Brain & Language 172 (2017) 9–15

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Short Communication

Neuroimaging genetic analyses of novel candidate genes associated with reading and language

Alessandro Gialluisi [a,b], Tulio Guadalupe [a], Clyde Francks [a,c], Simon E. Fisher [a,c,*]

[a] Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
[b] Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany
[c] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands

ARTICLE INFO

Article history:
Received 17 September 2015
Accepted 7 July 2016
Available online 27 July 2016

Keywords:
RBFOX2
Cortical thickness
Language
Reading
Brain
Imaging genetics

ABSTRACT

Neuroimaging measures provide useful endophenotypes for tracing genetic effects on reading and language. A recent Genome-Wide Association Scan Meta-Analysis (GWASMA) of reading and language skills (N = 1862) identified strongest associations with the genes *CCDC136*/*FLNC* and *RBFOX2*. Here, we follow up the top findings from this GWASMA, through neuroimaging genetics in an independent sample of 1275 healthy adults. To minimize multiple-testing, we used a multivariate approach, focusing on cortical regions consistently implicated in prior literature on developmental dyslexia and language impairment. Specifically, we investigated grey matter surface area and thickness of five regions selected *a priori*: middle temporal gyrus (MTG); pars opercularis and pars triangularis in the inferior frontal gyrus (IFG-PO and IFG-PT); postcentral parietal gyrus (PPG) and superior temporal gyrus (STG). First, we analysed the top associated polymorphisms from the reading/language GWASMA: rs59197085 (*CCDC136*/*FLNC*) and rs5995177 (*RBFOX2*). There was significant multivariate association of rs5995177 with cortical thickness, driven by effects on left PPG, right MTG, right IFG (both PO and PT), and STG bilaterally. The minor allele, previously associated with reduced reading-language performance, showed negative effects on grey matter thickness. Next, we performed exploratory gene-wide analysis of *CCDC136*/*FLNC* and *RBFOX2*; no other associations surpassed significance thresholds. *RBFOX2* encodes an important neuronal regulator of alternative splicing. Thus, the prior reported association of rs5995177 with reading/language performance could potentially be mediated by reduced thickness in associated cortical regions. In future, this hypothesis could be tested using sufficiently large samples containing both neuroimaging data and quantitative reading/language scores from the same individuals.

1. Introduction

Variability in speech, language and reading skills is known to be highly heritable (Graham & Fisher, 2015). Progress in molecular methods has begun to identify genetic polymorphisms that contribute not only to disorders of such abilities, but also to normal variation in the general population (reviewed by Deriziotis & Fisher, 2013; Graham & Fisher, 2015). Genetic association studies of behaviour and cognition can point to gene variants that are correlated with the distal phenotype of interest, but they do not uncover the intermediate pathways that explain the association. One potential route whereby a genetic polymorphism may ultimately impact on speech, language or reading performance is via effects on structural properties of relevant brain circuits. Indeed, Magnetic Resonance Imaging (MRI) studies of twins have shown that individual variation in language-related cortical structures is strongly influenced by genetic factors (Thompson et al., 2001), although the specific polymorphisms that are involved remain largely unknown. Neuroimaging genomics offers a way forward, by assessing cohorts in which both DNA and structural MRI data have been collected from the same individuals – structural measures are then used as endophenotypes for association testing (Thompson et al., 2014). In the current study, we took a number of candidate polymorphisms suggested by recent genetic screening of reading/language performance and assessed their effects on cortical structure for relevant brain regions.

To constrain our study, we first conducted a literature review to identify *a priori* which cortical regions to focus on. Neuroimaging studies have searched for structural differences in people with dyslexia (reading disability) and language impairment, compared to

* Corresponding author at: Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, Netherlands.
E-mail address: simon.fisher@mpi.nl (S.E. Fisher).

http://dx.doi.org/10.1016/j.bandl.2016.07.002
0093-934X/

matched controls (reviewed in Eicher & Gruen, 2013). Multiple investigations have reported effects on classic language-related regions of the cerebral cortex, namely Broca's and Wernicke's areas. The former roughly corresponds to pars opercularis and pars triangularis in the left inferior frontal gyrus (IFG-PO and IFG-PT, respectively), while the latter overlaps with the posterior part of the left superior temporal gyrus (STG) (Kennison, 2013).

Reduced leftward asymmetries in the posterior STG (also known as the planum temporale) have been associated with dyslexia in some studies; a meta-analysis revealed altered asymmetry of planum temporale surface area in male dyslexics, with a greater proportion of rightward asymmetrical cases compared to controls (Altarelli et al., 2014). Another study reported a significant correlation between asymmetry of STG white-matter density and a skill related to phonological processing in dyslexic subjects (Dole, Meunier, & Hoen, 2013), while cortical electrostimulation mapping has identified STG as a key area of activation in healthy subjects performing language-related tasks (Roux et al., 2014). Alterations of the STG have also been detected in children with specific language impairment (SLI), characterized by smaller white-matter volumes in the left hemisphere (Jäncke, Siegenthaler, Preis, & Steinmetz, 2007) and smaller grey-matter volumes bilaterally (Badcock, Bishop, Hardiman, Barry, & Watkins, 2012). Moreover, Childhood Apraxia of Speech (CAS), has been associated with a bilateral increase in STG grey-matter density (Belton, Salmond, Watkins, Vargha-Khadem, & Gadian, 2003; Watkins, Dronkers, & Vargha-Khadem, 2002).

