

ARTICLE

Received 4 Sep 2013 | Accepted 23 May 2014 | Published 8 Jul 2014

DOI: 10.1038/ncomms5204

OPEN

# The correlation between reading and mathematics ability at age twelve has a substantial genetic component

Oliver S.P. Davis<sup>1,2,\*</sup>, Gavin Band<sup>3,\*</sup>, Matti Pirinen<sup>3,\*</sup>, Claire M.A. Haworth<sup>2,4</sup>, Emma L. Meaburn<sup>5</sup>, Yulia Kovas<sup>6</sup>, Nicole Harlaar<sup>7</sup>, Sophia J. Docherty<sup>2</sup>, Ken B. Hanscombe<sup>2</sup>, Maciej Trzaskowski<sup>2</sup>, Charles J.C. Curtis<sup>2</sup>, Amy Strange<sup>3</sup>, Colin Freeman<sup>3</sup>, Céline Bellenguez<sup>3</sup>, Zhan Su<sup>3</sup>, Richard Pearson<sup>3</sup>, Damjan Vukcevic<sup>3</sup>, Cordelia Langford<sup>8</sup>, Panos Deloukas<sup>8</sup>, Sarah Hunt<sup>8</sup>, Emma Gray<sup>8</sup>, Serge Dronov<sup>8</sup>, Simon C. Potter<sup>8</sup>, Avazeh Tashakkori-Ghanbaria<sup>8</sup>, Sarah Edkins<sup>8</sup>, Suzannah J. Bumpstead<sup>8</sup>, Jenefer M. Blackwell<sup>9,10</sup>, Elvira Bramon<sup>11,12</sup>, Matthew A. Brown<sup>13</sup>, Juan P. Casas<sup>14,15</sup>, Aiden Corvin<sup>16</sup>, Audrey Duncanson<sup>17</sup>, Janusz A.Z. Jankowski<sup>18,19,20</sup>, Hugh S. Markus<sup>21</sup>, Christopher G. Mathew<sup>22</sup>, Colin N.A. Palmer<sup>23</sup>, Anna Rautanen<sup>3</sup>, Stephen J. Sawcer<sup>24</sup>, Richard C. Trembath<sup>22</sup>, Ananth C. Viswanathan<sup>25</sup>, Nicholas W. Wood<sup>26</sup>, Ines Barroso<sup>8</sup>, Leena Peltonen<sup>8</sup>, Philip S. Dale<sup>27</sup>, Stephen A. Petrill<sup>28</sup>, Leonard S. Schalkwyk<sup>2</sup>, Ian W. Craig<sup>2</sup>, Cathryn M. Lewis<sup>2</sup>, Thomas S. Price<sup>29</sup>, The Wellcome Trust Case Control Consortium 2<sup>†</sup>, Peter Donnelly<sup>3,30</sup>, Robert Plomin<sup>2,§</sup> & Chris C.A. Spencer<sup>3,§</sup>

Dissecting how genetic and environmental influences impact on learning is helpful for maximizing numeracy and literacy. Here we show, using twin and genome-wide analysis, that there is a substantial genetic component to children's ability in reading and mathematics, and estimate that around one half of the observed correlation in these traits is due to shared genetic effects (so-called Generalist Genes). Thus, our results highlight the potential role of the learning environment in contributing to differences in a child's cognitive abilities at age twelve.

<sup>1</sup> Department of Genetics, Evolution and Environment, UCL Genetics Institute, University College London, London WC1E 6BT, UK. <sup>2</sup> King's College London, Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London SE5 8AF, UK. <sup>3</sup> Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. <sup>4</sup> Department of Psychology, University of Warwick, Coventry CV4 7AL, UK. <sup>5</sup> Department of Psychological Sciences, Birkbeck, University of London, London WC1E 7HX, UK. <sup>6</sup> Department of Psychology, Goldsmiths, University of London, London SE14 6NW, UK. <sup>7</sup> Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, Colorado 80309-0345, USA. <sup>8</sup> Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK. <sup>9</sup> Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, Crawley, Western Australia, Australia. <sup>10</sup> Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK. <sup>11</sup> UCL Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, UK. <sup>12</sup> UCL Mental Health Sciences Unit, University College London, London W1W 7EJ, UK. <sup>13</sup> University of Queensland Diamantina Institute, Translational Research Institute, Princess Alexandra Hospital, University of Queensland, Brisbane, Queensland QLD 4102, Australia. <sup>14</sup> Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. <sup>15</sup> Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK. <sup>16</sup> Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Ireland. <sup>17</sup> Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE, UK. <sup>18</sup> Centre for Digestive Diseases, Blizard Institute, Queen Mary University of London, London E1 2AT, UK. <sup>19</sup> Wolfson College, Linton Road, Oxford OX2 6UD, UK. <sup>20</sup> Peninsula School of Medicine and Dentistry, Associate Deans Office, John Bull Building, Plymouth PL6 8BU, UK. <sup>21</sup> Clinical Neurosciences, Saint George's University of London, London SW17 0RE, UK. <sup>22</sup> Department of Medical and Molecular Genetics, King's College London, King's Health Partners Guy's Hospital, London SE1 9RT, UK. <sup>23</sup> Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. <sup>24</sup> Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK. <sup>25</sup> NIHR Biomedical Research Centre at Moorfields Eye Hospital NHSFT and UCL Institute of Ophthalmology, London EC1V 2PD, UK. <sup>26</sup> Department of Molecular Neuroscience, Institute of Neurology, University College London, London WC1N 3BG, UK. <sup>27</sup> Department of Speech and Hearing Sciences, University of New Mexico, Albuquerque, New Mexico 87131, USA. <sup>28</sup> Department of Family Science, Institute for Population Research, The Ohio State University, Columbus, Ohio 43210, USA. <sup>29</sup> Institute for Translational Medicine and Therapeutics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-5158, USA. <sup>30</sup> Department of Statistics, University of Oxford, Oxford OX1 3TG, UK. \* These authors contributed equally to this work. † These authors jointly supervised this work. Correspondence and requests for materials should be addressed to O.S.P.D. (email: oliver.davis@ucl.ac.uk) or to C.C.A.S. (email: chris.spencer@well.ox.ac.uk).

