European Neuropsychopharmacology (2014) **24**, 1567–1577

www.elsevier.com/locate/euroneuro

REVIEW

Genome-wide association studies of suicidal behaviors: A review

Marcus Sokolowski*, **Jerzy Wasserman**, **Danuta Wasserman**

National Centre for Suicide Research and Prevention of Mental Ill-Health (NASP), Karolinska Institute (KI), S-171 77 Stockholm, Sweden

Received 21 March 2014; received in revised form 24 July 2014; accepted 10 August 2014

KEYWORDS
Suicide attempt;
GWAS;
SNP;
Gene;
Pathway;
Network

Abstract
Suicidal behaviors represent a fatal dimension of mental ill-health, involving both environmental and heritable (genetic) influences. The putative genetic components of suicidal behaviors have until recent years been mainly investigated by hypothesis-driven research (of "candidate genes"). But technological progress in genotyping has opened the possibilities towards (hypothesis-generating) genomic screens and novel opportunities to explore polygenetic perspectives, now spanning a wide array of possible analyses falling under the term Genome-Wide Association Study (GWAS). Here we introduce and discuss broadly some apparent limitations but also certain developing opportunities of GWAS. We summarize the results from all the eight GWAS conducted up to date focused on suicidality outcomes; treatment emergent suicidal ideation (3 studies), suicide attempts (4 studies) and completed suicides (1 study). Clearly, there are few (if any) genome-wide significant and reproducible findings yet to be demonstrated. We then discuss and pinpoint certain future considerations in relation to sample sizes, the units of genetic associations used, study designs and outcome definitions, psychiatric diagnoses or biological measures, as well as the use of genomic sequencing. We conclude that GWAS should have a lot more potential to show in the case of suicidal outcomes, than what has yet been realized.

1. Introduction

A general genetic diathesis of suicidality has been shown to exist by family, adoption and twin studies, with heritability in the range between 30–55% (Voracek and Loibl, 2007). Such a general genetic diathesis mainly refers to the behavioral manifestations of suicidality (i.e. suicide attempts or completed

*Corresponding author. Tel.: +46 8 52 48 69 38; fax: +46 8 30 64 39.
E-mail address: marcus.sokolowski@ki.se (M. Sokolowski).

http://dx.doi.org/10.1016/j.euroneuro.2014.08.006
0924-977X/

suicides), rather than suicidal thoughts or ideation alone. However, this is not clear-cut, as ideation precede a proportion of suicide attempts or completed suicides (Baca-Garcia et al., 2011). Specific genetic components (i.e. particular candidate genes) are therefore often being studied, i.e. genes which are implicated in suicidality or its associated endophenotypes (Mann et al., 2009), selected on the basis of prior biological and/or pharmacological knowledge and observations (Currier and Mann, 2008; Wasserman et al., 2009). Selected genetic variants in serotonergic system genes have been the major historic focus of such investigations during the past decades, but the polygenetic perspective is nowadays also being addressed by e.g. the study of candidate genes in other neurosystems (Ben-Efraim et al., 2011, 2012, 2013; Ernst et al., 2009; Rujescu and Giegling, 2010; Sokolowski et al., 2012, 2010; Wasserman et al., 2005, 2007, 2008, 2009). A relatively novel approach for further elucidating the polygenetic perspective, which has become available in the last years, is to conduct a Genome-Wide Association Study (GWAS).

