HHS Public Access

Author manuscript

Genet Epidemiol. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

Genet Epidemiol. 2018 June ; 42(4): 405–414. doi:10.1002/gepi.22122.

Genetic associations with childhood brain growth, defined in two longitudinal cohorts

Eszter Szekely, PhD[1,2,*], **Tae-Hwi Linus Schwantes-An, PhD**[3,4,*], **Cristina M. Justice, PhD**[3,*], **Jeremy A. Sabourin, PhD**[3,*], **Philip R. Jansen, MD**[5,6], **Ryan L. Muetzel, MSc**[5,6], **Wendy Sharp, MSc**[1], **Henning Tiemeier, MD, PhD**[5,6,7], **Heejong Sung, PhD**[3], **Tonya J. White, MD, PhD**[5,6,†], **Alexander F. Wilson, PhD**[3,†], and **Philip Shaw, MD, PhD**[1,‡,†]

[1]Section on Neurobehavioral Clinical Research, Social and Behavioral Research Branch, National Human Genome Research Institute, Bethesda, MD, USA [2]Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, QC, Canada [3]Genometrics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, Baltimore, MD, USA [4]Department of Medical and Molecular Genetics, Indiana University School of Medicine, IN, USA [5]The Generation R Study Group, Erasmus Medical Center, Rotterdam, The Netherlands [6]Department of Child and Adolescent Psychiatry/Psychology, Sophia Children's Hospital—Erasmus Medical Center, Rotterdam, The Netherlands [7]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands

Abstract

Genome-wide association studies (GWAS) are unraveling the genetics of adult brain neuroanatomy as measured by cross-sectional anatomic magnetic resonance imaging (aMRI). However, the genetic mechanisms that shape childhood brain development are, as yet, largely unexplored. In this study we identify common genetic variants associated with childhood brain development as defined by longitudinal aMRI.

Genome-wide SNP data were determined in 2 cohorts: one enriched for attention-deficit/hyperactivity disorder (ADHD) (*LONG* cohort: 458 participants; 119 with ADHD) and the other from a population-based cohort (*Generation R*: 257 participants). The growth of the brain's major regions (cerebral cortex, white matter, basal ganglia and cerebellum) and one region of interest (the right lateral prefrontal cortex) were defined on all individuals from two aMRIs, and a genome-wide association study and a pathway analysis were performed. In addition, association between polygenic risk for ADHD and brain growth was determined for the LONG cohort.

For white matter growth, GWAS meta-analysis identified a genome-wide significant intergenic SNP (rs12386571, $p = 9.09 \times 10^{-9}$), near *AKR1B10*. This gene is part of the aldo-keto reductase superfamily and shows neural expression. No enrichment of neural pathways was detected and polygenic risk for ADHD was not associated with the brain growth phenotypes in the LONG

[‡]To whom correspondence should be addressed at: Blgd 31, B1B39, Bethesda 20892, Maryland, USA; shawp@mail.nih.gov.
[*]These authors contributed equally to this work.
[†]These authors contributed equally to this work.

Financial disclosure statement: All authors reported no biomedical financial interests or potential conflicts of interest.

cohort that was enriched for the diagnosis of ADHD. The study illustrates the use of a novel brain growth phenotype defined *in vivo* for further study.

Keywords

GWAS; pathway analysis; polygenic score; brain development; ADHD; children

INTRODUCTION

Cross-sectional studies show that variation in human brain structure, defined at a macroscopic level by anatomic magnetic resonance imaging (aMRI), is highly heritable with heritability estimates ~0.9 for the volumes of the prefrontal cortex and striatum (Neale et al., 2010; Peper, Brouwer, Boomsma, Kahn, & Hulshoff Pol, 2007; Wallace et al., 2006). Progress has been made in identifying common genetic variants associated with brain structure in adults (Bis et al., 2012; Hibar et al., 2013; Melville et al., 2012; Stein et al., 2012; Walters et al., 2013) but genetic factors involved in the developing brain are less well explored. Understanding how common genetic variation impacts brain growth is both inherently important and also has translational implications because neurodevelopmental disorders with a childhood-onset such attention-deficit/hyperactivity disorder (ADHD) are partly characterized by perturbed brain growth (F. X. Castellanos et al., 2002; Eric Courchesne, Campbell, & Solso, 2011; E. Courchesne, Carper, & Akshoomoff, 2003; Eric Courchesne & Pierce, 2005; Di Martino et al., 2014; Ducharme et al., 2012; El-Sayed, Larsson, Persson, Santosh, & Rydelius, 2003; Mackie et al., 2007; Redcay & Courchesne, 2005; Rubia, 2007; Shaw et al., 2011; Shaw, Gogtay, & Rapoport, 2010; Sripada, Kessler, & Angstadt, 2014). For example, the severity of inattention, impulsivity and hyperactivity symptoms has been associated with altered development of the cerebral cortex and basal ganglia in children and adolescents (F. X. Castellanos et al., 2002; Ducharme et al., 2012; Shaw et al., 2014; Shaw et al., 2011). Thus, uncovering the genetic pathways that impact brain development in childhood might also inform our understanding of psychiatric disorders with a neurodevelopmental origin.

Here, we attempted to identify the common genetic variants associated with brain growth. Two independent cohorts were studied—the Longitudinal Observations in Neuroimaging and Genetics (*LONG*) cohort at the NIH (with 458 participants) and data from the *Generation R* Study in Rotterdam (with 257 participants). All participants in both cohorts had two neuroanatomic scans which were used to determine brain growth. Age-related volumetric changes in the major regions of the brain (cerebral cortex and white matter, basal ganglia, and cerebellum) were defined. Longitudinal twin imaging studies have demonstrated heritability in the growth of these brain regions, ranging from 19% for white matter growth, to 37% for cerebral gray matter and 40% for the cerebellum (Brouwer et al., 2017). In addition, growth of the right lateral prefrontal cortex was also examined. This region was chosen because both cross-sectional and longitudinal studies tie age-related change in the right lateral prefrontal cortex to the severity of hyperactivity-impulsivity and inattention in all children, not just those with the diagnosis (Shaw, 2011; Ducharme, 2012). Right prefrontal volume loss is one of the more consistently replicated findings in cross-

sectional neuroimaging studies in ADHD and is tied to the atypical development of prefrontal lateralized processing seen in the disorder.

