The genetic architecture of autism spectrum disorder (ASD) is itself a diverse allelic spectrum that consists of rare \textit{de novo} or inherited variants in hundreds of genes and common polygenic risk at thousands of loci. ASD susceptibility genes are interconnected at the level of transcriptional and protein networks, and many function as genetic regulators of neurodevelopment or synaptic proteins that regulate neural activity. So that the core underlying neuropathologies can be further elucidated, we emphasize the importance of first defining subtypes of ASD on the basis of the phenotypic signatures of genes in model systems and humans.

**Introduction**

The autism spectrum is a clinically heterogeneous class of neurodevelopmental disorders that has a strong genetic basis. During the last decade a substantial proportion of the genetic architecture of autism spectrum disorder (ASD) has come to light. These studies have delivered a trove of susceptibility genes. However, this growing “parts list” does not translate immediately into a mechanistic understanding of ASD. That mechanistic understanding could emerge once the specific effects of genes on neurodevelopment are understood. Here, we review and discuss the current efforts to put gene discoveries into a mechanistic framework.

**An Emerging Spectrum of Genetic Risk**

At one extreme of the genetic architecture of ASD are the monogenic disorders, in which a major contributor to risk is a single gene mutation or copy number variation (CNV). At the other extreme is polygenic risk that is measured as the sum of thousands of common risk alleles with small effects. The gap between these two extremes spans a broad spectrum of alleles that, to date, have not been well characterized. Here, we summarize three major components of genetic risk, including \textit{de novo} mutations, rare inherited variants, and common polygenic variation and illustrate these in Figure 1.

**De Novo Mutations**

Early studies established that \textit{de novo} structural variants (SVs) (Brandler et al., 2016; Sebat et al., 2007) and protein-altering point mutations (Iossifov et al., 2012; Neale et al., 2012; O’Roak et al., 2012; Sanders et al., 2012) contribute to risk for ASD. It is estimated that \textit{de novo} mutations of genes contribute in approximately 30% of cases, including 25% of boys and 45% of girls (Iossifov et al., 2014). As would be expected for variants with large effects, subjects that carry \textit{de novo} mutations have lower nonverbal IQs than subjects who do not (Iossifov et al., 2014).

More recently, statistical methods such as the transmission and \textit{de novo} association (TADA) (He et al., 2013) have been developed for testing the disease association of genes based on the frequency of \textit{de novo} and transmitted mutations observed in parent-offspring trios. By this method, “high confidence” autism susceptibility genes are typically defined as those that meet a false discovery rate (FDR) threshold of < 10%. The application of this approach to exomes or whole genomes in large family samples has since become the most effective strategy for identifying ASD susceptibility genes (De Rubeis et al., 2014; Iossifov et al., 2014; Sanders et al., 2015; Satterstrom et al., 2018).

As each new study is combined with the publicly available data from its predecessors, the number of ASD susceptibility genes grows. A recent study in a combined sample of 35,584 subjects (11,986 with ASD), has brought the number of high confidence genes to 102 (Satterstrom et al., 2018). Fifty-three genes had a greater frequency in ASD than in the broader category of neurodevelopmental disorders (NDDs), suggesting that genes such as \textit{GIGYF1}, \textit{KDM6B}, \textit{PTEN}, \textit{ANK2}, \textit{KMT5B}, \textit{KMT2C}, and \textit{CHD8} might have a more direct influence on social behavior. Here, unless otherwise specified, the term “ASD susceptibility genes” refers to the high confidence (FDR < 0.1) genes reported in papers by Sanders et al., (2015) and Satterstrom et al., (2018). For reference, we provide a list of susceptibility genes and CNVs in Table S1.

Most germline mutations happen during mitotic cell division of spermatogonial cells, a process that occurs at a constant rate in the gonad (Crow, 2000). Thus, a majority (70%) of \textit{de novo} mutations originate from the father and total rate of new mutations in offspring increase by 1–2 mutations per year with his age (Kong et al., 2012; Michaelson et al., 2012). (Figure 1A)

**Rare Inherited Variants**

A portion of genetic risk for ASD consists of rare variants that are inherited from a mother or father (Figure 1A). Some of the first
examples identified were large CNVs, which occur as recurrent \textit{de novo} mutations in the population, but because of variable levels of cognitive impairments, can be transmitted from a parent with mild or no symptoms to their offspring. These include duplications of 15q11-13 and 16p11.2 and deletions of 15q11.2 (Doornbos et al., 2009). Recently, whole-genome analysis of SVs in large samples has succeeded in capturing a broader array of inherited variants including deletions that disrupt individual genes or \textit{cis}-regulatory elements (Brandler et al., 2018). Evidence for the contribution of inherited protein truncating variants to ASD has also been found by exome sequencing of families (Iossifov et al., 2015; Krumm et al., 2015) Thus, a portion of the genetic architecture consists of rare coding variants with incomplete penetrance that are inherited from parents who do not meet criteria for ASD.

