Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Neurobiology

Autism spectrum disorders: from genes to neurobiology

A Jeremy Willsey[1,2] and Matthew W State[1,2]

Advances in genome-wide technology, coupled with the availability of large cohorts, are finally yielding a steady stream of autism spectrum disorder (ASD) genes carrying mutations of large effect. These findings represent important molecular clues, but at the same time present notable challenges to traditional strategies for moving from genes to neurobiology. A remarkable degree of genetic heterogeneity, the biological pleiotropy of ASD genes, and the tremendous complexity of the human brain are prompting the development of new strategies for translating genetic discoveries into therapeutic targets. Recent developments in systems biology approaches that 'contextualize' these genetic findings along spatial, temporal, and cellular axes of human brain development are beginning to bridge the gap between high-throughput gene discovery and testable pathophysiological hypotheses.

Addresses
[1] Department of Psychiatry, University of California, San Francisco, San Francisco, California 94158, United States
[2] Institute for Human Genetics, University of California, San Francisco, San Francisco, California 94143, United States

Corresponding author: State, Matthew W (matthew.state@ucsf.edu)

Current Opinion in Neurobiology 2015, **30**:92–99

This review comes from a themed issue on **Neuropsychiatry**

Edited by **Steven Hyman** and **Raquel Gur**

For a complete overview see the Issue

Available online 28th November 2014

http://dx.doi.org/10.1016/j.conb.2014.10.015

Introduction

It is a very exciting time for the translational neuroscience of autism spectrum disorders (ASD). Advances in genomic technology, coupled with the availability of large study cohorts, are leading to a rapidly expanding pool of reliable ASD genes and presenting the field with new opportunities and challenges in conceptualizing the path forward from genetic data to an actionable understanding of pathophysiology. Conventional approaches to translating gene discoveries into therapeutic targets have developed largely in the context of Mendelian genetics, typically moving directly from gene identification to single gene manipulations in model systems. However, the extraordinary degree of etiological heterogeneity underlying common forms of ASD, the biological pleiotropy of the implicated genes, the difficulty in distinguishing primary from secondary effects in developmental syndromes, the finding that a myriad of divergent outcomes may be associated with a given ASD mutation, and the likely role that polygenic background plays in this diversity, all suggest that systems biological approaches that can integrate diverse genetic data and capture key dimensions of brain development and function will become an increasingly important complement to the existing armamentarium for moving from genes to neurobiology.

Gene discovery

ASDs are highly heritable, with a substantial contribution from common genetic variation [1•,2•]. However, while inherited polymorphisms are likely to account for the majority of heritability, genome wide association studies (GWAS) have yet to lead to replicable risk loci [3–6]. This is almost surely a consequence of small effects and an attendant lack of statistical power in current study cohorts. Conversely, while estimated to account for a smaller proportion of ASD population risk [1•], the search for large effect, rare mutations in the coding portion of the genome has proven to be a highly productive avenue of investigation.

The first rare variant discoveries in ASD involved genes contributing to monogenic syndromes characterized both by intellectual disability and an increased risk for social impairment. For example, the cloning of the genes *FMR1* [7–9], *PTEN* [10–12], and *TSC1* [13] and *TSC2* [14] are now appreciated to represent the earliest successes in ASD gene discovery. Subsequently, the first replicated findings in 'idiopathic' or 'non-syndromic' ASD emerged in the early part of this century. Rare point mutations in the genes *Neuroligin 3X* and *4X* (*NLGN3X* and *NLGN4X*) [15,16] were identified by targeted sequencing of segments of the X chromosome, based on the observation of rare chromosomal abnormalities in affected females. Importantly, the *NLGN4X* finding was quickly replicated in a linkage study of a large multiply affected pedigree [16]. The observation that neuroligins encode neuronal adhesion molecules at the post-synaptic density (PSD) led to an early interest in synaptic pathology in ASD [17–19].

In fact, the biological relevance of the first genes identified in ASD for the understanding of common, genetically complex forms of the syndrome has become increasingly apparent over time. However, the path forward toward a systematic approach to reliable gene discovery did not emerge until rare *de novo* mutations could be readily detected at high throughput, high resolution, and at a genome-wide scale. Using microarray technology, Sebat et al. [20] in 2007 showed that *de novo* variations in

submicroscopic chromosomal structure (a.k.a copy number variations or CNVs) carried substantial risk for ASD in simplex families (those with only a single affected child). This finding has now been widely replicated [21–26]. Moreover, this discovery was followed quickly by a series of observations that *de novo* CNVs cluster in specific segments of the genome in affected individuals [27,28], and that this could be leveraged to pinpoint risk regions. Dozens of ASD-associated CNVs have now been confirmed, including 16p11.2, 15q11.2-13.1, 15q13.2-q13.3, 7q11.23, NRXN1, 1q21.1, 22q11.2, 16p13.11, Xp22.1 (*PTCHD1-PTCHD1AS*), and *AGBL4* [21–30]. Current estimates suggest that several hundred ASD-related regions will eventually be identified [24,25].