† A full list of members is provided in Supplementary Note 1.

Understanding the aetiology of complex cognitive traits such as reading and mathematics ability is essential for helping children achieve their potential<sup>1</sup>. These traits are highly heritable<sup>2,3</sup> and have been shown to associate with quality of life including wealth and life expectancy<sup>4,5</sup>. In spite of their importance and well-established heritability, much remains to be understood about the genetic architecture of cognitive abilities and the genetic component to the correlation between them.

It has been shown<sup>6,7</sup> that population variation in cognitive abilities shares a substantial genetic component with learning difficulties such as dyslexia and dyscalculia (defined here as the low extreme of the distribution<sup>8</sup>). These difficulties affect more than 10% of the population of English-speaking countries<sup>9</sup>, with undiagnosed problems costing economies billions of dollars per year, as well as the less well-documented human cost of missed opportunities. Dyslexia is by far the most frequently diagnosed form of learning difficulty in school-age children<sup>10</sup>, it shows strong stability across childhood and adolescence<sup>10</sup>, and frequently co-occurs with other childhood learning difficulties and psychopathologies<sup>11,12</sup>. Although much less is known about dyscalculia, numeracy is as much a requirement as literacy in our increasingly technological world<sup>12,13</sup>.

Dyslexia was one of the first traits studied using QTL sib-pair linkage analysis<sup>14</sup>, and although it has been proven to be difficult to identify the genes responsible for these linkages, several candidate genes are under scrutiny<sup>14,15</sup>. The first steps towards genome-wide association studies (GWAS) of reading and mathematics ability, using pooled DNA on microarrays, concluded that it is likely that no common genetic variants of large effect influence either trait<sup>16,17</sup>. Until recently<sup>1,18,19</sup>, no common variants associated with the normal range of cognitive traits have been discovered with compelling levels of evidence, although some candidates have been reported.

Here we conduct a GWAS of Reading and Mathematics abilities in a sample of ~3,000 twin pairs. We find no replicable loci with convincing levels of evidence for association, consistent with a substantially polygenic contribution of genetics to these traits. Using bivariate twin- and population-level models, we estimate the heritability and genetic correlation between the two traits. We find a high genetic correlation (around 70%), indicating substantial pleiotropy, and accounting for a large proportion of phenotypic correlation.

