

# Evaluation of results from genome-wide studies of language and reading in a novel independent dataset

A. Carrion-Castillo<sup>†</sup>, E. van Bergen<sup>‡,§</sup>, A. VINO<sup>†</sup>,  
T. van Zuijlen<sup>¶</sup>, P. F. de Jong<sup>¶</sup>, C. Francks<sup>†,\*\*</sup> and  
S. E. Fisher<sup>†,\*\*,\*</sup>

Received 21 March 2016, revised 1 May 2016, 17 May 2016  
accepted for publication 17 May 2016

<sup>†</sup>Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, <sup>‡</sup>Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, <sup>§</sup>Department of Experimental Psychology, University of Oxford, Oxford, UK, <sup>¶</sup>Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, and <sup>\*\*</sup>Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands  
\*Corresponding author: Prof. S. E. Fisher, Language and Genetics Department, Max Planck Institute for Psycholinguistics, PO BOX 310, 6500 AH Nijmegen, the Netherlands. E-mail: simon.fisher@mpi.nl

**Recent genome-wide association scans (GWAS) for reading and language abilities have pin-pointed promising new candidate loci. However, the potential contributions of these loci remain to be validated. In this study, we tested 17 of the most significantly associated single nucleotide polymorphisms (SNPs) from these GWAS studies ( $P < 10^{-6}$  in the original studies) in a new independent population dataset from the Netherlands: known as Familial Influences on Literacy Abilities. This dataset comprised 483 children from 307 nuclear families and 505 adults (including parents of participating children), and provided adequate statistical power to detect the effects that were previously reported. The following measures of reading and language performance were collected: word reading fluency, nonword reading fluency, phonological awareness and rapid automatized naming. Two SNPs (rs12636438 and rs7187223) were associated with performance in multivariate and univariate testing, but these did not remain significant after correction for multiple testing. Another SNP (rs482700) was only nominally associated in the multivariate test. For the rest of the SNPs, we did not find supportive evidence of association. The findings may reflect differences between our study and the previous investigations with respect to the language of testing, the exact tests used and the recruitment criteria. Alternatively, most of the prior reported associations may have been false positives. A larger scale GWAS meta-analysis than those previously performed will likely be required to obtain robust insights into the genomic architecture underlying reading and language.**

Keywords: Association study, candidate SNPs, language, reading

It is well known that reading abilities have a genetic component, with reported heritability estimates ranging from 30% for a discriminant score of reading ability (DeFries *et al.* 1987) to 73% for word reading and 49% for reading comprehension in a recent meta-analysis (de Zeeuw *et al.* 2015). Until the last 2–3 years, most research on the genetics of reading ability and disability (developmental dyslexia) was focused on candidate genes (e.g. *ROBO1*, *KIAA0319*, *DCDC2* and *DYX1C1*) that were often identified through linkage analysis followed by fine-mapping association studies. Several candidate associations were reported, and some of these have shown further supportive evidence in independent samples (reviewed in Carrion-Castillo *et al.* 2013). However, there is also a considerable lack of consistency across studies (Becker *et al.* 2014; Tran *et al.* 2013; Zhong *et al.* 2013; Zou *et al.* 2012), which has hindered efforts to precisely define the role of any specific variant in affecting the neurobiological basis of reading (Carrion-Castillo *et al.* 2013).

More recently, several genome-wide association scan (GWAS) studies have tried to identify common genetic variants that influence language and reading abilities, without prior hypotheses with regard to specific candidate genes or regions of the genome (Eicher *et al.* 2013; Field *et al.* 2013; Gialluisi *et al.* 2014; Harlaar *et al.* 2014; Luciano *et al.* 2013; Nudel *et al.* 2014). This new wave of research for the field queries the whole genome for association in a relatively unbiased manner, which is appropriate for phenotypes when the vast majority of the underlying genetic architecture is unknown. An important consideration for this approach is that, ideally, dataset sizes must be in the order of thousands of participants or more, in order to detect the small effect sizes that are expected for individual polymorphisms, and in the context of a high degree of statistical correction for multiple testing over millions of genetic variants (Visscher *et al.* 2012).

The GWAS studies of reading and language performed so far have been based on a range of different designs and approaches. They have included cohorts ascertained through disorder [e.g. dyslexia or specific language impairment (SLI)] or sampled from the general population. Such cohorts have been tested for association of single nucleotide polymorphisms (SNPs) with either a categorically defined affection status or else with quantitative measures of performance for an array of reading/language-related skills. Here, we briefly summarize findings of the relevant GWAS reports published prior to October 2014, upon which we based the present study (see also Table 1).

**Table 1:** Summary of genome-wide association studies of language and reading traits published until October 2014, which provided the basis for selection of SNPs in this study. Studies are ordered by study type. Age range is specified per sample if it differs between the datasets included in each study

Reference	Design	Samples	Trait	Age (years)	Study type	Total <sub>max</sub>	Cases <sub>max</sub>
Field <i>et al.</i> (2013)	GWAS	Dys-Canada	Phonological coding dyslexia	8–18	Case–control	718	400
Eicher <i>et al.</i> (2013)	GWAS	ALSPAC	Dyslexia	7–9	Case–control	4291	353
Eicher <i>et al.</i> (2013)	GWAS	ALSPAC	SLI	7–9	Case–control	4291	163
Eicher <i>et al.</i> (2013)	GWAS	ALSPAC	Dyslexia/SLI	7–9	Case–control	4291	174
Nudel <i>et al.</i> (2014)	GWAS: basic, maternal, paternal	SLIC	SLI	5–19	Case–control	—	297
Gialluisi <i>et al.</i> (2014)	GWASMA	SLIC; UK-RD; CLDRC	Principal component of reading/language	5–19; 5–31; 8–19	Quantitative	1862	—
Luciano <i>et al.</i> (2013)	GWASMA	ALSPAC; BATS	Word reading	8–9; 12–25	Quantitative	6189	—
Luciano <i>et al.</i> (2013)	GWASMA	ALSPAC; BATS	Nonword reading	8–9; 12–25	Quantitative	6182	—
Luciano <i>et al.</i> (2013)	GWASMA	ALSPAC; BATS	Reading and spelling measure	8–9; 12–25	Quantitative	6182	—
Luciano <i>et al.</i> (2013)	GWASMA	ALSPAC; BATS	Nonword repetition	8–9; 12–25	Quantitative	6583	—
Harlaar <i>et al.</i> (2014)	GWAS	TEDS	Receptive language composite score	12	Quantitative	2329	—
St Pourcain <i>et al.</i> (2014)	GWASMA	ALSPAC; GenR	Expressive vocabulary CDI	1.25–1.5	Quantitative	8889	—
St Pourcain <i>et al.</i> (2014)	GWASMA	ALSPAC; GenR; Raine; TEDS	Expressive vocabulary CDI	2–2.5	Quantitative	10 819	—

ALSPAC, Avon Longitudinal Study of Parents and their Children; BATS, Brisbane Adolescent Twin Sample; CDI, Communicative Development Inventory; CLDRC, Colorado Learning Disability Research Center; GenR, Dutch Generation R; GWASMA, GWAS meta-analysis; Raine, Western Australian Pregnancy Cohort; SLIC, SLI Consortium; TEDS, UK Twins Early Development Study; UK-RD, UK-reading disability.

Field *et al.* (2013) tested for association with a trait defined as ‘phonological coding dyslexia’ in a family-based sample ( $n=718$ , with 400 cases from 101 families), using the transmission disequilibrium test to screen ~133 000 markers from the genome. An SNP within 5q35.1 (77 kb downstream of *FGF18*) was associated with dyslexia status at a borderline level of significance when considered against genome-wide multiple testing thresholds.

