Title Page

Uncovering the Genetic Profiles Underlying the Intrinsic Organization of the Human Cerebellum

Running Title: **Genetic Basis of Cerebellar Functional Organization**

Yaping Wang[1,2,3,4,15], Lin Chai[2,3,4,15], Congying Chu[2,3,4*], Deying Li[2,3,4], Chaohong Gao[1,2,3,4], Xia Wu[2,3,4], Zhengyi Yang[2,3,4], Yu Zhang[5], Junhai Xu[6], Jens Randel Nyengaard[1,7,8], Simon B. Eickhoff[9,10], Bing Liu[11], Kristoffer Hougaard Madsen[1,12,13], Tianzi Jiang[1,2,3,4,14], Lingzhong Fan[1,2,3,4,14*]

[1]Sino-Danish Center, [2]University of Chinese Academy of Sciences, 100190, Beijing, China

[3]Brainnetome Center, [4]National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 100190, Beijing, China

[5]Research Center for Healthcare Data Science, Zhejiang Lab, 311100, Hangzhou, China

[6]School of Computer Science and Technology, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, 300350, Tianjin, China

[7]Core Centre for Molecular Morphology, Section for Stereology and Microscopy, Department of Clinical Medicine, Aarhus University, 8000, Aarhus, Denmark

[8]Department of Pathology, Aarhus University Hospital, 8200, Aarhus, Denmark

[9]Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, 52425, Jülich, Germany

[10]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, 40225, Düsseldorf, Germany

[11]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, 100875, Beijing. China

[12]Informatics and Mathematical Modelling, Technical University of Denmark, 2800, Kongens Lyngby, Denmark

[13]Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital-Amager and Hvidovre, 2650, Hvidovre, Denmark

[14]CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, 100190, Beijing, China

[15]These authors contributed equally: Yaping Wang, Lin Chai.

* To whom correspondence should be addressed:

Email: lingzhong.fan@ia.ac.cn or chucongying@gmail.com

Abstract

The functional diversity of the human cerebellum is largely believed to be derived more from its extensive connections rather than being limited to its mostly invariant architecture. However, whether and how the determination of cerebellar connections in its intrinsic organization interact with microscale gene expression is still unknown. Here we decode the genetic profiles of the cerebellar functional organization by investigating the genetic substrates simultaneously linking cerebellar functional heterogeneity and its drivers, i.e., the connections. We not only identified 443 network-specific genes but also discovered that their co-expression pattern correlated strongly with intra-cerebellar functional connectivity (FC). Ninety of these genes were also linked to the FC of cortico-cerebellar cognitive-limbic networks. To further discover the biological functions of these genes, we performed a "virtual gene knock-out" by observing the change in the coupling between gene co-expression and FC and divided the genes into two subsets, i.e., a positive gene contribution indicator (GCI^{+}) involved in cerebellar neurodevelopment and a negative gene set (GCI^{-}) related to neurotransmission. A more interesting finding is that GCI^{-} is significantly linked with the cerebellar connectivity-behavior association and many recognized brain diseases that are closely linked with the cerebellar functional abnormalities. Our results could collectively help to rethink the genetic substrates underlying the cerebellar functional organization and offer possible micro-macro interacted mechanistic interpretations of the cerebellum-involved high order functions and dysfunctions in neuropsychiatric disorders.

Introduction

Converging evidence from animal and human studies is advancing our understanding of the human cerebellum, which is engaged in motor, complex cognitive, and emotional behaviors [1-4]. While such functional diversity of the cerebellum was believed to derive more from its extensive afferent and efferent connections to extra-cerebellar structures, rather than being limited to a regular lattice-like anatomic feature of cerebellar-cortical cytoarchitecture [1, 5-8]. Considering the widely accepted understanding that the macroscale functional organization of the human nervous system is ultimately regulated by the underlying microscale gene expression [9-12], it is thus intriguing to unravel the genetic profiles underlying the cerebellar functional organization. As yet, what remains unclear is whether and how the hypothesized determination of cerebellar connections in its intrinsic functional organization interact with microscale gene expression.

To date, the genetic mechanism supporting the functional organization of the human cerebellum is mainly unknown. Only a few studies have attempted to investigate the gene expression pattern of the human cerebellum, but they provided inconsistent results in gene expression variability. For instance, Hawrylycz et al. [13] and Negi and Guda [14] both found that gene expression is highly homogeneous across the anatomical regions of the healthy adult cerebellum. In contrast, Aldinger et al. [15] and Wang and Zoghbi [12] found that cerebellar development and function are governed by the precise regulation of molecular and cellular programs and that the gene expression pattern is heterogeneous across spatial and temporal scales. In addition, differences in gene expression patterns between the cerebellar gyri and sulci [16], and considerable cerebellar regional specializations containing specific cell types, as revealed by high-throughput single-

nucleus RNA-seq [17] have been found in the mouse cerebellum. Because relevant studies that showed homogeneity [13, 14] explored the overall cerebellar genetic expression pattern across its gross macro-anatomical boundaries (e.g., cerebellar lobules) and might have failed to fully reflect the functional organization of the human cerebellum [18, 19]. This inconsistency in the genetic variability of the cerebellum needs to be further explored.

In the past decade, functional topological maps describing the organization of the human cerebellum using task [20] and task-free functional magnetic resonance imaging (fMRI) [21, 22], specifically, separate cerebellar functional networks [21, 22] and intra-cerebellar functional gradients [23], have been proposed. In particular, Buckner et al. [21] employed resting-state functional connectivity (FC) of the cerebello-cortical circuit as a tool to map the intrinsic functional architecture of the human cerebellum and proposed a possible functional parcellation into 7 networks and 17 networks. It is thus possible to decode the genetic profiles of the cerebellar functional organization by investigating the molecular genetic substrates linking cerebellar functional heterogeneity and its drivers, i.e., the connections. One promising approach is imaging-transcriptomics analysis [24-26], which allows the brain-wide spatial analysis of microscopic transcriptome data combined with macroscopic neuroimaging phenotypes [9]. Moreover, it also offers the opportunity to link the transcriptome data with the behavior variations via the neuroimaging [27]. These cross-scale analyses could provide a better understanding of the interaction between microscale gene expression and macroscale functional network, and ultimately involved in the individual behaviors, and importantly, their putative multiscale interactions in cerebellar related diseases [28].

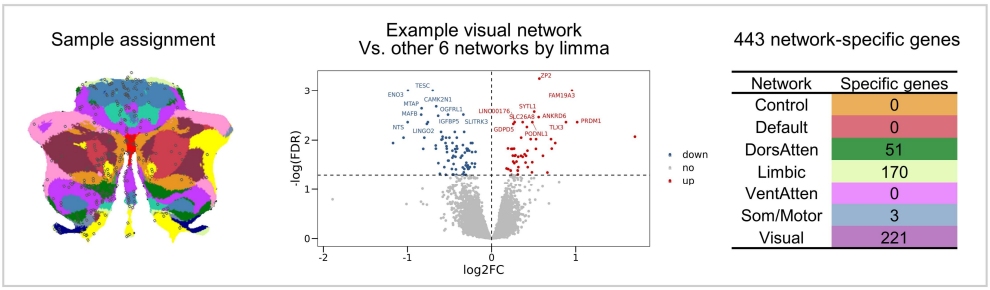
Thus, our goal was to investigate the neurobiological genetic substrates underlying the

101 functional organization of the human cerebellum. More specifically, we will seek to
102 address the following three progressive questions (Fig. S1). Are there differentially
103 expressed genes in the diverse intra-cerebellar functional networks with largely invariant
104 cytoarchitecture? How do gene expression and cerebellar connectivity relate to each other,
105 and is there any link between this association and human behavior, as well as brain disease?
106 In the current study, we first examined network-specific genes by differential gene
107 expression analysis across diverse cerebellar functional networks characterized by largely
108 invariant cytoarchitecture. This analysis allowed us to further investigate the genetic
109 explanations to the cerebellar inconsistency between functional heterogeneity and
110 cytoarchitecture near-homogeneity. Then to explore the relationship between these
111 network-specific genes with the cerebellar connection, we constructed the gene co-
112 expression matrix using the network-specific genes and found it is highly correlated with
113 both intra-cerebellar and cerebello-cortical cognitive-limbic FC. Furthermore, we tested
114 the contribution of each network-specific gene to this correlation by virtual gene knock-
115 out (KO) and divided the network-specific genes into two subsets, i.e., a positive gene
116 contribution indicator (GCI^+) and a negative gene set (GCI^-). We found that the GCI^+
117 appears to be mainly involved in cerebellar neurodevelopment. Whereas GCI^- seems to be
118 related to neurotransmission, emotion-cognitional behaviors and is significantly enriched
119 in various neuropsychiatric disorders that are closely linked with cerebellar functional
120 abnormalities. Together, the current exploration provides a starting point for associating
121 the genetic and behavior markers of the functional network to cerebellar involvements in
122 higher order non-motor functions and dysfunctions in various neuropsychiatric disorders
123 [1, 29, 30].

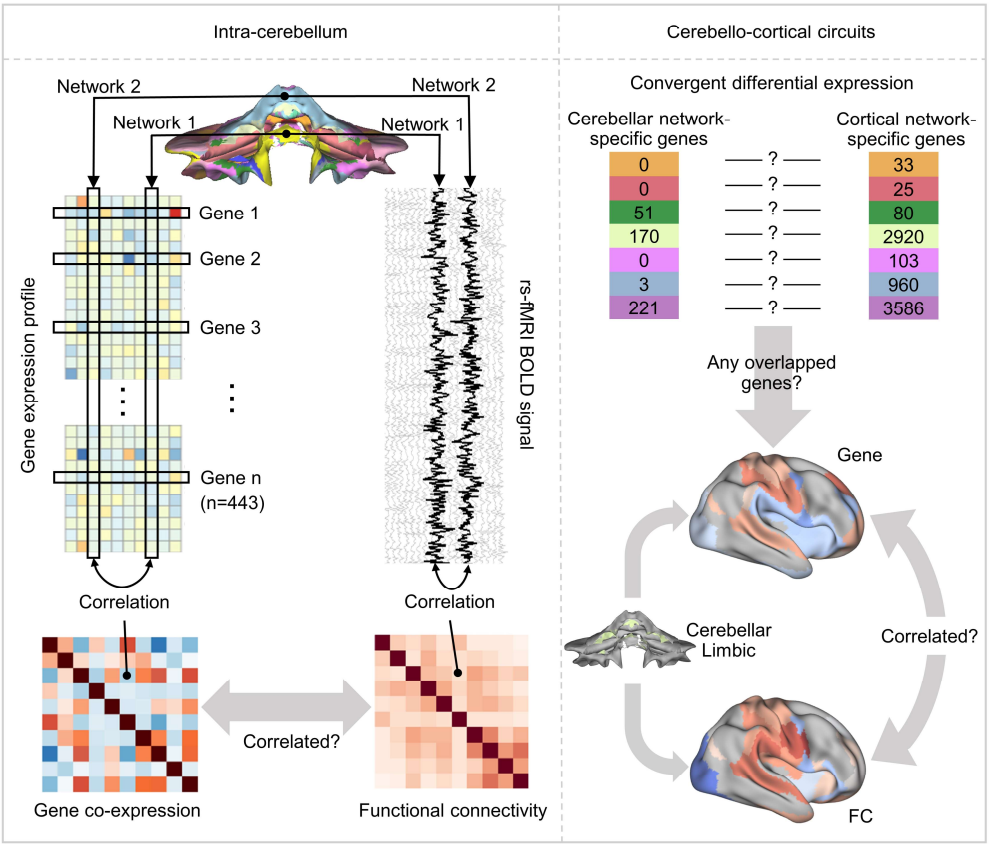
Materials and Methods

The schematic of the experimental design is depicted in Fig. 1 and includes three steps. Step 1 (Fig. 1A): to investigate whether network-specific genes occurred across cerebellar functional networks, we performed the differential gene expression analysis based on the combination of the Allen Human Brain Atlas (AHBA) transcriptome data [9] with a cerebellar functional parcellation atlas [21]. Step 2 (Fig. 1B): then, the co-expression matrix was constructed using the network-specific genes from step 1 and compared with FC to explore their overall correlation (referred to as the Gene-FC correlation for simplicity). Meanwhile, the relationship between gene and cerebello-cortical FC also has been explored. Step 3 (Fig. 1C): the “virtual gene knock-out” was leveraged to examine the direction of each gene's contribution to this Gene-FC correlation from step 2 and used to separate the network-specific genes into two subsets. Furthermore, we applied a series of functional annotation tools to explore the role of these genes, including gene enrichment analysis, Behavior-FC-Gene mapping analysis, disease enrichment analysis, and integrative temporal specificity analysis.