Our primary aim was to identify associations between Single Nucleotide Polymorphisms (SNPs) and the growth of five major brain regions in the *LONG* and *Generation R* cohorts. Secondly, we determined if the biological pathways implicated through SNPs or a child's aggregate (polygenic) risk for ADHD showed associations with brain growth.

METHODS

Subjects

'LONG cohort'—The *LONG* cohort includes a total of 458 participants (271 males), with 339 typically developing children and adolescents, and 119 youth with an ADHD diagnosis. Diagnosis was ascertained using the clinician-administered Parent Diagnostic Interview for Children and Adolescents (Reich, 2000). All participants had two aMRIs. The mean age of the first scan was 11.47 (SD 3.54) years and the mean age of the second scan was 16.13 (SD 4.72) years. Based on parent-reported race/ethnicity, there were 404 European Americans, 31 African Americans, 8 Asian Americans and 15 participants of mixed race. The sample included 146 sibling pairs (137 pairs of full siblings, 9 pairs of half siblings).

'Generation R'—The second cohort was a subsample of the *Generation R* Study, a population-based Dutch birth cohort (Kooijman et al., 2016), consisting of the 257 participants (135 males) with two aMRIs (White et al., 2013). The mean age of the participants was 7.71 (SD 0.86) years at the time of the first scan and 10.23 (SD 0.68) years at the second scan. Given its community-based nature, most children (n=244) were typically developing. Clinically significant levels of ADHD symptoms were identified using the Attention Problems and DSM-oriented ADHD Problems scales of the Child Behavior Checklist (CBCL) 1.5–5 (Achenbach & Rescorla, 2000), at a mean age 6.03 (SD 0.4). The preschool version of the CBCL was used because many of the children were younger than 6 years at the time of assessment. Using clinical thresholds established on Dutch norm populations there were 13 children who scored above the clinical threshold on attention problems or DSM-oriented ADHD problems (Tick, van der Ende, Koot, & Verhulst, 2007). Based on parent-reported race/ethnicity, 177 children were White/Caucasian and 80 non-Caucasian (these participants were of various ethnicities, such as African, primarily North African, Asian, and Caribbean).

The Institutional Review Board of each institute approved the study protocols at each site (the National Institutes of Health for the *LONG* cohort, and the Medical Ethics Committee of the Erasmus Medical Center for the *Generation R* Study). Written informed consent to participate in the study was obtained from parents.

Genotyping

LONG cohort—DNA was extracted from lymphoblastoid cell line (234 samples), saliva (124 samples), and blood (100 samples). Genotyping and initial quality control were performed at Center for Inherited Disease Research (http://www.cidr.jhmi.edu/) with the Illumina HumanOmniExpressExome-8v1-2 array. Initially, 959,200 SNPs were released.

After quality control procedures, 668,419 SNPs were available for analyses. Ungenotyped markers surrounding significant regions were imputed to facilitate fine mapping (see supplementary materials and methods).

Generation R—Genotype data from *Generation R* was available for 518,245 SNPs. Before imputation the *Generation R* and *LONG* cohort had 287,902 SNPs in common. SNPs from *Generation* R were imputed (to SNPs from the *LONG* cohort) to increase the SNP-overlap. Imputation resulted in an additional 365,128 SNPs with a quality mach.rsq > 0.3 in *Generation R* for final analysis, for a total of 653,030 SNPs for meta-analysis. Genotyping, quality control procedures and imputation analyses are described in detail elsewhere (Medina-Gomez et al., 2015).

The brain growth phenotypes

LONG cohort—Neuroanatomical MRI was acquired on 370 participants on a 1.5-T GE Signa scanner (Milwaukee, WI) at both baseline and follow up. Sixty-eight participants had aMRI acquired on a 3-T GE Signa scanner (Milwaukee, WI) at both time points. Twenty children were scanned on a 1.5-T scanner at baseline and a 3-T scanner at follow-up. Imaging parameters are given as supplemental information.

Generation R—Baseline MR images were acquired on a 3-T GE Discovery MR-750 scanner (Milwaukee, WI) and the second scan on a 3-T GE Discovery MR-750w scanner. Imaging parameters are provided as supplemental information.

Image preprocessing—Automatic volumetric segmentations were performed with the FreeSurfer image suite (version 5.3; https://surfer.nmr.mgh.harvard.edu). For quality assurance, two raters assessed the quality of all the initial images, blind to diagnosis and other demographic details, retaining only those judged to have no or minimal motion or other artifacts (determined using published guidelines (F. Castellanos et al., 2002)). The final segmentations of the cortical surface and mantle and deeper structures, provided by FreeSurfer, were also visually inspected by trained raters following published guidelines used by the Enhancing Imaging Genomics through Meta-Analysis consortium (http://enigma.ini.usc.edu/protocols/imaging-protocols/). At each stage, discrepant ratings of either the initial image or its final segmentations were reviewed and a consensus rating was given. Finally, a participant was retained only if both scans met these quality criteria. Because of these measures, 338 of the initial 1254 scans in the LONG cohort were excluded along with 378 of the original 892 in *Generation R*. Thus, the final analyses were comprised of 916 scans in the *LONG* cohort and 514 in *Generation R*.

Modeling growth

Growth was modeled for four major divisions of the brain (the cerebral cortex, the basal ganglia, the cerebellum, cerebral white matter) and the region of interest (the right lateral prefrontal cortex). Volume was used as the measure because it can be determined for all of these divisions: other measures such as thickness apply only to the cortex. Although many brain growth phenotypes can be determined (for example, considering multiple cortical and

subcortical regions), the four major divisions and the region of interest were selected as phenotypes prior to conducting any statistical analyses to limit the number of multiple tests.

