

# 43 Key Issues and Future Directions: Genes and Language

SIMON E. FISHER

Determining the molecular mechanisms that underlie human capacities for speech, language, and reading is a vibrant and exciting initiative. The prior chapters of this part comprehensively reviewed the state-of-the-art on mapping genetic factors that contribute to language- and reading-related skills (Luciano & Bates, chapter 39), discovery of mutations involved in speech disorders (Morgan, chapter 40), and deciphering the impacts of key genes and mutations on development and function of the nervous system, in different model systems (Vernes, chapter 41). Insights in this area come not only from language itself, but also by studying genetic influences on the associated neural architecture, most notably the lateralization of much of the relevant functional circuitry (Francks, chapter 42). Here, in this final chapter of the part, I lay out the promise and challenges of the coming years for tracing genetic pathways involved in language, in light of rapid improvements in the available research tools. In particular, I focus on the dramatically reduced costs and enhanced feasibility of large-scale genotyping and high-throughput DNA sequencing and what this will mean for the future landscape of studies in this area. I emphasize that the substantial advances in molecular technologies will need to be matched by increased sophistication in our approach to defining and characterizing language-related phenotypes. Thus, interdisciplinary connections among molecular biologists, psychologists, linguists, and neuroscientists are set to become ever more essential for progress in this fast-moving research endeavor. I will begin by considering rare mutations and the impact of next-generation DNA sequencing, before moving on to discuss common genetic variants and the prospect of systematic studies of interindividual variation in language proficiency.

## *1. Rare Gene Variants and How to Find Them*

Rare gene variants that disrupt aspects of speech and language development offer powerful windows into neurobiological pathways (see Vernes, chapter 41). The fundamental scientific value afforded by the discovery

of such mutations is clearly illustrated by the example of *FOXP2*<sup>1</sup> (Fisher & Scharff, 2009). The identification of *FOXP2* preceded the completion of the first full draft of the human genome sequence. It relied on analysis of an unusual family (the KE family) in which 15 relatives displayed apparent monogenic inheritance of developmental speech and language problems across three successive generations. Even with such a large family available to study, the process of uncovering the causative mutation was a lengthy one. It began with a screen for genomic regions linked to the disorder (Fisher, Vargha-Khadem, Watkins, Monaco, & Pembrey, 1998), followed by an intensive gene-by-gene search through one particular section of chromosome 7, containing many unknown genes that had never been sequenced before (Lai et al., 2000). Clues to limit the search came from an unrelated child with a similar disorder, along with a chromosomal rearrangement involving that same region. Eventually, researchers pinpointed the KE family mutation (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001), a change of a single base in a crucial part of *FOXP2* (at that time, a novel human gene that had not yet been described). Subsequent studies could then look specifically at *FOXP2* in other smaller families. Although these targeted investigations successfully confirmed an etiological role, by finding other mutations of this gene in independent cases/families with a similar profile of speech and language deficits, they also illustrated that such mutations are extremely rare (MacDermot et al., 2005). Indeed, more than a decade and a half after discovering *FOXP2*, most cases of speech and language disorders remain unexplained at the genetic level (Deriziotis & Fisher, 2017; Graham & Fisher, 2015).

In the meantime, the broader landscape of genetic research has been dramatically transformed by the development of novel technologies for reading off a person's genomic sequence (Metzker, 2010). Traditional methods, such as Sanger sequencing, are highly laborious, time-consuming, and expensive. These standard approaches were used to assemble the first draft sequence of the human genome, work that took the best

part of a decade, requiring the efforts of tens of thousands of researchers around the world, at a cost of billions of US dollars. So-called next-generation techniques (Goodwin, McPherson, & McCombie, 2016), massively parallel sequencing of a great many DNA fragments at the same time, allow for an almost entire human genome to be sequenced from a simple saliva sample in a matter of hours, at a cost of less than US\$1,000 (based on estimates in 2018, at time of writing). Another upcoming generation of technologies involves direct sequencing of individual intact molecules of DNA from start to finish, without shearing up into fragments (Loose, 2017; Zaaier et al., 2017). So, as it becomes ever cheaper and quicker to readout the genomes of people with speech and language disorders, how will this impact on our understanding of the genetic underpinnings?

In principle, the routine availability of whole genome sequencing should make it considerably easier to study pedigrees, such as the KE family, in which there are multiple relatives affected with speech-, language-, and/or reading-related disorders across several generations. Thus, one might expect a substantial acceleration in the discovery of novel causative mutations. In practice, even with these new molecular technologies, progress in uncovering susceptibility genes has been relatively slow, and it is worth considering some of the explanations for this. Crucially, the majority of cases of language-related disorders involve multifactorial etiology (Graham & Fisher, 2015), resulting from the combined effects of several, perhaps many, genetic risk factors of small effect (as will be discussed in section 2). Researchers rely on a certain amount of serendipity to successfully track down those families that have truly monogenic forms, in which the disorder can be fully explained by a single genetic mutation with a large effect size, also known as high penetrance. A pedigree may show what appears to be a simple inheritance pattern at the phenotypic level, but this surface level view will not necessarily be reflected in the underlying biology. Hence, it is not uncommon to come up empty-handed after searching for a high-penetrance genetic mutation in a family of interest, despite using the fastest and most powerful sequencing techniques.

Moreover, whole genome sequencing has revealed that each human individual carries a surprisingly large number of rare sequence variants that could potentially disrupt gene function. It can be challenging to distinguish causative mutations from more benign sequence variants in the genomic background (Derizotis & Fisher, 2017). When a rare disruptive variant cosegregates with a disorder in multiple affected relatives, this finding is consistent with an etiological role,

but often families with language-related disorders are not large enough to be sure that the observed pattern represents a statistically significant finding (i.e., that the cosegregation is beyond that expected by chance). The KE family is exceptional in this regard (Fisher et al., 1998). Sometimes a study will identify several different candidate mutations in a family, without concrete evidence to highlight a single true culprit. To have high confidence that a gene is implicated in a disorder, it is necessary to collate observations of this same gene being disrupted by independent mutations in different families and/or cases. These factors make it challenging to identify completely novel etiological genes that have not already been confirmed as potential contributors to language-related disorders.

