Title
Schizophrenia genetics complements its mechanistic understanding

Journal
Nature Neuroscience, 19(4)

ISSN
1097-6256

Authors
Ruzzo, Elizabeth K
Geschwind, Daniel H

Publication Date
2016-04-01

DOI
10.1038/nn.4277

Peer reviewed

HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2017 February 15.

Published in final edited form as:

Nat Neurosci. 2016 April ; 19(4): 523–525. doi:10.1038/nn.4277.

Schizophrenia genetics complements its mechanistic understanding

Elizabeth K Ruzzo and

Center for Neurobehavioral Genetics and Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA

Daniel H Geschwind

Center for Neurobehavioral Genetics and Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA

Abstract

Refinement of genetic variants within the major histocompatibility complex reveals three distinct genetic influences on schizophrenia risk and sheds light on the disease's neurobiology.

The rationale behind most genetic studies of psychiatric disease is to provide a causal mechanistic basis for the development of effective therapies[1,2]. However, this approach faces challenges at many levels, from the identification of bona fide risk genes, to modeling their effects in experimental systems, to translating experimental findings from these systems to focused human subject research, and eventually to the identification of targets with clinical utility. The advent of genome-wide association studies (GWAS) using high-density genotyping of single-nucleotide polymorphisms (SNPs) made it possible to scan large populations of affected individuals (cases) and controls to identify common genetic variants associated with common diseases. Once it became clear that enormous sample sizes were necessary to detect the relatively small effect sizes of common alleles, major collaborative efforts in psychiatric disease were initiated. For schizophrenia, this effort recently culminated in the identification of over 100 loci contributing to schizophrenia, a major milestone in the field[3]. The hope is that each of these hundred loci will guide us to the underlying gene or variant driving the disease association and that, when pieced together, these findings will expose the pathophysiological mechanisms of schizophrenia. A recent study published in *Nature* by Sekar *et al.*[4] takes this first critical step by identifying the functional alleles driving the signal in the most significantly associated schizophrenia GWAS locus, which lies within the major histocompatibility complex (MHC) locus on chromosome 6 (Fig. 1).

The MHC spans ~4 Mb and is one of the most gene-dense regions of the genome. It contains genes that function in the acquired immune system, including 18 highly polymorphic human

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

leukocyte antigen (HLA) genes. The prospect of identifying the functional alleles driving this MHC association is daunting for two main reasons: (i) the MHC is highly polymorphic and has many repetitive sequences and extensive linkage disequilibrium (large regions co-inherited as blocks) and (ii) the pattern of schizophrenia-associated markers is complex, with multiple associated markers in close proximity to hundreds of genes[3]. Given that the strongest association within the MHC lies near the complement component 4 (C4) genes and that a genome-wide significant association on chromosome 8 lies near *CSMD1* (ref. 3), which encodes a protein thought to regulate C4 (ref. 5), Sekar *et al.*[4] investigated whether variants in C4 could partially or fully explain this schizophrenia risk locus.

Human C4 is encoded by two genes, *C4A* and *C4B*, whose protein products target distinct molecules. A human endogenous retroviral (HERV) element that may function as an enhancer[6] is either present in or absent from the ninth introns of *C4A* and *C4B*, resulting in long or short forms of the genes, respectively. In addition to these four structural alleles for C4, humans also differ in the number of copies of each gene and have distinct SNP haplotypes in these genes. Regions containing repetitive elements or variable repeats are notoriously difficult to genotype. Crucially, Sekar *et al.*[4] developed a method to accurately genotype both the structure and copy number of the *C4A* and *C4B* genes. To assess how these aspects of *C4A* and *C4B* influence C4 gene expression, they analyzed 674 samples from among 5 regions in 245 postmortem human brains. The investigators found that the RNA expression of *C4A* and *C4B* increased proportionally with the number of genic copies, that *C4A* expression was two to three times higher than that of *C4B* (controlled for copy number) and that this ratio of *C4A* to *C4B* expression increased with copy number of C4-HERV.

Having established a model for how both the structure and the copy number of *C4A* and *C4B* influence C4 gene expression in the brain, Sekar *et al.*[4] were able to generate a genetic predictor of *C4A* and *C4B* expression without directly assessing RNA levels. Similarly, studying the SNP haplotypes across the MHC in 222 independent chromosomes from 111 unrelated individuals revealed that a given SNP haplotype was usually associated with a characteristic C4 locus structure, enabling the authors to impute the four most common C4 locus structural variants with reasonably high accuracy from SNP genotype data alone. Together, these two strategies enabled the researchers to predict C4 locus structure and subsequently infer C4 expression from the genome-wide SNP data of the Psychiatric Genomics Consortium's ~29,000 schizophrenia cases and ~36,000 controls, which were the basis of the original GWAS findings[3].

