

ORIGINAL ARTICLE

The *DCDC2* deletion is not a risk factor for dyslexia

TS Scerri¹, E Macpherson², A Martinelli², WC Wa³, AP Monaco⁴, J Stein⁵, M Zheng⁶, C Suk-Han Ho⁶, C McBride⁷, M Snowling⁸, C Hulme⁹, ME Hayiou-Thomas¹⁰, MMY Waye^{3,11}, JB Talcott¹² and S Paracchini²

Dyslexia is a specific impairment in learning to read and has strong heritability. An intronic deletion within the *DCDC2* gene, with ~8% frequency in European populations, is increasingly used as a marker for dyslexia in neuroimaging and behavioral studies. At a mechanistic level, this deletion has been proposed to influence sensory processing capacity, and in particular sensitivity to visual coherent motion. Our re-assessment of the literature, however, did not reveal strong support for a role of this specific deletion in dyslexia. We also analyzed data from five distinct cohorts, enriched for individuals with dyslexia, and did not identify any signal indicative of associations for the *DCDC2* deletion with reading-related measures, including in a combined sample analysis ($N = 526$). We believe we conducted the first replication analysis for a proposed deletion effect on visual motion perception and found no association ($N = 445$ siblings). We also report that the *DCDC2* deletion has a frequency of 37.6% in a cohort representative of the general population recruited in Hong Kong ($N = 220$). This figure, together with a lack of association between the deletion and reading abilities in this cohort, indicates the low likelihood of a direct deletion effect on reading skills. Therefore, on the basis of multiple strands of evidence, we conclude that the *DCDC2* deletion is not a strong risk factor for dyslexia. Our analyses and literature re-evaluation are important for interpreting current developments within multidisciplinary studies of dyslexia and, more generally, contribute to current discussions about the importance of reproducibility in science.

Translational Psychiatry (2017) **7**, e1182; doi:10.1038/tp.2017.151; published online 25 July 2017

INTRODUCTION

Dyslexia, or reading disability (RD), is a common neurodevelopmental disorder, with prevalence estimates ranging between 5–10% among school-age children in most countries.^{1,2} While a significant genetic component has been firmly established for this condition, only a small number of specific genetic factors have been identified to date.³ These include variants within the *DCDC2* gene, located on chromosome 6 (at 6p22.3). Single-nucleotide polymorphisms (SNPs) within *DCDC2* showed association with both dyslexia and reading abilities in independent studies,^{4–10} but the functional mechanisms underlying these associations remain unclear. Attention has increasingly focussed on an intronic deletion of 2,445 bp within *DCDC2* as a potential functional element (Table 1).

The original study describing the *DCDC2* deletion reported associations between this locus and quantitative measures.⁴ However, such associations were detected only when the deletion was analyzed in combination with a pool of 10 rare alleles of a short tandem repeat (STR; BV677278) marker at the same locus.⁴ Analyses of the deletion combined with the STR alleles, failed to replicate the association with dyslexia as a binary category.^{13,14} A separate study reported an association with a composite memory score, but the association appeared to be driven by one particular STR allele (allele 10 with frequency of 5%) rather than the deletion itself.¹⁵ This STR has been suggested to be a regulatory element.^{11,16} Powers *et al.* (2013)¹¹ reported that two 6-SNP haplotypes (allele frequency < 5%) were associated with reading

and language phenotypes in a large epidemiological sample (the Avon Longitudinal Study of Parents and Children (ALSPAC)) stratified for a severe phenotype (≥ 2 s.d. below the mean on a phoneme-deletion task). This stratification resulted in the comparison of 89 cases with the remaining 5,225 individuals who scored above the -2 s.d. diagnostic threshold. These two rare haplotypes were in linkage disequilibrium (LD) with alleles 5 and 6 of the STR. In a follow up study, allele 5 (3.6% frequency) and 6 (4.7% frequency), but not the deletion, showed association with reading and language measures.¹² Meng *et al.* (2011)¹⁶ showed that STR alleles 3, 4 and 5 have differential effects on gene expression regulation. In summary, there is no clear consensus as to which STR alleles are associated with dyslexia and regulate gene expression. Rare alleles, such as the STR variants, can increase the likelihood of false positives because their analysis is typically conducted in very small samples.¹⁷ Up to 40 alleles have been described for the STR,^{11,13} and therefore it is expected that some alleles would lead to significant P -values purely by chance.

The deletion has also been analyzed without the inclusion of the STR marker. Weak associations have been reported for various reading-related traits with significance values in the range of $0.02 < P < 0.05$ ^{14,15,18} (Table 1; Supplementary Table 1). Lack of replication has also been reported.^{12,14} Wilcke *et al.* (2009) reported the most compelling association to date ($P < 0.01$) in a case/control study of a German cohort (Table 1). In this sample, the deletion had a frequency of 10% in cases (8% in severe cases) and 4% in controls. This lower frequency in the controls, however,

¹Walter and Eliza Hall Institute of Medical Research & Melbourne University, Parkville, VIC, Australia; ²School of Medicine, University of St Andrews, St Andrews, UK; ³School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong; ⁴Wellcome Trust centre for Human Genetics, University of Oxford, Oxford, UK; ⁵Department of Physiology, University of Oxford, Oxford, UK; ⁶Department of Psychology, The University of Hong Kong, Hong Kong, Hong Kong; ⁷Department of Psychology, Shatin, NT, The Chinese University of Hong Kong, Shatin, Hong Kong; ⁸Department of Experimental Psychology and St John's College, University of Oxford, Oxford, UK; ⁹Department of Education, University of Oxford, Oxford, UK; ¹⁰Department of Psychology, University of York, York, UK; ¹¹The Nethersole School of Nursing, The Chinese University of Hong Kong, Shatin, Hong Kong and ¹²Aston Brain Centre, School of Life and Health Sciences, Aston University, Birmingham, UK. Correspondence: Dr S Paracchini, School of Medicine, University of St Andrews, North Haugh St. Andrews KY16 9TF, UK.