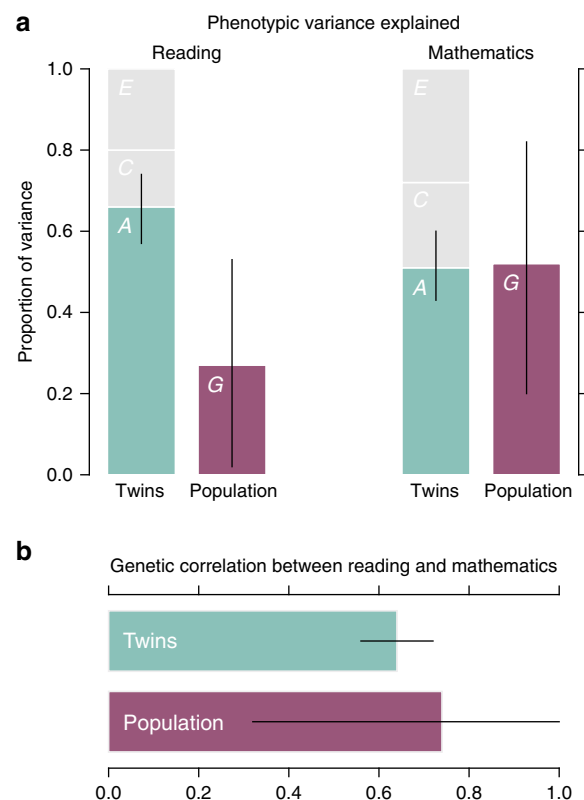
## Results

As part of the Wellcome Trust Case-Control Consortium 2 (WTCCC2), in collaboration with the Twins Early Development Study (TEDS), we performed a GWAS using 2,794 unrelated members of monozygotic (MZ) and dizygotic (DZ) twin pairs, measured for their reading and mathematics ability using a combination of web- and phone-based tests at age twelve. The scores were combined across tests and adjusted for age, while gender was used as a covariate in the analyses. Using genotype imputation we performed association analysis for 1,588,650 autosomal markers with reading and mathematics scores separately (see Methods). We followed up the strongest signals of association ( $P_{GWAS} < 5 \times 10^{-5}$ ; reading,  $N = 2,243$ ; mathematics,  $N = 2,772$ ) in a further 2,153 individuals, some of whom were co-twins of individuals in the discovery data. One region on chromosome 19 (rs349045) achieved a  $P$ -value of  $9.63 \times 10^{-9}$  (Merlin,  $N = 6,061$ ) for reading ability in the joint analysis of discovery and replication data. However, this association failed to replicate using a related phenotype (the Test of Word-Reading Efficiency (TOWRE)—one of four reading tests from the TEDS analysis) in the Avon Longitudinal Study of Parents and Children (ALSPAC,  $N = 2,077$ ). The results for the GWAS are shown in

Supplementary Figs 1 and 2 and Supplementary Tables 1 and 2. The results from loci previously reported to be associated with reading or mathematics ability or difficulties are reported in Supplementary Table 3.

One explanation for the lack of compelling evidence for association at individual single-nucleotide polymorphisms (SNPs), despite large sample sizes and high heritability estimates, is that the traits studied here are substantially polygenic, with each variant having a small effect. Recent studies have demonstrated that the genetic variants that determine measures of intelligence early and late in life overlap<sup>20</sup>. In our data, standardised reading and mathematics scores show a high correlation,  $r = 0.60$ . This is perhaps unsurprising given that many environmental influences (for example, parenting, schooling and socio-economic factors) will impact on both reading and mathematics ability. Twin studies have also identified a genetic contribution to the correlation<sup>21</sup>. Our data provide the opportunity to clarify the contribution of genetics to the strong correlation in these cognitive abilities using both twin and molecular data in the same sample.

To investigate the genetic contribution to the correlation, we first fit a bivariate version of the classical twin model using both MZ and DZ twin pairs for whom the reading and mathematics scores were available (Methods and Fig. 1). This method does not



**Figure 1 | Comparison of twin and population-level analyses.** Comparison of estimates of heritability and genetic correlation between reading and mathematics as estimated by twin and population-level models.

(a) comparison of heritability estimates. A, C and E stand for additive genetic, shared environment and non-shared environment effects, respectively, and G denotes the population-level estimate of additive genetic effects. (b) comparison of estimates of genetic correlation in reading and mathematics ability from twin ( $\rho_A$ ) and population-level ( $\rho_G$ ) models. Bars indicate point estimates (twin model:  $N = 2,794$  twin pairs; population-level model:  $N = 2,221$  individuals) with solid black lines indicating 95% confidence intervals.

use the genotype data, but assumes that genetic relatedness at the variants that affect the traits follows average relatedness of twins (one half for DZ twins and one for MZ twins). The approach estimates the phenotypic covariance explained by additive genetic effects (narrow-sense heritability and correlation), shared environmental influences and non-shared environmental influences across traits. Our analysis estimated the narrow-sense heritability of reading at 0.66 (95% confidence interval (CI): 0.57–0.74), with shared and non-shared environmental contributions of 0.14 (0.06–0.22) and 0.20 (0.18–0.23), respectively. Heritability of mathematics was 0.51 (0.43–0.60), with shared and non-shared environmental estimates of 0.21 (0.14–0.28) and 0.27 (0.25–0.30), respectively. Using the bivariate approach we estimated that the genetic correlation (denoted  $\rho_A$ ) between reading and mathematics is 0.64 (0.56–0.72), with shared and non-shared environmental correlations of 0.90 (0.67–1.00) and 0.30 (0.24–0.37), respectively.

To exploit the genome-wide data collected as part of this study, we next applied a population-level variance component model to assess polygenic contribution to these traits (see Methods) that bases inference on small differences in allele sharing between individuals who are not closely related<sup>22</sup>. In this model, we estimate the proportion of the phenotypic variance that can be explained by the autosomal SNPs available in the genotyping array data (see Methods) using only the individuals from the GWAS discovery phase (with estimated identity by descent <5%). Using this approach, we estimated the proportion of variance accounted for by the available SNPs as 0.27 for reading (95% CI: 0.02–0.53) and 0.52 (0.20–0.82) for mathematics, with genetic correlation (denoted  $\rho_G$ ) of 0.74 (0.32–1.00) (Fig. 1 and Supplementary Table 4). By both simulation- and permutation-based approaches we confirmed that the estimated genetic correlation was significantly larger than zero (empirical  $P < 0.02$ ,  $N = 2,221$ ; see Supplementary Fig. 3).

