

ORIGINAL ARTICLE

# The genetic architecture of pediatric cognitive abilities in the Philadelphia Neurodevelopmental Cohort

EB Robinson<sup>1,2</sup>, A Kirby<sup>1,2</sup>, K Ruparel<sup>3</sup>, J Yang<sup>4</sup>, L McGrath<sup>5</sup>, V Anttila<sup>1,2,6</sup>, BM Neale<sup>1,2</sup>, K Merikangas<sup>7</sup>, T Lehner<sup>8</sup>, PMA Sleiman<sup>9</sup>, MJ Daly<sup>1,2</sup>, R Gur<sup>3</sup>, R Gur<sup>3,10,11</sup> and H Hakonarson<sup>9,10,11</sup>

The objective of this analysis was to examine the genetic architecture of diverse cognitive abilities in children and adolescents, including the magnitude of common genetic effects and patterns of shared and unique genetic influences. Subjects included 3689 members of the Philadelphia Neurodevelopmental Cohort, a general population sample comprising those aged 8–21 years who completed an extensive battery of cognitive tests. We used genome-wide complex trait analysis to estimate the SNP-based heritability of each domain, as well as the genetic correlation between all domains that showed significant genetic influence. Several of the individual domains suggested strong influence of common genetic variants (for example, reading ability,  $h_g^2 = 0.43$ ,  $P = 4e - 06$ ; emotion identification,  $h_g^2 = 0.36$ ,  $P = 1e - 05$ ; verbal memory,  $h_g^2 = 0.24$ ,  $P = 0.005$ ). The genetic correlations highlighted trait domains that are candidates for joint interrogation in future genetic studies (for example, language reasoning and spatial reasoning,  $r(g) = 0.72$ ,  $P = 0.007$ ). These results can be used to structure future genetic and neuropsychiatric investigations of diverse cognitive abilities.

*Molecular Psychiatry* advance online publication, 15 July 2014; doi:10.1038/mp.2014.65

## INTRODUCTION

Twin and family studies have consistently evidenced strong genetic influence on 'g,' or general intelligence, as well as several specific cognitive domains.<sup>1–3</sup> Recent studies using population samples have estimated that common single-nucleotide polymorphisms (SNPs) alone explain 22–46% of variation in intelligence in childhood and 40–51% of the same differences in adulthood.<sup>4–7</sup>

There have been few molecular genetic investigations of the unique cognitive domains that may link brain systems to intelligence and behavior—for example: memory, social cognition, executive function, and reasoning. The purpose of these analyses was to examine the genetic architecture of diverse cognitive abilities in children, including the magnitude of common genetic effects and patterns of shared and unique genetic influences.

First, we use genome-wide complex trait analysis<sup>8</sup> (GCTA) to estimate the extent to which the traits are influenced by all common SNPs (chip heritability). The heritability of several of these traits has been previously estimated using twin and family studies,<sup>1,2,9–11</sup> but a consistent molecular genetic approach has not been used across domains in a single pediatric sample. In doing so, this study aims to highlight cognitive traits, and measurement methods, that are amenable to genetic study in pediatric samples. In addition to nominating phenotypes for

GWAS and sequencing analysis, traits with evidence of genetic influence may be strong candidates for endophenotypic research in neuropsychiatric disease. Schizophrenia and autism, for example, have been linked to familial deficits in several neuropsychiatric trait domains.<sup>12,13</sup> As knowledge of the genetic factors relevant to those disorders continues to improve,<sup>14</sup> heritable traits measured on samples from the general population can be examined for evidence of shared genetic influence with neuropsychiatric disease, and pursued for indicators of common underlying biology.

Second, we estimate genetic correlations between the cognitive traits that show evidence of common variant influence. While it has been shown that some genetic factors influence multiple cognitive abilities (that is, 'generalist genes'),<sup>15–17</sup> weak within-individual correlations between many domains suggest that other genetic effects are likely to be domain-specific or domain-preferential. A map of the genetic correlations between the traits could inform several types of neuropsychiatric research. For neuroscientists, the genetic correlation matrix, in conjunction with the phenotypic correlation matrix, may provide insight into ability domains most likely to share biological underpinnings. For example, there is a strong genetic correlation between schizophrenia and bipolar disorder, but no genetic correlation between either of those disorders and Crohn's disease.<sup>18</sup> The difference between the correlations is hypothesized to reflect greater shared

