Published in final edited form as:
*Neuroimage*. 2014 February 1; 86: 392–403. doi:10.1016/j.neuroimage.2013.10.006.

# Genetic effects on the cerebellar role in working memory: Same brain, different genes?

**Gabriëlla A.M. Blokland**[a,b,c,*], **Katie L. McMahon**[b], **Paul M. Thompson**[d], **Ian B. Hickie**[e], **Nicholas G. Martin**[a], **Greig I. de Zubicaray**[c], and **Margaret J. Wright**[a,c]

[a]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia

[b]Centre for Advanced Imaging, University of Queensland, Brisbane, Australia

[c]School of Psychology, University of Queensland, Brisbane, Australia

[d]Laboratory of Neuro Imaging, Department of Neurology, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

[e]Brain & Mind Research Institute, The University of Sydney, Sydney, Australia

## Abstract

Over the past several years, evidence has accumulated showing that the cerebellum plays a significant role in cognitive function. Here we show, in a large genetically informative twin sample ($n$ = 430; aged 16–30 years), that the cerebellum is strongly, and reliably ($n$ = 30 rescans), activated during an $n$-back working memory task, particularly lobules I–IV, VIIa Crus I and II, IX and the vermis. Monozygotic twin correlations for cerebellar activation were generally much larger than dizygotic twin correlations, consistent with genetic influences. Structural equation models showed that up to 65% of the variance in cerebellar activation during working memory is genetic (averaging 34% across significant voxels), most prominently in the lobules VI, and VIIa Crus I, with the remaining variance explained by unique/unshared environmental factors. Heritability estimates for brain activation in the cerebellum agree with those found for working memory activation in the cerebral cortex, even though cerebellar cyto-architecture differs substantially. Phenotypic correlations between BOLD percent signal change in cerebrum and cerebellum were low, and bivariate modeling indicated that genetic influences on the cerebellum are at least partly specific to the cerebellum. Activation on the voxel-level correlated very weakly with cerebellar gray matter volume, suggesting specific genetic influences on the BOLD signal. Heritable signals identified here should facilitate discovery of genetic polymorphisms influencing cerebellar function through genome-wide association studies, to elucidate the genetic liability to brain disorders affecting the cerebellum.

### Keywords

Cerebellum; Heritability; Genetics; Functional MRI; Working memory; Twin study



*Corresponding author at: Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Locked Bag 2000, Royal Brisbane Hospital, Herston, Qld 4029, Australia. gabriella.blokland@uqconnect.edu.au (G.A.M. Blokland)..

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.10.006.

## Introduction

Traditionally, the cerebellum has been associated primarily with the coordination of movement. However, in many human lesion and functional neuroimaging studies, the cerebellum clearly also subserves emotional control and higher cognitive operations, such as working memory (Chen and Desmond, 2005; Gottwald et al., 2004; Hokkanen et al., 2006; Ravizza et al., 2006; Scott et al., 2001). Recently, a meta-analysis of working memory neuroimaging studies reported significant activation in the cerebellum, in bilateral lobules VI and VIIa Crus I, and right lobule VIIIa (Stoodley and Schmahmann, 2009). The cerebellar role in working memory and other non-motor functions is attributed to cerebro-cerebellar pathways that link the cerebellum with motor, association, and paralimbic cortices (Stoodley and Schmahmann, 2009). Histological studies and investigations in non-human primates reveal connections from vermal and hemispheral parts of cerebellar lobule VII Crus II to mainly lateral portions of dorsolateral prefrontal cortex (DLPFC). Their co-activation is consistent with the operation of a functional network: the 'prefrontal' cortico-cerebellar loop (Kelly and Strick, 2003; Middleton and Strick, 2001; Ramnani, 2006). Cerebellar activation during working memory may reflect the automated simulation of cognitive operations that initially rely on interactions between prefrontal areas (Hayter et al., 2007).

Abnormal cerebellar activation and disruption of cortico-cerebellar circuits, in addition to abnormal cerebral activation, have been observed in neuropsychiatric disorders such as schizophrenia and attention-deficit hyperactivity disorder (Andreasen et al., 1996; Bor et al., 2011; Massat et al., 2012; Valera et al., 2005). This disruption may be an early diagnostic indicator of psychopathology. As susceptibility to these disorders is strongly genetically determined, finding the genetic influences on the underlying brain processes may greatly increase our understanding of the etiology of these disorders. Hypo-activation of the fronto-cerebellar loop observed in these disorders may actually be due to specific genetic risk factors affecting the cerebellum, and not the frontal cortex. Although it is reasonable to assume that if genetic effects on cerebral working memory blood oxygenation level dependent (BOLD) signal exist (~33%; Blokland et al., 2011), they would also exist for the cerebellar involvement in working memory, it is of interest to study the cerebellum separately because of its unique cyto-architecture. The cerebellum may be influenced by genes specific to this unique cyto-architecture, and/or may be influenced by genes to a different extent.

Even though some twin studies have shown high heritability for cerebellar gray matter (GM) volume (~64–93%; Kremen et al., 2010; van Soelen et al., 2013), so far no functional MRI (fMRI) twin studies have examined the heritability of task-related neural activity in the cerebellum (Blokland et al., 2008, 2011; Côté et al., 2007; Koten et al., 2009; Matthews et al., 2007; Park et al., 2012; Polk et al., 2007). The focus of imaging genetics studies using working memory fMRI as an endophenotype has been on the prefrontal cortex (e.g., Bertolino et al., 2006, 2010; Egan et al., 2001; Krug et al., 2008), but if cerebellar activation is heritable, and particularly if different sets of genes influence the BOLD signal of the cerebrum and cerebellum, using activation of the cerebellum as an endophenotype in gene finding studies, provides an opportunity to detect novel gene variants that may be relevant to brain (dys)function, whose association would not be detected when using activation of the cerebrum as an endophenotype. This could significantly increase our knowledge of the brain, and these gene variants could be of critical importance to brain disorders.

Therefore, the goal of this study was two-fold: First, we investigated the usefulness of working memory cerebellar activation as an endophenotype by computing voxel-wise heritability maps. Even with a rigidly standardized *n*-back task, the extent and intensity of

cerebellum activation differs greatly between individuals. To disentangle genetic and environmental contributions to this individual variability in cerebellar BOLD response, we measured cerebellar response to an established spatial *n*-back task (Callicott et al., 2003), using fMRI in a large sample that includes both monozygotic (MZ, who share all their genes) and dizygotic (DZ, who share 50% of their genes, on average) twin pairs. We improved the localization of cerebellar activation in response to a working memory task by normalizing the individual functional scans to the spatially unbiased infra-tentorial template (Diedrichsen, 2006).

Second, we investigated the phenotypic and genetic relationship between BOLD percent signal change in the cerebellum and cerebrum on a region-of-interest (ROI) basis. Using bivariate genetic modeling we investigated whether the same set of genes influences BOLD percent signal change in both the cerebrum and the cerebellum, i.e., whether there is genetic overlap between cerebellar areas of peak BOLD response and cerebral areas of peak BOLD response. In line with the existence of a cortico-cerebellar loop, we expected to find significant phenotypic correlations between prefrontal BOLD signal and cerebellar BOLD signal, but given the different cyto-architecture in the cerebral and cerebellar cortex, we expected partly specific genetic influences on the cerebellar BOLD signal.

## Materials and methods

### Participants

Twins between 16 and 30 years of age participated in the present study. The large majority of these individuals also participated in the Brisbane Twin Cognition study at age 16 (Wright and Martin, 2004). Thus, additional information was available on general intellectual ability (Full-scale intelligence quotient; FIQ; mean interval between cognitive testing and MRI scanning = 4.4 years; range 0–14 years), gestational duration, birth weight, and parental socio-economic status (rated according to the Australian Socioeconomic Index (SEI) 2006; McMillan et al., 2009). Gestational duration, birth weight, and parental socio-economic status were obtained from parental report, usually the mother. Prior to inclusion, twins were assessed for handedness, and screened (by self-report) for their suitability for imaging. Individuals with significant medical, psychiatric or neurological conditions, including head injuries, a current or past diagnosis of substance abuse, or current use of medication that could affect cognition were excluded. Zygosity was determined by genome-wide single nucleotide polymorphism genotyping (Illumina 610K chip) in ~90% of participants. If this was not available zygosity was established objectively by typing nine independent DNA microsatellite polymorphisms by polymerase chain reaction, and cross-checked with blood group results and phenotypic data, as described elsewhere (Wright and Martin, 2004).

A sample of 430 twins, which included 60 MZ pairs (47 female, 13 male pairs), 76 DZ pairs (39 female, 12 male, 25 opposite sex pairs), and 158 unpaired twins (114 female, 44 male pairs), aged 20.8 ± 3.1 years (mean ± s.d.), and all right handed, was included in our analyses. Unpaired twins did not contribute to the estimation of the genetic and environmental parameters, but they did contribute to the estimation of mean and variance effects, allowing more accurate estimation of phenotypic correlations and phenotypic effects. The majority of twin pairs (both MZ and DZ) were scanned on the same day (86%), with the remainder, on average, within 13 days of each other. A subset of this sample was included in our previous study on working memory brain activation in the cerebrum (60 of 430, 14%) (Blokland et al., 2011). However, in that study we did not explicitly examine the cerebellum. That is, the sample was not selected according to cerebellar coverage in their fMRI scans, whereas the sample here is a select subset of the cohort with full cerebellar coverage.

Human Research Ethics Committees of the Queensland Institute of Medical Research, University of Queensland, and Uniting Health Care approved the study. Written informed consent was obtained from each participant, and from a parent or legal guardian if the participant was under 18. Each participant received a $100 gift voucher in appreciation of their time.

## Experimental procedure

Imaging was conducted on a 4 Tesla Bruker Medspec whole body scanner (Bruker, Germany) in Brisbane, Australia. Functional images were acquired using a T2*-weighted gradient echo planar imaging (EPI) sequence, sensitive to blood oxygen level-dependent (BOLD) contrast (interleaved; repetition time, TR = 2100 ms; echo time, TE = 30 ms; flip angle = 90°; field of view, FOV = 230 × 230 mm). Geometric distortions in the EPI images caused by magnetic field inhomogeneities at high-field were corrected using a point-spread mapping approach (Zeng and Constable, 2002). Over a continuous imaging run, 127 axial brain volumes were acquired, with 36 coronal slices of 3mm thickness (64×64 matrix; voxel size 3.6×3.6×3.0mm), and with a 0.6mm slice gap. In addition to the functional scans, 3D T1-weighted anatomical images were acquired (MP-RAGE; TR = 1500 ms; TE = 3.35 ms; TI = 700 ms; pulse angle = 8°; coronal orientation; FOV = 230 mm; 256 × 256 × 256 matrix; slice thickness = 0.9 mm).