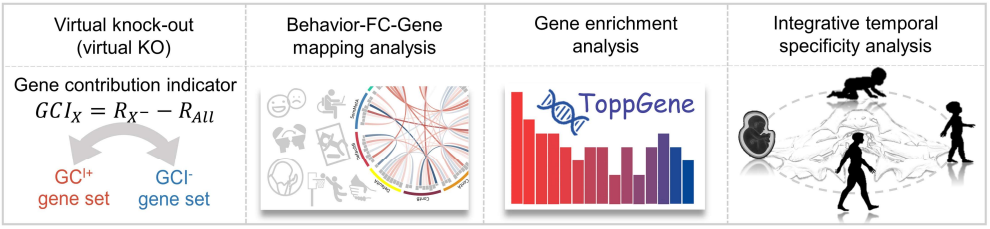
A. Step 1: Differential gene expression analysis



B. Step 2: Correlation between gene co-expression and FC



C. Step3: Functional annotation



139

140 **Fig. 1. Analysis pipeline. A Step 1**: differential gene expression analysis. We assigned the
141 AHBA cerebellar samples into seven cerebellar functional networks (left) [21] and
142 averaged each gene's expression within the same network individually. Then we compared
143 the gene expression in each network with all the other networks by limma [31] (middle)
144 with a fold change > 1 and $p < 0.05$ (FDR corrected) as an indicator (Red indicates that the
145 genes were found significantly positively expressed in the visual network.). Thus, we
146 obtained the network-specific genes for the seven networks (right). **B Step 2:** correlations
147 between the gene co-expression and the FC included intra-cerebellar and cerebello-cortical
148 circuit. Intra-cerebellum: for each pair of networks, we calculated the gene expression
149 similarity between them using the cerebellar network-specific genes and then constructed
150 the gene co-expression matrix. The FC matrix was constructed by correlating the BOLD
151 signal for all pairs. Then the relationship between the genetic correlation and functional
152 correlation was evaluated. Cerebello-cortical circuit: we first defined the cortical network-
153 specific genes as we did for the cerebellum and tested whether any convergently
154 differentially expressed genes occurred. Then we used the overlapping genes to obtain the
155 cortical genetic correlation for each cerebellar network and evaluated the relationship
156 between the cortical genetic and functional correlation for each cerebellar network. **C Step**
157 **3:** functional annotation includes virtual gene knock-out (KO), Behavior-FC-Gene
158 mapping, gene enrichment analysis, and integrative temporal specificity analysis.

159 **AHBA preprocessing**

160 The AHBA [9] is a publicly available transcriptome dataset (http://www.brain-map.org),
161 which provides normalized microarray gene expression data from six adult donors (ages
162 24, 31, 34, 49, 55, and 57 years old; $n = 4$ left-hemisphere only, $n = 2$ both left and right

163 hemispheres). Table S1 shows the demographic information. Tissue collection was
164 approved by Institutional Review Boards of the Maryland Department of Health and
165 Hygiene, University of Maryland Baltimore and University of California Irvine, and the
166 informed consents were obtained from decedent next-of-kin [9]. Although AHBA provides
167 gene expression from only six adult donors, it still has many unprecedented advantages.
168 While some existing human gene expression atlases just cover multiple brain regions, only
169 the AHBA delivers high-resolution coverage of nearly the entire brain [32]. It includes the
170 expression of more than 20,000 genes taken from 3,702 spatially distinct tissue samples [9]
171 ranging from the cerebral cortex to the cerebellum.

172 Referred to Anderson et al. [26], the current preprocessing pipeline included data
173 filtering, probe selection, sample selection, and assignment. We first filtered the probes
174 with the AHBA binary indicator to mitigate the background noise and excluded probes
175 without an Entrez ID. Then for the genes that corresponded to two or more probes, we
176 chose the probe with the maximum summed adjacency to represent the corresponding gene
177 expression; otherwise, we retained the probe with the highest mean expression, using the
178 CollapseRows function [33] in R. The first two steps generated 20,738 unique mRNA
179 probes, which provided expression data for 20,738 genes. As suggested by Arnatkeviciute
180 et al. [32] and given the known transcriptional differences [13] between the cortical and
181 sub-cortical regions and the cerebellum, we separated the cortical and cerebellar samples a
182 priori based on the slab type and structure name provided by AHBA and processed them
183 separately later. In the end, 337 samples were retained for the cerebellar cortex and 1,701
184 samples for the cortical cortex.

185 Finally, we respectively assigned these 337 cerebellum samples and 1,701 cortical
186 samples into the cerebellar functional networks atlas [21] and cortical functional networks
187 atlas [34], both of which have 7- and 17-network parcellation strategies. For each cerebellar
188 sample, we first generated a single $1 \times 1 \times 1$ mm^3 region of interest (ROI) at the MNI
189 coordinate for each sample using the AFNI 3dmaskdump -nbox function [35]. The network
190 label from either 7- or 17-network parcellation was assigned if the ROI fell within a
191 cerebellar network of the Buckner atlas. Considering the uneven and discrete sampling of
192 the AHBA data [9], if the $1 \times 1 \times 1$ mm^3 ROI did not overlap with any network, the
193 associated ROI was expanded to $3 \times 3 \times 3$ mm^3. And if the $3 \times 3 \times 3$ mm^3 ROI overlapped
194 with the functional atlas, the network that had the maximum number of shared voxels with
195 the ROI was assigned. Otherwise, the steps above were repeated for a $5 \times 5 \times 5$ mm^3 ROI.
196 The cerebellar samples were excluded if the $5 \times 5 \times 5$ mm^3 ROI did not overlap with any
197 cerebellar networks. Tables S2 and S3 show the distributions of the cerebellar sample
198 assignment for the 7-network and 17-network atlases. The assignment of the AHBA
199 cortical samples into the cortical functional network atlas was consistent with the method
200 used for the cerebellum, and the cortical sample distributions are shown in Tables S4 and
201 S5.

202 **Differential gene expression analysis across functional networks**

203 The gene expressions of the cerebellar samples within the same network were averaged for
204 each gene across the samples individually, resulting in 20,738 genes $\times$ 7 or 17 networks
205 matrices for each donor. Here we used as many of the most original overall genes as
206 possible, because some existing methods [13, 32] for screening genes do not consider the
207 specificity of gene expression within the human cerebellum. Then we calculated the

208 differentially expressed genes across the 7 networks using the R limma package [31] by
209 comparing the gene expression in one network (e.g., control) with the remaining 6
210 networks (e.g., default, limbic, visual, etc.). Since the gene expression vectors of the same
211 network derived from different donors can be treated as biological replicates [26], the
212 limma's duplicatedCorrelation tool [36] was leveraged to evaluate the independence
213 between replicates. Specifically, the correlations between replicates were estimated using
214 the duplicateCorrelation function, which individually fits a mixed linear model by
215 restricted maximum likelihood for each gene and returns a consensus correlation [36]. This
216 consensus correlation between replicates would then be incorporated into the limma's
217 linear model to preserve more information than simply averaging the biological replicates
218 [36]. The traditional minimum fold change threshold was unsuitable for determining
219 biologically meaningful but subtly different expressions [37], in particular for the
220 cerebellum, due to the highly homogenous cytoarchitecture. Instead, we applied the
221 Benjamini-Hochberg (BH) method to control the false discovery rate (FDR), and the
222 statistical threshold $p \leq 0.05$ (FDR corrected) combined with a fold change > 1 was used
223 as the key indicator for differentially expressed genes. For simplicity, the differentially
224 expressed genes across cerebellar networks are referred to as cerebellar network-specific
225 genes throughout this paper. The cortical network-specific genes were identified in the
226 same way. The only difference was that the gene expression of the cortical samples was
227 first averaged within each parcel (51 and 114 parcels, which corresponded to the 7- and
228 17-networks, respectively) [34] and then averaged within each network.

229 To test the specificity of these network-specific genes derived based on task-free 7
230 functional networks, we also applied the differential gene expression analysis with the

same procedure in other different cerebellar atlases. There are the different resolution of the same task-free functional atlas [21], i.e., 17-network parcellation, independent task-based multi-domain task battery (MDTB) [20] functional atlas and different resolutions, i.e., 11-, 28-lobular parcellations, of the anatomical atlas [38]. Then we estimated the overlapping genes between the network-specific genes derived based on the task-free 7 functional networks with other atlases. Moreover, we applied the differential gene expression analysis in only 4 left-hemisphere donors with the same procedure to test the sensibility of these network-specific genes.

Cerebellar resting-state functional connectivity (FC)

The minimally preprocessed [39, 40] Human Connectome Project (HCP) S1200 release dataset [41], which has 1,018 subjects (aged from 22 to 37 years old) with both structural MRI and resting-state functional MRI (rs-fMRI, HCP S1200 manual), was used. The preprocessing pipeline includes artifact correction (correction of gradient nonlinearity distortion, realignment for head motion, registration of fMRI data using structural data, reduction of geometric distortions due to B0 field inhomogeneity, etc.) as well as denoising by ICA-FIX [42, 43]. Time courses were extracted from these CIFTI grayordinate-format preprocessed rs-fMRI images, and the global signal was regressed as well. The resting-state blood oxygen level-dependent time series were averaged within each cortical parcel of the 7- or 17-network cortical atlases and within each cerebellar network of the 7- or 17-network cerebellar atlases [21], separately. The FC within the cerebellum was computed using Pearson's correlation for the averaged time courses for each interest ROI. Because four runs were performed for each subject, the correlation values were separately calculated for each run, Fisher's z-transformed, and averaged across the runs, resulting in a 17 $\times$ 17

254 cerebellar networks matrix. The same process was used to calculate the correlations
255 between each functional cerebellar network and each cortical parcel, resulting in a 114
256 cortical parcels × 7 cerebellar networks functional correlation matrix, which represents
257 the FC across the cerebello-cortical circuit. Regardless of whether the FC was within the
258 cerebellum or across the cerebello-cortical circuit, both categories of FC were defined
259 using the more fine-grained 17-network parcellation to increase the spatial resolution. The
260 only exception was that the cerebellar 7-network was applied while calculating the FC
261 across the cerebello-cortical circuit to compare each cerebellar network directly.

262 **Correlation between gene co-expression with intra-cerebellar FC**

263 To fully capture the genetic correlation with the FC within the cerebellum, we leveraged
264 the genetic samples of the two bi-hemisphere donors when constructing the gene co-
265 expression matrix since the FC of the cerebellum is bilateral. Therefore, the gene co-
266 expression was analyzed for the two bi-hemisphere donors using the network-specific
267 genes derived from all six donors across 7 networks, using a finer 17-network parcellation
268 to increase the spatial resolution. Ten networks that contained samples from both bi-
269 hemisphere donors were retained (Table S3). For each bi-hemispheric donor, the log2 gene
270 expression of the cerebellar samples was mean-normalized and then averaged within each
271 network. The cerebellar 10 × 10 networks correlation matrix was calculated using
272 Spearman's correlations individually (since the non-normality of gene expression data),
273 then Fisher's z-transformed, and finally averaged to construct the final 10 networks gene
274 co-expression matrix. The correlation significance level of the gene co-expression was
275 evaluated using the overlap between the correlation significance matrix for these two
276 individuals after being adjusted by Bonferroni correction. Meanwhile, we transformed the

277 17 × 17 networks FC matrix into a 10 × 10 networks size to be consistent with the gene
278 co-expression matrix. Finally, the relationship between the 10 × 10 networks gene co-
279 expression and the 10 × 10 networks FC matrix was computed using Pearson's
280 correlation. The correlation between the gene co-expression and FC is referred to as the
281 Gene-FC correlation throughout the present paper for simplicity.

282 To test the significance of the Gene-FC relationships while taking into account the
283 spatial autocorrelation (SA), we leveraged Brain Surrogate Maps with Autocorrelated
284 Spatial Heterogeneity (BrainSMASH) [44] to generate 10,000 surrogate maps which
285 persevered the genetic SA of the 443 genes × 106 samples matrix. The 106 samples were
286 equivalent to the amount of samples for the 10 networks that were used to construct the
287 gene co-expression matrix since they contained samples from both bi-hemisphere donors.
288 The parameter "knn" was set to 100 to keep it roughly consistent with the number of
289 samples. These 10,000 surrogate maps were used to construct the gene co-expression
290 matrix and thus generate an empirical null distribution of the Gene-FC correlation while
291 taking SA into account. The *p* value (p_{SA}) was defined as the proportion of correlation
292 values produced by the surrogate maps that exceeded the correlation coefficient for the real
293 data.

294 In addition, to evaluate the robustness of the verified Gene-FC correlation within the
295 cerebellum, we also recalculated it using several different parcellations, i.e., task-free 7-
296 network parcellation, and independent task-based MDTB functional parcellation [20]. The
297 criteria for each step were consistent with our primary method. Moreover, we employed a
298 control test to learn whether the Gene-FC correlation could be obtained using only the
299 network-specific genes. That is, no Gene-FC correlation while using other genes. We

300 randomly selected 443 genes from the full gene set without the network-specific genes and
301 referred to them as non-network-specific genes. Then we calculated the Gene-FC
302 correlation using the non-network-specific genes and ran this step randomly 10,000 times.
303 In addition to these thresholdless non-network-specific genes, we applied a set of
304 thresholds to the averaged original log2 gene expression data to confirm that these non-
305 network-specific genes were expressed in the cerebellum. Doing this can also help us test
306 whether the gene co-expression pattern constructed using these threshold non-network-
307 specific genes was correlated with FC.

308 Lastly, we also verified the robustness of this Gene-FC correlation using the 218
309 unrelated participants from the HCP S1200 release [41]. Besides, to constitute an
310 independent neuroimaging validation dataset, we leveraged 296 participants with four runs
311 (aged 38−58 years old) from preprocessed HCP-Aging Lifespan 2.0 Release [45]. This
312 dataset, together with the 1,018 participants from the HCP S1200 release of our primary
313 strategy, neatly covered the age ranges of AHBA (extends from 24 to 57 years old). The
314 analysis strategy for these two additional datasets is similar to our primary approach.

315 **Correlation between gene co-expression and FC across the cerebello-cortical circuit**

316 To thoroughly investigate the cerebellar functional organization, we also explored the
317 relationship between the cerebello-cortical FC and the genetic correlation based on the
318 strategy used in Anderson et al. [26]. First, we defined the network-specific genes in the
319 cortex using the same procedure as we had for the cerebellum and examined the genes that
320 overlapped within the same network of the cerebellum and the cortex. Then the gene co-
321 expression matrix was constructed between 6 cerebellar networks and 59 cortical parcels
322 from the two bi-hemisphere donors, using the 90 unique genes derived from the overlap

323 between the cortical network-specific genes and the cerebellar network-specific genes.
324 Here, the cerebellar 7-network parcellation was selected to compare the different cerebellar
325 networks directly. The visual network was excluded because it only had two samples from
326 one of the 2 bi-hemisphere donors. For the cerebral cortex, 59 cortical parcels that
327 contained samples from both bi-hemisphere donors were estimated. The log2 mean-
328 normalized expression within each cerebellar network and each cortical parcel was
329 estimated individually and correlated using Spearman's ρ, Fisher's z-transformed, and
330 averaged. We transformed the 114 cortical parcels $\times$ 7 cerebellar networks FC matrix into
331 59 cortical parcels $\times$ 6 cerebellar networks size to be consistent with the gene co-
332 expression matrix. Finally, the relationship between the cortical genetic correlation and the
333 cerebello-cortical FC matrix was computed using Pearson's correlation across six
334 cerebellar networks and adjusted by the BH method to correct for multiple comparisons.