<sup>1</sup>Analytic and Translational Genetics Unit, Massachusetts General Hospital and Department of Medicine, Harvard Medical School, Boston, MA, USA; <sup>2</sup>Stanley Center for Psychiatric Research and Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA; <sup>3</sup>Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; <sup>4</sup>Queensland Brain Institute, University of Queensland, Brisbane, Queensland, Australia; <sup>5</sup>Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research and Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA; <sup>6</sup>Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland; <sup>7</sup>Genetic Epidemiology Research Branch, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA; <sup>8</sup>Office of Genomics Research Coordination, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA and <sup>9</sup>Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. Correspondence: Dr EB Robinson, Analytic and Translational Genetics Unit, Massachusetts General Hospital and Department of Medicine, Harvard Medical School, 6th Floor, 185 Cambridge Street, Boston, MA 02114, USA. E-mail: erobinson@atgu.mgh.harvard.edu

<sup>10</sup>Information about the Philadelphia Neurodevelopmental Cohort can be found at raquel@mail.med.upenn.edu and hakonarson@email.chop.edu.

<sup>11</sup>These authors contributed equally to this work.

Received 12 December 2013; revised 29 April 2014; accepted 13 May 2014

biology between schizophrenia and bipolar disorder, brain diseases associated with aberrant behavior, as compared to Crohn's disease, a disease of the gastrointestinal tract and autoimmune system. A similar gradient in biological proximity may be predicted by genetic correlations *among* cognitive phenotypes. For researchers conducting GWAS and sequencing analysis, traits with strong genetic correlations can be jointly interrogated,<sup>19</sup> thereby increasing the probability of collaborative analysis across cohorts with different phenotypic and measurement approaches. As very large sample sizes are required for association research in cognitive and neuropsychiatric disease,<sup>5</sup> cross-cohort collaborations will be needed to identify informative genetic signals.

## SUBJECTS AND METHODS

### The Philadelphia Neurodevelopmental Cohort

The Philadelphia Neurodevelopmental Cohort (PNC) is a population-based sample from the greater Philadelphia area, including over 9000 individuals aged 8–21 years who received medical care within the Children's Hospital of Philadelphia network. The participants presented for a diverse set of medical needs, ranging from a general health checkup and minor problems (for example, sports-related bruise, rash) to chronic condition management (for example, asthma, type 1 diabetes) to potentially life-threatening health problems (for example, cancer). They were initially enrolled in the genetic study at the Center for Applied Genomics at Children's Hospital of Philadelphia. Upon assent/consent, the participants were genotyped during the time of their clinical visit and provided written permission to be recontacted for studies of complex pediatric disorders. The PNC participants were selected at random after stratification by sex, age and ethnicity. The overall inclusion criteria included: (1) ability to provide signed informed consent (parental consent was required for participants under age 18), (2) English language proficiency, and (3) physical and cognitive ability to participate in computerized cognitive testing.

All PNC participants completed the Computerized Neurocognitive Battery (CNB) and were assessed psychiatrically with a structured interview. The CNB consists of tests that have been used in functional neuroimaging to probe specific brain systems and is administered with a web browser. It assesses performance on a range of cognitive tasks. The CNB was designed to capture variation in four ability domains, and includes three specific tasks within each domain: (1) executive control (abstraction and mental flexibility, attention, working memory); (2) episodic memory (verbal, facial, spatial); (3) reasoning (verbal, nonverbal, spatial); and (4) social cognition (emotion identification, emotion differentiation, age differentiation). The specific measurement strategy employed for each of the 12 tasks has been described elsewhere,<sup>20</sup> but a summary of the measures and their psychometric properties in the PNC is included in Supplementary Table S1. The battery also included the reading items from the Wide Range Achievement Test (WRAT).<sup>21</sup>

### Cleaning and imputation of genotype data

This study employed genomewide complex trait analysis (GCTA) to estimate the fractional contribution of common SNPs to phenotypic variation in cognitive ability in the general population. One can reduce bias in values estimated through GCTA by minimizing ancestral heterogeneity in the sample.<sup>8,22</sup> As the PNC cohort was drawn from a diverse US urban population, these analyses were limited at the outset to the subset of participants who identified themselves as white non-hispanic ( $n=5141$ ). All samples were genotyped on one of three Illumina arrays: the HumanHap550, HumanHap610 or OmniExpress v2. Within the self-described white non-hispanic group, population outliers were further excluded based on directly genotyped SNP data, prior to imputation. Data were cleaned using a standard approach<sup>23</sup> (Supplementary Material), which reduced the sample by 584 individuals. Over half (62%) of these individuals were excluded for excess relatedness (the PNC included siblings). We conducted a PCA in PLINK<sup>24</sup> (Supplementary Figures S1a–S1d), which identified 527 individuals with outlying ethnicity, who were subsequently removed. An additional 341 individuals were removed in further phenotypic and genotypic exclusions, described below, resulting in a final analytic sample of 3689 individuals.