## Discussion

As discussed elsewhere<sup>23,24</sup>, the difference between the twin and population-level models and their underlying assumptions complicates direct comparison of estimated parameters (see Methods). For example, unlike the twin model the population-level approach assumes that all environmental influences are independent among individuals. If there are geographically structured determinates of ability (for example, quality of teaching) that correlate with the genetic differences, then this can inflate population-based estimates of heritability<sup>25</sup>. To address potential confounding by population structure, we fit the model both with and without the leading principal components (PCs) of genetic structure as covariates with similar results (Supplementary Table 4). The population-level approach is also influenced by the coverage of the SNPs used to estimate allele sharing; if a proportion of heritability is due to variants that are not in linkage disequilibrium with typed variants, the population-level model will underestimate heritability. Twin model estimates in principle capture all genetic variation but interpretation of the parameters depends on assumptions regarding the presence of dominance or interaction effects, correlation or interaction between genetics and environment, and putative genetic influences on the shared environmental component. We note that factors affecting additive genetic variance are likely to similarly affect genetic covariance between traits, so that estimates of genetic correlation may be more robust to these effects. The observation that the two different approaches, using different information in the data, estimate a substantial correlation in the genetic component of reading and mathematics ability strongly supports a shared genetic basis.

This observation can be interpreted in at least two ways. First, as a decomposition of the correlation in reading and mathematics ability (see Supplementary Methods), where the twin model estimates that 62% of the observed phenotypic correlation is due to additive genetic factors, and the population-level model estimates that 47% of the observed phenotypic correlation is captured by the available SNPs. Second, by assuming that the genetic variants that affect these traits can be classified into either trait-specific or pleiotropic effects, with similar distribution of effect sizes (see Supplementary Methods), we estimate that at least 10%, and probably around a half, of genetic variants that affect at least one of the traits contribute to both traits. These results suggest substantial pleiotropy, in line with the Generalist Genes Hypothesis<sup>7</sup>.

Our results support previous evidence that common learning abilities and their associated disabilities are unlikely to be affected by common genetic variants of large effect; even with a sample of thousands of individuals and 1,588,650 genetic variants, our most convincing signal of association failed to replicate in an independent sample (although it may still be of interest to future studies). However, we do find suggestive evidence in favour of some previously reported associations (see Supplementary Table 3) for reading ability, most notably rs807701 ( $P_{GWAS} = 0.0084$ ,  $N = 2,243$ ) in the *DCDC2* gene, which has been implicated in neuronal development<sup>26–28</sup>. As is the case for other complex traits such as height<sup>29</sup> larger sample sizes and meta-analyses are needed to pinpoint individual genetic variants<sup>19</sup>. The comparison of our population-level and twin-based variance components analysis, conducted in the same cohort using identical phenotypes, shows that the GWAS data were able to explain a significant proportion of the variance in cognitive abilities (Fig. 1). This is particularly true for mathematics, where the population-level model estimate of heritability is very close to the twin model estimate.

Importantly, our analyses show that a substantial proportion of the observed correlation in reading and mathematics abilities is due to genetics. If a large proportion of the genetic factors that affect these traits are pleiotropic, then the factors that lead to differences in an individual's abilities (or disabilities) are relatively more likely to be environmental. Understanding the aetiology of these patterns increases our chances of developing effective learning environments that will help individuals attain the highest level of literacy and numeracy, increasingly important skills in the modern world.

## Methods

**Twins Early Development Study.** TEDS recruited over 15,000 families of twins born in England and Wales<sup>30</sup> in 1994, 1995 and 1996 and the sample remains representative of the UK population<sup>2</sup> (Supplementary Note 2). Ethical approval for TEDS has been provided by the Institute of Psychiatry ethics committee, reference number 05/Q0706/228. We excluded from the analyses children with severe current medical problems and children who had suffered severe problems at birth or whose mothers had suffered severe problems during pregnancy. We also excluded twins whose zygosity was unknown or uncertain, whose first language was other than English, and included only twins whose parents reported their ethnicity as 'white', which is 93% of this UK sample.

At age 12, the TEDS twins participated in web- and telephone-based testing, as described previously<sup>31</sup>. Four measures of reading ability were used: two measures of reading comprehension and a measure of reading fluency presented on the web, and a fourth measure (TOWRE) administered over the telephone. Mathematics ability was assessed using a web-based battery of tests that included questions from three components of mathematics, based on the UK national curriculum. Both the reading and mathematics phenotypes comprised an equally weighted combination of the quantile-normalized scales. For each phenotype, we regressed out the effect of age before further analyses. Further details are provided in Supplementary Note 2.

Phenotypic measurements were available for 2,243 (reading) and 2,772 (maths) of these, with 2,794 samples having at least one measurement and 2,221 samples having both. The sample genotyped on the ImmunoChip ( $N = 2,432$ ) included  $N = 2,153$  individuals with at least one phenotypic measurement, of which



$N = 1,388$  were DZ co-twins of individuals in the discovery sample. For analyses, taking into account family structure, we additionally included untyped MZ and DZ co-twins of individuals typed in the discovery or Immunochip phases ( $N = 1,737$ ) for a combined sample of  $N = 7,323$ . For twin analyses, to parallel the population-based variance estimates as closely as possible, we used a sub-sample of the full TEDS cohort: the 2,794 informative samples with genome-wide data plus their co-twins, for a sample size of  $N = 2,794$  twin pairs.