Recessive variants account for a small fraction of idiopathic ASD (Gamsz et al., 2013) and developmental delay, but they account for a larger proportion (up to 30%) of cases in consanguineous families (Martin et al., 2018). DNA sequencing in such families is an effective approach for identifying cases of ASD that are attributable to rare homozygous variants (Morrow et al., 2008; Scott et al., 2016). Sequencing of consanguineous families has identified novel candidate genes associated with autism (Yu et al., 2013). For example, Novarino et al., (2012) identified inactivating mutations in the gene branched-chain ketoacid dehydrogenase kinase (BCKDK) in consanguineous families with autism, epilepsy, and intellectual disability. Bckdk knockout mice show abnormal brain amino acid profiles and neurobehavioral deficits that respond to dietary supplementation.

**Common Polygenic Risk**

Current data from genome-wide association studies (GWASs) is consistent with a liability threshold model, in which many risk alleles contribute additively to the overall risk. Estimates of the heritability of ASD explained by common SNPs range from 17% (Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, 2017; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013) to 52% (Gaugler et al., 2014). A recent meta-analysis of 18,381 ASD cases and 27,969 controls has found credible evidence for five risk loci and seven additional loci shared with other psychiatric disorders (Grove et al., 2019). However, there are likely to be thousands of loci with very small effects. The contribution of all common SNPs to a trait can be summarized as a polygenic risk score (PRS) (Dudbridge, 2013). A PRS constructed from previous GWASs of ASD has been shown to be significantly associated with cases in independent cohorts (Grove et al., 2017). Surprisingly, multiple studies have found that the PRS for ASD correlates with higher educational attainment and IQ in the population (Brainstorm Consortium et al., 2018), opposite to what has been reported for PRSs of other neurodevelopmental disorders (Niemi et al., 2018; Sniekers et al., 2017). This result has led to fascinating speculation as to the potential causative relationship of high intelligence to ASD (Crespi, 2016), but a rigorous dissection of the relationship of ASD PRS with intelligence in the general population is needed.

**Complex Genetic Inheritance**

In subjects who carry \textit{de novo} CNVs or gene mutations, psychiatric diagnoses are variable, and typically only a subset meet criteria for ASD (De Rubeis and Buxbaum, 2015; Malhotra and Sebat, 2012). This suggests that the determinants of psychiatric traits are multifactorial even in the context of a large-effect variant. Here, we summarize the evidence for the joint contributions of multiple rare or common genetic variants to psychiatric risk to the individual.

**Oligogenic Effects (≥ 2 hits)**

Analysis of rare variants in ASD and NDD cohorts has found evidence for an oligogenic model in which variation in the clinical severity of known pathogenic mutations is influenced by the genomic burden of additional rare variants. A series of studies have demonstrated that the clinical outcomes of rare genetic disorders are influenced by rare variants in the genetic background. This was first observed for a large (500 kb) deletion of 16p12.1 that carries moderate risk for developmental disorders (Girirajan et al., 2010). Deletion carriers with developmental delay had an increased burden of secondary CNVs compared with deletion carriers in controls. Subsequent studies have shown that the clinical severity of subjects who carry CNVs at 1q21.1,
7q11.23, and 16p11.2 correlates with the number of secondary rare variants, and similar results have been observed for ASD subjects who carry gene-disrupting de novo mutations (Pizzo et al., 2018). These results are consistent with a joint contribution of multiple rare variants to risk for ASD and other NDDs.

### The Joint Effects of Rare Mutation and Polygenic Risk

The clinical outcome of individuals that carry a rare variant of large effect can also be influenced by the background of common polygenic variation. A recent study reported that compared with typically developing controls, individuals with ASD that carry de novo mutations have significantly increased PRSs for ASD (Weiner et al., 2017). A second study examined the polygenic contribution to risk in a large cohort of NDDs (Niemi et al., 2018). This study reported that individuals with NDDs who carry a reportable clinical diagnostic variant have a significantly increased PRS. The joint contributions of rare CNVs and PRSs are evident for other psychiatric disorders, such as schizophrenia. Bergen et al., (2018) have shown that the polygenic contribution in CNV carriers is inversely proportional to the effect size of the CNV, consistent with an attenuated contribution of common variation in subjects who carry a highly penetrant rare variant.

