DE NOVO MUTATIONS IN AUTISM IMPLICATE THE SYNAPTIC ELIMINATION NETWORK*

GUHAN RAM VENKATARAMAN

Department of Bioengineering, Stanford University, 318 Campus Drive
Stanford, CA 94305, USA
Email: guhan@stanford.edu

CHLOE O'CONNELL, FUMIKO EGAWA, DORNA KASHEF-HAGHIGHI, DENNIS P. WALL

Department of Pediatrics and Biomedical Data Science, 1265 Welch Road
Stanford, CA 94305, USA
Email: dpwall@stanford.edu

Autism has been shown to have a major genetic risk component; the architecture of documented autism in families has been over and again shown to be passed down for generations. While inherited risk plays an important role in the autistic nature of children, *de novo* (germline) mutations have also been implicated in autism risk. Here we find that autism *de novo* variants verified and published in the literature are Bonferroni-significantly enriched in a gene set implicated in synaptic elimination. Additionally, several of the genes in this synaptic elimination set that were enriched in protein-protein interactions (CACNA1C, SHANK2, SYNGAP1, NLGN3, NRXN1, and PTEN) have been previously confirmed as genes that confer risk for the disorder. The results demonstrate that autism-associated *de novos* are linked to proper synaptic pruning and density, hinting at the etiology of autism and suggesting pathophysiology for downstream correction and treatment.

* This work was supported in part by the Hartwell Foundation's Autism Research and Technology Initiative.

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that impairs social skills, communication, and normal behavior. About 1% of the world's population has ASD, and this number is rapidly rising: the prevalence of autism more than doubled between 2002 and 2012[1]. ASD-linked impairment leads to higher lifespan costs[2] and a significant reduction in ability to procure both postgraduate education and jobs[3].

Both inherited (present in mother or father) and *de novo* (germline) mutations have been shown to contribute to the disease[4]. Although several of each type appear to contribute to ASD risk, there is still not a clear picture or full map of what leads to ASD[5]. Several hypotheses exist for the genetic etiology of autism; one of note is referred to as the "synaptic elimination hypothesis," the exploration of which is the focus of this paper.

Synaptic elimination is a normal neurodevelopmental process, starting in the fifth week of development and continuing throughout life. The process occurs in parallel with synaptic formation, which relies on input from both presynaptic and postsynaptic neurons. Elimination eventually outpaces formation in adolescence and adulthood[6,7]. During the development of the central nervous system, neurons form multiple synapses in excess of functional need. These redundant synapses are later eliminated through various means: 1) loss of signals necessary from either presynaptic or postsynaptic neurons to maintain synaptic stability[8]; 2) apoptosis of synapses; 3) ubiquitination of synaptic proteins for proteosomal degradation[9]; 4) macro-autophagy[7]; or 5) phagocytosis of synapses as a result of opsonization by synaptic elimination-mediating complement factors, microglia, and astrocytes[10-13].

Previous work has shown that faulty synaptic formation and maturation contribute to ASD[6]. However, given that increases in both dendritic spine density and brain weight (both of which are characteristic of autism) can be caused by mutations in genes regulating synaptic elimination, the hypothesis developed that autism could *also* be a disease of abnormal synaptic elimination[8,14-16].

Currently, the major pathway and ontology databases (KEGG, GO, Panther, and Reactome) do not contain any gene sets that pertain to synaptic elimination or synaptic pruning. As part of this study, we endeavored to create a robust and manually curated list of genes contributing to synaptic elimination; our goal was to test the hypothesis that the curated gene set would be enriched for *de novo* mutations (see Supplemental Materials 2 and 3 for list of genes and references used to generate this list, respectively). We hypothesized that increased burden of mutations in synaptic elimination genes would lead to the synaptic pruning abnormalities observed in autism, such as increased dendritic spine density and increased brain weight.

We used the *dnenrich* package[17], a network burden analysis tool, to test for enrichment in the synaptic elimination gene set on a comprehensive set of exomes from family-based trios having one child with autism. The package has been shown to be particularly powerful for identifying *de novo* mutations with small individual association to phenotype, but large effect in combination. We used dnenrich on previously documented autism-associated gene sets and autism-associated *de novos* as a pilot. This was done to verify that the program was suitable for use with autism *de novos* and that our list of *de novos* was large enough to provide sufficient power to detect enrichment of certain gene sets. After doing so, we tested the hypothesis that our list of genes

involved in synaptic elimination would have a higher burden of autism *de novos* than would be expected by chance.

