multivariate (relative to univariate) test-statistics. Limiting the sample to participants of European ancestry was deemed necessary considering the current state of the multivariate GWAS methods used but may limit the generalizability of our findings. Second, while the current sample size is large in comparison with previous imaging genetics studies, it is still relatively small in comparison to GWASs of other complex human phenotypes (e.g., intelligence, with a current n of > 3 million[71]). Moreover, very rare genetic variants (MAF < 0.005) were excluded from the current multivariate GWAS, but are likely to include a number of variants with relatively large effect sizes on complex human traits [72]. Thus, future studies using larger and more diverse samples – as well as whole exome and/or genome sequencing – are likely to discover more of the genetic variants associated with cerebellar structure. Finally, while multivariate GWAS increases the power to detect genetic variants associated with brain phenotypes relative to univariate approaches, it also places important limitations on the possible follow-up analyses (e.g., genetic correlations or Mendelian randomization), which require directional effects. Although comprehensive follow-up analyses of univariate GWAS results fall beyond the scope of the current report, we have made all summary statistics (univariate and multivariate) publicly available for the research community.

In conclusion, the current results enhance our understanding of the genetic architecture of human cerebellar morphology, provide supporting evidence for cerebellar morphological changes during the last ~6-8 million years of human evolution, and reinforce the notion of cerebellar involvement in several mental disorders by demonstrating significant genetic overlap.

642 **Materials and Methods:**

643

644 **Participants**

645 For our main analyses, T1-weighted MR images, demographic and genetic data from 39,178 UK
646 Biobank participants were accessed using access number 27412. After removing 1043 participants
647 who either had missing genetic data or had withdrawn consent (as of 19.11.2019), data from 38,135
648 participants remained for the main analysis (age range: 44.6-82.1; mean age: 64.1, 51.9% female).
649 Following quality control procedures (QC, see below), 28,212 UK Biobank participants of European
650 descent remained for the main analyses (age range: 45.1-82.1; mean age: 64.1, 55.1% female), 910
651 of which were identified as close relatives and removed prior to genetic analyses (see below), leaving
652 a final sample for the primary analyses of 27,302 (age range: 45.1-82.1; mean age: 64.1, 54.9%
653 female). For the replication sample, we accessed a newer release of UK Biobank participants (n =
654 48,045) and removed the 27,302 participants included in the primary analyses. After running through
655 identical QC procedures as for the main sample (although applied only to the cerebellar features of
656 primary interest and covariates included when analyzing these), the replication sample consisted of
657 11,264 unrelated UK Biobank participants of European descent (age range: 46.1-83.7; mean age:
658 66.8, 46.9% female). The UK Biobank was approved by the National Health Service National
659 Research Ethics Service (ref. 11/NW/0382),

660

661 **Initial MR image processing**

662 MRI data was first processed using the recon-all pipeline in Freesurfer 5.3[73], yielding a large number
663 of brain features. Of these, we retained measures of estimated total intracranial volume (eTIV), total
664 cerebro-cortical surface area, average cerebro-cortical thickness as well as the volumes of seven
665 subcortical structures (hippocampus, amygdala, thalamus, pallidum, putamen, caudate nucleus and
666 nucleus accumbens) as cerebral (or global, in the case of eTICV) comparison regions for our main
667 cerebellar analyses. These anatomical features were averaged across hemispheres, yielding a total
668 of ten comparison phenotypes.

669 Next, the bias-field corrected T1-images from the FreeSurfer analyses were analyzed using
670 the cerebellum-optimized SUIT-toolbox[74]. In brief, SUIT isolates the cerebellum and brain stem from
671 T1-images, segments cropped images into grey and white matter, and normalizes these tissue
672 probability maps to a cerebellum-specific anatomical template. After multiplying the grey matter
673 maps with the Jacobian of the transformation matrix (i.e. preserving information about absolute
674 volumes), we extracted grey matter intensity values overlapping the 28 cerebellar lobular labels in
675 the probabilistic SUIT-atlas. Since cerebellar volumetric indices showed high correlations across the
676 two hemispheres (range: .87-.96; mean: .94), we created mean volumetric measures by averaging
677 across hemispheres. Finally, we also combined the two smallest regions in the SUIT-atlas (Vermis
678 Crus I and II located in the midline, average volumes: 2.9 and 293.4 mm^2, respectively) to create a
679 new Vermis Crus region (average volume: 296.2 mm^2). This procedure reduced the 28 cerebellar

lobular volumes to 16 morphological indices. Total cerebellar grey matter volume was defined as the sum of all 28 lobular labels in the SUIT-atlas.

Quality control procedures

After excluding 4,818 UKBB participants of non-European origin, anatomical indices from the remaining 33,317 participants went through an iterative quality control (QC) procedure. First, we excluded 639 subjects with a mean Euler number below -217, indicating poor MRI quality[31], as well as 12 subjects with missing and 90 subjects with zero values for any of the key cerebellar or cerebral brain measures of interest. Next, we used general additive models (GAM, implemented in the R-package "mgcv") in order to model the effects of age (estimated as smooth functions for males and females separately, using cubic splines with 10 knots), sex, and scanner site on estimated total intracranial volume (eTIV) and mean thickness of the cerebral cortex. Specifically, we used the following formula:

GAM_model<-gam(Anatomical feature ~ s(Age, bs = "cs", by = Sex, k = 10) + as.factor(Sex) + as.factor(Scanner), data = UKBB_cerebellum_GWAS, select = TRUE, method = "REML")

Adjusted eTIV and mean cortical thickness indices were then created by reconstructing the data using the intercept and residuals from this model (i.e., removing effects of age, sex and scanner), before identifying and rejecting potential outliers, defined as +/- 3 median absolute deviations (MAD)[75] from the median of these adjusted values. Data from 350 subjects were rejected based on these criteria.

For the remaining cerebral anatomical measures, as well as total cerebellar volume, this procedure was then repeated, with scaled eTIV as an additional predictor. Since previous studies have demonstrated that the relationship between regional brain volumes and intracranial volume is not strictly allometric[76,77], we estimated the effect of eTIV using cubic splines with 10 knots. Specifically, we used the following formula:

GAM_model<-gam(Anatomical feature ~ s(Age, bs = "cs", by = Sex, k = 10) + as.factor(Sex) + as.factor(Scanner) + s(eTIV_scaled, bs = "cs", k = 10), data = UKBB_cerebellum_GWAS, select = TRUE, method = "REML")

In order to be maximally sensitive to outliers in relative cerebellar volumes, we replaced eTIV with total cerebellar volume in the GAM models of cerebellar regions of interest (i.e., SUIT atlas regions), using the following formula:

GAM_model<-gam(Anatomical feature ~ s(Age, bs = "cs", by = Sex, k = 10) + as.factor(Sex) + as.factor(Scanner) + total cerebellar volume, data = UKBB_cerebellum_GWAS, select = TRUE, method = "REML")

Adjusted cerebral and cerebellar indices were then created by reconstructing the data using the intercept and residuals from these models (i.e., removing estimated effects of age, sex and eTIV or total cerebellar volume), before rejecting 1792 participants with potential outlier cerebral indices and 2222 participants with potential outlier cerebellar indices (MAD > +/- 3).

This iterative QC procedure resulted in the rejection of 5,105 (i.e., 15.3%) of the original 33,317 datasets, leaving 28,212 datasets for further analysis.