The genotype data were imputed in a separate phase of the study at Children's Hospital of Philadelphia. Unobserved genotypes from each chip set were imputed using the IMPUTE2 package and the reference haplotypes in Phase I of the 1000 genomes data (June 2011 release) that included ~37 138 905 variants from 1094 individuals from Africa, Asia, Europe and the Americas. Methodological details regarding the imputation are provided in the Supplementary Material. The imputed genotype data were used in the GCTA analyses.

### Phenotypic analyses

We created an index of the overall severity of participant medical conditions based on information obtained from parent or adult proband (ages 18–21) interviews and electronic medical records. The index included the following levels: (1) none or minor (43.4% of the analytic sample)—no ongoing medical conditions requiring sustained intervention (for example, well-child visit or sprained ankle); (2) mild (33.5% of the analytic sample)—conditions requiring pediatric visits and at times medications but mild in severity (for example, asthma, allergies); (3) moderate (22.8% of the analytic sample)—medical conditions requiring standing medications and monitoring (for example, diabetes, lupus); (4) severe medical conditions requiring multiple procedures and monitoring that can be life threatening. Individuals with medical rating 4 ( $n=274$ ) were excluded from the analysis to avoid conflation of their cognitive ability with the physical symptoms of their medical phenotype, which may influence test-taking performance.

Age of the participants was distributed uniformly between 8 and 21 years (mean = 13.7 years). Both age and sex were associated with mean differences in CNB performance,<sup>20</sup> and were regressed out of all CNB variables, as well as the WRAT, prior to analysis. Individuals with scores more than four standard deviations from the mean of any CNB variable were designated missing for that variable. Using the age- and sex-regressed scores, we conducted common factor and principal components analyses to examine their correlation structure. Common factor analysis analyzes the structure of the shared variance of a variable set; principal components analysis (PCA) examines all observed variance. We generated common factor and principal components scores—continuous variables that indicate individual rank within the standardized distribution of the factor or principal component—for use in the genetic analyses. Both the common factor analysis and PCA analyses used promax rotation, allowing the factors to be correlated. Phenotypic analyses were conducted in SAS version 9.3.<sup>25</sup> An unrotated PCA was conducted for purposes of comparison and yielded similar results (Supplementary Table S2). Given the strong correlation between the variables, only the rotated factor and principal components scores were included in the genetic analyses. There were 46 individuals missing the cognitive battery who were not included in the analysis.

### GCTA

The approach underlying GCTA has been described previously.<sup>5,8,22</sup> In brief, one uses common SNPs to create a genomic relatedness matrix (GRM) that includes pairwise relatedness estimates for all individuals in a sample. The GRM can be estimated using all typed SNPs (across the genome) or using a particular subset of SNPs—for example, those on a specific chromosome. Using the GCTA software tool,<sup>8</sup> one can compare phenotypic similarity with genotypic similarity, which provides an estimate of the fraction of phenotypic variation attributable to the additive effects of common SNPs (SNP-based heritability). For example, a univariate GCTA estimate of 0.35 suggests that 35% of variation in a trait is attributable to the SNPs used to make the GRM, as well as additional variants with which the typed markers are in linkage disequilibrium (LD; correlated variants). If the GRM was estimated using all genotyped SNPs, that fraction reflects the contribution of all common SNPs; if the GRM was estimated using just the SNPs on a single chromosome, that fraction reflects the specific contribution of common SNPs on that chromosome.