335 **Gene functional annotation**

336 *Virtual gene knock-out (KO)*

337 To extend our investigation of the overall relationship between gene co-expression and FC
338 within the cerebellum, we referred to a similar previous approach [46, 47] and termed it
339 the "Virtual Gene Knock-out (KO)" to evaluate each gene's contribution to the Gene-FC
340 correlation. In brief, we deleted each of the 443 cerebellar network-specific genes one-by-
341 one to simulate the gene KO, then constructed the gene co-expression matrix without that
342 gene, analyzed the correlation between the FC and the gene co-expression, and finally
343 calculated the difference in the correlation coefficient between before and after the
344 simulated deletion, with the result being defined as the gene contribution indicator (GCI)
345 [46]. Based on the GCI, we identified two different gene sets that had opposite effects on

346 the Gene-FC correlation: a GCI positive gene set (GCI^+) and a GCI negative gene set
347 (GCI^-). The virtual KO of GCI^+ increased the Gene-FC correlation, and, accordingly, its
348 expression decreased the Gene-FC correlation; in contrast, the virtual KO of GCI^-
349 decreased the correlation, and, accordingly, its expression increased the Gene-FC
350 correlation.

351 *Behavior-FC-Gene mapping analysis*

352 To evaluate the relationship between these genes and human behaviors, we leveraged the
353 stable brain-behavior space strategy proposed by Ji et al. [27], which could be briefly
354 divided into the following three sections. These procedures were called Behavioral-FC-
355 Gene mapping analysis to facilitate a better understanding and fluency of the paper.

356 (1). Principal component analysis (PCA) of HCP behavioral data: 59 behavioral
357 measures [48] (Table S6) covering six behavior categories that represent the general
358 domains of human behavior were used, including alertness, cognition, emotion, motor,
359 personality, and sensory. The 218 HCP unrelated participants were selected to avoid
360 violation of the exchangeability assumption of permutation tests, and subjects with missing
361 data were excluded, resulting in 211 participants. The significant principal components
362 (PCs) were calculated by a permutation test, which randomly and independently shuffled
363 subjects' order for each behavior measure to re-run PCA 10,000 times to establish the null
364 model. The PCs that explained variance exceeded chance ($p < 0.05$ across permutation test)
365 were considered significant and retained for the next step. The nomination of PCs was
366 based on the pattern of loadings on the original 59 behavior measures.

367 (2). Mass univariate Behavior-FC map: the relationship between significant PCs scores
368 and individual intra-cerebellar FC between 17 networks was quantified by a mass

369 univariate regression procedure called PALM tool [49] across 211 unrelated participants.
370 For each significant PC, the derived regression coefficients on the total 136 functional
371 connections were then Z-scored and referred to as the Behavior-FC map for this PC. The
372 significances of the regression coefficients were assessed by a permutation test, which
373 randomly shuffled the order of the subjects for each significant PC score 10,000 times.

374 (3). Behavior-FC-Gene co-expression mapping analysis: to evaluate the relationship
375 between the Behavior-FC map and the co-expression map of GCI^+ and GCI^-, we first
376 constructed these maps across the ten cerebellar networks that corresponded to the
377 networks containing genetic samples from both AHBA bi-hemisphere donors. The
378 relationship between the Behavior-FC maps of eight significant PCs and the co-expression
379 patterns of GCI^+ and GCI^- was calculated by the Pearson's correlation respectively, and the
380 significance was assessed by a permutation test which randomly shuffled the gene co-
381 expression pattern 10,000 times.

382 *GO, pathway, and disorder enrichment analysis (ToppGene portal)*

383 To characterize the biological role of GCI^+ and GCI^-, we applied the *ToppGene portal* [50]
384 to conduct a gene ontology (GO), pathway, and disorder enrichment analysis. The GO [51]
385 enrichment analysis provides ontologies to describe accumulated knowledge of genes in
386 three biological domains: biological process, cellular component, and molecular function.
387 The BH method for FDR (FDR-BH correction) ($p < 0.05$) was chosen to correct for
388 multiple comparisons.

389 *Integrative temporal specificity analysis*

390 To investigate the overall temporal expression features of these genes, we applied an online
391 cell type-specific expression analysis (CSEA) tool [52] to do the enrichment analysis of

392 the genes within the cerebellum during different lifespan windows. Here, a specificity
393 index probability (pSI = 0.05, 0.01, 0.001, and 0.0001, permutation corrected) was used to
394 define the probability of a gene being expressed in each time window relative to all other
395 time windows to represent the varying stringencies for enrichment. The significance of the
396 overlap between the interest gene set and those enriched in a specific time window was
397 evaluated by Fisher's exact test, and the BH method for FDR (FDR-BH correction) was
398 chosen to correct for multiple comparisons.

399 **Data and availability**

400 R 3.6.1 and custom scripts were used to perform statistical analysis. All R packages were
401 mentioned explicitly in the text where the package was used. The code is freely available
402 at https://github.com/FANLabCASIA/CerebellarGeneFCCorrelation. The ToppGene
403 (https://toppgene.cchmc.org) and CSEA tool (http://genetics.wustl.edu/jdlab/csea-tool-2/)
404 which used to do the functional annotation of genes were all freely accessible. All data
405 needed to evaluate the conclusions in this study are present in the article and the
406 Supplementary Materials.

Results

The cerebellar network-specific genes derived based on the functional segregation within the cerebellum

The genes that were expressed much more in one network than in all the other six networks in the cerebellum and cerebral cortex were identified based on the differential gene expression analysis and are referred to as cerebellar network-specific genes and cortical network-specific genes, respectively. We identified 443 cerebellar network-specific genes (Supplementary Sheets 1 and 3) using all samples from six donors across 7 networks. The distribution of these network-specific genes is shown in Table 1 (Supplementary Sheet 2), which shows that these were mainly expressed in the limbic (*n* = 170), dorsal attention (*n* = 51), somatomotor (*n* = 3), and visual (*n* = 221) networks.

Table 1. Counts of significant differentially expressed genes within each network compared to other networks, referred to as the network-specific genes.

Networks	Cerebellum	Cortex	Overlapping genes
Control	0	33	0
Default	0	25	0
Dorsal Attention	51	80	0
Limbic	170	2920	56
Ventral Attention	0	103	0
SomatoMotor	3	960	2
Visual	221	3586	33
Total (unique)	443	6987	90

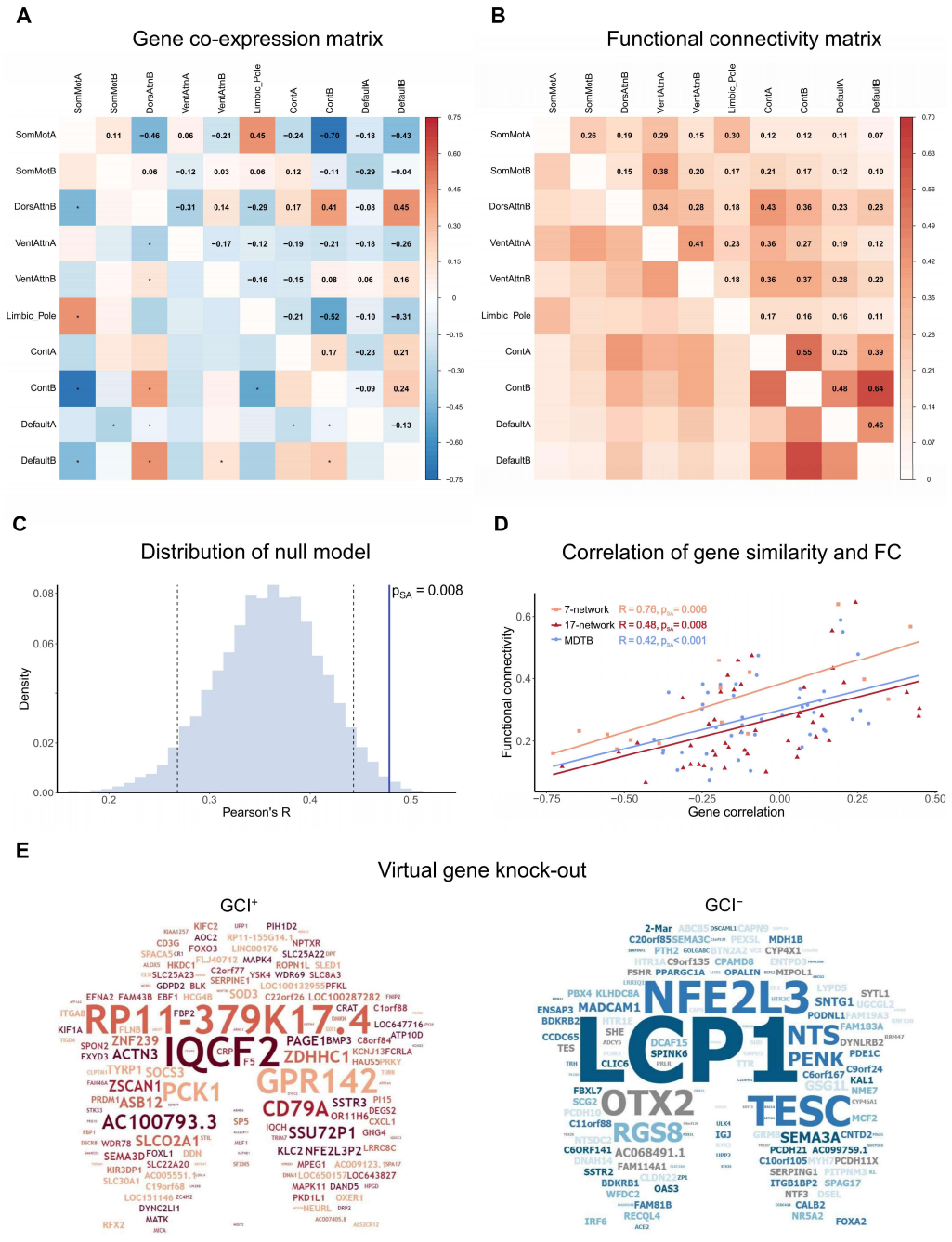
The cerebellar (*n* = 443, left column, Supplementary Sheets 2 and 3) and cortical network-specific genes (*n* = 6987, middle column, Supplementary Sheets 7 and 8) were defined across the cerebellar [21] and cortical [34] 7-network strategies, respectively. The rightmost column measures the overlap between the cerebellar and cortical network-specific genes for each network (Supplementary Sheets 9 and 10).

To test the specificity of these 443 network-specific genes in reflecting the difference of gene expression across task-free cerebellar functional atlas, we applied the differential expression analysis with the same procedure in other different cerebellar atlases (see more details in "Methods"). More overlaps between these 443 network-specific genes with the differentially expressed genes were observed in the different resolution of the same task-free functional atlas, whereas less in the independent task-based functional atlas and different resolutions of the lobular atlas (Supplementary Sheet 4). We also tested the sensitivity of network-specific genes in only four left-hemisphere donors with the same procedure and observed 230 cerebellar network-specific genes (Supplementary Sheet 5) with preferentially expressed in the limbic (n = 49, overlap = 44) and visual (n = 181, overlap = 97) networks.

Meanwhile, we obtained 6,987 cortical network-specific genes (Supplementary Sheets 6−8 and Table 1) using the same strategy and found that the cerebellar and cortical network-specific genes distribution patterns across 7-network parcellation were highly correlated (r = 0.95, p = 0.001). Moreover, we found that 90 of these 443 cerebellar network-specific genes (~20%) (Supplementary Sheets 9 and 10 and Table 1) were convergently expressed in the cerebral cortex (overlap in limbic = 56, somatomotor = 26, visual = 33). These results mean that the 56 limbic genes were differentially expressed in the limbic cortex and the limbic cerebellum and that the two somatomotor genes were differentially expressed in the somatomotor cortex and somatomotor cerebellum, as well as the 33 overlap genes in the visual network.

The co-expression of the cerebellar network-specific genes highly correlated with intra-cerebellar FC

449 Using the 443 cerebellar network-specific genes, we constructed the gene co-expression
450 matrix for the two bi-hemisphere donors and explored the relationship between gene
451 correlation and FC within the cerebellum. Noted, the 443 network-specific genes were
452 defined based on 7-nework parcellation, while the genetic co-expression was constructed
453 using 17-network parcellation to increase the spatial precision. Across all the available
454 network-network pairs, the genetic co-expression correlates with the FC within the
455 cerebellum ($r = 0.48$, $p_{SA} = 0.008$, Fig. 2), which remains significant when assessed by the
456 spatial autocorrelation (SA) preserving strategy. This correlation between gene co-
457 expression and FC was referred to as Gene-FC correlation throughout the present paper for
458 simplicity. To validate the Gene-FC correlation within the cerebellum, we also leveraged
459 the different resolution of the same task-free functional atlas and the independent task-
460 based MDTB functional parcellation [20] to re-perform the aforementioned steps (Fig. 2D).
461 The gene co-expression and FC within the cerebellum also correlated when analyzed based
462 on the 7-network parcellation (Figs. 2D and S2): $r = 0.76$, $p_{SA} = 0.006$, and the MDTB
463 functional parcellation (Figs. 2D and S3): $r = 0.42$, $p_{SA} < 0.001$. Moreover, this Gene-FC
464 correlation was also obtained using unrelated participants from HCP S1200 release
465 (Supplementary Sheet 17): $r = 0.49$, $p_{SA} = 0.010$, and another independent HCP-Aging
466 neuroimaging dataset (Supplementary Sheet 18): $r = 0.45$, $p_{SA} = 0.013$.



