ORIGINAL ARTICLE
Childhood intelligence is heritable, highly polygenic and associated with FNBP1L

B Benyamin1,2,26, BST Pourcain1,26, OS Davis1,26, G Davies1,26, NK Hansell2, M-JA Brion3,6, RM Kirkpatrick7, RAM Cents8,9, S Franic10, MB Miller1, CMA Haworth4, E Meaburn11, TS Price5, DM Evans3, N Timpson1, J Kemp3, S Ring3, W McArdrile3, SE Medland1, J Yang12, SE Harris13,14, DC Liewald5,14, P Scheet16, X Xiao15, JJ Hudziak16, EJC de Geus10, Wellcome Trust Case Control Consortium 2 (WTCCC2), DI Boomsma10, D Posthuma9,23,24, M McGue7,25, MJ Wright2, G Davey Smith3,27, IJ Deary5,14,27, R Plomin4,27 and PM Visscher1,2,12,14,27

VWV Jaddoe8,17,18, JM Starr14,19, FC Verhulst9, C Pennell6, H Tiemeier9,17,20, WG Iacono7, LJ Palmer21,22, GW Montgomery2, NG Martin2, DI Boomsma10, D Posthuma9,23,24, M McGue7,25, MJ Wright6, G Davey Smith12,7, U Deary5,14,27, R Plomin4,27 and PM Visscher1,2,12,14,27

Intelligence in childhood, as measured by psychometric cognitive tests, is a strong predictor of many important life outcomes, including educational attainment, income, health and longevity, and intelligence in childhood is a predictor of those outcomes.1 Twin, family and adoption studies have shown that intelligence, as measured using validated psychometric cognitive tests (Intelligence Quotient (IQ)-type tests), is one of the most heritable behavioural traits.2 These findings have been consistently replicated, but the molecular basis of intelligence remains poorly understood. The supporting evidence from the molecular findings has not been consistent and many reported candidate–gene associations have not been replicated.2 A recent study suggested that most reported associations between candidate genes and intelligence are likely to be false.3 The recent successes of genome-wide association studies (GWAS) for many complex traits, where >1200 genetic variants

INTRODUCTION
Intelligence in human populations is associated with a wide range of important life outcomes, including educational attainment, income, health and longevity, and intelligence in childhood is a predictor of those outcomes.1 Twin, family and adoption studies have shown that intelligence, as measured using validated psychometric cognitive tests (Intelligence Quotient (IQ)-type tests), is one of the most heritable behavioural traits.2 These findings have been consistently replicated, but the molecular basis of intelligence remains poorly understood. The supporting evidence from the molecular findings has not been consistent and many reported candidate–gene associations have not been replicated.2 A recent study suggested that most reported associations between candidate genes and intelligence are likely to be false.3 The recent successes of genome-wide association studies (GWAS) for many complex traits, where >1200 genetic variants