We constructed the primary GRM from all imputed, autosomal SNPs with minor allele frequency >1%, and Impute2 info score (an imputation quality metric) >0.6 ( $n=7\,635\,695$  SNPs).<sup>26</sup> This GRM was used in the primary GCTA analyses. After constructing the GRM, we removed the remaining individuals with relatedness >0.05 ( $n=20$ ),<sup>26</sup> which yielded the final analytic sample of 3689 individuals. In the univariate analyses, we first estimated the fraction of phenotypic variation in each cognitive trait explained by all common SNPs. Secondly, we estimated genetic correlations between each of the traits with nominally significant ( $P < 0.05$ ) univariate values. Genetic correlations provide an estimate of the degree to

which the common variants' influence on two traits is shared. A correlation of 1 suggests that the traits share all of their common genetic influences; a correlation of 0 suggests their influences are independent.

The second set of analyses examined the genomic distribution of genetic influence on the factor scores. We segmented the 7.6 million SNPs used in the primary GRM into: (a) 22 chromosome-specific GRMs and (b) genic or intergenic space using three genic window definitions.<sup>22</sup> For each gene, all genic footprints were defined using the RefSeq transcript with the longest spliced protein-coding DNA sequence.<sup>27</sup> All genome files were adjusted for incomplete LD between tagged and causal SNPs, as described by Yang *et al.*<sup>8</sup> All analyses controlled for the first 10 principal components of ancestry, which collectively explained < 2% of phenotypic variation in the common factor score.

## RESULTS

### Phenotypic analyses

The common factor analysis suggested that the measures have a single-factor solution; only one factor had an eigenvalue > 1 (Supplementary Figure S2a). In PCA, accuracy had a dominant first component but two additional interpretable components with eigenvalues > 1. The first principal component consisted of the reasoning and executive function items, along with the WRAT. The second principal component included the social cognition items; the third principal component included the memory items (Supplementary Figure S2b). Factor loadings for both the full sample and age-specific subgroups, as well as the phenotypic correlations, are included in Supplementary Tables S3a and b).

### Univariate GCTA analysis

Table 1 presents the univariate GCTA results. The individual trait scores suggested a wide range of common variant influence, ranging from 0% (emotion differentiation) to 43% (WRAT) of the phenotypic variation explained. The reasoning traits had the strongest SNP-based heritabilities, with estimates between 0.30 (s.e. = 0.10;  $P = 2e - 04$ ) and 0.41 (s.e. = 0.10;  $P = 4e - 06$ ). The heritability estimates of the factor and principal component scores reflected the results of their component traits—the first principal component score (reasoning/executive function) was the most genetically influenced of the summary measures (0.46; s.e. = 0.11;  $P = 8e - 06$ ). Strong common variant influence was also suggested for the common factor score (0.36; s.e. = 0.11;  $P = 5e - 04$ ). As indicated in Table 1, the univariate GCTA estimates for 7 of the outcome variables remained significantly different from 0 following a Bonferroni correction for 17 independent tests, which is conservative given that the outcome variables are correlated.

### Bivariate GCTA analysis

First, we estimated bivariate correlations between each of the variables with nominally significant ( $P < 0.05$ ) univariate heritability (Table 2). The genetic correlations ranged from -0.25 (emotion identification and nonverbal reasoning; s.e. 0.20;  $P(r_g = 0) = 0.09$ ) to 1.00 (WRAT and language reasoning; s.e. 0.15;  $P(r_g = 0) = 2e - 07$ ). More than 40% of the variation in the genetic correlations (40.8%) was explained by the phenotypic correlations ( $P = 0.01$ ), which led the variables to cluster genetically in a manner similar to that predicted by the phenotypic factor and PCAs. As shown in Supplementary Figure S3, the genetic correlations were, on average, greater than the phenotypic correlations, reflecting substantial pleiotropy across cognitive traits.

Second, we correlated the individual tasks (for example, language reasoning) with the common factor score. These correlations indicate the strength of the genetic relationship between the variable and the factor, and the extent of genetic influences on the variable that are distinct from those on the factor (Supplementary Table S4). With the exception of the association between emotion identification and the factor score

**Table 1.** Contribution of common genetic variation to population differences in the Computerized Neurocognitive Battery