Non-negative matrix factorization

Cerebellar grey matter maps from 28,212 subjects passing the iterative QC procedure were smoothed with a 4mm full-with-half-maximum gaussian kernel in SPM12[78], concatenated across all subjects and multiplied with a binary mask to isolate voxels located in the cerebellar cortex. This cerebellar cortical mask was constructed by multiplying a binary mask containing all 28 cerebellar lobules of the SUIT-atlas with the thresholded (at a value of 0.1) mean (unsmoothed) grey matter segmentation across all 28,212 participants.

The smoothed, concatenated and masked grey matter maps were then subjected to orthogonal projective non-negative matrix-factorization (OPNMF)[26], in order to derive data-driven parcellations of regional cerebellar grey matter volume.

Non-negative matrix factorization (NNMF) is a blind source separation technique that allows structural brain networks to be described in a hypothesis-free, data-driven way by identifying patterns of covariation in the data. In contrast to alternative techniques, such as principal component analysis and independent component analysis, which yield components with both positive and negative weights that are often difficult to interpret, NNMF produces a sparse, positive-only, parts-based representation of the data. Importantly, NNMF has previously been proven effective in estimating covariance patterns in neuroimaging data while providing an easier interpretation of the results than other matrix decomposition techniques such as principal component analysis (PCA) or independent component analysis (ICA)[26-30].

Briefly, NNMF decomposes an input matrix (voxels×subjects) into two matrices; a component matrix W (voxels×k) and a weight matrix H (k×subjects) where *k* is the number of components that needs to be specified by the user. Here, we used an implementation of orthogolnal projective non-negative matrix factorization previously used in a number of publications[26-30] and downloaded from https://github.com/asotiras/brainparts. Given the large number of participants, we used the opnmf_mem.m function, which has been optimized for high-dimensional data. The function was run

753 with default parameter settings, exept for maximal number of iterations, which was increased to
754 200k, in order to ensure full convergence across all tested model orders.
755
756 **Model order selection**
757 Since the resulting parcellations are highly dependent on the requested number of components (or
758 model order) specified, we tested model orders ranging from 2 to 30 in steps of 1, as well as from
759 30 to 100 in steps of 10. Binarized winner-winner-takes-all (i.e. assigning each voxel to the NNMF
760 component with the highest loading) maps of the resulting decompositions projected onto a flattened
761 representation of the cerebellar cortex are shown in Supplementary Figure 1.
762 Resulting NNMF decompositions were then evaluated based on two criteria; 1) how much of the
763 variance in the original input data a given NNMF solution (i.e., component maps and subject weights)
764 could explain and on 2) how reproducible the component maps were. As a metric of change in
765 explained variance we used the change of the Frobenius norm of the reconstruction error. With
766 increasing model orders the variance explained will always increase and the reconstruction error
767 decrease, but if the decrease in the reconstruction error (or gradient) levels off, this indicates that
768 the intrinsic dimensionality of the data might have been approximated (and that the subsequent
769 increase in explained variance can largely be attributed to fitting random noise in the input data). In
770 order to assess reproducibility, we split the full dataset into two equal sets (matched with respect to
771 scanner site, n = 14,105 and 14,107, respectively), and ran NNMF on each split-half sample. For
772 each set of independent NNMF parcellations, we computed two reproducibility indices. First, for each
773 model order we matched components across split-half runs using the Hungarian algorithm[79], before
774 computing the spatial correlations between matched component maps, and extracting the median
775 correlation across all matched components as our first reproducibility index. For our second
776 reproducibility index, we first computed one overall – and categorical – component map using a
777 "winner-takes-all"-approach, i.e., assigning each voxel to the NNMF component with the highest
778 loading. Next, we calculated the adjusted Rand index (ranging from 0 to 1, with higher values
779 indicating greater similarity[80], across the two categorical parcellation maps for each model order as
780 our second metric of reproducibility.
781 As can be seen in Supplementary Figure 2A, for lower model orders, increasing the model
782 order resulted in a sharp decrease in the reconstruction error, indicating that models with more
783 components resulted in a significantly better fit to the data. However, after reaching model orders
784 between 15 and 30, the incremental improvement in fit from adding another component appeared to
785 level off. As expected, the reproducibility results showed a largely inverse pattern, with both
786 reproducibility indices decreasing with increasing model orders (Supplementary Figure 2B). Of note,
787 for model orders up to 8, median spatial correlations across all matched components was above .99,
788 indicating almost identical parcellations derived from the two independent samples. However, even
789 for a model order of 100, the median pairwise correlation was still pretty high (.85), with 60% of
790 components showing pairwise spatial correlations above 0.8, suggesting a reasonable level of

reproducibility even for the most fine-grained parcellation. Our reproducibility index for the categorical parcellations showed very similar results; the adjusted Rand index remained above .9 for model orders up to 8, and then decreased to 0.56 at a model order of 100.

Given that the change in reconstruction error appeared to stabilize between 15 and 30 components, indicating that the intrinsic dimensionality of the data had been approximated, we searched for the most reproducible parcellations within this range. Since the reliability estimates for model orders of 16 and 23 were very similar, with the 23-component solution explaining more of the variance in the original data, we used the 23-component parcellation for all further analyses.

Adjustment and rank-order normalization of anatomical indices

Prior to being subjected to genome-wide association analyses, we adjusted all anatomical indices for effects of age, sex, estimated total intracranial volume, scanner site, 40 genetic population components, genetic batch and mean Euler number (i.e, an index of MRI image quality[31], averaged across hemispheres), using general additive models. Finally, all adjusted anatomical indices were inverse rank normalized[81].

Pre-processing of genetic data

For all genetic analyses we made use of the UKB v3 imputed data, which has undergone extensive quality control procedures as described by the UKB genetics team[82]. After converting the BGEN format to PLINK binary format, we additionally carried out standard quality check procedures. We first selected White Europeans, as determined by a combination of self-identification as 'White British' and similar genetic ancestry based on genetic principal components, that had undergone the neuroimaging protocol. We then filtered out individuals with more than 10% missingness, removed SNPs with low imputation quality scores (INFO <.5), SNPs with more than 10% missingness, and SNPs failing the Hardy-Weinberg equilibrium test at p=1*10-9. We further set a minor allele frequency threshold of 0.005 leaving 9,061,238 SNPs. After estimating the genetic relationship matrix (GRM) using genetic complex trait analysis (GCTA[33]), we finally removed 910 participants defined as close relatives using a threshold of 0.05 (approximately corresponding to 3rd cousins).

Heritability estimation and genetic correlation analyses

SNP-based heritability estimates for all morphological features – as well as the pairwise genetic correlations between cerebellar features - were estimated using genetic complex trait analysis (GCTA[33]).

Univariate genome-wide association analyses

Univariate analyses of total cerebellar grey matter volume and the ten cerebral comparison phenotypes were conducted using Plink v1.9.