E-mail: sp58@st-andrews.ac.uk

Received 26 January 2017; revised 9 May 2017; accepted 13 June 2017

Table 1. Previous studies that analyzed the DCDC2 deletion for association with dyslexia

Reference	Country	Sample	Type of analysis	Allele	Quantitative phenotype	P-value	Deletion Frequency
Meng et al., 2005	USA	153 nuclear families	Quantitative	Deletion+STR rare alleles	Homonym Choice Untimed PIAT Word Recognition Word Recognition OC OC	0.00002 0.023 0.04 0.0035 0.0385	8.5%
Harold et al., 2006	England	264 nuclear families 126 severe families subgroup	Quantitative	Deletion	IWR READ PA	Not associated 0.03 0.04 0.05	10% 9%
	Wales	350 cases 273 controls	Categorical	Deletion Deletion+STR rare alleles		Not associated Not associated	9% cases 9% controls
Brkanac et al., 2007	USA	191 trios	Quantitative	Deletion	SWE Spelling	0.05 0.04	7.5%
Ludwig et al., 2008	Germany	396 trios	Categorical	Deletion+STR rare alleles		Not associated	8.6%
Wilcke et al., 2009	Germany	72 cases 25 severe cases 184 controls	Categorical	Deletion		$P < 0.01$	10% cases 8% severe cases 4% controls
Marino et al., 2012	Italy	303 nuclear families	Quantitative	Deletion Deletion STR allele 10 Deletion+STR rare alleles	SWR PD Memory Composite Score Memory Composite Score	0.05 0.04 0.00001 0.00001	6%
Powers et al., 2016	UK	89 cases 4428 controls ^a	Categorical	Deletion	Severe RD Severe LI	Not associated	8.3%

Abbreviations: IWR, irregular word reading; LI, language impairment; OC, orthographic coding; PA, phonological awareness; PD, non-word reading; PIAT, Peabody individual achievement test; READ, single-word reading; RD, reading disability; SWE, sight word efficiency. ^aThe sample sizes were estimated by combining information from different studies.^{1,1,2} these studies reported association between the 5 and 6 STR alleles with reading-related measures.

Table 2. Frequency of the *DCDC2* deletion reported in this study

Cohort	Structure	Recruitment	<i>DCDC2</i> deletion frequency
Oxford Family Dyslexia	219 families	Clinical—children with reading difficulties	8.9% parents (<i>N</i> = 414)
			9.8% probands (<i>N</i> = 153) ^a
Oxford Cases Dyslexia	272 singletons	Clinical—children with reading difficulties	6.1%
Aston Dyslexia	105 singletons	Clinical—children with reading difficulties	7.6%
York	103 families	Longitudinal—children with language difficulties, family history of dyslexia and typically developing	7.4% parents (<i>N</i> = 174)
			7.7% all probands (<i>N</i> = 103)
			5.5% RD group (<i>N</i> = 18)
			10.7% TD group (<i>N</i> = 72)
Combined	203 singletons	Individuals scoring READ < -1.5s.d. in the above cohorts	9.6%
Hong Kong	220 singletons/twins	Typically developing	37.6%
1000 Genomes	503 singletons	General population	8.05% Europe
	504 singletons		33.5% East Asia

Abbreviations: RD, reading difficulties; TD, typically developing. ^aThe number of probands is lower than the number of families because of missing data.

is not consistent with what is observed in other studies of European samples. In fact, the deletion had a frequency of 8.6% among the parents of probands with dyslexia in a different German sample.¹³ A frequency of 9% was reported for both cases with dyslexia and controls in a UK cohort,¹⁴ similar to the 8.3% frequency reported in a UK epidemiological sample.¹²

Taken together, these studies do not support a role of the *DCDC2* deletion as a specific risk factor for dyslexia. Nevertheless, the *DCDC2* deletion is becoming increasingly used in research aimed at identifying the neuronal correlates of dyslexia. For example, the deletion has been suggested to directly affect the development of brain structures relevant to language and reading, through quantitative imaging studies of anatomical variability in healthy individuals.^{19,20} The deletion has also been linked to functional mechanisms, such as reduced sensitivity to visual coherent motion.^{21,22} Perception of motion is an important component of visual development, for which impairment has been demonstrated to occur in several developmental disorders,²³ including dyslexia.²⁴ Observed associations between visual motion sensitivity and reading skills contributed to the development of the magnocellular hypothesis for dyslexia.²⁵ However, motion perception deficits may be restricted to subgroups of individuals with dyslexia, and typically those with the most severe phenotypes who might also present comorbidities with other developmental disorders.²⁶ Cicchini *et al.*²² reported that a decreased ability to detect and discriminate visual motion correlated with the presence of the *DCDC2* deletion and suggested that motion sensitivity was more strongly associated with the presence of the deletion than with the presence of dyslexia. A major limitation of these imaging and behavioral studies, however, is the very small sample size, which suffers from both low power of detecting real effects as well as the increased chance of generating false positives.¹⁷