**Avon Longitudinal Study of Parents and Children.** ALSPAC recruited more than 14,000 pregnant women in the former Avon area of the UK (around Bristol and Bath), with estimated dates of delivery between April 1991 and December 1992 (ref. 32). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. The ALSPAC study website contains details of all the data that are available through a fully searchable data dictionary (<http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/>). The sample used for replication here is a population-representative group of participants who were tested for word-reading efficiency (TOWRE) at the age of 12.5 years. After combining TOWRE scores across two subtests (see Supplementary Note 2) a total of  $N = 2,140$  samples were available for analysis.

**GWAS genotyping and imputation.** Samples were genotyped at the Affymetrix's service laboratory on the Genome-Wide Human SNP Array 6.0. For all samples passing Affymetrix's laboratory quality control, raw intensities (from the CEL files) were renormalized within collections using CelQuantileNorm (<http://sourceforge.net/projects/outmodedbonsai/files/CelQuantileNorm/>). These normalized intensities were used to call genotypes with an updated version of the Chiamo software<sup>33</sup> adapted for Affymetrix 6.0 SNP data.

As is the standard practice for GWAS studies, we excluded sets of individuals whose genome-wide patterns of diversity are outliers compared with the majority of those in the study<sup>34</sup>, and we excluded SNPs for which there is evidence that genotype calls do not provide precise estimates of genotype frequencies. Details of the quality control methods used are published elsewhere<sup>34,35</sup>. In total, 465 of 3,665 samples were excluded from the analyses by these criteria (see Supplementary Table 5). Genotypes and phenotypes from the discovery sample will be made available through the European Genome-Phenome Archive (<https://www.ebi.ac.uk/ega/>).

To assess relatedness among study individuals, we compared each individual with the 100 individuals they were most closely related to (on the basis of genome-wide levels of allele sharing) and used a hidden Markov model (HMM) to decide, at each position in their genome, whether the two individuals shared 0, 1 or 2 chromosomes identical by descent. We obtained a set of 'unrelated' individuals with identity by descent  $< 5\%$  by iteratively removing the member of each pair of putatively related individuals with more missing genotypes. A total of 3,154 individuals were included in subsequent analyses.

In addition to standard SNP filters (Supplementary Table 6), we considered a measure of the statistical information (the IMPUTE info measure) carried by the genotype calls for the underlying allele frequency<sup>36</sup>. SNPs were removed prior to imputation if this information measure was below 0.98 or if the estimated minor allele frequency was below 1%. In total 84,029 (9%) of SNPs were removed by these criteria.

We imputed additional genotypes from a combined reference panel of the 120 CEU trios in HapMap2 and HapMap3 and the common control group of WTCCC2. As an additional quality control step, prior to imputation we re-imputed each typed SNP using IMPUTE version 1 and removed any SNP where the concordance between typed and imputed genotypes was  $< 0.965$ . We used this high-confidence subset of 736,939 SNPs from the array to impute additional genotypes using IMPUTE2 (refs 36,37).

IMPUTE2 adopts a two-stage approach using both a haploid reference panel and a diploid reference panel. For the haploid reference panel, we used HapMap2 and HapMap3 SNP data on the 120 unrelated CEU trios; and for the diploid reference, we used a merged set of genotype calls from Affymetrix 6.0 and Illumina 1.2M genotyping chip typed on 5000 1958 Birth Cohort (58C) and National Blood Service (NBS) individuals forming the common control group of WTCCC2. For association testing we included SNPs with info measure of at least 0.98 (if imputed from HapMap) or 0.9 (if imputed from the WTCCC2 controls) and having an estimated minor allele frequency of at least 1%.

**Immunochip genotyping.** Replication samples were typed on the Illumina 'Immunochip', a custom chip designed by the Immunochip Consortium and WTCCC2, at the Wellcome Trust Sanger Institute. Bead intensity data were processed and normalized for each sample in BeadStudio. Data for successfully genotyped samples were extracted and genotypes were called using the Illuminus algorithm<sup>38</sup>. Samples and SNPs were subject to similar quality control procedures as described above.

**ALSPAC genotyping.** We obtained data for 194 SNPs from the region 48891732-49091732 (NCBI build 36) on chromosome 19 for ALSPAC participants by application to the ALSPAC executive. Details on ALSPAC genotyping and imputation are described elsewhere<sup>39</sup>. Samples were included in the analysis if they had

attended the TOWRE test session and completed both parts of the test.  $N = 63$  individuals who were recorded as having scored zero on either part of the test were removed, leaving a total of  $N = 2,077$  individuals for analysis.

