Genome-wide screening for DNA variants associated with reading and language traits


© 2014 The Authors.

Reading and language abilities are heritable traits that are likely to share some genetic influences with each other. To identify pleiotropic genetic variants affecting these traits, we first performed a genome-wide association scan (GWAS) meta-analysis using three richly characterized datasets comprising individuals with histories of reading or language problems, and their siblings. GWAS was performed in a total of 1862 participants using the first principal component computed from several quantitative measures of reading- and language-related abilities, both before and after adjustment for performance IQ. We identified novel suggestive associations at the SNPs rs59197085 and rs5995177 (uncorrected \( P \approx 10^{-7} \) for each SNP), located respectively at the CCDC136/FLNC and RBFOX2 genes. Each of these SNPs then showed evidence for effects across multiple reading and language traits in univariate association testing against the individual traits. FLNC encodes a structural protein involved in cytoskeleton remodelling, while RBFOX2 is an important regulator of alternative splicing in neurons. The CCDC136/FLNC locus showed association with a comparable reading/language measure in an independent sample of 6434 participants from the general population, although involving distinct alleles of the associated SNP. Our datasets will form an important part of ongoing international efforts to identify genes contributing to reading and language skills.

Keywords: pleiotropic variants, CLDRC, developmental dyslexia, GWAS, language, meta-analysis, reading, reading disability, SLIC, specific language impairment

Received 24 March 2014, revised 24 June 2014, 14 July 2014 and 20 July 2014, accepted for publication 24 July 2014

Reading disability (RD, also known as developmental dyslexia) refers to a significant difficulty in reading that cannot be explained by obvious causes, such as sensory impairments or lack of educational opportunity (Shaywitz et al. 1990). Specific language impairment (SLI) is diagnosed as an unexpected difficulty or delay in acquiring spoken language abilities, despite normal hearing and intelligence, and in absence of overt neurological deficits (Bishop 1994). RD and SLI are among the most prevalent neurocognitive disorders of school-aged children, with prevalence 5–8% in many populations (Shaywitz et al. 1990; Tomblin et al. 1997). Both are complex disorders with moderate to high heritabilities (30–70%) as assessed by studies of families and twins (Barry et al. 2007; Fisher & DeFries 2002).

RD and SLI display high comorbidity: 43% of SLI children are later diagnosed with RD and up to 55% of dyslexic children meet criteria for SLI (McArthur et al. 2000; Snowling et al. 2000). Moreover, RD and SLI show comorbidity with other neurodevelopmental traits including attention deficit hyperactivity disorder (ADHD) (Pennington 2006; Wilcutt et al. 2010) and speech sound disorders (Newbury & Monaco 2010; Pennington & Bishop 2009). It is likely that these disorders arise due to some shared genetic/neurobiological...
mechanisms, as well as non-shared causal factors (Newbury et al. 2011; Paracchini 2011). A study of twins by Harlaar et al. (2008) indicated that an association between early language and later reading is underpinned by common environmental and genetic influences, and a family study by Logan et al. (2011) also found significant genetic correlations of reading and language measures.

Variants of several genes have previously been associated with RD, most notably DDX1C1 (15q21, Taipale et al. 2003), KIAA0319 and DCDC2 (6p22, Cope et al. 2005; Francks et al. 2004; Meng et al. 2005), MRPL19/C2ORF3 (2p12, Anthoni et al. 2007) and ROBO1 (3p12, Bates et al. 2011; Hannula-Jouppi et al. 2005). Similarly, some loci have been implicated in SLI; variants in genes, such as CNTNAP2 (7q36, Vernes et al. 2008) and CMIP and ATP2C2 (16q23.2-24, Newbury et al. 2009) show associations with quantitative traits in children with typical SLI, while rare mutations of FOXP2 (7q31, Fisher & Scharff 2009) cause a monogenic speech and language disorder. These genes were mostly identified through linkage analysis followed by either positional cloning or else targeted association mapping. Functional analyses suggest that some of these genes mediate important processes in central nervous system (CNS) development, such as neuronal migration, axonal guidance and neurite outgrowth (Carroll-Castillo et al. 2013; Vernes et al. 2011). A subset of the candidate genes may contribute to both RD and SLI, again indicating a partial genetic overlap for these traits (Bates et al. 2011; Newbury et al. 2011; Scerri et al. 2011b). Crucially, an overwhelming majority of the heritable variance in reading and language skills is unexplained, and the molecular mechanisms that contribute to RD and SLI remain largely unknown (Newbury & Monaco 2010; Peterson & Pennington 2012).

Some of the genetic variation contributing to RD and SLI is also likely to impact on reading/language skills in the general population (Bates et al. 2011; Luciano et al. 2007; Paracchini et al. 2008, 2011; Scerri et al. 2011b; Whitehouse et al. 2011). To detect previously undiscovered associations of common genetic variants with reading and language skills, it is therefore appropriate to sample broad ranges of the trait distributions in study datasets, while screening over the entire genome.

In recent years, a small number of studies have tried to identify genes involved in reading and/or language through genome-wide association scanning (GWAS). An early GWAS for reading ability used DNA pooling of low vs. high reading ability groups in ~1500 7-year-old children, and a relatively low density SNP microarray with ~107,000 SNPs (Meaburn et al. 2008). The SNPs showing the largest allele frequency differences between low and high ability groups were further genotyped and tested in an additional sample of 4,258 children, with 10 SNPs finally showing nominally significant association with continuous variation in reading ability (Meaburn et al. 2008). A GWAS on mismatch negativity, which is a putative endophenotype of dyslexia derived from electroencephalography, has also been reported based on 386 dyslexic children, and showed replicable association of the SNP rs4234898 on 4q32 along with the haplotype rs4234898-rs1100040 (Roeske et al. 2011). These were shown to affect mRNA expression levels of SLC2A3 (12p13), which codes for a neuronal glucose transporter, suggesting a possible role of glucose levels in memory performance necessary for speech perception in dyslexia (Roeske et al. 2011). More recently, a genome-wide linkage and association scan using ~133,000 SNPs, in 718 subjects from 101 dyslexia-affected families, reported a borderline significant association with dyslexia status at rs9313548, near FGFR1 (5q35.1), which is a gene involved in laminar positioning of cortical neurons during development (Field et al. 2013).

Two GWAS studies have directly attempted to identify shared genetic contributions to reading and language abilities. Luciano et al. (2013), in a GWAS on quantitative reading and language traits in two population datasets (N ~6500), found the strongest association between rs2192161, in the ABCC13 pseudogene (21q11.2), and a nonword repetition measure (p ~7 x 10^-6), while rs4807927 (DAZAP1, 19p13.3) showed association with both word reading and a composite reading–spelling factor score (p ~10^-4 for both traits). In the same study, CDC2L1, CDC2L2, LOC728661 (1p36.33) and RCAN3 (1p36.11) showed significant gene-based associations with the reading–spelling factor (Luciano et al. 2013). A case-control GWAS using a relatively small number of RD (N =353), language impairment (N =163) and comorbid cases (N =174), in comparison to general population controls (N =4117), identified nominally significant associations for the comorbid cases at rs12636438 and rs1679255 in ZNF385D (3p24.3) (Eicher et al. 2013). These SNPs also showed associations with a vocabulary measure and white matter volumes of brain fibre tracts previously implicated in language, in an independent dataset (Eicher et al. 2013).

In the present study, we carried out a GWAS meta-analysis for genetic variants influencing reading and language abilities. We included three long-established datasets comprising children with reading or language problems, along with their siblings. This approach complemented other recent GWAS studies of reading/language performance (Eicher et al. 2013; Luciano et al. 2013) because it included continuous trait variance across a broad range of reading and language abilities, but also involved a pronounced enrichment for poor performance while not applying an arbitrary dichotomy between RD/SLI cases and controls.

Within each dataset, we tested single nucleotide polymorphisms (SNPs), along with single base insertions/deletions (indels), for association with the first Principal Component (PC) derived from a range of reading- and language-related quantitative traits. We then meta-analysed the GWAS results from the separate datasets, followed by gene- and pathway-level analysis, and we checked the most significant associations arising from our analysis within the GWAS results generated by Luciano et al. (2013).

Although we used PC-based analysis as a form of data reduction for the purposes of GWAS, we also investigated the two most significant SNP associations arising from our meta-analysis by using multivariate association modelling in each dataset, and by testing of these SNPs against the individual measures separately. This approach would help to understand the cross-phenotypic effects involved. In other words, the PC-based GWAS was used to identify potential genetic effects on shared variance between multiple reading and language measures, and then pleiotropy was

investigated in more detail through univariate analysis and multivariate modelling, for individual SNPs implicated by the PC-based GWAS meta-analysis. In addition, in order to more closely match the trait measurement across all datasets, we repeated the GWAS and meta-analysis using the first PC of only single word reading and spelling ability, because these were the only two measures available in all datasets.

