IMAGING GENETICS STUDY OF SPECIFIC READING DISABILITY: THE ROLE OF *SEMAPHORIN6D*

by
Christina Irene Chen Thomas

A dissertation submitted to the Psychology Department,
College of Liberal Arts and Social Sciences
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Clinical Psychology

Chair of Committee: Elena Grigorenko, Ph.D.

Committee Member: Paul Cirino, Ph.D.

Committee Member: David Francis, Ph.D.

Committee Member: Fumiko Hoeft, M.D., Ph.D.

University of Houston
August 2021

ABSTRACT

The current thesis focused on imaging genetics of specific reading disability (SRD), to better understand the biological risk factors that contribute to SRD. Part 1 of the thesis was a systematic review focusing on summarizing the current imaging genetics literature and characterizing effect sizes of these results, revealing significant associations between reading disability risk genes and brain phenotypes. A Fisher's test revealed promising results for risk genes that had been replicated, including as *DCDC2, KIAA0319, FOXP2, SLC2A3,* and *ROBO1*. Part 2 of the thesis specifically examined associations of the novel candidate gene *Semaphorin 6d (SEMA6D)* on reading-related regions of interest and reading, revealing associations with white matter volume in the left transverse temporal gyrus, which was significantly associated with reading performance measures. Other phenotypes related to *SEMA6D* SNPs included cortical thickness in the fusiform gyrus and gyrification in the supramarginal gyrus, but these regions were not related to reading in the current sample. It is possible that the transverse temporal gyrus was more related to reading in this sample due to young participants who are still developing reading skills, with greater reliance on auditory processing for reading. Part 3 applied the imaging genetics literature to a clinical case to determine whether risk factors were related to SRD in a family consisting of twins discordant for SRD and an older sibling with reading difficulty. Results indicated that SNPs in the genes *ZNF385D, LPHN3, CNTNAP2, FGF18, NOP9, CMIP, MYO18B,* and *RBFOX2* corresponded with SRD. Furthermore, cortical thickness in reading-related regions of interest was more similar among the siblings with SRD compared to the twins, with specific asymmetry

differences in the transverse temporal and superior temporal gyri. The siblings with SRD also demonstrated reduced leftward asymmetry of grey matter volume and cortical surface area in the fusiform gyrus, supramarginal gyrus, and transverse temporal gyrus. Overall, the current thesis summarized and added to the imaging genetics of SRD literature, demonstrated promising effects of a previously unstudied gene, *SEMA6D*, and used a clinical application to understand factors that may be related to SRD risk within a single family. Future research in this field using integrative imaging genetics methodology to understand and predict risk of SRD can better early identification and intervention to have a valuable clinical impact on children with SRD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Specific reading disability (SRD) is a complex neurobehavioral disorder affecting up to 10% of individuals (National Institute of Child Health & Human Development, 2018) in the United States. SRD is a member of a family of developmental disorders involving difficulties acquiring specific academic skills; these disorders are referred to as learning disabilities (LD). An SRD is diagnosed when an individual has a persistent reading difficulty that cannot be explained by sensory deficits, severe cognitive difficulties, lack of motivation, or lack of reading instruction (Lyon et al., 2003). Current definitions focus on response to intervention (RTI) for classification, defining reading disability as a specific deficit in reading despite adequate education and intervention (Fletcher et al., 2019; Miciak & Fletcher, 2020). Reading deficits tend to persist over time, as individuals with reading difficulties early in education tend to consistently have reading difficulties in later grades as well (Francis et al., 1996).

SRD is heterogeneous and arises from diverse and interrelated cognitive, neural, and genetic influences. Determining the interrelated neurobiological mechanisms involved in the development of these disorders will increase understanding of the risk of SRD, and could potentially lead to early identification and classification of SRD. SRD research in particular and LD research in general have involved the use of imaging or genetic approaches since the 1980s, but few studies have combined imaging and structural genetics data to understand the underlying interactions of these contributing biomarkers and neural endophenotypes.

Heritability of SRD

Specific reading disability (SRD) has been shown to have a clear genetic component in its mechanism, with heritability estimates ranging from 0.18 to 0.72 according to a recent review (Mascheretti et al., 2017). In addition, SRD tends to cluster in families, as family members of individuals with SRD have an increased risk of SRD themselves. However, SRD is a genetically heterogeneous disorder, meaning the same phenotype can be produced through different genetic mechanisms. Therefore, there are likely multiple genes that have small effects and interact to cause susceptibility to SRD.

Genetics and Reading

Much of the analysis of the genetic mechanisms of SRD has been focused on nine loci that potentially harbor genetic variation that may be associated with the variation in reading skills; they are named DYX1-DYX9 for their association with dyslexia, and reside on eight different autosomal chromosomes, in 9 chromosomal locations (Mascheretti et al., 2017). There are specific genes within these loci that are related to reading. Within these loci, the genes *DYX1C1*, *DCDC2*, *KIAA0319*, and *ROBO1* have been most replicated by more than one group in more than one sample. In addition, genes not at a DYX locus have been hypothesized to be involved as well, including genes *FOXP2* and *CNTNAP2*, that also play a role in speech and language (Raskind et al., 2013). Genome-wide association scans, which scan markers across the entire genome to find genes associated with a certain disorder, have been used to identify additional risk variants, such as *VEPH1* or *LOC388780* (Gialluisi et al., 2020). Imaging genetics studies may improve detection of genes associated with reading disability because brain structure and function are closer to the level of the gene than behavior and may help improve understanding of the mechanisms and pathways by which genes affect behavior.

Many of the genes associated with SRD affect processes such as neuronal migration, cortical morphogenesis, and ciliary structure and function (Mascheretti et al., 2017). Genetic variants of some genes associated with SRD, including *DCDC2* and *DOCK4*, are involved in neuronal migration and neurite outgrowth (Shao et al., 2015). A genome-wide scan also revealed the involvement of genes *GABARAP*, *NEGR1*, *ACCN1*, *DCDC5*, and *CNTNAP2*, which are involved in learning and reading through their effects on processes such as dendritic spinal plasticity, synaptic transmission, axon guidance, and cell adhesion (Veerappa et al., 2013). Another SRD-risk gene, *ROBO1*, is an axon guidance receptor that helps to regulate the connections between brain hemispheres (Hannula-Jouppi et al., 2005). The association of these genes with reading ability, along with our knowledge of the roles the genes have in the brain and development, provide information about the processes in the brain that may be important for the development of reading.

Early theories suggested that SRD is caused by defects in the migration of neurons in the cortex during development (Galaburda, 1985; Galaburda, 1993). However, Guidi and colleagues (2018) suggest that a general process such as neuronal migration cannot specifically affect SRD and note that some risk variants are also found in non-SRD populations. In addition, the original research by Galaburda and colleagues used small sample sizes and included atypical examples of individuals with SRD (Galaburda, 1985; Galaburda & Kemper, 1979; Humphreys et al., 1990). Other studies have demonstrated that other processes such as axon growth, modulation of synaptic transmission, and structure and function of cilia are associated with SRD as well (Guidi et al., 2018). Therefore, while neuronal migration is one process that may be related to SRD, it is not solely responsible for SRD and may be related to other disorders as well. Recent research

suggested that many genes with functions in various biological pathways (e.g., neuron migration, dendrite development) tend to interact, although neuronal migration was the most common type of biological process pathway implicated (Lancaster et al., 2020).

Genetic studies of reading also suggest that related genes tend to be pleiotropic, that is, tend to influence two or more phenotypic traits. For example, genes associated with reading have also been found to be associated with language ability as well as disorders such as ADHD. A recent genome-wide association study established associations between rapid automatized naming of letters and the genes *MIR924HG* and *NKAIN3*, which were also related to other cognitive traits, educational attainment, and ADHD (Gialluisi et al., 2019). Similarly, another recent GWAS using polygenic risk scores to predict reading found overlap in relevant genes for SRD, ADHD, and ASD (Price et al., 2020). Another SRD-risk gene, *CNTNAP2*, which encodes a cell-surface presynaptic protein implicated in neuronal connectivity, synaptic organization, and migration of neurons in the developing brain, has been implicated in autism spectrum disorder, schizophrenia, intellectual disability, and language impairment (Rodenas-Cuadrado et al., 2014). These examples provide evidence that genes associated with SRD may have an impact on brain processes that underlie other phenotypic traits and neurodevelopmental disorders. Therefore, genes that have been associated with related deficits in language or attention may have an impact on reading as well.

SEMA6D and Cognition

One gene for which the variation, or alteration in DNA sequence at a certain locus, may be associated with individual differences in reading is *Semaphorin 6D (SEMA6D*), located within the DYX1 locus. Semaphorin proteins were originally identified because of their role in

regulating axon guidance. Transmembrane semaphorin proteins provide repulsive axon guidance cues that help axons navigate to the correct target cells by directing them away from certain regions, through interactions with plexin A1 (*plxna1*), a receptor for semaphorins (Leslie et al., 2011). A recent study in rodents demonstrated that mice with mutations in *SEMA6D* had defects in positioning of proprioceptive axons, which carry information about the position and movement of the body, and oligodendrites, which are likely to inhibit synapse formation (Leslie et al., 2011). Semaphorin signaling has been associated with establishing the identity of neuronal cell processes as well as synapse formation, including processes such as synaptic partner choice (which neurons end up forming connections with each other), synapse development, axon pruning, and regulation of dendrite development (Alto & Terman, 2018). In addition, semaphorins may affect synaptic physiology and plasticity in children and adults (Alto & Terman, 2018), as well as immune system functioning (O'Connor et al., 2008; Ito et al., 2014).

More recently, the semaphorin family has been associated with cognition and neurodevelopmental disorders. A boy with autism spectrum disorder and intellectual disability was demonstrated to have a microdeletion of *Semaphorin 5a (SEMA5A)*, which, like *SEMA6D*, also acts as an axon guidance cue (Mosca-Boidron et al., 2016). Another study found that *plexin B3*, a receptor for semaphorins during axon guidance, was related to verbal performance and white matter volume (Rujescu et al., 2007). Regarding *SEMA6D* specifically, a male child with a developmental language disorder was characterized to have a t(10;15)(q24.1;q21.1) translocation, in which a portion of the genetic material on the long arms (indicated by q) of chromosomes 10 and 15 were exchanged, interrupting the *SEMA6D* gene (Ercan-Sencicek et al., 2012). The gene was further investigated in a sample of Russian patients with developmental

language disorder; 2 of 368 were found to have missense mutation, a mutation that results in a substitution of a different amino acid in the protein, in *SEMA6D* while these variants were not found in 1230 control participants (Ercan-Sencicek et al., 2012).

While there is evidence to suggest *SEMA6D* has a role in language and related to other neurodevelopmental disorders, there has been little research on the association between *SEMA6D* and reading ability. However, there is evidence for overlapping genetic risk factors for speech disorders, language impairment, and reading disability, meaning that there are genes that may increase risk for all these disorders and likely affect cognitive skills underlying all of these domains (Pennington & Bishop, 2009). *SEMA6D* has a similar function as many of the other genes implicated in SRD and other neurodevelopmental disorders, in that they are all involved in axonal guidance and synapse formation. In addition, language impairment and SRD share some of the same cognitive deficits, including phonological processing and language fluency (Pennington & Bishop, 2009). Children at family risk of SRD tend to experience delayed language development and demonstrate poor phonological awareness and literacy skills (Snowling & Melby-Lervag, 2016). Genome-wide screenings for genes related to SRD have implicated genes that are involved in learning, cognition, and memory in general (Eicher et al., 2013; Veerappa et al., 2013; Gialluisi et al., 2014). Many genes previously associated with SRD have been demonstrated to influence language skills as well, suggesting the presence of generalist genes that contribute to multiple related traits (Eicher & Gruen, 2015). Therefore, *SEMA6D*, given its association with language and neurodevelopmental disorders such as autism, as well as its critical role in axon positioning, is likely to have an impact on reading as well.

Brain Structure and Reading

To fully elucidate the neurobiological mechanisms contributing to reading, neural substrates should be considered in addition to genetic effects. It is likely that genes are related to reading ability through their effects on endophenotypes such as brain structure. Previous neuroimaging and lesion studies have investigated the structural and functional brain correlates associated with measures of reading ability corresponding with the dual routes of reading, the phonological and orthographic pathways. According to the dual route model of reading, there are two routes involved in reading, one involving semantic and phonological processing, or the processing of basic word sounds and syllables, and the other involving orthographic processing, or the automatic, visual processing involved in sight recognition (Coltheart, 1985). Early readers typically rely on phonological processing to sound out and recognize words at first, and then transition to orthographic processing as their reading becomes more fluent and automatic. Alternatively, the 'triangle' model of reading (Seidenberg & McClelland, 1989) proposed that reading involves distributed patterns of activity across orthographic, phonological, and semantic processing units. This model assumes that there can be a direct pathway between orthography and phonology, as well as an indirect pathway that connects orthography and phonology through semantics (Seidenberg & McClelland, 1989).

Evidence for separate neural routes associated with different types of processing was first noted by Dejerine & Dejerine (1895), who distinguished between effects of ventral and dorsal lesions, which were associated with different forms of SRD. The ventral pathway tends to be associated with orthographic processing during reading, while the dorsal pathway is associated with phonological processing (Fiez & Petersen, 1998). An important part of the ventral pathway

is the visual word form area (VWFA), which is thought to process a representation of letter patterns that allows for automatic recognition of words (Schlaggar & McCandliss, 2007). Recent research on the ventral pathway found that the ventral occipitotemporal cortex could be broken down further into two areas responsible for different aspects of automatic recognition of words, a posterior part responsible for visual word extraction and an anterior part involved in integrating information with the language network (Lerma-Usabiaga et al., 2018). The dorsal pathway, known as the perisylvian region, includes the superior temporal cortex, and supramarginal gyrus, is responsible for linking phonology to orthography to allow the recognition and manipulation of the sounds that make up words (Schlaggar & McCandliss, 2007). In addition, Broca's area in the inferior frontal cortex has been associated with reading as well, responsible for speech production or silent articulatory behaviors (Price, 2012). Activation in the inferior frontal cortex is also related to functions such as verbal working memory, phonological and semantic processing, silent reading, and speech planning (D'Mello & Gabrieli, 2018). Consistent with models of reading, varying regions in the brain are associated with different parts of the reading process, and altogether comprise the reading network.

Imaging work has also compared individuals with SRD to normal readers. Functionally, individuals with SRD tend to have reduced activation in regions that compose of the neural reading network, even when compared with younger readers matched for reading ability (Hoeft et al., 2007; Schlaggar & McCandliss, 2007). Structurally, studies have found atypical brain morphology in individuals with SRD as compared to controls, in regions such as the left temporo-parietal cortex, occipital temporal cortex, cerebellum (Kronbichler et al., 2008), as well as the parietal operculum, corona radiata, and internal capsule (Eckert et al., 2017). Results of a

meta-analysis indicated that there was reduced gray matter volume in the right superior temporal gyrus and left superior temporal sulcus in dyslexic readers, which was consistent with functional imaging studies finding atypical activity in those same regions (Richlan et al., 2013). A meta-analysis of voxel-based morphometry SRD studies demonstrated low gray matter volume in left posterior superior temporal sulcus/middle temporal gyrus regions and left orbitofrontal gyrus/pars orbitalis regions, which was related to reading comprehension ability (Eckert et al., 2016). Across studies, there tended to be reduced gray matter volume in temporal regions identified as part of the reading network for individuals with SRD (although right-lateralized temporal differences were also found by Richlan and colleagues (2013), with corresponding deficits in phonological processing, automatic reading, and other reading related cognitive processes. Across different languages and writing systems, there tends to be underactivation in the reading network (including the occipito-temporal, temporo-parietal, and inferior frontal regions), but the degree of under or over activation in a certain region may differ based on the type of language (Richlan 2020). While the reading network emerging in typically developing individuals has been well characterized, more research using an integrative imaging genetics approach, using both genetic and imaging measures, should be completed to better understand the interrelations between genetic variation and brain morphology, and how these neurobiological substrates affect reading ability.

Imaging Genetic Studies of Reading

Previous research using an imaging-genetic approach to understand the pathway from genes to reading behavior and skills have established genetic associations with brain structure, including white matter and grey matter structure, and function. The literature demonstrates that

there is a relationship between brain structure and function and genes related to SRD. Most imaging genetic studies used a candidate gene approach using genes already found to be associated with SRD or reading. Therefore, the genes most studied include genes from the DYX1 or DYX2 loci, specifically genes such as *DCDC2* and *KIAA0319*.

Some studies have focused on genes and their relationship with voxel-based morphometry, examining outcomes such as grey matter volume or cortical thickness, which will be the focus of the proposed study. For example, one study found that individuals heterozygous for a deletion in intron 2 of *DCDC2* had significantly higher grey matter volumes in reading and language related brain regions, in medial and inferior temporal, fusiform, inferior occipito-parietal, inferior and middle frontal gyri, and hippocampal areas in the left hemisphere (Meda et al., 2008). Eicher and colleagues (2016) examined the effects of DYX2 and DYX3 markers, demonstrating associations between genes (*KIAA0319, FAM65B*, and *ACOT13*) and both fractional anisotropy and cortical thickness in reading regions, including the left orbitofrontal area and left pars opercularis. Another study by Skeide and colleagues (2016) found that the variation in the gene *NRSN1* was associated with the volume of the left fusiform gyrus, or visual word form area. In a study of German participants with or without SRD, there were two grey matter clusters in the left posterior temporal cortex related to verbal working memory, which was correlated with genetic risk variants in *TNFRSF1B* (Männel et al., 2015). Furthermore, Mascheretti and colleagues (2021) demonstrated that the gene *BDNF*'s effect on reading was fully mediated by brain activation, demonstrating the path from gene to brain to reading. Taken together, these studies demonstrate the relationship between genetic variation and brain structure in regions involved in different reading pathways, according to the dual route model. Other

studies focused on genetic associations with other measures of brain structure or function, including variation in white matter volume or changes in activation on functional magnetic resonance imaging (fMRI) or electroencephalography (EEG).

However, a recent imaging genetics review noted that structural findings tend to be mixed for various reasons, including which single nucleotide polymorphisms (variation in a single base pair in DNA) were investigated for a particular gene, the structural metrics used, the age of participants, and the types of samples used (Landi & Perdue, 2019). More research and replication are needed to identify how each gene may affect both structure *and* function in the brain (as often only one is measured), how effects may vary with age of the participants, and how large effect sizes actually are for imaging genetic associations. The current thesis expanded on previous imaging genetics reviews by including a discussion of effect sizes and the differing effects of the genes related to reading ability. Additionally, genes linked to language or genes that may affect general cognitive processes such as memory and attention should be further investigated to understand how they may impact reading through their effects on structure and function in relevant regions of the brain.

Considering the number of genes already identified to be associated with SRD in the genetic literature, the number of studies using an integrative imaging genetics approach is small. However, using a neural endophenotype rather than a behavioral one, closer to the underlying causative genetic code, may help us to understand the pathway from gene to brain structure to behavior, and allow us to more easily detect contributing genes that have small effects. Rose & Donohoe (2012) demonstrated that genetic associations with neural measures tended to have larger effects as compared to cognitive investigations in studies of schizophrenia, but that this

may depend on imaging modality, cognitive functions, and risk variants. Using an integrative imaging genetics approach to identify the effects of *SEMA6D*, a gene known to have effects on language and cognition, allows us to better understand the mechanisms by which *SEMA6D* may influence reading in addition to language, through intermediate effects on neural endophenotypes that contribute to reading ability. Multiple structural phenotypes will be used, including cortical thickness, gyrification, curvature, grey matter volume, and white matter volume in the following studies. Many brain phenotypes have been demonstrated to be genetically related (i.e., to emit genetic correlation), and often are associated with genes involved in brain development and plasticity (Elliott et al., 2018). Individual brain phenotypes, such as cortical thickness and surface area, have been shown to be genetically and phenotypically independent (Winkler et al., 2010). Furthermore, as an example, cortical thickness as well as changes in cortical thickness throughout the lifespan have been asserted to be affected by genetic factors (Fjell et al., 2015). The investigation of effects of this gene's impact on brain structure and reading is highly relevant in understanding the biological correlates of reading.

Clinical Application of Research Findings

Translating research findings about the biological risk factors underlying SRD to a clinical application for individuals or families is highly important. While the literature about imaging genetics of SRD has been growing, there have been fewer studies focused on application of these findings to help individuals and families. Understanding the biological etiology may have many beneficial clinical applications, including early detection of risk, prediction of outcomes, assisting with individualized intervention, and even serving as targets of intervention (Breen et al., 2016). Combining effects of multiple SNPs using a polygenic risk

score, as well as imaging phenotypes through machine learning methods could help to increase confidence in predicting risk of SRD and differentiating it from other related neurodevelopmental disorders (Price et al., 2020; Scarpazza et al., 2020). Preliminary research applying knowledge about the neural and genetic risk factors of SRD and determining if expected regions and genes are related to SRD risk within individuals, an aim of the current study, can improve understanding of the translation of research findings to actual clinical risk. This would help the research literature transition into a more applied, impactful domain, and lead to more literature informing intervention, policy, and evidence-based practices (Khoury et al., 2007).

Aims of empirical studies

SRD are complex neurobehavioral disorders with diverse cognitive, neural, and genetic profiles. The overall objective of the current thesis was to better characterize the underlying biological and neural mechanisms and patterns related to SRD. There has been evidence linking genes to SRD, as well as studies of brain structure and SRD, but imaging genetics related to SRD is a newer field. Therefore, it is relevant to summarize and understand what the current state of the field indicates about the pathway from genes to neural mechanism, to reading phenotype, add to the literature by examining a relevant gene that has not been previously investigated in relation to reading, *SEMA6D*, and apply research findings to better understand biological risk of SRD in a clinical family case study.

Three studies were conducted to complete this objective. The first study was a systematic review and meta-analysis aiming to summarize and interpret current imaging genetics research on SRD. The second study added to this literature through examination of the association of

SEMA6D, as a candidate gene, with variation in brain structure, related to aspects of reading in a sample of children with a broad range of reading and language abilities. The third was a case study examining whether *SEMA6D* and other reading-related genes are associated with reading-related brain structures and reading disability diagnosis in a family with history of reading difficulty. Correspondingly, the following three aims of the thesis are discussed:

The aim of study 1 was to conduct a systematic review on current imaging genetics studies related to reading, with a focus on effect sizes of genes on brain structure and function. The first aim was to conduct a systematic review summarizing overall findings of imaging genetics studies of reading disability, and to examine which genes had the largest effects on brain structure and function. The second aim was to identify future directions for research based on gaps in the literature.

The aim of study 2 was to investigate the association between *SEMA6D* and imaging phenotypes in brain regions of the reading network and investigate whether these imaging phenotypes were related to reading ability. It was hypothesized that variation in *SEMA6D* would be statistically significantly associated with variation in cortical thickness, gyrification, and white matter volume in regions involved in reading and language, including the perisylvian region (the superior temporal cortex, angular gyrus, and supramarginal gyrus). In addition, it was hypothesized that cortical thickness, gyrification, and white matter volume in the regions described would modulate the relationship between genetic variation and performance measures of reading skills, such as word reading, reading fluency, and reading comprehension.

The aim of study 3 was to characterize the role of *SEMA6D* and other reading-related genes in a family case study with a pair of twins and a sibling, where one twin has SRD and the

other does not, and the third sibling has reading difficulties as well, to translate the research literature into a clinical application. The aim was to investigate whether variation in *SEMA6D* and other reading-related genes were related to imaging phenotypes as well as association with the diagnosis of SRD within this clinical family case study. It was hypothesized that variation in *SEMA6D* and other reading-related genes would be related to SRD status and individual differences in reading-related processes. It was also hypothesized that cortical thickness, surface area, and grey matter volume in regions involved in reading language (including the perisylvian region (the superior temporal cortex, angular gyrus, and supramarginal gyrus) would be related to *SEMA6D* and other reading-related genes and reading ability within this family.

PART 1

A Systematic Review and Meta-Analysis of Imaging Genetics Studies of Specific Reading Disorder

Tina Thomas[1]

Shiva Khalaf[2]

Elena L. Grigorenko[1, 3]

Author note
[1]Department of Psychology, University of Houston, Houston, TX, USA. [2]Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX, USA. [3]Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

The preparation of this article was supported by P50 HD052117 (PI: Fletcher), P20 HD091005 (PI: Grigorenko), and P50 HD052120 (PI: Wagner). Grantees undertaking such projects are encouraged to freely express their professional judgment. Therefore, this article does not necessarily reflect the position or policies of the abovementioned agencies, and no official endorsement should be inferred. We thank Irina Ovchinnikova (University of Houston) for her assistance in the literature search.

Manuscript submitted for publication.

Abstract

Specific reading disabilities (SRD) affect 7-10% of individuals in the United States, with significant impacts on academic performance and psychosocial factors. They have complex neurobiological etiologies, characterized by atypical brain structure and function and distinct genetic architecture. The imaging genetics of SRD is an emerging field that aims to better characterize the disabilities' neurobiological causes using both genetic and imaging factors. The aims of the present review were summarizing the current imaging genetics studies of reading disability, characterizing the effect sizes of reported results by calculating Cohen's *d*, completing a Fisher's Combined Probability Test for genes that have been featured in multiple studies, and determining gaps and relevant areas for future research. Results of the review demonstrate associations between reading disability risk genes and reading network brain phenotypes, and the Fisher's test revealed promising results for risk genes such as *DCDC2, KIAA0319, FOXP2, SLC2A3,* and *ROBO1.* Future research should focus on more exploratory approaches to identify previously undiscovered genes, as most of the current studies were focused on candidate genes. Additionally, the use of more comprehensive neuroimaging (e.g., functional and effective connectivity) and genetic (e.g., sequencing and epigenetic analyses) techniques, as well as the use of larger samples, diverse stages of development, and longitudinal investigations, would help researchers better understand the neurobiological correlates of SRD to improve its early identification.

Keywords: imaging genetics, reading disability, imaging, genetics

Introduction

Specific reading disabilities (SRD) are complex neurodevelopmental disorders affecting 7-10% of individuals in the United States, significantly impacting academic and psychosocial functioning (Hulme & Snowling, 2016). Individuals with SRD typically have difficulties with accurate or fluent word recognition and reading (Grigorenko et al., 2020). Reading disabilities can be separated into difficulties in reading domains at the word level, i.e., decoding reading disability, and at the text level, i.e., reading-comprehension reading disability (Fletcher & Grigorenko, 2017). Word reading initially requires phonological awareness or identifying and manipulating the sound structures of speech. Children must then link the orthography of written language to speech structures for automatic recognition of words and fluent reading. At the text level, reading comprehension involves interpretation and understanding of written language; it capitalizes on word reading, but engages many other cognitive processes. Current standards for diagnosing reading disability emphasize response to intervention, which includes provision of increasing intensity of reading instruction to determine if a student improves (Grigorenko et al., 2020). Non-responders to intervention are considered to have a true reading disability (Grigorenko et al., 2020). This approach allows differentiation between having a true reading disability from having a lack of adequate instruction.

Reading disabilities are shaped by diverse and interacting cognitive, genetic, and neural influences. Because of these interacting influences on reading and language difficulties, a multi-level-analysis framework should be used to better understand reading disability and its treatment (Fletcher, 2009; Peterson & Pennington, 2015). These underlying levels of influencing factors include neurobiology, cognitive processes (e.g., phonological awareness and sequential naming),

behavior, psychosocial factors, and the environmental context (e.g. schooling), which must be integrated to understand their interacting effects on the development of reading skills, as well as impairment in the acquisition of reading in SRD (Fletcher & Grigorenko, 2017). Developing understanding of the etiology of reading disorders is important for improving the early identification of those who are at risk. A meta-analysis revealed that 41-74% of the variance in reading achievement and 90% of the variance in reading-related processes such as phonological awareness could be attributed to genetic factors (Grigorenko, 2005). Therefore, the contribution of genetic factors and their interactions with other contributing phenotypes such as measures of brain function or structure, or underlying cognitive processes, may be very informative in understanding risk and etiology of reading disorders, as well as who may be likely to respond to intervention.

Contributing genes may be identified through genome wide association studies (GWAS), which scan the whole genome for relationships between common variants and specific phenotypes, and candidate gene studies, which examine the relationship between a single or few genes with relevant phenotypes. Becker and colleagues (2017) proposed that GWAS and candidate gene studies of reading disability have revealed genes falling under four major categories: 1) Signal transduction (e.g., *NEDD4L*); 2) Neuronal migration and axon guidance (e.g., *DYX1C1*, *DYX2*, *DCDC2*, *KIAA0319*, *ROBO1*, and *SEMA6D*); 3) Cell adhesion, cytoskeleton and division (e.g., *CEP63*, *CNTNAP2*, and *KIAA0319*); and 4) Developmental biology (e.g., *FOXP2*, related to language development). However, most of these genes are likely to contribute small effects to the etiology of reading disabilities in the general population; moreover, inclusion criteria and diagnosis of reading difficulties tends to vary among studies.

Both genetic and phenotypic heterogeneity make it difficult to compare and replicate findings (Becker et al., 2017).

Because reading is a polygenic trait, it is likely that many genes have small effects, meaning there may be additional genes that have not been identified. For example, in a more recent study, Truong and colleagues (2019) conducted a GWAS investigating associations with rapid automatised naming (RAN), which is a predictor of reading disability. Significant associations were found between RAN and rs1555839 (10q23.31), associated with *RPL7P34*, a non-protein coding transcript with unknown function, close to the gene *RNLS* that is related to schizophrenia; furthermore, this SNP was associated with cortical volume in the right inferior parietal lobule (Truong et al., 2019). Another recent GWAS examining genetic effects on RAN found effects of variants within *MIR924HG* (18q12.2), a trend for an association within *NKAIN3* (18q12.3), and an association of rs17663182 (18q12.2) (Gialluisi et al., 2019). Finally, a third recent GWAS suggested a possible link to reading in the *ARHGAP23* gene, which also has functions in neuronal migration and axon pathfinding (Price et al., 2020). Other identified genes included *LINC00935* and *CCNT1* (Price et al., 2020). These recent GWASs have identified more SNPs and genes that were not previously known to be related to reading, suggesting even more complex associations between genetic variation, imaging, and reading.

Neuroimaging studies have investigated associations of brain structure and function with reading, and whether there are differences in brain structure or activity related to reading disability. Structural phenotypes may include grey matter volume, cortical thickness, or cortical surface area, as well as measures of white matter structure, including white matter volume and fractional anisotropy (measuring diffusion of water molecules). Brain activation may be

measured with functional MRI (fMRI), electroencephalography, or magnetoencephalography. Brain activation is typically measured during reading-related tasks, including tasks of word or non-word reading, phonological tasks such as rhyming, or auditory discrimination of sounds.

Neuroimaging studies have generally identified three left hemisphere regions associated with reading that are under-activated or have structural variation in individuals with reading disorders, including (1) a ventral stream involving the occipital region and posterior temporal lobe (occipitotemporal region), (2) a dorsal stream involving the posterior superior and middle temporal gyri, and temporoparietal areas (temporoparietal), and (3) the inferior frontal lobe (Dehaene, 2009; Richlan, 2020), implicated and replicated in multiple meta-analyses and reviews of the literature. For example, results of a meta-analysis comparing functional imaging studies of children and adults with SRD revealed underactivation of a left ventral occipital-temporal region in both children and adults, while underactivation of the left superior temporal gyrus and left inferior frontal gyrus were only found in adults (Richlan et al., 2011). Consistently, a meta-analysis examining grey matter abnormalities in individuals with SRD revealed reduced grey matter volume in the right superior temporal gyrus and left superior temporal sulcus (Richlan et al., 2013). A subsequent review examining white matter structure also found lower fractional anisotropy in left temporoparietal and frontal areas, specifically with involvement of the left arcuate fasciculus and corona radiata (Vandermosten et al., 2012). Wise Younger and colleagues (2017) theorized, based on their results, that the dorsal stream, associated with phonological processing, tends to exhibit higher connectivity earlier during reading when reading is more characterized by phonological processing and sounding out of words, followed by decreases in connectivity as reading ability improves (Wise Younger, Tucker-Drob, & Booth, 2017).

Alternatively, the ventral stream, associated with rapid visual processing of orthographic patterns, tends to increase in connectivity over time, indicating more automatic orthographic processing and improved reading fluency (Wise Younger, Tucker-Drob, & Booth, 2017). While these findings identifying reading network regions have been replicated in many studies, there have also been studies implicating other regions. For example, Jednoróg and colleagues (2015), in their large-scale voxel-based morphometry study across three countries (France, Poland, and Germany), found significant differences in the left thalamus between individuals with SRD and controls, as well as correlations with reading accuracy in the left supramarginal gyrus and left cerebellum. They suggest that these differing results may be due to heterogeneity in languages and orthographies, but also that there may be publication bias and insufficient correction for multiple tests in imaging data analysis (Jednoróg et al., 2015).

While there have been many studies of reading disability investigating either its genetic or neuroimaging bases, there have been fewer studies using an integrative imaging genetics approach. However, imaging genetics approaches may be highly beneficial to the study of the neurobiology underlying reading disability by allowing researchers to understand how genetic influences and neural correlates interact and influence risk. Increasing understanding of the pathway from genetic risk to neural correlates to phenotype can help with early identification of risk factors for reading disability and provide insight into who may benefit from intervention. This is particularly relevant for reading disability because it is influenced by a variety of genes with small effects (Becker et al., 2017). Using neural structure or function as an endophenotype can improve the ability of genome wide association studies (GWAS) to detect these small effects

that may have been missed when using GWAS to directly detect gene associations with reading phenotypes and may provide a clearer understanding of genetic risk for reading disability.

The current imaging genetics of reading disability literature is still an advancing field, with more studies being conducted and published, and with different methodologies being developed. The aim of this systematic review is to (1) summarize the existing findings of the field of imaging-genetics related to reading disability, with summaries of evidence on each identified gene, (2) examine the effect sizes and significance of the studied genes, (3) identify current gaps in the literature in terms of methodology and predictors used, and (4) propose future directions for the use of imaging-genetic techniques for the characterization of reading disability.

Methods

Search Strategy

A systematic literature search was performed to identify studies using an imaging genetic approach to study reading disability or language disorders (Figure 1). PSYCINFO and PubMed databases were searched using combinations of the following key words: “neuroimaging,” “fmri,” “mri,” “functional magnetic resonance imaging,” “brain structure,” “diffusion tensor imaging” “electroencephalography” “neurogenetics,” “genetics,” “genes,” “reading disability,” “reading disabilities,” “reading difficulties,” “dyslexic,” “dyslexia,” and “language disorder.” Combinations included (a) reference to neuroimaging, (b) reference to genetics, and (c) clinical condition (“reading disability”). This search was carried out from April 2018 to April 2019.

Studies were included if they were empirical articles using a neuroimaging genetics approach to investigate the biomarkers related to reading, and to determine associations with

imaging phenotypes. These studies could include candidate gene studies investigating the neural correlates of previously determined reading disability risk genes in the DYX loci in typically developing or reading impaired individuals, or they could include genome-wide association studies aiming to identify genes and neural markers in individuals with reading difficulties. Exclusion criteria were: (a) studies that included participants with another comorbid disorder (e.g., schizophrenia), (b) studies focused on language without a measure of reading, (c) animal studies, (d) studies comparing imaging phenotypes in those "at risk" or "not at risk" for reading disability, without investigating associations with genes, and (e) articles focused only on language without reading measures. There was no date restriction on publication date of articles. Overall, 26 articles were retained.

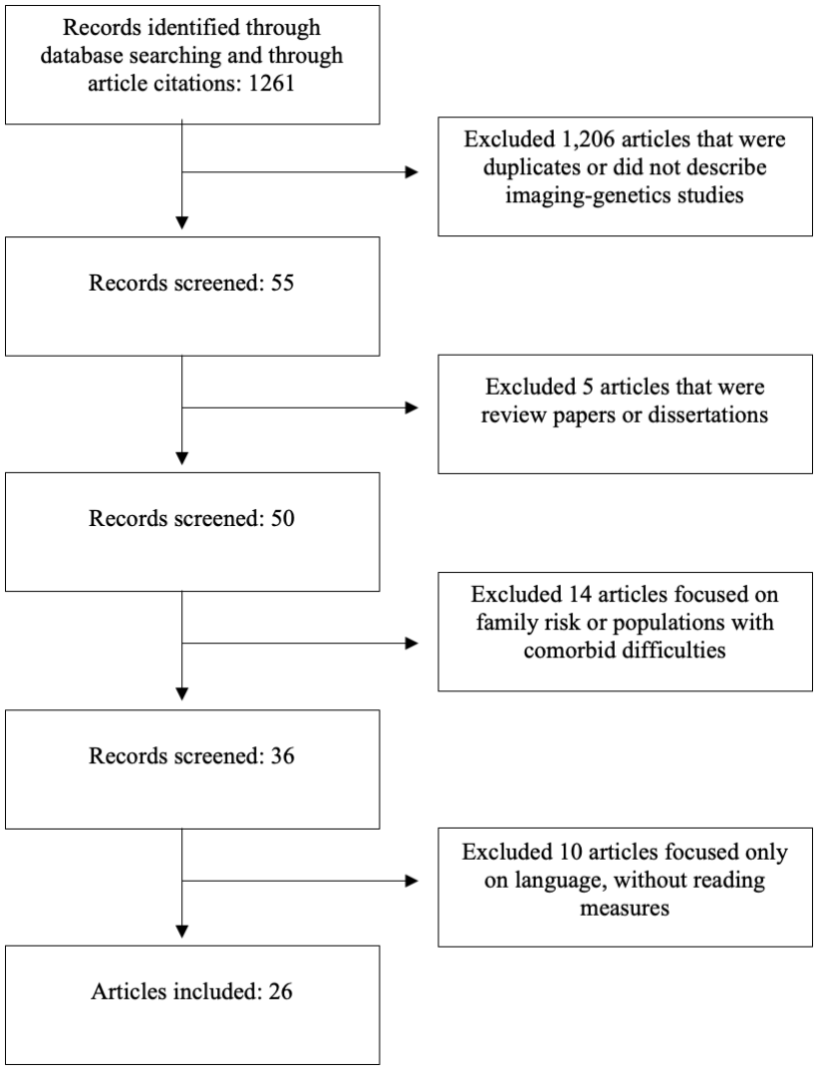

Figure 1. Flow Diagram of Search Strategy.

Interpretation and Analyses

Effect sizes were coded for the associations between imaging measures and genetic variation for each article. When possible, effect sizes were converted into Cohen's *d* to allow comparison of effects across studies (Supplementary Table 1), using The Practical Meta-Analysis Effect Size Calculator (Wilson, 2018) or calculating by hand. A Fisher's combined

probability test was performed to determine overall significance of results. Fisher's method is used to statistically combine the results from independent tests, and compared with a chi-squared distribution with 2k degrees of freedom (k=number of tests). The analysis was carried out first using both significant and reported non-significant p-values. Fisher's method was then used to find overall significance for each gene for genes with results reported by 2 or more studies. However, because the reported p-values are likely not independent, and Fisher's method tends to be more sensitive to smaller p-values, these results should be interpreted with caution (Heard & Rubin-Delanchy, 2017). Additionally, not all studies reported p-values for non-significant results. Due to lack of independence, combination of p-values over varying methodologies and imaging phenotypes, and publication bias, Fisher's analysis results are likely to be an overestimation of significance. When possible, additional Fisher's analyses were conducted focused on specific genes and brain regions to better examine specific effects, but due to publication bias and variable SNPs and phenotypes, these results are still likely to be an overestimation of significance. Gene functions were identified using GeneCards - the human gene database (www.genecards.org; Stelzer et al., 2016) or the Genetics Home Reference (National Library of Medicine).

