INTRODUCTION

One of the best-established findings in cognitive science is that individual differences in performance on diverse cognitive tasks correlate about 0.30 and that a general factor explains about 40% of the total variance.1 This general cognitive ability factor, usually called general intelligence (‘g’), is one of the best predictors of important life outcomes including education, occupation, and mental and physical health.2 General intelligence is also one of the most heritable behavioural traits, with heritability increasing from 40% in childhood to 80% in later adulthood.3

Identifying some of the many DNA differences that account for its heritability is key for advancing research in intelligence.4 Throughout the life sciences, genome-wide association (GWA) studies have been successful in identifying genes associated with complex traits and common disorders. They have also shown that many DNA variants of very small effect size contribute to heritability.5 Although the small effect size of individual DNA variants detracts from their utility in neurocognitive research, polygenic scores can be created that aggregate the effects of DNA variants to predict genetic propensities for individuals.6,7 For example, the current strongest polygenic score prediction of a quantitative trait is for height, which predicts nearly 20% of the variance explained by GWA studies of intelligence with substantially larger sample sizes. The gene familyplexins, members of which are mutated in several monogenic neurodevelopmental disorders, was significantly enriched for associations with high IQ. This study shows the utility of extreme trait selection for genetic study of intelligence and suggests that extremely high intelligence is continuous genetically with normal-range intelligence in the population.

We used a case–control genome-wide association (GWA) design with cases consisting of 1238 individuals from the top 0.0003 (~170 mean IQ) of the population distribution of intelligence and 8172 unselected population-based controls. The single-nucleotide polymorphism heritability for the extreme IQ trait was 0.33 (0.02), which is the highest so far for a cognitive phenotype, and significant genome-wide genetic correlations of 0.78 were observed with educational attainment and 0.86 with population IQ. Three variants in locus ADAM12 achieved genome-wide significance, although they did not replicate with published GWA analyses of normal-range IQ or educational attainment. A genome-wide polygenic score constructed from the GWA results accounted for 1.6% of the variance of intelligence in the normal range in an unselected sample of 3414 individuals, which is comparable to the variance explained by GWA studies of intelligence with substantially larger sample sizes. The gene familyplexins, members of which are mutated in several monogenic neurodevelopmental disorders, was significantly enriched for associations with high IQ. This study shows the utility of extreme trait selection for genetic study of intelligence and suggests that extremely high intelligence is continuous genetically with normal-range intelligence in the population.

Molecular Psychiatry (2018) 23, 1226–1232; doi:10.1038/mp.2017.121; published online 4 July 2017
intelligence and thus yield increased power to detect associations for alleles that operate throughout the normal distribution.\textsuperscript{20}

The quantitative genetic model could be construed to suggest that the extreme low end of the distribution of intelligence is the mirror image of the extreme high end; however, new mutations may be drivers of very low IQ, as they can more easily disrupt than improve finely tuned neurocognitive performance. Recent quantitative genetic research supports the hypothesis that extremely high intelligence is caused by the same DNA variants responsible for individual differences in intelligence throughout the normal distribution,\textsuperscript{2,21} whereas extremely low intelligence is caused by DNA variants that are not associated with individual differences in intelligence in the normal distribution.\textsuperscript{22}

Here we capitalize on the increased power of association at the high extreme of intelligence as a strategy to facilitate the discovery of alleles that contribute to genetic variation in intelligence throughout the distribution. We conducted a case–control GWA analysis with cases consisting of 1238 individuals from the top 0.0003 (mean IQ score $\sim$170) of the population distribution of intelligence\textsuperscript{23,24} and 8172 unselected population-based controls.

**MATERIALS AND METHODS**

**Participants**

Participants were recruited from two separate US studies. The project received ethical approval from the King’s College London Research Ethics Committee (reference number PNM/11/12–51) and from the European Research Council Executive Agency (reference number Ares (2012)6321). Informed consent was obtained from all subjects. All methods were performed in accordance with relevant guidelines and regulations.