Longitudinal data was used to chart individual growth (Kraemer, Yesavage, Taylor, & Kupfer, 2000). However, modeling growth at the level of individuals presents several challenges. First, it is not feasible to model non-linear growth at an individual level as this requires a minimum of three, and ideally more, observations, and all participants in the *Generation R* cohort had only 2 images. Thus individual-level growth was taken to be linear. Volumetric change was measured controlling for the age at initial assessment and has the same unit of time (i.e., change per annum) through adjustment for inter-scan interval. Such linear modeling is a reasonable approximation for growth. The proportional change in volume for each region was calculated as follows: change in volume between the two scans divided by the baseline volume. Then, using linear regression, this proportional change was adjusted for the age at baseline scan, the interscan interval, sex and ADHD status. A positive value indicates that the brain region showed overall (adjusted) increase in volume; a negative value indicates age-related decrease in volume. After adjustment, residualized values were normalized using rank-based inverse normal transformation (Blom's formula: (rank-3/8)/(n +1/4). Normalization is necessary to ameliorate the well known inflated type I error rate (Schwantes-An et al., 2016).

Residualized change in the cerebral cortex and right lateral prefrontal cortex were highly correlated (r = 0.86), but correlation between all other phenotypes was modest (mean overall correlation of 0.27; maximum r = 0.39 between the cerebellum and white matter, and minimum of r = 0.11 for the cerebellum and basal ganglia).

Data analysis

GWAS—Tests of association were performed assuming an additive linear model. Due to the degree of relatedness (N=146 sibling pairs) in the *LONG* cohort, Efficient Mixed-Model Association eXpedited (EMMAX) (H. M. Kang et al., 2010), a method which accounts for population structure and relatedness among the participants, was used to perform the GWAS for each trait. EMMAX has a flexible variance component approach that enables correcting for a wide range of sample structures by explicitly accounting for pair-wise relatedness between individuals, using high-density SNPs. Furthermore, admixture-adjusted marker-based kinship matrices were used as input for EMMAX, calculated using the Relationship Estimation in Admixed Population (REAP) method (Thornton et al., 2012). Both simulations and analysis of real data have previously demonstrated that REAP can adequately account for population structure and ancestry-related assortative mating by using individual-specific allele frequencies at SNPs that are calculated on the basis of ancestry derived from whole-genome analysis.) (Thornton et al., 2012). To this end, ADMIXTURE (Alexander, Novembre, & Lange, 2009) was used to estimate the number of subpopulations in each cohort, assuming k = 3 subpopulations for *LONG* and k = 2 subpopulations for *Generation R* based on the best model fit criteria comparing models with k = 1 through 5 subpopulations in both cohorts. Output from the ADMIXTURE analysis provided the individual-specific allele frequencies used in REAP. To illustrate population structure within

the LONG cohort, principal components analysis and Scree plots are provided in supplementary Figures S1a and S1b.

Replication was performed with two different methods. In the first, a meta-analysis based on the EMMAX associations obtained from these two cohorts was performed. This was done with the rmeta package in R, which is based on an inverse variance weighted meta-analysis conducted for each of the five phenotypes assuming a random effect. Nominal genome-wide significance was determined by the standard significance criteria of $5x10^{-8}$. Although five growth phenotypes were tested, multiple test correction was based on 4 traits (Li & Ji, 2005) because of the high correlation between cerebral cortex and right lateral prefrontal cortex (r = 0.86) resulting in an adjusted genome-wide significance level of $1.25x10^{-8}$.

The second approach was to use ComPaSS-GWAS [Sabourin et al., submitted], a method based on complimentary pairs stability selection (Shah & Samworth, 2013) applied to a traditional regression based GWAS selection procedure. ComPaSS-GWAS approximates replication by randomly splitting the data in half (into a pseudo-discovery and pseudo-replication set) multiple times and looking for corroboration of GWAS results within the random splits. For a given critical value, ComPaSS-GWAS returns a corroboration score between 0 and 1, indicating the proportion of random splits where each SNP was corroborated at a given threshold. SNPs that are corroborated at a high rate are less sensitive to sampling variability, i.e., their significance is less likely to be an artifact of the sample analyzed, and therefore are more likely to be replicated. Simulations show that a ComPaSS-GWAS score of at least 0.25 and 0.5 based on a within split critical value parameter of 10^{-3} results in type I error rates of less than $6.6x10^{-7}$ and $5x10^{-8}$, respectively. For this study, ComPaSS-GWAS was used only on the larger cohort (the *LONG* cohort). ComPaSS-GWAS scores based on a within split critical value parameter of 10^{-3} and a corroboration score of at least 0.4 are reported. Significant results from ComPaSS-GWAS for the *LONG* cohort that were not supported by the meta-analysis but which had evidence of heterogeneity ($P_{het} < 0.05$) were also examined.

Pathway analysis—Pathway-based analyses can detect pathophysiological mechanisms that may be missed by focusing on single alleles, especially when considering only those attaining genome-wide significance (Mooney et al., 2016; Wang, Li, & Hakonarson, 2010). Pathway analyses was based on SNPs associated with the brain growth phenotypes in the meta-analyses at p-values < $1x10^{-5}$. SNPs meeting these conditions were annotated to genes using wANNOVAR(Chang & Wang, 2012). Using Ingenuity Pathway Analysis (IPA, QIAGEN, Redwood City, www.qiagen.com/ingenuity), enrichment for canonical pathways among these variants was tested. Pathways that were significantly enriched at a Benjamini-Hochberg (B-H) corrected q < 0.05, were taken to be significant.