Results

A total of 26 articles were eligible for inclusion in this review. These have been summarized and organized by the candidate gene under investigation (Supplementary Table 1). Information regarding the study population, genes studied and neuroimaging technique are summarized in Table 1. Of the articles identified, 25 out of 26 used a candidate gene approach or a study of a sample carrying a specific mutation or polymorphism. Only one study used an

exploratory GWAS approach to identify new genes that may be associated with imaging phenotypes in reading (Roeske et al., 2011). In regard to neuroimaging techniques, twelve studies used functional imaging (including fMRI and EEG studies), twelve used structural imaging, and two included both functional and structural imaging. Thirteen studies used a whole-brain approach, which was accompanied by further analysis of regions of interest in four studies. All other studies focused on specific brain regions of interest. Sample numbers ranged from 26 to 1,275. Only two studies included samples larger than 500, and 16 had samples smaller than 100. Seven of the studies were conducted using an adult sample, thirteen with children, and six with both children and adults or young adults. Overall, studies were variable in methods used, as well as which genes were analyzed.

Cohen's *d* was calculated for all studies that reported statistics that were able to be converted (Supplementary Table 1). In total, 126 statistics reported were converted to Cohen's *d*, which are reported throughout the following text. Depending on the study and gene, Cohen's *d* ranged from very small to very large. However, because many of the studies only reported p-values, the Fisher's combined probability test was used to determine overall p-values for genes investigated by at least two studies. Overall, a total of 185 associations between genes and imaging phenotypes were used in the combined analyses. Both the results of the imaging-genetic studies of reading disability and the results of the Fisher's combined probability test are summarized below, organized by candidate gene in order of position on the chromosome. When a gene or a marker has been interrogated in a single publication only, the results from the corresponded papers are presented.

Table 1

Candidate genes examined in imaging-genetics studies. Only first author and year reported (See reference list for full reference). GM = grey matter. WM = white matter. CT = cortical thickness.

Locus	Gene	Location	Reference	Imaging Phenotype	N	Age Range (years)	Language
DYX8	*TNSFRSF1B*	1p36.22	Männel 2015	GM volume	32	Not reported: Mean 26	German
DYX8	*RCAN3*	1p36.11	Skeide 2016	GM & WM volume	141	9-12	German
DYX3	*MRPL19* and *GCFC2*	2p12	Eicher 2016	Volume, CT, WM structure	332	3-20	English
DYX3	*MRPL19*	2p12	Scerri 2012	WM structure	76	6-25	English
			Müller 2017	Brain Activation - EEG	67	9-10	German
			Eicher 2016	CT & GM volume	332	3-20	English
DYX5	*ROBO1*	3p12.3	Lamminmäki 2012	Brain activation - MEG	20	18-51	Finnish
			Skeide 2016	GM and WM volume	141	9-12	German
			Sun 2017	WM structure	115	10-15	Mandarin
None	*CEP63*	3q22.2	Einarsdottir 2015	WM volume	62	7-92	Swedish
None	*CLSTN2*	3q23	Roeske 2011	Brain activation – EEG	386	8-19	German
DYX2	*DCDC2*	6p22	Meda 2008	GM volume	69	20-85	English
			Darki 2012	WM volume	76	6-25	Swedish
			Darki 2014	WM volume, DTI, CT	76	6-25	Swedish
			Marino 2014	WM structure	47	16-21	Italian
			Eicher 2016	WM structure & CT	332	3-20	English
			Männel 2015	GM volume	32	Not reported: Mean 26	German
			Müller 2017	Brain activation – EEG	67	9-10	German
			Cope 2012	Brain activation – fMRI	82	7-12	English
			Czamara 2011	Brain activation – EEG	200	8-19	German
			Neef 2017	Brain activation – EEG	159	4-7, 11-13	German
			Skeide 2016	GM and WM volume	141	9-12	German
			Su 2015	Brain activation – EEG	60	Close to 12	Chinese
DYX2	*KIAA0319*	6p22	Eicher 2016	CT and WM structure	332	3-20	English
			Centanni 2018	Brain activation – MEG	32	7-14	English
			Darki 2012	WM volume	76	6-25	Swedish
			Darki 2014	WM volume, DTI, CT	76	6-25	Swedish
			Männel 2015	GM volume	32	Not reported: Mean 26	German
			Cope 2012	Brain activation – fMRI	82	7-12	English
			Neef 2017	Brain stem responses	159	4-7, 11-13	German
			Skeide 2016	GM and WM volume	141	9-12	German
			Pinel 2012	Brain activation – fMRI	94	Not reported: Mean 25	French
DYX2	*TTRAP*	6p22	Cope 2012	Brain activation – fMRI	82	7-12	English
			Pinel 2012	Brain activation – fMRI	94	Not reported: Mean 25	French
			Müller 2017	Brain activation – EEG	67	9-10	German
			Skeide 2016	GM and WM volume	141	9-12	German
DYX2	*THEM2*	6p22	Pinel 2012	Brain activation – fMRI	94	Not reported: Mean 25	French
			Eicher 2016	CT	332	3-20	English
DYX2	*FAM65B*	6p22	Eicher 2016	CT & WM structure	332	3-20	English
DYX2	*NRSN1*	6p22.3	Skeide 2016	GM and WM volume	141	9-12	German

DYX2	Non-gene associated	7p12	Gialluisi 2016	Surface area and CT	1275	18-83	Dutch
None	*FOXP2*	7q31.1	Müller 2017	Brain activation – EEG	67	9-10	German
			Skeide 2016	GM and WM volume	141	9-12	German
			Pinel 2012	Brain activation – fMRI	94	Not reported: Mean 25	French
None	*CCDC136*	7q32.1	Gialliusi 2016	Surface area, CT	1275	18-83	Dutch
None	*FLNC*	7q32.1	Gialliusi 2016	Surface area, CT	1275	18-83	Dutch
None	*DGKI*	7q33	Skeide 2016	GM and WM volume	141	9-12	German
None	*CREB3L2*	7q33	Skeide 2016	GM and WM volume	141	9-12	German
None	*CNTNAP2*	7q35-36.1	Müller 2017	Brain activation - EEG	67	9-10	German
			Skeide 2016	GM and WM volume	141	9-12	German
None	*BDNF*	11p14.1	Jasińska 2016	Brain activation – fMRI	81	6-10	English
None	*SLC2A3*	12p13.31	Roeske 2011	Brain activation – EEG	386	8-19	German
			Skeide 2015	WM structure	34	9-12	German
None	*COL4A2*	13q34	Skeide 2016	GM and WM volume	141	9-12	German
None	Deletion	15q11.2(BP1-BP1)	Ulfarsson 2017	GM and WM volume and activation – fMRI	714	18-65	Icelandic
DYX1	*CYP19A1*	15q21.2	Skeide 2016	GM and WM volume	141	9-12	German
DYX1	*DYX1C1*	15q21.3	Darki 2012	WM volume	76	6-25	Swedish
			Darki 2014	WM volume, DTI, CT	76	6-25	Swedish
			Müller 2017	Brain activation – EEG	67	9-10	German
			Männel 2015	Grey matter probability	32	Not reported: Mean 26	German
			Skeide 2016	GM and WM volume	141	9-12	German
None	*CMIP*	16q23.2	Skeide 2016	GM and WM volume	141	9-12	German
			Müller 2017	Brain activation – EEG	67	9-10	German
None	*ATP2C2*	16q24.1	Müller 2017	Brain activation – EEG	67	9-10	German
			Skeide 2016	GM and WM volume	141	9-12	German
DYX6	*EPB41L3*	18p11.31	Skeide 2016	GM and WM volume	141	9-12	German
None	*SETBP1*	18q12.3	Perdue 2019	Brain activation – fMRI	135	5-12	English
None	*DYM*	18q21.1	Skeide 2016	GM and WM volume	141	9-12	German
None	*MYO5B*	18q21.1	Müller 2017	Brain activation – EEG	67	9-10	German
			Skeide 2016	GM and WM volume	141	9-12	German
None	*NEDD4L*	18q21.31	Müller 2017	Brain activation – EEG	67	9-10	German
			Skeide 2016	GM and WM volume	141	9-12	German
None	*COMT*	22q11.21	Landi 2013	Brain activation – fMRI	86	6-10	English
None	*RBFOX2*	22q12.3	Gialliusi 2016	CT	1275	18-83	Dutch
None	*SYN1*	Xp11.23	Cabana 2018	NODDI – microstructure	26	18-67	French

Chromosome 1

TNSFRSF1B

TNSFRSF1B codes for a protein that has a role in preventing apoptosis in neurons, specifically through stimulating antioxidative pathways (Stelzer et al., 2016). Männel and

colleagues (2015) investigated the role of *TNFRSF1B*-rs496888 (chr1:12172749), an intron variant, in reading disability by first doing a whole-brain analysis of grey matter volume associated with reading difficulty, followed by a region of interest (ROI) analysis of the association between identified brain regions and *TNFRSF1B*-rs496888 in 32 adult German males. Results revealed a trend suggesting that the presence of risk alleles in *TNFRSF1B*-rs496888, an intron variant, was related to reduced gray matter volume in the left Heschl's gyrus/posterior superior temporal gyrus (STG; Cohen's d = -0.70), and increased gray matter volume in the left posterior superior temporal sulcus (Cohen's d = 0.65; Männel et al. 2015). This pattern of gray matter volume was related to successful classification of participants with reading disability against control participants based on genetic risk (Männel et al., 2015).

RCAN3

RCAN3 is a protein coding gene that has been associated with Down Syndrome (Stelzer et al., 2016). The SNP rs196402 (chr1:24512003), an intron variant of *RCAN3* on chromosome 1, was investigated for associations with grey and white matter volume clusters, but there were no significant findings (Skeide et al., 2016).

Chromosome 2

MRPL19/GCFC2 locus

MRPL19 codes for a mitochondrial ribosomal protein that helps with protein synthesis, while *GCFC2* codes for a protein that suppresses transcription (Stelzer et al., 2016). Both genes are coregulated and have been associated with reading disability (Anthoni et al., 2007). Scerri and colleagues (2012) examined the associations of 7 SNPS related to the *MRPL19* and *GCFC2* genes with whole-brain white matter volume in a Swedish sample of 76 children and adults. The

SNP rs917235 (chr2:75598693), related to verbal IQ, was associated with variation in white matter volume in the posterior part of the corpus callosum and the cingulum, connecting the right and left postcentral gyrus, superior parietal lobule, precuneus, lateral occipital cortex, and fusiform gyrus (Scerri et al., 2012). While this gene is considered a "reading disability risk variant," its role in reading disability is unclear because it was associated with the phenotype of verbal IQ and not with reading measures in this sample (Scerri et al., 2012). A multi regression analysis of genotypes associated with fiber tracts and ROIs previously implicated in reading and language in a sample of 332 children revealed a suggestive association of *GCFC2*-rs2298948 (chr2:75699439) with cortical thickness and volume in right temporal regions, including cortical thickness in the middle temporal gyrus and cortical volume in the right inferior temporal region (Eicher et al., 2016). An analysis of another SNP in the same chromosomal region, rs1000585 (chr2:75596036), using a multifactorial linear regression model adjusted for poor spelling in a sample of 67 children, registered no significant associations between rs1000585 and an EEG mismatch response signal in a frontal ROI (F3, Fz, F4) related to auditory discrimination capabilities (Müller, 2017). Whereas there were 3 imaging genetics studies investigating the role of the *MRPL19/GCFC2* locus in reading disability, there was no overlap in specific SNPs examined, and imaging phenotypes were variable. The current meta-analysis of p-values across studies, phenotypes (16 cortical volume phenotypes, 1 cortical thickness measure, 1 EEG MMR signal, and 1 white matter volume phenotype) and brain regions indicated that the overall p-value for associations between MRPL19-C2ORF and imaging phenotypes across the three studies was p=0.04 using Fisher's combined probability test. This value summarized 19 reported p-values

(17 reported by Eicher et al., 2016, 1 reported by Müller et al., 2017, and 1 reported by Scerri et al., 2012).

Non-gene associated

Some studies investigated SNPs that were not associated with a specific gene. Eicher and colleagues (2016) used a regression to investigate effects of two such SNPs, rs917235 and rs6732511, in the DYX3 locus on chromosome 2p12-2p13. Significant associations were found between rs917235 and cortical thickness in the left middle temporal gyrus, as well as between rs6732511 and cortical volume in the right fusiform gyrus (Eicher et al., 2016).

Chromosome 3

ROBO1

ROBO1, a gene involved in axon growth across the midline of the central nervous system, has been identified as a susceptibility gene for reading disability (Hannula-Jouppi et al., 2005). A significant association was found using a general linear model between the rs4535189/rs6803202 haplotype of *ROBO1* and fractional anisotropy (T/T>C/T p=0.006; T/T>C/C p=0.007) and radial diffusivity (C/T>T/T p=0002; C/C>T/T p=0.011) near the posterior part of the corpus callosum and axial diffusivity (C/T>C/C p=0.001; T/t>C/C p=0.017) in the anterior portion (Sun et al., 2017). Further, the axial diffusivity cluster around the genu significantly mediated the effect of the *ROBO1* gene on reading performance (Sun et al., 2017). Consistently, another study had previously demonstrated that *ROBO1* expression was correlated with reduced ipsilateral suppression of auditory cortex steady-state responses, indicating issues with axonal midline crossing across the corpus callosum (Lamminmäki et al., 2012). *ROBO1* had no significant association with intracranial volume, cortical thickness, or surface area (Sun et al.,

2017). Results of the current meta-analysis indicate that across both of these studies (2 phenotypes by Lamminmäki et al., 2012 and 12 by Sun et al., 2017), the overall p-value was $p<0.00001$ overall for reported associations between the gene and imaging phenotypes. Imaging phenotypes included MEG ipsilateral suppression (n=2), fractional anisotropy (n=4), radial diffusivity (n=4), and axial diffusivity (n=4). Skeide and colleagues (2016) also examined the effects of 9 SNPs related to *ROBO1* on grey matter and white matter volume and found no significant effects.

CEP63

The *CEP63* gene encodes a protein involved in microtubule-organizing and DNA damage response in the centrosome (Stelzer et al., 2016) and has been proposed as a candidate risk gene for reading disability (Einarsdottir et al., 2015). In a sample of 76 healthy controls, the AA and AC genotypes of the SNP rs761945 in *CEP63* were significantly correlated with higher white matter volume in the right and left hemisphere, in the right superior longitudinal fasciculus and the posterior part of the corpus callosum (Einarsdottir et al., 2015). The cluster identified in the right temporo-parietal region partially overlaps with a region previously found significant for *DYX1C1* and *KIAA0319* (Einarsdottir et al., 2015). In addition, the association between rs7619451 and reading comprehension was found to be significant using a mixed linear model (Einarsdottir et al., 2015).

CLSTN2

The *CLSTN2* gene has been associated with calcium ion binding and motor system functioning (Stelzer et al., 2016). In an initial sample of 200 children with dyslexia and a replication sample of 186 children with dyslexia, an analysis of variance revealed significant

associations between two SNPs (both intron variants) and mismatch negativity signaling (MMN ERP) reflecting automatic speech deviance processing (rs1365152 p=1.44e-04; rs2114167 p=6.82e-04; Roeske et al., 2011).

Chromosome 6

DCDC2

DCDC2 is the candidate gene most favored in the imaging-genetic articles identified, part of the DYX2 locus. *DCDC2* encodes a protein that is thought to be involved in neuronal migration through its functions binding tubulin and enhancing microtubule polymerization (Stelzer et al., 2016). Meda and colleagues (2008) examined the association between an intron 2 deletion in *DCDC2* and grey matter volume in a sample of 56 adults, finding a positive correlation between suprathreshold volume and *DCDC2* expression levels (Cohen's $d = 1.5$). Significant differences in grey matter volume for individuals heterozygous for the deletion compared to individuals homozygous for no deletion were found in the inferior temporal cortex, as well as other frontal, parietal, and temporal regions of the reading network, with Cohen's *d* ranging from 0.66-0.96 for significant associations (Meda et al., 2008). Männel and colleagues (2015) focused their analysis of *DCDC2* on associations of grey matter probability with the SNP rs71745442 in a sample of 32 adult males, half with a reading disability. A regression analysis revealed non-significant associations with the right lateral occipital cortex (LOC; Cohen's $d = 0.45$), left Heschl's gyrus/posterior superior temporal gyrus (HG/pSTG; Cohen's $d = 0.43$), and left superior temporal sulcus (pSTS; Cohen's $d = 0.22$; Männel et al., 2015).

Other studies focused on other types of imaging phenotypes, including white matter volume and structure and cortical thickness. Marino and colleagues (2014) investigated the

association between the gene's intron 2 deletion and white matter structure in 47 children and young adults in 4 groups, individuals with dyslexia with and without the *DCDC2* deletion, and normal readers with and without the *DCDC2* deletion. Results suggested significant differences in fractional anisotropy of the left temporal segment of the arcuate fasciculus and the splenium of the corpus callosum between the four groups, with pairwise comparisons ranging from Cohen's *d*= 1.06-2.89 (Marino et al., 2014). Darki and colleagues (2012) investigated associations of 7 intron variants in *DCDC2* (rs793842, rs793862, rs807701, rs2328819, rs2792682, rs7751169, and rs9460974) with white matter volume and cortical thickness using a flexible factorial design, in a sample of 76 children and young adults ranging from 6-25 years. The SNP rs793842 was associated with white matter volume in a left temporo-parietal region (Cohen's *d* = 1.11; Darki et al., 2012). Darki and colleagues (2014) further investigated the associations between the SNP rs793842 and imaging regions of interest, including white matter volume, fractional anisotropy, and cortical thickness, using a mixed linear model using age, sex, their interactions by the SNPs, and handedness as covariates. Significant associations were found with white matter in the superior longitudinal fasciculus and posterior corpus callosum (Cohen's *d* = 1.11) and cortical thickness in the left lateral cortical region (Cohen's *d* = 0.81), the left supramarginal gyrus (SMG; Cohen's *d* = 0.84), the left angular gyrus (AG; Cohen's *d* = 0.70), and the left lateral occipital cortex (LOC; Cohen's *d* = 0.89; Darki et al., 2014). Further, there was a significant interaction with age on thickness in the left SMG and left LOC (Darki et al., 2014).

However, conflicting evidence from other studies found a lack of association between *DCDC2* with grey and white matter structure. In a sample of 141 children, 6 *DCDC2* SNPs (rs793842, rs807702, rs807724, rs1091047, rs6922023, and rs1087266) were not significantly

associated with grey matter or white matter volume clusters using a multi-locus model (Skeide et al., 2016). In a sample of 332 children aged 3-20 years, the SNP rs707864 was not significantly associated with cortical volume, cortical thickness, or fractional anisotropy in reading-related ROIs (Eicher et al., 2016). However, all of the studies conducted used different methods, SNPs, and populations of varying ages, which may have contributed to this variation.

Other studies investigated relationships between *DCDC2* and functional activation of the brain. For example, Cope and colleagues (2012) investigated the associations between 4 intron variants (rs793862, rs87701, rs807724, rs1087266), an intronic single tandem repeat (BV677278), and an intron 2 deletion (BV677278) with fMRI brain activation during auditory categorization, print categorization, word rhyming, and nonword rhyming tasks. Associations were found with reading-related ROIs during rhyming tasks, with the most significant associations between *DCDC2* and the superior anterior cingulate gyrus (SAC w/ allele 5 of BV677278; Cohen's d = 0.19) during word rhyming, and the posterior cingulate gyrus (PC w/ allele 8 of BV677278: Cohen's d = -0.04), left paracentral lobule (LPC w/ deletion of BV677278: Cohen's d = 0.08), and left inferior frontal gyrus, inferior aspect (LIFGI w/ rs1087266: Cohen's d = 0.04) during non-word rhyming tasks, but these associations did not remain after correction for multiple testing (Cope et al., 2012). During categorization tasks, the strongest associations were between the BV677278 complex tandem repeat and activation with the left anterior inferior parietal lobe (LAIPL w/ allele 4 of BV677278: Cohen's d = 0.02 and the right lateral occipital temporal gyrus (RLOTG w/ allele 8 of BV677278: auditory categorization Cohen's d = -0.67, print categorization Cohen's d = -0.55; Cope et al., 2012).

Müller and colleagues (2017) used EEG to investigate the association between 4 *DCDC2* SNPs (rs1419228, rs7765678, rs793862, and rs807701) with a mismatch response signal (MMR) in the frontal region of the brain, which is also representative of auditory discrimination capabilities, but found no significant results (Cohen's *d* range -0.20-0.30). Another EEG study used MMN, a negative ERP component that is generated by the automatic response of the brain to a mismatch in auditory stimulation, to investigate reading disability (Czamara et al., 2011). This MMN component is the brain's response to an auditory mismatch of an odd syllable in a sequence of stimuli, and may be a neural marker of phonological processing deficits (Stoodley et al., 2006). This study revealed that rare variants in an intron of *DCDC2* and in an intergenic region between *DCDC2* and *KIAA0319* were associated with attenuated late MMN amplitude during a passive oddball paradigm listening to standard (/da/) and deviant (/ba/) syllables, indicating difficulty with auditory discrimination (Czamara et al., 2011). In contrast, genetic risk of *DCDC2* has been associated with increased brainstem response consistency when listening to a target syllable (/da/), suggesting that *DCDC2* risk variants may be protective for early auditory sensing (Neef et al., 2017).

Another EEG study investigated associations between two intron variants (rs1419228 and rs1091047) and an N170 component during a visual-word color decision tasks, related to orthographic processing, and demonstrated no significant genotype effects on peak amplitude and mean amplitude, Cohens *d* range 0.23-0.48 (Su et al., 2015). However, there was a significant interaction between rs1091047 and home literacy on changes of N170 in the left hemisphere (Cohen's $d = 0.74$) and mean amplitude in the left hemisphere (Cohen's $d = 0.71$; Su et al., 2015).

The results of the current meta-analysis indicated genetic associations of *DCDC2* with imaging phenotypes was found to be p<.00001, across 10 studies with various methodologies and in various regions of the brain. This statistic summarized 40 p-values of *DCDC2* associations with phenotypes, including fMRI activation (n=13), EEG signals (n=11), brain stem response consistency (n=3), white matter volume (n=2), cortical thickness (n=5), grey matter volume (n=3), and grey matter probability (n=3). When possible, p-values of genetic associations between *DCDC2* and specific replicated brain regions were summarized for a more accurate understanding of combined effects within specific regions and genes. For example, the overall summary *p*-value for genetic association with the right temporal-occipital region and left temporal-occipital regions were p<.0001 across 5 statistics in 3 studies and 6 statistics in 3 studies, respectively. In addition, when p-values of associations broadly in the temporo-parietal region were combined, the significance was also p<.00001.

KIAA0319

KIAA0319, another of the genes on the DYX2 locus, codes for a protein that is involved with brain development through regulation of neuronal migration and cell adhesion, and has also been identified as a candidate risk gene for reading disability (Stelzer et al., 2016). Imaging genetics studies have been conducted on both brain structure and function related to *KIAA0319*. For example, a regression analysis revealed non-significant associations between an intron variant (rs6935076) of *KIAA0319* and grey matter probability in regions of interest in the brain (Männel et al., 2015). Eicher and colleagues (2016) examined cortical volume, cortical thickness, and fractional anisotropy in a sample of 332 children, but only found significant effects on cortical thickness and fractional anisotropy. The SNP rs9461945 in the *KIAA0319* gene was

associated with fractional anisotropy (FA) in the corpus callosum in a regression analysis, but this association was non-significant when overall FA was included, indicating that the gene could be associated with global FA effects rather than regional (Eicher et al., 2016). The same SNP has been associated with cortical thickness in the left orbitofrontal region (Eicher et al., 2016). Another SNP in *KIAA0319*, rs6935076 (intron variant), was associated with white matter volume in the left temporo-parietal region (Cohen's d = 1.07) in a sample of 76 children and young adults, and this increased white matter volume was associated with improved reading (Darki et al., 2012). The same intron variant was associated with white matter volume in the bilateral superior longitudinal fasciculus and posterior corpus callosum (Cohen's d = 1.54 and 1.04 respectively; Darki et al., 2014). However, Skeide and colleagues (2016) examined the effects of 3 intron variants (rs2179515, rs761100, and rs6935076), and did not find any significant effects on grey or white matter volume in a sample of 141 children.

Studies also investigated how variation in *KIAA0319* affects brain activation, as measured by MRI, EEG, or MEG. For example, Centanni and colleagues (2018) used an analysis of variance to examine the relationship between two *KIAA0319* SNPs (rs6935067 and rs761100) and MEG-measured neural response to visual and auditory stimuli, specifically looking at neural variability in the inferior frontal gyrus and superior temporal gyrus in a sample of 32 children (20 with SRD, 12 without). A significant main effect was found for rs6945067 (Cohen's d = 0.88), but not for rs761100 (Cohen's d = 0.19; Centanni et al., 2018). In addition, another study revealed that when listening to a target syllable (/da/), genetic risk of *KIAA0319* (determined with a principal components analysis to account for linkage disequilibrium) was associated with brainstem response inconsistency (Cohen's d = -0.39), indicating that carriers of the *KIAA0319*

risk gene may be likely to have difficulty processing of speech stimuli at an early age (Neef et al., 2017). Conversely, Müller and colleagues (2017) found that variance in 4 *KIAA0319* SNPs was not significantly associated with an EEG mismatch response signal related to auditory discrimination in a sample of 67 children. KIAA0319 was also not related to MRI brain activation in a sample of 82 children (Cope et al., 2012) or EEG brain activation in a sample of 200 children (Czamara et al., 2011).

In the current meta-analysis, across all reported analyses in 7 studies investigating *KIAA0319*, the overall p-value was p<.00001. This statistic summarized 47 associations with various imaging phenotypes, including MEG neural variability (n=6), EEG signal (n=4), brain stem response consistency (n=1), white matter volume (n=2), fractional anisotropy (n=16), cortical thickness (n=15), and grey matter probability (n=3). When possible, p-values of genetic associations between *KIAA0319* and specific replicated brain regions were summarized as well. For example, when results were narrowed to only those in the left superior temporal gyrus, the overall p-value was 0.012, combining 8 statistics over 3 studies. Summarized associations with the superior longitudinal fasciculus were significant at p<.00001, combining 5 statistics over 2 studies. However, associations with the right occipital region were summarized as p=.41, which was not statistically significant, across two statistics from two studies.

TTRAP/TDP2

TDP2, also known as *TTRAP*, is a gene that codes for a phosphodiesterase, an enzyme that breaks phosphodiester bonds, which is thought to be involved with DNA repair. Research on imaging genetic associations of *TDP2* has generally produced non-significant associations. In a sample of 82 children, Cope and colleagues (2012) examined correlations between the SNP

rs2143340, an intron variant, and fMRI brain activation in 16 regions of interest, finding nominal associations with activation in the right (Cohen's d = 0.06) and left anterior inferior parietal lobe (Cohen's d = 0.05) during a word rhyming task. However, this association was non-significant after FDR correction (Cope et al., 2012). Müller and colleagues (2017) found no significant regression associations between an EEG mismatch response signal in frontal regions of interest (indicative of auditory discrimination capabilities) with rs9461045 (Cohen's d=0.41) and rs32122336 (Cohen's d = 0.41) of *TDP2* in a sample of 67 children. In the current meta-analysis, summarizing 4 phenotypes (2fMRI and 2 EEG) across 2 studies (Cope et al., 2012 and Müller et al., 2017), the overall p-value for associations with brain activation was p=0.0007. However, two other studies, one investigating the effects of *TDP2* SNPs on fMRI brain activation in 94 adults (Pinel et al., 2012), and another examining effects of a *TDP2* intron variant on grey and white matter volume in a sample of 141 children (Skeide, 2016) also did not find significant associations and did not report p-values.

THEM2/ACOT13

Two studies investigated *ACOT13*, also called *THEM2*, which encodes a protein involved in cell proliferation (Stelzer et al., 2016). A SNP, rs17243157, in *ACOT13* has been associated with asymmetry of activation in the posterior superior temporal sulcus (pSTS) during a reading and listening task (Cohen's d = 1.13; Pinel et al., 2012). In addition, a different SNP in *ACOT13*, rs3777662, was associated with cortical thickness in the left pars opercularis (Eicher et al., 2016). In the current meta-analysis, the overall significance calculated through Fisher's method was p=0.000015. This statistic summarized associations with 19 phenotypes (14 cortical thickness and 5 fMRI phenotypes) across 2 studies (Eicher et al., 2016 and Pinel et al., 2012).

Combining statistics of associations focused in the left superior temporal region resulted in a combined p-value of 0.002.

FAM65B

FAM65B is a gene that codes for a protein important for hearing, as it is required for the development of hair cell stereocilia. In an imaging genetics analysis in a sample of 332 children by Eicher and colleagues (2016), *FAM65B* was not significantly associated with cortical thickness, but the SNP rs9348646 was associated with fractional anisotropy in the left superior longitudinal fasciculus (SLF) and right temporal SLF. This effect was non-significant when overall FA was included, indicating that *FAM65B* may have global FA effects rather than regional (Eicher et al., 2016).

NRSN1

NRSN1 is a gene that codes for a protein that is thought to play a role in transduction of nerve signals, nerve growth, and neurite extension, and may play a role in memory (Stelzer et al., 2016). In a sample of 141 children, *NRSN1* has been associated with gray matter volume in the right dorsal parieto-occipital cortex, left occipital cortex, and the left temporo-occipital fusiform cortex, including the visual word form area, VWFA (Skeide et al., 2016). In addition, *NRSN1* and volume of the VWFA were significantly associated with reading performance (Skeide et al., 2016). *NRSN1* has also been associated with the white matter volume of the left postcentral cortex (Skeide et al., 2016). In a later abstract by the same group, *NRSN1* was significantly associated with the volume of the left FFG, and using these volumetric profiles determined by genetic association could predict reading disability 10 months before school entry with a classification accuracy of 75% (Skeide et al., 2017).

Non-gene associated

Gialluisi and colleagues (2016) investigated effects of LOC105375496, 2kb downstream of rs59197085, but associations with surface area and cortical thickness were not significant in their sample of 1,275 adults.

Chromosome 7

FOXP2

FOXP2 is a gene thought to be important for the development of language, having an effect during embryogenesis and influencing other biological pathways associated with language (Stelzer et al., 2016). Three imaging studies were conducted investigating associations with imaging in relation to reading disability. A structural study in a sample of 141 children indicated an association between *FOXP2* with gray matter volume in the left medial superior frontal gyrus (Skeide et al., 2016). Another study investigated associations of 31 SNPs in *FOXP2* with brain activation during a reading task in 94 adults, finding that 5 SNPS of *FOXP2* were significantly associated with reading-related activation in frontal regions, specifically in the left inferior frontal gyrus and dorsolateral left precentral gyrus (Pinel et al., 2012). Finally, Müller and colleagues (2017) investigated the association of the SNP rs12533005 with an EEG mismatch response signal (indicative of auditory discrimination skills) and found a non-significant association (Cohen's d =0.52). In the current meta-analysis, across 3 studies examining effects of *FOXP2* on 12 phenotypes (1 EEG, 1 grey matter volume, and 10 fMRI phenotypes), overall significance was found to be $p<.00001$.

CCDC136

CCDC136 is a gene that has been associated with epilepsy and acrosome formation (Stelzer et al., 2016), that has also been thought to affect reading and language skills (Adams et al., 2017). In a sample of 1,275 adults, the SNP rs59197085 did not reach significance for association with cortical surface area in the inferior frontal gyrus (pars triangularis) bilaterally, but results indicated that continued consideration is warranted (Gialluisi et al., 2016).

FLNC

FLNC is a gene that codes for filamin proteins that participate in anchoring of membrane proteins, involved in signaling and cell junction organization (Stelzer et al., 2016). One study examined the effects of *FLNC* on grey matter surface area and thickness in a sample of 1,275 adults but no significant associations were found (Gialluisi et al, 2016).

DGKI

DGKI is a protein coding gene associated with retinal degeneration as well as reading disability (Stelzer et al., 2016). One study examined the associations between 4 SNPs of *DGKI* and grey and white matter volume in a sample of 141 children, but no significant results were found (Skeide et al., 2016).

CREB3L2

CREB3L2 encodes a transcription factor associated with DNA binding and transcription (Stelzer et al., 2016). Skeide and colleagues (2016) examined association of the SNP rs273933 with grey and white matter volume in 141 children but found no significant associations.

CNTNAP2

CNTNAP2 is a gene associated with language difficulties and autism spectrum disorder (Whalley et al., 2011). Skeide and colleagues (2016) investigated associations of 7 SNPS of

CNTNAP2 with grey and white matter volume and found significant associations with white matter volume in the left cerebral and cerebellar peduncles, as well as reading comprehension, in a sample of 141 children. In a sample of 67 children, three SNPS were not significantly associated with a mismatch response EEG brain activation representative of auditory discrimination capabilities (Cohen's d = 0.15-0.46; Müller et al., 2017). In the current meta-analysis, the overall significance level across 5 phenotypes (3 EEG and 2 white matter volume phenotypes) from both studies investigating CNTNAP2 was $p<.00001$.

Chromosome 11

BDNF

BDNF, coding for a family of proteins involved in nerve growth factor, has been associated with stress responses, mood disorders, and diseases such as Alzheimer's, Parkinson's, and Huntington's disease (Stelzer et al., 2016). Jasińska and colleagues (2016) investigated the association of a Val66Met polymorphism, a missense variant, on brain activation in a sample of 81 children, finding that there were regions were there was greater activation for the Met allele carriers in several regions during reading tasks ($t=1.993$, Cohen's d = 0.47). Regions of differential activation included frontal and temporal regions, as well as the cerebellum (Jasińska et al., 2016).

Chromosome 12

SLC2A3

SLC2A3 codes for a glucose transporter protein (Stelzer et al., 2016). A genome wide association study in a sample of 386 dyslexic children identified a SNP rs4234898 and a

neighboring SNP rs11100040, related to expression levels of *SLC2A3*, to be significantly associated with the late EEG mismatch negativity (MMN) component during a passive oddball paradigm of listening to syllables, indicative of alterations in phonological processing (Roeske et al., 2011). Another EEG study in a sample of 67 children found significant associations between an EEG MMR signal in frontal regions related to auditory discrimination with both rs11100040 (Cohen's $d = 0.88$) and rs4234898 (Cohen's $d = 0.84$; Müller et al., 2017). Skeide and colleagues (2015) found that the SNP rs11100040, but not rs4234898, was associated with decreased fronto-temporal functional connectivity at resting state (Cohen's $d = 0.58$), and reduced fractional anisotropy in the left arcuate fasciculus, and this reduction in fractional anisotropy was related to the performance of children in phonological awareness tasks. The current meta-analysis resulted in an overall significance level of $p<.00001$, across the 10 phenotypes (3 EEG and 7 fMRI) in the three studies examining *SLC2A3*.

Chromosome 13

COL4A2

COL4A2, a gene coding for a membrane protein and previously identified as a candidate risk gene for reading disability (Eicher et al., 2013), had a significant association with a gray matter cluster in the right cerebellum in a sample of 141 children (Skeide et al., 2016).

Chromosome 15

Rare Copy Number Variant - 15q11.2(BP1–BP2) deletion

A rare copy number variant (CNV), a 15q11.2(BP1–BP2) deletion, has been found to be related to risk of neuropsychiatric disorders, including specific learning difficulties (Stefansson

et al., 2014). In a sample of adults, 71 with the deletion and 643 controls, carriers of the 15q11.2(BP1–BP2) deletion tended to have less gray matter volume in the left fusiform gyrus extending into the parahippocampal gyrus and increased gray matter volume in the superior occipital gyrus and superior frontal regions (Ulfarsson et al., 2017). In addition, deletion carriers tend to have less white matter volume in the right cerebellum, right paracentral lobule, and the left superior temporal lobe, but increased white matter volume in the anterior corpus callosum and right amygdala (Ulfarsson et al., 2017). Deletion carriers also had less activation in the left fusiform gyrus when reading pseudowords as compared to real words (Ulfarsson et al., 2017).

CYP19A1

CYP19A1 encodes a protein involved in drug metabolism and synthesis of lipids such as cholesterol and steroids (Stelzer et al., 2016). In a sample of 141 children, 3 SNPs associated with *CYP19A1* were not significantly associated with grey or white matter volume (Skeide et al., 2016).

DYX1C1

DYX1C1 is a gene encoding a protein involved in neuronal migration, and it has been proposed to be a dyslexia risk gene (Stelzer et al., 2016). Darki and colleagues (2012) examined associations between 3 *DYX1C1* SNPs and imaging phenotypes, including white matter volume and cortical thickness, in a sample of 76 children and young adults. The SNP rs3743204 in the *DYX1C1* gene was associated with white matter volume in a left temporo-parietal region (Cohen's $d = 0.98$) and a similar cluster in the right hemisphere, as well as global white matter volume, which in turn was related to reading ability (Darki et al., 2012). Darki and colleagues (2014) further investigated rs3743204 and its association with white matter volume, fractional

anisotropy, and cortical thickness in the same sample, finding associations with bilateral white matter volume in the superior longitudinal fasciculus and posterior corpus callosum (Cohen's *d* = 1.07). However, in a sample of 32 adult males, rs3743204 was not significantly associated with grey matter probability in regions of interest including the right lateral occipital complex, left Heschl's gyrus/posterior superior temporal gyrus, and left posterior superior temporal sulcus (Cohen's *d* = 0.14-0.65; Männel et al., 2015). An EEG study examining 4 *DYX1C1* SNPs found significant associations with the late component of EEG MMR signal in the frontal regions (related to auditory discrimination) for rs17819126 (Cohen's *d* = 1.74), rs3743204 (Cohen's *d* = -0.88), and nonsignificant associations for rs3743205 (Cohen's *d* = 0.41; Müller et al., 2017). Across these three studies, the current meta-analysis resulted in an overall significance of $p<.0001$ for the 9 reported associations (3 white matter volume, 3 grey matter probability, and 3 EEG phenotypes). Summary of analyses focused in the temporo-parietal region were significant at p=.0006, across 4 statistics in 2 studies. Skeide and colleagues (2016) also examined associations of 7 SNPs at the *DYX1C1-CCPG1* locus but found no significant relationships with grey or white matter volume.

Chromosome 16

CMIP

The gene *CMIP* codes for a protein that plays a role in T-cell signaling (Stelzer et al., 2016), and has been previously associated with language impairment (Scerri et al., 2011). With regard to neuroimaging, *CMIP* has been associated with bilateral portions of cerebellar white matter volume as well as reading comprehension in a sample of 141 children (Skeide et al., 2016). Müller and colleagues (2017) investigated associations with EEG MMR signal in an

anterior ROI in a sample of 67 children, but results were not significant (Cohen's *d* = 0.2-0.41). Results of the meta-analysis indicated that across 5 reported associations (3 EEG MMR and 2 white matter volume phenotypes) across both studies, overall significance was found to be p<.00001.

ATP2C2

3 SNPS in *ATP2C2,* a gene involved in maintaining transmembrane gradients and Ca2+ signaling, have been associated with decreased mismatch response (MMR) levels during a passive oddball paradigm requiring syllable discrimination, indicating decreased auditory discrimination capabilities, in a sample of 67 children (Cohen's *d*= 0.73=0.99; Müller et al., 2017). Skeide and colleagues (2016) examined associations of 5 SNPs with grey and white matter volume in 141 children, but results were not significant.

Chromosome 18

EPB41L3

EPB41L3 is a gene coding for a protein involved in protein interactions at synapses, chemical transmission, and apoptosis (Stelzer et al., 2016). A SNP in *EPB41L3* was not found to be significantly associated with grey or white matter volume in 141 children (Skeide et al., 2016).

SETBP1

SETBP1 is a gene associated with brain and nervous system development (Piazza et al., 2018). With regard to reading, an investigation of 32 SNPs within *SETBP1* in a sample of 135 children revealed significant associations with the SNP rs7230525 with brain activation in the right inferior lobule (Cohen's *d* = 0.78; Perdue et al., 2019).

DYM

DYM is a gene thought to play a role during early brain development (Stelzer et al., 2016). Skeide and colleagues (2016) investigated effects of a SNP (rs11873029) on grey and white matter volume in 141 children, but no significant associations were found.

MYO5B

MYO5B codes for a protein involved in cell movement and transport of materials between cell (National Library of Medicine, 2020). Skeide and colleagues (2016) investigated the association of a *MYO5B* SNP with grey and white matter volume in 141 children, but findings were not significant. Müller and colleagues (2017) also did not find significant associations between a SNP in *MYO5B* and the EEG MMR in the frontal regions in a sample of 67 children (Cohen's $d = 0.25$).