829 **Genetic correlation analyses across discovery and replication samples and with previous**
830 **published GWASs**
831 Genetic correlation analyses across different samples were conducted using LD-score
832 regression[32].
833
834 **Multivariate genome-wide association analyses**
835 For our main analysis, we used a recently developed multivariate analysis method (MOSTest[36]), to
836 conduct a multivariate genome-wide association (GWA) analysis on cerebellar morphological
837 features. MOSTest identifies genetic effects across multiple phenotypes, yielding a multivariate
838 GWAS summary statistic across all 23 features, and provides robust (permutation based) test
839 statistics. For mathematical details of the implementation, see van der Meer et al. (2020)[36], for details
840 on the software implementation see github.com/precimed/mostest. MOSTest has been extensively
841 validated in the original methods paper, including simulations and comparisons with other methods
842 that have confirmed its solid discovery performance as well as an order of magnitude shorter runtime
843 compared to other tools[36]. For comparison to standard univariate approaches, we also performed
844 univariate GWASs (extracted from the univariate stream of MOSTest[36] and identical to results from
845 analyses using Plink).
846
847 **Multivariate replication analysis**
848 To ensure that not only single locus lead SNP associations replicate, but that also the multivariate
849 pattern of these associations are consistent in the discovery and replication sample, we implemented
850 a multivariate replication procedure established in Loughnan et al.[39]. In brief, for each locus lead
851 SNP identified in the multivariate analysis in the discovery sample, this procedure derives a
852 composite score from the mass-univariate *z*-statistics and tests for associations of the composite
853 score with the genotype in the replication sample (for mathematical formulation see Loughnan et
854 al.[39]). 12 of the 351 locus lead SNPs could not be tested as they were not available in the replication
855 sample after QC. For the remaining SNPs we report the percent of loci replicating at $P < 0.05$, the
856 percent remaining significant after Bonferroni correction for 339 conducted tests, and the percent of
857 lead SNPs showing the same effect direction.
858
859 **Multivariate cerebral comparison phenotypes**
860 To compare key multivariate cerebellar results with other multivariate brain phenotypes, we
861 downloaded summary statistics from two recent studies on cerebrocortical[44] and hippocampal[45]
862 regional morphology. These comparison summary statistics were next analyzed using FUMA[37] as
863 described for the main cerebellar results below.
864
865 **Locus identification and SNP annotation**

866 To identify genetic loci we uploaded summary statistics to the FUMA platform v1.4.1[37] Using the
867 1000GPhase3 EUR as reference panel, we identified independent SNPs as SNPs below the
868 significance threshold of $P < 5e{-}8$ that were also in linkage equilibrium with each other at $r^2 < 0.6$.
869 For each independent SNP, all candidate variants are identified as variants with LD r2≥ 0.6 with the
870 independent SNP. A fraction of the independent significant SNPs in approximate linkage equilibrium
871 with each other at $r^2 < 0.1$ were considered as lead SNPs. For a given lead SNP, the borders of the
872 genomic locus are defined as min/max positional coordinates over all corresponding candidate
873 SNPs. Finally, loci are merged if they are separated by less than 250kb.
874 FUMA further annotates associated SNPs based on functional categories, Combined
875 Annotation Dependent Depletion (CADD) scores which predicts the deleteriousness of SNPs on
876 protein structure/function[54], RegulomeDB scores which predicts regulatory functions[83]; and
877 chromatin states that shows the transcription/regulation effects of chromatin states at the SNP
878 locus[84]. For all these analyses, we used default FUMA parameters.
879
880 **Genome-wide gene-based association and gene-set analyses**
881 We conducted genome-wide gene-based association and gene-set analyses using MAGMA v.1.10[49]
882 (http://ctg.cncr.nl/software/magma). MAGMA performs multiple linear regression to map the input
883 SNPs to 19,190 protein coding genes and estimates the significance value of that gene. Genes were
884 considered significant if the *P* value was <0.05 after Bonferroni correction for 19,190 genes. The
885 same procedure was used for MAGMA analysis of summary statistics for the three multivariate
886 cerebral comparison phenotypes (cerebrocortical thickness and surface area[44] and regional
887 hippocampal volumes[45]).
888 MAGMA gene-level statistics were next used as input to gene-property and gene-set
889 analyses in MAGMA. Gene-property analyses test for associations between tissue specific gene
890 expression profiles and disease-gene associations. The gene-property analysis is based on the
891 regression model: $Z \sim \beta_0 + E_t\beta_E + A\beta_A + B\beta_B + \epsilon$, where Z is a gene-based Z-score converted from
892 the gene-based P-value, B is a matrix of several technical confounders included by default (e.g.,
893 gene size, gene density, sample size), E_t is the gene expression value of a testing tissue type and A
894 is the average expression across all tissue types in a data set (ensuring a test of expression
895 specificity). We performed a one-sided test ($\beta_E > 0$) which is essentially testing the positive
896 relationship between tissue specificity and genetic association of genes.
897 We tested associations with two regional brain gene expression datasets (Allen Human
898 Brain Atlas[35] and GTEx) and one developmental brain gene expression dataset (BrainSpan). For
899 extraction and processing of gene expression data, see below.
900
901 **Extraction and processing of gene expression data**
902 *Allen Human Brain Atlas:* Regional microarray expression data were obtained from 6 post-mortem
903 brains (1 female, ages 24.0--57.0, 42.50 +/- 13.38) provided by the Allen Human Brain Atlas[35]. Data

were processed with the abagen toolbox (version 0.1.3) [34] using two volumetric atlases; 1) the binarized 23-region NNMF-derived parcellation of the cerebellar cortex; and 2) a modified version of the Desikan atlas were ROIs were merged to construct 9 bilateral regions: cerebellum, cerebral cortex, pallidum, caudate, putamen, thalamus, amygdala, nucleus accumbens and hippocampus.

First, microarray probes were reannotated using data provided by Arnatkeviciute, Fulcher and Fornito[85]; probes not matched to a valid Entrez ID were discarded. Next, probes were filtered based on their expression intensity relative to background noise[86], such that probes with intensity less than the background in >=50.00% of samples across donors were discarded, yielding 31,569 probes. When multiple probes indexed the expression of the same gene, we selected and used the probe with the most consistent pattern of regional variation across donors (i.e., differential stability[87]), calculated with:

$$\Delta_S(p) = \frac{1}{\binom{N}{2}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho[B_i(p), B_j(p)]$$

where *p* is Spearman's rank correlation of the expression of a single probe, p, across regions in two donors B_i and B_j, and N is the total number of donors.

Here, regions correspond to the structural designations provided in the ontology from the AHBA. The MNI coordinates of tissue samples were updated to those generated via non-linear registration using the Advanced Normalization Tools (ANTs; https://github.com/chrisfilo/alleninf). Samples were assigned to brain regions in the provided atlas if their MNI coordinates were within 2 mm of a given parcel. To reduce the potential for misassignment, sample-to-region matching was constrained by hemisphere and gross structural divisions (i.e., cortex, subcortex/ brainstem, and cerebellum, such that e.g., a sample in the left cortex could only be assigned to an atlas parcel in the left cortex[85]). All tissue samples not assigned to a brain region in the provided atlas were discarded. \n\nInter-subject variation was addressed by normalizing tissue sample expression values across genes using a robust sigmoid function[88]:

$$x_{norm} = \frac{1}{1+\exp(-\frac{(x-\langle x \rangle)}{\mathrm{IQR}_x})}$$

where (x) is the median and IQR_z is the normalized interquartile range of the expression of a single tissue sample across genes. Normalized expression values were then rescaled to the unit interval:

$$x_{scaled} = \frac{x_{norm} - \min(x_{norm})}{\max(x_{norm}) - \min(x_{norm})}$$