Polygenic risk score analysis—Polygenic risk scores (PRS) for ADHD has already been shown to predict neuropsychological and clinical characteristics pertinent to the disorder (Hamshere et al., 2013). However, the ability of PRS for ADHD to predict brain growth has been untested. PRS was calculated for ADHD in the LONG cohort based on the recent release of ADHD GWAS mega-analysis results by the Psychiatric Genomics Consortium and the Lundbeck Foundation Initiative for Integrative Psychiatric Research

(iPSYCH) (Demontis et al., 2017). These meta-analyses did not include the *LONG* or *Generation R* cohorts. PRS were generated using the PRSice software (Euesden, Lewis, & O'Reilly, 2014). We excluded SNPs with an $r^2 > 0.1$, which were within 250 kilobases (kb) of each other. The following p-value thresholds to define PRS: 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5. Brain growth phenotypes within *LONG* were recalculated so that proportional change in volume was adjusted for age at baseline scan, interscan interval and sex but not ADHD status, as we were interested in the genetic overlap between ADHD and brain growth in a sample that was enriched for the diagnosis of ADHD. Residualized brain growth phenotypes were normalized using rank-based inverse normal transformation and regressed onto the standardized polygenic risk scores using linear mixed-models, which included a random term to account for relatedness among some participants and the first three principal components reflecting population structure.

RESULTS

Phenotype

Demographic differences were found between the two cohorts: the *LONG* cohort had an older mean age and a higher proportion of Caucasian participants (Table I). These demographic differences were accounted for in all analyses. Considering the brain growth phenotypes, the overall pattern of age-related change in regional volumes was similar across the cohorts (supplementary Figure S2). There was no significant difference in brain growth phenotypes between individuals with and without ADHD in any of the cohorts (Table II). Nevertheless, due to the high frequency of ADHD diagnosis within *LONG*, ADHD status was adjusted for in the main analyses. There was no significant effect of study site on the overall growth phenotypes in a repeated-measures ANOVA (between-subjects main effect of site ($p = 0.24$).

GWAS results

When the cohorts were analyzed separately, no SNP attained genome-wide significance ($p < 5x10^{-8}$) based on the GWAS with EMMAX in either the *LONG* or *Generation R* cohorts' analyses. When combining the results from the two cohorts in a meta-analysis, one SNP was found to be genome-wide significant on chromosome 7 (rs12386571, $p = 9.09x10^{-9}$) for white matter (Figure 1). This SNP is 20kb upstream from the aldo-keto reductase family 1 member B10 (*AKR1B10*) gene and 48kb upstream from the *AKR1B1* gene. Regional association plots of the most significant SNPs are shown in Figure 2. QQ and Manhattan plots from the meta-analyses can be found in the supplementary Figures S3–S7. The SNPs that were suggestive for association (meta-analysis $p < 1x10^{-5}$, or in *LONG* cohort alone at $p < 1x10^{-7}$ are presented in suppplementary Tables S1–S5 and Figures S8–9. Fine mapping through imputation of the region surrounding rs12386571 for white matter growth was performed (Figure S10), and one SNP, rs13237016, was found to be slightly more significant ($p = 5.42 \times 10^{-9}$). Fine mapping was also performed for the region encompassing rs10056913 for right lateral prefrontal cortical growth, but no SNPs reached nominal genomewide significance.

The application of ComPaSS-GWAS to the *LONG* cohort resulted in multiple SNPs with corroboration scores greater than 0.5, which corresponds to genome-wide significant type I error rates of $< 5x10^{-8}$. SNPs that were found to be at least suggestively associated by both ComPaSS-GWAS and the meta-analysis are considered. For white matter, ComPaSS-GWAS corroborates the meta-analysis' most significant SNP on chromosome 7, rs12386571, with a ComPaSS-GWAS score of 0.56. ComPaSS-GWAS identified multiple SNPs in the *LONG* cohort, which were not replicated in the meta-analysis but had significant meta-analysis heterogeneity p-values. The SNPs that were found strongly suggestive based on ComPaSS-GWAS on the *LONG* cohort (corroboration score ≥ 0.4) are also given in Tables S1–S5, and forest plots of the top associations for white matter and right lateral prefrontal cortical growth are in Figures S8–S9.

Pathway and polygenic risk analyses

No neural pathways were enriched by genes implicated by nominally significant genome-wide associations with any of the brain growth phenotypes. Polygenic risk for ADHD was also not significantly associated with the brain growth phenotypes (all $p > 0.1$)

DISCUSSION

We detected a genome-wide significant common genetic variant that was associated with white matter growth. This SNP was located on chromosome 7 between *AKR1B1* and *AKR1B10*. Both genes are part of the nicotinamide adenine dinucleotide (phosphate)-dependent aldo-keto reductase (AKR) 1B subfamily (Mindnich & Penning, 2009). Candidate gene and gene expression studies have also associated *AKR1B10* with eating-disorders and nicotine dependence (M. W. Kang et al., 2011; Rohde et al., 2015). This is consonant with the gene's decreased expression in the reward circuitry of rat strains that are more sensitive to the reinforcing properties of cocaine, morphine, and ethanol (Higuera-Matas et al., 2011). *AKR1B1* is involved in the biosynthesis of neurotransmitters dopamine and serotonin (Friedman et al., 2012) and elevated peripheral blood expression has been reported in adults with schizophrenia (de Jong et al., 2012).

The polygenic risk for ADHD determined from meta-analyses of existing GWAS was not associated with brain growth in a sample enriched for the diagnosis of ADHD. We had speculated that by taking into account prior knowledge about complex molecular networks and biological pathways, we could identify genes and mechanisms involved in brain growth, through considering 'promising' genetic assocaitions in pathway analyses. However, no pathways pertinent to brain structure or function were detected that survived our adjustment for multiple trait comparisons.

The importance of using longitudinal data when studying developmental processes such as brain growth has been remarked upon by many, but presents several practical challenges as discussed earlier (Kraemer et al., 2000). The approach we used modeled individual level change. There was considerable similarity between the cohorts in growth rates despite their differences in age range. Even when specific growth parameters differed, the pattern of overall change did not and there was no overall effect of cohort on growth estimates.