### The Nature of Gene Action in ASDs

As will be discussed in detail in the following section, multiple ASD genes are connected within transcriptional and protein interaction networks. Consequently, a genetic effect that originates from a single gene mutation has the potential to influence the function of other ASD genes and can thereby fan out broadly through a gene regulatory network. For example, a mutation that affects a protein that regulates gene expression in the brain can have numerous downstream effects that are mediated through other regulatory genes in trans (Figure 2A). Other types of ASD risk alleles, such as CNV or common polygenic risk, might differ in how they directly affect gene function (Figure 2B and 2C). However, all 3 categories of genetic risk are similar in that their effects can be broadly distributed across a gene regulatory network.

### Large CNVs Influence Traits through the Dosage Effects of Multiple Genes

Large pathogenic CNVs typically affect the copy number of dozens of genes. There are a few known microdeletion syndromes where clinical features are attributable to a major driver gene within the region. For example, clinical features of Angelman, Phelan-McDermid, and Smith-Magenis syndromes, which are caused by large deletions of multiple genes, can be recapitulated by point mutations in UBE3A (Fang et al., 1999), SHANK3 (Bonaglia et al., 2006), and RAI1 (Slager et al., 2003), respectively. However, for the majority of large CNVs, studies have not found evidence for the existence of a single driver gene (De Rubeis et al., 2014; Iossifov et al., 2014). Sanders et al., (2015) examined rates of gene mutations within large CNVs and without and found no evidence for an increased rate of de novo protein truncation mutations inside of large CNVs that contained > 7 genes, whereas there was a significantly increased rate of de novo mutation within smaller CNVs.

Evidence from transgenic model organisms suggests that developmental phenotypes associated with CNVs are influenced by the dosage effects of multiple genes. Oligogenic effects have been characterized in detail for one CNV that is frequently observed in ASD: the deletion and duplication of 16p11.2. Pizzo et al., 2018).

---

Figure 2. Forms of Gene Action in ASD

Network graphs represent interactions within gene regulatory networks that are affected by causal variants. Red nodes represent genes affected directly by a risk variant in an individual case of ASD (de novo mutations, CNVs, or common variants). Pink nodes represent other ASD susceptibility genes that are not mutated in the same individual but do interact closely with the primary gene mutation(s) in the network.

(A) A de novo mutation in a key regulatory gene. Many ASD genes encode regulatory proteins that control the expression of genes in the developing brain. Those target genes include other monogenic ASD genes. Thus, the effect of a single gene mutation can fan out quite broadly through a gene regulatory network.

(B) Large CNVs directly alter the dosage of dozens of genes. Thus, the network-level effects of a CNV are distributed more broadly than the example shown in (A).

(C) Polygenic risk is very broadly distributed across the genome and throughout gene regulatory networks.
et al., (2018) tested the effect of RNA-knockdown of genes on > 20 classic phenotypes used to assess development in Drosophila. This study found that combinations of multiple 16p11.2 genes could elicit a variety of effects on development of the eye, wing, and nervous system. Our group has studied the effects of 16p11.2 genes on craniofacial development and found that the CNV has significant effects on craniofacial structure that are conserved in humans, mice, and zebrafish and several genes have significant effects on growth of the skull (Qiu et al., 2019). These observations are consistent with developmental phenotypes associated with a CNV being attributable to the dosage effects of multiple genes.

If the influence of a CNV on a trait is attributable to multiple genes, this suggests that effects of the CNV on genetic regulatory networks could be to some degree more complex than the monogenic example described above. In contrast to a gene mutation in which the network level effects radiate outward from a single point, a large CNV contributes multiple rare “hits” that have direct effects that are distributed more broadly across a gene regulatory network (Figure 2B).

Polygenic Effects
The nature of polygenic risk is fundamentally different from that of rare variants of large effect. Polygenic risk consists of thousands of small effects that segregate independently in the population. Thus, a PRS that is defined for a specific trait represents genetic effects that originate from thousands of different points within a genetic regulatory network (Figure 2C) and propagate even more broadly through effects mediated in trans through downstream regulatory genes.

ASD Susceptibility Genes Are Interconnected within Gene Regulatory Networks
Here, we describe how ASD genes show significant functional convergence within the context of gene networks, synaptic function, and signaling pathways; see Figure 3.