To distinguish true causative mutations from benign variation in an affected family, cosegregation with the disorder is important, but not sufficient. We also need to draw inferences about the likely functional relevance of any variant of interest (i.e., determining whether it will disturb gene function). Although it is sometimes possible to make confident *in silico* predictions about the disruptive potential of a genetic variant, based on DNA sequence data alone, the ability to do this varies considerably for different parts of the genome. We are on firmest ground when it comes to the genomic regions that directly code for proteins. For example, a DNA sequence variant in a protein-coding gene may lead to substitution of one amino acid for another at a key point of the encoded protein, and if that substituted amino acid has different chemical properties it can distort the shape of the protein, preventing it from carrying out its normal cellular functions. (This is the kind of mutation that was found in *FOXP2* in the affected members of the KE family; Lai et al., 2001.) Other well-studied types of mutation lead to truncation of the encoded protein, so that it is missing crucial parts, or may even stop the protein from being expressed at all (MacDermot et al., 2005). Importantly, the protein-coding regions (collectively known as the *exome*) make up only 1% to 2% of the genetic makeup of a person—the vast majority of the genome comprises noncoding DNA. Variants in noncoding regions can also play etiological roles in a disorder, by disturbing the way that expression of a gene is regulated (Devanna et al., 2018), but our understanding of such effects is lagging behind that for coding variants. Thus, even as it becomes trivial to obtain whole genome sequences from people affected with language-related disorders, and to accurately identify all the rare DNA variants that they carry, we still face difficulties in interpreting the functional significance of the variants, and this will impede

progress in the field. Chapter 41 by Vernes shows how experiments in model systems, such as neurons grown in the laboratory, will become increasingly important for making sense of the avalanche of sequence data we face.

Finding families with multiple affected relatives is not the only way to pinpoint rare mutations involved in neurodevelopmental disorders. Perhaps counterintuitively, we may sometimes learn about genetic underpinnings by investigating so-called *simplex* families in which there is just a single affected proband with unaffected parents and siblings and no obvious family history (Fischbach & Lord, 2010). In this kind of strategy, we compare the genome (or exome) sequence of the proband to those of the parents and search for any new mutations that have arisen *de novo* in his or her genome (i.e., changes that are present in the proband, but absent from both parents) (Veltman & Brunner, 2012). The natural mutation rate in humans is low, so that a typical child carries fewer than a hundred *de novo* changes, with only one or two of these being located in the protein-coding parts of the genome. In the rare cases where a *de novo* mutation disrupts an important gene in a substantive way, this may be sufficient to lead to a disorder in the child who carries it. Using next-generation sequencing, it has been shown that severe neurodevelopmental disorders such as autism and intellectual disability (ID) are enriched for these types of causative *de novo* mutations, explaining a significant proportion of cases (Gilissen et al., 2014; Iossifov et al., 2014). By studying large enough numbers of simplex families, it becomes possible to find genes that are recurrently mutated in independent probands. Interestingly, a number of genes thereby implicated in autism and/or ID phenotypes, such as *FOXP1* and *TBRI*, are connected to language-related genetic pathways through protein-protein interactions with *FOXP2* (Deriziotis et al., 2014; Sollis et al., 2016).

## 2. Next-Generation Sequencing in Language-Related Disorders: The Story So Far

It is early days for applying next-generation DNA sequencing to speech and language disorders (Deriziotis & Fisher, 2017). Only a handful of studies in this field have so far taken advantage of these techniques. Most have faced limitations in reliably determining etiological relevance of identified variants, in line with the issues discussed in section 1. For example, exome sequencing of people from Robinson Crusoe Island, an isolated population with a high incidence of specific language impairment (SLI, also known as

developmental language disorder, DLD), revealed a rare variant of the *NFXL1* gene that was associated with the disorder, but within a complex model, since only 39% of affected cases carried the putative risk variant, and it was found in 10% of people with typical language ability (Villanueva et al., 2015). Another study of a geographical isolate with elevated levels of SLI/DLD, from a remote cluster of villages in the north of Russia, combined data from genome-wide association screening (see section 3) and exome sequencing in a subset of individuals, and again was unable to find a high-penetrant mutation accounting for the phenotype, although a pathway-based analysis identified enrichment for variants affecting *MEF2*-regulated genes (Kornilov et al., 2016). The emerging view that most cases of SLI/DLD involve complex genetic architecture is further supported by an exome sequencing study of rare variants in 43 unrelated probands, which found possible pathogenic variants in known speech-/language-related genes *ERCI*, *GRIN2A*, and *SRPX2*, as well as in novel candidates including *OXRI*, *SCN9A*, and *KMT2D*, and potential “multiple-hit” cases (i.e., rare variants in two different genes in the same proband) involving genes such as *AUTS2*, *STARD9*, *SEMA6D*, and *SYNPR* (X. S. Chen et al., 2017). However, it was difficult to draw firm conclusions about the etiological significance of the variants, without clear-cut evidence of segregation with disorder, recurrent independent mutations, or data from functional experiments.

Typically cases of SLI/DLD are not as severe as autism or ID and may not be enriched for high-penetrance mutations. Interestingly, the monogenic disorder caused by *FOXP2* mutation has a distinct phenotype from common forms of SLI/DLD, in that it involves childhood apraxia of speech (CAS) accompanied by impairments affecting most aspects of language development (Watkins, Dronkers, & Vargha-Khadem, 2002). Indeed, next-generation sequencing studies suggest that by targeting a disorder such as CAS it is possible to increase the yield of causative variants. A recent investigation used whole genome sequencing to analyze 19 unrelated probands with a robust diagnosis of CAS (Eising et al., 2018). For nine of the probands, whole genome sequences were also obtained from both unaffected parents, allowing systematic screening for *de novo* coding variants, the first time that this had been done for a primary speech disorder. Three probands carried *de novo* disruptions with high likelihood of being pathogenic, affecting the genes *CHD3*, *SETD1A*, and *WDR5*. Analyses of the genome sequences from the 10 isolated probands (for whom parental

DNA was not available) identified novel loss-of-function variants affecting *KAT6A*, *SETBP1*, *ZFHX4*, *TNRC6B*, and *MKL2*. The sample size in this study was too small for observation of recurrent mutations (same gene independently affected in different cases) within the cohort. However, these new potential CAS candidates are regulatory genes with known links to neurodevelopment, and analyses of RNA-sequencing data revealed that they are clustered in a single module of genes showing concordant spatiotemporal expression in the developing human brain (Eising et al., 2018). Another notable feature of these findings is that, despite the lack of recurrent mutation within the 19 CAS probands, most of the new candidate genes have been implicated in neurodevelopmental phenotypes in prior human studies.