**High-intelligence cases (TIP).** Individuals with extremely high intelligence were recruited from the Duke University Talent Identification Program (TIP), a non-profit organization established in 1980 and dedicated to identifying and fostering the development of academically gifted children. Individuals were selected from the United States for participation in TIP on the basis of performance on the Scholastic Assessment Test or American College Test taken at age 12 years rather than the usual age of 18 years. A composite that aggregates verbal and mathematics Scholastic Assessment Test and American College Test scores correlated $\geq$0.80 with intelligence test scores\textsuperscript{25} and it is estimated that the TIP programme recruits from the top 3\% of the intelligence distribution.\textsuperscript{23–26} For the present study, TIP cases were selected and DNA solicited from the top 1\% of these TIP individuals, representing approximately the top 0.03\% of the intelligence distribution. Illumina Infinium OmniExpress (Illumina.com) genotypes were available for 1409 white European Caucasian individuals (1238 subjects post quality control (QC) see below and Supplementary Table S1). The TIP sample was 62\% male post QC, which is to be expected for a sample of individuals with high IQ.\textsuperscript{27} There was no significant difference in intelligence test scores between males and females. This sample was previously used in a case–control analysis of putative functional exonic variants assayed on the Illumina HumanExome BeadChip.\textsuperscript{28}

**Unselected controls (Health and Retirement Study).** The controls for this study were from The University of Michigan Health and Retirement Study (http://hrsonline.isr.umich.edu), which is a longitudinal panel study that surveyed a representative sample of $\sim$20,000 people in the United States over the age of 50 years every 2 years. Saliva was collected for DNA extraction and genotyping was performed for 12,807 subjects. The genotyping was performed by the NIH Center for Inherited Disease Research using the Illumina Human Omni-2.5 Quad Beadchip, with coverage of 2.5 million single-nucleotide polymorphisms (SNPs). Genotype data were obtained through dbGaP. Matching to TIP and after QC (see below and Supplementary Table S1) genotypes were available for 8185 white Caucasian individuals.

**Extension sample: The Twins Early Development Study.** Twins Early Development Study (TEDS) is a longitudinal UK-based population sample of over 15,000 families with twins born in England and Wales 1994, 1995 and 1996, and identified from birth records.\textsuperscript{29} At age 12 years, individuals were assessed on 16 cognitive tests. Individuals with severe recurrent medical problems or severe perinatal medical problems were excluded, as were individuals whose first language was not English and individuals who reported their ethnicity as other than ‘White’. Phenotypic and genotypic data were available for individuals after QC in cognitive tests for $g$ and educational achievement. The $g$ measure was available at age 12 years as a core four-test composite version, $n$=3414, and an extended 16-test version, $n$=4731, which includes some items that can be regarded as educational achievement measures such as reading. The primary educational achievement phenotype we used in TEDS were the grades achieved in the English General Certificate of School Education at age 16 years, $n$=3584. Full details are given in Supplementary Text.

**Quality control**

For the TIP and Health and Retirement Study data, before SNP imputation and genetic association analyses, SNPs were excluded based on standard criteria such as call rate ($<98\%$), minor allele frequency (MAF $<0.5\%$) and Hardy–Weinberg equilibrium test ($P$-value $<10^{-9}$). Further details are given in the Supplementary Text and Supplementary Table S2.

**SNP imputation**

Genotype imputation was carried out based on the Haplotype Reference Consortium reference panel (vr1.1) (www.haploreference-consortium.org) using PBWT\textsuperscript{30} as implemented in the Sanger Imputation Server (imputation.sanger.ac.uk). Post imputation QC was carried out using QCTool v2 (see URLs). Further details are given in the Supplementary Text.

**Single SNP association analysis**

The single SNP association analysis was performed using a logistic regression for the imputed SNPs with MAF $>0.01$ and imputation info score $>0.90$ using SNPTTEST v2.5.2.\textsuperscript{31} These conservative MAF and info score thresholds were applied to accommodate the fact that the two cohorts have been genotyped separately. An additive model was used after adjusting for sex and the first 10 principal components to control for population structure.\textsuperscript{32} In SNPTTEST we used method ‘expected’ to account for genotype uncertainty of the genotype dosage scores.

**Replication**

Lookups were performed of the genome-wide significant SNP-based findings in four published studies. We also investigated replication of published genome-wide significant findings from related traits from these four studies in the TIP sample.