Cis-Regulatory Targets of ASD Genes Are Enriched for Other ASD Genes
A subset of ASD susceptibility genes encode proteins that bind directly to DNA or RNA and are involved in the regulation of gene expression, referred to as DNA binding proteins (DBPs) and RNA binding proteins (RBPs). For example, TBR1 encodes a transcription factor that binds DNA and regulates gene expression in the developing brain (Bedogni et al., 2010; Hevner et al., 2001). Chromatin immunoprecipitation sequencing (ChIP-seq) of TBR1 in the developing neocortex found that this protein binds next to high-confidence ASD genes more frequently than other

---

Figure 3. Points of Convergence among ASD Susceptibility Genes
Multiple ASD genes interact within the context of gene regulatory networks. These are highlighted as biological processes within a single neuron (cytoplasm in pink and nucleus in blue). Convergence is evident at multiple levels of interaction including DNA binding, RNA binding, and protein-protein interactions. Biological processes that are associated with ASD genes include the regulation of gene expression and synaptic function. ASD genes are expressed preferentially in the developing brain. Rare gene mutations in ASD also converge upon specific signaling pathways involved in the regulation of cell proliferation and differentiation including mTOR, MAPK, and Wnt signaling.

1290 Cell 178, September 5, 2019
transcription factors that are expressed in the brain (Notwell et al., 2016). Similarly, the chromatin remodeling factor CHD8 binds and positively regulates other ASD genes (Cotney et al., 2015; Sugathan et al., 2014).

Some ASD susceptibility genes encode RBPs that are involved in post-transcriptional regulation of messenger RNA abundance, splicing, or translation. A seminal example is the fragile X mental retardation protein (FMRP) (Darnell et al., 2011) and other RBPs that are associated with ASD include CELF4 (Wagnon et al., 2012), ELAVL3 (Ince-Dunn et al., 2012), and GIGYF1 (Peter et al., 2017). RNA targets of both FMRP (Darnell et al., 2011) and CELF4 (Wagnon et al., 2012) were reported to be significantly enriched for ASD genes, and ASD candidate genes were further enriched among the common targets of FMRP and CELF4 (Wagnon et al., 2012). By compiling the binding sites for 26 regulatory proteins associated with ASD, a recent study reported that the ASD genes that are enriched among these regulatory targets consist of other regulatory genes. By contrast, they did not find a significant enrichment for ASD genes involved in “neuronal communication” (Satterstrom et al., 2018).

**ASD Gene Mutations Lead to Dysregulation of Other ASD Genes In Trans**

The studies described above suggest that gene mutations that disrupt DBPs or RBPs can directly affect the regulation of numerous target genes in cis. The effects of a rare gene mutation can be propagated further by downstream regulatory proteins. Indeed, evidence for the enrichment of ASD genes among genes that are dysregulated in induced pluripotent stem cell (iPSC) models of CHD8 (Sugathan et al., 2014) or FOXP1 (Araujo et al., 2015) was strongest for genes that were dysregulated in trans (i.e., their respective transcripts were differentially expressed in mutant lines, but the genes were not identified as a direct binding site of the transcription factor). Similarly, transcriptional analysis of developing brain in transgenic mouse models of SETD5 (Pizzo et al., 2018), FOXP1 (Araujo et al., 2015), and TBR1 (Notwell et al., 2016) suggests that mutations in these genes leads to the dysregulation of ASD genes in trans.

Gene expression alterations from idiopathic ASD overlap with developmentally regulated genes implicated in cortical patterning, cell cycle, proliferation, and neural differentiation. Transcriptome analysis of neural progenitor cells (NPCs) derived from idiopathic ASD with macrocephaly found that differentially expressed genes were significantly enriched for ASD genes and genes within ASD-associated CNVs (Marchetto et al., 2017). Up-regulation of immune genes and downregulation of synaptic genes was observed in postmortem brain from idiopathic ASD (Voineagu et al., 2011), and this finding has been replicated in a larger ASD sample (Gupta et al., 2014). Recent analyses of much larger cohorts of postmortem brain collections from major psychiatric disorders such as autism, schizophrenia, bipolar disorder, and depression assembled by the PsychEncode Consortium et al., (2015) have lent further support to previous observation by identifying upregulation of immune-microglia and mitochondrial modules, and downregulation of neuronal and synaptic modules in ASD and schizophrenia (Gandal et al., 2018).

**Protein-Protein Interactions**

Analysis of ASD genes within protein-protein interaction (PPI) networks (Corominas et al., 2014; Gilman et al., 2011; O’Roak et al., 2012; Pinto et al., 2014; Sakai et al., 2011) has shown that proteins encoded by ASD genes have “high connectivity,” meaning that, compared with genes that are selected at random or genes that are mutated in controls, these proteins are more closely connected in a PPI network. Studies of large CNVs in ASD have shown that genes within ASD CNVs are highly connected with each other (Corominas et al., 2014) and with other ASD genes (Pinto et al., 2014). Pathways implicated in the above studies included synaptic transmission, chromatin remodeling, transcriptional regulation, translational regulation, ion transport, and cell adhesion.