Some genetic effects on reading and language may be pleiotropic for IQ, whereas other effects may be largely or wholly independent of IQ (Bishop & Snowling 2004; Pennington & Bishop 2009). To detect the latter type of effect it is advantageous to remove the shared variance with IQ that is present in measures of reading and language, prior to association testing. We therefore performed our GWAS analyses both with and without IQ-adjustment of the reading and language measures. In addition, Luciano et al. (2013) analysed only IQ-adjusted data, so that for cross-comparing of results an IQ-adjustment was desirable to include in this study.

Subjects and methods

Datasets

UK Reading Disability (UK-RD)

This dataset comprised children diagnosed with RD, and their siblings, collected at the Dyslexia Research Centre clinics in Oxford and Reading, or the Aston Dyslexia and Development Clinic in Birmingham, UK. Ethical approval was acquired from the Oxfordshire Psychiatric Research Ethics Committee (OPREC O01.02) and written informed consent of the participants (or their parents) was obtained. The total number of participants was 983, mean age 11.7 years, age range 5–31, from 608 independent nuclear families. All children, regardless of diagnosis, were administered psychometric tests of verbal and non-verbal IQ (details further below). A subset of this dataset has been analysed in previous studies on reading (Becker et al. 2009). To detect the latter type of effect it is advantageous to remove the shared variance with IQ that is present in measures of reading and language, prior to association testing. We therefore performed our GWAS analyses both with and without IQ-adjustment of the reading and language measures. In addition, Luciano et al. (2013) analysed only IQ-adjusted data, so that for cross-comparing of results an IQ-adjustment was desirable to include in this study.

SLI Consortium (SLIC)

The SLI Consortium dataset comprised children affected by SLI, along with their siblings, recruited from five centres across the UK; The Newcomen Centre at Guy’s Hospital, London (now called Evelina Children’s Hospital); the Department of Child Health at the University of Edinburgh; the Cambridge Language and Speech Project Newcomen Centre at Guy’s Hospital, London (now called Evelina with their siblings, recruited from five centres across the UK; The

The SLI Consortium dataset has been used for prior linkage studies (Ceroni et al. 2006; The SLI Consortium 2002, 2004). This sample included 49 families from the Guy’s Hospital, London cohort which had not been included in previous SLI Consortium studies. Ethical agreement was given by local ethics committees of the hospitals involved in the consortium, and all subjects provided informed consent. All children in this sample were assessed for a number of reading- and language-related traits (see below) regardless of their language ability. For this study, for MZ twin pairs, we selected one child per pair based on the maximum availability of reading- and language-related trait data, or otherwise randomly. The sample of twins and siblings available for this study comprised 749 participants in total, mean age 11.7 years, age range 8–19, from 343 unrelated twinships/sibships. Of these, 266 of the twinships/sibships (a total of 582 participants) were originally recruited via a proband with a history of RD, and 77 of the twinships/sibships (164 participants in total) were originally recruited via a proband with a history of ADHD. We analysed these two subsets separately for GWAS before meta-analysing the results together with those from the other datasets listed above. The two subsets are indicated hereafter as CLDRC-RD and CLDRC-ADHD. As for the other datasets, no prior GWAS has been reported.

Genotype data generation, quality control (QC) and imputation

DNA was extracted from whole blood or buccal swab samples and prepared for genotyping using standard protocols. Genome-wide genotype data were generated for each dataset using Illumina® SNP arrays. These were the HumanHap 550k for a first genotyping wave of 200 subjects from UK-RD, and the Human OmniExpress (730k SNPs) for SLIC, CLDRC and the remaining UK-RD samples. Data were processed using Illumina’s BeadStudio®/GenomeStudio® software, following the manufacturer’s guidelines. All datasets then underwent a first round of quality control, using functions in the software PLINK v1.07 (Purcell et al. 2007; http://pngu.mgh.harvard.edu/purcell/plink/) in which all SNPs deviating from Hardy–Weinberg Equilibrium (HWE, P < 1 × 10−6), with Minor Allele Frequency (MAF) < 1% and call frequency < 99%, were filtered out. In addition, samples were excluded if they showed inconsistencies in genome-wide identity-by-descent sharing with their siblings and unrelated individuals, or sex mismatches, or call rates < 95%. Multi-dimensional scaling (MDS) analysis of genome-wide genotype data was used to identify any subjects that did not cluster together with the majority of the dataset, and these were discarded, as were any outliers for genome-wide homozygosity. These QC steps were followed by genotype phasing using MACH v1.0 (Li et al. 2010; http://www.sph.umich.edu/csg/abecasis/MACH/index.html) and imputation of SNPs and single-base indels using Minimac (Hovier et al. 2012; http://genome.sph.umich.edu/wiki/Minimac), with the 1000 Genomes Project reference dataset (GIANT all populations panel, Phase 1; v3; The 1000 Genomes Project Consortium, 2012; http://www.1000genomes.org). We excluded poorly imputed polymorphisms (with r2 < 0.3), and deleted individual genotypes with imputation quality scores < 0.9. A final quality control procedure was then run on the imputed data, using PLINK, in which we discarded SNPs with HWE P < 5 × 10−8, MAF < 1%, and call frequency < 95%. Key features of the QC are shown in Table 1. Further details are reported in Appendix S4.

At the end of the genotype QC process, we had data for 959 participants and 6,190,549 polymorphisms in UK-RD, 729 participants and 6,440,000 polymorphisms in CLDRC, and 502 participants and 6,240,842 polymorphisms in SLIC, with 5,518,498 polymorphisms shared across all three datasets.

Effects on SLI (Nudel et al. 2014). However, no GWAS for continuous language and reading scores has yet been reported for this (or any other) SLI sample.
Table 1: Genotype quality control (QC) filters used and number of samples/markers discarded at each step (see Methods and Appendix S4 for details)

<table>
<thead>
<tr>
<th>QC step</th>
<th>CLDRC (749)†</th>
<th>UK-RD (200 + 818)‡</th>
<th>SLIC (548)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HWE (P &lt; 1 \times 10^{-6}) (SNPs)</td>
<td>57</td>
<td>12,631; 191</td>
<td>54</td>
</tr>
<tr>
<td>MAF &lt; 1% (SNPs)</td>
<td>74,770</td>
<td>23,467; 77,342</td>
<td>1,718</td>
</tr>
<tr>
<td>Call freq &lt; 99% (SNPs)</td>
<td>0 †</td>
<td>82,052; 0†</td>
<td>72,043</td>
</tr>
<tr>
<td>Call rate &lt; 98% (samples)</td>
<td>0 ‡</td>
<td>3; 0 ‡</td>
<td>9</td>
</tr>
<tr>
<td>IBD sharing (samples)</td>
<td>11</td>
<td>1; 7</td>
<td>17</td>
</tr>
<tr>
<td>Sex mismatch (samples)</td>
<td>3</td>
<td>0; 8**</td>
<td>13††</td>
</tr>
<tr>
<td>Homozygosity outlier (samples)</td>
<td>6</td>
<td>1; 3</td>
<td>2</td>
</tr>
<tr>
<td>MDS outlier (samples)</td>
<td>0</td>
<td>0; 2</td>
<td>5</td>
</tr>
<tr>
<td>HWE (P &lt; 5 \times 10^{-6}) (SNPs)*</td>
<td>2,166</td>
<td>2,779</td>
<td>2,096</td>
</tr>
<tr>
<td>MAF &lt; 1% (SNPs)*</td>
<td>3,640,742</td>
<td>1,980,500</td>
<td>3,260,639</td>
</tr>
<tr>
<td>Call freq &lt; 95% (SNPs)*</td>
<td>1,729,493</td>
<td>1,704,412</td>
<td>1,766,376</td>
</tr>
<tr>
<td>Call rate &lt; 95%, MDS outliers, IBD sharing (samples)*</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Passing QC</td>
<td>729 (6,427,200)</td>
<td>959 (6,190,549)</td>
<td>502 (6,240,842)</td>
</tr>
</tbody>
</table>

Final number of samples (and SNPs in brackets) passing the genotype QC are reported in the bottom row. Note that these numbers do not also account for QC of the trait scores.