Overall, the first data on rare variants from next-generation sequencing studies of speech-/language-related disorders are consistent with the emerging view of convergence in the neurogenetic pathways underlying multiple distinct brain-related disorders (e.g., ID, autism, epilepsy, schizophrenia) (Deriziotis & Fisher, 2017). Putative risk genes often do not seem to respect the clinical boundaries that we impose on definitions of disorder, which should not be surprising given the complex relationships among genes, brains, and behavior (Fisher, 2006). Thus, a major goal for the future is to gain a better understanding of the mappings between rare high-penetrance mutations and the potential cognitive and behavioral (as well as physiological and anatomical) outcomes observed in the children that carry them. Although the individual mutation events are by definition rare, through national and international networks of geneticists and clinicians, it is becoming easier to pull together data from multiple unrelated cases for a particular candidate gene of interest, enabling a genotype-driven approach to phenotype definition. For example, much of our initial knowledge on the consequences of *FOXP2* mutation came from studies of the etiological substitution found in members of one family (the KE family), demonstrating disproportionate effects on speech and language skills (Watkins, Dronkers, & Vargha-Khadem, 2002). Recent years have seen acceleration in the identification of new point mutations disrupting *FOXP2* (e.g., Reuter et al., 2017). Systematic in-depth studies of these independent cases, to characterize the range of profiles of speech, language, and other aspects of behavior and cognition, promise valuable new insights into the relationship between *FOXP2* disruptions and brain development. More broadly, this kind of genotype-driven strategy, applied to language-related candidate genes, is likely to enhance our studies of rare disruptive variants in the coming decade, and it will reshape

concepts concerning the boundaries between neurodevelopmental disorders involving speech and language deficits.

### 3. Learning from Common Genetic Architecture

Rare mutations are just one source of evidence for connecting genotypes to phenotypes. Studies of common genetic variation provide a complementary strategy for dissecting biological pathways that are important for language skills. Moving beyond a view that language proficiency is a fixed unvarying trait across all human beings, the idea here is to capture variability in aspects of language proficiency and try to relate that to the well-documented interindividual variation in people's genomes. Again, technological advances are making the molecular side of this rather trivial to carry out—for less than US\$100 per sample, DNA microarrays can simultaneously read out the alleles of hundreds of thousands of positions across the genome. Using these variants, known as SNPs, together with our knowledge of population genetics, it is possible to recover the majority of common genetic variation that a person carries, without the need to directly sequence the whole genome. (But note that rare variation is more difficult to reliably reconstruct with these techniques.) Hence, as summarized in the present volume (Luciano & Bates, chapter 39) and elsewhere (Deriziotis & Fisher, 2017; Graham & Fisher, 2015), the past few years have seen the first genome-wide association scans (GWASs) of speech-, language-, and reading-related traits, either studying disorders (Eicher et al., 2013; Field et al., 2013; Gialluisi et al., 2014; Kornilov et al., 2016; Nudel et al., 2014; Roeske et al., 2011) or normal variation (Harlaar et al., 2014; Luciano et al., 2013; St. Pourcain et al., 2014). Some interesting suggestive findings have emerged from these studies. However, beyond an association of variants in the *ROBO2* gene with a measure of early vocabulary, which didn't persist for language measures at later stages of development (St Pourcain et al., 2014), robust significant evidence of association has eluded researchers, and there are not yet any positive independent replications of GWAS hits. Moreover, it turns out that few, if any, of the claims for common variant association from earlier studies (prior to advancements in GWAS technologies) are borne out by these newer large-scale screens (e.g., Becker et al., 2014). Why have GWAS approaches failed to deliver major insights so far for the language sciences, and what are the prospects for turning this situation around in the future?

A key constraint here is genetic complexity. Most language-related traits are highly heritable, meaning that the combined effects of all genetic variations

together account for a significant proportion of phenotypic variability. However, the underlying genetic architecture is nonetheless complex, involving multiple different loci (Graham & Fisher, 2015). For language phenotypes, just as with other complex traits, the individual effect size of any single common genetic variant is usually very small and can be difficult to detect. Moreover, a standard GWAS involves testing of many hundreds of thousands of largely independent SNPs spread across the genome. The value of this approach is that the search is unbiased and does not depend on prior assumptions about the nature of the genes involved (Visscher et al., 2017), which is important for language traits where we are only just beginning to glimpse the underlying biology. But the downside of a systematic search is that it imposes a huge multiple-testing burden and renders a traditional threshold of  $P < .05$  useless, since it would yield a hugely inflated number of false positive associations. This issue is usually dealt with through a rigorously adjusted  $P$  value threshold for declaring genome-wide significance, of  $5 \times 10^{-8}$ , ideally along with an independent replication in another sample. Taken together, the small effect sizes of individual SNPs and the substantive multiple-testing involved, mean that cohorts of thousands of people (preferably tens of thousands of people) are required to give adequate basic power for a GWAS (Visscher et al., 2017). A classic and often-cited example of the impact of sample size is schizophrenia, where early GWAS efforts struggled to identify significant associations. Through combining available cohorts into a shared sample of almost 37,000 cases, the field moved from only around 30 potential associated loci to more than 128 robust independent genome-wide significant signals (Schizophrenia Working Group of the Psychiatric Genomics, 2014). As sample sizes continue to increase, more and more replicable loci are being uncovered. This is just one of many potentially compelling illustrations from the field over the past few years (Visscher et al., 2017). Note that it can be hard to predict the sample size at which one can be confident of capturing associations, since that depends on multiple unknown factors, including the precise nature of the genomic architecture. But it is clear that the sample sizes used thus far for language-related phenotypes are on the edge of what is desirable, to say the least. To give an extreme example, one study performed GWAS case-control analyses with only 163 language-impaired cases (Eicher et al., 2013), which is a highly underpowered study design for seeking out common variant associations in a genome-wide context.

Clearly, a single laboratory is unlikely to by itself have the capacity to recruit and study (phenotype and

genotype) cohorts of tens of thousands of participants. The success of GWAS approaches for biomedical traits has largely stemmed from researchers embracing highly collaborative models of team science, in which data from multiple independent smaller cohorts are brought together for meta-analysis (or even mega-analysis, depending on issues over data alignment, sharing policies, and ethical approvals). To return to the much-cited schizophrenia GWAS example (Schizophrenia Working Group of the Psychiatric Genomics, 2014), this was only possible through the efforts of a large international network of clinical and research groups, known as the Psychiatric Genomics Consortium, which also supports genetic analyses of several other psychiatric phenotypes (Sullivan, 2010). To coordinate large-scale consortium investigations of this kind, which may depend on work of hundreds of clinicians and scientists at different sites around the world, is by no means a trivial task. It involves harmonized approaches to phenotypic definition, integrating data from different genotyping platforms, standardizing analytical procedures, and so on, to ensure optimal signal-to-noise ratios in any combined genetic analysis. Still, if the language sciences are to take full advantage of the promise offered by GWAS approaches, then this field will also have to move toward the team science model, otherwise we will continue to come up against limitations of insufficient power in our studies. Indeed, this move is already beginning. For example, the new GenLang Consortium ([www.genlang.org](http://www.genlang.org)) has been established to try to integrate existing cohorts (both disorder-based and epidemiological) that have available speech, language, and/or reading phenotype data coupled to genome-wide genotype information, and the first meta-GWAS studies are underway.

