The human lexinome: Genes of language and reading

Christopher J. Gibson [a], Jeffrey R. Gruen [a,b,*]

[a] Department of Pediatrics, Yale Child Health Research Center, Yale University School of Medicine, New Haven, CT, USA

[b] Department of Genetics and Investigative Medicine Program, Yale University School of Medicine, New Haven, CT, USA

Abstract

Within the human genome, genetic mapping studies have identified 10 regions of different chromosomes, known as DYX loci, in genetic linkage with dyslexia, and two, known as SLI loci, in genetic linkage with Specific Language Impairment (SLI). Further genetic studies have identified four dyslexia genes within the DYX loci: *DYX1C1* on 15q, *KIAA0319* and *DCDC2* on 6p22, and *ROBO1* on 13q. *FOXP2* on 7q has been implicated in the development of Speech-Language Disorder. No genes for Specific Language Impairment have yet been identified within the two SLI loci. Functional studies have shown that all four dyslexia genes play roles in brain development, and ongoing molecular studies are attempting to elucidate how these genes exert their effects at a subcellular level. Taken together, these genes and loci likely represent only a fraction of the human lexinome, a term we introduce here to refer to the collection of all the genetic and protein elements involved in the development of human language, expression, and reading.

Learning outcomes: The reader will become familiar with (i) methods for identifying genes for complex diseases, (ii) the application of these methods in the elucidation of genes underlying disorders of language and reading, and (iii) the cellular pathways through which polymorphisms in these genes may contribute to the development of the disorders.
© 2008 Elsevier Inc. All rights reserved.

* Corresponding author at: Department of Pediatrics, Yale Child Health Research Center, 464 Congress Avenue, New Haven, CT 06520-8081, USA. Tel.: +1 203 737 2202; fax: +1 203 737 5972.

E-mail address: jeffrey.gruen@yale.edu (J.R. Gruen).

1. Introduction

The idea that differences in language and reading abilities are partially attributable to genetics is not new. As early as the 19th century, for instance, educators and physicians described families in which more than one member had difficulty learning to read (Hinshelwood, 1917). With the evolution of more sophisticated techniques of genetic analysis, our understanding of the biologic basis of these language disorders continues to grow. This article summarizes some of the current understanding of the genes and proteins that underlie three language disorders: developmental dyslexia, also known as Reading Disability (RD: OMIM 127700, 605755; McKusick, 2000), Specific Language Impairment (SLI: OMIM 606711, 606712), and Speech-Language Disorder (SLD: OMIM 602081). It is our hope that continued study of these and other disorders may eventually elucidate all the components of the human lexinome, a term we introduce here to refer to the entire collection of genes and proteins involved in the development of speech, language, and reading.

2. The identification of language disorder genes

To understand how the genes for language disorders are identified, it is essential to understand the types of studies geneticists utilize. This can be visualized as a multi-step process of increasingly narrow scope, starting with heritability studies, proceeding to classical karyotype analysis, then to genetic linkage analysis followed by high-resolution genetic association studies in a process termed ''positional cloning'' (Collins, 1992), and ending with functional assays of candidate genes.

2.1. Heritability studies

The simple observation that a trait runs in families is not enough to establish that it is genetic, since its familial nature may instead be due to shared environment. The most common method used to confirm that a trait is at least partly heritable is a twin study, in which the concordance of the trait in monozygotic twins, who are genetically identical, is compared to concordance in dizygotic twins, who are not (Farrer, 2004). A significantly higher concordance in monozygotic twins implies that the trait has a genetic component. Statistical analysis can estimate the degree of variation in a trait due solely to genetics; this is known as the trait's *heritability*.

For RD, twin studies have shown a heritability of 44–77% (DeFries, Fulker, & LaBuda, 1987). This approaches the heritability of other complex diseases, such as schizophrenia (70–90%; Sullivan, Kendler, & Neale, 2003) and shows that RD has a large genetic component has well as an environmental component. The heritability of several language traits within SLI has been estimated at close to 100%, indicating that this disorder may be almost entirely genetic (Bishop, North, & Donlan, 1995).