During functional imaging, participants performed the 0-back and 2-back versions of a block design spatial, numerical *n*-back working memory task based on Callicott et al. (2003). In this task, a number (1–4, randomized) was presented in a fixed position in one of the four corners of a diamond-shaped square. For the 0-back condition, the task required a simple button press in response to the number displayed, using a response box with four buttons arranged in the same configuration as the numbers presented on the screen. For the 2-back condition, participants pressed the key corresponding to the number presented two trials before the current one, thus requiring both the maintenance of the last 2 numbers in memory and the updating of these remembered stimuli as each new stimulus was presented. See Blokland et al. (2008, 2011) for a full task description. Participants were fully trained on the task, prior to being positioned in the scanner. The importance of effort and commitment to the task was emphasized. Task performance was measured as the percentage of correct responses (accuracy) and average response time (RT; across correct trials) for each of the task conditions separately.

It should be noted that there is more than one strategy that can be applied to perform this particular type of *n*-back task. Given that the response buttons were arranged in the same configuration as the numbers presented on the screen, a spatial or motor strategy can be applied. Participants could also apply a more verbal, numerical strategy where the number sequence is repeated using internal speech. For a subset of 312 participants, we ascertained their task strategy from self-report, divided into four categories: numerical, spatial, both numerical and spatial, or no strategy.

## Image pre-processing

Images were processed and analyzed using Statistical Parametric Mapping software (SPM8; Wellcome Trust Centre for Neuroimaging, Queen Square, London, UK). The first five EPI volumes were discarded to ensure that steady state tissue magnetization was reached. Time-series volumes were realigned and unwarped using a robust rigid-body transformation procedure (Freire et al., 2002). A mean image generated during realignment was then co-registered with the participant's 3D T1 image. After this, image processing diverged into separate pipelines for the cerebellum (A) and cerebrum (B).

**Cerebellum processing pipeline (A)—**Global signal effects were estimated and removed using a voxel-level general linear model (detrending; Macey et al., 2004). High pass (Discrete Cosine Transform cut-off 128s) and an approximate AR1 serial autocorrelation model were used to account for the drift and structured temporal noise. First-level (within-subject) analysis consisted of block design fixed effects model fitting. Separate regressors were constructed for the 0- and 2-back conditions, comprising a boxcar reference waveform convolved with a canonical hemodynamic response function. Then, on the T1-weighted image, the cerebellum was isolated and spatially normalized to the Spatially Unbiased Infra-tentorial Template (SUIT; Diedrichsen, 2006). The SUIT normalization was applied because the SPM nonlinear normalization algorithm to the MNI template can lead to an elongated cerebellum, which may alter the lobular localization of certain coordinates (Diedrichsen, 2006). The nonlinear transformations were then applied to the beta contrast images, resampling to 2 × 2 × 2 mm voxels. Images were smoothed with an 8 × 8 × 8 mm full width half maximum (FWHM) isotropic Gaussian kernel.

In addition, the inverse deformation field (with respect to the SUIT template) calculated in the segmentation step was applied to a cerebellar gray matter region of interest obtained from the SPM Anatomy Toolbox (all regions in Supplementary Fig. 1 together) (Diedrichsen et al., 2009; Eickhoff et al., 2006) to create a native space ROI and obtain cerebellar GM volume.

**Cerebrum processing pipeline (B)—**The T1-weighted image was segmented using the "New Segment" procedure in SPM8. The Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL; Ashburner, 2007) toolbox was then employed to create a custom group template from the gray and white matter images. The individual flow fields were used to normalize the realigned fMRI volumes to the Montreal Neurological Institute (MNI) atlas T1 template. The resulting images were resampled to 3×3×3mm voxels and smoothed with an 8×8×8mm FWHM isotropic Gaussian kernel. Global signal effects were estimated and removed using a voxel-level general linear model (Macey et al., 2004); high pass filtering (cut-off 128 s) and an approximate AR1 serial autocorrelation model were applied; and the first-level fixed effects model was fitted.

Additionally, intracranial volume was computed from the native space gray matter, white matter, and cerebrospinal fluid tissue class images obtained from the "New Segment" procedure.

## Image analysis

**Voxel-based analysis—**Second-level (group) analysis consisted of random effects analysis (one-sample *t*-test) of all single-subject 2-back > 0-back beta contrast images, irrespective of zygosity. Total cerebellar GM volume was included as a covariate in the random effects model to control for individual differences in overall cerebellar size. A brain mask was created by thresholding the group map at $p<0.05$, FWE-corrected (minimum cluster size 25 voxels) to determine which voxels were analyzed further (16,926 voxels).

Finally, to prepare the BOLD phenotype for genetic analysis, beta contrast values were extracted from single-subject 2 > 0-back contrast images in each of the 16,926 voxels specified by the mask using the Volumes Toolbox (Freiburg Brain Imaging Center, Germany; sourceforge.net/projects/spmtools). Voxel values from the contrast images approximated normal distributions and did not require transformation. For each case, in each voxel, the square of the Mahalanobis distance was calculated to screen for and remove outliers. The Mahalanobis distance was converted to a probability via the chi-square ($\chi^2$) statistic. The significance of the Mahalanobis distances was assessed against the critical value of the $\chi^2$ distribution. At a *p*-value threshold of 0.01, corresponding to a squared

Mahalanobis distance cut-off of 7.1, the number of outliers per voxel ranged between 0 and 22 (≤5%), averaging 9.4 ± 3.0 outliers. Percentages were as follows: 0–5 outliers: 7.5%; 6–10 outliers: 59.3%; 11–15 outliers: 30.6%; N15 outliers: 2.6%.

**ROI analysis—**To enable bivariate genetic modeling of cerebellar and cerebral activation, an additional ROI analysis was performed. The goal of these bivariate genetic analyses was not to estimate heritability of cerebellar activation, but to identify any genetic correlations between BOLD signal change in the cerebellum and cerebrum. To reduce the number of tests, and given that, by default, non-heritable regions, non-reliable regions, and phenotypically uncorrelated regions would not show genetic correlations, definition of ROIs was informed by heritability, reliability, and group activation. Firstly, 5 mm sphere ROIs were created around the local maxima from the group random effects analysis for the cerebellum, using the significant heritability and reliability clusters from the present voxel-based analysis as explicit masks. Secondly, 5 mm sphere ROIs were created around the local maxima from the group random effects analysis for the cerebrum, using the significant heritability and reliability clusters from Blokland et al. (2011) as explicit masks. Then, for each participant and each task condition, we extracted the average percent BOLD signal across all voxels in each of the ROIs using the MarsBaR Toolbox for SPM (Brett et al., 2002), and computed the difference in percent BOLD signal for the 2-back minus 0-back contrast.

## Reproducibility

Test–retest reliability of task-related brain activation across sessions was assessed by calculating voxel-wise intra-class correlations (Shrout and Fleiss, 1979) between contrast images from two time points for 30 participants (3 MZF pairs, 2 DZF pairs, 1 DZM pair, 15 unpaired twins and 3 singletons; 22 females and 8 males) who were rescanned 97±28days (range: 35–173days) after their initial scan. Corresponding *p*-value maps were height- and cluster-thresholded in the same manner as the variance component *p*-value maps (see below). Additionally, intra-class correlations were obtained for the functional ROI measures and cerebellar GM volume.

## Phenotypic correlations

In each group-activated voxel, we computed Pearson correlations between the beta-difference-score in that voxel on the one hand and age, sex, task performance measures, FIQ, and total cerebellar GM volume on the other hand. When computing phenotypic correlations, we included only one co-twin for MZ twin pairs (total $n = 377$). Corresponding *p*-value maps were height- and cluster-thresholded in the same manner as the variance component *p*-value maps (see below). Additionally, phenotypic correlations were computed among BOLD ROI measures and cerebellar GM volume.

## Genetic modeling

**Voxel-based genetic modeling—**Using the statistical package Mx (Neale and Maes, 2004; Neale et al., 2002), maximum likelihood twin correlations were estimated and univariate structural equation models examined the means and genetic (A) and environmental (E) sources of variance for task-related cerebellar activation extracted from single-subject 2 > 0-back contrast images in each of the voxels specified by the group analysis. Age, sex, 2-back performance accuracy and RT, were included as covariates, leaving their effects free to vary in each voxel. Correlations between genetic factors are fixed at 1 for MZ twin pairs, as they share 100% of their genes, and 0.5 for DZ pairs as they share, on average, 50% of their genes. Thus, if the patterns and intensities of cerebellar activity associated with the recruitment of working memory are significantly more similar in

MZ twins than in DZ twins, this is strong evidence of heritability. By definition, variance due to non-shared environmental factors (e.g., illness, prenatal or postnatal traumas, peer groups), which also includes measurement error, is left uncorrelated.

The significance of A was determined by testing whether dropping this parameter resulted in a significant ($p<0.05$) decrease in the goodness-of-fit chi-squared statistic ($\chi^2$). Probability maps were height thresholded per voxel at an uncorrected $p > 0.05$, then cluster thresholded to correct for multiple statistical tests. Cluster thresholding of *p*-value maps was carried out using FSL (www.fmrib.ox.ac.uk/fsl/) and AFNI (afni.nimh.nih.gov) command line tools. At 10,000 iterations, a voxel level threshold of $p < 0.05$, and an estimated heritability image smoothness of 6.81 × 6.72 × 6.02 mm (3dFWHMx), Monte Carlo simulation (3dClustSim) determined a minimum cluster size of 282 voxels for FWE corrected cluster significance at $p < 0.05$.