While CNV analyses represented a critical step in clarifying the genomic architecture of ASD, they have proven challenging for translational neuroscience, in large part due to the multigenic nature of many of the relevant risk regions. However, more recent exome sequencing studies have now confirmed that *de novo* coding single nucleotide variants (SNVs) and *de novo* insertion deletions (indels) similarly contribute substantial risk for ASD [31,32•,33,34•,35•,36,38••,39••]. Moreover, a subset of these *de novo* mutations—specifically premature stop codons, canonical splice-site mutations, and frameshift indels (a.k.a likely gene-disrupting or LGD/loss of function or LoF)—has been found to carry the largest effects. Similar to CNVs, the search for multiple *de novo* SNVs and indels at the same locus in unrelated individuals has proven to be a powerful approach to gene discovery [33,37••,38••,39••,40]. Currently, based on de novo LGD/LoF mutations alone, 13 genes have been strongly associated with ASD [31,32•,33,34•,35•,36,37••,38••,39••,41,42] (Table 1; false discovery rate < 0.05), with another 24 genes likely to validate as true risk genes (FDR < 0.2). The yield of well-associated genes can be further increased by incorporating de novo missense mutations, inherited mutations, case-control data, and per-gene estimates of mutability [39••,43], as well as gene expression data from the human brain [39••,44•]. Overall, current estimates suggest that up to, or perhaps more than, 1000 risk genes will carry *de novo* coding mutations [33,35•,38••,39••,43]. Importantly, the path to their identification is now readily apparent [40], dependent only on available cohorts and the resources necessary to sequence and analyze them.

From genes to neurobiology

Single gene syndromes and ASD models

In depth studies of model systems that recapitulate monogenic ASD syndromes have yielded critical insights into neurobiological mechanisms. The details of these efforts have been well reviewed elsewhere (e.g. [17,19,45,46,47,48,49•,50•]) and will not be considered here in depth. It is useful, however, to briefly consider some of the key conceptual advances that have attended studies of, for instance, Fragile X and Tuberous Sclerosis complex, among the earliest single gene syndromes found to be associated with an increased risk for autism.

In 1991 the relationship between the fragile X mental retardation 1 (*FMR1*) gene, and fragile X syndrome (FXS) [7–9] was discovered, followed subsequently by the finding that *FMR1* encodes an RNA-binding protein (Fragile X Mental Retardation Protein; FMRP) [51]. Over the last several decades, a host of studies have confirmed that individuals with FXS demonstrate markedly elevated rates of ASD (e.g. [52–55]). Studies in model systems have shown that a key function of FMRP is regulation of protein translation at the synapse, acting in opposition to metabotropic glutamate receptors (mGluRs). These insights led to the mGluR theory of FXS, postulating that the CNS pathology results from an inability to regulate protein synthesis downstream of mGluR activation (reviewed in Refs. [47,49•]). A series of seminal subsequent findings have revealed that both genetic and pharmacological inhibition of mGluR5 rescues key aspects of the phenotype in FXS models, even in adulthood (reviewed in Refs. [47,49•]). These observations have led to several clinical trials of mGluR5 antagonists and negative allosteric modulators, both with regard to

Table 1

ASD risk genes discovered through recurrent *de novo* loss of function (dnLoF) mutations identified by whole-exome and whole-genome sequencing

Recurrent dnLoF[a]	False discovery rate[b]	Genes
7	<0.0005	*CHD8*
5	<0.005	*ARID1B, DYRK1A, SYNGAP1*
4	<0.01	*ADNP, ANK2, DSCAM, SCN2A*
3	<0.05	*CHD2, GRIN2B, KDM5B, POGZ, SUV420H1*
2	<0.2	*ANKRD11, ASXL3, ASH1L, BCL11A, CACNA2D3, CUL3, DIP2A, FOXP1, GIGYF1, ILF2, KATNAL2, KDM6B, MED13L, NCKAP1, PHF2, RANBP17, RIMS1, SPAST, TBR1, TCF7L2, TNRC6B, WAC, WDFY3, ZC3H4*

[a] De novo mutations were identified in Refs. [31,32•,33,34•,35•,36,37••,38,39,41,42].
[b] False discovery rate estimated as in Ref. [37••].

FXS as well as idiopathic ASD (http://www.clinicaltrials.gov) [47,49•].

Similarly, it has long been observed that individuals with Tuberous Sclerosis complex (TSC) show elevated rates of ASD (e.g. [56–60]). Subsequent to the cloning of causal mutations in the genes *TSC1* [13] and *TSC2* [14], it was determined that the encoded protein products function as negative regulators in the Akt/mTOR pathway, also involved in protein synthesis. Notably, an increased incidence of ASD has also been found in patients carrying rare mutations in the genes *PTEN* [61–64] and *NF-1* [57,65], which are likewise negative regulators of the Akt/mTor signaling pathway. Mouse models of these mutations recapitulate many of the phenotypes observed in affected individuals, some of which are reversed or ameliorated by treatment with inhibitors of the mTor pathway, such as rapamycin or lovastatin (reviewed in Ref. [47]). Accordingly, clinical trials with these drugs are similarly underway (http://www.clinicaltrials.gov).