1The University of Queensland, Queensland Brain Institute, St Lucia, Queensland, Australia; 2Queensland Institute of Medical Research, Brisbane, Queensland, Australia; 3Medical Research Council Centre for Causal Analyses in Translational Epidemiology, University of Bristol, Bristol, UK; 4King’s College London, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London, UK; 5Department of Psychology, University of Edinburgh, Edinburgh, Scotland, UK; 6School of Women’s and Infants’ Health, The University of Western Australia, Perth, Western Australia, Australia; 7Department of Psychology, University of Minnesota, St Paul, MN, USA; 8The Generation R Study Group, Erasmus MC-University Medical Centre Rotterdam, Rotterdam, The Netherlands; 9Department of Child and Adolescent Psychiatry, Erasmus MC-University Medical Centre Rotterdam, Rotterdam, The Netherlands; 10Netherlands Twin Register, Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; 11Department of Psychology, Birbeck University of London, London, UK; 12The University of Queensland Diamantina Institute, Princess Alexandra Hospital, Brisbane, Queensland, Australia; 13Molecular Medicine Centre, Institute for Genetics and Molecular Medicine Centre, University of Edinburgh, Edinburgh, UK; 14Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; 15Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, TX, USA; 16Department of Psychiatry, College of Medicine, University of Vermont, Burlington, VT, USA; 17Department of Epidemiology, Erasmus MC-University Medical Centre Rotterdam, Rotterdam, The Netherlands; 18Department of Pediatrics, Erasmus Medical Center, Rotterdam, Rotterdam, The Netherlands; 19Alzheimer Scotland Dementia Research Centre, Department of Psychology, University of Edinburgh, Edinburgh, Scotland, UK; 20Department of Psychiatry, Erasmus MC-University Medical Centre Rotterdam, Rotterdam, The Netherlands; 21Genetic Epidemiology and Biostatistics Platform, Ontario Institute for Cancer Research, University of Toronto, Toronto, Ontario, Canada; 22Samuel Lunenfeld Research Institute, University of Toronto, Toronto, Ontario, Canada; 23Department of Functional Genomics, Center for Neurogenomics and Cognitive Research (CNCR), Neuroscience Campus Amsterdam (NCA), VU University Amsterdam and VU Medical Centre, Amsterdam, The Netherlands; 24Department of Clinical Genetics, Section Medical Genomics, VU Medical Centre, Amsterdam, The Netherlands and 25Department of Epidemiology, University of Southern Denmark, Odense, Denmark. Correspondence: Dr PM Visscher, Queensland Brain Institute, University of Queensland, QBI Building (#79), St Lucia, Queensland 4072, Australia.
E-mail: peter.visscher@uq.edu.au
26These authors contributed equally to this work.
27These authors contributed equally to this work.
Received 25 May 2012; revised 28 October 2012; accepted 12 November 2012
have been associated with complex traits, have not been achieved for behavioural traits, including intelligence. The most plausible reason for this failure is that the effect size of individual genetic variants is so small that the current experimental sample sizes are not large enough for detection. For example, using ~3500 individuals, Davies et al. did not find any genome-wide significant single-nucleotide polymorphism (SNPs) associated with intelligence in adults. However, when the combined effects of SNPs were analysed simultaneously, they found that common SNPs accounted for 40–51% of the variation for different measures of intelligence.

Whereas intelligence shows remarkable stability throughout the life course, age-to-age change also occurs. A correlation between intelligence score from age 11 to age 17 of 0.63 has been reported. It has also been consistently shown that heritability increases from childhood ($h^2 = 0.41$) to young adulthood ($h^2 = 0.66$). To date, several GWAS results on intelligence have been reported but the only published GWAS results for childhood intelligence were based on DNA pooling, where SNP genotyping was carried out on pools made of DNA from many individuals.

Understanding individual differences in childhood intelligence can contribute to dissecting its observed association with important outcomes later in life. The aim of this study is to elucidate the genetic and environmental bases of childhood intelligence by identifying associated genetic variants and estimating their contribution to the variation in childhood intelligence.

MATERIALS AND METHODS

Participants

We established the CHIC (Childhood Intelligence Consortium) to combine our efforts in elucidating the genetic and environmental bases of childhood intelligence. CHIC currently consists of six discovery ($N = 12,441$) and three replication ($N = 5548$) cohorts with a total sample size of 17,989 children of European ancestry for whom genome-wide SNP genotypes and intelligence scores are available (Table 1). The discovery cohorts are the Avon Longitudinal Study of Parents and Children (ALSPAC, $N = 5517$), the Lothian Birth Cohorts (LBC1921, $N = 464$; LBC1936, $N = 947$), the Brisbane Adolescent Twin Study, Queensland Institute of Medical Research (QIMR, $N = 1752$), the Western Australian Pregnancy Cohort Study (Raine, $N = 936$) and the Twins Early Development Study (TEDS; $N = 2825$). The three studies that formed the replication cohorts are the Generation Rotterdam study (Generation R, $N = 1442$), the Netherlands Twin Registry study (NTR; $N = 739$) and the University of Minnesota study (UMN; $N = 3367$). Each study obtained ethical approval from the relevant institution. The age of the children ranged between 6 and 18 years. More details about the cohorts are provided in the Supplementary Note of the Supplementary Online Information.