Gene expression values were then normalized across tissue samples using an identical procedure. Samples assigned to the same brain region were averaged separately for each donor

935 and then across donors, yielding two regional expression matrices with 23 and 9 rows, corresponding
936 to brain regions, and 15,631 and 15,633 columns, corresponding to the retained genes. Prior to
937 inclusion in MAGMA gene property analyses, we converted gene names for the modified Desikan
938 atlas to ENSMBL IDs, and calculated the mean expression value across tissue types (in order to
939 include this as a covariate in MAGMA analyses testing for gene expression specificity), resulting in
940 a 10 (regions) by (15,490) gene expression matrix.
941
942 *GTeX:* Text files containing median transcript per millimeter (TPM) values for 53 tissue types
943 (GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_median_tpm.gct.gz) were down-loaded
944 from the GTEx portal (https://gtexportal.org/home/datasets/). After selecting only expression data
945 from the seven (out of nine) comparison brain regions present in the GTEx dataset (i.e., amygdala,
946 caudate, cerebellum, cortex, hippocampus, nucleus accumbens and putamen), we filtered the data
947 by only including genes with median TPM values above 1 for at least one of these tissue type
948 (retaining 19,578 of 56,200 annotated genes). Following the procedure used by FUMA[37], we next
949 winsorized median TPM values at 50 (i.e., replaced TPM>50 with 50), before log transforming TPM
950 with pseudocount 1 (log2(RPKM+1)). Finally, we calculated the mean expression value across tissue
951 types (in order to include this as a covariate in MAGMA analyses testing for gene expression
952 specificity).
953
954 *BrainSpan data:* The analysis of BrainSpan data testing for developmentally specific brain
955 expression was performed entirely within FUMA v1.4.1., using default parameters.
956
957 **Analysis of Human Dating Genome Data.** The Atlas of Variant Age for chromosomes 1-22 was
958 downloaded from the Human Genome Dating (HGD) website: https://human.genome.dating/. This
959 atlas contains more than 45 million SNPs which has been assigned dates of origin based on a
960 recombination clock and mutation clock applied to two large-scale sequencing datasets (1000
961 Genomes Project[89] and The Simons Genome Diversity Project[90]), with no assumptions made about
962 demographic or selective processes[46]. The current study used the median joint age estimates from
963 both clocks when analyzing SNPs present in both datasets in combination (i.e., 13,694,493 SNPs).
964 After merging dated SNPs with the 40,405,505 SNPs also present in the Haplotype
965 Reference Consortium reference data (to add minor allele frequencies, MAF) and removing 715,083
966 (5.2%) SNPs with missing MAF values as well as the very few (14,549, 0.1%) SNPs dated older
967 than 2 million years, 12,960,066 SNPs remained.
968 548 of these dated SNPs were matched to lead SNPs linked to cerebellar morphology
969 (defined as being in mutual LD at an r^2 threshold < 0.1), hereafter referred to as cerebellar-SNPs.
970 Partially because very rare variants (Minor Allele Frequency/MAF < 0.005) had been removed prior
971 to the multivariate GWAS analysis, MAF was not equally distributed between these cerebellar-SNPs
972 and the full range of dated SNPs in the HGD dataset. Importantly, MAF has been shown to be

973 systematically related to the estimated age of SNPs, with a higher proportion of low-MAF SNPs in
974 more recent than in older time-bins[46,47]. Consequently, following the analysis approach established
975 by Libedinsky et al. [47], we first determined the MAF-distribution of cerebellar-SNPs (across eight
976 bins: <0.0001; 0.0001-0.001; 0.001-0.01; 0.01-0.1; 0.1-0.2; 0.2-0.3; 0.3-0.4; 0.4-0.5) and selected
977 random set of 2,892,270 SNPs from the HGD dataset that were matched to the cerebellar-SNPs in
978 terms of MAF-bin distribution. Similar MAF-matched HDG subsets were created for lead SNPs
979 asociated with the two cerebrocortical comparison phenotypes, i.e., regional cerebrocortical surface
980 area and thickness (862 and 714 lead SNPs, respectively).
981 For statistical inference we constructed null models (separate for each brain phenotype) by
982 randomly drawing sets of SNPs (of equal size to the number of phenotype-linked lead SNPs) from
983 the MAF-matched HGD-datasets and computing the histograms of estimated dates from 0 to 2
984 million years ago (divided into 100 bins of 20.000 years) over 10,000 iterations. From these null
985 models we extracted bin means as well as significance thresholds (defined as the upper and lower
986 99.95th percentile of the null model (i.e. corresponding to a two-tailed threshold of 0.05 Bonferroni
987 corrected across 100 bins).
988 For the validation analyses, we ran the same analyses based on the larger number (range
989 across phenotypes: 1574-2883) of individual significant SNPs (defined as being in mutual LD at an
990 r^2 threshold < 0.6).
991
992 **Positional and eQTL mapping of SNPs to plausible causal genes**
993 In addition to using MAGMA, we also mapped candidate SNPs to plausible causal genes using two
994 complementary gene mapping strategies implemented ion FUMA[37]: 1) Positional mapping of
995 deleterious SNPs (defined as having a CADD-score > 12.37) and 2) eQTL-mapping of SNPs
996 previously shown to alter gene expression in cerebellar tissue (from the BRAINEAC and GTEx v8
997 databases). These analyses were run with default FUMA parameters. For the three multivariate
998 cerebral comparison phenotypes (i.e., cerebrocortical thickness, cerebrocortical surface area and
999 hippocampal regional volumes), we employed identical gene mapping procedures to our cerebellar
1000 morphology results, except for the tissues chosen for eQTL mapping (cerebrocortical and
1001 hippocampal, respectively).
1002
1003 **Gene set analyses using lists of mapped genes**
1004 All gene set analyses using mapped genes were conducted using the hypeR R-package [91]. This
1005 package implements the hypergeometric test (also known as Fisher's exact test), which assigns a
1006 p-value to gene-set overlaps given gene set sizes and the number of background genes. This R-
1007 package also contains functions for downloading and/or formatting gene sets. The following gene
1008 sets were accessed using hypeR: "Allen_Brain_Atlas_up" (regional overexpression in the Allen
1009 Muse Brain Atlas), "MGI_Mammalian_Phenotype_Level_4_2021", "DisGeNet" (all from the enrichR
1010 platform. In addition, we downloaded sets of genes with regional overexpression in the Allen Human