*After imputation QC. Before this step, imputed SNPs with \(r^2 < 0.3\) were filtered out, and all the genotypes with quality score <0.9 were set to missing.
†As CLDRC-RD and CLDRC-ADHD were processed together and drawn from the same population, we treated them as a single dataset in the genotype QC.
‡As UK-RD samples had been genotyped on two different Illumina® platforms (see Methods) the subsets were analysed separately before imputation, and pre-imputation QC details are therefore reported for both the subsets (first genotyping wave with HumanHap 550k and second genotyping wave with Human OmniExpress). Note that 35 samples were genotyped on both of the arrays, and one of these samples showed inconsistent genotyping and was therefore discarded in both subsets.
††The high number of SNPs discarded at this stage was due to the fact that no quality filter had been applied on this subset during genotype call process (see Appendix S4).
¶In these cases, SNPs with call frequency <99% and samples with call rate <98% had already been discarded during genotype call process (see Appendix S4).
**Includes three sex chromosome abnormalities carriers.
†††Includes nine samples with sex chromosome abnormalities and one with X chromosome call rate <95%.

Reading and language measures

Table 2 lists the reading- and language-related traits that were assessed in the different datasets, as detailed in prior publications (Compton et al. 2001; Francks et al. 2004; Friend & Olson 2010; The SLI Consortium 2002, 2004). Further information on these measures is given in Tables 3–5. To remove outliers, trait scores were otherwise age-adjusted according to normative data (Compton et al. 2001; Francks et al. 2004; Friend & Olson 2010; The SLI Consortium 2002, 2004). When a measure differed significantly from normality, we performed a within-dataset rank-normalization to attain normality and excluding subjects with any missing measure (missing listwise option). A Kaiser–Meyer–Olkin measure of sampling adequacy and a Bartlett’s test of sphericity were run in all the PCAs. These tests revealed a high common variance (KMO = 0.8–0.9) and a significant interdependence (Bartlett’s test \(P < 0.05\)) among the variables examined in each dataset, justifying the PCAs. The proportion of total variance explained by PC1 was 75.3% in UK-RD, 68.6% in SLIC, 64.5% in CLDRC-RD and 52.0% in CLDRC-ADHD. In all the datasets PC2 explained no more than 13% of the total variance. All of the PC1s showed a broad pattern of loadings across the traits (Table 2). The total number of participants for which we finally obtained genotype and PC1 data (i.e. all datasets combined) was 1862. We also obtained residuals from regressing PC1 against performance IQ (which had not been included in PC1 computation), again separately within each dataset. A measure of performance IQ was not available for 36 of the 1862 participants, and therefore the total sample size for IQ-adjusted PC1 was 1826. We also derived a first principal component score within each dataset from only word reading and spelling, because these were the only measures available in all datasets and therefore provided a possibility to match traits as closely as possible across datasets. The first PC derived from word reading and spelling is referred to as PC1read hereafter. The proportion of variance in word reading and spelling explained by PC1read was 86.9% in UK-RD, 88% in CLDRC-RD, 93.4% in SLIC and 80.1% in CLDRC-ADHD. As only two measures were correlated measures into a smaller set of latent variables (factors or principal components) that can explain the maximum amount of shared variance (Field 2005). In each dataset, only linear components with Eigenvalue >1 were extracted, allowing for correlation among the components (oblique rotation, direct oblim method) and excluding subjects with any missing measure (missing listwise option). A Kaiser–Meyer–Olkin measure of sampling adequacy and a Bartlett’s test of sphericity were run in all the PCAs. These tests revealed a high common variance (KMO = 0.8–0.9) and a significant interdependence (Bartlett’s test \(P < 0.05\)) among the variables examined in each dataset, justifying the PCAs. The proportion of total variance explained by PC1 was 75.3% in UK-RD, 68.6% in SLIC, 64.5% in CLDRC-RD and 52.0% in CLDRC-ADHD. In all the datasets PC2 explained no more than 13% of the total variance. All of the PC1s showed a broad pattern of loadings across the traits (Table 2). The total number of participants for which we finally obtained genotype and PC1 data (i.e. all datasets combined) was 1862. We also obtained residuals from regressing PC1 against performance IQ (which had not been included in PC1 computation), again separately within each dataset. A measure of performance IQ was not available for 36 of the 1862 participants, and therefore the total sample size for IQ-adjusted PC1 was 1826. We also derived a first principal component score within each dataset from only word reading and spelling, because these were the only measures available in all datasets and therefore provided a possibility to match traits as closely as possible across datasets. The first PC derived from word reading and spelling is referred to as PC1read hereafter. The proportion of variance in word reading and spelling explained by PC1read was 86.9% in UK-RD, 88% in CLDRC-RD, 93.4% in SLIC and 80.1% in CLDRC-ADHD. As only two measures were

First Principal Component score computation

The First Principal Component from all of the language- and reading-related traits available (PC1, Table 2) was derived in each dataset, through the SPSS® 20.0 Factor Analysis (Principal Component extraction method, hereafter called PCA). This reduced our...
used to construct PC1\textsubscript{read}, then the measures loaded equally onto this component, and the loadings were high in all datasets (≥0.9). PC1\textsubscript{read} was therefore a highly comparable construct across datasets (see Appendix S3). Moreover, the correlation between PC1 and PC1\textsubscript{read} was high in each dataset (Pearson’s \( r = 0.925 \) in CLDRC-RD, \( 0.947 \) in UK-RD, \( 0.914 \) in SLIC and \( 0.917 \) in CLDRC-ADHD), so that PC1 itself could also be regarded as highly comparable across datasets. Note that these correlations were based on repeat random sampling of one member from each unrelated sibling (as for all pairwise trait correlations; see above). The total number of subjects across all datasets for PC1\textsubscript{read} was 1913, and for IQ-adjusted PC1\textsubscript{read} it was 1875. We primarily focused on PC1 for our subsequent genetic analysis (below), because this would maximize the chance of identifying SNPs that affect variance shared between both reading and language measures. However, we also repeated GWAS meta-analysis using PC1\textsubscript{read} to provide a comparable analysis that would be minimally affected by the heterogeneity of available measures across datasets.

### Genetic association analyses

#### Sibling-pair GWAS

Sibling-based genome-wide association analyses were conducted using PC1 and PC1\textsubscript{read} scores separately within each dataset, both before and after IQ-adjustment, and using the ‘total’ association option of the QFAM function implemented in PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/; Purcell et al. 2007). This method tests for association at each SNP by regressing trait scores on genotypes in an additive linear model. To correct for non-independence of siblings, permutations were run (i.e. label-swapping of phenotypes/genotypes) to obtain empirical significance levels (further details in Appendix S4).

GWAS meta-analysis (GWASMA)

The results from GWAS in the separate datasets were then meta-analysed together. This was implemented in the programme METAL (http://www.sph.umich.edu/csg/abecasis/Metal/index.html; Willer et al. 2010). We chose an approach that does not assume equivalence of allele effect sizes between datasets, which was appropriate given the heterogeneity of study recruitment and assessment. Put briefly, the GWAS meta-analysis tested each SNP for a genetic effect, across the contributing datasets, computing an overall \( z \)-score for that SNP determined by the \( P \) value, the direction of the allele effect and the sample size of each study involved in the meta-analysis.

Gene-based analysis

The results of the GWASMA on PC1 were used as input for gene-based association analyses using VEGAS v0.8.27 (http://gump.qimr.edu.au/VEGAS/; Liu et al. 2010). This software performs association tests for \(~10,000\) autosomal genes, by assigning multiple SNPs to each individual gene according to their genomic locations, and then combining the evidence for association across all SNPs assigned to a given gene, while taking into account the linkage disequilibrium (LD) structure between SNPs. Each tested gene also included potentially regulatory regions located up to 50kb beyond the 5′- and 3′-untranslated regions (UTRs). A Bonferroni-corrected significance threshold was set at \( P < 2.8 \times 10^{-8} \) to account for the number of genes tested (see Appendix S4 for details).

