Molecular Psychiatry (2014) 19, 168–174

www.nature.com/mp

IMMEDIATE COMMUNICATION

Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT)

T Lencz[1,2,3], E Knowles[4], G Davies[5,6,7], S Guha[1], DC Liewald[5,6], JM Starr[5,8], S Djurovic[9,10], I Melle[9,10,11], K Sundet[9,11], A Christoforou[12,13], I Reinvang[14], S Mukherjee[1,2], Pamela DeRosse[1,2], A Lundervold[15 16 17], VM Steen[12,13], M John[1,2], T Espeseth[14,18], K Räikkönen[19,20], E Widen[21], A Palotie[21,22,23], JG Eriksson[24,25,26,27,28], I Giegling[29], B Konte[29], M Ikeda[30], P Roussos[31], S Giakoumaki[32], KE Burdick[31], A Payton[33], W Ollier[33], M Horan[34], G Donohoe[35], D Morris[35], A Corvin[35], M Gill[35], N Pendleton[36], N Iwata[30], A Darvasi[37], P Bitsios[38], D Rujescu[29], J Lahti[19,20], SL Hellard[12,13], MC Keller[39], OA Andreassen[9,10,11], IJ Deary[5,6], DC Glahn[4] and AK Malhotra[1,2,3]

It has long been recognized that generalized deficits in cognitive ability represent a core component of schizophrenia (SCZ), evident before full illness onset and independent of medication. The possibility of genetic overlap between risk for SCZ and cognitive phenotypes has been suggested by the presence of cognitive deficits in first-degree relatives of patients with SCZ; however, until recently, molecular genetic approaches to test this overlap have been lacking. Within the last few years, large-scale genome-wide association studies (GWAS) of SCZ have demonstrated that a substantial proportion of the heritability of the disorder is explained by a polygenic component consisting of many common single-nucleotide polymorphisms (SNPs) of extremely small effect. Similar results have been reported in GWAS of general cognitive ability. The primary aim of the present study is to provide the first molecular genetic test of the classic endophenotype hypothesis, which states that alleles associated with reduced cognitive ability should also serve to increase risk for SCZ. We tested the endophenotype hypothesis by applying polygenic SNP scores derived from a large-scale cognitive GWAS meta-analysis (~5000 individuals from nine nonclinical cohorts comprising the Cognitive Genomics consorTium (COGENT)) to four SCZ case-control cohorts. As predicted, cases had significantly lower cognitive polygenic scores compared to controls. In parallel, polygenic risk scores for SCZ were associated with lower general cognitive ability. In addition, using our large cognitive meta-analytic data set, we identified nominally significant cognitive associations for several SNPs that have previously been robustly associated with SCZ susceptibility. Results provide molecular confirmation of the genetic overlap between SCZ and general cognitive ability, and may provide additional insight into pathophysiology of the disorder.

Molecular Psychiatry (2014) **19,** 168–174; doi:10.1038/mp.2013.166; published online 17 December 2013

Keywords: endophenotype; general cognitive ability; GWAS; polygenic; schizophrenia

[1]Division of Psychiatry Research, Zucker Hillside Hospital, Glen Oaks, NY, USA; [2]Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset, NY, USA; [3]Hofstra North Shore–LIJ School of Medicine, Departments of Psychiatry and Molecular Medicine, Hempstead, NY, USA; [4]Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA; [5]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; [6]Department of Psychology, University of Edinburgh, Edinburgh, UK; [7]Medical Genetics Section, University of Edinburgh Molecular Medicine Centre and MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK; [8]Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK; [9]NorMent, KG Jebsen Centre, Oslo, Norway; [10]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; [11]Psychosis Research Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway; [12]K.G. Jebsen Centre for Psychosis Research, Dr. Einar Martens Research Group for Biological Psychiatry, Department of Clinical Medicine, University of Bergen, Bergen, Norway; [13]Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway; [14]Department of Psychology, University of Oslo, Oslo, Norway; [15]K.G. Jebsen Centre for Research on Neuropsychiatric Disorders, University of Bergen, Bergen, Norway; [16]Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway; [17]Kavli Research Centre for Aging and Dementia, Haraldsplass Deaconess Hospital, Bergen, Norway; [18]K.G. Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; [19]Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland; [20]Institute of Genetics, Folkhälsan Research Centre, Helsinki, Finland; [21]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland; [22]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK; [23]Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland; [24]National Institute for Health and Welfare, Helsinki, Finland; [25]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; [26]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland; [27]Folkhälsan Research Centre, Helsinki, Finland; [28]Vasa Central Hospital, Vasa, Finland; [29]Department of Psychiatry, University of Halle, Halle, Germany; [30]Department of Psychiatry, School of Medicine, Fujita Health University, Toyoake, Japan; [31]Department of Psychiatry, The Mount Sinai School of Medicine, New York, NY, USA; [32]Department of Psychology, School of Social Sciences, University of Crete, Heraklion, Greece; [33]Centre for Integrated Genomic Medical Research, University of Manchester, Manchester, UK; [34]School of Community-Based Medicine, Neurodegeneration Research Group, University of Manchester, Manchester, UK; [35]Neuropsychiatric Genetics Research Group, Department of Psychiatry and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland; [36]Institute of Brain, Behaviour and Mental Health, University of Manchester, Manchester, UK; [37]Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel; [38]Department of Psychiatry and Behavioral Sciences, Faculty of Medicine, University of Crete, Heraklion, Greece and [39]Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA. Correspondence: Dr AK Malhotra, Division of Psychiatry Research, Zucker Hillside Hospital, 75-59 263rd Street, Glen Oaks 11004, NY, USA.
E-mail: malhotra@lij.edu
Received 16 July 2013; revised 20 September 2013; accepted 24 October 2013; published online 17 December 2013

INTRODUCTION

Schizophrenia (SCZ) is a brain disorder characterized by, on average, a reduction in general cognitive abilities of >1 s.d. below the population mean. Given the long-standing observation of subtle cognitive abnormalities in undiagnosed and unmedicated relatives of patients with SCZ, it has been suggested that cognitive deficits may serve as an endophenotype, permitting identification of SCZ risk genes using a quantitative phenotype more closely reflecting gene function.[1] Extensive family and twin data support the role of shared additive genetic factors underpinning both SCZ and cognitive deficits.[2] A recent population-scale study of siblings and twins further suggested that the overlap between these phenotypes is largely genetic, but questioned the overall magnitude of the phenotypic correlation.[3] However, these family-based studies have two primary limitations: (1) they lack direct molecular assays of genetic variance and (2) they have relied upon measurement of cognitive abilities in patients with SCZ, which are subject to potential confounds relating to illness process and state.