2.2. Karyotype analysis

This type of genetic analysis has been used since the 1950s and involves light microscopic analysis of peripheral white blood cell chromosomes arrested in metaphase, and stained with giemsa to distinguish characteristic banding patterns for each chromosome. Classical karyotype analysis can detect chromosomal deletions or duplications, as well as the exchange of large chromosomal segments, called translocations, on the order of 10 Mb. With higher resolution labeling, such as that used in Fluorescence In Situ Hybridization, smaller chromosomal abnormalities on the order of 100 kb can be identified. No method of karyotype analysis, however, can provide data on the underlying DNA sequence of the chromosomes.

Although karyotype analysis is routinely used in clinical genetics, it is not commonly used in large research studies. Nevertheless, we mention it here because two of the four RD genes, *DYX1C1* and *ROBO1*, and the *SLD* candidate gene *FOXP2*, were discovered through the serendipitous discovery of chromosomal translocations by karyotype analysis. We will detail its application when we discuss the specific genes below.

2.3. Genetic linkage studies

Traditionally, once a trait or disease has been shown to have a genetic component, the next step has been genetic linkage analysis, which determines the chromosomal regions, usually 1–20 Mb in size, that contribute to the development of the trait (Pericak-Vance, 1996a, 1996b). In performing these studies, researchers compare genotypes from multiple members across several generations of families affected by the trait in question. In genome-wide linkage studies, these genotypes are of markers distributed across all chromosomes. For both linkage studies and genetic association studies (see below), a ''marker'' is a fragment of DNA sequence at a unique location within the entire genome that varies with a known frequency within a population. The markers most commonly used in modern genetic studies are Single Nucleotide Polymorphisms (SNPs, pronounced ''snips''), which are variations in single bases that occur on the order of one per 100 bases of DNA (Gregory & Gilbert, 2005). The variation of single SNPs has been determined by compiling the results of many sequencing studies in open-source databases; thus it is possible to know, for instance, that SNP x on chromosome y exists as adenine (A) in 60% of Western Europeans and as guanine (G) in 40%. Geneticists call each unique variant of a marker an ''allele;'' thus, the adenine variant of SNP x might be called ''allele 1,'' and the guanine variant might be called ''allele 2.'' The frequency that an individual SNP in a defined population has at least two alleles is called the *heterozygosity value*, which is critical to their usefulness in mapping loci for diseases or traits.

By tracing the lineage of marker alleles across generations and comparing it with the lineage of the trait in the family, researchers can identify markers within a specific chromosomal region that are inherited in the same pattern as the trait. These markers are said to be ''in genetic linkage'' with the trait, and the chromosomal region in which they are encoded is termed a ''susceptibility locus.''

Numerous linkage studies have been performed for RD and have identified a total of 10 susceptibility loci, which are collectively designated by the label *DYX*. Of these, four have

been most consistently replicated: DYX1 on 15q (Grigorenko et al., 1997; Wigg et al., 2004), DYX2 on 6p (Fisher et al., 1999; Gayán et al., 1999), DYX5 on 3p (Nopola-Hemmi et al., 2000), and DYX6 on 18p (Fisher et al., 2002). These regions are likely to encode genes responsible for the bulk of heritable susceptibility to RD across the general population, while the other six loci likely encode genes with RD polymorphisms unique to small populations. Similar linkage analyses have identified two loci for Specific Language Impairment: SLI1 on 16q, and SLI2 on 19q (Consortium, 2002).

SLD is a complex disease that involves both severe developmental dyspraxia, in which lack of coordination of the mouth and tongue render speech poorly coherent, and difficulties with conventional grammar and usage. It was originally characterized in a single multi-generational family and localized by linkage analysis to a region of 7q31, termed SPCH1 (Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998; Lai et al., 2000).

2.4. Genetic association studies

After a linkage study has implicated a chromosomal region in the development of a disease, the next step has traditionally been to identify possible candidate genes using genetic association studies, which compare the frequency of marker alleles in affected subjects (cases) to the frequency in matched unaffected controls. The premise of these studies is that, due to historical recombination events over many generations, marker alleles found significantly more often in affected individuals must be in close physical proximity to the disease-causing mutation. Whereas genetic linkage analysis can sensitively localize the location of a disease-related gene to within 5–10 million base pairs, genetic association studies, while less sensitive, are more precise.