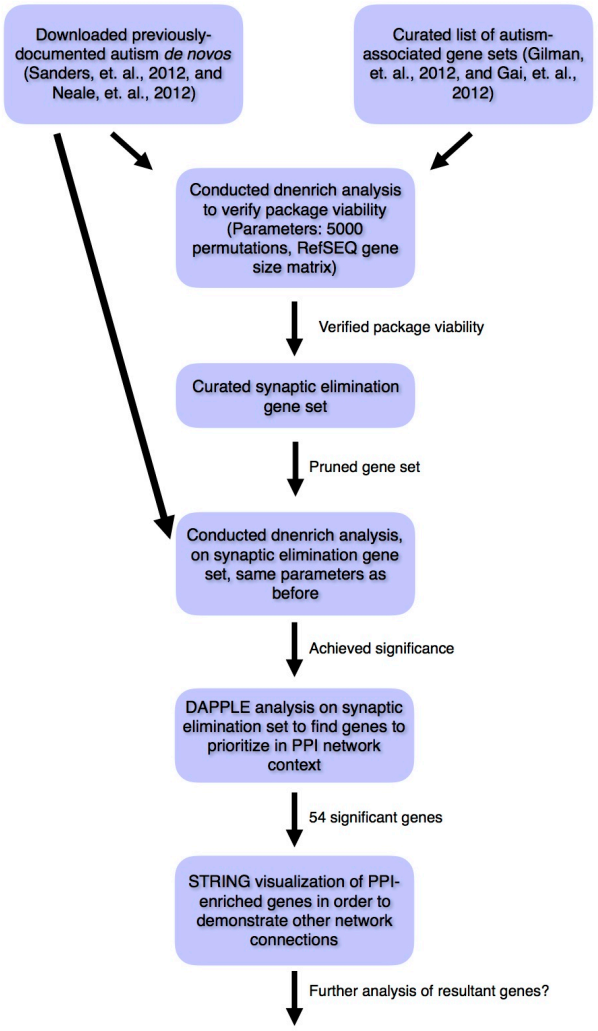

Fig 1. Schematic describing the overall flow of our experiment.

2. Methods

2.1. Autism *de novo* variants

We downloaded genomes of 3982 autism family trios from the Autism Sequencing Consortium (ASC) and the Simons Simplex Collection (SSC)[18-20]. These cohorts have been studied from a single-variant perspective, but have not yet been examined for their potential relationship to the synaptic elimination network.

Focusing on previously published full exome data, we built a comprehensive database of genomic variants to test for enrichment of synaptic elimination[18-21]. Specifically, we selected 189

autism trios and 31 unaffected siblings from SSC and then filtered out samples known to carry large *de novo* CNVs. Whole-exome sequencing was completed for 238 families selected from SSC, 200 of which included an unaffected sibling[23]. 15,480 DNA samples in 16 sample sets were analyzed, integrating *de novo*, inherited, and case-control loss-of-function counts and *de novo* missense variants predicted to be damaging. *De novos* were called using enhancements of previously published methods[18].

The full variant list from this collection, which also includes ASC cohorts, were compared to a larger set of 1,779 other exomes to confirm their putative roles in autism, and all *de novo* events were validated via PCR amplification and Sanger sequencing[22]. Family quads selected from SSC were sequenced with enrichment for higher functioning probands[19,20]. These *de novos* were interpreted using pipeline tools at each respective participating data center.

2.2. *Dnenrich Pilot Study*

Dnenrich simulates peppering the genome with random *de novos* by taking into account tri-nucleotide contexts, gene sizes, sequencing coverage, and functional effects of mutations. After permuting this process for a user-defined number of times, it then calculates one-sided P values, testing whether the observed number of mutations (in each gene set) is greater than the average simulated number of mutations (again, in each gene set).

We assembled 37 candidate gene sets to test their enrichment in our curated list of autism de novo variation. These 37 gene sets included Gene Ontology sets from previous autism network analyses[24,25], as well as genes shown to interact with FMRP (a mutation in which causes Fragile X syndrome, one of the most common causes of autism spectrum disorders)[21,26]. A full list of genes in each set tested for enrichment can be found in Supplemental Materials 1, along with their sizes.