A direct test of the endophenotype hypothesis would examine molecular genetic variants associated with cognitive performance in the general (not clinically referred) population, in order to see if these variants are also associated with SCZ. However, this approach has not been adequately tested because of the lack of strongly replicated cognition-associated single-nucleotide polymorphisms (SNPs). By contrast, a number of candidate gene studies have applied a 'reverse endophenotype' approach, in which SCZ risk variants are tested for association with cognition in the general population.[4,5]

Large-scale genome-wide association studies (GWAS) of SCZ have demonstrated that a substantial proportion of the heritability of the disorder is explained by a polygenic component consisting of thousands of common SNPs of extremely small effect.[6] Similarly, recent GWAS of general cognitive ability have indicated that a polygenic architecture accounts for a majority of the heritability, consistent with other normally distributed traits in the general population, such as height.[7–9] The relationship of the underlying genetic architecture between two phenotypes can be examined using polygenic score tests, in which large numbers of alleles demonstrating subtle (not statistically significant) levels of association with a given phenotype are combined to produce a genetic risk profile.[10] The association of these alleles in a different cohort (measured on a different phenotype) can then be utilized to estimate the degree of molecular overlap between phenotypes.

Very recently, two reverse endophenotype studies have examined the ability of SCZ polygenic risk scores to predict cognitive abilities in independent cohorts.[11,12] Both studies draw from the work of the Psychiatric GWAS Consortium on SCZ (PGC-SCZ), a large-scale mega-analytic GWAS of 17 Caucasian cohorts encompassing >9000 SCZ cases and >12 000 controls.[13] In one study,[12] polygenic risk scores for SCZ were significantly associated with IQ in a cohort of patients with SCZ and controls, but these associations were not observed in the smaller ($n = 322$) healthy cohort alone. However, in a much larger study,[11] polygenic risk scores for SCZ were found to be significantly associated with general cognitive ability (as assessed by the first principal component of cognitive tests assessing multiple domains) in two moderately sized (total N ~ 1500) cohorts of psychiatrically normal older adults.

In this study, we sought to test the endophenotype hypothesis by comparing SCZ cases to controls on polygenic SNP scores derived from a large-scale meta-analysis of cognitive GWAS. We focused on general cognitive ability as the primary endophenotype for several reasons that we have discussed in detail previously.[14] Although a variety of specific cognitive traits have been proposed as SCZ endophenotypes, research to date has failed to conclusively demonstrate any specific cognitive signature of SCZ.[15] Overall, deficits in general cognitive ability follow the pattern required of an endophenotype[1]: they are observed in SCZ patients before illness onset, are largely independent of clinical state and medication status, and are familial in nature.[16,17]

It has been recognized for more than 100 years that most cognitive abilities in humans are not orthogonal traits, but instead tend to covary; the underlying factor accounting for this pattern of intercorrelation is termed general cognitive ability (or Spearman's *g*).[18] General cognitive ability is a robust phenotype, accounting for nearly half of the variance on the broad range of specific neurocognitive traits identifiable in humans.[19] Moreover, general cognitive ability in the population is itself highly heritable,[9,20,21] and heritability studies have demonstrated substantial genetic overlap across most specific cognitive domains tested.[18,19,21] Crucially, general cognitive ability can be reliably estimated by extracting the first principal component of any appropriately diverse set of neurocognitive test scores, regardless of the specific components of the battery; it has been empirically demonstrated that estimates of *g* derived from varying batteries tend to be very highly correlated, with correlation coefficients often approaching 1.[22,23]

The present study represents the first empirical report of an international collaborative effort entitled, 'The Cognitive Genomics consorTium (COGENT)'.[14] COGENT aims to bring together human genetic data sets with both: (1) high-density genome-wide genotype data and (2) phenotype data on cognitive function in individuals drawn from the general population. At the time of the first data freeze, COGENT consists of nine sites across seven countries with approximately 5000 individuals with available genotype and phenotype data. Although genotyping platforms and phenotype measures vary by site, genetic imputation and factor analysis of cognitive scores were used to harmonize data across sites. Because generalized cognitive ability (*g*) can be robustly estimated from a variety of test batteries, we were able to perform meta-analysis of GWAS associations to *g* across the nine COGENT cohorts. From this meta-analysis, we derived polygenic allele scores associated with general cognitive ability. These allele scores were then applied to four SCZ case-control cohorts consisting of more than 11 000 independently ascertained subjects (>5000 cases and >5000 controls), as described in detail below. We in addition performed 'reverse endophenotype' analyses, examining the effects of SCZ risk alleles (derived from PGC-SCZ) on cognitive scores in the nine COGENT cohorts.

MATERIALS AND METHODS

Subjects–cognitive GWAS cohorts

Volunteers for cognitive studies were drawn from nine cohorts, for which study investigators agreed to share data as part of the COGENT. Details on subject recruitment procedures for each cohort are described in the Supplementary Materials; summaries of each cohort are presented in Table 1. Although screening procedures differed somewhat across cohorts, subjects were drawn from the general population, either as epidemiologically representative cohorts or as recruited control cohorts for studies of SCZ and/or other mental illnesses. All subjects were of Caucasian descent (as confirmed by principal components analysis (PCA) of genetic data). All subjects provided written, informed consent to protocols approved by their institutional ethics boards in accordance with the Helsinki declaration.