Eicher *et al.* (2013) performed case–control genome-wide association analyses with three different affection statuses for reading and language disorders defined from a general population sample of 4291 children, the ALSPAC (Avon Longitudinal Study of Parents and their Children) cohort. Available quantitative traits were first used to define cases of reading disability ( $n=353$ ), language impairment ( $n=163$ ) as well as comorbid cases showing both reading and language deficits ( $n=174$ ). Case–control analyses were then performed to evaluate associations of variants across the genome. Suggestive associations were found at several loci (3p24.4, 4q26 and within the *COL4A2* gene on 13q34), some of which could be tentatively linked to variation in brain white matter tracts in follow-up analysis of diffusion tensor imaging (DTI) data in a separate sample of 332 healthy participants.

In another case–control GWAS study, Nudel *et al.* (2014) investigated a family-based sample recruited on the basis of probands with an SLI diagnosis; the SLI Consortium (SLIC) cohort, including 297 cases from 278 nuclear families. Also,

using standard association analyses, the study tested for parent-of-origin effects and found two paternal effects: a significant association for a locus on 7p14.1 and a suggestive association within the *NOP9* gene on 14q11.2. For maternal parent-of-origin analyses, they identified a suggestive association in the 5p13.1 region.

Moving away from case–control designs, Gialluisi *et al.* (2014) tested for pleiotropic effects of common genetic variants by carrying out GWAS meta-analysis of quantitative data from three different cohorts enriched for participants with reading/language disorders: the SLIC families, a UK-reading disability (UK-RD) dataset and the Colorado Learning Disability Research Center (CLDRC) dataset (total  $n=1862$ ). The available quantitative traits variously included spelling, word reading, nonword reading, nonword repetition, phonological awareness (PA), expressive and receptive language scores, and these were used to derive first principal components in each dataset. The GWAS meta-analysis identified two loci showing suggestive associations with the principal component, one within 7q32.1 (~10 kb upstream of *FLNC*) and the other on 22q12.3 within the *RBFOX2* gene.

Several GWAS studies have focused on normal variation in reading and language in epidemiological cohorts. Luciano *et al.* (2013) meta-analysed GWAS results based on various quantitative traits in two general population cohorts, ALSPAC (maximum  $n=5472$ ) and the Brisbane Adolescent Twin Sample ( $n=1177$  from 538 families). They reported suggestive

associations with a compound score of reading and spelling (19p13.3 within *DAZAP1*), word reading (1p13.1, 16q22.2 and 19p13.3) and nonword repetition (16q23.2 and 21q11.2 within *ABCC13*). A subsequent study tested for association with a composite score of receptive language measures at 12 years in the UK Twins Early Development Study (TEDS) (Harlaar *et al.* 2014). Suggestive associations were identified within 2q31.2 in a primary discovery sample ( $n=2329$ ), but did not show any evidence of association in replication samples (total  $n=2639$ ).

Finally, St Pourcain *et al.* (2014) focused on expressive vocabulary at an early age (15–18 months,  $n=8889$ ) and later age (24–30 months,  $n=10819$ ) in four general population datasets: ALSPAC, TEDS, the Dutch Generation R cohort and Western Australian Pregnancy Cohort (Raine) Study. One SNP within 3p12.3 (~19 kb downstream of the *ROBO2* gene) was significantly associated with early expressive vocabulary in GWAS meta-analysis. This SNP is roughly 1 megabase from the dyslexia susceptibility locus *DYX5* (where the dyslexia candidate gene *ROBO1* lies; which is also a paralogue of *ROBO2*). Other SNPs were reported as suggestively associated in chromosomal bands 11p15.2, 12q15 and 19p13.3.

The sample sizes used for most of these studies of reading and language abilities are below what is optimal for GWAS of highly polygenic human traits, with the possible exception of two of the reports (Luciano *et al.* 2013; St Pourcain *et al.* 2014). Most of the reported top findings are statistically 'suggestive', and all remain to be validated in other independent samples, although they provide promising new candidate genes that could potentially play roles in the neurobiology of language and reading.

For this study, we have tested the most significantly associated SNPs from the above GWAS studies, targeting those that have shown association  $P < 10^{-6}$  in the original reports. Through direct genotyping of the key SNPs, we attempted to find supportive evidence in a new Dutch general population sample in which there are several reading-related quantitative traits available, known as the FIOLA (Familial Influences on Literacy Abilities) dataset (Fig. 1). While not all of the GWAS studies used datasets that are independent of one another, all of the top associations were study specific. The GWAS studies encompassed a broad range of phenotypic traits, including some measures, such as expressive vocabulary (St Pourcain *et al.* 2014) and receptive language ability (Harlaar *et al.* 2014), which are not available in FIOLA. However, there is abundant evidence that reading and language performance measures have partly overlapping genetic contributions (Gialluisi *et al.* 2014; Luciano *et al.* 2013; Newbury *et al.* 2011). Several studies show evidence supporting the hypothesis that vocabulary size at different ages predicts later reading ability (van Bergen *et al.* 2014; Duff *et al.* 2015). A study of Dutch children with and without the risk for dyslexia found that expressive vocabulary at 4.5 years is correlated ( $r=0.26$ ) with reading scores at age 8 (van Bergen *et al.* 2014). Similarly, Duff *et al.* (2015) found that a latent variable comprising expressive and receptive vocabulary at 16–24 months predicts school age (age range: 4–9 years) PA, reading accuracy and reading comprehension in a British sample (accounting for 4% of the variance in PA, 11% in

reading accuracy and 18% in reading comprehension). Furthermore, these phenotypic relationships are supported by correlations at the genetic level as well, as twin studies have shown that genetic factors account for part of the correlation between language proficiency and later reading ability (Harlaar *et al.* 2008; Hayiou-Thomas *et al.* 2010).

Therefore, it was reasonable to use FIOLA to test all top associations from the GWAS studies, whether or not a given SNP had been originally reported to associate with a trait for which there was a matching measure available in the FIOLA dataset. Note that our goal was not strict replication, for which identical measures and recruitment strategies should be used across datasets. Rather, we aimed to test for supportive evidence that might help to validate any of the GWAS findings, and extend our understanding of them with respect to pleiotropy across measures of reading and language performance.