We assessed whether the AE model was appropriate in several steps. Firstly, we compared the maximum likelihood twin correlation maps for MZ and DZ pairs. DZ correlations were low, and approximately half the MZ correlation. Then, from the MZ and DZ twin correlations we used Falconer's estimator, rather than the structural equation model, to calculate initial estimates of heritability [$h^2 = 2 * (r_{MZ} - r_{DZ})$] and common environmental influence [$c^2 = (2 * r_{DZ}) - r_{MZ}$] (Falconer and Mackay, 1996). These were plotted against the expected sampling distributions under the null hypothesis of no heritability and no common environmental influence; where $h^2$ and $c^2$ expected are normal distributions with a mean of zero (Supplementary Fig. 2). The differences between the observed and expected sampling distributions are consistent with the presence of genetic influences (heritability) (i.e., the midline of the observed $h^2$ sampling distribution is well above zero) and the absence of common environmental influences and genetic dominance (i.e., the midline of the observed $c^2$ distribution is at zero). At a voxel-level threshold of $p < 0.05$, in less than 1% of group-activated voxels, the variance component C was significant. This C component did not survive cluster thresholding. The same was observed for the D variance component. Based on this result, an AE model was fitted for all group-activated voxels.

### ROI-based genetic modeling

Univariate genetic model fitting was also applied to BOLD percent signal change cerebellar and cerebral ROI measures and cerebellar GM volume. Similar to the voxel-based analysis, for the BOLD percent signal change cerebellar and cerebral ROI measures, an AE model was fitted.

Based on the twin correlations for cerebellar GM volume, a classical twin model was fitted, which separates the variance into additive genetic (A), common or shared environmental (C; e.g., socioeconomic status or parental rearing style), and unique or non-shared environmental (E) components. Nested sub-models (AE, CE, E) were compared to the full ACE/AE model by testing whether dropping a parameter resulted in a significant increase in the $\chi^2$ statistic.

Bivariate AE models examined the sources and pattern of covariation for reliable and significantly correlated cerebellar and cerebral ROI measures of BOLD percent signal change. Specifically, this allowed us to assess whether there were any genetic influences common to cerebral and cerebellar activation. Phenotypic correlations among ROI measures and 95% confidence intervals were computed by maximum likelihood (ML).

# Results

## Demographic characteristics and task performance

Consistent with our previous report (Blokland et al., 2011), there were no significant age, FIQ, gestation duration, birth weight, or socioeconomic index differences between MZ and DZ groups, but males had slightly higher FIQ (115.8 ± 12.4 vs. 111.8 ± 12.0, $d = 0.33$, $p < 0.01$), and birth weight (2617.0 ± 585.0 vs. 2436.2 ± 509.9, $d = 0.33$, $p < 0.01$) than females.

The observed higher mean FIQ (112.8±12.2) is likely due to the fact that the Multidimensional Aptitude Battery (Jackson, 1984) was created and normalized for Canadian samples, so results on this test may differ when used in a different country. However, ascertainment bias cannot be excluded, as more intelligent and more highly educated people tend to volunteer for studies more frequently. However, the higher FIQ mean did not affect the representativeness of this sample, because FIQ follows a normal distribution, with scores ranging from 82 to 149, thus showing wide variability.

Similar to our prior report (Blokland et al., 2011), the mean (±s.d.) accuracy percentages in the 0-back and 2-back conditions were 87.0% (±10.9) and 71.1% (±18.9), respectively. On the 0-back condition, 98.4% of the sample had >50% accuracy, and 100% had >40% accuracy. On the 2-back condition, 82.8% of the sample had >50% accuracy, and 91.6% had >40% accuracy. The mean RTs in the 0-back and 2-back conditions were 444.9 (±61.4) ms and 235.0 (±116.9) ms, respectively. Also consistent with our prior report, males had slightly higher 0-back performance accuracy (90.5 ± 8.6 vs. 85.7 ± 11.4, $d = 0.48$, $p < 0.001$), and 2-back performance accuracy (76.7 ± 18.2 vs. 68.9 ± 18.8, $d = 0.42$, $p < 0.001$) than females.

Strategy use for a subsample of 312 twins was as follows: 28.7% used a numerical strategy, 19.9% a spatial strategy, 4.4% used a combination of numerical and spatial strategies, and 47.2% used no particular strategy. This distribution was approximately the same for MZ and DZ twins. Importantly, MZ co-twins were somewhat more likely to use the same strategy ($r = 0.68$, $p < 0.001$; 34 pairs) than DZ co-twins ($r = 0.41$, $p < 0.01$; 60 pairs).

## Group activation and test–retest reliability

The group level random effects analysis showed the most significant increase in BOLD signal during the 2-back compared to the 0-back condition (pb0.05, family wise error [FWE]-corrected, cluster threshold 25 voxels) in the left and right lobules VIIa Crus I ([−32; −62; −32], [−8; −78; −28], [34; −60; −32], [38; −62; −28]) and VIIa Crus II ([−36; −64; −46], [32; −64; −44]), left and right lobule IX ([−12; −54; −46], [16; −54; −46]), left lobules I–IV [0; −50; −18], and right Vermis lobule VIIa Crus II [6; −80; −30] (Fig. 1A). This significant group activation, which covers about a quarter of the cerebellar cortex (16,926 voxels), was used as a brain mask. Our subsequent voxel-wise analyses were restricted to these regions. Only the right VIIb, right Vermis VIIa Crus I, and left and right Vermis X were completely excluded from the group mask. All other lobules were at least partly activated during the working memory task.

We found high reproducibility of task-related BOLD signal for the rescanned subsample of 30 individuals. Voxel-wise intra-class correlations between *t*-scores at time points 1 and 2 ranged between 0.7 and 0.9 for most activated areas as shown in Fig. 1B. A large cluster of 6537 voxels and a smaller cluster of 421 voxels were significant after thresholding, with the first cluster extending into the following lobules: VI L [−4; −86; −18], VI R [38; −66; −22], VIIa Crus I L ([−6; −86; −21], [−38, −64, −25]); VIIa Crus I R [36; −64; −39], VIIa Crus II L [−44; −60; −47], VIIa Crus II R [32; −70; −42], VIIb L [−32; −68; −45], VIIb R [28; −74; −44], VIIIb L [−16; −46; −43], X R [28; −36; −48], and Vermis VI L [−2; −84; −21]. The

second cluster covered lobules I–IV ([0; −44; −17], [−6; −44; −15], [−6; −48; −9]) and lobule V [−6; −54; −9]. The reliable activation areas covered almost half (41.1%) of the group activation area. The average reliability in these significant areas was $r = 0.61$ (range: 0.39–0.93).

## Genetic modeling

**Voxel-based univariate modeling—**To decide on the genetic model to test, we calculated voxel-by-voxel maximum likelihood twin correlations for task-related brain activation, as shown in Fig. 1C. For the regions activated by this task, overall, MZ twin correlations were more than twice the size of the DZ correlations, suggesting that individual variation in cerebellar working memory activation is genetically influenced. DZ twin correlations in many voxels did not differ significantly from zero.

Based on the twin correlations, we fitted a model that includes genetic (A), and unique environmental factors (E) to explain the variance in task-related cerebellar activation. The A and E estimates and the A clusters surviving height- ($p < 0.05$) and cluster-thresholding (FWE; $p < 0.05$) are shown in Fig. 1D. At a cluster threshold of 282 voxels, one large cluster of 8759 voxels was significant (51.7% of group-activated voxels), with the highest heritability estimates (50–65%) in right lobule VIIa Crus I [45; −73; −37], left lobule VIIa Crus I [−36; −69; −27], and left lobule VI [−18; −77; −20]. The significantly heritable area approximately matched the significantly reliable area. Lobules with the highest average heritability (regardless of significance of estimate) were Vermis lobule X at 32% (range 0–43%), right lobule VIIa Crus II at 29% (range 0–63%), Vermis lobule IX at 26% (range 0–47%), Vermis lobule VIIIb at 25% (range 0–49%), and left lobule VIIa Crus I at 25% (range 0–48%). Heritability across significant voxels (within the cluster of 8759 voxels) ranged between 20% and 65%, averaging 33.7%. A comparison of the group activation map (Fig. 1A) with the heritability map (Fig. 1D) demonstrates that the most strongly activated regions are not necessarily the most heritable regions.

## ROI-based univariate modeling

**Functional ROI measures.—**The area of significant heritability and reliability yielded 15 cerebellar local maxima and 14 cerebral local maxima. Similar to the voxel-based activation, the Mahalanobis distance was computed across all 29 reliable ROIs, and assessed against the critical value of the $\chi^2$ distribution, to screen for and remove outliers. At d.f. = 29, $p < 0.01$, and critical $\chi^2$ value = 49.6, less than 5% multivariate outliers were identified. After outlier removal the ROI measures approximated normal distributions so no transformations were required. The univariate results for the BOLD ROI measures are shown in Supplementary Table 1.

**Cerebellar GM volume.—**Cerebellar GM volume ranged between 95.2 and 182.2 ml, with a mean (±s.d.) of 126.5 (±15.5) ml. Males had significantly greater cerebellar GM volumes than females (140.4 ± 14.6 vs. 121.1±12.2, $d$=1.44, $p$ <0.001). Cerebellar GM volume showed high test–retest reliability, with an intra-class correlation of 0.89 (95% CI: 0.78, 0.95; $p < 0.001$). Because the distribution showed a very minor positive skew, a log10-transform was applied to the data. After this, the distribution approached normality, and showed no outliers at >4 s.d. from the mean.

Assumption testing supported homogeneity of means and variances across birth order and zygosity. Maximum likelihood twin correlations (corrected for sex, age, and intracranial volume) for cerebellar GM volume were consistent with the presence of both additive genetic and common environmental influences. That is, the MZ twin correlation of 0.77 (95% CI: 0.67, 0.85) was greater than the DZ correlation of 0.58 (95% CI: 0.41, 0.71), but

less than twice the DZ correlation. Under the full ACE model, the additive genetic contribution (A) was estimated at 38.3% (95% CI: 8.7, 73.9), the common environmental influence (C) at 39.1% (95% CI: 4.9, 63.9), and the unique environmental variance component (E) at 22.6% (95% CI: 15.4, 33.5). Maximum likelihood modeling showed that the ACE model was in fact the best fitting model (ACE: $\chi^2 = 542.5$, d.f. = 265; AE: $\chi^2 = 547.4$, $\Delta\chi^2 = 4.9$, $\Delta$d.f. = 1, p b 0.05; CE: $\chi^2 = 548.8$, $\Delta\chi^2 = 6.3$, $\Delta$d.f. = 1, p b 0.05; E: $\chi^2 = 631.0$, $\Delta\chi^2 = 88.5$, $\Delta$d.f. = 2, p b 0.001).

