Brain Research Bulletin 83 (2010) 93–102

Contents lists available at ScienceDirect

Brain Research Bulletin

journal homepage: www.elsevier.com/locate/brainresbull

Review

Genome-wide approaches to schizophrenia

Jubao Duan*, Alan R. Sanders, Pablo V. Gejman

Center for Psychiatric Genetics, Department of Psychiatry and Behavioral Sciences, Northshore University HealthSystem Research Institute, 1001 University Place, Evanston, IL 60201, USA

ARTICLE INFO

Article history:
Received 20 October 2009
Received in revised form 20 April 2010
Accepted 21 April 2010
Available online 28 April 2010

Keywords:
Genome-wide association
Schizophrenia
Polygenic
CNV
Rare variant
Gene expression
Systems biology

ABSTRACT

Schizophrenia (SZ) is a common and severe psychiatric disorder with both environmental and genetic risk factors, and a high heritability. After over 20 years of molecular genetics research, new molecular strategies, primarily genome-wide association studies (GWAS), have generated major tangible progress. This new data provides evidence for: (1) a number of chromosomal regions with common polymorphisms showing genome-wide association with SZ (the major histocompatibility complex, MHC, region at 6p22–p21; 18q21.2; and 2q32.1). The associated alleles present small odds ratios (the odds of a risk variant being present in cases vs. controls) and suggest causative involvement of gene regulatory mechanisms in SZ. (2) Polygenic inheritance. (3) Involvement of rare (<1%) and large (>100 kb) copy number variants (CNVs). (4) A genetic overlap of SZ with autism and with bipolar disorder (BP) challenging the classical clinical classifications. Most new SZ findings (chromosomal regions and genes) have generated new biological leads. These new findings, however, still need to be translated into a better understanding of the underlying biology and into causal mechanisms. Furthermore, a considerable amount of heritability still remains unexplained (missing heritability). Deep resequencing for rare variants and system biology approaches (e.g., integrating DNA sequence and functional data) are expected to further improve our understanding of the genetic architecture of SZ and its underlying biology.

© 2010 Elsevier Inc. All rights reserved.

Contents

1. Overview

The GWAS approach is based on linkage disequilibrium (LD), which is the non-random association of alleles (alternative forms of a polymorphism) at different loci, and is implemented with single nucleotide polymorphism (SNP) arrays that interrogate common variation across the genome. GWAS experiments have been remarkably successful. Over 900 genes have been reported to be associated at genome-wide significant levels ($p < 5 \times 10^{-8}$) [22] to one or more of ~200 complex phenotypes (http://www.genome.gov/gwastudies as of 02/12/2010 [45,77]). Current array platforms capture ~80% of the common variation in the genome for European ancestry (EA) samples [70]. Imputation, the computational prediction of genotypes of untyped SNPs, extends GWAS map coverage [41,96], and enables the combined analysis of samples genotyped with different platforms. The same arrays used for testing SNP associations carry probes designed for the detection of CNVs, but lower accuracy, particularly for duplications.

2. Main GWAS findings in SZ

Table 1 summarizes the published GWAS for SZ [68,125,25,93,107,124,118]. No single study detected genome-wide significant association with individual SNPs. Only a

* Corresponding author. Tel.: +1 224 364 7564; fax: +1 224 364 7570.
E-mail address: jduan69@gmail.com (J. Duan).

0361-9230/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.brainresbull.2010.04.009

Table 1
Top genes or genomic regions identified in recent SZ GWAS.

First author and year	Sample	Gene or region	Lowest *p*-values	OR	Reference
Lencz 2007	178/144 (EA)	*CSF2RA, SHOX*	3.7×10^{-7}	3.23	[68]
Sullivan 2008	738/733 (EA)	*AGBL1*	1.71×10^{-6}	6.01	[125]
O'Donovan 2008	Discovery: 479/2937 (EA); follow up: 6829/9897 (EA)	*ZNF804A*	1.61×10^{-7}	1.12	[97]
Need 2009	Discovery: 871/863 (EA); follow up: 1460/12,995 (EA)	*ADAMTSL3*	1.35×10^{-7}	0.68	[93]
Purcell 2009 (ISC)	3322/3587 (EA)	MHC region[a]	9.5×10^{-9}	0.82	[107]
		MYO18B	3.4×10^{-7}		
Stefansson 2009 (SGENE)	Discovery: 2663/13,498 (EA)	MHC region[b]	1.4×10^{-12}	1.16[c]	[124]
	Follow up: 4999/15,555 (EA)	*NRGN*[b]	2.4×10^{-9}	1.15	
		TCF4[b]	4.1×10^{-9}	1.23	
Shi 2009 (MGS)	2681/2653 (EA)	MHC region[a]	9.5×10^{-9}	0.88	[118]
	1286/973 (AA)	*CENTG2* (in EA only)	4.59×10^{-7}	1.23	
		ERBB4 (in AA only)	2.14×10^{-6}	0.73	

[a] Combined analysis of ISC, SGENE, and MGS GWAS.
[b] Combined analysis of ISC, SGENE (including SGENE follow up samples) and MGS.
[c] OR is for common allele of the associated SNP, which is different from that in ISC and MGS.