**LD score regression**

Linkage disequilibrium (LD) score regression\textsuperscript{33} was used to estimate SNP heritability $h_{\text{SNP}}^2$ from GWA summary statistics. Estimates of $h_{\text{SNP}}^2$ on the liability scale were assessed with a prevalence estimate of 0.0003. In addition, genetic correlations ($r_g$) were estimated with several traits from GWA summary statistics, as well as partitioned LD score $h_{\text{SNP}}^2$ to assess enrichment of heritability in SNPs with specific function annotations and tissue-specific genomic annotations, including the central nervous system. For a full description of these procedures, see Supplementary Text.

We also used LD score regression to estimate genetic correlations between high intelligence estimated from our TIP GWA analysis with other studies of intelligence and a range of other traits.\textsuperscript{34,35} The intent of these comparisons was to evaluate similarities and differences in IQ genetic architecture, to suggest hypotheses about the fundamental genetic basis of intelligence, between very high IQ and IQ in the normal range.

**Gene-based association and pathway analysis**

P-values quantifying the degree of association of genes and gene-sets in TIP were generated using MAGMA (vr1.03).\textsuperscript{35} Our approach was guided by rigorous method comparisons of type I error rates of different algorithms.\textsuperscript{36,37} Further details are given in Supplementary Text.

**Heritability estimation**

We estimated the contribution of all common SNPs (MAF $\geq$0.01) in this sample by performing a linear mixed-model analysis to fit all genotyped SNPs simultaneously in the model as implemented in GCTA.\textsuperscript{38} Heritability
was estimated on the liability scale with prevalence = 0.0003 and adjusted for sex and the first 10 PCs for the genotyped SNPs. Heritability was also estimated using the LD score\(^3\) method; estimates were from imputed SNPs and LD was calculated from HapMap 3.

Polygenic scores

We used GWA results from TIP to create polygenic scores for the UK-representative TEDS\(^3\) (Supplementary text). Individuals were tested at 12 years and the phenotypes g-4 and g-16 were calculated from 4 and 16 tests at age 12 years, respectively, where the 4 tests are a subset of the 16. More details about these phenotypes is in Supplementary Text. The two phenotypes were adjusted for age within each testing period and first principal component scores were derived using principal component analysis implemented in R. Imputed genotypes were available for 6710 samples. Stringent QC procedures were applied followed by imputation of SNPs using the Haplotype Reference Consortium reference panel\(^4\) (Supplementary Text). After QC, 7 581 516 genotyped or well-imputed (info \(>0.70\)) variants were available for the polygenic score analyses; 4 657 119 overlapped with variants tested in the independent TIP GWA analysis\(^1\) after exclusions due to nucleotide inconsistencies and MAF < 0.01. We created genome-wide polygenic scores for each individual.
in the TEDS sample using summary statistics from the TIP GWA analyses and all SNPs. To avoid a reduction in predictive accuracy and loss of information due to pruning markers by LD, we used LDpred,41 which infers the posterior mean effect size of each marker using a prior based on the proportion of real effects and the distribution of possible effect sizes and reference LD data (we used our sample as the LD reference), and SNPs are weighted accordingly. This was used to assess the association of these reference LD data (we used our sample as the LD reference), and SNPs are proportion of real effects and the distribution of possible effect sizes and the posterior mean effect size of each marker using a prior based on the

RESULTS
Analysis of genome-wide SNPs
After performing the GWA single SNP analysis using an additive model for 6773 587 SNPs, three intronic SNPs: rs4962322, rs4962520 and rs10794073 located in ADAM12 on chromosome 10 reached genome-wide significant P-values of 8.0 × 10⁻⁸, 1.2 × 10⁻⁸ and 2.0 × 10⁻⁸, respectively. The three SNPs are in high LD with r² ranging between 0.9 and 1.0, with rs4962322 representing the sentinel SNP after clumping. ADAM12 encodes a member of the ADAM (a disintegrin and metalloprotease) protein family. Members of this family are membrane-anchored proteins and have been implicated in a variety of biological processes involving cell-cell and cell-matrix interactions, including fertilization, muscle development and neurogenesis (www.ncbi.nlm.nih.gov/gene). Results for all SNPs are presented in Figure 1 and details of the top three SNPs are in Table 1. A regional plot for rs4962322 is shown in Figure 2 and a comparison of the MAF for rs4962322 from the different samples is in Supplementary Table S3.

