

The Genetics of Specific Language Impairment

Ron Nudel, Fabiola Ceroni, Nuala H. Simpson & Dianne F. Newbury

University of Oxford

Abstract

Specific language impairment (SLI) is diagnosed when a child has major problems in learning to talk and/or understand language, despite showing normal development in all other areas (Bishop, 2006). In this chapter we will discuss recent research on the genetics of SLI. We will introduce methods used in human genetics and describe studies in which these methods were applied. We will discuss how emerging genetic technologies may be used to investigate the molecular basis of neurodevelopmental disorders and elucidate the biological mechanisms influencing linguistic traits.

Basic concepts in genetics

We each carry two sets of 23 chromosomes, one set from each of our parents. These consist of two copies of chromosomes 1-22 and two sex chromosomes (females carry two X chromosomes; males have one X chromosome and one Y chromosome). A gene is a physical stretch of DNA sequence that holds the instructions for the synthesis of a functional molecule, usually a protein. Each gene may have several variants in a population, called alleles. As we have two copies of each chromosome (chromosomes 1-22), for any given variant, we can carry two identical alleles (in which case we would be homozygous for said allele), or two different alleles (in which case we would be heterozygous).

Different alleles in a given gene do not necessarily have functional differences: the proteins they encode can be identical, or so similar as to have only subtle, functional differences. However, sometimes, an allele can lead to a dysfunctional protein, to various extents, and this may result in a disease or a disorder. A disorder is called recessive when it only affects people who carry two deleterious alleles in the relevant gene (or one, when the individual does not carry another allele), or dominant, if one deleterious allele is enough to cause it. In contrast, many common disorders, including SLI, are complex disorders. A complex disorder involves chance combinations of functional, but subtly different, alleles which combine across relevant biological pathways and interact with environmental factors. Different combinations confer different risks of developing the disorder. The inheritance pattern of complex disorders cannot be classified as dominant or recessive.

During the creation of the sex cells, homologous chromosomes (i.e., a pair of chromosomes of the same type) cross-over and exchange DNA between them (a process called recombination), so that when one of them is passed onto the child, it is not identical to either of the copies the parent has. Instead, the chromosomes that are passed on consist of mosaics of randomly defined “parental blocks”. The probability of two ancestral alleles that are on the same chromosome being inherited together depends upon the physical distance between them. The closer the genes (or, indeed any DNA features) are, the less likely they are to be separated by a cross-over. Traditional linkage-based methods of identifying disease-causing genes (discussed below) use information on recombination events which occurred in the pedigree in order to assess linkage.

Investigating the Heritability of SLI

Before genetic studies of SLI can be performed, one must first evaluate its heritability using familial aggregation and twin studies. Familial aggregation studies compare the incidence of a disorder in families containing an affected proband (the individual through whom the family was identified) with that in control families. However, familial aggregation studies alone cannot exclude environmental factors. This detail requires twin studies. In a twin study, the incidence of disorder in identical co-twins of probands (monozygotic (MZ) twin concordance) is compared with that in non-identical co-twins of probands (dizygotic (DZ) twin concordance). MZ twins are assumed to be 100% genetically similar, while DZ twins are assumed to be 50% similar. Thus, under the assumption that both twin types have a shared environment, if the MZ concordance is significantly higher than the DZ concordance, we can conclude that the disorder has a genetic component. Stromswold (1998 and 2001) reviewed several familial aggregation and twin studies of SLI and concluded that language disorders have a strong genetic component.

Methods in human genetics

If a disorder has a genetic component, the next step is to look for the contributory genes. We will discuss two traditional approaches to gene identification: linkage and association, and follow up with information on newer approaches. Both linkage and association methods involve the use of genetic markers, of which there are various kinds (e.g. microsatellites and Single-Nucleotide Polymorphisms (SNPs)). Markers provide points of reference in the DNA sequence allowing the tracking of parental blocks between individuals, and the identification of the genetic signature of an individual.

Linkage analyses aim to identify chromosome regions that co-segregate with a disorder within family units. Linkage methods can be divided into parametric and non-parametric approaches. Parametric analyses are more powerful, but require the specification of a genetic model of the disease. This includes the inheritance pattern (e.g. whether the disease is recessive or dominant), the expected allele frequencies in the population, and the penetrance, that is, the probability of expressing a trait (the phenotype) given the set of alleles a person has (the genotype) at a given position on the chromosome (locus). When these parameters are unknown, which is often the case with complex disorders, non-parametric linkage analyses provide an alternative. Results of linkage analyses are reported as Logarithm Of Odds (LOD) scores. The LOD was originally formulated as a function of the recombination frequency (which is correlated with the distance between the marker and the conditional place of the disease-causing gene) and was defined as the base-10 logarithm of the ratio between the probability of obtaining the observed data (in terms of the number of recombination events detected in the pedigree) assuming linkage at a distance defined by r and the probability of obtaining the data assuming no linkage. However, there are different ways to calculate LOD scores depending on the test used. The higher the LOD score, the more likely it is that the disease-causing gene is located near the marker, at a distance correlated with r in the tested hypothesis. Thus, LOD scores can be maximized over r to produce a maximum LOD score for each marker (i.e. one could test linkage at different distances from each marker to find the distance which would give the highest LOD score). A LOD score of 3 is usually required to claim significant linkage, but a higher score may be required, depending on the number of markers tested and the type of analysis used. Lower LOD scores (~ 2) are sometimes

reported as suggestive linkage. For a review of linkage methods, see Teare & Barrett (2005).

In an association study, one looks for a direct statistical correlation between a marker allele and a disease in a population (assuming that the disease is caused by the marker itself or a variant close to it). These days, case-control association cohorts usually consist of large numbers (thousands) of unrelated cases and controls. However, some association methods have been developed for family-based studies. Results of association studies are usually P values, or $-\log_{10}(P)$, when they are plotted, indicating whether a certain marker allele is significantly more frequent in cases than controls or *vice versa*. Association significance thresholds depend on the number of markers tested: as the number of markers increases, the required number of tests increases and the significance threshold becomes more stringent. Genome-wide association studies (GWAS) can involve in excess of a million markers, and therefore the accepted significance threshold is 5×10^{-8} . In general, association studies require a higher density of markers than linkage studies, but they allow a better resolution of the identified chromosome segments. For a review of association studies for complex traits, see McCarthy et al. (2008).

Whereas the methods described above use genetic markers to locate disease-causing genes, newer technologies allow researchers to obtain the sequence of a person's entire genome. A whole-genome sequence can provide much more information about variants that may be influencing a certain trait or increase the risk of having a disorder. In fact, an ambitious project called the "1000 Genomes Project" aims to produce a catalog of most of the genetic variants with a frequency of at least 1% in several human populations studied (1000 Genomes Project Consortium, 2012). Such information may have many applications. It could be used, for example, to obtain

frequencies of genetic variants in the general population, which can then be compared against a sample of individuals affected by some disorder. These data may also be used to find statistical correlations between different alleles at specific locations in the genome (i.e. the patterns in which they co-occur in different populations), which can increase the number of markers to be used in association studies in which only some of these genetic markers have been genotyped. That said, genome-wide sequencing may at times provide so much information that it becomes difficult to find the cause of a genetic disorder. Therefore, another approach, called exome sequencing, may be more appropriate in some studies. Exome sequencing entails sequencing only the parts of the genomes that contain genes (which make up about 1% of the genome). Apart from the fact that it is cheaper to sequence only the exome, compared with the entire genome, the information obtained through exome sequencing is more specific, as variants detected through exome sequencing (especially if they are in the coding regions of genes) may have a more substantial effect on the gene in which they are found. It is, therefore, a good method for identifying rare mutations that may be deleterious, as opposed to common variants that are used in association studies.

The next sections will discuss applications of the above methods to the study of the genetics of SLI.