Subjects–SCZ case-control cohorts

The primary test of the endophenotype hypothesis was performed in the Molecular Genetics of Schizophrenia (MGS) European–American case-control cohort. This data set was selected for several reasons: it is large (n>5000), publicly available, has been extensively studied,[24–26] and contains an ethnic distribution that is comparable to nine COGENT cohorts (primarily Northern European in ancestry but with a non-negligible Southern European component as well). To replicate and extend our findings, we secondarily tested three additional SCZ case-control cohorts

Table 1. Description of COGENT cohorts

Data set[a]	*Genotyping platform*	*Concordance*[b]	*Genotypes after QC*	*N*	*% Male*	*Mean age (s.d.)*	*Lambda*[c]
Germany	Illumina OE[d]	99.50%	1,078,289	594	51%	54.0 (15.0)	1.01
LOGOS	Illumina OE	99.40%	835,287	802	100%	22.3 (3.8)	1.03
IBG	Affymetrix 6.0	99.59%	938,800	299	77%	15.9 (1.5)	1.00
LBC1936	Illumina 610	99.60%	1,058,722	1005	51%	69.5 (0.8)	1.01
TOP	Affymetrix 6.0	94.23%	917,315	351	48%	34.2 (9.8)	1.01
NCNG	Illumina 610	99.40%	944,135	629	32%	47.6 (18.3)	1.00
Manchester	Illumina 610	99.60%	1,059,916	697	30%	67.7 (2.8)	1.01
HBCS	Illumina 610	99.60%	1,043,380	318	100%	67.7 (2.3)	1.00
ZHH	Illumina OE	99.40%	1,043,785	201	47%	39.1 (1.8)	1.06

Abbreviations: COGENT, Cognitive Genomics consorTium; GWAS, genome-wide association studies; QC, quality control. [a]Detailed descriptions of each cohort provided in Supplementary Text. [b]Concordance between imputed and genotyped SNPs. [c]Lambda to refers λ_{GC}, a measure of the degree of statistical inflation in GWAS. [d]OE refers to the Illumina OmniExpress genotyping bead chip.

Table 2. Description of SCZ case-control cohorts

SCZ data set	N *cases*	N *controls*	*GWAS platform*
MGS European–American, Shi *et al.*[25]	2681	2653	Affymetrix 6.0
Japan, Ikeda *et al.*[27]	575	564	Affymetrix 5.0
Ashkenazi Jewish (Israel), Guha et al.[28]	904	1640	Illumina Omni1-Quad
MGS African–American, Shi *et al.*[25]	1286	973	Affymetrix 6.0
Total	5446	5830	

Abbreviations: GWAS, genome-wide association studies; SCZ, schizophrenia.

of varying ethnicities: (1) a Japanese cohort with >1000 subjects;[27] (2) an Ashkenazi Jewish cohort with >2500 subjects;[28] and the African–American subcohort (n>2000) of the MGS sample.[25] Demographic details of these cohorts are presented in Table 2. It should be noted that increasing evidence suggests substantial common architecture of complex traits (including SCZ) across populations,[29,30] but it would still be anticipated that replication samples would demonstrate attenuated results due to residual differences in allele frequencies and effect sizes.[31]

Genotyping, quality control and imputation

As described in detail in the Supplementary Materials, all COGENT subjects were genotyped on one of three microarray platforms: Affymetrix 6.0 (Santa Clara, CA, USA; ~900 K SNPs), Illumina 610 K or Illumina OmniExpress (San Diego, CA, USA; ~770 K SNPs). A standardized quality control (QC) pipeline was applied to each COGENT GWAS data set: SNP call rate >95%; sample call rate >90%; SNP Hardy–Weinberg equilibrium $P>10^{-6}$; and X chromosome sex match with reported gender. For any pair of subjects with cryptic relatedness (pi-hat >0.125 in PLINK (ref. 32) 1.07), the sample with the lower call rate was eliminated. For each data set, a PCA was performed (in SVS 7.7.4, GoldenHelix, Bozeman, MT, USA), and samples demonstrating non-Caucasian ancestry were eliminated.

After QC, all SNPs within a given cohort were strand-aligned to HapMap3 and phased using SHAPEIT[33] before imputation with IMPUTE2.[34] As recently recommended to increase imputation accuracy,[35] a large, cosmopolitan HapMap3 reference panel (n = 1011 individuals from Africa, Asia and Europe, and the Americas) was utilized (except for NCNG, which was the only data set for which imputation was not performed centrally). Because our phenotype is a quantitative trait, we sought to avoid potentially spurious findings introduced by random association of rare alleles with a few extreme scores.[36] Therefore, imputed SNPs receiving a probability call >0.90 were retained and converted to PLINK-format genotype calls. The imputed data were then recleaned using the same call rate and Hardy–Weinberg equilibrium criteria described above; in addition, SNPs with minor allele frequency < 2.5% were dropped. For each cohort, ~3000 randomly selected genotyped SNPs were held out for concordance analysis with imputation results. As shown in Table 1, concordance exceeded 99% for eight COGENT cohorts, and ~1 M SNPs were available for analysis in each cohort.

Neurocognitive assessment

Details of neurocognitive batteries for each cohort are provided in the Supplementary Materials. Although the specific instruments varied across cohorts, each cohort was required to have test scores available across at least three domains of cognitive ability for computation of Spearman's *g*,[18,37] an estimate of general cognitive ability derived from PCA.[22,23] (For one cohort, a validated estimate of general cognitive ability derived from two subscales of the Wechsler Adult Intelligence Scale was utilized.) For each of the cohorts, available measures were entered into PCA and the first unrotated component was extracted. Any variable with more than 5% of missing data was dropped from the analysis. Normality is not a strict requirement of PCA implemented for the purpose of data reduction and so no variable was subject to transformation.[38] Moreover, inspection of box plots indicated that variables were generally normally distributed and no noticeable outliers were observed. In each cohort, as expected based on hundreds of prior studies,[39] the first principal component significantly loaded all measures and accounted for ~40% of the variance on average.