## Methods

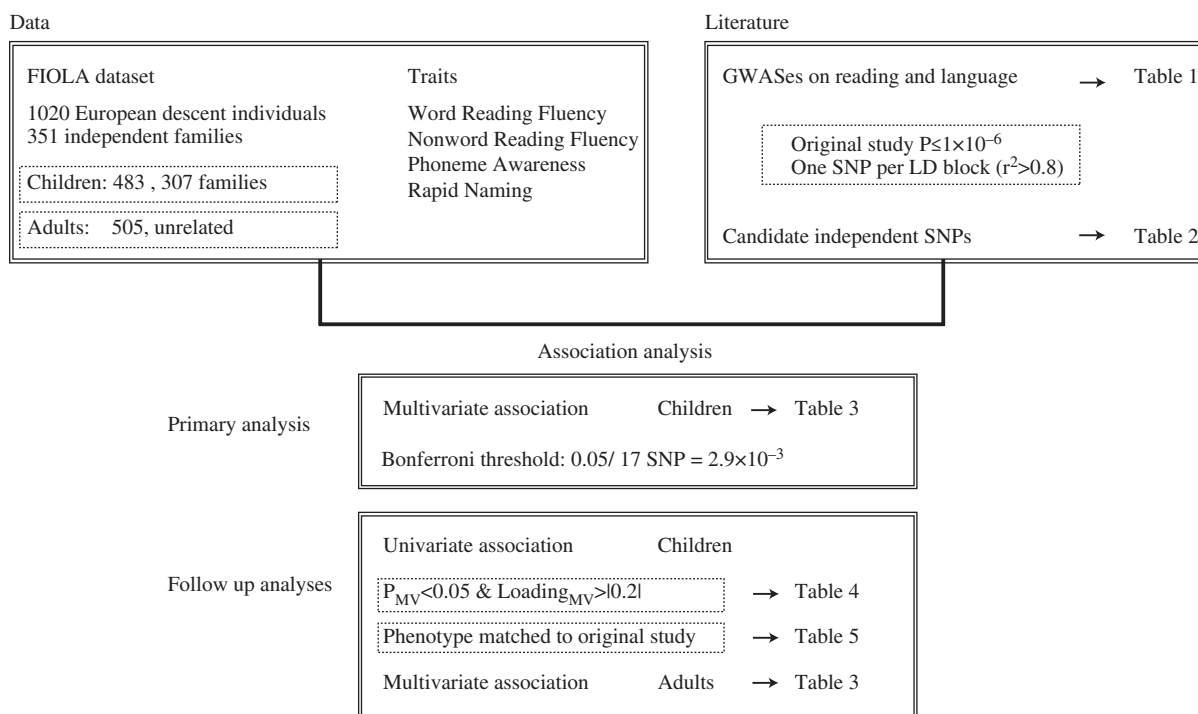
### Sample

The FIOLA project consists of a general population, family-based, Dutch sample that has been assessed with reading-related tests (van Bergen *et al.* 2015). Families that visited the Amsterdam Science Museum NEMO were invited to take part. This research is part of Science Live, the innovative research programme of Science Museum NEMO that enables scientists to carry out real, publishable, peer-reviewed research using NEMO visitors as volunteers. The museum offered quiet rooms for one-to-one testing. Ethical approval for this study was provided by the University of Amsterdam ethics committee, file number 2011-OWI-1882, and written informed consent of the participants (or their parents) was obtained. Two indicators of sample representativeness for the general population were evaluated: The level of education of the parents was 0.49 SD above the national average, and children's reading scores were 0.40 SD above the national average (van Bergen *et al.* 2015).

For this study, only individuals of European descent were included, in order to reduce the possible impact of genetic stratification (ancestry was assessed using an ethnicity questionnaire which queried as far back as the four grandparents and 60 individuals were excluded as a result). Other inclusion criteria were that participants had Dutch as their first language (no exclusion based on second language learning), and had attended Dutch primary education.

Our primary analysis was conducted on children, i.e. participants aged less than 200 months (16.6 years), because the previous GWAS studies have been carried out on children and teenagers (see Table 1), and genetic effects on reading and language performance may be partially age dependent, targeting specific developmental stages (St Pourcain *et al.* 2014). In total, there were 483 children from 307 independent families, which comprised 149 singletons, 140 sibships for which two siblings were available and 18 sibships for which three children were available. The minimum age was 6 years (mean age = 10.06 years, SD = 1.97 years).

As a secondary step, we separately analysed all of the unrelated adults available ( $n=505$ , aged 33–68 years, mean = 43.27 years, SD = 4.58 years), who included parents of the children (when available) and 50 other unrelated individuals who had been tested at their homes using the same measures, including parallel form tests to compute parallel form reliability. The inclusion of parents meant that this analysis was not entirely statistically independent from the analysis of children, but provided a useful complementary analysis available in this dataset. While programmes do exist to analyse the child and parent information all together, we were unaware of an option that would do this in a multivariate context, to support simultaneous association analysis with multiple phenotypic measures (see below). Possible age dependence of genetic effects on language-related traits also supports an approach of analysing adults separately from children.



**Figure 1: FIOLA genetics study design.**

### Traits of interest

The traits of interest for this study were reading fluency of words and nonwords, phoneme awareness (PA), and serial rapid automatized naming (RAN). These are well-established indices of reading ability and related cognitive performance. Phonological awareness and rapid naming are correlated with reading ability, and may each be linked to independent fractions of the variance in reading performance (Landerl *et al.* 2013; van der Leij *et al.* 2013). Both have often been included in previous studies of the genetics of dyslexia and reading (Francks *et al.* 2004; Rubenstein *et al.* 2014).

Word reading fluency was assessed using the One-Minute-Test (in Dutch, Een-Minuu-Test or EMT), (Brus & Voeten 1972) and nonword reading fluency was assessed using the Klepel test (van den Bos *et al.* 1994). Participants were asked to correctly read as many (non)words as possible within 1 min (word reading) or 2 min (nonword reading). To avoid a ceiling effect in adults, the original lists were extended, resulting in lists of 145 words or nonwords (van Bergen *et al.* 2015).

Phonological awareness was measured with a phoneme deletion task (van Bergen *et al.* 2015). For each test item, a phoneme (always a consonant) had to be deleted from a nonword, resulting in another nonword. Accuracy and speed were combined into a fluency measure.

Rapid automatized naming was measured by naming a matrix of 50 digits as quickly as possible (van den Bos 2003). The time to completion was transformed to the number of digits per second, to normalize the score distribution.

The reliability of all tasks was high. Reliability measures were available from manuals for children: parallel form reliabilities for word reading (0.76–0.96) and nonword reading (0.89–0.95), and split-form and test–retest reliabilities for RAN (0.78–0.92). Test–retest reliabilities were also estimated in a sample of 66 independent children for RAN (0.80) and PA (0.75–0.81) (van Bergen *et al.* 2015). A subset of 50 adults from the present sample were tested with parallel forms to estimate reliabilities for adults, which were 0.96 for word reading, 0.94 for nonword reading and 0.94 for RAN. Internal consistency for PA in this adult sample had a Cronbach's  $\alpha$  of 0.71–0.89 (for accuracy and reaction time) [for details and full task descriptions, see van

Bergen *et al.* (2015)]. In order to account for the linear and quadratic effects of age (in months), age and age<sup>2</sup> were regressed out from the scores of children (age < 200 months), and the standardized residuals were used as trait scores. Scores for adults (age > 200 months) were z-standardized to bring them on the same scale. Extreme values for each trait were identified as outliers when they were below/above 1.5 times the interquartile range from the first or third quartiles. As a result, 15 datapoints from the children's dataset were removed (word reading = 3, nonword reading = 1, PA = 6, RAN = 5) and 31 from the adults' dataset (word reading = 7, nonword reading = 7, PA = 8, RAN = 9).

All of the phenotypic traits were normally distributed and significantly correlated with each other both for children and adults [see Fig. S1 (Supporting information) and van Bergen *et al.* 2015]. The correlations ranged from moderate ( $r_{\text{PA-RAN}} = 0.34$  for children and 0.39 for adults) to high ( $r_{\text{EMT-KL}} = 0.84$  for children and 0.68 for adults).

### Power analysis

Power estimates were performed using the Genetic Power Calculator (Purcell *et al.* 2003), corresponding to univariate association analysis. Power was calculated for a range of type I error rates ( $\alpha$  values) and QTL effect sizes, given two different experimental designs that were roughly comparable to analysis of either the children or the adults from the FIOLA dataset: (1) 240 nuclear families with sibship size of 2, with both parents genotyped, but with only offspring phenotyped and (2) 500 unrelated individuals.

