Somewhere Between Chomsky and Tomasello: The Genetic Component of Human Language

Morten R. S. Andreasen

1. Introduction

Is language something we learn by listening to adults as children or does it lie encoded within our DNA so that we can't help but learn it under the right circumstances? These are the extremes of Michael Tomasello's Usage-Based theory of language and Noam Chomsky's model of Universal Grammar.

Ever since Noam Chomsky put his theory of Universal Grammar (UG) forth, which posits an innate, genetically determined language faculty (Chomsky 1986, 3), it has been a topic of debate. Could language really be as innate as learning to walk? Maybe even a singular language gene could be identified? If only it were as simple as some thought when the FOXP2 gene was pinpointed in the KE family in the 1990's. Roughly half of the family exhibited verbal dyspraxia that "affected expression and articulation of language" (Vargha-Khadem 20025, 131). The hereditary pattern found in the family helped researchers trace the mutation to the FOXP2 gene (Vargha-Khadem 2005, 132). But, as with most things, language turned out to be more complicated, as the FOXP2 gene began showing up in other non-language related areas. Even though the KE family was not the magical key to a specific language gene, they highlight the importance of family and twin studies by showcasing the heritability of language and language impairment. By looking at the occurrence of traits in mono- (MZ) and dizygotic (DZ) twins, heritability of different traits can be estimated. While twin studies, which will be explained in further detail under section 3.0, hint at there being some genetic component in language, they do not provide any specific candidate genes, nor do they show any support for a polygenic model of language in which different genes across the genome contribute to the language faculty. The arrival of more modern analytical tools like Genome-Wide Association Studies (GWAS) help uncover otherwise unknown associations between traits and genetic variants (Visscher et. al. 2017, 5). In language impairment studies GWAS helps to identify genetic variations involved in specific language impairments (SLI) and in nonspecific language impairments (NLI) in otherwise healthy individuals. These impairments are defined by Rice et. al. as SLI being a "language disorder that delays the mastery of language skills in children who have no hearing loss, intellectual impairment, or other developmental delays" (Rice et. al. 2020, 794) and NLI "children with impairment in both language and nonverbal cognition" (Rice et. al. 2020, 794). Understanding the genetic side of such impairments might help broaden the understanding of the genetic underpinnings of language.

Leviathan: Interdisciplinary Journal in English (ISSN: 2446-3981), No. 11, 2026.

2. Key questions on the genetic component of human language

The following will investigate concepts of heritability, candidate genes and Genome-Wide Association Studies in language and language impairments to uncover whether human language seems to be genetically dependent and if so, whether anything can be said of the degree to which language is genetically dependent. It will also investigate how specific genes seem to be related to SLI and if they seem to be linked to other bodily functions[1] or solely to SLI. Lastly it will look at how multiple loci and genes across the genome seem to be integral in normal functioning language as impairment seem to stem from multiple genetic areas. Lastly these insights will be used to discuss the ideas of Noam Chomsky's Universal Grammar and Michael Tomasello's Usage-Based theory arriving at a polygenic view on the genetic component of human language.

3. Evidence for genetic influence in twins

MZ and DZ twins can be used to look for whether a trait seems to be genetic or environmentally conditioned by looking at the variation in behaviour between individuals. If a trait is genetically influenced, the behaviour of MZ twins, who share an identical genetic make-up, should be similar every time (Mountford and Newbury 2019, 34). I.e. if one twin has blue eyes the other should also have blue eyes. As DZ twins only share about half their genetic make-up, as regular non-twin siblings, their outcome should be repeated half of the time (Mountford and Newbury 2019, 34). If, on the other hand, a trait is purely produced by a specific environment shared by a set of twins, the pattern of results should be similar every time no matter the zygosity of a given twin pair. And lastly if a trait stems from environmental factors unique to the single child there should be no correlation between trait and zygosity (Mountford and Newbury 2019, 34). Twin studies make it possible to bypass the problem of genetic difference between individuals and make it possible to draw precise conclusions on which parts of our behaviour can be explained as nature and what might be nurture.