Linkage on chromosomes 16 and 19

A study by the SLI Consortium (SLIC) (2002) included 98 nuclear families (consisting of a father, a mother, and their children), each with at least one child affected by SLI, and used non-parametric sib (sibling) pair methods. The children were assessed using the Clinical Evaluation of Language Fundamentals (CELF-R), the Wechsler Scales of Intelligence (WISC) and Non-Word Repetition (NWR). All

probands had Expressive and Receptive Language Scores (ELS and RLS, obtained from CELF-R) >1.5 standard deviations (SD) below the normative mean for their chronological age and a Performance IQ (PIQ) of >80 . For any given marker, sib-pairs are expected to share 0, 1 or 2 parental alleles with probabilities of 25%, 50% and 25%, respectively. However, if both sibs in a given pair are affected by the same disorder that is directly caused by a single gene, then the same ancestral chromosome region containing the disease-causing allele will always be shared between them. Thus, by identifying regions of the genome in which affected sib-pairs show increased similarity, we can narrow the search for candidate genes. In the case of complex disorders, we expect the regions containing major contributory genes to be shared between some affected sib-pairs, and this can be assessed statistically. The SLIC study (2002) did not use pairs of affected siblings, but instead performed quantitative analyses known as Haseman-Elston (HE) (Haseman & Elston, 1972) and Variance-Components (VC) methods (Pratt et al., 2000). The HE method implements a regression of the squared phenotype differences on the estimated IBD sharing (Identity By Descent – the property of two alleles being identical copies of the same ancestral allele) for a given sib-pair. The hypothesis is that if a marker is close to a gene influencing the trait, then there should be an inverse relation between the sib-pair trait differences and the IBD sharing at the marker locus (Haseman & Elston, 1972). The VC method separates the trait variability between siblings into a major gene variance component, a background genetic variance component, and an environmental variance component. It is a model-based method that uses family data and compares the likelihood of two models: one in which there is a major gene component and another without this effect. At a marker which is close to the gene influencing the trait, the first model will be most likely. In SLIC (2002), two

significant linkage regions were found: one on chromosome 16 (designated SLI1) with NWR (maximum LOD score of 3.55, HE), and another on chromosome 19 (designated SLI2) with ELS (maximum LOD score of 3.55, HE).

In their next study, SLIC (2004) increased their sample size to 184 families and looked for linkage in the regions previously implicated. They replicated the linkage between chromosome 16 and NWR performance, with a maximum LOD score of 7.46, using the combined sample, confirming the validity of the previous result. The linkage on chromosome 19 also replicated, but was linked to different traits: in one wave (the 2002 cohort) the trait was ELS, in the second wave (the 2004 sample) it was NWR. The maximum LOD score on chromosome 19 with NWR, taking into account both samples, was 1.4 (HE).

Further evidence for the linkage region on chromosomes 16 and 19

Falcaro et al. (2008) used an independent cohort of 93 nuclear families to try to replicate the SLIC linkage to chromosomes 16 and 19. They focused on NWR and a Past Tense marking trait (PT) but also had data available for ELS. Linkage was assessed using the HE method, as described for the SLIC studies, as well as another method that is based on the classic DeFries–Fulker approach (DeFries & Fulker, 1985). The DeFries–Fulker approach is usually applied to twin data to assess the heritability of a given trait. It is based on the idea that if a trait is genetically endowed then the performance of MZ and DZ co-twins of selected probands will regress to the population mean to different extents. The method used by Falcaro et al. (2008) is an extension of this algorithm that allows the detection of linkage in sib-pairs, using quantitative traits (Fulker et al., 1991), hereafter referred to as the DF-linkage method. The DF-linkage method assumes that a sibling's phenotypic score will regress back to

the population mean as a function of environmental and genetic effects (estimated by IBD sharing) shared with the proband. The authors found weak linkage with NWR on chromosome 16 (maximum LOD score of 1.69, DF-linkage) and significant linkage with ELS on chromosome 19 (maximum LOD score of 5.8, DF-linkage). When PT was measured as a continuous trait, some linkage was found on both chromosome 16 (maximum LOD score of 1.8, DF-linkage) and chromosome 19 (maximum LOD score of 2.2, DF-linkage). When PT was represented as a binary trait (affected or unaffected), linkage was found only on chromosome 19 (maximum LOD score of 1.66, HE). The authors argue that PT competence is either acquired, or not acquired, by a certain age, and it is therefore best analyzed qualitatively.

A targeted association study of chromosome 16

Newbury et al. (2009) performed a high density association screen of SLI across the chromosome 16 region of linkage using 211 SLIC families. They did a case-control analysis (as described previously) and used a family-based quantitative association method known as QTDT (Abecasis et al., 2000). QTDT is a regression-based method that tests for a significant and consistent correlation between a trait and the number of copies of a given marker allele carried by a child (taking into account the expected value of this number). For a review of this and similar methods see Ewens et al. (2008). The case-control analysis used a binary measure under which cases were defined as children with low NWR (>2 SD below the SLIC cohort mean, n=79). Controls were defined as children with above-average NWR performance (>0.5 SD above population mean, n=71). Only one case or one control was selected from each family to maintain independence between cases and controls. Markers falling in the *CMIP* (c-Maf inducing protein, $\text{minP}=5\times 10^{-7}$) and *ATP2C2* (ATPase, Ca^{++}

transporting, type 2C, member 2, $\text{minP}=2\times 10^{-5}$) genes were significantly associated ($P < 0.01$) with NWR. Both genes are expressed in the brain. *ATP2C2* is involved in calcium ion transport (Missiaen et al., 2007). Many signaling pathways use calcium as a messenger, and calcium regulation is crucial to various nerve cell functions and processes (Zheng & Poo, 2007). *CMIP* interacts with a member of the nuclear factor-kappaB family of molecules (Kamal et al., 2009), which are involved in synaptic activity and plasticity. Alterations of this pathway are found in several nervous system diseases, including Alzheimer's and Parkinson's (Mémet, 2006).

The associations with *CMIP* and *ATP2C2* were independent, meaning that both these genes could contribute to SLI susceptibility. Both associations were followed up in a replication sample, selected from the ALSPAC cohort (Jones et al., 2000) on the basis of low language measures (490 cases). ALSPAC is a population study that follows the development of ~14,000 children born in the southwest of England between 1991 and 1992. The same association techniques were applied but the case-control analyses used different case/control cut-offs, due to differences in the distributions of NWR between SLIC and ALSPAC. Regression trends for markers in *ATP2C2* were significant ($P < 0.05$) in the replication sample and replicated those observed by SLIC ($\text{minP}=0.0058$). However, the association with *CMIP*, although significant ($\text{minP}=0.0182$), was in the opposite direction, i.e., the marker alleles associated with high NWR scores in SLIC were associated with low NWR scores in the replication cohort. This could reflect differences in the relationship between the markers and the causal variant in the two populations (Lin et al., 2007), but may also indicate that this association is a false positive. The authors also investigated the relationship between *ATP2C2* and *CMIP* markers and NWR performance in the general population ($n=3612$) using the entire ALSPAC cohort (Jones et al., 2000), but found no evidence

for association. They suggest that these two genes affect NWR only in language-impaired individuals.

Extended-pedigree studies

The SLIC studies used small nuclear families. However, linkage studies traditionally use extended pedigrees, in which multiple affected individuals provide increased power to identify chromosome regions that segregate with a disease. Following this strategy, Bartlett et al. (2002) performed a linkage screen of five Canadian SLI families, two nuclear and three extended (86 individuals).

These families were identified during a linkage study of schizophrenia, when many individuals were noted to have severe language or reading deficits (Brzustowicz et al., 2000). Seven individuals in the SLI study were schizophrenic, but the largest pedigree (34 individuals) was connected to a schizophrenia family only by marriage. Thus, the authors hypothesize that these subjects should not influence the identification of SLI susceptibility loci.