467

468 **Fig. 2. Network-specific gene co-expression correlates with functional connectivity**
469 **(FC) within the cerebellum. A** Genetic correlation was shown by the co-expression matrix
470 (Supplementary Sheet 11) constructed for two bi-hemisphere donors across ten cerebellar
471 networks using 443 cerebellar network-specific genes derived from all six donors. The 10
472 cerebellar networks corresponded to the networks containing samples from both bi-
473 hemisphere donors (Table S3). $*p \leq 0.05$ (Bonferroni corrected). **B** The FC matrix
474 (Supplementary Sheet 12) shows the functional correlation for the ten cerebellar networks
475 using 1,018 subjects from the HCP S1200 release [41]. All passed the significant threshold

$p \leq 0.01$ (Bonferroni corrected). **C** The probability distribution of null models preserving genetic spatial autocorrelation generated by BrainSMASH. The vertical black dashed lines correspond to the p values of 0.05 and 0.95; the blue vertical line showed the practical observed Gene-FC correlation ($r = 0.48$) and the corresponding $p_{SA} = 0.008$. **D** The overall intra-cerebellar Gene-FC correlation using different atlases: task-free 7-network (orange), 17-network (red) parcellation of the cerebellar functional atlas based on the cerebello-cortical FC, and task-based MDTB functional parcellation (blue) based on the task activation pattern. The Pearson's correlation r and p_{SA} values are shown by the corresponding colors. **E** The GCI^{+} ($n = 246$, left, red) and GCI^{-} ($n = 197$, right, blue) gene lists were displayed on the flattened shape of the cerebellum. The word size and color intensity both indicate the GCI value.

Therefore, the 443 cerebellar network-specific genes that we derived based on the functional segregation of the cerebellum also correlated with the functional integration of the cerebellum. This Gene-FC correlation was not generated by the genetic SA, so it was consistently significant whenever using unrelated participants from HCP, independent HCP-Aging dataset, different parcellation resolution, or independent cerebellar functional atlas for the calculation. Moreover, the control test exhibited no Gene-FC correlation when the gene co-expression was constructed using non-network-specific genes (Supplementary Sheet 19) regardless of whether using a threshold or not. These findings confirmed that these 443 network-specific genes play a crucial role in intra-cerebellar functional organization.

Convergently expressed genes among the cerebellar and cortical network-specific genes correlated with the FC across the cerebello-cortical cognitive-limbic networks

Since 90 of the 443 cerebellar network-specific genes were convergently expressed across the cerebello-cortical circuit, we wanted to know whether these ~20% genes correlated with the FC across the cerebello-cortical circuit. A correspondence between the genetic and functional correlations was identified for the limbic (Fig. 3A): $r = 0.36$, $p = 0.030$ (FDR corrected), and control networks (Fig. 3B): $r = -0.33$, $p = 0.034$ (FDR corrected), but was

not significant for the somatomotor: $r = -0.15$, $p = 0.394$ (FDR corrected), dorsal attention: $r = -0.19$, $p = 0.281$ (FDR corrected), ventral attention: $r = -0.04$, $p = 0.779$ (FDR corrected), or default: $r = 0.10$, $p = 0.544$ (FDR corrected) networks. The high cortical genetic similarity between the limbic system and the adjacent control network: $r = -0.90$, $p < 0.001$ (FDR corrected), somatomotor network: $r = -0.55$, $p < 0.001$ (FDR corrected), and ventral attention network: $r = -0.72$, $p < 0.001$ (FDR corrected) indicates that the gene co-expression between the cerebellar limbic network and the cortex reflects a gradual genetic gradient rather than genetic dissimilarity between the cerebellar limbic network and the other cerebellar networks. In addition, while controlling the effect of the cortical genetic similarity between the limbic and control networks, the partial correlation showed no cortical Gene-FC correlation for the control network: $r = -0.13$, $p = 0.316$. It implies that the significant cortical Gene-FC correlation for the control network could be explained by the high cortical genetic similarity between the cerebellar limbic and control networks. This observation is also consistent with the finding that convergently expressed genes were only observed in the limbic network but not in the control network (Table 1). Overall, these 443 cerebellar network-specific genes not only correlated with the intra-cerebellar FC, but ~20% of them were also linked with the cerebello-cortical cognitive-limbic networks.

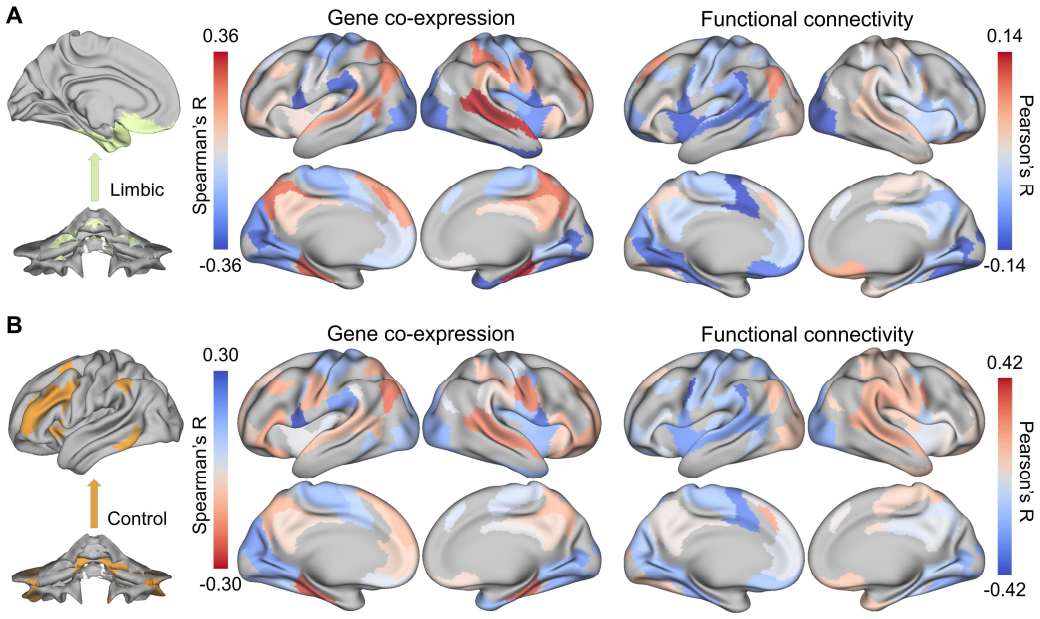


Fig. 3. Cortical genetic and functional correlation of cerebellar limbic and control networks seeds. Both were calculated for two bi-hemisphere donors across six cerebellar networks and 59 cortical parcels that contained samples from both bi-hemisphere donors. **A** Limbic: the cortical gene co-expression (Supplementary Sheet 20) was calculated using the 90 overlapping genes between the cerebellar and cortical network-specific genes by Spearman's correlation. The FC across each cerebellar network with each cortical parcel was calculated using Pearson's correlation (Supplementary Sheet 21). The cortical limbic genetic and functional correlations were correlated with each other: $r = 0.36$, $p = 0.030$ (FDR corrected). **B** Control: the cortical gene co-expression and the FC for the control network were correlated with each other: r = -0.33, p = 0.034 (FDR corrected). Notice that the color bar of gene co-expression was inverted considering the negative Gene-FC correlation for the control network.

Functional annotation revealed distinct biological properties of GCI$^+$ and GCI$^-$ separated by virtual KO

In addition to the overall correlation between gene co-expression and the functional integration of the human cerebellum, we investigated each gene's importance to the intra-cerebellar Gene-FC correlation by scoring the 443 cerebellar network-specific genes based on the GCI. Using the virtual gene KO procedure (see more details in "Methods"), we were able to classify the 443 network-specific genes that linked cerebellar functional segregation and integration into two groups: a 246 GCI positive gene set (GCI$^+$, Fig. 2E left and

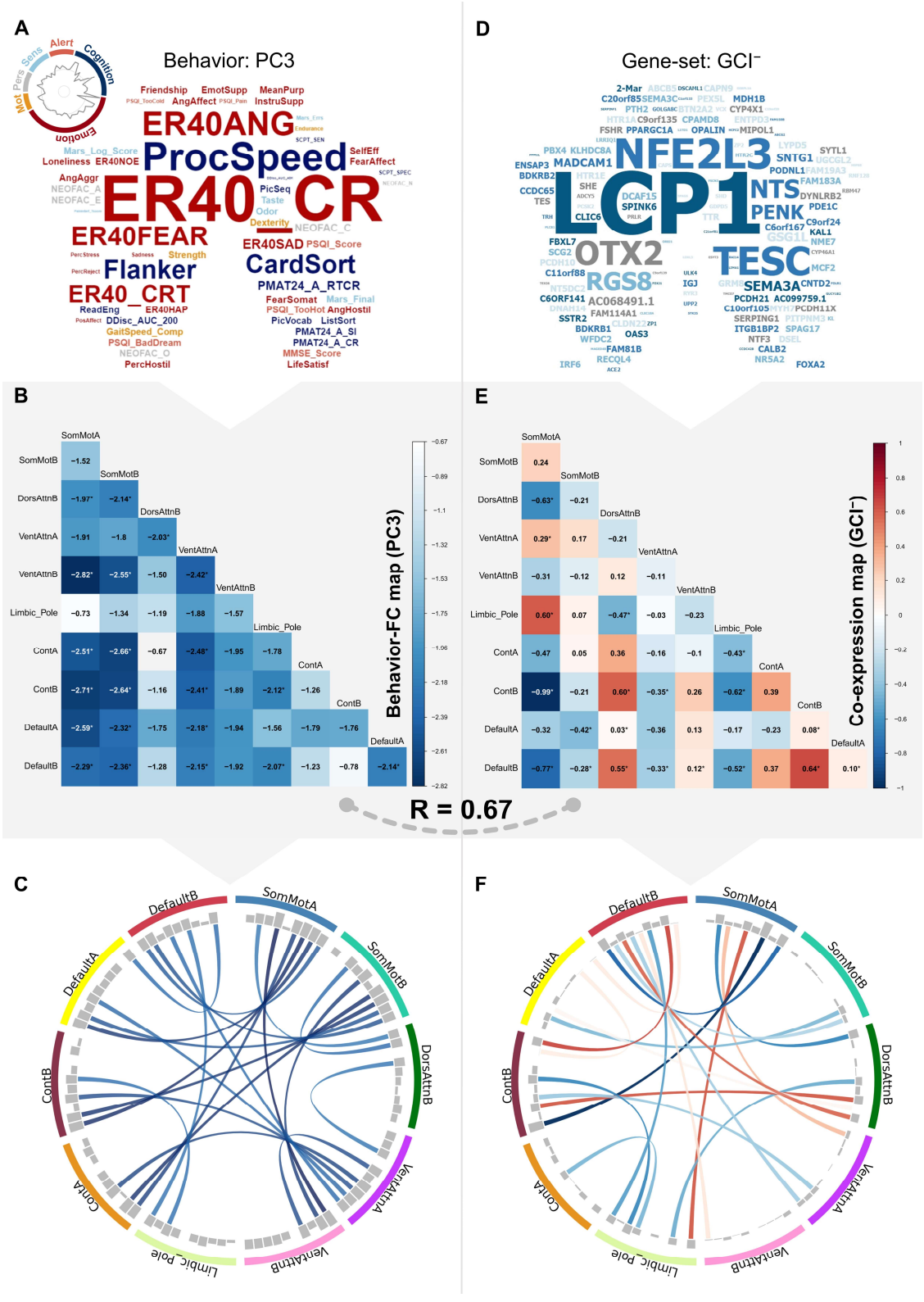
542 Supplementary Sheet 22) and a 197 GCI negative gene set (GCI^{-}, Fig. 2E right and
543 Supplementary Sheet 23). The distinction between the two sets is that the virtual KO of
544 GCI^{+} genes increased the Gene-FC correlation, whereas the virtual KO of GCI^{-} genes
545 decreased the Gene-FC correlation. Based on the winner-take-all principle, GCI^{-} genes
546 may have a critical impact on the functional organization of the cerebellum. An example
547 is that the top genes, LCP1 and TESC, enable GTPase binding and calcium binding,
548 respectively [53], which are key functions within signaling transduction and consequently
549 brain functions. Therefore, we applied a range of functional annotation tools to further
550 explore the underlying roles of the GCI^{+} and GCI^{-}.

551 *Behavior-FC-Gene mapping analysis*

552 Since these genes largely engaged in the cerebellar functional network which could predict
553 human behaviors, whether they also involved in human behaviors with the FC acting as an
554 intermediary role? We leveraged the Behavior-FC-Gene mapping analysis (see more
555 details in methods) to explore the relationship between the co-expression of GCI^{+} and GCI^{-}
556 with the human behaviors. Based on the PCA of 59 behavior measures (Table S6) across
557 211 HCP unrelated participants, 8 significant PCs were derived (Supplementary Sheet 24).
558 Then the Behavior-FC map of each significant PC calculated by mass univariate regression
559 procedure was compared with the co-expression pattern of GCI^{+} and GCI^{-} respectively
560 (Supplementary Sheet 25). Among the eight significant PCs, we observed correlations of
561 the Behavior-FC maps of PC1, PC3, and PC4 (Supplementary Sheet 25) with the GCI^{-} co-
562 expression, rather than the GCI^{+}.

