Editorials

The contribution of epidemiology to defining the most appropriate approach to genetic research on schizophrenia

SARAH TOSATO and ANTONIO LASALVIA

Abstract. Psychosis is thought to have a strong genetic component, but many efforts to discover the underlying putative schizophrenia genes have yielded disappointing results. In fact, no strong associations emerged in the first genome-wide association studies in psychiatry and weakly observed associations were not related to the candidate genes identified in previous studies. These partially successful findings may be explained by the fact that genetic research in psychiatry suffers from confounding issues related to phenotype definition, the considerable degree of phenotypic variability and diagnostic uncertainty, absence of specific neuropathological features and environmental influences. To make progress it is first necessary to deconstruct psychosis based on symptomatology, and then to correlate particular phenotypes with genetic variants. Moreover, it is time to conduct studies that define persistent aspects of the schizophrenic profile that are more likely to represent an underlying biological pathogenesis, as opposed to fluctuating symptoms that are possibly environmentally mediated. In fact, progress in understanding the etiology of schizophrenia will depend upon the availability of good measures of genetic liability as well as relevant environmental exposures during critical periods of an individual’s life. If environmental and/or genetic factors are not precisely measured, it is impossible to study their independent effects or interactions.

Declaration of Interest: None.

Address for correspondence: Dr. S. Tosato, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, Policlinico G.B. Rossi, P.le L.A. Scuro 10, 37134 Verona (Italy).
Fax: +39-045-8027498
E-mail: sarah.tosato@univr.it

INTRODUCTION

Twin and adoption studies (Cardno et al., 1999) have established that schizophrenia has a strong genetic component. Schizophrenia is a complex genetic disorder consisting of multiple genes which exert small effects. Although the heritability of schizophrenia is approximately 80% (Cardno & Gottesman, 2000; Sullivan et al., 2003) and many efforts to discover the underlying putative schizophrenia genes have been made, no specific susceptibility gene has been clearly identified.

During the last two decades, two complementary approaches have been used in the search for susceptibility genes: linkage studies, which do not rely on specific biological hypotheses but seek to identify chromosomal regions containing susceptibility loci; and association studies, which are sensitive enough to detect small gene effects, but have to rely on plausible candidate genes (Norton et al., 2006; Owen et al., 2007). More recently, two major developments contributed to the transformation of prospects for searching for genes involved in complex trait disorders:

1) the HapMap project (http://www.hapmap.org/) has permitted a better knowledge of the patterns of the genome sequence landscape and variation in human populations, and
2) the rapid advances in genotyping capabilities makes it possible to assay affordably most genome sequence variations attributable to common single nucleotide polymorphisms (SNPs) and copy number variants (CNVs) (St Clair, 2009; Stefansson et al., 2008).

In spite of these advanced techniques, the results are often disappointing. In fact, the first genome-wide association studies in schizophrenia and bipolar disorder produced no strong findings and weakly observed
associations were not related to candidate genes identified in previous studies. Moreover, no consistent results were replicated by different studies (Crow, 2008). Until now a thousand association studies involving 700 candidate genes (Allen et al., 2008) supported the role of some genes, including Neuregulin 1 (NRG1), dysbindin (DTNBP1), dopamine receptors D1-4 (DRD1-4) and Disrupted-in-Schizophrenia-1 (DISC1) in inducing vulnerability to schizophrenia. However, even for these “promising genes”, there is a remarkable failure to replicate exactly the same markers and haplotypes across studies and a lack of consistency in implicating particular alleles in the development of schizophrenia (Alkelai et al., 2008; Sanders et al., 2008; Sullivan, 2008). Consequently, it appears that the identification of polymorphisms for psychiatric disorders is more difficult than for other complex trait disorders such as cardiovascular diseases, diabetes or cancer (Welcome Trust Case Control Consortium, 2007).

In spite of this, the advantages in genome scanning technology and the discovery of CNVs with autism and schizophrenia (Walsh et al., 2008; Weiss et al., 2008) have raised a number of fascinating new clinical questions concerning the phenotypic boundaries among major neuropsychiatric disorders, the genetic and environmental factors that influence phenotype and the relationship between behavior and genomic evolution.