3.1 Vocabulary development

There seems to be higher correlation between MZ twins compared to DZ twins in the development of vocabulary (Stromswold 2001, 668) hinting at there being some heritable component in vocabulary development. As much as 29% of vocabulary size in late infancy, up to 24 months of age in this review, could be attributed to genetic factors while shared environment accounts for 66% and non-shared environment accounts for just 5% (Stromswold 2001, 669). According to the same meta-review, the genetic component appears to gain influence as children progress and acquire a larger vocabulary from late infancy up to 12 years of age. At this stage, genetic factors account for 53% of the variance in children's vocabulary, while shared environment accounts for 18% and non-shared environment for 29% (Stromswold 2001, 669). Although the analytical methods used in the meta-review vary in how they partition genetic and

[1] Bodily functions understood as the fundamental physical or psychological processes necessary for sustaining homeostasis, enable perception, movement and cognitive ability.

environmental influences, they converge on the same overall pattern reflected in the percentages presented here; a substantial proportion of the variance in children's vocabulary is attributable to heritable factors.

3.2 Phonological abilities

Five studies in the review investigated the heritability of phonological capabilities. Among others the phonological awareness of 6–7-year-old twins were tested. These tests included the ability to make up words when a phoneme is removed, for example: children hear the word /sid/ and then repeat it without the final phoneme /d/ (Stromswold 2001, 671). Another tested the ability to find the odd one out in a sequence of rhyming words such as *cat*, *hat*, *car*, *bat* (Stromswold 2001, 671). This, along with a list of additional tests was used to estimate how much of the variance observed could be attributed to genetic factors. Not only was the genetic component found to play a significant role in the phonological capabilities of 6–7-year-old twins, it also seems to increase with age. In 6-year-olds the genetic component was found to account for anywhere between 52% and 56% (Stromswold 2001, 671, 673). In 7-year-olds it seemed to account for anywhere between 62% and 78% (Stromswold 2001, 674), suggesting that there is not only a strong genetic component to the phonological capabilities in children but that it also seems to increase somewhat with age. It was also assessed whether the phonological capabilities were independent from the non-verbal intelligence. In 6-year-olds it was found that 29% of the phonological awareness could be linked to the same genetic factors affecting IQ, 23% were linked to vocabulary and morphosyntactic genetics and 9% were specific to genetics related to phonological awareness (Stromswold 2001, 673). In the 7-year-olds none of the variance was related to genetic factors involved in phonological awareness. Instead, it was mostly explained by genetics related to vocabulary and morphosyntax (Stromswold 2001, 673).

3.3 Morphology and syntax

Findings from 11 studies examining the morphosyntactic abilities in children aged 20 months to 13 years (Stromswold 2001, 676) suggested the presence of a genetic component. Across all studies, except four in which the difference between MZ and DZ twins were not statistically significant, the correlation was greater in monozygotic twins than dizygotic ones (Stromswold 2001, 680). Generally, Stromswold found that the more twin pairs a study included the more likely the researchers were to report significant MZ-DZ differences regarding abilities in morphosyntax. The tests used in the different studies on the heritability of morphosyntax reviewed by Stromswold vary considerably. One study analysed data collected from diaries recorded by parents in which everything their child said within an hour was recorded (Stromswold 2001, 675). Another study administered the Berko Wug[2] test monthly over the course of a year in a longitudinal study to assess the heritability of inflectional morphology (Stromswold 2001, 678). When filtering out the influence of non-verbal IQ, the genetic factors were still significant hinting at there being genetic factors for morphosyntax that are not related to the genetic factors that influence non-verbal intelligence (Stromswold

[2] In the Burko Wug test children are presented with nonce words, i.e. words that do not exist, they are then asked to inflect these nonce words, e.g. in the plural , such as *wug* to *wugs* (Levy 1987, 71).

2001, 681). Genetic factors influencing the syntactic abilities are seemingly strongly related to the ones that are partly responsible for vocabulary development (Stromswold 2001, 681). These findings are consistent with the findings of Hohnen and Stevenson on phonological abilities in children aged 6-7 (Hohnen & Stevenson 1999, 598), as mentioned earlier, whose findings indicated that IQ related factors account for part of the variance in phonological awareness, but are diminished from age 6 to 7 (Stromswold 2001, 673) supporting the idea that genetic factors affecting phonological awareness as well as morphosyntax are, to some extent, independent of general intelligence.

In general, it seems that in studies done on normal twins somewhere between 25% and 50% of the variance in linguistic performance can be said to be because of the genetic make-up. Stromwold's meta-review shows support for the hypothesis that genetics seem to play a vital role in the language of children without any disabilities or language impairments.