NEDD4L

NEDD4L encodes a protein involved with sodium transport and protein degradation (National Library of Medicine, 2020). Skeide and colleagues (2016) investigated the association of two *NEDD4L* SNPs with grey and white matter volume in 141 children, but findings were not significant. Müller and colleagues (2017) also did not find significant associations between two SNPs in *MYO5B* and the EEG MMR in the frontal regions in a sample of 67 children (Cohen's $d = 0.20$).

Chromosome 22

COMT

COMT, a gene implicated in dopamine regulation and skills such as attention and working memory, has variation at codon 158 (rs4680) consisting of a valine (val)-to-methionine

(Met) substitution, associated with improved performance on tasks recruiting prefrontal regions (Egan et al., 2001). In a sample of 86 children, individuals with a Met/Met genotype demonstrated greater activation than those with a Val/Val genotype in a region including the left occipitotemporal junction (OT) and fusiform gyrus (including the VWFA), the left middle temporal gyrus (MTG), a region of the right frontal cortex, and right parietal cortex on a reading task (Landi et al., 2014). Those with the Met/Met genotype showed significantly greater activation than those with the Val/Met genotype in similar regions, including the left OT, left superior temporal gyrus, and left middle temporal gyrus (Landi et al., 2014). Val/Met carriers had greater activation in the left precentral gyrus and right occipital temporal gyrus compared to Val/Val carriers, while Val/Val carriers demonstrated greater activity in the parahippocampal gyrus, regions of the frontal cortex, and cerebellum (Landi et al., 2014).

RBFOX2

RBFOX2 is thought to encode an alternative splicing regulator important for brain development and mature neuronal function (Gehman et al., 2012), and has been associated with reading and language skills (Gialluisi et al., 2016). In a sample of 1,275 adults, the SNP rs5995177 in *RBFOX2* was associated with cortical thickness, particularly in the left postcentral parietal gyrus (Cohen's d = -0.14), right middle temporal gyrus (Cohen's d = -0.15), right inferior frontal gyrus (pars opercularis Cohen's d = -0.14 and planum temporale Cohen's d = -0.14), and superior temporal gyrus bilaterally, with the minor allele being associated with reduced cortical thickness in those regions (Gialluisi et al., 2016).

X Chromosome

SYN1

SYN1 is a member of the synapsin gene family, coding for neuronal phosphoproteins on synaptic vesicles, involved in axonogenesis and synaptogenesis (Stelzer et al., 2016). Cabana and colleagues (2018) investigated the effects of a mutation in *SYN1* in 13 mutation carriers and 13 age and sex-matched controls, finding that the mutation was associated with increased fractional anisotropy and lowered mean diffusivity in reading-related brain regions.

Table 2

Summary of candidate genes investigated in more than one study

Gene	Mechanism of Action	Significant Results	Association with SRD	Effect Sizes of Reported Significant Associations	Mediation
MRPL19 /GCFC2 Locus	Protein synthesis, mRNA translation (*MRPL19*); transcription suppression (*GCFC2*)	Cortical thickness in the right middle temporal region; cortical volume in the right inferior temporal; white matter volume in the posterior corpus callosum and cingulum	Associated with SRD (Anthoni et al., 2007) and rapid naming (Rubenstein et al., 2014)	Not able to be calculated	Not examined
ROBO1	Axon guidance and growth across the midline of the CNS	Ipsilateral binaural suppression; fractional anisotropy, radial diffusivity, and axial diffusivity of the midline corpus callosum	Associated with SRD (Hannula-Jouppi et al., 2005)	0.76-2.48	Axial diffusivity of corpus callosum cluster significantly mediates genotype and reading scores
DCDC2	Neuronal migration	Grey matter volume in frontal, parietal, and temporal regions; fractional anisotropy in the left arcuate fasciculus corpus callosum; white matter volume in left temporo-parietal region, superior longitudinal fasciculus, and posterior corpus callosum; cortical thickness in the left supramarginal gyrus, angular gyrus, and lateral occipital cortex; functional activation in the cingulate gyrus, left	Associated with reading disability (Meng et al., 2005)	0.016-2.87	Not examined

		paracentral lobule, left inferior frontal gyrus, left anterior inferior parietal lobe, right lateral occipital temporal gyrus			
KIAA0319	Neuronal migration and cell adhesion	White matter volume in the left temporo-parietal region, superior longitudinal fasciculus, and posterior corpus callosum; global fractional anisotropy; cortical thickness in the left orbitofrontal region; neural variability in the inferior frontal gyrus and superior temporal gyrus in response to visual and auditory stimuli	Associated with reading skills (Paracchini et al., 2008) and SRD (Cope et al., 2005)	0.33-1.54	Not examined
TTRAP	DNA repair	Brain activation in the anterior inferior parietal lobule	Related to variation in reading and spelling (Luciano et al., 2007)	0.05-0.41	Not examined
THEM2	Cell proliferation	Cortical thickness in the left pars opercularis and brain activation in the posterior superior temporal sulcus during reading	Likely related to reading disability (Cope et al., 2005)	0.46-1.13	Not examined
FOXP2	Transcription factor affecting development of language; embryogenesis	Activation in the inferior frontal and precentral regions during reading and speech listening, and grey matter volume in the left superior frontal gyrus	Related to reading disability (Wilcke et al., 2011)	0.50-1.25	Not examined
CNTNAP2	Cell adhesion receptor; involved in nervous system development	White matter volume in cerebral peduncle and left inferior cerebellar peduncle	Related to reading and language (Newbury et al., 2011)	0.15-0.46	Not examined
SLC2A3	Glucose transporter	Mismatch negativity event-related potential reflecting automatic speech deviance processing; functional connectivity in the left inferior frontal gyrus and posterior superior temporal gyrus; fractional anisotropy in the left arcuate fasciculus	Related to mismatch negativity event-related potential in children with SRD (Roeske et al., 2011)	0.01-1.5	Not examined
DYX1C1	Neuronal migration	White matter volume in left temporo-parietal regions and the superior longitudinal fasciculus; Late component	Associated with reading (Bates et al., 2010)	0.14-1.74	Not examined

		of EEG MMR related to auditory discrimination			
CMIP	T-cell signaling pathway	White matter in the cerebellum	Related to reading ability (Scerri et al., 2011)	0.2-0.41	Not examined
ATP2C2	ATPase Secretory Pathway; Calcium transporter	Late component of EEG MMR related to auditory discrimination	Related to SRD in Chinese population (Wang et al., 2015), and language (Newbury et al., 2009)	0.73-0.98	Not examined
MYO5B	Plasma membrane recycling; vesicular trafficking	No significant results	Related to reading in a German population (Mueller et al., 2014)	0.25	Not examined
NEDD4L	E3 Ubiquitin-protein ligase	No significant results	Related to reading in a German population (Mueller et al., 2014)	0-0.20	Not examined

Discussion

The results of this review suggest the existence of a relationship between SRD risk genes, brain structure and function, and reading disability or language impairments. Regions of brain function or structure that were associated with reading disability risk genes tended to be in frontal and temporal areas previously determined to be part of the reading network, as expected. Studies indicated that genetic risk variants are associated with variations in gray matter volume, white matter volume and structure, and activation in the brain during reading and language tasks in these reading related regions. In current analyses, many of the genes that were investigated had a overall Fisher's p-value level of $p<.00001$, indicating that there are overall significant genetic associations with brain structure and function for the candidate genes most replicated.

Overall, genes with significant evidence for associations with both reading and imaging phenotypes included SRD-candidate genes on the DYX2 locus, *DCDC2* and *KIAA0319,* and the DYX1 locus, *DYX1C1*, all involved in neuronal migration. Other genes identified through both amount of replication and significance of results included *FOXP2*, a language-related gene, *SLC2A3*, a glucose transporter, and *ROBO1*, related to axon growth in the corpus callosum. For both *DCDC2* and *KIAA0319*, significant associations were found with grey matter and white matter structural phenotypes, as well as brain activation during tasks related to phonological processing and reading, in reading related frontal and temporal regions of the brain. There were also studies that did not find significant associations, but this may have been due to differences in methodology (e.g. which SNP was used). Additionally, the Fisher's combined probability tests completed in the current analysis for these genes suggest overall significant results for associations between these genes and the relevant imaging phenotypes examined in these studies. However, due to concerns related to publication bias, inconsistent correction for multiple comparisons, and summaries across many different phenotypes and SNPs, these results should be interpreted with caution. More research replicating these analyses would be necessary to be able to better understand effect sizes and overall significance for imaging phenotypes within specific brain regions, as well as for overall effects of specific SNPs or genes.

Other relevant genes were identified as well, although evidence was more mixed for some of the genes. There were four studies conducted on the nearby gene *TTRAP/TDP2*, but no significant associations were found by the study authors (Cope et al., 2012; Müller et al., 2017; Pinel et al., 2012; Skeide et al., 2016). In studies of both children and adults, the language-related gene *FOXP2* was found to be related to structure and brain activation in frontal regions, with the

current analysis demonstrating significance when studies were combined ($p<.00001$; Müller et al., 2017; Pinel et al., 2012; Skeide et al., 2016). *SLC2A3* tended to be associated with brain activation related to phonological processing tasks, reduced fronto-temporal functional connectivity, and reduced fractional anisotropy in the arcuate fasciculus, also with overall significant effects in the current analysis (Müller et al., 2017; Roeske et al., 2011; Skeide et al., 2015). *DYX1C1* has been linked to white matter volume in structure in the left temporo-parietal region, superior longitudinal fasciculus, and posterior corpus callosum (Darki et al., 2012; Darki et al., 2014), but other studies found no associations with grey or white matter volume (Männel et al. 2015; Skeide et al., 2016) or frontal activation related to auditory discrimination (Müller et al., 2017). Studies on *ROBO1* suggested effects on white matter structure in the corpus callosum, influencing processes that involved crossing between hemispheres (Lamminmäki et al., 2012; Sun et al., 2017). The genes described here had the most related imaging genetic studies published, and therefore the most evidence describing their associations with imaging phenotypes. Overall, these genes, with functions such as neuronal migration, glucose transport, axon growth, and language development, likely impact reading through primary effects on brain structure and function in the reading network. These results may help researchers to identify more genes related to reading – through genes with common functions (e.g., neuronal migration) or affecting related processes (e.g., language), leading to a greater understanding of which genes contribute to variance in reading ability. However, for many of the other genes described, there was only one or two imaging-genetic studies identifying neural correlates of that gene. Even within studies replicating results on the same gene, there was significant variability in which specific SNPs were tested, as well as potential differences in the allelic directions of effects on

phenotypes. Studies differed in which imaging phenotypes they used as well. In addition, many studies within the field have small sample sizes, and run into multiple testing difficulties with such a large number of potential SNPs and brain regions examined. These limitations can potentially lead to an increased susceptibility to false positive results, particularly due to the complexity of both genetic and imaging variables. These limitations in power and the significant variability between studies and replicated studies make it difficult to draw firm conclusions about which genes are most effective at this point in the field's development, despite promising results. Furthermore, the existing current imaging genetics studies of reading have been carried out in many different countries and languages. However, because various languages differ in orthographic depth, or the extend of deviation from simple letter to phoneme correspondence, the process of learning reading and neural correlates may differ depending on the language (Becker et al., 2017). This variability also makes it hard to compare different studies.

Because the imaging genetics field, particularly in relation to reading disability, is still advancing, much more research is needed to understand the relationship between genetic risk and the neural correlates of reading disability. Even for the most studied genes, *DCDC2* and *KIAA0319*, there was a lack of comprehensiveness in the neuroimaging techniques used or replication of the same results. With the neuroimaging techniques available, it is possible to study factors such as gray matter volume, white matter volume, white matter structure (e.g., fractional anisotropy), cortical thickness, activation, functional connectivity, and more. So far, these imaging-genetics studies have only scratched the surface of possible techniques to be used. In addition, because some of these studies used ROI approaches, which may or may not be the same, it is difficult to compare results across studies, even for one gene. In addition, within all

the imaging genetics studies included, there was only one functional connectivity study conducted, and no studies attempted to do an effective connectivity analysis to better understand how genetic risk is related to the entire reading network. It is clear that more research and replication is needed to better understand effects of these genes, and other genes, on imaging phenotypes and reading.

Another missing component to the imaging-genetics literature so far is the use of GWAS or other whole-genome methods (WGS or WES). Only one of the studies identified used a genome-wide approach (Roeske et al., 2011). The use of GWAS to investigate genetic associations with reading phenotypes has been helpful in identifying candidate genes but is likely missing small effects due to the rigid threshold required for multiple comparisons corrections (Becker et al., 2017). Because the genetic risk variants contributing to reading disability are likely to have small effects individually (Becker et al., 2017), the ability to use a neural endophenotype such as brain structure or function would improve power to detect genetic risk variants that may not have been previously associated with reading disability or language difficulties. Specifically, these imaging genetics studies helped to demonstrate relationships between genetic risk variants and phenotypes such as grey matter volume, white matter volume, cortical thickness, as well as brain activation in the already defined reading network of the brain. Starting from reading network regions, such as the temporo-parietal or inferior frontal regions, and determining associations with genes, may help researchers to identify genes that have indirect effects on reading through affecting brain structure and function within this network. Furthermore, it may be helpful for future studies to compare effect sizes between genetic effects

on imaging phenotypes as compared to genetic effects on behavioral phenotypes, to help support the use of imaging as a relevant endophenotype.

Imaging genetics approaches can also help with the missing heritability problem – the fact that single genetic variations cannot account for all of the estimated heritability of traits. The missing heritability problem affects understanding of genetic causes of many complex diseases and behaviors, including reading disability. Even combining all known marker associations with reading only explain a small portion of the phenotypic variance of reading disability (Mascheretti et al., 2015). The use of imaging genetics methodology may help with the missing heritability problem because it may help to identify genes that have small effects on reading, with main effects on underlying brain structure or function. Therefore, the use of GWAS along with imaging phenotypes may improve identification of additional genes contributing to the phenotypic variance of reading disability. In addition, there are additional genetic techniques that can be used to account for some of this unexplained variance, such as the investigation of gene-by-gene or gene-by-environment interactions and their association with neural phenotypes. These interactions may help to elucidate some of the non-additive effects that genes may have, helping to explain part of the missing heritability.

Epigenetics, or heritable changes in gene expression that do not involve changes to the underlying DNA sequence, may explain additional variance as well. Research has established that epigenetic mechanisms can modulate the effect of environmental influences on risk of disease and neurodevelopmental disorders (Feil & Fraga, 2012; Miyake et al., 2012; Salinas, Connolly, & Song, 2020), and play a role in learning and cognitive functioning (Grigorenko, Kornilov, & Naumova, 2016). However, there has been a lack of research investigating how

epigenetic mechanisms are associated with the neural phenotypes associated with reading disability. In addition, because epigenetic mechanisms can modulate the effect of the environment on outcomes, the study of epigenetics may be a valuable resource in the study of response to intervention in reading disability.

In addition to the use of more comprehensive imaging and genetic techniques, the use of longitudinal approaches and replication in different populations is needed. All of the imaging-genetic studies identified in this review used a cross-sectional approach. However, brain activation in the reading network changes during development and as reading skills are gained (Wise Younger, Tucker-Drob, & Booth, 2017). Moreover, heritability estimates change with age over, with the contribution of genetic factors increasing after onset of school and with age (Olson et al., 2014). Therefore, it is probable that the association of genetic risk and neural structure and function is variable during different developmental stages as well. For example, Darki and colleagues (2014) demonstrated that there was a significant interaction between rs793842 and age on the thickness of the left supramarginal gyrus and left lateral occipital cortex, with genotype having varying effects on cortical thickness depending on age. However, many of the studies included in this review had significant age ranges (e.g., 3-20 years or 6-25 years), but most did not investigate how these imaging-genetic associations varied by age. Therefore, more replication or expansion of results in different age brackets, examination of interacting effects, or the use of longitudinal approaches to examine the variation in these relationships over time is needed. Further, the use of imaging-genetic approaches in longitudinal, intervention studies would be valuable in determining the biomarkers associated with improvement of reading or may help to differentiate those individuals who do or do not respond to intervention. Because

response to intervention is the gold standard method for identifying reading disabilities, response to intervention should be better incorporated into future research. Using imaging genetics methods to determine biomarkers related to differentiation of individuals who do or do not respond to intervention is an important direction for research to turn as definitions of reading disability evolve.

Conclusion

This review aimed to summarize the findings of imaging-genetic studies of reading disability to date and suggest directions for future research. Because this is a new field, there is huge potential for imaging-genetics research to generate new insights into biological and neural contributors to reading disability. Both the articles and results of the Fisher's combined probability tests suggest significant associations between reading- and language- related genes (e.g., *DCDC2, KIAA0319, FOXP2, SLC2A3, ROBO1)* and brain structure and function in parts of the reading network (i.e., temporal and frontal regions). The use of more comprehensive neuroimaging techniques, including functional and effective connectivity, as well as genetic techniques, such as genome-wide association studies, gene-by-gene or gene-by-environment interaction studies, and epigenetic studies will provide critical insight into the underlying biomarkers and correlates of reading disability. More exploratory approaches focused on using imaging genetics techniques to identify previously undiscovered genes, as well as the use of larger samples, the inclusion of diverse stages of development, and longitudinal investigations, will help researchers to understand the underlying mechanisms contributing to the development of reading disability. In addition, future research should focus on identifying genetic and imaging associations with response to intervention to better differentiate those with true reading

disabilities as compared to those with inadequate instruction. Understanding these mechanisms will lead to better characterization of reading disability for improved early identification and evaluation of treatment response in affected individuals. Overall, this is an exciting and promising area of research with clearly demonstrated relationships between genetics, brain structure and function, and reading.

PART 2

Neuroimaging genetic associations between *SEMA6D*, brain structure, and reading skills

Tina Thomas [1]

Meaghan V. Perdue [2, 3]

Shiva Khalaf [4]

Nicole Landi [2, 3]

Fumiko Hoeft [2, 5]

Kenneth Pugh [2, 3]

Elena L. Grigorenko [1, 4, 6]

Author note:
[1]Department of Psychology, University of Houston, Houston, TX, USA. [2]University of Connecticut Dept. of Psychological Sciences, Storrs, CT, USA. [3]Haskins Laboratories, University of Connecticut, New Haven, CT, USA. [4]Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Houston, TX, USA. [5]Department of Psychiatry, University of California, San Francisco, CA, USA. [6]Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA.

The preparation of this article was supported by P20 HD091005 (PI: Grigorenko) and P50 HD052120 (PI: Wagner). Grantees undertaking such projects are encouraged to freely express their professional judgment. Therefore, this article does not necessarily reflect the position or policies of the abovementioned agencies, and no official endorsement should be inferred.

This article will appear in the Journal of Clinical and Experimental Neuropsychology published by Taylor and Francis.

Abstract

Specific reading disability (SRD) is defined by genetic and neural risk factors that are not fully understood. The current study used imaging genetics methodology to investigate relationships between *SEMA6D*, brain structure, and reading. *SEMA6D*, located on SRD risk locus DYX1, is involved in axon guidance, synapse formation, and dendrite development. *SEMA6D*'s associations with brain structure in reading-related regions of interest (ROIs) were investigated in a sample of children with a range of reading performance, from sites in Connecticut, CT (n=67, 6-13 years, mean age=9.07) and San Francisco, SF (n=28, 5-8 years, mean age = 6.5). Multiple regression analyses revealed significant associations between *SEMA6D*'s rs16959669 and cortical thickness in the fusiform gyrus and rs4270119 and gyrification in the supramarginal gyrus in the CT sample, but this was not replicated in the SF sample. Significant clusters were not associated with reading. For white matter volume, combined analyses across both samples revealed associations between reading and the left transverse temporal gyrus, left pars triangularis, left cerebellum, and right cerebellum. White matter volume in the left transverse temporal gyrus was nominally related to rs1817178, rs12050859, and rs1898110 in *SEMA6D*, and rs1817178 was significantly related to reading. Haplotype analyses revealed significant associations between the whole gene and brain phenotypes. Results suggest *SEMA6D* likely has an impact on multiple reading-related neural structures, but only white matter volume in the transverse temporal gyrus was significantly related to reading in the current sample. As the sample was young, the transverse temporal gyrus, involved in auditory perception, may be more strongly involved in reading because phonological processing is still being learned. The relationship between *SEMA6D* and reading may change as

different brain regions are involved during reading development. Future research should examine mediating effects, use additional brain measures, and use an older sample to better understand effects.

Keywords: *SEMA6D, reading ability, brain structure*

Introduction

Specific reading disability (SRD), diagnosed when individuals have significant difficulty with reading words or text, affects about 7% of the population, putting them at risk for poor academic performance (Hulme & Snowling, 2016). SRD has diverse, interacting risk factors at various levels (Miciak & Fletcher, 2020): neurobiological (genetic factors and brain structure and function), cognitive (e.g., phonemic and morphological awareness), behavioral (e.g., attentiveness and motivation), and environmental (e.g., socioeconomic and schooling contexts). The focus of the current study is on the neurobiological level, specifically how genetic and brain factors interact to influence reading ability or disability.

Relationships between brain structure and reading have been clearly established in the field. Most research has focused on cortical structures, with three major reading circuits identified; the dorsal temporo-parietal pathways associated with phonological processing, a ventral occipito-temporal pathway associated with word identification and automatic word recognition, and an anterior frontal region involved in articulation and higher order reading processes (D'Mello & Gabrieli, 2018; Richlan 2020). The basal ganglia and cerebellum have also been associated with procedural learning related to reading, as well as articulation (D'Mello & Gabrieli, 2018; Hancock, Richlan, & Hoeft, 2017; Ullman, Earle, Walenski, & Janacsek, 2020).

The integrative use of imaging and genetics is referred to as imaging genetics. It is a field that attempts to improve understanding of the connections between genes and behavior through the investigation of brain imaging as an intermediate phenotype, which is argued to be closer to the level of the gene (Flint, Timpson, & Munafó, 2014). For example, a study comparing effect sizes between gene-brain associations and gene-behavior demonstrated that imaging studies were

generally associated with larger effects (Rose & Donohoe, 2013). While most existing imaging genetics studies have focused on candidate genes that have already been associated with reading disability, imaging genetics may also allow better detection of genes that may have small effects on behavioral phenotypes. By investigating the relationship between genes and imaging endophenotypes, imaging genetics methods may be used to identify additional relevant genes that affect phenotypes through their effects on the brain (Flint et al., 2014).

Therefore, in the current study, we focus on a gene that has not been previously investigated for its relation to reading disability, *Semaphorin 6D (SEMA6D)*. However, based on its location in a dyslexia locus DYX1, on chromosome 15q21 (Schumacher, Hoffmann, Schmäl, Schulte-Körne, & Nöthen, 2007), and other studies about its function and association with various disorders, it is likely to have an impact on reading. First, *SEMA6D* is part of a family of genes coding for proteins that regulate axon guidance. Semaphorins mediate many other functions, including processes such as establishing the identity of neuronal cell processes, synapse formation, axon pruning, and regulation of dendrite development (Leslie et al., 2011; Alto & Terman, 2018). Many of the genes already associated with reading disability have similar functions, in processes such as neuronal migration, cortical morphogenesis, dendritic spinal plasticity, and axon guidance (Guidi et al., 2018; Hannula-Jouppi et al., 2005; Mascheretti et al., 2017). Importantly, neuronal migration and axon guidance have been proposed to lead to small cortical malformations, which can affect left hemispheric neural circuits involved in reading and learning (Galaburda, Sherman, Rosen, Aboitiz, & Geschwind, 1985; Galaburda, LoTurco, Ramus, Fitch, & Rosen, 2006). A more recent study suggested that focusing on just neuronal migration is limiting, and other processes such as axon growth, synaptic transmission, and ciliary

function may affect reading disorders as well (Guidi et al., 2018). These processes, which through their effects on axons can lead to changes in white matter structure, affect the reading network. Children with reading disability tend to exhibit white matter differences in left temporo-parietal areas and frontal regions, with involvement of the left arcuate fasciculus and corona radiata (Vandermosten, Boets, Wouters, & Ghesquière, 2012) and have different developmental trajectories, with delayed white matter development in the reading network (Christodoulou et al., 2017; Lebel et al., 2019).

Second, semaphorin genes have been associated with other neurodevelopmental disorders, including autism spectrum disorder (Mosca-Boidron et al., 2016), language disorder (Ercan-Sencicek et al., 2012; Chen et al., 2017), attention deficit hyperactivity disorder (Demontis et al., 2019; Hawi et al., 2018), and schizophrenia (Arion, Horvath, Lewis, & Mirnics, 2010), all of which have underlying genetic and neural risk factors. It is likely that many neurodevelopmental disorders may have some common underlying genetic causal factors, due to pleiotropy (i.e., the phenomenon that genes can influence two or more phenotypic traits). For example, the gene *CNTNAP2*, which has been associated with reading disability (Peter et al., 2011; Gu et al., 2018), has also been implicated in autism spectrum disorder, schizophrenia, intellectual disability, and language impairment (Rodenas-Cuadrado, Ho, & Vernes, 2014). Language impairment and reading disability share common deficits in underlying cognitive processes, such as phonological processing and language fluency (Pennington & Bishop, 2009), so there is likely overlap in genetic contributions to these disorders. Genome-wide association studies searching for genes associated with SRD have implicated genes that are involved in learning in general (Eicher et al., 2013; Veerappa, Saldanha, Padakannaya, & Ramachandra, 2013; Gialluisi et al., 2014), and

many genes that were previously associated with SRD have been demonstrated to influence language skills also (Eicher & Gruen, 2015). *SEMA6D* has also been associated with educational attainment (Okbay et al., 2016), indicating that the variation in this gene might be associated with overall learning or cognition, directly or indirectly through reading or other academic functions. All of this is evidence that there are generalist genes that can contribute to multiple related traits (Kovas & Plomin, 2006). Therefore, because *SEMA6D* has been associated with other developmental disorders, including autism spectrum disorder and language disorder, there is a strong possibility it could be involved in specific reading disability as well.

Few imaging genetics studies have focused on *SEMA6D*. However, there is evidence that SNPs in *SEMA6D* are associated with brain phenotypes as well, as Klein and colleagues (2019) demonstrated that *SEMA6D* was related to both ADHD risk, as well as intracranial volume and volume of the putamen of the basal ganglia (Klein et al., 2019). The basal ganglia have been shown to be important in procedural learning, impacting learning of language and reading and underlying many neurodevelopmental disorders including reading disability (Ullman et al., 2020). Variation in other genes in the semaphorin family with related functions have been shown to affect brain structure as well. For example, specific mutations in *SEMA6A*, also involved in axon guidance, have been found to affect brain cellular organization and connectivity in mice, which investigators reported modeled brain changes in other neurodevelopmental disorders such as autism spectrum disorder and schizophrenia (Rünker et al., 2011). Alterations in expression of the *SEMA* gene family have been linked to structural changes in the prefrontal cortex and synapse function associated with schizophrenia (Arion et al., 2010). Further research needs to be

conducted to determine if *SEMA6D* may have effects on other brain imaging phenotypes and how this may relate to reading.

The aim of the current study was to investigate the association between variants in *SEMA6D* and reading-related brain and behavior phenotypes. Specifically, associations between SNPs in or close to *SEMA6D* with various imaging phenotypes (cortical thickness, gyrification, and white matter volume) in reading related regions of interest in the brain were explored. In turn, relationships between brain structure and reading measures of word reading fluency and phonological processing were investigated.

Method

Participants

Data were collected from two different sites: University of California, San Francisco (UCSF), CA, and Haskins Laboratories in New Haven, CT. These studies were approved by the Yale University Institutional Review Board (Original IRB #1208010711, Re-analysis IRB # 2000021826) and the Stanford University Institutional Review Board (Original IRB #96574, Re-analysis University of Connecticut IRB# H18-200). Written informed consent was obtained from the parent or legal guardian of minor participants, and assent was obtained from participants age 8 years and older. Due to significant variability across different scanner strengths and other parameters and the relevance of this variability to the analyses of gyrification and cortical thickness (Han et al., 2006), the data from the two sites could not be merged. Because of the differences in the sample size, data collected from Haskins laboratories were used as the primary sample, and data collected from UCSF were used as a replication sample. For volume measures

of regions of interest, which tend to be more comparable across different scanner strengths (McCarthy et al., 2015), the samples were combined, and site was used as a covariate in the analysis. Inclusion criteria for both samples required native English language, an IQ above 75, no history of severe developmental or neuropsychological disorders, normal or corrected to normal vision, and normal hearing. Participants from both samples had a broad range of reading abilities. The sample size for the Haskins Laboratories data was 67, collected between 2006 and 2012, which was a subset of participants with both imaging and genetic data from within a larger dataset. The age range of the participants was 6-13 years (mean = 9.07), and all participants were Caucasian (white) within the subset used in the current study. The sample size for the UCSF data was 28, collected between 2008 and 2012. The age range of these participants was 5-8 years (mean=6.50). Race/ethnicity data for the UCSF sample was 63% White/Caucasian, 4% Asian, 7% American Indian/Alaskan Native, 11% multiracial, and 15% unreported.

Behavioral Assessments

Assessments of word reading included the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), a timed measure of an individual's ability to read printed words (TOWRE: Sight Word Efficiency) and pseudowords (TOWRE: Phonemic Decoding Efficiency) accurately and fluently. The Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) Elision subtest was used to assess phonological awareness and processing. These measures were used at both sites. For all reading measures, analyses were done with raw scores.

Genetic Data

Oragene saliva kits (DNA Genotek, Inc.) were used to obtain saliva samples during behavioral testing sessions and DNA was extracted from the samples according to manufacturer's protocol. DNA libraries were prepared for microarray genotyping on Illumina's HumanCoreExome v1 panel according to the manufacturer's protocol, and genotyping was carried out by Illumina, Inc., San Diego, CA, U.S.A. Illumina's GenomeStudio for Windows software was used for allele calling. Following quality assurance procedures, call rates were evaluated. All samples had a call rate, or the fraction of called SNPs per sample over the total number of SNPs in the data set, above 95%, and SNP markers with a call rate below 95% were excluded from the dataset. For the UCSF data, samples were collected using peripheral blood. Genotyping was done using the Illumina Core Exome v1.2 according to the manufacturer's protocol.

Overall, 67 SNPs in *SEMA6D* were common to both data collected at UCSF and at Haskins laboratories. SNPs were analyzed for linkage disequilibrium (Figure 1) and those in high linkage disequilibrium (R^2>.90) were removed, leaving 55 SNPs remaining. Because of the small sample size, genotype was coded by the presence of the derived, or nonancestral, allele (0 without the presence of the derived allele, and 1 with the presence of the derived allele). The ancestral allele is the allelic state of the last common ancestor, while the derived allele is the one that arose due to mutation.

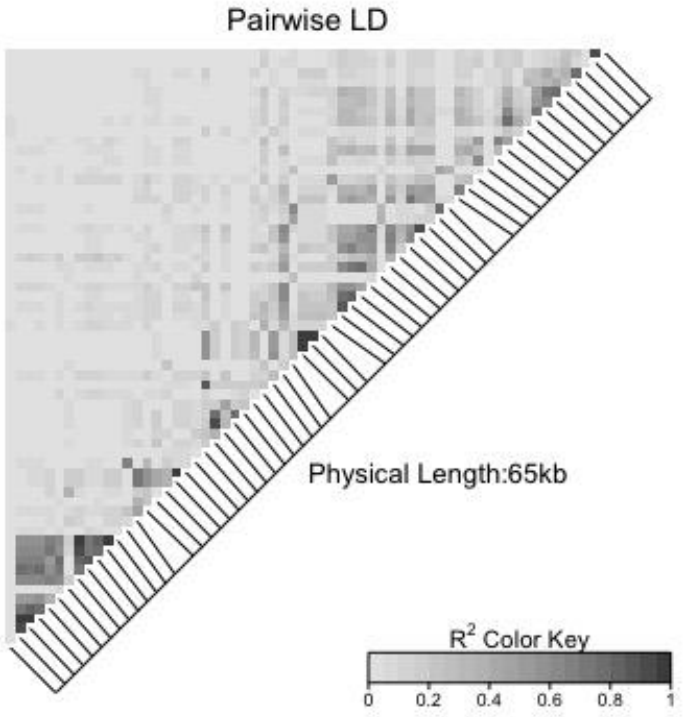

Figure 1. Linkage Disequilibrium for *SEMA6D* SNPs.

Imaging Data

T2 structural MRI data were analyzed in order to obtain data on cortical thickness, gyrification, and white matter volume. These data were collected on two different MRI scanners. Acquisition of brain images by Haskins Laboratories was conducted using a Siemens Sonata 1.5-Tesla MRI scanner with a 12 channel head coil. High-resolution anatomical images were acquired (sagittal MPRAGE acquisition, FA 8°; TE 3.65 ms; TR 2000 ms; FOV 256× 256 mm; 1mm slice thickness, no gap; 256× 256×160, 1 NEX). Acquisition at UCSF was conducted using a 3T GE Signa scanner with an 8 channel head coil (FSPGR3D-1nex Acquisition; FA 15°; TE

3.4 ms; TR 8.5 ms; FOV 220x220 mm; 1 mm slice thickness, no gap; 256x192x128 matrix, 1 NEX).

MRI preprocessing was conducted in order to prepare MRI data for analysis using Freesurfer v6.0.0 software (Fischl, Sereno, & Dale, 1999; Dale, Fischl, & Sereno, 1999). The preprocessing pipeline involves reconstruction of a two-dimensional cortical surface into a three-dimensional volume, skull stripping, classification of white and grey matter, and correction for motion. Manual visual inspection and preprocessing was done to correct errors in pial boundary (between grey matter and skull), the white matter surface boundary, and to correct intensity normalization errors. Cortical thickness was calculated as the shortest path between vertices on pial and white matter boundaries. Spatial smoothing was conducted using a 10 mm FWHM Gaussian filter. A local gyrification index (LGI), measuring the amount of folding in the brain, was calculated using Freesurfer as well; specifically, the LGI quantifies the amount of cortex buried within the sulcal folds as compared to the cortex on the outer cortex (Schaer et al., 2012). For the gyrification analysis, spatial smoothing was conducted using only a 5 mm FWHM Gaussian filter because smoothing was already done as part of the automated calculation of the local gyrification index, with gyrification estimations based on 15 mm diameter spheres.

Regions of interest (ROIs) were isolated by using pre-existing labels of the Destrieux Atlas (Destrieux, Fischl, Dale, & Halgren, 2010). These regions of interest were chosen based on their previous association with SRD (Ma et al., 2015), which included the left hemisphere inferior frontal gyrus (pars opercularis, pars orbitalis, and pars triangularis), superior temporal gyrus, Heschl's gyrus, inferior occipital gyrus, planum polare and planum temporale, fusiform gyrus, supramarginal gyrus, and angular gyrus. Similar to the methods used by Ma and

colleagues (2015), a cortical mask was created over the ROIs, and vertex-based analysis of cortical thickness and gyrification was conducted within the mask. The white matter volume associated with each grey matter ROI was extracted from each subject and analyses were conducted in Freesurfer software with the Killiany/Desikan parcellation atlas (Desikan et al., 2006). For white matter volume, regions of interest were the white matter underlying the supramarginal gyrus, pars opercularis, pars triangularis, pars orbitalis, fusiform gyrus, transverse temporal gyrus, the superior temporal gyrus, and the left and right cerebellum. The names of the ROIs included are based on the cortex because of the way Freesurfer labels white matter, but the regions of interest were, actually, the underlying white matter associated with these grey matter ROIs.

Statistical Analyses

Statistical analyses were completed first for gyrification and cortical thickness phenotypes, then for white matter volume. For cortical thickness and gyrification phenotypes, the steps were (1) a vertex-based analysis for genetic associations with imaging phenotypes was done within a mask over all relevant ROIs, (2) significant clusters were extracted and associations with reading were examined, (3) significant SNPs were regressed on reading phenotypes, and (4) a haplotype analysis was completed combining the effects of all SNPs within *SEMA6D,* which were regressed on significant clusters of gyrification or cortical thickness. Following these analyses, significant clusters were not found to be related to reading. Therefore, for white matter volume analyses, for which the measure of white matter volume was calculated over the whole ROI, the analysis was done first to determine brain regions associated

with reading. This helped to focus the analysis on the ROIs that were relevant to reading in this sample and increase power to detect genetic effects by reducing the number of imaging genetic analyses. Therefore, for part 2 of the analysis, focused on white matter volume, the steps were (1) white matter volume was regressed on reading measures, (2) white matter volume of regions significantly related to reading was associated with SNPs, and (3) significant SNPs were regressed on reading phenotypes, and (4) a haplotype analysis was completed with white matter volume in the same brain regions, but combining effects of all SNPs within *SEMA6D*. These steps are described in further detail below.

Part 1: Gyrification and Cortical Thickness Analyses

Because the data were collected across two sites that differ on a number of factors (including scanner strength, USCF =3T and Haskins = 1.5T scanner), we elected to first analyze data from Haskin's Laboratories (1.5T scanner, n=67), and then replicate the analysis with the UCSF sample (3T scanner, n=28). This was done for both cortical thickness and gyrification analyses. A multiple regression analysis was conducted to analyze the associations between SNPs in *SEMA6D* and cortical thickness and gyrification in the brain. After accounting for linkage disequilibrium and removing 12 SNPs, each of the remaining 55 SNPs was used in a separate analysis. All analyses included age and sex as covariates. Correction for multiple comparisons accounting for spatial correlation was done using a Monte Carlo simulation on Freesurfer software with a cluster-forming p-value set at $p<.01$, and clusterwise p-value set at .05, to reduce false positive rates with cluster-wise corrections. Cortical thickness or gyrification means from significant clusters with any significant SNPs were then used as predictors in a multiple regression analysis to determine whether the local gyrification index (LGI) and cortical

thickness in significant clusters predicted reading. Following analyses of individual SNPs, significant clusters identified in individual SNP analyses were used in a haplotype analysis to determine effects of the entire gene within the Haskins sample. The haplo.stats package in R (Sinnwell & Schaid, 2016) was used to quantify effects of all SNPs in *SEMA6D*, taking into account the fact that SNPs tend to be inherited together, using the haplo.glm function. Haplotype analyses were adjusted for age and gender.

Part 2: White Matter Volume Analyses

For analysis of white matter volume, analyses were done across the entire sample because broader measures of volume in ROIs have been found to be more comparable across different scanner strengths (McCarthy et al., 2015). For these analyses, data collection site was included as a covariate. The Shapiro-Wilk test (Shapiro & Wilk, 1965) was carried out to determine normality, and non-normal variables were scaled for further analysis. Partial correlations were first completed to examine whether white matter volume in regions of interest predicted performance on reading measures, controlling for age, gender, and data collection site. Brain associations with reading were examined first to focus this analysis on the brain regions known to be involved in reading. The regions that predicted reading were then further analyzed to determine genetic associations, using multiple regression, controlling for age, gender, and site. Whole brain volume was not controlled for in these regressions because whole brain volume was not significantly correlated with white matter volume in each of the ROI's investigated or any of the reading measures. Each SNP was used in a separate multiple regression for each brain region, and then for the analysis of each SNP, correction for multiple comparisons of brain regions was done using the false discovery rate (FDR; Benjamini & Hochberg, 1995). FDR correction for

multiple comparisons was used rather than the Monte Carlo cluster-wise correction for multiple comparisons because the analysis was done with average volume over ROIs rather than a vertex-wise analysis. Lastly, multiple regressions were used to analyze associations between SNPs and reading measures, controlling for age, gender, and site, using FDR correction for multiple comparisons. The false discovery rate correction for multiple comparisons results in a q-value, or the expected proportion of false positives among all positive results. Following analyses of individual SNPs, the white matter volume in the same brain regions were used in a haplotype analysis to determine effects of the whole gene, using the haplo.glm function in the haplo.stats R package. Within the haplo.stats R package, most common haplotypes (frequency of greater than .05) are each analyzed separately for effects on phenotypes, while rare haplotypes are combined and analyzed together. Analyses were adjusted for age, gender, and site.