### SNP selection and genotyping

From the GWAS studies summarized in Table 1, and as described above, we selected all SNPs reported to be associated with reading or language traits with  $P < 10^{-6}$ . We then pruned the candidate variant list to reduce redundancy by selecting only one SNP per linkage disequilibrium (LD) block ( $r^2 > 0.8$ , CEU (Utah residents with ancestry from northern and western Europe) population) using the SNP annotation and proxy search (SNAP) (Johnson *et al.* 2008),

**Table 2:** Candidate SNPs from the top hits of the recent GWAS literature on reading and language abilities, ordered by genomic coordinates. SNPs that failed genotyping assay design are marked in italics. SNP position is given for the human reference hg19

Reference	Phenotype	SNP	Risk	P-value	Effect	Chromosome	Position (bp)	MAF <sub>eur</sub>	Min	Type	Gene	Closest gene
Luciano <i>et al.</i> (2013)	Word reading	rs4839516	C	3.62e-07	0.08 (0.02)*	1	116 740 700	0.29	A	—	—	MAB21L3
Harlaar <i>et al.</i> (2014)	Receptive language	rs12474600	G	4.57e-07	-0.24 (0.05)*	2	179 885 136	0.10	A	Intronic	CCDC141	SESTD1
Eicher <i>et al.</i> (2013)	Case (dyslexia + SLI) control	rs12636438	G	5.45e-07	1.81†	3	22 063 527	0.17	G	Intronic	ZNF385D	ZNF385D-AS2
St Pourcain <i>et al.</i> (2014)	Early expressive vocabulary	rs7642482	G	1.3e-08	0.11 (0.02)*	3	77 718 006	0.15	G	—	—	ROBO2
Eicher <i>et al.</i> (2013)	Case-control	rs482700	G	1.4e-07	1.83†	4	116 067 740	0.28	G	—	—	NDST4
Nudel <i>et al.</i> (2014)	Maternal-SLI	rs10447141	G	1.16e-07	2.77†	5	398 174 17	0.29	T	—	—	LINC00603
Field <i>et al.</i> (2013)	Case (dyslexia) control	rs9313548	T	6.2e-07	3.08†	5	170 961 550	0.45	C	—	—	FGF18
Nudel <i>et al.</i> (2014)	Paternal-SLI	rs7801303	-	3.51e-07	—	7	40 918 320	0.42	A	—	—	LINC01450
Gialluisi <i>et al.</i> (2014)	Principal component	rs59197085	A	3.86e-07	0.001-0.033‡	7	128 461 006	0.09	A	Intronic	CCDC136	FLNC
Harlaar <i>et al.</i> (2014)	Receptive language	rs1326167	T	9.59e-07	0.17 (0.04)*	10	92 101 276	0.25	C	—	—	LOC101926942
St Pourcain <i>et al.</i> (2014)	Early expressive vocabulary	rs10734234	T	1.9e-07	-0.14 (0.03)*	11	15 466 110	0.10	C	—	—	LOC102724957
St Pourcain <i>et al.</i> (2014)	Early expressive vocabulary	rs11176749	T	7.2e-07	0.12 (0.03)*	12	67 853 034	0.14	T	—	—	LOC100507175
Eicher <i>et al.</i> (2013)	Case (dyslexia + SLI) control	rs9521789	C	7.59e-07	1.71†	13	111 119 870	0.41	C	Intronic	COL4A2	COL4A2-AS1
Nudel <i>et al.</i> (2014)	Paternal-SLI	rs4280164	G	3.74e-08	3.87†	14	24 771 535	0.20	A	Missense	NOP9	CIDEB
Luciano <i>et al.</i> (2013)	Word reading	rs764255	T	1.8e-07	0.25 (0.05)*	16	73 713 933	0.37	C	—	—	LOC100506172
Luciano <i>et al.</i> (2013)	Nonword repetition	rs7187223	G	9.9e-08	-0.08 (0.02)*	16	82 457 983	0.04	G	—	—	CDH13
St Pourcain <i>et al.</i> (2014)	Early expressive vocabulary	rs1654584	G	3.4e-07	0.08 (0.02)*	19	3 970 933	0.22	G	Intronic	DAPK3	EEF2
Luciano <i>et al.</i> (2013)	Nonword repetition	rs2192161	G	7.34e-08	0.20 (0.04)*	21	15 682 244	0.07	A	Intronic	ABCC13	ABCC13
Gialluisi <i>et al.</i> (2014)	Principal component	rs5995177	A	5.01e-07	0.006-0.025‡	22	36 309 803	0.07	A	Intronic	RBFOX2	RBFOX2
Gialluisi <i>et al.</i> (2014)	Principal component	rs12158565	G	7.57e-07	0.004-0.036‡	22	36 317 093	0.10	G	Intronic	RBFOX2	RBFOX2

MAF, minor allele frequency; Min, minor allele.

\*Effect sizes are given as  $\beta$  coefficient (SE).

†Effect sizes are given as the odds ratio for case-control studies.

‡Effect sizes are given as  $r^2$  for the quantitative trait studies.

which resulted in 20 SNPs as listed in Table 2. The SNPs were directly genotyped using the Kompetitive Allele Specific PCR (KASP) genotyping assay (LGC Ltd., Teddington, UK). Three of the SNPs failed at either the assay design or validation phase (rs59197085, rs5995177 and rs11176749), and were therefore excluded from the study. As a result, 17 SNPs were included in this study (Table 2).

Mendelian inheritance errors were detected using pedstats 0.6.12 (Wigginton & Abecasis 2005). PLINK 1.07 (Purcell *et al.* 2007) was used to calculate missing genotype rates and to test for Hardy–Weinberg disequilibrium. Three individuals (two children and one adult) with more than 2/17 missing genotypes (missing genotype rate > 12%) were excluded. The total genotyping rate in the remaining individuals ( $n_{\text{children}} = 481$ ,  $n_{\text{adults}} = 505$ ) was 99.8%, with a missing genotype rate < 5% for all the SNPs. All SNPs were in Hardy–Weinberg equilibrium ( $P > 0.05$ ).

### Association analysis

We used multivariate association analysis as our primary means of testing for association of the reading and language traits with each SNP, using the 'mqfam' option of PLINK-multivariate (Ferreira & Purcell 2009). This approach was appropriate given the correlation structure of our traits of interest in both subsets of the dataset (children and adults, see Fig. S1), and it had the advantage that only one statistical test was performed per SNP in primary testing, thus reducing multiple testing (Bonferroni-adjusted significance threshold of  $P = 3 \times 10^{-3}$  for 17 SNPs,  $\alpha = 0.05$ ). The method finds the linear combination of a set of correlated measures which is most strongly associated with the given SNP genotypes. We performed the multivariate association analysis separately for children and parents. For children, 10 000 permutations were used in PLINK-multivariate to account for the family-based structure of the dataset (including children and parental genotypes when available), in order to obtain an empirical  $P$  value per SNP.

The multivariate analysis was followed by testing 'total' univariate association with PLINK 1.07, for three SNPs (see below) which showed multivariate  $P < 0.05$  in the children, in order to explore the evidence for pleiotropy across measures. The 'total' test of association may potentially be affected by population stratification. Therefore, a population stratification test was also performed using the  $-ap$  model in QTDT (Linkage Disequilibrium Analyses for Quantitative and Discrete Traits) 2.6.1, which assesses the equivalence of the 'within-family' and 'between-family' mean allelic effects (Abecasis *et al.* 2000). Additionally, for two of the SNPs, the FIOLA dataset

contained a particularly closely matched phenotypic measure to that which showed association in the original GWAS study (see Table 2), and therefore univariate association testing was performed for these SNPs using the closest matching measure, as a relatively direct attempt at replication.