4. SLI, candidate genes and genetic mutations

4.1 Distinguishing SLI, NLI & nonverbal intelligence

Twin studies distinguishing between children with Specific Language Impairment (SLI) and Nonverbal Language Impairment (NLI) allow for research into whether specific language difficulties arise from specific genetic factors that cannot be reduced to deficits in the nonverbal intelligence. Furthermore, research into SLI also provides a unique opportunity for identifying candidate genes associated with language impairment, offering further insight into the genetic components of human language.

If language were only a result of general intelligence, studying children with SLI, with average intelligence, should show little to no genetic influence, whereas lower nonverbal intelligence should correlate directly with language impairment. However, research into SLI and NLI uncovered that "nonverbal IQ is not the same causal pathway as language impairment" (Rice et. al. 2020, 793). It was also found that low scores in the nonverbal cognition phenotype does not equate to lower relative scores in language performance, supporting the idea that language, to an extent, is separate from nonverbal IQ (Rice et. al. 2020, 810). Heritability of nonverbal cognition varies across language phenotypes (Rice et. 2020, al. 810) indicating partially independent genetic contributions to language ability. In other words, low nonverbal IQ does not necessarily lead to language impairment. On the other hand, clear heritability has been found among children with SLI across multiple language phenotypes (Rice et. al. 2020, 810) supporting the idea that language to some extent is genetically independent from nonverbal intelligence.

4.2 FOXP2

After the discovery of the FOXP2 gene, it was put in relation to developmental verbal dyspraxia as it seemed to impair speech in both expressive and receptive ability (Fisher and Scharff 2009, 166). In individuals in which the FOXP2 gene is mutated, fine coordinated movements of specific mouth muscles like the tongue and lips are impaired, resulting in verbal dyspraxia where individuals cannot produce specific sounds correctly (Fisher and Scharff 2009, 166). This happens as the mutated FOXP2, as shown in the infamous

KE family, impairs "orofacial musculature, particularly for movement sequences" (Vargha-Khadem et. al. 2005, 136). Not only were the affected members of the KE family impaired in areas of pronunciation, they were also impaired in areas of grammar, semantics and nonverbal intelligence (Vargha-Khadem et. al. 2005, 132). This indicates that the mutation of the FOXP2 gene also seems to be implicated in the language faculties as well as having implications for the orofacial coordination as all the affected KE members showed signs of impairment in this domain exemplified in speech as it requires "precise selection, coordination and timing of sequences of orofacial movements" (Vargha-Khadem et. al. 2005, 132). Because all affected individuals displayed the same pattern of impairment, FOXP2 may play an important role in the "development of (…) networks that are involved in the learning, planning and execution of orofacial (…) sequences" (Vargha-Khadem et. al. 2005, 133). And not only that. The FOXP2 gene also seems to be implicated in functions related to other organs such as the lungs, heart and gut (Vargha-Khadem et. al. 2005, 134). To take it even further, the FOXP2 has also been found in relation to illnesses such as schizophrenia to which it may contribute (Salmón-Gómez et. al. 2025, 210).

4.3 SETBP1

Another seemingly important gene in human language is SETBP1. In a 2012 case study a child was described as having severe oral dyspraxia and generally hit developmental milestones slightly later than in normal development (Marseglia et. al. 2012, 217). At preschool age, his fine motor skills were impaired, at three his communication was impaired in the sense that he started producing communicative sounds while his nonverbal skills were very poor. Lastly, at five his language was strongly impaired with verbal dyspraxia but generally showed a better comprehension of language than his expression (Marseglia et. al. 2012, 217). The receptive language was estimated to that of a four-year-old when he was 13 years of age while the spoken language was pretty much gone at the age of 16 (Marseglia et. al. 2012, 217). This all could support the hypothesis that SETBP1 is implicated in the expressive aspects of language. The case study also found that the gene seems to be implicated in other areas of human development as the child also showed intellectual disability, autistic features and "peculiar facial features" (Marseglia et. al. 2012, 217).

4.4 CNTNAP2

Finally, the CNTNAP2 gene seems to be implicated in syntactic processing and in early language development (Kos et. al. 2012, 2). When mutated, the gene has been found to be implicated in developmental language impairment in both autism as well as SLI (Kos et. al. 2012, 1). However, CNTNAP2 is not limited to functions in language. Much like the genes previously discussed, the mutated CNTNAP2 have been associated with other disorders such as Tourette's syndrome and more rarely in disorders linked to abnormal brain development, epilepsy and autistic spectrum disorders (Vernes et. al. 2008, 2341).