Results

Behavioral Results

Demographic and behavioral data are presented in Table 1. The Haskins sample was overall older than the UCSF sample. Intelligence tended to be within the average range as well, but differed significantly across groups, with the UCSF sample tending to have a higher IQ. On reading measures, raw scores varied due to the differing age ranges of the two samples.

Table 1.
Demographic and reading descriptive statistics

Variable	Haskins (n=67) *M(SD)*	UCSF (n=28) *M(SD)*	T-Test
Years at age of MRI	9.07 (range 6-13)	6.5 (range 5-8)	t = -11.74, p<2.2e-16
Sex (% Male)	64	55	χ^2=89, p=0.33

IQ Measure	WASI 111.68(13.12)	WJ BIA 118.90(10.97)	t = 2.78, p=.007
CTOPP Elision Raw Scores	13.48(4.98)	10.17(678)	t = -1.09, p=0.04
TOWRE Sight Word Raw Scores	57.85(18.48)	30.17(27.10)	t=-4.09, p=9.74e-05
TOWRE Phonemic Decoding Raw Scores	27.30(13.26)	12.97(12.34)	t=-3.16, p=0.0022

SEMA6D-Cortical Thickness Analyses

To assess associations between *SEMA6D* and cortical thickness in the reading network, multiple regression analyses were done for each SNP using age and gender as a covariate, first in the Haskins sample, followed by the UCSF sample. These analyses were done across the reading network using a mask. Correction for multiple comparisons within the brain were done with a Monte Carlo simulation. In the Haskins sample, one SNP (rs16959669) demonstrated a significant association with cortical thickness in the fusiform gyrus after correcting for cluster-wise comparisons. Specifically, the presence of the nonancestral (derived) allele (denoted by 2; n=5) was associated with greater cortical thickness as compared to SNPs that were homozygous for the ancestral allele (n=58). This SNP, which has a sample frequency of 0.06 for the derived allele, has been previously shown to be related to skin pigmentation, but has not been previously investigated for relation to reading.

Cluster	Maximum p-value	Peak Vertex (Vertex number at which maximum p-value)	Size (mm^2)	MNI Coordinates	Cluster-wise p-value
1	0.0002	28473	706.26	(-41 -67, -13)	0.0016

Figure 2. Significant cortical thickness cluster in fusiform gyrus for rs16959669, with the presence of the non-ancestral allele associated with greater cortical thickness. Peak vertex describes the location of the vertex where the maximum effect was identified for the cluster.

Results of haplotype analyses combining effects from all SNPs in *SEMA6D* revealed significant effects of one haplotype (t=2.55, p=0.014) on cortical thickness in the fusiform gyrus cluster. The haplotype frequency was 0.024 and it included the following alleles for each SNP (2 indicating derived allele), in order by position in gene: 21111111111111121111121122111-11111111121111111121112211. Other haplotypes were not significantly related to cortical thickness.

When replication analyses were performed in the UCSF sample, using the same SNP (rs16959669; ancestral allele n=22, presence of derived allele n=5) there was a small cluster nearby (in the lateral occipital region) with an uncorrected significant at p<.01, along with other

small clusters throughout the reading network (pars opercularis, supramarginal gyrus, and the superior temporal region), but none were significant following cluster-wise correction for multiple comparisons.

Cortical Thickness and Reading

A multiple regression analysis, controlling for age and gender, was completed to assess the association between thickness in the significant cluster and reading in the Haskins sample in separate regression analyses for each reading phenotype. Results of the multiple regression revealed no significant associations between thickness in this cluster and measures of phonological processing (CTOPP Elision), word reading (TOWRE Sight Word Efficiency), or nonword reading (TOWRE Phonemic Decoding).

The association between the SNP rs16959669 and reading was examined in a separate regression but was not found to be significantly related to the CTOPP Elision raw score ($p=0.26$), the TOWRE sight word efficiency ($p=0.56$), or the TOWRE phonemic decoding efficiency ($p=0.47$).

SEMA6D-Gyrification Analyses

Multiple regressions, controlling for age and gender, were carried out first in the Haskins sample, followed by the UCSF sample, to determine genetic associations with local gyrification in the reading network (specified by a mask over all of the included ROIs). In the Haskins sample, one SNP, rs4270119 (ancestral allele n=25; presence of derived allele n=38), was

significantly associated with local gyrification in the supramarginal gyrus after correction for cluster-wise multiple comparisons. Specifically, the presence of the ancestral allele was associated with a smaller local gyrification index. This SNP, rs4270119, has not been previously studied in the literature. The general population frequency of the derived allele is 0.30.

Cluster	Maximum p-value	Peak Vertex (Vertex number at which maximum p-value)	Size (mm^2)	MNI Coordinates	Cluster-wise p-value
1	0.00027	5031	878.68	(-59, -47, 33)	0.027

Figure 3. Significant cluster of gyrification in supramarginal gyrus associated with rs4270119

Two other SNPs (rs1369645 [ancestral allele n=25; presence of derived allele n=38; general population frequency of derived allele 0.32] and rs16952896 [ancestral allele n=25; presence of derived allele n=38; general population frequency of derived allele 0.30]) had very similar significant clusters in the supramarginal gyrus at p<.01, but these did not survive

correction for multiple comparisons. Results of haplotype analyses combining effects from all SNPs in *SEMA6D* revealed significant effects of two haplotypes (haplotype 1 [11111111111111111111211211212222122-22111221112112221111]: t=-3.13, p=0.003, frequency=0.073; haplotype 2 [22221222121222211222111211121222212222111221112112-221111]: t=-2.88, p=0.006, frequency=0.040) on gyrification in the supramarginal gyrus cluster. Other common haplotypes were not significantly associated with gyrification.

When replication analyses were performed in the UCSF sample, using the same SNPs (rs4270119 [ancestral allele n=12; presence of derived allele n=15], rs1369645 [ancestral allele n=13; presence of derived allele n=14], and rs16952896 [ancestral allele n=11; presence of derived allele n=16]), there were no significant clusters following correction for multiple comparisons.

Gyrification and Reading

A multiple regression was carried out in the Haskins sample to determine associations between local gyrification values in the significant supramarginal gyrus cluster and reading measures, controlling for age and gender. Results revealed that gyrification in the supramarginal gyrus was not significantly associated with measures of phonological processing (CTOPP Elision), word reading (TOWRE Sight Word Efficiency), or nonword reading (TOWRE Phonemic Decoding).

The significant SNP, rs4270119, was also regressed on reading, controlling for age, gender, and site, but was not significantly related to the CTOPP Elision raw score (p=.60), TOWRE sight word efficiency (p=.70), or TOWRE phonemic decoding efficiency (p=0.87).

White Matter Volume and Reading

White matter volume values in each region of interest were assessed for normality. ROIs that were not normally distributed (pars opercularis and transverse temporal) were scaled, and these scaled variables were used in the following multiple regression analyses. For white matter volume analyses, white matter volume under grey matter ROIs (left hemisphere supramarginal, pars triangularis, pars opercularis, pars orbitalis, transverse temporal, superior temporal, right and left cerebellum) were partially correlated with reading measures, using age, gender, and scanner as covariates and using FDR correction for multiple comparisons of brain regions. The white matter volume of the left transverse temporal gyrus was significantly associated with all reading phenotypes, including the CTOPP Elision raw scores ($r=0.38$, $p=0.00037$, $q=0.0015$), TOWRE sight word efficiency ($r=0.24$, $p=0.026$, $q=.026$), and TOWRE phonemic decoding ($r=0.24$, $p=0.026$, $q=0.026$).

Other white matter volume ROIs were also significantly related to the CTOPP Elision raw scores, but not other reading measures. These included the pars triangularis ($r=0.31$, $p=0.005$, $q=0.02$), the left cerebellum ($r=0.36$, $p=0.00075$, $q=0.0030$), and the right cerebellum ($r=0.34$, $p=0.0016$, $q=0.0064$). Only these four regions were considered for further analyses.

SEMA6D-White Matter Volume and SNP-Reading Analyses

Because it had the strongest association with reading, white matter volume of the transverse temporal gyrus was further analyzed for association with SNPs from *SEMA6D* using

multiple regressions controlling for gender, age, and scanner. Three SNPs, rs1817178 (ancestral allele n=53; presence of derived allele n=38; general population frequency of derived allele 0.21), rs12050859 (ancestral allele n=77; presence of derived allele n=12; general population frequency of derived allele 0.07), and rs1898110 (ancestral allele n=20; presence of derived allele n=70; general population frequency of derived allele 0.49) were all significantly associated with white matter volume in the transverse temporal region. For rs1817178, presence of the non-ancestral allele was associated with decreased white matter volume. For rs12050859 and rs1898110, the presence of the non-ancestral allele was associated with increased white matter volume in the transverse temporal region. These three SNPs have not been previously studied for phenotypic associations in the literature.

These SNPs were further analyzed for association with other brain regions that were associated with reading (CTOPP Elision raw scores), i.e., the pars triangularis, left cerebellum, and right cerebellum. Rs1817178 was significantly related to white matter volume of the transverse temporal region, but not after correction for multiple comparisons ($p=0.03$, $q=0.11$), and was not related to white matter in the pars triangularis ($p=0.17$, $q=0.17$), left cerebellum ($p=0.08$, $q=0.12$), or the right cerebellum ($p=0.09$, $q=0.12$). Rs12050859 was nominally related to the white matter of the transverse temporal ($p=0.0135$, $q=0.054$) and left cerebellum ($p=0.033$, $q=0.066$) after correction for multiple comparisons, but not the pars triangularis ($p=0.35$, $q=0.47$) or right cerebellum ($p=0.92$, $q=0.92$). Rs1898110 was significantly related to white matter in the transverse temporal region before correction for multiple comparisons but was not related to white matter in the transverse temporal ($p=0.046$, $q=0.18$), pars triangularis ($p=0.82$, $q=0.82$), left cerebellum ($p=0.17$, $q=0.23$), or right cerebellum ($p=0.16$, $q=0.23$) after correction.

Further analysis of these SNPs in a separate model examining genetic associations with reading, controlling for age, gender, and site and using FDR correction for multiple comparisons, revealed that the SNP rs1817178 also predicted reading, including CTOPP Elision raw scores ($p=0.004$, $q=0.013$), TOWRE sight word efficiency raw scores ($p=0.03$, $q=0.04$), and TOWRE phonemic decoding efficiency ($p=0.05$, $q=0.05$). Both rs12050859 and rs1898110 were not significantly associated with reading.

Haplotype analyses were completed for white matter volume in each brain region associated with reading in order to determine effects of the whole gene, taking into account the fact that alleles are linked and inherited together, controlling for age, gender, and scanner. Results of the haplotype analysis revealed significant associations of all of the most common haplotypes (with expected counts greater than 5) with all four investigated brain regions. For the transverse temporal gyrus, the most frequent haplotype [11111111111111111111221111112-11111111111112121211111122], with a frequency of 0.057, was significantly associated with white matter volume ($t=-1.85x10^3$, $p=0.00$). For the pars triangularis white matter volume, the association with the same haplotype was also significant ($t=-2.20x10^3$, $p=0.00$). For the left cerebellum white matter volume, the same, the same haplotype also demonstrated the significant association ($t=4.73x10^{17}$, $p=0.00$). Finally, the same was observed for the right cerebellum white matter ($t=1.810 \times 10^{18}$, $p=0.00$).

Discussion

The current study investigated whether the analyzed SNPs in the *SEMA6D* gene were related to brain structure and reading in a sample of children at various reading levels, using an integrative imaging genetic approach. Overall, we found relationships between SNPs in

SEMA6D and brain structure indicators of gyrification, cortical thickness, and white matter volume in the reading network. Taking into consideration the whole gene, there were also strongly significant results between the most common haplotypes and white matter volume in reading-related regions, as well as associations between several haplotypes and cortical thickness in a fusiform gyrus cluster and gyrification in a supramarginal gyrus cluster. However, gyrification and cortical thickness findings, which were found in the Haskins sample, were not replicated in the UCSF sample, potentially due to a small sample size in the UCSF sample.

SEMA6D, with its role in axon guidance and synapse formation, likely affects brain structure and function during brain development. When examining gene networks *SEMA6D* is involved in, *SEMA6D* has been shown to work in tandem with the *PLXN* family of genes, as plexin proteins act as receptors for semaphorin proteins (Alto & Terman, 2018). *PLXN* genes have been associated with dyslexia, dyspraxia, and language impairment (Rudov et al., 2013), as well as autism (Suda et al., 2011). Similarly, *SEMA6D* has been associated with autism (Mosca-Boidron et al., 2016) and language disorder (Ercan-Sencicek et al., 2012; Chen et al., 2017), and based on our current results, likely has an effect on reading as well. Furthermore, these related *PLXN* genes have been shown to have effects on white matter structure (Belyk, Kraft, Brown, & Pediatric Imaging, Neurocognition and Genetics Study, 2017), consistent with our results for *SEMA6D* in the current study. Other reading disability related genes, including *KIAA0319* and *ROBO1* are related to axon growth and guidance as well (Franquinho et al., 2017). Therefore, the current results are consistent with expectations based on *SEMA6D*'s functions, gene networks, and effects on other related disorders such as language disorder.

Regarding the link between brain structure and reading, the strongest associations were between white matter volume in the left transverse temporal region (Heschl's gyrus) and measures of phonological processing, word reading, and decoding of nonwords. Previous studies have demonstrated the importance of Heschl's gyrus in reading, and particularly phonologically based learning. For example, Welcome and Joanisse (2014) demonstrated that white matter volume in Heschl's gyrus predicted nonword reading skills in adults. In addition to white matter volume, other studies reveal corresponding associations between grey matter volume of Heschl's gyrus and reading. The size of the left hemisphere Heschl's gyrus, along with differences in planum temporale asymmetry and cerebral volume size, have also been shown to help distinguish between children with phonologically-based reading disability and children with language impairment, with SRD children having a larger Heschl's gyrus (Leonard et al., 2002). Furthermore, the pattern of larger Heschl's gyrus predicted phonological decoding skills in typically developing children (Leonard et al., 2002). Similarly, Wong and colleagues (2008) found that the volume of left Heschl's gyrus was negatively related to ability to learn pitch patterns, important when learning spoken language. An increased white matter volume, indicating more or stronger connections with a grey matter region of interest, corresponding with a smaller gray matter volume, reflective of increased grey matter density and more efficient processing, tend to be associated with improvements in cognition. Therefore, these corresponding findings demonstrate the importance of Heschl's gyrus in reading. Furthermore, cortical thickness studies provide additional evidence, as thicker cortical thickness in relevant brain regions tends to be associated with improved cognition. In a Norwegian sample, children who later developed dyslexia had thinner cortex in the left hemisphere Heschl's gyrus (along

with other primary auditory and visual regions) prior to learning how to read (Clark et al., 2014). Similarly, cortical thickness in the left superior temporal cortex, partially overlapping with Heschl's gyrus, has been shown to be positively correlated with word and pseudoword reading in typically developing children (Perdue, Mednick, Pugh, & Landi, 2020). Overall, across phenotypes, corresponding increases in white matter volume, decreases in grey matter volume, and thicker cortex tend to be related to improvements in reading.

Because the children in the current sample are young (some only 5 and 6), and measures that were used involved basic word reading and decoding and phonological processing, the white matter under Heschl's gyrus may have been more important in affecting reading ability, rather than white matter under other structures such as the fusiform gyrus that become more important as there is development of fluent reading, automatic recognition of words and higher order processing of meaning (Devlin, Jamison, Gonnerman, & Matthews, 2006). In our study, *SEMA6D* SNPs and haplotypes had effects on several reading-related regions like the supramarginal gyrus and fusiform gyrus, but only the white matter underlying the transverse temporal gyrus was related to reading. Therefore, it is possible that *SEMA6D* may have a different impact in affecting reading as children develop, because the effects of *SEMA6D* on reading-related structures may have a greater impact on reading during different stages. In other words, these regions may have a more important moderating effect on the link between *SEMA6D* and reading as children age and reading becomes more developed. Earlier development of connectivity between regions tends to predict later functions of reading regions of interest (Saygin et al., 2016), suggesting that the role of certain regions of interest that develop later in reading may not be evident at early ages, with early connectivity developing first. During reading

development, children demonstrate changing patterns of functional activation, with readers having more involvement in the inferior frontal gyrus, precentral and postcentral gyrus, and fusiform gyrus during reading when compared to pre-readers (Chyl et al., 2018), while activation in superior temporal regions related to speech processing is evident in both pre-readers and emergent readers (Chyl et al., 2018). Therefore, using a sample of older children may show differential impacts of *SEMA6D* on reading as functional networks change and develop. This may also help to explain why the results for cortical thickness and gyrification in the Haskin's sample did not replicate in the UCSF sample, as the age ranges were variable.

Three SNPs in *SEMA6D* (rs1817178, rs12050859, rs1898110) were significantly associated with white matter volume in the left hemisphere transverse temporal region, and white matter volume in the left transverse temporal region was significantly related to all three measures of phonological processing, word reading, and decoding of nonwords. Of the three SNPs that were significantly related to white matter volume in the left transverse temporal region, rs1817178 was the only one significantly associated with the reading measures when controlling for age, site, and gender. These findings are novel, because these SNPs have not been previously studied for phenotypic associations in the literature. All three of these SNPs were intron variants of *SEMA6D*, which can have effects on gene expression due to regulatory elements. Therefore, these SNPs may influence expression of *SEMA6D*, which can then lead to changes in brain structure or function. Furthermore, results of the haplotype analysis indicated strong associations between the entire gene with white matter volume in all four regions that were related to reading. However, due to the low frequency of each haplotype when considering all available markers in the gene, the effects of haplotypes should be studied in larger samples to

better understand these promising preliminary results. Future studies with larger samples would also benefit from an analysis of how these relevant haplotypes are associated with reading scores as well. Because *SEMA6D* has known functions in axon guidance and synapse formation, it likely influences the development of the brain and white matter structure in the brain.

Other findings were that SNPs in *SEMA6D* were significantly associated with gyrification in the left hemisphere supramarginal gyrus and cortical thickness in the left hemisphere fusiform gyrus in the Haskins sample. However, these results did not replicate in the UCSF sample, potentially due to its small sample size. Additionally, gyrification in the supramarginal gyrus and cortical thickness in the fusiform gyrus were not significantly related to reading. However, gyrification and cortical thickness in the reading network have been shown to be related to reading in previous research. For example, Blackmon and colleagues (2010) demonstrated that better pronunciation of irregular words (more representative of the orthographic components of word recognition) was associated with thinner cortex in reading network regions including the supramarginal gyrus. Alternatively, in a sample of Chinese children, cortical thickness in the left supramarginal gyrus was positively correlated with oral word reading, and also predicted phonological awareness (Xia et al., 2018). One study looking at both gyrification and cortical thickness found corresponding increased gyrification and thinner cortex in left occipitotemporal region, where the fusiform gyrus is located, in children with dyslexia (Williams, Juranek, Cirino, & Fletcher, 2018). Gyrification and cortical thickness tend to be negatively related to each other, with greater gyrification and thinner cortex related to more efficient processing (White, Schmidt, Kao, & Shapiro, 2010). While we did not find these relationships to reading in the current study, this may have been due to the young age of the

participants and less developed reading skills that are likely still relying on phonological processes rather than orthographic recognition during reading. Future studies may benefit from doing analyses to examine whether there is an interaction with this relationship and age. Furthermore, this sample generally consisted of children who were low-average to above average readers, not meeting criteria for reading disability. Therefore, there may not have been enough variability in the sample to detect significant relationships between cortical thickness and gyrification and reading. In addition, the smaller sample size of using the UCSF sample alone may have made it more difficult to detect significant effects.

Future research may benefit from using a sample of children with reading disability. While effects of SNPs tended to be small and sometimes did not survive after correction for multiple comparisons, the study had an overall small sample size and also used a sample of children with a wide range of reading ability. Therefore, results may have been stronger using a sample of children with diagnosed reading disability compared to typically developing children. Research should also consider gene-by-gene or gene-by-environment interactions, as taking interacting effects into account will help us better understand the relationship between genes and brain structure and reading (Gilbert-Diamond & Moore, 2011). Genes often work as part of a pathway or network, so having a full understanding of interacting effects can help improve our understanding of the strength of the relationship between genes and phenotypes. Additionally, while we examined effects on three different phenotypes in the current study, gyrification, cortical thickness, and white matter volume, future research should expand on these findings to better understand *SEMA6D*'s effect on the brain. There are limitations to measuring global white matter volume underlying grey matter regions of interest because it limits our understanding of

the whole white matter pathway. Therefore, effects on white matter could be better understood by using fractional anisotropy or radial or axial diffusivity for a finer look at impact on white matter structure. In addition, the use of functional methodology, such as fMRI, would improve understanding of how *SEMA6D* influences brain activity, and how this may affect reading. Lastly, future research may benefit from examining other subcortical brain regions, such as the basal ganglia, as well as the corpus callosum and brain regions involved in visual processing, as *SEMA6D* is associated with axon guidance during development of the corpus callosum and retinal mapping (Alto & Terman, 2018).

Conclusions:

Overall, results of the current study suggest that the variation in *SEMA6D* is associated with the variation in the brain structure within the reading network. Specifically, SNPs in *SEMA6D* were associated with gyrification in the supramarginal gyrus, cortical thickness in the fusiform gyrus, and white matter volume in the transverse temporal gyrus. With respect to brain-behavior relations, regardless of genotype, white matter volume in the transverse temporal gyrus was most strongly related to reading, possibly due to the young age of the participants and their still developing reading skills, likely relying mostly on phonological processing. While *SEMA6D* was associated with several reading-related brain regions, these regions fluctuate in their role in reading development depending on the stage of reading, whether it is phonological processing or automatic recognition of words. Therefore, *SEMA6D*, through its associations with various reading-related brain regions, may indirectly impact reading at various stages of reading development. *SEMA6D* has known functions in axon guidance and synapse formation, likely influencing the development of the brain, white matter structure, and synaptic connections. The

results show that *SEMA6D* gene and its variation appear to be associated with individual differences in performance on language and reading, and further research should focus on an older population, more phenotypes focused on white matter, and potential gene interactions with other genes and the environment.

PART 3

Genetic and neural risk of SRD within a family case study: A clinical application

Tina Thomas[1]

Department of Psychology
University of Houston, Houston, TX, USA

The preparation of this article was supported by P50 HD052117 (PI: Fletcher), P20 HD091005 (PI: Grigorenko), and P50 HD052120 (PI: Wagner). Grantees undertaking such projects are encouraged to freely express their professional judgment. Therefore, this article does not necessarily reflect the position or policies of the abovementioned agencies, and no official endorsement should be inferred.

Abstract

Imaging and genetic studies have helped to characterize the biological risk factors contributing to development of specific reading disability (SRD). The current study aimed to apply this research literature to an individual family consisting of twins discordant for SRD and an older sibling with reading difficulty, to examine how this knowledge could be applied to understand clinical risk within a clinical case. Intraclass correlations were used to understand similarity of imaging phenotypes between each pair. Reading-related genes and brain region phenotypes, including asymmetry indices representing the relative size of structures in the left as compared to right hemispheres, were then descriptively examined to determine which risk factors were related to reading within the family. For genetics, there were SNPs that corresponded between the SRD siblings and not the typically developing (TD) sibling in the genes *ZNF385D*, *LPHN3*, *CNTNAP2*, *FGF18*, *NOP9*, *CMIP*, *MYO18B*, and *RBFOX2*. Overall imaging phenotypes tended to be similar among all sibling pairs for grey matter volume and surface area, but cortical thickness in reading-related ROIs was more similar among the siblings with SRD, followed by the twins, and then the TD twin and older sibling, suggesting cortical thickness may be an important differentiator of risk for this family. Siblings with SRD tended to have more symmetry in cortical thickness in the transverse temporal and superior temporal gyri, while the TD sibling had greater rightward asymmetry. The TD sibling had a greater leftward asymmetry of grey matter volume and cortical surface area in the fusiform gyrus, supramarginal gyrus, and transverse temporal gyrus. Overall, this exploratory study demonstrated that reading-related risk factors appeared to correspond with SRD within this particular family, suggesting that examination of biological risk factors early on may be beneficial for early identification and

intervention. Future studies may benefit from the use of polygenic risk scores or machine learning to better combine and integrate imaging genetics knowledge to understand SRD risk.

Introduction

Specific reading disability (SRD) has both genetic and neural risk factors, characterized by an evolving research literature. However, there has been less focus on the application of this research literature to better understand risk within individuals or families. As understanding of genetics and neural factors increases, it is likely to impact the development of precision medicine where diagnosis and treatment can be catered to the individual based on their specific risk factors using methodology such as polygenic risk scores (Breen et al., 2016; Dick, 2019). Reiss (2010) suggested that genetic studies are beneficial for intervention because they help to identify important target variables for intervention, provide information about mechanisms of effects, may differentiate individual response to intervention, and help to reveal the best timing for intervention. Genetic risk factors interact with various environmental variables in affecting reading-related outcomes, and this information can be used to help cater treatment to specific individuals as research develops (Belsky & van Ijzendoorn, 2015). The incorporation of neural factors into calculation and understanding of risk would further improve application of this knowledge into helping individual families understand diagnoses and intervention options (Wandell & Le, 2017), as neural function and structure have been shown to predict and differentiate response to intervention (Aboud et al., 2018; Nugiel et al., 2019; Odegard et al., 2008; Rezaie et al., 2011) and change with intervention (Barquero et al., 2014; Romeo et al., 2018; Simos et al., 2002). Response to intervention is the gold standard for diagnosis of SRD, differentiating those individuals who are able to respond given adequate intervention compared

to those who do not (Miciak & Fletcher, 2020), and these latter individuals have a different biological risk profile compared to those who do. The use of imaging and genetic factors to understand individual risk may provide valuable information about the level or type of intervention that is most appropriate for an individual, and how genetic risk interacts with the environment (Belsky & van Ijzendoorn, 2015; Dick, 2019).

Phase 1 genetic translational research focuses on the transition from genome-based discovery of relevant genes to into the application of these findings to individuals (Khoury et al., 2007). The current study examines biological risk factors within an individual family case study, using the SRD imaging genetics literature to select genes and brain structures that may be relevant to risk within the family, and determining whether there is evidence of SRD risk within the family. Because the family consists of a pair of twins, one with SRD and one without, as well as an older sibling with reading difficulties, the risk factors that vary with SRD within the family, and are more similar among those siblings with SRD, are likely important for understanding risk within this family. A recent meta-analysis synthesizing results of twin studies of SRD suggested that the heritability of reading ability is thought to be 66%, with a shared environment effect of 13%, and a non-shared environment effect of 21% (Andreola et al., 2021). Given the similar environment of each of these siblings, particularly the twins, the biological factors contributing to development of SRD may be particularly important in understanding risk. To better understand risk within this family, reading-related genes and brain structures were examined.

Genetic risk factors likely influence brain structure and function in networks associated with reading development. SRD risk genes have been identified through both candidate gene studies and genome-wide association studies (GWAS). SRD is genetically heterogeneous,

meaning the same phenotype can be produced through different genetic mechanisms. Furthermore, it is influenced by small effects from many genes. Much of the genetic research focused on SRD has been focused on SRD risk loci, named DYX1-DYX9, which reside on eight different autosomal chromosomes in 9 chromosomal locations (Mascheretti et al., 2017). Within these SRD risk loci, genes such as *DYX1C1*, *DCDC2*, *KIAA0319*, and *ROBO1* have been replicated in most studies. Other genes outside of these specific loci have been identified as well (Mascheretti et al., 2017; Landi & Perdue, 2019). Some genes, such as *FOXP2* and *CNTNAP2*, have also been investigated in SRD due to their relationship to language (Raskind et al., 2013). Genome wide association studies have also been conducted to identify other genes that may be associated with reading skills. For example, a recent genome-wide association study revealed a nominal association of a SNP near the gene *FGF18* with SRD (Field et al., 2013). Results of a linkage analysis suggested relationships between SRD and SNPs within the *MSI2* gene and upstream of the ADHD-related *LPHN3* gene (Field et al., 2013). Similarly, another recent GWAS implicated an uncharacterized gene, *LOC388780*, and the gene *VEPH1*, related to brain development (Gialluisi et al., 2020). As this literature and methodology develops, it improves the ability to use polygenic risk scores combining effects of multiple risk variants in predicting SRD, which can lead to more accurate early identification and intervention for individuals with SRD (Dick, 2019).

Incorporating the use of imaging methodology has also improved understanding of risk factors contributing to SRD by examining associations between genes and neural factors related to reading, which are closer to the level of the gene. For example, another GWAS using an EEG mismatch negativity signal, related to processing of speech sounds, revealed associations with

SNPs related to mRNA-expression levels of the gene *SLC2A3* (Roeske et al., 2011). Because the reading brain network has been well characterized, it can be helpful to examine how these reading-related regions are associated with genes, or how brain structure or function relates to SRD risk. Reading-related brain regions include left hemispheric networks, including a dorsal pathway related to phonological processing in occipital-temporal and the inferior parietal lobe, a second ventral pathway related to more automatic word reading in the left temporal lobe and fusiform gyrus, and a frontal network, including Broca's area, involved in attention and mental verbalization (Cattinelli et al., 2013). Meta-analyses of structure and function have also demonstrated consistent structural and functional changes within these regions in individuals with SRD. For example, adults with SRD demonstrated underactivation in superior temporal regions, while children with RD demonstrated underactivation in inferior parietal regions (Richlan et al., 2011). Similarly, nine studies found reduced grey matter volume in the right superior temporal gyrus and left superior temporal sulcus (Richlan et al., 2013). Diffusion tensor imaging studies have indicated that there are lower fractional anisotropy levels in left temporoparietal and frontal regions, including the left arcuate fasciculus and corona radiata (Vandermosten et al., 2012).

While the reading network tends to be localized in the left hemisphere, atypical structure and function in these reading-related brain regions may also be related to atypical lateralization or brain asymmetry in individuals with SRD. For example, children with SRD have been shown to have atypical asymmetry in the inferior frontal-occipital fasciculus (leftward) and superior longitudinal fasciculus (rightward), which in turn was related to reading skills (Zhao et al., 2016). In turn, children with SRD may have more compensatory activation in right hemispheric

regions such as the right superior temporal gyrus (Waldie et al., 2017). Children without SRD have been shown to have greater rightward asymmetry of the cerebellum compared to those with SRD (Kibby et al., 2008). These differences in lateralization may exist even before the development of reading, as Guttorm and colleagues (2009) showed that pre-reading children under age 5 with a family risk of SRD tended to have atypical speech processing in the right hemisphere on EEG. Similarly, children with no family risk for SRD tended to have greater leftward asymmetry of the planum temporale compared to those with family risk (Vanderauwera et al., 2018). It has been proposed that reading-related genes such as *DYX1C1*, *ROBO1*, and *DCDC2*, which have functions that contribute to ciliogenesis and cilia function, may contribute to processes contributing to asymmetry such as development of the corpus callosum or the direction of neuronal migration (Paracchini, Diaz, & Stein, 2016).

A greater understanding of these genetic and imaging factors may hold promise for using these methodologies for clinical assessment of disorders such as SRD. Some imaging tools have been developed for assessment of disorders such as Alzheimer's and Multiple Sclerosis, but there has been less application to other types of learning or psychiatric disorders, potentially because of the smaller effects of many genes and structures, a lack of established biomarkers, and overlap across various learning and psychiatric disorders (Scarpazza et al., 2020). Therefore, the clinical application of imaging research is still underdeveloped. Similarly, while understanding of genetic risk factors for SRD has been increasing, there is often variability across genome wide association studies, there are likely genetic factors that have not yet been identified. There may also be significant interaction with environmental factors in affecting risk (Lancaster et al., 2020). Similarly, the use of a polygenic risk score to predict SRD revealed significant overlap in

etiology with word reading and other neurodevelopmental disorders such as ADHD and ASD, as well as general cognitive ability (Price et al., 2020). Using both imaging and genetics methodology in combination may increase ability to understand risk. Therefore, while there is no current diagnostic tool, it may be beneficial to use the current literature basis on imaging and genetic factors contributing to SRD and apply it to clinical cases to better understand the specific genetic and imaging risk factors that may be related to development of SRD in these cases. The current study services as an exploration of impact within the individual family as part of phase 1 translational research. Future research may focus on using additional methodologies for application of genetic testing and intervention, incorporating genetics into evidence-based guidelines and health practice, and evaluating the beneficial clinical impact of this type of research (Khoury et al., 2007).

The aim of the current study was to apply the research literature in a clinical case study by examining the genetic and imaging factors contributing to SRD in a single family, including a pair of twins, one with reading difficulty and one without, and an older sibling with reading difficulty as well. This allowed us to examine how the previous imaging genetics literature of SRD is related to risk of SRD within a clinical family case, allowing the translation of research-based data into more clinical relevance. The unique profile of this family allowed examination of how genetic and imaging patterns, including structural asymmetry, covary with SRD. Specifically, patterns were descriptively examined to identify relevant brain regions and single nucleotide polymorphisms that were more similar among the siblings with SRD, as compared to the twins. This also allowed us to determine which specific genes and structures, within the

larger group of known reading-related genes and brain structures, were most related to SRD within this sample.

Methods

Participants

Participants were members of white/Caucasian family, consisting of a pair of dizygotic twins (age 9), one with a diagnosed SRD and one without, an older sibling (age 11) with a reported history of reading difficulty, and their parents. All siblings were female. Imaging and behavioral data was collected from the three siblings, and genetic data was collected from all five family members. The collection of this data was approved by the University of Houston Institutional Review Board (IRB# HSC-MS-12-0259 and CR00001300). Written informed consent was obtained from the parent or legal guardian of minor participants.

Behavioral Assessments

Assessments of word reading included the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), a timed measure of an ability to read printed words, the Kaufman Test of Educational Achievement, 3rd edition (KTEA-3; Kaufman & Kaufman, 2014), including the subtests Letter Word Recognition, measuring word reading ability, Nonsense Word Decoding, measuring decoding abilities, Reading Comprehension, and Listening Comprehension.

Genetic Data

Oragene saliva kits (DNA Genotek, Inc.) were used to obtain saliva samples during behavioral testing sessions. Genomic DNA was extracted from the samples using the FlexiGene DNA Kit (Qiagen) per the manufacturer's protocol. Genotyping was carried out at the Human Genome Sequencing Center of the Baylor College of Medicine according to the manufacturer's instructions (Illumina). Genotyping was completed with the Infinium CoreExome-24 v1.4 BeadChip, which contains 567,218 probes. Illumina's GenCall algorithm was run on the raw genetic data to cluster and call genotypes and assign confidence scores. QC filtering was applied to each sample separately, with a no-call threshold of 0.15. Therefore, all genotypes with a GenCall score less than or equal to 0.15 were assigned as missing, since they are considered too far from the cluster centroid to be reliably genotyped. Variants for 28 genes of interest, reading-related genes, were extracted to examine whether genotypes varied with reading difficulty within the family.

Imaging Data

MRI data were collected to obtain data on brain morphometry. The data were collected at the Core for Advanced Magnetic resonance Imaging (CAMRI) at Baylor College of Medicine. Structural data covering the whole brain was obtained, using a Siemens 3T scanner with a 64 channel head coil and slice-accelerated, simultaneous multislice imaging sequence (0.8mm3 slice thickness, FOV=256x256, TR/TE=2400/2.22ms, α=8 (Nugiel et al., 2019). Freesurfer v6.0.0 software (Fischl, Sereno, & Dale, 1999; Dale, Fischl, & Sereno, 1999) was used to obtain high-resolution anatomical images with an accelerated 3dT1-weighted sequence (Nugiel et al., 2019). Imaging data was registered to fsaverage on Freesurfer to allow comparisons across participants. Imaging phenotypes were grey matter volume, cortical thickness, and cortical surface area to

better understand the underlying grey matter factors that may be related to reading disability within a clinical case study. Heritability tends to be higher for cortical surface area and volume, while thickness is more likely to be impacted by environmental factors (Hofer et al., 2020), helping to differentiate phenotypes that may represent underlying risk factors and those that may be related to compensation.

Analyses

For genetic analyses, reading-related risk genes were identified from the literature, and included *MRPL19, ZNF385D, ROBO1, VEPH1, LPHN3, FGF18, DCDC2, KIAA0319, TTRAP, THEM2, RIPOR2, CMAHP, FOXP2, CCDC136, CNTNAP2, SLC2A3, COL4A2, NOP9, TUBGCP5, CYFIP1, NIPA2, SEMA6D, DNAAF4, CMIP, ATP2C2, MSI2, MYO18B,* and *RBFOX2*. Genotypes of SNPs associated with these previously identified reading-related genes were determined. SNPs that were the same within all family members were removed, leaving 684 SNPs remaining for the analysis. Genetic patterns were examined to identify SNPs for which the genotypes were the same between the sibling with reading difficulties, but for which the twin differs. The proportion of genotypes following this pattern was compared between SRD risk genes and the rest of the genome.

For imaging analyses, multiple phenotypes were used to examine patterns in cortical thickness, surface area, and grey matter volume in the reading network. Pairwise intraclass correlations (ICCs) were used to determine the degree of similarity of imaging phenotypes between each pair of twins, allowing us to determine how closely the siblings resemble each other. The ICCs were calculated for each phenotype (e.g. cortical thickness) using average values for each region of interest covering the whole cortex, and then for just the reading-related regions

(fusiform gyrus, inferior parietal, banks superior temporal sulcus, inferior frontal gyrus pars opercularis and pars triangularis, supramarginal gyrus, and transverse temporal gyrus).

Asymmetry indices were calculated for phenotypes in specific regions of interest as a measure of lateralization of regions. Asymmetry indices were calculated as (L-R)/((L+R)/2). A positive value indicates leftward asymmetry, while a negative value indicates rightward asymmetry. Imaging phenotypes and asymmetry were plotted to visualize patterns between the siblings, and identify regions where phenotypes may be related to SRD.

Results

Behavioral data (see Table 1) revealed low average reading abilities for one twin and the older sibling, and average reading abilities for the second twin. Scores tended to be lower in those siblings with SRD in both word reading and decoding, as well as reading comprehension.

Table 1. Sibling behavioral data

Participant	**KTEA Letter/Word Identification**	**KTEA Nonsense Word Decoding**	**KTEA Reading Comprehension**	**KTEA Listening Comprehension**
Twin 1	82	90	82	106
Twin 2	105	113	103	109
Older Sibling	90	84	83	93

Genetic Analysis

In the genetic analysis, there were 79 SNPs where the siblings with SRD had the same genotype, but the typically developing twin did not. This was about 10% of the overall SNPs

within the reading genes examined, consistent with the rest of the genome. Overall, there were 33 SNPs identified in *ZNF385D,* 17 SNPs in *LPHN3,* 9 SNPs in *CNTNAP2*, 2 SNPs in *FGF18,* 2 SNPs in *NOP9,* 11 SNPs in *CMIP*, 4 SNPs in *MYO18B,* and 1 SNP in *RBFOX2.* Fifty-seven of these SNPs were intronic, 1 SNP was exonic (in *NOP9*), 1 SNP was in the three prime untranslated region, and 20 SNPs were intergenic.

Cortical Thickness

Pairwise ICCs were used to determine the degree of similarity of average cortical thickness across regions of interest across the whole brain, and then across just reading-related regions. For cortical thickness in regions of interest across the whole brain, the ICCs were similar across each pair of siblings (0.85 for the twins, 0.87 for the siblings with SRD, and 0.82 for the last pair of siblings). Across reading-related regions, the ICC for cortical thickness was highest for the siblings with SRD (ICC = 0.76). The ICC was 0.66 for cortical thickness in reading-related ROIs in the twins, and the ICC was 0.54 for the TD twin and older sibling.

For asymmetry indices calculated for cortical thickness, regions that showed greater correspondence between the two siblings with reading difficulty, as compared to the twins, were visualized (see Figure 1). In the transverse temporal gyrus, siblings with SRD showed symmetric or slightly leftward asymmetry, while the twin without SRD had greater thickness in the right hemispheric transverse temporal gyrus. Three was a similar pattern, with greater symmetry in the impacted siblings and greater thickness in the right hemisphere in the typically developing (TD) sibling in the superior temporal gyrus. These regions were examined in more detail through creation of plots of average cortical thickness values in these regions.

