Published in final edited form as:

Life Sci. 2012 April 9; 90(13-14): 469–475. doi:10.1016/j.lfs.2012.01.016.

Defining the Genetic Architecture of Human Developmental Language Impairment

Ning Li^a and Christopher W. Bartlett^{a,b,*}

^aThe Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital The Ohio State University, Columbus, OH

^bDepartment of Pediatrics, The Ohio State University, Columbus, OH

Abstract

Language is a uniquely human trait, which poses limitations on animal models for discovering biological substrates and pathways. Despite this challenge, rapidly developing biotechnology in the field of genomics has made human genetics studies a viable alternative route for defining the molecular neuroscience of human language. This is accomplished by studying families that transmit both normal and disordered language across generations. The language disorder reviewed here is specific language impairment (SLI), a developmental deficiency in language acquisition despite adequate opportunity, normal intelligence, and without any apparent neurological etiology. Here, we describe disease gene discovery paradigms as applied to SLI families and review the progress this field has made. After review the evidence that genetic factors influence SLI, we discuss methods and findings from scans of the human chromosomes, including the main replicated regions on chromosomes 13, 16 and 19 and two identified genes, ATP2C2 and CMIP that appear to account for the language variation on chromosome 16. Additional work has been done on candidate genes, i.e., genes chosen a priori and not through a genome scanning studies, including several studies of CNTNAP2 and some recent work implicating BDNF as a gene $\times$ gene interaction partner of genetic variation on chromosome 13 that influences language. These recent developments may allow for better use of post-mortem human brain samples functional studies and animal models for circumscribed language subcomponents. In the future, the identification of genetic variation associated with language phenotypes will provide the molecular pathways to understanding human language.

Keywords

language; specific language impairment; genetic linkage; association; reading; dyslexia

Introduction

Language is perhaps the defining trait of human beings as biological entities. Homology with other organisms is powerful in most biological contexts, but with regard to language,

© 2012 Elsevier Inc. All rights reserved

*Corresponding Author: Christopher W. Bartlett, PhD, Assistant Professor of Pediatrics Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital & The Ohio State University 700 Children's Drive (JW3926) Columbus OH 43205 Phone: (614) 722-2688, Fax: (614) 355-2728 Christopher.Bartlett@nationwidechildrens.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

animal models have yet to become universal research tools as is the case for most brain research. How then, does one find a molecular basis of language and once “found” how can it be validated? Over the course of this review a basic overview of the relevant paradigms used for discovery of human disease genes and their application to understanding the molecular basis of human language will be presented. By the end of this review, the reader will have an appreciation for how complex human behaviors are rigorously studied at the molecular level and understand how molecular biologists and human geneticists can synergize to further the understanding of language unique to humans. We use specific language impairment (SLI, defined in more detail below) as a human “model system” to understand the molecular basis of language and present the relevant human disease gene discovery paradigms with exposition of the SLI literature (see Figure 1 for an overview of the paradigm).

What is Specific Language Impairment?

Developmental communication disorders are a common cause of parental concern and medical referral yet this natural diagnostic grouping actually includes many distinct disorders with unique etiologies and symptoms (Harel et al. 1996). The standard psychiatric nosology found in the Diagnostic and Statistical Manual IV (DSM-IV) defines five primary speech and language disorders: expressive language disorder, mixed receptive-expressive language disorder, phonological disorder (pronunciation problems), stuttering, and communication disorder not-otherwise-specified (American Psychiatric Association 1994). SLI researchers typically include both expressive language disorder and mixed receptive-expressive language disorder in the definition of SLI as a super-category for research purposes, without further distinction. Therefore for research purposes, the SLI diagnosis, refers to a neurodevelopmental failure to use and/or acquire normal language given adequate education and environment, and in the absence of other explanations such as mental retardation, or sensory deficits. This disorder affects 5%-8% of preschool children (Law et al. 2000, Tomblin et al. 1997). SLI is often a persistent impairment that can affect later academic performance (Bishop and Edmundson 1987a, b, Catts et al. 2002, Rice et al. 2008, Snowling et al. 2000, Young et al. 2002).

In many cases, the speech and language problems manifest as part of a global developmental syndrome, such as autistic spectrum disorder, or mental retardation. However, by definition, SLI is not part of a larger diagnosable neurological syndrome. These exclusionary criteria are the source of the word “specific” in SLI, and with this specificity, studies of SLI are studies of language with far fewer complications than if other syndromes were also considered. However, it is shown that SLI subjects do have deficits outside the domain of language (Alloway and Archibald 2008, Conti-Ramsden et al. 2008, Leonard et al. 2007, Misyak et al. 2010, Zelaznik and Goffman 2010), prompting some researchers simply to use the term LI instead. We use the term SLI for the purposes of continuity with the majority of the relevant literature, with the understanding that SLI does not exclude *every* other kind of low performance in cognition.

Measurement of SLI symptoms

Quantitative measures of language ability have a long history and are commonly applied in educational policy as well as in intelligence testing and research. This literature defines how language is not a unitary phenomenon. Language can be viewed from different perspectives, each with a different emphasis. One may emphasize the difference between understanding language that one receives from others (i.e., receptive language) or one's use of language for expression (i.e., expressive language). There are commonly accepted tests that measure these omnibus domains of language, the two most common in SLI genetics being the Test of

Language Development (TOLD) (Hammil et al. 1987, Hammil and Newcomer 1988, Newcomer and Hammil 1988) and the Clinical Evaluation of Language Fundamentals (CELF) (Semel et al. 2003, 2004, Wiig et al. 2004).