The phenotypic correlations between cerebellar GM volume on the one hand and the cerebellar BOLD ROI measures on the other hand did not exceed 0.2 (average $r = 0.08$). Therefore, bivariate modeling of these phenotypes was not justified.

**ROI-based bivariate modeling—**A heat map of the phenotypic correlations between the most reliable ROI measures of BOLD percent signal change in the cerebellum and cerebrum ($r > 0.4$), is shown in Fig. 2. Overall, correlations between BOLD percent signal change in the cerebellum and cerebrum were low. Of the $n = 210$ possible phenotypic correlations among 15 cerebellar and 14 cerebral ROIs (average $r = 0.05$), n = 7 (3.3%) were significant at $p < 0.05$ after correcting for multiple inference using Holm's method (Holm, 1979) (average $r$=0.21). Only 4 of those 7 significant correlations exceeded 0.2. The highest correlations between the cerebellum and cerebrum were observed among right lobule VIIa Crus I and left middle frontal gyrus ($r = 0.27$; $p = 0.004$, corrected), and among left lobule VI and right precentral gyrus ($r = 0.24$; $p = 0.004$, corrected).

A bivariate AE Cholesky decomposition was applied to the cerebellar and cerebral ROI measures that correlated most strongly (Fig. 3). Bivariate modeling showed that the phenotypic correlations between cerebellar and cerebral activation during working memory are partly accounted for by genetic factors shared between the two major brain divisions and partly specific to each brain division.

The two bivariate models showed rather striking differences. The first model represents a correlation on the cognitive level, between right VII Crus I and left middle frontal gyrus. Although the genetic correlation, $r_g$, was estimated at 0.27, the corresponding path coefficient did not significantly differ from zero, indicating that genetic influences on these phenotypes are specific to each ROI. The environmental correlation $r_e$, was of the same magnitude and significant, indicating that the largest proportion of the covariance between the two ROIs (61%) is due to shared E-influences and/or correlated measurement error.

The second model represents a correlation on the motor level, between left lobule VI and right precentral gyrus. Most of the phenotypic correlation between these two areas is accounted for by shared genetic influences. An A-factor shared with the cerebral ROI explained about 93% of the phenotypic correlation, and most of the genetic variance in the cerebellar ROI ($r_g = 0.85$). An additional specific genetic factor explained the remaining genetic variance in the cerebellar ROI. E-influences were mainly specific to each ROI. The environmental correlation was non-significant.

## Voxel-based phenotypic correlations

Fig. 4 shows the phenotypic correlations of 9 variables – age, sex, task performance measures, FIQ, and cerebellar GM volume – with cerebellar activation. Summary statistics and coordinates of peak correlation for significant correlations are reported in Table 1. Absolute correlations with cerebellar activation for all variables did not exceed 0.30.

Higher age was mostly associated with lower activation. Being male was mostly associated with lower activation as well, with the caveat that the number of males included in the

analyses was relatively small, making it difficult to get an accurate estimate of the effect of sex. The correlation between 2-back performance accuracy and BOLD signal was rather striking, and quite a bit stronger than in the cerebral cortex. Higher 2-back performance accuracy was associated with significantly lower activation in part of the cerebellar cortex, and with higher activation in other areas. Slower RT during the 2-back condition was associated with higher activation. FIQ correlated with activation in the same approximate areas as 2-back performance accuracy. Zero-back performance, both accuracy and RT, correlated very weakly with activation, as did cerebellar GM volume, although the latter correlation was somewhat more significant. Sex, age, 2-back performance accuracy and 2-back RT were included as covariates in our genetic analyses. Zero-back performance, FIQ, and cerebellar GM volume were non-significant predictors of brain activation after correcting for the other four variables.

## Discussion

The cerebellum is known to play an important role in cognition, possibly via a direct connection with the prefrontal cortex (Ramnani, 2006). Our results confirm that the cerebellum is actively involved in working memory, showing a highly reliable group activation pattern that includes the bilateral lobules VIIa Crus I and II, left lobules I–IV and bilateral IX, and right Vermis lobule VIIa Crus II. We found strong regional differences in the heritability of cerebellar activation, with the highest heritabilities (50–65%) in bilateral lobules VIIa Crus I, and left lobule VI. Region-of-interest-based bivariate genetic modeling showed that the phenotypic correlation between cerebellar and cerebral BOLD percent signal change is partly accounted for by genetic influences shared between the cerebrum and cerebellum, but there are also specific genetic factors influencing cerebellar activation during working memory. Furthermore, cerebellar activation correlated weakly with cerebellar GM volume, suggesting specific genetic influences on these phenotypes.

Relatively few fMRI studies have focused on the role of the cerebellum in cognitive function. Consistent with the crossed cerebrocerebellar fiber pathways, the left cerebellar hemisphere is more involved in visuo-spatial functions, whereas the right cerebellar hemisphere is more involved in language functions (Gottwald et al., 2004; Hokkanen et al., 2006; Scott et al., 2001). This is reflected in the greater task-related activation of the left cerebellar hemisphere compared to the right. Cerebellar activation patterns during working memory differed from recent meta-analysis results, which showed only activation of bilateral lobules VI and VIIa Crus I, and right lobule VIIIa (Stoodley and Schmahmann, 2009). This is most likely because the meta-analysis combined studies carried out using different working memory paradigms (spatial *n*-back plus verbal *n*-back plus Sternberg), task conditions, and small, heterogeneous samples, that detected cerebellar activation as part of whole-brain investigations. SPM nonlinear normalization of brain data to the MNI template alters the lobular localization of certain coordinates, and alignment errors may increase the spatial variability of activation (Diedrichsen, 2006). Here, we isolated the cerebellum prior to normalization, and analyzed a much larger sample, providing us with greater confidence in the localization of activation and of heritability.

Prior studies point to the existence of a DLPFC-cerebellar loop (Kelly and Strick, 2003; Middleton and Strick, 2001), which would explain the co-activation of the cerebellum and DLPFC. In the present study cerebellar activation did not correlate strongly with DLPFC activation or with task strategy (in which the DLPFC is considered to be involved), but this does not mean that the cerebellum is not involved in the *n*-back encoding, maintenance, or retrieval processes. A common aspect of encoding and complex working memory demands is inner speech. It has been suggested that encoding newly presented verbalizable information engages cerebellar activity for the purpose of creating an internal motor

(articulatory) representation, and once this process has been completed, cerebellar activity subsides. Inner speech would be required during the encoding phase upon initial presentation of the stimulus, and may generally be used to represent, maintain, and organize task-relevant information (Hayter et al., 2007; Owen et al., 2005; Strick et al., 2009). In fact, a review of cerebellar involvement in verbal working memory (Ben-Yehudah et al., 2007) suggests that the superior/lateral cerebellum participates in encoding and retrieval phases of verbal working memory tasks, rather than maintenance of information, as would be suggested by this region's preferential response to verbal stimuli (e.g., letters) over non-verbal stimuli (e.g., symbols) (Ravizza et al., 2004). Two event-related fMRI studies parsed the encoding, maintenance, and retrieval phases in their verbal working memory tasks and found superior cerebellar activity only for the encoding (Chen and Desmond, 2005) or encoding and retrieval phases (Chein and Fiez, 2001). This role in the *n*-back manipulations is supported by the correlations between the cerebellar BOLD signal and task performance measures, which were of similar magnitude to the cerebral BOLD correlation with task performance (Blokland et al., 2011). Bearing in mind that there is an opposite lateralization of function in the cerebellum compared to the cerebrum (Hu et al., 2008), the pattern of correlations for task performance accuracy is also quite similar in the cerebellum when compared to the cerebrum (Blokland et al., 2011), with positive correlations between task performance and activity in the visuospatially dominant left hemisphere, and negative correlations between task performance and activity in the verbally dominant right hemisphere. Although the phenotypic correlation between task strategy and cerebellar activity did not survive cluster thresholding, there was a slight trend towards higher activity in the left cerebellar hemisphere with a self-reported spatial task strategy and towards higher activity in the right cerebellar hemisphere with a numerical/verbal strategy. It is possible that participants who use a predominantly verbal (“inner speech”) task strategy are more efficient in their task performance, requiring fewer neural resources to perform the task well. The twin correlations suggested a heritable component to *n*-back task strategy. Thus, an alternative explanation for the observation that MZ co-twins have more similar task-related brain activation is because they are more likely to use the same task strategy than DZ co-twins. The observation that right hemisphere regions, such as the right lobule VIIa Crus I, are more highly heritable could suggest that genetic effects on verbal skills are more heritable. This would be consistent with studies that investigated the heritability of verbal versus visuospatial ability (Plomin et al., 1997; Posthuma et al., 2005), and with a study by Nandagopal et al. (2010) who demonstrated that the estimated heritability of performance on cognitive tasks is mediated, at least in part, by the use of specific cognitive strategies. It should be noted that the number of participants that reported using one strategy over another was relatively small (19.9% spatial versus 28.7% numeric/verbal), so this needs to be investigated further.

Evaluating the difference in BOLD signal between the 0-back and 2-back condition is assumed to exclude the raw motor component from the activation, but another explanation for greater activation of the cerebellum during the 2-back condition compared to the 0-back condition is that it at least partly represents a preparatory response, where the participant is planning their next move. This would be consistent with the faster reaction times during the 2-back condition. Presumably, participants are trying to respond as fast as they can, before they forget the number. Due to its smaller memory requirement, this is less imperative for the 0-back condition. This is an interesting possibility, given the similar heritability of the cerebellum compared to the cerebral cortex, suggesting the influence of genetic factors on planned or anticipated motor responding.

Heritability of cerebellar activation found here (up to 65%, averaging 34% across significant voxels) is highly comparable to the heritability we found for the cerebrum (Blokland et al., 2011), regardless of the different neuronal cell types in the cerebrum (pyramidal cells) and

cerebellum (Purkinje cells). A comparison of the peak heritability regions with the peak activation regions showed that the most consistently activated regions are not necessarily the most heritable regions, probably because some of the group activation local maxima have low variance (Koten et al., 2009). This underlines the importance of voxel-based approaches in fMRI twin studies and has implications for gene finding studies. Gene finding studies have often focused on those areas that are most consistently activated during a task, and areas that show the greatest increase/decrease in activation between cases and controls. Our findings suggest that genetic association studies should move their focus to the areas identified as being most heritable.