Variable	N	V(G)/Vp	s.e.	P
<i>Domain scores</i>				
Abstraction and mental flexibility	3657	0.064	0.096	0.3
Attention	3618	0.148	0.097	0.07
Working memory	3577	0.108	0.096	0.1
Facial memory	3652	0.064	0.093	0.2
Spatial memory	3643	0.028	0.090	0.4
Verbal memory	3638	0.244	0.097	0.005
Language reasoning	3651	0.302	0.098	0.0008 <sup>a</sup>
Nonverbal reasoning	3654	0.406	0.096	4e - 06 <sup>a</sup>
Spatial reasoning	3581	0.357	0.101	0.0002 <sup>a</sup>
Age differentiation	3637	0.039	0.098	0.4
Emotion differentiation	3650	0.000	0.092	> 0.5
Emotion identification	3661	0.357	0.093	1e - 05 <sup>a</sup>
Wide Range Achievement Test (reading)	3680	0.433	0.098	4e - 06 <sup>a</sup>
<i>Factor and principal component scores</i>				
Common factor score	3380	0.360	0.108	0.0005 <sup>a</sup>
PC1 (reasoning and executive function)	3380	0.460	0.107	8e - 06 <sup>a</sup>
PC2 (memory)	3380	0.124	0.102	0.1
PC3 (social cognition)	3380	0.149	0.104	0.08

Abbreviations: PC, principal component; V(G)/Vp, estimate of the percentage of phenotypic variation attributable to the additive influence of common single-nucleotide polymorphisms; s.e. = standard error of V(G)/Vp.  $P = [P(V(G)/Vp = 0)]$ . Blue shading = executive function; beige = memory; green = complex cognition; Purple = social cognition; and gray = reading. <sup>a</sup> $P < 0.05$  following conservative correction for 17 independent tests.

( $r_g = 0.18$ ; s.e. 0.18;  $P(r_g = 0) = 0.2$ ), each of the genetic correlations exceeded 0.5. However, of the six item-to-factor genetic correlations, four were significantly different from 1 ( $P(r_g = 1) < 0.05$ ). This further suggests that while many of the genetic influences on cognitive ability are shared, cognitive domains often also have unique genetic influences.

### Genomic partitioning

We estimated the association between chromosome length and phenotypic variation explained per chromosome for the common factor score, which suggested that the factor is highly polygenic. Chromosome length explained 31.1% ( $P = 0.007$ ) of the variation in GRM prediction (Figure 1). The polygenic model is further supported by the regional analyses. SNPs in both genic and intergenic space were significantly associated with variation in cognitive ability ( $P(\text{joint association with genic and intergenic regions}) = 0.009 - 0.08$ ; Supplementary Figure S4).

## DISCUSSION

These analyses suggest that many cognitive domains are influenced by common genetic variants, and that the domains' degree of genetic overlap is strongly predicted by their phenotypic associations. The more genetically proximal traits in the analysis (for example, language reasoning and reading) are better candidates for joint interrogation in future genetic and neurobiological studies; the traits with limited association (for example, emotion identification and nonverbal reasoning) are more likely to reflect largely independent physiological influences.

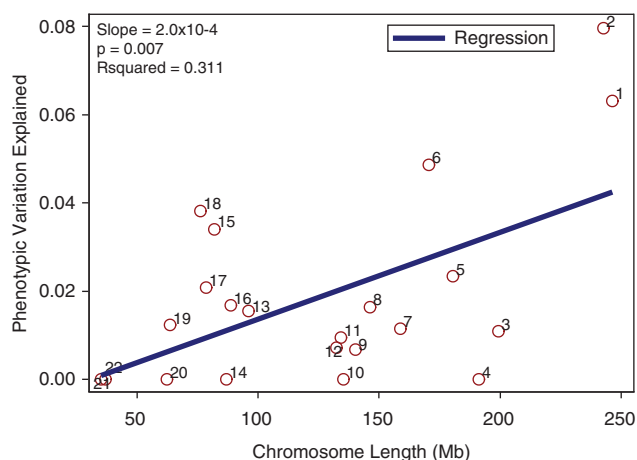
Several aspects of these analyses may introduce variation between the findings reported here and heritability assessed in a



**Table 2.** Genetic correlations between variables with nominally significant univariate heritability

<i>r</i> ( <i>g</i> ) (s.e.)	WRAT	Spatial reasoning	Nonverbal reasoning	Language reasoning	Verbal memory
Spatial reasoning	0.43 (0.16), <i>P</i> =0.01				
Nonverbal reasoning	0.23 (0.16), <i>P</i> =0.1	0.67 (0.15) 0.0003			
Language reasoning	1.00 (0.15), <i>P</i> =2e-07	0.72 (0.17), <i>P</i> =0.0007	0.43 (0.16), <i>P</i> =0.02		
Verbal memory	0.48 (0.20), <i>P</i> =0.01	0.57 (0.23), <i>P</i> =0.009	0.53 (0.21), <i>P</i> =0.008	0.41 (0.23), <i>P</i> =0.05	
Emotion identification	0.40 (0.16), <i>P</i> =0.01	0.26 (0.18), <i>P</i> =0.08	-0.25 (0.20), <i>P</i> =0.09	0.45 (0.20), <i>P</i> =0.01	0.17 (0.22), <i>P</i> =0.2

Abbreviations: *P* = probability of no genetic correlation between the variables; s.e., standard error of genetic correlation; WRAT, Wide Range Achievement Test.