Language consists of many building blocks, each with possibly many neural substrates with complex interrelations. One domain that has been demonstrably important to SLI genetics is phonological short-term memory (PSTM), the working memory buffer for speech sounds (Baddeley 2003). This buffer is critical for determining incoming words and storing the word context so that other language operations may be performed. Limitations in encoding these memories, processing speed, storage capacity and memory retrieval all have the potential to limit language ability at the sub-word or word level (Baddeley and Wilson 1993, Bishop et al. 1999, Gathercole et al. 1997, Gathercole et al. 1994, Montgomery and Evans 2009, Willis and Gathercole 2001).

Reading is a skill that heavily co-opts the language system so tests of reading are not purely tests of reading ability, rather they can also be informative about aspects of language. The ability to sound out unfamiliar words, often tested using nonsense words that follow the phonotactics rules of English but convey no meaning (e.g., fancalipation or grocky), may provide insight into reading and the representation of speech sounds (phonemes) in the brain. Additionally, reading comprehension (understanding what one reads) requires both reading skills and language comprehension skills. Unsurprisingly, the SLI genetic literature also utilizes these tests.

The Genetic Hypothesis and Complexity of the Disease

If a common disorder has a genetic component, then genetics will cause the disorder to aggregate in families. Stromswold (Stromswold 1998) reviewed eighteen studies of spoken language impairment familial aggregation, where families are included in the study through an SLI “proband”, or index case, with the disease used to enroll the rest of the family into the study. The proportion of additional affected family members (i.e., aggregation in the family) was significantly greater for families with an SLI proband than families with unaffected control probands in all seven studies that included control proband families. In the remaining 11 studies without control families, the prevalence among additional family members was greater than the population rate of SLI (see also, Barry et al. 2007). Since nuclear families share both environment and genes, familial aggregation is suggestive of genetics influencing a human trait though not conclusive.

When using the categorical diagnosis of SLI, twin studies report increased monozygotic (identical) twin concordance rates relative to dizygotic (fraternal) twin rates, suggesting that the observed familial aggregation can be partially attributed to genetic influences (Bishop 2002, Bishop et al. 1995, Tomblin and Buckwalter 1998). Quantitative language scores have also been used to determine heritability of language difficulty with DeThorne et al. (2005) estimating that 54% of the variation in low language is attributable to genetics, which is congruent with other previously reported heritability rates (Bishop 2002, Kovas et al. 2005). A different research group, the Twin Early Development Study, determined the heritability of normal language in a set of 4-year-old twins to be 0.39 (Colledge et al. 2002). Taken together, twin studies and additional, related extended family designs (Logan et al. 2011) indicate that both normal language variation and language disability have attributable genetic factors.

Molecular genetic studies for SLI susceptibility

Finding and enrolling large families that have a mix of people with SLI and without SLI (i.e., normal language) creates a “natural experiment” with which to determine the genetic

architecture of SLI. If genotypes segregating through the family across generations co-segregate (correlate) systematically with SLI status in the family members, then linkage exists between the genetic locus and SLI. This is known as “linkage analysis,” and numerous statistical methods for assessing the strength of the evidence for linkage are available. The same principle applies to linkage of genotypes with quantitative language variation, where linkage implies similar quantitative trait scores co-segregates with similar genotypes. Specific examples are discussed in context as they arise, sometimes in the context of examining specific regions, and other times in the context of scanning the autosomes or the whole genome (including X and part of Y).

Regardless of statistical metric chosen, genome-wide linkage analysis is performed along the length of every human chromosome. Regions of linkage often extend over many million base pairs (Mb) and may encompass hundreds of genes. Linkage analysis has identified three chromosomal loci as linked in SLI families and, therefore, harboring genes with genetic variation that negatively influences language (summarized in Table 1). The initial studies were published in 2002 when two independent groups presented linkage studies using complementary study designs (Bartlett et al. 2002, The SLI Consortium 2002).

Initial Extended Pedigree Studies

Bartlett et al. (2002) performed a genome-wide linkage scan of five nuclear and extended Canadian pedigrees consisting of 73 subjects, requiring each family to have one person with SLI and another with a language and/or a reading impairment. They assessed linkage with three categorical affection status phenotypes: 1) language impairment, 2) reading impairment -- both based on direct testing and, 3) clinical impairment that included both direct testing of language and historical information about speech/language treatment during childhood. Language was evaluated using an age-appropriate version of the TOLD and reading was assessed with the Woodcock Johnson Reading Mastery Battery. Significant evidence for linkage between chromosome 13 and reading impairment was found using a model based method where recessive inheritance was assumed ($LOD = 3.92$, or $\log_{10}$ of the odds ratio of linkage versus no linkage). They also calculated the posterior probability of linkage (PPL), a Bayesian statistical method to quantify evidence that can be directly interpreted as a probability that naturally lies on a 0 to 1 scale (though commonly reported as a percentage). In this case, the PPL was 53%. In 2004, Bartlett et al. recruited another 22 nuclear and extended families from the United States, and found evidence for linkage to chromosome 13 ($LOD = 2.61$), where only a subset of the new families contributed to this finding (Bartlett et al. 2004). Combining the new US families with the previous Canadian families increased the PPL evidence from the 2002 and 2004 studies to 92% in favor of an SLI locus at that location.