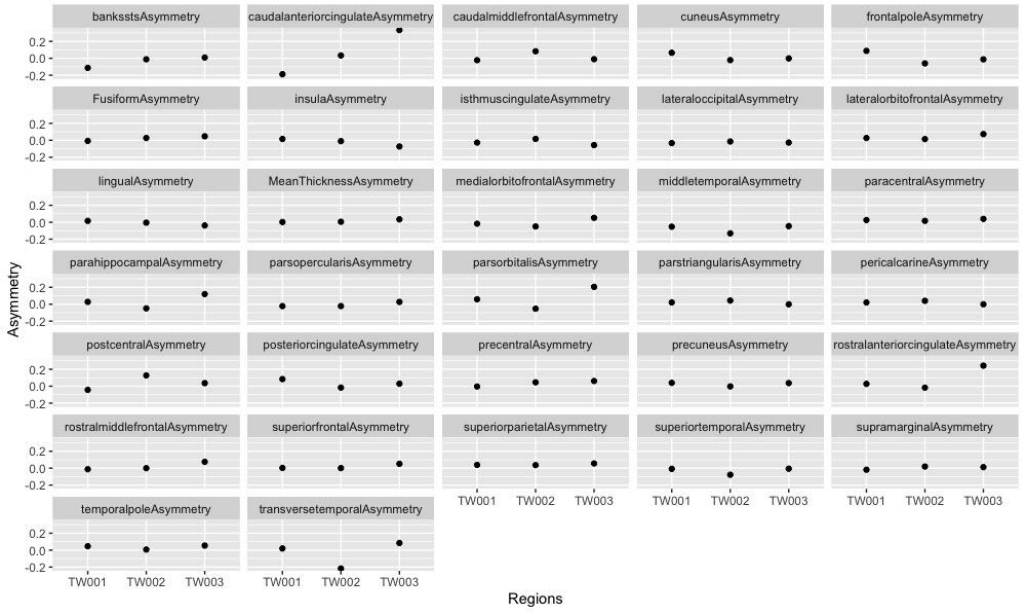

Figure 1. Plots of cortical thickness asymmetry for each sibling

When average cortical thickness values were examined in these regions, the TD sibling appeared to have lower cortical thickness in the left hemisphere as compared to the SRD siblings, and higher cortical thickness in the right hemispheric transverse temporal gyrus as compared to the SRD siblings.

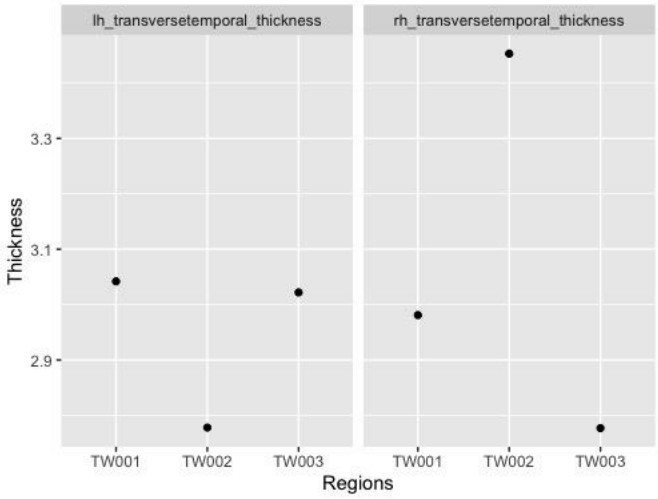

Figure 2. Plots of cortical thickness in the transverse temporal gyrus

In the superior temporal gyrus, siblings with SRD tended to have greater cortical thickness in the left hemisphere as compared to the TD sibling, but the twin with SRD had the greatest cortical thickness. In the right hemisphere, the SRD twin had the highest cortical thickness, followed by the TD twin, and then the older SRD sibling. Therefore, the left hemisphere may have been a greater driver of the asymmetry pattern.

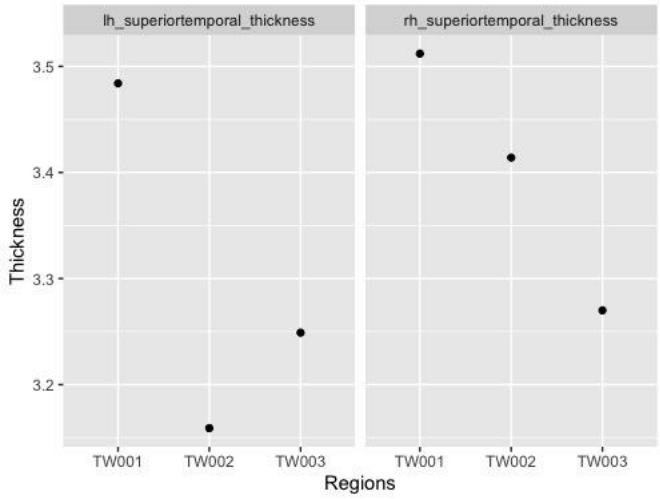

Figure 3. Plots of cortical thickness in the superior temporal gyrus

Grey Matter Volume

Pairwise ICCs were used to determine the degree of similarity of grey matter volume across regions of interest across the whole brain, and then across just reading-related regions. ICCs were similar across each pair when examining grey matter volume of regions, with high levels of correspondence. For grey matter volume in regions across the whole brain, the ICC was 0.97 for the twins, 0.98 for the SRD siblings, and 0.97 for the TD twin and older siblings. Across reading-related regions, the ICCs were similar, with ICCs of 0.97 for the twins, 0.99 for the SRD siblings, and 0.95 for the TD sibling and older sibling.

Regions where volume asymmetry was more similar between SRD siblings as compared to the twins include the rostral anterior cingulate, frontal pole, fusiform gyrus, hippocampal gyrus, pars orbitalis, postcentral gyrus, precentral gyrus, paracentral gyrus, and the supramarginal gyrus (Figure 2). Of these regions, the fusiform gyrus and the supramarginal

gyrus are within the reading network. For both of these regions, the TD sibling had greater leftward asymmetry, indicating greater volume in the left hemisphere.

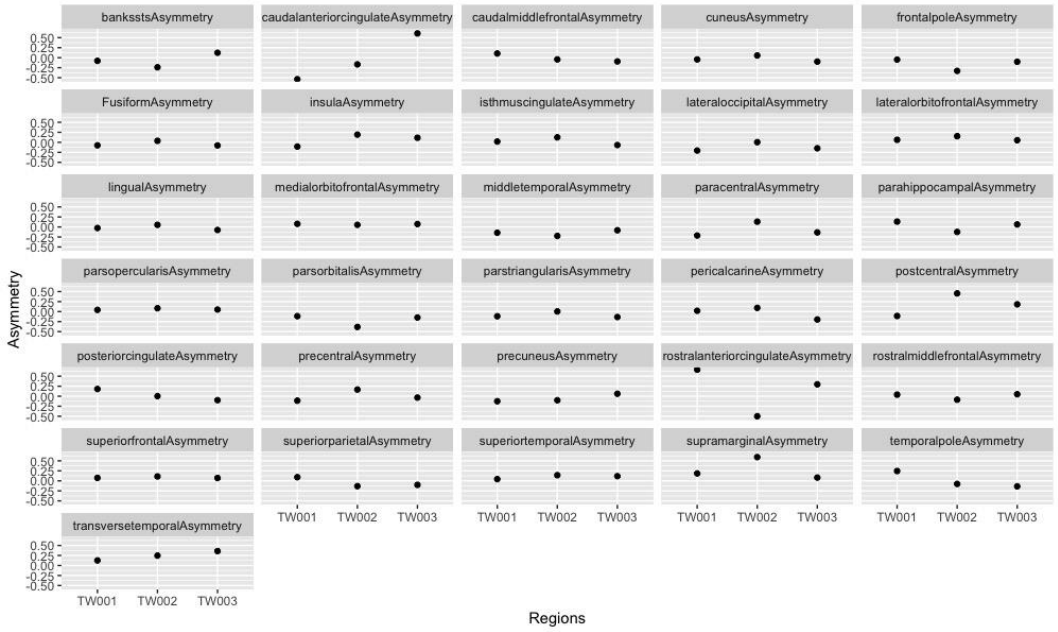

Figure 4. Plots of grey matter volume asymmetry for each sibling

In the fusiform gyrus, the SRD twin appeared to have the highest volume in both the right and the left hemispheres. When looking at global grey matter volume in these regions, the siblings do not fit the pattern of SRD siblings being more similar in each hemisphere. However, both the SRD siblings have a higher volume in the right hemisphere as compared to the left, while the TD sibling has a higher volume in the left hemisphere.

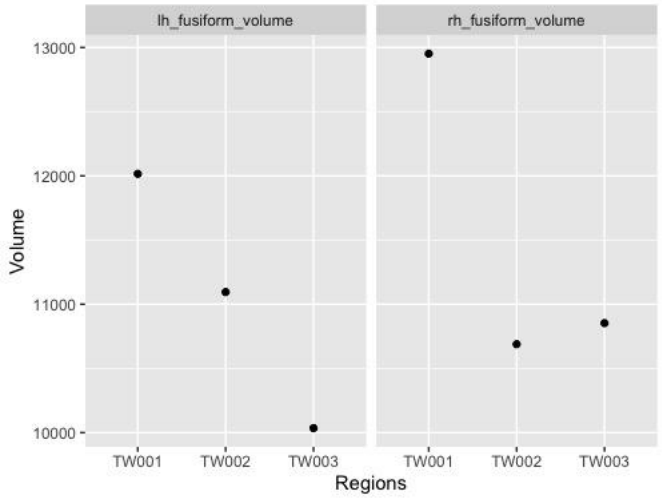

Figure 5. Plots of grey matter volume in the fusiform gyrus

In the supramarginal gyrus, the SRD siblings had a higher volume than the TD sibling in both the left hemisphere and right hemisphere, with greater overall differences in the right hemisphere.

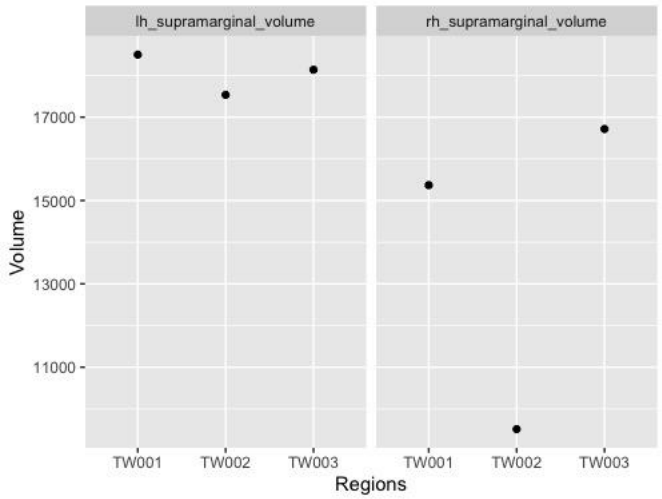

Figure 6. Plots of grey matter volume in the supramarginal gyrus

Cortical Surface Area

Pairwise ICCs were used to determine the degree of similarity of cortical surface area across regions of interest across the whole brain, and then across just reading-related regions. ICCs for cortical surface area were similar across each sibling pair. For cortical surface area in regions of interest across the whole brain, the ICC was 0.97 for the twins, 0.99 for the SRD siblings, and 0.96 for the TD sibling and older sibling. For cortical surface area in reading-related regions of interest, the ICC was 0.96 for the twins, 0.99 for the SRD siblings, and 0.95 for the TD sibling and older sibling.

Regions where surface area asymmetry was more similar between SRD siblings as compared to the twins include the banks of the superior temporal sulcus, cuneus, fusiform gyrus, lateral occipital gyrus, paracentral gyrus, postcentral gyrus, precentral gyrus, pars triangularis, rostral anterior cingulate, supramarginal gyrus, and the transverse temporal gyrus (Figure 3). Reading-related regions include the fusiform gyrus, supramarginal gyrus, and the transverse temporal gyrus. For all of these regions, the TD sibling had a more leftward asymmetry, indicated greater cortical surface area in the left hemisphere as compared to the right.

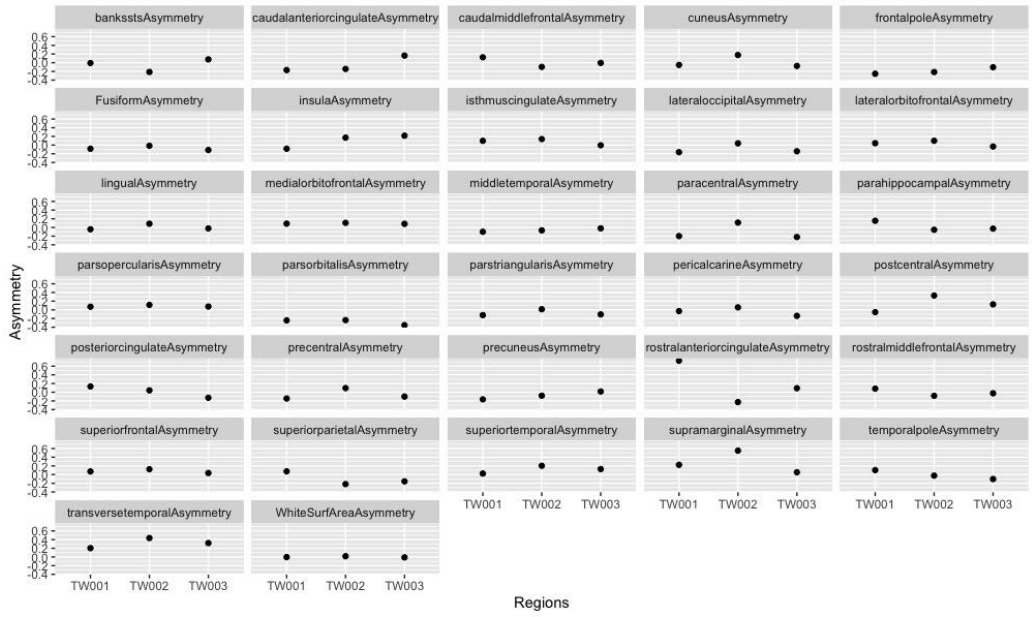

Figure 7. Plots of cortical surface area asymmetry for each sibling

In a comparison of global cortical surface area in each region, surface area of the fusiform gyrus was observed to be higher for the SRD twin. In the left hemisphere, the TD twin and SRD older sibling had more similar cortical surface area. In the right hemisphere, the SRD twin had the highest cortical surface area, the older sibling with SRD had the next highest area, and the TD twin had the lowest surface area. Both siblings with SRD appeared to have greater differences in surface area between hemispheres, with higher surface area in the right hemisphere.

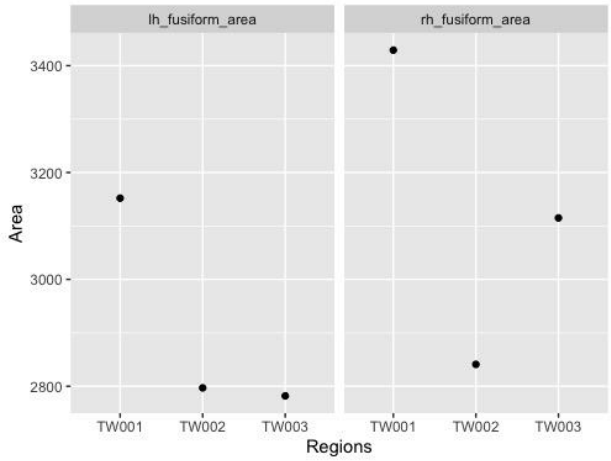

Figure 8. Plots of cortical surface area in the fusiform gyrus

In the supramarginal gyrus, the TD sibling was observed to have a lower cortical surface area than the SRD siblings in both hemispheres. The magnitude of difference was observed to be greater in the right hemisphere.

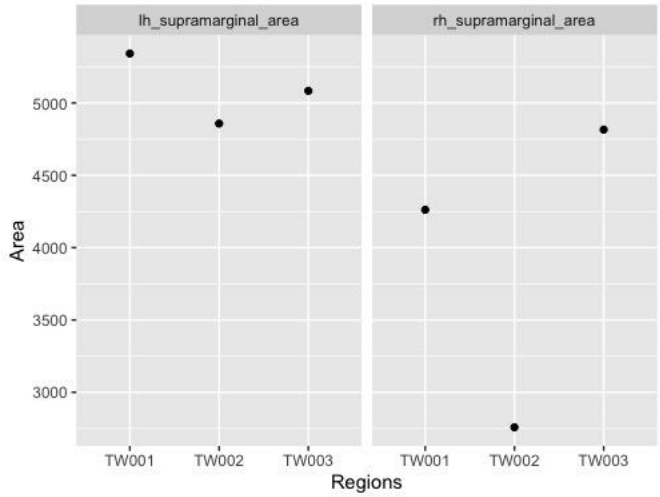

Figure 9. Plots of cortical surface area in the supramarginal gyrus

For the transverse temporal gyrus, the TD sibling had lower cortical surface are in both hemispheres as well.

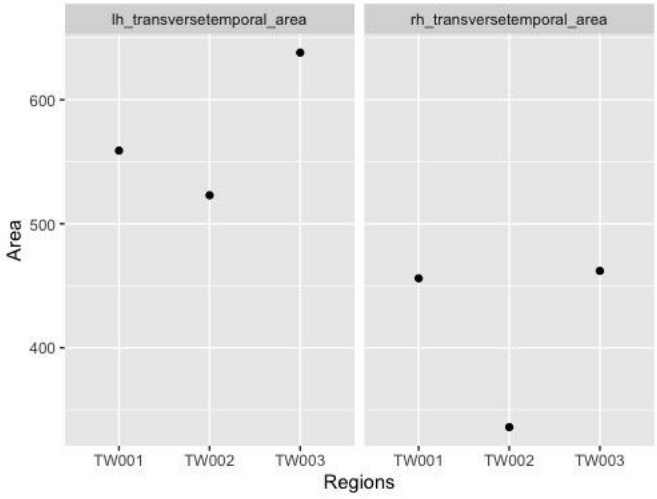

Figure 10. Plots of cortical surface area in the transverse temporal gyrus

Discussion

The current study used a family case study to better understand the translation of research-based findings to application in a clinical case. While many studies have helped identify brain structures comprising the reading network, and the genetic risk factors contributing to development of SRD, little research has been conducted on how these results may translate to an individual family. To achieve this goal, intraclass correlations were calculated to understand similarity of brain phenotypes between each sibling pair and how neural risk factors are related to SRD. In addition, the reading-related genes and brain structures that were more similar among siblings with SRD, as compared to twins with differing reading abilities, were descriptively examined by comparison of genotypes and visualization of imaging phenotypes to determine

which regions were most relevant to SRD within this family. Overall, there were many SNPs and brain structures within reading-related networks that fit this pattern, consistent with expectations based on the current imaging and genetics literature.

Relevant SNPs within reading-related genes were first identified by examining which genotypes were more similar between the SRD siblings as compared to the twins. The gene with the most SNPs fitting this pattern was *ZNF385D*, which has been previously associated with SRD as well as overall fiber tract volumes and global brain volume (Eicher et al., 2013). *ZNF385D*'s functions are in nucleic acid binding (Stelzer et al., 2016). *LPHN3* and *FGF18*, identified to be related to SRD from a GWAS (Field et al., 2013) also had SNPs following the selected pattern. *LPHN3*, also previously associated with ADHD (Bruxel et al., 2020), is involved in cell adhesion and signal transduction (Stelzer et al., 2016). *FGF18* codes for a fibroblast growth factor, involved in processes such as mitogenesis, cell proliferation, cell differentiation, and cell migration, and has been shown to be related to development of cerebellar structures in mice (Stelzer et al., 2016).

Another gene with SNPs fitting this pattern was *CNTNAP2*, which codes for a member of the neurexin family and functions in cell adhesion (Stelzer et al., 2016). *CNTNAP2* has been shown to be related to nonword reading and language abilities (Peter et al., 2011). In addition to SRD, it has also been associated with autism spectrum disorder, intellectual disability, language impairment, and schizophrenia (Rodenas-Cuadrado et al., 2013). In a Chinese sample, the relationship between *CNTNAP2* and risk of SRD was higher in females (Gu et al., 2018). *NOP9*, which had 2 SNPs following the chosen pattern, has been associated with language through effects of paternal SNPs on child reading (Nudel et al., 2014; Pettigrew et al., 2016). *NOP9*

codes for a binding protein which may regulate cell processes such as transcription and translation (Stelzer et al., 2016).

CMIP, which contained 11 SNPs varying with SRD, has been related to reading skills (Scerri et al., 2011), as well as short term memory skills in language impairment (Newbury et al., 2009). It has also been weakly related to SRD in a Chinese sample (Wang et al., 2015). *CMIP* is involved in T-cell signaling (Stelzer et al., 2016). 4 SNPs in *MYO18B,* which has been associated with math skills in children with SRD (Ludwig et al., 2013) fit the relevant pattern as well. *MYO18B* has functions in nucleotide binding, intracellular trafficking, and motor activity (Stelzer et al., 2016). There was one SNP in *RBFOX2* that varied with SRD in the siblings. *RBFOX2*, which is a regulator of alternative splicing in neurons (Stelzer et al., 2016), has been associated with reading and language in a genome wide association study (Gialluisi et al., 2014). Further imaging genetic investigations have revealed associations with cortical thickness, particularly in the left parahippocampal gyrus, right middle temporal gyrus, right inferior frontal gyrus, and bilateral superior temporal gyrus (Gialluisi et al., 2017).

The described SNPs in reading-related genes were all shown to be more similar between the siblings with SRD as compared to the twins, who would be expected to be more similar, suggesting that these specific SNPs may be important for adding to reading-related risk within this family. However, it is also important to realize that there may be SNPs fitting this pattern that are not related to reading. Therefore, while this exploratory analysis identified SNPs that may be particularly important for understanding SRD in this particular family, future research should further investigate these SNPs using statistical analyses, specifically examining their relationship to reading measures and brain structures. Furthermore, future research may use a

polygenic risk score to better understand the cumulative and interacting effects of all of these identified SNPs in conveying risk of SRD.

Imaging results, when comparing the overall similarity between sibling pairs using ICCs, suggest that measures of grey matter volume and cortical surface area were highly similar among all three siblings. There was more variability when examining similarity of cortical thickness between sibling pairs. The ICCs for cortical thickness between sibling pairs ranged from 0.82-0.87 when calculated across regions covering the whole cortex. However, when the focus was on cortical thickness in only reading-related regions, the siblings with SRD tended to have a greater degree of similarity as compared to the twins, with the TD twin and older sibling having the lowest similarity. This suggests that cortical thickness may be a key differentiator of SRD risk within individual families, as measures of grey matter volume and cortical surface area may be less sensitive.

Visualization of individual brain regions revealed differences in cortical asymmetry of cortical thickness in the transverse temporal gyrus and superior temporal gyrus. The siblings with SRD were observed to have more cortical symmetry, while the typically developing symmetry and greater rightward asymmetry of cortical thickness. Regarding volume, the typically developing sibling tended to have greater leftward asymmetry of the fusiform gyrus and supramarginal gyrus. There was a similar pattern with cortical surface area, as the typically developing sibling had greater leftward asymmetry of cortical surface area in the fusiform gyrus, supramarginal gyrus, and transverse temporal gyrus as compared to the siblings with SRD. Overall, the typically developing sibling had more leftward asymmetry of grey matter volume and cortical surface area in reading-related regions, and more rightward asymmetry of cortical

thickness. While typically, a reduced leftward asymmetry would be expected in SRD (Zhao et al., 2016), the greater rightward asymmetry of cortical thickness in the typically developing sibling may reflect compensation, as cortical thickness tends to be more impacted by environmental factors (Hofer et al., 2020). Grey matter volume and cortical surface area, conversely, fit the pattern of having a greater leftward asymmetry in the typically developing sibling, which may be more reflective of genetic risk of SRD given the higher genetic association of these phenotypes (Hofer et al., 2020). Therefore, examination of the degree of cortical asymmetry in reading-related brain regions may be an important measure for understanding risk within individuals or families.

These results corresponded with previous literature investigating asymmetry of cortical structures in relation to SRD. For example, a prior study found that children who had no family risk of SRD had greater leftward asymmetry of cortical surface area in the planum temporale, in the superior temporal region posterior to Heschl's gyrus, as compared to those with risk (Vanderauwera et al., 2018). Leonard and colleagues (2002) created an anatomical risk index based on leftward asymmetry of the planum temporale, combined plana, the cerebellar anterior lobe, rightward asymmetry of cerebral volume, and larger overall values for cerebral volume and surface areas of Heschl's gyri. They found that negative risk indices, indicated by smaller and more symmetrical brain structures, were related to comprehension deficits, while positive risk indices, or having larger, asymmetrical brain structures, were related to poor word reading (Leonard et al., 2002; Leonard et al., 2006; Leonard & Eckert, 2009). In our sample, the superior temporal region was characterized by greater symmetry in SRD siblings, which is typically more related to comprehension deficits according to Leonard and colleagues (2002) anatomical risk

index. However, the siblings with SRD in our sample had difficulties with both reading comprehension and word reading, so the symmetry in the current sample may provide additional information about asymmetry in children with deficits in both areas. Atypical asymmetry has been shown in relevant white matter structures as well, as children with SRD were shown to have reduced leftward lateralization of white matter structure in the inferior frontal-occipital fasciculus, and greater rightward asymmetry of the superior longitudinal fasciculus (Zhao et al., 2016).

Reading-related genes, such as those involved in cilia function, may affect brain development and asymmetry through their effect on processes such as neuronal migration (Paracchini et al., 2016), which could affect these structures even before reading develops. For example, children with a family risk of SRD show atypical lateralized speech processing measured through EEG, which is further related to reading (Guttorm et al., 2009). While these studies demonstrate the link between asymmetry of brain structures and reading, it is difficult to determine directionality or causality of these effects, and whether asymmetry influences reading or vice versa, or whether this interaction may change over time (Bishop et al., 2013). Future longitudinal studies using larger samples may directly examine how asymmetry influences reading development, which may later contribute to clinical understanding as well.

Further examination of the patterns in the left and right hemisphere for those regions that were identified as having different asymmetry in the TD and SRD siblings provided further information about the brain phenotypes in these specific regions. For example, the TD sibling had lower cortical thickness than the SRD siblings in the left hemispheric transverse temporal gyrus, and higher cortical thickness in the right hemisphere. Regarding grey matter volume in the

transverse temporal gyrus, the TD sibling appeared to have smaller volumes in both hemispheres. Grey matter volume results were consistent with prior literature, as an increased size of Heschl's gyrus has been previously associated with decreased reading performance (Phinney et al., 2007), and the overall size has been used to distinguish children with difficulties in phonological decoding as compared to overall verbal ability (Leonard et al., 2001). However, the pattern of lower cortical thickness in the TD sibling in the left hemisphere in the current study conflicts with prior literature that found reduced cortical thickness in the left Heschl's gyrus in children who later developed SRD (Clark et al., 2014). The current results also conflict with previous literature demonstrating that cortical thickness in the left superior temporal cortex is positively associated with word reading (Perdue, Mednick, Pugh, & Landi, 2020). Given these differences, it may be helpful to consider other types of measures, such as asymmetry, to better understand the contribution of left hemispheric and right hemispheric structures to reading-related risk.

For the fusiform gyrus, involved in visual word recognition (Dahaene et al., 2005), the TD sibling had greater leftward asymmetry as compared to the SRD siblings, although this pattern was not evident when specifically examining global grey matter volume in only the left hemisphere. Regarding cortical surface area, the TD sibling had a similar cortical surface area to the older SRD sibling in the left hemispheric fusiform gyrus but was observed to have a lower surface area than the SRD siblings in the right hemisphere. Therefore, it may be important to take both hemispheres into account to better understand the reading network and how it relates to reading-related risk. In the previous literature, better reading skills have been associated with a larger fusiform gyrus (Torre & Eden, 2019), and there have been findings of reduced grey matter

in both the left and right fusiform gyrus in children with SRD (Kronbichler et al., 2007), indicating that there may be effects in both hemispheres, although larger volumes tend to be related to better reading. In addition, following intervention, children with SRD demonstrated increases in grey matter volume in the left anterior fusiform gyrus (Krafnick et al., 2011). This relative change in structure in an individual may be an important indicators of reading development, particularly if these changes are compared relative to other brain regions. This may also contribute to a clinical application in better understanding how an individual may be responding to intervention or changing as their reading develops.

In the supramarginal gyrus, the TD sibling had a lower overall grey matter volume and cortical surface area as compared to the SRD siblings in both the right and left hemispheres, but a greater leftward asymmetry in this region. Therefore, while grey matter volumes were relatively smaller in the TD sibling, the left was larger relative to the right. While previous literature demonstrates positive associations between grey matter volume in the left supramarginal gyrus and reading (Jednoróg et al., 2015), and smaller gray matter volume in the right supramarginal gyrus in children with SRD relative to TD children (Kronbichler et al., 2007), again, it may be important to take into account the relative size of the left and right hemispheric structures when showing relationships to reading. While global volume and surface area measured in the current study conflicted from previous literature, a leftward asymmetry in size of the supramarginal gyrus may be a more relevant indicator of association with reading in an individual. Using measures such as asymmetry can help consider sizes or characteristics of structures relative to other structures with an individual's brain, which improves understanding of relationships between various brain structures. Using asymmetry as a measure allows a

comparison of relative size of structures, which may be more meaningful in understanding an individual's SRD risk, given differences in factors such as overall brain size making it more difficult to interpret overall size of a region.

The current study used measures of asymmetry as well as global measures of grey matter volume, cortical thickness, and cortical surface area to examine the clinical application of research findings to better understand risk factors related to SRD within a single family, a promising start to phase 1 of genetic translational research (Khoury et al., 2007). The findings demonstrated that some of the critical brain regions and genes that have been previously associated with SRD also co-varied with SRD in the current family, including the transverse temporal gyrus, supramarginal gyrus, and fusiform gyrus. In addition, cortical thickness in reading-related regions tended to be overall more similar among the SRD siblings as compared to the twins, demonstrating that this may be a critical variable in understanding risk. Genotype patterns were also used to understand which specific reading-related SNPs may be related to SRD in the family as well. Given the brain regions and genes that covaried with SRD were generally consistent with the previous literature, the current study demonstrates the promise of using biological risk factors to better identify risk. As knowledge of the genetic factors contributing to certain disorders is increased, the idea of "precision medicine" and catering intervention based on an individual's specific risk factors has been developing (Belsky & van Ijzendoorn, 2015; Dick, 2019; Reiss, 2010). Earlier identification of these risk factors in clinical cases may help clinicians and families understand who may benefit from early reading intervention, allowing earlier access to services. In addition, as response to intervention research increases, it may be helpful to specifically investigate the biological factors that differentiate

those that respond compared to those that do not respond to reading intervention through understanding of gene-environment interactions. For example, SNPs close to the *DRD2* gene were associated with the level of improvement from a working memory intervention (Söderqvist et al., 2014). Similarly, various imaging phenotypes have been shown to predict or differentiate those who respond to reading interventions (Aboud et al., 2018; Nugiel et al., 2019; Odegard et al., 2008; Rezaie et al., 2011). Using response to intervention in research on biological factors may help to identify those genes that better differentiate those individuals with a true SRD, which can further help with diagnosis and catering of individualized intervention (Dick, 2019). Furthermore, biological risk factors may help us to better understand the population of individuals with SRD, including subgroups, as there is significant variability in the phenotype of SRD as well as overlap with other neurodevelopmental disorders (Price et al., 2020).

Neural factors, including structure and function, are also highly relevant for understanding risk and intervention response, as well as potential targets for intervention. A recent study demonstrated that brain activation mediated the relationship between a polymorphism in the gene *BDNF* and reading (Mascheretti et al., 2021). Furthermore, neural activation may be an active target to improve reading performance for individuals who do not respond to intervention, using techniques such as transcranial direct current stimulation (tDCS). For example, tDCS applied to the left posterior temporal cortex resulted in an increase in word reading efficiency in adult poor readers (Turkeltaub et al., 2012). In children, tDCS applied to temporo-parietal regions led to long-lasting improvements in reading 6 months afterwards (Costanzo et al., 2019). Across studies, these effects were particularly effective in poor readers, with specific improvements in word decoding for adults and non-word and low-frequency word

reading in younger children (Cancer & Antonietti, 2018), although there have been mixed results depending on the parameters used (Westwood & Romani, 2017). Other methodologies, including transcutaneous auricular vagus nerve stimulation, also led to improvements in decoding and automaticity of learning novel letter-sound relationships in a new language (Thakkar et al., 2020). This literature, while still developing, is also a promising clinical application for imaging genetics research, particularly for individuals who may be biologically identified as being at risk or likely to not respond to intervention.

While this study helped demonstrate the clinical relevance of research literature, findings also identified targets of future research. In larger samples used in future research, it would also be helpful to increase use other measures, such as functional connectivity measures, that give insight into how networks work together. Examining the relationships between multiple structures using asymmetry may provide additional insight into neural factors related to SRD. Using larger samples, particularly in a longitudinal study, to examine asymmetry would also allow more statistical analyses that may help to better understand causality and relationship of these phenotypes to specific reading skills. Analyses of lateralization and asymmetry may provide further insight into the development of reading and could serve as measures of progress in reading or predisposition to reading difficulties. Furthermore, the specific genes and brain structures examined in the current study are promising targets for future research. In terms of translational research, it is also important to investigate clinical impact in diverse populations, as allele frequencies and genetic risk may differ among individuals with different backgrounds (Dick, 2019). As these future research directions are tackled, they will provide increased knowledge about the biological etiology of SRD and how these biological risk factors interact

with the environment, improving our ability to understand interacting risk factors and contribute to early diagnosis and intervention.

GENERAL DISCUSSION

The present thesis aimed to improve understanding of the underlying neural and genetic risk factors affecting SRD, through summarizing the current literature, examining the effects of a previously unstudied gene, *SEMA6D*, and applying research findings to an individual family clinical case study to better understand the biological risk factors related to SRD within the individual family. Examination of both genetic and imaging factors helped to identify biological risk factors that may interact to influence differences in reading ability. Improving knowledge of these underlying biological factors that contribute to development of SRD may allow early detection of those individuals at risk as the relationship between genes, brain structure and function, and reading skills are better understood. The findings of the current thesis, integrated across all three studies, as well as a discussion of limitations and targets for future research are summarized below.

Genes and Reading

Part 1 of the current thesis, a review summarizing the current literature, revealed the strongest support and replication for SRD-related genes such as *DCDC2, KIAA0319, FOXP2, SLC2A3, ROBO1*, the *MRPL19/GCFC2* locus, *TTRAP, THEM2, CNTNAP2, DYX1C1, CMIP*, and *ATP2C2* which had overall significant effects on brain functioning. Part 2 of the thesis focused on a previously understudied gene, *SEMA6D*, which had significant reading-related effects on white matter volume in the transverse temporal gyrus. Lastly, part 3 of the thesis used a family case study to determine which reading-related genes had genotypes that co-varied with SRD in the family, with promising effects for *ZNF385D, CNTNAP2, NOP9, MYO18B*, and *RBFOX2*. These relevant genes identified across the three parts of the thesis are located on

chromosomes 2 (*MRPL19/GCFC2* 2p12), 3 (*ROBO1* 3p12.3, *ZNF385D* 3p24.3), 6 (*DCDC2* 6p22.3, *KIAA0319* 6p22.3, *TTRAP* 6p22.3, *THEM2* 6p22.3), 7 (*FOXP2* 7q31.1, *CNTNAP2* 7q35-7q36.1), 12 (*SLC2A3* 12p13.31), 14 (*NOP9* 14q12), 15 (*SEMA6D* 15q21.1, *DYX1C1* 15q21.3), 16 (*CMIP* 16q23.2-q23.3, *ATP2C2* 16q24.1), and 22 (*RBFOX2* 22q12.3, *MYO18B* 22q12.1).

The genes identified in the current thesis are likely to influence reading skills through their functions in brain development. For example, *ROBO1* and *SEMA6D* are both involved in axon guidance (Stelzer et al., 2016), influencing the development of synapses and location of neurons in the brain. Similarly, *DCDC2*, *KIAA0319*, and *DYX1C1* are also specifically related to neuronal migration during development (Stelzer et al., 2016). Similarly, *THEM2* is related to cell proliferation during development (Stelzer et al., 2016). *DCDC2, KIAA0319,* and *ROBO1* also code for proteins that are also all part of the same interacting network (Szklarczyk et al., 2019). *FOXP2*, *NOP9,* and the *MRPL19/GCFC2* locus are all regulators of transcription, while *CNTNAP2* codes for a protein which helps to mediate interactions between neurons and glia during development, and impacts formation of neuronal domains (Stelzer et al., 2016). Additionally, *FOXP2* and *CNTNAP2* are part of the same gene network, interacting in the same pathway to influence brain development. *CMIP,* which is involved in T-cell signaling (Stelzer et al., 2016), codes for a protein which also interacts in networks with proteins coded for by *KIAA0319, DCDC2, CNTNAP2,* and *FOXP2* (Szklarczyk et al., 2020). Both *ZNF385D* and *MYO18B* are involved in nucleic acid or nucleotide binding, while *SLC2A3* codes for a transmembrane transporter (Stelzer et al., 2016). *RBFOX2* is a regulator of alternative splicing (Stelzer et al., 2016). Many of the relevant genes identified through both the meta-analysis, as well as the exploratory descriptive analysis in the family case study, are part of the same gene

networks, with functions influencing various biological pathways that are particularly important during brain development. Lancaster and colleagues (2020) similarly identified multiple important biological pathways, including dendrite development and regulation, brain development, and particularly neuron migration, that likely influence the development of reading skills through effects on brain structure and function.

Brain Structures and Reading

The reading network has already been well-characterized by previous imaging studies and meta-analyses. The reading network includes a ventral occipitotemporal region (ventral stream) associated with automatic word recognition, a temporoparietal region (dorsal stream) associated with phonological processing, and an inferior frontal region associated with verbal working memory and articulation (Dehaene, 2009; Richlan, 2020). Meta-analyses have revealed structural or functional abnormalities within this network in children with SRD, including underactivation in a left ventral occipital-temporal region (Richlan et al., 2011), reduced grey matter volume in the right superior temporal gyrus and left superior temporal sulcus (Richlan et al., 2013), and lower fractional anisotropy in left temporoparietal and frontal regions (Vandermosten et al., 2012). The results by Richlan and colleagues (2013) suggest that, while the reading network is primarily in the left hemisphere, regions in the right hemisphere might also be atypical in individuals with SRD. Right hemispheric regions may be atypical due to compensation for left hemispheric deficits as well (Yu, Zuk, & Gaab, 2018). Part 1 of the current thesis, summarizing the current imaging genetics literature, suggested that these reading network regions are related to both reading-related genes and reading performance, and also revealed other associated regions in the cerebellum (Jasińska et al., 2016; Jednoróg et al., 2015; Landi et

al., 2013; Skeide et al., 2016; Ulfarsson et al., 2017) and corpus callosum (Einarsdottir et al., 2015; Darki et al., 2012; Darki et al., 2014; Lamminmäki et al., 2012; Marino et al., 2014; Scerri et al., 2012; Sun et al., 2017). Therefore, reading is not only affected by the regions characterized as part of the left-hemispheric reading network, but also by other subcortical regions or white matter structures that further interact with the left hemispheric reading network. Furthermore, part 1 of the current thesis built on the previous imaging results and integrated findings using imaging genetics, confirming and replicating findings regarding the reading brain network and connecting this network to genetic variables as well.

Part 2 of the current thesis built upon these findings through investigation of which of these reading-related regions may be associated with reading in a sample of young typical readers, in addition to examining the role of *SEMA6D*. The region that was identified as most associated with reading in the sample was white matter volume in the transverse temporal gyrus, or Heschl's gyrus, related to phonological processing, likely due to the young age of the participants and still developing reading skills. These results are consistent with the idea that different regions may be more involved at different stages of reading development, depending on the process being used for reading (e.g., phonological processing vs. orthographic processing; (Zhou et al., 2021). Therefore, atypical brain structure or function in these phonological processing related regions, such as Heschl's gyrus, may be an important predictor of later reading ability.