**Figure 1.** Genetic influences on complex cognition are distributed across the genome. Note: Mb = megabases; Rsquared = fraction of phenotypic variation attributable to chromosome length.

different population. First, SNP-based heritabilities, like any association metric, are influenced by the approach taken to assess the phenotype. These results accordingly reflect both the common genetic influences on cognitive domains, as well as the domains' methods of measurement in the PNC. The measures in this analysis that have lower reliability may accordingly have higher estimated heritability in a sample in which the constructs have been assessed differently. While the measures used in this analysis have been previously validated,<sup>28</sup> it will be important to evaluate the genetic influences on these constructs using additional assessment approaches.

In a pediatric population, age of the participants is also likely to introduce variation in SNP-based heritabilities. An association between age and the heritability of IQ has been shown in pediatric samples;<sup>29</sup> specifically, genetic variation appears to have a greater influence on IQ as children mature. The present analyses should accordingly be interpreted in the context of the participants' age distribution, and mean age of 13.7 years. Despite the broad age distribution of the PNC sample, the univariate heritability results we present for the common factor score, and the domains that are most correlated with intelligence (for example, reasoning), are similar to those reported for IQ in pediatric populations of similar mean age.<sup>5</sup>

The extent to which the genetic correlations may be associated with age is less clear. The phenotypic correlations were strongly associated with the genetic correlations in this analysis; age-based

stability in the genetic correlations is accordingly suggested by the consistent phenotypic associations we report across age groups. Further, both twin and molecular genetic studies have suggested that, despite increasing heritability estimates, the genetic influences on intelligence are largely stable over time.<sup>6,30</sup> If the genetic influences on cognition are themselves stable, it would follow for the genetic correlations to be relatively stable as well.

While this study introduces the largest data set available with this array of cognitive assessments, it was limited by the size of the standard errors in both the univariate heritability and genetic correlation analyses. The strong relationship between the phenotypic and genetic correlations suggests a nonrandom ranking of genetic correlations. However, the standard errors associated with the genetic correlation estimates make them difficult to compare statistically, either independently or as a set. After multiple testing correction, it is unlikely that many of the genetic correlations would be significantly different from the others specified in Table 2. The results should accordingly be primarily used to highlight pairs of variables with high degrees of genetic sharing (for example, reading and language reasoning) or variables with strong common genetic influence (for example, the reasoning and executive function principal component).

These analyses were also limited by the need to reduce the sample to individuals of European ancestry. Future studies will be needed to assess the extent to which the present findings extend to non-European populations. Further analysis will also be necessary to replicate the present findings and examine these patterns for age and sex effects. This sample was not large enough to stratify by age or sex, as we would not have had the statistical power necessary to meaningfully compare any group-specific estimates.

These analyses are among the first to use a molecular genetic approach to evaluate a system of complex cognitive traits in a pediatric sample. The development of genetic relationship matrices among cognitive and behavioral phenotypes introduces an opportunity to more effectively study those traits, and gain insight into the neuropsychiatric disorders existing at the tail of the impairment distributions.<sup>31</sup>

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGMENTS

We greatly appreciate the efforts of the participants and families involved in the Philadelphia Neurodevelopmental Cohort. The PNC was supported by NIH RC2 grants MH089983 and MH089924. EBR was funded by grant 1K01MH099286-01A1 from the National Institutes of Mental Health.