Pathway-based analysis

Finally, a pathway/network-based association analysis was run using the PC1 GWASMA results, with the programme INRICH v1.0 (http://atgu.mgh.harvard.edu/inrich/started.html; Lee et al. 2012).
GWAS of reading and language

Table 3: Language-reading-related traits available in the UK-RD dataset

<table>
<thead>
<tr>
<th>Trait</th>
<th>Test</th>
<th>Test description*</th>
<th>Statistical elaboration†</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRead</td>
<td>British Ability Scale (BAS) Wide Range Achievement Test-Revised (WRAT-R)</td>
<td>Reading aloud a series of real words presented on a card</td>
<td>A, S, R</td>
</tr>
<tr>
<td>WSpell</td>
<td>BAS/WRAT-R</td>
<td>Writing words that are dictated by the test administrator</td>
<td>A, S, R</td>
</tr>
<tr>
<td>PD</td>
<td>Castles &amp; Coltheart (C&amp;C)</td>
<td>Reading aloud nonsense words of increasing difficulty, according to English grapheme-phoneme conversion rules</td>
<td>A, S, R</td>
</tr>
<tr>
<td>PA</td>
<td>Spoonerism test</td>
<td>Simple phoneme deletion and substitution (e.g. replace the first sound in <em>dog</em> with <em>V</em> to make <em>log</em>)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>OC</td>
<td>C&amp;C</td>
<td>Complex phoneme deletion and substitution Spoonerism (swapping the first sounds of two words, e.g. from <em>spoon</em> to <em>doon</em>, <em>sppogi</em>)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>vIQ</td>
<td>BAS/Wechsler Adult Intelligence Scale – Revised (WAIS-R)</td>
<td>Reading aloud irregular words of increasing difficulty (i.e. words whose pronunciation does not follow the English grapheme-phoneme conversion rules, e.g. <em>yacht</em>)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>pIQ</td>
<td>BAS</td>
<td>Similarities subtest only (explaining how two/three words are similar or go together)</td>
<td>A, S, R</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Matrices subtest only (predicting missing components of increasingly complex matrices containing abstract symbols)</td>
<td>A, S</td>
</tr>
</tbody>
</table>

Superscript numbers after each test indicate the initial reference for it, where further details on the test can be found. 1 Elliot et al. 1979; 2 Jastak & Wilkinson 1984; 3 Castles & Coltheart 1993; 4 Coltheart & Leahy 1996; 5 Gallagher & Frederickson 1995; 6 Frederickekson 1995; 7 Wechsler 1981.

*Where more than one battery is administered, the total score is usually computed as a sum of the raw scores from each subtest.

†Legend of trait adjustments: A, age-adjusted; S, standardized against the normative mean of the population of reference; R, further rank-normalized (using Blom’s formula), because the trait distribution after standardization differed from normality (Shapiro–Wilks test \( P < 0.05 \)).

This tool tests for an enrichment of association within predefined gene sets, through a permutation-based approach. We defined associated genomic intervals as those containing an individual association \( P < 0.001 \) in the GWASMA results. Gene boundaries were again defined as extending 50 kb beyond the 5’- and 3’-UTRs. Three candidate gene lists, based on the gene sets of the Gene Ontology Database (http://www.geneontology.org/), were tested for an enrichment of association. These represented three distinct neurobiological hypotheses on the aetiology of reading and language disabilities: axon guidance (including all the GO sets containing the term ‘axon guidance’), neuronal migration (including all the GO sets containing the term ‘neuron migration’) and steroid sex hormone biology (including all the GO sets containing the terms ‘steroid’, ‘androgen’, ‘oestrogen’, ‘progestosterone’ and ‘testosterone’). Further details on the analysis can be found in Appendix S4.

Further analysis of top association signals

Effect sizes on different traits
We repeated the regressions of PC1 and IQ-adjusted PC1 on the genotypes of our two most significantly associated SNPs from GWAS meta-analysis, in an additive linear model, in order to conveniently obtain the regression \( r^2 \) as indicative measures of effect sizes. To generate measures unbiased by sample relatedness, regression \( r^2 \) were calculated in \( R \) (R Core Team 2013, http://www.r-project.org/) as the median \( r^2 \) over 100 repeat random samplings of one individual from each independent sibship, separately in each dataset.

We further investigated each of our top two association signals by running QFAM univariate association tests in PLINK v1.07 (Purcell et al. 2007) for each individual trait that was used in constructing PC1, and separately in each dataset. This analysis provided an initial assessment of pleiotropy for these loci. We also performed multivariate association analysis for these two loci, in PLINK Multivariate v1.06 (https://genepi.qimr.edu.au/staff/manuelP/multivariate/main.html; Ferreira & Purcell 2009), again separately in each dataset and using each of the reading/language traits that were used in constructing PC1. PLINK multivariate extracts the linear combination of traits that explains the largest possible amount of covariance between the SNP and all of the traits. The loading produced for each trait represent its contribution to the multivariate association. MIGFAM ‘total’ association was run, with adaptive permutations to adjust for sample relatedness (see Appendix S4 for details).

Assessment of top association signals in two additional datasets
Our two most significant association signals from PC1 meta-analysis were checked against published and unpublished results from the recent GWASMA of reading and language abilities reported by Luciano et al. (2013). This prior study analysed two population datasets, the Brisbane Adolescent Twin Sample (BATS) and the Avon Longitudinal Study of Parents and their Children (ALSPAC). BATS is a cohort of twins and their non-twin siblings recruited from ongoing studies of melanoma risk factors and cognition in an Australian population-based sample (Wight et al. 2001). Subjects had been administered psychometric tests assessing regular-, irregular-, and nonword reading, and spelling, together with the Schonell graded word reading test, and nonword repetition (see Luciano et al. 2013). ALSPAC is a longitudinal, population-based sample recruited from the county of Avon, UK (Boyd et al. 2013). The study website contains details of all the data available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary). Ethical approval was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Participants (all free of neurological/psychiatric conditions) had been tested for word reading, nonword reading, spelling and nonword repetition (see Luciano et al. 2013). BATS and ALSPAC had been generated using Illumina® 610k Quad Bead and HumanHap 550k Quad chips, respectively, and imputed using the HapMap Phase II CEU reference panel (NCBI build 36) (The International HapMap 3
Table 4: Language-/reading-related traits available in the SLIC dataset

<table>
<thead>
<tr>
<th>Trait</th>
<th>Test</th>
<th>Test description*</th>
<th>Statistical elaboration†</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRead</td>
<td>Wechsler Objectives of Reading Dimensions (WORD)¹</td>
<td>Reading single real words of increasing difficulty</td>
<td>A, S, R</td>
</tr>
<tr>
<td>WSPELL</td>
<td>WORD¹</td>
<td>Spelling of single real words</td>
<td>A</td>
</tr>
<tr>
<td>NWR</td>
<td>Gathercole &amp; Baddeley²</td>
<td>Repeating tape-recorded nonsense words of increasing length and complexity</td>
<td>A, S, R</td>
</tr>
<tr>
<td>ELS</td>
<td>Clinical Evaluation of Language Fundamentals Revised (CELF-R)³</td>
<td>Formulating sentences (formulating sentences about visual stimuli using a targeted word or phrase)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>ELS</td>
<td>CELF-R³</td>
<td>Recalling sentences (imitating sentences presented by the examiner)</td>
<td></td>
</tr>
<tr>
<td>ELS</td>
<td>CELF-R³</td>
<td>Sentence assembly (producing two semantically/grammatically correct sentences from visually and orally presented words/groups of words)</td>
<td></td>
</tr>
<tr>
<td>RLS</td>
<td>CELF-R³</td>
<td>Oral directions (pointing to pictured objects in response to oral directions)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>RLS</td>
<td>CELF-R³</td>
<td>Semantic relations (listening to a sentence and selecting the two choices that answer a target question, out of four possible answers)</td>
<td></td>
</tr>
<tr>
<td>RLS</td>
<td>CELF-R³</td>
<td>Word classes (choosing two related words and describing their relationship)</td>
<td></td>
</tr>
<tr>
<td>vIQ</td>
<td>Wechsler Intelligence Scale for Children (WISC)/WAIS⁴</td>
<td>Arithmetic (solving orally administered arithmetic word problems)</td>
<td>A</td>
</tr>
<tr>
<td>vIQ</td>
<td>Wechsler Intelligence Scale for Children (WISC)/WAIS⁴</td>
<td>Comprehension (explaining situations, actions, or activities that the examinee is expected to be familiar with)</td>
<td></td>
</tr>
<tr>
<td>vIQ</td>
<td>Wechsler Intelligence Scale for Children (WISC)/WAIS⁴</td>
<td>Digit span (reciting a sequence of digits presented by the examiner by recalling them in the same/reverse order)</td>
<td></td>
</tr>
<tr>
<td>vIQ</td>
<td>Wechsler Intelligence Scale for Children (WISC)/WAIS⁴</td>
<td>Information (general cultural knowledge test)</td>
<td></td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Similarities (explaining how two words are alike/similar)</td>
<td></td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Vocabulary (defining a provided word)</td>
<td></td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Block design (arranging blocks to duplicate a given image/design)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Coding (marking rows of shapes with different lines/transcribing symbols under digits, according to a given code)</td>
<td></td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Object assembly (correctly assembling the parts that an object is divided into, like a puzzle)</td>
<td></td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Picture arrangement (arranging a number of given pictures from left to right to tell the intended story)</td>
<td></td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC/WAIS⁴</td>
<td>Picture completion (identifying the missing part in a series of pictures representing common objects)</td>
<td></td>
</tr>
</tbody>
</table>

Superscript numbers after each test indicate the initial reference for it, where further details on the test can be found: ¹Rust et al. 1993; ²Gathercole et al. 1994; ³Semel et al. 1992; ⁴Wechsler et al. 1992.