73 subjects were assessed using several tests, including: an age-appropriate version of the Test of Language Development (TOLD), an age-appropriate Wechsler Intelligence Test and the Token Test, subtests from the Woodcock Reading Mastery Test and the Oral Speech Mechanism Screening Examination, and self- or parental-report questionnaires to assess history of hearing problems. Each family included at least two language-impaired individuals, with Spoken Language Quotient Standard Score (SLQ), taken from the TOLD, of $>1SD$ below expected, PIQ of ≥ 80 and PIQ \geq SLQ. Three binary diagnoses were derived for all family members: language impairment (SLQ ≤ 85); reading impairment (single non-word reading score of $>1SD$ below PIQ); clinical impairment (language or reading impairment or a self-reported

history of language difficulties or dyslexia). These diagnoses were not mutually exclusive.

Parametric linkage analyses were performed under two separate models, assuming either dominant or recessive inheritance, for each diagnostic category, giving a total of six analyses. The highest LOD scores were further characterized with a Bayesian statistical tool, called Posterior Probability of Linkage (PPL). PPL adjusts for the structure of the dataset in hand and allows for different effects between individuals (Vieland et al., 2001). It is therefore particularly suited to the study of complex disorders and provides an increased sensitivity to detect false positives. PPL values range from 0 to 1, but are commonly reported as percentages, which can be interpreted as the probability that a trait gene is present within the chromosome region, linked to a marker.

Significant evidence of linkage was found on chromosome 13 (LOD=3.92, recessive, reading). The maximum PPL was 0.53, suggesting a 53% chance that a risk factor for SLI falls in this region. Suggestive evidence for linkage was also detected on chromosomes 2 (LOD=2.79, recessive, language) and 17 (LOD=2.19, dominant, reading) with maximum PPL values of 5.8% and 3.6%, respectively. No linkage was found on chromosomes 16 or 19, in contrast to the results obtained by SLIC.

Further evidence for linkage to chromosome 13

Bartlett et al. (2004) added 22 nuclear and extended families (279 individuals from the United States) to their previous sample, in a targeted replication of chromosome 13. In this study, they also investigated chromosomes 2 and 7, which had previously been implicated in autism. The assessment tools, criteria for diagnosis and model parameters were the same as those described above. Both sample sets were combined

for PPL analysis, and an additional linkage method known as the heterogeneity LOD score method (HLOD) was used. The HLOD takes into account the fact that several loci can contribute to a complex disorder. Two HLOD variants were generated: HLOD-P, which is calculated from pooled datasets, and HLOD-S, which calculates HLOD scores for each sample separately and then sums them across datasets.

Weak evidence for linkage to chromosomes 2 and 7 was observed, but chromosome 13 yielded evidence of linkage in each separate sample set (US sample HLOD=2.616, PPL=16.8%; Canadian sample HLOD=3.565, PPL=54.2%) and in the combined sample (HLOD-P=6.031; HLOD-S=6.181, PPL=92.3%), under the recessive reading impairment model. This region has been designated SLI3.

Robinson Crusoe Island study

Extended pedigrees were also used in the unique case of the Robinson Crusoe Island population (Villanueva et al., 2011). The Robinson Crusoe Island is located 677 km west of Chile and is the only inhabited island in the Juan Fernández archipelago, with 633 residents (based on the 2002 Chilean census). The island was colonized in the late 19th century by eight families, and 77% of the current population have a colonizing surname (Villanueva et al., 2008). The prevalence of SLI among island children related to one of the founder families is very high (35%), compared to that among children who are not (3.8%), the latter being close to that reported in mainland Chile (4%) (Villanueva et al., 2008). Given the structure of this population, Villanueva et al. (2011) hypothesized that the disorder on the island could be the result of a mutation in a single gene. They performed non-parametric and parametric (under different models) linkage analyses. Several chromosome regions (on chromosomes 6, 7, 12, 13 and 17) reached genome-wide significance. No linkage to chromosomes 16 or 19 was

observed, but the region on chromosome 13 was close to SLI3. The most consistently linked region across analyses was a 48 Mb (megabases, i.e. 1000000 base pairs of DNA) region on chromosome 7. This overlaps with a chromosome region that has been linked to autism and contains several interesting candidate genes, including *FOXP2* and *CNTNAP2*, discussed in more detail below. Further studies involving additional markers will be required to ascertain which of the genes in this region contribute to the language disorder on the island.

A comparison between the different linkage studies of SLI

The linkage studies described above differ in several ways. SLIC performed non-parametric analyses, which do not require a specification of a model, using nuclear families; the Bartlett studies applied a parametric approach to a small number of large families, and the Robinson Crusoe study used both parametric and non-parametric methods with an extremely large pedigree. Each approach has advantages and disadvantages, and these design differences may explain much of the lack of overlap in results. Parametric models provide increased power, if the model parameters are correctly specified. Although mis-specification of parameters can lead to false positives, previous studies (Greenberg et al., 1998; Abreu et al., 1999) have demonstrated that parametric analyses have sufficient power to detect genomic regions linked to complex disorders. Non-parametric methods are affected by the distribution of marker allele frequencies and the traits. Discrepant results may also reflect the way in which traits are measured, especially for disorders such as SLI, which has no standardized diagnostic criteria. Bartlett and Villanueva used binary categories of language impairment by applying an arbitrary threshold for affection. SLIC considered a continuous distribution of language abilities using standard scores,

calibrated against the general population. Three quantitative traits (ELS, RLS and NWR), with strong genetic components, were used. Thus, it is possible that these alternative aspects of language are under the influence of different genes. The linkage regions on chromosomes 16 and 19 (SLIC, 2002) were linked to NWR and ELS, respectively, while chromosome 7 (Villanueva et al., 2011) was linked to language impairment, and chromosome 13 (Bartlett et al., 2002) was linked to reading impairment. Reading impairment was not interpreted as a dyslexia phenotype, but as a consequence of language deficits, since failure in phonological processing is often observed in SLI (Bishop & Snowling, 2004). These studies indicate that the investigation of alternative traits may be useful in dissecting the genetic susceptibility in SLI.

Comorbidity of SLI with other neurodevelopmental disorders

In SLI, language impairment occurs as a specific phenotype, in the absence of any explanatory medical conditions. Nevertheless, delayed language and deficits in communication are often present in developmental conditions with a strong genetic component, including Autism Spectrum Disorders (ASDs) and Attention Deficit Hyperactivity Disorder (ADHD).

ASDs are a group of neurodevelopmental disorders, autism being the most severe form. Family studies suggest that ASDs can be dissected into three heritable, potentially distinct, core components (social interaction, language, and repetitive behavior), or ‘endophenotypes’, determined in a genetically complex manner.

In autism, language abilities can be extremely varied. About 50% of autistic children do not develop any verbal language and others usually show a delayed development of spoken language. The most evident linguistic deficit in autistic individuals is an

inappropriate use of language, but structural aspects (phonology and syntax) can be affected, in a way similar to that in SLI (e.g. Bishop, 2003; Tager-Flusberg & Joseph, 2003; Rapin & Dunn, 2003). First and second-degree relatives of autistic patients display a higher prevalence of language impairment than the general population. These similarities have stirred a debate regarding the diagnostic boundaries between SLI and ASD (as reviewed by Taylor et al., 2012) leading to the hypothesis that SLI and autism might share some genetic susceptibility factors. The alternative hypothesis argues that differing patterns of linguistic performance may reflect alternative, distinct causes.

Genetic overlaps between SLI and autism

Several whole-genome linkage studies have been performed for autism, but replicated regions between samples are rare. The most consistently linked loci are on chromosomes 2 (designated AUTS5) and 7 (designated AUTS1) (IMGSAC 1998, IMGSAC 2001). A possible strategy to understand the divergence at both the phenotypic and genetic levels and to increase the chances of identifying contributory genetic variants is the study of single endophenotypes.