We then performed an extensive process of literature mining and curation, through combined database search, hand-search, and related reference review, to assemble the synaptic elimination gene set. Our Pubmed search (conducted between May and June., 2016) included use of the terms "synapse," "synaptic," "elimination," "pruning," and "gene." We then performed additional hand-searches of *Nature* and *Cell* using the same terms. References of included studies, review articles, and related references were screened for additional relevant studies based on title and abstract review. Our screening criteria for inclusion in the synaptic elimination set was the presence of the following terms: "synaptic elimination," "synaptic pruning," "synaptic stabilization," "synaptic destabilization," and "synaptic plasticity." In all searches we excluded the following terms: "axon scaling," "viral infection," "axon repulsion," "axon retraction," and "neuromuscular junction." Studies pertaining to synaptic formation, maintenance, and/or elimination within the peripheral nervous system were excluded on the basis of arising from separate embryologic origin than the central nervous system. Abstracts and unpublished data were excluded. The synaptic elimination gene set was curated through careful review of 120 selected studies and related reviews yielding 274 genes related to synaptic formation or elimination. Gene function was cross-referenced in ClinVar (accessed July 11, 2016), and 213 genes of interest were selected based on their role in synaptic elimination (see Supplemental Materials 2). After its curation, we tested the synaptic elimination gene set for enrichment for autism *de novos*.

Table 1. The 213 genes in the synaptic elimination gene set.

C1QA	PROS1	TYROBP	CD200	IGFBP4	TLR4	NFKB1
C1QB	CXC3L1	NFKB1	ITGAX	EDNRB	BDNF	NFKB2
C1QC	CX3CR1	CREB	ITGB2	TIMP2	C5	CAMK2G
C3	DAP12	MAPK14	GDNF	COL1Q2	H2-D	NCKAP5L
Mac-2	TREM2	NPTX2	CSF1	FN1	CCL7	NRXN2
CRK	CR-1	NGFR	CNTF	IRF8	CCL2	NRXN3
ELMO1	PGRN	APP	PTGER2	TGRBR2	CDC42	NRTK2
RAC1	CD68	PILRB	C1QBP	CFB/MHCIII	MBP	CRMP1
BAI1	CASP8	CD247	CALR	FCGR1B	CXCL13	CRK
MEGF10	CASP3	B2M	CR2	AIF1	Uba1	PLXA3
GULP1	CASP6	KLRA1	CD33	IL10BR	Mov34	PLXA4
ABCA1	CLU	TAP1	TNFRSF19	NOS1AP	Rpn6	TBR1
TYRO3	HLA-DR	C4	PDGFRA	MASP1	USP2	DPYSL2
AXL	HLA-C	CR-3	LEP	CD46	UFD2A	ADNP
MERTK	HLA-A	CD22	LEPR	CD55	MEF2	SPARC
GAS6	DR6	CD47	IGFBP3	TLR2	MEF2A	DYRK1A
EN2	GDA	TSPAN7	PAK3	CTNNB1	WNT2	FOXP1
MEF2B	REL	NLGN1	SEMA3A	BDNF	CHN2	RCAN1
MEF2C	RELA	NLGN2	SEMA3F	DHCR7	MAPK3	CHD8
MEF2D	RELB	NLGN3	NRP1	FMR1	MAPK1	RAC1
PARK2	SERPINA3	NLGN4	NRP2	AUTS10	TSC1	OPHN1
caspases	CUL3	SHANK1	RhoA	LAMC3	TSC2	FOXP2
hdc	ESCRT-I	SHANK2	ROCK1	MECP2	DOCK1	ARC
MIB1	shrub	SHANK3	OTX1	THBS1	EPHA4	CASK
UBE3A	ESCRT-III	CNTN4	DISC1	THBS2	EPHB3	DLG4
UBE3B	CHMP2B	CNTNAP2	KATNAL2	THBS4	EFNA4	HOMER1
PCDH10	mop	CNTNAP4	NTNG1	MAP2	EFNB3	PTEN
ATG5	Kat60L	CACNA1C	SYNGAP1	KALRN	NCK2/GRB4	
Atg7	IKBKG	SCN1A	Mek-1	KALRN	EB3	
LC3-II	Mical	SCN2A	Mek-2	CDC42	NGFR	
p62	NRXN1	RELN	SPARCL1	PPP1R9B	GRM5	