Cerebellar BOLD signal did not show significant common/shared environmental influences, in accordance with the repeated finding that the influence of shared environment on cognitive function dissipates in adulthood (Haworth et al., 2010). Prominent unique environmental influences on the cerebellum are consistent with postnatal neurogenesis of cerebellar Purkinje cells, and its preferential susceptibility to insults such as alcohol, lead, or anoxia (Welsh et al., 2002). High 3-month test– retest reliability for *n*-back task-related cerebellar activation, similar to that found in the cerebrum (Blokland et al., 2011), suggests that it is unlikely that the large E variance component reflects only measurement error. Screening for trauma, illness, drug and medication use prior to inclusion allows us to largely rule out these factors as possible unique environmental influences.

A major issue in the interpretation of cerebellar functional imaging studies is that it is still unclear what type of neural activity is reflected in the cerebellar BOLD signal (Diedrichsen et al., 2010). In the cerebral cortex, the BOLD signal appears to reflect mainly postsynaptic activity caused by efferent and recurrent excitatory input to a cortical region (Logothetis et al., 2001), whereas the cerebellar cortex has two input systems: First, mossy fibers project to granule cells, each of which gives rise to a long parallel fiber, which in turn synapses on the dendritic trees of thousands of Purkinje cells. Second, each Purkinje cell also receives strong synaptic input from a single climbing fiber, arising from the inferior olive (Purves et al., 2001). As such, the different neuronal cell types in the cerebrum and cerebellum may be associated with slightly different BOLD mechanisms (Iadecola et al., 1996), perhaps with different genetic underpinnings. We addressed the latter question concerning genetic overlap through bivariate genetic modeling of cerebellar and cerebral BOLD ROI measures.

The two bivariate genetic models for cerebellar and cerebral BOLD percent signal change showed rather striking differences. The first model, representing a correlation on the cognitive level, between right VII Crus I and left middle frontal gyrus, indicated that genetic influences on these phenotypes are specific to each ROI, and that the largest proportion of the covariance between the two ROIs is due to shared environmental influences and/or correlated measurement error. In the second model, representing a correlation on the motor level, most of the phenotypic correlation between the left lobule VI and right precentral gyrus was accounted for by shared genetic influences. Environmental influences were mainly specific to each ROI. A possible explanation for the finding that there is little genetic or environmental overlap between cerebellar and cerebral BOLD signal on the cognition level, while there is a significant genetic correlation between cerebellar and cerebral ROIs on the motor level, is because motor functions and the corresponding brain regions develop earlier in life, and may therefore be more similarly genetically wired. The meaning of the negative path coefficient for the specific cerebellar A-factor is unknown. It indicates that this genetic factor affects the cerebellar ROI in the opposite direction compared to the genetic factor shared with the cerebral ROI. It should be noted that confidence intervals in the bivariate models were wide, and that phenotypic correlations among cerebellum and cerebrum were low, so the actual proportion of variance accounted for by genes shared between the brain's major subdivisions is quite small. Thus, further investigation and

replication of these findings in larger samples are required. However, the present results are encouraging with regard to the usefulness of cerebellar activation during cognition in imaging genetics studies. The significant heritability, partly accounted for by genetic factors specific to the cerebellum, suggests that gene finding studies of the cerebellar BOLD phenotype have the potential to detect novel gene variants, which may be of critical importance to neuropsychiatric disorders in which cerebellar function is affected. A disruption of prefrontal–thalamic–cerebellar circuits is thought to underlie heritable neuropsychiatric disorders such as schizophrenia, autism, and dyslexia (Allen and Courchesne, 2003; Andreasen et al., 1996; Beneventi et al., 2010; Bor et al., 2011). Abnormal cerebellar processing can lead to alterations in mental functions. Cerebellar patients often show cognitive disorders, in addition to mood disorders and personality change. This constellation of symptoms is called cerebellar cognitive affective syndrome (Schmahmann and Sherman, 1998) and may be integrated into the larger pathological frameworks of schizophrenia, autism, dyslexia, and depression (D'Angelo and Casali, 2012). However, clear functional mechanisms that are able to account for cerebellar involvement in these neuropsychiatric disorders remain to be identified. Gene finding studies may provide insight into the functional pathways and lead to a broader understanding of these psychiatric disorders.

Finally, our heritability estimate of total cerebellar GM volume (38%) was lower than those of other studies that extracted cerebellar GM volumes (Kremen et al., 2010; van Soelen et al., 2013), likely because the common environmental factor reached significance, but similar to the average significant heritability estimate for cerebellar activation (34%). However, the low phenotypic correlation between cerebellar activation and cerebellar GM volumes suggests that different neurobiological mechanisms, reflecting a different underlying genetic architecture, may be responsible for the variation in cerebellar volume and cerebellar activation, and is the reason for not testing bivariate genetic models to assess the genetic correlation between activation and GM volume. Further studies are needed to clarify relationships between cerebellar structure and function.

In summary, this is the first fMRI study to investigate cerebellar involvement in spatial and verbal working memory, and its reliability, by normalizing to a spatially unbiased infra-tentorial template (Diedrichsen, 2006). Most importantly, this is the first twin fMRI study to investigate genetic and environmental influences on cerebellar activation. Consistent with earlier work on the cerebrum (Blokland et al., 2011), these genetic cerebellar maps demonstrate a significant influence of genetic factors on working-memory-related cerebellar activation, accounting for up to 65% of the variance. This genetic influence may partly be shared with the cerebrum, and partly be specific to the cerebellum. The possibility that there may be several clusters of genes that play an important role in how the brain retains information is very interesting and merits urgent further investigation.

The next step will be identifying which genetic variants are responsible for individual variation in the cerebellar BOLD response to working memory. Highly heritable areas identified here should facilitate discovery of gene variants influencing cognitive cerebellar function through genome-wide association studies. A greater understanding of the forces that guide normal cerebellar functioning may shed light on the genetic liability to brain disorders affecting the cerebellum. It should also open up new avenues for developing and implementing more effective diagnosis and treatment of brain disorders.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This study was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development, USA, and National Health and Medical Research Council (NHMRC), Australia. The collection of IQ data and zygosity typing were supported by the Australian Research Council (ARC). G.A.M.B. is supported by an ANZ Trustees PhD Scholarship in Medical Research, Queensland, Australia. G.I.Z. is supported by an ARC Future Fellowship. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, The National Institutes of Health, NHMRC, or ARC. We are very grateful to the twins for their generosity of time and willingness to participate in our studies. We thank research nurses Marlene Grace and Ann Eldridge for twin recruitment, research assistants Kori Johnson, Lachlan Strike, Angus Wallace, Aaron Quiggle, and Natalie Garden and radiographers Matthew Meredith, Peter Hobden, Kate Borg, Aiman Al Najjar, and Anita Burns for data acquisition, and Daniel Park for IT support.

## References

Allen G, Courchesne E. Differential effects of developmental cerebellar abnormality on cognitive and motor functions in the cerebellum: an fMRI study of autism. Am. J. Psychiatry. 2003; 160(2):262–273. [PubMed: 12562572]

Andreasen NC, O'Leary DS, Cizadlo T, Arndt S, Rezai K, Ponto LL, Watkins GL, Hichwa RD. Schizophrenia and cognitive dysmetria: a positron-emission tomography study of dysfunctional prefrontal–thalamic–cerebellar circuitry. Proc. Natl. Acad. Sci. U. S. A. 1996; 93(18):9985–9990. [PubMed: 8790444]

Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage. 2007; 38(1):95–113. [PubMed: 17761438]

Beneventi H, Tonnessen FE, Ersland L, Hugdahl K. Working memory deficit in dyslexia: behavioral and FMRI evidence. Int. J. Neurosci. 2010; 120(1):51–59. [PubMed: 20128672]

Ben-Yehudah G, Guediche S, Fiez JA. Cerebellar contributions to verbal working memory: beyond cognitive theory. Cerebellum. 2007; 6(3):193–201. [PubMed: 17786815]

Bertolino A, Caforio G, Petruzzella V, Latorre V, Rubino V, Dimalta S, Torraco A, Blasi G, Quartesan R, Mattay VS, Callicott JH, Weinberger DR, Scarabino T. Prefrontal dysfunction in schizophrenia controlling for COMT Val158Met genotype and working memory performance. Psychiatry Res. 2006; 147(2–3):221–226. [PubMed: 16952445]

Bertolino A, Taurisano P, Pisciotta NM, Blasi G, Fazio L, Romano R, Gelao B, Lo Bianco L, Lozupone M, Di Giorgio A, Caforio G, Sambataro F, Niccoli-Asabella A, Papp A, Ursini G, Sinibaldi L, Popolizio T, Sadee W, Rubini G. Genetically determined measures of striatal D2 signaling predict prefrontal activity during working memory performance. PLoS One. 2010; 5(2):e9348. [PubMed: 20179754]

Blokland GAM, McMahon KL, Hoffman J, Zhu G, Meredith M, Martin NG, Thompson PM, de Zubicaray GI, Wright MJ. Quantifying the heritability of task-related brain activation and performance during the N-back working memory task: a twin fMRI study. Biol. Psychol. 2008; 79(1):70–79. [PubMed: 18423837]

Blokland GAM, McMahon KL, Thompson PM, Martin NG, de Zubicaray GI, Wright MJ. Heritability of working memory brain activation. J. Neurosci. 2011; 31(30):10882–10890. [PubMed: 21795540]

Bor J, Brunelin J, Sappey-Marinier D, Ibarrola D, d'Amato T, Suaud-Chagny MF, Saoud M. Thalamus abnormalities during working memory in schizophrenia. An fMRI study. Schizophr. Res. 2011; 125(1):49–53.

Brett M, Anton J-L, Valabregue R, Poline J-B. Region of interest analysis using an SPM toolbox, 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan. Neuroimage. 2002; 16(2) abstract 497.