The need for a prior linkage study can be eliminated, however, if the entire genome could be interrogated by genetic association, as opposed to a single susceptibility locus, for disease-associated marker alleles. The advent of gene-chip technology, in which one million markers spanning the genome can be tested for association with a trait or disease, has made such ''genome-wide association studies'' (GWASs) possible. Several well-publicized GWAS have already been performed for common complex disorders (due to a combination of multiple genes and environmental factors), such as coronary artery disease (Samani et al., 2007), breast cancer (Easton et al., 2007), Type 2 diabetes (Scott et al., 2007) and multiple sclerosis (Hafler et al., 2007).

The major drawback to GWAS is the huge number of markers needed to cover the genome, which greatly increases the likelihood of false-positive associations due to multiple testing. This requires the recruitment of several thousand case subjects and at least as many controls, making GWAS expensive and complicated, and effectively impossible for rare diseases (Zondervan & Cardon, 2007).

Within the three most-replicated *DYX* loci, four candidate genes for dyslexia have been identified: *DYX1C1* in DYX1 (Taipale et al., 2003), *DCDC2* (Meng et al., 2005b) and *KIAA0319* (Cope et al., 2005) in DYX2, and *ROBO1* in (Hannula-Jouppi et al., 2005) DYX5. The SPCH1 locus was found to contain the gene *FOXP2*, which has been proposed as a risk gene for SLD (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). No studies have yet been conducted that identify candidate genes within the two SLI loci.

3. Candidate genes for language and reading disorders

The methods described above have thus far resulted in the identification of four genes for RD and one gene for SLD, which we describe here.

3.1. DCDC2

Meng et al. (2005b) first identified *DCDC2* in a genetic association study of 220 RD families from Colorado, which was confirmed by an independent study from Germany by Schumacher et al. (2006) shortly thereafter. Meng et al. examined 149 SNPs over 1.2 Mb of 6p22 and found a peak of association in six non-coding SNPs within an intron of *DCDC2*. Since introns are spliced out of the RNA transcript and do not affect the sequence of the ultimate protein product, these results imply that the susceptibility polymorphisms for *DCDC2* occur in a regulatory region of the gene. That is, the known polymorphisms do not affect the amino acid sequence of the *DCDC2* protein product, but rather affect where (e.g., specific brain region or type of neuron), when (e.g., during brain development), and/or how much of the protein is produced.

3.2. KIAA0319

Cope et al. (2005) were the first to identify *KIAA0319* by interrogating 5.3 million bases (Mb) spanning 6p21 with 57 SNPs in a sample of 223 RD siblings in the United Kingdom. The study found association not within the gene itself, but in a region immediately adjacent to the gene's transcription start site. These regions, called 5′ untranslated regions (5′ UTRs), regulate the timing and degree of transcription. In this case, the SNPs associated with the development of RD appeared to decrease the transcription of *KIAA0319*. The results of the initial association study for *KIAA0319* have been replicated in two further studies, one in a cohort from Wales (Harold et al., 2006) and one in a cohort from the United States (Luciano et al., 2007).

3.3. DYX1C1

As previously mentioned, *DYX1C1* (originally called *EKN1*) was initially discovered by karyotype analysis, which showed a translocation segregating with RD in a family in Finland (Nopola-Hemmi et al., 2000). In this case, the q11 portion of chromosome 15, which had already been identified by linkage analysis and designated DYX1, was translocated onto chromosome 2, and the q21 portion of chromosome 2 was translocated onto chromosome 15 (notated t(2;15)(q11;q21)); analysis of the exact position of the chromosome 15 breakpoint showed that it lay within a previously described gene called *EKN1*. A small genetic study of families in Finland showed association with *EKN1/DYX1C1* (Taipale et al., 2003), as did a study in Canada (Wigg et al., 2004), but studies in the United States (Meng et al., 2005a) and Italy (Marino et al., 2005) failed to show an association. This implies that the importance of *DYX1C1* as a susceptibility gene for RD may be limited to specific populations.

3.4. ROBO1

Similar to *DYX1C1*, *ROBO1* was first identified by finding a translocation in an RD family (Nopola-Hemmi et al., 2001) through karyotype analysis. In this case, the translocation was between the p12 region of chromosome 3 and the q11 region of chromosome 8 (notated t(3;8)(p12;q11)). Since 3p12 had already been identified as DYX5 in a previous linkage study, an association study using SNPs was conducted on the original linkage cohort, which showed an association between certain SNPs within *ROBO1* and decreased expression of the gene (Hannula-Jouppi et al., 2005). However, there has not yet been independent validation of this gene in a separate RD cohort.