It is important to note that these genes are not specific to linguistic function, nor are they uniquely associated with any single disorder nor impairment. Rather, they serve to illustrate that language seems to be dependent on the involvement of a multitude of genes. Mutations in specific genes do not only affect linguistic domains such as vocalisation or syntactical understanding; their effects seem to be multifaceted and can affect other areas of the body such as facial features or nonverbal intelligence. This leaves implications for singular genes being influential in multiple areas of bodily function. Or more plainly: it is not so simple as a gene being language specific. Rather, they seem to be multifaceted and important for multiple functions.

5. Insights offered by Genome-Wide Association Studies

Even though the genetics of language might not be as simple as the Danish newspaper Berlingske Tidende made it out to be in their 2001 headline "Forskere finder sprog-gen" or "Researchers find language-gene" (Svennevig 2001, 1). The notion that language seems to be both heritable and that the genetic component seems to be quite large across different linguistic phenotypes seems to support the view that human language largely relies on a vast genetic infrastructure to function normally.

The employment of Genome-Wide Association Studies (GWAS) allows for the investigation into how multiple genes and loci contribute to language impairment and thus how language is a product of a cumulative effect of interacting genes, by looking at genetic influences across the genome rather than at singular candidate genes like FOXP2. Although GWAS has not been able to map every gene involved in language impairment, more and more genes get identified as somehow being involved in language impairments. A 2013 GWAS study presented the ZNF385D and COL4A2 as candidate genes implicated in reading disability as well as language impairment, while the gene NDST4 also showed strong relations to language impairment more specifically (Eicher et. al. 2013, 796). Seven other genes were also identified: all but one showed higher frequencies of the minor allele, meaning that the less common genetic variant occurred more often in affected individuals with reading disabilities, language impairment or comorbid cases than in unaffected controls (Eicher et. al. 2013, 796-797). Importantly this should not be interpreted as evidence for which genes are in control of language impairment but rather to show that a variety of different genes across chromosomes seem to be implicated.

The same multi-genetic influence affecting SLI was found in a 2019 Genome-Wide Association Study which revealed a list of new loci involved in SLI. By analysing the DNA of 14 Pakistani families who to some degree suffered from SLI, Andres et. al. found new loci across different chromosomes that are seemingly implicated in SLI across family members in the 14 different families (Andres et. al. 2019, 1274). The identified loci further helped identify new candidate genes (Andres et. al. 2019, 1283) of interest for future research that could help broaden the genetic picture of the causal relationship between language and genetics.

This leads to the idea that a variety of genes across chromosomes seem to be involved in the impairment of language, i.e. when language fails partly because of genetic mutations, it suggests that normally functioning language is dependent on a massive infrastructure of genetics that all contribute and in conjunction make up the genetic capability for language.

6. Discussion: implications for Universal Grammar and Usage-Based theory

The evidence of heritability and specific genetics seemingly involved in language and language impairment carry implications for both Noam Chomsky's idea of a universal grammar as well as Michael Tomasello's Usage-Based theory. In *Knowledge of Language as a Focus of Inquiry,* Chomsky defines Universal Grammar as "a characterization of the genetically determined language faculty" (Chomsky 1986, 3), i.e. an innate system that guides the acquisition of any language. This stands in contrast to Tomasello's Usage-Based model in which language is learned firstly via holophrases in which single linguistic symbols are used as complete utterances such as *Give!* Afterwards holophrases get combined to simple phrases like *Ball roll* (Tomasello 2001, 65-66). As children progress, they learn via imitative learning and entrenchment (Tomasello 2001, 72) i.e. using central parts of sentences uttered by adults and thus minimizing syntactic complexity such as *she close it* (Tomasello 2001, 71). After hearing the same things multiple times children seem to have a hard time going beyond what they have heard meaning sentences have become entrenched (Tomasello 2001, 72). As development progresses, children begin to form linguistic schemas as they begin to identify recurring patterns in the input they get from adult speakers. They categorize utterances according to communicative function enabling them to produce utterances not modelled in the input (Tomasello 2001, 73, 77).