Consistently, part 3 of the current thesis, suggested that cortical thickness in the transverse temporal gyrus helped to differentiate the children with SRD with a TD sibling, despite the expectation that the twins would be more similar in their brain structures. Grey matter

volume and cortical surface area in the supramarginal and fusiform gyri also co-varied with SRD in the family case study. These results help to demonstrate the clinical application of the research literature, because knowledge of the brain networks and genes that have been shown to be related to reading informed the target regions investigated and identified as relevant within the single family. Within the overall network, specific regions were identified that may be critical parts of the reading network associated with reading skill within this particular family, as they were regions that differed between the twins but were more similar in the siblings with SRD. As reading skills develop in families, in response to intervention, relevant regions for a family may be monitored over time with intervention to understand the brain's response to reading development.

Furthermore, in part 3 of the current thesis, when brain phenotypes were generally compared in similarity between the siblings using intraclass correlations, rates of similarity between each pair were similar when comparing cortical thickness in regions of interest across the whole brain. However, when ICCs were calculated across average cortical thicknesses for only reading-related regions, the siblings with SRD were more similar (ICC = 0.76) as compared to the twins (ICC = 0.66) and the TD twin and older sibling (ICC = 0.54). ICCs for volume and cortical surface area tended to be similar across each pair. This indicates that cortical thickness may be a particularly important variable for understanding SRD risk within this family.

The results of all three parts of the current thesis implicate both white matter and grey matter structures within the reading network. It is likely that reading-related genes, which affect processes such as neuronal migration and axon guidance, have impacts on the development of both white matter and grey matter structures in the brain, and even specific types of phenotypes

like cortical thickness. Future research may benefit from examining the effects of other types of brain phenotypes, such as functional connectivity or tractography, to better understand how these interacting structures and pathways are influenced by reading-related genes and in turn affect reading skills.

Imaging Genetics

Part 1 of the current thesis, summarizing the current imaging genetics of SRD field, demonstrated that many reading-related genes are associated with brain structures within the reading network. For example, *DCDC2*, the most replicated gene in the imaging genetics literature, has been associated with grey matter volume in frontal, parietal, and temporal regions, fractional anisotropy in the left arcuate fasciculus and corpus callosum, white matter volume in the left temporo-parietal region, superior longitudinal fasciculus, and corpus callosum, cortical thickness in the left supramarginal gyrus, angular gyrus, and lateral occipital cortex, and functional activation in reading-related regions (Cope et al., 2012; Czamara et al., 2011; Darki et al., 2012; Darki et al., 2014; Eicher et al., 2016; Männel et al., 2015; Marino et al., 2014; Meda et al., 2008; Müller et al., 2017; Neef et al., 2017; Skeide et al., 2016; Su et al., 2015). *KIAA0319* has also been shown to have effects on structures across frontal, temporal, and parietal regions of the brain (Centanni et al., 2018; Darki et al., 2012; Darki et al., 2014; Eicher et al., 2016; Männel et al., 2015; Neef et al., 2017), with some mixed effects depending on specific SNPs and imaging phenotypes examined (Cope et al., 2012; Czamara et al., 2011; Müller et al., 2017; Skeide et al., 2016). Other genes have more specific known relationships to brain structures in single regions, such as *ROBO1,* which has been associated with white matter structure and activation related to the corpus callosum (Lamminmäki et al., 2012; Sun et al.,

2017), or *FOXP2*, which has been related to structure and function in the frontal lobe (Pinel et al., 2012; Skeide et al., 2016). Other genes, such as *CNTNAP2*, were associated with subcortical structures, such as white matter volume in the cerebral peduncle and left inferior cerebellar peduncle (Skeide et al., 2016). Therefore, some genes may lead to broader effects across the brain, while some may have their effects more localized to specific brain regions. In addition, results sometimes varied depending on the specific SNPs examined, and often times results were not replicated with specific SNPs.

Part 2 of the current thesis, focused on *SEMA6D*, revealed associations with cortical thickness in the fusiform gyrus, gyrification in the supramarginal gyrus, and white matter volume in the transverse temporal gyrus, although only white matter volume of the transverse temporal gyrus was related to reading in the sample. Therefore, genes may have effects on various reading-related brain regions that have different effects on reading during development. While part 3 of the current thesis identified both genes and brain regions that co-varied with reading, due to the small sample size, statistical analyses were not able to be completed to determine genetic associations with brain structures. However, the examination of genes and structures related to SRD within the family, informed by the previous literature, helped to provide information about genes and structures that were more essential for impacting SRD risk within the specific family investigated. Future studies using larger samples could help to better clarify the types of relationship between the genes and brain regions that determine the presence of reading disability in twins or siblings who partially share a genetic background.

Limitations

Part 1 of the current thesis helped to identify some of the limitations of the literature as a whole. For example, many of the studies had small sample sizes, and involved multiple testing due to many brain regions and SNPs being examined. This multiple testing problem was common, despite many of the studies focusing on one or a few genes, because each gene has multiple SNPs. Therefore, many imaging genetics studies of SRD are limited by a lack of statistical power. Furthermore, because the literature is still developing, often times there are few studies replicating effects on a certain gene. In addition, there may be variability in the imaging phenotypes examined as well, making it difficult to replicate the specific types of results found by previous studies. Therefore, more replication and the use of larger samples is necessary to better understand which genes have the largest effects on brain structures and function, and the impact of imaging genetics on the development of reading skills.

Part 2 and part 3 of the current thesis have some of the same limitations as the field as a whole. Part 2 was limited by a small sample size, particularly because parts of the analyses had to be done with subsets of the whole sample due to varying scanner strengths from the two different data collection sites. Because of limitations in sample size and a specific hypothesis of interest, the focus was on one gene of interest, *SEMA6D*, which helped to reduce multiple comparisons. Furthermore, analyses were focused on three phenotypes, including cortical thickness, gyrification, and white matter volume, but there may be other more fine-grained measures that would provide additional information about *SEMA6D*'s effect on the brain. Part 3 of the current thesis was also limited due to the family case study format. While this format allowed for a detailed understanding of the genes and brain structures that co-varied with SRD in a particular family, allowing for a clinical application of the research literature, the results may

not be generalizable to other families with SRD. In addition, the family case study format did not provide enough power to do frequentist statistical analyses to determine if effects were statistically significant to understand the relationship between genes and structures. Therefore, this study may be considered a promising starting point to identify the critical brain structures and genes that may be of interest in further studies, as well as providing an understanding of how research findings may inform understanding of risk in an individual family case study. Associations between genes and brain structures, which were not examined in part 3, which would be an important target of future studies. Lastly, these studies are limited by their cross-sectional nature, making it more difficult to understand the different brain structures or genes that may have greater impacts at different ages or stages of reading development and how this may change over time.

Areas of Future Research

The current thesis provided support for the importance of better understanding the brain and genome etiology of SRD and revealed many promising avenues of future research. Many of the genes identified as relevant to risk of SRD, including *SEMA6D*, should be investigated further in larger samples. While most studies focused on SNPs, other genetic factors such as epigenetics, which are changes in DNA that can be affected by the environment and may be passed on to offspring, may be important in understanding SRD. For example, the study of epigenetics may be relevant for understanding the underlying biological changes that are associated with reading development and response to intervention. In addition, the study of gene-environment interactions, focused on some of these new gene targets, may provide additional knowledge about how genetic factors are related to environmental factors, such as

socioeconomic status or birth weight (Mascheretti et al., 2017). These environmental factors are important to understand in relation to genes and brain structures as they interact with biological risk factors in affecting risk of SRD.

In addition, part 1 of the current thesis revealed that there has not been very much replication of investigation of specific imaging phenotypes and relationships to specific genes. There are many additional genes and imaging phenotypes that may be associated, that have yet to be studied, suggesting that there are a plethora of opportunities for better understanding the relationship between various brain measures and genes. In addition, part 2 of the current thesis revealed a strong relationship between *SEMA6D* and white matter volume in the transverse temporal gyrus. These results suggest that using white matter tractography or measures such as fractional anisotropy may also be revealing about more specific ways that the brain is related to *SEMA6D.* The use of imaging phenotypes related to functional activation may provide further information about how active brain processing may be affected by *SEMA6D*, and whether this is further associated with reading skills. In addition, *SEMA6D* may have impacts on subcortical structures as well, as it has been shown to be associated with the corpus callosum and basal ganglia (Alto & Terman, 2018). Similarly, there are many newly identified genes related to SRD through genome wide association studies that likely impact brain structure and function as well, opening up many potential avenues of future investigation through the use of different imaging methodologies and examination of various brain regions.

Exploratory research, particularly as new methodologies are developed, is also important for improving understanding of the etiology of SRD. Using genome-wide association studies with reading-related imaging measures as phenotypes may help to identify additional genes that

impact SRD through their effects on the brain. Furthermore, genome-wide association studies may be used to investigate the SNPs that are related to specific reading skills, such as phonological processing or rapid automatic naming. These association studies may help to identify additional reading-related genes, which can then be studied in more detail in relation to a variety of imaging phenotypes.

Part 3 of the current thesis focused on the clinical application of research findings to understand SRD risk within a specific family. These clinical applications have been largely understudied in the current literature, but new developing methodologies may lead to significant improvements in ability to understand risk in the individual. For example, machine learning methods may combine knowledge about genes and imaging to better predict and understand risk given an individual's specific biological factors. While this avenue of research is still developing, the use of multimodal imaging techniques incorporating multiple types of imaging measures, as well as genetic factors, can be combined with the use of machine learning to predict or classify neurodevelopmental or psychiatric disorders (Tulay et al., 2018). Using a polygenic approach, taking into account multiple genetic risk factors to predict reading disability, may also provide additional power for SRD classification. Using this type of approach may also help to differentiate unique versus overlapping risk factors that affect multiple neurodevelopmental disorders (Gialluisi et al., 2020). While these approaches have been largely focused on a group level so far, these approaches may be developed further to have clinical utility for individuals or families in the future. Discordant twin families are a useful subpopulation to help apply these results when disentangling genetic and neural factors. Furthermore, using longitudinal study

designs may improve the ability to understand how biological factors predict reading development over time.

Conclusions

The current thesis demonstrated the promising developments in the field of imaging genetics studies of SRD. Part 1 of the current thesis summarized the current literature and emphasized that reading-related genes have been shown to have highly significant relationships with brain structure and function, and provided support for the relationship between specific genes (e.g. *DCDC2*) and reading. In part 2 of the current thesis, *SEMA6D*, previously not studied in relation to SRD, was shown to be related to structural phenotypes within the reading network, as well as reading performance. Furthermore, reading-related genes and brain structures were examined within a clinical family case study to determine how the current field of imaging genetics literature could inform understanding of the biological risk factors that covary with SRD within a specific clinical family case study. Part 3 of the current thesis demonstrated that many of the brain regions and genes that have been shown to be highly related to SRD also covaried with SRD within this specific family. Overall, the current thesis summarized and added to the rich imaging genetics of SRD literature, and identified many avenues of future research, including emphasis on a novel gene, *SEMA6D*, and the translation of the research literature into clinical application. As knowledge of the biological etiology of SRD improves, the next important step will be using this knowledge to better understand risk and improve early identification and early intervention, to have a beneficial clinical impact on children with SRD.

BIBLIOGRAPHY

Aboud, K., Barquero, L., & Cutting, L. (2018). Prefrontal mediation of the reading network predicts intervention response in dyslexia. *Cortex, 101,* 96-106.

Adams, A. K., Smith, S. D., Truong, D. T., Willcutt, E. G., Olson, R. K., DeFries, J. C., … & Gruen, J. R. (2017). Enrichment of putatively damaging rare variants in the DYX2 locus and the reading-related genes CCDC136 and FLNC. *Human Genetics, 136*(11-12), 1295-1405.

Alto, L. T. & Terman, J. R. (2018). Semaphorins and their signaling mechanisms. *Methods in Molecular Biology, 1493,* 1-25.

Andreola, C., Mascheretti, S., Belotti, R., Ogliari, A., Marino, C., Battaglia, M., & Scaini, S. (2021). The heritability of reading and reading-related neurocognitive components: A multi-level meta-analysis. *Neuroscience & Biobehavioral Reviews, 121,* 175-200.

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

Arion, D., Horvath, S., Lewis, D., & Mirnics, K. (2010). Infragranular gene expression disturbances in the prefrontal cortex in schizophrenia: Signature of altered neural development? *Neurobiology of Disease, 37*(3), 738-746.

Barquero, L., Davis, N., & Cutting, L. (2014). Neuroimaging of Reading Intervention: A Systematic Review and Activation Likelihood Estimate Meta-Analysis. *Plos ONE, 9*(1), e83668.

Bates, T. C., Lind, P. A., Luciano, M., Montgomery, G. W., Martin, N. G., & Wright, M. J. (2009). Dyslexia and *DYX1C1*: deficits in reading and spelling associated with a missense mutation. *Molecular Psychiatry, 15*, 1190-1196.

Becker, N., Vasconcelos, M, Oliveira, V., Dos Santos, F. C., Bizarro, L., De Almeida, R. M. M., … & Carvalho, M. R. S. (2017). Genetic and environmental risk factors for developmental reading disability in children: systematic review of the last decade. *Developmental Neuropsychology*, *42*(7-8), 423-445.

Belsky, J., & van Ijzendoorn, M. (2015). What works for whom? Genetic moderation of intervention efficacy. *Development Psychopathology, 27*(1), 1-6.

Belyk, M., Kraft, S. J., Brown, S., & Pediatric Imaging, Neurocognition and Genetics Study. (2017). PlexinA polymorphisms mediate the developmental trajectory of human corpus callosum microstructure. *Journal of Human Genetics, 60*(3), 147-150.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B, 57*, 289–300.

Bishop, D. (2013). Cerebral asymmetry and language development: Cause, correlate, or consequence? *Science, 340*(6138), 1230531-1230531.

Blackmon, K., Barr, W. B., Kuzniecky, R., DuBois, J., Carlson, C., Quinn, B. T., … & Thesen, T. (2010). Phonetically irregular word pronunciation and cortical thickness in the adult brain. *NeuroImage*, *51*(4), 1453-1458.

Breen, G., Li, Q., Roth, B., O'Donnell, P., Didriksen, M., Dolmetsch, R., … & Edenberg, H. J. (2016). Translating genome-wide association findings into new therapeutics for psychiatry. *Nature Neuroscience, 19*(11), 1392-1396.

Bruxel, E., Moreira-Maia, C., Akutagava-Martins, G., Quinn, T., Klein, M., Franke, B. … & Hutz, M. H. (2020). Meta-analysis and systematic review of ADGRL3 (LPHN3) polymorphisms in ADHD susceptibility. *Molecular Psychiatry*.

Cabana, J., Gilbert, G., Létourneau-Guillon, Safi, D., Rouleau, I., Cossette, P., & Nguyen, D. K. (2017). Effects of SYN1Q555X mutation on cortical gray matter microstructure. *Human Brain Mapping, 39*(8), 3428-3448.

Cancer, A., & Antonietti, A. (2018). tDCS Modulatory Effect on Reading Processes: A Review of Studies on Typical Readers and Individuals With Dyslexia. *Frontiers In Behavioral Neuroscience, 12*.

Cattinelli, I., Borghese, N. A., Gallucci, M., Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics, 26*(1), 214-238.

Centanni, T. M., Pantazis, D., Truong, D. T., Gruen, J. R., Gabrieli, J. D. E., & Hogan, T. P. (2018). Increased variability of stimulus-driven cortical responses is associated with genetic variability in children with and without dyslexia. *Developmental Cognitive Neuroscience, 34*, 7-17.

Chen, X. S., Reader, R. H., Hoischen, A., Veltman, J. A., Simpson, N. H., Francks, C., Newbury, D.F., & Fisher, S. (2017). Next-generation DNA sequencing identifies novel gene

variants and pathways involved in specific language impairment. *Scientific Reports, 7,* 46105.

Christodoulou, J. A., Murtagh, J., Cyr, A., Perrachione, T. K., Chang, P., Halverson, K., … & Gabrieli, J. D. E. (2017). Relation of white-matter microstructure to reading ability and disability in beginning readers. *Neuropsychology, 31*(5), 508-515.

Chyl, K., Kossowski, B., Dębska, A., Łuniewska, M., Banaszkiewicz, A., Żelechowska, A., … & Jednoróg, K. (2018). Prereader to beginning reader: changes induced by reading acquisition in print and speech brain networks. *Journal of Child Psychology and Psychiatry, 59*(1), 76-87.

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

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 3-37). Hillsdale, NJ: Erlbaum.

Cope, N., Eicher, J. D., Meng, H., Gibson, C. J., Hager, K., Lacadie, C., … & Gruen, J. R. (2012). Variants in the *DYX2* locus are associated with altered brain activation in reading-related brain regions in subjects with reading disability. *Neuroimage, 63,* 148-156.

Cope, N., Harold, D., Hill, G., Moskvina, V., Stevenson, J., Holmans, P., … & Williams, J. (2005). Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *American Journal of Human Genetics, 76*(4), 581-591.

Costanzo, F., Rossi, S., Varuzza, C., Varvara, P., Vicari, S., & Menghini, D. (2019). Long-lasting improvement following tDCS treatment combined with a training for reading in children and adolescents with dyslexia. *Neuropsychologia, 130,* 38-43.

Czamara, D., Bruder, J., Becker, J., Bartling, J., Hoffmann, P., Ludwig, K. U., … & Schulte-Körne, G. (2011). Association of a rare variant with mismatch negativity in a region between KIAA0319 and DCDC2 in reading disability. *Behavioral Genetics, 41*(1), 110-119.

D'Mello, A. M. & Gabrieli, J. D. E. (2018). Cognitive neuroscience of dyslexia. *Language, Speech, and Hearing Services in Schools, 49*(4), 798-809.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences, 9*(7), 335-341.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage, 9*(2), 179-194.

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

Darki, F., Peyrard-Janvid, M., Matsson, H., Kere, J., & Klingberg, T. (2014). *DCDC2* polymorphism is associated with left temporoparietal gray and white matter structures during development. *The Journal of Neuroscience, 34*(43), 14455-14462.

Dehaene, Stanislas (2009). *Reading in the Brain: The New Science of How We Read*. New York: Penguin.

Dejerine, J., & Dejerine-Klumpke, A. (1895). Anatomie des centres nerveux (Vol. 1). Rueff.

Demontis, D., Walters, R. K., Martin, J., Mattheisen, Mn., Als, T. D., Agerbo, E., … & Neale, B. J. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics, 51*(1), 63-75.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31,* 968-980.

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage, 53*(1), 1-15.

Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006). The role of the posterior fusiform gyrus in reading. *Journal of Cognitive Neuroscience, 18*(6), 911-922.

Dick, D. M. (2019). Commentary for special issue of prevention science: "Using genetics in prevention: Science fiction or science fact?" *Prevention Science, 19*(1), 101-108.

Eckert, M.A., Berninger, V.W., Vaden, K.L. Jr., Gebregziabher, M., & Tsu, L. (2016). Gray Matter Features of Reading Disability: A Combined Meta-Analytic and Direct Analysis Approach(1,2,3,4). *eNeuro, 3*(1), ENEURO.0103-15.2015.

Eckert, M.A., Vaden Jr., K. I., Maxwell, A. B., Cute, S. L., Gebregziabher, M., Berninger, V. W., & Dyslexia Data Consortium. (2017). Common brain structure findings across children with varied reading disability profiles. *Scientific Reports, 7,* 6009.

Egan, M.F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M. & Straub, R. E. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of the Sciences, 98,* 6917-6922.

Eicher, J. D. & Gruen, J. R. (2015). Language impairment and dyslexia genes influence language skills in children with autism spectrum disorders. *Autism Research, 8*(2), 229-234.

Eicher, J. D., Montgomery, A. M., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Libiger, O., … & Gruen, J. R. (2016). Reading disability and language impairment associated genetic markers influence cortical thickness and white matter in typically developing children. *Brain Imaging and Behavior, 10,* 272-282.

Eicher, J. D., Powers, N. R., Miller, L. L., Akshoomoff, N., Amaral, D. G., Bloss, C. S., … & Gruen, J. R. (2013). Genome-wide association study of shared components of reading disability and language impairment. *Genes, Brain, and Behavior, 12*(8), 792-801.

Einarsdottir, E., Svensson, I., Darki, F., Peyrard-Janvid, M., Lindvall, J. M., Ameur, A. (2015). Mutation in CEP63 co-segregating with developmental reading disability in a Swedish family. *Human Genetics, 134*(11-12), 1239-1248.

Elliott, L. T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K. L., Douaud, G., … & Smith, S. M. (2018). Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature, 562,* 210-216.

Ercan-Sencicek, A., Wright, N. R. D., Sanders, S. J., Oakman, N., Valdes, L., Bakkaloglu, B., … & Grigorenko, E. L. (2012). A balanced t(10;15) translocation in a male patient with developmental language disorder. *European Journal of Medical Genetics, 55*(2), 128-131.

Feil, R., & Fraga, M. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics, 13*(2), 97-109. doi: 10.1038/nrg3142

Field, L. L., Shumansky, K., Ryan, J., Truong, D., Swiergala, E., & Kaplan, B. J. (2013). Dense-map genome scan for dyslexia supports loci at 4q13, 16p12, 17q22; suggests novel locus at 7q36. *Genes, Brain and Behavior, 12*, 56-69.

Fiez, J. A. & Peterson, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 914-921.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2), 195-207.

Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., … & Walhovd, K. B. (2015). Development and aging of cortical thickness correspond to genetic organization patterns. *PNAS, 112*(50), 15462-15467.

Fletcher, J. M. (2009). Reading disability: The evolution of a scientific concept. *Journal of International Neuropsychology Society, 15*(4), 501–508.

Fletcher, J. M. & Grigorenko, E. L. (2017). Neuropsychology of learning disabilities: The past and future. *Journal of the International Neuropsychological Society, 23*, 930-940.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M.A. (2019). Learning Disabilities, Second Edition: From Identification to Intervention. New York, NY: The Guilford Press.

Flint, J., Timpson, N., & Munafó, M. (2014). Assessing the utility of intermediate phenotypes for genetic mapping of psychiatric disease. *Trends in Neurosciences, 37*(12), 733-741.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B.A., & Fletcher, J. M. (1996). Development lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3-17.

Franquinho, F., Nogueira-Rodrigues, J., Duarte, J. M., Esteves, S. S., Carter-Su, C., Monaco, A. P., … & Sousa, M. M. (2017). The dyslexia-susceptibility protein *KIAA0319* inhibits axon growth through *SMAD2* signaling. *Cerebral Cortex, 27*(3), 1732-1747.

Galaburda A. M., & Kemper T. L. (1979). Cytoarchitectonic abnormalities in developmental dyslexia: A case study. *Annals of Neurology, 6*, 94–100.

Galaburda, A. M. (1985). Developmental dyslexia: A review of biological interactions. *Ann Dyslexia, 35*, 19-33.

Galaburda, A. M. (1993). Neurology of developmental dyslexia. *Current Opinion in Neurobiology, 3*(2), 237-242.

Galaburda, A. M., LoTurco, J., Ramus, F., Fitch, R. H., & Rosen, G. D. (2006). From genes to behavior in developmental dyslexia. *Nature Neuroscience, 9*(10), 1213-1217.

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., & Geschwind, N. (1985). Developmental dyslexia: Four consecutive patients with cortical anomalies. *Annals of Neurology, 18*, 222-233.

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

Gialluisi, A., Andlauer, T. F. M., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., … & Schulte-Kome, G. (2019). Genome-wide association scan identifies new variants associated with a cognitive predictor of dyslexia. *Translational Psychiatry, 9*, 77.

Gialluisi, A., Andlauer, T. F. M., Mirza-Schreiber, N., Moll, K., Becker, J., Hoffmann, P., … & Schulte-Körne, G. (2020). Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Molecular Psychiatry.*

Gialluisi, A., Guadalupe, T., Francks, C., & Fisher, S. E. (2016). Neuroimaging genetic analyses of novel candidate genes associated with reading and language. *Brain & Language, 172*, 9-15.

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

Gilbert-Diamond, D. & Moore, J. H. (2011). Analysis of gene-gene interactions. *Current Protocols in Human Genetics, 70*(1), 1.14.1-1.14.12.

Glahn, D. C., Thompson, P. M., & Blangero, J. (2007). Neuroimaging phenotypes: strategies for finding genes influencing brain structure and function. *Human Brain Mapping, 28*(6), 488-501.

Grigorenko, E. L., Fuchs, L. S., Willcutt, E. G., Compton, D. L., Wagner, R. K., & Fletcher, J. M. (2020). Understanding, education, and supporting children with specific learning disabilities: 50 years of science and practice. *American Psychologist, 75*(1), 37-51.

Grigorenko, E.L. (2005). A conservative meta-analysis of linkage and linkage-association studies of developmental reading disability. *Scientific Studies of Reading, 9*, 285–316.

Grigorenko, E., Kornilov, S., & Naumova, O. (2016). Epigenetic regulation of cognition: A circumscribed review of the field. *Development And Psychopathology, 28*(4pt2), 1285-1304.

Gu, H., Hou, F., Liu, L., Luo, X., Nkomola, P. D., Xie, X., … & Song, R. (2018). Genetic variants in the *CNTNAP2* gene are associated with gender differences among dyslexic children in China. *EBioMedicine, 34*, 165-170.

Guidi, L. G., Velayos-Baeza, A., Martinez-Garay, I., Monaco, A. P., Paracchini, S., Bishop, D. V. M., & Molnár, Z. (2018). The neuronal migration hypothesis of dyslexia: A critical evaluation 30 years ago. *European Journal of Neuroscience, 48*(10), 3212-3233.

Guttorm, T. K., Leppänen, T., Hämäläinen, J., Eklund, K. M., & Lyytinen, H. J. (2009). Newborn event-related potentials predict poorer pre-reading skills in children at risk for dyslexia. *Journal of Learning Disabilities, 43*(5), 391-401.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., … & Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage, 32*, 180-194.

Hancock, R., Richlan, F., & Hoeft, F. (2017). Possible roles for fronto-striatal circuits in reading disorder. *Neuroscience & Biobehavioral Reviews, 72*, 243-260.

Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kääriäinen, H., & Kere, J. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLOS Genetics, 1*(4), e50.

Hawi, Z., Yates, H., Pinar, A., Arnatkeviciute, A., Johnson, B., Tong, J., … & Bellgrove, M. A. (2018). A case-control genome-wide association study of ADHD discovered a novel association with the tenascin R (TNR) gene. *Translational Psychiatry, 8*(1), 284.

Heard, N. A. & Rubin-Delanchy, P. (2018). Choosing between methods of combining *p*-values. *Biometrika, 105*(1), 239-246.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., … & Gabrieli, J. D. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *PNAS, 104*(10), 4234-4239.

Hofer, E., Roshchupkin, G. V., Adams, H. H. H., Knol, M. J., Lin, H., Li, S., … & Seshadri, S. (2020). Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. *Nature Communications, 11*(1).

Hulme, C. & Snowling, M. J. (2016). Reading disorders and dyslexia. *Current Opinion in Pediatrics, 28*(6), 731-735.

Humphreys P., Kaufmann W. E., & Galaburda A. M. (1990). Developmental dyslexia in women: Neuropathological findings in three patients. *Annals of Neurology, 28,* 727–738.

Ito, D., Nojima, S., & Kumanogoh, A. (2014). The role of semaphorin family in immune systems. *Nihon Rinsho Meneki Gakkai Kaishi, 37*(1), 1-10.

Jasińska, K. K., Molfese, P. J., Kornilov, S. A., Einar Mencl, W., Frost, S. J., Lee, M., … & Landi, N. (2016). The BDNF Val66Met polymorphism influences reading ability and patterns of neural activation in children. *PLoS One, 11*(8), e0157449.

Jednoróg, K., Marchewka, A., Altarelli, I., Lopez, A. K. M., van Ermingen-Marbach, M., Grande, M., … & Ramus, F. (2015). How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? Insights from a large-scale voxel-based morphometry study. *Human Brain Mapping, 36*(5), 1741-1754.

Kaufman, A. S., Kaufman, N. L. (2014). *Kaufman Test of Educational Achievement–Third Edition (KTEA-3)*. Bloomington, MN: Pearson.

Khoury, M., Gwinn, M., Yoon, P., Dowling, N., Moore, C., & Bradley, L. (2007). The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genetics In Medicine, 9*(10), 665-674.

Kibby, M. Y., Fancher, J. B., Markanen, R., & Hynd, G. W. (2008). A quantitative magnetic resonance imaging analysis of the cerebellar deficit hypothesis of dyslexia. *Journal of Child Neurology, 23*(4), 368-380.

Klein, M. Walters, R. K., Demontis, D., Stein, J. L., Hibar, D. P., Adams, H. H., … & Franke, B. (2019). Genetic markers of ADHD-related variations in intracranial volume. *American Journal of Psychiatry, 176*(3), 228-238.

Kovas, Y. & Plomin, R. (2006). Generalist genes: implications for the cognitive sciences. *Trends in Cognitive Sciences, 10*(5), 198-203.

Krafnick, A. J., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *NeuroImage, 57*(3) 733-741.

Kronbichler, M., Wimmer, H., Staffen, W., Hutzler, F., Mair, A., & Ladurner, G. (2007). Developmental dyslexia: Gray matter abnormalities in the occipitotemporal cortex. *Human Brain Mapping, 29*(5), 613-625.

Lamminmäki, S., Massinen, S., Nopola-Hemmi, J., Kere, J., Hari, R. (2012). Human ROBO1 regulates interaural interaction in auditory pathways. *Journal of Neuroscience, 32*(3), 966-971.

Lancaster, H. S., Liu, X., Dinu, V., & Li, J. (2020). Identifying interactive biological pathways associated with reading disability. *Brain and Behavior, 10*(8), e01735.

Landi, N. & Perdue, M. V. (2019). Neuroimaging genetics studies of specific reading disability and developmental language disorder: A review. *Language and Linguistics Compass, 13*(9), e12349.

Landi, N., Frost, S. J., Mencl, W. E., Preston, J. L., Jacobsen, L. K., Lee, M., … & Grigorenko, E. L. (2014). The COMT Val/Met polymorphism is associated with reading related skills and consistent patterns of functional neural activation. *Developmental Science, 16*(1), 13-23.

Lebel, C., Benischek, A., Geeraert, B., Holahan, J., Shaywitz, S., Bakhshi, K., Shaywitz, B. (2019). Developmental trajectories of white matter structure in children with and without reading impairments. *Developmental Cognitive Neuroscience, 36,* 100633.

Leonard, C. M. (2001). Imaging brain structure in children: Differentiating language disability and reading disability. *Learning Disability Quarterly, 24*(3), 1158-176.

Leonard, C. M. & Eckert, M. A. (2009). Asymmetry and dyslexia. *Developmental Neuropsychology, 33*(6), 663-681.

Leonard, C. M., Lombardino, L. J., Walsh, K., Eckert, M. A., Mockler, J. L., Rowe, L. A., … & DeBose, C. B. (2002). Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children. *Journal of Communication Disorders, 35*(6), 501-531.

Leonard, C., Eckert, M., Given, B., Berninger, V., & Eden, G. (2006). Individual differences in anatomy predict reading and oral language impairments in children. *Brain, 129*(12), 3329-3342.

Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *PNAS, 115*(42), E9981-E9990.

Leslie, J. R., Imai, F., Fukuhara, K., Takegahara, N., Rizvi, T. A., Friedel, R. H., … & Yoshida, Y. (2011). Ectopic myelinating oligodendrocytes in the dorsal spinal cord as a consequence of altered Semaphorin 6D signaling inhibit synapse formation. *Development, 138*(18), 4085-4095.

Luciano, M., Lind, P. A., Duffy, D. L., Castles, A., Wright, M. J., Montgomery, G. W., … & Bates, T. C. (2007). A haplotype spanning *KIAA0319* and *TTRAP* is associated with normal variation in reading and spelling ability. *Biological Psychiatry, 62*(7), 811-817.

Ludwig, K. U., Sämann, P., Alexander, M., Becker, J., Bruder, J., Moll, K., … & Czamara, D. (2013). A common variant in *Myosin-18B* contributes to mathematical abilities in children with dyslexia and intraparietal sulcus variability in adults. *Translational Psychiatry, 3,* e229.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, b. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*(1), 1-14.

Ma, Y., Koyama, M. S., Milham, M. P., Castellanos, F. X., Quinn, B. T., Pardoe, H., … & Blackmon, K. (2015). Cortical thickness abnormalities associated with dyslexia, independent of remediation status. *NeuroImage: Clinical, 7*, 177-186.

Männel, C., Meyer, L., Wilcke, A., Boltze, J., Kirsten, H., & Friederici, A. D. (2015). Working-memory endophenotype and reading disability-associated genetic variant predict reading disability phenotype. *Cortex, 71,* 291-305.

Marino, C., Scifo, P., Della Rosa, P. A., Mascheretti, S., Facoetti, A., Lorusso, M. L., … & Perani, D. (2014). The DCDC2/intron 2 deletion and white matter disorganization: focus on developmental reading disability. *Cortex, 57,* 227-243.

Mascheretti, S., Bureau, A., Trezzi, V., Giorda, R., & Marino, C. (2015). An assessment of gene-by-gene interactions as a tool to unfold missing heritability in reading disability. *Human Genetics, 134*(7), 749-760.

Mascheretti, S., De Luca, A., Trezzi, V., Peruzzo, D., Nordio, A., Marino, C., & Arrigoni, F. (2017). Neurogenetics of developmental dyslexia: from genes to behavior through brain neuroimaging and cognitive and sensorial mechanisms. *Translational Psychiatry, 7*(1), e987.

Mascheretti, S., Perdue, M., Feng, B., Andreola, C., Dionne, G., & Jasińska, K. et al. (2021). From *BDNF* to reading: Neural activation and phonological processing as multiple mediators. *Behavioural Brain Research, 396,* 112859.

McCarthy, C. S., Ramprashad, A., Thompson, C., Botti, J., Coman, I. L., & Kates, W. R. (2015). A comparison of FreeSurfer-generated data with and without manual intervention. *Frontiers in Neuroscience, 9,* 379.

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

Meng, H., Smith, S. D., Hager, K., Held, M., Liu, J., Olson, R. K., … & Gruen, J. R. (2005). *DCDC2* is associated with reading disability and modulates neuronal development in the brain. *PNAS, 102*(47), 17053-17058.

Miciak, J. & Fletcher, J. M. (2020). The critical role of instructional response for identifying dyslexia and other learning disabilities. *Journal of Learning Disabilities, 53*(5), 343-353.

Miyake, K., Hirasawa, T., Koide, T., & Kubota, T. (2012). Epigenetics in Autism and Other Neurodevelopmental Diseases. *Advances In Experimental Medicine And Biology,* 91-98.

Mosca-Boidron, A., Gueneau, L., Huguet, G., Goldenberg, A., Henry, C., Gigot, N., … & Bourgeron, T. (2016). A *de novo* microdeletion of *SEMA5A* in a boy with autism spectrum disorder and intellectual disability. *European Journal of Human Genetics, 24*(6), 838-843.

Mueller, B., Ahnert, P., Burkhardt, J., Brauer, J., Czepezauer, I., Quente, E., … & Kirsten, H. (2014). Genetic risk variants for dyslexia on chromosome 18 in a German cohort. *Genes Brain & Behavior, 13*(3), 350-356.

Müller, B., Schaadt, G., Boltze, J., Emmrich, F., LEGASCREEN Consortium, Skeide, M. A., & Wilcke, A. (2017). ATP2C2 and DYX1C1 are putative modulators of reading disability-related MMR. *Brain and Behavior, 7*(11), e00851.

National Institute of Child Health & Human Development (2018). *Reading Difficulty and Disability.* Retrieved from https://report.nih.gov/nihfactsheets/ViewFactSheet.aspx?csid=114

National Library of Medicine (US). (2020, May, 26). *MYO5B gene*. Genetics Home Reference. Retrieved June 2, 2020, from https://ghr.nlm.nih.gov/gene/MYO5B

National Library of Medicine (US). (2020, May, 26). *NEDD4L gene*. Genetics Home Reference. Retrieved June 2, 2020, from https://ghr.nlm.nih.gov/gene/NEDD4L

Neef, N. E., Müller, B., Liebig, J., Schaadt, G., Grigutsch, M., Gunter, T. C., … & Friederici, A. D. (2017). Reading disability risk gene relates to representation of sound in the auditory brainstem. *Developmental Cognitive Neuroscience, 24,* 63-71.

Newbury, D. F., Paracchini, S., Scerri, T. S., Winchester, L., Addis, L., Richardson, A. J., … & Monaco, A. P. (2011). Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behavior Genetics, 41,* 90-104.

Newbury, D. F., Winchester, L., Addis, L., Paracchini, S., Buckingham, L., Clark, A., … & Monaco, A. P. (2009). *CMIP* and *ATP2C2* modulate phonological short-term memory in language impairment. *The American Jounral of Human Genetics, 85*(2), 264-272.

Nudel, R., Simpson, N. H., Baird, G., O'Hare, A., Conti-Ramsden, G., Bolton, P. F., … & Newbury, D. F. (2014). Genome-wide association analyses of child genotype effects and parent-of-origin effects in specific language impairment. *Genes Brain, and Behavior, 13*(4), 418-429.

Nugiel, T., Roe, M., Taylor, W., Cirino, P., Vaughn, S., Fletcher, J., … & Church, J. A. (2019). Brain activity in struggling readers before intervention relates to future reading gains. *Cortex, 111,* 286-302.

O'Connor, B., Eun, S., Ye, Z., Zozulya, A., Lich, J., & Moore, C. et al. (2008). Semaphorin 6D regulates the late phase of CD4+ T cell primary immune responses. *Proceedings Of The National Academy Of Sciences, 105*(35), 13015-13020.

Odegard, T., Ring, J., Smith, S., Biggan, J., & Black, J. (2008). Differentiating the neural response to intervention in children with developmental dyslexia. *Annals Of Dyslexia, 58*(1), 1-14.

Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. L., Pers, T. H., Rietveld, C. A., … & Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533*(7604), 539-542.

Olson, R.K., Keenan, J.M., Byrne, B., & Samuelsson, S. (2014). Why do children differ in their development of reading and related skills? *Scientific Studies of Reading, 18,* 38–54.

Paracchini, S., Diaz, R., & Stein, J. (2016). Advances in dyslexia genetics – New insights into the role of brain asymmetries. *Advances in Genetics, 96,* 53-97.

Paracchini, S., Steer, C. D., Buckingham, L., Morris, A. P., Ring, S., Scerri, T., … & Monaco, A. P. (2008). Association of the KIAA0319 dyslexia susceptibility gene with reading skills in the general population. *American Journal of Psychiatry, 165*(12), 1576-1584.

Pennington, B.F. & Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60,* 283-306.

Perdue, M. V., Mascheretti, S., Kornilov, S. A., Jasinska, K. K., Ryherd, K., Einar Mencl, W., … & Landi, N. (2019). Common variation within the SETBP1 gene is associated with reading-related skills and patterns of functional neural activation. *Neuropsychologia, 130,* 44-51.

Perdue, M. V., Mednick, J., Pugh, K. R., & Landi, N. (2020). Gray matter structure is associated with reading skill in typically developing young readers. *Cerebral Cortex, 30*(10), 5449-5459.

Peter, B., Raskind, W. H., Matsushita, M., Lisowski, M., Vu, T., Berninger, V. W., … & Brkanac, Z. (2011). Replication of *CNTNAP2* association with nonword repetition and support for *FOXP2* association with timed reading and motor activities in a dyslexia family sample. *Journal of Neurodevelopmental Disorders, 3*(1), 39-49.

Peterson, R. L. & Pennington, B. F. (2015). Developmental Reading disability. *Annual Review of Clinical Psychology, 11,* 283-307.

Pettigrew, K. A., Frinton, E., Nudel, R., Chan, M. T. M., Thompson, P., Hayiou-Thomas, M. E., … & Paracchini, S. (2016). Further evidence for a parent-of-origin effect at the *NOP9* locus on language-related phenotypes. *Journal of Neurodevelopmental Disorders, 8,* 24.

