Schizophrenia Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Schizophrenia Research

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

Genetic analysis of deep phenotyping projects in common disorders

Elliot S. Gershon [a,*], Godfrey Pearlson [b,n], Matcheri S. Keshavan [c], Carol Tamminga [d], Brett Clementz [e], Peter F. Buckley [f], Ney Alliey-Rodriguez [g], Chunyu Liu [h], John A. Sweeney [d,i], Sarah Keedy [g], Shashwath A. Meda [b], Neeraj Tandon [j], Rebecca Shafee [k,m], Jeffrey R. Bishop [l], Elena I. Ivleva [d]

[a] Department of Psychiatry, Department of Human Genetics, University of Chicago, United States
[b] Yale University Departments of Psychiatry & Neuroscience, Hartford, CT, United States
[c] Harvard Medical School, Department of Psychiatry, Boston, MA, United States
[d] Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, United States
[e] Department of Psychology, University of Georgia, Athens, GA, United States
[f] School of Medicine Virginia Commonwealth University (VCU), Richmond, VA, United States
[g] University of Chicago, Department of Psychiatry and Behavioral Neurosciences, Chicago, IL, United States
[h] University of Illinois at Chicago, Chicago, IL, United States
[i] University of Cincinnati, Department of Psychiatry and Behavioral Neuroscience, Cincinnati, OH, United States
[j] Beth Israel Deaconess Medical Center, Dept of Psychiatry, Harvard Medical School, United States
[k] Broad Institute of MIT and Harvard, Cambridge, MA, United States
[l] Department of Clinical and Experimental Pharmacology, University of Minnesota, Minneapolis, MN, United States
[m] Department of Genetics, Harvard Medical School, United States
[n] Olin Neuropsychiatry Research Center, Institute of Living, Hartford, Connecticut, USA

ARTICLE INFO

Article history:
Received 2 July 2017
Received in revised form 19 September 2017
Accepted 22 September 2017
Available online xxxx

Keywords:
Genetic analysis
Phenotype
Multiple testing
Functional genomics
Imputation

ABSTRACT

Several studies of complex psychotic disorders with large numbers of neurobiological phenotypes are currently under way, in living patients and controls, and on assemblies of brain specimens. Genetic analyses of such data typically present challenges, because of the choice of underlying hypotheses on genetic architecture of the studied disorders and phenotypes, large numbers of phenotypes, the appropriate multiple testing corrections, limited numbers of subjects, imputations required on missing phenotypes and genotypes, and the cross-disciplinary nature of the phenotype measures. Advances in genotype and phenotype imputation, and in genome-wide association (GWAS) methods, are useful in dealing with these challenges. As compared with the more traditional single-trait analyses, deep phenotyping with simultaneous genome-wide analyses serves as a discovery tool for previously unsuspected relationships of phenotypic traits with each other, and with specific molecular involvements.

The major medical successes of genetic association have been with patient-control comparisons, initially in single-gene diseases and more recently in common (multifactorial/polygenic) diseases and syndromes. In these analyses, disease is often used as a categorical phenotype, present or absent, ignoring the underlying clinical, genetic, and biological complexity of many medical and psychiatric disorders. Recently there have been successful associations with continuous complex phenotypes, such as height, and educational attainment (Rietveld et al., 2013; Wood et al., 2014), and with components of disease, such as blood pressure (Liu et al., 2016; Surendran et al., 2016). In psychiatric disorders, systematic studies of disease aimed at component phenotypes and their biological basis, based on diverse phenotypic measurements, may offer special promise for illuminating their genetics. This view has led to the development of several studies of multiple phenotypes in which many clinical, behavioral, neurophysiological, and neuroanatomic phenotypes, as well as genotypes, are assessed for each studied person in large samples of individuals. We refer to these as deep phenotyping studies; the NIMH-supported Bipolar- Schizophrenia Network on Intermediate Phenotypes (B-SNIP), which has generated this paper, is one such project. The genetic analysis of a deep phenotyping study presents its own challenges, because of the large number of tests, limited numbers of subjects, imputations of phenotypes and genotypes, methodological diversity of phenotype measures, and because of the cross-disciplinary nature and multiple collaborators involved in such studies. The diverse approaches to investigating genotype- phenotype relationships can lead to a confusing array of findings which may often appear contradictory; a comprehensive

* Corresponding author at: 5841 S. Maryland Ave. MC3077, Chicago, IL 60637, United States.
E-mail address: egershon@yoda.bsd.uchicago.edu (E.S. Gershon).

https://doi.org/10.1016/j.schres.2017.09.031
0920-9964/

perspective on the advantages, limitations and tradeoffs involved in the different approaches can provide a clearer perspective for the field.

1. Genetic architecture of common traits and diseases

Quantitative neurobiological traits related to common neuropsychiatric diseases have become of particular interest since the publications on Research Domain Criteria (RDoCs) (Insel et al., 2010; Insel and Cuthbert, 2009), which are expansions of the endophenotype concept that had been proposed decades earlier by Gottesman (Gottesman and Gould, 2003; Gottesman and Shields, 1973; Gottesman and Shields, 1972). Biological markers, phenotypes, and underlying neurobiological functions related to disease are conceptualized as continuous "domains" that are commonly show varying degrees of abnormality in patients. The implicit theory on the genetic architecture of disease is that there are multiple genetic variants that are correlated with trait markers, and with the right genetic combination the trait markers' quantitative value crosses a threshold for genetic liability to disease. An alternative architecture would be when, the *endo*- and/or clinical phenotype appears continuous, but in ill people different genes are operating to produce extreme values, or illness overrides the genes that influence the trait in well people. Yet another complexity of genetic architecture would be when single rare genetic variants can by themselves produce very substantial risk of disease and trait abnormality.