The posterior part of the IFG is known to have a prominent role in phonological processing (Lu et al., 2007; Salo, Rinne, Salonen, & Alho, 2013). Reduced grey-matter volumes in left IFG and decreased leftward asymmetry have been reported both in dyslexia (Brambati et al., 2006; Hoeft et al., 2007) and in CAS (Belton et al., 2003), while larger grey-matter volumes in the left IFG were found in SLI cases (Badcock et al., 2012).

Beyond the IFG and STG findings described above, reduced grey-matter leftward asymmetry has been observed in the middle temporal gyrus (MTG) of dyslexic people (Dole et al., 2013), in partial contrast with a previous study (Brambati et al., 2004); and white matter anomalies of this region have been associated with SLI (Soriano-Mas et al., 2009). More recently, reduced grey-matter volumes were reported in the right postcentral parietal gyrus (PPG) in both comorbid dyslexia-SLI cases and SLI-only cases (Girbau-Massana, Garcia-Marti, Marti-Bonmati, & Schwartz, 2014). Interestingly, earlier investigations identified a bilateral reduction of PPG in dyslexia cases versus controls (Hoeft et al., 2007), and an atypical bilateral activation of this region in CAS subjects while performing language tasks (Liegeois et al., 2003).

These (and other) regions are known to be structurally connected, and functionally linked during language- and reading-related tasks (as reviewed in Friederici, 2011; Vandermosten, Boets, Wouters, & Ghesquière, 2012). Moreover, the degree of connectivity of such regions has been associated with variability in reading and language performance (Boets et al., 2013; Verly et al., 2014), further supporting their involvement in these cognitive domains.

Several candidate genes for susceptibility to speech, language and reading disorders (reviewed in Deriziotis & Fisher, 2013) have been tested using neuroimaging genetics, in prior attempts to bridge gaps between genes, brain and behaviour/cognition. Variants in *KIAA0319*, *DCDC2*, *ACOT13*, *DYX1C1*, *DYX3*, *FOXP2,* and *CNTNAP2* have been reported to show associations with structural variation in language-related brain regions, including some of the cortical regions mentioned above (Darki, Peyrard-Janvid, Matsson, Kere, & Klingberg, 2012; Eicher et al., 2016; Jamadar et al., 2011, 2013; Meda et al., 2008). However, sample sizes thus far have been small in most of the studies of common genetic variants, yielding susceptibility to false-positive findings, and non-replications have been reported (e.g. Hoogman et al., 2014).

The present study focused on novel candidate genes recently reported to be associated with reading and language skills: *CCDC136* (coiled-coil domain containing 136) and *FLNC* (filamin C) on 7q32.1, and *RBFOX2* (RNA-binding protein, fox-1 homolog 2) on 22q12.3. These new candidates were identified by Gialluisi et al. (2014) in a Genome-Wide Association Scan Meta-Analysis (GWASMA) of a principal component score derived from reading and language traits, using three well characterized datasets comprising individuals with histories of reading or language problems, and their siblings (N = 1862). In the GWASMA, the strongest associations ($p \sim 10^{-7}$) were observed for the single-nucleotide polymorphisms (SNPs) rs59197085, within *CCDC136* and ~10 kb upstream of *FLNC*, and rs5995177, within *RBFOX2* (Gialluisi et al., 2014). Recently, a new gene was annotated to the 7q32.1 region, namely *LOC105375496* (long noncoding RNA, uncharacterized), ~2 kb downstream of rs59197085.

Here, we assessed whether rs59197085, rs5995177 and/or other SNPs in *CCDC136, FLNC* or *RBFOX2*, have detectable effects on brain architecture, tightly constraining our hypotheses by targeting the cortical regions from the literature review above. Specifically, we used structural MRI data from a large dataset of 1275 healthy adults to analyse genetic association with grey-matter measures of the five cortical regions highlighted by our literature search; MTG, PPG, STG, IFG-PO and IFG-PT (Fig. 1). For these regions, we analysed surface area and thickness for left and right hemispheres, and carried out multivariate association analysis with these correlated measures (see Table 1), circumventing the increased multiple-testing burden of a separate region-by-region approach. Moreover, our multivariate approach enabled detection of potentially multi-regional genetic effects on the cortical language-related networks constituted by these brain structures, allowing for genetic effect sizes to vary across regions.

2. Results

Our study used a staged approach. We first analysed the two most significantly associated SNPs from the recent Gialluisi et al. (2014) GWASMA of reading and language skills, namely rs59197085 (7q32.1) and rs5995177 (22q12.3). Multivariate analyses with cortical surface area and thickness of the ten selected

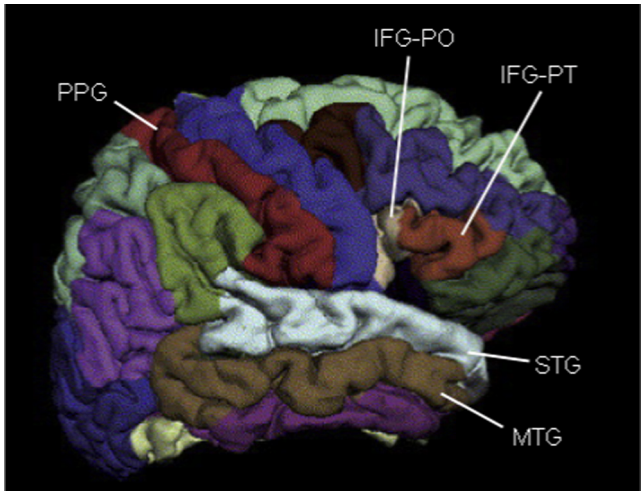

Fig. 1. Cortical brain regions tested for association in this study. MTG = middle temporal gyrus; IFG-PO = inferior frontal gyrus - pars opercularis; IFG-PT = inferior frontal gyrus - pars triangularis; PPG = postcentral parietal gyrus; STG = superior temporal gyrus. Figure adapted by authors from Desikan et al. (2006).