Nuclear Families Studies

In 2002, The SLI Consortium recruited 98 nuclear families with an average of 2.2 children. The study required families to have at least one child with SLI, but all siblings were administered language tests regardless of SLI status (The SLI Consortium 2002). Rather than use SLI status as the phenotype of interest, they assessed genome-wide linkage with two quantitative scores derived from the CELF, expressive language and receptive language, and with a test of nonword repetition as a measure of short-term memory for speech sounds known as phonemes (PSTM). Their analysis revealed two linked regions, one on chromosome 16 with nonword repetition ($LOD = 3.55$) and another on chromosome 19 with expressive language ($LOD = 3.55$). The SLI Consortium observed no linkage signals on chromosome 13, and conversely Bartlett and colleagues observed no linkage signals on chromosomes 16 and 19.

In 2004, The SLI Consortium increased their sample to a combined total of 184 nuclear families (393 sibling pairs), including the original 2002 genome screen cohort, to attempt to replicate their findings on chromosome 16 and 19 (The SLI Consortium 2004). Chromosome 16 showed stronger linkage in the larger dataset ($LOD = 7.46$), essentially doubling the evidence for linkage with nonword repetition as well as showing weaker signals with several tests of reading. However, the linkage signal on chromosome 19 was markedly diminished ($LOD = 1.4$), though it was shown that the two waves of data were both linked to chromosome 19 but only with different phenotypes (CELF expressive and nonword repetition, respectively, for the two cohorts). A multivariate analysis, where multiple phenotypes are analyzed simultaneously, indicated that common variation between the CELF expressive and nonword repetition phenotypes was linked to chromosome 19 ($P = .0008$) indicating that this locus was replicated.

Recent Extended Pedigree Studies

In 2011, Villanueva et al. performed a genome-wide linkage analysis of an isolated Chilean population with high rates of SLI. They applied complementary analysis methods that only utilize affected pedigree members but make no explicit assumption about dominance versus recessiveness (called a nonparametric linkage analysis, NPL) as well as the familiar LOD score that uses data from all family members but does make such assumptions. The results included genome-wide significant peaks on chromosomes 6q, 7 and 12 ($NPL > 4.99$, $P < 3.0 \times 10^{-7}$), which were significant after stringent Bonferroni correction. Two additional peaks on chromosomes 13 and 17 were also significant ($NPL > 4.08$, $P < 2.2 \times 10^{-5}$). From these five peaks, the linkages to chromosomes 7, 13 and 17 were observed across all 3 non-parametric analyses performed even after overconservative Bonferroni correction, the other 2 only survived under single NPL analysis. None of the linkage peaks align with previous SLI linkage scans. The follow-up linkage analyses indicated that 4 pedigrees contributed to the linkage signals on chromosome 7 under all three non-parametric models (max $NPL = 6.73$, $P = 4.0 \times 10^{-11}$). Under recessive parametric model of inheritance, this region achieved a LOD of 1.24. Within the candidate region, there is a single nucleotide polymorphism (SNP) (rs727714) inducing a synonymous base substitution in exon 3 of the NOBOX gene. The two-SNP haplotype (rs727714/rs969356, AG) reconstructed for chromosome is significantly more rare in unaffected individuals than affected ones ($P = 0.04$ after Bonferroni correction). This haplotype showed moderate long-range correlation, known as linkage disequilibrium, ($D' > 0.4$ and $LOD > 2$) with other SNPs that covered a 74 kb genomic region.

A Rare Pedigree with SLI and Auditory Processing Deficits—In 2010, Addis et al. performed a genome-wide linkage scan of the NE family, which included several family members who presented with an auditory processing deficit and language impairment. Linkage analysis yielded two peaks on chromosomes 4 and 12, respectively, with a maximum LOD score of 1.5. The follow-up fine mapping analysis with additional family members and greater marker density increased the LOD score on chromosome 12 to 2.1. However, the peak on chromosome 4 decreased. Reconstructed haplotypes indicated that the risk haplotype on chromosome 12 fully co-segregated with the disorder. The candidate region on chromosome 12 covered a 58.5-Mb region between markers D12S99 at 12p13.31 and D12S1686 at 12q14.3, which contains almost 600 RefSeq genes. Six candidate genes (CNTN1, FOXJ2, GRIN2B, NAB2, NELL2 and SRGAP1) were then sequenced, but no coding changes were found. Copy-number variations (CNV) analysis showed six candidate regions, however, all of them were located in regions of common structural variation unlikely to be involved in SLI.

Discussing the different linkage approaches

Different experimental design strategies and different language phenotypes contribute to the lack of replication across the groups. Use of quantitative traits, as done by The SLI Consortium, circumvents diagnosing an affection status for every family member using an a priori threshold for affection. A nuclear family with only one affected person does not have the same genetic risk loading as that a family with multiple affected persons. As a consequence of the data type, only common risk factors are mappable in small families since for rare effects the majority of families would be unlinked. Therefore the SLI Consortium design is most sensitive to common genetic variants that reduce language ability but may not be related to SLI per se. Large extended pedigrees with many affected individuals (Addis et al. 2010, Bartlett et al. 2004, Bartlett et al. 2002, Simmons et al. 2010, Villanueva et al. 2011), can have high power to detect loci despite the small overall sample size. Effects due to rare alleles are mappable in single (or few) large pedigrees as are effects from common alleles, and single families control for genetic background effectively limiting the number of relevant genetic variants within the family. Hence both designs display differential strengths for locating the candidate regions, and it is quite likely that both designs have revealed loci that contribute to the SLI phenotype.