The aim of the present study was to provide additional data to evaluate the role of the *DCDC2* deletion as a contributor to risk for dyslexia and impairment on related measures. In addition to a re-evaluation of the published literature (Table 1), we report new association analyses for the deletion in four distinct cohorts

characterized for reading measures, including, for the first time, to our knowledge, a cohort of Asian origin. We also present a replication analysis for association between the *DCDC2* deletion and a measure of visual motion sensitivity. The evidence we have gathered does not support an association between the *DCDC2* deletion and dyslexia or quantitative measures of reading abilities. This work is important for the interpretation of a developing body of literature and for framing the direction of future research studies of dyslexia. Moreover, our results are relevant to the current discussion about the importance of reproducibility in science,²⁷ including the observation that it may be very easy to canonize weak findings as scientific facts unless a substantial number of negative findings are published.²⁸

MATERIALS AND METHODS

Study participants

Our analysis included four distinct UK cohorts, referred to here as the Oxford Family Dyslexia cohort, Oxford Cases Dyslexia Cohort, Aston Dyslexia cohort, York cohort, and a fifth cohort, the Hong Kong cohort from the Chinese-English Twin Study of Bilingualism (Table 2).

The Oxford Family Dyslexia cohort was recruited by research clinics in Oxford and Reading. Genotype and phenotype data were available for 219 families (445 siblings, age range 6–27 years). Parental genotypes were also available. The same research clinics recruited the Oxford Cases Dyslexia Cohort, a collection of 272 unrelated individuals (age range 8–18 years). The psychometric measures used to assess these cohorts included the British Ability Scales (BAS) single-word reading (READ) and spelling (SPELL) tests,²⁹ irregular word reading (OC-irreg)³⁰ and phonological decoding (PD).³⁰ An orthography measure using the forced choice task (OC-choice)³¹ and phoneme awareness (PA)³² scores were available for the Family cohort. Both cohorts have been described before³³ and the Family cohort was also previously analyzed for association between the *DCDC2* deletion and reading abilities.¹⁴ Here, we analyzed the Oxford Family cohort for association with a visual motion (VMOT) measure described previously³⁴ (See also Supplementary Materials for details) and as part of a combined sample (see below).

The Aston Dyslexia cohort includes 105 unrelated individuals, recruited through the Aston Dyslexia and Developmental Assessment Unit in Birmingham, UK (age range 7–16 years). Children attending this

educational clinic have been referred for assessments of reading difficulties by local schools. On the basis of their educational assessments, children are invited back to participate in research studies. The phenotypes obtained for this cohort included single-word reading (READ) and single-word spelling (SPELL), measured by standardised assessments including using the Wechsler Individual Achievement Test,³⁵ the British Abilities Scales, and the Test of Word Reading Efficiency (TOWRE).³⁶ Sub-scales of TOWRE were used to measure pseudo-word reading (or phonological decoding, TOWRE_PD) and sight word efficiency (TOWRE-SWE).

The York cohort is a longitudinal cohort designed to study language and literacy development³⁷ and has been analyzed previously for genetic studies.³⁸ DNA was available for 103 families and a total of 318 individuals. Of these, phenotype data were available for $N=103$ probands and $N=41$ siblings. Phenotypes were available from the same children at different time points. Reading abilities were measured at the ages of 4½, 5½, 6 and 8 years, representing the years when reading skills are in ascendancy. At age 5½ years, the measure was word reading from the York Assessment of Reading Comprehension (READ);³⁹ at age 6 years, the measures were word reading efficiency from the Test of Word Reading Efficiency Word (TOWRE_READ) and Pseudoword (TOWRE_PD) tests³⁶ and spelling from the Wechsler Individual Achievement Test (WIAT-SPELL);³⁵ at age 8 years the measures were irregular word reading (IWR) from the Diagnostic Test of Word Reading Processes (DTWRP),⁴⁰ phonological awareness measured by a phoneme-deletion task (PA) and rapid automatic naming (RAN).⁴¹ The factor scores were latent variables derived following confirmatory factor analysis: a phonology factor score (PHON_FS) was derived from word recall and phoneme awareness tasks at 4½ years and from phoneme awareness (deletion) at ages 5½, 6 and 8 years; a literacy factor score (LIT_FS) was derived from measures of single-word reading, reading accuracy, word and non-word reading fluency and spelling at age 8 years. At age 8 years, children were assigned to a formal diagnosis of dyslexia when they scored >1.5 s.d. below the control group mean on the literacy composite measure. This approach identified $N=18$ children with dyslexia, including $N=9$ who also had language impairment (LI). Another $N=11$ had LI only and $N=72$ children were classified as typical developing (TD).

Exclusion criteria for the UK cohorts were non-European ethnicity, signs of other neurological conditions in the probands or performance IQ < 80 .