Nonetheless, in common with all previous multi-site imaging genetic studies, the imaging data were acquired on different scanners and integrating such data is a challenge common to all multi-center studies. This is particularly true for legacy data, which did not include the collection of data that can aid cross-scanner calibration (such as the use of human phantoms etc). However, there are several mitigating factors. First, there is a similarity in patterns of brain growth across the two cohorts, despite the different scanners. Second, some effects of cohort heterogeneity can be attenuated by processing all data on the same servers with the same version of FreeSurfer and by having the same raters conduct all quality assurance procedures (Gronenschild et al., 2012). Most importantly, association analyses were conducted separately for each cohort, and only then combined results using meta-analytic techniques.

Given the differences in the two cohorts, it is not unexpected to have regions of heterogeneity in the meta-analysis, thus explaining some of the result discrepancies found between the ComPaSS-GWAS and meta-analysis results. These genomic regions with contrasting SNP associations may be due to population differences between the cohorts, not just relating to ethnicity, but may also be due to the difference between cohort in regards to the ages at which the first and second scan were taken.

There were other limitations to our study. First, we were underpowered to detect SNPs with small effect size involved in brain growth using conventional single-marker association analysis. However, only the currently available childhood cohorts that combine longitudinal neuroimaging with genome-wide SNP data were included. Attempts to replicate these preliminary findings will be possible as similar longitudinal neuroimaging studies (e.g., IMAGEN, ABCD or the Saguenay Youth Study) begin to report their findings (NIDA, September 2016; Nymberg, Jia, Ruggeri, & Schumann, 2013; Pausova et al., 2007). Finally, as is often the case in GWAS meta-analyses, the two cohorts were genotyped on different platforms (~42% of the SNPs genotyped in the *LONG* cohort were also genotyped in *Generation R*); however, allele frequencies were similar in both studies. To maximize the number of overlapping SNPs across the two sites, we included both genotyped and imputed SNPs in analyses of *Generation R*. We also used recommended approaches to reduce the possibility of type 1 errors by retaining only imputed SNPs with high quality imputation scores (de Bakker et al., 2008; Sinnott & Kraft, 2012). The study included sibling pairs, and relatedness was accounted for in the analyses using EMMAX. However, we did not have enough sibling sets to provide heritability estimates for our phenotypes, nor to conduct linkage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The *LONG* Study was funded by the Intramural Programs of the National Human Genome Research Institute (NHGRI) and National Institute of Mental Health (NIMH). The authors gratefully acknowledge the participation of all children and families in the study. Genotyping services for the *LONG* cohort were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268201200008I. This project was also supported

in part by the Division of Intramural Research Program of NHGRI. Analyse The *Generation R* Study is conducted by the Erasmus Medical Center in close collaboration with the Erasmus University Rotterdam, the Municipal Health Service in the Rotterdam area, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of general practitioners, hospitals, midwives and pharmacies in Rotterdam. Neuroimaging studies within the *Generation R* are supported through the Netherlands Organization for Health Research and Development (NWO) (ZonMw TOP 40-00812-98-11021), the European Community's 7th Framework Program (FP7/2008–2013) under grant agreement 212652 (NUTRIMENTHE), the Stichting Sophia Kinderziekenhuis Fonds, and General Electric Healthcare. The *Generation R* Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, ZonMw (ZonMW 10.000.1003), NWO, the Ministry of Health, Welfare and Sport, and the Ministry of Youth and Families.

References

Achenbach, TM., Rescorla, L. Manual for the ASEBA preschool forms & profiles : an integrated system of multi-informant assessment. Burlington, VT: ASEBA; 2000.

Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Research. 2009; 19(9):1655–1664. DOI: 10.1101/gr.094052.109 [PubMed: 19648217]

Bis JC, DeCarli C, Smith AV, van der Lijn F, Crivello F, Fornage M, … Seshadri S. Common variants at 12q14 and 12q24 are associated with hippocampal volume. Nature genetics. 2012; 44(5):545–551. DOI: 10.1038/ng.2237 [PubMed: 22504421]

Brouwer RM, Panizzon MS, Glahn DC, Hibar DP, Hua X, Jahanshad N, … Hansell NK. Genetic influences on individual differences in longitudinal changes in global and subcortical brain volumes: Results of the ENIGMA plasticity working group. Hum Brain Mapp. 2017

Castellanos F, Lee P, Sharp W, Jeffries N, Greenstein D, Clasen L, … Rapoport J. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. JAMA. 2002; 288:1740–1748. [PubMed: 12365958]

Castellanos FX, Lee PP, Sharp W, Jeffries NO, Greenstein DK, Clasen LS, … Rapoport JL. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. JAMA. 2002; 288(14):1740–1748. [PubMed: 12365958]

Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. J Med Genet. 2012; 49(7):433–436. DOI: 10.1136/jmedgenet-2012-100918 [PubMed: 22717648]

Courchesne E, Campbell K, Solso S. Brain growth across the life span in autism: Age-specific changes in anatomical pathology. Brain Research. 2011; 1380:138–145. doi:http://dx.doi.org/10.1016/j.brainres.2010.09.101. [PubMed: 20920490]

Courchesne E, Carper R, Akshoomoff N. Evidence of brain overgrowth in the first year of life in autism. JAMA. 2003; 290(3):337–344. DOI: 10.1001/jama.290.3.337 [PubMed: 12865374]

Courchesne E, Pierce K. Brain overgrowth in autism during a critical time in development: implications for frontal pyramidal neuron and interneuron development and connectivity. International Journal of Developmental Neuroscience. 2005; 23(2–3):153–170. doi:http://dx.doi.org/10.1016/j.ijdevneu.2005.01.003. [PubMed: 15749242]

de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet. 2008; 17(R2):R122–128. DOI: 10.1093/hmg/ddn288 [PubMed: 18852200]

de Jong S, Boks MPM, Fuller TF, Strengman E, Janson E, de Kovel CGF, … Ophoff RA. A Gene Co-Expression Network in Whole Blood of Schizophrenia Patients Is Independent of Antipsychotic-Use and Enriched for Brain-Expressed Genes. PLoS ONE. 2012; 7(6):e39498.doi: 10.1371/journal.pone.0039498 [PubMed: 22761806]

Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, … Cerrato F. Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. bioRxiv. 2017:145581.

Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, … Milham MP. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry. 2014; 19(6):659–667. DOI: 10.1038/mp.2013.78 [PubMed: 23774715]

Ducharme S, Hudziak JJ, Botteron KN, Albaugh MD, Nguyen TV, Karama S, Evans AC. Decreased Regional Cortical Thickness and Thinning Rate Are Associated With Inattention Symptoms in

Healthy Children. Journal of the American Academy of Child & Adolescent Psychiatry. 2012; 51(1):18–27. e12. doi:http://dx.doi.org/10.1016/j.jaac.2011.09.022. [PubMed: 22176936]

El-Sayed E, Larsson JO, Persson HE, Santosh PJ, Rydelius PA. "Maturational lag" hypothesis of attention deficit hyperactivity disorder: an update. Acta Pædiatrica. 2003; 92(7):776–784. DOI: 10.1111/j.1651-2227.2003.tb02531.x

Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. Bioinformatics. 2014; 31(9):1466–1468. [PubMed: 25550326]

Friedman J, Roze E, Abdenur JE, Chang R, Gasperini S, Saletti V, … Blau N. Sepiapterin reductase deficiency: a treatable mimic of cerebral palsy. Ann Neurol. 2012; 71(4):520–530. DOI: 10.1002/ana.22685 [PubMed: 22522443]

Gronenschild EH, Habets P, Jacobs HI, Mengelers R, Rozendaal N, Van Os J, Marcelis M. The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. PLoS ONE. 2012; 7(6):e38234. [PubMed: 22675527]

Hamshere ML, Langley K, Martin J, Agha SS, Stergiakouli E, Anney RJ, … Thapar A. High loading of polygenic risk for ADHD in children with comorbid aggression. Am J Psychiatry. 2013; 170(8): 909–916. DOI: 10.1176/appi.ajp.2013.12081129 [PubMed: 23599091]

Hibar DP, Medland SE, Stein JL, Kim S, Shen L, Saykin AJ, … Thompson PM. Genetic clustering on the hippocampal surface for genome-wide association studies. Medical image computing and computer-assisted intervention : MICCAI … International Conference on Medical Image Computing and Computer-Assisted Intervention. 2013; 16(2):690–697.

Higuera-Matas A, Montoya GL, Coria SM, Miguens M, Garcia-Lecumberri C, Ambrosio E. Differential gene expression in the nucleus accumbens and frontal cortex of lewis and Fischer 344 rats relevant to drug addiction. Curr Neuropharmacol. 2011; 9(1):143–150. DOI: 10.2174/157015911795017290 [PubMed: 21886580]

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-y, Freimer NB, … Eskin E. Variance component model to account for sample structure in genome-wide association studies. Nat Genet. 2010; 42(4): 348–354. doi:http://www.nature.com/ng/journal/v42/n4/suppinfo/ng.548_S1.html. [PubMed: 20208533]

Kang MW, Lee ES, Yoon SY, Jo J, Lee J, Kim HK, … Kim H. AKR1B10 is associated with smoking and smoking-related non-small-cell lung cancer. J Int Med Res. 2011; 39(1):78–85. [PubMed: 21672310]

Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IMH, … Jaddoe VW. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016; 31(12):1243–1264. DOI: 10.1007/s10654-016-0224-9 [PubMed: 28070760]

Kraemer HC, Yesavage JA, Taylor JL, Kupfer D. How Can We Learn About Developmental Processes From Cross-Sectional Studies, or Can We? Am J Psychiatry. 2000; 157(2):163–171. DOI: 10.1176/appi.ajp.157.2.163 [PubMed: 10671382]

Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb). 2005; 95(3):221–227. DOI: 10.1038/sj.hdy.6800717 [PubMed: 16077740]

Mackie S, Shaw P, Lenroot R, Pierson R, Greenstein DK, Nugent TF 3rd, … Rapoport JL. Cerebellar development and clinical outcome in attention deficit hyperactivity disorder. Am J Psychiatry. 2007; 164(4):647–655. DOI: 10.1176/ajp.2007.164.4.647 [PubMed: 17403979]

Medina-Gomez C, Felix JF, Estrada K, Peters MJ, Herrera L, Kruithof CJ, … Rivadeneira F. Challenges in conducting genome-wide association studies in highly admixed multi-ethnic populations: the Generation R Study. Eur J Epidemiol. 2015; 30(4):317–330. DOI: 10.1007/s10654-015-9998-4 [PubMed: 25762173]

Melville SA, Buros J, Parrado AR, Vardarajan B, Logue MW, Shen L, … Farrer LA. Multiple loci influencing hippocampal degeneration identified by genome scan. Ann Neurol. 2012; 72(1):65–75. DOI: 10.1002/ana.23644 [PubMed: 22745009]

Mindnich RD, Penning TM. Aldo-keto reductase (AKR) superfamily: Genomics and annotation. Human Genomics. 2009; 3(4):362–370. DOI: 10.1186/1479-7364-3-4-362 [PubMed: 19706366]

Mooney MA, McWeeney SK, Faraone SV, Hinney A, Hebebrand J, Nigg JT, Wilmot B. Pathway analysis in attention deficit hyperactivity disorder: An ensemble approach. Am J Med Genet B Neuropsychiatr Genet. 2016; 171(6):815–826. DOI: 10.1002/ajmg.b.32446 [PubMed: 27004716]

Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch KP, … Nelson S. Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2010; 49(9):884–897. DOI: 10.1016/j.jaac.2010.06.008 [PubMed: 20732625]