563 But the PC1 and PC4 had few regression coefficients (1.47% and 0.74%, respectively)
564 in the Behavior-FC map that survived the permutation test (Supplementary Sheet 25)

565 compared with PC3 (Figs. 4B and S4). These significant Behavior-FC regression
566 coefficients mainly occurred in the connections between default, control, and somatomotor
567 networks with others (Fig. 4C), such as the connection between default B network with
568 somatomotor A, somatomotor B, ventral attention A, limbic pole, and default A networks.
569 Taking the intra-cerebellar FC as an intermediate role, PC3 had a significant correlation
570 with the GCI^{-} co-expression (Fig. 4E, F): $r = 0.67$, permutation test $p < 0.001$. In contrast,
571 no correlation between GCI^{+} co-expression and PC3 Behavior-FC map: $r = 0.27$, $p > 0.05$.
572 Based on the loading pattern of PC3 for the original 59 behavior measures, the PC3 reflects
573 the variation in the emotion-cognitional behavior measures (Fig. 4A). Specifically, PC3
574 has a large positive loading for the number of correct responses in emotion recognition,
575 executive function, and fluid ability of cognition tasks, whereas it has a large negative
576 loading for the response time in emotion recognition and fluid intelligence tasks. It suggests
577 a high PC3 score might reflect high emotion and cognition ability. This Behavior-FC-Gene
578 mapping analysis revealed that the GCI^{-} not only plays an important profiling role in the
579 functional organization of the cerebellum but also implicates emotion-cognitional behavior
580 vicariously through the intra-cerebellar FC.


A
Behavior: PC3
Alert
Cognition
Emotion
Mot
Pers
Sens
Friendship
EmotSupp
MeanPurp
AngAffect
InstruSupp
ER40ANG
ProcSpeed
SelfEff
FearAffect
Loneliness
ER40NOE
AngAggr
ER40
CR
PicSeq
Taste
Odor
Dexterity
ER40FEAR
ER40SAD
PSQI_Score
Flanker
CardSort
ER40_CRT
PMAT24_A_RTCR
ReadEng
ER40HAP
DDisc_AUC_200
GaitSpeed_Comp
PSQI_BadDream
PercHostil
FearSomat
PSQI_TooHot
AngHostil
PicVocab
ListSort
PMAT24_A_SI
PMAT24_A_CR
MMSE_Score
LifeSatisf
D
Gene-set: GCI⁻
NFE2L3
LCP1
NTS
PENK
OTX2
TESC
RGS8
SEMA3A
B
Behavior-FC map (PC3)
SomMotA
SomMotB
DorsAttnB
VentAttnA
VentAttnB
Limbic_Pole
ContA
ContB
DefaultA
DefaultB
E
Co-expression map (GCI⁻)
R = 0.67
C
F

582 **Fig. 4. The Behavior-FC-Gene mapping analysis. A** The loading pattern of principal
583 component 3 (PC3) for the original 59 behavior measures (Table S6) was obtained from
584 principal component analysis (PCA). The word sizes represent the absolute values of the
585 loading coefficients, and the different colors indicate the six categories of the behaviors
586 (Table S6). **B** Behavior-FC map for PC3, which are the z-scored regression coefficients of
587 PC3 in each intra-cerebellar functional connection via mass univariate regression
588 procedure. **C** The chord plot shows the significant regression coefficients after the
589 permutation test consisting of 10,000 permutations. The circularly arranged gray histogram
590 shows the absolute magnitudes of the regression coefficients, which correspond to the
591 values of the regression coefficients shown in **B**. The color intensity of chords is the same
592 as the color bar of **B**. **D** The GCI^{-} gene set which used to construct the GCI^{-} co-expression
593 map **E**. **F** The chord plot shows the significant correlations: $p \leq 0.05$ (Bonferroni corrected).
594 The circularly arranged grey histogram shows the absolute magnitudes of the correlation
595 coefficients, which correspond to the values of the correlation coefficients shown in **E**. The
596 color intensity of chords is the same as the color bar of **E**.

597 *GO, pathway, and disorder enrichment analysis*

598 The GO enrichment analysis of the GCI^{+} and GCI^{-} is shown in Fig. 5A. The GCI^{+} was
599 mainly enriched in microtubule-related terms, including the microtubule associated
600 complex (ID: 0005875), motile cilium (ID: 0031514), and dynein complex (ID: 0030286).
601 Compared with GCI^{+}, the GCI^{-} was not only enriched in microtubule-related terms but was
602 also significantly enriched in terms related to neurotransmitter transport, such as calcium
603 ion binding (ID: 0005509), regulation of hormone levels (ID: 0010817), response to
604 catecholamine (ID: 0071869), response to monoamine (ID: 0071867), and regulation of
605 neurotransmitter receptor activity (ID: 0099601). These findings are consistent with their
606 different pathway enrichment results (Supplementary Sheets 26 and 27) in that the GCI^{+}
607 was primarily enriched in some basic biological pathways: proximal tubule bicarbonate
608 reclamation (ID: M4361) and glycolysis/gluconeogenesis (ID: M39474), which provides
609 the energy needed during microtube-related processes. In contrast, the GCI^{-} was primarily
610 involved in signaling transduction, especially in some neurotransmission pathways, such
611 as the neuroactive ligand-receptor interaction (ID: M13380).

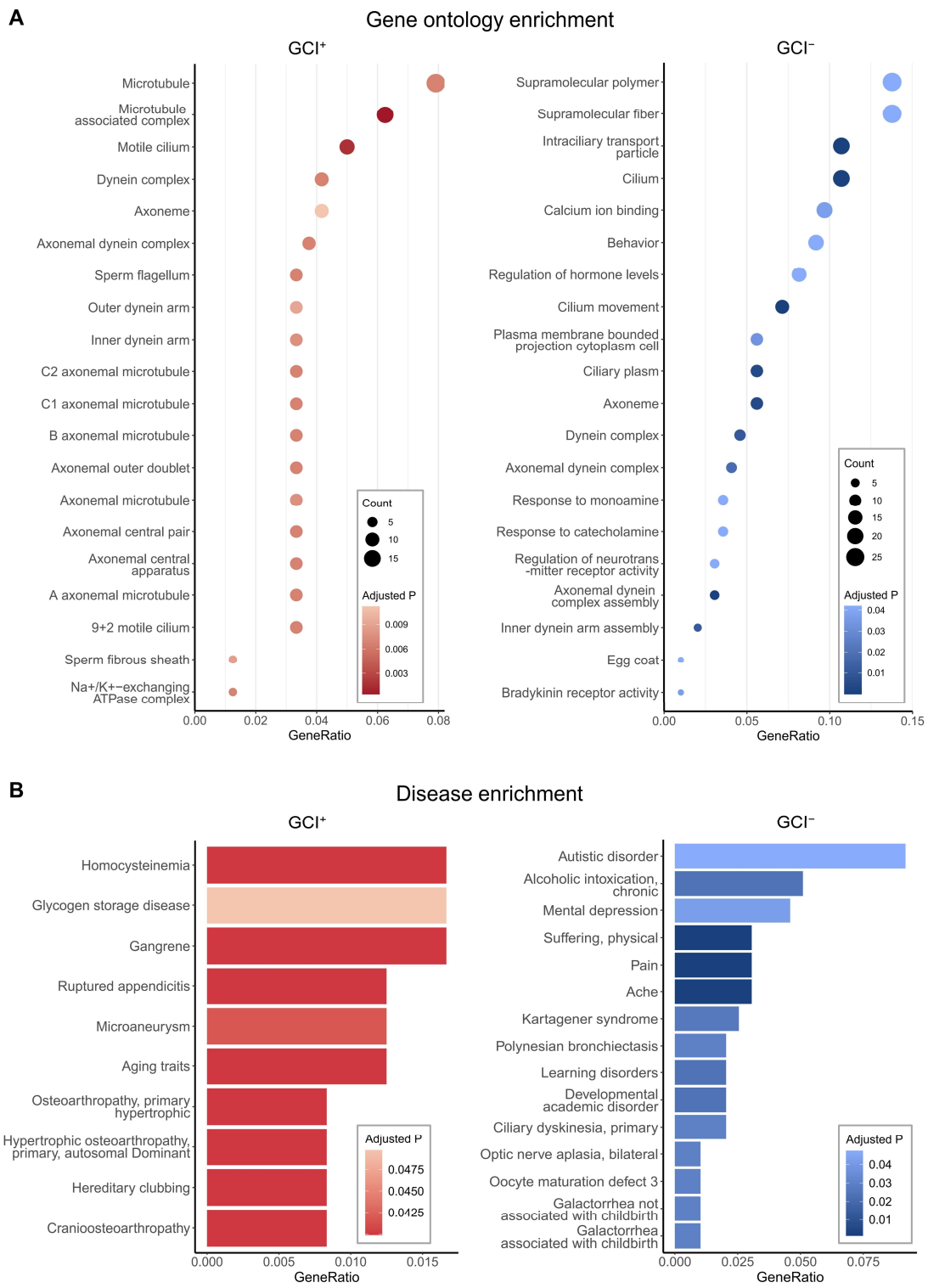


Fig. 5. The gene ontology (GO) and disease enrichment analysis for GCI$^+$ and GCI$^-$. **A** Bubble plots showing the GO enrichment top 20 terms for GCI$^+$ (left) and GCI$^-$ (right) (Complete results are shown in Supplementary Sheets 26 and 27, respectively). The

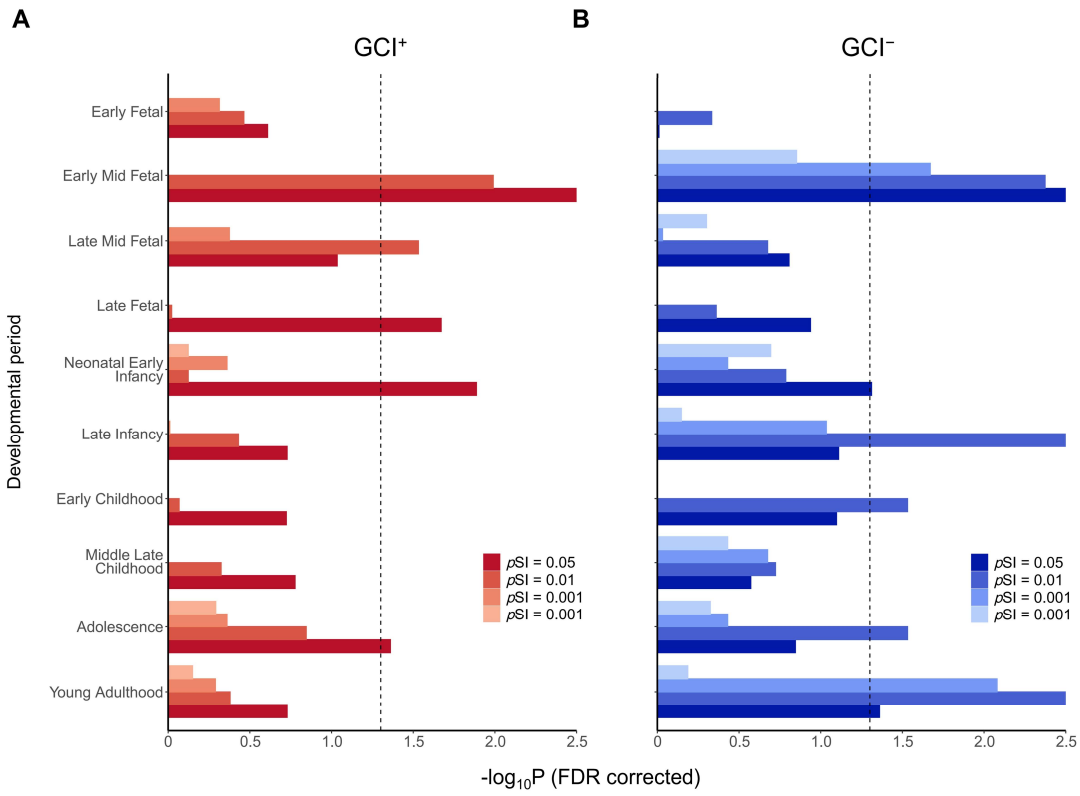
616 biological process (BP), cellular component (CC), and molecular function (MF) are
617 displayed together. The dot size (count) represents the number of genes that are within the
618 interest GCI^+ or GCI^- gene panels as well as a specific GO term (*y*-axis). The different
619 color intensities indicate the *p* value (FDR corrected). **B** Gradient bar plots showing the
620 disease enrichment for all representative results for GCI^+ and top 15 representative terms
621 for GCI^-. The different color intensities indicate the *p* value (FDR corrected).

622 Since the GCI^+ and GCI^- are involved in different biological processes, we
623 hypothesized that they also play different roles in brain disease or are related to different
624 brain diseases. Unexpectedly, we found no link between GCI^+ and any brain-related
625 illnesses (Fig. 5B, left) but observed involvement of GCI^- in various neuropsychiatric
626 disorders (Fig. 5B, right), including autistic disorder (ID: C0004325), alcoholic
627 intoxication (ID: C0001973), mental depression (ID: C0011570), pain (ID: C0030193),
628 learning disorders (ID: C0023186) and others (Supplementary Sheet 27). Many of these,
629 especially mental depression and autistic disorder, have a close relationship with the human
630 cerebellum, in which patients have shown FC abnormalities [54, 55]. The mental
631 depression- and autistic disorder-associated genes were TRH, PENK, TTR, ADCY5,
632 NRXN1, HTR1A, HTR2C, NTS, PEX5L (*n* = 9, Supplementary Sheet 27), and DLGAP2,
633 TRH, PENK, RYR3, SEMA3A, NRXN1, TESC, ABCG2, PCDH10, CNTN4, HTR1A,
634 CALB2, HTR2C, DNAAF4, FOLR1, NTS, GRM8, UPP2 (*n* = 18, Supplementary Sheet
635 27), respectively, and the overlapping genes were TRH, PENK, NRXN1, HTR1A, HTR2C,
636 NTS (*n* = 6).