The genomic inflation λ for these analyses was equal to 1.096 and the LD score intercept is 1.020. This is consistent with inflation largely driven by polygenic variation rather than population stratification.33 Although there is no similar study in which to carry out replication for the top SNP, we performed a look up of all SNPs with P-values < 10⁻⁷ including rs4962322 in other studies with similar phenotypes, as described below.

Replication of single SNP variation
We have performed a look-up of the three top SNPs in four studies with related phenotypes: CHIC9 (childhood intelligence), CHARGE11 (normal range adult intelligence), educational attainment17 and VNR-Ukr12 (numerical and verbal reasoning) (Supplementary Table S4). The direction of effect was concordant with two of the two studies for rs4962322 and rs4962520, and only for one study for rs10794073. None of the SNPs replicated in significance. Summary data for rs4962520 were not available for CHIC or CHARGE. We also compared the associations of the four published GWAS with our results; a summary of these is in Supplementary Table S5. There were 95 published genome-wide significant SNPs from the four studies and we found comparable SNP data for 63 of these. Ninety-four percent of the TIP-GWA SNPs showed concordance in direction of effect with the published SNPs, although none reached genome-wide significance.

Gene-based association and pathway analyses
To test whether the combined effect of SNPs within a gene has a significant effect, we conducted a gene-based analysis using MAGMA v1.05. The gene with the genome-wide significant marker, ADAM12, also ranked first in the gene-based analysis with P-value = 5.82 × 10⁻⁶ (Bonferroni threshold = 2.57 × 10⁻⁶). The second top gene, SH2D1A, with P-value = 7.66 × 10⁻⁶ is located on the X chromosome. Both genes are suggestive but not significant. Top results from the gene-based analyses are in Supplementary Table S6; the full results are in Supplementary Table S7. Pathway analyses were applied to three sets of pathways: gene families, gene intolerance gene sets based on Residual Variation Intolerance Score percentiles and biological pathways. Complete results with links to source data are provided in Supplementary Tables S8–S10 and S10a. Information on gene sets and pathway analysis procedures is given in Supplementary Text. One gene family showed a significant association (P-value = 6.43 × 10⁻⁵, Bonferroni threshold = 1.26 × 10⁻⁴); plexins. Plexins are transmembrane proteins which act as receptors to semaphorins. We also determined the enrichment of intolerant genes in the GWAS by testing gene-sets with decreasing or increasing gene intolerance. The gene intolerance results are given in Supplementary Table S9.

Heritability and genetic correlation estimation
The proportion of variance explained by all common variants using GCTA was 0.33 (0.02); using LD score regression (with unconstrained intercept), heritability was 0.42 (0.06). These estimates are compatible with a highly polygenic architecture of this extreme IQ trait, given the few genome-wide significant hits we observe. Using LD score regression we also find large and significant genetic correlations between high IQ and other phenotypes from other studies (Table 2). Genetic correlations ranged from 0.75 (0.13) with childhood IQ9 to 0.86 (0.10) with verbal and numerical reasoning.12

We carried out partitioned heritability analysis using LD score regression. The results show that the central nervous system category accounts for the largest proportion of the heritability: 68.8% from 14.8% of SNPs (enrichment P-value = 4.07 × 10⁻⁵, Bonferroni threshold = 8.06 × 10⁻⁶). A summary table with the proportions explained by the 62 categories is in Supplementary Table S11.