1011 Brain Atlas from the Harmonizome platform[92] and accessed a list of genes mapped to human
1012 accelerated regions[93]. For all gene-set analyses we employed Bonferroni-correction by dividing the
1013 p-value threshold of 0.05 by the number of gene sets included in each analysis.
1014
1015 **Genetic overlap between cerebellar morphology and brain disorders**
1016 We accessed GWAS summary statistics for attention deficit hyperactivity disorder (ADHD)[94], autism
1017 spectrum disorder (ASD)[95], bipolar disorder (BIP)[96] and major depressive disorder (MDD)[97] from the
1018 Psychiatric Genomics Consortium. In order to avoid sample overlap, for MDD we used summary
1019 statistics based on a sample with the UK Biobank participants removed. 23&me participants included
1020 in the original MDDGWAS were also excluded, since these data are not freely available). Finally,
1021 we included data from a recent study of schizophrenia (SCZ)[98]. Shared variants associated with
1022 cerebellar morphology and each of the above-mentioned brain disorders were identified using
1023 conjunctional FDR statistics (FDR < 0.05)[99,100]. In contrast to genetic correlation analysis,
1024 conjunctional FDR does not require effect directions and can therefore be applied to summary
1025 statistics from multivariate GWAS, which do not contain effect directions. Two genomic regions, the
1026 extended major histocompatibility complex genes region (hg19 location Chr 6: 25119106–33854733)
1027 and chromosome 8p23.1 (hg19 location Chr 8: 7242715–12483982) for all cases and *APOE* region
1028 for ASD, were excluded from the FDR-fitting procedures because complex correlations in regions
1029 with intricate LD can bias FDR estimation. We further controlled for spurious enrichment by
1030 calculating all conditional Q-Q plots after random pruning averaged over 500 iterations. At each
1031 iteration, one SNP in every LD block (defined by an r2 >0.1) was randomly selected and the empirical
1032 cumulative distributions were computed using the corresponding p-values. Finally, we submitted
1033 the results from conjunctional FDR to FUMA v1.3.7[37] to annotate the genomic loci with conjFDR
1034 value < 0.10 having an $r^2 \geq 0.6$ with one of the independent significant lead SNPs. Genetic
1035 correlations between univariate results for the 23 cerebellar features, the 10 comparison brain
1036 phenotypes, and each of the five mental disorders were computed using LD-score regression as
1037 described above.
1038

References for Online Methods:

1. Herculano-Houzel, S. (2016). The human advantage : a new understanding of how our brain became remarkable (The MIT Press).
2. Barton, R.A., and Venditti, C. (2014). Rapid Evolution of the Cerebellum in Humans and Other Great Apes. Curr. Biol.
3. Sereno, M.I., Diedrichsen, J., Tachrount, M., Testa-Silva, G., d'Arceuil, H., and De Zeeuw, C. (2020). The human cerebellum has almost 80% of the surface area of the neocortex. Proc Natl Acad Sci U S A *117*, 19538-19543. 10.1073/pnas.2002896117.
4. Harrison, P.W., and Montgomery, S.H. (2017). Genetics of Cerebellar and Neocortical Expansion in Anthropoid Primates: A Comparative Approach. Brain Behav Evol *89*, 274-285. 10.1159/000477432.
5. Leiner, H.C., Leiner, A.L., and Dow, R.S. (1993). Cognitive and language functions of the human cerebellum. Trends in Neurosciences.
6. Strick, P.L., Dum, R.P., and Fiez, J.A. (2009). Cerebellum and Nonmotor Function. Annu. Rev. Neurosci.
7. Buckner, R.L. (2013). The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. Neuron.
8. King, M., Hernandez-Castillo, C.R., Poldrack, R.A., Ivry, R.B., and Diedrichsen, J. (2019). Functional boundaries in the human cerebellum revealed by a multi-domain task battery. Nat Neurosci *22*, 1371-1378. 10.1038/s41593-019-0436-x.
9. Wang, S.S., Kloth, A.D., and Badura, A. (2014). The cerebellum, sensitive periods, and autism. Neuron *83*, 518-532. 10.1016/j.neuron.2014.07.016.
10. Moberget, T., Doan, N.T., Alnaes, D., Kaufmann, T., Cordova-Palomera, A., Lagerberg, T.V., Diedrichsen, J., Schwarz, E., Zink, M., Eisenacher, S., et al. (2018). Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: a multisite mega-analysis of 983 patients and 1349 healthy controls. Mol Psychiatry *23*, 1512-1520. 10.1038/mp.2017.106.
11. Adams, H.H., Hibar, D.P., Chouraki, V., Stein, J.L., Nyquist, P.A., Renteria, M.E., Trompet, S., Arias-Vasquez, A., Seshadri, S., Desrivieres, S., et al. (2016). Novel genetic loci underlying human intracranial volume identified through genome-wide association. Nat Neurosci *19*, 1569-1582. 10.1038/nn.4398.
12. van der Meer, D., Rokicki, J., Kaufmann, T., Cordova-Palomera, A., Moberget, T., Alnaes, D., Bettella, F., Frei, O., Doan, N.T., Sonderby, I.E., et al. (2018). Brain scans from 21,297 individuals reveal the genetic architecture of hippocampal subfield volumes. Mol Psychiatry. 10.1038/s41380-018-0262-7.
13. Chambers, T., Escott-Price, V., Legge, S., Baker, E., Singh, K.D., Walters, J.T.R., Caseras, X., and Anney, R.J.L. (2022). Genetic common variants associated with cerebellar volume and their overlap with mental disorders: a study on 33,265 individuals from the UK-Biobank. Mol Psychiatry *27*, 2282-2290. 10.1038/s41380-022-01443-8.
14. Tissink, E., de Lange, S.C., Savage, J.E., Wightman, D.P., de Leeuw, C.A., Kelly, K.M., Nagel, M., van den Heuvel, M.P., and Posthuma, D. (2022). Genome-wide association study of cerebellar volume provides insights into heritable mechanisms underlying brain development and mental health. Commun Biol *5*, 710. 10.1038/s42003-022-03672-7.
15. Macri, S., Savriama, Y., Khan, I., and Di-Poi, N. (2019). Comparative analysis of squamate brains unveils multi-level variation in cerebellar architecture associated with locomotor specialization. Nat Commun *10*, 5560. 10.1038/s41467-019-13405-w.
16. Smaers, J.B., Turner, A.H., Gomez-Robles, A., and Sherwood, C.C. (2018). A cerebellar substrate for cognition evolved multiple times independently in mammals. Elife *7*. 10.7554/eLife.35696.
17. E, K.-H., Chen, S.-H.A., Ho, M.-H.R., and Desmond, J.E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. Hum Brain Mapp.
18. Stoodley, C.J., and Schmahmann, J.D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. NeuroImage.

