Behav Genet
DOI 10.1007/s10519-010-9402-9

ORIGINAL RESEARCH

Genetic Variance in a Component of the Language Acquisition Device: *ROBO1* Polymorphisms Associated with Phonological Buffer Deficits

Timothy C. Bates · Michelle Luciano · Sarah E. Medland · Grant W. Montgomery · Margaret J. Wright · Nicholas G. Martin

Received: 27 January 2010 / Accepted: 28 September 2010

Abstract The region containing *ROBO1* (Chromosome 3p12.3) has been implicated as a susceptibility gene for reading disorder and language deficit by translocation and linkage data. No association studies have yet been reported supporting any candidate gene. Here we report the first association of this gene with language deficits, specifically with phonological buffer deficits (a phenotype implicated in language acquisition, Specific Language Impairment and Speech Sound Disorder) and dyslexia (reading and spelling ability traits) in an unselected sample of adolescent twins and their siblings. Family-based analyses were performed on 144 tag SNPs in *ROBO1*, typed in 538 families with up to five offspring and tested for association with a developmental marker of language impairment (phonological buffer capacity, assessed using non word repetition).A reading and spelling ability measure—based on validated measures of lexical processing (irregular word) and grapheme–phoneme decoding (pseudo word)—and measures of short-term and working memory were also analysed. Significant association for phonological buffer capacity was observed for 21 of 144 SNPs tested, peaking at 8.70×10^{-05} and 9.30×10^{-05} for SNPs rs6803202 and rs4535189 respectively for nonword repetition, values that survive correction for multiple testing. Twenty-two SNPs showed significant associations for verbal storage (forward digit span)—a trait linked to phonological span. By contrast, just 5 SNPs reached nominal significance for working-memory, not surviving correction, and, importantly, only one SNP in the 144 tested reached nominal significance (0.04) for association with reading and spelling ability. These results provide strong support for *ROBO1* as a gene involved in a core trait underpinning language acquisition, with a specific function in supporting a short-term buffer for arbitrary phonological strings. These effects of *ROBO1* appear to be unrelated to brain mechanisms underpinning reading ability, at least by adolescence. While replication will be critical, the present results strongly support *ROBO1* as the first gene discovered to be associated with language deficits affecting normal variation in language ability. Its functional role in neuronal migration underlying bilateral symmetry and lateralization of neuronal function further suggests a role in the evolution of human language ability.

Keywords Specific language impairment · Dyslexia · *ROBO1* · Association · Language · Reading ability · Spelling ability · Family-based

Edited by Elena Grigorenko and Brett Miller.

T. C. Bates · M. Luciano
Centre for Cognitive Aging and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

T. C. Bates (✉) · M. Luciano
Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, UK
e-mail: tim.bates@ed.ac.uk

T. C. Bates · S. E. Medland · G. W. Montgomery · M. J. Wright · N. G. Martin
Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia

Introduction

Both language and reading are normally varying, neurodevelopmental traits (Bates et al. 2007d; Olson and Byrne 2005). Genetic studies suggest that these disorders are highly heritable but also genetically complex (Paracchini et al. 2007), with multiple genes influencing variance, and,

Published online: 15 October 2010

likely, some genes with generalist effects influencing a broad range of cognitive ability (Kovas and Plomin 2006)—corresponding to the genetic basis of what Spearman termed *g*. These studies also show a considerable specificity of gene effects, restricted to particular components of cognition in language and dyslexia (Bates et al. 2007a). Importantly for theories of neurodevelopmental variation, the same genes appear to account for both normal variation and the majority of disorder (Bates et al. 2007b, c, d; Plomin et al. 2007).

To date, molecular research has been considerably more successful for dyslexia than language impairment, with some dozen regions implicated by linkage studies in both affected (Fisher et al. 2002; Paracchini et al. 2007) and normally varying (Bates et al. 2007b) samples. Support for specific genes, namely, *KIAA0319* (Cope et al. 2005; Luciano et al. 2007; Paracchini et al. 2008), *DCDC2* (Lind et al. 2009; Meng et al. 2005; Schumacher et al. 2006), and *DYX1C1* (Bates et al. 2009a; Dahdouh et al. 2009) has also been found. To date for language impairment one gene, *FOXP2*, is indicated as causing a Mendelian form of language and speech disorder, and with no apparent mutations in this gene affecting either Specific Language Impairment (SLI) cases outside the affected family or normal variation in language (Newbury et al. 2002). A second candidate has been reported—*CNTNAP2*—which may affect nonword repetition (Vernes et al. 2008). Other genes under selection in human evolution have been implicated in language evolution (Dediu and Ladd 2007), but have yet to be linked to particular language phenotypes (Bates et al. 2008).