The Hong Kong cohort included 220 unrelated individuals who were not selected for reading difficulties (age range 6–10.5 years old).⁴² Study participants were recruited either as singletons ($N=104$) or as twin pairs with only one twin per pair retained for the study ($N=116$). All participants speak Cantonese as their first language and have not been diagnosed with either intellectual disabilities or neurological conditions. Three reading-related phenotypes were available from a standardized test battery described previously: Chinese Word Reading (CWR), Chinese one-minute (COM) word reading and Chinese Digit Rapid Automatized Naming (CDRAN).⁴³ Mean scores are shown in Supplementary Table 2.

DNA was extracted from blood or buccal swabs using standard procedures, and from saliva using Oragene kit (prepITL2P) (DNA Genotek, Ottawa, Canada). Ethical approval was obtained from the local ethical research committees, including the Oxfordshire Psychiatric Research Ethics Committee, the NHS Research Ethics (Yorkshire & The Humber – Humber Bridge), the University of York Department of Psychology Ethics Committee and the Aston University Ethics Committee, and the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee. Written informed consent was obtained from all participants and/or their caregivers.

Genotype and statistical analysis

Genotype data were generated using a previously described protocol⁴ and checked for Mendelian errors and deviation from Hardy-Weinberg equilibrium using PEDSTATS.⁴⁴ Family-based cohorts were tested using the Quantitative Transmission Disequilibrium Test (QTDT) under the total association model⁴⁵ after estimating the identical-by-descent (IBD) sharing with MERLIN.⁴⁶ No population stratification was observed for the UK cohorts following analysis of genome-wide genotype data.^{47,48} Singletons analysis was conducted with PLINK (Purcell et al., 2007) under an allelic model in each cohort as well as in combined samples including the Oxford Cases cohort, the Aston cohort and the probands of the Oxford Family cohort. Analysis was also conducted in a subgroup derived from this combined sample for meeting a stringent criteria of dyslexia defined by a score below -1.5 s.d. from the normative population mean for a single-word reading test. Power calculations were conducted with the genetic Power Calculator⁴⁹ (Supplementary Material).

RESULTS

The DCDC2 deletion and reading abilities

The frequency of the deletion ranged 6.1–9.8% in the UK cohorts (Table 2), which is consistent with the range reported previously, including that for UK controls (9%; Table 1). Differences in frequency across the cohorts and in specific subgroups likely resulted from sampling bias and illustrates how small samples can potentially affect this type of analysis. For example, the longitudinal design of the York cohort enabled the identification of children with dyslexia on the basis of very detailed assessments. The overall frequency of the deletion, across the 103 probands was 7.7%, similar to that observed in the parents (7.4%, $N=174$). On the basis of the final diagnosis, the deletion had a frequency of 10.7% ($N=72$) in the TD group compared to 5.5% in the RD ($N=18$) group (Table 2). The deletion frequency in a combined group of individuals ($N=203$) from the UK cohorts meeting a stringent definition of dyslexia (READ < -1.5 s.d. from the standardised population mean) was 9.6%. Strikingly, the deletion had a frequency of 37.6% ($N=220$; Table 2) in the Hong Kong cohort which was sampled from a population with typical literacy achievement (Supplementary Table 2). Given the substantial difference in allele frequency between the UK and Hong Kong population we queried the data from the 1000 Genomes project.⁵⁰ The deletion (ID=Esv3608367) presents a global frequency of 11.5% across 2504 individuals, and consistently with our data, shows a frequency of 8.05% and 33.53% in the European and East Asian populations (Table 2; Supplementary Table 3). The frequency observed in the European population is important in interpreting the results by Wilcke et al. (2009) that were based on comparisons with a control cohort of German origin that presented a deletion frequency of 4% (Table 1).

The DCDC2 deletion was tested for association with quantitative reading-related measures in the Oxford Cases, Aston, York and Hong Kong cohorts. The DCDC2 deletion was tested for association with reading measures in the Oxford Family Dyslexia cohort previously;¹⁴ no associations were detected in the entire cohort, and only marginally significant associations were observed in the subgroup of 126 families selected for severity of the phenotype (Supplementary Methods and Supplementary Table 1). We then conducted quantitative association analyses in the combined sample including individuals with directly comparable phenotypes. We used a combined sample including probands from the Oxford Family cohorts and the Oxford Cases ($N=425$) which had four comparable phenotypes. We then analyzed this sample with the addition of the Aston cohort ($N=526$) which was characterised with the same READ and SPELL phenotypes. In both combined samples we analyzed subgroups of individuals who met a stringent definition of dyslexia defined by a score below -1.5 s.d. from the standardised mean for READ of the population. No association was detected with any of the phenotypes tested (Supplementary Table 4).

The DCDC2 deletion and visual motion

The distribution of VMOT thresholds (threshold = $1/\text{sensitivity}$) approximated normality across $N=701$ individuals (including individuals for whom genotype data were not available) with a mean of 18.1 (s.d. = 12.2) in the Oxford Dyslexia cohort. VMOT thresholds showed a small, but statistically significant correlation with READ ($r = -0.16$; one-tailed $P = 1.0 \times 10^{-5}$), indicating that lowered motion sensitivity is associated with poorer reading scores.