3. Results

We tested the 37 initial gene sets with dnenrich with the default gene size matrix provided on the dnenrich website (as adjusting for per-trio joint sequencing coverage "[does] not have a noticeable effect on results"[17]). We ran the simulation on the downloaded autism *de novos* for 5000 permutations without weighting any genes. Of the 37 gene sets tested, 10 were significantly enriched for *de novos* after Bonferroni adjustment for 37 hypotheses. These sets are listed in Table

2. Given the enrichment of *de novos* in known autism networks calculated by dnenrich, we felt confident in using both this set of previously-published *de novos* and dnenrich to test the single hypothesis that synaptic elimination genes would have an exceedingly high burden of *de novos*.

Table 2. Bonferroni-significant gene sets enriched for autism *de novos* using dnenrich. Unadjusted p-values were obtained directly from dnenrich; adjusted p-values were Bonferroni-corrected by the number of sets tested.

	p-value		Number of Mutations		Location	
Gene Set Name	Unadjusted	Adjusted	Observed	Expected	Reference to Autism	Source
Developmental Process	1.9996×10^{-4}	8.798×10^{-3}	731	648.659	Gai et. al. (2012)	GO
FMRP	1.9996×10^{-4}	8.798×10^{-3}	412	285.33	Darnell et. al. (2011)	Paper
Learning and/or Memory	1.9996×10^{-4}	8.798×10^{-3}	78	44.1032	Gilman et. al. (2011)	GO
M3	1.9996×10^{-4}	8.798×10^{-3}	206	151.955	Parikshak et. al. (2013)	Paper
Protein modification process	1.9996×10^{-4}	8.798×10^{-3}	577	496.748	Gai et. al. (2012)	GO
Synaptic transmission	1.9996×10^{-4}	8.798×10^{-3}	163	114.367	Gai et. al. (2012)	GO
Axonogenesis	3.9992×10^{-4}	1.7596×10^{-2}	136	93.8364	Gilman et. al. (2011)	GO
Cell-cell signaling	3.9992×10^{-4}	1.7596×10^{-2}	241	188.513	Gai et. al. (2012)	GO
Neuron development	5.9988×10^{-4}	2.6395×10^{-2}	253	207.273	Gilman et. al. (2011)	GO
Axon	9.998×10^{-4}	4.3991×10^{-2}	121	87.635	Gilman et. al. (2011)	GO

Consistent with the synaptic elimination hypothesis, the synaptic elimination set also proved to be significantly enriched for autism *de novo* mutations ($p = 1.9996*10^{-4}$). It exceeded the observed-to-expected mutation ratio of all other significantly enriched gene sets (Figure 2).

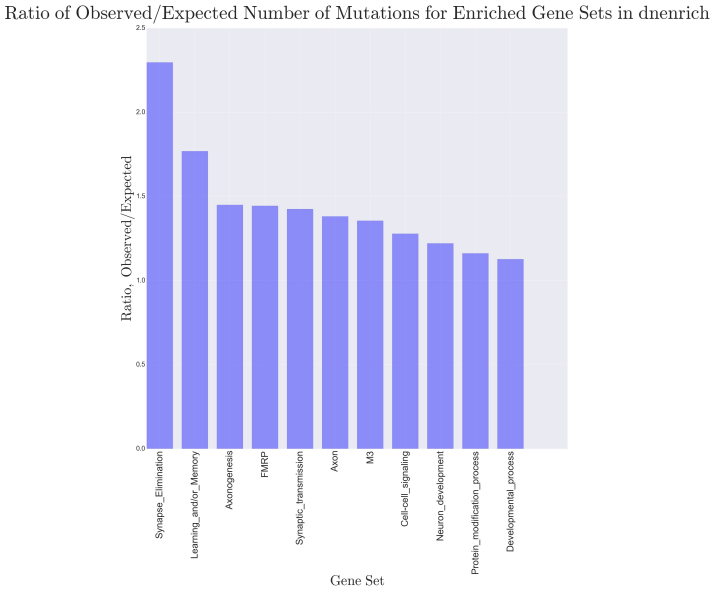

Fig 2. Ratio of observed-to-expected mutations per enriched gene set. The dnenrich software calculated expected number of mutations by simulating and averaging the number of *de novo* events in each gene set using information like tri-nucleotide context, gene size, etc. The systematically-generated synaptic elimination set has the highest ratio of observed-to-expected mutations by a significant margin.