Polygenic scores
The results from the TIP GWA analyses of extremely high intelligence were used to build polygenic scores in the unselected and representative TEDS sample. Highly significant predictions were observed for individual differences in intelligence (r² = 2.4%) for the broader 16-test composite g score at age 12 years and 1.6% for the 4-test composite g score at age 12 years. Although the TEDS 16-test composite was better predicted by the polygenic score, this composite includes educationally relevant skills such as

---

**Table 1.** Genome-wide significant SNPs (P < 5 × 10⁻⁷) from the logistic regression

| SNP      | CHR | BP    | OR    | s.e.  | P-values | A1  | A2  | MAF   | Function   | Gene     |
|----------|-----|-------|-------|-------|----------|-----|-----|-------|------------|----------|
| rs4962322a | 10  | 127 932 765 | 1.489 | 0.070 | 8.05 × 10⁻⁶ |  A  |  C  | 0.089 | Intronic   | ADAM12  |
| rs4962520  | 10  | 127 917 302 | 1.469 | 0.070 | 1.18 × 10⁻⁶ |  C  |  T  | 0.091 | Intronic   | ADAM12  |
| rs10794073 | 10  | 127 945 261 | 0.672 | 0.070 | 2.02 × 10⁻⁶ |  C  |  A  | 0.085 | Intronic   | ADAM12  |

Abbreviations: MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism. *Sentinel SNP after clumping.
significant for cognitive phenotypes, ranging from 0.75 (0.13) with childhood IQ to 0.86 (0.10) for verbal and numerical reasoning.9,12 In our pathway analysis, the gene family plexins, members of which are mutated in several monogenic neurodevelopmental disorders, was significantly enriched for associations with high IQ (P-value = 6.43 × 10⁻⁸). The plexin-semaphorin pathway has been linked to axon guidance,62 mental disability and neural connectivity,43 axon regeneration in the central nervous system, bone disorders, cancer and inflammatory diseases.44 Noticeably, the top biological pathway was GO:SEMAPHORIN_RECEPTOR_ACTIVITY with P-value = 5.82 × 10⁻⁸; however, it was not significant. Partitioned LD score analyses also demonstrated a highly significant enrichment of SNP heritability in the central nervous system-annotated portion of the genome—68.8% of the SNP heritability results from 14.8% of SNPs—suggesting a strong role for variation in brain-expressed genes and their regulatory regions.

The primary single-variant findings from our GWA association analysis with high IQ were three intronic SNPs in high LD, r² ≥ 0.90: rs4962322 (P = 8.05 × 10⁻⁸) and rs4962520 (P = 1.18 × 10⁻⁸), and rs10794073 (P = 2.02 × 10⁻⁷), all genome-wide significant and located in ADAM12 on chromosome 10. Gene-based analyses also pointed to the same locus, with a gene-wise P-value = 5.82 × 10⁻⁸. Similar to another study using this same case cohort,49 there is no available replication cohort with extremely high IQ. Therefore, we compared results from our case–control design to comparable phenotypes in GWA studies of individual differences in the normal range. The three single variants identified in ADAM12 did not replicate in published data for the normal range of intelligence,11—childhood intelligence,9 educational attainment,17 and verbal and numerical reasoning,12 although the effect was in the same direction in two of the studies for rs4962322 and rs4962520, and only for one study for rs10794073 (Supplementary Table S2). Conversely, lookup of GWA-significant SNPs from these same GWA studies in TIP were likewise not significant (Supplementary Table S5). One possibility is that ADAM12 affects high IQ but not IQ in the normal range.

We also used polygenic scores to test the generalizability of TIP GWA results to individual differences in intelligence and educational achievement. For this, we used TEDS, which is a UK-based study also of white European Caucasian individuals. We used GWA

### Table 2. Genetic correlations estimated through LD score regression with summary statistics from publicly available data

| Phenotype1 | Phenotype2 | r_g | s.e. | P-values |
|------------|------------|-----|------|----------|
| TIP_IQ     | Childhood IQ³ | 0.75 | 0.13  | 1.6 × 10⁻⁹ |
| TIP_IQ     | Cognitive performance¹¹ | 0.80 | 0.08  | 3.1 × 10⁻²⁶ |
| TIP_IQ     | College completion | 0.79 | 0.08  | 2.0 × 10⁻²³ |
| TIP_IQ     | Years of education | 0.79 | 0.07  | 1.9 × 10⁻²⁶ |
| TIP_IQ     | VNR-UKB reasoning¹² | 0.86 | 0.10  | 1.6 × 10⁻¹⁶ |

Abbreviation: LD, linkage disequilibrium.
results from TIP to create polygenic scores, in order to predict variance in normal-range intelligence and educational achievement in TEDS for g based on four cognitive tests at age 16 years, as well as a broader measure of g that includes 16 tests, some of which could be considered as assessing educational performance such as reading, mathematics and language. A polygenic score created from the TIP GWA results accounted for 1.6% of the variance in individual differences in intelligence in TEDS at age 16 years for our 4-test composite of intelligence and 2.4% for our 16-test measure of g.