#### *4. Toward Large-Scale Genomic Screens of Speech, Language, and Reading Skills*

With the formation of consortium efforts focused on language and genetics, and advances in methods for analyzing and integrating high-throughput molecular data, there are good grounds for optimism. Nonetheless, it is worth recognizing that phenotypes related to speech, language, and reading pose some special challenges. Large-scale biomedical studies often tap into routine clinical diagnostic pipelines in order to recruit sufficient numbers of affected participants for genetic analyses of adequate power. For instance, a wealth of information about the genetic architecture of ID and autism (and many other more medically orientated phenotypes) has come through study of clinically ascertained cohorts. In contrast, language-related neurodevelopmental difficulties, such as dyslexia, SLI, and

CAS, are less likely to be considered within a clinical framework, especially when they occur against a background of preserved general cognitive function. This makes it harder to assemble large cohorts through opportunistic routes involving medical screening systems that are already in place (but see the following discussion for some possibilities). There is also a degree of resistance in certain quarters to “medicalization” of problems in this area. Even if considering speech and language therapists as a potential source, it is not yet common practice for therapists to posit a potential role for genetic factors, despite growing evidence of their relevance (see Morgan, chapter 40). It is sobering to note that while developmental disorders that disturb speech, language, and/or reading are estimated to be some of the most common problems facing children in the modern world (Newbury & Monaco, 2010), the available cohorts collected for genetic studies over the past few decades remain disappointingly small (Deriziotis & Fisher, 2017). Thus, even with an international consortium like GenLang bringing together the existing data sets for meta-/mega-analysis, the combined sample sizes still fall short of that which would be optimal for high-powered GWASs.

Another complication when characterizing language ability for large-scale genetic studies is the complex multifaceted nature of the relevant capacities. There is no single agreed test that can give a simple readout of a person’s skills in this area, since to be linguistically competent one must draw on a wide range of distinct processes, acquiring and manipulating items at multiple levels from phonemes to words to sentences and discourse, using different modalities (sign, speech, writing), expressing, and comprehending. Test batteries do exist and have been used in prior genetic studies for a variety of purposes. Some batteries have been designed for diagnosing children with problems in acquiring aspects of language or for indexing impaired linguistic competence in adults with neurological lesions and or age-related decline. Other batteries were developed to identify proficiency of language use in adult second language learners, primarily as tools for teaching and assessment purposes. Few, if any, available batteries are effective for capturing interindividual differences across the full suite and spectrum of language-related abilities in the general population. Where particular tests are available and have already been used to collect phenotypes from different cohorts with matched genotype data, aligning the phenotypes is often not straightforward. For example, tests of “single-word reading” (typically used to diagnose dyslexia; Carrion-Castillo, Franke, & Fisher, 2013) exist in multiple different forms and can be coded in different

ways for accuracy and speed, among other things. Moreover, consortium efforts usually depend on bringing together cohorts from different countries across the world—if your phenotype of interest is a biometric one such as blood pressure, height, or head circumference, then your test instruments can be easily aligned across international boundaries, regardless of the host language. However, if the phenotype of interest is itself a language-related skill, and there are different native languages spoken by the distinct cohorts being tested, then harmonization of test data becomes much more challenging, for reasons that are self-evident. This difficulty applies not only to tests of spoken language and grammar processing where the surface properties of the languages involved might be highly distinct, but also to assessments of reading and spelling, given obvious differences in orthographic depth of alternative writing systems (Ziegler & Goswami, 2006).

Thus, for the language sciences to fully take advantage of the potential for consortium-based efforts to perform productive large-scale genetic analyses across existing cohorts, there is much work to do in the area of phenotype alignment over the coming years. Success in studying genetic architecture that underpins language will depend on collaborations across disciplines, in particular interactions among geneticists, psychologists, and linguists. More crucially, perhaps, it could be valuable to develop novel test batteries that are designed to reliably capture the true extent of interindividual differences in language-related skills at multiple levels, in both expressive and receptive domains. The degree of variation in such skills in randomly selected people from the general population, the relationships between supposedly different skills, how much they draw on general cognitive processes, and so on, are all important unaddressed questions in their own right, worthy of attention by the field. Indeed, this is the focus of a major ongoing project by the Dutch Language in Interaction Consortium. If new batteries assessing language-related skills are designed in such a way that they can be reliably administered via remote testing, through web portals or apps, that could revolutionize genetic studies of such traits. There are more and more large-scale population cohorts available with genome-wide genotype data already collected (e.g., Ge, Chen, Neale, Sabuncu, & Smoller, 2017), but in which no language-relevant phenotypes are yet available. Web-/app-based collection of the missing phenotypes offers a highly cost-effective way to gather standardized data for association analyses with the pre-existing genetic information.

Another approach that is being used for clinical traits makes use of electronic health records (EHRs) in the medical databases of hospital biobanks (Wei &

Denny, 2015). These are biobanks in which hundreds of thousands of DNA samples have been collected from all patients, and genome-wide genotypes have been obtained. By searching through the matching EHRs from patients for particular medical codes that designate a particular disorder, it is possible to virtually assemble a sample of cases from the biobank and run a GWAS against suitable controls also taken from the same biobank. A number of investigations are currently underway using *International Statistical Classification of Diseases and Related Health Problems, 10th Revision* codes for disorders related to speech, language, and reading. These types of disorders are likely to be under-represented in hospital-based biobanks, and the EHR approach may lack the rigor of actively recruited samples, but it will be interesting to see how this strategy fares.