637 *Integrative temporal specificity analysis*

638 In light of the distinct properties of GCI^+ and GCI^-, we wanted to know whether the roles
639 played by these two gene sets showed variable prevalence at different ages. By leveraging
640 the BrainSpan dataset [56] and applying the analysis strategy of the CSEA tool [52], we

641 found that GCI^{+} showed significant overexpression in early middle fetal, late middle fetal,
642 late fetal, neonatal early infancy, and adolescence compared with GCI^{-} (Fig. 6A). These
643 stages neatly correspond to the timeline of the protracted development of the human
644 cerebellum [57], which extends from the early embryonic period until the end of the first
645 postnatal year. This result appears to be consistent with the observation that the GCI^{+} is
646 involved in some fundamental biological processes, especially microtubule-related activity,
647 whose dynamics play a key role in cerebellar neurodevelopment [58]. In contrast,
648 compared with the GCI^{+}, the GCI^{-} was significantly expressed in early mid fetal, neonatal
649 early infancy, late infancy, early childhood, adolescence, and young adulthood (Fig. 6B).
650 These periods include the highest neurodevelopmental risk windows for autism spectrum
651 disorder (ASD) [59] and major depression disorder (MDD) [60], both of which also existed
652 in the disease enrichment analysis of GCI^{-}.



653

Fig. 6. Integrative temporal specificity analysis within the human cerebellum for GCI^+ and GCI^-. A Bar plot showing the temporal specificity results for GCI^+. The specificity index probability (pSI = 0.05, 0.01, 0.001, and 0.0001, permutation corrected, shown as different color intensities) was used to determine how likely a gene was to be expressed in a given time window relative to all other time windows [52]. The *x*-axis corresponds to the $-\log_{10}p$ (FDR corrected), and for esthetics if $-\log_{10}p$ (FDR corrected) > 2.5, $-\log_{10}p$ (FDR corrected) = 2.5; the y-axis represents the 10 development windows collected by BrainSpan [56]. The vertical dark dashed line corresponds to the *p* (FDR corrected) = 0.05. Complete results are shown in Supplementary Sheet 28. **B** Bar plot showing the temporal specificity results for GCI^-.

Discussion

The current study provided a tentative exploration of the genetic differential and co-expression linked with the functional organization of the human cerebellum and has the potential for elaborating and rethinking the neurobiological underpinnings of the cerebellar functional organization. Furthermore, we identified two gene sets involved in cerebellar neurodevelopment and neurotransmission and found indirect but interesting genetic evidence supporting a key role played by the cerebellar functional network in emotion-cognitional behaviors, as well as many neuropsychiatric disorders. These findings hint at a micro-macro interacted mechanistic possibility for the cerebellar contributions to emotion-cognitional behaviors, and consequently to related neuropsychiatric disorders.

The genetic profiles underlying cerebellar functional segregation correlate with intra-cerebellar and cerebello-cerebral connections

In this study, we found correlations between the identified cerebellar network-specific genes with the intra-cerebellar connection and cerebello-cerebral FC. These findings could provide possible empirical genetic support for the hypothesized decisive role of cerebellar connectivity in the functional heterogeneity of the cerebellum. First, while obtaining the network-specific genes, we found significant differences in the number of identified genes between the functional specificity (i.e., limbic, visual networks) and functional diversity networks (i.e., the control, default networks); specifically, more differentially expressed genes were in the former and vice versa in the latter [61]. This observation was also found in a previous cortical gene expression homogeneity analysis [9] that showed that a relatively high differential expression pattern was observed in the primary sensory cortex, area 38, and the primary visual cortex. But the findings related to the inconsistency in the

687 amount of somatomotor cerebellar (n = 3) and somatomotor cortical network-specific
688 genes (n = 960) were unclear. One possible explanation may be that the preferential links
689 between the cerebellar representations of body space and the motor, somatosensory, and
690 premotor cortices are challenging to distinguish [21]. The finding of cerebellar network-
691 specific genes is consistent with the elaborate regional difference in cerebellar-cortical
692 cytoarchitecture [2, 62] in addition to the uniform cell types and their connectivity, such as
693 regional variation in the cell size and packing density of Purkinje, granule, and Golgi cells.
694 Second, the overall distribution patterns of the cerebellar and cortical network-specific
695 genes were highly correlated, a finding in line with a similar macroscale principle identified
696 in the cerebellar and cortical functional organization [23, 63]. These correlated patterns
697 may be related to defining the cerebellar networks, which were by projecting the cerebral
698 cortical networks onto the cerebellum by computing the functional connections between
699 the two structures [21]. More interestingly, the molecular genetic substrates simultaneously
700 linking functional heterogeneity and integration could be observed across different
701 functional subdivisions, regardless of whether the parcellation was based on the task-free
702 cerebello-cortical FC [21] or the intra-cerebellar task-based activation pattern [20]. These
703 interpretations are further supported by the widely accepted notion about the human
704 cerebellum that its functional specialization is dominated more by its connection with
705 extra-cerebellar structures than within its generally homogeneous cytoarchitecture [8].
706 Although no intra-cerebellar anatomical fiber connections linking adjacent or distant
707 cerebellar regions have been observed [64, 65], it is well recognized that the intra-
708 cerebellar functional map is a consequence of the topological arrangement of its extra-
709 cerebellar anatomical connections [8]. This fact is especially important to note when

710 explaining the Behavior-FC results that although our results were obtained through intra-
711 cerebellar connections, the cerebellar-cortical connections must also be involved actually.
712 The proposed relationship between extra- and intra-cerebellar connectivity, in turn, can be
713 expected to affect the resting-state activity between cerebellar regions [23]. Furthermore,
714 for the interaction between gene and FC, they influence one another bidirectional [66]
715 through gene-environment (G-E) interplay [67] with the development as an essential
716 element [68], instead of a unidirectional determination role of the genes in the brain
717 connection. Here we observed the Gene-FC interaction from the adult perspective (AHBA
718 donors aged from 24 to 57 years old), which lacks exploration in the G-E interplay
719 compared with the early development [68].

720 Third, in addition to the intra-cerebellar Gene-FC correlation, we observed a direct
721 correlation between genes underlying the cerebellar functional specialization and
722 cerebello-cerebral FC of the limbic and control networks. The Gene-FC correlation in the
723 control network was mainly caused by the genetic similarity between these two networks;
724 this interaction between limbic-emotion and control-cognition has been confirmed
725 anatomically and behaviorally [69]. For instance, the integrated processing by the emotion
726 and cognition areas has been identified solely based on their anatomical connections [70].
727 This relationship can also be observed in that when looking at the top of a hill, a sad mood
728 induces a steeper perception of the hill than a happy one [71]. One possible reason why we
729 only obtained this correspondence in the limbic network may be the low functional
730 heterogeneity [61] and inter-individual functional variability [72] of the limbic network
731 compared with others as well as the complexity of gene expression; i.e., the Gene-FC
732 correlation is not fully portrayed by the differentially expressed genes [26]. Considering

733 the indirect connection between cerebellum and cortex, and large differences between the
734 cerebellum and cortex in terms of their gene expression patterns [13] and structure-function
735 relationships [73], as well as the individual variability of their functional networks [75],
736 the convergently expressed genes associated with the cerebello-cortical cognitive-limbic
737 networks were evident through simple analysis of co-expression, demonstrating these
738 genes are very significant and may hold clues to the molecular underpinnings of the
739 cognitive-emotion roles played by the cerebello-cortical circuit. For example, the HTR1A
740 and HTR2C, both preferentially expressed in the cerebellar and cortical limbic network,
741 are pivotal genes in serotonin transmission and play a modulation role in the limbic system
742 act as important therapeutic targets in limbic system-related disorders [74]. This
743 observation was further confirmed by the intra-cerebellar Behavior-FC-Gene mapping
744 analysis, which showed that some of the genes were strongly associated with the emotion-
745 cognitional behaviors.

746 **Cerebellar neurodevelopment feature of GCI$^+$, cerebellar neurotransmission,**
747 **emotion-cognitional behavior, and neuropsychiatric related features of GCI$^-$**

748 Interestingly, we identified two gene subsets with pronouncedly different characteristics
749 based solely on the direction in which each gene influenced the intra-cerebellar Gene-FC
750 correlation by applying a simple virtual KO approach on the 443 cerebellar network-
751 specific genes. By using a series of bioinformatic tools, we found converging evidence for
752 GCI$^+$ and GCI$^-$ involvement in cerebellar neurodevelopment and cerebellar
753 neurotransmission, respectively. It is also interesting to speculate that these 443 network-
754 specific genes that link both cerebellar functional segregation and integration have a
755 relationship with some brain disorders since prior evidence showed that the cerebellar

756 functional organization plays a key role in various neurological [30, 75] and psychiatric
757 disorders [29], most of which possess common underlying genetic risks [76]. But a tricky
758 problem emerged in that the genes we are interested in were derived from healthy
759 individuals. It could be tackled to some extent by using the virtual KO method, which can
760 simulate the different expression levels of each gene and thus coarsely corresponds to a
761 fraction of the expression level under normal health and disease situations. This is why we
762 thought that we might be able to see whether the GCI^+ and GCI^- are related to a specific
763 disease even though the genes were derived from healthy individuals.

764 The GCI^+ is involved in many microtubule-related terms and is overexpressed
765 throughout the protracted development of the cerebellum. The dynamics and flexibility of
766 microtubules were found to be essential throughout cerebellar development via leading the
767 morphological alterations of Purkinje cells [58]. In addition, some genes of the GCI^+, such
768 as GTPBP2 [77] and Lin28b [78], were found to play a key role in neurodevelopment;
769 overexpression of the Lin28b gene can induce the development of pathological lobulation
770 in the cerebellum [78]. This converging evidence prompts our speculation that the GCI^+ is
771 engaged in cerebellar neurodevelopment. Unexpectedly, the GCI^+ showed no link to brain-
772 related diseases, which appears to be consistent with its primary involvement in many
773 fundamental biological functions. However, this lack of disease linkage is inconsistent with
774 the significant overexpression of GCI^+ genes during the protracted development of the
775 cerebellum. Many researchers pointed out that this protracted development increased the
776 susceptibility of the cerebellum to many psychiatric disorders [57]. It is likely
777 complemented by the overexpression of GCI^- in the early middle fetal and neonatal early
778 infancy periods. Other possible explanations include few genetic studies of the cerebellum

779 compared with the cerebral cortex, and large gene expression differences between the
780 cerebellum and extra-cerebellar structures [13], so the related datasets may lack sufficient
781 information specific to the cerebellum. It calls for future studies to provide a complete
782 explanation by considering multiple perspectives.

783 The GCI^{-} was found to be related to emotion-cognitional behaviors, involved in many
784 neurotransmission processes, enriched in various neurological and psychiatric disorders,
785 and significantly overexpressed in late infancy, early childhood, adolescence, and young
786 adulthood compared with GCI^{+}. These results are mutually supportive. The
787 neurotransmission has long been believed to play a crucial role in emotional [79] and
788 cognitional behaviors [80], and its abnormality has been extensively associated with
789 various neurological [81] and psychiatric disorders [82, 83] with the characteristic of
790 cognitive and emotional impairments. For example, the abnormal transmission of
791 monamines and catecholamines, such as serotonin and dopamine, has been widely linked
792 with many psychiatric disorders, and these transmitters have thus become potential
793 treatment targets [84]. The period through which the GCI^{-} genes are expressed includes
794 the high-risk time windows for GCI^{-} enriched disorders, such as mental depression (aged
795 18–29) [60] and autistic disorder (from infancy to childhood) [59]. And the high expression
796 of GCI^{-} in early middle fetal life might be associated with the prenatal risk factors
797 associated with depression [85] and autism [86].

798 **A possible micro-macro interacted mechanistic explanation for the**
799 **function/dysfunction of the human cerebellum**

800 The cerebellum regulates motor, emotion, and cognition by functioning as an oscillation
801 dampener that combines multiple internal representations with external stimuli and

802 appropriate responses to maintain the behavior around a homeostatic baseline
803 automatically to optimize performance according to the cortex [87]. Interestingly, we found
804 that the GCI^{-} associated with the emotion-cognitional behaviors via the intra-cerebellar FC,
805 mainly includes the connections between control, default, and somatomotor networks
806 which have been widely confirmed to be involved in these behaviors via human clinical
807 data or functional imaging studies [4, 88-90]. This result suggests that the interaction
808 between genes and FC is also involved in various higher functions of the cerebellum,
809 providing new evidence and molecular substrates for the progressively complete notion
810 that the cerebellum modulates cognition and emotion as it does to the motor control [1].
811 Beyond that, the specifics of how the cerebellum is involved in emotional and cognitive
812 functions are explained by many compatible theories [87]. Such as the internal model [91],
813 in which the cerebellum shows increased activations for negative feedback compared to
814 the positive feedback in the reversal-learning task [92], which is maybe consistent with the
815 negative Behavior-FC map found in the present study.

816 Meanwhile, the regions that showed a significant correlation between emotion-
817 cognitional behaviors and GCI^{-} via connections, correspond exactly to the cerebellar
818 posterior lobe whose lesions produce the dysmetria of thought and emotion, i.e., the
819 cerebellar cognitive affective syndrome, which is characterized by the deficits in executive
820 function, visual-spatial processing, linguistic skills and affect regulation [93]. And no
821 significant difference in the performance on newly established objective cognitive tests
822 battery between patients with isolated cerebellar lesions versus complex cerebello-cortical
823 connection disorders, such as schizophrenia, mental depression, and autistic disorder [94].
824 It is thus reasonable and mutually validating that we found that the GCI^{-} enriched in many

825 neuropsychiatric disorders involving alternation in these two categories of behaviors,
826 including mental depression, autistic disorder, pain, alcoholic intoxication, learning
827 disorder, and others. Moreover, these disorders are all closely related to the alterations of
828 the cerebellar FC [29, 54, 95-97].