With the development of genetic technologies and analytic methods in the past decade, these and other hypotheses on the roles of genes in trait biomarkers related to disease can be tested directly. In this paper, we discuss genetic analysis of traits that may underlie psychosis syndromes, and detection of genetic architectures of sub-phenotypic traits. We pay particular attention to comprehensive analyses of a wide range of neurobiological phenotypes, and of genomic events associated with these phenotypes (variations in genotype, gene expression, or epigenomic measures), and their relation to individual case/control status.

Among the types of genetic architectures of illness and neurobiological traits, as discussed in the preceding paragraph, there are single gene disorders/traits, quantitative traits with threshold for disease, different genetic events in patients vs. controls, and different genetic events in different ethnic backgrounds. For illness and other traits, a general principle was elucidated by Manolio et al. (Manolio et al., 2009) (Fig. 1).

With the exception of rare variants with a large effect size (on disease risk), a polygenic model of common diseases, at the lower right of Fig. 1, with multiple common genes of weak effect, fits schizophrenia and bipolar disorder, and may be applicable to some marker phenotypes we discuss in this paper.

The term polygenic originated decades before there was a genetic map, to refer to genetic influences that are each too small to identify, but which can have a net cumulative influence on a phenotype. Polygenic inheritance with normally distributed genetic liability was first proposed by Sewall Wright (Wright, 1934), to account for inherited discrete phenotypes. The first applications were to inheritance of the number of digits in guinea pig paws, which could take one of two values. This type of inheritance was recently depicted graphically by Felsenstein (Felsenstein, 2005) (Fig. 2). By inspection, the graphic makes it clear that the phenotype as well as the genotype can be continuous, with a threshold for a binary trait (such as illness). For genetic association analysis, there is a statistical appeal in phenotypically continuous traits, because there would be enhanced statistical power to detect genetic associations as compared with the same trait analyzed as a binary phenotype. For many common disorders, including Type 2 diabetes, schizophrenia, and bipolar disorder, there are sub-threshold diagnostic states found in persons at increased risk of illness, and in some family members of patients, suggesting that a quantitative trait analysis of some phenotypes across patients and controls would be appropriate.

At the present time, when there is a dense map of human genetic markers, it is possible to use genotypes to directly assess polygenic risk to a phenotype. In the psychiatric diseases, there is a mixture of common variants with low effect, which can be summarized in a polygenic risk score (PGRS) (Purcell et al., 2009) which has a larger effect than any single variant, but is only currently applicable in Caucasian ancestry persons. The same study showed the schizophrenia PGRS to be applicable to bipolar disorder but not to several medical disorders. The schizophrenia PGRS has also been associated with cognitive function in healthy individuals but not in patients with a history of psychosis (Shafee et al., 2017), suggesting that disease related factors may in some circumstances overwhelm the genetic influences typically influencing a trait in the general population [cf. (Hochberger et al., in press)]. A mirror image example is found in genetic effects on intellectual ability and disability. The normal range of intelligence is polygenic (Sniekers et al., 2017), but intellectual disability (IQ < 50) is overwhelmingly associated with single rare variants (Gilissen et al., 2014).

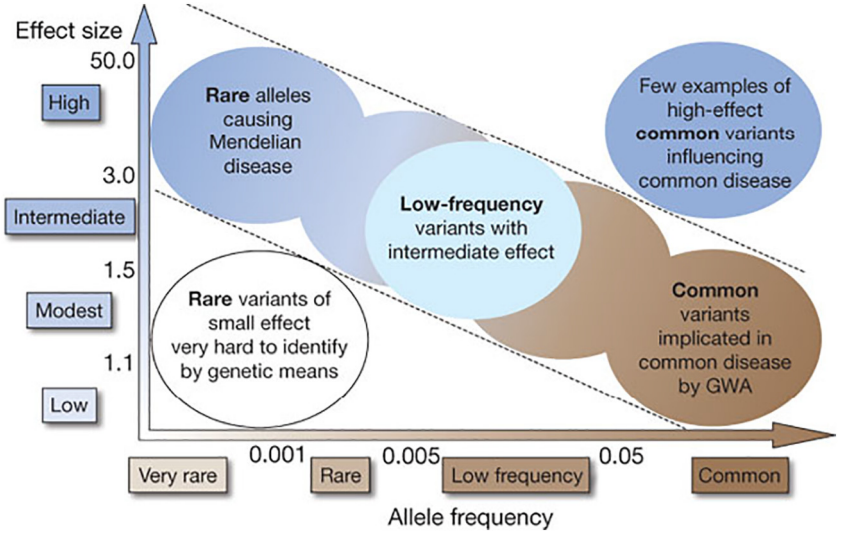

Fig. 1. Feasibility of identifying genetic variants by risk allele frequency and strength of genetic effect (odds ratio). From Manolio (Manolio et al., 2009).

Please cite this article as: Gershon, E.S., et al., Genetic analysis of deep phenotyping projects in common disorders, Schizophr. Res. (2017), https://doi.org/10.1016/j.schres.2017.09.031