Genetics in ASD

The state of research on the genetics of autism spectrum disorder: methodological, clinical and conceptual progress

Short Title: Genetics in ASD
Anne B. Arnett, Ph.D.\textsuperscript{a}
Sandy Trinh, M.S.\textsuperscript{a}
Raphael A. Bernier, Ph.D.\textsuperscript{a}

\textsuperscript{a}University of Washington Department of Psychiatry & Behavioral Sciences, Center on Human Development and Disability, Box 357920, Seattle, WA 98195, USA

Declarations of interest: none.

**Corresponding Author:**

Raphael A. Bernier, Ph.D.
University of Washington
Department of Psychiatry & Behavioral Sciences
Box 357920, Seattle, WA 98195
Rab2@uw.edu
206-685-7585

The state of research on the genetics of autism spectrum disorder: methodological, clinical and conceptual progress

**Highlights**

- Heterogeneity in autism spectrum disorder presents a challenge to clinical care and research.
- Genetic sequencing indicates disruptive gene variants and CNVs account for over 30% of ASD cases.
- A genotype-first approach offers scientific, clinical and psychosocial benefits.
- Whole genome sequencing promises further advances in identifying genetic causes of ASD.
- Advancements in genomic sequencing precipitate progress toward precision medicine care for ASD.
Abstract
Autism spectrum disorder (ASD) is a behaviorally heterogeneous disorder with a strong genetic component, as evidenced by decades of twin and family studies. In recent years, enhanced methods of genomic sequencing have revealed that structural variation and mutations to both coding and non-coding regions of single, candidate genes may account for more than 30% of ASD cases. The current review highlights a genotype-first approach that builds upon these molecular findings to parse the heterogeneity of ASD. Advantages of this approach include strong potential for precision medicine diagnosis and treatment, as well as opportunity to advance basic science research on neurodevelopmental disorders. Psychosocial benefits of identifying genetic subtypes of ASD have already been realized through social networking, comprehensive clinical phenotyping, and increased awareness among providers of rare genetic mutations.

Introduction
Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction and communication, as well as a presence of restricted or repetitive behaviors and interests [1]. Core social communication deficits range from difficulty with reciprocal conversations and reduced nonverbal communication to impaired understanding of social norms and lack of insight into social relationships [1]. The development and presentation of symptoms varies widely across individuals, rendering the ASD population etiologically and phenotypically heterogeneous [2]. In an effort to meet the individual diagnostic, treatment and prognostic needs of those with ASD, research has recently shifted focus from phenotypic subtypes to a genotype-first approach. The current review describes the reasons for this shift, as well as methodological advancements in genome sequencing, advantages of the genotype-first method, and future directions in the ultimate goal of precision medicine care for individuals with ASD.
A Failed Behavioral Subtype Approach to ASD

Attempts to parse ASD into meaningful subtypes began with the work of Wing and Gould [2], who identified three distinct phenotypes characterized by “aloof”, “active but odd”, and “passive” social profiles. Autism subtype classification has since undergone a number of modifications, most notably the inclusion of official nomenclature for four diagnostic subtypes in the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) [3]. Regrettably, these heuristic behavioral categorizations failed to demonstrate divergent validity with regard to course of illness or treatment response [4] and had poor clinician agreement [5]. Given these shortcomings, the current diagnostic guidelines provide a single diagnostic label of autism spectrum disorder [1], and efforts to find alternative ways of defining ASD subtypes have accelerated. In particular, biology has emerged as a powerful tool; and given significant advances into the genetics of ASD, the ability to parse subtypes based on genetics has amplified in recent years.

Heritability of ASD

Evidence for genetic influences on ASD has been established through twin and family studies, with heritability estimates ranging from approximately 50-90% [6-8] and rates of recurrence among non-twin siblings approaching 20% [9-12]. Heritability of a broader autism phenotype, including mild differences in social communication, cognition and executive functioning, suggests additive polygenic factors explain many ASD cases [13,14]. Genome-wide association studies (GWAS) provide further support for contributions of common allelic variance to the broader ASD phenotype, including single nucleotide polymorphisms (SNPs) and common CNVs [15].

De Novo Genetic Variants

Though crucial to establishing the role of genetics in ASD, GWAS and family studies largely capture variants of very small effect and thus lack power to elucidate the impact of
Genetics in ASD

specific genes and molecular genetic pathology on atypical neurodevelopment. To address this hurdle, Sebat and colleagues [16] utilized high resolution array sequencing to detect de novo copy number variations (CNVs) in affected children. This novel approach, which relied on genomic sequencing of family trios (i.e. an affected proband and two unaffected biological parents), increased statistical power to find an association between the ASD phenotype and meaningful, likely causal spontaneous mutations across a broad range of genetic loci. Initial results indicated a strong role of de novo CNVs in ASD, with rates up to ten times higher among affected versus unaffected individuals [16-18]. The de novo method was subsequently applied to whole exome sequencing, which identified mutations specifically within coding regions of the genome [19-21]. Further methodological enhancement using molecular inversion probe (MIP) sequencing of targeted genes has resulted in cost-effective, recurrent identification of hundreds of high-confidence ASD genes in large cohorts [18,20,22-25]. At the time of this publication, de novo gene mutations and CNVs are estimated to account for approximately 30% of simplex ASD cases [22], with notable enrichment among genes involved in chromatin remodeling, fragile X mental retardation protein binding, postsynaptic density protein encoding, CHD8 binding, and embryonic expression [22,26]. In addition, rare, inherited single nucleotide variants (SNVs) and CNVs explain unique phenotypic variance particularly among cognitively impaired males, consistent with an additive risk model [27].