One study focusing on language-related endophenotypes in autism (Bradford et al., 2001) found linkage to chromosomes 7 and 13, the latter overlapping with that identified by Bartlett et al. (2002) in SLI families. Both signals were strongest in families with ASD probands with Phrase Speech Delay (PSD) beyond 36 months of age, and parents with a history of language difficulties. In another ASD linkage screen, three endophenotypes from the Autism Diagnostic Interview (ADI) were examined: "age at first word", "age at first phrase" and "repetitive and stereotyped behavior" (Alarcón et al., 2002). The strongest evidence of linkage was obtained for

"age at first word", on chromosome 7, close to that reported by IMGSAC (1998).

The other main linkage locus found by the IMGSAC (2001), on chromosome 2, has also been studied in specific categories of autism sub-phenotypes. One study (Buxbaum et al., 2001), found that linkage in this region was strongest in families which included two or more autistic children with PSD beyond 36 months of age. Subsequently, Shao et al. (2002) also detected an increase in linkage signals on chromosome 2, when they stratified their sample by PSD. Note, however, that this region has not been implicated in SLI studies. This does not exclude the possibility that this locus harbors a risk factor for language impairment, but may instead reflect the different language categories used in other disorders.

The chromosome 19 locus, found to be linked to ELS in the SLIC study (2002), also overlaps with an autism linkage region (Liu et al., 2001). This linkage was driven by a strictly-defined group of autistic families, in which the probands had impairments in all three core domains, and an age at onset of <3 years. One would expect that the greatest overlap with SLI would be found in the less severely affected families. However, these families include Asperger patients, who are characterized by relatively high linguistic capabilities. Thus, it remains possible that variants that contribute to susceptibility to SLI and autism may be found on chromosome 19.

More recently, linkage studies with higher resolution and larger cohorts of ASD families (AGP et al., 2007; Weiss et al., 2009) have been carried out, but the linkage regions described above did not reach genome-wide significance. In the study realized by the Autism Genome Project (AGP, 2007), in order to create more homogeneous categories, the samples were stratified in categories, but, even in these subsets, the loci on 7q and 2q reached only suggestive evidence of linkage, confirming the genetic heterogeneity underlying the ASDs. Genome-wide association studies (Weiss et al.,

2009; Wang et al., 2009, Anney et al., 2010) have been also carried out for autism, but the common variants identified explain only a small fraction of the genetic heritability of this disorder. On the other hand, a growing body of evidence suggests that genetic contribution to ASDs includes *de novo* or rare inherited variants of intermediate-high penetrance (as reviewed by Devlin & Scherer, 2012).

Recent high-throughput CNV screenings and exome-sequencing studies have begun to uncover a large number of individually rare variants, sequence mutations and structural rearrangements (CNVs, explained in the next section), that might contribute to ASDs, implicating hundreds of autism risk genes (Devlin & Scherer, 2012). The next challenge will be establishing which of the potentially deleterious variants actually play a role in the disease: investigating whether they affect genes involved in interconnected pathways that influence synaptogenesis and related biological processes, will narrow down the list of candidate genes and clarify the molecular mechanisms impaired in ASDs.

Exome-sequencing or CNV studies for SLI have not been published yet, but these kind of studies will help to elucidate the genetic architecture of SLI, establishing the contribution of rare variants and structural rearrangements and determining the genes involved in the disorder. Network analyses might provide new evidence of potential genetic overlaps between SLI and other neurodevelopmental conditions like ASDs.

Copy Number Variants (CNVs)

It has recently been shown that a large part of human variation does not involve changes to the genetic sequence *per se*, but rather structural rearrangements or submicroscopic copy number variations (CNVs) (Sebat et al., 2004; Redon et al., 2006). CNVs can involve the loss (deletions) or gain (duplications) of DNA segments

or can be more complex. They can be inherited from a parent or may be *de novo* if they originate in the parent germ line or during early fetal development. CNVs can directly disrupt genes or regulatory elements, or alter the number of gene copies, leading to phenotypic consequences. However, CNVs affecting the same gene are associated with different disorders and found in unaffected individuals (as reviewed by Cook & Scherer, 2008). Presumably, some events may have lower impact due to their location in the gene or their consequences may be modified by additional genetic variants, or genetic background. Although they are not yet fully understood, the study of CNVs has helped to highlight biologically relevant genetic pathways in neurodevelopmental disorders. They play an important role in susceptibility to many complex disorders, including: autism, schizophrenia and ADHD (Cook & Scherer, 2008).

An example of potential genetic overlap between autism and SLI given by CNVs can be found in a study of sporadic autistic trios, (i.e. mother, father and proband from a family in which the parents and siblings are all unaffected), where a deletion in *ATP2C2*, a candidate gene for SLI discussed in a previous section, was found in an autistic child. However, this reported deletion was inherited from an unaffected mother (O'Roak et al., 2012). Similarly, a *de novo* deletion of two genes on chromosome 16, including the SLI candidate gene *CMIP*, was identified in an autistic child with severe receptive and expressive language deficits (Van der Aa et al., 2012).

Genetic overlap between SLI and ADHD

ADHD is a common childhood neuropsychiatric disorder, characterized by a triad of impairments (inappropriate inattentiveness, increased impulsivity, and hyperactivity), (DSM IV-TR - (American Psychiatric Association (APA) 2000)). Although ADHD is

highly variable among individuals, patients often have deficits in executive and self-regulatory functions, such as: temporal organization, holding and manipulating information, which require working memory (Barkley, 1997) and often exhibit language impairments (Tirosh & Cohen, 1998).

Genome wide association studies of ADHD (Neale et al., 2008; Mick et al., 2010; Neale et al., 2010; Stergiakouli et al., 2012) have failed to detect statistically significant association results for any SNP. However, sub-threshold signals might indicate potentially relevant loci. A GWAS of ADHD (Lesch et al., 2008) detected association with *ATP2C2* ($P=8 \times 10^{-7}$), among the top 30 markers located in gene regions. Interestingly, considering the fact that the SLIC *ATP2C2* association was with NWR, studies report that reduced performance in working memory tasks are correlated with verbal deficits in children with ADHD (Cohen et al., 2000). Therefore, impairments in working memory have been proposed as a shared mechanism between ADHD and SLI. The SLIC studies did not exclude individuals with ADHD, and a high proportion of their samples were predicted to have ADHD, developmental coordination disorder, or reading problems (Newbury et al., 2009). The causal variants in *ATP2C2* have yet to be identified, but these findings support the idea that *ATP2C2* may be involved in neurological processes important for phonological short-term memory and may be relevant to developmental disorders that involve working memory deficits, such as SLI and ADHD.

Genetic pathways

There is a growing body of evidence that shared or related genetic mechanisms exist between disorders such as SLI, autism and ADHD. Given this, researchers have begun studying genetic risk factors at the level of the biological pathway rather than

specific candidate genes. Under this hypothesis, related neurodevelopmental disorders are caused by combinations of genetic risk variants across overlapping pathways that serve the developmental mechanisms important for brain development. Support for this hypothesis has been largely driven by investigations into autism and learning disability, and by studies of biological pathways mediated by *FOXP2* (Fisher & Scharff, 2009).

FOXP2* and *CNTNAP2

Mutations disrupting *FOXP2* cause a rare and severe speech and language disorder (Lai et al., 2001). The *FOXP2* gene encodes a transcription factor protein which regulates the expression of other genes (referred to as its targets) by binding to specific DNA sequences in their regulatory regions. *FOXP2* regulates hundreds of genes, many of which are known to be important during brain development (Vernes et al., 2007; Spiteri et al., 2007). One such gene is called *CNTNAP2*, which codes for a protein called CASPR2 and is found on chromosome 7 (Vernes et al., 2008). The CASPR2 protein belongs to the neurexin family, whose members establish interactions between brain cells during nervous system development. In the human fetal brain, *CNTNAP2* has an unusual enrichment in regions that are important for language e.g., the prefrontal cortex (Abrahams et al., 2007). Levels of *CNTNAP2* are highest at low *FOXP2* levels and *vice versa*, indicating that *FOXP2* negatively regulates *CNTNAP2*.