Here we focus on the peri-centromeric region of chromosome 3, implicated in language function by four independent groups studying diverse populations. A Finnish pedigree study first implicated a 34-megabase region on C312p identified as being transmitted in the majority of affected individuals in a four-generation family segregating developmental dyslexia in an autosomal-dominant fashion (Nopola-Hemmi et al. 2001). Fisher et al. next reported evidence for linkage in this region in UK and US samples selected for dyslexia (Fisher et al. 2002) and Bates et al. using an unselected Australian twin sample also supported linkage at 3p12 to a broad dyslexia phenotype (Bates et al. 2007b). Importantly for understanding the phenotype underlying this linkage signal, a fourth report using a US sample selected for Speech Sound Disorder (SSD) (rather than dyslexia) also found strong support for linkage in this region (Stein et al. 2004). In this latter case, the phenotype most strongly implicated was reception and storage of phonological information (markers D32465 empirical $P = 2 \times 10^{-5}$, and D3S3716, empirical $P = 4.6 \times 10^{-4}$). Within this domain the critical phenotype appeared to be nonword repetition. Turning back to original reports (Nopola-Hemmi et al. 2000, 2001) the critical phenotypes linked to the C3 translocation were not dyslexia per se, but rather involved short-term memory (linked to nonword repetition), phonological awareness (possibly reliant on adequate phonological storage), and phonological naming (an articulatory task similar to that found to be linked to C3p12 by Stein et al. (2004)). Evidence to date, then, is compatible with a gene or genes in the C3p12 region being linked to either or both of dyslexia and SLI/SSD. Difficulties in reading have been argued to follow as a consequence of SSD (Pennington and Lefly 2001), and it may be that damage to some, if not all mechanisms implicated in language acquisition has wide-spread effects on communication, manifesting in the severe dyslexia seen by Nopola-Hemmi et al. in their Finnish pedigree study.

As the search in this region turned to candidate genes, attention focused on *ROBO1* as a candidate dyslexia gene (Hannula-Jouppi et al. 2005); it was initially suggested as being disrupted by the t(3;8)(p12;q11) translocation (Hannula-Jouppi et al. 2005), and was located in the region implicated in linkage scans for SSD (Stein et al. 2004) and dyslexia (Nopola-Hemmi et al. 2001). *ROBO1* is involved in bilateral symmetry of the nervous system due to its influence on axons that project across the midline in their migration along long distance chemo-gradients. A member of a family of similar genes (cf. Barber et al. 2009), *ROBO1* has roles in both axon guidance and neuronal precursor guidance (Andrews et al. 2006), and is mediated by SLIT-family proteins (Andrews et al. 2008).

Several lines of evidence thus prompted us to explore the possibility that the *ROBO1* gene is related not to the core phenotypes of reading—lexical storage and grapheme-phoneme decoding—but rather to core traits underpinning language acquisition. Firstly, studies of SSD implicate phonological storage as a core phenotype linked to the DYX5 region (Stein et al. 2004), and this is supported by the phonological basis of many phenotypes affected by a disrupting translocation in *ROBO1* (Nopola-Hemmi et al. 2000, 2001). Secondly, the mechanism of *ROBO1* is distinct from that of other neuronal migration genes implicated in dyslexia, which have been shown to affect shorter-range migration and do not involve symmetry (Galaburda et al. 2006; LoTurco and Bai 2006). Third, the finding that *ROBO1* is implicated in disorders involving language deficits (Anitha et al. 2008) provides direct support for a role in language. Here we hypothesized that the nonword repetition measure of the phonological buffer mechanism (Gathercole 2006) has evolved as a component in a language acquisition device (Baddeley et al. 1998) and will be associated with *ROBO1* function but not directly with reading and spelling ability or manipulation of working memory.

Methods and materials

Subjects

Twins were initially recruited from primary schools in the greater Brisbane area, by media appeals and word of mouth, as part of ongoing studies of melanoma risk factors and cognition (McGregor et al. 1999; Wright et al. 2001). Data were also gathered from non-twin siblings of twins, with families comprising up to five siblings (including twins). There were two waves to this study: the first in which reading and spelling measures were collected and the second in which language measures (including short term memory) were added to the existing test battery. Exclusion criteria were parental report of significant head injury, neurological or psychiatric illness, substance abuse or dependence, or current use of psychoactive medication in either twin. Participants had normal or corrected-to-normal vision (better than 6/12 Snellen equivalent).

Data were available for 1177 individuals (1111 for the language assessment) with relevant phenotype and genotyping data. In the case of the reading phenotype, this comprised 538 families, and 136 MZ pairs, 343 DZ twin pairs and a further 11 triplets ranging in age from 12.3 to 25.1 years (mean = 17.9, SD = 2.9), 54.5% of whom were female. Numbers were slightly lower for the language phenotype (1111 individuals, 505 families, 126 MZ and 326 DZ pairs, and 9 triplets, mean age 20.1 (SD = 3.4), again 54.5% female).

The sample is 98% Caucasian by self-report, predominantly Anglo-Celtic ($\sim$82%) and is typical of the Queensland population on a range of traits including intellectual ability (Schousboe et al. 2003). Blood samples were collected at the end of testing sessions from participants and if possible from their parents. Zygosity was initially based on self-report and revised where necessary using the genotyping results. Ethical approval for this study was received from the Human Research Ethics Committee, Queensland Institute of Medical Research. Written informed consent was obtained from each participant and their parent/guardian (if younger than 18 years) prior to phenotype and blood collection.