These discoveries, along with multiple additional examples of phenotypic rescue in adult models of monogenic intellectual disability (ID)/ASD syndromes (reviewed in Refs. [17,19,46,48,50•]), have led to a marked shift in thinking about the opportunities for treating neurodevelopmental disorders. Certainly a decade ago, Fragile X, TSC, and other ID/ASD syndromes were considered largely static entities offering little hope for substantial reversal after birth. These are now widely thought of as reflecting a combination of developmental as well as ongoing functional deficits, opening the door to somatic treatments across the lifespan. Not surprisingly, such discoveries have had a profound influence on thinking about the possibility that common forms of ASD may reflect similar mechanisms and opportunities.

Idiopathic ASD, genetic complexity and neurobiology

The extraordinary degree of locus heterogeneity so far identified in idiopathic ASD presents clear challenges. However, in addition, the combination of biological pleiotropy of risk genes, the (incompletely characterized) cellular diversity of human brain, the difficulty in distinguishing primary from secondary effects in developmental syndromes, and the dynamic nature of brain development, all underscore the challenge of pinpointing key pathological mechanisms in ASD. The issue of characterizing the relevant dysfunction in animal models of ASD mutations has become particularly pressing in light of findings demonstrating that an identical variant carrying risk for ASD may also carry substantial risk for phenotypes with strikingly distinct symptom profiles and natural histories, ranging from schizophrenia, to specific language impairment [66]. This overall complexity is forcing a reevaluation of the strategies that will be necessary to tackle the next phases of translational neuroscience of ASD.

Not surprisingly, in the face of an over-abundance of rare, large effect ASD-related mutations, there have been multiple attempts to identify unifying characteristics among risk genes. Initially this has involved querying sets of genes against pre-existing databases of biological information to determine if particular characteristics, processes, or relationships are overrepresented. For example, in ASD, recent results of gene ontology and protein-protein interaction analyses have focused attention on chromatin biology, neuronal development, Wnt/beta-catenin signaling, FMRP binding, and synaptic functioning (e.g. [26,32•,35•,67,68••]).

These findings, while clearly important, provide limited information with regard to when and where to look in model systems for specific ASD related mechanisms. Moreover, based on the hypotheses that developmental processes are important in ASD pathophysiology, it would follow that being able to specify spatial and temporal variables while modeling specific mutations could help constrain the search for relevant molecular, cellular and/or circuit level phenotypes.

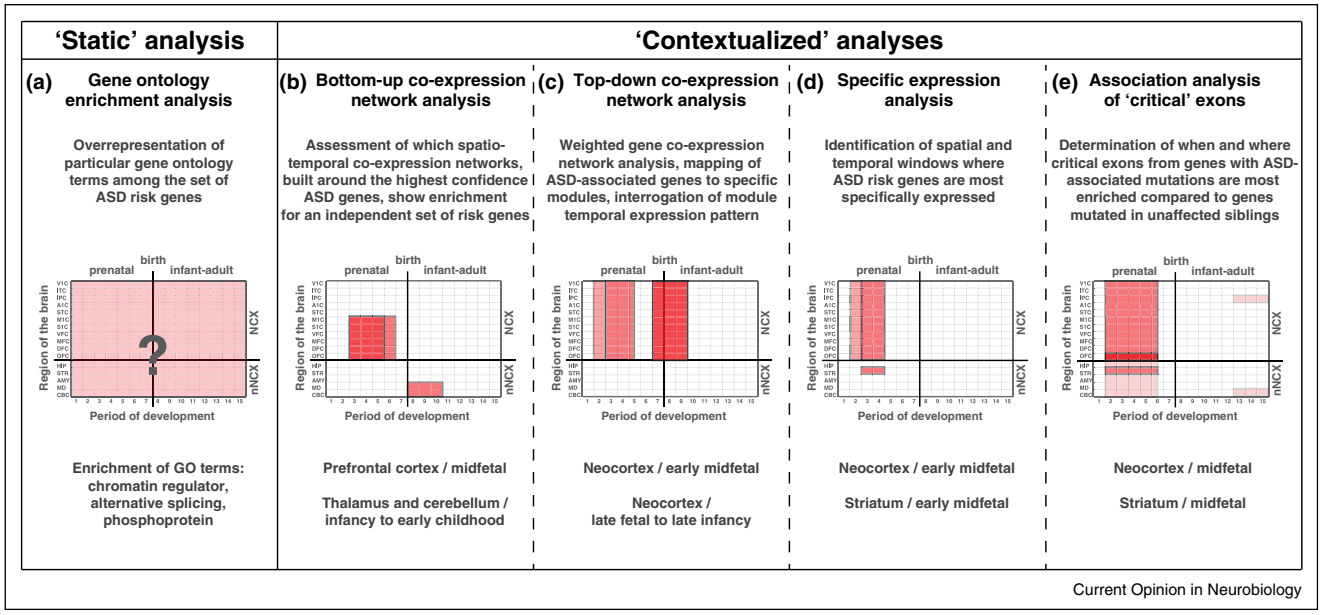
Until recently, such 'contextualized' analyses have not been possible, due to the fact that most available biological databases draw information from a wide variety of sources, but without allowing for the parsing of relevant variables such as cell or tissue type, brain region or developmental stage. Consequently, the results of many current systems biological analyses have tended to define a general and static topology of ASD pathology (Figure 1a). However, recent foundational efforts are rapidly leading to the creation of multidimensional 'omics' datasets, reflecting the entire course of brain development. For example, the BrainSpan developmental transcriptome project summarizes gene expression from early fetal to late adult stages across multiple distinct anatomical regions in 57 typically developing human brains [69••,70••]. Similar datasets are now becoming available in primate[1] and mouse.[2] These provide the substrate to begin to capture developmental dimensionality in systems biology analyses of ASD-related genes.