Variants in *FOXP2* itself are not thought to contribute to SLI (Newbury et al., 2002), but markers in *CNTNAP2* have been shown to be associated with language impairment. 184 SLIC families were used in a quantitative association analysis with ELS, RLS and NWR (Vernes et al., 2008). Several markers showed association,

primarily to NWR ($\text{minP}=5\times 10^{-5}$) and, to a lesser extent, to RLS ($\text{minP}=0.003$) and ELS ($\text{minP}=0.008$). Markers in *CNTNAP2* have also been associated with endophenotypes of early language development in the general population ($\text{minP}=0.0239$) (Whitehouse et al., 2011) and with "age at first word" in an autistic cohort ($\text{minP}=0.002$) (Alarcón et al., 2008). Complex chromosomal rearrangements and CNVs, in particular deletions, involving *CNTNAP2*, have also been reported in studies of other neurodevelopmental disorders including autism (Bakkaloglu et al., 2008; Poot et al., 2010), stuttering (Petrin et al., 2010), ADHD (Elia et al., 2010), Tourette's syndrome (Verkerk et al., 2003), schizophrenia and epilepsy (Friedman et al., 2008, Sehested et al., 2010) and mental retardation (Zweier et al., 2009), supporting the idea of shared pathways between related disorders. The conditions related to *CNTNAP2* mutations are numerous and variable, but they present a certain degree of comorbidity, perhaps representing shared or intersected neuronal mechanisms that converge on common genes like *CNTNAP2*.

A loss-of-function mutation, which is predicted to prevent the production of functional *CNTNAP2* protein, was found in children from an isolated population (Old Order Amish), affected by a syndrome associated with ASDs, called Cortical Dysplasia-Focal Epilepsy syndrome (CDFE) (Strauss et al., 2006). Screening of non-syndromic ASD patients also found mutations which altered the sequence of *CNTNAP2*, but did not prevent the production of the protein (Bakkaloglu et al., 2008). These changes were not present in controls, but were always inherited from an apparently unaffected parent. Thus, as discussed with regards to CNVs above, it is likely that some alterations of *CNTNAP2* may need to occur in conjunction with mutations in related genes to cause neurological dysfunction. The outcome is likely to be determined by the molecular pathways affected and the genetic and environmental

backgrounds of subjects. A recent exome-sequencing study identified a *CNTNAP2* mutation that is predicted to be damaging in an autistic child, which supports this hypothesis (O’Roak et al., 2011). In this study, the mutation was inherited from the mother, and was present in an unaffected sister. The exome sequence of the proband, who was severely affected with regression, language delay, moderate intellectual disability, revealed an additional deleterious mutation in *FOXP1* that was absent in the rest of the family (O’Roak et al., 2011). Further experiments are required to clarify the role of *FOXP1*, but it is known to be closely related to *FOXP2*, and has been implicated in speech and language impairments and intellectual disability (reviewed by Bacon & Rappold, 2012). Furthermore, like *FOXP2*, the expression levels of *FOXP1* are inversely correlated with those of *CNTNAP2* (O’Roak et al., 2011).

Taken together, these findings support the hypothesis of a widespread effect of *CNTNAP2*. New technologies and the falling costs of high-throughput sequencing mean that the identification of similar candidate genes will become increasingly more common.

Conclusions

SLI is a complex disorder with a strong genetic component. Therefore, the study of the genetics of SLI can complement SLI research done in other disciplines: behavioral studies of SLI are important in order to better characterize the disorder itself and find differences between and similarities with other developmental disorders. Genetic studies of SLI require good classifications of disease status and/or appropriate (linguistic) measures in order to be meaningful (see Figure 1). When those are available, genetic studies may find similarities between disorders at a molecular level,

which otherwise might not have been discovered. Recent studies have allowed the identification of several genes that may be involved in SLI, and, indeed, some of these genes have also been implicated in other neurodevelopmental disorders in which language may be impaired, such as ASD and ADHD. We believe that a good choice of endophenotypes and the application of new technologies hold the key to understanding the genetic basis of SLI. Emerging links between comorbid developmental disorders suggest the existence of overlapping biological pathways. Researchers continue to study these genes and biological pathways in an attempt to further elucidate the molecular mechanisms that are impaired in SLI and related disorders.

References

- 1000 Genomes Project Consortium, Abecasis G.R., Auton A., Brooks L.D., DePristo M.A., Durbin R.M., Handsaker R.E., Kang H.M., Marth G.T., McVean G.A. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56-65.
- Abecasis G.R., Cardon L.R., Cookson W.O. 2000. A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66(1):279-292.
- Abrahams B.S., Tentler D., Perederiy J.V., Oldham M.C., Coppola G., Geschwind D.H. 2007. Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc Natl Acad Sci U S A.* 6;104(45):17849-17854.
- Abreu P.C., Greenberg D.A., Hodge S.E. 1999. Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. *Am J Hum Genet* 65(3):847-857.
- Autism Genome Project Consortium, Szatmari P., Paterson A.D., Zwaigenbaum L., Roberts W., Brian J., Liu X.Q., Vincent J.B., Skaug J.L., Thompson A.P., Senman L., Feuk L., Qian C., Bryson S.E., Jones M.B., Marshall C.R., Scherer S.W., Vieland V.J., Bartlett C., Mangin L.V., Goedken R., Segre A., Pericak-Vance M.A., Cuccaro M.L., Gilbert J.R., Wright H.H., Abramson R.K., Betancur C., Bourgeron T., Gillberg C., Leboyer M., Buxbaum J.D., Davis K.L., Hollander E., Silverman J.M., Hallmayer J., Lotspeich L., Sutcliffe J.S., Haines J.L., Folstein S.E., Piven J., Wassink T.H., Sheffield V., Geschwind D.H., Bucan M., Brown W.T., Cantor R.M., Constantino J.N., Gilliam T.C., Herbert M., Lajonchere C., Ledbetter D.H., Lese-Martin C., Miller J., Nelson S., Samango-Sprouse C.A., Spence S., State M., Tanzi R.E., Coon H., Dawson G., Devlin B., Estes A.,

Flodman P., Klei L., McMahon W.M., Minshew N., Munson J., Korvatska E., Rodier P.M., Schellenberg G.D., Smith M., Spence M.A., Stodgell C., Tepper P.G., Wijisman E.M., Yu C.E., Rogé B., Mantoulan C., Wittemeyer K., Poustka A., Felder B., Klauck S.M., Schuster C., Poustka F., Bölte S., Feineis-Matthews S., Herbrecht E., Schmötzer G., Tsiantis J., Papanikolaou K., Maestrini E., Bacchelli E., Blasi F., Carone S., Toma C., Van Engeland H., de Jonge M., Kemner C., Koop F., Langemeijer M., Hijmans C., Staal W.G., Baird G., Bolton P.F., Rutter M.L., Weisblatt E., Green J., Aldred C., Wilkinson J.A., Pickles A., Le Couteur A., Berney T., McConachie H., Bailey A.J., Francis K., Honeyman G., Hutchinson A., Parr J.R., Wallace S., Monaco A.P., Barnby G., Kobayashi K., Lamb J.A., Sousa I., Sykes N., Cook E.H., Guter S.J., Leventhal B.L., Salt J., Lord C., Corsello C., Hus V., Weeks D.E., Volkmar F., Tauber M., Fombonne E., Shih A., Meyer K.J. 2007. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 39(3):319-328.

Alarcón M., Cantor R.M., Liu J., Gilliam T.C., Geschwind D.H., Autism Genetic Research Exchange Consortium. 2002. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet* 70(1):60-71.