We did not identify patterns of association between the DCDC2 deletion and VMOT thresholds, either in the whole cohort or in the same subset of 126 families selected for severity as described previously¹⁴ (Supplementary Table 1 and Supplementary Material). The subgroup analysis has the limitation of reducing the power of the analysis by decreasing the sample size, but was aimed at

Table 3. Mean score of READ and VMOT stratified by the *DCDC2* deletion genotype in the Oxford dyslexia cohort

Number of deletion alleles	READ (s.d.;N)		VMOT % (s.d.;N)	
	All	Probands	All	Probands
0	46.2 (9.7; 417)	38.6 (6.8; 126)	18.8 (12.0; 374)	20.7 (13.5; 112)
1	45.0 (10.8; 73)	36.6 (7.6; 24)	17.5 (9.8; 66)	21.3 (14.2; 22)
2	46.4 (14.5; 7)	36.7 (8.7; 3)	16.7 (6.9; 6)	19.2 (5.8; 3)

Abbreviations: READ, single-word reading; VMOT, visual motion. The READ scores are standardized with a population mean of 50 and s.d. = 10. VMOT scores are in % coherent motion at detection threshold. Poorer reading is represented by lower standard scores; reduced sensitivity to VMOT is indicated by higher scores. See Supplementary Figure 1 for the distribution of the raw scores.

addressing the alternative hypothesis that the effect of the deletion on visual motion abilities is restricted to individuals within the lower tail of the normal distribution of reading ability.²⁶

We stratified the READ and VMOT scores for the *DCDC2* deletion genotypes across all siblings and for the probands only (Table 3). As expected, the mean scores for the probands showed lower performance for both READ and VMOT compared to the whole data set, but their raw scores were similarly distributed across the three genotype groups (Supplementary Figure 1). Accordingly, genetic association analysis run under different linear models (dominant, recessive and genotypic conducted using PLINK) did not show any trends of association (data not shown).

DISCUSSION

With this study, we aimed to investigate the role of the *DCDC2* deletion in contributing to dyslexia or measures of reading and component skills. Our work was prompted by recent studies that based their hypotheses and research designs on the assumption that the *DCDC2* deletion is an established susceptibility factor for dyslexia. However, as our re-evaluation of the literature shows, the accumulated evidence in support of this assumption is statistically weak (Table 1). Most prior studies either did not show association with the deletion^{4,12–14} or showed only moderate signals of association ($0.01 < P\text{-values} < 0.5$)^{15,18,51} (Table 1). This latter group included also our previous study that reported a weak association signal ($P=0.04$) with one phenotype (IWR) in a subgroup of families (Supplementary Table 1). Only one study reported a significantly higher frequency of the deletion in cases with dyslexia (10%) compared to controls (4%).⁷ However, through comparisons with other studies^{12–14} and with the data derived from the 1000 Genome project for European reference populations (Supplementary Table 3), we suggest that the low frequency reported in the control cohort analyzed by Wilcke *et al.*⁷ most likely resulted from an artifact. Stronger association signals were observed only when the deletion was combined with different rare STR alleles.^{4,12,15} We did not test the STR in this study because our focus was to address the role of the *DCDC2* deletion specifically as a reliable marker for dyslexia. Our goal was to verify the theoretical foundation of the design of imaging^{19,20} and behavioral^{21,22} studies for which the deletion alone, and not in combination with STR alleles, has been used as a dyslexia marker. These studies suggested that the deletion is directly associated with structural brain changes and sensory deficits regardless of the presence of dyslexia, implying a causative effect for this genetic marker. Furthermore, the STR harbors at least 40 (mostly extremely rare) alleles, making it challenging to distinguish genuine signals from noise and to test specific hypotheses. In fact, associations so far have been reported for different STR alleles (i.e. deletion+pool of rare alleles, allele 10, allele 5 and allele 6; Table 1).

We tested the deletion for association with reading abilities in four cohorts from the UK analysed both separately and as

combined samples, including analysis in a subgroup meeting stringent criteria for dyslexia. The Oxford Family Dyslexia cohort, analyzed previously,¹⁴ was re-analyzed here as part of this approach and for association with a visual motion test. Furthermore, we tested a cohort recruited in Hong Kong, which represents the first cohort of non-European ancestry to be analyzed for this marker in the context of reading abilities. None of our analyses showed any statistically (or nominally) significant associations between the deletion and reading measures. Power calculations indicate that our largest sample used for genetic analysis ($N=526$) has the power to detect only relatively large effects ($>1.5\%$ of the phenotypic variance; Supplementary Material). Although this is larger than what would generally be predicted for common genetic factors contributing to complex traits, our data are important for the interpretation of the current literature. The analysis in the Hong Kong cohort revealed a frequency of 37.6% which is consistent with frequency data reported for the 1000 Genomes Project Asian populations. The Hong Kong cohort is representative of the general Hong Kong population of Chinese origin and is not enriched for individuals presenting reading difficulties, as demonstrated by the distribution of their language and reading abilities scores (Supplementary Table 2). It is debated whether dyslexia in Chinese is the manifestation of a different deficit compared to European populations as a result of culture.⁵² However both behavioral⁵³ and neuroimaging⁵⁴ studies predict commonalities in the neuronal markers for dyslexia that can be dissociated from language-specific effects in both English and Chinese individuals. The incidence of dyslexia in Hong Kong, similar to the UK populations, is about 10%.² However, visual processing may be more relevant for decoding Chinese compared to alphabetic scripts because of the complex visual composition of Chinese characters.⁵⁵ The role of the visual magnocellular pathway in reading has been suggested to be particularly relevant for decoding Chinese scripts.⁵⁶ The visual motion perception tests that compared groups on the basis of the *DCDC2* deletion genotypes were designed to assess the role of this pathway in dyslexia.^{21,22} Genetic association analysis for VMOT measures in the Hong Kong cohort, or other Chinese speaking groups, would allow direct comparisons with these studies.