GWAS efforts will not only fuel the identification of significantly associated genetic variants that can give insights into the biological pathways that underlie language. The expanding tool set of molecular genetic epidemiology also offers innovative ways to answer important questions about how our various language-related skills change during development, how they relate to each other and to other traits, as well as their connections to evolutionary data (see section 6). In such studies, we can directly use molecular data from cohorts to empirically assess genetic correlations, that is, how much of the shared variance between two different traits is accounted for by shared effects at the genetic level. To give just one example, one can ask broadly whether the same genetic factors are contributing to a language-related skill when monitored at different stages of children's development (St. Pourcain, Eaves, et al., 2018; St. Pourcain, Robinson, et al., 2018). These kinds of empirical observations can then constrain future theories about how nature and nurture interact during language acquisition, which again indicates the increased need for interdisciplinary exchanges as this field matures.

### *5. Investigating Language Neurobiology with Neuroimaging Genetics*

As discussed extensively in section 4, the small effect size of each individual genetic variant is a limiting factor for studies of any complex multifactorial trait. Researchers have attempted to overcome this problem in a variety of ways, highlighted in sections 3 and 4. One complementary avenue that initially held particular promise centered on the idea of moving from more peripheral traits of behavioral/cognitive performance toward the structural and functional neural correlates (Bigos & Weinberger, 2010). The original argument

was that brain-based measures, for example collected using methods such as MRI, would be closer to the underlying biological mechanisms, such that the effect sizes of associated genetic variants would become larger and easier to detect in small samples. A new interdisciplinary field—now known as neuroimaging genetics—crystallized around this hope at the turn of the century. Early prominent studies targeted a handful of well-studied candidate genes in samples with only tens of participants and reported significant associations that, if true, would account for a substantial proportion of variation in the brain-based measures (e.g., Egan et al., 2003; Hariri et al., 2002). However, as neuroimaging genetics has become more systematic and rigorous and has moved away from candidate gene studies toward unbiased GWASs, early optimism about effect sizes has proved unfounded, at least for variations in neuroanatomy (de Zubicaray et al., 2008; Hibar et al., 2015). Sadly, genes really do not care about how much cost and effort it has taken researchers to collect a phenotype. It turns out that even if that phenotype has been measured using expensive sophisticated tools of MRI, it is likely to face just the same issues as other multifactorial traits.

To help illustrate this point, consider a fairly standard brain trait that can be reliably captured in living participants through MRI scanning, such as the volume of a subcortical structure. In 2015, GWAS work by the ENIGMA (Enhancing NeuroImaging Genetics through Meta-Analysis) Consortium analyzed the volumes of seven structures, including the hippocampus, nucleus accumbens, caudate, putamen, pallidum, amygdala, and thalamus, in almost 31,000 people from 50 different cohorts around the world and identified a number of significant associations (Hibar et al., 2015). The strongest association observed in that study, for a SNP with a significant replicable influence on the volume of the putamen, accounted for only 0.52% in the variance of the measured trait. Even with tens of thousands of people, most of the variance in subcortical volumes is still to be accounted for, but larger follow-on consortium studies (as yet unpublished) are uncovering additional significant replicable loci and moving toward a more complete account of the relevant genetic architecture. Perhaps for a consortium like ENIGMA, involving a great many different sites (Bearden & Thompson, 2017), it could be that the diversity of scanning machines and parameters, and of types of cohorts across the world, leads to increased heterogeneity in the genetic analyses, reducing power. However, there are now various neuroimaging genetics cohorts, such as that collected by UK Biobank, in which the same scanners and test paradigms have been used for relatively homogenous

populations and the sample sizes are rapidly approaching similar levels to that for ENIGMA (Alfaro-Almagro et al., 2018). The first neuroimaging genetics analyses of resources such as UK Biobank are drawing the same conclusions as prior consortium studies; just as with behavioral and cognitive traits, interindividual variations in brain structure have a highly multifactorial genetic basis, involving a great many common variants spread across the genome, most with a tiny effect size.

Thus, to be successful, neuroimaging genetics studies need to confront the exact same issues as more basic genetic association studies, with designs that have adequate power for reliably detecting modest effect sizes, and that factor in sufficient replication and validation steps to guard against false positives. Thus far, most language-related neuroimaging genetics studies have failed to fulfil these requirements, and they may have already distorted the literature of this emerging field. A brief example, to highlight the issues, concerns *FOXP2*, the most well-studied of language-related candidate genes (Fisher & Scharff, 2009). Neuroimaging studies of people with rare high-penetrance mutations that cause severe speech and language disorders have revealed differences in brain structure and function; these alterations are subtle yet significant and they are distributed across various regions, with inferior frontal gyrus, striatum, and cerebellum being of particular interest (Watkins, Vargha-Khadem, et al., 2002). Subsequent to identification of rare etiological mutations that obviously disrupt *FOXP2*, research groups around the world sought evidence that common variants (of unknown molecular effect) might be associated with variability in aspects of behavior, brain structure or function in other disorders, or in the general population (Mueller et al., 2016; Padovani et al., 2010; Peter et al., 2011; Pinel et al., 2012; Premi et al., 2012; Wilcke et al., 2012). Despite the fact that common variants are expected to have (at best) much smaller effect sizes than the high-penetrance mutations causing monogenic disorders, the group sizes in neuroimaging genetics studies of common *FOXP2* SNPs were low—as few as 14 healthy participants in one report (Wilcke et al., 2012). A later study of 1,300 people from the general population targeted each of the *FOXP2* SNPs that had been claimed to show associations from the earlier smaller investigations and found no evidence of any effects on neuroanatomy in that larger independent sample (Hoogman et al., 2014). Moreover, no significant associations of *FOXP2* SNPs with aspects of brain structure have so far emerged in larger GWAS studies of tens of thousands of people, even when those studies included measurement of striatal volumes (e.g., Hibar et al., 2015), a key structure of interest from studies of

rare mutations of this gene. This does not discount the possibility that effects of common *FOXP2* variants will eventually be found, perhaps depending on the granularity of the structural phenotypes that we are studying. But the main point here is that the prior positive claims of *FOXP2* SNP effects based on only tens of individuals are unlikely to be valid.