A Genotype-First Approach to ASD

Identification of recurrently observed CNVs and disruptive gene variants with shared ontology has led to reappraisal of the heterogeneity problem in ASD. Where DSM-IV failed to parse variability at the phenotypic level, a new genotype-first [28] characterization of individuals at the etiological level provides a clear path to precision medicine approaches to diagnosis, treatment, and prognosis for those with ASD [29-31] (Figure 1). Accordingly, within the last several years, multidisciplinary collaborations among psychologists, physicians, geneticists and affected families have resulted in comprehensive phenotyping of the most common high-
Genetics in ASD

confidence ASD gene mutations. This genotype-first approach has promoted basic science and clinical research on ASD through identification of 1) ASD-specific genetic subtypes, 2) variant-specific phenotypic range and reciprocal phenotypes, 3) promising pharmacological treatment targets, and 4) psychosocial benefits to affected families.

**CHD8: ASD-specific subtype**

Heterogeneous disruptive mutations to *CHD8* were among the first to be described as an ASD subtype [32]. *CHD8* is involved in chromatin remodeling and its targets include many other genes that have been associated with ASD [33]. The *CHD8* genetic subtype is unique in that it is more specific to ASD (≥ 87%) than intellectual disability (~60%). Physical characteristics include macrocephaly, dysmorphic facial features, and gastrointestinal discomfort. Critically, the significance of this constellation of features would have been obscured in a clinical setting amidst a group of individuals with diverse etiologies. In contrast, the phenotype became clear when the *CHD8* cohort was examined together. A zebrafish model of *chd8* disruption demonstrates uncanny similarities, with increased head size and reduced gastrointestinal motility, absent of global embryonic delays [26,32]. Close evaluation of the morphant zebra fish and available human microarray revealed tissue-specific patterns of accelerated neural growth during the prenatal and postnatal periods [32].

**16p11.2 CNV: Reciprocal phenotype**

In contrast to the relatively narrow phenotype associated with variable disruptions to *CHD8*, the manifestation of the 16p11.2 subtype appears to depend on the type of CNV involved [34]. While 16p11.2 deletions are associated with hyporeflexia, macrocephaly and speech sound disorder, duplications show a contrasting pattern of hyperreflexia, microcephaly and tremor. Additional features, such as seizures, motor difficulties, sacral dimples and ASD, are common across both genotypes, with severity of comorbidities driven in part by additional genetic risk [35]. Variant specificity among ASD subtypes is invaluable to identification of meaningful genotype-phenotype correlations. The 16p11.2 subtype, which evidences both
Genetics in ASD

pleiotropy and equifinality, as well as interactive and additive effects of multiple genetic risk, is therefore a key player in efforts to further elucidate functional gene ontology [34].

SCN2A: Pharmacotherapy target

The SCN2A subtype is notable due to its striking potential for a precision medicine approach to treatment of ASD. SCN2A encodes a subunit of the neuronal voltage-gated sodium channel and is therefore critical to proper functioning of neurons through propagation of action potentials. Disruptions to SCN2A can lead to gain-of-function or loss-of-function effects, with the former primarily associated with infantile-onset seizures and the latter predominantly associated with ASD [36]. Of all the genetic ASD subtypes, the path from genes to brain and behavior is the clearest in SCN2A and thus presents a specific target for pharmacological intervention. Moreover, knowledge about the mechanism of impact in SCN2A disruptions has enormous potential to inform targeted neurodevelopmental research in other ASD populations, who may share common neurological differences despite differing etiology [37].

Psychosocial benefits

While genetically informed therapies are the ultimate goal, clinical and psychosocial benefits of the genotype-first approach have already been realized. Family groups associated with genetic subtypes of ASD have mobilized to collect and report their own data through social media and annual meetings, supporting one another with anecdotes about treatment strategies, stages of development, and shared experiences. These networks of exceptionally rare gene events have also been crucial in efforts to identify, recruit and enroll participants and have led to further characterization of genetic subtypes beyond those already described, including DYRK1A [38], ADNP [39,40], and FOXP1 [41]. With increased awareness and availability of literature about affected individuals, clinicians have growing access to resources to aid genetic counseling, clinical prognosis and treatment recommendations.