19. Buckner, R.L., Krienen, F.M., Castellanos, A., Diaz, J.C., and Yeo, B.T.T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. Journal of Neurophysiology.
20. Sang, L., Qin, W., Liu, Y., Han, W., Zhang, Y., Jiang, T., and Yu, C. (2012). Resting-state functional connectivity of the vermal and hemispheric subregions of the cerebellum with both the cerebral cortical networks and subcortical structures. NeuroImage. Elsevier Inc.
21. Krienen, F.M., and Buckner, R.L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. Cerebral Cortex.
22. Bernard, J.A., Seidler, R.D., Hassevoort, K.M., Benson, B.L., Welsh, R.C., Wiggins, J.L., Jaeggi, S.M., Buschkuehl, M., Monk, C.S., Jonides, J., and Peltier, S.J. (2012). Resting state cortico-cerebellar functional connectivity networks: a comparison of anatomical and self-organizing map approaches. Front Neuroanat.
23. Kipping, J.A., Grodd, W., Kumar, V., Taubert, M., Villringer, A., and Margulies, D.S. (2013). Overlapping and parallel cerebello-cerebral networks contributing to sensorimotor control: an intrinsic functional connectivity study. NeuroImage.
24. Wang, C., Kipping, J., Bao, C., Ji, H., and Qiu, A. (2016). Cerebellar Functional Parcellation Using Sparse Dictionary Learning Clustering. Front Neurosci *10*, 188. 10.3389/fnins.2016.00188.
25. Kipping, J.A., Tuan, T.A., Fortier, M.V., and Qiu, A. (2016). Asynchronous Development of Cerebellar, Cerebello-Cortical, and Cortico-Cortical Functional Networks in Infancy, Childhood, and Adulthood. Cereb Cortex. 10.1093/cercor/bhw298.
26. Sotiras, A., Resnick, S.M., and Davatzikos, C. (2015). Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization. Neuroimage *108*, 1-16. 10.1016/j.neuroimage.2014.11.045.
27. Sotiras, A., Toledo, J.B., Gur, R.E., Gur, R.C., Satterthwaite, T.D., and Davatzikos, C. (2017). Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. Proc Natl Acad Sci U S A *114*, 3527-3532. 10.1073/pnas.1620928114.
28. Varikuti, D.P., Genon, S., Sotiras, A., Schwender, H., Hoffstaedter, F., Patil, K.R., Jockwitz, C., Caspers, S., Moebus, S., Amunts, K., et al. (2018). Evaluation of non-negative matrix factorization of grey matter in age prediction. Neuroimage *173*, 394-410. 10.1016/j.neuroimage.2018.03.007.
29. Nassar, R., Kaczkurkin, A.N., Xia, C.H., Sotiras, A., Pehlivanova, M., Moore, T.M., Garcia de La Garza, A., Roalf, D.R., Rosen, A.F.G., Lorch, S.A., et al. (2019). Gestational Age is Dimensionally Associated with Structural Brain Network Abnormalities Across Development. Cereb Cortex *29*, 2102-2114. 10.1093/cercor/bhy091.
30. Pehlivanova, M., Wolf, D.H., Sotiras, A., Kaczkurkin, A.N., Moore, T.M., Ciric, R., Cook, P.A., Garcia de La Garza, A., Rosen, A.F.G., Ruparel, K., et al. (2018). Diminished Cortical Thickness Is Associated with Impulsive Choice in Adolescence. J Neurosci *38*, 2471-2481. 10.1523/JNEUROSCI.2200-17.2018.
31. Rosen, A.F.G., Roalf, D.R., Ruparel, K., Blake, J., Seelaus, K., Villa, L.P., Ciric, R., Cook, P.A., Davatzikos, C., Elliott, M.A., et al. (2018). Quantitative assessment of structural image quality. Neuroimage *169*, 407-418. 10.1016/j.neuroimage.2017.12.059.
32. Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet *47*, 291-295. 10.1038/ng.3211.
33. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet *88*, 76-82. 10.1016/j.ajhg.2010.11.011.
34. Markello, R.D., Arnatkeviciute, A., Poline, J.B., Fulcher, B.D., Fornito, A., and Misic, B. (2021). Standardizing workflows in imaging transcriptomics with the abagen toolbox. Elife *10*. 10.7554/eLife.72129.

35. Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., et al. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. Nature *489*, 391-399. 10.1038/nature11405.
36. van der Meer, D., Frei, O., Kaufmann, T., Shadrin, A.A., Devor, A., Smeland, O.B., Thompson, W.K., Fan, C.C., Holland, D., Westlye, L.T., et al. (2020). Understanding the genetic determinants of the brain with MOSTest. Nat Commun *11*, 3512. 10.1038/s41467-020-17368-1.
37. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat Commun *8*, 1826. 10.1038/s41467-017-01261-5.
38. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res *38*, e164. 10.1093/nar/gkq603.
39. Loughnan, R.J., Shadrin, A.A., Frei, O., van der Meer, D., Zhao, W., Palmer, C.E., Thompson, W.K., Makowski, C., Jernigan, T.L., Andreassen, O.A., et al. (2022). Generalization of cortical MOSTest genome-wide associations within and across samples. Neuroimage *263*, 119632. 10.1016/j.neuroimage.2022.119632.
40. van der Sluis, S., Posthuma, D., and Dolan, C.V. (2013). TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies. PLoS Genet *9*, e1003235. 10.1371/journal.pgen.1003235.
41. O'Reilly, P.F., Hoggart, C.J., Pomyen, Y., Calboli, F.C., Elliott, P., Jarvelin, M.R., and Coin, L.J. (2012). MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. PLoS One *7*, e34861. 10.1371/journal.pone.0034861.
42. Zhao, B., Luo, T., Li, T., Li, Y., Zhang, J., Shan, Y., Wang, X., Yang, L., Zhou, F., Zhu, Z., et al. (2019). Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. Nat Genet *51*, 1637-1644. 10.1038/s41588-019-0516-6.
43. Smith, S.M., Douaud, G., Chen, W., Hanayik, T., Alfaro-Almagro, F., Sharp, K., and Elliott, L.T. (2021). An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. Nat Neurosci *24*, 737-745. 10.1038/s41593-021-00826-4.
44. van der Meer, D., Kaufmann, T., Shadrin, A.A., Makowski, C., Frei, O., Roelfs, D., Monereo-Sanchez, J., Linden, D.E.J., Rokicki, J., Alnaes, D., et al. (2021). The genetic architecture of human cortical folding. Sci Adv *7*, eabj9446. 10.1126/sciadv.abj9446.
45. Bahrami, S., Nordengen, K., Shadrin, A.A., Frei, O., van der Meer, D., Dale, A.M., Westlye, L.T., Andreassen, O.A., and Kaufmann, T. (2022). Distributed genetic architecture across the hippocampal formation implies common neuropathology across brain disorders. Nat Commun *13*, 3436. 10.1038/s41467-022-31086-w.
46. Albers, P.K., and McVean, G. (2020). Dating genomic variants and shared ancestry in population-scale sequencing data. PLoS Biol *18*, e3000586. 10.1371/journal.pbio.3000586.
47. Libedinsky, I., Wei, Y., Leeuw, C.d., Rilling, J., Posthuma, D., and Heuvel, M.P.v.d. (2023). Genetic timeline of human brain and cognitive traits. bioRxiv, 2023.2002.2005.525539. 10.1101/2023.02.05.525539.
48. Davies, S. (2021). Behavioral Modernity in Retrospect. Topoi *40*, 221-232. 10.1007/s11245-019-09671-4.
49. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol *11*, e1004219. 10.1371/journal.pcbi.1004219.
50. Wei, Y., de Lange, S.C., Scholtens, L.H., Watanabe, K., Ardesch, D.J., Jansen, P.R., Savage, J.E., Li, L., Preuss, T.M., Rilling, J.K., et al. (2019). Genetic mapping and evolutionary analysis of human-expanded cognitive networks. Nat Commun *10*, 4839. 10.1038/s41467-019-12764-8.
51. Won, H., Huang, J., Opland, C.K., Hartl, C.L., and Geschwind, D.H. (2019). Human evolved regulatory elements modulate genes involved in cortical expansion and neurodevelopmental disease susceptibility. Nat Commun *10*, 2396. 10.1038/s41467-019-10248-3.

52. Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-425. 10.1016/j.cels.2015.12.004.
53. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550. 10.1073/pnas.0506580102.
54. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet *46*, 310-315. 10.1038/ng.2892.
55. Andreassen, O.A., Thompson, W.K., Schork, A.J., Ripke, S., Mattingsdal, M., Kelsoe, J.R., Kendler, K.S., O'Donovan, M.C., Rujescu, D., Werge, T., et al. (2013). Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. PLoS Genet *9*, e1003455. 10.1371/journal.pgen.1003455.
56. King, M.Z., Zonglei; Ivry, Richard B.; Weiner, Kevin S. Transcriptomic Gradients Of The Human Cerebellum.
57. Guell, X., Gabrieli, J.D.E., and Schmahmann, J.D. (2018). Triple representation of language, working memory, social and emotion processing in the cerebellum: convergent evidence from task and seed-based resting-state fMRI analyses in a single large cohort. Neuroimage *172*, 437-449. 10.1016/j.neuroimage.2018.01.082.
58. Chen, C.H., Gutierrez, E.D., Thompson, W., Panizzon, M.S., Jernigan, T.L., Eyler, L.T., Fennema-Notestine, C., Jak, A.J., Neale, M.C., Franz, C.E., et al. (2012). Hierarchical genetic organization of human cortical surface area. Science *335*, 1634-1636. 10.1126/science.1215330.
59. Chen, C.H., Fiecas, M., Gutierrez, E.D., Panizzon, M.S., Eyler, L.T., Vuoksimaa, E., Thompson, W.K., Fennema-Notestine, C., Hagler, D.J., Jr., Jernigan, T.L., et al. (2013). Genetic topography of brain morphology. Proc Natl Acad Sci U S A *110*, 17089-17094. 10.1073/pnas.1308091110.
60. Guissart, C., Latypova, X., Rollier, P., Khan, T.N., Stamberger, H., McWalter, K., Cho, M.T., Kjaergaard, S., Weckhuysen, S., Lesca, G., et al. (2018). Dual Molecular Effects of Dominant RORA Mutations Cause Two Variants of Syndromic Intellectual Disability with Either Autism or Cerebellar Ataxia. Am J Hum Genet *102*, 744-759. 10.1016/j.ajhg.2018.02.021.
61. Chi, C.L., Martinez, S., Wurst, W., and Martin, G.R. (2003). The isthmic organizer signal FGF8 is required for cell survival in the prospective midbrain and cerebellum. Development *130*, 2633-2644. 10.1242/dev.00487.
62. Li, S., Qiu, F., Xu, A., Price, S.M., and Xiang, M. (2004). Barhl1 regulates migration and survival of cerebellar granule cells by controlling expression of the neurotrophin-3 gene. J Neurosci *24*, 3104-3114. 10.1523/JNEUROSCI.4444-03.2004.
63. de Sousa, A.A., Beaudet, A., Calvey, T., Bardo, A., Benoit, J., Charvet, C.J., Dehay, C., Gomez-Robles, A., Gunz, P., Heuer, K., et al. (2023). From fossils to mind. Commun Biol *6*, 636. 10.1038/s42003-023-04803-4.
64. Neubauer, S., Hublin, J.J., and Gunz, P. (2018). The evolution of modern human brain shape. Sci Adv *4*, eaao5961. 10.1126/sciadv.aao5961.
65. Valence, S., Garel, C., Barth, M., Toutain, A., Paris, C., Amsallem, D., Barthez, M.A., Mayer, M., Rodriguez, D., and Burglen, L. (2016). RELN and VLDLR mutations underlie two distinguishable clinico-radiological phenotypes. Clin Genet *90*, 545-549. 10.1111/cge.12779.
66. Larouche, M., Beffert, U., Herz, J., and Hawkes, R. (2008). The Reelin receptors Apoer2 and Vldlr coordinate the patterning of Purkinje cell topography in the developing mouse cerebellum. PLoS One *3*, e1653. 10.1371/journal.pone.0001653.
67. Wolfers, T., Doan, N.T., Kaufmann, T., Alnaes, D., Moberget, T., Agartz, I., Buitelaar, J.K., Ueland, T., Melle, I., Franke, B., et al. (2018). Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. JAMA Psychiatry *75*, 1146-1155. 10.1001/jamapsychiatry.2018.2467.

68. Wolfers, T., Rokicki, J., Alnaes, D., Berthet, P., Agartz, I., Kia, S.M., Kaufmann, T., Zabihi, M., Moberget, T., Melle, I., et al. (2021). Replicating extensive brain structural heterogeneity in individuals with schizophrenia and bipolar disorder. Hum Brain Mapp *42*, 2546-2555. 10.1002/hbm.25386.
69. Armstrong, N.C., Anderson, R.C., and McDermott, K.W. (2019). Reelin: Diverse roles in central nervous system development, health and disease. Int J Biochem Cell Biol *112*, 72-75. 10.1016/j.biocel.2019.04.009.
70. Li, M., Huang, L., Grigoroiu-Serbanescu, M., Bergen, S.E., Landen, M., Hultman, C.M., Forstner, A.J., Strohmaier, J., Hecker, J., Schulze, T.G., et al. (2016). Convergent Lines of Evidence Support LRP8 as a Susceptibility Gene for Psychosis. Mol Neurobiol *53*, 6608-6619. 10.1007/s12035-015-9559-6.
71. Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S.M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., et al. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nat Genet *54*, 437-449. 10.1038/s41588-022-01016-z.
72. Wang, Q., Dhindsa, R.S., Carss, K., Harper, A.R., Nag, A., Tachmazidou, I., Vitsios, D., Deevi, S.V.V., Mackay, A., Muthas, D., et al. (2021). Rare variant contribution to human disease in 281,104 UK Biobank exomes. Nature *597*, 527-532. 10.1038/s41586-021-03855-y.
73. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron *33*, 341-355. 10.1016/s0896-6273(02)00569-x.
74. Diedrichsen, J. (2006). A spatially unbiased atlas template of the human cerebellum. NeuroImage.
75. Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology *49*, 764-766. https://doi.org/10.1016/j.jesp.2013.03.013.
76. de Jong, L.W., Vidal, J.S., Forsberg, L.E., Zijdenbos, A.P., Haight, T., Alzheimer's Disease Neuroimaging, I., Sigurdsson, S., Gudnason, V., van Buchem, M.A., and Launer, L.J. (2017). Allometric scaling of brain regions to intra-cranial volume: An epidemiological MRI study. Hum Brain Mapp *38*, 151-164. 10.1002/hbm.23351.
77. Mankiw, C., Park, M.T.M., Reardon, P.K., Fish, A.M., Clasen, L.S., Greenstein, D., Giedd, J.N., Blumenthal, J.D., Lerch, J.P., Chakravarty, M.M., and Raznahan, A. (2017). Allometric Analysis Detects Brain Size-Independent Effects of Sex and Sex Chromosome Complement on Human Cerebellar Organization. J Neurosci *37*, 5221-5231. 10.1523/JNEUROSCI.2158-16.2017.
78. Friston, K.J. (2007). Statistical parametric mapping : the analysis of funtional brain images, 1st Edition (Elsevier/Academic Press).
79. Kuhn, H.W. (1955). The Hungarian method for the assignment problem. Naval Research Logistics Quarterly *2*, 14.
80. Santos, J.M., and Embrechts, M. (2009). On the Use of the Adjusted Rand Index as a Metric for Evaluating Supervised Classification. held in Berlin, Heidelberg, (Springer Berlin Heidelberg), pp. 175-184.
81. Beasley, T.M., Erickson, S., and Allison, D.B. (2009). Rank-based inverse normal transformations are increasingly used, but are they merited? Behav Genet *39*, 580-595. 10.1007/s10519-009-9281-0.
82. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203-209. 10.1038/s41586-018-0579-z.
83. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res *22*, 1790-1797. 10.1101/gr.137323.112.
84. Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., and Yang, J. (2016). Integration of summary data

from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet *48*, 481-487. 10.1038/ng.3538.