The dependent measure for the cognitive GWAS in each cohort was this first PC score, corrected for the following (using linear regression before GWAS): age, sex, age × sex, age^2, and age^2 × sex, based on consistent evidence demonstrating the presence of both linear and quadratic effects of age on general cognitive ability across the life span.[40]

Statistical analysis

Genome-wide association analysis of the quantitative cognitive phenotype was performed in each COGENT cohort using linear regression (additive model) in SVS7.7.4. As shown in Table 1, lambda (genomic control) values for each cohort were at or near 1, indicating no significant effect of subtle population structure on association results. Fixed effects meta-analysis of β-weights from the linear regression analyses was performed in PLINK 1.07,[32] using data from all available cohorts possessing high-QC genotyped or imputed data for each given SNP. Only SNPs with data available in three or more cohorts were retained. By convention, a positive β-weight for a given allele indicated an additive (allele-dose) relationship in the direction of higher cognitive phenotype scores.

Based on results of the meta-analysis, polygenic scores were computed in PLINK using β-weights of alleles at varying statistical thresholds (nominal $P < 0.10$, 0.20, 0.30, 0.40 and 0.50), following the procedure originally described by Purcell *et al.*[10] For each statistical threshold, the clump procedure in PLINK was utilized to prune the set of SNPs for linkage disequilibrium (using r^2 threshold of 0.50 within a 250-kb window), so as to avoid redundancy of SNPs representing a given association signal.

For each of the five statistical thresholds, the weighted allele scores of each 'clumped' SNP were summed for each subject in each SCZ case-control cohort, thus creating a 'cognitive polygene score.' Thus, each subject in each SCZ case-control data set had a set of five cognitive polygene scores: one for each of the statistical thresholds applied to the original COGENT meta-analytic results. For each SCZ case-control data set, five logistic regression analyses were then used to compare cases and controls on cognitive polygene scores at each threshold. Nagelkerkes'

pseudo R^2 was utilized to reflect estimated percent variance in the SCZ phenotype accounted for by cognitive polygene scores at each threshold.[10]

RESULTS

GWAS results for general cognitive ability

Meta-analytic results of the nine cognitive GWAS cohorts are presented as a Manhattan plot in Figure 1; no SNPs reached genome-wide significance, and the overall lambda was 1.031 (see QQ plot in Supplementary Figure 1). This lambda value is higher than that for any individual cohort (Table 1), indicative of polygenic signal,[41] but is lower than reported in similar recent studies of cognitive ability.[7,9] A list of top SNPs ($P < .001$) emerging from the meta-analysis are presented in Supplementary Table 1.

Primary test of endophenotype hypothesis–cognitive polygenic score analysis

We utilized a large, publicly available, Caucasian SCZ GWAS case-control cohort to test our primary endophenotype hypothesis. Using 'clumped' SNPs at five different thresholds (nominal $P < 0.10$, 0.20, 0.30, 0.40 and 0.50), polygenic cognitive scores were computed for each of 2886 cases and 2056 controls in the MGS/GAIN European–American cohort. As predicted, cases had significantly lower cognitive polygenic scores across each of the five thresholds (*P*-values ranging from 6.56×10^{-6} to 3.73×10^{-7}; see Figure 2). In other words, SCZ cases had fewer alleles associated with good cognitive performance and more alleles associated with poorer cognitive performance in the COGENT meta-analysis. The overall magnitude of the polygenic effect was small ($R^2 < 0.01$), but comparable to the total variance explained by a similar approach applied from one cognitive cohort to another[7] and from a reverse endophenotype study.[11]

As replication, the same cognitive SNPs were tested in three smaller SCZ case-control cohorts of differing ethnicities. Because of the apparent dip in variance explained that appears at $P < 0.2$ in the MGS EA cohort (Figure 1), we applied a $P < 0.3$ threshold in our replication data sets. As depicted in Table 3, nominally significant results were observed in two cohorts (Japanese and Ashkenazi Jewish). Meta-analysis of these results across the four SCZ case-control cohorts yielded a strongly significant effect of cognitive polygenic scores on prediction of SCZ risk ($P = 3.6 \times 10^{-7}$). Very similar results were observed using a clump threshold of $P < 0.5$ (Supplementary Table 2; meta-analytic $P = 3.8 \times 10^{-7}$), and somewhat stronger results were observed when the IBG cohort (the only non-adult COGENT cohort) was removed before calculation of polygenic allele weights (Supplementary Table 3; meta-analytic $P = 4.0 \times 10^{-9}$).

Reverse endophenotype approach–examination of prior SCZ GWAS hits

As an additional examination of the relationship between cognitive GWAS results and the SCZ phenotype, we applied the commonly employed 'reverse endophenotype' approach described

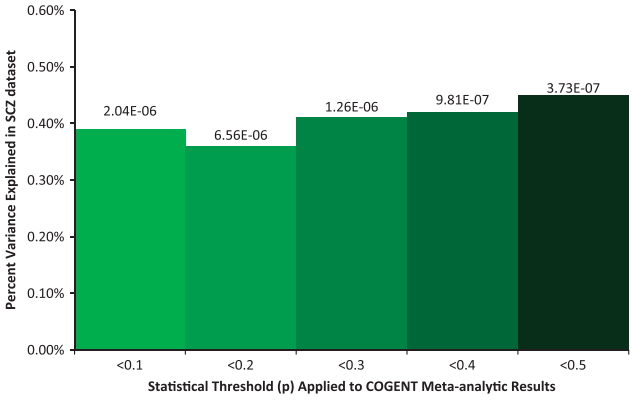

Figure 2. Polygenic overlap between cognitive allele scores (derived from the Cognitive Genomics consorTium (COGENT) meta-analysis thresholded at varying *P*-values) and schizophrenia (SCZ) case-control status in the Molecular Genetics of Schizophrenia (MGS) European–American cohort.