**ASD Genes Are Co-Expressed During Fetal Brain Development**

Analysis of spatiotemporal expression of genes has revealed that ASD genes as a group are preferentially expressed in late mid-fetal prefrontal cortex, and have concentrated expression in layer 5/6 cortical projection neurons (Willesey et al., 2013). Another study has identified several developmentally co-regulated modules of gene expression in the fetal brain that are enriched in ASD risk genes, with layer 2-4 and glutamatergic projection neurons showing the strongest enrichment (Parikshak et al., 2013). Likewise, co-expression of genes within ASD CNVs was reported to be enriched in the developing cortex (Lin et al., 2015). Altogether, these observations highlight the relevance of early fetal brain development in the pathophysiology of ASD. Recently, common polygenic heritability of ASD was found to be significantly enriched in genes and enhancer marks that are expressed in the developing brain and had a similar enrichment of cortical cell types (Grove et al., 2017). Thus, common variation in ASD also preferentially affects genes expressed in the fetal cortex during embryonic development.

**Convergence of ASD Genes within Cell Signaling Pathways**

The regulation of cell-proliferation pathways in the developing brain has become another point of convergence in ASD (Ennist, 2016), specifically mammalian target of rapamycin (mTOR), mitogen activated protein kinase (MAPK) and Wnt signaling. Early studies made the observation that syndromes caused by the disruption of genes involved in mTOR signaling, including PTEN, TSC1, TSC2, and NF1 were frequently associated with brain overgrowth and ASD (Wang and Doering, 2013). Subsequently, inhibitors of mTOR have been proposed as treatments for such syndromes and one drug Everolimus (Novartis) has been approved for treatment of seizures in tuberous sclerosis (French et al., 2016).

Mutations in a set of genes involved in RAS-MAPK signaling form a set of syndromes known as “Rasopathies,” which are frequently associated with features of ASD (Adviento et al., 2014). MAPK (Pinto et al., 2010) and Rho GTPase (Lin et al., 2015) signaling has been implicated by gene set enrichment analysis of CNVs associated with ASD. The dysregulation of MAPK signaling has also been reported in transgenic mice carrying the ASD-associated microdeletion of 16p11.2 (Blizinsky et al., 2016; Pucilowska et al., 2015) and in other mouse models of social impairment (Faridar et al., 2014).

Dysregulation of Wnt signaling has been reported in multiple transgenic models of ASD genes. For example, knockdown of CHD8 in adult mice leads to a disruption of Wnt signaling and abnormal NPC proliferation (Durak et al., 2016). Mutations that
disrupt \textit{ARID1B} result in activation of Wnt signaling and dysregulation of B-catenin transcription (Vasileiou et al., 2015). Recent studies of \textit{SETDS} in mouse fetal tissue and human cells have reported that loss of \textit{SETDS} results in the activation of neuronal genes and Wnt signaling (Deliu et al., 2018).

Getting to the “Core” of Autism

Defining the Core ASD genes

Our current knowledge of the neurobiological basis of ASD is derived almost exclusively from studies of rare variants. Some have cautioned that the genes and pathways that are implicated by rare variants might not represent the core biology of a complex trait (Wray et al., 2018). However, it is not yet clear whether common variant associations in ASD will tell a different story. Indeed, this seems unlikely. From the perspective of both rare variants and common variants, our current understanding of the genetics of ASD is generally compatible with an “omnigenic” model of gene action recently proposed by Boyle et al. (2017). This model posits that the genetic basis of complex traits is polygenic to such an extent that it becomes difficult to distinguish “core” genes that have direct effects on a trait from the many “peripheral” genes that act indirectly and which might have a distant relationship to the core biological processes. Core genes, according to Liu et al. (2018), are defined as those that have direct effects on a trait and which are not mediated through the regulation of other genes. For psychiatric traits, core genes might include those that have direct effects of neural activity such as neurotransmitter receptors and ion channels. By contrast, peripheral genes are defined as all genes that have effects that are mediated indirectly, for example through gene regulation. One of the major predictions of the omnigenic model is that a majority of trait heritability is explained by peripheral genetic effects that are mediated indirectly, for example through gene networks and genetic variation in core genes explains a small fraction of the overall heritability (Liu et al., 2018). The authors further assert that studies of rare variants might be more effective than GWAS for identifying core genes.