To narrow the list of 213 genes in the synaptic elimination set down to a shorter list of genes to prioritize, we used DAPPLE. Developed by Elizabeth Rossin of the Broad Institute, the algorithm marks genes that are ripe for further study[28]. DAPPLE relies on protein-protein interaction databases such as InWeb (populated with hundreds of thousands of known protein-protein interactions). When researchers input a network, the algorithm compares the network it to what would be expected by pure probability by permuting proteins (linked to the inputted genes) many times. It determines if genes in all of the inputted regions could play a role in disease, and tests whether or not the network is more connected than would be expected by chance. For our purposes, this analysis would point to genes of interest within our synaptic elimination network that have higher levels of interconnectivity than expected.

Using DAPPLE (Figure 3) on the synaptic elimination gene set yielded fifty-four genes significantly enriched for protein-protein interactions (PPI), which are listed in Table 3. Six of these fifty-four (CACNA1C, SHANK2, SYNGAP1, NLGN3, NRXN1, and PTEN) have already been confirmed as genes associated with autism risk[27]. Those genes that were enriched were visualized using the STRING database (Figure 4) in order to examine other known (and predicted) gene interactions. Further inquiry into these resultant genes involved in synaptic elimination could elucidate etiology and shed light on related ASD risk.

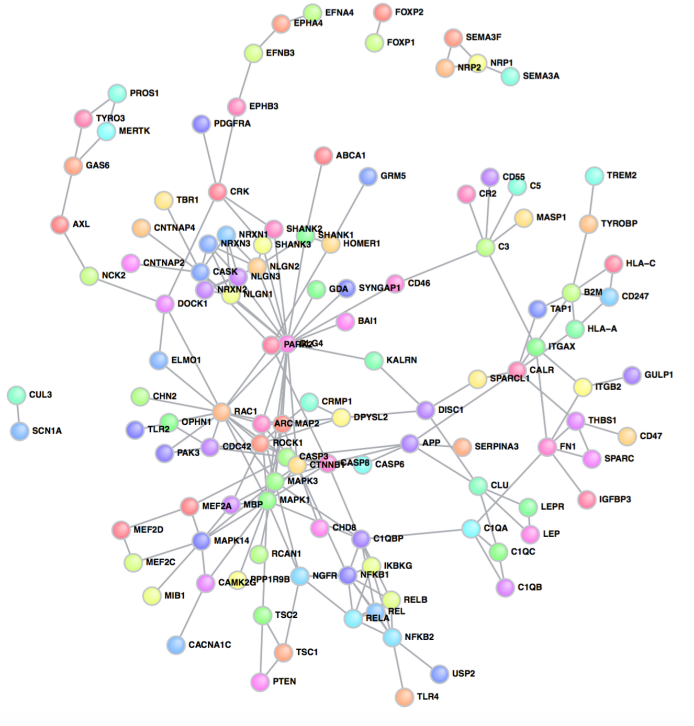

Fig 3. DAPPLE visualization of the synaptic elimination gene set. DAPPLE analyzes the protein-protein interaction network generated by the genes in the set; it marks the genes that are significantly more connected in the network than by chance (PPI-enriched). The nodes represent genes in the network, and the edges represent interactions between proteins downstream of the connected genes. The graphic is arbitrarily colored and is meant to show connectivity only.

Table 3. DAPPLE PPI-Enriched Genes in the synaptic elimination gene set. The table pairs genes with their DAPPLE significances.