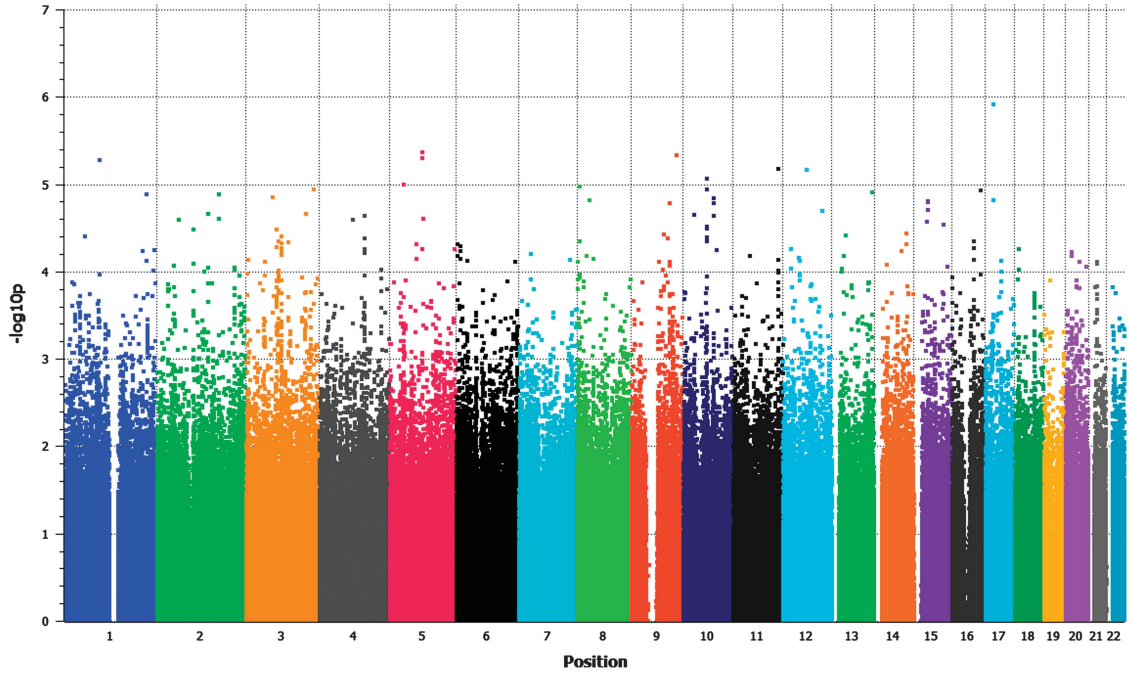

Figure 1. Manhattan plot depicting results of the Cognitive Genomics consorTium (COGENT) meta-analysis.

above. First, we selected all SNPs that have demonstrated genome-wide significance ($P < 5 \times 10^{-8}$) in large-scale (total n>5000) published SCZ GWAS obtained from the NHGRI GWAS catalog (http://www.genome.gov/gwastudies/, downloaded January 10, 2013). As shown in Table 4, four of the 13 independent SNPs (excluding the major histocompatibility complex) associated with SCZ have nominally significant ($P < 0.05$) associations with general cognitive ability in volunteers from the general population. Although effect sizes for these SNPs on cognitive phenotypes are extremely modest, the number of SNPs achieving nominal significance exceeded that expected by chance (binomial test, $P = 0.006$, two-tailed). Although extensive linkage disequilibrium in the major histocompatibility complex has prohibited clear discernment of the source of SCZ GWAS signal, Supplementary Table 4 demonstrates results for five SNPs in the region derived from published SCZ GWAS.

Reverse endophenotype approach–SCZ susceptibility polygenic score analysis

Finally, we performed a polygenic score analysis, comparable to the one presented in Table 2, but in the 'reverse' direction. Specifically, we downloaded "clumped" SNPs derived from the publicly available PGC[13] data set (https://pgc.unc.edu/Sharing.php#SharingOpp). Because polygenic score approaches to SCZ have tended to demonstrate increasing variance explained at higher *P*-value thresholds,[10] we utilized a threshold of $P < 0.50$ on the resulting clumped SNPs. These polygenic scores were then computed for each subject in each of the nine COGENT cohorts separately; scores were then compared with cognitive ability (*g*) using linear regression in each cohort. As shown in Table 5, SCZ-derived polygenic scores were correlated with cognitive ability in the predicted direction (greater SCZ load associated with lower cognitive scores) in seven of the nine cohorts, with three cohorts demonstrating nominally significant (or nearly so) results. As with the primary analysis, total variance accounted for was small (R^2 range from 0 to 2%) but strongly significant ($P = 1.4 \times 10^{-4}$) in the combined analysis.

DISCUSSION

To our knowledge, this is the first study to present molecular genetic evidence supporting general cognitive ability as a true endophenotype for SCZ susceptibility. In a large Caucasian SCZ case-control cohort, we demonstrated that a set of polygenic alleles associated with lower general cognitive ability strongly predicted increased likelihood for the disorder. Consistent, though less strong, associations were observed in additional cohorts,

Table 3. Polygenic overlap between cognitive alleles (derived from COGENT meta-analysis using $P < 0.3$ threshold) and SCZ in four case-control cohorts

SCZ data set	*# overlapping SNPs*	*R^2 for SCZ*	P-*value*	*Direction*
MGS European–American	17,237	0.41%	1.3×10^{-6}	Negative
Japan	6468	0.38%	0.039	Negative
Ashkenazi Jewish	15,151	0.16%	0.041	Negative
MGS African–American	17,382	0.00%	0.958	Positive
Meta-analysis			3.6×10^{-7}	Negative

Abbreviations: COGENT, Cognitive Genomics consorTium; MGS, Molecular Genetics of Schizophrenia; SCZ, schizophrenia; SNPs, single-nucleotide polymorphisms.