Intelligence measure

We used the best available measure of general cognitive ability (g) or intelligence quotient (IQi), derived from diverse tests that assess both verbal and non-verbal ability (Table 1). In some studies, this was derived from an IQ-type test; in other studies, it was derived from the first unrotated factor of a factor analysis. Much research has shown that g is robust to the composition of the test battery. More details about the intelligence measure from each cohort are provided in the Supplementary Note.

Genotyping and quality controls

Individual sample quality control. Within each cohort, individuals were removed based on missingness, heterozygosity, relatedness, population and ethnic outliers, and other cohort-specific quality control (QC) steps. There were variations of QC between participating studies as the exact choice of QC thresholds depends on genotyping platform and study. More details on the QCs for each cohort are described in the Supplementary Note.

SNP QC. For the meta-analysis, SNPs were removed based on missingness (call rate <95%), minor allele frequency (<1%), Hardy–Weinberg (P-value < 10$^{-6}$), Mendelian errors (if family data were available) and other QC, such as the mean of GenCall score for Illumina arrays. As part of the QC procedure, we also calculated the average effective sample size (N) per cohort as a function of the allele frequency (p) and the standard error of the effect size (se) from the association test as $N = \frac{1}{1 - \frac{\sum m}{\sum m^2}}$, where $m$ is the number of SNPs and $Rsq$ is the imputation quality score. This formula was derived from linear regression theory (Supplementary Note). This calculated N is a useful measure to check for the consistency of the reported sample size and the actual sample size that was used in the association analysis. We found that the calculated Ns were consistent with the reported Ns in all cohorts (Supplementary Table 1). We also checked for the consistency of the SNP allele frequencies between cohorts (Supplementary Figures 1 and 2).

Statistical analyses

Imputation. To facilitate the meta-analysis, the imputation of unobserved genotypes from the HAPMAP II CEU Panel (Release 22, NCBI Build36, dbSNP b126) was conducted within each cohort. This imputation was conducted on QC-ed data using the positive strand as the reference. We conducted the imputation using either BEAGLE, IMPUTE or MACH. We excluded imputed SNPs when the quality score (IMPUTE) or Rsq (MACH) was <0.3.

Association analysis. The association analysis was performed separately within each cohort. Except for the family data from QIMR, the analysis used the dosage score (the estimated counts of the reference allele in each individual; these estimates could be fractional and ranged from 0.0 to 2.0). An additive model was used on the standardised residuals (Z-score, transformed to normality if the phenotype is highly skewed) of the trait after adjusting for known covariates (age, sex, cohort, etc., including subtle population stratification effects, that is, the first four multi-dimensional
scaling or PC (principal component) scores for each individual from a stratification analysis) on both genotyped and imputed SNPs. Both the directly genotyped and imputed SNPs were aligned to the HapMap reference strand. The Manhattan and Q–Q plots of the association P-values for each discovery cohort are presented in Supplementary Figures 3 and 4.

Meta-analysis. The results for associations between SNPs and childhood intelligence from the discovery samples were meta-analysed in the Meta-analysis. We also conducted this gene-based analysis in each of the independent samples, that is, gene-based analysis is likely to be conservative. A detailed plot of the meta-analysis results from SNPs that survived QC in all cohorts (Total SNPs: 138093). The meta-analysis results from SNPs that survived QC in all cohorts were used for subsequent analyses, that is, gene-based analysis and profile scoring for the genetic prediction analysis. The detailed plot of the most significantly associated SNP in the meta-analysis is presented in Supplementary Figure 5.

Gene-based analysis. By considering all SNPs within a gene as a unit for the association analysis, a gene-based analysis can be a powerful complement to the single SNP–trait association analysis. We performed this gene-based analysis in the Metascore (Queensland Institute of Medical Research, Brisbane, Australia) using the P-values of the association between SNPs and childhood intelligence generated from the meta-analysis. We also conducted this gene-based analysis in each of the replication cohorts. Since there were ~17 000 genes, the genome-wide P-value threshold for declaring statistical significance following the Bonferroni correction for multiple testing was 0.05/17 000 = 3 × 10−6. Given the overlap between genes, the actual number of independent genes tested is likely to be smaller. Therefore, the Bonferroni correction for gene-based analysis is likely to be conservative. A detailed plot of the most significant gene from this gene-based analysis, FNPB1L, is presented in Supplementary Figure 6.