If Chomsky's idea of Universal Grammar is to hold, one should be able to find genetic traces of it, and conversely, if Tomasello's behavioural approach is to hold, there should be no indication of genetic traces connected to language directly. As this paper argues, however, there firstly seems to be a highly and well documented heritable component to language as has been observed across numerous phenotypes such as vocabulary and morphosyntax. By testing the heritability of regular language and language impairment in twins, upwards of 60-70% of language ability is in some areas documented as owing to a genetic component rather than the environment. Furthermore, it is also argued that these percentages, to some extent, are separate from the nonverbal intelligence i.e. language can be impaired without also scoring lower than average on nonverbal intelligence tests. This further supports the argument of there being an innate or genetic component to human language faculties that is otherwise separate from other factors such as general intelligence. Furthermore, it is argued that singular genes have been positively correlated to specific language impairments in individuals who otherwise showcase normal intelligence and are not affected by other impairments such as hearing loss, developmental disorders etc. These correlations should not be interpreted as evidence that a single gene causes SLI, nor that such a gene is restricted to linguistic functions. However, the fact that single candidate genes can be associated with SLI seems to support the idea of there being genetic underpinnings involved in human language to some extent. Lastly, the use of GWAS broadens the scope in trying to uncover how genetics contribute to the language faculties. By mapping entire genomes of many people and looking for common variants in affected individuals, numerous loci and candidate genes have been identified as implicated in language impairment. Implicated loci have been identified across chromosomes providing both various candidate genes possibly involved in language as well as showcasing how the genetic component of language is seemingly spread across the entire genome. These findings seem to support the idea that the genetic component of human language is a result of a myriad of genetic influences in which no gene is responsible for singular aspects of language, nor is any gene specifically a language gene. Rather, single genes seem to be implicated, in multiple functions as seen with FOXP2, SETBP1 and CNTNAP2, as they are seemingly also implicated in other bodily functions.

The capacity for language does not seem to stem from the Usage-Based learning mechanisms that Tomasello describes, in which children begin with holophrases, imitate adult utterances and gradually construct linguistic schemas as discussed earlier. However, heritability estimates below 100% indicate that environmental factors necessarily contribute to language development. This does not imply that Tomasello's usage-based mechanisms must be adopted to account for the variance that genetics cannot explain, but simply that some aspects of language are shaped by environmental factors. Although the arguments presented here seemingly make a strong case for Chomsky's idea of a Universal Grammar, the notion of single, innate language faculty that uniformly guides acquisition does not receive direct support from the evidence reviewed in this article. Instead, the points made here are in support of a polygenic foundation for language. A notion of UG in which language is dependent on a polygenic structure that is partly independent of nonverbal intelligence but still allows room for environmental factors to play a part in aspects such as vocabulary, morphology etc. seems to be the most likely model of human language.

Reference List

Andres, Erin M., Huma Hafeez, Adnan Yousaf, Sheikh Riazuddin, Mabel L. Rice, Muhammad Asim Basra and Muhammad Hashim Raza. 2019. "A Genome-Wide Analysis in Consanguineous Families Reveals New Chromosomal Loci in Specific Language Impairment (SLI)." *European Journal of Human Genetics 27 (8): 1274-1285*. doi.org/10.1038/s41431-019-0398-1.

Chomsky, Noam. 1986. "Knowledge of Language as a Focus of Inquiry." *University of Pittsburgh*. https://sites.pitt.edu/~perfetti/PDF/Chomsky.pdf.

Eicher, J. D., N. R. Powers, L. L. Miller, N. Akshoomoff, D. G. Amaral, C. S. Bloss, O. Libiger, N. J. Schork, B. F. Darst, B. J. Casey, L. Chang, T. Ernst, J. Frazier, W. E. Kaufmann, B. Keating, T. Kenet, D. Kennedy, S. Mostofsky, S. S. Murray, E. R. Sowell, H. Bartsch, J. M. Kuperman, T. T. Brown, D. J. Hagler, A. M. Dale, T. L. Jernigan, B. St. Pourcain, G. Davey Smith, S. M. Ring and J. R. Gruen. 2013. "Genome-wide Association Study of Shared Components of Reading Disability and Language Impairment." *Genes, Brain and Behavior* 12 (8): 792-801. doi.org/10.1111/gbb.12085.

Fisher, Simon E. and Constance Scharff. 2009. "FOXP2 as a Molecular Window into Speech and Language." *Trends in Genetics* 25 (4): 166-177. doi.org/10.1016/j.tig.2009.03.002.