*Where more than one battery is administered, the total score is usually computed as a sum of the raw scores from each subtest.

†Legend of statistical elaborations: A, age-adjusted; S, standardized against the normative mean of the population of study, when required (Shapiro–Wilk test \( P < 0.05 \)); R, further rank-normalized (using Blom’s formula), because the trait distribution after standardization differed from normality (Shapiro–Wilk test \( P < 0.05 \)).

Consortium, 2010. A total of 6434 subjects (962 from BATS and 5472 from ALSPAC) were meta-analysed by Luciano et al. (2013), for three different traits: word reading, nonword repetition and a composite/component score of reading and spelling (called hereafter the reading–spelling factor).

Results

GWAS meta-analysis

Table 6 describes the most significant associations from the meta-analyses on PC1 (\( N = 1862 \)) and IQ-adjusted PC1 (\( N = 1826 \)). Figure 1 shows genome-wide Manhattan Plots. QQ-plots revealed no evidence of population stratification affecting the meta-analysis results, or of genome-wide significant associations [Fig. S1a,b (Appendix S1)]. The most significant association was observed for rs59197085 in PC1 and IQ-adjusted PC1 meta-analyses (\( P = 3.86 \times 10^{-7} \) for PC1, and \( P = 3.01 \times 10^{-7} \) for IQ-adjusted PC1; A/G, MAF ∼ 8%). This SNP is located at 7q32.1, within CCDC136 (coiled-coil domain containing 136, or NAG6) and ∼10 kb upstream of FLNC [filamin C; Fig. S1c (Appendix S1)]. The second most significantly associated region, before IQ-adjustment, was

Table 5: Language-reading-related traits available in the CLDRC dataset

<table>
<thead>
<tr>
<th>Trait</th>
<th>Test</th>
<th>Test description*</th>
<th>Statistical elaboration†</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRead</td>
<td>Peabody Individual Achievement Test (PIAT)</td>
<td>Reading aloud in sequence single real words increasing in semantic and phonetic difficulty, until errors are made in five out of any seven consecutive items (untimed)</td>
<td>C, A, S, R</td>
</tr>
<tr>
<td></td>
<td>Timed oral reading2,3</td>
<td>Reading aloud a series of single real words within 2 seconds of their presentation, until errors are made in 10 out of any 20 consecutive items</td>
<td></td>
</tr>
<tr>
<td>WSpell</td>
<td>PIAT1</td>
<td>Choosing the correct spelling of a series of real words (of increasing difficulty) orally presented, among four orthographically and often phonologically similar alternatives printed on a card (for each word), until errors are made in five out of seven consecutive responses</td>
<td>A, S</td>
</tr>
<tr>
<td>PD</td>
<td>Oral Nonword Reading Task2,3</td>
<td>Reading aloud a series of two-syllable nonsense words (structure ranging from vvcv to cccvcv)</td>
<td>C, A, S, R</td>
</tr>
<tr>
<td></td>
<td>Phonological Choice (Silent Nonword Reading Task)2,3</td>
<td>Choosing which of three nonsense words would sound like a real word if read aloud (for n triplets of nonwords)</td>
<td>A, S, R</td>
</tr>
<tr>
<td>PA</td>
<td>Phoneme Segmentation and Transposition Task2</td>
<td>Taking the first phoneme of a word, putting it at the end and add the sound /ay/ (for n words, e.g. rope → ope-ray)</td>
<td>C, A, S, R</td>
</tr>
<tr>
<td></td>
<td>Phoneme Deletion Task2</td>
<td>Repeating nonwords within 2 seconds of their oral presentation, then removing a specified phoneme and pronouncing the resulting words within another 4 seconds (e.g. “say prot..now say prot without the k” “pot”)</td>
<td></td>
</tr>
<tr>
<td>OC</td>
<td>Word-Pseudohomophone Choice2,4</td>
<td>Speeded forced-choice to distinguish a real word from a phonologically similar nonword (for n pairs of words-nonsense words; e.g. rane vs. rain)</td>
<td>C, A, S, R</td>
</tr>
<tr>
<td></td>
<td>Homophone Choice2,4</td>
<td>Selecting which of two homophones visually presented answers a question asked orally by the tester (for n pairs of words, e.g. “Which is a flower?” rose rows)</td>
<td></td>
</tr>
<tr>
<td>NWR</td>
<td>Gathercole &amp; Baddeley5</td>
<td>Repeating tape-recorded nonsense words of increasing length and complexity</td>
<td>A, S, R</td>
</tr>
<tr>
<td>vIQ</td>
<td>WISC-R/WAIS-R6</td>
<td>Comprehension (explaining situations, actions, or activities that the examinee is expected to be familiar with) Information (general cultural knowledge test) Similarities (explaining how two words are alike/similar) Vocabulary (defining a provided word)</td>
<td>None</td>
</tr>
<tr>
<td>pIQ</td>
<td>WISC-R/WAIS-R6</td>
<td>Block design (arranging blocks to duplicate a given image/design) Object assembly (correctly assembling the parts that an object is divided into, like a puzzle) Picture arrangement (arranging a number of given pictures from left to right to tell the intended story) Picture completion (identifying the missing part in a series of pictures representing common objects)</td>
<td>None</td>
</tr>
</tbody>
</table>

Superscript numbers after each test indicate the initial reference for it, where further details on the test can be found: 1 Dunn & Markwardt 1970; 2 Olson et al. 1989; 3 Olson et al. 1994a; 4 Olson et al. 1994b; 5 Gathercole et al. 1994; 6 Wechsler 1974.

*Where more than one battery is administered, the total score is computed as a sum of the raw scores from each subtest (IQ measures), as an average of z-scores derived from accuracy scores (% of correct responses) and median correct reaction times of the two subtests (nonword reading), or as the arithmetic average of the raw scores from each subtest (all the other measures).

†Legend of statistical elaborations: C, composite score; A, age-adjusted (score regressed against age and age2); S, standardized against the normative mean of a control population; R, further rank-normalized (using Blom’s formula), because the trait distribution after standardization differed from normality (Shapiro–Wilk test P value < 0.05).
located on 22q12.3, SNP rs5995177 (\(P=5.01 \times 10^{-7}\), A/G, MAF \(\sim 8\%\), within RBFOX2 [also known as RNA-binding motif protein 9, or RBM9, Fig. S1d (Appendix S1)]. The association was less significant after IQ-adjustment of PC1 (\(P=1.5 \times 10^{-5}\)), and this difference was not merely due to the loss of 36 subjects in the IQ-adjusted analysis (investigated by performing a repeat PC1 analysis in the same reduced set of subjects as were available for IQ-adjusted PC1, data not shown).

Table S2a,b (Appendix S2) shows all SNPs with association \(P<1 \times 10^{-5}\) in GWAS meta-analysis of PC1 or IQ-adjusted PC1. No genome-wide significant associations were observed in the GWAS in the individual datasets (data not shown).

The results of our complementary PC1\(_{\text{read}}\) meta-analysis (Appendix S3) were consistent with the PC1 meta-analysis, with rs59197085 and rs5995177 among the top suggestive associations (\(P \sim 10^{-6}\)). This was expected given the high correlations between PC1 and PC1\(_{\text{read}}\) in each dataset (all correlations \(>0.9\), see above).

### Effect sizes and profiles of top associations

rs599517085 (CCDC136/FLNC) explained 3% of PC1 variance and 3.2% of IQ-adjusted PC1 variance in our largest GWAS dataset (UK-RD), and 1.3% of PC1 variance and 1.5% of IQ-adjusted PC1 variance in the next largest dataset (CLDRC-RD). The estimated effect sizes in the smaller datasets were \(\lesssim 0.2\%\). Estimated effect sizes for rs5995177 (RBFOX2) were more consistent across datasets. This SNP explained 1.2% of PC1 and IQ-adjusted PC1 variance in UK-RD, and 1.8% of PC1 variance and 1.2% of IQ-adjusted PC1 variance in CLDRC-RD, while estimated effect sizes in the smaller datasets were between 0.6% and 1.6% of variance.