Table 1
Cross-trait correlations of the brain measures tested, corrected for covariates used in the analysis (gender, age, TBV, and field strength of the MRI). The upper part of the matrix (above the diagonal) shows correlations across measures of cortical surface area, while the lower part (below the diagonal) refers to measures of cortical thickness. All correlations were significant ($p < 0.05$), unless otherwise indicated.

Brain measure[a]	MTG_L	IFG-PO_L	IFG-PT_L	PPG_L	STG_L	MTG_R	IFG-PO_R	IFG-PT_R	PPG_R	STG_R
MTG_L	1	0.143	0.069	0.225	0.26	0.577	0.124	0.076	0.232	0.355
IFG-PO_L	0.456	1	0.374	0.085	0.201	0.176	0.367	0.299	0.139	0.239
IFG-PT_L	0.419	0.554	1	0.045[b]	0.186	0.031[b]	0.212	0.437	0.077	0.187
PPG_L	0.357	0.367	0.378	1	0.241	0.26	0.111	0.053[b]	0.478	0.254
STG_L	0.58	0.517	0.457	0.445	1	0.351	0.188	0.218	0.254	0.578
MTG_R	0.707	0.443	0.396	0.342	0.556	1	0.132	0.08	0.283	0.388
IFG-PO_R	0.398	0.551	0.404	0.338	0.491	0.432	1	0.212	0.094	0.206
IFG-PT_R	0.4	0.525	0.563	0.351	0.449	0.425	0.509	1	0.054[b]	0.169
PPG_R	0.326	0.342	0.334	0.705	0.422	0.309	0.317	0.329	1	0.291
STG_R	0.561	0.531	0.465	0.47	0.758	0.615	0.506	0.47	0.438	1

[a] MTG = middle temporal gyrus; IFG-PO = inferior frontal gyrus - pars opercularis; IFG-PT = inferior frontal gyrus - pars triangularis; PPG = postcentral parietal gyrus; STG = superior temporal gyrus. Suffixes "L" and "R" indicate left and right hemisphere, respectively.
[b] Non-significant correlation ($p \geqslant 0.05$).

regions (left and right MTG, IFG-PO, IFG-PT, PPG and STG) revealed a significant association of rs5995177 ($p \sim 0.012$, A/G, minor allele A, MAF ~ 7.8%) with thickness (Table 2). This association survives correction for multiple testing of 2 independent SNPs and 2 types of measure (surface area and thickness, $\alpha = 0.0125$). We used univariate analyses to further characterize this result, assessing the relative contributions of the different regions to this significant multivariate finding, and investigating direction of effect. The multivariate association was mainly driven by associations with left PPG, right MTG, right IFG (PO and PT), and STG bilaterally (Table 3). The minor allele (A) was associated with a reduction of grey matter thickness in these regions (Fig. Sa–f). The allelic trend was consistent across all regions tested, and all which were not significantly associated with rs5995177 showed p-values trending towards significance, with the exception of right PPG (see Table 3 for details).

After focusing on the top SNPs from the Gialluisi et al. (2014) reading/language GWASMA, we next extended our multivariate association analysis to all 682 SNPs falling within or close to the candidate genes *CCDC136*, *LOC105375496*, *FLNC* and *RBFOX2* (up to 50 kb beyond the 5′- and 3′-UTRs). This exploratory gene-wide analysis did not reveal any significant association withstanding correction for multiple testing of 2 multivariate tests and the effective number of independent SNPs (78) tested in the two genes ($\alpha = 3.2 \times 10^{-4}$, see Section 4 for further details). The most significant multivariate associations were observed with cortical thickness for four polymorphisms in *RBFOX2*, namely rs78563107, rs6000084, rs6000085, and rs144006011 ($p = 4.3–7.1 \times 10^{-3}$). Targeted follow-up of these top SNPs with univariate analyses of cortical thickness suggests that the effects were mainly driven by the IFG-PO and the STG, bilaterally ($p = 5.1 \times 10^{-4}$ - 0.041). These SNPs were all in perfect linkage disequilibrium (LD) among themselves ($r^2 = 1$) and in moderate LD with rs5995177 ($r^2 \sim 0.5$).

Similarly, we observed suggestive associations with cortical surface area, at the SNPs rs56184882, rs339054 and rs339046 ($p = 7.3–9.9 \times 10^{-3}$), located ~41–48 kb upstream of *CCDC136*. These SNPS showed significant univariate associations in the IFG-PT bilaterally ($p = 0.8–2.4 \times 10^{-3}$), were in moderate/high reciprocal LD ($r^2 \sim 0.7–1$) and in low LD with rs59197085 ($r^2 < 0.1$).

The gene-based analysis did not reveal any significant enrichment of multivariate SNP associations in the candidate genes tested (see Table Sa and b).

Table 2
Multivariate analyses of rs59197085 (7q32.1) and rs5995177 (22q12.3), the two most significant SNPs from a recent reading/language GWASMA (Gialluisi et al., 2014). Here, these SNPs were investigated for multivariate association with measures of cortical surface area and thickness of pre-selected brain regions, based on prior literature review of neuroimaging findings in dyslexia and language impairment (see Sections 1 and 4). P-values as computed by the software TATES are reported (no beta value was produced in the output). Significant multivariate associations ($p < 0.0125$) are highlighted in bold.