Indeed, several groups have utilized the BrainSpan data set to evaluate genetic findings and these studies have already begun to point to convergent neurobiology. For example, Ben-David et al. [68••] found that a large proportion of the set of genes identified by whole-exome sequencing [32•,33,34•,35•] are present in a single co-expression module that is most highly expressed before birth. Several analyses have now extended and refined

[1] Website: ©2014 Allen Institute for Brain Science. NIH Blueprint Non-Human Primate (NHP) Atlas [Internet]. Available from: http://www.blueprintnhpatlas.org/.

[2] Website: ©2013 Allen Institute for Brain Science. Allen Developing Mouse Brain Atlas [Internet]. Available from: http://developingmouse.brain-map.org.

Figure 1



Systems biology approaches in ASD. Two main types of systems biology analyses, which we refer to as 'static' versus 'contextualized', have been recently applied to ASD genes. **(a)** A set of 131 ASD risk genes [37••] evaluated for enrichment of gene ontology terms using DAVID [74,75] identifies 'chromatin regulator', 'alternative splicing', and 'phosphoprotein' as enriched terms. The grids illustrate temporal (*X*-axis; developmental periods as defined in Ref. [69••]) and spatial (*Y*-axis; anatomical brain regions) variables, and the intensity of red indicates the strength of evidence in each panel. The '?' in (a) reflects the absence of data on these dimensions from static enrichment analyses. **(b–e)** Illustrate several recent approaches to contextualized analyses: (b) BrainSpan exon array expression dataset was used to create co-expression networks based on a small number of high confidence ASD genes [37••]. Each of the spatially and temporally defined networks was then evaluated for the presence of additional ASD associated genes. Two spatiotemporal windows were found to be significantly enriched for ASD genes: midfetal prefrontal cortex and thalamus/cerebellar cortex during neonatal to early childhood. The enriched networks were then assessed for additional properties with regard to cortical layer and cell type [69••]. (c) Weighted gene co-expression network analysis was conducted with BrainSpan RNA-seq expression data from the neocortex, and resultant modules were assessed for enrichment of ASD-associated genes [71••]. Based on the temporal expression pattern and biological properties of the genes within enriched modules, biological processes likely key during early fetal development and late fetal to early infancy were identified. Laminar and cellular specificity were also assessed. (d) A specificity index was calculated by brain region and developmental epoch for expression of ASD risk genes identified from whole-exome sequencing [31,32•,33,34•,35•]. The authors determined that expression of the set of ASD genes is particularly specific to early midfetal neocortex as well as the early midfetal striatum [72••]. (e) A fourth recent analysis utilized exon transcriptome-mutation contingency indices [73••]. After summarizing exon level mutation rates based on population controls, the authors identified critical exons (defined as having high expression but low mutational burden) and observed that enrichment of critical exons in genes mutated in ASD cases versus unaffected sibling controls [33,35•] is strongest in prenatal neocortex and striatum.

these findings, implicating specific brain regions, developmental epochs, and cell types in pathology (Figure 1b–e). For instance, by sub-setting the BrainSpan exon array expression data into smaller spatiotemporal windows, constructing gene co-expression networks around only the strongest set of ASD genes, and then assessing each for enrichment of additional ASD risk genes, our group found that deep layer (layers V–VI) cortical projection neurons in midfetal prefrontal cortex are one important nexus for a subset of ASD risk genes [37••]. In a similar vein, Parikshak et al. [71••] used BrainSpan RNA-seq data, a broad range of input genes, and weighted gene co-expression network analysis, and found evidence for the importance of developmental epochs encompassing both early fetal and late fetal to early infancy, as well as projection neurons. In this case, their findings pointed to both deep and superficial (layers II–IV) cortical layers, however the latter were found to be most relevant for ASD pathology.

Xu et al. subsequently developed an approach, known as specific expression analysis (SEA) [72••], for associating candidate gene lists with a particular context based on expression profiles across multiple contexts (e.g. cell type, brain region, time period). They applied SEA to interrogate the genes identified in exome studies [32•,33,34•,35•] and, similar to the findings noted above, found enrichment in the developing midfetal cortex, as well as midfetal striatum. Finally, Uddin and colleagues [73••] utilized exon level expression data from BrainSpan set to assess the relationship between exon expression and relevance to ASD. The authors observed a strong relationship between brain-expressed exons under purifying selection and ASD risk—an effect strongest in

prenatal development, and in particular the orbital frontal cortex.

These studies collectively begin to give a sense of new opportunities in addressing the path from rare variant discovery in idiopathic ASD. The findings with regard to cell-type, regional, and developmental variables should begin to narrow the search for mechanisms and treatment targets in ASD. One can envision either focused investigations of single genes in particular cell types, developmental stages and brain regions, or, alternatively, using this spatiotemporal information as the basis for the simultaneous examination of multiple mutations in (apparently) functionally distinct genes.