Studies of Specific Genes

A linked region may be rather large, on the order to tens of Mb and containing up to hundreds of genes, so other designs are useful to narrow down the region of interest. As a complementary approach to linkage analysis, association analysis assesses if an allele occurs more often in persons with SLI than persons without SLI. Associations are only detected at a short distance, tens to hundreds of kilobases (kb) from the functional susceptibility allele; therefore, association analysis is a suitable method for refining localization of a linkage signal down to one or a few genes. Families are not required for this analysis since segregation from a parent to a child is not the research question. Unrelated persons can be used for this analysis i.e. a sample of unrelated case individuals with SLI and a sample of unrelated control individuals without SLI (Cardon and Bell 2001). However, family data are also informative, especially the families initially defining a linkage region and association studies using family data have been much more common in the literature. Though despite the difference in data structure, unrelated persons versus families, association analysis still retains the same key genetic properties and interpretations with minimal nuanced differences (Sham 1998).

Contactin-association protein-like 2 (CNTNAP2)

Lai et al. (2001) positionally cloned a point mutation in a forkhead-domain containing transcription factor, *FOXP2*, and implicated this mutation in a rare and severe oral motor dyspraxia with some language involvement by demonstrating perfect co-segregation with the disease in that family. On the hypothesis that variation in related genes would be likewise involved in language ability, Vernes et al. (2008) examined genes near DNA binding targets of *FOXP2*. One gene found in the chromatin immunoprecipitation screen, *CNTNAP2*, was dramatically down-regulated by *FOXP2*. Additionally, microarray and in situ hybridization experiments demonstrated enrichment of *CNTNAP2* expression in temporal and frontal lobe language areas (Abrahams et al. 2007). An association study of 184 SLI Consortium families indicated association ($P = 5.0 \times 10^{-5}$) of *CNTNAP2* genotypes and nonword repetition performance (Vernes et al. 2008). Replication of the association of *CNTNAP2* with language and reading traits has been documented in the literature (Peter et al., 2011, Whitehouse et al., 2011). Interestingly, *CNTNAP2* is also implicated by many independent studies of autism, a disorder that includes language deficits as part of the

diagnosis (Alarcon et al. 2008, Arking et al. 2008, Bakkaloglu et al. 2008, Li et al. 2010, O'Roak et al. 2011, Poot et al. 2011).

ATPase, Ca (2+)-transporting, type 2C, member 2 (ATP2C2) and C-MAF-inducing protein (CMIP)

In 2009, the SLI Consortium performed a follow-up association study of 211 nuclear families on chromosome 16 to elucidate the specific genetic variants accounting for the observed linkage (Newbury et al. 2009). Two nearby genes, ATP2C2 and CMIP, were implicated by association analysis ($P=5.5 \times 10^{-7}$ at rs6564903 and $P=2.0 \times 10^{-5}$ at rs11860694, respectively) in the families used to define the linkage signal. Both genes are associated with PSTM in what appeared to be independent (non-interacting) effects on the phenotype.

Brain-derived neurotrophic factor (BDNF)

Bartlett and colleagues (Simmons et al. 2010) performed follow-up analysis of their chromosome 13 findings in a sample of 4 Canadian families. They first established that in their sample, the observed linkage to a reading impairment phenotype was strongly influenced by underlying spoken language deficits, thus establishing more firmly the connection to SLI over a purely reading-based impairment. In a screen of known functional variants related to human cognition, though not necessarily directly related to language, the *BDNF* gene showed a trend towards association with SLI. A gene $\times$ gene interaction (epistasis) analysis indicated a strong interaction between *BDNF* and the uncloned genetic risk variants previously found on chromosome 13. Inclusion of *BDNF* genotypes into the statistical model increased the evidence for linkage to chromosome 13 by four orders of magnitude (going from LOD=4 to LOD=8). Examination of the estimated model parameters indicated that two methionine BDNF alleles are equivalent to a single risk allele from chromosome 13. Hence persons who are heterozygous on chromosome 13 have little risk for SLI unless they also have two methionine BDNF alleles, in which case they are almost certain to have SLI. Therefore, *BDNF* did not show a direct association with SLI but was only a component of an interaction that increases the risk for SLI. It is possible to speculate that having two methionine BDNF alleles may prevent or attenuate compensation in language learning for some persons at risk of developing SLI, since verbal short term memory is an essential component of language processing and *BDNF* having been associated with memory (Egan et al. 2003, Hariri et al. 2003, Komulainen et al. 2008, Miyajima et al. 2008, Pezawas et al. 2004).

Functional research

Once specific alleles in/near individual genes show strong statistical evidence for involvement in SLI, there remains the fact that the data were not derived by experiment therefore, causality is not assured. Moving a research program forward in many human studies is challenging, where randomization and group assignment are not possible. However, gene expression, including cellular localization in tissues and cell types, as well as protein expression and subcellular localization are steps that increase confidence in the validity of the findings. Additionally, these initial characterizations present a convenient point for human genetics to hand off findings to molecular biologists that can extract specific mechanistic information using traditional model systems.