**Genome-wide association analysis.** After quality control, 1,588,650 SNPs were analyzed in the GWAS using SNPTEST ([https://mathgen.stats.ox.ac.uk/genetics\\_software/snpctest/snpctest.html](https://mathgen.stats.ox.ac.uk/genetics_software/snpctest/snpctest.html)), fitting an additive linear model to the data, with sex as a covariate. We used the missing data likelihood score test implemented in SNPTEST to compute  $P$ -values, and refer to these as  $P_{\text{GWAS}}$  above. SNPs with a  $P$ -value  $< 5 \times 10^{-5}$  were further analyzed in a sample combining the GWAS and Immunochip participants, along with informative co-twins with phenotype data but no genotypes available ( $N = 1,737$ ). To take account of the relatedness in the combined sample, we fitted the association model using Merlin software<sup>40</sup>, again with sex as a covariate. The Merlin software also infers the posterior probabilities of missing genotype information from available pedigree information to increase power. For the region on chromosome 19 we analyzed the ALSPAC data using SNPTEST again with sex as a covariate.

**Variance component analysis.** For twin and population-level analyses, we consider a general partitioning of a quantitative phenotype  $Y$  (either reading or mathematics ability in our study) into five components  $Y = A + D + I + C + E$ , where  $A$ ,  $D$  and  $I$  correspond to additive, dominance and interaction genetic effects over the whole genome, respectively, and  $C$  and  $E$  are within-family and individual environmental effects, respectively. We assume that these components are defined to be uncorrelated with each other and thus the phenotypic variance is also partitioned into five components  $V_Y = V_A + V_D + V_I + V_C + V_E$ .

**Bivariate twin analysis.** We consider the traditional ACE twin model (see Supplementary Methods) assuming that dominance and interaction effects are zero ( $D = I = 0$ ). To extend the model to bivariate phenotype, we introduce three parameters  $\rho_A$ ,  $\rho_C$  and  $\rho_E$  to describe the correlation between additive genetic, shared environmental and individual environmental effects, respectively, between reading and mathematics abilities (see Supplementary Methods). We use the model to estimate the variance components  $V_A$ ,  $V_C$ ,  $V_E$  and the three correlation parameters. The narrow-sense heritability is then defined as the ratio  $-V_A/V_Y$ .

For the twin analyses, standardized residuals correcting for age and sex were used because the age of twins is perfectly correlated across pairs, which means that, unless corrected, variation within each age group at the time of testing would contribute to the correlation between twins and be misrepresented as shared environmental influence. The same applies to the sex of the twins, since MZ twins are always of the same sex. The model was fitted using full information maximum likelihood analysis of raw data in the structural equation modelling R package *OpenMx*<sup>41</sup>, estimating the variance parameters for both phenotypes together with the correlation parameters.

**Bivariate population-level analysis.** As opposed to the twin model, the population-level model considers only individuals who are not closely related. The univariate version of this model was recently introduced to study human height<sup>22,25</sup> and subsequently further assessed<sup>24</sup>. The bivariate extension was also recently considered<sup>20,42</sup>.

This model decomposes the variance into an additive genetic component ( $G$ ) that is due to the available panel of SNPs, and the residual component which in principle includes  $D$ ,  $I$ ,  $C$  and  $E$  as defined previously, together with the part of the additive component  $A$  that is not captured by  $G$ . Thus it can be used for estimating a lower bound for the variance and covariance (or correlation) between the additive genetic components of the two phenotypes (Supplementary Methods). A caveat of this model is that environmental effects that correlate with genetics can act as potential confounders.

As for the twin model, we extend the population-level model to the bivariate case by introducing parameters  $\rho_G$  and  $\rho_e$  to describe the correlation between genetic and residual effects between phenotypes (see Supplementary Methods). This model was previously used elsewhere<sup>20</sup>.

For population-level analysis, we included one member of each of those 2,221 twin pairs for which both the reading and the mathematics ability was measured. After quality control 686,458 autosomal SNPs from the Affymetrix array were included in the analysis. We used linear regression to adjust the phenotypes for age, sex and population structure (using 10 PCs) before the variance component analysis. We checked that the programme GCTA<sup>22</sup> gave very similar results for the variance parameters as our own implementation (see Supplementary Methods).

We implemented a Metropolis-Hastings random walk algorithm to explore the posterior distribution on the parameters (here proportional to the likelihood due to the use of uniform priors, see Supplementary Figs 4 and 5 and Supplementary Methods), and compute credible intervals. Finally, we computed a Bayes factor to quantify the evidence for this model relative to the model where  $\rho_G = 0$ , and used both permutations and simulations to obtain a  $P$ -value for this model comparison (Supplementary Fig. 5 and Supplementary Methods).

**Interpretation of twin and population-level estimates.** Although the twin and population models are similar in spirit, they differ in modelling assumptions

and parameter interpretation. Here we list factors that could potentially lead to differences between the model estimates.

Population-level estimates of heritability and genetic correlation take into account only those genetic factors which are tagged by variants present on the genotyping platform. Consequently the population-level model gives a lower-bound estimate of heritability. In particular, we chose not to include genotypes of SNPs on the sex chromosomes in order to help simplify the model and its interpretation. In principle, twin model estimates capture all genetic factors which produce differences between trait values for siblings.

Twin model estimates of narrow-sense heritability may be biased upwards by the presence of dominance or interaction effects<sup>23</sup>. By contrast, because the population-level model employs only distantly related individuals, dominance and interaction effects are expected to be much less important effects.