GCTA analysis. We estimated the contribution of all common SNPs on childhood intelligence by performing a linear mixed-model analysis to fit all genotyped SNPs simultaneously in the model, as implemented in the GCTA program. We excluded close relatives in the analysis by removing an individual from a pair where the estimated genetic coefficient of relatedness was > 0.025. One of the reasons for this exclusion is to eliminate bias due to common environmental factors. We conducted this analysis in the three largest cohorts, that is, ALSPAC, TEDS and UMN. The numbers of individuals used for this analysis were 5517, 2794 and 1736 children in the ALSPAC, TEDS and UMN cohorts, respectively.

Genetic prediction analysis. We used the estimates of SNP effect size from the meta-analysis to build a multi-SNP prediction model. We used this model to estimate the proportion of the phenotypic variation in independent samples that is due to genotypic information alone. To do this, we first identified independent SNPs from the meta-analysis using a P-value informed linkage disequilibrium (LD) clumping approach in PLINK with a cutoff of pairwise R² ≤ 0.25 within a 200-KB window. Using this approach, we identified all independent SNPs that are significant at various P-value thresholds (Pt) that is, 0.001, 0.005, 0.01, 0.05, 0.10, 0.25, 0.5 and 1. From groups of SNPs at each Pt threshold, we then calculated a quantitative genetic score or multi-SNP predictors in each of the three independent samples, that is, Generation R, NTR and UMN. We then performed a linear regression analysis between the quantitative genetic score and the observed measure of childhood intelligence, and quantified the precision of the predictor as the R² measure of variance explained in the phenotype by the predictor.

RESULTS AND DISCUSSION

To discover specific genetic variants affecting variation in childhood intelligence, we performed a meta-analysis of the SNPs–intelligence association results from the six discovery cohorts. Within each cohort, we tested for an association between (genotyped and imputed) SNPs that passed stringent quality controls and childhood intelligence, using an additive model. In our meta-analysis, we did not find any SNPs that reached a genome-wide significant P-value of 5 × 10−8 or, equivalent, a SNP that explained >0.24% of phenotypic variation (Figure 1). Q–Q plots and a genomic-control l value of 1.078 are consistent with evidence for population stratification and/or polygenic variation. Population stratification appears unlikely because the phenotype was adjusted for at least the first four PCs or multi-dimensional scaling factors derived from population stratification analysis in each cohort (except for the Raine cohort). Therefore, the association results are consistent with many variants having small effects.

We conducted in silico replication for the top 100 independently significant SNPs identified from the meta-analysis of the discovery samples in three independent samples, the Generation Rotterdam Study (Generation R, N = 1442), the Netherlands Twin Register (NTR, n = 739) and the UMN Study (N = 3367). There were no SNPs that reached nominal significance after a Bonferroni adjustment. When we meta-analysed the results from the discovery and replication cohorts, there were also no SNPs that reached a genome-wide significant threshold (Supplementary Table 2). However, when we plotted the estimated regression coefficients for the top 100 SNPs from the replication samples vs those from the discovery samples, we observed a positive correlation in Generation R (r = 0.340, P = 0.0005) and NTR (r = 0.219, P = 0.028), but not in UMN (r = −0.074, P = 0.46) samples (Supplementary Figure 7). These results demonstrated that the direction of the top 100 SNPs in the replication samples were generally consistent with those from the discovery samples. Again, these results are consistent with small effects with small effects.