Critiques of the omnigenic model have been published elsewhere (Cox, 2017; Wray et al., 2018), but these have not addressed the model specifically with respect to ASDs or NDDs. From the perspective of ASD, major predictions of the model hold true. With respect to psychiatric traits, the most difficult aspect of the model to reconcile in our view is the rigid dichotomy of core and peripheral genes. Unlike some other complex traits, for example serum lipid levels, which are controlled by a set of proteins involved in lipid transport and metabolism (Cox and Garcia-Palmieri, 1990), psychiatric traits cannot be easily reduced to dysfunction in a single biological pathway. Traits like social motivation, anxiety, and aggression don’t emerge directly from a handful of gene products. Psychopathology emerges from dysfunction at the level of neural circuitry, and circuit function arises through a process of development in which multiple biological processes are tightly regulated (Polleux et al., 2007). Given that the regulation of neurodevelopment is central to ASD pathogenesis, we might argue that a master regulator of gene expression could be as near to the center of the hypothetical ASD gene network as any neurotransmitter receptor or voltage-gated sodium channel.

Defining the Core Neuropathologies of ASD

The diversity of clinical phenotypes across the autism spectrum is in part a reflection of the underlying genetic heterogeneity of ASD (McClellan and King, 2010). To the extent that there exist specific neuropathologies that are common among multiple ASD etiologies, identifying these could help to define the biological processes that influence social behavior. However, in conceptualizing the relationship between gene networks and cognitive dysfunction, we must consider the hypothesis that the autism spectrum consists of multiple clinical subtypes, each having characteristics that are driven by a subset of gene mutations and common risk alleles. To the extent that there exist multiple clinical entities within ASD that have a distinct neurobiological basis, defining the genetic and clinical components of the autism spectrum is necessary to achieving a mechanistic understanding of it.

In principle, multiple core genes and regulators that are tightly connected in a genetic regulatory network might also have correlated phenotype profiles that underlie their predisposition to ASD (Figure 4). Conversely, mutations of distantly related genes or reciprocal deletions or duplications of the same genes could have contrasting molecular and cellular phenotypes. Thus, relationships between genes and biological processes could be defined on the basis of the phenotype profiles of multiple genes across a common battery of assays. Modules of core genes could be defined on the basis of their trait correlations in a manner not unlike the detection of modules based on gene co-expression (Langfelder and Horvath, 2008). Genotype-phenotype correlation has been proposed for clinical subtyping of common disease (Luo et al., 2018), and factor analysis of phenotype data have been proposed as a means to define homogeneous clinical subtypes of ASD (Georgiades et al., 2013). We support the systematic phenotyping of a broad collection of ASD gene models as an experimental approach to defining subsets of interrelated genes and PRSs and to define how biological processes are influenced by these gene sets.

Recent efforts have been initiated to systematize the phenotyping of model systems and to apply a standard battery of assays to gene mutations on a single isogenic background (Deneault et al., 2018 and https://bit.ly/2Tc0lQ4). Such efforts are needed to compare the strength and directionality of effects of genes and CNVs on traits such as cell proliferation, dendritic arborization and synapse numbers, and electrophysiological network signals.

In order to characterize the effects of gene variants on neurodevelopment, model systems are needed that can elucidate the origins, migration, fate, and activity of human cells. iPSCs are attractive human models for understanding complex diseases, and iPSC models have been generated for a variety of ASD-associated variants. A key advantage of human cellular models is the ability to model the effects of CNVs or gene mutations in isogenic cell lines (Deneault et al., 2019b) or to model naturally occurring variation in cell lines derived from patients (Chailangkarn et al., 2016). Furthermore, results from recent studies highlight how different gene mutations can have different (sometimes opposite) effects on cellular phenotypes, which highlight how ASD might result from different types of neuropathology.
Cell Proliferation and Brain Growth. A robust finding from human clinical studies of ASD has been the observation of significantly accelerated brain growth during the first three years of life in a subset of cases (Courchesne et al., 2011). As genetic findings have emerged, it has become evident that macrocephaly is a characteristic of some genetic subtypes of ASD, whereas the opposite phenotype (microcephaly) is associated with other genetic disorders within the autism spectrum. For example, for CNVs at two loci, 1q21.1 and 16p11.2, reciprocal deletions and duplications have opposite effects on brain growth (Malhotra and Sebat, 2012). Among the ASD genes that have effects on head size, mutations in CHD8 (Bernier et al., 2014) and PTEN (Butler et al., 2005) are associated with macrocephaly, and mutations of DYRK1A (Courcet et al., 2012) and CDKL5 (Archer et al., 2006) are associated with microcephaly.