Callicott JH, Egan MF, Mattay VS, Bertolino A, Bone AD, Verchinksi B, Weinberger DR. Abnormal fMRI response of the dorsolateral prefrontal cortex in cognitively intact siblings of patients with schizophrenia. Am. J. Psychiatry. 2003; 160(4):709–719. [PubMed: 12668360]

Chein JM, Fiez JA. Dissociation of verbal working memory system components using a delayed serial recall task. Cereb. Cortex. 2001; 11(11):1003–1014. [PubMed: 11590110]

Chen SH, Desmond JE. Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. Neuroimage. 2005; 24(2):332–338. [PubMed: 15627576]

Côté C, Beauregard M, Girard A, Mensour B, Mancini-Marie A, Perusse D. Individual variation in neural correlates of sadness in children: a twin fMRI study. Hum. Brain Mapp. 2007; 28(6):482–487. [PubMed: 17437293]

D'Angelo E, Casali S. Seeking a unified framework for cerebellar function and dysfunction: from circuit operations to cognition. Front. Neural Circuits. 2012; 6116

Diedrichsen J. A spatially unbiased atlas template of the human cerebellum. Neuroimage. 2006; 33(1): 127–138. [PubMed: 16904911]

Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N. A probabilistic MR atlas of the human cerebellum. Neuroimage. 2009; 46(1):39–46. [PubMed: 19457380]

Diedrichsen J, Verstynen T, Schlerf J, Wiestler T. Advances in functional imaging of the human cerebellum. Curr. Opin. Neurol. 2010; 23(4):382–387. [PubMed: 20581682]

Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, Goldman D, Weinberger DR. Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. Proc. Natl. Acad. Sci. U. S. A. 2001; 98(12):6917–6922. [PubMed: 11381111]

Eickhoff SB, Heim S, Zilles K, Amunts K. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage. 2006; 32(2):570–582. [PubMed: 16781166]

Falconer, DS.; Mackay, TFC. Introduction to Quantitative Genetics. Longmans Green; Harlow, Essex, UK: 1996.

Freire L, Roche A, Mangin JF. What is the best similarity measure for motion correction in fMRI time series? IEEE Trans. Med. Imaging. 2002; 21(5):470–484. [PubMed: 12071618]

Gottwald B, Wilde B, Mihajlovic Z, Mehdorn HM. Evidence for distinct cognitive deficits after focal cerebellar lesions. J. Neurol. Neurosurg. Psychiatry. 2004; 75(11):1524–1531. [PubMed: 15489381]

Haworth CM, Wright MJ, Luciano M, Martin NG, de Geus EJ, van Beijsterveldt CE, Bartels M, Posthuma D, Boomsma DI, Davis OS, Kovas Y, Corley RP, Defries JC, Hewitt JK, Olson RK, Rhea SA, Wadsworth SJ, Iacono WG, McGue M, Thompson LA, Hart SA, Petrill SA, Lubinski D, Plomin R. The heritability of general cognitive ability increases linearly from childhood to young adulthood. Mol. Psychiatry. 2010; 15(11):1112–1120. [PubMed: 19488046]

Hayter AL, Langdon DW, Ramnani N. Cerebellar contributions to working memory. Neuroimage. 2007; 36(3):943–954. [PubMed: 17468013]

Hokkanen LS, Kauranen V, Roine RO, Salonen O, Kotila M. Subtle cognitive deficits after cerebellar infarcts. Eur. J. Neurol. 2006; 13(2):161–170. [PubMed: 16490047]

Holm S. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 1979; 6(2):65–70.

Hu D, Shen H, Zhou Z. Functional asymmetry in the cerebellum: a brief review. Cerebellum. 2008; 7(3):304–313. [PubMed: 18418670]

Iadecola C, Li J, Xu S, Yang G. Neural mechanisms of blood flow regulation during synaptic activity in cerebellar cortex. J. Neurophysiol. 1996; 75(2):940–950. [PubMed: 8714666]

Jackson, DN. Multidimensional Aptitude Battery. SIGMA Assessment Systems Inc; 1984.

Kelly RM, Strick PL. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J. Neurosci. 2003; 23(23):8432–8444. [PubMed: 12968006]

Koten JW Jr. Wood G, Hagoort P, Goebel R, Propping P, Willmes K, Boomsma DI. Genetic contribution to variation in cognitive function: an fMRI study in twins. Science. 2009; 323(5922): 1737–1740. [PubMed: 19325117]

Kremen WS, Prom-Wormley E, Panizzon MS, Eyler LT, Fischl B, Neale MC, Franz CE, Lyons MJ, Pacheco J, Perry ME, Stevens A, Schmitt JE, Grant MD, Seidman LJ, Thermenos HW, Tsuang MT, Eisen SA, Dale AM, Fennema-Notestine C. Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. Neuroimage. 2010; 49(2):1213–1223. [PubMed: 19786105]

Krug A, Markov V, Eggermann T, Krach S, Zerres K, Stocker T, Shah NJ, Schneider F, Nothen MM, Treutlein J, Rietschel M, Kircher T. Genetic variation in the schizophrenia-risk gene neuregulin1 correlates with differences in frontal brain activation in a working memory task in healthy individuals. Neuroimage. 2008; 42(4):1569–1576. [PubMed: 18606232]

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature. 2001; 412(6843):150–157. [PubMed: 11449264]

Macey PM, Macey KE, Kumar R, Harper RM. A method for removal of global effects from fMRI time series. Neuroimage. 2004; 22(1):360–366. [PubMed: 15110027]

Massat I, Slama H, Kavec M, Linotte S, Mary A, Baleriaux D, Metens T, Mendlewicz J, Peigneux P. Working memory-related functional brain patterns in never medicated children with ADHD. PLoS One. 2012; 7(11):e49392. [PubMed: 23166657]

Matthews SC, Simmons AN, Strigo I, Jang K, Stein MB, Paulus MP. Heritability of anterior cingulate response to conflict: an fMRI study in female twins. Neuroimage. 2007; 38(1):223–227. [PubMed: 17707125]

McMillan J, Beavis A, Jones FL. The AUSEI06: a new socioeconomic index for Australia. J. Sociol. 2009; 45(2):123–149.

Middleton FA, Strick PL. Cerebellar projections to the prefrontal cortex of the primate. J. Neurosci. 2001; 21(2):700–712. [PubMed: 11160449]

Nandagopal K, Roring RW, Ericsson KA, Taylor J. Strategies may mediate heritable aspects of memory performance: a twin study. Cogn. Behav. Neurol. 2010; 23(4):224–230. [PubMed: 21042210]

Neale, MC.; Maes, HHM. Methodology for Genetic Studies of Twins and Families. Kluwer Academic Publishers B.V.; Dordrecht, The Netherlands: 2004.

Neale, MC.; Baker, SM.; Xie, G.; Maes, HH. Mx: Statistical Modeling. Richmond, VA 23298. Department of Psychiatry, University of Virginia; VCU Box 900126: 2002.

Owen AM, McMillan KM, Laird AR, Bullmore E. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. 2005; 25(1):46–59. [PubMed: 15846822]

Park J, Shedden K, Polk TA. Correlation and heritability in neuroimaging datasets: a spatial decomposition approach with application to an fMRI study of twins. Neuroimage. 2012; 59(2): 1132–1142. [PubMed: 21763433]

Plomin R, Fulker DW, Corley R, DeFries JC. Nature, nurture, and cognitive development from 1 to 16 years: a parent-offspring adoption study. Psychol. Sci. 1997; 8(6):442–447.

Polk TA, Park J, Smith MR, Park DC. Nature versus nurture in ventral visual cortex: a functional magnetic resonance imaging study of twins. J. Neurosci. 2007; 27(51):13921–13925. [PubMed: 18094229]

Posthuma D, Luciano M, Geus EJ, Wright MJ, Slagboom PE, Montgomery GW, Boomsma DI, Martin NG. A genomewide scan for intelligence identifies quantitative trait loci on 2q and 6p. Am. J. Hum. Genet. 2005; 77(2):318–326. [PubMed: 16001363]

Purves, D.; Augustine, GJ.; Fitzpatrick, D.; Katz, LC.; Lamantia, A-S.; McNamara, JO.; Williams, SM., editors. Neuroscience. Sinauer Associates; Sunderland, MA: 2001.

Ramnani N. The primate cortico-cerebellar system: anatomy and function. Nat. Rev. Neurosci. 2006; 7(7):511–522. [PubMed: 16791141]

Ravizza SM, Delgado MR, Chein JM, Becker JT, Fiez JA. Functional dissociations within the inferior parietal cortex in verbal working memory. Neuroimage. 2004; 22(2):562–573. [PubMed: 15193584]

Ravizza SM, McCormick CA, Schlerf JE, Justus T, Ivry RB, Fiez JA. Cerebellar damage produces selective deficits in verbal working memory. Brain. 2006; 129(Pt 2):306–320. [PubMed: 16317024]

Schmahmann JD, Sherman JC. The cerebellar cognitive affective syndrome. Brain. 1998; 121(Pt 4): 561–579. [PubMed: 9577385]

Scott RB, Stoodley CJ, Anslow P, Paul C, Stein JF, Sugden EM, Mitchell CD. Lateralized cognitive deficits in children following cerebellar lesions. Dev. Med. Child Neurol. 2001; 43(10):685–691. [PubMed: 11665825]

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 1979; 86(2):420–428. [PubMed: 18839484]

Stoodley CJ, Schmahmann JD. Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. Neuroimage. 2009; 44(2):489–501. [PubMed: 18835452]

Strick PL, Dum RP, Fiez JA. Cerebellum and nonmotor function. Annu. Rev. Neurosci. 2009:32413–434.

Valera EM, Faraone SV, Biederman J, Poldrack RA, Seidman LJ. Functional neuroanatomy of working memory in adults with attention-deficit/hyperactivity disorder. Biol. Psychiatry. 2005; 57(5):439–447. [PubMed: 15737657]

van Soelen ILC, Brouwer RM, van Baal GCM, Schnack HG, Peper JS, Chen L, Kahn RS, Boomsma DI, Hulshoff Pol HE. Heritability of volumetric brain changes and height in children entering puberty. Hum. Brain Mapp. 2013; 34(3):713–725. [PubMed: 22140022]

Welsh JP, Yuen G, Placantonakis DG, Vu TQ, Haiss F, O'Hearn E, Molliver ME, Aicher SA. Why do Purkinje cells die so easily after global brain ischemia? Aldolase C, EAAT4, and the cerebellar contribution to posthypoxic myoclonus. Adv. Neurol. 2002:89331–359.