THE CLINICAL HETEROGENEITY OF THE “PHENOTYPE” DEFINITION

The inconsistent results and disappointing findings of genetic research into schizophrenia may be due to the intrinsic characteristics of the phenotype under investigation: the “diagnosis of schizophrenia”. Its manifest clinical heterogeneity, combined with an association of CNVs, has led to the conceptualization of schizophrenia as a heterogeneous disorder. Many efforts have been made to link levels of heterogeneity, particularly between pathophysiology and phenomenology. However, the extent to which the clinical, pathophysiological and etiological components are interrelated is still largely unknown. Progress in understanding the disorder may have been hampered by the heterogeneous groups – at clinical, etiological and pathological levels – of patients whom researchers have studied under the name of “schizophrenia” (Peralta & Cuesta, 2003).

The heterogeneity issue in schizophrenia is reflected by the term ‘schizophrenias’ coined by Bleuler (1911) to describe the complex phenomenological picture at the clinical level. Symptoms of schizophrenia affect multiple psychological domains, including perception, inferential thinking, language, attention, social interaction, emotional expression, and volition. While the description of symptoms and signs of schizophrenia has remained mainly unchanged over the years, the way in which authors have articulated the varied phenomenological manifestations has been very unequal, thus rendering different views of schizophrenia across periods and countries. During the last forty years, psychopathologists have tried to simplify this complex array of symptoms and signs in several ways. Firstly, the concept of positive and negative symptoms has been developed in an attempt to integrate the various aspects (symptoms, pathophysiology and outcome) of schizophrenia (Strauss et al., 1974; Andreasen & Olsen, 1982). Subsequently, a 3-factor model was proposed by Liddle (1987), including positive, negative and disorganised dimensions, which has since been confirmed through several factor analytic studies (Lenzenweger et al., 1991; Malla et al., 1993; Andreasen et al., 1994). In recent years more complex multidimensional models have been reported: a four-syndrome model based on the dimensions of psychosis, disorganization, negative and social dysfunction has been proposed (Peralta et al., 1994; Dollfus & Everitt, 1998). Kay & Sevy (1990), suggest seven dimensions, although the consensus at this point is that just five factors, positive, negative, depressive, disorganization, and excitement, are the ones that best represent the scale’s factor structure (Lindenmayer et al., 1994). In conclusion, there is still little agreement about the number of dimensions necessary for an adequate representation of schizophrenic psychopathology, and views remain largely dependent on the rating scale employed (Peralta & Cuesta, 2000).

THE TRAJECTORIES OF SCHIZOPHRENIA: THE HETEROGENEITY OF OUTCOME

Over the last decades research has consistently found that, contrary to kraepelinian dogma, schizophrenia shows considerable heterogeneity in both course and outcome (Hegarty et al., 1994; Davidson & McGlashan, 1997). A recent review on the long-term outcome of schizophrenia found that between 21% and 57% of patients have a good outcome, depending on the outcome dimension and the strictness of the diagnostic criteria used (Jobe & Harrow, 2005). Some of this heterogeneity...
may be due to patient-related factors, such as age of onset, acuteness of onset, and to other factors linked to the severity of the disorder, such as the duration of untreated psychosis, cognition and early treatment response (Emsley et al., 2008).

Long-term follow-up studies have also established heterogeneity in the levels of symptoms and functioning within individuals with schizophrenia as well as across individuals. The earlier work of Strauss & Carpenter (1977) demonstrates the “loosely linked” nature of the relationships between different outcome dimensions. Building on this work, later research has consistently found that levels of psychopathology, specifically positive symptoms and social functioning, show only weak relationships to each other (Lasalvia et al., 2007a). Indeed, the multiple outcome dimensions in schizophrenia represent an “open-linked system” of outcome, since they all show relatively independent trajectories.

The negative symptom cluster, once established, is more stable over time and is more likely to be associated with neurocognitive impairments (Harvey et al., 2006), brain abnormalities (Shenton et al., 2001) and work and social incapacity (Lysaker & Bell, 1995; Bowie et al., 2006). Positive symptoms have consistently been found to show an independent pattern of evolution over the course of the disorder (Eaton et al., 1995), with a tendency to a reduction in severity levels over time (Lasalvia et al., 2007a). On the other hand, once established, cognitive deficits, similarly to negative symptoms, are relatively stable over time (Hyde et al., 1994; Mockler et al., 1997; Hijman et al., 2003); in addition they are associated with incapacity to work (Dickinson et al., 2007; Bowie et al., 2008).