Human behavioral traits are difficult to evaluate in animal models, and as language is unique to humans, the process of how to verify the biological role of specific genetic findings is made considerably more complex. Despite the many caveats and qualifications they require, it may be that postmortem human brain samples are the best hope for knowing if a genetic

variant identified statistically in families has in vivo functionality. However, there are at least two components of language, vocal learning and auditory processing of complex sounds, which are not unique to humans. Vocal learners are animals with exceptional ability to modify innate vocalizations to imitate sounds from the environment or create new sounds. Genetic manipulation of vocal learning animal models may provide insight into some aspects of language learning in humans (Bolhuis et al. 2010). For example, White et al. (2006) found that when the birds are highly actively learning to sing, there is more *FoxP2*, bilaterally in Area X, the striatal region known to be vital for song learning. In this way, human speech and birdsongs are exemplars of vocal learning, and the experimentally tractable songbird has provided molecular and physiological insights.

Studies of deficits in processing rapid frequency changes in the time window relevant for distinguishing human speech sounds (<250ms) have been conducted on rats with genetic manipulations of a dyslexia (reading impairment) gene (Threlkeld et al. 2009, Threlkeld et al. 2007). It may be possible that short-term memory for speech sounds (PSTM described in the section entitled Measurement of SLI symptoms) is amenable to animal studies using these similar paradigms which would be a significant development for language genetics since several of the SLI susceptibility loci discussed above have PSTM as the currently known “best” fit phenotype. It remains to be seen if transgenic mice are truly amenable for use in these paradigms.

Conclusion

Ultimately the goal of human disease gene finding is to have a catalog of genetic variants associated with phenotypes in humans for the purposes of disease risk assessment and treatment when disease occurs. As SLI involves many loci, it follows that not everyone will have the same genetic risk factors or the same underlying neurobiological etiology. This presents an opportunity for an analog of “personalized medicine” to be applied in SLI though it is a behavioral disorder that is not normally treated by physicians but instead by speech-language pathologists. However, the principles are clinical practice goals of personalized medicine are fully aligned in this case. They are 1) predict risk of language impairment, 2) develop strategies to prevent language impairment, 3) develop the optimal treatment plan and 4) predict the course of how language impairment and related cognition will change, for the better or worse in persons with language impairment. All of these goals are currently being addressed through strictly behavioral markers, but genetics may provide additional information to further refine progress on each of these clinical goals.

The most recent trend is to understand the genetic underpinnings of language genetics to reading phenotypes and understanding the genetics of reading as it relates to language phenotypes, which brings to fruition ideas from research started a decade ago (Bartlett et al. 2002). Though still new as a widely researched topic, there is evidence for associations of language loci with reading and reading loci with language (Newbury et al. 2011, Rice et al. 2009, Scerri et al. 2011, Simmons et al. 2010). This new area of research is noteworthy since it may stimulate novel theories on the relationship between language and reading while further clarifying the connections already being explored between these and other related processes such as speech (Lewis et al. 2006, Newbury et al. 2010, Paracchini 2011, Pennington and Bishop 2009, Peterson et al. 2007).

Acknowledgments

The authors would like to thank Dr. Judy Flax for helpful discussion on earlier versions of the manuscript; Dr. Will Ray for his vision and crucial assistance in creating Figure 1; two anonymous reviewers that provided detailed and thoughtful feedback. We gratefully acknowledge NIH funding R01DC009453 (NL and CWB) and RC1MH088288 (CWB).

References

- Abrahams BS, Tentler D, Perederiy JV, Oldham MC, Coppola G, Geschwind DH. Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc Natl Acad Sci U S A*. 2007; 104:17849–54. [PubMed: 17978184]
- Addis L, Friederici AD, Kotz SA, Sabisch B, Barry J, Richter N, et al. A locus for an auditory processing deficit and language impairment in an extended pedigree maps to 12p13.31-q14.3. *Genes, brain, and behavior*. 2010; 9:545–61.
- Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, et al. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *American journal of human genetics*. 2008; 82:150–9. [PubMed: 18179893]
- Alloway TP, Archibald L. Working memory and learning in children with developmental coordination disorder and specific language impairment. *Journal of learning disabilities*. 2008; 41:251–62. [PubMed: 18434291]
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV. 4th ed.. American Psychiatric Association; Washington, DC: 1994.
- Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, et al. A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *American journal of human genetics*. 2008; 82:160–4. [PubMed: 18179894]
- Baddeley A. Working memory and language: an overview. *J Commun Disord*. 2003; 36:189–208. [PubMed: 12742667]
- Baddeley A, Wilson BA. A developmental deficit in short-term phonological memory: implications for language and reading. *Memory*. 1993; 1:65–78. [PubMed: 7584260]
- Bakkaloglu B, O’Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, et al. Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *American journal of human genetics*. 2008; 82:165–73. [PubMed: 18179895]
- Barry JG, Yasin I, Bishop DV. Heritable risk factors associated with language impairments. *Genes Brain Behav*. 2007; 6:66–76. [PubMed: 17233642]
- Bartlett CW, Flax JF, Logue MW, Smith BJ, Vieland VJ, Tallal P, et al. Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Hum Hered*. 2004; 57:10–20. [PubMed: 15133308]
- Bartlett CW, Flax JF, Logue MW, Vieland VJ, Bassett AS, Tallal P, et al. A major susceptibility locus for specific language impairment is located on 13q21. *Am J Hum Genet*. 2002; 71:45–55. [PubMed: 12048648]
- Bishop DV. The role of genes in the etiology of specific language impairment. *Journal of Communication Disorders*. 2002; 35:311–28. [PubMed: 12160351]
- Bishop DV, Bishop SJ, Bright P, James C, Delaney T, Tallal P. Different origin of auditory and phonological processing problems in children with language impairment: evidence from a twin study. *Journal of Speech, Language, & Hearing Research*. 1999; 42:155–68.
- Bishop DV, Edmundson A. Language-impaired 4-year-olds: distinguishing transient from persistent impairment. *Journal of Speech & Hearing Disorders*. 1987a; 52:156–73. [PubMed: 3573746]
- Bishop DV, Edmundson A. Specific language impairment as a maturational lag: evidence from longitudinal data on language and motor development. *Dev Med Child Neurol*. 1987b; 29:442–59. [PubMed: 2445609]
- Bishop DV, North T, Donlan C. Genetic basis of specific language impairment: evidence from a twin study. *Developmental Medicine & Child Neurology*. 1995; 37:56–71. [PubMed: 7828787]
- Bolhuis JJ, Okanoya K, Scharff C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci*. 2010; 11:747–59. [PubMed: 20959859]
- Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet*. 2001; 2:91–9. [PubMed: 11253062]
- Catts HW, Fey ME, Tomblin JB, Zhang X. A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, & Hearing Research*. 2002; 45:1142–57.