Table 5. Polygenic overlap between schizophrenia risk alleles (derived from PGC meta-analysis using $P < 0.5$ threshold) and general cognitive ability in nine COGENT cohorts

COGENT data set	*No. of overlapping SNPs*	*R^2 for g*	P-*value*	*Direction*
Germany	89,360	0.08%	0.475	Positive
LOGOS	69,069	0.00%	0.985	Positive
IBG	89,353	0.04%	0.72	Negative
LBC1936	96,820	1.17%	0.0006	Negative
TOP	88,946	0.10%	0.5481	Negative
NCNG	87,934	0.04%	0.5925	Negative
Manchester	96,907	1.61%	0.0007	Negative
HBSC	93,890	0.11%	0.5467	Negative
ZHH	85,681	1.90%	0.0532	Negative
Meta-analysis			1.4×10^{-4}	Negative

Abbreviations: COGENT, Cognitive Genomics consorTium; GWAS, genome-wide association studies; PGC, psychiatric GWAS Consortium.

Table 4. Examination of cognitive associations (in COGENT meta-analysis) for SNPs identified in published GWAS for schizophrenia (excluding major histocompatibility complex)

Source	*SNP*	*Region*	*Gene*	*Risk-allele frequency*[a]	*OR_Sz*[a]	*COGENT studies*	*P_COGENT meta-analysis*
Ripke (PGC) Nat Gen 2011	rs1625579	1p21.3	MIR137	0.8	1.12	8	0.4789
Shi (China) Nat Gen 2011	rs10489202	1q24.2	MPC2	0.141	1.23	9	0.9024
Ripke (PGC) Nat Gen 2011	rs6703335	1q43	SDCCAG8	0.56	1.09	6	0.9127
O'Donovan Nat Gen 2011	rs1344706	2q32.1	ZNF804A	0.59	1.12	9	0.4048
Ripke (PGC) Nat Gen 2011	rs17662626	2q32.3	PCGEM1	0.91	1.2	5	0.4509
Bergen (Swe) Mol Psy 2012	rs7709645	5q12.1	ZSWIM6	0.475	1.11	9	0.7207
Bergen (Swe) Mol Psy 2012[b]	rs12666575	7p22.3	MAD1L1	0.673	1.12	8	0.0320
Ripke (PGC) Nat Gen 2011	rs10503253	8p23.2	CSMD1	0.19	1.16	7	0.5874
Shi (China) Nat Gen 2011[c]	rs16887244	8p11.23	LSM1	0.683	1.19	9	0.0171
Ripke (PGC) Nat Gen 2011[b]	rs7914558	10q24.32	CNNM2	0.59	1.10	9	0.0368
Ripke (PGC) Nat Gen 2011	rs11191580	10q24.33	NT5C2	0.91	1.15	9	0.6575
Decode Nature 2009[c]	rs12807809	11q24.2	NRGN	0.83	1.15	8	0.0399
Ripke (PGC) Nat Gen 2011	rs12966547	18q21.2	TCF4	0.58	1.09	8	0.7327

Abbreviations: COGENT, Cognitive Genomics consorTium; GWAS, genome-wide association studies; PGC, psychiatric GWAS Consortium; SNPs, single-nucleotide polymorphisms. [a]Based on source publication. [b]Schizophrenia risk allele associated with lower cognitive ability. [c]Schizophrenia risk allele associated with higher cognitive ability.

despite the fact that they were drawn from populations of differing ethnicities.

Although the statistical evidence (*P*-values ranging from ~10^{-6} to ~10^{-9}) for association was strong, the overall amount of variance explained, even in the Caucasian SCZ cohort (~0.5%) was modest, a result that must be interpreted in the context of other studies of polygenic overlap. For example, in a recent study of cognition,[7] polygenic SNP scores accounted for only ~1% of the variance in a separate cognitive cohort (these cohorts are now included in COGENT). Nevertheless, these authors were able to demonstrate that nearly half of the variation in cognitive ability could ultimately be accounted for by common SNPs. Similarly, the initial study of polygenic effects in SCZ identified only ~2–3% overlap between any two SCZ cohorts, and ~1–2% overlap between SCZ and bipolar cohorts.[10] Again, extensive simulations demonstrated that the polygenic SNPs tagged common genetic variation accounting for approximately one-third of the total variance in SCZ risk; this estimate has been replicated, indicating that the empirically observed cross-sample allelic overlap is a substantial under-estimate of the total polygenic effect.[6] Moreover, observed polygenic overlap tends to rapidly increase as a function of sample size; for example, the empirically observed variance explained by polygenic effects in SCZ has increased 10-fold with larger sample sizes.[42] Although the COGENT cohort represents the largest genetic sample of cognition to date, sample size remains small relative to polygenic studies of other quantitative traits such as height.[43,44]

Thus, the results of the present study are consistent with a model in which a substantial fraction of the molecular basis of general cognitive ability is shared with genetic risk for SCZ. These results are consistent with a large body of evidence from family-based studies that use twin and sibling correlations to model genetic effects.[2] Notably, a recent population-based study using similar modeling strategies[3] found only limited phenotypic overlap between intelligence and psychosis; even so, the source of this overlap was estimated to be largely (~90%) comprised of additive genetic variation. The present study is unique in directly testing molecular genetic variation, and utilizing non-clinical volunteers for the estimation of the cognitive genetic component. Our demonstration of parallel effects when examining genome-wide overlap in the reverse direction (SCZ risk alleles predicting cognitive scores, Table 5) add further confidence to our conclusions, and are also consistent with a recent genome-wide reverse endophenotype study, which demonstrated significant overlap between polygenic SCZ risk alleles and cognitive decline in two aging cohorts.[11]