Wright MJ, Martin NG. The Brisbane Adolescent Twin Study: outline of study methods and research projects. Aust. J. Psychol. 2004:5665–78.

Zeng H, Constable RT. Image distortion correction in EPI: comparison of field mapping with point spread function mapping. Magn. Reson. Med. 2002; 48(1):137–146. [PubMed: 12111941]

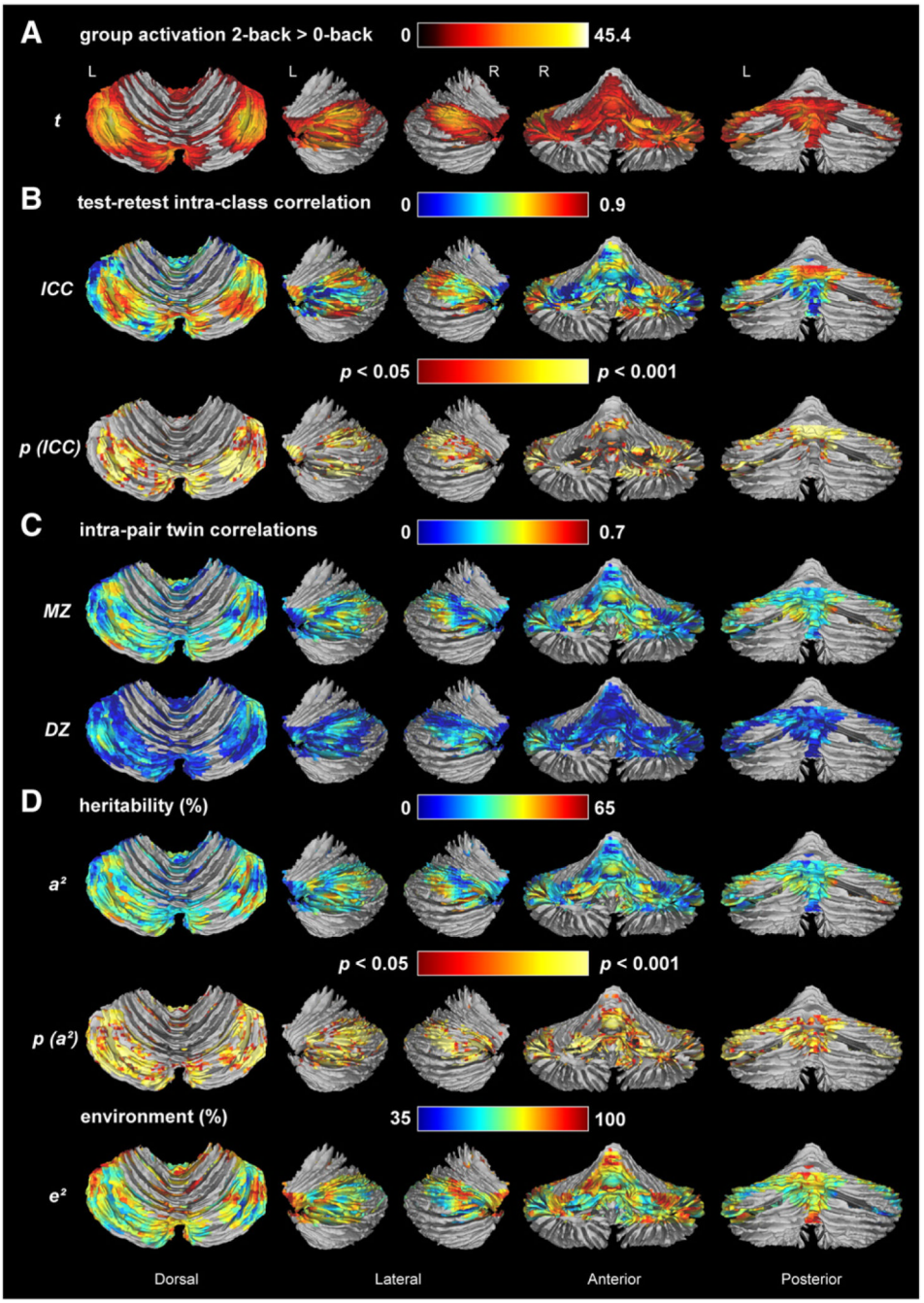


**Fig. 1.**

*N*-back task-related group cerebellar activation (A), test–retest reliability (B), twin correlations (C), and variance component estimates for *n*-back task-related cerebellar activation (D). (A) Group random effects analysis for the 2-back > 0-back t-contrast ($p<0.05$, FWE-corrected); (B) intra-class correlations and corresponding height- ($p<0.05$) and cluster- (>540 voxels) thresholded *p*-values for the group activation; (C) intra-pair MZ and DZ twin correlations; (D) percentage of variance explained by genetic ($a^2$) and unique environmental factors ($e^2$) factors and probability map for $a^2$, indicating which genetic estimates were significant after height- ($p<0.05$) and cluster- (>540 voxels) thresholding. Twin correlations and variance component estimates were estimated using maximum

likelihood procedures and corrected for sex, age, 2-back performance accuracy and RT. Green, yellow, and red areas are those with the highest twin correlations and heritability. Assumption testing supported homogeneity of means across birth order for ~86% of voxels, homogeneity of variances across birth order for 100% of voxels, homogeneity of means across zygosity for ~90% of voxels and homogeneity of variances across zygosity for ~96% of voxels. There were no significant mean or variance differences between twins and siblings. Therefore, sibling pairs were included in the DZ group.

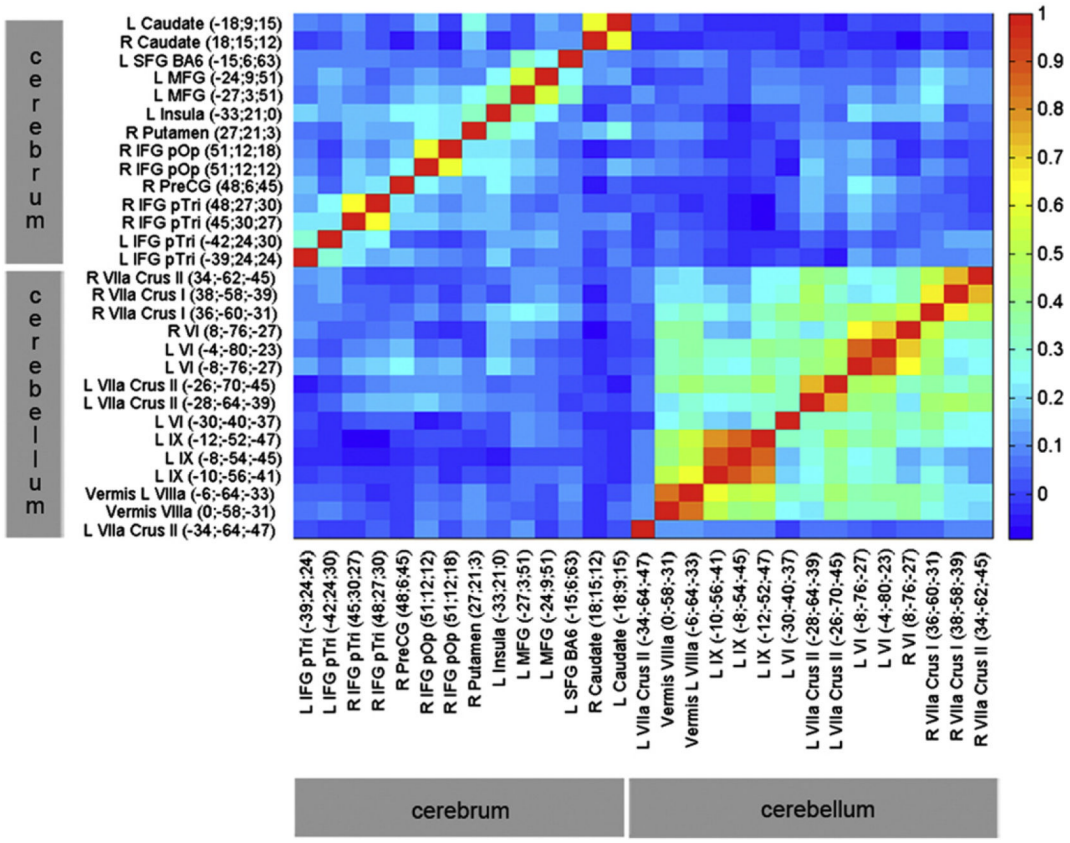


**Fig. 2.**

Heat map of phenotypic correlations among reliable BOLD ROI measures. Phenotypes that are most highly positively correlated are shown in orange–red. Specific r values are presented in Supplementary Table 2. Correlations within the cerebellum are much higher than correlations within the cerebrum, most likely due to the closer spatial proximity. Correlations are highest among equivalent contralateral ROIs, and among ipsilateral ROIs from the same anatomical region. Correlations are the highest among cerebral areas with other cerebral areas, and among cerebellar areas with other cerebellar areas. Overall, there is a stronger correlation between left cerebellar and right cerebral regions, and between right cerebellar and left cerebral regions, consistent with a cross-over between hemispheres for the cerebrum, but not the cerebellum. The highest correlations between the cerebellum and the cerebrum are between right lobule VIIa Crus I and left middle frontal gyrus (0.27), and between left lobule VI and right precentral gyrus (0.24).