Gene Name	*p*-value	Gene Name	*p*-value
DOCK1	0.005985024	CASK	0.001997004
HLA-A	0.045426102	TYRO3	0.005985024
CLU	0.00797604	HOMER1	0.025805363
SHANK1	0.00797604	SEMA3F	0.00797604
CD33	0.005985024	AXL	0.001997004
PARK2	0.003992012	SYNGAP1	0.00996506
ITGAX	0.005985024	CASP3	0.003992012
MERTK	0.045426102	TREM2	0.001997004
ELMO1	0.037601759	MAPK1	0.001997004
OPHN1	0.003992012	CACNA1C	0.035640683
CASP6	0.033677611	SHANK2	0.003992012
MAPK3	0.001997004	SHANK3	0.001997004
NRP1	0.01790118	CALR	0.013937112
NLGN1	0.001997004	NLGN2	0.001997004
GDA	0.011952084	NFKB2	0.013937112
NRXN3	0.001997004	DLG4	0.001997004
C3	0.001997004	TLR2	0.049326298
CTNNB1	0.001997004	ROCK1	0.00996506
CASP8	0.001997004	NRXN1	0.00797604
NGFR	0.001997004	MAP2	0.001997004
REL	0.023832312	EFNB3	0.037601759
ARC	0.001997004	MEF2A	0.01988022
RAC1	0.001997004	B2M	0.011952084
GAS6	0.001997004	NLGN3	0.001997004
NRP2	0.027776419	PTEN	0.00797604
GRM5	0.00996506	THBS1	0.039560839
NRXN2	0.001997004	NFKB1	0.001997004