Chr	SNP	Position	MAF (%)	Surface area (p)	Thickness (p)
7	rs59197085	128460756	8.52	0.663	0.724
22	rs5995177	36309553	7.82	0.996	**0.012**

Table 3
Univariate associations of rs5995177 (22q12.3) with cortical thickness of the brain regions tested. Association p-values are reported, with beta values of the minor allele (A) in brackets. Nominally significant univariate associations ($p < 0.05$) are highlighted in bold.

Brain structure[a]	MTG	IFG-PO	IFG-PT	PPG	STG
L	0.143 (−0.019)	0.061 (−0.021)	0.117 (−0.019)	**0.021 (−0.021)**	**2.4×10^{-3} (−0.037)**
R	**0.049 (−0.025)**	**0.015 (−0.029)**	**9×10^{-3} (−0.032)**	0.313 (−0.01)	**2.3×10^{-3} (−0.038)**

[a] MTG = middle temporal gyrus; IFG-PO = inferior frontal gyrus - pars opercularis; IFG-PT = inferior frontal gyrus - pars triangularis; PPG = postcentral parietal gyrus; STG = superior temporal gyrus. Suffixes "L" and "R" indicate left and right hemisphere, respectively.

3. Discussion

In the present study, we analysed association of variants in the genes *CCDC136* (coiled-coil domain containing 136), *LOC105375496* (long noncoding RNA, uncharacterized), *FLNC* (filamin C, 7q32.1) and *RBFOX2* (RNA-binding protein, fox-1 homolog 2, 22q12.3), with structural brain measures in a sample of 1275 healthy participants. These variants and genes had shown the strongest associations in a prior GWASMA of reading and language traits that we previously carried out in three other cohorts (Gialluisi et al., 2014). Here, we used a multivariate approach to test association with grey-matter surface area and thickness of five selected perisylvian cortical regions implicated in reading and language disorders by previous neuroimaging literature; MTG, IFG-PO, IFG-PT, PPG and STG. We preferred a multivariate approach over data collapsing methods (such as testing association with a principal component score derived from the brain measures tested), as it allowed us to analyze both the variance shared among structural brain measures in our regions of interest and region-specific genetic effects, while keeping multiple-testing correction to a minimum.

A focused analysis of the top hits from the reading/language GWASMA (Gialluisi et al., 2014), namely rs59197085 (within *CCDC136*, ~2 kb upstream of *LOC105375496* and ~10 kb upstream of *FLNC*) and rs5995177 (within *RBFOX2*), revealed a significant

multivariate association of rs5995177 with grey-matter thickness. This suggested a generalized effect of rs5995177 on cortical thickness in the reading/language-related brain regions analysed, mainly driven by associations with left PPG, right MTG and IFG (both pars opercularis and pars triangularis), and bilaterally in the STG. The minor allele (A) showed a negative effect on grey-matter thickness. The same allele was associated with reduced reading/language scores in our GWASMA (Gialluisi et al., 2014). Reduced grey-matter volumes have been observed in reading and/or language impaired children for several of the associated brain regions, including STG (Badcock et al., 2012) and PPG (Girbau-Massana et al., 2014). Reduced thickness of cortical areas including IFG-PO, MTG and STG has been reported as a neuroanatomical correlate of dyslexia (Clark et al., 2014). As lower cortical thickness in localized brain regions has been identified as a predictor of dyslexia at pre-reading ages (Clark et al., 2014), we may speculate that the contribution of the minor allele of rs5995177 to poor reading/language performance could be mediated by reduced thickness in associated cortical regions. Since *RBFOX2* encodes an alternative splicing regulator with important roles in brain development (Gehman et al., 2012), it seems plausible that alterations in brain morphology and architecture may result from developmental effects of this gene. Data on reading and language performance were not available for the brain imaging dataset analysed here. Thus, this mediation hypothesis will need to be tested in future in large samples containing both MRI data and quantitative reading/language phenotypes from the same individuals.

Changes in brain morphology can be the result of experience-dependent plasticity in the central nervous system (Dole et al., 2013; Zatorre, Fields, & Johansen-Berg, 2012), which could point to an alternative account of our findings. Thus, it is possible that the associated structural variations in language-related cortex which we observed in this study reflect genotype effects on neural function in those regions. In this view, effects of the minor allele of rs5995177 on reading/language behaviour could indirectly lead to reduced thickness in those cortical areas which play key roles in reading- and language-related cognitive tasks. Interestingly, rs5995177 is located only ~1 kb from a potential binding site for FOXP2, a transcription factor known to be involved in speech and language development (Fisher & Scharff, 2009). Other SNPs close to this binding site showed nominally significant multivariate associations with both cortical thickness and surface area (Table Sc), but none showed an effect as significant as rs5995177 for reading and language skills in our previous GWASMA, and as broad in relation to structural brain measures in the present study. Based on the accumulated evidence thus far, we consider rs5995177 more likely to tag a biological effect on reading/language capacities and on brain architecture in this region, possibly mediated by FOXP2 binding.