85. Arnatkeviciute, A., Fulcher, B.D., and Fornito, A. (2019). A practical guide to linking brain-wide gene expression and neuroimaging data. Neuroimage *189*, 353-367. 10.1016/j.neuroimage.2019.01.011.

86. Quackenbush, J. (2002). Microarray data normalization and transformation. Nat Genet *32 Suppl*, 496-501. 10.1038/ng1032.

87. Hawrylycz, M., Miller, J.A., Menon, V., Feng, D., Dolbeare, T., Guillozet-Bongaarts, A.L., Jegga, A.G., Aronow, B.J., Lee, C.K., Bernard, A., et al. (2015). Canonical genetic signatures of the adult human brain. Nat Neurosci *18*, 1832-1844. 10.1038/nn.4171.

88. Fulcher, B.D., Little, M.A., and Jones, N.S. (2013). Highly comparative time-series analysis: the empirical structure of time series and their methods. J R Soc Interface *10*, 20130048. 10.1098/rsif.2013.0048.

89. Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68-74. 10.1038/nature15393.

90. Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., et al. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature *538*, 201-206. 10.1038/nature18964.

91. Federico, A., and Monti, S. (2020). hypeR: an R package for geneset enrichment workflows. Bioinformatics *36*, 1307-1308. 10.1093/bioinformatics/btz700.

92. Rouillard, A.D., Gundersen, G.W., Fernandez, N.F., Wang, Z., Monteiro, C.D., McDermott, M.G., and Ma'ayan, A. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford) *2016*. 10.1093/database/baw100.

93. Doan, R.N., Bae, B.I., Cubelos, B., Chang, C., Hossain, A.A., Al-Saad, S., Mukaddes, N.M., Oner, O., Al-Saffar, M., Balkhy, S., et al. (2016). Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. Cell *167*, 341-354 e312. 10.1016/j.cell.2016.08.071.

94. Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E., Baldursson, G., Belliveau, R., Bybjerg-Grauholm, J., Baekvad-Hansen, M., et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet *51*, 63-75. 10.1038/s41588-018-0269-7.

95. Grove, J., Ripke, S., Als, T.D., Mattheisen, M., Walters, R.K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O.A., Anney, R., et al. (2019). Identification of common genetic risk variants for autism spectrum disorder. Nat Genet *51*, 431-444. 10.1038/s41588-019-0344-8.

96. Mullins, N., Forstner, A.J., O'Connell, K.S., Coombes, B., Coleman, J.R.I., Qiao, Z., Als, T.D., Bigdeli, T.B., Borte, S., Bryois, J., et al. (2021). Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. Nat Genet *53*, 817-829. 10.1038/s41588-021-00857-4.

97. Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A., Adams, M.J., Agerbo, E., Air, T.M., Andlauer, T.M.F., et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet *50*, 668-681. 10.1038/s41588-018-0090-3.

98. Pardinas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet *50*, 381-389. 10.1038/s41588-018-0059-2.

99. Andreassen, O.A., Djurovic, S., Thompson, W.K., Schork, A.J., Kendler, K.S., O'Donovan, M.C., Rujescu, D., Werge, T., van de Bunt, M., Morris, A.P., et al. (2013). Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. Am J Hum Genet *92*, 197-209. 10.1016/j.ajhg.2013.01.001.

100. Smeland, O.B., Frei, O., Shadrin, A., O'Connell, K., Fan, C.C., Bahrami, S., Holland, D., Djurovic, S., Thompson, W.K., Dale, A.M., and Andreassen, O.A. (2020). Discovery of shared genomic

loci using the conditional false discovery rate approach. Hum Genet *139*, 85-94. 10.1007/s00439-019-02060-2.

Acknowledgements

This work was performed on the Tjeneste for Sensitive Data (TSD) facilities, owned by the University of Oslo, operated and developed by the TSD service group at the University of Oslo, IT-Department (USIT) and on resources provided by UNINETT Sigma2—the National Infrastructure for High Performance Computing and Data Storage in Norway. The research has been conducted using the UK Biobank Resource (access code 27412) and using summary statistics for various brain disorders. We would like to thank the research participants and employees of UK Biobank and the consortia contributing summary statistics for making this work possible.

Funding: The authors were funded by the South-Eastern Norway Regional Health Authority (TM: 2021-040; OAA: 2013-123, 2017-112, 2019-108. LTW: 2014-097, 2015-073, 2016-083), the Research Council of Norway (TK: 276082, 323961. OAA: 213837, 223273, 248778, 273291, 262656, 229129, 283798, 311993. LTW: 204966, 249795, 273345), Norwegian Health Association (SB: 22731), Stiftelsen Kristian Gerhard Jebsen, and the European Research Council (LTW: ERCStG 802998). The funding bodies had no role in the analysis or interpretation of the data; the preparation, review or approval of the manuscript; nor in the decision to submit the manuscript for publication.

Author contributions: TM originally conceived of the project. TM, DvdM and SB performed the analyses. TM, LTW and OAA wrote the initial draft of the manuscript. All authors discussed the results and contributed to the final manuscript.

Competing interests: O.A.A. has received speaker's honorarium from Lundbeck and is a consultant to HealthLytix. A.M.D. is the founder of and holds equity in CorTechs Labs Inc. and serves on its Scientific Advisory Board. A.M.D. is also a member of the Scientific Advisory Board of Human Longevity Inc. and receives funding through research agreements with General Electric Healthcare and Medtronic Inc. The terms of these arrangements have been reviewed and approved by UCSD in accordance with its conflict-of-interest policies. A.M.D. is an inventor on a patent related to this work, filed by CorTechs Labs Inc. (9 US-7324842B2, filed 22 January 2002, published 29-01-2008). The other authors declare that they have no competing interests.

Data availability

In this study we used brain imaging and genetics data from the UK Biobank [https://www.ukbiobank.ac.uk/], and GWAS summary statistics obtained from the Psychiatric Genomics Consortium [https://www.med.unc.edu/pgc/shared-methods/] and GWAS catalog (https://www.ebi.ac.uk/gwas/). The summary statistics for cerebellar morphology derived in this study are available both in the GWAS Catalog [https://www.ebi.ac.uk/gwas/] and in our github repository [https://github.com/norment/open-science] (will make this public upon acceptance). FUMA results are available online (will make this public upon acceptance). The Human Genome Dating Dataset (HGD) is available at https://human.genome.dating.

Code availability

All code and software needed to generate the results is available as part of public resources, specifically FreeSurfer (https://surfer.nmr.mgh.harvard.edu), SUIT (https://github.com/jdiedrichsen/suit), OPNMF (https://github.com/asotiras/brainparts), MOSTest (https://github.com/precimed/mostest), FUMA (https://fuma.ctglab.nl/), MAGMA (https://ctg.cncr.nl/software/magma), abagen (https://github.com/rmarkello/abagen), conjunctional FDR (https://github.com/precimed/pleiofdr/) and LD score regression (https://github.com/bulik/ldsc).

Supplementary Materials:

1. Supplementary Figures 1-15.
2. Supplementary Information – Data Tables 1-37.