Alarcón M., Abrahams B.S., Stone J.L., Duvall J.A., Perederiy J.V., Bomar J.M., Sebat J., Wigler M., Martin C.L., Ledbetter D.H., Nelson S.F., Cantor R.M., Geschwind D.H. 2008. Linkage, association, and gene-expression analyses identify *CNTNAP2* as an autism-susceptibility gene. *Am J Hum Genet* 82(1):150-159.

American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition, revised. American Psychiatric Press: Washington, DC.

Anney R., Klei L., Pinto D., Regan R., Conroy J., Magalhaes T.R., Correia C., Abrahams B.S., Sykes N., Pagnamenta A.T., Almeida J., Bacchelli E., Bailey A.J., Baird G., Battaglia A., Berney T., Bolshakova N., Bölte S., Bolton P.F., Bourgeron T., Brennan S., Brian J., Carson A.R., Casallo G., Casey J., Chu S.H., Cochrane L., Corsello C., Crawford E.L., Crossett A., Dawson G., de Jonge M., Delorme R., Drmic I., Duketis E., Duque F., Estes A., Farrar P., Fernandez B.A., Folstein S.E., Fombonne E., Freitag C.M., Gilbert J., Gillberg C., Glessner J.T., Goldberg J., Green J., Guter S.J., Hakonarson H., Heron E.A., Hill M., Holt R., Howe J.L., Hughes G., Hus V., Iglizzi R., Kim C., Klauck S.M., Kolevzon A., Korvatska O., Kustanovich V., Lajonchere C.M., Lamb J.A., Laskawiec M., Leboyer M., Le Couteur A., Leventhal B.L., Lionel A.C., Liu X.Q., Lord C., Lotspeich L., Lund S.C., Maestrini E., Mahoney W., Mantoulan C., Marshall C.R., McConachie H., McDougle C.J., McGrath J., McMahon W.M., Melhem N.M., Merikangas A., Migita O., Minshew N.J., Mirza G.K., Munson J., Nelson S.F., Noakes C., Noor A., Nygren G., Oliveira G., Papanikolaou K., Parr J.R., Parrini B., Paton T., Pickles A., Piven J., Posey D.J., Poustka A., Poustka F., Prasad A., Ragoussis J., Renshaw K., Rickaby J., Roberts W., Roeder K., Roge B., Rutter M.L., Bierut L.J., Rice J.P., Salt J., Sansom K., Sato D., Segurado R., Senman L., Shah N., Sheffield V.C., Soorya L., Sousa I., Stoppioni V., Strawbridge C., Tancredi R., Tansey K., Thiruvahindrapduram B., Thompson A.P., Thomson S., Tryfon A., Tsiantis J., Van Engeland H., Vincent J.B., Volkmar F., Wallace S., Wang K., Wang Z., Wassink T.H., Wing K., Wittemeyer K., Wood S., Yaspan B.L., Zurawiecki D., Zwaigenbaum L., Betancur C., Buxbaum J.D., Cantor R.M., Cook E.H., Coon H., Cuccaro M.L., Gallagher L., Geschwind D.H., Gill M., Haines J.L., Miller J., Monaco A.P., Nurnberger J.I. Jr, Paterson A.D., Pericak-Vance M.A.,

- Schellenberg G.D., Scherer S.W., Sutcliffe J.S., Szatmari P., Vicente A.M., Vieland V.J., Wijisman E.M., Devlin B., Ennis S., Hallmayer J. 2010. A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet* 19(20):4072-4082.
- Bacon C.& Rappold G.A. 2012. The distinct and overlapping phenotypic spectra of *FOXP1* and *FOXP2* in cognitive disorders. *Hum Genet* 131(11):1687-1698.
- Bakkaloglu B., O'Roak B.J., Louvi A., Gupta A.R., Abelson J.F., Morgan T.M., Chawarska K., Klin A., Ercan-Sencicek A.G., Stillman A.A., Tanriover G., Abrahams B.S., Duvall J.A., Robbins E.M., Geschwind D.H., Biederer T., Gunel M., Lifton R.P., State M.W. 2008. Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *Am J Hum Genet* 82(1):165-173.
- Barkley R.A. 1997. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull* 121(1):65-94.
- Bartlett C.W., Flax J.F., Logue M.W., Vieland V.J., Bassett A.S., Tallal P., Brzustowicz L.M. 2002. A major susceptibility locus for specific language impairment is located on 13q21. *Am J Hum Genet* 71(1):45-55.
- Bartlett C.W., Flax J.F., Logue M.W., Smith B.J., Vieland V.J., Tallal P., Brzustowicz L.M. 2004. 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* 57(1):10-20.
- Bishop, D. V. 2003. Autism and specific language impairment: Categorical distinction or continuum? *Novartis Foundation Symposium* 251, 213–226.
- Bishop D.V. & Snowling M.J. 2004 Developmental dyslexia and specific language

- impairment: same or different? *Psychol Bull.* 130(6):858-886.
- Bishop D.V. 2006. What causes specific language impairment in children? *Curr Dir Psychol Sci* 15(5):217-221.
- Bradford Y., Haines J., Hutcheson H., Gardiner M., Braun T., Sheffield V., Cassavant T., Huang W., Wang K., Vieland V., Folstein S., Santangelo S., Piven J. 2001. Incorporating language phenotypes strengthens evidence of linkage to autism. *Am J Med Genet* 105(8):539-547.
- Brzustowicz L.M., Hodgkinson K.A., Chow E.W., Honer W.G., Bassett A.S. 2000. Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* 288(5466):678-682.
- Buxbaum J.D., Silverman J.M., Smith C.J., Kilifarski M., Reichert J., Hollander E., Lawlor B.A., Fitzgerald M., Greenberg D.A., Davis K.L. 2001. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am J Hum Genet* 68(6):1514-1520.
- Cohen N.J., Vallance D.D., Barwick M., Im N., Menna R., Horodezky N.B., Isaacson L. 2000. The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. *J Child Psychol Psychiatry* 41(3):353-362.
- Cook E.H. Jr. & Scherer S.W. 2008. Copy-number variations associated with neuropsychiatric conditions. *Nature* 455(7215):919-923.
- DeFries J.C. & Fulker D.W. 1985. Multiple regression analysis of twin data. *Behav Genet* 15(5):467-473.
- Devlin B. & Scherer S.W. 2012. Genetic architecture in autism spectrum disorder. *Curr Opin Genet Dev* 22(3):229-237.

Elia J., Gai X., Xie H.M., Perin J.C., Geiger E., Glessner J.T., D'arcy M., deBerardinis R., Frackelton E., Kim C., Lantieri F., Muganga B.M., Wang L., Takeda T., Rappaport E.F., Grant S.F., Berrettini W., Devoto M., Shaikh T.H., Hakonarson H., White P.S. 2010. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol Psychiatry* 15(6):637-646.

Ewens W.J., Li M., Spielman R.S. 2008. A review of family-based tests for linkage disequilibrium between a quantitative trait and a genetic marker. *PLoS Genet* 4(9):e1000180.

Falcaro M., Pickles A., Newbury D.F., Addis L., Banfield E., Fisher S.E., Monaco A.P., Simkin Z., Conti-Ramsden G.; SLI Consortium 2008. Genetic and phenotypic effects of phonological short-term memory and grammatical morphology in specific language impairment. *Genes Brain Behav* 7(4):393-402.

Fisher S.E. & Scharff C. 2009. *FOXP2* as a molecular window into speech and language. *Trends Genet* 25(4):166-177.

Friedman J.I., Vrijenhoek T., Markx S., Janssen I.M., van der Vliet W.A., Faas B.H., Knoers N.V., Cahn W., Kahn R.S., Edelmann L., Davis K.L., Silverman J.M., Brunner H.G., van Kessel A.G., Wijmenga C., Ophoff R.A., Veltman J.A. 2008. *CNTNAP2* gene dosage variation is associated with schizophrenia and epilepsy. *Mol Psychiatry* 13(3):261-266

Fulker D.W., Cardon L.R., DeFries J.C., Kimberling W.J., Pennington B.F., Smith S.D. 1991. Multiple regression analysis of sib-pair data on reading to detect quantitative trait loci. *Reading Writing Interdiscip J* 3(3):299-313.