If high-powered cohorts with thousands of people fail to find effects, why are there some reports in the literature that show supposedly significant associations in much smaller samples? First, although it is not often appreciated, a study design that is severely underpowered is also more prone to spurious false positive findings; for discussion of these issues, see Button et al. (2013) and follow-up commentaries. Second, neuroimaging genetic approaches involve bringing together complex data sets of high multidimensionality in which there are many choice points regarding how the data are processed and tested. At the genetic level there are choices about which markers to choose from a gene of interest, which modes of inheritance to consider (additive, recessive/dominant), how to deal with rare alleles or markers with multiple different alleles, what direction of effect to expect, whether to test genders separately, and so on. At the neuroimaging level, there are choices about whether to look bilaterally or at individual hemispheres, which regions of the brain to test, how to best define those regions of interest (e.g., data-driven or atlas-based, which atlas to choose), which processed features to assess (e.g., gray matter or white matter density, volume, thickness, surface area, shape, some measure of connectivity). The excessive choice points of even the simplest neuroimaging genetics study makes this area of work especially susceptible to inflation bias due to implicit (or even explicit) selective reporting, without suitable levels of multiple-testing adjustment. Third, and in a related point, given the broad possibilities regarding the (typically unknown) relationship between a genetic variant of interest and its impact on brain measures, exploratory testing and hypothesis-driven testing are often conflated. Some studies engage in post hoc hypothesis testing using the same data set that was originally used for exploratory analyses, rather than carrying out a formal replication in an independent cohort. These, and other issues facing neurogenetic association studies, are part of the broader discussions currently underway in psychology and neuroscience around the so-called *replicability crisis* (Grabitz et al., 2018). Future neuroimaging genetics studies of language should pay close attention.

Notwithstanding the need for increased rigor moving forward, the prospects for this area of work remain exciting. There is little doubt that neuroimaging

genetics will be a useful strategy for understanding effects of molecular factors on brain circuits involved in language, especially with the growing data sets that are becoming available, such as those from large-scale consortia like ENIGMA (Bearden & Thompson, 2017) and population resources such as UK Biobank (Alfaro-Almagro et al., 2018). It is already feasible to run high-powered GWAS studies for cortical thickness and surface area of language-related areas, such as the inferior frontal gyrus and superior temporal gyrus, and it will soon be possible to thoroughly investigate more sophisticated characterizations of language systems, such as structural and functional connectivity of the key circuits, aspects of gyral/sulcal morphology, and degree of lateralization of key phenotypes (Guadalupe et al., 2017). One area where it will remain difficult to obtain suitably high-powered data sets for genetic association studies is in task-based functional neuroimaging, tapping into particular aspects of language processing. Often, due to scanning time limitations, the task-based paradigms used in establishing large-scale public biobanks are not designed to tackle research questions about language (or to go deeply into any particular cognitive trait). It is an open question as to how much of the relevant genetic architecture involved in language processing could be deciphered through (for example) resting-state functional MRI data or some other more generic source. Nonetheless, several different groups around the world have been using language tasks for functional MRI and/or magnetoencephalography in hundreds of participants and have obtained DNA samples to match (Mahowald & Fedorenko, 2016; Schoffelen et al., 2017); it should be possible to harmonize data across these groups for meta-analyses. If these combined sample sizes are not large enough to support GWASs themselves, they could still attain sufficient power for targeted follow-up investigations of top SNPs (or a polygenic risk score—a weighted cumulative score based on SNP associations at multiple loci) from other GWAS efforts. Finally, as discussed by Francks (chapter 42), there is also considerable potential to integrate association findings with data from analyses of gene expression in different brain regions, another area that is rapidly expanding with advances in molecular technology.

## 6. Concluding Remarks

Due to limitations of space, this chapter has primarily focused on methods for capturing rare and common variation in the genomes of modern humans and relating that to brain-based and behavioral/cognitive measures. Remarkably, the impact of advanced molecular technologies is also extending its reach into the distant

past, giving us novel means to answer fundamental questions about the evolution of our unusual communicative capacities. With next-generation sequencing, it has become possible to read the genome sequences not only of our extant primate cousins but also of extinct archaic hominins, such as Neanderthals and Denisovans (Meyer et al., 2012; Prufer et al., 2014; Slon et al., 2017). For more information on this, I refer the reader to recent reviews discussing the promise of molecular anthropology (Paabo, 2014) and how comparative genomics of different hominins can help shed light on language origins (Fisher, 2017). The researchers of tomorrow are incredibly fortunate to have access to catalogues of genetic differences between *Homo sapiens* and other extant and extinct apes, along with information about parts of the genome that have undergone selection, or introgressed from one hominin to another, at distinct stages of evolution (Fisher, 2017; Paabo, 2014). Perhaps the most exciting work of the future will be that which integrates these historical findings with the new wave of genome-wide screening data from studies of language-related phenotypes in modern human populations (Deriziotis & Fisher, 2017). Do genes that are implicated in speech and language disorders, or interindividual variability in language-related skills, overlap with those that changed during the emergence of our species, and how do such genes influence the structure and function of key brain circuits, in humans and other species? Moreover, with the rise of gene-editing tools, such as CRISPR/Cas9 (J. S. Chen et al., 2017) and novel systems for modeling human brain development in the lab, such as cerebral organoids (Lancaster et al., 2013), we are in an unprecedented position for tracing out pathways at multiple levels, unpicking effects of language-related gene variants on neurons and circuits, to help explain this elusive human trait.

## NOTE

1. Following standard nomenclature, genes are denoted in italics, proteins in regular font. Uppercase letters denote the human version of the gene (i.e., *FOXP2*), lowercase the mouse version of the gene (i.e., *Foxp2*).

## REFERENCES

- Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., ... Smith, S. M. (2018). Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage*, *166*, 400–424.
- Bearden, C. E., & Thompson, P. M. (2017). Emerging global initiatives in neurogenetics: The Enhancing Neuroimaging Genetics through Meta-analysis (ENIGMA) Consortium. *Neuron*, *94*(2), 232–236.