Both rs59197085 and rs5995177 showed broad profiles of association across the measures that were used to construct PC1, as assessed from the PLINK multivariate loadings and corresponding OFAM univariate association \(P\) values shown in Table 7. These findings suggest pleiotropic effects of the two SNPs on reading and language.

### Gene-based meta-analysis

The strongest gene-based associations inferred from the PC1 and IQ-adjusted PC1 meta-analyses are reported in Table S2c,d (Appendix S2). While no gene exceeded the appropriate genome-wide significance threshold for this analysis (\(P<2.8 \times 10^{-6}\), CCDC136, FLNC and RBFOX2 were among the most significantly associated genes, with the latter approaching the significance threshold in the PC1 analysis (\(P=5 \times 10^{-6}\)). However, after conditioning on the most significant association signal within each gene, no other SNP within each of these genes showed significant evidence for having an independent residual effect, after correction for multiple testing [Table S2e,f (Appendix S2)]. For this analysis the gene boundaries were defined in the same way as for gene-based analysis (see above).

### Pathway-based meta-analysis

We assessed evidence for an excess of association signals within the genes of three neurobiological pathways that are prominent in prior literature on reading and language: axon guidance, neuronal migration and steroid sex hormone biology (see Discussion for the relevant citations). None of the three tested sets showed significant associations with PC1 or IQ-adjusted PC1 [Table S2g,h (Appendix S2)], although the association between PC1 and the steroid-related pathway approached significance (\(P=0.051\)).

### Assessment of top associations within previous GWAS results

We assessed our most significant associations from PC1 meta-analyses within published and unpublished results from the previous GWAS study of the BATS/ALSPAC datasets, for which the reading and language measures were IQ-adjusted (Luciano et al. 2013). FLNC and CCDC136 showed nominally significant associations in gene-based (VEGAS) analyses of reading-related traits in BATS/ALSPAC (CCDC136 \(P=0.034\) for reading-spelling factor and \(P=0.003\) for word-reading; FLNC \(P=0.009\) for word-reading; see Table S3 of Luciano et al. 2013). The reading–spelling factor in the BATS/ALSPAC datasets was the most comparable trait to the IQ-adjusted PC1 score of this study. As the study of Luciano et al. 2013 had used the HapMap2 reference dataset for genotype imputation, it was not possible to directly investigate the most highly associated SNPs from this study in the BATS/ALSPAC datasets. We therefore investigated association for two HapMap2 SNPs that were closest to our top hits on 7q32 and 22q12.3. rs3734972 (PC1 \(P=5.66 \times 10^{-7}\), IQ-adjusted PC1 \(P=4.68 \times 10^{-7}\); T/C, minor allele T, MAF \(\approx 8\%\)) lies \(\sim 10\) kb away from rs59197085 on 7q32 and is in high LD with it (\(R^2=0.89\), see local association plot, Fig. S1c (Appendix S1)). rs3734972 showed a \(P\) value of 0.032 with the IQ-adjusted reading-spelling factor in BATS/ALSPAC. The allelic trend was in the opposite direction to that observed in the UK-RD/SLIC/CLDRC datasets, with the T allele having a positive effect on the trait score in the BATS/ALSPAC cohorts. rs12158665 (PC1 \(P=7.57 \times 10^{-7}\), IQ-adjusted PC1 \(P=4.65 \times 10^{-5}\); C/G, minor allele G, MAF \(\approx 13\%\)) was the second most significant association in 22q12.3, mapping \(\sim 7\) kb from the top SNP at this locus rs5995177, and in low LD with it (\(R^2=0.083\)), as are all the other suggestively associated SNPs in 22q12.3 [see local association plot, Fig. S1d (Appendix S1)]. rs12158665 showed no evidence of association in BATS/ALSPAC (\(P=0.81\)).

### Discussion

This study aimed to identify pleiotropic variants having effects on reading and language abilities by analysing continuous traits in multiple datasets. Our study is complementary to two recently published GWAS: one using a similar approach but in general population samples (Luciano et al. 2013), and another contrasting a relatively small number of categorically defined RD-SLI comorbid cases with unaffected controls (Eicher et al. 2013).
GWAS of reading and language

Table 6: Top association signals ($P < 1 \times 10^{-6}$) in the PC1 and IQ-adjusted PC1 meta-analyses

<table>
<thead>
<tr>
<th>Chr</th>
<th>SNP*</th>
<th>Position (hg19)</th>
<th>Allele1</th>
<th>Allele2</th>
<th>Freq Allele1 (%)</th>
<th>P value</th>
<th>Direction†</th>
<th>Gene (distance)‡</th>
<th>Variant type</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>7</td>
<td>rs59197085</td>
<td>128460756</td>
<td>a</td>
<td>g</td>
<td>7.971</td>
<td>3.86 × 10^{-7}</td>
<td>––––</td>
<td>FLNC(–9.726)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>rs58845495</td>
<td>128462847</td>
<td>t</td>
<td>c</td>
<td>92.029</td>
<td>4.09 × 10^{-7}</td>
<td>+++++</td>
<td>FLNC(–7.635)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7:128439695:i</td>
<td>128439695</td>
<td>i</td>
<td>r</td>
<td>7.94</td>
<td>4.99 × 10^{-7}</td>
<td>––––</td>
<td>CCDC136(0)</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>rs5995177</td>
<td>36309553</td>
<td>a</td>
<td>g</td>
<td>8.049</td>
<td>5.01 × 10^{-7}</td>
<td>––––</td>
<td>RBFOX2(0)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>rs3734972</td>
<td>128470838</td>
<td>t</td>
<td>c</td>
<td>7.983</td>
<td>5.66 × 10^{-7}</td>
<td>+++++</td>
<td>FLNC(0)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>rs3800560</td>
<td>128461094</td>
<td>t</td>
<td>c</td>
<td>7.971</td>
<td>6.25 × 10^{-7}</td>
<td>––––</td>
<td>FLNC(–9.388)</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>rs12158565</td>
<td>36316843</td>
<td>c</td>
<td>g</td>
<td>87.23</td>
<td>757 × 10^{-7}</td>
<td>+++++</td>
<td>RBFOX2(0)</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>rs5755979</td>
<td>36290707</td>
<td>t</td>
<td>c</td>
<td>12.77</td>
<td>9.05 × 10^{-7}</td>
<td>––––</td>
<td>RBFOX2(0)</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>rs5790202</td>
<td>36339542</td>
<td>t</td>
<td>c</td>
<td>12.77</td>
<td>9.06 × 10^{-7}</td>
<td>––––</td>
<td>RBFOX2(0)</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>rs5750203</td>
<td>36339998</td>
<td>a</td>
<td>t</td>
<td>87.23</td>
<td>9.72 × 10^{-7}</td>
<td>+++++</td>
<td>RBFOX2(0)</td>
</tr>
<tr>
<td>IQ-adjusted PC1</td>
<td>7</td>
<td>rs59197085</td>
<td>128460756</td>
<td>a</td>
<td>g</td>
<td>7.971</td>
<td>3.01 × 10^{-7}</td>
<td>––––</td>
<td>FLNC(–9.726)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>rs58845495</td>
<td>128462847</td>
<td>t</td>
<td>c</td>
<td>92.029</td>
<td>3.23 × 10^{-7}</td>
<td>+++++</td>
<td>FLNC(–7.635)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7:128439695:i</td>
<td>128439695</td>
<td>i</td>
<td>r</td>
<td>7.94</td>
<td>3.95 × 10^{-7}</td>
<td>––––</td>
<td>CCDC136(0)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>rs3800560</td>
<td>128461094</td>
<td>t</td>
<td>c</td>
<td>7.971</td>
<td>4.48 × 10^{-7}</td>
<td>––––</td>
<td>CCDC136(0)</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>rs3734972</td>
<td>128470838</td>
<td>t</td>
<td>c</td>
<td>7.983</td>
<td>4.68 × 10^{-7}</td>
<td>––––</td>
<td>FLNC(0)</td>
</tr>
</tbody>
</table>

*Single-base indels were not filtered out from the imputed polymorphisms since they were reliably called in the imputation reference (1000 Genomes, Phase 1 v3), and were tested for association as they could represent coding frameshift variants of biological interest.
†The direction of effect of Allele1 is reported for datasets in the following order: CLDRC-RD, UK-RD, SLIC, CLDRC-ADHD.
‡Physical distance (kb) from closest genes (in a ±10 kb range from each marker) is indicated, along with orientation based on the direction of transcription (‘-’, upstream of 5′-UTR; ‘+’, downstream of 3′-UTR).