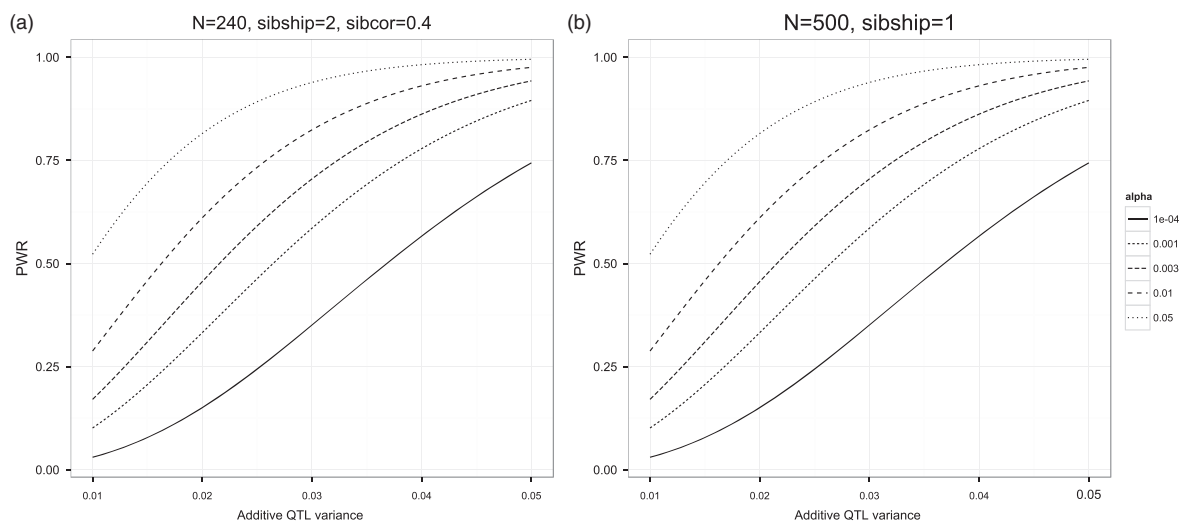
The results from the univariate tests for selected SNPs (in Tables 4 and 5) were meta-analysed together with the results from the original studies. This was implemented in the programme METAL (version March 2011) (<http://www.sph.umich.edu/csg/abecasis/Metal/index.html>; Willer *et al.* 2010). We chose an approach that does not assume equivalence of allelic effect sizes between datasets, which was appropriate given the heterogeneity of study recruitment, assessment and trait definitions. Put briefly, the meta-analysis tested each SNP for a genetic effect, across the two contributing datasets, computing an overall  $z$ -score for that SNP determined by the  $P$  value, the direction of the allelic effect and the sample size of each study involved in the meta-analysis. For the SNPs that had first been reported by studies with unequal numbers of cases and controls, we computed the effective sample size as recommended by Willer *et al.* (2010) ( $N_{\text{eff}} = 4 / (1/N_{\text{cases}} + 1/N_{\text{ctrls}})$ ).

Finally, for three SNPs that had been associated with parent-of-origin SLI transmission to children in one GWAS study, we modelled the effects of maternally and paternally derived alleles separately in a parent-of-origin analysis using QTDT (Abecasis *et al.* 2000) with the  $-ao -of$  options.

## Results

### Power analysis

Our study provided power of < 60% to detect nominally significant results ( $\alpha = 0.05$ ) for effects of 1%, for both the children and parent analysis. However, the effect sizes that were reported in the GWAS papers, for the SNPs we investigated here, ranged from 0.5 to 12% (e.g. rs2192161 explained 3.9% of the variance in the original study; Luciano *et al.* 2013). We had high power to detect effects that explain roughly 5% of the phenotypic variance ( $\beta > 0.80$  with  $\alpha = 0.003$ , i.e. Bonferroni correction for 17 SNPs) and moderate power to detect smaller effects (e.g.  $\beta \approx 0.60$  for effects that explain 3% of the



**Figure 2: Power estimation.** The x-axis is the additive proportion attributable to the quantitative trait locus (range: 0–5%). (a) Children, family-based design and (b) adults, unrelated sample design.

**Table 3:** Plink multivariate association results (total test) from children and adults

SNP	Children					Adults							
	$N_{fam}$	$N_{ind}$	$F$	Loadings: WRF, NWRF, PA, RAN		$P_{emp}$	MAF	$N_{fam}$	$N_{ind}$	$F$	Loadings: WRF, NWRF, PA, RAN		$P$
rs4839516	307	945	1.82			0.104	0.27	481	481	1.31			0.265
rs12474600	307	943	0.13			0.967	0.09	479	479	2.02			0.090
rs12636438	307	945	2.37	0.6044, 0.847, 0.2881, 0.745		0.045	0.20	479	479	0.24			0.916
rs7642482	307	941	0.32			0.856	0.21	479	479	0.92			0.451
rs482700	307	946	2.31	-0.2751, 0.21, 0.3467, 0.3788		0.047	0.26	481	481	0.75			0.560
rs10447141	307	945	1.61			0.144	0.33	481	481	1.28			0.278
rs9313548	307	942	0.42			0.759	0.48	480	480	0.74			0.564
rs7801303	307	945	2.18			0.052	0.44	480	480	0.41			0.800
rs1326167	307	940	0.75			0.507	0.24	481	481	1.94			0.102
rs10734234	307	944	2.01			0.097	0.11	481	481	0.84			0.501
rs9521789	307	944	1.19			0.291	0.40	481	481	0.21			0.935
rs4280164	307	943	0.36			0.822	0.21	481	481	0.27			0.896
rs764255	306	940	0.09			0.982	0.36	480	480	1.36			0.249
rs7187223	307	943	2.40	0.2892, 0.2347, 0.7195, 0.6826		0.020	0.03	481	481	2.41	-0.4112, -0.5619, -0.9859, -0.2483		0.049
rs1654584	307	945	2.13			0.074	0.25	481	481	2.03			0.089
rs2192161	307	940	0.64			0.592	0.06	478	478	1.12			0.348
rs12158565	307	945	0.26			0.891	0.13	481	481	0.51			0.727

Loadings reflect the contribution of that trait to the multivariate test, and are only shown when  $P < 0.05$ . Multiple testing correction for 17 SNPs sets the threshold to  $2.9 \times 10^{-3}$ . NWRF, nonword reading fluency;  $P_{emp}$ , empirical  $P$ -values (10 000 permutations); PA, phonological awareness; RAN, rapid naming; WRF, word reading fluency.

phenotypic variance; see Fig. 2). Nevertheless, it is likely that the effect sizes in the GWAS studies where the associations were first reported are overestimates of any real effects, due to the winner's curse (Ioannidis 2008), and hence our power to detect these effects may be only roughly 60%.

**Association analysis**

We performed multivariate association analysis first in the sample of children and then in adults, followed by univariate analysis for the SNPs that were nominally significant in the multivariate analysis for each sample.