829 Therefore, the GCI^{-} provides a possible micro-macro interacted mechanistic
830 explanation for the functions and dysfunctions of the human cerebellum. The genes
831 underlying the functional organization of the human cerebellum are likely involved in
832 emotion-cognitional behaviors through their interaction with the cerebellar connection.
833 One of the possible ways the risk genes act in the pathogenesis of corresponding diseases
834 maybe through their interactions with the cerebellar FC, which results in dysregulation of
835 cerebellar FC and thus pathologically manifested as its functional connection abnormalities
836 and consequently transient or long-term impairment at the cognitive and emotional
837 behavioral levels of these diseases. For instance, the fluctuation in the correspondence of
838 the Gene-FC relationship found in the present study and widespread altered cerebellar FC
839 excavated previously in diverse neuropsychiatric disorders [29]. The GCI^{-} also provides a
840 promising genetic resource for investigating the cerebellar involvements in emotion-
841 cognitional behaviors and related brain diseases. For example, the observed overlapping
842 genes, i.e., NRXN1, are associated with mental depression and autistic disorder supported
843 previous clinical studies showing that rare and common variants in NRXN1 carried risks
844 for MDD [98], ASD, and schizophrenia [99]. And HTR1A, which has a high expression in
845 the cerebellum [14] and has been linked to the cognitive process in both animal models and
846 human [79], was found to be involved in pain, mental depression, autistic disorder,
847 alcoholic intoxication, learning disorder, and other conditions (Supplementary Sheet 27).

848 **Methodological considerations**

849 The interpretation of our findings is not without limitations. First, we validated our results
850 across independent imaging datasets that cover the age range of AHBA, and various
851 cerebellar atlases with different parcellation criteria and resolutions. But we are unable to
852 test our results in an independent gene dataset as there is no other publicly available dataset
853 with a detailed sampling of subregions of the human cerebellum. Second, the gene co-
854 expression we constructed only considered one small part of the relationship between the
855 genes and FC. Thereby it did not fully recapitulate the complexity of the brain
856 transcriptome, such as gene-gene interactions [100]. That is one possible reason we only
857 found a cerebello-cortical Gene-FC correlation for the cognitive-limbic networks. Third,
858 the simple correlation approach [25, 61] used in this study and other linear regression
859 models like partial least squares analysis [101], can only prioritize genes for further
860 investigation and cannot fully explore the causal relationship between genes and functional
861 organization. As a result, further exploration is hindered by the intricacies of genetic and
862 epigenetic regulation. It makes the discussion and explanation of the different directions of
863 this correlation challenging. For example, why the direction of influence on the Gene-FC
864 correlation could separate these 443 genes into two distinct gene sets with different
865 functions remains unclear. Hence further related exploration is necessary but very
866 challenging. Fourth, since the 443 genes were not derived based on the correlation with a
867 spatially-defined phenotype, the newly proposed strategy [102] that mitigates the bias of
868 leveraging gene enrichment approach in the spatial transcriptomic data cannot be directly
869 utilized. Nonetheless, it cannot be determined whether this type of bias affects the present
870 results, which calls for future efforts to develop methods to answer this question.

871 Briefly, more donors, the overall pattern of gene expression, gene regulation,
872 epigenomics, and improved cellular resolution are needed and imperative for developing
873 more appropriate and ingenious approaches to understand the bidirectional relationship
874 between genes and functional organization, which is a greater challenge for neuroscience
875 than just identifying a link between genetic and imaging data. Nevertheless, in light of the
876 currently limited understanding of how microscale genes contribute to macroscale brain
877 functional organization, the prioritization of genes and the related functional annotations
878 presented here are still necessary [24, 26].

879 **Conclusions**

880 In summary, we found that the network-specific genes underlying cerebellar functional
881 heterogeneity correlated with the intra-cerebellar and cerebello-cerebral FC. It indicates
882 that the genetic infrastructure associated with functional segregation coalesces to form a
883 collective system, which closely relates to the functional integration of these functional
884 subregions. The current study has thus unveiled part of the neurobiological genetic
885 substrates underlying the cerebellar functional organization. We also identified important
886 indirect genetic markers that support the key role played by the cerebellar functional
887 network in emotion-cognitional behaviors and many brain disorders. These findings hint
888 at the possibility of establishing a cerebellar "gene—connection—function/dysfunction"
889 chain, as well as of helping to bridge the knowledge gap between the genetic mechanisms
890 driving the cerebellar functional organization and the predictive markers of behaviors and
891 heritable risks of disorders, especially major depression and autistic disorder.

Acknowledgements

This work was partially supported by the Science and Technology Innovation 2030—Brain Science and Brain-Inspired Intelligence Project (Grant No. 2021ZD0200203), the Natural Science Foundation of China (Grant Nos. 82072099 and 91432302), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB32030214), the Youth Innovation Promotion Association, the Beijing Advanced Discipline Fund and the Major Scientific Project of Zhejiang Lab (Grant No. 2021ND0PI01). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen, and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The authors appreciate the English language and editing assistance of Rhoda E. and Edmund F. Perozzi, PhDs.

Author Contributions

YW and LF designed the research; YW and LC performed the experiments; YW, LF, CC, TJ, BL, and ZY contributed new analytic tools; YW, LC, DL, XW and CG analyzed data; YW and LF wrote the paper; and LF, CC, KHM, JRN, YZ, ZY, JX, SBE, BL, and TJ contributed analytic expertise, theoretical guidance, paper revisions, and informed interpretation of the results; and LF and CC supervised the project.

Conflict of Interest

All authors report no biomedical financial interests or potential conflicts of interest.

912 **References**

913 1. Schmahmann JD, Guell X, Stoodley CJ, Halko MA. The Theory and Neuroscience
914 of Cerebellar Cognition. *Annu Rev Neurosci* 2019; **42:** 337-364.

915 2. De Zeeuw CI, Lisberger SG, Raymond JL. Diversity and dynamism in the
916 cerebellum. *Nat Neurosci* 2021; **24**(2)**:** 160-167.

917 3. Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. *Annu Rev*
918 *Neurosci* 2009; **32:** 413-434.

919 4. Buckner RL. The cerebellum and cognitive function: 25 years of insight from
920 anatomy and neuroimaging. *Neuron* 2013; **80**(3)**:** 807-815.

921 5. Sathyanesan A, Zhou J, Scafidi J, Heck DH, Sillitoe RV, Gallo V. Emerging
922 connections between cerebellar development, behaviour and complex brain
923 disorders. *Nat Rev Neurosci* 2019; **20**(5)**:** 298-313.

924 6. Guell X, Schmahmann JD, Gabrieli JDE. Functional Specialization is Independent
925 of Microstructural Variation in Cerebellum but Not in Cerebral Cortex. *bioRxiv*
926 2018**:** 424176.

927 7. Voogd J, Glickstein M. The anatomy of the cerebellum. *Trends Cognit Sci* 1998;
928 **2**(9)**:** 307-313.

929 8. Schmahmann JD. The role of the cerebellum in affect and psychosis. *J*
930 *Neurolinguistics* 2000; **13**(2)**:** 189-214.

931 9. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA *et al.*
932 An anatomically comprehensive atlas of the adult human brain transcriptome.
933 *Nature* 2012; **489**(7416)**:** 391-399.

934 10. Johnson MB, Kawasawa YI, Mason CE, Krsnik Z, Coppola G, Bogdanovic D *et al.*
935 Functional and evolutionary insights into human brain development through global
936 transcriptome analysis. *Neuron* 2009; **62**(4)**:** 494-509.

937 11. Guevara EE, Hopkins WD, Hof PR, Ely JJ, Bradley BJ, Sherwood CC.
938 Comparative analysis reveals distinctive epigenetic features of the human
939 cerebellum. *PLoS Genet* 2021; **17**(5)**:** e1009506.

940 12. Wang VY, Zoghbi HY. Genetic regulation of cerebellar development. *Nat Rev*
941 *Neurosci* 2001; **2**(7)**:** 484-491.

942 13. Hawrylycz M, Miller JA, Menon V, Feng D, Dolbeare T, Guillozet-Bongaarts AL
943 *et al.* Canonical genetic signatures of the adult human brain. *Nat Neurosci* 2015;
944 **18**(12)**:** 1832-1844.

945 14. Negi SK, Guda C. Global gene expression profiling of healthy human brain and its
946 application in studying neurological disorders. *Sci Rep* 2017; **7**(1)**:** 897.

947 15. Aldinger KA, Thomson Z, Phelps IG, Haldipur P, Deng M, Timms AE *et al.* Spatial
948 and cell type transcriptional landscape of human cerebellar development. *Nat*
949 *Neurosci* 2021; **24**(8)**:** 1163-1175.

950 16. Zeng T, Chen H, Fakhry A, Hu X, Liu T, Ji S. Allen mouse brain atlases reveal
951 different neural connection and gene expression patterns in cerebellum gyri and
952 sulci. *Brain Struct Funct* 2015; **220**(5)**:** 2691-2703.

953 17. Kozareva V, Martin C, Osorno T, Rudolph S, Guo C, Vanderburg C *et al.* A
954 transcriptomic atlas of the mouse cerebellum reveals regional specializations and
955 novel cell types. *bioRxiv* 2020**:** 2020.2003.2004.976407.

956 18. Ren Y, Guo L, Guo CC. A connectivity-based parcellation improved functional
957 representation of the human cerebellum. *Sci Rep* 2019; **9**(1)**:** 9115.

958 19. Bernard JA, Seidler RD, Hassevoort KM, Benson BL, Welsh RC, Wiggins JL *et al.*
959 Resting state cortico-cerebellar functional connectivity networks: a comparison of
960 anatomical and self-organizing map approaches. *Front Neuroanat* 2012; **6:** 31.

961 20. King M, Hernandez-Castillo CR, Poldrack RA, Ivry RB, Diedrichsen J. Functional
962 boundaries in the human cerebellum revealed by a multi-domain task battery. *Nat*
963 *Neurosci* 2019; **22**(8)**:** 1371-1378.

964 21. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BT. The organization of
965 the human cerebellum estimated by intrinsic functional connectivity. *J*
966 *Neurophysiol* 2011; **106**(5)**:** 2322-2345.

967 22. Ji JL, Spronk M, Kulkarni K, Repovs G, Anticevic A, Cole MW. Mapping the
968 human brain's cortical-subcortical functional network organization. *Neuroimage*
969 2019; **185:** 35-57.

970 23. Guell X, Schmahmann JD, Gabrieli JDE, Ghosh SS. Functional gradients of the
971 cerebellum. *eLife* 2018; **7:** e36652.

972 24. Fornito A, Arnatkeviciute A, Fulcher BD. Bridging the Gap between Connectome
973 and Transcriptome. *Trends Cogn Sci* 2019; **23**(1)**:** 34-50.

974 25. Richiardi J, Altmann A, Milazzo AC, Chang C, Chakravarty MM, Banaschewski
975 T *et al.* BRAIN NETWORKS. Correlated gene expression supports synchronous
976 activity in brain networks. *Science* 2015; **348**(6240)**:** 1241-1244.

977 26. Anderson KM, Krienen FM, Choi EY, Reinen JM, Yeo BTT, Holmes AJ. Gene
978 expression links functional networks across cortex and striatum. *Nat Commun* 2018;
979 **9**(1)**:** 1428.

980 27. Ji JL, Helmer M, Fonteneau C, Burt JB, Tamayo Z, Demšar J *et al.* Mapping brain-
981 behavior space relationships along the psychosis spectrum. *eLife* 2021; **10:** e66968.

982 28. van den Heuvel MP, Scholtens LH, Kahn RS. Multiscale Neuroscience of
983 Psychiatric Disorders. *Biol Psychiatry* 2019; **86**(7)**:** 512-522.

984 29. Phillips JR, Hewedi DH, Eissa AM, Moustafa AA. The cerebellum and psychiatric
985 disorders. *Front Public Health* 2015; **3:** 66.

986 30. Rajendran AG, Nutakki C, Sasidharakurup H, Nair eBB, Diwakar S. Cerebellum
987 in Neurological Disorders: A Review on the Role of Inter-Connected Neural
988 Circuits. *J neurol stroke* 2017; **6**(2)**:** 00196.

989 31. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al.* limma powers
990 differential expression analyses for RNA-sequencing and microarray studies.
991 *Nucleic Acids Res* 2015; **43**(7)**:** e47.

992 32. Arnatkeviciute A, Fulcher BD, Fornito A. A practical guide to linking brain-wide
993 gene expression and neuroimaging data. *Neuroimage* 2019; **189:** 353-367.

994 33. Miller JA, Cai C, Langfelder P, Geschwind DH, Kurian SM, Salomon DR *et al.*
995 Strategies for aggregating gene expression data: the collapseRows R function. *BMC*
996 *Bioinformatics* 2011; **12:** 322.

997 34. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M *et al.*
998 The organization of the human cerebral cortex estimated by intrinsic functional
999 connectivity. *J Neurophysiol* 2011; **106**(3)**:** 1125-1165.

1000 35. Cox RW. AFNI: software for analysis and visualization of functional magnetic
1001 resonance neuroimages. *Comput Biomed Res* 1996; **29**(3)**:** 162-173.

1002 36. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing
1003 differential expression in microarray experiments. *Bioinformatics* 2005; **21**(9)**:**
1004 2067-2075.

1005 37. Dalman MR, Deeter A, Nimishakavi G, Duan Z-H. Fold change and p-value cutoffs
1006 significantly alter microarray interpretations. *BMC Bioinformatics* 2012; **13**(2)**:**
1007 S11.

1008 38. Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N. A probabilistic MR
1009 atlas of the human cerebellum. *Neuroimage* 2009; **46**(1)**:** 39-46.

1010 39. Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G *et*
1011 *al.* Resting-state fMRI in the Human Connectome Project. *Neuroimage* 2013; **80:**
1012 144-168.

1013 40. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL *et*
1014 *al.* The minimal preprocessing pipelines for the Human Connectome Project.
1015 *Neuroimage* 2013; **80:** 105-124.

1016 41. Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz R *et al.*
1017 The Human Connectome Project: a data acquisition perspective. *Neuroimage* 2012;
1018 **62**(4)**:** 2222-2231.