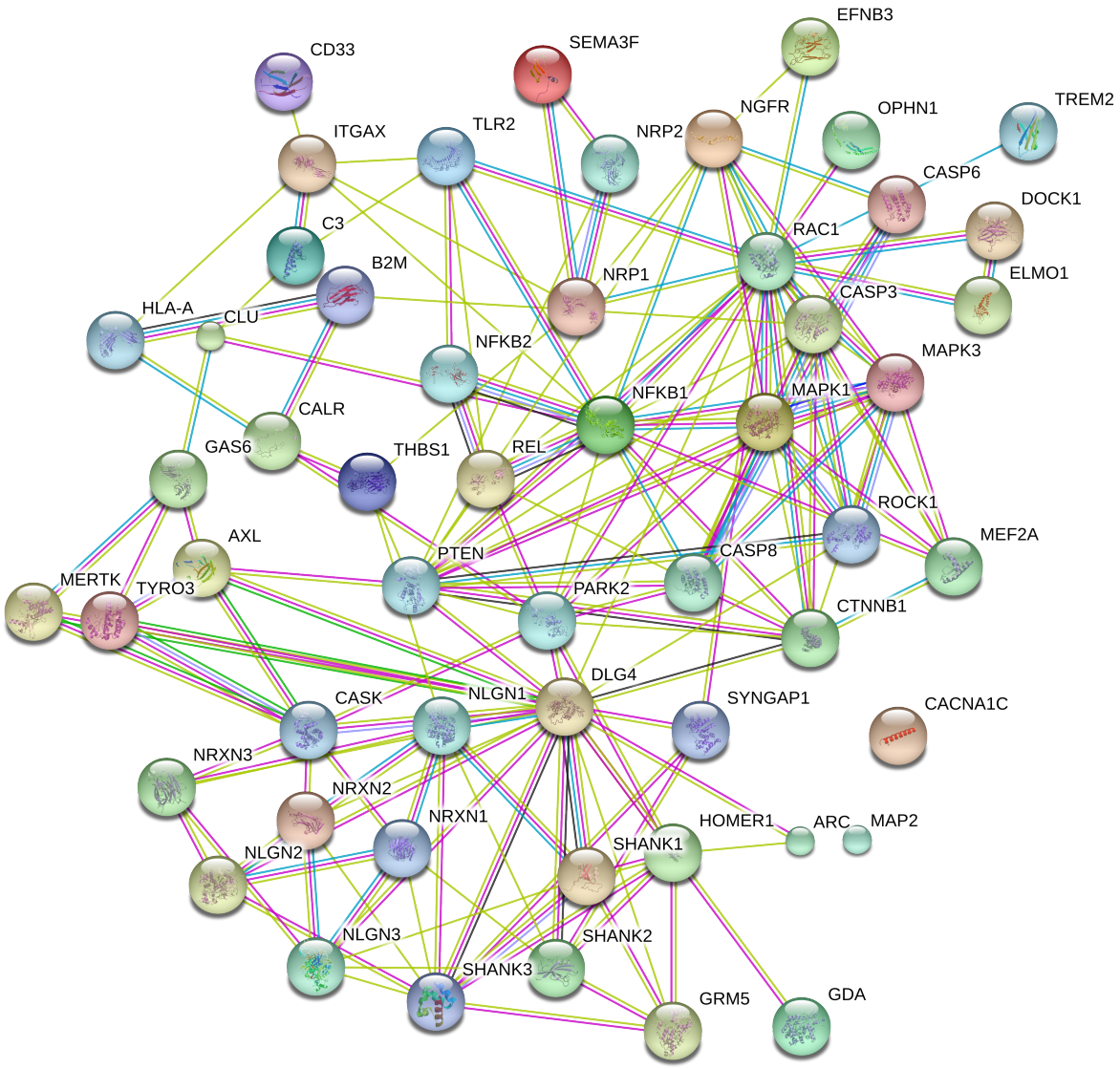

Fig 4. STRING visualization of DAPPLE PPI-enriched genes in the synaptic elimination gene set. Colored nodes represent query proteins and the first shell of interactors. White nodes represent the second shell of interactors. Cyan edges represent known interactions from curated databases; purple edges represent known interactions that are experimentally determined; green edges represent a gene neighborhood predicted interaction; red edges represent a gene fusion predicted interaction; blue edges represent a gene co-occurrence predicted interaction; yellow edges represent textmining; black edges represent co-expression; light blue edges represent protein homology.

4. Discussion

Overall, our results supported the hypothesis that genes involved in synaptic elimination are significantly enriched for autism *de novo* mutations, pointing to deregulation in synaptic elimination as a potential pathogenic mechanism for ASD. Synaptic elimination, as part of the larger synaptic homeostatic mechanism, contributes to higher structural and functional connectivity underlying cognitive functions through the removal of synaptic structures. Several of

the genes that were PPI-enriched in the gene set were confirmed autism disease genes, suggesting that the genes central to the synaptic elimination network may play an important role in influencing genetic risk for autism. Given the biological plausibility of this pathway, along with the enrichment for *de novos* in known autism cases, the additional genes in this pathway may serve as candidate genes in the future investigation of the genetic etiology of autism spectrum disorder.

Within the context of the hypothesis that *de novo* mutations contribute to the risk of developing ASD in families with no previous history, previous gene enrichment studies have focused on identification of these *de novo* mutations and their interconnections as a multifaceted network without exploration of specific neurodevelopmental processes[22]. In the present study, we took advantage of a large collection of full exomes from trios with one affected child. This enabled us to explore the role of de novo mutations in synaptic density and pruning, confirming that there is a strong link and supporting the potential value of these de novos for use in increasing precision in early diagnosis/prognosis.

The lack of a validation cohort is a drawback of this study. A new set of *de novos* is currently undergoing quality control procedures; we will attempt to replicate this signal in a much larger collection of families. A consortium that includes our group has amassed over 5000 whole genomes (30x coverage) in multiplex families containing 2 or more children with autism. This is the largest database of its kind and valuable for determining whether the de novo signal seen replicates across siblings and families with varying levels of autism severity. In addition, it may be worthwhile to consider the genes involved in synaptic formation or maintenance in addition to those involved in elimination. Gene sets like the GO Neuron Development or Cell-cell Signaling sets, which showed significant *de novo* mutation enrichment, provide a good starting point for future studies, as neuronal activity and signaling play a definitive role in determining synapse strength and number.

More work is necessary to determine the biological implications of the association between synaptic elimination and autism. For the PPI-enriched genes in the synaptic elimination network, many of which have validated associations with autism, the exact process by which they affect brain development leading to behavioral change is unclear. The true role of these genes in the pathophysiology of autism must be elucidated by future science.

Network analyses like these have successfully been able to identify and validate gene sets that contribute risk to ASD and other neuropsychiatric disorders. The high likelihood that these findings are reproducible in the context of newer, more complete, and more specific datasets bolsters the hope of eventually having a more complete picture of ASD risk factors that impact precision care of this complex disorder. Such a map would be invaluable to both the diagnosis and subsequent treatment of ASD; synaptic elimination may play a key role in that map.

Supplemental Materials

Supplemental Materials 1 – Gene set sizes and gene/gene set mappings – https://drive.google.com/open?id=0B4nOSzAytcrBdlBLVWJWNzFpcGc

Supplemental Materials 2 – synaptic elimination genes and sources – https://drive.google.com/open?id=0B2UCU6mZg1CuSXNKaEJOODNLcmc

Supplemental Materials 3 – synaptic elimination curation references – https://drive.google.com/open?id=0B2UCU6mZg1CudVNiYzJmWGxTWjA

Acknowledgements

This work was supported in part by the Hartwell Foundation's Autism Research and Technology Initiative. It was also conducted with support from a TL1 Clinical Research Training Program of the Stanford Clinical and Translational Science Award to Spectrum (NIH TL1 TR 001084).