Multivariate analysis in children resulted in three nominally significant associations (rs12636438,  $P_{emp} = 0.045$ ; rs482700,  $P_{emp} = 0.047$  and rs7187223,  $P_{emp} = 0.020$ ) (see Table 3). The loadings were  $>|0.2|$  for all four phenotypic measures in the multivariate models for all three SNPs, although for rs482700 word reading fluency had the opposite direction of effect to the rest of the measures. One of these markers, rs7187223, was the only SNP to show nominally significant association in the adults, with the highest loading for phoneme awareness (loading<sub>PA</sub>  $>|0.9|$ ,  $P = 0.049$ ; see Table 3). These associations did not survive statistical correction for multiple testing.

Univariate analyses were performed for these three SNPs in children, confirming that the putative effects of rs12636438 and rs7187223 encompassed several traits. However, there was no nominally significant univariate association with rs482700 (Table 4). The most significant associations with rs12636438 were with nonword reading fluency and rapid naming ( $P_{NWRF} = 0.01411$  and  $P_{RAN} = 0.02376$ ; see Fig. S3). There was also a consistent but non-significant trend of association with word reading fluency. However, the direction

of these effects in the FIOLA dataset was opposite to that previously reported: the allele associated with poorer performance was major allele (A) in our study, but the minor allele (G) was overrepresented in comorbid cases with SLI and dyslexia in the original GWAS study of Eicher *et al.* (2013). This inconsistency in the direction of effect was reflected by the meta-analysed  $P$ -values which were all higher than the originally reported one (Table 4). In the children, rs7187223 was nominally associated with rapid naming and there was a trend with the same direction of effect with phoneme awareness ( $P_{RAN} = 0.03897$ ,  $P_{PA} = 0.03584$ ). The minor allele G was associated with lower performance, which was consistent with the effect reported by Luciano *et al.* (2013), and reflected by meta-analysed  $P$  values  $< 1.4e-08$  for RAN and PA (Table 4). Because this SNP was nominally significant in the multivariate analysis of the adults, we tested association with univariate traits within this sample too, and found that it was also associated with phoneme awareness ( $P_{PA} = 0.0026$ ). However, the result on the adults had the opposite direction of effect to that within the children, and also opposite to that previously reported for nonword repetition (Luciano *et al.* 2013) (see Fig. S3). The tests for population stratification did not show significant differences of the between-family and within-family components of association for any SNP (all  $P > 0.05$ ), suggesting that population stratification was unlikely to be a substantial confounding factor in the analysis.

Word reading fluency in FIOLA is closely matched to the 'word reading' phenotype from Luciano *et al.* (2013), and hence we tested association of this trait within the children with the two SNPs (rs4839516 and rs764255) that were reported in Luciano *et al.* (2013). However, we did not find supportive evidence for associations of any of these SNPs

**Table 4:** PLINK univariate total association results for children. Empirical  $P$  values (10 000 permutations) are shown

SNP	Phenotype	$N_{ind}$	Risk <sub>FIOLA</sub>	$\beta$	$P_{emp}$	$P_{meta}$	Effect
rs12636438	Word reading fluency	480	A	0.16	0.08	0.00732	-, +
	Nonword reading fluency	481	A	0.22	0.01	0.02519	-, +
	Phoneme awareness	470	A	0.07	0.39	0.001007	-, +
	Rapid naming	478	A	0.18	0.02	0.01759	-, +
rs7187223	Word reading fluency	478	G	-0.18	0.36	7.298e-08	-, -
	Nonword reading fluency	478	G	-0.10	0.63	1.35e-07	-, -
	Phoneme awareness	467	G	-0.42	0.04	1.3e-08	-, -
	Rapid naming	475	G	-0.41	0.04	1.338e-08	-, -
rs482700	Word reading fluency	480	G	-0.87	0.42	7.17e-06	-, -
	Nonword reading fluency	481	A	0.56	0.60	0.0002978	-, +
	Phoneme awareness	470	A	1.01	0.33	0.000828	-, +
	Rapid naming	478	A	1.15	0.26	0.001237	-, +

Effect, consistency of the direction of effect for the original and present studies;  $P_{meta}$ , meta-analysed  $P$ -value from the original report and this study.

with word reading fluency (see Table 5). Finally, because rs7801303, rs4280164 and rs10447141 had been reported to be associated with parent-of-origin-specific transmission to children with SLI (Nudel *et al.* 2014), we tested for such effects using QTDT for these SNPs and our traits of interest, but there were no significant parent-of-origin effects ( $P > 0.1$ ).

## Discussion

The aim of this study was to evaluate the most significant associations from a recent wave of the first GWAS studies of language and reading performance, by testing for their reported effects on reading-related quantitative traits in FIOLA, a new and independent Dutch dataset from the general population. We adopted an inclusive criterion for the selection of candidate SNPs, with the primary hypothesis that genetic effects might be pleiotropic and shared among several language and reading traits (Gialluisi *et al.* 2014; Luciano *et al.* 2013; Newbury *et al.* 2011). Seventeen candidate SNPs were selected, for which the FIOLA dataset provided sufficient statistical power to detect most of the previously reported effects, although it was underpowered to detect effects smaller than 1%. Three nominally significant associations were observed in the multivariate analysis of the children, and two of these (rs12636438 and rs7187223) showed consistent signals in univariate analysis.

The association of rs12636438 loaded on all four traits, but more strongly on nonword reading fluency and rapid naming,

with the major allele A being associated with lower performance in both cases. This SNP lies within the *ZNF385D* gene, which codes for a zinc finger protein that may act as a transcriptional regulator (Eicher *et al.* 2013). However, Eicher *et al.* (2013) found that the minor allele (G) of this SNP was over-represented in comorbid cases of SLI and dyslexia, as well as associated with lower scores on a measure of receptive vocabulary (picture vocabulary test) in a follow-up sample. Thus, our findings show an opposite direction of effect to the original study. The incongruence of the allelic direction might be because of the different genetic backgrounds of the populations analysed. The haplotype structure in a region can differ between populations, changing the LD pattern between the tag SNP (where the association is detected) and the causal SNP that is driving the association, which is as yet unknown. Such factors could in principle lead to contrasting directions of effect of the same tag SNP in different studies when substantial population stratification is present (Lin *et al.* 2007), a phenomenon that has been suggested to potentially explain previous inconsistencies in the literature of the genetics of reading (Gialluisi *et al.* 2014; Luciano *et al.* 2007). It is difficult to evaluate this hypothesis without genome-wide information to assess the genetic homogeneity of all the samples (only directly genotyped SNP information is currently available for FIOLA). Despite the fact that only individuals of European descent were included in the analysis, different LD patterns between our sample and the ALSPAC sample may account for the contrasting result for rs12636438. Alternatively, it might reflect type I error, representing a false positive. Eicher *et al.* (2013) also looked at the relationship of rs12636438

**Table 5:** PLINK univariate total association results for children, where SNPs and phenotypes matched those of the original report

SNP	Reference	Phenotype <sub>ref</sub>	Risk <sub>ref</sub>	Phenotype <sub>FIOLA</sub>	$N_{ind}$	$\beta$	$P_{emp}$	$P_{meta}$	Effect
rs4839516	Luciano <i>et al.</i> (2013)	Word reading	C	Word reading fluency	480	-0.104	0.188	4.812e-06	+, -
rs764255	Luciano <i>et al.</i> (2013)	Word reading	T	Word reading fluency	477	0.033	0.650	8.507e-07	-, +

Effect, consistency of the direction of effect for the original and present studies;  $P_{meta}$ , meta-analysed  $P$ -value from the original report and this study.



and white matter volumes and reported association with volumes of several DTI fibre tracts that are important for reading and language (the inferior longitudinal fasciculus, the inferior fronto-occipital fasciculus and the temporal superior longitudinal fasciculus). Of note, FIOLA did not involve any neuroimaging of participants, so we are unable to evaluate the reliability of those findings in this study.