- Becker, J., Czamara, D., Scerri, T. S., Ramus, F., Csepe, V., Talcott, J. B., ... Schumacher, J. (2014). Genetic analysis of dyslexia candidate genes in the European cross-linguistic NeuroDys cohort. *European Journal of Human Genetics*, 22(5), 675–680.
- Bigos, K. L., & Weinberger, D. R. (2010). Imaging genetics—Days of future past. *NeuroImage*, 53(3), 804–809.
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376.
- Carrion-Castillo, A., Franke, B., & Fisher, S. E. (2013). Molecular genetics of dyslexia: an overview. *Dyslexia*, 19(4), 214–240.
- Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., ... Doudna, J. A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*, 550(7676), 407–410.
- Chen, X. S., Reader, R. H., Hoischen, A., Veltman, J. A., Simpson, N. H., Francks, C., ... Fisher, S. E. (2017). Next-generation DNA sequencing identifies novel gene variants and pathways involved in specific language impairment. *Scientific Reports*, 7, 46105.
- Deriziotis, P., & Fisher, S. E. (2017). Speech and language: Translating the genome. *Trends in Genetics*, 33(9), 642–656.
- Deriziotis, P., O’Roak, B. J., Graham, S. A., Estruch, S. B., Dimitropoulou, D., Bernier, R. A., ... Fisher, S. E. (2014). De novo TBRI mutations in sporadic autism disrupt protein functions. *Nature Communications*, 5, 4954.
- Devanna, P., Chen, X. S., Ho, J., Gajewski, D., Smith, S. D., Gialluisi, A., ... Vernes, S. C. (2018). Next-gen sequencing identifies non-coding variation disrupting miRNA-binding sites in neurological disorders. *Molecular Psychiatry*, 23(5), 1375–1384.
- de Zubicaray, G. I., Chiang, M. C., McMahon, K. L., Shattuck, D. W., Toga, A. W., Martin, N. G., ... Thompson, P. M. (2008). Meeting the challenges of neuroimaging genetics. *Brain Imaging Behavior*, 2(4), 258–263.
- Egan, M. F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., ... Weinberger, D. R. (2003). The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*, 112(2), 257–269.
- Eicher, J. D., Powers, N. R., Miller, L. L., Akshoomoff, N., Amaral, D. G., Bloss, C. S., ... Pediatric Imaging, Neurocognition and Genetics Study. (2013). Genome-wide association study of shared components of reading disability and language impairment. *Genes, Brain and Behavior*, 12(8), 792–801.
- Eising, E., Carrion Castillo, A., VINO, A., Strand, E. A., Jakielski, K. J., Scerri, T. S., ... Fisher, S. E. (2018). A set of regulatory genes co-expressed in embryonic human brain is implicated in disrupted speech development. *Molecular Psychiatry*. Advance online publication. doi: 10.1038/s41380-018-0020-x.
- Field, L. L., Shumansky, K., Ryan, J., Truong, D., Swiergala, E., & Kaplan, B. J. (2013). Dense-map genome scan for dyslexia supports loci at 4q13, 16p12, 17q22; suggests novel locus at 7q36. *Genes, Brain and Behavior*, 12(1), 56–69.
- Fischbach, G. D., & Lord, C. (2010). The Simons Simplex Collection: A resource for identification of autism genetic risk factors. *Neuron*, 68(2), 192–195.
- Fisher, S. E. (2006). Tangled webs: Tracing the connections between genes and cognition. *Cognition*, 101(2), 270–297.
- Fisher, S. E. (2017). Evolution of language: Lessons from the genome. *Psychonomic Bulletin and Review*, 24(1), 34–40.
- Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics*, 25(4), 166–177.
- Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics*, 18(2), 168–170.
- Ge, T., Chen, C. Y., Neale, B. M., Sabuncu, M. R., & Smoller, J. W. (2017). Phenome-wide heritability analysis of the UK Biobank. *PLOS Genetics*, 13(4), e1006711.
- Gialluisi, A., Newbury, D. F., Wilcutt, E. G., Olson, R. K., DeFries, J. C., Brandler, W. M., ... Fisher, S. E. (2014). Genome-wide screening for DNA variants associated with reading and language traits. *Genes, Brain and Behavior*, 13(7), 686–701.
- Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W., Willemsen, M. H., ... Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, 511(7509), 344–347.
- Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), 333–351.
- Grabitz, C. R., Button, K. S., Munafò, M. R., Newbury, D. F., Pernet, C. R., Thompson, P. A., & Bishop, D. V. M. (2018). Logical and methodological issues affecting genetic studies of humans reported in top neuroscience journals. *Journal of Cognitive Neuroscience*, 30(1), 25–41.
- Graham, S. A., & Fisher, S. E. (2015). Understanding language from a genomic perspective. *Annual Review of Genetics*, 49, 131–160.
- Guadalupe, T., Mathias, S. R., vanErp, T. G. M., Whelan, C. D., Zwiers, M. P., Abe, Y., ... Francks, C. (2017). Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. *Brain Imaging Behavior*, 11(5), 1497–1514.
- Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., ... Weinberger, D. R. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science*, 297(5580), 400–403.
- Harlaar, N., Meaburn, E. L., Hayiou-Thomas, M. E., Wellcome Trust Case Control, C., Davis, O. S., Docherty, S., ... Plomin, R. (2014). Genome-wide association study of receptive language ability of 12-year-olds. *Journal of Speech, Language, and Hearing Research*, 57(1), 96–105.
- Hibar, D. P., Stein, J. L., Renteria, M. E., Arias-Vasquez, A., Desrivieres, S., Jahanshad, N., ... Medland, S. E. (2015). Common genetic variants influence human subcortical brain structures. *Nature*, 520(7546), 224–229.
- Hoogman, M., Guadalupe, T., Zwiers, M. P., Klarenbeek, P., Francks, C., & Fisher, S. E. (2014). Assessing the effects of common variation in the FOXP2 gene on human brain structure. *Frontiers in Human Neuroscience*, 8, 473.
- Iossifov, I., O’Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., ... Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, 515(7526), 216–221.
- Kornilov, S. A., Rakhlin, N., Kopolov, R., Lee, M., Yrigollen, C., Caglayan, A. O., ... Grigorenko, E. L. (2016). Genome-wide association and exome sequencing study of language