We tested for effects of the deletion on a visual motion task (Table 3; Supplementary Table 1) in the Oxford Dyslexia cohort and found no associations, thereby failing to replicate the findings by Cicchini *et al* (2015),²² albeit we used a different version of the VMOT test (see Supplementary Material for details). Cicchini and colleagues used a contrast sensitivity paradigm with a very short presentation time that varied the contrast of a moving stimulus to the detection threshold. Our VMOT task involved much longer presentation times for random dot kinematograms (RDK) in which the signal to noise ratio was manipulated to the detection threshold. Although we cannot directly compare the two tasks, the observation that the probands in our sample have lower sensitivity to motion compared to the entire sample (Table 3)

provides cross-validation of previous work with this task in dyslexia⁵⁷ and as a correlate with reading skills in the normal population.³⁴ The sample we used for the VMOT analysis was substantially larger ($N=374$ non-carriers, $N=66$ heterozygotes; $N=6$ deletion homozygotes; Table 3) compared to the analysis reported by Cicchini and colleagues in 11 deletion carriers and 10 non-carriers. Therefore, while we cannot exclude an effect dependent on specific phenotypes or of smaller size, our study provides a new data set for the interpretation of previously reported results. Regardless of the differences in the tests used, our genetic analysis does not support the design of behavioral studies based on the stratification for *DCDC2* deletion genotypes.

While the *DCDC2* deletion is interesting for the potential regulatory effect, other *DCDC2* markers, mainly SNPs, showed more consistent associations across independent studies.^{4–10} Our current study, therefore, does not undermine the overall evidence in support of *DCDC2* as a candidate gene for dyslexia or reading abilities.

Replication studies are essential but present particular challenges. The acquisition of cognitive measures is only possible through expensive and time consuming one-on-one sessions, making it extremely difficult to obtain large sample sizes. Heterogeneity of phenotypic and behavioral measures makes direct comparison across studies very difficult. Overall, these observations reinforce the importance of collecting high quality cognitive data in general population samples to enable follow up studies of genetic associations reported for neurodevelopmental disorders.⁵⁸

In summary, through both a re-evaluation of the published literature and new genetic association analyses, our results show that there is no strong evidence in support of the *DCDC2* deletion as a risk factor for dyslexia. These data are important for guiding the future direction of dyslexia research and more generally highlight the caveats associated with over-generalizing from analyses conducted in small sample sizes and from misinterpretations of the literature. Most importantly, we provide a useful example of the importance of publishing null results toward avoiding the potential of canonizing weak evidence as fact.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

SP is a Royal Society University Research Fellow. EM was the recipient of a Wolfson Intercalated Award. Support to the analysis was provided by the St Andrews Bioinformatics Unit funded by the Wellcome Trust [grant 105621/Z/14/Z]. Analysis of the York cohort was supported by a Wellcome Trust Programme Grant to MJS [WT082032MA]. Assessment of the Aston cohort was supported by funding from The Waterloo Foundation to JBT and SP [797–1720]. The work on the Chinese-English Twin Study of Biliteracy was supported by a Hong Kong Collaborative Research Fund (CUHK8/CRF/13G) award to CM, CH, and MW, The Chinese University of Hong Kong VCF2014023 fund awarded to CM, and a SFC/RGC Hong Kong/Scotland Collaborative Partnership award to SP and CM (2014_HK-Scot-0020; X-CUHK404/14). We are grateful to Dianne Newbury and Dorothy Bishop for useful comments during the preparation of this manuscript.

REFERENCES

- Katusic SK, Colligan RC, Barbareasi WJ, Schaid DJ, Jacobsen SJ. Incidence of Reading Disability in a Population-Based Birth Cohort, 1976–1982, Rochester, Minn. *Mayo Clin Proc* 2001; **76**: 1081–1092.
- Chan DW, Ho CS, Tsang S, Lee S, Chung KKH. Prevalence, gender ratio and gender differences in reading-related cognitive abilities among Chinese children with dyslexia in Hong Kong. *Educ Stud* 2007; **33**: 249–265.
- Paracchini S, Diaz R, Stein J. Chapter two – advances in dyslexia genetics—new insights into the role of brain asymmetries. *Adv Genet* 2016; **96**: 53–97.