Finally, while the genes so far discovered have all pointed to some degree to fetal cortical development, they also have shown some divergence (Figure 1b–e). Given the heterogeneity of ASD, this is not surprising as there will likely be multiple regions and time periods involved, some of which have yet to be identified. Finally, while current studies have found that risk mutations appear to converge in terms of their function in early brain development, this does not necessarily argue against the notion, supported by studies of monogenic syndromes, that the phenotypic manifestations may be a consequence both of developmental as well as ongoing functional deficits.

Future directions

There are, of course, multiple areas of investigation that will be essential to support continued progress. First, ongoing gene discovery is essential. As noted above, many of the promising recent findings in translational neuroscience have relied on identifying either functional or spatiotemporal convergence of multiple risk loci. It follows that the more rare, large effect ASD mutations that are confirmed, the more inputs there will be for future studies. Similarly, it is clear that a substantial portion of ASD risk is carried in common polymorphisms. While these variants will undoubtedly confer smaller biological effects, clarifying this source of risk will be critical, including for determining the role of polygenic factors in dictating widely divergent outcomes from identical high effect mutations. Finally, as gene discovery in ASD continues to advance, the likelihood that population level sequencing will lead to the discovery of protective alleles increases, as does the potential that these will provide another path to therapeutics development.

An additional area of tremendous need is the development of additional 'omics' datasets that capture key aspects of brain development across species, including, but not limited to, isoform-level gene expression, ChIP-seq, methyl-seq, and proteomics data. These resources must ultimately include cell-type specificity to allow for the type of high resolution analyses that are likely to be essential for the next steps in clarifying when and where ASD pathology may be most productively studied.

Finally models systems will clearly be a critical resource. The addition to this armamentarium of induced pluripotent stem cells, brain organoids, genome editing tools, and methods to capture data from increasingly large groups of neurons will play a key role in allowing the field to move forward. While there are currently well-known limitations to any given system, the increasing ability to model multiple mutations simultaneously, across the range of systems, from fly to human cells holds real promise for the successful navigation of the difficult path from genes to neurobiology.

Conflict of interest

Nothing to declare.

Acknowledgements

The authors acknowledge financial support from the National Institute of Mental Health (RC2 MH089956 to M.W.S.), the Simons Foundation (to M.W.S), the Overlook International Foundation (to M.W.S.), and the Canadian Institutes of Health Research (Doctoral Foreign Study Award to A.J.W.). We would also like to thank Stephan J Sanders, for estimating the false discovery rates summarized in Table 1.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. • Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, Mahajan M, Manaa D, Pawitan Y, Reichert J *et al.*: **Most genetic risk for autism resides with common variation**. *Nat Genet* 2014, **46**:881-885.

Assessed data from Sweden's universal health registry to determine the relative contributions of rare and common variation to narrow-sense heritability. Estimated that rare inherited variants and de novo mutations account for a small proportion of population-level ASD risk.

2. • Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW, Fombonne E, Geschwind D *et al.*: **Common genetic variants, acting additively, are a major source of risk for autism**. *Mol Autism* 2012, **3**:9.

Established that a large number of common variant loci additively contribute a substantial proportion of ASD risk. Clarified the contribution of common variation to ASD liability in both simplex and multiplex families.

3. Weiss LA, Arking DE, Daly MJ, Chakravarti A: **A genome-wide linkage and association scan reveals novel loci for autism**. *Nature* 2009, **461**:802-808.

4. Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, Salyakina D, Imielinski M, Bradfield JP, Sleiman PM *et al.*: **Common genetic variants on 5p14.1 associate with autism spectrum disorders**. *Nature* 2009, **459**:528-533.

5. Anney R, Klei L, Pinto D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Sykes N, Pagnamenta AT *et al.*: **A genome-wide scan for common alleles affecting risk for autism**. *Hum Mol Genet* 2010, **19**:4072-4082.

6. Devlin B, Melhem N, Roeder K: **Do common variants play a role in risk for autism? Evidence and theoretical musings**. *Brain Res* 2011, **1380**:78-84.

7. Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI: **Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n**. *Science* 1991, **252**:1711-1714.

8. Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP *et al.*: **Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome**. *Cell* 1991, **65**:905-914.

9. Pieretti M, Zhang FP, Fu YH, Warren ST, Oostra BA, Caskey CT, Nelson DL: **Absence of expression of the FMR-1 gene in fragile X syndrome**. *Cell* 1991, **66**:817-822.

10. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, Puc J, Miliaresis C, Rodgers L, McCombie R *et al.*: **PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer**. *Science* 1997, **275**:1943-1947.

11. Li DM, Sun H: **TEP1, encoded by a candidate tumor suppressor locus, is a novel protein tyrosine phosphatase regulated by transforming growth factor beta**. *Cancer Res* 1997, **57**:2124-2129.

12. Steck PA, Pershouse MA, Jasser SA, Yung WK, Lin H, Ligon AH, Langford LA, Baumgard ML, Hattier T, Davis T *et al.*: **Identification of a candidate tumour suppressor gene MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers**. *Nat Genet* 1997, **15**:356-362.

13. van Slegtenhorst M, de Hoogt R, Hermans C, Nellist M, Janssen B, Verhoef S, Lindhout D, van den Ouweland A, Halley D, Young J *et al.*: **Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34**. *Science* 1997, **277**:805-808.