Measures

A quantitative measure of reading and spelling ability was formed as the principal component of the Irregular-word and non-word scales for reading and for spelling assessed using the CORE (Bates et al. 2004), a reliable 120-word extended version of the Castles and Coltheart (1993) test with additional items included to increase the difficulty level for an older sample. The SLI endophenotype of phonological storage efficiency was assessed using a standardised battery combining scores on the Gathercole et al. (1994) and Dollaghan and Campbell (1998) measures of nonword repetition, a heritable trait, previously linked to at least two genetic loci involved in specific language impairment (SLI Consortium 2004). Also included were the WAIS III (Wechsler 1997) Digit-Span Forwards task, as a measure of verbal short-term memory and letter number sequencing task as a measure of working memory.

Scores on the nonword repetition (NWR) tasks were normally distributed, and were simply standardized and summed to produce a composite measure of phonological memory. Each of the three reading sub-tests and three spelling tests were calculated as a simple sum of correct items and were Box–Cox (Box and Cox 1964) transformed to normalize their distributions. Intelligence was used as a covariate in all cases, as controlling for general cognitive ability has been shown to increase sensitivity for reading ability (Luciano et al. 2007). Because scores on the verbal IQ are confounded with reading ability, performance IQ was used as a covariate, using performance scales from the Multidimensional Aptitude Battery (Jackson 1984) completed by subjects as close as possible to their 16th birthday (siblings were a year older on average than twins at time of testing). Therefore, some participants will have been tested on IQ several years before or after their reading and language measures were collected; but note that IQ is a very stable trait. The distribution of performance IQ was normal (mean 113, SD 16.2; range: 64–151) and no exclusions were made for participants with low IQ.

Genotyping

DNA was extracted from blood samples and genotyped with Illumina 610K chips. Data-checking procedures were based on exclusion of unreliable samples and SNPs, as described in Benyamin et al. (2009). These included deviation from Hardy–Weinberg equilibrium at $P < 10^{-6}$, minor allele frequency <0.01, and Mendelian errors. Subjects found to be of non-European ancestry by principal components analysis of the genotyping data were also excluded. One hundred and sixty-six tag SNPs covering the *ROBO1* region were selected via Tagger (de Bakker et al. 2005). Of these, 144 were available on the chip and passed quality checks.

Association analysis of SNPs in the *ROBO1* region was conducted using MERLIN (Chen and Abecasis 2007). These analyses model the additive genotypic effect at each SNP as a fixed effect in the means model. These analyses concurrently account for the relatedness among participants by explicitly modeling the covariance structure assuming an AE background genetic model. These

analyses also included adjustments for the fixed effects of age (and age squared), sex, performance IQ, and tester within the means model.

For monozygotic twin pairs the mean for the two participants was used in preference to the random selection of one member of the pair, this has the effect of increasing the test measurement reliability. Visualization and annotation of the genome-wide association results—focusing only on the *ROBO1* SNPs identified by Tagger—were performed in WGA Viewer (Ge et al. 2008). Other analyses (correlations, estimation of genotypic means) were done using R (R Development Core Team 2009).For a SNP explaining 1% of variance in our traits, under an additive model and against a background sibling correlation of ~ 0.30, we have >95% power ($\alpha = 0.05$) to detect association for a SNP with minor allele frequency above 0.05 (Purcell et al. 2003). Three SNPs had frequencies <5% (rs324755, 0.018; rs9839572, 0.018; rs11925648, 0.025). Power for these SNPs was 0.87 for a 1% effect (0.99 for a 2% effect). Effects of multiple testing are critical in association studies. Therefore significance levels were controlled using matSpD to identify independent SNPs within the *ROBO1* gene (Nyholt 2004).

Results

Descriptive

The physical locations of the SNPs (and their inter-marker linkage disequilibrium) are shown in Fig. 1.

SNPs with nominal *P* values <0.05 for association with the NWR phenotype are listed in Table 1. A schematic of chromosome 3 showing the location of *ROBO1* and the association values of all 144 tested SNPs for NWR is shown in Fig. 2. As can be seen, 21 SNPs exceeded nominal significance, with two—rs6803202 and rs4535189—reaching values surviving correction for multiple testing (P 8.70×10^{-05} and 9.30×10^{-05} respectively).

Results of association testing for digit-span forward are shown in Table 2 for the 22 SNPs reaching nominal significance with three SNPs achieving *P*-values <0.0006, including rs7653197 which was significant for NWR. Both the top two SNPs associated with NWR were also significant for digits-forward.

Finally, the top-10 SNPs for association with reading and spelling ability and with working memory are shown in Table 3. For working memory, five SNPs had nominal *P*-values < 0.05, including rs333491 with *P* of 0.004 for working memory. By contrast, for reading, only one SNP in the 144 tested was even nominally significant. No SNPs for either trait survived correction for multiple testing.