The investigators then examined ~8,000 SNPs in the extended MHC locus, testing the C4 structural alleles, the HLA sequence polymorphisms (imputed) and the *C4A* and *C4B* expression levels for their association with schizophrenia. This analysis identified two independent genetic association signals in the MHC: one at the distal end of the extended MHC locus, with associated SNPs spanning ~2 Mb, and the other centered at C4. Interestingly, SNPs within the C4 association peak were associated with schizophrenia risk in proportion to their effects on *C4A* expression. Controlling for both of these genetic signals revealed another genome-wide-significant association peak proximal to MHC, near *SYNGAP1*.

Evaluating the risk conferred by each C4 allele, the investigators established a ranking for the four common structural alleles, with relative risks for schizophrenia ranging from 1.0 to 1.27. Furthermore, the predicted *C4A* expression level for each C4 structural allele mirrored the allelic risk, with the highest *C4A* gene expression predicted in samples carrying the C4 allele with the highest schizophrenia risk. Measurement of *C4A* RNA expression in brain tissue from 35 patients with schizophrenia and 70 unaffected controls revealed that, as predicted from this genetic finding, *C4A* expression was higher in brains from patients than in those from controls. The elevation in *C4A* expression was observed in all five brain regions examined, including the cerebellum—which, although not a classical schizophrenia-associated brain region, shows evidence of dysfunction in schizophrenia[7]. How the focal cognitive and behavioral abnormalities present in schizophrenia arise from what appears to be a relatively widespread change in gene expression remains to be determined.

In the adolescent and early adult human brain, mature neural circuits are formed through a process of synapse refinement. Previous characterization of brains from patients with schizophrenia identified a reduction in the number of synapses[8], suggesting that this process is altered in the disease state. Earlier studies have shown that members of the classic complement cascade, of which C4 is a key member, are critical mediators of this synaptic pruning[9–11]. In the immune system, C4 is known to promote activation of complement component 3 (C3); in the developing mouse brain, C3 can target specific synapses and is required for synapse elimination by microglia[9,10]. Sekar *et al.*[4] investigated the localization of C4 in the human brain, finding C4-positive neurons and synapses in hippocampal slices. The investigators then studied an available C4 knockout mouse (only one copy of *C4* exists in mice) and found that C3 was greatly reduced in the presynaptic terminals of the homozygous deletion mice and that synaptic refinement (as modeled in the retinogeniculate pathway of the visual system) was altered in both the heterozygous and homozygous C4 knockout mice.

The polygenicity of common diseases, such as schizophrenia, poses many challenges to disease modeling and mechanistic studies. In addition to the difficulty of identifying the causal loci tagged by these alleles, even when identified, causal common alleles typically have individually small effect sizes. Thus, most disease models derived from genetic findings are based on rare alleles with large effect sizes. However, investigating large-effect-size mutations in model systems is not necessarily a panacea; a good example is the case of the fragile X gene (*FMR1*) in autism-spectrum disorder (ASD), where, although a large-effect risk gene has been identified, the development of associated therapeutics has been challenging[12]. Nevertheless, understanding how small-effect-size alleles such as *C4A* act combinatorially with other risk variants to cause schizophrenia is a major challenge. Furthermore, moving from understanding risk alleles in the population to understanding the specific risk profile of an individual remains a formidable task[1]. However, if these small-effect-size alleles do converge into a shared pathway(s), then it is conceivable that, because of their modifiability either by genetic background or environment, they are more amenable than major risk alleles to drug targeting of key molecules within the pathway.

This study provides an excellent model for the meticulous identification and characterization of causal, small-effect-size, common loci identified in GWAS. This work suggests that

excessive dosage of C4 leads to increased postnatal synaptic pruning, providing the first potential mechanism for the previously observed loss of grey matter[13] and reduced numbers of synaptic structures in the neurons of patients with schizophrenia[8]. Importantly, if this aberrant synaptic pruning is critical to disease pathogenesis, then this also raises the question of whether other schizophrenia susceptibility loci found in an individual patient act in the same biological pathway as *C4A* to affect synaptic pruning, act in multiple distinct pathways leading to increased synaptic pruning or, more complicated still, create vulnerability by limiting the activities of other, potentially compensatory pathways. Certainly, connecting excessive synaptic pruning to schizophrenia pathogenicity by no means proves that this is the only pathogenic mechanism at work in schizophrenia, but it is intriguing to consider whether clinical subtypes can be clustered by their underlying mechanistic origins. Disentangling these possibilities assumes not only that we are able to identify the causal loci themselves, but also that we can recognize and group them into biological pathways once identified. Given that the *C4A* gene was not previously annotated with a nonimmune function (via gene ontology databases), it is almost certain that many incomplete functional gene annotations exist, and the amount of biological characterization required for each identified gene and gene-gene interaction may be extensive.