14. Consortium ECTS: **Identification and characterization of the tuberous sclerosis gene on chromosome 16**. *Cell* 1993, **75**:1305-1315.

15. Jamain S, Quach H, Betancur C, Råstam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C *et al.*: **Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism**. *Nat Genet* 2003, **34**:27-29.

16. Laumonnier F, Bonnet-Brilhault F, Gomot M, Blanc R, David A, Moizard MP, Raynaud M, Ronce N, Lemonnier E, Calvas P *et al.*: **X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family**. *Am J Hum Genet* 2004, **74**:552-557.

17. Bourgeron T: **A synaptic trek to autism**. *Curr Opin Neurobiol* 2009, **19**:231-234.

18. Zoghbi HY: **Postnatal neurodevelopmental disorders: meeting at the synapse?** *Science* 2003, **302**:826-830.

19. Walsh CA, Morrow EM, Rubenstein JLR: **Autism and brain development**. *Cell* 2008, **135**:396-400.

20. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J *et al.*: **Strong association of de novo copy number mutations with autism**. *Science* 2007, **316**:445-449.

21. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y *et al.*: **Structural variation of chromosomes in autism spectrum disorder**. *Am J Hum Genet* 2008, **82**:477-488.

22. Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, Eichler EE: **De novo rates and selection of large copy number variation**. *Genome Res* 2010, **20**:1469-1481.

23. Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS *et al.*: **Functional impact of global rare copy number variation in autism spectrum disorders**. *Nature* 2010, **466**:368-372.

24. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA *et al.*: **Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism**. *Neuron* 2011, **70**:863-885.

25. Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, Marks S, Lakshmi B, Pai D, Ye K *et al.*: **Rare de novo and transmitted copy-number variation in autistic spectrum disorders**. *Neuron* 2011, **70**:886-897.

26. Pinto D, Delaby E, Merico D, Barbosa M, Merikangas A, Klei L, Thiruvahindrapuram B, Xu X, Ziman R, Wang Z *et al.*: **Convergence of genes and cellular pathways dysregulated in autism spectrum disorders**. *Am J Hum Genet* 2014, **94**:677-694.

27. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T *et al.*: **Association between microdeletion and microduplication at 16p11.2 and autism**. *N Engl J Med* 2008, **358**:667-675.

28. Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH, Dobyns WB *et al.*: **Recurrent 16p11.2 microdeletions in autism**. *Hum Mol Genet* 2008, **17**:628-638.

29. Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, Senman L *et al.*: **Mapping autism risk loci using genetic linkage and chromosomal rearrangements**. *Nat Genet* 2007, **39**:319-328.

30. Moreno-De-Luca D, Sanders SJ, Willsey AJ, Mulle JG, Lowe JK, Geschwind DH, State MW, Martin CL, Ledbetter DH: **Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts**. *Mol Psychiatry* 2013, **18**:1090-1095.

31. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C *et al.*: **Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations**. *Nat Genet* 2011, **43**:585-589.

32. • O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD *et al.*: **Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations**. *Nature* 2012, **485**:246-250.

One of the initial four whole-exome sequencing studies in simplex ASD families. Together with the other three, established the contribution of de novo point mutations and insertion/deletions to ASD risk. The combined set of mutations from these four papers underlies most of the recent systems biology approaches outlined in this review.

33. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL *et al.*: **De novo mutations revealed by whole-exome sequencing are strongly associated with autism**. *Nature* 2012, **485**:237-241.

One of the initial four whole-exome sequencing studies in simplex ASD families. Together with the other three, established the contribution of de novo point mutations and insertion/deletions to ASD risk. Established rigorous statistical framework for associating genes with ASD based on recurrent de novo loss of function mutations. The combined set of mutations from these four papers underlies most of the recent systems biology approaches outlined in this review.

34. • Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V *et al.*: **Patterns and rates of exonic de novo mutations in autism spectrum disorders**. *Nature* 2012, **485**:242-245.

One of the initial four whole-exome sequencing studies in simplex ASD families. Together with the other three, established the contribution of de novo point mutations and insertion/deletions to ASD risk. The combined set of mutations from these four papers underlies most of the recent systems biology approaches described in this review.

35. • Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee Y-h, Narzisi G, Leotta A *et al.*: **De novo gene disruptions in children on the autistic spectrum**. *Neuron* 2012, **74**:285-299.

One of the initial four whole-exome sequencing studies in simplex ASD families. Together with the other three, established the contribution of de novo point mutations and insertion/deletions to ASD risk. The combined set of mutations from these four papers underlies most of the recent systems biology approaches outlined in this review.

36. Dong S, Walker M, Carrerio N, DiCola M, Willsey A, Ye A, Waqar Z, Gonzalez L, Overton J, Frahm S *et al.*: **De novo insertions and deletions of predominantly paternal origin are associated with autism spectrum disorder**. *Cell Rep* 2014, **9**:16-23 http://dx.doi.org/10.1016/j.celrep.2014.08.068.

37. •• Willsey AJ, Sanders Stephan J, Li M, Dong S, Tebbenkamp Andrew T, Muhle Rebecca A, Reilly Steven K, Lin L, Fertuzinhos S, Miller Jeremy A *et al.*: **Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism**. *Cell* 2013, **155**:997-1007.