Discussion

Based on translocation and association data, and its role in bilateral asymmetry and neuronal migration, we examined the association of SNPs in *ROBO1* with theoretically motivated measures of language impairment, reading and spelling ability, and short-term verbal information storage and manipulation in a large sample representative of the general population.

Significant support was found for association of *ROBO1* polymorphisms with phonological buffer capacity peaking

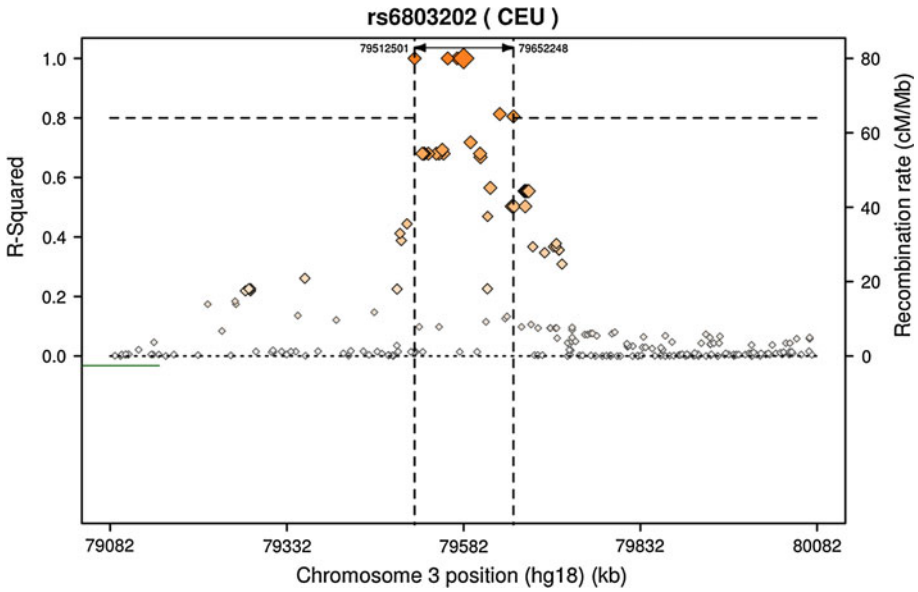

Fig. 1 *ROBO1* linkage disequilibrium plot centered on SNP rs6803202. Figure generated using SNAP (Johnson et al. 2008)

Table 1 Nominally significant associations of polymorphisms in *ROBO1* to nonword repetition

	SNP	Location	*P* value	Allele	Ancestral allele
1	rs6803202	79499153	8.70×10^{-05}	C/T	C
2	rs4535189	79489971	9.30×10^{-05}	G/A	A
3	rs7653197	79470784	0.0001	C/A	C
4	rs1387665	79429811	0.0001	G/A	G
5	rs7628757	79522372	0.0002	G/A	A
6	rs4130991	79469022	0.0002	A/G	A
7	rs4680960	79449566	0.0002	C/T	C
8	rs6548628	79569558	0.0007	A/C	A
9	rs6548621	79550373	0.0011	A/G	G
10	rs6802848	79411014	0.0014	A/G	A
11	rs7612183	79536912	0.0033	C/T	T
12	rs7356113	79753563	0.0035	A/G	G
13	rs6548651	79784416	0.0037	G/A	A
14	rs7622444	79557927	0.0038	A/G	A
15	rs7429525	79678609	0.0052	T/C	C
16	rs9853895	79585158	0.0052	T/C	C
17	rs9857859	79591307	0.0053	T/C	T
18	rs7623728	79597035	0.0160	T/C	C
19	rs7637338	79560604	0.0190	G/A	G
20	rs7614913	79567629	0.0220	C/T	C
21	rs4264688	79546348	0.0460	C/T	N/A

Note: Analyses performed in Merlin (Chen and Abecasis 2007). Trait was residualised for the effects of age, age^2, sex, interviewer, and performance IQ. *P* values as observed, uncorrected

at 8.70×10^{-05} and 9.30×10^{-05} for SNPs rs6803202 and rs4535189 respectively, values that survive correction for multiple testing. Support for association was also found for the genetically related trait of short-term storage and recall of meaningful verbal sequences (assessed by forward digit span), but not for working memory function (assessed by the letter number sequencing task). In contrast, only one SNP reached nominal significance for association with the reading and spelling ability phenotype, despite this trait and sample showing adequate power in previous studies of association to detect genetic associations for reading including for *KIAA0319* (Luciano et al. 2007), *DCDC2* (Lind et al. 2009), and *DYX1C1* (Bates et al. 2009a).