In addition, following the conventional 'reverse endophenotype' approach, multiple GWAS identified SCZ risk SNPs (4/13, or 31%) demonstrated nominal evidence of association to general cognitive ability in a large meta-analytic cohort of nonclinical volunteers. Three of these are intronic SNPs, in *MAD1LI, LSM1* and *CNNM2*, and the present study represents the first report of human neurocognitive correlates of variants at these loci. Little is known about the functions of these genes in the central nervous system, and the structural and functional properties of their associated proteins vary widely. The fourth nominally significant locus is < 5 kb, 5′ to *NRGN* (and is in a linkage disequilibrium block encompassing the gene). *NRGN* encodes neurogranin, a well-characterized postsynaptic protein that binds to calmodulin and thereby modulates postsynaptic calcium signaling.[45] Although this locus did not show significant association with cognitive variables in a smaller prior study,[46] neuroimaging studies have previously associated this locus with structural and functional variation in the frontal cortex, cingulate and hippocampus.[47–50] Although these results were significant in the aggregate ($P = 0.006$ by binomial test), it should be emphasized that the effect sizes for individual SNPs were extremely small, and *P*-values would not survive Bonferroni correction.

Based on the present results, it is likely that larger sample sizes will afford greater ability to predict SCZ using cognitive polygenic scores. This overlap can then be utilized to enhance SCZ gene detection using recently developed 'pleiotropic enrichment' techniques.[51] At the same time, the study of the genetic basis of normal variation in cognition is likely to further our understanding of the mechanisms by which SCZ risk genes affect the central nervous system.

Several caveats should be placed on the interpretation of this study. First, the present study did not directly evaluate the genetic source of cognitive deficits in patients with SCZ. It could be argued that GWAS of cognitive ability in SCZ cohorts would be required to test whether this putative endophenotype actually mediates the relationship between cognitive polygene score and SCZ risk. However, cognitive performance in patients with SCZ can be influenced by potential confounds such as effects of medication or acute symptomatology, which would tend to attenuate any genetic signal. Consequently, our approach of utilizing nonclinical samples was designed to maximize the potential power of GWAS.

Finally, it should be noted that no genome-wide significant loci for cognition were identified in the present study, despite being the largest GWAS of cognitive ability in predominately adult cohorts. This result was anticipated based on recent large-scale GWAS results for childhood intelligence,[9] as well as early GWAS (with comparable sample size to the present study) of potentially comparable quantitative traits, height[52] and weight.[53] It is also possible that power to detect genetic signals was reduced due to unavoidable heterogeneity in cognitive assessment methods across cohorts; such an interpretation is consistent with the relatively low lambda observed in our study. Future studies, ideally with prospectively collected cohorts utilizing harmonized approaches to phenotype assessment, will be required to tease out genome-wide significant loci for cognitive ability.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work has been supported by grants from the National Institutes of Health (R01 MH079800 and P50 MH080173 to AKM; RC2 MH089964 to TL; R01 MH080912 to DCG; K23 MH077807 to KEB; K01 MH085812 to MCK). Dr Donohoe is generously funded by the Health Research Board (Ireland) and Science Foundation Ireland. Data collection for the TOP cohort was supported by the Research Council of Norway, South-East Norway Health Authority. The NCNG study was supported by Research Council of Norway Grants 154313/V50 and 177458/V50. The NCNG GWAS was financed by grants from the Bergen Research Foundation, the University of Bergen, the Research Council of Norway (FUGE, Psykisk Helse), Helse Vest RHF and Dr Einar Martens Fund. The Helsinki Birth Cohort Study has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation. For the LBC cohort, phenotype collection was supported by Research Into Ageing (continues as part of Age UK The Disconnected Mind project). Genotyping was funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (G0700704/84698). Funding from the BBSRC, Engineering and Physical Sciences Research Council (EPSRC), Economic and Social Research Council (ESRC), and MRC is gratefully acknowledged. We are grateful to investigators, led by Pablo Gejman, who have made data for the MGS cohort publicly available through dbGAP (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000167.v1.p1; /cgi-bin/study.cgi?study_id=phs000021.v2.p1).

REFERENCES

1 Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; **160**: 636–645.