- Colledge E, Bishop DV, Koeppen-Schomerus G, Price TS, Happe FG, Eley TC, et al. The structure of language abilities at 4 years: a twin study. *Developmental Psychology*. 2002; 38:749–57. [PubMed: 12220052]
- Conti-Ramsden G, Botting N, Durkin K. Parental perspectives during the transition to adulthood of adolescents with a history of specific language impairment (SLI). *Journal of speech, language, and hearing research : JSLHR*. 2008; 51:84–96.
- DeThorne LS, Petrill SA, Hayiou-Thomas ME, Plomin R. Low expressive vocabulary: higher heritability as a function of more severe cases. *J Speech Lang Hear Res*. 2005; 48:792–804. [PubMed: 16378474]
- Egan MF, Kojima M, Callicott JH, Goldberg TE, Kolachana BS, Bertolino A, et al. The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*. 2003; 112:257–69. [PubMed: 12553913]
- Gathercole SE, Hitch GJ, Service E, Martin AJ. Phonological short-term memory and new word learning in children. *Dev Psychol*. 1997; 33:966–79. [PubMed: 9383619]
- Gathercole SE, Willis CS, Baddeley AD, Emslie H. The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory*. 1994; 2:103–27. [PubMed: 7584287]
- Hammil, DD.; Brown, VL.; Larsen, SC.; Wiederhold, JL. Test of Adolescent Language-2. Pro-Ed.; Austin, TX: 1987.
- Hammil, DD.; Newcomer, PL. Test of Language Development-2, Intermediate. Pro-Ed; Austin, TX: 1988.
- Harel S, Greenstein Y, Kramer U, Yifat R, Samuel E, Nevo Y, et al. Clinical characteristics of children referred to a child development center for evaluation of speech, language, and communication disorders. *Pediatr Neurol*. 1996; 15:305–11. [PubMed: 8972529]
- Hariri AR, Goldberg TE, Mattay VS, Kolachana BS, Callicott JH, Egan MF, et al. Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance. *J Neurosci*. 2003; 23:6690–4. [PubMed: 12890761]
- Komulainen P, Pedersen M, Hanninen T, Bruunsgaard H, Lakka TA, Kivipelto M, et al. BDNF is a novel marker of cognitive function in ageing women: the DR's EXTRA Study. *Neurobiol Learn Mem*. 2008; 90:596–603. [PubMed: 18707012]
- Kovas Y, Hayiou-Thomas ME, Oliver B, Dale PS, Bishop DV, Plomin R. Genetic influences in different aspects of language development: the etiology of language skills in 4.5-year-old twins. *Child Dev*. 2005; 76:632–51. [PubMed: 15892783]
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*. 2001; 413:519–23. [PubMed: 11586359]
- Law J, Boyle J, Harris F, Harkness A, Nye C. Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord*. 2000; 35:165–88. [PubMed: 10912250]
- Leonard LB, Ellis Weismer S, Miller CA, Francis DJ, Tomblin JB, Kail RV. Speed of processing, working memory, and language impairment in children. *Journal of speech, language, and hearing research : JSLHR*. 2007; 50:408–28.
- Lewis BA, Shriberg LD, Freebairn LA, Hansen AJ, Stein CM, Taylor HG, et al. The genetic bases of speech sound disorders: evidence from spoken and written language. *Journal of speech, language, and hearing research : JSLHR*. 2006; 49:1294–312.
- Li X, Hu Z, He Y, Xiong Z, Long Z, Peng Y, et al. Association analysis of CNTNAP2 polymorphisms with autism in the Chinese Han population. *Psychiatric genetics*. 2010; 20:113–7. [PubMed: 20414140]
- Logan J, Petrill SA, Flax J, Justice LM, Hou L, Bassett AS, et al. Genetic covariation underlying reading, language and related measures in a sample selected for specific language impairment. *Behavior Genetics*. 2011; 41:651–9. [PubMed: 21193955]
- Misyak JB, Christiansen MH, Tomblin JB. On-line individual differences in statistical learning predict language processing. *Front Psychol*. 2010; 1:31. [PubMed: 21833201]
- Miyajima F, Ollier W, Mayes A, Jackson A, Thacker N, Rabbitt P, et al. Brain-derived neurotrophic factor polymorphism Val66Met influences cognitive abilities in the elderly. *Genes Brain Behav*. 2008; 7:411–7. [PubMed: 17973920]