Together with the positive attributes of this study such as its size, utilization of normally-varying continuous-trait measures assessed over multiple instruments, and examination of a full suite of tagging SNPs across this large gene derived from the HapMap (Altshuler et al. 2005), and tested in a population of homogenous and appropriate ethnic origin, these findings suggest that *ROBO1* is a strong candidate for involvement in normal variation in language acquisition, and that *ROBO1* is functionally specialized to support the phonological buffer component (Baddeley et al. 1998; Gathercole 2006) of a language acquisition system. Independently, the results suggest that deficits in this system do not impact on the ability to either form the rules of decoding novel letter strings, or store the lexical entries of a written language (Bates et al. 2007a).

This conclusion is compatible with linkage studies in speech sound disorder, which implicated this region in nonword repetition (Stein et al. 2004), and with the original Nopola-Hemmi report in which again, translocation status segregated with verbal short term memory, brief phonological storage and manipulation traits (Nopola-Hemmi et al. 2001), as shown here at a molecular level for the first time, highlighting the importance of phenotypic specificity,

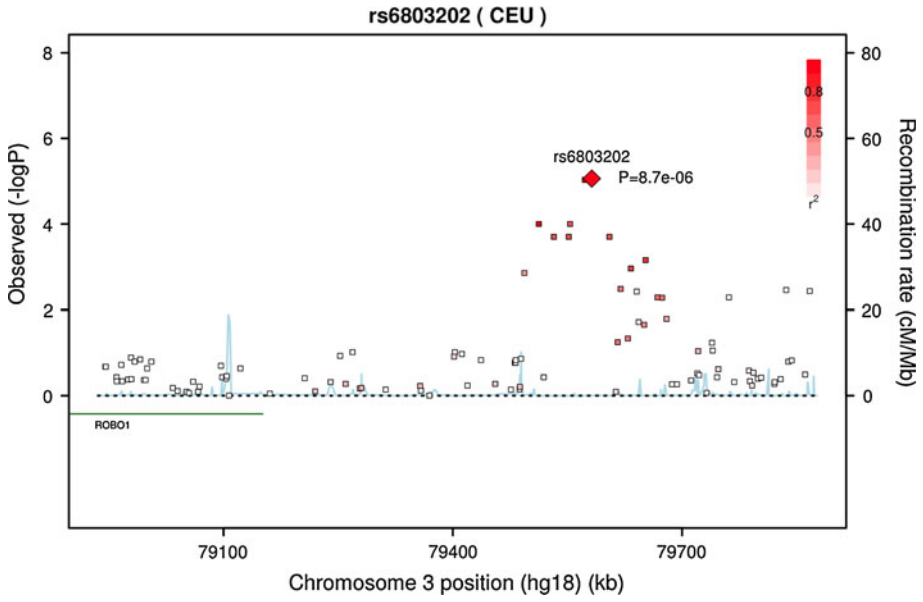

Fig. 2 Plot of association for nonword repetition. Figure generated using SNAP (Johnson et al. 2008)

Table 2 Nominally significant associations of polymorphisms in *ROBO1* to digits-forward memory span

	SNP	Location	*P*	Allele	Ancestral allele
1	rs7644521	79784534	0.0002	T/C	T
2	rs4564923	79533095	0.0054	G/A	G
3	rs7653197	79470784	0.0057	C/A	C
4	rs7629503	79813292	0.0062	G/T	G
5	rs4680960	79449566	0.0070	C/T	C
6	rs4130991	79469022	0.0074	A/G	A
7	rs4264688	79546348	0.0074	C/T	N/A
8	rs7628757	79522372	0.0078	G/A	A
9	rs4535189	79489971	0.0084	G/A	A
10	rs6803202	79499153	0.0088	C/T	C
11	rs1387665	79429811	0.0111	G/A	G
12	rs7612183	79536912	0.0124	C/T	T
13	rs7614913	79567629	0.0160	C/T	C
14	rs9849596	79756484	0.0160	A/G	G
15	rs6802848	79411014	0.0170	A/G	A
16	rs9309828	79760774	0.0250	G/A	G
17	rs6548621	79550373	0.0380	A/G	G
18	rs7623728	79597035	0.0380	T/C	C
19	rs1447833	79039406	0.0420	T/C	T
20	rs9853895	79585158	0.0440	T/C	C
21	rs9857859	79591307	0.0440	T/C	T
22	rs723766	78657774	0.0470	T/C	C

Note: Analyses performed in Merlin (Chen and Abecasis 2007). Trait was residualised for the effects of age, age^2, sex, interviewer, and performance IQ. *P* values as observed, uncorrected

Table 3 Top ten associations for reading and spelling principal component and working memory (letter number sequencing)

Reading and spelling PC		Working memory	
SNP	*P* value	SNP	*P* value
rs1995402	0.040	rs333491	**0.004**
rs2608021	0.062	rs9309819	0.021
rs1489848	0.064	rs11127636	0.023
rs12107379	0.077	rs162870	0.040
rs11127664	0.080	rs7644521	0.046
rs7431092	0.109	rs328049	0.053
rs2304503	0.113	rs1025946	0.055
rs1865862	0.130	rs7629503	0.057
rs333491	0.132	rs12054167	0.058

Note: Analyses performed in Merlin (Chen and Abecasis 2007). Trait was residualised for the effects of age, age^2, sex, interviewer, and performance IQ. *P* values as observed, uncorrected

as found in other genetic associations with complex phenotypes—for instance memory and forgetting (Bates et al. 2009b).