This study presents a novel, bottom-up co-expression network analysis leveraging the BrainSpan exon array developmental transcriptome

dataset and the combined set of ASD genes identified via accumulated exome-wide and genome-wide sequencing data from simplex ASD families. Identifies deep layer cortical projection neurons in prefrontal and primary motor-somatosensory cortical regions during midfetal development as a critical nexus of ASD risk.

38. •• Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE *et al.*: **The contribution of de novo coding mutations to autismspectrum disorder**. *Nature* 2014, **515**:216-221 http://dx.doi.org/10.1038/nature13908 PubMed PMID: 25363768.
One of the two most recent and largest whole exome sequencing studies in simplex ASD. The entire Simons Simplex Collection is sequenced (>2,500 families), resulting in the identification of 27 genes with recurrent de novo LGF/LoF mutations.

39. •• De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Ercument Cicek A, Kou Y, Liu L, Fromer M, Walker S *et al.*: **Synaptic, transcriptional and chromatin genes disrupted inautism**. *Nature* 2014, **515**:209-215 http://dx.doi.org/10.1038/nature13772 PubMed PMID: 25363760.
One of the two most recent and largest whole exome sequencing studies in simplex ASD. 3,871 ASD cases and 9,937 ancestry-matched or parental controls are sequenced, resulting in the identification of 22 genes with a false-discovery rate of <0.05

40. Buxbaum JD, Daly MJ, Devlin B, Lehner T, Roeder K, State MW, Consortium TAS: **The autism sequencing consortium: large-scale high-throughput sequencing in autism spectrum disorders**. *Neuron* 2012, **76**:1052-1056.

41. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Wong WS *et al.*: **Rate of de novo mutations and the importance of father's age to disease risk**. *Nature* 2012, **488**:471-475.

42. O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K *et al.*: **Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders**. *Science* 2012, **338**:1619-1622.

43. He X, Sanders SJ, Liu L, De Rubeis S, Lim ET, Sutcliffe JS, Schellenberg GD, Gibbs RA, Daly MJ, Buxbaum JD *et al.*: **Integrated model of de novo and inherited genetic variants yields greater power to identify risk genes**. *PLOS Genet* 2013, **9**:e1003671.

44. • Liu L, Lei J, Sanders SJ, Willsey AJ, Kou Y, Cicek AE, Klei L, Lu C, He X, Li M *et al.*: **DAWN: a framework to identify autism genes and subnetworksusing gene expression and genetics**. *Mol Autism* 2014, **5**:22 http://dx.doi.org/10.1186/2040-2392-5-22 PubMed PMID: 24602502; PubMed Central PMCID: PMC4016412.
Integrated genetic data with gene expression data from the developing human brain to predict additional ASD risk genes, as well as putative ASD subnetworks

45. Moy SS, Nadler JJ: **Advances in behavioral genetics: mouse models of autism**. *Mol Psychiatry* 2008, **13**:4-26.

46. Ehninger D, Li W, Fox K, Stryker MP, Silva AJ: **Reversing neurodevelopmental disorders in adults**. *Neuron* 2008, **60**:950-960.

47. Krueger DD, Bear MF: **Toward fulfilling the promise of molecular medicine in fragile X syndrome**. *Annu Rev Med* 2011, **62**:411-429.

48. Castren E, Elgersma Y, Maffei L, Hagerman R: **Treatment of neurodevelopmental disorders in adulthood**. *J Neurosci* 2012, **32**:14074-14079.

49. • Bhakar AL, Dolen G, Bear MF: **The pathophysiology of fragile X (and what it teaches us about synapses)**. *Annu Rev Neurosci* 2012, **35**:417-443.
A comprehensive review summarizing the molecular pathology underlying fragile X syndrome (FXS), the large body of work in various FXS model systems including phenotypic rescue, and how these findings are being translated to therapeutic development.

50. • van der Voet M, Nijhof B, Oortveld MA, Schenck A: **Drosophila models of early onset cognitive disorders and their clinical applications**. *Neurosci Biobehav Rev* 2014 http://dx.doi.org/10.1016/j.neubiorev.2014.01.013. pii: S0149-7634(14)00057-8; [Epub ahead of print].
An excellent review summarizing the history and utility of Drosophila as a model system for understanding neurodevelopmental disorders.

51. Ashley CT Jr, Wilkinson KD, Reines D, Warren ST: **FMR1 protein: conserved RNP family domains and selective RNA binding**. *Science* 1993, **262**:563-566.

52. Kielinen M, Rantala H, Timonen E, Linna SL, Moilanen I: **Associated medical disorders and disabilities in children with autistic disorder: a population-based study**. *Autism* 2004, **8**:49-60.

53. Reddy KS: **Cytogenetic abnormalities and fragile-X syndrome in Autism Spectrum Disorder**. *BMC Med Genet* 2005, **6**:3.

54. Hatton DD, Sideris J, Skinner M, Mankowski J, Bailey DB Jr, Roberts J, Mirrett P: **Autistic behavior in children with fragile X syndrome: prevalence, stability, and the impact of FMRP**. *Am J Med Genet A* 2006, **140a**:1804-1813.