- Montgomery JW, Evans JL. Complex sentence comprehension and working memory in children with specific language impairment. *J Speech Lang Hear Res.* 2009; 52:269–88. [PubMed: 18723601]
- Newbury DF, Fisher SE, Monaco AP. Recent advances in the genetics of language impairment. *Genome Med.* 2010; 2:6. [PubMed: 20193051]
- Newbury DF, Paracchini S, Scerri TS, Winchester L, Addis L, Richardson AJ, et al. Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behavior genetics.* 2011; 41:90–104. [PubMed: 21165691]
- Newbury DF, Winchester L, Addis L, Paracchini S, Buckingham LL, Clark A, et al. CMIP and ATP2C2 modulate phonological short-term memory in language impairment. *Am J Hum Genet.* 2009; 85:264–72. [PubMed: 19646677]
- Newcomer, PL.; Hammil, DD. *Test of Language Development-2, Primary.* Pro-Ed; Austin, TX: 1988.
- O’Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nature Genetics.* 2011; 43:585–9. [PubMed: 21572417]
- Paracchini S. Dissection of genetic associations with language-related traits in population-based cohorts. *Journal of neurodevelopmental disorders.* 2011; 3:365–73. [PubMed: 21894572]
- Pennington BF, Bishop DV. Relations among speech, language, and reading disorders. *Annual review of psychology.* 2009; 60:283–306.
- Peter B, Raskind WH, Matsushita M, Lisowski M, Vu T, Berninger VW, et al. Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. *J Neurodev Disord.* 3:39–49. [PubMed: 21484596]
- Peter B, Raskind WH, Matsushita M, Lisowski M, Vu T, Berninger VW, et al. 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.* 2011; 3:39–49. [PubMed: 21484596]
- Peterson RL, McGrath LM, Smith SD, Pennington BF. Neuropsychology and genetics of speech, language, and literacy disorders. *Pediatr Clin North Am.* 2007; 54:543–61. vii. [PubMed: 17543909]
- Pezawas L, Verchinski BA, Mattay VS, Callicott JH, Kolachana BS, Straub RE, et al. The brain-derived neurotrophic factor val66met polymorphism and variation in human cortical morphology. *J Neurosci.* 2004; 24:10099–102. [PubMed: 15537879]
- Poot M, van der Smagt JJ, Brilstra EH, Bourgeron T. Disentangling the Myriad Genomics of Complex Disorders, Specifically Focusing on Autism, Epilepsy, and Schizophrenia. *Cytogenetic and genome research.* 2011
- Rice ML, Smith SD, Gayan J. Convergent genetic linkage and associations to language, speech and reading measures in families of probands with Specific Language Impairment. *J Neurodev Disord.* 2009; 1:264–82. [PubMed: 19997522]
- Rice ML, Taylor CL, Zubrick SR. Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *J Speech Lang Hear Res.* 2008; 51:394–407. [PubMed: 18367685]
- Scerri TS, Morris AP, Buckingham LL, Newbury DF, Miller LL, Monaco AP, et al. DCDC2, KIAA0319 and CMIP Are Associated with Reading-Related Traits. *Biological psychiatry.* 2011; 70:237–45. [PubMed: 21457949]
- Semel, E.; Wiig, EH.; Secord, WA. *Clinical evaluation of language fundamentals, fourth edition (CELF-4).* The Psychological Corporation/A Harcourt Assessment Company; Toronto, Canada: 2003.
- Semel, E.; Wiig, EH.; Secord, WA. *Clinical evaluation of language fundamentals, fourth edition—Screening test (CELF-4 screening test).* The Psychological Corporation/A Harcourt Assessment Company; Toronto, Canada: 2004.
- Sham, PC. *Statistics in Human Genetics.* John Wiley & Sons, Inc.; New York, NY: 1998. p. 83
- Simmons TR, Flax JF, Azaro MA, Hayter JE, Justice LM, Petrill SA, et al. Increasing genotype-phenotype model determinism: application to bivariate reading/language traits and epistatic