Our study is novel and distinct for several reasons:

First, we analysed continuous variation in reading and language skills while also having an enrichment of participants with low abilities (i.e. through analysing poor performing probands together with their siblings), and without applying a dichotomous classification into cases and controls that necessarily involves arbitrary thresholding. Our design was therefore suited to detect genetic effects on susceptibility to RD and SLI that also act across the entire distribution of reading and language skills.

Second, we specifically focused on shared neurobiological mechanisms underlying language and reading, by analysing the first principal component of all of the reading- and language-related measures available in each dataset, followed by investigating the cross-phenotypic effects of the resulting top GWAS hits through univariate association analysis using each individual measure. We additionally followed this with a confirmatory analysis focused only on word reading and spelling, because these measures provided the closest matching possibility across our datasets. The first principal component (PC1) of all available measures extracted a large proportion of shared trait variance across the two domains of reading and language, and was highly correlated with the component derived from only reading and spelling (PC1\text{read}).

Third, we performed GWAS both before and after IQ-adjustment of PC1. This was done in order to identify both genetic variants having effects broadly across reading, language and general cognitive abilities, and variants having effects on reading and language but independently of general cognitive ability. This approach also facilitated a comparison of our top results with those from datasets investigated in Luciano et al. (2013).

We checked within our GWASMA results 18 specific SNPs that had been highlighted to show the most promising candidate associations by the authors of previous GWAS studies of reading and/or language (Eicher et al. 2013; Field et al. 2013; Luciano et al. 2013; Meaburn et al. 2008; Roeske et al. 2011). Seventeen of these SNPs showed no nominally significant association within our GWASMA results (data not shown). Only rs10485609 (Meaburn et al. 2008) showed a nominally significant association ($P = 0.013$ for PC1, $P = 0.015$ for IQ-adjusted PC1; allele A was associated with lower performance, which was a consistent allelic direction of effect with that reported by Meaburn et al. 2008), but this was not significant after multiple testing correction for 18 tests.

Like the other recently published GWAS efforts in this field, our study did not find any individual associations that achieved genome-wide significance (threshold $P = 5 \times 10^{-8}$). However, we did identify two novel, suggestive results of particular interest, on 7q32.1 and 22q12.3, with the most significant associations at rs59197085 and rs5995177, respectively. As shown in Table 7, both SNPs displayed a broad pattern of association across multiple reading and language traits, consistent with effects on neurobiological processes shared
Figure 1: Manhattan plots of the meta-analyses of the first principal component scores. (a) PC1; (b) IQ-adjusted PC1. The blue line represents the nominal suggestive significance threshold ($P = 1 \times 10^{-5}$).

between reading and language cognition. In the regression model these SNPs explained a notable proportion (up to 3.2%) of variance in PC1 and IQ-adjusted PC1 scores, particularly in the largest datasets (CLDRC-RD and UK-RD), although these effect sizes are likely to be overestimated since this is the first report of these associations (Ioannidis 2008). Gene based-tests were consistent with the results of the SNP-based analysis for FLNC, CCDC136 and RBFOX2, and the gene-based $P$ values were found to be largely or wholly reflective of the individual top associations within each of these genes.

rs5995177 is an intronic variant localized within RBFOX2 (RNA-binding protein, fox-1 homologue 2, also known as RBM9), a protein that regulates alternative splicing and is active in neurons. RBFOX2 is highly expressed in the foetal brain and has important roles in CNS development (Gehman et al. 2012). The homologous gene RBFOX1 has been implicated in several neurodevelopmental disorders, including Rolandic Epilepsy (Lal et al. 2013) and Autism Spectrum Disorder (Voineagu et al. 2011), and is a downstream target of FOXP2, a transcription factor implicated in monogenic speech and language disorders (Ayub et al. 2013). The high comorbidity between Rolandic Epilepsy and RD (Clarke et al. 2007) and the presence of a FOXP2 binding site ∼5 kb from rs5995177 (The ENCODE Project Consortium, 2012), further support a link of RBFOX2 with reading and language abilities. Thus convergent evidence from multiple lines of research makes RBFOX2 an intriguing candidate gene for future studies. There was no evidence of association of this locus with reading and language measures in the results of the population-based study of Luciano et al. (2013).

rs59197085 is located in CCDC136 (coiled-coil domain containing 136, or NAG6) and ∼10 kb upstream of FLNC
GWAS of reading and language

Table 7: Effect of the top association signals rs59197085 (7q32.1) and rs5995177 (22q12.3) on the single reading and language traits used in constructing PC1

<table>
<thead>
<tr>
<th>Trait</th>
<th>CLDRC-RD</th>
<th>UK-RD</th>
<th>SLIC</th>
<th>CLDRC-ADHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs59197085</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRead</td>
<td>-0.66 (0.024)</td>
<td>-0.87 (5.3 x 10^{-5})</td>
<td>-0.29 (0.626)</td>
<td>-0.5 (0.427)</td>
</tr>
<tr>
<td>WSpell</td>
<td>-0.89 (3.8 x 10^{-3})</td>
<td>-0.75 (1.1 x 10^{-2})</td>
<td>0.08 (0.862)</td>
<td>-0.1 (0.871)</td>
</tr>
<tr>
<td>PD</td>
<td>-0.76 (7.9 x 10^{-3})</td>
<td>-0.86 (1.6 x 10^{-1})</td>
<td>0.37 (0.549)</td>
<td>0.13 (0.854)*</td>
</tr>
<tr>
<td>PA</td>
<td>-0.65 (0.029)</td>
<td>-0.49 (0.018)</td>
<td>0.35 (0.588)</td>
<td>0.04 (0.95)</td>
</tr>
<tr>
<td>OC</td>
<td>-0.64 (0.036)</td>
<td>-0.89 (3 x 10^{-6})</td>
<td>-0.57 (0.32)</td>
<td>-0.28 (0.686)</td>
</tr>
<tr>
<td>NWR</td>
<td>-0.34 (0.269)</td>
<td>-0.57 (0.32)</td>
<td>0.08 (0.821)</td>
<td></td>
</tr>
<tr>
<td>ELS</td>
<td>-0.25 (0.807)</td>
<td>-0.25 (0.807)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RLS</td>
<td>0.08 (0.821)</td>
<td>0.08 (0.821)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs5995177</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRead</td>
<td>-0.66 (0.027)</td>
<td>-0.81 (2 x 10^{-3})</td>
<td>-0.71 (0.116)</td>
<td>0.01 (0.98)</td>
</tr>
<tr>
<td>WSpell</td>
<td>-0.81 (6.9 x 10^{-3})</td>
<td>-0.77 (1.8 x 10^{-1})</td>
<td>-0.52 (0.262)</td>
<td>-0.33 (0.359)</td>
</tr>
<tr>
<td>PD</td>
<td>-0.65 (0.026)</td>
<td>-0.79 (8.9 x 10^{-3})</td>
<td>-0.46 (0.158)</td>
<td>-0.37 (0.26)*</td>
</tr>
<tr>
<td>PA</td>
<td>-0.72 (0.023)</td>
<td>-0.72 (2.5 x 10^{-3})</td>
<td>-0.65 (0.046)</td>
<td>-0.02 (0.968)</td>
</tr>
<tr>
<td>OC</td>
<td>-0.68 (0.026)</td>
<td>-0.57 (0.017)</td>
<td>-0.23 (0.676)</td>
<td>0.06 (0.976)</td>
</tr>
<tr>
<td>NWR</td>
<td>-0.04 (0.922)</td>
<td>-0.23 (0.676)</td>
<td>-0.82 (0.057)</td>
<td></td>
</tr>
<tr>
<td>ELS</td>
<td>-0.61 (0.206)</td>
<td>-0.61 (0.206)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RLS</td>
<td>0.08 (0.821)</td>
<td>0.08 (0.821)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These were computed for each trait as PLINK Multivariate MOFAM loadings and PLINK univariate QFAM association P values (in brackets) and refer to the minor alleles (A for both SNPs).

WRead, word reading; WSpell, word spelling; PD, phonological decoding; PA, phoneme awareness; OC, orthographic coding; NWR, nonword repetition; ELS/RLS, expressive/receptive language score.

*Loading on nonword reading and phonological choice (respectively).
†Although PA had been excluded from the PCA in UK-RD (due to the low number of measures available), it was tested in this case to have a term of comparison to the other datasets.