55. Clifford S, Dissanayake C, Bui QM, Huggins R, Taylor AK, Loesch DZ: **Autism spectrum phenotype in males and females with fragile X full mutation and premutation**. *J Autism Dev Disord* 2007, **37**:738-747.

56. Gillberg IC, Gillberg C, Ahlsen G: **Autistic behaviour and attention deficits in tuberous sclerosis: a population-based study**. *Dev Med Child Neurol* 1994, **36**:50-56.

57. Fombonne E, Du Mazaubrun C, Cans C, Grandjean H: **Autism and associated medical disorders in a French epidemiological survey**. *J Am Acad Child Adolesc Psychiatry* 1997, **36**:1561-1569.

58. Lewis JC, Thomas HV, Murphy KC, Sampson JR: **Genotype and psychological phenotype in tuberous sclerosis**. *J Med Genet* 2004, **41**:203-207.

59. Wong V: **Study of the relationship between tuberous sclerosis complex and autistic disorder**. *J Child Neurol* 2006, **21**:199-204.

60. Muzykewicz DA, Newberry P, Danforth N, Halpern EF, Thiele EA: **Psychiatric comorbid conditions in a clinic population of 241 patients with tuberous sclerosis complex**. *Epilepsy Behav* 2007, **11**:506-513.

61. Goffin A, Hoefsloot LH, Bosgoed E, Swillen A, Fryns JP: **PTEN mutation in a family with Cowden syndrome and autism**. *Am J Med Genet* 2001, **105**:521-524.

62. Butler MG, Dasouki MJ, Zhou XP, Talebizadeh Z, Brown M, Takahashi TN, Miles JH, Wang CH, Stratton R, Pilarski R *et al.*: **Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations**. *J Med Genet* 2005, **42**:318-321.

63. Varga EA, Pastore M, Prior T, Herman GE, McBride KL: **The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly**. *Genet Med* 2009, **11**:111-117.

64. McBride KL, Varga EA, Pastore MT, Prior TW, Manickam K, Atkin JF, Herman GE: **Confirmation study of PTEN mutations among individuals with autism or developmental delays/mental retardation and macrocephaly**. *Autism Res* 2010, **3**:137-141.

65. Garg S, Lehtonen A, Huson SM, Emsley R, Trump D, Evans DG, Green J: **Autism and other psychiatric comorbidity in neurofibromatosis type 1: evidence from a population-based study**. *Dev Med Child Neurol* 2013, **55**:139-145.

66. State MW, Šestan N: **The emerging biology of autism spectrum disorders**. *Science* 2012, **337**:1301-1303.

67. Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D: **Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses**. *Neuron* 2011, **70**:898-907.

68. •• Ben-David E, Shifman S: **Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism**. *Mol Psychiatry* 2013, **18**:1054-1056.
One of the first studies to leverage whole-exome sequencing data from simplex ASD families in a spatially and temporally aware analysis. Identifies chromatin regulation during prenatal development as an important component of ASD pathogenesis.

69. •• Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AMM, Pletikos M, Meyer KA, Sedmak G *et al.*: **Spatio-temporal transcriptome of the human brain**. *Nature* 2011, **478**:483-489.
This is the initial paper characterizing the BrainSpan developmental transcriptome that underlies most of the systems biology approaches described in this review.

70. •• Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, Riley ZL, Royall JJ, Aiona K *et al.*: **Transcriptional landscape of the prenatal human brain**. *Nature* 2014, **508**:199-206.
This is the newest publication from the BrainSpan developmental transcriptome. Updates previous results and adds laminar specific in situ hybridization, MRI, and gene expression data from laser micro-dissected human prenatal cortex. This dataset also underlies some of the systems biology approaches reviewed here.

71. •• Parikshak Neelroop N, Luo R, Zhang A, Won H, Lowe Jennifer K, Chandran V, Horvath S, Geschwind Daniel H: **Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism**. *Cell* 2013, **155**:1008-1021.
This study presents a novel, top-down weighted gene co-expression network analysis utilizing the BrainSpan RNA-seq developmental transcriptome dataset and a broad range of ASD-associated genes and intellectual disability-associated genes. Early fetal as well as late fetal to early infancy were identified as key developmental periods. Both deep and superficial cortical layers (III–IV) were also implicated, with the strongest signal observed for the latter.

72. •• Xu X, Wells AB, O'Brien DR, Nehorai A, Dougherty JD: **Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders**. *J Neurosci* 2014, **34**:1420-1431.
Utilized specific expression analysis (SEA) to demonstrate that, as a group, the set of ASD genes identified in the initial four whole-exome sequencing studies have expression patterns highly specific to developing midfetal cortex and striatum.

73. •• Uddin M, Tammimies K, Pellecchia G, Alipanahi B, Hu P, Wang Z, Pinto D, Lau L, Nalpathamkalam T, Marshall CR *et al.*: **Brain-expressed exons under purifying selection are enriched for de novo mutations in autism spectrum disorder**. *Nat Genet* 2014, **46**:742-747.
Identified that exons with low mutational burden in population controls but high brain brain expression are associated with neuropsychiatric disorders. This association was strongest in prenatal neocortex.

74. Huang da W, Sherman BT, Lempicki RA: **Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists**. *Nucleic Acids Res* 2009, **37**:1-13.

75. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources**. *Nat Protoc* 2009, **4**:44-57.