NIDA. Recruitment begins for landmark study of adolescent brain development [Press release]. Sep, 2016. Retrieved from https://www.drugabuse.gov/news-events/news-releases/2016/09/recruitment-begins-landmark-study-adolescent-brain-development

Nymberg C, Jia T, Ruggeri B, Schumann G. Analytical strategies for large imaging genetic datasets: experiences from the IMAGEN study. Annals of the New York Academy of Sciences. 2013; 1282(1):92–106. DOI: 10.1111/nyas.12088 [PubMed: 23488575]

Pausova Z, Paus T, Abrahamowicz M, Almerigi J, Arbour N, Bernard M, … Watkins K. Genes, maternal smoking, and the offspring brain and body during adolescence: Design of the Saguenay Youth Study. Hum Brain Mapp. 2007; 28(6):502–518. DOI: 10.1002/hbm.20402 [PubMed: 17469173]

Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE. Genetic influences on human brain structure: a review of brain imaging studies in twins. Hum Brain Mapp. 2007; 28(6):464–473. DOI: 10.1002/hbm.20398 [PubMed: 17415783]

Redcay E, Courchesne E. When is the brain enlarged in autism? A meta-analysis of all brain size reports. Biological Psychiatry. 2005; 58(1):1–9. [PubMed: 15935993]

Reich W. Diagnostic Interview for Children and Adolescents (DICA). Journal of the American Academy of Child & Adolescent Psychiatry. 2000; 39(1):59–66. doi:http://dx.doi.org/10.1097/00004583-200001000-00017. [PubMed: 10638068]

Rohde K, Federbusch M, Horstmann A, Keller M, Villringer A, Stumvoll M, … Bottcher Y. Genetic variants in AKR1B10 associate with human eating behavior. BMC Genet. 2015; 16:31.doi: 10.1186/s12863-015-0189-9 [PubMed: 25887478]

Rubia K. Neuro-anatomic evidence for the maturational delay hypothesis of ADHD. Proceedings of the National Academy of Sciences. 2007; 104(50):19663–19664. DOI: 10.1073/pnas.0710329105

Schwantes-An TH, Sung H, Sabourin JA, Justice CM, Sorant AJ, Wilson AF. Type I error rates of rare single nucleotide variants are inflated in tests of association with non-normally distributed traits using simple linear regression methods. BMC Proc. 2016; 10(Suppl 7):385–388. DOI: 10.1186/s12919-016-0060-7 [PubMed: 27980666]

Shah RD, Samworth RJ. Variable selection with error control: another look at stability selection. Journal of the Royal Statistical Society Series B-Statistical Methodology. 2013; 75(1):55–80. DOI: 10.1111/j.1467-9868.2011.01034.x

Shaw P, De Rossi P, Watson B, Wharton A, Greenstein D, Raznahan A, … Chakravarty MM. Mapping the development of the basal ganglia in children with attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2014; 53(7):780–789. e711. DOI: 10.1016/j.jaac.2014.05.003 [PubMed: 24954827]

Shaw P, Gilliam M, Liverpool M, Weddle C, Malek M, Sharp W, … Giedd J. Cortical development in typically developing children with symptoms of hyperactivity and impulsivity: support for a dimensional view of attention deficit hyperactivity disorder. Am J Psychiatry. 2011; 168(2):143–151. DOI: 10.1176/appi.ajp.2010.10030385 [PubMed: 21159727]

Shaw P, Gogtay N, Rapoport J. Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. Hum Brain Mapp. 2010; 31(6):917–925. DOI: 10.1002/hbm.21028 [PubMed: 20496382]

Sinnott JA, Kraft P. Artifact due to differential error when cases and controls are imputed from different platforms. Human genetics. 2012; 131(1):111–119. [PubMed: 21735171]

Sripada CS, Kessler D, Angstadt M. Lag in maturation of the brain's intrinsic functional architecture in attention-deficit/hyperactivity disorder. Proceedings of the National Academy of Sciences. 2014; 111(39):14259–14264. DOI: 10.1073/pnas.1407787111

Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, … Thompson PM. Identification of common variants associated with human hippocampal and intracranial volumes. Nat Genet. 2012; 44(5):552–561. DOI: 10.1038/ng.2250 [PubMed: 22504417]

Thornton T, Tang H, Hoffmann TJ, Ochs-Balcom HM, Caan BJ, Risch N. Estimating kinship in admixed populations. Am J Hum Genet. 2012; 91(1):122–138. DOI: 10.1016/j.ajhg.2012.05.024 [PubMed: 22748210]

Tick NT, van der Ende J, Koot HM, Verhulst FC. 14-year changes in emotional and behavioral problems of very young Dutch children. J Am Acad Child Adolesc Psychiatry. 2007; 46(10):1333–1340. DOI: 10.1097/chi.0b013e3181337532 [PubMed: 17885575]

Wallace GL, Eric Schmitt J, Lenroot R, Viding E, Ordaz S, Rosenthal MA, … Giedd JN. A pediatric twin study of brain morphometry. J Child Psychol Psychiatry. 2006; 47(10):987–993. DOI: 10.1111/j.1469-7610.2006.01676.x [PubMed: 17073977]

Walters JTR, Rujescu D, Franke B, Giegling I, Vásquez AA, Hargreaves A, … Owen MJ. The Role of the Major Histocompatibility Complex Region in Cognition and Brain Structure: A Schizophrenia GWAS Follow-Up. American Journal of Psychiatry. 2013; 170(8):877–885. DOI: 10.1176/appi.ajp.2013.12020226 [PubMed: 23903335]

Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. Nat Rev Genet. 2010; 11(12):843–854. [PubMed: 21085203]

White T, El Marroun H, Nijs I, Schmidt M, van der Lugt A, Wielopolki PA, … Verhulst FC. Pediatric population-based neuroimaging and the Generation R Study: the intersection of developmental neuroscience and epidemiology. Eur J Epidemiol. 2013; 28(1):99–111. DOI: 10.1007/s10654-013-9768-0 [PubMed: 23354984]