At a physiological level, the results suggest that while reading and spelling ability is likely dependent on local neuron-migration abnormalities as reflected in molecular layer ectopias and microgyria (Galaburda et al. 2006), neuronal migration in language is perhaps dependent on selection affecting longer range axonal guidance underpinning anatomical asymmetry, and, perhaps, related local dendritic guidance functions. The consistent finding of linkage around the region containing *ROBO1* for dyslexia, combined with the surprising absence of even nominal support for association to reading and spelling ability in the present study suggests also that there may be a second gene, related to reading, in this region.

The weaker but significant link to short-term memory, but not for reading, further supports the independence of neuronal mechanisms supporting grapheme and lexical item storage from phoneme and syllabic manipulation. The finding is also relevant for understanding pleiotropy across cognitive mechanisms. For instance we recently showed that *DYX1C1* is related to reading, but that an independent effect (tagged by SNP rs3743204) was associated with short-term memory, but not reading in the same sample (Bates et al. 2009a). This suggests that gene-level analyses may suggest pleiotropy, but that distinct functional effects of mutations in this gene may mediate this gene-level pleiotropy. In the case of *DYX1C1* it seems that TPR-domain mediated neuronal migration (Wang et al. 2006) may be related to short-term storage.

How these SNPs operate at psychological level is clearly important. The specificity of the association for nonword repetition and lack of association to reading, digits span, and letter number sequencing allow a (necessarily speculative) account of the mental mechanisms affected by *ROBO1* polymorphisms. Each of these tests of reading and language have significant perceptual and articulatory demands: nonword repetition, digit span, and letter number sequencing all depend on adequate perception of graphemic or phonemic stimuli, and the articulation of responses. While it may be that NWR s more sensitive to these processes, it seems likely that these input and output systems are not central to the role of the SNPs identified here. In specifying the Working Memory model, Baddeley was clear to distinguish the sub-system he calls the 'phonological buffer' (Baddeley et al. 1998) from other elements of working memory (such as the central executive). Moreover the NWR task was designed by Gathercole and Baddeley precisely to be a sensitive indicator of the function of this phonological loop as a brief store of meaningless phonemes, independent of lexical support and the integrity of rehearsal mechanisms (Gathercole 2006). It is therefore likely that *ROBO1* SNPs identified here act primarily to reduce the size of the phonological buffer, without impacting significantly on rehearsal, lexical (as

opposed to phoneme-level) storage or activation), executive tasks, or more peripheral perceptual or articulatory systems.

Finally, as the present findings are derived from normally varying language and short-term and working memory ability, they have additional implications for the etiology and biology of cognitive disability. The argument (Bates et al. 2007a; Shaywitz et al. 1992) that dyslexia represents the low-tail of a normal distribution of reading ability in the population is strengthened, as supported by the utility of normal samples in genetic research on reading ability and dyslexia in linkage (Bates et al. 2007b) and association (Bates et al. 2009a; Luciano et al. 2007; Paracchini et al. 2008) studies. The implication then is that the *ROBO1* influences both poor (Hannula-Jouppi et al. 2005; Stein et al. 2004) and normally varying language ability, as shown here.

These results provide the first support for *ROBO1* as a gene involved in a core trait underpinning acquisition of language, with a specific function in supporting a short-term buffer for arbitrary phonological strings. Reduced effects of *ROBO1* were observed for the processing of verbal meaningful strings and working with the stored contents of memory. These effects of *ROBO1* appear to be unrelated to brain mechanisms underpinning reading ability, at least by adolescence. The results strongly support *ROBO1* as the first gene discovered to be associated with language deficits affecting normal variation in language ability. Its functional role in neuronal migration underlying bilateral symmetry and lateralization of neuronal function suggests a role in the evolution of human language ability. Additional research is warranted to identify the specific mechanism of this association and to identify the basis of these effects including pleiotropy for short-term verbal memory.

Acknowledgments We wish to thank the twins and their parents for their co-operation; Anjali Henders and Megan Campbell for managing sample processing and DNA extraction; Alison Mackenzie for coordinating the reading test battery mail-out; Thanuja Gunasekera for coordinating the language test battery mail-out; Marlene Grace, Ann Eldridge and the research interviewers for data collection. We thank The Royal Society of Edinburgh, The Office of the Chief Scientist of Scotland, The Australian Research Council and National Health and Medical Research Council and for present and past support of this research. ML is a Royal Society of Edinburgh/Lloyds TSB Foundation for Scotland Personal Research Fellow.