Advancements in Genomic Sequencing
Advancements in scientific methods continue to enhance our understanding of genotype-phenotype associations. Whole genome sequencing (WGS) has improved detection of previously unmeasured variants, particularly among non-coding, regulatory regions of DNA [42]. Recent studies of ASD cohorts report relatively high rates of de novo mutations in non-coding regions of the genome, as well as very small exome mutations (e.g. indels) involving both known [43] and previously undetected [44] ASD candidate genes. Importantly, these findings underscore the relevance of genes expressed in striatal circuitry [43]. Polygenic factors may also contribute to the impact of regulatory DNA mutations; this is suggested by a recent finding that paternal transmission of non-coding variants to offspring was associated with increased risk for ASD among two large, independent samples [45].

Additionally, postzygotic mosaic mutations (PMMs) may account for 1-2% of ASD cases in which a single gene mutation is not identified [46]. PMMs occur after fertilization, are rarely inherited, and may occur in only a subset of cells. For all these reasons, they are prime candidates for analysis of additive and mechanistic impacts of genomic differences. Currently, reports suggest that risk for ASD is associated with greater PMM density specifically proximal to splice sites [46], as well as PMMs involving genes highly expressed during the embryonic period and/or in the amygdala [47].

Future Directions

Progress in clinical care for ASD will continue to be contingent on multidisciplinary, collaborative research efforts. Efforts are currently underway to identify functional, genetic-ontological subtypes that may provide additional utility with regard to clinical intervention. This includes investigation of broader phenotypes associated with gene disruptions that share molecular properties, e.g. genes regulated by the CHD8 protein [26,33]. Identification of measurable neurological effects of gene disruptions, such as an electrophysiological (EEG) signature, could translate to meaningful ASD biomarkers that are essential for clinical treatment trials [48]. Meanwhile, development of pharmacotherapies for genetic subtypes, such as
Genetics in ASD

SCN2A, will require increased knowledge about the timing of genetic expression and reversibility of neurodevelopmental impairment. Thus, continued comprehensive phenotyping will be essential to the success of clinical trials for genetic subtypes of ASD [37].

Conclusions

Altogether, the behaviorally based DSM-5 diagnosis of ASD describes a formidably heterogeneous population whose medical and neuropsychological risks may be vastly different from one individual to the next. The genotype-first approach builds upon the massive gains in our understanding of ASD over the past decade and presents an opportunity to narrow the phenotype in a way that is scientifically and clinically useful. Of course, there is still a lot of work to be done. Few of the identified molecular pathways are thoroughly understood, and both gene-gene and gene-environment interactions are likely contributory, but not easily quantified among small samples [49]. Perhaps most germane, one cannot underestimate the value of the DSM-5 ASD diagnosis for accessing appropriate therapies at this time. While many individuals with high-confidence ASD gene disruptions have symptoms of social deficits, repetitive behaviors, and cognitive impairment that would benefit from behavioral strategies designed for ASD [50], these individuals may not fit well into the DSM-5 diagnostic “boxes” that facilitate access to evidence-based treatments. The clinical utility of a genotype-first approach will thus depend on future work examining variant genetic expression and function, as well as systemic change toward a precision medicine approach to treatment.

The authors, Anne Arnett, Sandy Trinh, and Raphael Bernier, declare no conflicts of interest.

Acknowledgments

This work was supported by a grant from the National Institutes of Health: R01MH100047 to R.A.B.
References


Genetics in ASD


Genetics in ASD


*This detailed investigation of 11 individuals with SCN2A mutations highlights phenotypic specificity associated with gain-of-function versus loss-of-function genetic variants. Specifically, the authors report that variants conveying increased Na,1.2 channel function are associated with infant-onset seizures, while variants conveying diminished channel function are associated with ASD.*


**Functional specificity of the SCN2A gene makes it a strong candidate for development of gene therapy for symptoms of ASD, cognitive impairment, and seizures. This article reviews recent progress toward this goal and outlines future directions.**


**Using whole genome sequencing (WGS), the authors identified *de novo* mutations to putative regulatory DNA, i.e. non-protein-coding regions, among individuals with ASD. Interestingly, these regulatory region mutations were located proximal to genes previously linked to ASD. Additionally, WGS revealed small *de novo* mutations and CNVs previously missed using traditional exome sequencing.**


*The authors utilized advancements in genomic sequencing, namely whole genome sequencing (WGS) to increase statistical power to detect new candidate genes for the pathology of ASD. A genotype-phenotype correlation was established between disruptions to ASD-risk genes and lower adaptive functioning.*

*Paternal, but not maternal, inheritance of structural variants in cis-regulatory genetic elements was found among offspring with ASD more frequently than from siblings. The authors describe a pattern wherein these variants were common among “intolerant genes.” The results were replicated across two large, independent cohorts (N=829 and N=1,771 families).


*Data from the Simons Simplex Collection was re-analyzed to identify postzygotic mosaic mutations (PMMs) in a large cohort of families. Results indicated a significant burden of PMMs among affected children, with high density of PMMs located close to splice sites and involving genes previously implicated in ASD.


Genetics in ASD

Fig 1