meta-analysis of 8008 EA cases and 19,077 EA controls, in regions with individual study *p*-values < 0.001, detected a genome-wide significant association at the MHC locus on chromosome 6p [107,124,118]. The most significant SNP (rs13194053) reached $p = 9.54 \times 10^{-9}$ (Table 1). The odds ratios, a measure of effect size, are small (e.g., the strongest associations at the MHC had ORs ranging 1.14–1.16; other associations with common SNPs also show low ORs). The MHC region is very gene-dense, containing over 200 genes. The low recombination rate at the MHC causes long LD blocks within the region [130]. Because of this long range LD and high gene density, it has not yet been determined whether one or more genes or intergenic regions are implicated. The MHC region contains many genes with roles in immunity and self-recognition, and has been implicated by GWAS in multiple common immune diseases, including type 1 diabetes (T1D), multiple sclerosis, Crohn's disease (CD), and rheumatoid arthritis (RA) (see review [45]). Furthermore, one of the three studies also found a suggestive association for SZ with the *FAM69A-EVI-RPL5* gene cluster (1p22) [118], which was previously reported as associated with multiple sclerosis [99]. Consistent with an immune hypothesis of SZ [42], a recent Danish registry study indicated that autoimmune disorders increase the risk for SZ [24]. However, final proof of immune abnormalities in SZ is still missing.

TCF4 (*transcription factor 4*), located in chromosome 18q21, has also been identified as another novel SZ susceptibility locus [124]. *TCF4* is a neuronal transcription factor essential for neurogenesis [39]. Mutations in *TCF4* cause Pitt–Hopkins syndrome, a disorder characterized by severe motor and mental retardation [27,54,6].

Another GWAS reported an association of SZ with *zinc finger protein 804A* (*ZNF804A*) in 2q32.1 [97]. Although not genome-wide significant for SZ ($p = 1.61 \times 10^{-7}$ with rs1344706), the association reached genome-wide significance in the combined analysis of SZ and BP ($p = 9.96 \times 10^{-9}$) [97]. In a subsequent meta-analysis of a much larger dataset with 18,945 SZ cases, 2329 BP cases, and 38,675 controls, support strengthened: $p = 2.5 \times 10^{-11}$ for SZ and $p = 4.1 \times 10^{-13}$ for SZ and BP [146]. *ZNF804A* was reported associated with altered neuronal connectivity in the dorsolateral prefrontal cortex in a functional magnetic resonance imaging study of healthy controls [25].

Thus, several SZ susceptibility loci have emerged from GWAS and meta-analyses thereof. None of the genome-wide significant loci implicated by GWAS span leading pathophysiological SZ candidate genes (e.g., those examined in Ref. [111]), and thus may represent new biological leads.

3. Evidence for a polygenic model

Purcell et al. [107] tested the polygenic hypothesis of SZ [34] by evaluating whether many common variants with small effects could explain a large proportion of the variation in disease risk. Based on different thresholds of association *p*-values in the International SZ Consortium (ISC) dataset, sets of common variants with small effects ("score" alleles") were first defined. For each set, an aggregate risk score was then generated for each subject from the Molecular Genetics of SZ (MGS) EA and African American (AA) datasets, and a SZ UK sample [118,97]. Aggregate risk scores in SZ cases were found to be higher than in controls in SZ samples and also for BP cases from two BP samples [121,148]. Performing the same basic analysis for T1D, T2D, hypertension, CD, RA, and coronary artery disease [107,148] subsequently showed that the result was not an artifact of population stratification, genotyping quality, or other potential systematic biases. Although the variance explained by the observed score alleles derived from ISC study was only ~3%, with simulated datasets, the variance explained by a set of score alleles reached ~1/3. A set of causal alleles with minor allele frequency (MAF) <5% did not fit the model well [107], but a role for rare susceptibility variants could not be excluded. Like for other complex disorders, a large proportion of SZ heritability still remains unexplained by GWAS. Further increasing GWAS sample sizes, including meta-analyses (e.g., the ongoing Psychiatric GWAS Consortium, PGC, for SZ [12]), is expected to provide incremental evidence for the known and as yet undiscovered common loci.

4. Rare CNVs implicated in SZ

About one-quarter of the human genome harbors CNVs, stretches of DNA deletions or duplications ranging from 1 kb to several Mb [49]. Multiple CNVs have been suggested associated with SZ (Table 2). So far, only rare (<1%) and large (>100 kb) CNVs have been implicated in SZ [93,62,139,150,123,52,61] as reflected by overall CNV burden (aggregate) and individual CNV loci. Initial genome-wide CNV scans using 200–400 subjects observed 3- to 8-fold over-representation in SZ cases [139,150], but subsequent studies with larger sample sizes (>2000) revealed smaller effects of CNVs in aggregate (Table 2) [123,52]. A 3 Mb 22q11.21 deletion, known as 22q11 deletion syndrome (22qDS), was increasingly linked to SZ in the 1990s [55,119,3]. 22qDS still is the only CNV that has reached genome-wide significance in a single GWAS of SZ [52]. 22qDS causes DiGeorge or velocardiofacial syndrome (DGS/VCFS), one of the most common anomaly syndromes with >180 variable clinical features, both physical and behavioral, such as congenital

Download English Version:

https://daneshyari.com/en/article/4319056

Download Persian Version:

https://daneshyari.com/article/4319056

Daneshyari.com