References

Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P (2005) A haplotype map of the human genome. Nature 437:1299–1320

Andrews W, Liapi A, Plachez C, Camurri L, Zhang J, Mori S et al (2006) Robo1 regulates the development of major axon tracts and interneuron migration in the forebrain. Development 133(11):2243–2252

Andrews W, Barber M, Hernadez-Miranda LR, Xian J, Rakic S, Sundaresan V et al (2008) The role of Slit-Robo signaling in the generation, migration and morphological differentiation of cortical interneurons. Dev Biol 313(2):648–658

Anitha A, Nakamura K, Yamada K, Suda S, Thanseem I, Tsujii M et al (2008) Genetic analyses of roundabout (ROBO) axon guidance receptors in autism. Am J Med Genet B Neuropsychiatr Genet 147B(7):1019–1027

Baddeley A, Gathercole S, Papagno C (1998) The phonological loop as a language learning device. Psychol Rev 105(1):158–173

Barber M, Di Meglio T, Andrews WD, Hernandez-Miranda LR, Murakami F, Chedotal A et al (2009) The role of Robo3 in the development of cortical interneurons. Cereb Cortex 19(Suppl 1): i22–i31

Bates TC, Castles A, Coltheart M, Gillespie N, Wright M, Martin NG (2004) Behaviour genetic analyses of reading and spelling: a component processes approach. Aust J Psychol 56(2):115–126

Bates TC, Castles A, Luciano M, Wright M, Coltheart M, Martin N (2007a) Genetic and environmental bases of reading and spelling: a unified genetic dual route model. Read Writ 20(1–2):147–171

Bates TC, Luciano M, Castles A, Coltheart M, Wright MJ, Martin NG (2007b) Replication of reported linkages for dyslexia and spelling and suggestive evidence for novel regions on chromosomes 4 and 17. Eur J Hum Genet 15(2):194–203

Bates TC, Luciano M, Wright M, Montgomery GW, Martin NG (2007c) Genetic architecture of language: phonological and lexical storage capacity deficits in specific language impairment & dyslexia. Paper presented at the 37th annual meeting of the behavior genetics association, Amsterdam, The Netherlands, 3–6 June 2007

Bates TC, Luciano M, Wright MJ, Montgomery GW, Martin NG (2007d) Heritability of phonological storage deficits related to specific language impairment in reading ability. Paper presented at the Behavior Genetics Association 37th Annual Meeting, Amsterdam. Behav Genet 37:734–809

Bates TC, Luciano M, Lind PA, Wright M, Montgomery GW, Martin NG (2008) Recently-derived variants of brain-size genes ASPM, MCPH1, CDK5RAP and BRCA1 not associated with general cognition, reading or language. Intelligence 36:689–693

Bates TC, Lind PA, Luciano M, Montgomery GW, Martin NG, Wright MJ (2009a) Dyslexia and DYX1C1: deficits in reading and spelling associated with a missense mutation. Mol Psychiatry. doi:10.1038/mp.2009.120

Bates TC, Price JF, Harris SE, Marioni RE, Fowkes FG, Stewart MC et al (2009b) Association of KIBRA and memory. Neurosci Lett 458(3):140–143

Benyamin B, McRae AF, Zhu G, Gordon S, Henders AK, Palotie A et al (2009) Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. Am J Hum Genet 84(1):60–65

Box GEP, Cox DR (1964) An analysis of transformations. J R Stat Soc B 26:211–246

Castles A, Coltheart M (1993) Varieties of developmental dyslexia. Cognition 47(2):149–180

Chen WM, Abecasis GR (2007) Family-based association tests for genomewide association scans. Am J Hum Genet 81(5):913–926

Cope NA, Hill G, van den Bree M, Harold D, Moskvina V, Green EK et al (2005) No support for association between Dyslexia Susceptibility 1 Candidate 1 and developmental dyslexia. Mol Psychiatry 10(3):237–238

Dahdouh F, Anthoni H, Tapia-Paez I, Peyrard-Janvid M, Schulte-Korne G, Warnke A et al (2009) Further evidence for DYX1C1 as a susceptibility factor for dyslexia. Psychiatr Genet 19(2):59–63

de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D (2005) Efficiency and power in genetic association studies. Nat Genet 37(11):1217–1223

Dediu D, Ladd DR (2007) Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. Proc Natl Acad Sci USA 104(26): 10944–10949

Dollaghan C, Campbell TF (1998) Nonword repetition and child language impairment. J Speech Lang Hear Res 41(5):1136–1146

Fisher SE, Francks C, Marlow AJ, MacPhie IL, Newbury DF, Cardon LR et al (2002) Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. Nat Genet 30(1):86–91

Galaburda AM, LoTurco J, Ramus F, Fitch RH, Rosen GD (2006) From genes to behavior in developmental dyslexia. Nat Neurosci 9(10):1213–1217

Gathercole SE (2006) Nonword repetition and word learning: The nature of the relationship. Appl Psycholinguist 27(4):513–543

Gathercole SE, Willis CS, Baddeley AD, Emslie H (1994) The Children's Test of Nonword Repetition: a test of phonological working memory. Memory 2(2):103–127