Focusing on the core 4-test g measure, we think that accounting for 1.6% of the variance of intelligence is exciting for several reasons. First, this effect size is greater than the effect sizes from previous GWA studies of the normal range of intelligence.45 The effect size of 1.6% represents > 2.5% of the heritable variance of intelligence, which is comparable to the most robust effect sizes in behavioural research.46 Second, the result demonstrates that genetic effects on extremely high intelligence are similar to those responsible for the heritability of the normal range of intelligence in the population. That is, extremely high intelligence is quantitatively, not qualitatively, different from the rest of the distribution.19 Third, it shows the utility of extreme trait ascertainment for genetic analysis of neurocognitive traits ascertained for genetic analysis of neurocognitive traits as in the population. That is, extremely high intelligence is responsible for the heritability of the normal range of intelligence. It may be that although the TIP sample has extremely high IQ, individuals were nominated initially by their schools. They were then screened using college entrance examinations, which are a mix of intelligence and educational achievement. For this reason, we speculate that the TIP GWA results represent a mix of intelligence and educational achievement. It may also be the case that additional elements of high IQ are captured by educational achievement.

Although not the focus of this study, we acknowledge the role of the environment, which has been shown to be particularly important in childhood studies of IQ with, conversely, increasing heritability of IQ with age.27 This work has several limitations. As mentioned previously, there is a lack of a similar replication sample of extremely high IQ individuals to validate our findings. Furthermore, cases and controls were collected separately, which is why our QC protocols were very stringent, resulting in a reduced number of individuals and SNPs. In addition, our case–control design resulted in some loss of power, in contrast to GWA analysis using a quantitative scale. Finally, another limitation is that this study, similar to the other studies mentioned above, is restricted to white Europeans.

In summary, we have shown that extremely high intelligence is a polygenic trait and its high heritability indicates that GWA analysis captures a large portion of the genetic variance. The novel aspect of the present study is that it represents a complementary strategy to the ‘brute force’ approach of increasing sample sizes of GWA studies of IQ variation in the normal range (and is an example for quantitative trait genetics in general). It demonstrates the utility of a ‘positive genetics’ strategy of focusing on the extremely high end of the distribution of IQ. Larger scale studies focusing on either high IQ or IQ in the normal range are likely to be successful in the identification of many significant loci and biological pathways.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGMENTS
This research was supported by a European Research Council Advanced Investigator award (295366) to RP. Collecting DNA from the highest-scoring TIP individuals was supported by an award from the John Templeton Foundation (13575) to RP. The TEDS, which was used to test the extent to which the extremely high intelligence case–control GWA yields a polygenic score that also accounts for genetic influence throughout the normal distribution, is supported by the UK Medical Research Council (MR/M021475/1 and previously G0901245), with additional support from the US National Institutes of Health (HD044454; AG046938) and the European Commission (602768/295366). EK is supported by the MRC/IoPPN Excellence Award. RP is supported by a Medical Research Council Research Professorship and a European Research Council Advanced Investigator award (295366). GB and the laboratory work were supported in part by the NIHR Biomedical Research Centre (‘BRC’) and NIHR Dementia Biomedical Research Unit (‘BRU’) hosted at King’s College London and South London and Maudsley NHS Foundation Trust, and funded by the National Institute for Health Research under its Biomedical Research Centres funding initiative. The views expressed are those of the authors and not necessarily those of the BRC or the BRU, the NHS, the NIHR or the Department of Health or King’s College London. We gratefully acknowledge capital equipment funding from the Maudsley Charity (Grant Reference 980) and Guy’s and St Thomas’s Charity (Grant Reference STR130505).