3.5. FOXP2

Like *DYX1C1* and *ROBO1*, *FOXP2* was found by karyotype analysis, when a translocation present in an individual with SLD was found to lie within *FOXP2*. With this starting point, members of the family in which SLD was originally described were screened for mutations (Hurst, Baraitser, Auger, Graham, & Norell, 1990). This analysis found a novel SNP in *FOXP2* inherited in the same pattern as the SLD language defects that was not present in any of 364 control individuals, providing strong evidence that *FOXP2* is responsible for the development of SLD (Lai et al., 2001). A more recent study found an association between a different *FOXP2* polymorphism and the development of isolated verbal dyspraxia in 46 children (MacDermot et al., 2005), indicating a broad role for *FOXP2* in the formation of words and language.

4. Validation of candidate genes: functional studies

Once candidate genes have been identified by genetic association studies, the final step in elucidating the pathway through which they contribute to the disease is the determination of their physiologic function. The disease variants of the gene, whether full translocations or smaller sequence variations, can then be evaluated in the context of the gene's function to understand how they confer disease susceptibility. In the case of RD, three genes (*DCDC2*, *KIAA0319*, and *DYX1C1*) appear to influence the migration of developing neurons during early embryogenesis, while *ROBO1* appears to affect the extension of axons from neuron cell bodies.

4.1. DCDC2

Surveys of *DCDC2* messenger RNA (mRNA) levels have shown that it is expressed in the brain, especially in the temporal cortex and cingulate gyrus, areas known to be involved in reading. Functional studies of the gene implied it may be necessary for the proper migration of neurons from the region around the brain ventricles, where they originate during early embryogenesis, to the outermost layer of the cerebral cortex, where they reside in maturity (Meng et al., 2005b). In these studies, called RNAi knockdown assays, *DCDC2* mRNA in embryonic rat brain is specifically prevented from being translated into protein in

early neural progenitor cells. These RNAi experiments showed that knockdown of *DCDC2* expression arrested neurons in mid-migration in the intermediate layers of the brain, and prevented them from assuming their normal positions in the outer cortex.

At a molecular level, *DCDC2*'s effect on neuronal migration may occur via interactions between the *DCDC2* protein product and the cellular scaffolding known as the cytoskeleton; one of the functional domains of the *DCDC2* protein product, known as the doublecortin domain, has been shown in other studies to stabilize the assembly of important cytoskeletal components called microtubules (Francis et al., 1999). If this proposed pathway is correct, decreased levels of the *DCDC2* protein product caused by dyslexia-associated polymorphisms may confer disease susceptibility by destabilizing microtubule structures and impairing the ability of neurons to migrate through the developing brain.

Since it is not usually possible to obtain brain tissue from living humans, it is unknown whether the same types of neuronal malformations observed in rats during RNAi experiments exist in the brains of human dyslexics with *DCDC2* polymorphisms. However, it is easy to imagine how aberrant development of neurons could impact the normal development of neuronal circuits that connect discontinuous reading centers in the brain, and therefore affect higher-order functions such as reading. With the advent of new imaging modalities like functional MRI, we are gaining the ability to non-invasively study the relationships of brain structure, neuronal circuits, and function to genetics in human subjects.

4.2. KIAA0319

This gene's mRNA is present at even greater levels in brain than that of *DCDC2*, particularly in the visual and parietal cortices. RNAi knockdown studies of *KIAA0319* have also resulted in aberrant migration of affected neurons, though in a pattern different from that of *DCDC2* knockdown (Paracchini et al., 2006). The molecular mechanism by which *KIAA0319* exerts its neurostructural effects has not yet been elucidated, though recent studies have shown that the full protein product of the gene resides in the plasma membrane of neurons and may mediate interactions between these and the supporting cells of the brain, known as glial cells (Velayos-Baeza, Toma, Paracchini, & Monaco, 2007).

4.3. DYX1C1

Like *DCDC2* and *KIAA0319*, *DYX1C1* has been shown to be highly expressed in brain (Taipale et al., 2003). Similar to both those genes, furthermore, RNAi knockdown of *DYX1C1* significantly impaired neuronal migration, implying that it likely plays a role in early brain development (Wang et al., 2006).