Cellular models have the potential to elucidate the basis for these different effects on cell proliferation. IPSC models of rare microcephaly syndromes have been shown to undergo loss of neural progenitor cells and premature neural differentiation (Lancaster et al., 2013). Human-iPSC-derived NPCs from subjects with ASD and early developmental brain enlargement have been found to display rapid proliferation, consistent with altered cell cycle regulation underlying neuroanatomical traits (Marchetto et al., 2017). Neurons derived from this cohort formed fewer excitatory synapses and matured into defective neuronal networks with less bursting. Cellular models of reciprocal deletions

Figure 4. Defining Core Gene Sets That Regulate Neurodevelopment Based on Trait Correlations in Human Cell Models
We propose the characterization of genotype-trait correlations in cell-based models as means to define sets of genes and CNVs that have common phenotype profiles that might reflect common effects on neuronal function. We illustrate an experimental pipeline where the effects of multiple gene mutations and CNVs are tested in relation to isogenic controls across a series of cellular assays. Likewise, trait correlations with PRS could be tested in patient-derived lines. In this manner, genes, CNVs, and common variants could be clustered into discrete groups. Comparing patterns of trait correlation in cell models and in clinical phenotype data (not shown) could help to identify clinical subtypes of ASD that share common neuropathologies.
and duplications of 16p11.2 recapitulate mirror effects on cell proliferation and reveal mirror effects on synaptic density (Deshpande et al., 2017). These results suggest that standard assays for cell proliferation and differentiation could detect genetic effects on development in specific gene models. **Dendritic Arborization and Synapse Number.** Early studies of human cellular models were performed on neural cell lines derived from Rett syndrome patients with loss-of-function mutations in the gene MECP2. Human cortical neurons displayed reduced arborization and lower glutamatergic synaptic puncta, leading to defects on neural networks (Marchetto et al., 2010). Reductions in neurite outgrowth, dendritic arborization, and excitatory synapses are neuronal phenotypes that have been commonly observed in cellular models of gene mutation, including models of SHANK3 (Scheglovitov et al., 2013; Yi et al., 2016), FMR1 (Doers et al., 2014), and CACNA1C (Krey et al., 2013). Importantly, cellular models of the reciprocal MECP2 gain of function (MECP2 duplication syndrome) (Lugtenberg et al., 2009) exhibit the opposite effects of increased synapses and dendrites (Nageshappa et al., 2016). An increase in dendrite length and synaptogenesis has also been reported in neuron models of Williams syndrome (Chialangkarn et al., 2016) and SHANK2 (Zaslavsky et al., 2019), further highlighting how different mutations of the same gene or different genes can have opposing effects on a cellular phenotype.

**Electrophysiology.** In a systematic screen of ten ASD susceptibility genes (FMR2, ANOS1, ASTN2, ATRX, CACNA1C, CHD8, DLGAP2, KCNQ2, SCN2A, and TENM1) by using CRISPR/Cas9-created iPSC lines and isogenic controls (Deneault et al., 2018), patch-clamp recordings detected reduced excitatory post-synaptic potentials (EPSPs) in a subset of gene models including ATRX, AFF2, KCNQ2, SCN2A, and ASTN2. Other studies have reported that neuronal models of SHANK2 (Zaslavsky et al., 2019), CNTNAP5, and EHMT1 (Deneault et al., 2019a) display hyperconnectivity. Electrophysiological behavior of neurons thus represents a readout that could potentially distinguish genes with different effects on neural activity. **Models of Early Fetal Brain Development.** An attractive approach to systematically record multiple phenotypes within a single system is to use brain organoids, three-dimensional self-assembled multi-cellular structures that mimic the organization, transcriptional, and epigenetic signature of a developing human brain (Lancaster et al., 2013). Brain organoids can be used to dynamically study neural proliferation of progenitor cells, migration, differentiation, and maturation (Trujillo et al., 2018). Brain organoid models have been shown to recapitulate cell proliferation effects in microcephaly (Thomas et al., 2017) and have been used to study neural migration defects in Timothy syndrome, a rare ASD form caused by mutations in the Ca_{1.2} calcium channel gene (Birey et al., 2017). Long-term maturation of the neural networks on brain organoids can lead to electroencephalography (EEG)-like oscillatory waves that could be directly compared with EEG signals from ASD individuals. However, it is currently unknown whether brain organoids can functionally mature to a stage in which network effects could be modeled. As methodologies improve, we might be able to identify additional or more specific in vitro correlates of behaviors and network-level changes that accompany them in human iPSC-derived neurons.