Greenberg D.A., Abreu P., Hodge S.E. 1998. The power to detect linkage in complex

- disease by means of simple LOD-score analyses. *Am J Hum Genet* 63(3):870-879.
- Haseman J.K. & Elston R.C. 1972. The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet.* 2(1):3-19.
- International Molecular Genetic Study of Autism Consortium (IMGSAC). 1998. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum Mol Genet* 7(3):571-578.
- International Molecular Genetic Study of Autism Consortium (IMGSAC). 2001. A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet* 69(3):570-581.
- Jones R.W., Ring S., Tyfield L., Hamvas R., Simmons H., Pembrey M., Golding J., ALSPAC Study Team. 2000. A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). *Eur J Hum Genet* 8(9):653-660.
- Kamal M., Valanciute A., Dahan K., Ory V., Pawlak A., Lang P., Guellaen G., Sahali D. 2009. C-mip interacts physically with RelA and inhibits nuclear factor kappa B activity. *Mol Immunol* 46(5):991-998.
- 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.
- Lesch K.P., Timmesfeld N., Renner T.J., Halperin R., Röser C., Nguyen T.T., Craig D.W., Romanos J., Heine M., Meyer J., Freitag C., Warnke A., Romanos M., Schäfer H., Walitza S., Reif A., Stephan D.A., Jacob C. 2008. Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. *J Neural Transm* 115(11):1573-1585.

- Lin P.I., Vance J.M., Pericak-Vance M.A., Martin E.R. 2007. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* 80(3):531-538.
- Liu J., Nyholt D.R., Magnussen P., Parano E., Pavone P., Geschwind D., Lord C., Iversen P., Hoh J., Ott J., Gilliam T.C., Autism Genetic Resource Exchange Consortium. 2001. A genomewide screen for autism susceptibility loci. *Am J Hum Genet* 69(2):327-40.
- McCarthy M.I., Abecasis G.R., Cardon L.R., Goldstein D.B., Little J., Ioannidis J.P., Hirschhorn J.N. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 9(5):356-369.
- Mémet S. 2006. NF-kappaB functions in the nervous system: from development to disease. *Biochem Pharmacol* 72(9):1180-1195.
- Mick E., Todorov A., Smalley S., Hu X., Loo S., Todd R.D., Biederman J., Byrne D., Dechairo B., Guiney A., McCracken J., McGough J., Nelson S.F., Reiersen A.M., Wilens T.E., Wozniak J., Neale B.M., Faraone S.V. 2010. Family-based genome-wide association scan of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 49:898–905.
- Missiaen L., Dode L., Vanoevelen J., Raeymaekers L., Wuytack F. 2007. Calcium in the Golgi apparatus. *Cell Calcium* 41(5):405-416.
- Neale B.M., Lasky-Su J., Anney R., Franke B., Zhou K., Maller J.B., Vasquez A.A., Asherson P., Chen W., Banaschewski T., Buitelaar J., Ebstein R., Gill M., Miranda A., Oades R.D., Roeyers H., Rothenberger A., Sergeant J., Steinhausen H.C., Sonuga-Barke E., Mulas F., Taylor E., Laird N., Lange C., Daly M., Faraone S.V. 2008. Genome-wide association scan of attention deficit hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 147B:1337–1344.

Neale B.M., Medland S., Ripke S., Anney R.J., Asherson P., Buitelaar J., Franke B., Gill M., Kent L., Holmans P., Middleton F., Thapar A., Lesch K.P., Faraone S.V., Daly M., Nguyen T.T., Schafer H., Steinhausen H.C., Reif A., Renner T.J., Romanos M., Romanos J., Warnke A., Walitza S., Freitag C., Meyer J., Palmason H., Rothenberger A., Hawi Z., Sergeant J., Roeyers H., Mick E., Biederman J. 2010. Case-control genome-wide association study of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 49:906–920.

Neale B.M., Medland S.E., Ripke S., Asherson P., Franke B., Lesch K.P., Faraone S.V., Nguyen T.T., Schafer H., Holmans P., Daly M., Steinhausen H.C., Freitag C., Reif A., Renner T.J., Romanos M., Romanos J., Walitza S., Warnke A., Meyer J., Palmason H., Buitelaar J., Vasquez A.A., Lambregts-Rommelse N., Gill M., Anney R.J., Langley K., O'Donovan M., Williams N., Owen M., Thapar A., Kent L., Sergeant J., Roeyers H., Mick E., Biederman J., Doyle A., Smalley S., Loo S., Hakonarson H., Elia J., Todorov A., Miranda A., Mulas F., Ebstein R.P., Rothenberger A., Banaschewski T., Oades R.D., Sonuga-Barke E., McGough J., Nisenbaum L., Middleton F., Hu X., Nelson S. 2010. Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 49:884–897.

Newbury D.F., Bonora E., Lamb J.A., Fisher S.E., Lai C.S., Baird G., Jannoun L., Slonims V., Stott C.M., Merricks M.J., Bolton P.F., Bailey A.J., Monaco A.P., International Molecular Genetic Study of Autism Consortium. 2002 *FOXP2* is not a major susceptibility gene for autism or specific language impairment. *Am J Hum Genet.* 70(5):1318-1327.

Newbury D.F., Winchester L., Addis L., Paracchini S., Buckingham L.L., Clark A., Cohen W., Cowie H., Dworzynski K., Everitt A., Goodyer I.M., Hennessy E.,

- Kindley A.D., Miller L.L., Nasir J., O'Hare A., Shaw D., Simkin Z., Simonoff E., Slonims V., Watson J., Ragoussis J., Fisher S.E., Seckl J.R., Helms P.J., Bolton P.F., Pickles A., Conti-Ramsden G., Baird G., Bishop D.V., Monaco A.P. 2009. *CMIP* and *ATP2C2* modulate phonological short-term memory in language impairment. *Am J Hum Genet* 85(2):264-272.
- O'Roak B.J., Deriziotis P., Lee C., Vives L., Schwartz J.J., Girirajan S., Karakoc E., Mackenzie A.P., Ng S.B., Baker C., Rieder M.J., Nickerson D.A., Bernier R., Fisher S.E., Shendure J., Eichler E.E. 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 43(6):585-9.
- O'Roak B.J., Vives L., Girirajan S., Karakoc E., Krumm N., Coe B.P., Levy R., Ko A., Lee C., Smith J.D., Turner E.H., Stanaway I.B., Vernot B., Malig M., Baker C., Reilly B., Akey J.M., Borenstein E., Rieder M.J., Nickerson D.A., Bernier R., Shendure J., Eichler E.E. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485(7397):246-250.
- Petrin A.L., Giacheti C.M., Maximino L.P., Abramides D.V., Zanchetta S., Rossi N.F., Richieri-Costa A., Murray J.C. 2010. Identification of a microdeletion at the 7q33-q35 disrupting the *CNTNAP2* gene in a Brazilian stuttering case. *Am J Med Genet A*. 152A(12):3164-3172.
- Poot M., Beyer V., Schwaab I., Damatova N., Van't Slot R., Prothero J., Holder S.E., Haaf T. 2010. Disruption of *CNTNAP2* and additional structural genome changes in a boy with speech delay and autism spectrum disorder. *Neurogenetics* 11(1):81-89.
- Pratt S.C., Daly M.J., Kruglyak L. 2000. Exact multipoint quantitative-trait linkage analysis in pedigrees by variance components. *Am J Hum Genet* 66(3):1153-1157.