We found that rs7187223 was nominally associated with rapid naming and PA in the univariate analysis of children. The minor allele (G) was associated with lower performance in both cases, supporting the original association by Luciano *et al.* (2013) that reported that this allele was associated with lower nonword repetition scores. rs7187223 was also the only SNP that was nominally significant for the multivariate analysis in the sample of unrelated adults. This effect loaded most strongly on phoneme awareness, which was also reflected by the univariate association with this trait. Surprisingly, the association in the adults had the opposite direction of effect to that on children, with the major allele (A) being associated with lower performance. Haplotype structure differences as discussed above cannot be responsible for this contrasting effect, as we are comparing two subsets of the same population. The main difference between the children and adult subsamples was obviously their age. There is some evidence suggesting that age-varying effects may also contribute to differences in allelic effects (Lasky-Su *et al.* 2008), and a longitudinal study found that an SNP in the *DYX1C1* dyslexia candidate gene is associated with children's orthographic judgments at ages 7 and 8, but not at age 6 (Zhang *et al.* 2012). Moreover, in the GWAS study of St Pourcain *et al.* (2014), an association of expressive vocabulary with rs7642482 near to *ROBO2* was also age dependent: it was genome-wide significant for an early developmental stage (15–18 months) but showed no evidence of association for a similar measure a few months later (24–30 months). Nevertheless, because language and reading acquisition are developmental processes, most of the studies that look at genetic effects on these traits tend to focus on datasets consisting of children (see Table 1), and very little is known about the stability of genetic effects over time into adulthood. Alternatively, it is also possible that the opposite allelic effect between children and adults could be because of the differences underlying the measured trait. Despite the fact that the same phoneme deletion task was used to measure PA in both samples, the across-age invariance of this measurement has not been tested. Taken together, our results for rs7187223 do not provide compelling evidence to support the previously reported association.

Since one of the reading measures available for the FIOLA dataset (i.e. word reading fluency) was conceptually similar to some reported by Luciano *et al.* (2013), we specifically attempted to replicate their top hits with word reading (rs4839516 and rs764255), using univariate association tests. However, we did not replicate these associations. Similarly, we did not find any parental effect on the SNPs that had been reported to be associated with parent-of-origin transmission in relation to SLI (Nudel *et al.* 2014).

To our knowledge, this has been the first attempt to independently and systematically assess the results of recent GWAS reports for language and reading traits. We

were only able to find limited support for three of the SNPs that had been previously associated with reading- and language-related traits in these GWAS studies. The associations with the other SNPs were not supported in our study. However, this does not translate into a complete rejection of these SNPs as potentially relevant, given the range of the studies that we have considered when selecting the SNPs, and the between-study heterogeneity of the measures even in the cases where similar phenotypes were available (e.g. word reading accuracy vs. word reading fluency). The FIOLA dataset was ascertained in an unconventional manner through a public venue. This is an advantageous set-up to test large numbers of individuals and has the potential to contribute to the ascertainment of large cohorts in the future. However, the FIOLA dataset is not entirely representative of the Dutch population: it is biased towards a higher-than-average educational level for adults (0.49 SD above the average), and towards higher-than-average word reading scores for children (0.4 SD above national norms) (van Bergen *et al.* 2015). Moreover, there is a small but significant correlation of the spousal reading ability (0.16,  $P=0.019$ ; van Bergen *et al.* 2015), which may indicate assortative mating. Environmental factors such as parental education and socio-economic status have been suggested to moderate genetic influences on reading disabilities, in terms of both heritability estimates and single SNP effects (Friend *et al.* 2008; Mascheretti *et al.* 2013). Hence, non-random sampling from the population, as is apparently the case for FIOLA, adds yet another potential source of heterogeneity across different datasets and studies. IQ, which correlates phenotypically and genetically with educational attainment, should also be considered (Davies *et al.* 2016). Several of the GWAS studies have adjusted their reading-related measures for IQ prior to genetic association testing (Gialluisi *et al.* 2014; Luciano *et al.* 2013), and one showed that some association signals are sensitive to this manipulation (Gialluisi *et al.* 2014). In other words, some genetic effects may be more pleiotropic for reading-related cognition and IQ than others. The previously reported GWAS results remain to be further studied in other samples, and ideally meta-analysed across all available samples together, while accounting for the issues addressed in this study.

## References

- Abecasis, G.R., Cardon, L.R. & Cookson, W.O. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* **66**, 279–292.
- Becker, J. *et al.* (2014) Genetic analysis of dyslexia candidate genes in the European cross-linguistic NeuroDys cohort. *Eur J Hum Genet* **22**, 675–680.
- van Bergen, E., Bishop, D., van Zuijlen, T. & de Jong, P.F. (2014) IQ of four-year-olds who go on to develop dyslexia. *J Learn Disabil* **47**, 475–484.
- van Bergen, E., Bishop, D., van Zuijlen, T. & de Jong, P.F. (2015) How does parental reading influence children's reading? a study of cognitive mediation. *Sci Stud Read* **19**, 325–339.
- van den Bos, K.P. (2003) *Serieel benoemen en woorden lezen [serial naming and word reading]*. Rijksuniversiteit Groningen, Groningen, The Netherlands.
- van den Bos, K.P., Lutje Spelberg, H.C., Scheepstra, A.J.M. & de Vries, J.R. (1994) *De klepel: Een test voor de leesvaardigheid*