- disorder in an isolated population. *Pediatrics*, 137(4), e20152469.
- Lai, C. S., Fisher, S. E., Hurst, J. A., Levy, E. R., Hodgson, S., Fox, M., ... Monaco, A. P. (2000). The SPCH1 region on human 7q31: Genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *American Journal of Human Genetics*, 67(2), 357–368.
- Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413(6855), 519–523.
- Lancaster, M. A., Renner, M., Martin, C. A., Wenzel, D., Bicknell, L. S., Hurles, M. E., ... Knoblich, J. A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, 501(7467), 373–379.
- Loose, M. W. (2017). The potential impact of nanopore sequencing on human genetics. *Human Molecular Genetics*, 26(R2), R202–R207.
- Luciano, M., Evans, D. M., Hansell, N. K., Medland, S. E., Montgomery, G. W., Martin, N. G., ... Bates, T. C. (2013). A genome-wide association study for reading and language abilities in two population cohorts. *Genes, Brain and Behavior*, 12(6), 645–652.
- MacDermot, K. D., Bonora, E., Sykes, N., Coupe, A. M., Lai, C. S., Vernes, S. C., ... Fisher, S. E. (2005). Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *American Journal of Human Genetics*, 76(6), 1074–1080.
- Mahowald, K., & Fedorenko, E. (2016). Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. *NeuroImage*, 139, 74–93.
- Metzker, M. L. (2010). Sequencing technologies—The next generation. *Nature Reviews Genetics*, 11(1), 31–46.
- Meyer, M., Kircher, M., Gansauge, M. T., Li, H., Racimo, F., Mallick, S., ... Paabo, S. (2012). A high-coverage genome sequence from an archaic Denisovan individual. *Science*, 338(6104), 222–226.
- Mueller, K. L., Murray, J. C., Michaelson, J. J., Christiansen, M. H., Reilly, S., & Tomblin, J. B. (2016). Common genetic variants in FOXP2 are not associated with individual differences in language development. *PLOS One*, 11(4), e0152576.
- Newbury, D. F., & Monaco, A. P. (2010). Genetic advances in the study of speech and language disorders. *Neuron*, 68(2), 309–320.
- Nudel, R., Simpson, N. H., Baird, G., O'Hare, A., Conti-Ramsden, G., Bolton, P. F., ... Newbury, D. F. (2014). Genome-wide association analyses of child genotype effects and parent-of-origin effects in specific language impairment. *Genes, Brain and Behavior*, 13(4), 418–429.
- Paabo, S. (2014). The human condition—A molecular approach. *Cell*, 157(1), 216–226.
- Padovani, A., Cosseddu, M., Premi, E., Archetti, S., Papetti, A., Agosti, C., ... Borroni, B. (2010). The speech and language FOXP2 gene modulates the phenotype of fronto-temporal lobar degeneration. *Journal of Alzheimer's Disease*, 22(3), 923–931.
- Peter, B., Raskind, W. H., Matsushita, M., Lisowski, M., Vu, T., Berninger, V. W., ... Brkanac, Z. (2011). Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. *Journal of Neurodevelopmental Disorders*, 3(1), 39–49.
- Pinel, P., Fauchereau, F., Moreno, A., Barbot, A., Lathrop, M., Zelenika, D., ... Dehaene, S. (2012). Genetic variants of FOXP2 and KIAA0319/TTRAP/THEM2 locus are associated with altered brain activation in distinct language-related regions. *Journal of Neuroscience*, 32(3), 817–825.
- Premi, E., Pilotto, A., Alberici, A., Papetti, A., Archetti, S., Seripa, D., ... Borroni, B. (2012). FOXP2, APOE, and PRNP: New modulators in primary progressive aphasia. *Journal of Alzheimer's Disease*, 28(4), 941–950.
- Prufer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., ... Paabo, S. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, 505(7481), 43–49.
- Reuter, M. S., Riess, A., Moog, U., Briggs, T. A., Chandler, K. E., Rauch, A., ... Zweier, C. (2017). FOXP2 variants in 14 individuals with developmental speech and language disorders broaden the mutational and clinical spectrum. *Journal of Medical Genetics*, 54(1), 64–72.
- Roeske, D., Ludwig, K. U., Neuhoff, N., Becker, J., Bartling, J., Bruder, J., ... Schulte-Korne, G. (2011). First genome-wide association scan on neurophysiological endophenotypes points to trans-regulation effects on SLC2A3 in dyslexic children. *Molecular Psychiatry*, 16(1), 97–107.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510), 421–427.
- Schoffelen, J. M., Hulten, A., Lam, N., Marquand, A. F., Udden, J., & Hagoort, P. (2017). Frequency-specific directed interactions in the human brain network for language. *Proceedings of the National Academy of Sciences*, 114(30), 8083–8088.
- Slon, V., Hopfe, C., Weiss, C. L., Mafessoni, F., de la Rasilla, M., Lalueza-Fox, C., ... Meyer, M. (2017). Neandertal and Denisovan DNA from Pleistocene sediments. *Science*, 356(6338), 605–608.
- Sollis, E., Graham, S. A., Vano, A., Froehlich, H., Vreeburg, M., Dimitropoulou, D., ... Fisher, S. E. (2016). Identification and functional characterization of de novo FOXP1 variants provides novel insights into the etiology of neurodevelopmental disorder. *Human Molecular Genetics*, 25(3), 546–557.
- St. Pourcain, B., Cents, R. A., Whitehouse, A. J., Haworth, C. M., Davis, O. S., O'Reilly, P. F., ... Davey Smith, G. (2014). Common variation near ROBO2 is associated with expressive vocabulary in infancy. *Nature Communications*, 5, 4831.
- St. Pourcain, B., Eaves, L. J., Ring, S. M., Fisher, S. E., Medland, S., Evans, D. M., & Davey Smith, G. (2018). Developmental changes within the genetic architecture of social communication behavior: A multivariate study of genetic variance in unrelated individuals. *Biological Psychiatry*, 83(7), 598–606.
- St. Pourcain, B., Robinson, E. B., Anttila, V., Sullivan, B. B., Maller, J., Golding, J., ... Davey Smith, G. (2018). ASD and schizophrenia show distinct developmental profiles in common genetic overlap with population-based social communication difficulties. *Molecular Psychiatry*, 23(2), 263–270.
- Sullivan, P. F. (2010). The psychiatric GWAS consortium: Big science comes to psychiatry. *Neuron*, 68(2), 182–186.

- Veltman, J. A., & Brunner, H. G. (2012). De novo mutations in human genetic disease. *Nature Reviews Genetics*, *13*(8), 565–575.
- Villanueva, P., Nudel, R., Hoischen, A., Fernandez, M. A., Simpson, N. H., Gilissen, C., ... Newbury, D. F. (2015). Exome sequencing in an admixed isolated population indicates NFXL1 variants confer a risk for specific language impairment. *PLOS Genetics*, *11*(3), e1004925.
- Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: Biology, function, and translation. *American Journal of Human Genetics*, *101*(1), 5–22.
- Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002). Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, *125*(3), 452–464.
- Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., ... Gadian, D. G. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain*, *125*(3), 465–478.
- Wei, W. Q., & Denny, J. C. (2015). Extracting research-quality phenotypes from electronic health records to support precision medicine. *Genome Medicine*, *7*(1), 41.
- Wilcke, A., Ligges, C., Burkhardt, J., Alexander, M., Wolf, C., Quente, E., ... Kirsten, H. (2012). Imaging genetics of FOXP2 in dyslexia. *European Journal of Human Genetics*, *20*(2), 224–229.
- Zaaijer, S., Gordon, A., Speyer, D., Piccone, R., Groen, S. C., & Erlich, Y. (2017). Rapid re-identification of human samples using portable DNA sequencing. *eLife*, *6*, e27798.
- Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*, *9*(5), 429–436.