In the prior GWASMA, the association of rs5995177 with reading-language scores was weaker after IQ-adjustment (see Gialluisi et al., 2014). Other candidate SNPs in dyslexia/SLI susceptibility genes have been reported to be associated with decreased general cognitive abilities and with reduced volumes in specific brain regions. Scerri et al. (2012) reported association of the SNPs rs917235 and rs714939 in the *MRPL19/GCFC2* locus (2p12) with lower verbal IQ and with bilateral decreases of white-matter volume in the corpus callosum and in the cingulum. These SNPs had been previously associated with dyslexia susceptibility in other samples (Anthoni et al., 2007). In a very recent study, suggestive associations with structural MRI measures were reported at the same locus, namely for rs917235 and rs2298948 with cortical thickness of left and right MTG, respectively, in a sample of 332 healthy participants (Eicher et al., 2016). In the same study, the authors reported associations with cortical thickness at the dyslexia susceptibility locus DYX2, including a suggestive association at rs3777663 (*ACOT13*) in the left IFG-PO, and a significant association at rs9461045 (*KIAA0319*) in the left orbitofrontal cortex (Eicher et al., 2016). These two SNPs had been previously found to be associated with reading, language and cognitive performance in other samples. Although Eicher et al. (2016) investigated a substantially smaller sample than we did here (332 subjects rather than 1275), their findings are consistent with ours, in that risk alleles of SNPs implicated in poor reading/language performance showed negative effects on cortical thickness in both studies.

When we extended our analysis to all SNPs annotated to the candidate genes on 7q32.1 and to *RBFOX2*, we did not detect additional significant associations surviving Bonferroni correction, nor did we find significant effects in a gene-based analysis.

In the current study, we focused tightly on a small well-defined set of cortical areas, which represent five of the most prominently implicated cerebral cortical regions of the perisylvian language network. Although our selection of regions of interest was not exhaustive for all regions of the cerebral cortex that have been implicated in healthy or disordered language or reading cognition, restricting our analysis to these five perisylvian regions of interest meant that the statistical power of our multivariate approach did not suffer adversely from extensive testing over multiple poorly correlated regional measures. Additional brain structures - other than those investigated here - have been implicated in reading and language, including not only cortical areas but also the cerebellum, thalamus, basal ganglia and multiple fiber bundles which are thought to be important in connecting language-related areas (reviewed in Eicher & Gruen, 2013; Klostermann, 2013; Mariën et al., 2014; Vandermosten et al., 2012). Future neuroimaging genetic analyses should investigate *CCDC136/FLNC*, *RBFOX2* and other reading/language-related candidate genes in larger samples with sufficient power for a broader scope, including all brain structures with relevance to reading and language cognition. Ideally such work should make use of longitudinal data to assess genetic effects on brain endophenotypes across development, and include behavioural measures on reading and language performance, combined with the analysis of structural and functional brain imaging data in the same cohorts. Such a design would help to elucidate potential effects of susceptibility genes on reading and language cognition, and to correlate these effects with changes in the architecture of the critical brain circuits, possibly building causality links among these factors.

4. Methods

4.1. Dataset and genotype quality control (QC)

The Brain Imaging Genetics (BIG) resource has been described elsewhere (Franke et al., 2010). It comprises healthy volunteer subjects, including many university students, recruited in Nijmegen, The Netherlands. At the time of this study the BIG subject pool consisted of 2337 self-reported healthy individuals (1248 females) who had undergone anatomical (T1-weighted) MRI scans and who had given their consent to participate in BIG. The participants in BIG have not been assessed using reading/language measures, so such data were not available for this study.

This dataset has been used in other imaging genetics studies (Guadalupe, Zwiers, et al., 2014; Guadalupe et al., 2015), where a detailed description of genotyping of BIG can be found. Briefly, genotype calls for 1303 subjects were generated using the Birdseed algorithm (Rabbee & Speed, 2006) on raw data from the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix Inc., Santa Clara, CA, USA). Future phases of genome-wide genotyping are anticipated, but were not yet available for this study. Samples were

excluded that had call rates <90% and that showed deviant values of genome-wide heterozygosity. SNPs with a Minor Allele Frequency (MAF) <1% or that failed the Hardy-Weinberg equilibrium test ($p \leqslant 10^{-6}$) were also excluded. The resulting markers were then adjusted to the forward strand. A two-step imputation protocol was followed. We used the software MACH for haplotype phasing and Minimac for the final imputation (Howie, Fuchsberger, Stephens, Marchini, & Abecasis, 2012; Li, Willer, Ding, Scheet, & Abecasis, 2010), with the 1000 Genomes Phase 1 v3 EUR reference panel (The 1000 Genomes Project Consortium, 2012). All monomorphic markers were removed from the reference dataset. Individual genotype calls with an imputation certainty <90% were removed, as were markers with an overall quality score (r^2) <0.3. Finally, only markers with ⩽5% missing data were selected. At the end of these procedures, genotypes were available for 1276 subjects (748 females) from BIG. Their mean age was 22.9 years (SD 3.8; range 18–35).

4.2. Phenotype elaboration and QC

4.2.1. Image acquisition

MRI data were acquired in BIG as described elsewhere (Guadalupe, Willems et al., 2014; Guadalupe, Zwiers et al., 2014; Guadalupe et al., 2015). MRI data acquisition was carried out with either a 1.5 T Siemens Sonata or Avanto scanner or a 3 T Siemens Trio or Tim Trio scanner (Siemens Medical Systems, Erlangen, Germany). For the genotyped sample, 634 subjects were scanned at 1.5 T, and 642 subjects at 3 T.