Phinney, E., Pennington, B. F., Olson, R., Filley, C. M., & Filipek, P. A. (2007). Brain structure correlates of component reading processes: Implications for reading disability. *Cortex, 43*(6), 777-791.

Piazza, R., Magistroni, V., Redaelli, S., Mauri, M., Massimino, L., Sessa, A., … & Gambacorti-Passerini, C. (2018). *SETBP1* induces transcription of a network of developmental genes by acting as an epigenetic hub. *Nature Communications, 9,* 2192.

Pinel, P., Fauchereau, F., Moreno, A., Barbot, A., Lathrop, M., Zelenika, D., … & Dehaene, S. (2012). Genetic variants of FOXP2 and KIAA0319/TTRAP/THEM2 locus are associated with altered brain activation in distinct language-related regions. *Journal of Neuroscience, 32*(3), 817-825.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage, 62*(2), 816-847.

Price, K. M., Wigg, K. G., Feng, Y., Blokland, K., Wilkinson, M., He, G., … & Barr, C. L. (2020). Genome-wide association study of word reading: Overlap with risk genes for neurodevelopmental disorders. *Genes, Brain and Behavior, 19*(6), e12648.

Raskind, W. H., Peter, B., Richards, T., Eckert, M. M., & Berninger, V. W. (2013). The genetics of reading disabilities: From phenotypes to candidate genes. *Frontiers in Psychology, 3*, 601.

Reiss, D. (2010). Introduction to the special issue: Genetics, personalized medicine, and behavioral intervention – Can this combination improve patient care? *Perspectives on Psychological Science, 5*(5), 499-501.

Rezaie, R., Simos, P., Fletcher, J., Cirino, P., Vaughn, S., & Papanicolaou, A. (2011). Temporo-parietal brain activity as a longitudinal predictor of response to educational interventions among middle school struggling readers. *Journal Of The International Neuropsychological Society, 17*(05), 875-885.

Richlan, F. (2020). The functional neuroanatomy of developmental dyslexia across languages and writing systems. *Frontiers in Psychology, 11*, 155.

Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage, 56,* 1735-1742.

Richlan, F., Kronbichler, M., & Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies. *Human Brain Mapping, 34*, 3055-3065.

Rodenas-Cuadrado, P., Ho, J., & Vernes, S. C. (2014). Shining a light on *CNTNAP2:* complex functions to complex disorders. *European Journal of Human Genetics, 22*, 171-178.

Roeske, D., Ludwig, K. U., Neuhoff, N., Becker, J., Bartling, J., Bruder, J., … & Schulte-Körne, G. (2011). First genome-wide association scan on neurophysiological endophenotypes points to trans-regulation effects on SLC2A3 in dyslexic children. *Molecular Psychiatry, 16*(1), 97-107.

Romeo, R., Christodoulou, J., Halverson, K., Murtagh, J., Cyr, A., Schimmel, C., … & Gabrieli, J. D. E. (2017). Socioeconomic Status and Reading Disability: Neuroanatomy and Plasticity in Response to Intervention. *Cerebral Cortex, 28*(7), 2297-2312.

Rose, E. J & Donohoe, G. (2013). Brain vs behavior: An effect size comparison of neuroimaging and cognitive studies of genetic risk for schizophrenia. *Schizophrenia Bulletin, 39*(3), 518-526.

Rubenstein, K.B., Raskind, W. H., Berninger, V. W., Matsushita, M. M., & Wijsman, E. M. (2014). Genome scan for cognitive trait loci of dyslexia: Rapid naming and rapid switching of letters, numbers, and colors. *Neuropsychiatric Genetics, Part B of the American Journal of Medical Genetics, 165B*(4), 345-356.

Rudov, A., Rocchi, M. B. L., Accorsi, A., Spada, G., Procopio, A. D., Olivieri, F., … & Albertini, M. C. (2013). Putative miRNAs for the diagnosis of dyslexia, dyspraxia, and specific language impairment. *Epigenetics, 8*(10), 1023-1029.

Rujescu, D., Meisenzahl, E. M., Krejcova, S., Giegling, I., Zetzsche, T., Reiser, M., … & Finckh, U. (2006). Plexin B3 is genetically associated with verbal performance and white matter volume in human brain. *Molecular Psychiatry, 12*, 190-194.

Rünker, A. E., O'Tuathaigh, C., Dunleavy, M., Morris, D. W., Little, G. E., Corvin, A. P., … & Mitchell, K. J. (2011). Mutation of *Semaphorin-6A* disrupts limbic and cortical

connectivity and models neurodevelopmental psychopathology. *PLoS One, 6*(11), e26488.

Salinas, R., Connolly, D., & Song, H. (2020). Invited Review: Epigenetics in neurodevelopment. *Neuropathology And Applied Neurobiology, 46*(1), 6-27. doi: 10.1111/nan.12608

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., … & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience, 19*(9), 1250-1255.

Scarpazza, C., Ha, M., Baecker, L., Garcia-Dias, R., Pinaya, W. H. L., Vieira, S., & Mechelli, A. (2020). Translating research findings into clinical practice: a systematic and critical review of neuroimaging-based clinical tools for brain disorders. *Translational Psychiatry, 10*, 107.

Scerri, T. S., Darki, F., Newbury, D. F., Whitehouse, A. J. O., Peyrard-Janvid, M., Matsson, H., … & Paracchini, S. (2012). The reading disability candidate locus on 2p12 is associated with general cognitive ability and white matter structure. *PLoS One, 7*(11), e50321.

Scerri, T. S., Morris, A. P., Buckingham, L. L., Newbury, D. F., Miller, L. L., Monaco, A. P., … & Paracchini, S. (2011). *DCDC2*, *KIAA0319* and *CMIP* are associated with reading-related traits. *Biological Psychiatry, 70*(3), 237-245.

Schaer, M., Cuadra, M. B., Schmansky, N., Fischl, B., Thiran, J., & Eliez, S. (2012). How to measure cortical folding from MR images: A step-by-step tutorial to compute local gyrification index. *Journal of Visualized Experiments, 59*, e3417.

Schlaggar, B., & McCandliss, B. (2007). Development of Neural Systems for Reading. *Annual Review Of Neuroscience, 30*(1), 475-503.

Schumacher, J., Hoffmann, P., Schmäl, C., Schulte-Körne, G., & Nöthen, M. M. (2007). Genetics of dyslexia: the evolving landscape. *Journal of Medical Genetics, 44*(5), 289-297.

Seidenberg, M.S. & McClelland J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Shao, S., Kong, R., Zou, L., Zhong, R., Lou, J., Zhou, J., … & Song, R. (2015). The Roles of Genes in the Neuronal Migration and Neurite Outgrowth Network in Developmental Dyslexia: Single- and Multiple-Risk Genetic Variants. *Molecular Neurobiology, 53*(6), 3967-3975.

Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3-4), 482-483.

Simos, P., Fletcher, J., Bergman, E., Breier, J., Foorman, B., Castillo, E. … & Papanicolaou, A. C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58*(8), 1203-1213.

Sinnwell, J. P., & Schaid, D. J. (2016). Haplo Stats (version 1.7.7) Statistical Methods for Haplotypes When Linkage Phase is Ambiguous.

Skeide, M. A., Kirsten, H., Kraft, I., Schaadt, G., Müller, B., Neef, N., … Friederici, A. D. (2015). Genetic reading disability risk variant is related to neural connectivity patterns underlying phonological awareness in children. *Neuroimage, 118,* 414-421.

Skeide, M. A., Kraft, I., Muller, B. Schaadt, G., Neef, N. E., Brauer, J., … & Friederici, A. D. (2016). *NRSN1* associated grey matter volume of the visual word form area reveals dyslexia before school. *Brain, 139,* 2792-2803.

Snowling, M. J. & Melby-Lervag, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin, 142*(5), 498-545.

Söderqvist, S., Matsson, H., Peyrard-Janvid, M., Kere, J., & Klingberg, T. (2014). Polymorphisms in the Dopamine Receptor 2 Gene Region Influence Improvements during Working Memory Training in Children and Adolescents. *Journal Of Cognitive Neuroscience, 26*(1), 54-62.

Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., … & Stefansson, K. (2014). CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature, 505*(7483), 361-366.

Stelzer, G., Rosen, R., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., … & Lancet, D. (2016). The GeneCards Suite: From gene data mining to disease genome sequence analysis. *Current Protocols in Bioinformatics, 54,* 1.30.1-1.30.33.

Stoodley, C.J., Hill, P.R., Stein, J.F., & Bishop, D.V. (2006). Auditory event-related potentials differ in dyslexics even when auditory psychophysical performance is normal. *Brain Research, 1121,* 190–199.

Su, M., Wang, J., Maurer, U., Zhang, Y., Li, J., McBride-Chang, C., … & Shu, H. (2015). Gene-environment interaction on neural mechanisms of orthographic processing in Chinese children. *Journal of Neurolinguistics, 33,* 172-186.

Suda, S., Iwata, K., Shimmura, C., Kameno, Y., Anitha, A., Thanseem, I., … & Mori, N. (2011). Decreased expression of axon-guidance receptors in the anterior cingulate cortex in autism. *Molecular Autism, 2*(1), 14.

Sun, X., Song, S., Liang, X., Xie, Y., Zhao, C., Zhang, Y., … & Gong, G. (2017). ROBO1 polymorphisms, callosal connectivity, and reading skills. *Human Brain Mapping, 38*(5), 2616-2626.

Szklarczyk, D., Gable, A., Lyon, D., Junge, A., Wyder, S., & Huerta-Cepas, J. et al. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research, 47*(D1), D607-D613.

Thakkar, V., Engelhart, A., Khodaparast, N., Abadzi, H., & Centanni, T. (2020). Transcutaneous auricular vagus nerve stimulation enhances learning of novel letter-sound relationships in adults. *Brain Stimulation, 13*(6), 1813-1820.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency (TOWRE).* Austin, TX: Pro-ed.

Torre, G. A., & Eden, G. F. (2019). Relationships between gray matter volume and reading ability in typically developing children, adolescents, and young adults. *Developmental Cognitive Neuroscience, 36,* 100636.

Truong, D. T., Adams, A. K., Paniagua, S., Frijters, J. C., Boada, R., Hill, D. E., … & Gruen, J. R. (2019). Multivariate genome-wide association study of rapid automatized naming and rapid alternating stimulus in Hispanic American and African-American youth. *Journal of Medical Genetics, 56,* 557-566.

Tulay, E., Metin, B., Tarhan, N., & Arıkan, M. (2018). Multimodal neuroimaging: Basic concepts and classification of neuropsychiatric diseases. *Clinical EEG And Neuroscience, 50*(1), 20-33.

Turkeltaub, P., Benson, J., Hamilton, R., Datta, A., Bikson, M., & Coslett, H. (2012). Left lateralizing transcranial direct current stimulation improves reading efficiency. *Brain Stimulation, 5*(3), 201-207.

Ulfarsson, M. O., Walters, G. B., Gustafsson, O., Steinberg, S., Silva, A., Doyle, O. M., … & Stefansson, K. (2017). 15q11.2 CNV affects cognitive, structural and functional correlates of reading disability and dyscalculia. *Translational Psychiatry, 7*(4), e1109.

Ullman, M. T., Earle, F. S., Walenski, M., & Janacsek, K. (2020). The neurocognition of developmental disorders of language. *Annual Review of Psychology, 71*, 389-417.

Vanderauwera, J., Altarelli, I., Vandermosten, M., De Vos, A., Wouters, J., & Ghesquière, P. (2018). Atypical structural asymmetry of the planum temporale is related to family history of dyslexia. *Cerebral Cortex, 28*(1), 63-72.

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.

Veerappa, A. M., Saldanha, M., Padakannaya, P., & Ramachandra, N. B. (2013). Family-based genome-wide copy scan number scan identifies five new genes of dyslexia involved in dendritic spinal plasticity. *Journal of Human Genetics, 58*, 539-547.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing.* Austin, TX: Pro-Ed.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing–2nd ed. (CTOPP-2).* Austin, TX: Pro-Ed.

Waldie, K. E., Wilson, A. J., Roberts, R. P., & Moreau, D. (2017). Reading network in dyslexia: similar, yet different. *Brain and Language, 174,* 29-41.

Wandell, B., & Le, R. (2017). Diagnosing the Neural Circuitry of Reading. *Neuron, 96*(2), 298-311.

Wang, G., Zhou, Y., Gao, Y., Chen, H., Xia, J., Xu, J., … & Sun, Y. (2015). Association of specific language impairment candidate genes *CMIP* and *ATP2C2* with developmental dyslexia in Chinese population. *Journal of Neurolinguistics, 33,* 163-171.

Welcome, S. E. & Joanisse, M. F. (2014). Individual differences in white matter anatomy predict dissociable components of reading skill in adults. *NeuroImage, 96,* 261-275.

Westwood, S., & Romani, C. (2017). Transcranial direct current stimulation (tDCS) modulation of picture naming and word reading: A meta-analysis of single session tDCS applied to healthy participants. *Neuropsychologia, 104*, 234-249.

Whalley, H. C., O'Connell, G., Sussmann, J. E., Peel, A., Stanfield, A. C., Hayiou-Thomas, M.E., … & Hall, J. (2011). Genetic variation in CNTNAP2 alters brain function during linguistic processing in healthy individuals. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 156B*(8), 941-948.

White, T., Su, S., Schmidt, M., Kao, C., & Shapiro, G. (2010). The development of gyrification in childhood and adolescence. *Brain and Cognition, 72*(1), 36-45.

Wilcke, A., Ligges, C., Burkhardt, J., Alexander, M., Wolf, C., Quente, E., … & Kirsten, H. (2011). Imaging genetics of *FOXP2* in dyslexia. *European Journal of Human Genetics, 20*, 224-229.

Williams, V. J., Juranek, J., Cirino, P., & Fletcher, J. M. (2018). Cortical thickness and local gyrification in children with developmental dyslexia. *Cerebral Cortex, 28*(3), 963-973.

Wilson, D. (2018). Practical Meta-Analysis Effect Size Calculator. Available from: http://cebcp.org/practical-meta-analysis-effect-size-calculator/standardized-mean-difference-d/unstandardized-regression-coefficient/.

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., … & Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage, 53,* 1135-1146.

Wise Younger, J., Tucker-Drob, E., Booth, J. R. (2017). Longitudinal changes in reading network connectivity related to skill improvement. *NeuroImage, 158,* 90-98.

Wong, P. C. M., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's gyrus and linguistic pitch learning. *Cerebral Cortex, 18*(4), 828-836.

Xia, Z., Zhang, L., Hoeft, F., Gu, B., Gong, G., & Shu, H. (2018). Neural correlates of oral word reading, silent reading comprehension, and cognitive subcomponents. *International Journal of Behavioral Development, 42*(3), 342-356.

Yu, X., Zuk, J., & Gaab, N. (2018). What factors facilitate resilience in developmental dyslexia? Examining protective and compensatory mechanisms across the neurodevelopmental trajectory. *Child Development Perspectives, 12*(4), 240-246.

Zhao, J., Thiebaut de Schotten, M., Altarelli, I., Dubois, J., & Ramus, F. (2016). Altered hemispheric lateralization of white matter pathways in developmental dyslexia: Evidence from spherical deconvolution tractography. *Cortex, 76,* 51-62.

APPENDICES

Part 1 Supplementary Material

Part 1 Supplementary Table 1. Summary of findings and effect sizes.

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
TNSFRSF1B 1p36.22	Männel 2015	rs496888: intron variant	Grey matter probability	32 adult males (16 with RD, 16 without)	Significant	Converted ß to r, and then converted to Cohen's *d*	rs496888 regression analyses: Right lateral occipital cortex (LOC) ß=.11, p=.51; Cohen's *d* = 0.32 Left Heschl's gyrus/posterior superior temporal gyrus (HG/pSTG) ß=-.28, p=.07; Cohen's *d* = -0.70 Left posterior superior temporal sulcus (pSTS) ß=.26, p=.07; Cohen's *d* = 0.65 **Difference HG/pSTG-pSTS ß=-.36, p=.03); Cohen's *d* = -0.90**
RCAN3 1p36.11	Skeide 2016	rs196402: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
MRPL19-C2ORF3/ GCFC2 2p12	Eicher 2016	rs2298948: intron variant	Cortical volume, cortical thickness, and fractional anisotropy	332 children in pediatrics imaging neurocognition genetics (PING) study ages 3-20	Some significant	No standard deviation of dependent variable reported, so unable to convert to Cohen's *d*	Multiple regression: <u>Cortical thickness:</u> **Right middle temporal region: Effect = 0.032, *p*=3.96x10^{-3}** <u>Cortical volume</u>: **Right inferior temporal: effect = -548.75, *p*=7.21x10^{-3}** Left inferior temporal: effect = -379.33, p=0.08 Right middle temporal: effect = 141.21, p=0.49 Left middle temporal: effect=-152.4, p=0.46 Right superior temporal: effect = -40.11, p=0.81 Left superior temporal: effect=-24.96, p=0.90 Right temporal pole: effect=7.01, p=0.86 Left temporal pole: effect=6.96, p=0.87 Right transverse temporal: effect=-2.47, p=0.92 Left transverse temporal: effect=-9.74, p=0.75

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Right fusiform: effect=-271.73, p=0.10 Left fusiform: effect=79.41,p=0.65 Right parahippocampal: effect=21.41, p=0.59 Left parahippocampal: effect=-13.49, p=0.70 Right lingual: effect=45.73, p=0.72 Left lingual: effect=-132.6, p=0.32
	Müller 2017	rs1000585: none	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs1000585: beta=0, p=0.98; Cohen's *d* = 0
	Scerri 2012	Swedish sample: rs3088180: *GCFC2* intron variant rs4853169: *GCFC2* intron variant rs917235: none rs6732511: none rs714939: none rs17689640: *MRPL19* intron variant rs17689863: *MRPL19*: synonymous variant, *GCFC2*: 500B Downstream Variant	White matter structure	Swedish sample: imaging genetics analysis: 76 children and young adults aged 6-25 years)	Some significant	Only p-values reported	**rs917235 w/ white matter volume (P corrected = $1.27x10^{-3}$ at cluster level with threshold of P<0.01)** G allele associated with lower bilateral white matter volume (confined to posterior part of corpus callosum and cingulum)
DYX3 locus (non-gene associated) 2p12	Eicher 2016	rs917235 (Chromosome 2) rs6732511 (Chromosome 2)	Cortical volume, cortical thickness, and fractional anisotropy	332 children in pediatrics imaging neurocognition genetics (PING) study ages 3-20	Some significant	Unstandardized regression coefficients reported, but no standard deviation of dependent variable, so Cohen's *d* unable to be calculated	rs917235 w/ thickness in left middle temporal: **Effect = 0.059, p=$3.96x10^{-3}$** rs6732511 w/ cortical volume: **Right fusiform: Effect = 478.22, p=$3.15x10^{-3}$** Right inferior temporal: effect = 112.11, p=0.63 Left inferior temporal: effect = 344.21, p=0.17 Right middle temporal: effect = 513.96, p=0.03 Left middle temporal: effect=341.3, p=0.16

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Right superior temporal: effect = 233.50, p=0.20 Left superior temporal: effect=39.05, p=0.86 Right temporal pole: effect=-100.36, p=0.03 Left temporal pole: effect=-99.27, p=0.04 Right transverse temporal: effect= 28.73, p=0.31 Left transverse temporal: effect=16.41, p=0.64 Left fusiform: effect=-70.36, p=0.73 Right parahippocampal: effect=-29.13, p=0.52 Left parahippocampal: effect=19.82, p=0.61 Right lingual: effect=70.41, p=0.62 Left lingual: effect=194.05, p=0.20 Non-significant not reported
ROBO1 3p12.3	Lamminmaki 2012	Haplotype situated between microsatellite markers D3S3039 and D3S3045 (includes two SNPs, one exonic insertion/deletion polymorphism, and four SNPs in 3' UTR	MEG	20 adults, 10 with SRD (19-51 years) from a Finnish family sharing haplotype, and 10 healthy, typical readers matched by age and gender (ages 18-49)	Significant	Calculated Cohen's *d* from correlation coefficients	Binaural suppression as a measure of axonal crossing: **Correlation b/w ROBO1 expression and left hemisphere ipsilateral suppression: r=0.75, p<0.02; Cohen's *d*=2.27** **Correlation b/w ROBO1 expression and right hemisphere ipsilateral suppression: r=0.78, p<0.01; Cohen's *d* = 2.49**
	Skeide 2016	rs162870: intron variant rs331142: intron variant rs12495133: intron variant rs11127636: intron variant rs4535189: intron variant rs7614913: intron variant rs6548628: intron variant rs9853895: intron variant rs1995402: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
	Sun 2017	rs4535189: intron variant rs6803202: intron variant	Fractional anisotropy (FA), axial diffusivity (AD), radial diffusivity	115 typically developing children aged 10-15 (average 8 = 12.87)	Some significant	Only p-values reported Path a in mediation analysis: Beta	Midline Corpus Callosum Fractional anisotropy: **T/T>C/T (p=0.006)** **T/T>C/C (p=0.007)** C/T and C/C did not differ (p=0.538)

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
			(RD), cortical thickness and surface area			converted to r, r converted to Cohen's *d*	Radial diffusivity: **C/T>T/T (p=0.002)** **C/C>T/T (p=0.011)** C/T and C/C did not differ (p=0.898) Axial diffusivity: **C/T>C/C (p=0.001)** **T/T>C/C (p=0.017)** C/T and T/T did not differ (p=0.513) From mediation analysis: <u>Path a (genotype → CC cluster):</u> **AD: Beta = 0.305, p=0.001, 95% CI: 0.121-0.489; Cohen's *d* = 0.76** FA: p=0.157 RD: p=0.26
CEP63 3q22.2	Einarsdottir 2015	Two-base mutation, causing a p.R229L amino acid substitution in CEP63 protein; chr3:134264558-9; variant in exon 7 of CEP63 Imaging genetics analysis: rs7619451	White matter volume	Three generation family, 7-92 years. 22 out of 62 family members with SRD. Imaging genetics analysis done with 76 healthy controls	Significant	Only p-values reported for imaging genetics analysis	rs7619451 (AA/AC genotypes) significantly associated with higher white matter volume in the right hemisphere, overlapping with right superior longitudinal fasciculus and posterior part of corpus callosum (peak coordinate: 28, −55, 29; p = 0.0076, corrected at the cluster level with p < 0.01). rs7619451 significantly associated with white matter volume in left (p = 0.006) and right (p = 0.003) regions of interest. Effect on reading comprehension: Main effect of rs7619451 [F(1, 88.98) = 5.938, p = 0.017] and interaction by age [F(1,91.62) = 5.518, p = 0.021] significant for reading comprehension.
CLSTN2 3q23	Roeske 2011	Genome-wide association scan Significant after correction: rs1365152: intron variant rs2114167: intron variant	EEG – mismatch negativity (MMN) reflecting automatic speech deviance processing	Initial sample of 200 dyslexic children aged 8-19 (mean age = 12.53) and a replication sample of 186 dyslexic children aged 8-18 (mean age = 11.40)	Significant	Only p values reported	rs1365152 w/ MMa (allelic): 1.44e-04, (Carrier A): 1.14e-03 rs2114167 w/ MMNa (allelic): 6.82e-04, (Carrier G): 5.41e-03
DCDC2 6p22	Cope 2012	rs793862: intron 7	Brain activation – fMRI	82 children 7-12 years (mean=8.8	Significant	Cohen's *d* calculated	Genetic associations with auditory categorization task:

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		rs807701: intron 7 rs807724: intron 6 BV677278 *STR*: Intron 2 BV677278 deletion Intron 2 Rs1087266: intron 1		years; 50 with RD, 25 without, 7 unknown)		from correlation	**rsSTR (allele 8) w/ RLOTG: -0.32, p=0.0002, q=0.05; Cohen's *d* = -0.67** rs1087266 w/ LMOTG: 0.033, p=0.008, q=0.83; Cohen's *d* = 0.06 Genetic associations with print categorization: rsSTR (allele 1) w/ LPC: -0.026, p=0.003, q=0.78; Cohen's d = -0.05 rs793862 w/ LAIPL: 0.027, p=0.007, q=0.995; Cohen's *d* = 0.05 **rsSTR (allele 4) w/ LAIPL: 0.08, p=0.00003, q=0.04; Cohen's *d* = 0.016** **rsSTR (allele 8) w/ RLOTG: -0.263, p=0.00008, q=0.06; Cohen's *d* = -0.55** rs793862 w/ LMOTG: 0.023, p=0.0074, q=0.98; Cohen's *d* = 0.05 rsSTR (allele 4) w/ LMOTG: 0.046, p=0.0076, q=0.93; Cohen's *d* = 0.09 Genetic associations with word rhyming: rsSTR (allele 5) w/ SAC: 0.096, p=0.004, q=0.74; Cohen's *d* = 0.19 Genetic associations with nonword rhyming: rs1087266 w/ LIFGI: 0.02, p=0.004, q=0.74; Cohen's *d* = 0.04 rsSTR (deletion) w/ LPC: 0.042, p=0.005, q=0.81; Cohen's *d* = 0.08 rsdel w/ LPC: 0.039, p=0.008, q=0.89; Cohen's *d* = 0.08 rsSTR (allele 8) w/ PC: -0.02, p=0.0016, q=0.59; Cohen's *d* = -0.04
	Czamara 2011	Chr6:24459391 Intergenic region b/w *DCDC2* and *KIAA0319*: Chr6:24564881 Chr6:24571041 Chr6:24581378	Brain activation – EEG	200 children with RD 8-19 years (mean=12.53)	Significant (Non-significant effect sizes not reported)	Only p-values reported	Chr6:24459391: Nominal p-value imputed: $3.69e^{-04}$; Nominal p-value genotyped: $6.28e^{-05}$ Chr6:24564881: Nominal p-value imputed: $6.63e^{-05}$; Nominal p-value genotyped: $1.40e^{-05}$ Chr6:24571041: Nominal p-value imputed: $4.55e^{-06}$; Nominal p-value genotyped: NA

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Chr6:24581378: Nominal p-value imputed: $9.81e^{-07}$; Nominal p-value genotyped: $1.40e^{-05}$
	Darki 2012	rs793842: intron variant rs793862: intron variant rs807701: intron variant rs2328819: intron variant rs2792682: intron variant rs7751169: intron variant rs9460974: intron variant	White matter volume and cortical thickness	76 children and young adults 6-25 years	Some significant	Cohen's *d* converted from Z scores	White matter volume in left temporo-parietal region: **rs793842 w/ MNI (-37, -49, 23): Pcorrected = 1.51×10^{-3}, Z = 4.23; Cohen's *d* = 1.11** Non-significant not reported
	Darki 2014	rs793842: intron variant	White matter volume, DTI (fractional anisotropy), cortical thickness	76 children and young adults in 9 age groups (6-25)	Some significant	Cohen's *d* calculated from Z and partial eta-squared	**WM (mainly superior longitudinal fasciculus and posterior corpus callosum: z=4.24, $p=8.19 \times 10^{-5}$; Cohen's *d* = 1.11** Cortical thickness: **L lateral cortical region: ($F(2,83.99)=9.39$, $p=2.09 \times 10^{-4}$, partial $eta^2=0.140$); Cohen's *d* = 0.81** **R hemisphere: p=0.037** **L SMG: $F(2,86.96)=5.05$, $p=2.68 \times 10^{-4}$, partial $eta^2=0.152$; Cohen's *d* = 0.84** **L AG: $F(2,88.78)=5.12$, $p=7.87 \times 10^{-3}$, partial $eta^2=0.112$; Cohen's *d* = 0.70** **L LOC: $F(2,84.21)=11.96$, $p=2.70 \times 10^{-5}$, partial $eta^2=0.165$; Cohen's *d* = 0.89** Significant interaction w/ age on thickness: Left SMG: $F(2,114,78)=7.61$, $p=7.88 \times 10^{-4}$ Left LOC: $F(2,110.38)=7.77$, $p=6.94 \times 10^{-4}$
	Eicher 2016	rs707864: intron variant	Cortical volume, cortical thickness, and fractional anisotropy	332 children in pediatrics imaging neurocognition genetics (PING) study ages 3-20	Not significant	Supplementary checked	Not reported
	Männel 2015	rs71745442: none	Grey matter probability	32 adult males (16 with	Not significant	Beta converted into r, which	rs71745442 regression analyses:

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
				dyslexia, 16 without)		was converted into Cohen's *d*	Right LOC ß=.17, p=.25; Cohen's *d* = 0.45 Left HG/pSTG ß= .16, p=.45; Cohen's *d* = 0.43 Left pSTS ß=.06, p=.69; Cohen's *d* = 0.22 Difference HG/pSTG-pSTS ß=.13, p=.47; Cohen's *d* = 0.37
	Marino 2014	DCDC2d – *DCDC2* deletion/intron 2 – encompasses 'regulatory element associated with dyslexia 1 (READ1)	White matter structure	47 children to young adults ages 16-21 (mean age = 18.2) 4 groups: Dyslexia with/without DCDC2d and normal readers with/without DCDC2d	Some significant	Calculated Cohen's *d* from Z scores	F-contrast for significant differences among 4 groups: significant differences in left temporal segment of the arcuate fasciculus and in the splenium of the corpus callosum Range of Cohen's *d* for FA differences: 1.06-2.89 Decreased FA DYS+ versus NR- <u>Left hemisphere regions:</u> Superior longitudinal fasciculus, arcuate fasciculus, inferior longitudinal fasciculus, optic radiation, corpus callosum, inferior cerebellar pedunculus Right hemisphere regions: Superior longitudinal fasciculus, corpus callosum <u>Right hemisphere regions:</u> Superior longitudinal fasciculus, corpus callosum Decreased FA DYS- versus NR- <u>Left hemisphere regions</u>: Superior longitudinal fasciculus, arcuate fasciculus, inferior longitudinal fasciculus, corpus callosum, cerebellar pedunculus <u>Right hemisphere regions</u>: superior longitudinal fasciculus, corpus callosum, cerebellar pedunculus Decreased FA in DYS+ versus DYS- <u>Left hemisphere regions:</u> Inferior longitudinal fasciculus <u>Right hemisphere regions:</u> inferior longitudinal fasciculus, corpus callosum

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Decreased FA in NR+ versus NR- Left hemisphere regions: corpus callosum Right hemisphere regions: corpus callosum Increased FA in NR+ versus NR- Left hemisphere regions: arcuate fasciculus, inferior occipito-frontal fasciculus, inferior longitudinal fasciculus Right hemisphere regions: Inferior fronto-occipital fasciculus, corpus callosum
	Meda 2008	*DCDC2* – intron 2 deletion	Grey matter volume	56 adults 20-85 years (mean age = 41.69) homozygous for no deletion (1/1) and 13 adults (19-82) heterozygous for the deletion (1/2)	Some significant	Cohen's *d* reported by authors.	**Positive correlation b/w suprathreshold volume (right and left combined) and expression levels of *DCDC2* (r=0.6, p=0.28); Cohen's *d* = 1.5** Cohen's *d* for difference in GM volume for subjects with genotype ½ compared to genotype 1/1 **Inferior temporal gyrus: L = 0.82, R = 0.83** **STG: L = 0.9, R =0.74** **Uncus: L = 0.81, R = 0.86** **MTG: L = 0.79, R = 0.88** **Postcentral gyrus: L = 0.63, R = 0.7** **Fusiform Gyrus L = 0.83, R = 0.96** **Lentiform nucleus L = 0.8, R = 0.7** **Caudate L = 0.73**, R not significant **Parahippocampal gyrus L = 0.8, R = 0.78** **Inferior occipital gyrus: L = 0.8, R = 0.72** **Inferior frontal gyrus: L = 0.75, R = 0.9** **Sub-gyral: L = 0.85, R = 0.77** **Inferior parietal lobule: L = 0.7, R = 0.86** **Middle frontal gyrus: L = 0.67, R = 0.82** **Supramarginal gyrus: L = 0.7, R = 0.66** **Middle occipital gyrus: L = 0.81, R = 0.72** **Insula L = 0.76,** R = not significant **Lingual gyrus L = 0.73, R = 0.7** **Precuneus L = 0.67, R = 0.76**

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Superior frontal gyrus: L = 0.74, R = 0.68 **Superior parietal lobule L = 0.66, R = 0.91** **Angular gyrus** L = not significant, **R = 0.81** **Cuneus** L = not significant, **R = 0.71**
	Müller 2017	rs1419228: intron variant rs7765678: intron variant rs793862: intron variant rs807701: intron variant	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs793862: beta=-0.4, p=0.59; Cohen's d = -0.2 rs7765678: beta=0.6, p=0.65; Cohen's d = 0.3 rs807701: beta=-0.1, p=0.83; Cohen's d = -0.05 rs1419228: beta=-0.2, p=0.86; Cohen's *d* = -0.1
	Neef 2017	rs807724: intron variant rs1087266: intron variant (*KAAG1*: 2KB upstream variant) rs807701: intron variant rs793842: intron variant rs1091047: intron variant rs6922023: intron variant	Brain activity – Speech evoked brain-stem responses (cABRs)	159 children (95 preliterate children aged 4-7 and 64 literate children aged 11-13)	Some significant	Assumed equal sample size for conversion from F; sample size not reported because used PCA-components as within-subjects factor (*DCDC2* and *KIAA0319*) Converted from r to Cohen's *d*	**Repeated measures ANCOVA: gene x brainstorm interaction: F(2153)=4.35, p=0.014; Cohen's *d* = 0.33** **Response consistency correlation with principle component allocating genetic risk of *DCDC2* (rs807724, rs1087266, rs80771, rs793842): r=0.144, p=0.075, Cohen's *d* = 0.28** Not significant: Response consistency correlation with principle component allocating genetic risk of *DCDC2* (rs1091047, rs6922023, rs793843, rs807701): r=-0.02, p=0.801; Cohen's *d* = -0.04
	Skeide 2016	rs793842: intron variant rs807701: intron variant rs807724: intron variant rs1091047: intron variant rs6922023: intron variant rs1087266: intron variant (*KAAG1*: 2KB Upstream Variant)	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
	Su 2015	rs1419228: intron variant	ERP (N170) during implicit	60 children, all 12 years	Not significant	Calculated Cohen's *d* from F values	Occipital-temporal N170 Peak amplitude:

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		rs1091047: intron variant	visual-word color decision task				Genotype effect in left hemisphere: F=3.185, p=0.080; Cohen's *d* = 0.48 Genotype effect in right hemisphere: F=0.711, p=0.403; Cohen's *d* = 0.23 Mean amplitude: Genotype effect in left hemisphere: F=1.436, p=0.236; Cohen's *d* = 0.32 Genotype effect in right hemisphere: F=1.698, p=0.198; Cohen's *d* = 0.35 **Significant interaction between rs1091047 and home literacy on changes of N170 in left hemisphere: F=7.336, p=0.009; Cohen's *d* = 0.74** Rs1091047 and right hemisphere: F=0.228, p=0.635; Cohen's *d* = 0.13 Mean amplitude: **Interaction in left hemisphere: F=6.891, p=0.011; Cohen's *d* = 0.71** **Interaction in right hemisphere: F=1.751, p=0.191; Cohen's *d* = 0.36**
KIAA0319 6p22	Centanni 2018	rs6935067 - none rs761100: intron variant	MEG	32 children 7-14 years (20 with dyslexia, 12 without)	Significant main effect for rs6935067, not for rs761100	Cohen's *d* calculated from F	**Significant main effect for rs6945067 (F (1,20)=5.8, p=0.03; Cohen's *d* = 0.88)** but not for rs761100 (F (2,19)=0.25, p=0.62; Cohen's *d* = 0.19). Linear regression for relationship between number of minor alleles at **rs6945067 and neural variability = F(2,20)=4.47, p=0.04; Cohen's *d* = 0.77** **Effect of minor allele in response to auditory stimuli in the inferior frontal gyrus (IFG) – F(2,19)=2.72, p=0.09; Cohen's *d* = 0.6022; and superior temporal gyrus (STG) – F(2,19)=2.94, p=0.08); Cohen's *d* = 0.63**

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Effect of minor allele in response to visual stimuli in IFG – not significant and **STG – F(2,20)=3.57, p=0.07; Cohen's *d* = 0.69.**
	Cope 2012	rs4504469: exon 4 rs730860: intron 1 rs3033227: intron 1 rs6935076: intron 1 rs2038136: intron 1 rs2038135: exon 1 rs3212236: upstream	Brain activation – fMRI	82 children 7-12 years (8.8 years average; 50 with RD, 25 without, 7 unknown)	Not significant	Not reported	Not reported
	Czamara 2011	Genotyped and imputed SNPs (not specified)	Brain activation – EEG	200 children with RD 8-19 years (mean=12.53)	Not significant	Not reported	Not reported
	Darki 2012	rs4504469: exon 4 rs6935076: intron 1 rs2143340: (TDP2: intron variant)	White matter volume and cortical thickness	76 children and young adults 6-25 years	Some significant	Calculated Cohen's *d* from Z scores	White matter volume in left temporo-parietal regions: **rs6935076 w/ MNI(-38, -69, 38): Pcorrected = 5.51 x 10^{-4}, Z=4.10; Cohen's *d* = 1.07** Non-significant not reported
	Darki 2014	rs6935076: intron 1	White matter volume, DTI (fractional anisotropy), cortical thickness	76 children and young adults in 9 age groups (6-25)	Some significant	Calculated Cohen's *d* from Z scores	Bilateral WM (superior longitudinal fasciculus and posterior corpus callosum): **(-34,-58,31): Z=5.32, p=3.33x10^{-10}; Cohen's *d* = 1.54** **(36,-28,37): Z=4.01, p=3.32x10^{-10}; Cohen's *d* = 1.04**
	Eicher 2016	rs9295626: intron variant rs10456309: intron variant rs4576240: missense variation rs9461045: none	Cortical volume, cortical thickness, and fractional anisotropy	332 children in pediatrics imaging neurocognition genetics (PING) study ages 3-20	Some significant	Unstandardized regression coefficients reported but no standard deviation of dependent variables reported, so Cohen's *d* unable to be calculated	<u>rs9461045 cortical thickness:</u> **left orbitofrontal: effect = -0.048, p=4.89x10^{-4}** right occipital: effect=-0.01, p=0.33 Left occipital: effect=-0.01, p=0.27 Right anteromedial temporal: effect=-0.02, p=0.17 Left anteromedial temporal: effect=-0.02, p=-.12 Right posterolateral temporal: effect=-0.00, p=0.97 Left posterolateral temporal: effect=-0.01, p=0.69

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Left superior parietal: effect=-0.02, p=0.31 Left superior temporal: effect=-0.03, p=0.07 Left inferior parietal: effect=-0.03, p=0.08 Left dorsomedial frontal: Effect=-0.02, p=0.15 Left precuneus: effect=-0.02, p=0.08 Left dorsolateral prefrontal: effect=-0.03, p=0.10 Left pars opercularis: effect=-0.03, p=0.03 Left central: effect=-0.01, p=0.56 <u>rs9461045 fractional anisotropy:</u> **corpus callosum: effect = -0.0084, p=5.89x10^{-3}** All: effect=09.99, p=0.32 Right All: effect=-0.00, p=0.02 Left All: effect=-0.00, p=0.04 Right ILF: effect=-0.00, p=0.66 Left ILF: effect=-0.00, p=0.50 Right IFO: effect=-0.00, p=0.72 Left IFO: Effect=-0.00, p=0.72 Right SLF: effect=-0.01, p=0.05 Left SLF: Effect=-0.01, p=0.03 Right tSLF: effect=-0.00, p=0.11 Left tSLF: effect=-0.01, p=0.02 Right pSLF: effect=-0.01, p=0.04 Left pSLF: Effect=-0.00, p=0.08 Right SIFC: effect=-0.00, p=0.46 Left SIFC: effect=0.00, p=0.77
	Männel 2015	rs6935076: intron variant	Grey matter probability	32 adult males (16 with dyslexia, 16 without)	Not significant	Converted beta to r, and then converted r to Cohen's *d*	rs6935076 regression analyses: Right LOC ß=-.16, p=.41; Cohen's *d* = -0.43 Left HG/pSTG ß= .02, p=.93; Cohen's *d* = 0.14 Left pSTS ß=-.08, p=.60; Cohen's *d* = -0.26 Difference HG/pSTG-pSTS ß=.07, p=.68; Cohen's *d* = 0.24
	Müller 2017	rs2143340: intron variant TDP2	EEG – mismatch response	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent	rs6935076: beta=-0.6, p=0.38; Cohen's *d* = -0.31