GWAS is performed using many genetic markers across the whole genome to analyze for association with a trait (Ikegawa, 2012). It has been termed as a hypothesis-free approach (with regard to that any of the markers/genes are being hypothesized to show association), which is in contrast to the candidate gene approach (Goldstein et al., 2003). One main goal of GWAS is to suggest novel candidate genes which could not be hypothesized a priori by current knowledge. The majority of markers analyzed in GWAS are Single Nucleotide Polymorphisms (SNPs), in addition to certain Copy Number Variants (CNVs). Technological progress of genotyping chips (with Illumina Inc leading the market) combined with the efforts to map all human SNPs and their patterns of linkage disequilibrium (LD) in different populations (see e.g. http://www.hapmap.org and http://www.1000genomes.org), have resulted in a continuous increase in the number of SNPs analyzed for each year, leading up to the present $\sim$5 million low and high allele-frequency SNPs assayed on one GWAS chip (e.g. the Illumina HumanOmni5-Quad chip). Given the multiple association tests subsequently conducted, one today usually requires a P-value in the range of 10^{-8} to declare a certain SNP significant at alpha 0.05 after accounting for LD (Li et al., 2012b). In addition, samples in the sizes of tens of thousands are the goal to obtain sufficient power to detect the usually small SNP-by-SNP effect sizes, as have been observed for many complex traits. The initial enthusiasms about GWAS were somewhat damped (Need and Goldstein, 2010; Rowe and Tenesa, 2012) as "genome-wide significant" GWAS SNP-associations were found to be of small effect, difficult to generalize and thus not directly applicable to real life, e.g. in the clinic or in public health prevention (Eichler et al., 2010). But GWAS data nevertheless remain a rich resource for continuing to elucidate polygenetic etiology. Its success (Teslovich et al., 2010) or failure (Sebastiani et al., 2010, 2011) to generate new hypotheses and insights seem to depend on the trait under study, the study design, analysis approaches and avoidance of previous mistakes (Califano et al., 2012; Lambert and Black, 2012; Need and Goldstein, 2010; Rowe and Tenesa, 2012; Terwilliger and Goring, 2009). Promising novel ways to analyze GWAS data on poly-SNP, gene-, pathway- or network-levels are being developed and applied (complementing the classical SNP-by-SNP approach), which are (i) less penalized by multiple comparison corrections and (ii) attempt to better capture a polygenetic biology more efficiently. Therefore, there are now different analytical approaches to GWAS data available and their utility is ultimately decided by the mainly unknown genetic architecture one tries to map.

1.1. Classical SNP-by-SNP GWAS

The wide majority of GWAS conducted on many complex traits up to date, including suicidality (Table 1), has mainly focused on the identification of significant, single SNP-associations. The premise is that a subset of (usually higher frequency, "common") SNPs capture the overall genetic involvement in a phenotypic effect, an effect which deviates from the rest of the genome. But complex traits are believed to involve the interplay of many different genetic perturbations. For psychiatric disorders, results from single SNP-analyses have yet resolved only a subset of consistent findings (see e.g. (Bhat et al., 2012; Hamshere et al., 2013)) among the presumably thousands of SNP-associations of small effect that are hypothesized to be involved (Binder, 2012; Moore et al., 2011), and which has as indeed also been demonstrated for e.g. schizophrenia (Purcell et al., 2009). Consequently, poly-SNP associations represent a valuable evolving approach, giving insight about the overall heritability mediated by many thousands SNPs at once, enabling comparisons of genetic overlaps between traits and being more powerful than single-SNP associations (Dudbridge, 2013).

1.2. Gene-wide GWAS

For the case of genetic heterogeneity in a single gene, different random mutations can cause the same phenotypic (biological) effect, as simplistically exemplified by e.g. the mendelian phenylketonuria disorder (Scriver, 2007). Genetic heterogeneity may cause each SNP per se to have a lower power to be detected in a classical SNP-by-SNP GWAS as the relative importance of each SNP is diluted. Different strategies are continuously being published about how to best compile a gene-wide association signals from multiple SNP-signals. No golden standard exists in this domain, but the general procedure is to select the most significant SNP located within, or in LD with a gene, or to combine the signal from many different SNPs (Lehne et al., 2011; Li et al., 2012a). By doing so, the focus is shifted to association of known genes with the suicidality phenotype in question, facilitating biological interpretations and reducing the multiple comparisons burden (down to 20–30,000 tests).

1.3. Pathway / network GWAS

In the case of complex polygenetic traits such as suicidality phenotypes, genetic heterogeneity can become problematic also at the gene-level. Namely, if different genes can drive the same phenotypic effect (pleiotrophy), or if multiple genes, each of small marginal effects (but with large, compounded epistatic effects) underlie the phenotype, then also each gene-wide effect will be small *per se*. Therefore, novel methods are being developed and applied to (re)analyze the associations from a SNP-GWAS in relation to (poly-)gene groupings based on prior knowledge, e.g. pathways (Holmans, 2010) and/or protein-protein interaction (PPI)