4.4. ROBO1

As opposed to the other three RD susceptibility genes, *ROBO1* has not been shown to affect neuronal migration. Rather, the gene has been shown to encode an axonal guidance receptor, that is, a protein involved in receiving cellular signals to help direct the projection

of axons, which carry electrical signals outward from the neuron cell body (Hannula-Jouppi et al., 2005). As in the case of *DCDC2*, understanding part of the molecular role of the *ROBO1* protein product lends credibility to assertions of its role in reading and of *ROBO1* mutants in the development of some cases of RD.

4.5. FOXP2

Although the function of the *FOXP2* protein product is not entirely known, the gene is thought to encode a transcription factor, a type of protein that binds to and regulates the transcription of DNA into RNA. This is mainly based on the presence of a "forkhead" domain, which has been shown to have a DNA-binding role in other transcription factors, in the *FOXP2* protein product (Lai et al., 2001).

Although it is not yet known how mutations in this transcription factor could lead to the disruption of speech production, recent studies of songbirds have shown an acute down-regulation of *FOXP2* RNA during singing in singing-associated regions of the brain (Teramitsu & White, 2006). Furthermore, RNAi knockdown of *FOXP2* in songbirds impaired their ability to correctly imitate the songs of other birds (Haesler et al., 2007). These studies may indicate an evolutionary role for *FOXP2* in the development of communication and language, and invite further study of the mechanism by which this gene exerts its effects.

5. Future directions

The results summarized above have contributed to our understanding of the pathways underlying reading and language, and to the ways in which these pathways can be disrupted. Ultimately, researchers are hopeful that this knowledge can be used for the early detection of children at risk for developing language and reading disorders. We also hope that our knowledge will permit the development of interventions, either pharmacological or behavioral, to ameliorate these disorders. Currently, however, substantial gaps in our understanding of these disorders prevent us from achieving these goals.

For instance, we have only just begun to understand the molecular and cellular mechanics that underlie RD. Functional studies of *DCDC2*, *KIAA0319*, and *DYX1C1* indicate that they all play a role in the migration of neurons during brain development. This implies a common pathway in which the protein products of all three genes – and other, as-yet-unknown elements – interact to induce the proper directionality and distance of cellular movement. Yet the details of these interactions remain mostly unknown, and their discovery will require strong collaboration between geneticists, cellular biologists, and neuroscientists.

We also do not understand how the cellular dysfunction that may underlie these disorders lead to disruption of higher-order neurological functions like speech and reading. How is it, for instance, that changes in the migration of certain neurons disrupt the development of reading ability while preserving overall intelligence? Although we have only just begun to investigate this question, intriguing new data from functional MRI studies of dyslexics show that certain polymorphisms of *DCDC2* and *KIAA0319* are associated with unique circuits in

the brain, compared with fluent readers. Further studies like these may help us understand how mutations in all the genes described here alter brain structure and function.

Finally, we must recognize that for most of the disorders described here, elucidating the elements of the lexinome will explain only part of the problem. With the exception of SLI, which appears to be almost entirely genetic, most language and reading disorders appear to involve both genetic and environmental components. Comprehensively addressing these disorders will require detailed, large-scale studies of how they are produced by interactions between genes and environment, which in turn will require cooperation between biologists, physicians, psychologists, and educators. Some such studies are already underway, and the knowledge we gain from them, in combination with our improving understanding of basic biology and neuroscience, will help us make great strides in understanding and treating these complex disorders of reading, speech, and language.

Acknowledgements

Appendix A. Continuing education

1. Compared to genetic association studies, genetic linkage analysis is characterized by greater:
   a. Precision.
   b. Sensitivity.
   c. Both precision and sensitivity.
   d. False negative rate.
2. *DCDC2* was identified as an RD risk gene through
   a. Karyotype analysis, followed by a genetic association study.
   b. A genetic association study, followed by linkage analysis.
   c. Genome-wide association study.
   d. Genetic linkage analysis, followed by a genetic association study.
3. *ROBO1* is thought to encode a protein that
   a. Helps to guide axonal projections out from the neuron body.
   b. Is involved in building the protective sheath around axons.
   c. Interacts with cell structure to mediate neuronal migration during early brain development.
   d. Interacts with glial cells, the supporting cells of the nervous system.
4. Which of the following genes has been associated with Specific Language Impairment?
   a. *KIAA0319*.
   b. *DYX1C1*.
   c. *FOXP2*.
   d. There are currently no candidate genes proposed for Specific Language Impairment.