Even if a risk gene is functionally annotated and preliminary neurobiology studies have been conducted, a complete understanding of mechanism in the context of disease pathophysiology in humans requires appropriate modeling. To interpret how elevated *C4A* expression increases schizophrenia risk, Sekar *et al.*[4] studied the retinogeniculate pathway of the visual system in C4 knockout mice. The most obvious question raised by this approach is whether conclusions drawn from knocking out C4 can be translated to understanding the function of C4 overexpression. In mice, the isotypes of *C4A* and *C4B* are encoded by *C4* alone; and although sequence similarity suggests parallel function, the implication that *C4A* dosage alone increases risk in humans is difficult to scrutinize in this model system. Finally, although they are undoubtedly relevant, we have no formal rubric to guide translation of the results obtained from modeling in a primary sensory circuit of a divergent species to synaptic pruning in higher association areas of the human brain that are implicated in schizophrenia.

Furthermore, since synaptic pruning is implicated in other neurological disorders, we need to reach a more refined understanding of which aspects of this process provide disorder specificity. One possibility is that synaptic pruning in these other disorders is caused by perturbation of a pathway other than C4 and that this other pathway shows temporal, spatial or cell type differences. Support for this theory comes from ASD, where synaptic abnormalities have been reported[14], but preliminary results from the Psychiatric Genomics Consortium's ASD GWAS (unpublished data) show no hint of association in the MHC locus. This preliminary GWAS included ~6,500 ASD cases, a sample size on the same order of magnitude as that with which MHC was detected in schizophrenia[15], suggesting that ASD synaptic pruning defects are likely not caused by the C4 locus. Given that antipsychotic drugs blocking the dopamine D2 receptor (DRD2) are particularly effective in some schizophrenia patients, one obvious avenue to explore is the possibility that C3-mediated synaptic pruning preferentially targets dopaminergic neurons or the frontal lobe circuits that they project to, and thus clinical responsiveness may also correlate with C4 allele genotype.

Finally, one very important observation is the approximately twofold elevation in concordance for schizophrenia in dizygotic twins as compared to non-twin siblings, which, along with other epidemiological data, implicates shared environmental influences (in addition to genetics) in schizophrenia risk. Given the dosage sensitivity of synaptic pruning to C4 levels and their potential environmental regulation (for example, via maternal infection or immune activation), C4 alleles provide a potential target for gene-environment interactions to modulate risk for schizophrenia by shaping brain circuit development. So, although we still lack a comprehensive view of the pathophysiological basis of schizophrenia, Sekar *et al.*[4] have taken a major step toward building a solid foundation for understanding of the biological mechanisms leading to schizophrenia and provide hope that such neurobiological understanding will soon catch up to the recent progress in genetics.

Acknowledgments

We thank members of the Geschwind laboratory for their critical reading of this manuscript. We apologize to any authors whose work we were unable to cite. DHG is supported with NIH funding from 5R01MH100027 and U01MH105578.

References

1. Geschwind DH, Flint J. Science. 2015; 349:1489–1494. [PubMed: 26404826]
2. Nestler EJ, Hyman SE. Nat Neurosci. 2010; 13:1161–1169. [PubMed: 20877280]
3. Ripke S, et al. Nature. 2014; 511:421–427. [PubMed: 25056061]
4. Sekar A, et al. Nature. 2016; 530:177–183. [PubMed: 26814963]
5. Kraus DM, et al. J Immunol. 2006; 176:4419–4430. [PubMed: 16547280]
6. Suntsova M, et al. Proc Natl Acad Sci USA. 2013; 110:19472–19477. [PubMed: 24218577]
7. Baumann O, et al. Cerebellum. 2015; 14:197–220. [PubMed: 25479821]
8. Glausier JR, Lewis DA. Neuroscience. 2013; 251:90–107. [PubMed: 22546337]
9. Stevens B, et al. Cell. 2007; 131:1164–1178. [PubMed: 18083105]
10. Schafer DP, et al. Neuron. 2012; 74:691–705. [PubMed: 22632727]
11. Bialas AR, Stevens B. Nat Neurosci. 2013; 16:1773–1782. [PubMed: 24162655]
12. Jeste SS, Geschwind DH. Sci Transl Med. 2016; 8:321fs1.
13. Cannon TD, et al. Biol Psych. 2015; 77:147–157.
14. Tang G, et al. Neuron. 2014; 83:1131–1143. [PubMed: 25155956]
15. International Schizophrenia Consortium. et al. Nature. 2009; 460:748–752. [PubMed: 19571811]