- van pseudowoorden [the klepel: A test for the reading skills of pseudowords]. Swets & Zeitlinger, Lisse, The Netherlands.
- Brus, B.T. & Voeten, M.J.M. (1972) *Een-minuut-test [one-minute-test]*. Swets & Zeitlinger, Lisse, The Netherlands.
- Carrion-Castillo, A., Franke, B. & Fisher, S.E. (2013) Molecular genetics of dyslexia: an overview. *Dyslexia* **19**, 214–240.
- Davies, G. et al. (2016) Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol Psychiatry* **21**, 758–767.
- DeFries, J.C., Fulker, D.W. & LaBuda, M.C. (1987) Evidence for a genetic aetiology in reading disability of twins. *Nature* **329**, 537–539.
- Duff, F.J., Reen, G., Plunkett, K. & Nation, K. (2015) Do infant vocabulary skills predict school-age language and literacy outcomes? *J Child Psychol Psychiatry* **56**, 848–856.
- Eicher, J.D. et al. (2013) Genome-wide association study of shared components of reading disability and language impairment. *Genes Brain Behav* **12**, 792–801.
- Ferreira, M.A. & Purcell, S.M. (2009) A multivariate test of association. *Bioinformatics* **25**, 132–133.
- Field, L.L., Shumansky, K., Ryan, J., Truong, D., Swiergala, E. & Kaplan, B.J. (2013) Dense-map genome scan for dyslexia supports loci at 4q13, 16p12, 17q22; suggests novel locus at 7q36. *Genes Brain Behav* **12**, 56–69.
- Francks, C., Paracchini, S., Smith, S.D., Richardson, A.J., Scerri, T.S., Cardon, L.R., Marlow, A.J., MacPhie, I.L., Walter, J., Pennington, B.F., Fisher, S.E., Olson, R.K., DeFries, J.C., Stein, J.F. & Monaco, A.P. (2004) A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. *Am J Hum Genet* **75**, 1046–1058.
- Friend, A., DeFries, J.C. & Olson, R.K. (2008) Parental education moderates genetic influences on reading disability. *Psychol Sci* **19**, 1124–1130.
- Gialluisi, A., Newbury, D.F., Wilcutt, E.G., Olson, R.K., DeFries, J.C., Brandler, W.M., Pennington, B.F., Smith, S.D., Scerri, T.S., Simpson, N.H., Luciano, M., Evans, D.M., Bates, T.C., Stein, J.F., Talcott, J.B., Monaco, A.P., Paracchini, S., Francks, C. & Fisher, S.E. (2014) Genome-wide screening for DNA variants associated with reading and language traits. *Genes Brain Behav* **13**, 686–701.
- Harlaar, N., Hayiou-Thomas, M.E., Dale, P.S. & Plomin, R. (2008) Why do preschool language abilities correlate with later reading? A twin study. *J Speech Lang Hear Res* **51**, 688–705.
- Harlaar, N., Meaburn, E.L., Hayiou-Thomas, Wellcome Trust Case Control Consortium 2, M.E., Davis, O.S., Docherty, S., Hanscombe, K.B., Haworth, C.M., Price, T.S., Trzaskowski, M., Dale, P.S., Plomin, R. (2014) Genome-wide association study of receptive language ability of 12-year-olds. *J Speech Lang Hear Res* **57**, 96–105.
- Hayiou-Thomas, M.E., Harlaar, N., Dale, P.S. & Plomin, R. (2010) Preschool speech, language skills, and reading at 7, 9, and 10 years: etiology of the relationship. *J Speech Lang Hear Res* **53**, 311–332.
- Ioannidis, J.P. (2008) Why most discovered true associations are inflated. *Epidemiology* **19**, 640–648.
- Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J. & de Bakker, P.I. (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938–2939.
- Landerl, K. et al. (2013) Predictors of developmental dyslexia in European orthographies with varying complexity. *J Child Psychol Psychiatry* **54**, 686–694.
- Lasky-Su, J. et al. (2008) On the replication of genetic associations: timing can be everything!. *Am J Hum Genet* **82**, 849–858.
- van der Leij, A., van Bergen, E., van Zuijen, T., de Jong, P., Maurits, N. & Maassen, B. (2013) Precursors of developmental dyslexia: an overview of the longitudinal Dutch Dyslexia Programme study. *Dyslexia* **19**, 191–213.
- Lin, P.I., Vance, J.M., Pericak-Vance, M.A. & Martin, E.R. (2007) No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* **80**, 531–538.
- Luciano, M., Lind, P.A., Duffy, D.L., Castles, A., Wright, M.J., Montgomery, G.W., Martin, N.G. & Bates, T.C. (2007) A haplotype spanning KIAA0319 and TTRAP is associated with normal variation in reading and spelling ability. *Biol Psychiatry* **62**, 811–817.
- Luciano, M., Evans, D.M., Hansell, N.K., Medland, S.E., Montgomery, G.W., Martin, N.G., Wright, M.J. & Bates, T.C. (2013) A genome-wide association study for reading and language abilities in two population cohorts. *Genes Brain Behav* **12**, 645–652.
- Mascheretti, S., Bureau, A., Battaglia, M., Simone, D., Quadrelli, E., Croteau, J., Cellino, M.R., Giorda, R., Beri, S., Maziade, M. & Marino, C. (2013) An assessment of gene-by-environment interactions in developmental dyslexia-related phenotype. *Genes Brain Behav* **12**, 47–55.
- Newbury, D.F., Paracchini, S., Scerri, T.S., Winchester, L., Addis, L., Richardson, A.J., Walter, J., Stein, J.F., Talcott, J.B. & Monaco, A.P. (2011) Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behav Genet* **41**, 90–104.
- Nudel, R., Simpson, N.H., Baird, G., O'Hare, A., Conti-Ramsden, G., Bolton, P.F., Hennessy, E.R., Ring, S.M., Davey Smith, G., Francks, C., Paracchini, S., Monaco, A.P., Fisher, S.E. & Newbury, D.F. (2014) Genome-wide association analyses of child genotype effects and parent-of-origin effects in specific language impairment. *Genes Brain Behav* **13**, 418–429.
- Purcell, S., Cherny, S.S. & Sham, P.C. (2003) Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. & Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575.
- Rubenstein, K.B., Raskind, W.H., Berninger, V.W., Matsushita, M.M. & Wijsman, E.M. (2014) Genome scan for cognitive trait loci of dyslexia: rapid naming and rapid switching of letters, numbers, and colors. *Am J Med Genet B Neuropsychiatr Genet* **165B**, 345–356.
- St Pourcain, B. et al. (2014) Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nat Commun* **5**, 4831.
- Tran, C., Gagnon, F., Wigg, K.G., Feng, Y., Gomez, L., Cate-Carter, T.D., Kerr, E.N., Field, L.L., Kaplan, B.J., Lovett, M.W. & Barr, C.L. (2013) A family-based association analysis and meta-analysis of the reading disabilities candidate gene DYX1C1. *Am J Med Genet B Neuropsychiatr Genet* **162**, 146–156.
- Visscher, P.M., Brown, M.A., McCarthy, M.I. & Yang, J. (2012) Five years of GWAS discovery. *Am J Hum Genet* **90**, 7–24.
- Wigginton, J.E. & Abecasis, G.R. (2005) PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* **21**, 3445–3447.
- Willer, C.J., Li, Y. & Abecasis, G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191.
- de Zeeuw, E.L., de Geus, E.J. & Boomsma, D.I. (2015) Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends Neurosci Educ* **3**, 69–76.
- Zhang, Y., Li, J., Tardif, T., Burmeister, M., Villafuerte, S.M., McBride-Chang, C., Li, H., Shi, B., Liang, W., Zhang, Z. & Shu, H. (2012) Association of the DYX1C1 dyslexia susceptibility gene with orthography in the Chinese population. *PLoS One* **7**, e42969.
- Zhong, R., Yang, B., Tang, H., Zou, L., Song, R., Zhu, L.Q. & Miao, X. (2013) Meta-analysis of the association between DCDC2 polymorphisms and risk of dyslexia. *Mol Neurobiol* **47**, 435–442.
- Zou, L., Chen, W., Shao, S., Sun, Z., Zhong, R., Shi, J., Miao, X. & Song, R. (2012) Genetic variant in KIAA0319, but not in DYX1C1, is associated with risk of dyslexia: an integrated meta-analysis. *Am J Med Genet B Neuropsychiatr Genet* **159B**, 970–976.

## Acknowledgments

This research was supported by the Max Planck Society (Germany). E.vB. is supported by a Rubicon Fellowship (grant no. 446-12-005) from the Netherlands Organisation for Scientific Research (NWO). Many thanks to all of the participants in the study.

## Supporting Information

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

**Figure S1: Correlation panel of z-scores for the four phenotypes analysed in samples 1 and 2.** The lower panel

contains the scatter plot of the raw data for each pair of phenotypes. The values in the upper panel correspond to the Pearson's correlation coefficient, and the significance of the correlation. The diagonal shows the histogram of each of the scores. EMT, word reading fluency; KL; nonword reading fluency.

**Figure S2: Rapid naming (zRAN) and nonword reading fluency (zKL) scores in relation to rs12636438 genotypes for the children.** The mean with SE per group is shown in red.

**Figure S3: Rapid naming (zRAN) and phoneme awareness (zPA) scores in relation to rs7187223 genotypes for children and adults.** The mean with SE per group is shown in red.