References

Bishop, D. V., North, T., & Donlan, C. (1995). Genetic basis of Specific Language Impairment: Evidence from a twin study. *Developmental Medicine and Child Neurology, 37*(1), 56–71.

Collins, F. S. (1992). Positional cloning: Let's not call it reverse anymore. *Nature Genetics, 1*(1), 3–6.

Consortium, S. (2002). A genomewide scan identifies two novel loci involved in Specific Language Impairment. *American Journal of Human Genetics, 70*(2), 384–398.

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

DeFries, J. C., Fulker, D. W., & LaBuda, M. C. (1987). Evidence for a genetic aetiology in Reading Disability of twins. *Nature, 329*(6139), 537–539.

Easton, D. F., Pooley, K. A., Dunning, A. M., Pharoah, P. D., Thompson, D., Ballinger, D. G., et al. (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature, 447*(7148), 1087–1093.

Farrer, L. (2004). Collection of clinical and epidemiological data for linkage studies. In Haines, J. L., Korf, B. R., Morton, C. C., Seidman, C. E., Seidman, J. G., & Smith, D. R. Eds. *Current protocols in human genetics* (Vol. S40). New York, NY: Wiley and Sons, Inc.. pp. 1.1.1–1.1.17.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics, 18*(2), 168–170.

Fisher, S. E., Marlow, A. J., Lamb, J., Maestrini, E., Williams, D. F., Richardson, A. J., et al. (1999). A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *American Journal of Human Genetics, 64*(1), 146–156.

Fisher, S. E., Francks, C., Marlow, A. J., MacPhie, I. L., Newbury, D. F., Cardon, L. R., et al. (2002). Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nature Genetics, 30*, 86–91.

Francis, F., Koulakoff, A., Boucher, D., Chafey, P., Schaar, B., Vinet, M. C., et al. (1999). Doublecortin is a developmentally regulated, microtubule-associated protein expressed in migrating and differentiating neurons. *Neuron, 23*(2), 247–256.

Gayán, J., Smith, S. D., Cherny, S. S., Cardon, L. R., Fulker, D. W., Brower, A. M., et al. (1999). Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *American Journal of Human Genetics, 64*(1), 157–164.

Gregory, S., & Gilbert, J. (2005). Strategies for genotype generation. In Haines, J. L., Korf, B. R., Morton, C. C., Seidman, C. E., Seidman, J. G., & Smith, D. R. Eds. *Current protocols in human genetics* (Vol. S47). New York, NY: Wiley and Sons, Inc.. pp. 1.3.1–1.3.16.

Grigorenko, E. L., Wood, F. B., Meyer, M. S., Hart, L. A., Speed, W. C., Shuster, A., et al. (1997). Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *American Journal of Human Genetics, 60*(1), 27–39.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of *FOXP2* in songbird basal ganglia nucleus area X. *PLoS Biology, 5*(12), e321.

Hafler, D. A., Compston, A., Sawcer, S., Lander, E. S., Daly, M. J., De Jager, P. L., et al. (2007). Risk alleles for multiple sclerosis identified by a genomewide study. *New England Journal of Medicine, 357*(9), 851–862.

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

Harold, D., Paracchini, S., Scerri, T., Dennis, M., Cope, N., Hill, G., et al. (2006). Further evidence that the *KIAA0319* gene confers susceptibility to developmental dyslexia. *Molecular Psychiatry, 11*(12), 1085–1091.

Hinshelwood, J. (1917). *Congenital word-blindness*. Chicago, IL: Medical Book Co.

Hurst, J. A., Baraitser, M., Auger, E., Graham, F., & Norell, S. (1990). An extended family with a dominantly inherited speech disorder. *Developmental Medicine and Child Neurology, 32*(4), 352–355.

Lai, C. S., Fisher, S. E., Hurst, J. A., Levy, E. R., Hodgson, S., Fox, M., et al. (2000). The SPCH1 region on human 7q31: Genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *American Journal of Human Genetics, 67*(2), 357–368.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature, 413*(6855), 519–523.

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

MacDermot, K. D., Bonora, E., Sykes, N., Coupe, A. M., Lai, C. S., Vernes, S. C., et al. (2005). Identification of *FOXP2* truncation as a novel cause of developmental speech and language deficits. *American Journal of Human Genetics, 76*(6), 1074–1080.