To test whether the aggregate effect of a group of SNPs within a gene has a significant effect on childhood intelligence, we conducted a gene-based analysis and found that FNPB1L (formin binding protein 1-like) was the most strongly associated gene with childhood intelligence in the discovery samples (P = 4 × 10−5) (Table 2; Supplementary Table 3). This gene, previously known as Toca-1, encodes a protein that binds to both CDC42 and N-WASP and is involved in a pathway that links cell surface signals to the actin cytoskeleton. This is an interesting finding since a recent study also identified FNPB1L as the most significantly associated gene for adult intelligence from a gene-based analysis in ~3500 individuals. A subset of our sample (that is, the Lothian Birth Cohorts, LBC 1921 and 1936) was part of the previous study that found the association between FNPB1L and adult intelligence. Therefore, we performed the gene-based analysis again by excluding LBC samples. FNPB1L was now ranked 80th with a P-value of 0.0031. The combined gene-based P-value for FNPB1L–childhood intelligence associations in all (discovery + replication), but excluding the LBC samples, was 0.014. These results are consistent with one or more causal variants in FNPB1L being associated with childhood intelligence in the population.

Recently, we developed a method to estimate the contribution of common SNPs to the variation in complex traits using a linear model framework. We applied this method to estimate the contribution of all common SNPs to the variation of childhood intelligence by fitting all genotyped SNPs simultaneously in the model. This analysis requires substantial sample size to get estimates with small standard errors. Therefore, we re-estimated this analysis separately in the three largest cohorts, ALSPAC, TEDS and UMN. The estimated proportions of the variation in childhood intelligence explained by common SNPs were 0.46 (s.e. 0.06), 0.22 (0.10) and 0.40 (0.21) for ALSPAC, TEDS and UMN cohorts, respectively (Table 3). These significant results imply that childhood intelligence is heritable and highly polygenic, and confirm and extend, to a different period in the human life course, previous reports on the polygenic nature of adult intelligence.

We investigated evidence for the polygenic nature of childhood intelligence further by building a multi-SNP predictor from the meta-analysis to predict childhood intelligence in three independent replication cohorts (Generation R, NTR and UMN). We found significant correlation between predictors and childhood intelligence associated with FNPB1L

B Benyamin et al

© 2013 Macmillan Publishers Limited

Molecular Psychiatry (2013), 1 – 6
The maximum proportion of the phenotypic variance in childhood intelligence explained by genetic scores in each cohort were 1.2% ($P = 6 \times 10^{-5}$), 3.5% ($P = 0.001$) and 0.5% ($P = 6 \times 10^{-5}$) for Generation R, NTR and UMN, respectively. To verify whether these results were expected given the sample size, we performed the genetic prediction analysis on height and body mass index (BMI) using similar sample sizes. We used data from three population-based GWAS (the Health Professionals Follow-up Study (HPFS), the Nurses’ Health Study (NHS) and the Atherosclerosis Risk in Communities (ARIC) study as the discovery set ($N = 11,568$ unrelated individuals) for Table 2.

**Table 2.** Gene-based association analysis results for FNBP1L

<table>
<thead>
<tr>
<th>Gene or SNP</th>
<th>Cohort</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FNBP1L</td>
<td>Discovery</td>
<td>0.00004</td>
</tr>
<tr>
<td>FNBP1L</td>
<td>Discovery (excluding LBCs)</td>
<td>0.0031</td>
</tr>
<tr>
<td>FNBP1L</td>
<td>Discovery + replication (excluding LBCs)</td>
<td>0.0137</td>
</tr>
<tr>
<td>rs236330</td>
<td>Discovery</td>
<td>0.00015</td>
</tr>
<tr>
<td>(top SNP for FNBP1L)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs236330</td>
<td>Discovery (excluding LBCs)</td>
<td>0.0018</td>
</tr>
<tr>
<td>rs236330</td>
<td>Discovery + replication (excluding LBCs)</td>
<td>0.00045</td>
</tr>
</tbody>
</table>

Abbreviations: LBS, Lothian Birth Cohorts; SNP, single-nucleotide polymorphism.

**Table 3.** The proportion of the phenotypic variance of IQ explained by common SNPs ($h^2$) estimated using the GCTA software.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>N</th>
<th>$h^2$ (s.e.)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSPAC</td>
<td>5517</td>
<td>0.46 (0.06)</td>
<td>$3.9 \times 10^{-15}$</td>
</tr>
<tr>
<td>TEDS</td>
<td>2794</td>
<td>0.22 (0.10)</td>
<td>0.014</td>
</tr>
<tr>
<td>UMN</td>
<td>1736</td>
<td>0.40 (0.21)</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Abbreviations: IQ, Intelligence Quotient; SNP, single-nucleotide polymorphism.

