Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

First glimpses of the neurobiology of autism spectrum disorder

Stephan J Sanders

Rapid progress in identifying the genes underlying autism spectrum disorder (ASD) has provided the substrate for a first wave of analyses into the underlying neurobiology. This review describes the consensus across these diverse analyses, highlighting two distinct sets of genes: 1) Genes that regulate chromatin and transcription, especially in cortical projection neurons and striatal medium spiny neurons during mid-fetal development; and 2) Genes involved in synapse development and function, especially during infancy and early childhood, and differentially expressed in the *post mortem* ASD brain. Both gene sets are also regulatory targets of the ASD genes *CHD8* and *FMRP*. It remains to be seen whether these represent two independent paths to the ASD phenotype or two components of a common path.

Address
Department of Psychiatry, University of California, San Francisco, San Francisco, CA 94158, USA

Corresponding author: Sanders, Stephan J (stephansanders@gmail.com)

Current Opinion in Genetics & Development 2015, **33**:80–92

This review comes from a themed issue on **Molecular and genetic bases of disease**

Edited by **Dan E Arking** and **Johanna M Rommens**

http://dx.doi.org/10.1016/j.gde.2015.10.002

Introduction

Autism spectrum disorder (ASD) is characterized by impairment in social communication and restricted or repetitive behaviors or interests [1]. Twin, family, and genotype analyses demonstrate that ASD is highly heritable through a combination of common, rare, and *de novo* variation [2–5].

Initial steps in identifying the specific genetic risk factors underlying ASD were made through the association with syndromes (e.g. Fragile X), large-scale cytogenetic abnormalities (e.g. 15q11.2-13 duplication), and linkage analysis (*NLGN4X*) (described in detail by Willsey and State [6]), however, this progress has been greatly accelerated by recent advances in genomic technology, bioinformatics, and statistical methodology.

Microarray analysis improved the resolution for detecting cytogenetic abnormalities, specifically copy number variants (CNVs) in which several thousand DNA nucleotides are gained or lost. This enabled the systematic detection of *de novo* mutations, i.e. novel variants present in the child but not present in either parent, leading to the key observation that both *de novo* deletion and duplication CNVs are associated with ASD [7,8,9,10,11,12,13••]. Furthermore, specific ASD risk loci were identified through the observation of multiple overlapping *de novo* CNVs in independent ASD cases, including deletions or duplications at 16p11.2 [8,10,14,15,16] and duplications at 7q11.23 [10,17,18,19].