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		rs2179515: intron variant rs6935076: intron variant rs761100: intron variant	(MMR) in anterior region of interest (ROI) (F3, Fz, F4)			variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs2179515: beta=0.2, p=0.81; Cohen's *d* = 0.10 rs2143340: beta=0.1, p=0.91; Cohen's *d* = 0.05 rs761100: beta=0, p=0.99; Cohen's *d* = 0
	Neef 2017	rs761100: intron variant rs2179515: intron variant rs6935076: intron variant	Brain activity – Speech evoked brain-stem responses (cABRs)	159 children (95 preliterate children aged 4-7 and 64 literate children aged 11-13)	Significant	Assumed equal sample size for conversion from F; sample size not reported because used PCA-components as within-subjects factor (*DCDC2* and *KIAA0319*) Converted Cohen's *d* from r	**Repeated measures ANCOVA: gene x brainstorm interaction: F(2153)=4.35, p=0.014; Cohen's *d* = 0.33** **Response consistency correlation with principle component allocating genetic risk of KIAA0319: r=-0.190, p=0.018; Cohen's *d* = -0.39**
	Skeide 2016	rs2179515: intron variant rs761100: intron variant rs6935076: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
TTRAP (TDP2) 6p22	Pinel 2012	rs3756821: *KIAA0319* 2KB Upstream Variant rs9461045 rs707887: *TDP2* 3 prime UTR variant rs1047782: *TDP2* 3 prime UTR variant rs3087943: *TDP2* 3 prime UTR variant rs3181244: *TDP2* intron variant rs1129644: *TDP2* missense variant rs3212232: *TDP2* intron variant rs3212231: *TDP2* intron variant rs3756819: *TDP2* intron variant,	fMRI (reading 20 short sentences via visual and auditory stimulation)	94 adults (mean age = 24.7); typical readers	Not significant	Calculated Cohen's *d* from t and F-values	Not reported

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		ACOT13 2KB Upstream variant rs1061925: *TDP2* intron variant, *ACOT13* 2KB Upstream variant					
	Cope 2012	rs2143340: *TDP2* intron 2	Brain activation – fMRI	82 children 7-12 years (8.8 years average; 50 with RD, 25 without, 7 unknown)	Not significant	Calculated Cohen's *d* from correlation	**rs2143340 w/ RAIPL: -0.029, p=0.004, q=0.74; Cohen's *d* = -0.06** **rs2143340 w/ LAIPL: -0.024, p=0.003, q=0.85; Cohen's *d* = -0.05**
	Müller 2017	rs9461045 rs32122336	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs9461045: beta=0.8, p=0.33; Cohen's *d* = 0.41 rs32122336: beta=0.8, p=0.33; Cohen's *d* = 0.41
	Skeide 2016	rs3181238: *TDP2* intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
THEM2/ ACOT13 6p22	Eicher 2016	rs3777663: *ACOT13* intron variant	Cortical volume, cortical thickness, and fractional anisotropy	332 children in pediatrics imaging neurocognition genetics (PING) study ages 3-20	Some significant	Unstandardized regression coefficients reported, but no standard deviation of dependent variable reported, so Cohen's *d* unable to be calculated	<u>rs3777663 w/ cortical thickness:</u> **left pars opercularis: Effect = 0.037, p-value = $4.64x10^{-3}$** right occipital: effect=-0.00, p=0.72 Left occipital: effect=0.00, p=0.77 Right anteromedial temporal: effect=0.00, p=1.0 Left anteromedial temporal: Effect=0.01, p=0.71 Right posterolateral temporal: effect=0.00, p=0.85 Left posterolateral temporal: effect=0.02, p=0.13 Left superior parietal: effect=0.02, p=0.20 Left orbitofrontal: effect=-0.01, p=0.46

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Left superior temporal: effect=0.01, p=0.42 Left inferior parietal: effect=0.01, p=0.45 Left dorsomedial frontal: effect=0.00, p=0.77 Left precuneus: effect=0.00, p=0.74 Left dorsolateral prefrontal: effect=0.015, p=0.33 Left central: effect=0.02, p=0.28 Non-significant associations not reported
	Pinel 2012	rs17243157: *ACOT13* 5 Prime UTR Variant, *TDP2:* 2KB upstream variant rs3181227: *ACOT13* intron variant, *TDP2* 2KB upstream variant rs2223588: *ACOT13* intron variant rs6928074: *ACOT13* intron variant rs9461049: *ACOT13* intron variant rs926529: *ACOT13* intron variant rs1885211: *ACOT13* intron variant	fMRI (reading 20 short sentences via visual and auditory stimulation)	94 adults (mean age = 24.7); typical readers	Some significant	Calculated Cohen's *d* from t and F-values	**rs17243157 w/ temporal ROI: p=0.0009, p corrected (Bonferroni) = 0.0351** Association controlling for effects of other polymorphisms: **rs17243157 (CC>TC) w/ functional asymmetry in temporal ROI (T=4.43, voxel p=0.017 corrected for corresponding ROI); Cohen's *d* = 1.12** **rs17243157: association w/ asymmetry of pSTS activation during reading (F(2,90)=20.07, p=2.2x10^{-5}; Cohen's *d* = 1.13) w/ significant effect on left hemisphere (F(1,90)=10.59, p=0.002, CC>TC; Cohen's *d* = 0.82) during reading and a small effect during speech listening (F(1,90)=3.33, p=0.036, CC>TC; Cohen's *d* = 0.46)**
FAM65B (C6orf32, RIPOR2) 6p22	Eicher 2016	rs9348646: intron variant rs3756814: C6orf62: 3 Prime UTR variant	Cortical volume, cortical thickness, and fractional anisotropy	332 children in pediatrics imaging neurocognition genetics (PING) study ages 3-20	Some significant	Unstandardized regression coefficients reported, but no standard deviation of dependent variable reported, so Cohen's *d* unable to be calculated	rs9348646 w/ FA: **Right side all fiber tracts: Effect = -0.0039, p=9.20x10^{-3}** **Left SLF: effect = -0.00576, p=4.61x10^{-3}** **Right tSLF: effect = -0.0061, p=7.26x10^{-3}** left tSLF (temporal superior longitudinal fasciculus): effect = -0.0065, p=2.10x10^{-3} left pSLF (posterior superior longitudinal fasciculus): effect = -0.0053, p=1.00x10^{-2} All: effect=-0.00, p=0.02 Left All: effect=-0.00, p=0.04

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Right ILF: effect=-0.00, p=0.10 Left ILF: effect=-0.00, p=0.10 Right IFO: effect=-0.00, p=0.02 Left IFO: Effect=-0.00, p=0.47 Right SLF: effect=-0.00, p=0.05 Right pSLF: effect=-0.00, p=0.09 Right SIFC: effect=-0.00, p=0.127 Left SIFC: effect=-0.00, p=0.82 CC: effect=-0.01, p=0.02
NRSN1 6p22.3	Skeide 2016	rs9356928: intron variant rs4285310: intron variant rs3178: 3 Prime UTR variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Significant	Only p-values reported	Right pre- and postcentral gyri: (42, -18, 53): $p=9.92x10^{-5}$ Right lateral occipital cortex, superior division (35, -71, 41): $p=2.88x10^{-5}$ Right superior parietal lobule (29, -48, 57) $p=5.40x10^{-5}$ Left lateral occipital cortex: (-9, -83, 42): $p=5.28x10^{-6}$ Left temporal occipital fusiform cortex (-33, -63, -18): $p=3.19x10^{-4}$ Local white matter of the left postcentral cortex (-45, -23, 60): $p=1.90x10^{-5}$
Non-gene associated 7p12	Gialliusi 2016	LOC105375496	GM surface area and thickness	1,275 adults in Brain Imaging Genetics (BIG) dataset	Some significant	Only p-values reported	LOC105375496 gene-based analysis: Surface area: p=0.346 Cortical thickness: p=0.941
FOXP2 7q31.1	Müller 2017	rs12533005: intron variant	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs12533005: beta=-1, p=0.13; Cohen's *d* = -0.52
	Pinel 2012	rs6942634: intron variant rs2894699: intron variant rs1476535: intron variant rs10255943: intron variant rs10486026: intron variant	fMRI (reading 20 short sentences via visual and auditory stimulation)	94 adults (mean age = 24.7); typical readers	Some significant	Calculated Cohen's *d* from t and F-values	Association w/ inferior frontal ROI: rs10249234: p=0.0019, p corrected (Bonfer) = 0.0741 **rs6980093: p=0.0011, p corrected (Bonfer) = 0.0429**

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		rs10261780: intron variant rs10262103: intron variant rs4727799: intron variant rs17312686: intron variant rs2106900: intron variant rs17312861: intron variant rs10249234: intron variant rs12113612: intron variant rs10266297: intron variant rs10279936: intron variant rs6980093: intron variant rs7784315: intron variant rs7799109: intron variant rs12532920: intron variant rs17137124: intron variant rs10269986: intron variant rs7812028: intron variant rs17137135: intron variant rs1229761: intron variant rs1229758: intron variant rs12705966: intron variant rs10230087: intron variant rs7782412: intron variant rs1456029: intron variant rs12670585: intron variant rs6966051: intron variant rs17213159: intron variant rs1378771: intron variant rs12705971: intron variant rs12705973: intron variant rs2396766: intron variant rs12671330: 500B					Association w/ precentral ROI: **rs7784315: p=0.0007, p corrected (Bonfer) = 0.0273** rs7812028: p=0.0021, p corrected (Bonfer) = 0.0819 **rs17137135: p=0.0007, p corrected (Bonfer) = 0.0273** Association controlling for effects of other polymorphisms: **rs6980093 (AA>GG) w/ left IFG (t-4.64, voxel p=0.009 corrected); Cohen's *d* = 0.66 rs7784315 (TC>TT) w/ precentral activation (T-5.03, voxel p=0.003 corrected for** corresponding ROI); Cohen's *d* = 1.25 **rs6980093: Higher activation w/ A allele during reading condition (F(2,90)=12.89, p=1.2x10^{-5}; Cohen's *d* = 0.99) and speech listening (F(2,90)=3.32, p=0.040; Cohen's *d* = 0.50)** **rs7784315 associated w/ precentral activation during reading only (TC>TT): (F(1,91)=19.76, p=2.4x10^{-5}; Cohen's *d* = 1.10)**

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		Downstream Variant					
	Skeide 2016	rs923875: intron variant rs12533005: intron variant rs6980093: intron variant rs10230558: intron variant rs7782412: intron variant rs936146: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Significant	Only p-values reported	**Left superior frontal gyrus: (-3, 38, 53): $p=8.74x10^{-5}$**
CCDC136 7q32.1	Gialliusi 2016	rs59197085: intron variant	GM surface area and thickness	1,275 adults in Brain Imaging Genetics (BIG) dataset	Not significant	Only p-values reported	Multivariate association with cortical surface area: p=0.663 Multivariate association with cortical thickness: p=0.724 Gene-based analysis for CCDC136 Surface area: p=0.069 Cortical thickness: p=0.897
FLNC 7q32.1	Gialliusi 2016	None	GM surface area and thickness	1,275 adults in Brain Imaging Genetics (BIG) dataset	Not significant	Only p-values reported	Gene-based analysis for FLNC: Surface-area: p=0.495 Cortical thickness: p=0.926
DGKI 7q33	Skeide 2016	rs270891: intron variant rs270904: intron variant rs1991084: intron variant rs889869: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
CREB3L2 7q33	Skeide 2016	rs273933: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
CNTNAP2 7q35	Müller 2017	rs10246256: intron variant rs759178: intron variant	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d*	rs10246256: beta=-0.9, p=0.16; Cohen's *d* = -0.46 rs759178: beta=-0.3, p=0.50; Cohen's *d* = -0.15 rs2710102: beta=-0.4, p=0.50; Cohen's *d* = --0.20

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
						from beta coefficients	
	Skeide 2016	rs7794745: intron variant rs10246256: intron variant rs2710102: intron variant rs759178: intron variant rs17236239: intron variant rs4431523: intron variant rs2710117: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Some significant	Only p-values reported	Left cerebral peduncle: (-20, -27, -8): $p=5.66x10^{-7}$ Left inferior cerebellar peduncle: (-11, -41, -45): $p=1.70x10^{-6}$
BDNF 11p14.1	Jasinska 2016	Val66Met polymorphism (dbSNP:rs6265) – missense variant	fMRI	81 children ages 6-10 years (mean age – 8.1 years, typical reading ability)	Some significant	Calculated Cohen's *d* from t-scores For other comparisons: only p-values reported	Regions of greater activation for Met allele carriers in several regions during reading words and pseudowords: t=1.993, p=0.05, FEW corrected, cluster size = 309; Cohen's *d* = 0.47 Val/Met > Val/Val L.R. Precuneus, L. IPL: Peak activation = 0.31, p<<.01 L.R. Hippocampus, L.R. Parahippocampal gyrus, L.R. Fusiform Gyrus, Cerebellum: Peak activation = 0.39, p<<.01 L. mid frontal gyrus, L. inferior frontal gyrus, L. thalamus: peak activation = 0.27, p<.01 R. cingulate, R. mid frontal gyrus, R. sup frontal gyrus: peak activation = 0.31, p<.01 L. cingulate, L. medial frontal gyrus, L. mid frontal gyrus, L. precentral gyrus: peak activation = 0.28, p<.01 R. superior temporal gyrus, R. IPL, R. superior parietal lobule: peak activation = 0.34, p<.02
SLC2A3 12p13.31	Roeske 2011	Genome-wide association scan Significant after correction: rs4234898: none Also 2 marker haplotype of	EEG – mismatch negativity (MMN) reflecting automatic speech deviance processing	Initial sample of 200 dyslexic children aged 8-19 (mean age = 12.53) and a replication sample of	Some significant	Only p-values reported	Combined sample p-values w/ allelic, genotypic, or "carrier" models: rs2487742 (carrier T) w/ MMNa: 1.09e-04 (*Fam163A* – intron variant)

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		rs4234898 and rs11100040 analyzed (not associated w/ intron/exon)		186 dyslexic children aged 8-18 (mean age = 11.40)			rs11300 (allelic) w/ MMNb: 1.11e-05 (*ALS2* – noncoding transcript variant) **rs1365152 w/ Mma (allelic): 1.44e-04, (Carrier A): 1.14e-03 (see *CLSTN2*)** **rs2114167 w/ MMNa (allelic): 6.82e-04, (Carrier G): 5.41e-03 (see *CLSTN2*)** rs7683638 w/ MMNa (carrier G): 1.44e-07 (not gene-related) **rs4234898 w/ MMNb (allelic): 5.14e-08, (Carrier T): n/a** rs4704133 w/ MMNb (Carrier C): n/a in combined sample (not gene-related) rs9390586 w/ MMNb (Genotypic): 3.73e-02, (Carrier T): 1.05e-02 (not gene-related) rs7793973 w/ MMNa (Genotypic): 4.12e-04 (*LINC02587* – intron variant) rs1607924 w/ MMNb (Carrier A): 1.55e-03 (*SAMD12* – intron variant) rs965670 w/ MMNb (allelic): 1.55e-03 rs10996111 w/ MMNa (carrier G): n/a in combined sample rs4751178 w/ MMNb (genotypic): 1.04e-05, (carrier G): 6.55e-06 rs1777697 w/ MMNa (genotypic): 1.28e-03, (allelic): 6.05e-03, (carrier T): 3.69e-04 rs4238922 w/ MMNa (genotypic): 2.12e-03, (carrier A): 4.50e-04 rs11871364 w/ MMNa (carrier C): 2.84e-02 rs7217223 w/ MMNb (carrier C): n/a in combined sample rs2612570 w/ MMNb (carrier C): 4.87e-04 rs1736148 w/ MMNa (allelic): 5.05e-02
	Müller 2017	rs11100040 (chr4:156085802 rs4234898 (chr4:156077389)	EEG – mismatch response (MMR) in anterior region of interest	67 children (mean age = 9.63 years)	Significant	No standard deviation of dependent variable reported; estimated at 2	Associations with late component of MMR: rs11100040: Beta=1.6, p=0.0306, FDR = 0.14; Cohen's *d* = 0.88 rs4234898: Beta=1.6, p=0.03; Cohen's *d* = 0.84

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
			(ROI) (F3, Fz, F4)			based on graph Calculated Cohen's *d* from beta coefficients	
	Skeide 2015	rs11100040: none rs4234898: none -decrease expression of *SLC2A3*	Functional connectivity, fractional anisotropy	34 9-12 year old children from German Language Development Study	Some significant	Calculated from F values and *U*	Children w/o risk allele at rs11100040 showed stronger temporal correlations of activity in left IFT and left pSTG than children carrying at least one risk allele ($F(1,33)=2.81$, $P<0.05$, Bonferroni corrected); Cohen's $d = 0.58$ No significant effects between left IFG and TPJ ($F(1,33)=0.9$, $p=0.524$; Cohen's $d = 0.33$) or left pSTG and TPJ ($F(1,33)=0.69$, $p=0.682$; Cohen's $d = 0.28$). Children with a risk allele at rs4234898 did not differ from children carrying at least one risk allele in all three pairs of ROIs (IFG-pSTG: $U=67$, $p=0.233$; Cohen's $d = 0.53$; IFG-TPJ: $U=91$, $p=0.618$; Cohen's $d = 0.18$; pSTG-TPJ: $U=103$, $p=0.968$; Cohen's $d = 0.01$). No association with white matter skeleton. Children with risk allele in rs11100040 had reduced FA values in a cluster in the left arcuate fasciculus compared to non-carrier children (k=36, MNI (-34, -16, 34), $p<0.01$, cluster size Bonferroni corrected to $P<0.01$) Individual FA correlated with individual functional connectivity (partial $rs=0.6$, $p<0.005$; Cohen's $d = 1.5$)
COL4A2 13q34	Skeide 2016	rs9521789: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Some significant	Only p-values reported	Right Cerebellum, VIIb, Crus II, VIIIa: (17, -77, -54): $p=1.65 \times 10^{-4}$

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
15q11.2(BP1–BP2) deletion	Ulfarsson 2017	15q11.2(BP1–BP2) deletion	Grey and white matter volume, brain activation	Adults (18-65.71 years); 71 with the deletion and 643 controls	Significant	Effects reported as: (carrier status effect − mean)/|mean|. Unable to calculate Cohen's *d*	Carrier status-dependent changes (multiple regression) sMRI: gray matter volume **L fusiform gyrus (-35, -36, -15): Effect = 3.0 (95%CI 2.9-3.1), p=0.045** **L superior occipital (-22,-78,24): Effect = -4.8(95%CI:-5- -4.6), p=0.016** **R superior frontal (20,30,52): Effect = -5(95%CI: -5.2-4.8), p=0.016** L intraparietal sulcus (-20, -49, 39): Effect=4.2, p=0.09 R inferior frontal orbital (38, 45, -23): Effect=3.3, p=0.146 L posterior cingulate (-10, -12, 30): effect=2.6, p=0.201 R Cerebellum 8 (24, -63, -44): effect=2.7, p=0.372 L inferior frontal orbital (-39, -33, -14): effect=2.8, p=0.468 L Cerebellum 8 (-9, -63, -30): effect=2.1, p=0.683 R superior temporal (39, -34, 4): effect=1.8, p=0.694 R amygdala (16, 2, -15): effect=1.0, p=0.676 R hippocampus (27, -10, -20): effect=1.8, p=0.808 R superior occipital (26, -67, 28): effect=-4.4, p=0.236 L postcentral (-46, -10, 28): effect= -3.8, p=0.276 L thalamus (-15, -18, 15): effect=4.3, p= 0.434 L cuneus (-4, -70, 25): effect=2.5, p=0.615 R temporal inferior (51, -54, -9): effect=2.7, p=0.672 R putamen (30, 0, 10): effect=2.4, p=0.715 sMRI: white matter **R cerebellum cruz 1 (28, -72, -32): effect = 7.7(95%CI: 7.6-7.9), p=6.84x10^{-5}** **R paracentral lobule (10, -30, 54): Effect = 4.6 (95% CI 4.5-4.7), p=6.93x10^{-4}**

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							L superior temporal (-52,-12,13), Effect=5.0(95%CI 4.5-5.1), p=1.94x10^{-3} L fusiform (-30, -46, -15): effect=4.6, p=0.074 L precentral (62, 2, 25): effect=5.0, p=0.150 L supramarginal (-58, -37, 24): effect=5.1, p=0.224 R frontal mid orbital (33, 40, -17): effect=3.2, p=0.373 L frontal mid orbital (-42, 27, -15): effect=4.5, p=0.374 R supramarginal (56, -24, 39): effect=3.9, p=0.447 L paracentral lobule (-14, -34, 73): effect=4.5, p=0.526 L precuneus (-10, -58, 52): effect=3.5, p=0.634 **Anterior corpus callosum (4, 0, 22), effect=-4.6(95%CI -4.7 - -4.5), p=6.84x10^{-4}** **R amygdala (26, 2, -17), effect=-4.7(95%CI -4.8- -4.6), p=5.57x10^{-3}** L amygdala (-27, -3, -14): effect= -3.8, p=0.063 R superior frontal (22, 24, 52): effect=-4.8, p=0.446 fMRI word paradigm: PW vs. W L fusiform gyrus (-28, -36, -14), effect=68.2(95%CI 63.7-72.8), p=0.007 fMRI multiplication paradigm: C vs. F L angular gyrus (-50, -66, 24): effect=87.2(95% CI: 80.8-93.8), p=2.0810^{-4a}
CYP19A1 15q21.2	Skeide 2016	rs934634: *MIR4713HG*: intron variant, *CYP19A1*: 3 prime UTR variant rs10046: *MIR4713HG*: intron variant, *CYP19A1*: 3 prime UTR variant rs8034835: *CYP19A1*: intron variant, *MIR4713HG*: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
DYX1C1 15q21.3	Darki 2012	rs3743204 intron variant rs3743205: noncoding transcript variant rs17819126: noncoding transcript variant	White matter volume and cortical thickness	76 children and young adults 6-25 years	Some significant	Calculated Cohen's *d* from Z-scores	Association with white matter volume in left temporo-parietal regions: **rs3743204 w/ MNI (-15,-54,16): Pcorrected = $3.10x10^{-3}$, z=3.85; Cohens' d = 0.98 rs3743204 w/ MNI (13,-35,30): Pcorrected = $5.43x10^{-4}$, Z = 3.70; Cohen's *d* = 0.94** Nonsignificant effect sizes not reported
	Darki 2014	rs3743204: *DNAAF4:* intron variant, *DNAAF4-CCPG1:* intron variant	White matter volume, DTI (fractional anisotropy), cortical thickness	76 children and young adults in 9 age groups (6-25)	Some significant	Calculated Cohen's *d* from Z-scores	**Bilateral WM (superior longitudinal fasciculus and posterior corpus callosum): Z=4.11, $p=1.28x10^{-10}$; Cohen's *d* = 1.07**
	Müller 2017	rs17819126 – *DNAAF4*: missense variant, non coding transcript variant rs3743204: intron variant rs3743205: non coding transcript variant rs685935: intron variant	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Some significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	Associations with late component of MMR: **rs17819126: Beta=3.0, p-value = .0037, FDR = 0.05; Cohen's *d* = 1.74** rs3743204: Beta=-1.7, p=0.0157, FDR=0.11; Cohen's *d* = -0.88 rs3743205: Beta=-0.8, p=0.53; Cohen's *d* = -0.41
	Männel 2015	rs3743204: intron variant	Grey matter probability	32 adult males (16 with dyslexia, 16 without)	Not significant	Converted beta to r, and then converted r to Cohen's *d*	rs3743204 regression analyses: Right LOC ß=-.26, p=.10; Cohen's *d*=-0.65 Left HG/pSTG ß= .04, p=.80; Cohen's *d* = 0.18 Left pSTS ß=-.02, p=.91; Cohen's *d* = -0.14 Difference HG/pSTG-pSTS ß=.04, p=.80; Cohen's *d* = 0.18
DYX1C1 15q21.3	Skeide 2016	rs7174102: *DNAAF4* intron variant, *DNAAF4-CCPG1:* intron variant rs600753: *DNAAF4* missense variant, *DNAAF4-CCPG1* non coding transcript variant rs8037376: *DNAAF4* intron variant,	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not Significant	Not reported	Not reported

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
		DNAAF4-CCPG1: intron variant rs685935: *DNAAF4* intron variant, *DNAAF4-CCPG1:* intron variant rs11629841: *DNAAF4* intron variant, *DNAAF4-CCPG1:* intron variant rs8043049: *DNAAF4* intron variant, *DNAAF4-CCPG1:* intron variant rs3743204: *DNAAF4* intron variant, *DNAAF4-CCPG1:* intron variant					
CMIP 16q23.2	Müller 2017	rs3935802: intron variant rs6564903: intron variant rs7201632: intron variant	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs6564903: beta=0.8, p=0.16; Cohen's *d* = 0.41 rs3935802: beta=-0.7, p=0.24; Cohen's *d* = -0.36 rs7201632: beta=0.4, p=0.51; Cohen's *d* = 0.2
	Skeide 2016	rs12927866: intron variant rs6564903: intron variant rs3935802: intron variant rs4265801: *CMIP* intron variant, *LOC105371362* noncoding transcript variant rs16955705: intron variant rs7201632: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Some significant	Only p-values reported	Left middle cerebellar peduncle: (-9, -83, 42): $p=1.70x10^{-6}$ Local white matter of the right cerebellum, crus I (32, -68, -36): $p=6.86x10^{-5}$
ATP2C2 16q24.1	Müller 2017	rs16973771: intron variant rs2875891: intron variant rs8053211: intron variant	EEG – mismatch response (MMR) in anterior region of interest	67 children (mean age = 9.63 years)	Some significant	Standard deviation of dependent variable estimated at 2 from graph included in paper	Associations with late component of MMR: **rs8053211: Beta=-1.8, p = 0.0039, FDR = 0.05; Cohen's *d* = -0.98** rs2875891: Beta=-1.5, p=0.0146, FDR = 0.11; Cohen's *d* = -0.78

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
			(ROI) (F3, Fz, F4)			Calculated Cohen's *d* from beta coefficients	rs16973771: Beta=-1.4, p=0.0199, FDR = 0.11; Cohen's *d* = -0.73 rs11860694: Beta=1.7, p=0.0096; Cohen's *d* = 0.95 rs8045507: Beta=-1.4, p=0.02; Cohen's *d* = -0.75
	Skeide 2016	rs8053211: intron variant rs11860694: intron variant rs16973771: intron variant rs2875891: intron variant rs8045507: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
EPB41L3 18p11.31	Skeide 2016	rs11874896: *EPB4IL3* intron variant, *LOC100286986* intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
SETBP1 18q12.3	Perdue 2018	32 SNPs within *SETBP1* rs7230525: intron variant	fMRI	135 children 5-12 years with broad range of reading ability	Some significant	Calculated Cohen's *d* from F values	Brain activation patterns associated with rs7230525: Significant 3-way gene by lexicality by modality interaction (peak voxel: F = 20.546, p<.001; cluster size = 36 in the right inferior parietal lobule – activation increased for ancestral "t" allele homozygotes relative to derived "C" allele carriers; Cohen's *d* = 0.78
DYM 18q21.1	Skeide 2016	rs11873029: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
MYO5B 18q21.1	Müller 2017	rs555879: 3' prime UTR variant	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs555879:beta=0.5, p=0.42; Cohen's *d* = 0.25
	Skeide 2016	rs555879: 3' prime UTR variant	Voxel-based morphometry (GM and WM	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
			volume), MVPA				
NEDD4L 18q21.31	Müller 2017	rs12606138: intron variant rs809437:	EEG – mismatch response (MMR) in anterior region of interest (ROI) (F3, Fz, F4)	67 children (mean age = 9.63 years)	Not significant	Standard deviation of dependent variable estimated at 2 from graph included in paper Calculated Cohen's *d* from beta coefficients	rs12606138: beta=0.4, p=0.73; Cohen's *d* = 0.20 rs809437: beta=0, p=0.97; Cohen's *d* = 0
	Skeide 2016	rs8094327: intron variant rs12606138: intron variant	Voxel-based morphometry (GM and WM volume), MVPA	141 children (mean age = 6.4 years)	Not significant	Not reported	Not reported
COMT 22q11.21	Landi 2013	Val/Met substitution at codon 158 (rs4680)	fMRI	86 children 6-10 years (mean age = 8.28)	Some significant	Only p-values reported	Regions showing significant differences b/w the groups when viewing printed words or pseudowords: STG (56, -28, -1) p=0.0001 Parahippocampa gyrus (32, -50, -4) : p=0.0006 Precuneus (22, -56, 54) p=0.0015 Precuneus (18, -70, 30) p=0.0005 Lingual gyrus (-24, -74, -14) p=0.002 Uncus (-2, -10, -18) p=0.0002 Culmen (28, -54, 16) p=0.0001 Declive (6, -56, -43) p=0.0012 Precentral gyrus (-32, -18, 46) p=0.0005 Inferior frontal gyrus (12, 36, -22) p=0.0011 Inferior parietal lobule (40, -36, 32) p=0.0005 Precuneus (-20, -70, 40) = 0.0017 Uncus (20, 10, -24) p=0.001 Fusiform gyrus (24, -62, -12) p=0.0013 Superior frontal gyrus (-10, 66, -6) p=0.0007 Insula (-34, -38, 18) p=0.0016 Inferior frontal gyrus (-34, 14, -19) p=0.0004 Superior parietal lobule (-44, -62, 54) Uncus (2, -8, -40) p=-.—16

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Precentral gyrus (-32, -16, 64) p=0.0041 Precentral gyrus (20, -16, 54) p=0.0017 Middle frontal gyrus (-20, 44, -12) p=0.0019 Superior temporal gyrus (70, -24, 10) p=0.0021 Declive (24, -60, -30) p=0.0046 Middle temporal gyrus (64, -58, -6) p=0.0024 Cuneus (12, -96, 24) p=0.0014 Insula (-42, -34, 24) p=0.0062 Middle frontal gyrus (52, 29, 30) p=0.0033 Non-significant effect sizes not reported
RBFOX2 22q12.3	Gialliusi 2016	rs5995177: intron variant Other associated SNPS	GM surface area and thickness	1,275 adults in Brain Imaging Genetics (BIG) dataset	Some significant	Converted beta to r, and then converted r to Cohen's *d* If only p-values reported, no effect sizes calculated	Most significant associations b/w rs78563107, rs6000084, rs6000085, and rs144006011 w/ cortical thickness ($p=4.3\text{-}7.1x10^{-3}$) Gene-based analysis for RBFOX2 Surface area: p=0.173 Cortical thickness: p=0.135 SNPs rs56184882, rs339054, rs339046 suggestive associations with surface area ($p=7.3\text{-}99x10^{-3}$) Surface area: <u>rs5755988, rs6000032, rs5755989, rs5750198, rs5750199, rs5755992, rs5750200:</u> Multivariate: p=0.023 MTG_L: beta=069.26, p=0.0024 PPG_R: beta=-58.26, p=0.0353 Rs118119033 w/ cortical thickness: Multivariate: p=0.009 STG_L: beta=-0.04, p=0.0253 IFG-PO_R: beta=-0.03, p=0.0338 STG_R: beta=-0.05, p=0.0011 rs5995177: multivariate associations with: surface area: p=0.996

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							cortical thickness: p=0.012 rs599177 univariate associations with cortical thickness: Left MTG: beta = -0.019, p=0.143; Cohen's *d* = -0.14 Left IFG-PO: beta=-0.021, p=0.061; Cohen's *d* = -0.14 Left IFG-PT: beta=-0.019, p=0.117; Cohen's *d* = -0.14 **Left PPG: beta=-0.021, p=0.021; Cohen's *d* = -0.14** **Left STF: beta=-0.037, p=2.4x10^{-3}; Cohen's *d* = -0.17** **Right MTG: beta=-0.025, p=0.049; Cohen's *d* = -0.15** **Right IFG-PO: beta=-0.029, p=0.015; Cohen's *d* = -0.16** **Right IFG-PT: beta=-0.032, p=9x10^{-3}; Cohen's *d* = -0.16** Right PPG: beta=-0.01, p=0.313; Cohen's *d* = -0.12 **Right STG: beta=-0.038, p=2.3x10^{-3}; Cohen's *d* = -0.18**
SYN1 Xp11.23	Cabana 2018	Q555X mutation (on X-chromosome)	Neurite orientation and density imaging (NODDI) - microstructure	13 adults and 13 healthy subjects matched for age and sex (age 17-67 years)	Significant	Only p-values reported	Four clusters (insula, postcentral gyrus, supramarginal gyrus, and postcentral gyrus) with p<.005 found in fractional anisotropy Seven clusters (middle/inferior temporal gyrus, postcentral gyrus, superior frontal gyrus, precentral sulcus, lateral occipital sulcus, and superior parietal lobule) with p<.05. Orientation dispersion index significant decrease in only two clusters (insula, postcentral gyrus) Proton density significantly decreased in several clusters (middle/inferior temporal gyrus, lateral occipital sulcus, superior temporal gyrus, inferior parietal lobule/lateral occipital cortex, superior parietal lobule)

Gene	Reference	SNP	Imaging Phenotype	Population	Findings	Cohen's *d* Calculation	Effect Sizes
							Two clusters of decreased mean kurtosis (superior temporal gyrus and inferior parietal lobule/lateral occipital cortex)

*Bolded effect sizes were significant after correction for multiple comparisons

Part 2 Supplementary Material

Part 2 Supplementary Table 1. Significant associations between SNPs, brain phenotypes, and reading. Significant associations p<.05.

SNP	SNPs in linkage disequilibrium	Cortical Thickness	Gyrification	White Matter Volume	Reading
rs1817178		Non-significant	Non-significant	Significant association with transverse temporal gyrus; not significant after correction (p=0.03, q=0.11)	CTOPP Elision Raw (p=0.004, q=0.013) CTOPP Elision SS (p=0.0065, q=0.013) TOWRE sight Word Efficiency (p=0.03, q=0.04) TOWRE Phonemic Decoding Efficiency (p=0.054, q=0.054)
rs715693	rs7176566, rs8040414	Non-significant	Non-significant	Non-significant	
rs921764		Non-significant	Non-significant	Non-significant	
rs10459611		Non-significant	Non-significant	Non-significant	
rs2124132		Non-significant	Non-significant	Non-significant	
rs1378214		Non-significant	Non-significant	Non-significant	
exm2272238		Non-significant	Non-significant	Non-significant	
rs17358764	rs1466781, rs2045158	Non-significant	Non-significant	Non-significant	
rs6493270		Non-significant	Non-significant	Non-significant	
rs504729		Non-significant	Non-significant	Non-significant	
rs12050859		Non-significant	Non-significant	Significant association with transverse temporal gyrus (p=0.0135, q=0.054) and left cerebellum (p=0.033, q=0.066); trend after correction	CTOPP Elision Raw (p=0.63) CTOPP Elision SS (p=0.23) TOWRE sight Word Efficiency (p=0.73) TOWRE Phonemic Decoding Efficiency (p=0.05, q=0.20)
rs591143		Non-significant	Non-significant	Non-significant	
rs16959332		Non-significant	Non-significant	Non-significant	
rs281281	rs281279	Non-significant	Non-significant	Non-significant	
rs166837		Non-significant	Non-significant	Non-significant	

rs8040191		Non-significant	Non-significant	Non-significant	
rs281316		Non-significant	Non-significant	Non-significant	
rs281311	rs11070583	Non-significant	Non-significant	Non-significant	
rs2433019		Non-significant	Non-significant	Non-significant	
rs281215		Non-significant	Non-significant	Non-significant	
rs17311369		Non-significant	Non-significant	Non-significant	
rs281236		Non-significant	Non-significant	Non-significant	
rs4143629		Non-significant	Non-significant	Non-significant	
rs8039398		Non-significant	Non-significant	Non-significant	
rs16959499	rs16959504, rs1559677	Non-significant	Non-significant	Non-significant	
rs1390869		Non-significant	Non-significant	Non-significant	
rs2173093		Non-significant	Non-significant	Non-significant	
rs1496908		Non-significant	Non-significant	Non-significant	
rs2059475		Non-significant	Non-significant	Non-significant	
rs1369645		Non-significant	Cluster in supramarginal gyrus but not significant after correction	Non-significant	
rs16952896		Non-significant	Cluster in supramarginal gyrus but not significant after correction	Non-significant	
rs1618196		Non-significant	Non-significant	Non-significant	
rs890153		Non-significant	Non-significant	Non-significant	
rs16959669		Significant Cluster in Fusiform Gyrus (maximum p-value = 0.0002; cluster-wise p-value = 0.0016)	Non-significant	Non-significant	CTOPP Elision raw score (p=0.26), the CTOPP Elision standard score (p=0.97), the TOWRE sight word efficiency (p=0.56), TOWRE phonemic decoding efficiency (p=0.47)
rs1435749	rs1656631	Non-significant	Non-significant	Non-significant	
rs4270119		Non-significant	Significant cluster in supramarginal gyrus (Maximum p-value = 0.00027; cluster-wise p-value = 0.027)	Non-significant	CTOPP Elision raw score (p=.60), CTOPP Elision standard score (p=.77), TOWRE sight word efficiency (p=.70), TOWRE phonemic decoding efficiency (p=0.87)
rs11634974		Non-significant	Non-significant	Non-significant	
rs1898110		Non-significant	Non-significant	Significant association in transverse temporal gyrus but not after correction (p=0.046, q=0.18)	CTOPP Elision raw score (p=.50), CTOPP Elision standard score (p=.71), TOWRE sight word efficiency (p=.53), TOWRE

					phonemic decoding efficiency (p=0.63).
rs2117798		Non-significant	Occipital cluster but not significant after correction	Non-significant	
rs1435755		Non-significant	Non-significant	Non-significant	
rs2136897		Non-significant	Non-significant	Non-significant	
rs12898202		Non-significant	Non-significant	Non-significant	
rs1435742		Non-significant	Non-significant	Non-significant	
rs1865649		Non-significant	Non-significant	Non-significant	
rs1912637		Non-significant	Non-significant	Non-significant	
rs9673061		Non-significant	Non-significant	Non-significant	
rs12593611		Non-significant	Non-significant	Non-significant	
exm2267767		Non-significant	Non-significant	Non-significant	
rs4775702		Non-significant	Non-significant	Non-significant	
rs11070608		Non-significant	Non-significant	Non-significant	
rs1224656		Non-significant	Non-significant	Non-significant	
rs76739	rs765	Non-significant	Non-significant	Non-significant	
rs11629796		Non-significant	Non-significant	Non-significant	
rs501916		Non-significant	Non-significant	Non-significant	
rs532598		Non-significant	Non-significant	Non-significant	
exm1159243		Non-significant	Non-significant	Non-significant	

Part 2 Supplementary Analysis

Imaging genetic analyses of white matter volume were completed in the Haskins sample and UCSF sample separately as a comparison to cortical thickness and gyrification analyses. These analyses were carried out using the SNPs that had been identified as significant when using the whole sample, including rs1817178, rs12050859, and rs1898110. These SNPs were analyzed for association with white matter volume in the transverse temporal gyrus. For rs1817178, results were significant in the Haskins sample (t=-2.34, p=0.02), but not the UCSF sample (t=-0.15, p=0.88). For rs12050859, the results were significant in the Haskins sample (t=2.48, p=0.016) but not in the UCSF sample (t=0.68, p=0.50). Lastly, for rs1898110, the results were trending towards significance in the Haskins sample (t=1.85, p=0.07), but not significant in the UCSF

sample (t=1.27, p=0.22). These results were likely due to lower power in the UCSF sample due to a smaller sample size, and younger age range.