4.2.2. Image processing and phenotypic QC

Image processing has been described elsewhere (Guadalupe, Willems et al., 2014). Automated parcellation of cerebral cortical regions from T1-weighted images was done in FreeSurfer v5.3 (Fischl et al., 2004) according to the Desikan atlas (Desikan et al., 2006). Measures of surface area (in mm^2) were produced for the total cortical surface and for each of 68 cortical parcellations, in each hemisphere. Regional measures of cortical thickness were also generated and analysed, as there is evidence that cortical surface and thickness have independent sources of variation (Panizzon et al., 2009). Estimates of Total Brain Volume (TBV) were calculated as the voxel-wise sum of the grey-matter and white-matter probability maps produced by the VBM8 toolbox, in SPM8. In line with previous imaging genetic association studies on this dataset (Guadalupe, Zwiers et al., 2014; Guadalupe et al., 2015), the following covariates were controlled for in subsequent analyses: gender, age, TBV, and MRI field strength (1.5 or 3 T). Independent analyses of variance shared across the structural brain measures and of each single trait separately showed no batch effect of field strength on the subjects analysed, after residualizing our traits against the above mentioned covariates (data not shown). This suggested that there was no significant amount of residual variance due to scanner type, which had not already been taken into account in the analysis by our covariates. In the genotyped sample, brain measures were available for 1275 participants, which underwent subsequent analyses.

4.2.3. Cortical measures analysed

We analysed both cortical thickness and surface area of the following regions: middle temporal gyrus (MTG); pars opercularis and pars triangularis in the inferior frontal gyrus (IFG-PO and IFG-PT); postcentral parietal gyrus (PPG) and superior temporal gyrus (STG), as defined in the Desikan atlas (Desikan et al., 2006). For each region, we analysed left and right measures separately. These measures showed moderate to high repeatability in scan-rescan correlation analysis of 342 twice-scanned subjects (0.62–0.76 and 0.84–0.91 for measures of cortical thickness and of cortical surface area, respectively) and generally moderate cross-trait correlations (see Table 1). Their distributions were approximately normal (absolute values of skewness and kurtosis <1 and <1.4, respectively) making them suitable for genetic association testing. We also checked for the absence of phenotypic outliers (defined as subjects showing extreme values of both cortical surface area and thickness in at least 20% of the regions tested), in SPSS® 20.0.

4.3. Genetic association analyses

We first analysed the top two SNPs from a prior GWASMA of reading/language (Gialluisi et al., 2014). Both SNPs were imputed in the BIG dataset, with good imputation quality (r^2 = 0.9987 for rs59197085 and 0.952 for rs599177). Next, we extracted and analysed all SNPs falling within or close to *CCDC136*, *LOC105375496, FLNC* (7q32.1) and *RBFOX2* (22q12.13). To include potential regulatory regions in the analysis, we analysed also SNPs in the vicinity of these genes, up to 50 kb beyond the 5′- and 3′-Untranslated Regions (UTRs). The final number of SNPs available for the analysis was 264 for the genes on 7q32.1 and 418 for *RBFOX2*.

4.3.1. Association analysis with cortical surface area and thickness measures

We carried out multivariate genetic association tests of both left and right cortical thickness and surface area traits for selected regions (see Section 4.2.3) using TATES (van der Sluis, Posthuma, & Dolan, 2013). Thicknesses and areas were analysed in separate multivariate tests. The TATES method is claimed to be optimal for detecting multivariate genetic associations affecting some, but not necessarily all, of a set of correlated phenotypes, and is also powerful in the detection of contrasting genetics effects (e.g. to identify SNPs affecting some phenotypes positively, some negatively; van der Sluis et al., 2013).

TATES combines the p-values obtained in univariate genetic association analysis on multiple (correlated) phenotypes, to produce one multivariate association p-value per SNP, while correcting for the correlations between the phenotypes. The univariate associations needed as input for TATES analysis were produced through – *linear* analysis in PLINK v1.07 (Purcell et al., 2007), controlling for the covariates age, gender, TBV and field strength of the MRI.

We initially tested the top two independent SNPs from the Gialluisi et al. (2014) GWASMA; rs59197085 (7q32.1) and rs5995177 (22q12.3). Therefore, we performed 4 separate tests as our primary hypotheses for this study, i.e. each of two SNPs in each of two multivariate association tests (for thicknesses and for areas). This resulted in a corrected α threshold of 0.0125.

Then, as a follow-up exploratory analysis, we carried out multivariate association analysis (TATES) for all 682 SNPs within *CCDC136*, *LOC105375496, FLNC* and *RBFOX2* (including SNPs up to 50 kb beyond the 5′- and 3′-UTRs). To make an appropriate correction for multiple testing given the LD structure within each gene, we calculated the effective number of independent tests using the *Genetic Type I error calculator* (Li, Yeung, Cherny, & Sham, 2012), using our genotypes as input. The effective number of tests was 78 (36 in the 7q32.1 region and 42 in *RBFOX2*), further multiplied by a factor of two for testing separately for thicknesses and areas as above. This resulted in a corrected α threshold of 3.2×10^{-4}.

4.3.2. Gene-based analysis

Multivariate associations of the 682 SNPs analysed were tested for enrichment in our genes of interest, through VEGAS2 (Mishra & Macgregor, 2015). This tool performs gene-based association tests by assigning multiple SNPs to each individual gene according to their genomic locations, and then combining the evidence for asso-

ciation across all SNPs assigned to a given gene, while taking into account LD structure. Again, we assigned to genes all SNPs located up to 50 kb beyond their 5′- and 3′-UTRs, to include potential regulatory regions. For this analysis, we used a significance threshold $\alpha = 0.0063$ (corrected for four candidate genes and two traits tested).