Twin models take into account only those genetic factors which lead to differences in trait values between siblings: consider an unmeasured environmental variable  $S$  that depends only on the family (socio-economic status may be one such variable). If  $S$  is correlated with genetics—for example, through parental genotypes—its effect in the twin model, where  $S$  is unmodelled, will be to increase the estimated proportion of  $C^{43}$  and so to act to deflate estimates of genetic contributions. However, since shared environment is not modelled in the population-level model,  $S$  potentially contributes to the estimated proportion of  $G$ . In principle, including PCs in the model may help control for this effect. We find little difference in the estimates between the model including and not including PCs (Supplementary Table 4).

More complex effects, including interaction between genetic and environmental influences, could potentially have different effects on the two models. For example, twin model estimates of heritability may be affected by within-family (that is, shared) environmental effects that are more similar between MZ twins than DZ twins<sup>44,45</sup>. More generally, environmental effects that correlate or interact with genetics and other factors are not modelled directly<sup>43,46</sup>. However, several studies have shown that these assumptions of the twin model are usually reasonable in practice<sup>44,46</sup>. Population-level estimates appear to be remarkably robust to deviations from modelling assumptions<sup>24</sup>.

The population-level model assumes a particular dependency between the minor allele frequency at a SNP and the size of the effect of that SNP on the phenotype (see Supplementary Methods). Simulation studies<sup>24</sup> suggest the model is fairly robust to deviations from this assumption.

## References

- Butterworth, B. & Kovas, Y. Understanding neurocognitive developmental disorders can improve education for all. *Science* **340**, 300–305 (2013).
- Kovas, Y., Haworth, C. M., Dale, P. S. & Plomin, R. The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monogr. Soc. Res. Child. Dev.* **72** vii 1–144 (2007).
- Davies, G. *et al.* Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol. Psychiatry* **16**, 996–1005 (2011).
- Ritchie, S. J. & Bates, T. C. Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychol. Sci.* **24**, 1301–1308 (2013).
- Batty, G. D., Kivimaki, M. & Deary, I. J. Intelligence, education, and mortality. *BMJ* **340**, c563 (2010).
- Harlaar, N., Spinath, F. M., Dale, P. S. & Plomin, R. Genetic influences on early word recognition abilities and disabilities: a study of 7-year-old twins. *J. Child. Psychol. Psychiatry* **46**, 373–384 (2005).
- Plomin, R. & Kovas, Y. Generalist genes and learning disabilities. *Psychol. Bull.* **131**, 592–617 (2005).
- Stanovich, K. E. & Siegel, L. S. Phenotypic performance profile of children with reading disabilities—a regression-based test of the phonological-core variable-difference model. *J. Educ. Psychol.* **86**, 24–53 (1994).
- Snowling, M., Bishop, D. V. & Stothard, S. E. Is preschool language impairment a risk factor for dyslexia in adolescence? *J. Child. Psychol. Psychiatry* **41**, 587–600 (2000).
- Shaywitz, S. E. Dyslexia. *New Engl. J. Med.* **338**, 307–312 (1998).
- Beitchman, J. H. & Young, A. R. Learning disorders with a special emphasis on reading disorders: a review of the past 10 years. *J. Am. Acad. Child. Adolesc. Psychiatry* **36**, 1020–1032 (1997).
- Gersten, R., Jordan, N. C. & Flojo, J. R. Early identification and interventions for students with mathematics difficulties. *J. Learn. Disabil.* **38**, 293–304 (2005).
- Geary, D. C. Mathematics and learning disabilities. *J. Learn. Disabil.* **37**, 4–15 (2004).
- Cardon, L. R. *et al.* Quantitative trait locus for reading disability on chromosome 6. *Science* **266**, 276–279 (1994).
- Scerri, T. S. *et al.* DCDC2, KIAA0319 and CMIP are associated with reading-related traits. *Biol. Psychiatry* **70**, 237–245 (2011).
- Meaburn, E. L., Harlaar, N., Craig, I. W., Schalkwyk, L. C. & Plomin, R. Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Mol. Psychiatry* **13**, 729–740 (2008).
- Docherty, S. J. *et al.* A genome-wide association study identifies multiple loci associated with mathematics ability and disability. *Genes Brain Behav.* **9**, 234–247 (2010).
- Luciano, M. *et al.* A genome-wide association study for reading and language abilities in two population cohorts. *Genes Brain Behav.* **12**, 645–652 (2013).
- Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467–1471 (2013).
- Deary, I. J. *et al.* Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* **482**, 212–215 (2012).
- Markowitz, E. M., Willemsen, G., Trumbetta, S. L., van Beijsterveldt, T. C. & Boomsma, D. I. The etiology of mathematical and reading (dis)ability covariation in a sample of Dutch twins. *Twin Res. Hum. Genet.* **8**, 585–593 (2005).
- Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
- Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: genetic interactions create phantom heritability. *Proc. Natl Acad. Sci. USA* **109**, 1193–1198 (2012).
- Zaitlen, N. & Kraft, P. Heritability in the genome-wide association era. *Hum. Genet.* **131**, 1655–1664 (2012).
- Visscher, P. M., Yang, J. & Goddard, M. E. A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang *et al.* (2010). *Twin Res. Hum. Genet.* **13**, 517–524 (2010).
- Lind, P. A. *et al.* Dyslexia and DCDC2: normal variation in reading and spelling is associated with DCDC2 polymorphisms in an Australian population sample. *Eur. J. Hum. Genet.* **18**, 668–673 (2010).
- Meng, H. *et al.* DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Proc. Natl Acad. Sci. USA* **102**, 17053–17058 (2005).
- Marino, C. *et al.* DCDC2 genetic variants and susceptibility to developmental dyslexia. *Psychiatr. Genet.* **22**, 25–30 (2012).
- Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
- Haworth, C. M., Davis, O. S. & 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.* **16**, 117–125 (2013).
- Haworth, C. M. *et al.* Internet cognitive testing of large samples needed in genetic research. *Twin Res. Hum. Genet.* **10**, 554–563 (2007).
- Boyd, A. *et al.* Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
- Bellenguez, C., Strange, A., Freeman, C., Donnelly, P. & Spencer, C. C. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics* **28**, 134–135 (2012).
- Barrett, J. C. *et al.* Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat. Genet.* **41**, 1330–1334 (2009).
- Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
- Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* **457–470** (2011).
- Teo, Y. Y. *et al.* A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* **23**, 2741–2746 (2007).
- Paternoster, L. *et al.* Adult height variants affect birth length and growth rate in children. *Hum. Mol. Genet.* **20**, 4069–4075 (2011).
- Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30**, 97–101 (2002).
- Boker, S. *et al.* OpenMx: an open source extended structural equation modeling framework. *Psychometrika* **76**, 306–317 (2011).
- Korte, A. *et al.* A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat. Genet.* **44**, 1066–1071 (2012).
- Purcell, S. Variance components models for gene-environment interaction in twin analysis. *Twin Res.* **5**, 554–571 (2002).
- Boomsma, D., Busjahn, A. & Peltonen, L. Classical twin studies and beyond. *Nat. Rev. Genet.* **3**, 872–882 (2002).
- Martin, N., Boomsma, D. & Machin, G. A twin-pronged attack on complex traits. *Nat. Genet.* **17**, 387–392 (1997).
- Keller, M. C., Medland, S. E. & Duncan, L. E. Are extended twin family designs worth the trouble? A comparison of the bias, precision, and accuracy of parameters estimated in four twin family models. *Behav. Genet.* **40**, 377–393 (2010).