URLS
RegulomeDB, http://regulomedb.org.
1000 Genomes Project multi-ancestry imputation panel, https://mathgen.stats.ox.ac.uk/impute_data_download_1000G_phase1_integrated.html. Genotype-based checksums for relatedness determination, http://www.broadinstitute.org/~sripke/share_links/checksums_download. LD-Hub, http://ldsc.broadinstitute.org and https://github.com/mkanai/ggman

AUTHOR CONTRIBUTIONS
RP conceived the study. CC, LSH, HP and SN did the genotyping and initial QC. DZ, EK, HG and HMW did the statistical analysis. DZ, GB and RP wrote the manuscript. DZ, EK, HG, CC, LSH, HP, SN, MAS, MP, DL, RP and GB reviewed the manuscript.

REFERENCES
1 Carroll JB. Human Cognitive Abilities A Survey of Factor-Analytic Studies. Cambridge University Press: Cambridge, 1993.
2 Deary I. Intelligence. Annu Rev Psychol 2012; 63: 453–482.
3 Knopik VS, Neiderhiser JM, DeFries JC, Plomin R. Behavioral Genetics, 7th edn. Worth: New York, 2016.
4 Plomin R, Simpson MA. The future of genomics for developmentalists. Dev Psychopathol 2013; 25(4 Pt 2): 1263–1278.
5 Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet 2012; 90: 7–24.
6 Dudbridge F, Visscher P, Brown M, McCarthy M, Yang J, Wray N et al. Power and predictive accuracy of polygenic risk scores. PLoS Genet 2013; 9: e1003348.
7 Wray NR, Lee SH, Mehta D, Vinkhuizen AAE, Dudbridge F, Middeldorp CM et al. Review research: polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry 2014; 55: 1068–1087.
8 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S et al. Defining the role of common variation in the genomic and biological architecture of adult human height, Nat Genet 2014; 46: 1173–1186.
9 Benyamin B, Pourcain BS, Davies OS, Davies G, Hansell NK, Brion M-J et al. Childhood intelligence is heritable, highly polygenic and associated with FNB1P1. Mol Psychiatry 2014; 19: 253–258.
10 Butcher LM, Meaburn E, Knight J, Sham PC, Schalkwyk LC, Craig IW et al. SNPs, microarrays and pooled DNA identification of four loci associated with mild mental impairment in a sample of 6000 children. Hum Mol Genet 2005; 14: 1315–1325.
11 Davies G, Armstrong N, Bis JC, Bressler J, Chousiaki V, Giddaluru S et al. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). Mol Psychiatry 2015; 20: 183–192.
12 Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112151). Mol Psychiatry 2016; 21: 758–767.
13 Davis OSP, Butcher LM, Docherty SJ, Meaburn EL, Curtis CJ, Simpson MA et al. A three-stage genome-wide association study of general cognitive ability: hunting the small effects. Behav Genet 2010; 40: 759–767.

Molecular Psychiatry (2018), 1226–1232
GWA study for extremely high intelligence

D Zabaneh et al

14 Davis OSP, Band G, Pirinen M, Haworth CMA, Meaburn EL, Kovas Y et al. The correlation between reading and mathematics ability at age twelve has a substantial genetic component. Nat Commun 2014; 5: 4204.

15 de Vlaming R, Okbay A, Rietveld CA, Johannesson M, Magnusson PKE, Uitterlinden AG et al. Meta-GWAS Accuracy and Power (MetaGAP) calculator shows that hiding heritability is partially due to imperfect genetic correlations across studies. PLoS Genet 2017; 13: e1006495.

16 Rietveld CA, Esko T, Davies G, Pers TH, Turley P, Benyamin B et al. Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. Proc Natl Acad Sci USA 2014; 111: 13790–13794.

17 Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA et al. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 2016; 533: 539–542.

18 Selzam S, Krapohl E, von Stumm S, O'Reilly PF, Rimfeld K, Kovas Y et al. Predicting educational achievement from DNA. Mol Psychiatry 2017; 22: 267–272.

19 Plomin R, Haworth CMA, Davis OSP. Common disorders are quantitative traits. Nat Rev Genet 2010; 10: 872–878.

20 Pütter C, Pechlivanis S, Nöthen MM, Jöckel K-H, Wichmann HE, Scherag A. Missing heritability in the tails of quantitative traits? A simulation study on the impact of slightly altered true genetic models. Hum Hered 2011; 72: 173–181.