Marino, C., Giorda, R., Lorusso, M. L., Vanzin, L., Salandi, N., Nobile, M., et al. (2005). A family-based association study does not support *DYX1C1* on 15q21.3 as a candidate gene in developmental dyslexia. *European Journal of Human Genetics, 13*(4), 491–499.

McKusick, V. A. (2000). *OMIM—Online Mendelian Inheritance in Man*. Retrieved December 20, 2005, from, Johns Hopkins University, National Center for Biotechnology Information (NCBI) website: http://www.ncbi.nlm.nih.gov/omim/.

Meng, H., Hager, K., Held, M., Page, G. P., Olson, R. K., Pennington, B. F., et al. (2005a). TDT-association analysis of EKN1 and dyslexia in a Colorado twin cohort. *Human Genetics, 118*(1), 87–90.

Meng, H., Smith, S. D., Hager, K., Held, M., Liu, J., Olson, R. K., et al. (2005b). *DCDC2* is associated with Reading Disability and modulates neuronal development in the brain. *Proceedings of the National Academy of Sciences of the United States of America, 102*(47), 17053–17058.

Nopola-Hemmi, J., Taipale, M., Haltia, T., Lehesjoki, A. E., Voutilainen, A., & Kere, J. (2000). Two translocations of chromosome 15q associated with dyslexia. *Journal of Medical Genetics, 37*(10), 771–775.

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

Paracchini, S., Thomas, A., Castro, S., Lai, C., Paramasivam, M., Wang, Y., et al. (2006). The chromosome 6p22 haplotype associated with dyslexia reduces the expression of *KIAA0319*, a novel gene involved in neuronal migration. *Human Molecular Genetics, 15*(10), 1659–1666.

Pericak-Vance, M. A. (1996a). Analysis of genetic linkage data for Mendelian traits. In Haines, J. L., Korf, B. R., Morton, C. C., Seidman, C. E., Seidman, J. G., & Smith, D. R. Eds. *Current protocols in human genetics* (Vol. S9). New York, NY: Wiley and Sons, Inc.. pp. 1.4.1–1.4.21.

Pericak-Vance, M. A. (1996b). Overview of linkage analysis in complex traits. In Haines, J. L., Korf, B. R., Morton, C. C., Seidman, C. E., Seidman, J. G., & Smith, D. R. Eds. *Current protocols in human genetics* (Vol. S9). New York, NY: Wiley and Sons, Inc.. pp. 1.9.1–1.9.19.

Samani, N. J., Erdmann, J., Hall, A. S., Hengstenberg, C., Mangino, M., Mayer, B., et al. (2007). Genomewide association analysis of coronary artery disease. *New England Journal of Medicine, 357*(5), 443–453.

Schumacher, J., Anthoni, H., Dahdouh, F., Konig, I. R., Hillmer, A. M., Kluck, N., et al. (2006). Strong evidence of *DCDC2* as a susceptibility gene for dyslexia. *American Journal of Human Genetics, 78*(1), 52–62.

Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L., et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science, 316*(5829), 1341–1345.

Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait: Evidence from a meta-analysis of twin studies. *Archives of General Psychiatry, 60*(12), 1187–1192.

Taipale, M., Kaminen, N., Nopola-Hemmi, J., Haltia, T., Myllyluoma, B., Lyytinen, H., et al. (2003). A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11553–11558.

Teramitsu, I., & White, S. A. (2006). *FOXP2* regulation during undirected singing in adult songbirds. *Journal of Neuroscience, 26*(28), 7390–7394.

Velayos-Baeza, A., Toma, C., Paracchini, S., & Monaco, A. P. (2007). The dyslexia-associated gene *KIAA0319* encodes highly *N*- and *O*-glycosylated plasma membrane and secreted isoforms. *Human Molecular Genetics* .

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

Wigg, K. G., Couto, J. M., Feng, Y., Anderson, B., Cate-Carter, T. D., Macciardi, F., et al. (2004). Support for EKN1 as the susceptibility locus for dyslexia on 15q21. *Molecular Psychiatry, 9*(12), 1111–1121.

Zondervan, K. T., & Cardon, L. R. (2007). Designing candidate gene and genome-wide case–control association studies. *Nature Protocols, 2*(10), 2492–2501.