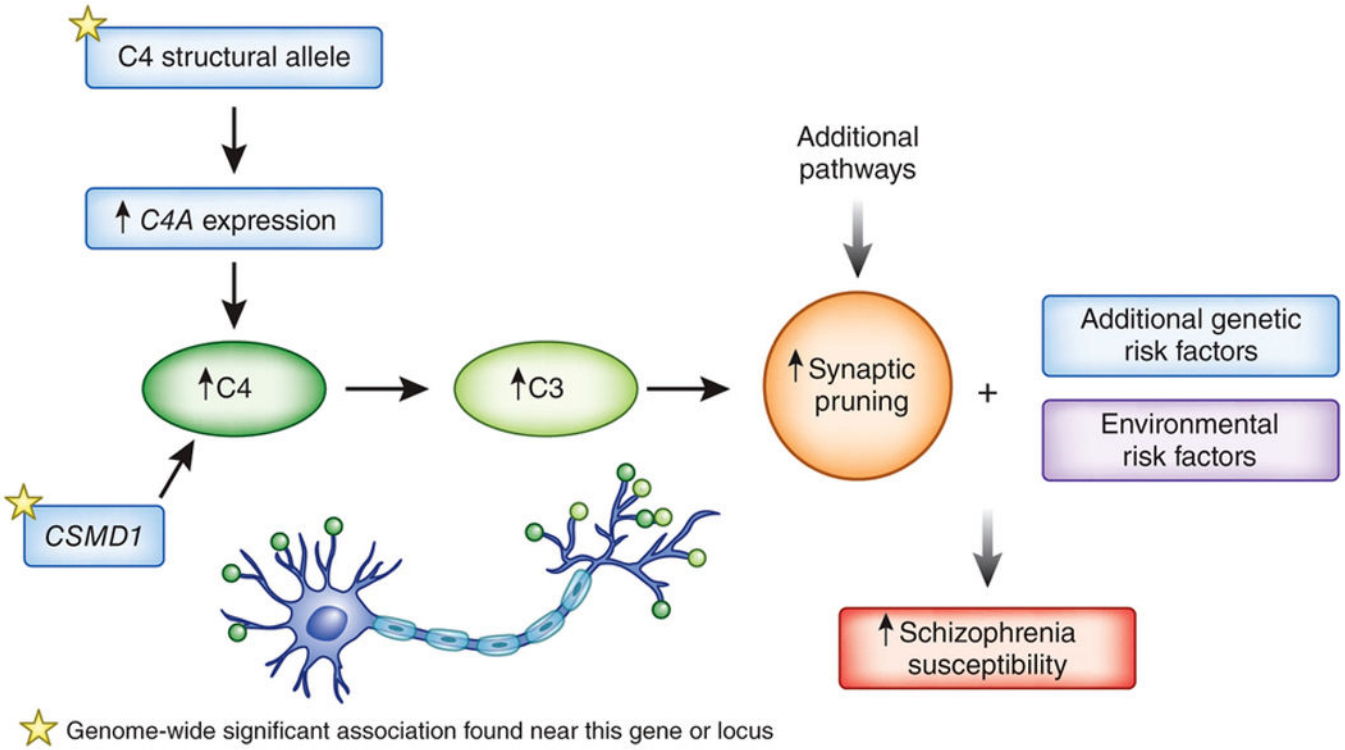

Figure 1.

Careful refinement of the schizophrenia GWAS locus in the MHC revealed that structural alleles of the C4 locus increase schizophrenia risk. These structural alleles increase *C4A* RNA levels in human brain, which predict a subsequent increase in C3, increasing synaptic pruning. A mouse knockout of C4 demonstrated that C3 levels decreased and synaptic pruning in the visual system was disrupted, which may be consistent with the model whereby an increase in human *C4A* expression results in increased synaptic pruning in schizophrenia. Points of potential convergence of other influences that may increase risk for this complex condition, such as environment and other genetic influences, are also indicated. Marina Corral Spence/Nature Publishing Group