## REFERENCES

- 1 Wadsworth SJ, Corley RP, DeFries JC. Cognitive abilities in childhood and adolescence. In: Finkel D, Reynolds CA (eds). *Behavior Genetics of Cognition Across the Lifespan*. Springer: New York, NY, 2014, pp 3–40.
- 2 Plomin R, DeFries JC, Knopik VS, Neiderhiser JM. *General Cognitive Abilities. Behavioral Genetics*, 6th edn. Worth: New York, NY, 2013.
- 3 Plomin R, DeFries JC, Knopik VS, Neiderhiser JM. *Specific Cognitive Abilities. Behavioral Genetics*, 6th edn. Worth: New York, NY, 2013.
- 4 Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald D *et al*. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psychiatry* 2011; **16**: 996–1005.
- 5 Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJ *et al*. Childhood intelligence is heritable, highly polygenic and associated with FBNP1L. *Mol Psychiatry* 2014; **19**: 253–8.
- 6 Deary IJ, Yang J, Davies G, Harris SE, Tenesa A, Liewald D *et al*. Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* 2012; **482**: 212–215.
- 7 Plomin R, Haworth CM, Meaburn EL, Price TS, Davis OS. Common DNA markers can account for more than half of the genetic influence on cognitive abilities. *Psychol Sci* 2013; **24**: 562–568.
- 8 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76–82.
- 9 Greenwood TA, Braff DL, Light GA, Cadenhead KS, Calkins ME, Dobie DJ *et al*. Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. *Arch Gen Psychiatry* 2007; **64**: 1242–1250.
- 10 Stins JF, de Sonneville LM, Groot AS, Polderman TC, van Baal CG, Boomsma DI. Heritability of selective attention and working memory in preschoolers. *Behav Genet* 2005; **35**: 407–416.
- 11 Rietveld MJ, Hudziak JJ, Bartels M, van Beijsterveldt CE, Boomsma DI. Heritability of attention problems in children: longitudinal results from a study of twins, age 3 to 12. *J Child Psychol Psychiatry* 2004; **45**: 577–588.
- 12 Lundstrom S, Chang Z, Rastam M, Gillberg C, Larsson H, Anckarsater H *et al*. Autism spectrum disorders and autisticlike traits: similar etiology in the extreme end and the normal variation. *Arch Gen Psychiatry* 2012; **69**: 46–52.
- 13 Gur RE, Nimgaonkar VL, Almasy L, Calkins ME, Ragland JD, Pogue-Geile MF *et al*. Neurocognitive endophenotypes in a multiplex multigenerational family study of schizophrenia. *Am J Psychiatry* 2007; **164**: 813–819.
- 14 Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 2012; **13**: 537–551.
- 15 Trzaskowski M, Davis OS, DeFries JC, Yang J, Visscher PM, Plomin R. DNA evidence for strong genome-wide pleiotropy of cognitive and learning abilities. *Behav Genet* 2013; **43**: 267–273.
- 16 Kovas Y, Plomin R. Generalist genes: implications for the cognitive sciences. *Trends Cogn Sci* 2006; **10**: 198–203.
- 17 Trzaskowski M, Shakeshaft NG, Plomin R. Intelligence indexes generalist genes for cognitive abilities. *Intelligence* 2013; **41**: 560–565.
- 18 Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH *et al*. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.
- 19 Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI *et al*. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
- 20 Gur RC, Richard J, Calkins ME, Chivacci R, Hansen JA, Bilker WB *et al*. Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology* 2012; **26**: 251–265.
- 21 Smith TD, Smith BL. Relationship between the Wide Range Achievement Test 3 and the Wechsler Individual Achievement Test. *Psychol Rep* 1998; **83**(3 Pt 1): 963–967.
- 22 Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM *et al*. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 2011; **43**: 519–525.
- 23 Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA *et al*. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
- 24 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 25 Inc SI. *CSAS/STAT<sup>®</sup> 9.3 User's Guide*. Care. SAS Institute Inc: Cary, NC, 2011.
- 26 Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME *et al*. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* 2012; **44**: 247–250.
- 27 Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A *et al*. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 2012; **485**: 242–245.
- 28 Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB *et al*. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Meth* 2010; **187**: 254–262.
- 29 Haworth CM, Wright MJ, Luciano M, Martin NG, de Geus EJ, van Beijsterveldt CE *et al*. The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol Psychiatry* 2010; **15**: 1112–1120.
- 30 Hoekstra RA, Barterls M, Boomsma DI. Longitudinal genetic study of verbal and nonverbal IQ from early childhood to young adulthood. *Learn Indiv Differ* 2007; **17**: 97–114.
- 31 Plomin R, Haworth CM, Davis OS. Common disorders are quantitative traits. *Nat Rev Genet* 2009; **10**: 872–878.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)