- Rapin I. & Dunn M. 2003. Update on the language disorders of individuals on the autistic spectrum. *Brain Dev* 25(3):166-72.
- Redon R., Ishikawa S., Fitch K.R., Feuk L., Perry G.H., Andrews T.D., Fiegler H., Shapero M.H., Carson A.R., Chen W., Cho E.K., Dallaire S., Freeman J.L., González J.R., Gratacòs M., Huang J., Kalaitzopoulos D., Komura D., MacDonald J.R., Marshall C.R., Mei R., Montgomery L., Nishimura K., Okamura K., Shen F., Somerville M.J., Tchinda J., Valsesia A., Woodwark C., Yang F., Zhang J., Zerjal T., Zhang J., Armengol L., Conrad D.F., Estivill X., Tyler-Smith C., Carter N.P., Aburatani H., Lee C., Jones K.W., Scherer S.W., Hurles M.E. 2006. Global variation in copy number in the human genome. *Nature* 444(7118):444-454.
- Sebat J., Lakshmi B., Troge J., Alexander J., Young J., Lundin P., Månér S., Massa H., Walker M., Chi M., Navin N., Lucito R., Healy J., Hicks J., Ye K., Reiner A., Gilliam T.C., Trask B., Patterson N., Zetterberg A., Wigler M. 2004. Large-scale copy number polymorphism in the human genome. *Science* 305(5683):525-528.
- Sehested L.T., Møller R.S., Bache I., Andersen N.B., Ullmann R., Tommerup N., Tümer Z. 2010. Deletion of 7q34-q36.2 in two siblings with mental retardation, language delay, primary amenorrhea, and dysmorphic features. *Am J Med Genet A*. 152A(12):3115-3119.
- Shao Y., Raiford K.L., Wolpert C.M., Cope H.A., Ravan S.A., Ashley-Koch A.A., Abramson R.K., Wright H.H., DeLong R.G., Gilbert J.R., Cuccaro M.L., Pericak-Vance M.A. 2002. Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *Am J Hum Genet* 70(4):1058-1061.
- SLI consortium (SLIC). 2002. A genomewide scan identifies two novel loci involved in specific language impairment. *Am J Hum Genet* 70(2):384-398.

- SLI consortium (SLIC). 2004. Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. *Am J Hum Genet* 74(6):1225-1238.
- Spiteri E., Konopka G., Coppola G., Bomar J., Oldham M., Ou J., Vernes S.C., Fisher S.E., Ren B., Geschwind D.H. 2007. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet* 81(6):1144-1157.
- Stergiakouli E., Hamshere M., Holmans P., Langley K., Zaharieva I., deCODE Genetics, Psychiatric GWAS Consortium, Hawi Z., Kent L., Gill M., Williams N., Owen M.J., O'Donovan M., Thapar A. 2012. Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *Am J Psychiatry* 169(2):186-194.
- Strauss K.A., Puffenberger E.G., Huentelman M.J., Gottlieb S., Dobrin S.E., Parod J.M., Stephan D.A., Morton D.H. 2006. Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N Engl J Med* 354(13):1370-1377.
- Stromswold K. 1998. Genetics of spoken language disorders. *Hum Biol* 70(2):297-324.
- Stromswold, K. 2001. The heritability of language: A review and metaanalysis of twin, adoption and linkage studies. *Language*. 77(4):647-723.
- Tager-Flusberg H. & Joseph R.M. 2003. Identifying neurocognitive phenotypes in autism. *Philos Trans R Soc Lond B BiolSci* 358(1430):303-314.
- Taylor L.J., Maybery M.T., Whitehouse A.J. 2012. Do Children with Specific Language Impairment have a Cognitive Profile Reminiscent of Autism? A Review

- of the Literature. *J Autism Dev Disord* (Epub ahead of print).
- Teare M.D. & Barrett J.H. 2005. Genetic linkage studies. *Lancet* 366(9490):1036-1044.
- Tirosh E. & Cohen A. 1998. Language deficit with attention-deficit disorder: a prevalent comorbidity. *J Child Neurol* 13(10):493-497.
- Van der Aa N., Vandeweyer G., Reyniers E., Kenis S., Dom L., Mortier G., Rooms L., Kooy R.F. 2012. Haploinsufficiency of *CMIP* in a Girl with autism spectrum disorder and developmental delay due to a de novo deletion on chromosome 16q23.2. *Autism Res* 5(4):277-281
- Verkerk A.J., Mathews C.A., Joosse M., Eussen B.H., Heutink P., Oostra B.A., Tourette Syndrome Association International Consortium for Genetics. 2003. *CNTNAP2* is disrupted in a family with Gilles de la Tourette syndrome and obsessive compulsive disorder. *Genomics* 82(1):1-9.
- Vernes S.C., Spiteri E., Nicod J., Groszer M., Taylor J.M., Davies K.E., Geschwind D.H., Fisher S.E. 2007. High-throughput analysis of promoter occupancy reveals direct neural targets of *FOXP2*, a gene mutated in speech and language disorders. *Am J Hum Genet* 81(6):1232-1250.
- Vernes S.C., Newbury D.F., Abrahams B.S., Winchester L., Nicod J., Groszer M., Alarcón M., Oliver P.L., Davies K.E., Geschwind D.H., Monaco A.P., Fisher S.E. 2008. A functional genetic link between distinct developmental language disorders. *N Engl J Med* 359(22):2337-2345.
- Vieland VJ, Wang K, Huang J. 2001. Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: comparative evaluation of model-based linkage methods for affected sib pair data. *Hum Hered* 51(4):199-208.

- Villanueva P., de Barbieri Z., Palomino H.M., Palomino H. 2008. High prevalence of specific language impairment in Robinson Crusoe Island. A possible founder effect. *Rev Med Chil* 136(2):186-192.
- Villanueva P., Newbury D.F., Jara L., De Barbieri Z., Mirza G., Palomino H.M., Fernández M.A., Cazier J.B., Monaco A.P., Palomino H. 2011. Genome-wide analysis of genetic susceptibility to language impairment in an isolated Chilean population. *Eur J Hum Genet* 19(6):687-695.
- Wang K., Zhang H., Ma D., Bucan M., Glessner J.T., Abrahams B.S., Salyakina D., Imielinski M., Bradfield J.P., Sleiman P.M., Kim C.E., Hou C., Frackelton E., Chiavacci R., Takahashi N., Sakurai T., Rappaport E., Lajonchere C.M., Munson J., Estes A., Korvatska O., Piven J., Sonnenblick L.I., Alvarez Retuerto A.I., Herman E.I., Dong H., Hutman T., Sigman M., Ozonoff S., Klin A., Owley T., Sweeney J.A., Brune C.W., Cantor R.M., Bernier R., Gilbert J.R., Cuccaro M.L., McMahon W.M., Miller J., State M.W., Wassink T.H., Coon H., Levy S.E., Schultz R.T., Nurnberger J.I., Haines J.L., Sutcliffe J.S., Cook E.H., Minshew N.J., Buxbaum J.D., Dawson G., Grant S.F., Geschwind D.H., Pericak-Vance M.A., Schellenberg G.D., Hakonarson H. 2009. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* 459(7246):528-533.
- Weiss L.A., Arking D.E.; Gene Discovery Project of Johns Hopkins & the Autism Consortium, Daly M.J., Chakravarti A. 2009. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 461(7265):802-808.
- Whitehouse A.J., Bishop D.V., Ang Q.W., Pennell C.E., Fisher S.E. 2011. *CNTNAP2* variants affect early language development in the general population. *Genes Brain Behav* 10(4):451-456.

Zheng J.Q. & Poo M.M. 2007. Calcium signaling in neuronal motility. *Annu Rev Cell Dev Biol* 23:375-404.

Zweier C., de Jong E.K., Zweier M., Orrico A., Ousager L.B., Collins A.L., Bijlsma E.K., Oortveld M.A., Ekici A.B., Reis A., Schenck A., Rauch A. 2009. *CNTNAP2* and *NRXN1* are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in *Drosophila*. *Am J Hum Genet* 85(5):655-666.