1019 42. Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM.
1020 Automatic denoising of functional MRI data: combining independent component
1021 analysis and hierarchical fusion of classifiers. *Neuroimage* 2014; **90:** 449-468.

1022 43. Griffanti L, Salimi-Khorshidi G, Beckmann CF, Auerbach EJ, Douaud G, Sexton
1023 CE *et al.* ICA-based artefact removal and accelerated fMRI acquisition for
1024 improved resting state network imaging. *Neuroimage* 2014; **95:** 232-247.

1025 44. Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD. Generative modeling of
1026 brain maps with spatial autocorrelation. *Neuroimage* 2020; **220:** 117038.

1027 45. Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL *et al.*
1028 The Lifespan Human Connectome Project in Aging: An overview. *Neuroimage*
1029 2019; **185:** 335-348.

1030 46. Seidlitz J, Vasa F, Shinn M, Romero-Garcia R, Whitaker KJ, Vertes PE *et al.*
1031 Morphometric Similarity Networks Detect Microscale Cortical Organization and
1032 Predict Inter-Individual Cognitive Variation. *Neuron* 2018; **97**(1)**:** 231-247 e237.

1033 47. Kong X-Z, Tzourio-Mazoyer N, Joliot M, Fedorenko E, Liu J, Fisher SE *et al.* Gene
1034 Expression Correlates of the Cortical Network Underlying Sentence Processing.
1035 *Neurobiology of Language* 2020; **1**(1)**:** 77-103.

1036 48. Tian Y, Zalesky A. Characterizing the functional connectivity diversity of the
1037 insula cortex: Subregions, diversity curves and behavior. *Neuroimage* 2018; **183:**
1038 716-733.

1039 49. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation
1040 inference for the general linear model. *Neuroimage* 2014; **92:** 381-397.

1041 50. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list
1042 enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;
1043 **37**(Web Server issue)**:** W305-311.

1044 51. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM *et al.* Gene
1045 ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat*
1046 *Genet* 2000; **25**(1)**:** 25-29.

1047 52. Dougherty JD, Schmidt EF, Nakajima M, Heintz N. Analytical approaches to RNA
1048 profiling data for the identification of genes enriched in specific cells. *Nucleic Acids*
1049 *Res* 2010; **38**(13)**:** 4218-4230.

1050 53. Huntley RP, Sawford T, Mutowo-Meullenet P, Shypitsyna A, Bonilla C, Martin
1051 MJ *et al.* The GOA database: gene Ontology annotation updates for 2015. *Nucleic*
1052 *Acids Res* 2015; **43**(Database issue)**:** D1057-1063.

1053 54. Zhu DM, Yang Y, Zhang Y, Wang C, Wang Y, Zhang C *et al.* Cerebellar-cerebral
1054 dynamic functional connectivity alterations in major depressive disorder. *J Affect*
1055 *Disord* 2020; **275:** 319-328.

1056 55. Verly M, Verhoeven J, Zink I, Mantini D, Peeters R, Deprez S *et al.* Altered
1057 functional connectivity of the language network in ASD: role of classical language
1058 areas and cerebellum. *Neuroimage Clin* 2014; **4:** 374-382.

1059 56. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A *et al.* Transcriptional
1060 landscape of the prenatal human brain. *Nature* 2014; **508**(7495)**:** 199-206.

1061 57. ten Donkelaar HJ, Lammens M, Wesseling P, Thijssen HO, Renier WO.
1062 Development and developmental disorders of the human cerebellum. *J Neurol* 2003;
1063 **250**(9)**:** 1025-1036.

1064 58. Munoz-Castaneda R, Diaz D, Peris L, Andrieux A, Bosc C, Munoz-Castaneda JM
1065 *et al.* Cytoskeleton stability is essential for the integrity of the cerebellum and its
1066 motor- and affective-related behaviors. *Sci Rep* 2018; **8**(1)**:** 3072.

1067 59. Beyreli I, Karakahya O, Cicek AE. Deep multitask learning of gene risk for
1068 comorbid neurodevelopmental disorders. *bioRxiv* 2020**:** 2020.2006.2013.150201.

1069 60. Villarroel MA, Terlizzi EP. Symptoms of Depression Among Adults: United States,
1070 2019. *NCHS Data Brief* 2020; (379)**:** 1-8.

1071 61. Anderson ML, Kinnison J, Pessoa L. Describing functional diversity of brain
1072 regions and brain networks. *Neuroimage* 2013; **73:** 50-58.

1073 62. Cerminara NL, Lang EJ, Sillitoe RV, Apps R. Redefining the cerebellar cortex as
1074 an assembly of non-uniform Purkinje cell microcircuits. *Nat Rev Neurosci* 2015;
1075 **16**(2)**:** 79-93.

1076 63. Margulies DS, Ghosh SS, Goulas A, Falkiewicz M, Huntenburg JM, Langs G *et al.*
1077 Situating the default-mode network along a principal gradient of macroscale
1078 cortical organization. *Proc Natl Acad Sci U S A* 2016; **113**(44)**:** 12574-12579.

1079 64. Schmahmann JD, Pandya DN. Disconnection syndromes of basal ganglia, thalamus,
1080 and cerebrocerebellar systems. *Cortex* 2008; **44**(8)**:** 1037-1066.

1081 65. Schmahmann JD. From movement to thought: anatomic substrates of the cerebellar
1082 contribution to cognitive processing. *Hum Brain Mapp* 1996; **4**(3)**:** 174-198.

1083 66. van den Heuvel MP, Yeo BTT. A Spotlight on Bridging Microscale and Macroscale
1084 Human Brain Architecture. *Neuron* 2017; **93**(6)**:** 1248-1251.

1085 67. Krubitzer L, Kaas J. The evolution of the neocortex in mammals: how is phenotypic
1086 diversity generated? *Curr Opin Neurobiol* 2005; **15**(4)**:** 444-453.

1087 68. Boyce WT, Sokolowski MB, Robinson GE. Genes and environments, development
1088 and time. *Proc Natl Acad Sci U S A* 2020; **117**(38)**:** 23235-23241.

1089 69. Storbeck J, Clore GL. On the interdependence of cognition and emotion. *Cogn*
1090 *Emot* 2007; **21**(6)**:** 1212-1237.

1091 70. Ghashghaei HT, Barbas H. Pathways for emotion: interactions of prefrontal and
1092 anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience*
1093 2002; **115**(4)**:** 1261-1279.

1094 71. Riener CR, Stefanucci JK, Proffitt DR, Clore G. An effect of mood on the
1095 perception of geographical slant. *Cogn Emot* 2011; **25**(1)**:** 174-182.

1096 72. Mueller S, Wang D, Fox MD, Yeo BT, Sepulcre J, Sabuncu MR *et al.* Individual
1097 variability in functional connectivity architecture of the human brain. *Neuron* 2013;
1098 **77**(3)**:** 586-595.

1099 73. Guell X, Schmahmann JD, Gabrieli JD. Functional Specialization is Independent
1100 of Microstructural Variation in Cerebellum but Not in Cerebral Cortex. *bioRxiv*
1101 2018**:** 424176.

1102 74. Hensler JG. Serotonergic modulation of the limbic system. *Neurosci Biobehav Rev*
1103 2006; **30**(2)**:** 203-214.

1104 75. Schmahmann JD. Disorders of the cerebellum: ataxia, dysmetria of thought, and
1105 the cerebellar cognitive affective syndrome. *J Neuropsychiatry Clin Neurosci* 2004;
1106 **16**(3)**:** 367-378.

1107 76. Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L *et al.*
1108 Analysis of shared heritability in common disorders of the brain. *Science* 2018;
1109 **360**(6395)**:** eaap8757.

1110 77. Bertoli-Avella AM, Garcia-Aznar JM, Brandau O, Al-Hakami F, Yuksel Z, Marais
1111 A *et al.* Biallelic inactivating variants in the GTPBP2 gene cause a
1112 neurodevelopmental disorder with severe intellectual disability. *Eur J Hum Genet*
1113 2018; **26**(4)**:** 592-598.

1114 78. Wefers AK, Lindner S, Schulte JH, Schuller U. Overexpression of Lin28b in Neural
1115 Stem Cells is Insufficient for Brain Tumor Formation, but Induces Pathological
1116 Lobulation of the Developing Cerebellum. *Cerebellum* 2017; **16**(1)**:** 122-131.

1117 79. Wang F, Yang J, Pan F, Ho RC, Huang JH. Editorial: Neurotransmitters and
1118 Emotions. *Front Psychol* 2020; **11:** 21.

1119 80. Svob Strac D, Pivac N, Muck-Seler D. The serotonergic system and cognitive
1120 function. *Transl Neurosci* 2016; **7**(1)**:** 35-49.

1121 81. Swoboda KJ, Hyland K. Diagnosis and treatment of neurotransmitter-related
1122 disorders. *Neurol Clin* 2002; **20**(4)**:** 1143-1161, viii.

1123 82. Brown RP, Mann JJ. A clinical perspective on the role of neurotransmitters in
1124 mental disorders. *Hosp Community Psychiatry* 1985; **36**(2)**:** 141-150.

1125 83. Kato TA, Yamauchi Y, Horikawa H, Monji A, Mizoguchi Y, Seki Y *et al.*
1126 Neurotransmitters, psychotropic drugs and microglia: clinical implications for
1127 psychiatry. *Curr Med Chem* 2013; **20**(3)**:** 331-344.

1128 84. Seo D, Patrick CJ, Kennealy PJ. Role of Serotonin and Dopamine System
1129 Interactions in the Neurobiology of Impulsive Aggression and its Comorbidity with
1130 other Clinical Disorders. *Aggress Violent Behav* 2008; **13**(5)**:** 383-395.

1131 85. Su Y, D'Arcy C, Meng X. Research Review: Developmental origins of depression
1132 – a systematic review and meta-analysis. *J Child Psychol Psychiatry* 2021; **62**(9)**:**
1133 1050-1066.

1134 86. Newschaffer CJ, Fallin D, Lee NL. Heritable and nonheritable risk factors for
1135 autism spectrum disorders. *Epidemiol Rev* 2002; **24**(2)**:** 137-153.

87. Guell X, Gabrieli JDE, Schmahmann JD. Embodied cognition and the cerebellum: Perspectives from the Dysmetria of Thought and the Universal Cerebellar Transform theories. *Cortex* 2018; **100:** 140-148.

88. Habas C. Functional Connectivity of the Cognitive Cerebellum. *Front Syst Neurosci* 2021; **15:** 642225.

89. Stoodley CJ, MacMore JP, Makris N, Sherman JC, Schmahmann JD. Location of lesion determines motor vs. cognitive consequences in patients with cerebellar stroke. *Neuroimage Clin* 2016; **12:** 765-775.

90. Saarimaki H, Ejtehadian LF, Glerean E, Jaaskelainen IP, Vuilleumier P, Sams M *et al.* Distributed affective space represents multiple emotion categories across the human brain. *Soc Cogn Affect Neurosci* 2018; **13**(5)**:** 471-482.

91. Ito M. Control of mental activities by internal models in the cerebellum. *Nat Rev Neurosci* 2008; **9**(4)**:** 304-313.

92. Peterburs J, Hofmann D, Becker MPI, Nitsch AM, Miltner WHR, Straube T. The role of the cerebellum for feedback processing and behavioral switching in a reversal-learning task. *Brain Cogn* 2018; **125:** 142-148.

93. Schmahmann JD. The role of the cerebellum in cognition and emotion: personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychol Rev* 2010; **20**(3)**:** 236-260.

94. Hoche F, Guell X, Vangel MG, Sherman JC, Schmahmann JD. The cerebellar cognitive affective/Schmahmann syndrome scale. *Brain* 2018; **141**(1)**:** 248-270.

95. Ramos TC, Balardin JB, Sato JR, Fujita A. Abnormal Cortico-Cerebellar Functional Connectivity in Autism Spectrum Disorder. *Front Syst Neurosci* 2018; **12:** 74.

1160 96. Baker KG, Harding AJ, Halliday GM, Kril JJ, Harper CG. Neuronal loss in
1161 functional zones of the cerebellum of chronic alcoholics with and without
1162 Wernicke's encephalopathy. *Neuroscience* 1999; **91**(2)**:** 429-438.

1163 97. Abdallah M, Zahr NM, Saranathan M, Honnorat N, Farrugia N, Pfefferbaum A *et*
1164 *al.* Altered Cerebro-Cerebellar Dynamic Functional Connectivity in Alcohol Use
1165 Disorder: a Resting-State fMRI Study. *The Cerebellum* 2021; **20**(6)**:** 823-835.

1166 98. Freire-Cobo C, Wang J. Dietary phytochemicals modulate experience-dependent
1167 changes in Neurexin gene expression and alternative splicing in mice after chronic
1168 variable stress exposure. *Eur J Pharmacol* 2020; **883:** 173362.

1169 99. Hu Z, Xiao X, Zhang Z, Li M. Genetic insights and neurobiological implications
1170 from NRXN1 in neuropsychiatric disorders. *Mol Psychiatry* 2019; **24**(10)**:** 1400-
1171 1414.

1172 100. Hua J, Yang Z, Jiang T, Yu S. Pairwise interactions in gene expression determine
1173 a hierarchical transcriptional profile in the human brain. *Sci Bull* 2021; **66**(14)**:**
1174 1437-1447.

1175 101. Ball G, Seidlitz J, Beare R, Seal ML. Cortical remodelling in childhood is
1176 associated with genes enriched for neurodevelopmental disorders. *Neuroimage*
1177 2020; **215:** 116803.

1178 102. Fulcher BD, Arnatkeviciute A, Fornito A. Overcoming false-positive gene-
1179 category enrichment in the analysis of spatially resolved transcriptomic brain atlas
1180 data. *Nat Commun* 2021; **12**(1)**:** 2669.

1181