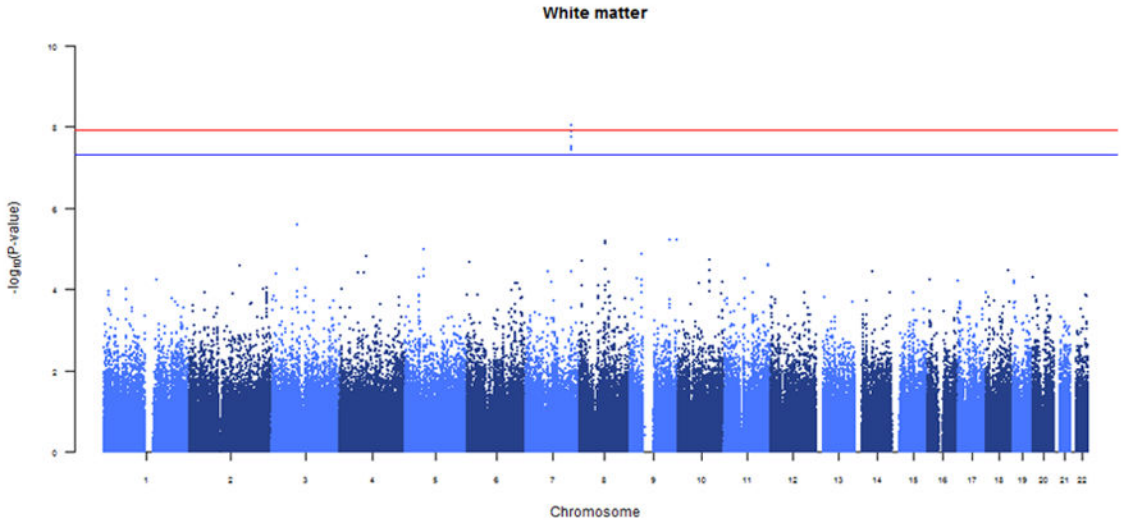

Figure 1.

Manhattan plot for the white matter growth phenotypes based on GWAS meta-analyses of *LONG* and *Generation R*. The blue line corresponds to $p=5x10^{-8}$ (genome wide significance), while the red line corresponds to $p=1.25x10^{-8}$ (genomewide significance adjusted for multiple traits).

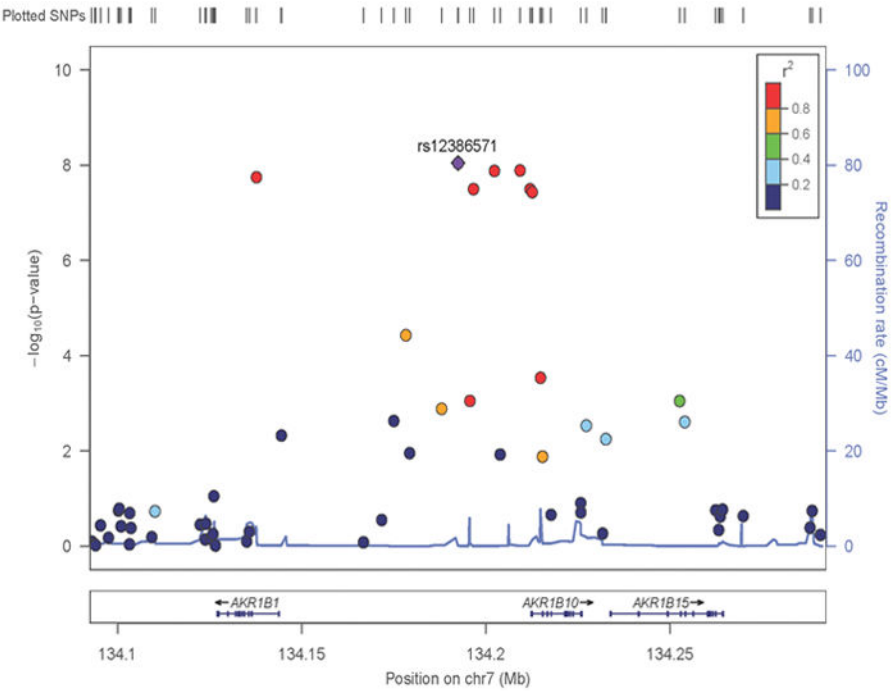

Figure 2.

Regional plot of lead signal for white matter. SNPs are depicted by circles, and the color of the circle indicates the amount of LD between that SNP and rs12386571. Plot created using LocusZoom (www.locuszoom.org).

Table I

Basic description of the two cohorts

	***LONG* cohort (N=458)**	***Generation R* (N=257)**	
	M(SD) or %	**M(SD) or %**	***p*-value**
Age (years) at scan 1	11.47 (3.54)	7.70 (0.86)	<0.001
Age (years) at scan 2	16.13 (4.72)	10.23 (0.68)	<0.001
Males	59.2	52.5	0.098
ADHD diagnosis	26.0	3.0	<0.001
White/Caucasian	88.2	68.9	<0.001

Note. Continuous variables were compared between the two cohorts using t-tests, categorical variables were compared using χ^2 tests.

Table II

Comparing brain growth phenotypes according to ADHD status

Residualized proportion change [a]	LONG cohort (N=458, ADHD=119)			Generation R (N=257, ADHD=13)		
	t	*df*	*p*	*T*	*df*	*p*
Cortex	0.541	211.655	0.589	−0.777	12.599	0.452
White matter	1.675	228.169	0.095	0.730	12.393	0.479
Right lateral prefrontal cortex	0.277	212.727	0.782	−0.454	12.665	0.658
Cerebellum	1.780	191.283	0.077	1.286	12.477	0.222
Basal ganglia	1.163	205.316	0.246	1.375	12.471	0.193

Note. Population variances of the two groups are assumed to be not equal; *t*-statistics are reported based on the un-pooled variances and corrected degrees of freedom.

[a]Proportion change in volume between first and last scan was adjusted for age at first scan, interscan interval and sex.