Hohnen, Bettina, and Jim Stevenson. 1999. "The Structure of Genetic Influences on General Cognitive, Language, Phonological, and Reading Abilities." *Developmental Psychology* 35 (2): 590-603. doi.org/10.1037/0012-1649.35.2.590.

Kos, Miriam, Danielle van den Brink, Tineke M. Snijders, Mark Rijpkema, Barbara Franke, Guillen Fernandez, and Peter Hagoort. 2012. "CNTNAP2 and Language Processing in Healthy Individuals as Measured with Erps." *PLoS ONE 7 (10)*. doi.org/10.1371/journal.pone.0046995.

Levy, Yonata. 1987. "The WUG Technique Revisited." *Cognitive Development 2* (1): 71-87. doi.org/10.1016/s0885-2014(87)90042-6.

Marseglia, Giuseppina, Maria Rosaria Scordo, Chiara Pescucci, Genni Nannetti, Elisabetta Biagini, Valeria Scandurra, Francesca Gerundino, Alberto Magi, Matteo Benelli and Francesca Torricelli. 2012. "372 KB Microdeletion in 18q12.3 Causing SETBP1 Haploinsufficiency Associated with Mild Mental Retardation and Expressive Speech Impairment." *European Journal of Medical Genetics* 55 (3): 216-221. doi.org/10.1016/j.ejmg.2012.01.005.

Mountford, Hayley S. and Dianne F. Newbury. 2019. "The Genetics of Language Acquisition." 1st ed. *Routledge*. https://www-taylorfrancis-com.ez.statsbiblioteket.dk/reader/read-online/22a17413-f2f0-4b43-b9a0-72d4f66fcf69/chapter/pdf?context=ubx.

Rice, Mabel L., Catherine L. Taylor, Stephen R. Zubrick, Lesa Hoffman, and Kathleen K. Earnest. 2020. "Heritability of Specific Language Impairment and Nonspecific Language Impairment at Ages 4 and 6 Years across Phenotypes of Speech, Language, and Nonverbal Cognition." *Journal of Speech, Language, and Hearing Research* 63 (3): 793-813. doi.org/10.1044/2019_jslhr-19-00012.

Salmón-Gómez, Gabriel, Paula Suárez-Pinilla, Esther Setién-Suero, Carlos Martínez-Asensi, and Rosa Ayesa-Arriola. 2025. "FOXP2 and Schizophrenia: A Systematic Review." *Journal of Psychiatric Research* 190: 205-215. doi.org/10.1016/j.jpsychires.2025.07.016.

Stromswold, Karin. 2001. "The Heritability of Language: A Review and Metaanalysis of Twin, Adoption, and Linkage Studies." *Language* 77 (4): 647-723. doi.org/10.1353/lan.2001.0247.

Svennevig, Birgitte. 2001. "Forskere Finder Sprog-Gen." *Berlingske Tidende*. https://apps-infomedia-dk.ez.statsbiblioteket.dk/mediearkiv/link?articles=AY268994.

Tomasello, Michael. 2001. "First Steps toward a Usage-Based Theory of Language Acquisition." *Cognitive Linguistics* 11 (1-2): 61-82. doi.org/10.1515/cogl.2001.012.

Vargha-Khadem, Faraneh, David G. Gadian, Andrew Copp, and Mortimer Mishkin. 2005. "Foxp2 and the Neuroanatomy of Speech and Language." *Nature Reviews Neuroscience* 6 (2): 131-138. doi.org/10.1038/nrn1605.

Vernes, Sonja C., Dianne F. Newbury, Brett S. Abrahams, Laura Winchester, Jérôme Nicod, Matthias Groszer, Maricela Alarcón, Peter L. Oliver, Kay E. Davies, Daniel H. Geschwind, Anthony P. Monanco and Simon E. Fisher. 2008. "A Functional Genetic Link between Distinct Developmental Language Disorders." *New England Journal of Medicine* 359 (22): 2337-2345. doi.org/10.1056/nejmoa0802828.

Visscher, Peter M., Naomi R. Wray, Qian Zhang, Pamela Sklar, Mark I. McCarthy, Matthew A. Brown, and Jian Yang. 2017. "10 Years of GWAS Discovery: Biology, Function, and Translation." *The American Journal of Human Genetics* 101 (1): 5-22. doi.org/10.1016/j.ajhg.2017.06.005.