(filamin C). This SNP, along with the nearby SNPs rs3800560, rs58845495 and rs3734972, forms roughly 10-kb haplotypes spanning the region between CCDC136 and FLNC and partially overlapping these genes [see local association plot, Fig. S1c (Appendix S1)]. CCDC136 encodes a poorly characterized tumour suppressor which has been found to be downregulated in gastric carcinoma (Zhang et al. 2004) and is highly expressed in the cerebellum and in the occipital cortex (Allen Human Brain Atlas, Hawrylycz et al. 2012; http://human.brain-map.org). Filamin C (or filamin gamma) is a structural protein that crosslinks actin filaments into orthogonal networks in the cortical cytoplasm and participates in cytoskeleton re-modelling, suggesting a possible role in cell motility and migration. Functions of FLNC have been demonstrated in muscle tissues, where mutations have been associated with several forms of myopathies (Duff et al. 2011). However, its pattern of expression includes spinal cord, cerebellum, corpus callosum, basal ganglia and some localized areas in the frontal, temporal and occipital cortex (Allen Human Brain Atlas, Hawrylycz et al. 2012). Its homologue FLNA (filamin A) is involved in neuronal migration and is implicated in an X-linked dominant form of periventricular heterotopia, a neurological disorder that sometimes involves reading and spelling problems (Robertson 2005).

Associations within the 7q32 region are particularly interesting in light of data from two previous independent studies that have each reported evidence for linkage between a microsatellite marker in this region (D7S530, located ∼650 kb from our peaks of association) and RD status (Kaminen et al. 2003) or else nonword spelling and irregular word reading (Bates et al. 2007). There was also evidence of association, at the gene level, with reading and language measures for FLNC, and CCDC136 in the BATS/ALSPAC datasets studied by Luciano et al. (2013). At the SNP level, one of our most significantly associated SNPs from GWASMA, rs3734972 also showed association with an IQ-adjusted reading–spelling score in the BATS/ALSPAC datasets. However, the allelic directions of effect on the traits in this study and the study by Luciano et al. were opposite.

We sought to detect an excess of association signals within genes belonging to each of three candidate gene sets based on different biological functions: axon guidance, neuronal migration and steroid hormone biology. Axon guidance and neuronal migration are functions linked to some of the previously identified candidate genes in RD and SLI; ROBO1 (Hannula-Jouppi et al. 2005), DCDC2 (Meng et al. 2005), KIAA0319 (Peschansky et al. 2010), DYZ1C1 (Tammimies et al. 2013) and FOXP2 (Vernes et al. 2011). A potential involvement of neuronal migration deficits in RD aetiology represents a longstanding hypothesis of the field (see Galaburda & Cestnick 2003). The steroid hypothesis was motivated by literature suggesting links between sex...
hormone biology, language performance and the brain architecture that subserves reading and language (Good et al. 2001; Lombardo et al. 2012; Shapleske et al. 1999; Whitehouse et al. 2012); and by evidence of interaction between Oestrogen Receptors and DXY1C1, both at the gene (Tammimies et al. 2012) and at the protein level (Massinen et al. 2009). None of the three gene sets showed a significant excess of association signals, although the steroid hormone biology set approached significance in this analysis.

In carrying out GWASMA studies of complex cognitive traits across multiple datasets collected by different research teams, an obvious limitation is that the specific trait measurements that are available may be quite diverse. Even when tests are similar, and hypothesized to measure corresponding cognitive processes, they may still create a substantial source of heterogeneity for a meta-analysis effort. In this study we sought to overcome this limitation by focusing on a principal component (PC1) capturing a majority of the shared variance between reading- and language-related traits. In spite of the phenotypic heterogeneity of our datasets, this measure can be considered comparable across datasets for a number of reasons. First, the loadings of the individual traits on PC1 scores were similar across the datasets. Second, dropping one or more traits from our PC1 computation did not substantially affect the resulting PC1 scores (data not shown). Third, the First Principal Component derived only from word reading and spelling (PC1\textsubscript{read}) was strongly correlated with PC1. Word reading and spelling were the only two measures available in all of the datasets and provided the closest phenotype matching possible across datasets. Not surprisingly, given the high correlations between PC1 and PC1\textsubscript{read} in all datasets, the association meta-analysis using PC1\textsubscript{read} (Appendix S3) produced results consistent with PC1-based meta-analysis. We therefore conclude that PC1 was a sufficiently well matched construct across datasets to support GWASMA, in which we nonetheless allowed for heterogeneity of effect sizes across datasets to avoid assuming a perfect matching. It is interesting that a single PC can capture comparable variation across a diverse range of reading and language traits and in the presence of heterogeneity of measurement across datasets. This indicates a robust unifying dimension to much of this variation, and supports a genetic approach framed around pleiotropy.

The use of a principal component can lead to some loss of information, both in terms of detecting trait-specific genetic effects, and of reducing the sample size (because individuals with one or more missing trait values were excluded from the analysis). However, as we aimed to identify shared genetic effects on reading and language, the use of PC1 scores, followed by investigating cross-phenotypic associations of the top SNPs at the level of individual traits, was an appropriate approach to analysing these multivariate datasets. There is now a need for a larger international meta-analysis effort that incorporates further datasets. This would improve the power to detect pleiotropic variants affecting reading and language.

References


GWAS of reading and language


GWAS of reading and language

Acknowledgments

The authors declare no conflicts of interest. This work was supported by the Max Planck Society, the University of St Andrews, the EU (Neurodyn, 018696), and the US National Institutes of Health (Grant ref: P50 HD027802). Genotyping at the Wellcome Trust Centre for Human Genetics was supported by the Wellcome Trust (090532/Z/09/Z) and a Medical Research Council Hub Grant (G0900747 91070). Silvia Paracchini is a Royal Society University Research Fellow. Dianne Newbury is an MRC Career Development Fellow and a Junior Research Fellow at St John’s College, University of Oxford. We thank Dr Manuel Ferreira and Dr Phil H Lee for useful advice regarding the use of PLINK Multivariate and INRICH tools. Dr Margaret Wright and Prof Nicholas Martin are the principal investigators for the BATS IQ data collection and genotyping.

Members of the SLI Consortium: Wellcome Trust Centre for Human Genetics, Oxford: D. F. Newbury, N. H. Simpson, R. Nudel, A. P. Monaco; Max Planck Institute for Psycholinguistics, Nijmegen: S. E. Fisher, C. Franks; Newcomen Centre, Guy’s Hospital, London: G. Baird, V. Stolmins, K Dworzynski; Child and Adolescent Psychiatry Department and Medical Research Council Centre for Social, Developmental, and Genetic Psychiatry, Institute of Psychiatry, London: P. F. Bolton; Medical Research Council Centre for Social, Developmental, and Genetic Psychiatry Institute of Psychiatry, London: E. Simonoff; Department of Reproductive and Developmental Sciences, University of Edinburgh: A. O’Hare; Molecular Medicine Centre, University of Edinburgh: J. Seckl; Department of Speech and Language Therapy, Royal Hospital for Sick Children, Edinburgh: H. Cowie; Speech and Hearing Sciences, Queen Margaret University College: A. Clark and J. Watson; Department of Educational and Professional Studies, University of Strathclyde: W. Cohen; Department of Child Health, the University of Aberdeen: A. Everitt, E. R. Hennessy, D. Shaw, P. J. Helms; Audiology and Deafness, School of Psychological Sciences, University of Manchester: Z. Simkin, G. Conti-Ramsden; Department of Experimental Psychology, University of Oxford: D. V. M. Bishop; Biostatistics Department, Institute of Psychiatry, London: A. Pickles.

ALSPAC: We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The project of ALSPAC is made possible by funding from the UK Medical Research Council and the Wellcome Trust (Grant ref: 092731) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and DME will serve as guarantor for the contents of this paper that are related to ALSPAC.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s web-site:

Appendix S1: QQ plots and association plots of the top association signals from analysis based on PC1 and IQ-adjusted PC1. Contribution of each dataset to the strength of the association in the PC1 and IQ-adjusted PC1 meta-analysis, for the top association signals.

Appendix S2: SNP-, gene- and pathway-based meta-analysis results from analysis based on PC1 and IQ-adjusted PC1. Test for independent associations within the top associated regions, after conditioning on the most significant local association signal.

Appendix S3: Analyses based on PC1 and IQ-adjusted PC1 (trait loadings on PC1, Manhattan plots, QQ plots and top associations)

Appendix S4: Genotype calls and QC protocols; correlation patterns of reading and language measures; statistical analyses, commands and parameters.