- interactions in language-impaired families. *Human Heredity*. 2010; 70:232–44. [PubMed: 20948219]
- Snowling M, Bishop DV, Stothard SE. Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology & Psychiatry & Allied Disciplines*. 2000; 41:587–600.
- Stromswold K. Genetics of spoken language disorders. *Human Biology*. 1998; 70:297–324. [PubMed: 9549241]
- A genome-wide scan identifies two novel loci involved in specific language impairment. *Am J Hum Genet*. 2002; 70:384–98. [PubMed: 11791209]
- Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *Am J Hum Genet*. 2004; 74:1225–38. Epub 2004 May 3. [PubMed: 15133743]
- Threlkeld SW, Hill CA, Rosen GD, Fitch RH. Early acoustic discrimination experience ameliorates auditory processing deficits in male rats with cortical developmental disruption. *Int J Dev Neurosci*. 2009; 27:321–8. [PubMed: 19460626]
- Threlkeld SW, McClure MM, Bai J, Wang Y, LoTurco JJ, Rosen GD, et al. Developmental disruptions and behavioral impairments in rats following in utero RNAi of *Dyx1c1*. *Brain Res Bull*. 2007; 71:508–14. [PubMed: 17259020]
- Tomblin JB, Buckwalter PR. Heritability of poor language achievement among twins. *Journal of Speech, Language, & Hearing Research*. 1998; 41:188–99.
- Tomblin JB, Records NL, Buckwalter P, Zhang X, Smith E, O'Brien M. Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, & Hearing Research*. 1997; 40:1245–60.
- Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, et al. A functional genetic link between distinct developmental language disorders. *N Engl J Med*. 2008; 359:2337–45. [PubMed: 18987363]
- Villanueva P, Newbury DF, Jara L, De Barbieri Z, Mirza G, Palomino HM, et al. Genome-wide analysis of genetic susceptibility to language impairment in an isolated Chilean population. *European journal of human genetics : EJHG*. 2011; 19:687–95. [PubMed: 21248734]
- White SA, Fisher SE, Geschwind DH, Scharff C, Holy TE. Singing mice, songbirds, and more: models for FOXP2 function and dysfunction in human speech and language. *J Neurosci*. 2006; 26:10376–9. [PubMed: 17035521]
- Whitehouse AJ, Bishop DV, Ang QW, Pennell CE, Fisher SE. CNTNAP2 variants affect early language development in the general population. *Genes Brain Behav*. 10:451–6. [PubMed: 21310003]
- Whitehouse AJ, Bishop DV, Ang QW, Pennell CE, Fisher SE. CNTNAP2 variants affect early language development in the general population. *Genes, brain, and behavior*. 2011; 10:451–6.
- Wiig, EH.; Secord, WA.; Semel, E. Clinical evaluation of language fundamentals—Preschool, second edition (CELF Preschool-2). The Psychological Corporation/A Harcourt Assessment Company; Toronto, Canada: 2004.
- Willis CS, Gathercole SE. Phonological short-term memory contributions to sentence processing in young children. *Memory*. 2001; 9:349–63.
- Young AR, Beitchman JH, Johnson C, Douglas L, Atkinson L, Escobar M, et al. Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *J Child Psychol Psychiatry*. 2002; 43:635–45. [PubMed: 12120859]
- Zelaznik HN, Goffman L. Generalized motor abilities and timing behavior in children with specific language impairment. *Journal of speech, language, and hearing research : JSLHR*. 2010; 53:383–93.

Figure 1.

The human genetics paradigm for disease gene discovery. Linkage analysis (left) uses molecular genetic data (genotypes) from families and correlates the transmission of genotypes with the apparent transmission of language status or quantitative measures (phenotypes). Linkage peaks cover large genomic regions, often 20–30Mb, so alternative methods/design may be applied to narrow down the critical region. Association studies (middle) can be conducted using unrelated persons with the disease (cases) and without the disease (controls) where any allele frequencies differences indicate that allele is associated (correlated) with disease status. This method usually narrows down the critical region to one or few genes. Functional work (right) can be done to understand how, as in this example, gene expression is altered by an associated allele. Ultimately when functional work provides clear evidence that an allele can be used to separate cases from controls, or even subgroups of cases, use of that DNA biomarker information can be applied in the personalized medicine framework.

Table 1

Region	Gene	Phenotype	Year	N		Design
				Family	Cohort	
16q24		PSTM	2002	98	473	sib-pair QTL genome linkage scan (STR) (The SLI Consortium 2002)
		PSTM	2004	184	840	sib-pair QTL genome linkage scan (STR)(The SLI Consortium 2004)
		LI	2010	31	115	pedigree candidate region association (SNP)(Villanueva et al. 2011)
	ATP2C2	PSTM	2011	181	847+213	sib-pair/case-control QTL candidate region association (Newbury et al. 2011)
19q13	ATP2C2	PSTM	2009	211	806	sib-pair QTL candidate region association (Newbury et al. 2009)
	CMIP	PSTM	2009	211	806	sib-pair QTL candidate region association (Newbury et al. 2009)
		EL	2002	98	473	sib-pair QTL genome linkage scan (STR) (The SLI Consortium 2002)
		PSTM	2004	184	840	sib-pair QTL genome linkage scan (STR) (The SLI Consortium 2004)
13q21	*BDNF	RI	2002	5	73	pedigree categorical genome linkage scan (STR)(Bartlett et al. 2002)
		RI	2004	22	279	pedigree candidate region linkage/association(Bartlett et al. 2004)
		RI/LI	2011	34	125	pedigree candidate region linkage (Simmons et al. 2010)
		PSTM	2010	1	14	pedigree genome linkage scan(Addis et al. 2010)
7q	CNTNAP2	PSTM, EL, RL	2008	184	847	sib-pair QTL candidate gene association(Vernes et al. 2008)
	CNTNAP2	PSTM, RL	2011	181	847+213	sib-pair/case-control QTL candidate region association (Newbury et al. 2011)
	7q34-36	LI	2011	34	125	pedigree candidate region association (Villanueva et al. 2011)

PSTM: short term memory, EL: expressive language impairment, RL: receptive language impairment, RI: reading impairment, LI: language impairment

* BDNF is not a candidate in the 13q21 but was shown to epistatically interact with 13q21