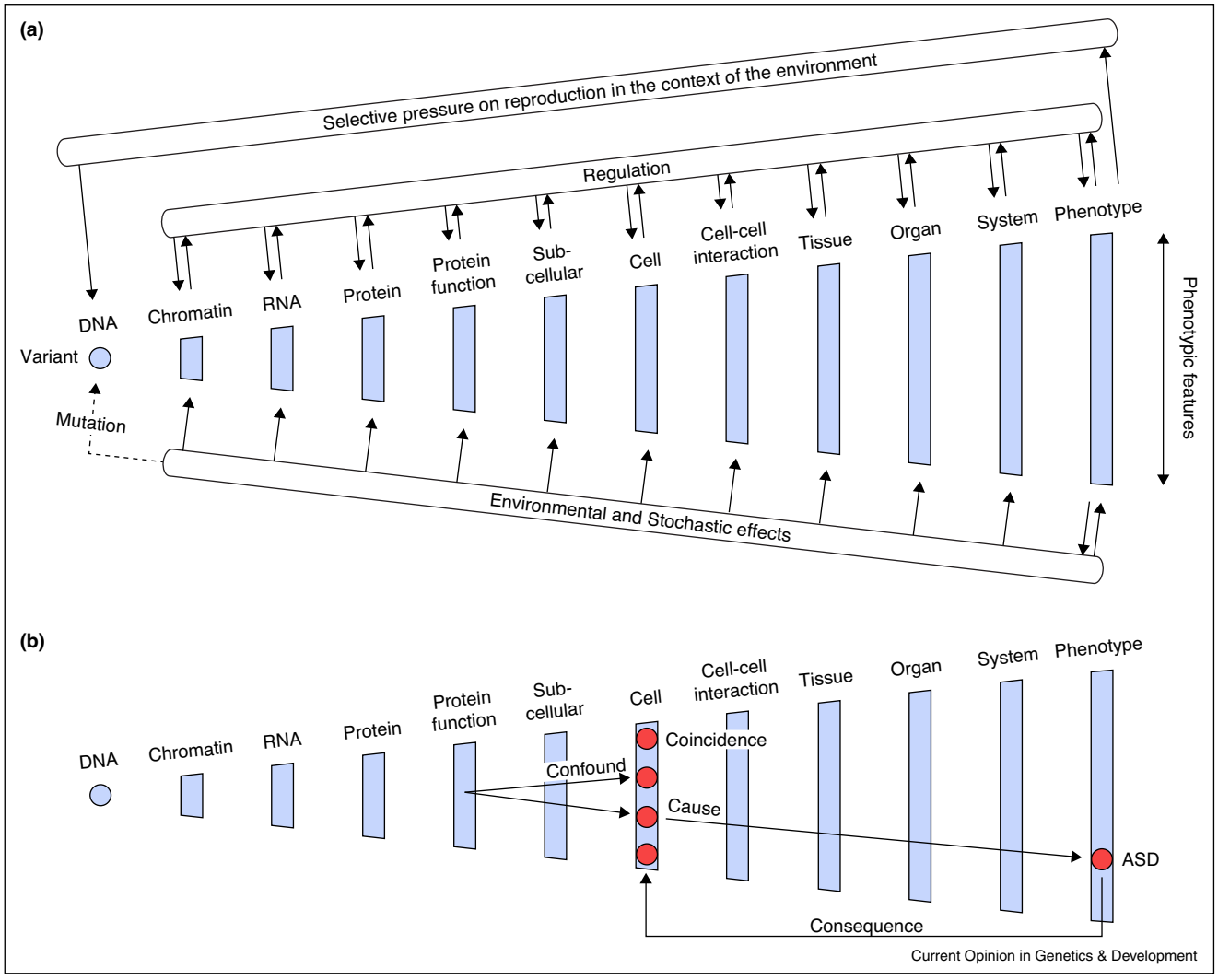
Similarly, high-throughput sequencing of protein-coding exons (whole-exome sequencing or WES) allowed systematic identification of small coding variants. Robust association was observed for *de novo* loss of function (LoF, also called Likely Gene Disrupting/LGD and including nonsense, splice-site, and frameshift variants) mutations [20,21,22••,23,24,25] that are predicted to trigger nonsense-mediated decay (NMD) [26], so that only one copy of the gene is translated into a protein. As with CNVs, the clustering of these mutations enabled the detection of ASD associated loci, specifically genes with multiple *de novo* LoF mutations in independent cases [20]. The Transmission And De novo Association (TADA) model [27] integrates this evidence from LoF mutations with the weaker evidence for ASD association from *de novo* missense mutations and inherited variants, yielding a per gene estimate of ASD association. Applying these methods to large cohorts has identified 65 ASD genes to date [13••,20,21,22••,23,24,25,28••,29••].

With gene discovery proceeding rapidly, a major challenge lies in understanding how these genetic variants lead to the ASD phenotype. Numerous analyses have focused on identifying points of convergence in the targets and functions of the ASD risk genes. In this review, we will consider these findings in the context of the information flow from genetic variants to clinical phenotypes (Figure 1a) and identify points of consensus between these approaches to build a model of ASD neurobiology from the perspective of genomic analyses.

Amplifying genetic information

The information encoded in molecular-scale genetic variants leads to an observable phenotype in an individual. This amplification occurs over multiple steps (Figure 1a), each of which also has the potential to diversify the information by affecting multiple downstream processes.

Figure 1



Information flow and amplification in biological systems. (a) Variation of a single DNA nucleotide can affect the phenotype of an individual. To achieve this feat, the information from the nucleotide is amplified through a series of molecular, cellular, anatomical, and physiological steps. Each step of amplification also has the potential to diversify the information across multiple downstream processes, for example a single protein kinase may influence multiple signaling molecules. The combination of amplification and diversification allows a genetic variant to result in numerous effects, a concept called pleiotropy. In addition, regulatory processes further diversify the information by influencing both upstream and downstream processes. This information flow is further modified by environmental or stochastic factors, some of which can be a direct response to the phenotype (e.g. a treatment). **(b)** The combination of regulatory, environmental, or stochastic factors adds considerable complexity to interpreting whether biological correlations with a phenotype are a cause, confound, consequence, or co-incidence. In contrast, the DNA variant almost always precedes the phenotype and, at an organism-wide level, is constant, making causal relationships easier to define.

Such diversification forms the basis of pleiotropy, the observation of multiple traits (or phenotypic features) as a consequence of a single genetic variant in a gene.

While the DNA nucleotides remain constant throughout an individual's life (with the exception of somatic mutation in individual cells), all the downstream amplification steps are dynamic, changing across development, between cell types, and through interaction with other genetic, environmental, and stochastic factors. Furthermore, through regulation the various steps may influence each other in a non-linear manner, including upstream passage of information (Figure 1a). This complex, interactive, and non-linear flow of information presents a major challenge in establishing how a variant (e.g. a *de novo* LoF mutation in an ASD risk gene) contributes to the risk of a given phenotype (e.g. ASD), since any characteristic observed in the presence of the variant could be a cause, confound, consequence, or coincidence (Figure 1b). For example, if a relative