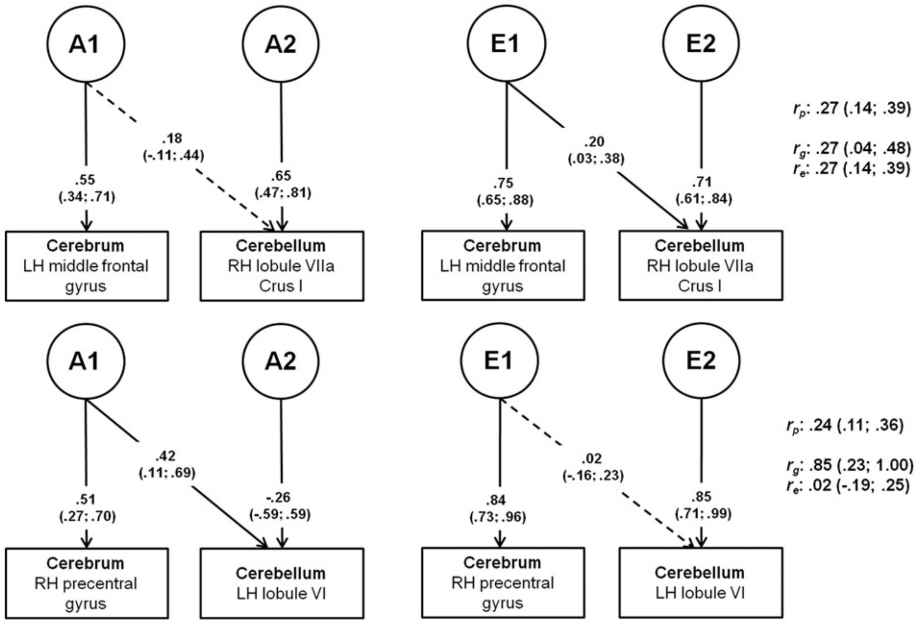


**Fig. 3.**

Cholesky decomposition for bivariate AE models for cerebral and cerebellar ROIs. Path coefficients (unstandardized) are displayed on their respective paths with the 95% confidence interval in parentheses. Solid paths indicate significant path coefficients; dotted paths indicate non-significant path coefficients. A1= first genetic factor (partly shared between cerebrum and cerebellum); A2 = second genetic factor (specific to cerebellum); E1 = first environmental factor (partly shared between cerebrum and cerebellum); E2 = second environmental factor (specific to cerebellum); LH = left hemisphere; RH = right hemisphere; $r_p$ = phenotypic correlation; $r_g$ = genetic correlation; $r_e$ = environmental correlation.

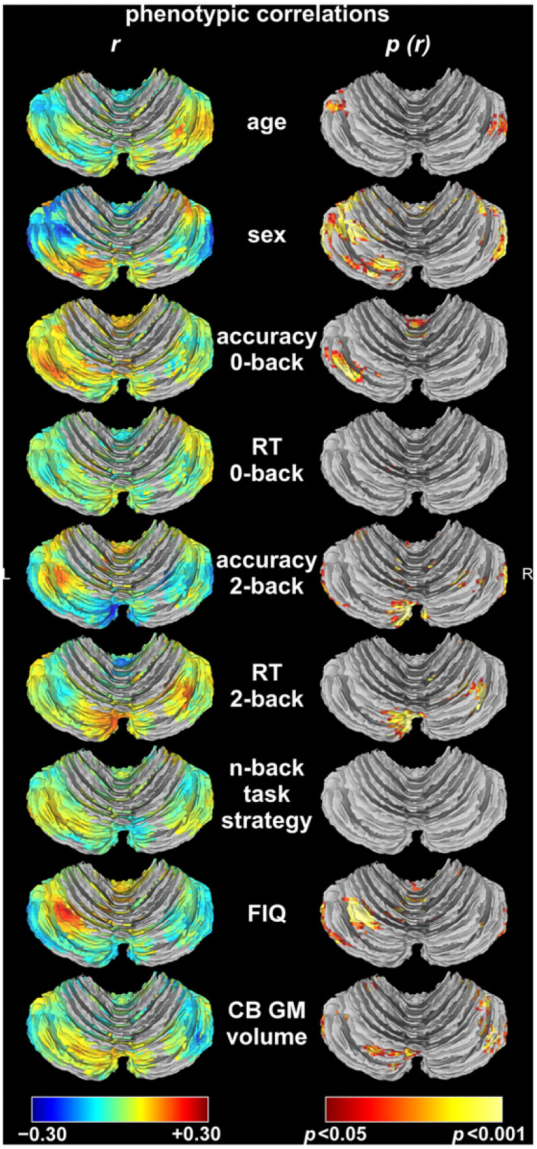


**Fig. 4.**

Phenotypic correlations with cerebellar activation. The first column shows the strength of the correlation (*r*). Positive correlations, i.e., greater activation in older participants, in males, in participants with higher task performance, slower response time (RT), with spatial task strategy as opposed to numerical task strategy, with higher full-scale intelligence quotient (FIQ), or larger cerebellar (CB) gray matter (GM) volumes, are represented by hot colors and negative correlations, i.e., greater activation in younger participants, in females, in participants with lower task performance, faster RT, with numerical as opposed to spatial task strategy, with lower FIQ, or smaller cerebellar GM volumes, are represented by cold colors. The second column shows height- ($p < 0.05$) and cluster-thresholded *p*-value maps corresponding to the correlations (p (*r*)). Based on the smoothness of the correlation maps, which was estimated at 5.79 × 4.45 × 4.09 mm for age, 7.16 × 4.76 × 3.92 mm for sex, 5.53 × 4.43 × 4.12 mm for RT on 0-back, 7.16 × 6.14 × 5.97 mm for RT on 2-back, 6.60 × 4.89 × 4.30 mm for accuracy on 0-back, 6.84 × 5.61 × 4.79 mm for accuracy on 2-back, 7.50 × 6.07 × 5.04 mm for FIQ, 6.39 × 6.35 × 5.86 mm for *n*-back task strategy, and 7.25 × 5.75 × 4.80 mm for cerebellar GM volume, Monte Carlo simulation determined a minimum cluster size

of 167 voxels, 193 voxels, 161 voxels, 275 voxels, 194 voxels, 225 voxels, 258 voxels, 258 voxels, and 237 voxels, respectively, for FWE corrected cluster significance at $p < 0.05$. The phenotypic correlations between *n*-back task strategy (spatial relative to numerical) and cerebellar activation did not survive cluster thresholding.

## Table 1

Areas of significant phenotypic correlation with cerebellar activation.[b]

| Phenotype | Positive correlation with activation | | | | | | Negative correlation with activation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cluster size in voxels | Peak MNI[a] coordinates | | | Lobule(s) | Mean *r* (range) | Cluster size in voxels | Peak MNI coordinates | | | Lobule(s) | Mean *r* (range) |
| | | x | y | z | | | | x | y | z | | |
| Age | 173 | 58 | –66 | –33 | R VIIa Crus I | 0.12 (0.10, 0.21) | 584 | –38 | –62 | –47 | L VIIb; L VIIa Crus II; L VIIa Crus I; L VIIIa | –0.14 (–0.10, –0.27) |
| Sex (males relative to females) | 633 | –44 | –60 | –47 | L VIIa Crus II; L VIIIa; L X; L VIIb | 0.14 (0.10, 0.24) | 3428 | –56 | –60 | –37 | L VIIa Crus I; R VIIa Crus I; R IX; L VI; Vermis R IX; Vermis R VIIIb | –0.15 (–0.10, –0.28) |
| | 522 | –24 | –82 | –25 | L VIIa Crus I; L VI | | | | | | | |
| | 147 | 42 | –44 | –31 | R VI; R V; R I-IV | | | | | | | |
| | 116 | –8 | –76 | –41 | L VIIIa; R VIIIa; R VIIb; L VIIa Crus II; Vermis R VIIIb; Vermis L VIIIa | | | | | | | |
| 0-Back accuracy | 481 | –38 | –72 | –25 | L VIIa Crus I; L VI | 0.12 (0.10, 0.18) | - | - | - | - | - | - |
| | 210 | –4 | –48 | –7 | L I-IV; R I-IV | | | | | | - | - |
| 0-Back response time | 222 | 34 | –60 | –39 | R VIIa Crus I/II; R VIIIa; Vermis R VIIIa; Vermis L VIIIa; Vermis L VIIb | 0.12 (0.10, 0.17) | - | - | - | - | - | - |
| | 196 | –10 | –62 | –31 | L IX; L VIIIb | | | | | | | |
| 2-Back accuracy | 575 | –32 | –66 | –47 | L VIIb; L VIIIb; L VIIIa; LX | 0.14 (0.10, 0.23) | 1306 | 30 | –66 | –35 | R VIIa Crus I; R VIIIa; R VI; L I-IV | –0.14 (–0.10, –0.26) |
| | 316 | –2 | –58 | –27 | Vermis L IX; Vermis R IX; L V | | 536 | –4 | –82 | –29 | Vermis L VIIb; Vermis L VIIa Crus II; R VIIa Crus II; L VI; R VIIa Crus I; L VIIa Crus II | |
| | | | | | | | 294 | –56 | –70 | –35 | R VIIa Crus I; R VIIa Crus II; L VIIa Crus I | |
| 2-Back response time | 1840 | 34 | –64 | –29 | R VIIa Crus I; Vermis L VI; Vermis L VIIb; Vermis L VIIIa | 0.14 (0.10, 0.24) | - | - | - | - | - | - |
| Full-scale IQ | 1478 | –30 | –60 | –23 | L VI; Vermis R IX | 0.14 (0.10, 0.25) | 578 | 42 | –62 | –43 | R VIIa Crus II; R VIIa Crus I | –0.14 (–0.10, –0.23) |

| Phenotype | Positive correlation with activation | | | | | | Negative correlation with activation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cluster size in voxels | Peak MNI[a] coordinates | | | Lobule(s) | Mean *r* (range) | Cluster size in voxels | Peak MNI coordinates | | | Lobule(s) | Mean *r* (range) |
| | | x | y | z | | | | x | y | z | | |
| | 680 | –18 | –44 | –43 | L X; L IX; L VIIIb; L VIIIa; L VIIa Crus II | | 408 | –58 | –62 | –41 | L VIIa Crus I; L VIIa Crus II | |
| Cerebellar gray matter volume | 799 | –18 | –50 | –47 | L VIIIb; L VIIIa; L IX; L VIIb | 0.13 (0.10, 0.21) | 948 | 44 | –66 | –29 | R VIIa Crus I; R VI | –0.13 (–0.10, –0.26) |
| | | | | | | | 325 | –42 | –58 | –41 | L VIIa Crus I; L VIIa Crus II | |
| | | | | | | | 282 | –22 | –30 | –29 | L I-IV and V | |

[a] Abbreviations: L, left hemisphere; MNI, Montréal Neurological Institute; R, right hemisphere.

[b] The phenotypic correlations between *n*-back task strategy (spatial relative to numerical) and activation did not survive cluster thresholding.