Figure 2. The proportion of the phenotypic variance in childhood intelligence explained by multi-single-nucleotide polymorphisms (SNPs) predictors ($R^2$) in three replication samples. Each bar represents $R^2$ for a given set of multi-SNPs predictors at a given $P$-value threshold.
validation (N = 3924 unrelated individuals).9 We found that the prediction accuracies for height and BMI were similar compared with that of childhood intelligence. Multi-SNP predictors explained 0.97% (P = 6.4 × 10^-10) and 0.01% (P = 0.56) of variation in height and BMI in the independent sample, respectively.

In summary, while we did not find any individual SNP associated with childhood intelligence that reached genome-wide significance, the aggregate effect of common SNPs on the variation of childhood intelligence was significant in all three of our biggest samples. The proportion of phenotypic variance explained by common SNPs ranged from 0.22 to 0.46. These values are close to the heritability of 0.41 for childhood intelligence estimated from twin data.12 Since the estimates of total heritability derived from common SNPs are a lower limit of narrow-sense heritability,8,23,24 these results imply that there are many common causal variants with small effects segregating in the population, because rare variants are not in sufficient linkage disequilibrium with the genotyped and imputed SNPs to be captured by our whole-genome method. The multi-SNP prediction model was significant in all replication cohorts.

Given the size of the discovery sample, the proportion of the childhood intelligence variation that can be explained by genetic predictors is consistent with previous findings on adult intelligence2 and with results from analyses on height and BMI. The variation in FNBPI1 was significantly associated with childhood intelligence. This gene was previously associated with adult intelligence.6 Our results suggest that childhood intelligence is heritable and highly polygenic. Any attempt to identify individual genetic variants requires a larger sample size than the current study, consistent with other quantitative traits in human populations, including height,25 BMI,26,27 lipids,12 blood pressure28 and platelet count.29

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
We acknowledge funding from the Australian National Health and Medical Research Council (Grants 552498, 613672, 613601 and 1011506) and the Australian Research Council (Grant DP1093502). Funding support for the GWAS of Gene and Environment Initiatives in Type 2 Diabetes was provided through the NIH Genes, Environment and Health Initiative (GEI) (U01HG004399). The human subjects participating in the GWAS derive from The Nurses Health Study and Health Professionals’ Follow-up Study and these studies are supported by National Institutes of Health Grants CA87969, CA55075 and DK58845. Assistance with phenotype harmonisation and genotype cleaning, as well as with general study coordination, was provided by the Gene Environment Associations Studies, GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Funding support for genotyping, which was performed at the Broad Institute of MIT and Harvard, was provided by the NIH GEI (U01HG004424). The data sets used for the analyses described in this manuscript were obtained from dbGaP at [http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gap] through dbGaP accession number [phs000091]. The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HSNS268201100005C, HSNS268201100006C, HSNS268201100007C, HSNS268201100008C, HSNS268201100009C, HSNS268201100101C, HSNS268201100102C, R01HL087641, R01HL59367 and R01HL 086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HSNS268200625226C. We thank the staff and participants of the ARC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. BB and PMV are the recipients of the Australian National Health and Medical Research Council (NHMRC) fellowships. Acknowledgements for individual study cohorts are presented in the Supplementary Note.

AUTHOR CONTRIBUTIONS
UD, GDS, RP and PMV designed the study and contributed to writing the paper. BB performed meta-analysis. BSP, OSPD, GD, NIKH, M-JAB, RMK, RMAC, SF performed statistical analyses for each study cohort. BB and PMV wrote the first draft of the paper. Other authors contributed phenotypic and genotypic information on individual cohorts.

REFERENCES
3 Chabris CF. Most reported genetic associations with general intelligence are probably false positives. Psychol Sci 2011; 23: 1314–1323.
19 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.
Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)