- Meng H, Smith SD, Hager K, Held M, Liu J, Olson RK et al. *DCDC2* is associated with reading disability and modulates neuronal development in the brain. *Proc Natl Acad Sci USA* 2005; **102**: 17053–17058.
- Couto JM, Gomez L, Wigg K, Ickowicz A, Pathare T, Malone M et al. Association of attention-deficit/hyperactivity disorder with a candidate region for reading disabilities on chromosome 6p. *Biol Psychiatry* 2009; **66**: 368–375.
- Schumacher J, Anthoni H, Dahdouh F, Konig IR, Hillmer AM, Kluck N et al. Strong genetic evidence of *DCDC2* as a susceptibility gene for dyslexia. *Am J Hum Genet* 2006; **78**: 52–62.
- Wilcke A, Weissfuss J, Kirsten H, Wolfram G, Boltze J, Ahnert P. The role of gene *DCDC2* in German dyslexics. *Ann Dyslexia* 2009; **59**: 1–11.
- Newbury DF, Paracchini S, Scerri TS, Winchester L, Addis L, Richardson AJ et al. Investigation of dyslexia and SLI risk variants in reading-and language-impaired subjects. *Behav Genet* 2011; **41**: 90–104.
- Zhang Y, Li J, Song S, Tardif T, Burmeister M, Villafuerte SM et al. Association of *DCDC2* polymorphisms with normal variations in reading abilities in a chinese population. *PLoS ONE* 2016; **11**: e0153603.
- Scerri TS, Morris AP, Buckingham LL, Newbury DF, Miller LL, Monaco AP et al. *DCDC2*, KIAA0319 and CMIP are associated with reading-related traits. *Biol Psychiatry* 2011; **70**: 237–245.
- Powers NR, Eicher JD, Butter F, Kong Y, Miller LL, Ring SM et al. Alleles of a polymorphic *ETV6* binding site in *DCDC2* confer risk of reading and language impairment. *Am J Hum Genet* 2013; **93**: 19–28.
- Powers NR, Eicher JD, Miller LL, Kong Y, Smith SD, Pennington BF et al. The regulatory element *READ1* epistatically influences reading and language, with both deleterious and protective alleles. *J Med Genet* 2016; **53**: 163–171.
- Ludwig KU, Schumacher J, Schulte-Korne G, Konig IR, Warnke A, Plume E et al. Investigation of the *DCDC2* intron 2 deletion/compound short tandem repeat polymorphism in a large German dyslexia sample. *Psychiatr Genet* 2008; **18**: 310–312.
- Harold D, Paracchini S, Scerri T, Dennis M, Cope N, Hill G et al. Further evidence that the KIAA0319 gene confers susceptibility to developmental dyslexia. *Mol Psychiatry* 2006; **11**(1061): 1085–1091.
- Marino C, Meng H, Mascheretti S, Rusconi M, Cope N, Giorda R et al. *DCDC2* genetic variants and susceptibility to developmental dyslexia. *Psychiatr Genet* 2012; **22**: 25–30.
- Meng H, Powers NR, Tang L, Cope NA, Zhang PX, Fuleihan R et al. A dyslexia-associated variant in *DCDC2* changes gene expression. *Behav Genet* 2011; **41**: 58–66.
- Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013; **14**: 365–376.
- Brkanac Z, Chapman NH, Matsushita MM, Chun L, Nielsen K, Cochrane E et al. Evaluation of candidate genes for *DYX1* and *DYX2* in families with dyslexia. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 556–560.
- Meda SA, Gelernter J, Gruen JR, Calhoun VD, Meng H, Cope NA et al. Polymorphism of *DCDC2* reveals differences in cortical morphology of healthy individuals—a preliminary voxel based morphometry study. *Brain Imaging Behav* 2008; **2**: 21–26.
- Marino C, Scifo P, Della Rosa PA, Mascheretti S, Facoetti A, Lorusso ML et al. The *DCDC2*/intron 2 deletion and white matter disorganization: Focus on developmental dyslexia. *Cortex* 2014; **57**: 227–243.
- Gori S, Mascheretti S, Giora E, Ronconi L, Ruffino M, Quadrelli E et al. The *DCDC2* intron 2 deletion impairs illusory motion perception unveiling the selective role of magnocellular-dorsal stream in reading (dis)ability. *Cereb Cortex* 2015; **25**: 1685–1695.
- Cicchini GM, Marino C, Mascheretti S, Perani D, Morrone MC. Strong motion deficits in dyslexia associated with *DCDC2* gene alteration. *J Neurosci* 2015; **35**: 8059–8064.
- Hadad B, Schwartz S, Maurer D, Lewis TL. Motion perception: a review of developmental changes and the role of early visual experience. *Front Integr Neurosci* 2015; **9**: 49.
- Benassi M, Simonelli L, Giovagnoli S, Bolzani R. Coherence motion perception in developmental dyslexia: a meta-analysis of behavioral studies. *Dyslexia* 2010; **16**: 341–357.
- Stein J, Walsh V. To see but not to read; the magnocellular theory of dyslexia. *Trends Neurosci* 1997; **20**: 147–152.
- Talcott JB, Witton C, Stein JF. Probing the neurocognitive trajectories of children's reading skills. *Neuropsychologia* 2013; **51**: 472–481.
- Baker M. 1,500 scientists lift the lid on reproducibility. *Nature* 2016; **533**: 452–454.
- Nissen SB, Magidson T, Gross K, Bergstrom CT, Arbesman S, Baker M et al. Publication bias and the canonization of false facts. *Elife* 2016; **5**: 452–454.
- Elliot C, Murray D, Pearson L. British abilities scales. *NFER-Nelson: Windsor* 1983.
- Castles A, Coltheart M. Varieties of developmental dyslexia. *Cognition* 1993; **47**: 149–180.