References

1 Keen, D. & Ward, S. Autistic spectrum disorder a child population profile. *Autism* **8**, 39-48 (2004).

2 Buescher, A. V., Cidav, Z., Knapp, M. & Mandell, D. S. Costs of autism spectrum disorders in the United Kingdom and the United States. *JAMA pediatrics* **168**, 721-728 (2014).

3 Shattuck, P. T. *et al.* Services for adults with an autism spectrum disorder. *The Canadian Journal of Psychiatry* **57**, 284-291 (2012).

4 Nord, A. S. *et al.* Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. *European Journal of Human Genetics* **19**, 727-731 (2011).

5 Sung, Y. J. *et al.* Genetic investigation of quantitative traits related to autism: use of multivariate polygenic models with ascertainment adjustment. *The American Journal of Human Genetics* **76**, 68-81 (2005).

6 Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nature Reviews Neuroscience* **16**, 551-563, doi:10.1038/nrn3992 (2015).

7 Tang, G. *et al.* Article Loss of mTOR-Dependent Macroautophagy Causes Autistic-like Synaptic Pruning Deficits. *Neuron* **83**, 1131-1143 (2014).

8 Ebert, D. H. & Greenberg, M. E. Activity-dependent neuronal signalling and autism spectrum disorder. *Nature* **493**, 327-337, doi:10.1038/nature11860.Activity-dependent (2013).

9 Toro, R. *et al.* Key role for gene dosage and synaptic homeostasis in autism spectrum disorders. *Trends in Genetics* **26**, 363-372, doi:10.1016/j.tig.2010.05.007 (2010).

10 Bialas, A. R. & Stevens, B. TGF-Beta Signaling Regulates Neuronal C1q Expression and Developmental Syanptic Refinement. *Nature Neuroscience* **16**, 1773-1782, doi:10.1038/nn.3560.TGF- (2013).

11 Stevens, B. *et al.* The Classical Complement Cascade Mediates CNS Synapse Elimination. *Cell* **131**, 1164-1178, doi:10.1016/j.cell.2007.10.036 (2007).

12 Paolicelli, R. C. & Gross, C. T. Microglia in development: linking brain wiring to brain environment. *Neuron glia biology* **7**, 77-83, doi:10.1017/S1740925X12000105 (2011).

13 Chung, W.-S., Allen, N. J. A. & Eroglu, C. Astrocytes Control Synapse Formation, Function, and Elimination. *Cold Spring Harb Perspect Biol* **7**, doi:10.1530/ERC-14-0411.Persistent (2015).
14 Siegel, A. & Sapru, H. N. *Essential neuroscience*. (Lippincott Williams & Wilkins, 2006).
15 Caglayan, A. O. Genetic causes of syndromic and non-syndromic autism. *Developmental Medicine and Child Neurology* **52**, 130-138, doi:10.1111/j.1469-8749.2009.03523.x (2010).
16 Geschwind, D. H. & Levitt, P. Autism spectrum disorders : developmental disconnection syndromes. *Current Opinion in Neurobiology* **17**, 103-111, doi:10.1016/j.conb.2007.01.009 (2007).
17 Fromer, M. *et al.* De novo mutations in schizophrenia implicate synaptic networks. *Nature* **506**, 179-184 (2014).
18 De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209-215 (2014).
19 Iossifov, I. *et al.* De novo gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285-299 (2012).
20 Iossifov, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**, 216-221 (2014).
21 Parikshak, N. N. *et al.* Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008-1021 (2013).
22 O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246-250 (2012).
23 Sanders, S. J. *et al.* De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237-241 (2012).
24 Gilman, S. R. *et al.* Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron* **70**, 898-907 (2011).
25 Gai, X. *et al.* Rare structural variation of synapse and neurotransmission genes in autism. *Molecular Psychiatry* **17**, 402-411, doi:10.1038/mp.2011.10 (2012).
26 Darnell, J. C. *et al.* FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism. *Cell* **146**, 247-261, doi:10.1016/j.cell.2011.06.013 (2011).
27 Sanders, S. J. *et al.* Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron* **87**, 1215-1233 (2015).
28 Rossin, Elizabeth J., et al. "Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology." *PLoS Genet* 7.1 (2011): e1001273.