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

The BIG database was established in Nijmegen in 2007. This resource is now part of Cognomics, a joint initiative by researchers of the Donders Centre for Cognitive Neuroimaging, the Human Genetics and Cognitive Neuroscience departments of the Radboud University Medical Centre, and the Max Planck Institute for Psycholinguistics. The Cognomics Initiative is supported by the participating departments and centres and by external grants, i.e. the Biobanking and Biomolecular Resources Research Infrastructure (Netherlands) (BBMRI-NL), the Hersenstichting Nederland and the Netherlands Organisation for Scientific Research (NWO). AG was funded by the Max Planck Society and the Munich Cluster for Systems Neurology (grant 9731). We also wish to thank all persons who kindly participated in this research.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bandl.2016.07.002.

References

Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., et al. (2014). Planum temporale asymmetry in developmental dyslexia: Revisiting an old question. *Human Brain Mapping, 35*(12), 5717–5735.

Anthoni, H., Zucchelli, M., Matsson, H., Muller-Myhsok, B., Fransson, I., Schumacher, J., et al. (2007). A locus on 2p12 containing the co-regulated MRPL19 and C2ORF3 genes is associated to dyslexia. *Human Molecular Genetics, 16*(6), 667–677.

Badcock, N. A., Bishop, D. V. M., Hardiman, M. J., Barry, J. G., & Watkins, K. E. (2012). Co-localisation of abnormal brain structure and function in specific language impairment. *Brain and Language, 120*(3), 310–320.

Belton, E., Salmond, C., Watkins, K., Vargha-Khadem, F., & Gadian, D. (2003). Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Human Brain Mapping, 18*, 194–200.

Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science, 342*(6163), 1251–1254.

Brambati, S. M., Ruffino, M., Danna, M., Lanzi, G., Stella, G., Cappa, S. F., et al. (2006). Neuropsychological deficits and neural dysfunction in familial dyslexia. *Brain Research, 1113*(1), 174–185.

Brambati, S. M., Termine, C., Ruffino, M., Stella, G., Fazio, F., Cappa, S. F., et al. (2004). Regional reductions of gray matter volume in familial dyslexia. *Neurology, 63*(4), 742–745.

Clark, K. A., Helland, T., Specht, K., Narr, K. L., Manis, F. R., Toga, A. W., et al. (2014). Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain, 137*(12), 3136–3141.

Consortium, T. G. P. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature, 491*(7422), 56–65.

Darki, F., Peyrard-Janvid, M., Matsson, H., Kere, J., & Klingberg, T. (2012). Three dyslexia susceptibility genes, DYX1C1, DCDC2, and KIAA0319, affect temporo-parietal white matter structure. *Biological Psychiatry, 72*(8), 671–676.

Deriziotis, P., & Fisher, S. E. (2013). Neurogenomics of speech and language disorders: the road ahead. *Genome Biology, 14*(4), 204.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage, 31*(3), 968–980.

Dole, M., Meunier, F., & Hoen, M. (2013). Gray and white matter distribution in dyslexia: A VBM study of superior temporal gyrus asymmetry. *PLoS ONE, 8*(10), e76823.

Eicher, J. D., & Gruen, J. R. (2013). Imaging-genetics in dyslexia: Connecting risk genetic variants to brain neuroimaging and ultimately to reading impairments. *Molecular Genetics and Metabolism, 110*(3), 201–212.

Eicher, J. D., Montgomery, A. M., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Libiger, O., et al. (2016). Dyslexia and language impairment associated genetic markers influence cortical thickness and white matter in typically developing children. *Brain Imaging and Behavior, 10*(1), 272–282.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex, 14*(1), 11–22.

Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics, 25*(4), 166–177.

Franke, B., Vasquez, A. A., Veltman, J. A., Brunner, H. G., Rijpkema, M., & Fernández, G. (2010). Genetic variation in CACNA1C, a gene associated with bipolar disorder, influences brainstem rather than gray matter volume in healthy individuals. *Biological Psychiatry, 68*(6), 586–588.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91*(4), 1357–1392.

Gehman, L. T., Meera, P., Stoilov, P., Shiue, L., O'Brien, J. E., Meisler, M. H., et al. (2012). The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. *Genes & Development, 26*(5), 445–460.

Gialluisi, A., Newbury, D. F., Wilcutt, E. G., Olson, R. K., DeFries, J. C., Brandler, W. M., et al. (2014). Genome-wide screening for DNA variants associated with reading and language traits. *Genes, Brain and Behavior, 13*(7), 686–701.

Girbau-Massana, D., Garcia-Marti, G., Marti-Bonmati, L., & Schwartz, R. G. (2014). Gray–white matter and cerebrospinal fluid volume differences in children with specific language impairment and/or reading disability. *Neuropsychologia, 56*, 90–100.

Graham, S. A., & Fisher, S. E. (2015). Understanding language from a genomic perspective. *Annual Review of Genetics, 49*, 131–160.

Guadalupe, T., Willems, R. M., Zwiers, M. P., Arias Vasquez, A., Hoogman, M., Hagoort, P., et al. (2014). Differences in cerebral cortical anatomy of left- and right-handers. *Frontiers in Psychology, 5*, 261.

Guadalupe, T., Zwiers, M. P., Teumer, A., Wittfeld, K., Vasquez, A. A., Hoogman, M., et al. (2014). Measurement and genetics of human subcortical and hippocampal asymmetries in large datasets. *Human Brain Mapping, 35*(7), 3277–3289.