- 31 Olson Forsberg H, Wise BRK Genes, environment, and the development of orthographic skills In: Berninger VW (ed) *The Varieties of Orthographic Knowledge, 1: Theoretical and Developmental Issues*. Kluwer Academic: Dordrecht, The Netherlands, 1994; 27–71.
- 32 Frederickson N, Frith U, Reason R. *Phonological Assessment Battery (Manual and Materials)*. NFER-Nelson: Windsor, 1997.
- 33 Scerri TS, Paracchini S, Morris A, MacPhie IL, Talcott J, Stein J *et al*. Identification of candidate genes for dyslexia susceptibility on chromosome 18. *PLoS ONE* 2010; **5**: e13712.
- 34 Talcott JB, Witton C, Hebb GS, Stoodley CJ, Westwood EA, France SJ *et al*. On the relationship between dynamic visual and auditory processing and literacy skills; results from a large primary-school study. *Dyslexia* 2002; **8**: 204–225.
- 35 Wechsler D. *Wechsler Individual Achievement Test*. 2nd edn. The Psychological Corporation: San Antonio, 2002.
- 36 Wagner R, Torgesen JK, Rashotte CA *Test of Word Reading Efficiency* 2nd edn. Pearson Assessment: Austin, 2011.
- 37 Nash HM, Hulme C, Gooch D, Snowling MJ. Preschool language profiles of children at family risk of dyslexia: continuities with specific language impairment. *J Child Psychol Psychiatry* 2013; **54**: 958–968.
- 38 Pettigrew KA, Frinton E, Nudel R, Chan MTM, Thompson P, Hayiou-Thomas ME *et al*. Further evidence for a parent-of-origin effect at the NOP9 locus on language-related phenotypes. *J Neurodev Disord* 2016; **8**: 24.
- 39 Hulme C, Stothard S, Clarke P, Bowyer-Crane C, Harrington A, Truelove E *et al*. YARC York assessment of reading for comprehension. Early reading. GL Assessment; 2009.
- 40 Stainthorp R Diagnostic test of word reading processes. GL Assessment 2012.
- 41 Hulme C, Nash HM, Gooch D, Lervag A, Snowling MJ. The foundations of literacy development in children at familial risk of dyslexia. *Psychol Sci* 2015; **26**: 1877–1886.
- 42 Wong SWL, Ho CS-H, McBride C, Chow BW-Y, Waye MMY, Chen J *et al*. Less is more in Hong Kong: investigation of biscriptal and trilingual development among Chinese twins in a (relatively) small city. *Twin Res Hum Genet* 2016; **13**: 1–6.
- 43 Ho CS-H, Chan D, Chung K, Tsang S-M, Lee S-H, Cheng RW-Y. *The Hong Kong Test of Specific Learning Difficulties in Reading and Writing for Primary School Students Second*. Hong Kong Specific Learning Difficulties Research Team: Hong Kong, 2007.
- 44 Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* 2005; **21**: 3445–3447.
- 45 Abecasis GR, Cardon LR, Cookson WO. A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 2000; **66**: 279–292.
- 46 Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97–101.
- 47 Pettigrew KA, Reeves E, Leavett R, Hayiou-Thomas ME, Sharma A, Simpson NH *et al*. Copy number variation screen identifies a rare de novo deletion at chromosome 15q13.1-13.3 in a child with language impairment. *PLoS One* 2015; **10**: e0134997.
- 48 Brandler WM, Morris AP, Evans DM, Scerri TS, Kemp JP, Timpson NJ *et al*. Common variants in left/right asymmetry genes and pathways are associated with relative hand skill. *PLoS Genet* 2013; **9**: e1003751.
- 49 Purcell S, Cherny SS, Sham PC. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; **19**: 149–150.
- 50 Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR *et al*. A global reference for human genetic variation. *Nature* 2015; **526**: 68–74.
- 51 Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P *et al*. Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am J Hum Genet* 2005; **76**: 581–591.
- 52 Siok WT, Perfetti CA, Jin Z, Tan LH. Biological abnormality of impaired reading is constrained by culture. *Nature* 2004; **431**: 71–76.
- 53 Ziegler JC, Goswami U. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol Bull* 2005; **131**: 3–29.
- 54 Hu W, Lee HL, Zhang Q, Liu T, Geng LB, Seghier ML *et al*. Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. *Brain* 2010; **133**: 1694–1706.
- 55 McBride-Chang C, Lam F, Lam C, Chan B, Fong CY-C, Wong TT-Y *et al*. Early predictors of dyslexia in Chinese children: familial history of dyslexia, language delay, and cognitive profiles. *J Child Psychol Psychiatry* 2011; **52**: 204–211.
- 56 Zhao J, Qian Y, Bi H-Y, Coltheart M, Rauschecker AM, Bar M *et al*. The visual magnocellular-dorsal dysfunction in Chinese children with developmental dyslexia impedes Chinese character recognition. *Sci Rep* 2014; **4**: 7068.
- 57 Hansen PC, Stein JF, Orde SR, Winter JL, Talcott JB. Are dyslexics' visual deficits limited to measures of dorsal stream function? *Neuroreport* 2001; **12**: 1527–1530.
- 58 Paracchini S. Dissection of genetic associations with language-related traits in population-based cohorts. *J Neurodev Disord* 2011; **3**: 365–373.



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

© The Author(s) 2017

Supplementary Information accompanies the paper on the *Translational Psychiatry* website (<http://www.nature.com/tp>)