21 Shakeshaft NG, Trzaskowski M, McMillan A, Krapohl E, Simpson MA, Reichenberg A et al. Thinking positively: the genetics of high intelligence. Intelligence 2015; 48: 123–132.

22 Reichenberg A, Cederlöf M, McMillan A, Trzaskowski M, Kapara O, Fruchter E et al. Discontinuity in the genetic and environmental causes of the intellectual disability spectrum. Proc Natl Acad Sci USA 2016; 113: 1098–1103.

23 Kell HJ, Lubinski D, Benbow CP. Who rises to the top? Early indicators. Psychol Sci 2013; 24: 648–659.

24 Makel MC, Kell HJ, Lubinski D, Putallaz M, Benbow CP. When lightning strikes twice: profoundly gifted, profoundly accomplished. Psychol Sci 2016; 27: 1004–1018.

25 Lubinski D, Webb RM, Morelock MJ, Benbow CP. Top 1 in 10,000: a 10-year follow-up of the profoundly gifted. J Appl Psychol 2001; 86: 718–729.

26 Clynes T. How to raise a genius: lessons from a 45-year study of super-smart children. Nature 2016; 537: 152–155.

27 Wai J, Cacchio M, Putallaz M, Makel MC. Sex differences in the right tail of cognitive abilities: a 30 year examination. Intelligence 2010; 38: 412–423.

28 Spain SL, Pedroso I, Kadeva N, Miller MB, Iacono WG, McGue M et al. A genome-wide analysis of putative functional and exonic variation associated with cognitive abilities: a 30 year examination. Twin Res Hum Genet 2013; 16: 117–125.

29 Haworth CMA, Davis OSP, Plomin R. Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavior development from childhood to young adulthood. Twin Res Hum Genet 2013; 16: 117–125.

30 Durbin R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). Bioinformatics 2014; 30: 1266–1272.

31 Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010; 11: 499–511.

32 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006; 38: 904–909.

33 Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N et al. LD Score regression distinguishes confounding from polygenic architecture in genome-wide association studies. Nat Genet 2015; 47: 291–295.

34 Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015; 47: 1236–1241.

35 de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2011; 7: e1002192.

36 Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study implicates neuronal, immune and histone pathways. Nat Neurosci 2015; 18: 199–209.

37 de Leeuw CA, Neale BM, Heskes T, Posthuma D. The statistical properties of gene-set analysis. Nat Rev Genet 2016; 17: 353–364.

38 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.

39 Haworth CMA, Davis OSP, Plomin R. Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. Twin Res Hum Genet 2013; 16: 117–125.

40 McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 2016; 48: 1279–1283.

41 Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindström S, Ripke S et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. Am J Hum Genet 2015; 97: 576–92.

42 Winberg ML, Noordermeer JN, Tamagnone L, Comoglio PM, Sprekks MK, Tessier-Lavigne M et al. Plexin A is a neuronal semaphorin receptor that controls axon guidance. Cell 1998; 95: 903–916.

43 Athanassakis E, Licastro D, Balestra F, Fabretto A, Dipresa S, Vozzi D et al. Next generation sequencing in nonsyndromic intellectual disability: from a negative molecular karyotype to a possible causative mutation detection. Am J Med Genet A 2014; 164A: 170–176.

44 Worzfeld T, Offermanns S. Semaphorins and plexins as therapeutic targets. Nat Rev Drug Discov 2014; 13: 603–621.

45 Krapohl E, Euesden J, Zabaneh D, Pingault J-B, Rimfeld K, von Stumm S et al. Phenomen-wide analysis of genome-wide polygenic scores. Mol Psychiatry 2016; 21: 1188–1193.

46 Richard FD, Bond CF, Stokes-Zoota JJ. One hundred years of social psychology quantitatively described. Rev Gen Psychol 2003; 7: 331–63.

47 Bouchard TJ. The Wilson Effect: the increase in heritability of IQ with age. Twin Res Hum Genet 2013; 16: 923–930.

© The Author(s) 2018

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