Guadalupe, T., Zwiers, M. P., Wittfeld, K., Teumer, A., Vasquez, A. A., Hoogman, M., et al. (2015). Asymmetry within and around the human planum temporale is sexually dimorphic and influenced by genes involved in steroid hormone receptor activity. *Cortex, 62*, 41–55.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., et al. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences, 104*(10), 4234–4239.

Hoogman, M., Guadalupe, T., Zwiers, M. P., Klarenbeek, P., Francks, C., & Fisher, S. E. (2014). Assessing the effects of common variation in the FOXP2 gene on human brain structure. *Frontiers in Human Neuroscience, 8*, 473.

Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics, 44*(8), 955–959.

Jamadar, S., Powers, N. R., Meda, S. A., Calhoun, V. D., Gelernter, J., Gruen, J. R., et al. (2013). Genetic influences of resting state fMRI activity in language-related brain regions in healthy controls and schizophrenia patients: A pilot study. *Brain Imaging and Behavior, 7*(1), 15–27.

Jamadar, S., Powers, N. R., Meda, S. A., Gelernter, J., Gruen, J. R., & Pearlson, G. D. (2011). Genetic influences of cortical gray matter in language-related regions in healthy controls and schizophrenia. *Schizophrenia Research, 129*(2–3), 141–148.

Jäncke, L., Siegenthaler, T., Preis, S., & Steinmetz, H. (2007). Decreased white-matter density in a left-sided fronto-temporal network in children with developmental language disorder: Evidence for anatomical anomalies in a motor-language network. *Brain and Language, 102*(1), 91–98.

Kennison, S. M. (2013). *Introduction to language development* : . SAGE Publications.

Klostermann, F. (2013). Functional roles of the thalamus for language capacities. *Frontiers in Systems Neuroscience, 7*, 32.

Li, Y., Willer, C. J., Ding, J., Scheet, P., & Abecasis, G. R. (2010). MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology, 34*(8), 816–834.

Li, M.-X., Yeung, J. Y., Cherny, S., & Sham, P. (2012). Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics, 131*(5), 747–756.

Liegeois, F., Baldeweg, T., Connelly, A., Gadian, D., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature Neuroscience, 6*, 1230–1237.

Lu, L., Leonard, C., Thompson, P., Kan, E., Jolley, J., Welcome, S., et al. (2007). Normal developmental changes in inferior frontal gray matter are associated with improvement in phonological processing: A longitudinal MRI analysis. *Cerebral Cortex, 17*(5), 1092–1099.

Mariën, P., Ackermann, H., Adamaszek, M., Barwood, C. S., Beaton, A., Desmond, J., et al. (2014). Consensus paper: Language and the cerebellum: An ongoing enigma. *The Cerebellum, 13*(3), 386–410.

Meda, S. A., Gelernter, J., Gruen, J. R., Calhoun, V. D., Meng, H., Cope, N. A., et al. (2008). Polymorphism of DCDC2 reveals differences in cortical morphology of healthy individuals—A preliminary voxel based morphometry study. *Brain Imaging and Behavior, 2*(1), 21–26.

Mishra, A., & Macgregor, S. (2015). VEGAS2: Software for more flexible gene-based testing. *Twin Research and Human Genetics, 18*(1), 86–91.
Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., et al. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cerebral Cortex, 19*(11), 2728–2735.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*(3), 559–575.
Rabbee, N., & Speed, T. P. (2006). A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics, 22*(1), 7–12.
Roux, F.-E., Durand, J.-B., Réhault, E., Planton, S., Draper, L., & Démonet, J.-F. (2014). The neural basis for writing from dictation in the temporoparietal cortex. *Cortex, 50*, 64–75.
Salo, E., Rinne, T., Salonen, O., & Alho, K. (2013). Brain activity during auditory and visual phonological, spatial and simple discrimination tasks. *Brain Research, 1496*, 55–69.
Scerri, T. S., Darki, F., Newbury, D. F., Whitehouse, A. J. O., Peyrard-Janvid, M., Matsson, H., et al. (2012). The dyslexia candidate locus on 2p12 is associated with general cognitive ability and white matter structure. *PLoS ONE, 7*(11), e50321.
Soriano-Mas, C., Pujol, J., Ortiz, H., Deus, J., López-Sala, A., & Sans, A. (2009). Age-related brain structural alterations in children with specific language impairment. *Human Brain Mapping, 30*(5), 1626–1636.
Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V. P., Huttunen, M., et al. (2001). Genetic influences on brain structure. *Nature Neuroscience, 4*(12), 1253–1258.
Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., et al. (2014). The ENIGMA consortium: Large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging and Behavior, 8*(2), 153–182.
van der Sluis, S., Posthuma, D., & Dolan, C. V. (2013). TATES: Efficient multivariate genotype-phenotype analysis for genome-wide association studies. *PLoS Genetics, 9*(1), e1003235.
Vandermosten, M., Boets, B., Wouters, J., & Ghesquière, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience & Biobehavioral Reviews, 36*(6), 1532–1552.
Verly, M., Verhoeven, J., Zink, I., Mantini, D., Peeters, R., Deprez, S., et al. (2014). Altered functional connectivity of the language network in ASD: Role of classical language areas and cerebellum. *NeuroImage: Clinical, 4*, 374–382.
Watkins, K., Dronkers, N., & Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain, 125*, 452–464.
Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience, 15*(4), 528–536.