2 Toulopoulou T, Goldberg TE, Mesa IR, Picchioni M, Rijsdijk F, Stahl D *et al.* Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Arch Gen Psychiatry* 2010; **67**: 905–913.
3 Fowler T, Zammit S, Owen MJ, Rasmussen F. A population-based study of shared genetic variation between premorbid IQ and psychosis among male twin pairs and sibling pairs from Sweden. *Arch Gen Psychiatry* 2012; **69**: 460–466.
4 Almasy L. The role of phenotype in gene discovery in the whole genome sequencing era. *Hum Genet* 2012; **131**: 1533–1540.
5 Bilder RM, Howe A, Novak N, Sabb FW, Parker DS. The genetics of cognitive impairment in schizophrenia: a phenomic perspective. *Trends Cogn Sci* 2011; **15**: 428–435.
6 Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME *et al.* Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ), International Schizophrenia Consortium (ISC). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* 2012; **44**: 247–250.
7 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.
8 Plomin R, Haworth CM, Meaburn EL, Price TS, Davis OS. Wellcome Trust Case Control Consortium 2 Common DNA markers can account for more than half of the genetic influence on cognitive abilities. *Psychol Sci* 2013; **24**: 562–568.
9 Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJ *et al.* Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol Psychiatry* advance online publication, 29 January 2013; doi:10.1038/mp.2012.184 (e-pub ahead of print).
10 International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
11 McIntosh AM, Gow A, Luciano M, Davies G, Liewald DC, Harris SE *et al.* Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age. *Biol Psychiatry* 2013; **73**: 938–943.
12 Terwisscha van Scheltinga AF, Bakker SC, van Haren NE, Derks EM, Buizer-Voskamp JE, Cahn W *et al.* Schizophrenia genetic variants are not associated with intelligence. *Psychol Med* 2013; **73**: 525–3.
13 Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
14 Donohoe G, Deary IJ, Glahn DC, Malhotra AK, Burdick KE. Neurocognitive phenomics: examining the genetic basis of cognitive abilities. *Psychol Med* 2012; **43**: 2027–2036.
15 Dickinson D, Ragland JD, Gold JM, Gur RC. General and specific cognitive deficits in schizophrenia: Goliath defeats David? *Biol Psychiatry* 2008; **64**: 823–827.
16 Keefe RS, Harvey PD. Cognitive impairment in schizophrenia. *Handb Exp Pharmacol* 2012; **213**: 11–37.
17 Keshavan MS, Kulkarni S, Bhojraj T, Francis A, Diwadkar V, Montrose DM, Seidman LJ, Sweeney J. Premorbid cognitive deficits in young relatives of schizophrenia patients. *Front Hum Neurosci* 2010; **3**: 62.
18 Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. *Nat Rev Neurosci* 2010; **11**: 201–211.
19 Deary IJ, Johnson W, Houlihan LM. Genetic foundations of human intelligence. *Hum Genet* 2009; **126**: 215–232.
20 Plomin R. The genetics of g in human and mouse. *Nat Rev Neurosci* 2001; **2**: 136–141.
21 Glahn DC, Almasy L, Blangero J, Burk GM, Estrada J, Peralta JM *et al.* Adjudicating neurocognitive endophenotypes for schizophrenia. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 242–249.
22 Johnson W, Bouchard TJ Jr, Krueger RF, McGue M, Gottesman II. Just one g: Consistent results from three test batteries. *Intelligence* 2004; **34**: 95–107.
23 Johnson W, Nijenhuis JT, Bouchard TJ. Jr. Still just 1 g: consistent results from five test batteries. *Intelligence* 2008; **36**: 81–95.
24 Sanders AR, Duan J, Levinson DF, Shi J, He D, Hou C *et al.* No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. *Am J Psychiatry* 2008; **165**: 497–506.
25 Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I *et al.* Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**: 753–757.
26 Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J *et al.* Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. *Am J Psychiatry* 2011; **168**: 302–316.
27 Ikeda M, Aleksic B, Kinoshita Y, Okochi T, Kawashima K, Kushima I *et al.* Genome-wide association study of schizophrenia in a Japanese population. *Biol Psychiatry* 2011; **69**: 472–478.
28 Guha S, Rees E, Darvasi A, Ivanov D, Ikeda M, Bergen SE *et al.* Implication of a rare deletion at distal 16p11.2 in schizophrenia. *JAMA Psychiatry* 2013; **70**: 253–260.
29 Fu J, Festen EA, Wijmenga C. Multi-ethnic studies in complex traits. *Hum Mol Genet* 2011; **20**: R206–R213.
30 de Candia TR, Lee SH, Yang J, Browning BL, Gejman PV, Levinson DF *et al.* Additive genetic variation in schizophrenia risk is shared by populations of african and European descent. *Am J Hum Genet* 2013; **93**: 463–470.
31 N'Diaye A, Chen GK, Palmer CD, Ge B, Tayo B, Mathias RA *et al.* Identification, replication, and fine-mapping of loci associated with adult height in individuals of african ancestry. *PLoS Genet* 2011; **7**: e1002298.
32 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.
33 Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods* 2011; **9**: 179–181.
34 Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3* 2011; **1**: 457–470.
35 Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012; **44**: 955–959.
36 Montgomery SB, Lappalainen T, Gutierrez-Arcelus M, Dermitzakis ET. Rare and common regulatory variation in population-scale sequenced human genomes. *PLoS Genet* 2011; **7**: e1002144.
37 Spearman C. General intelligence objectively determined and measured. *Am J Psychology* 1904; **15**: 201–293.
38 Jackson JE. *A User's Guide to Principal Components*. John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006.
39 Carroll JB. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press: Cambridge, UK, 1993.
40 Verhaeghen P, Salthouse TA. Meta-analyses of age-cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychol Bull* 1997; **122**: 231–249.
41 Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ *et al.* Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* 2011; **19**: 807–812.
42 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kähler AK, Akterin S *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.
43 Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565–569.
44 Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832–838.
45 Papaleo F, Lipska BK, Weinberger DR. Mouse models of genetic effects on cognition: relevance to schizophrenia. *Neuropharmacology* 2012; **62**: 1204–1220.
46 Donohoe G, Walters J, Morris DW, Da Costa A, Rose E, Hargreaves A *et al.* A neuropsychological investigation of the genome wide associated schizophrenia risk variant NRGN rs12807809. *Schizophr Res* 2011; **125**: 304–306.
47 Krug A, Krach S, Jansen A, Nieratschker V, Witt SH, Shah NJ *et al.* The effect of neurogranin on neural correlates of episodic memory encoding and retrieval. *Schizophr Bull* 2013; **39**: 141–150.
48 Ohi K, Hashimoto R, Yasuda Y, Nemoto K, Ohnishi T, Fukumoto M *et al.* Impact of the genome wide supported NRGN gene on anterior cingulate morphology in schizophrenia. *PLoS ONE* 2012; **7**: e29780.
49 Rose EJ, Morris DW, Fahey C, Robertson IH, Greene C, O'Doherty J *et al.* The effect of the neurogranin schizophrenia risk variant rs12807809 on brain structure and function. *Twin Res Hum Genet* 2012; **15**: 296–303.
50 Pohlack ST, Nees F, Ruttorf M, Witt SH, Nieratschker V, Rietschel M, Flor H. Risk variant for schizophrenia in the neurogranin gene impacts on hippocampus activation during contextual fear conditioning. *Mol Psychiatry* 2011; **16**: 1072–1073.
51 Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC *et al.* Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am J Hum Genet* 2013; **92**: 197–209.
52 Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JR *et al.* A common variant of HMGA2 is associated with adult and childhood height in the general population. *Nat Genet* 2007; **39**: 1245–1245.
53 Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007; **316**: 889–894.

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