**Mapping Core Genes onto Human Psychiatric Traits**

The last and arguably the most important piece to the precision psychiatry puzzle is to relate sets of core genes to specific dimensions of psychopathology in humans. This is not a task to be taken lightly. During the past decade, significant efforts were undertaken to define the genetic underpinnings of psychiatric “endophenotypes” (Facino et al., 2014). These studies, which were large for studies that apply biometrics like EEG but small by GWAS standards, were largely unsuccessful in achieving the goal of mapping genes to endophenotypes. This is unsurprising in retrospect given that the endophenotypes themselves were complex traits.

With hundreds of ASD susceptibility genes now identified, this problem can now be revisited with a reverse-genetic approach: starting with the genotype and determining how genes influence clinical phenotypes. Large cohort studies of neuroimaging (Franke et al., 2016) and medical health information (Collins and Varmus, 2015) and efforts to define dimensional psychiatric traits in large samples (https://bit.ly/2qTsged) and in rare genetic disorders (https://bit.ly/2VYSk1O) are providing opportunities conduct such studies. For rare mutations, a key challenge is in obtaining a large sample of subjects with a specific genetic disorder. Recruitment of large cohorts is being achieved for the most frequent genetic disorders (D’Angelo et al., 2016; Gur et al., 2017; Simons VIP Consortium, 2012). For instance, the International Brain Behavior Consortium (IBBC) has begun to apply a battery of dimensional measures of psychopathology systematically across multiple genetic disorders (Gur et al., 2017), with a particular focus on CNVs that carry significant risk. Results suggest that the developmental course of clinical phenotypic expression of CNV carriers resembles that of idiopathic populations (Monks et al., 2014). At the same time, there are measurable differences in phenotypic expression between different genetic disorders, which can serve as windows into distinct molecular subtypes of ASD. For example, reciprocal deletions of duplications of multiple loci associate with opposing brain phenotypes in neural circuits involved in social cognition, language, and reward (Deshpande and Weiss, 2018; Lin et al., 2017; Maillard et al., 2015). Cognitive and language studies in carriers of the reciprocal deletions and duplication of 16p11.2 have reported mirror effects on specific linguistic skills, deletion carriers having reduced performance, and duplication carriers having greater performance than IQ-matched controls (Hippolyte et al., 2016). This finding highlights the importance of fine-grained dimensional phenotyping and further suggests that the specificity of genotype-phenotype relationships, which are evident at the level of pathways, can be extended to human psychiatric traits.

**Conclusion**

Within the last decade, ASD has gone from being one of the most mysterious and misunderstood common disorders to one of the success stories of the post-genomics era. We attribute this success to the arrival of game-changing sequencing technologies and the creation of large genomic datasets via National Institutes of Health and privately funded resources (Fischbach and Lord, 2010; C Yuen et al., 2017; Spark Consortium, 2018). Bridging the gap between genes, neurodevelopment, and cognitive
function will likely require adoption of similar big data approaches in the fields of translational neuroscience and clinical psychiatry. In this review, we have not addressed the formidable technical challenges to doing translational science at scale, including the number of cellular assays or clinical measures that could be evaluated across a significant number of genes, and the number of samples or subjects required for adequate statistical power. Making precision psychiatric medicine a reality is a task for the next decade. Gene discoveries have accelerated this process by providing a foundation of knowledge that will form the basis for clinical sequencing and the development of new treatment strategies for ASD.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2019.07.037.

ACKNOWLEDGMENTS

The Beyster Center for Psychiatric Genomics is supported by a gift from the Beyster Family Foundation. J.S. is supported by grants from the NIH (R21MH113179, U01MH109501, and R01MH113715) and the Simons Foundation Autism Research Initiative (SFARI) (606768). L.I. is supported by NIH (R01MH109885, R01MH1108528, and R21MH104766 and SFARI (345469). A.M. is supported by SFARI (345469) and a NARSAD Independent Investigator Grant.

DECLARATION OF INTERESTS

Dr. Muotri is a co-founder and has equity interest in TISMOO, a company dedicated to genetic analysis focusing on therapeutic applications customized for autism spectrum disorder and other neurological disorders with genetic origins. The terms of this arrangement have been reviewed and approved by the University of California San Diego in accordance with its conflict of interest policies.

REFERENCES


Notwell, J.H., Hein, V.E., Darbandi, S., Katzman, S., McKenna, W.L., Ortiz-Londono, C.F., Tastad, D., Eckler, M.J., Rubenstein, J.L., McConnell, S.K.,