Ge D, Zhang K, Need AC, Martin O, Fellay J, Urban TJ et al (2008) WGAViewer: software for genomic annotation of whole genome association studies. Genome Res 18(4):640–643

Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmi J, Kaariainen H et al (2005) The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. PLoS Genet 1(4):50

Jackson DN (1984) Multidimensional aptitude battery manual, vol 1. Research Psychologists Press, Port Huron, MI

Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 24(24):2938–2939

Kovas Y, Plomin R (2006) Generalist genes: implications for the cognitive sciences. Trends Cogn Sci 10(5):198–203

Lind P, Luciano M, Duffy D, Castles A, Wright MJ, Martin NG et al (2009) Dyslexia and DCDC2: normal variation in reading and spelling is associated with DCDC2 polymorphisms in an Australian population sample. Eur J Hum Genet 18(6):668–673

LoTurco JJ, Bai J (2006) The multipolar stage and disruptions in neuronal migration. Trends Neurosci 29(7):407–413

Luciano M, Lind PA, Duffy DL, Castles A, Wright MJ, Montgomery GW et al (2007) A haplotype spanning KIAA0319 and TTRAP is associated with normal variation in reading and spelling ability. Biol Psychiatry 62(7):811–817

McGregor B, Pfitzner J, Zhu G, Grace M, Eldridge A, Pearson J et al (1999) Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. Genet Epidemiol 16(1):40–53

Meng H, Hager K, Held M, Page GP, Olson RK, Pennington BF et al (2005) TDT-association analysis of EKN1 and dyslexia in a Colorado twin cohort. Hum Genet 118(1):87–90

Newbury DF, Bonora E, Lamb JA, Fisher SE, Lai CS, Baird G et al (2002) FOXP2 is not a major susceptibility gene for autism or specific language impairment. Am J Hum Genet 70(5): 1318–1327

Nopola-Hemmi J, Taipale M, Haltia T, Lehesjoki AE, Voutilainen A, Kere J (2000) Two translocations of chromosome 15q associated with dyslexia. J Med Genet 37(10):771–775

Nopola-Hemmi J, Myllyluoma B, Haltia T, Taipale M, Ollikainen V, Ahonen T et al (2001) A dominant gene for developmental dyslexia on chromosome 3. J Med Genet 38(10):658–664

Nyholt DR (2004) A simple correction for multiple testing for SNPs in linkage disequilibrium with each other. Am J Hum Genet 74(4):765–769

Olson RK, Byrne B (2005) Genetic and environmental influences on reading and language ability and disability. In: Catts HW, Kamhi AG (eds) The connections between language and reading disabilities. Lawrence Erlbaum Associates, Mahwah, NJ, pp 173–200

Paracchini S, Scerri T, Monaco AP (2007) The genetic lexicon of dyslexia. Annu Rev Genomics Hum Genet 8:57–79

Paracchini S, Steer CD, Buckingham LL, Morris AP, Ring S, Scerri T et al (2008) Association of the KIAA0319 dyslexia susceptibility gene with reading skills in the general population. Am J Psychiatry 165(12):1576–1584

Pennington BF, Lefly DL (2001) Early reading development in children at family risk for dyslexia. Child Dev 72(3):816–833

Plomin R, Kovas Y, Haworth CMA (2007) Generalist genes: genetic links between brain, mind, and education. Mind Brain Educ 1(1):11–19

Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19(1):149–150

Schousboe K, Willemsen G, Kyvik KO, Mortensen J, Boomsma DI, Cornes BK et al (2003) Sex differences in heritability of BMI: a comparative study of results from twin studies in eight countries. Twin Res 6(5):409–421

Schumacher J, Anthoni H, Dahdouh F, Konig IR, Hillmer AM, Kluck N et al (2006) Strong genetic evidence of DCDC2 as a susceptibility gene for dyslexia. Am J Hum Genet 78(1):52–62

Shaywitz SE, Escobar MD, Shaywitz BA, Fletcher JM, Makuch R (1992) Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. N Engl J Med 326(3): 145–150

SLI Consortium (2004) Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. Am J Hum Genet 74(6):1225–1238

Stein CM, Schick JH, Gerry Taylor H, Shriberg LD, Millard C, Kundtz-Kluge A et al (2004) Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. Am J Hum Genet 74(2):283–297

R Development Core Team (2009) R: a language and environment for statistical computing: R Foundation for Statistical Computing. Vienna, Austria

Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M et al (2008) A functional genetic link between distinct developmental language disorders. N Engl J Med 359(22): 2337–2345

Wang Y, Paramasivam M, Thomas A, Bai J, Kaminen-Ahola N, Kere J et al (2006) DYX1C1 functions in neuronal migration in developing neocortex. Neuroscience 143(2):515–522

Wechsler D (1997) Wechsler adult intelligence scale III. Psychological Corporation, San Antonio

Wright MJ, De Geus E, Ando J, Luciano M, Posthuma D, Ono Y et al (2001) Genetics of cognition: outline of a collaborative twin study. Twin Res 4(1):48–56