## Acknowledgements

Funding for this study was provided by the Wellcome Trust Case Control Consortium 2 project (085475/B/08/Z and 085475/Z/08/Z) and TEDS is supported by a program grant from the UK Medical Research Council (G0901245). The UK Medical Research Council and the Wellcome Trust (Grant ref: 092731) and the University of Bristol provide core support for ALSPAC. O.S.P. Davis was supported by a Sir Henry Wellcome Fellowship from the Wellcome Trust (WT088984). M. Pirinen is supported by the Academy of Finland (257654). C.M.A. Haworth was supported by a research fellowship from the British Academy. R. Plomin was supported by a research professorship from the UK Medical Research Council (G19/2) and a European Research Council Advanced Investigator Award (295366). P. Donnelly was supported in part by a Wolfson-Royal Society Merit Award. C.C.A. Spencer was supported by a Wellcome Trust Fellowship [097364/Z/11/Z], and work was supported in part by Wellcome Trust Centre for Human Genetics core grants 090532/Z/09/Z and 075491/Z/04/B. We acknowledge use of the British 1958 Birth Cohort DNA collection funded by the Medical Research Council (grant G0000934) and the Wellcome Trust (grant 068545/Z/02), and the UK National Blood Service controls funded by the Wellcome Trust. E. Bramon holds a Medical Research Council new investigator award. We also thank the National Institutes of Health Research Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.

## Author contributions

Writing Group: O.S.P.D., G.B., M.P., P.D., R.P. and C.C.A.S. Data analysis: O.S.P.D., G.B., M.P., A.S., T.S.P., C.F., C.B., Z.S., R.P., D.V., P.D., C.C.A.S., K.B.H., M.T., L.S.S. and

C.M.L. Sample preparation and genotyping: E.L.M., S.J.D., C.J.C.C., I.W.C., C.L., P.D., S.H., E.G., S.D., S.C.P., A.T.G., S.E. and S.J.B. Phenotype data planning and collection: C.M.A.H., Y.K., N.H., S.A.P. and P.S.D. WTCCC2 Management Committee: P.D., I.B., J.M.B., E.B., M.A.B., J.P.C., A.C., P.D., A.D., J.A.Z.J., H.S.M., C.G.M., C.N.A.P., R.P., A.R., S.J.S., R.C.T., A.C.V. and N.W.W.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Davis, O. S. P. *et al.* The correlation between reading and mathematics ability at age twelve has a substantial genetic component *Nat. Commun.* 5:4204 doi: 10.1038/ncomms5204 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>