The marked heterogeneity in the outcomes of schizophrenia depends on which specific perception of outcome by treating clinicians or patients themselves is considered, and the stringency of defining a good or poor outcome. For instance, in a sample of long-term patients treated in the South-Verona Community-based Mental Health Service, it was found that overall psychopathology (clinician-rated) tend to worsen over time with a clear-cut deterioration in negative symptoms, whereas subjective quality of life showed no significant change over the same period (Ruggeri et al., 2004). With regards to clinician-assessed social functioning, a trend towards deterioration was found. The overall number of patient-rated needs for care showed a significant decrease in social and health domains and an increase in functioning needs. These findings lead us to the view that there may be partially overlapping but distinct domains that can be identified as legitimate outcomes for schizophrenia. Such different domains may not vary directly. They may be influenced at least partially by separate predictors that may reveal different rates of poor and good outcomes depending upon which we accord primacy. We therefore suggest that, for both mental health care provision and research, staff-rated and patient-rated outcome measures are not interchangeable, but should be separately considered. This opens up lines of scientific enquiry to investigate the heterogeneity of outcomes when measured across multiple dimensions and rated from different perspectives (Ruggeri et al., 2005; Lasalvia et al., 2007b).

**CROSSING THE DIAGNOSTIC BOUNDARIES**

Research in the field of genetic psychiatry suffers from confounding issues related to phenotype definition (Craddock et al., 2008), considerable phenotypic variability and diagnostic uncertainty, absence of specific neuropathological features or biomarkers and environmental influences (Kennedy et al., 2003).

One of the main issues is related to the validity of the construct of schizophrenia. Over the decade, the DSM-IV definition of schizophrenia has been the most influential classification in clinical practice and research. Its clear criterion-based definition facilitates diagnostic agreement and communication among practitioners and researchers. However, although the DSM definition of schizophrenia has an undoubtedly high clinical utility, it does not provide information about the fundamental nature of schizophrenia: it does not answer the basic taxonomic question: “Are the correlations of observed clinical characteristics corroborative of underlying latent phenotypic dimensions (continuous distributions), latent categories (composed of one or more class or sub-disorder, each with its own phenotypic presentation) or a mix of the two?” (Meehl, 1995). If our definition of schizophrenia does not represent a “real” construct in nature, then it will not delineate the true pathology and causal mechanisms underlying psychosis; it will obfuscate etiology (Allardyce et al., 2007).

Most of the genetic research into psychoses has been based on the “given” descriptive diagnostic categories of schizophrenia and bipolar disorder, notwithstanding the fact that their validity has been challenged by emerging data from many fields of psychiatric research (Craddock et al., 2006; Murray et al., 2004).

Recent findings provide evidence for an overlap in genetic susceptibility across traditional psychosis categories, which are entirely based on the presence of clinical symptoms with several dimensions. The clearest evidence

---

Epidemiologia e Psichiatria Sociale, 18, 2, 2009

83
was found in the familial co-aggregation of both schizophrenia and bipolar disorder (Baron et al., 1982; Gershon et al., 1988; Maier et al., 1993; 2005; Mortensen et al., 2003). In fact, the monozygotic (MZ) co-twins of probands with schizophrenia had increased chances of mania (8.2%) as well as schizophrenia (40.8%), while the MZ co-twins of manic probands had an increased risk of schizophrenia (13.6%) as well as mania (36.4%); the MZ co-twins of schizoaffective probands had identically increased rates (26.1%) of schizophrenia and mania (Cardno et al., 2002).

Several new reports on candidate genes implicate variations at the same loci that influence susceptibility to both schizophrenia and bipolar disorder (Craddock et al., 2006). In particular, the most convincing candidate gene for schizophrenia, NRG1 (Munafo et al., 2006; Tosato et al., 2005) has been associated with a clinical phenotype of bipolar disorder with mood-incongruent psychotic symptoms (Green et al., 2005). Moreover, there is quite impressive evidence supporting the association between DTNBP1 and schizophrenia in Caucasian, Chinese and Japanese populations (Straub et al., 2002; Schwab et al., 2003; Tang et al., 2003; Van den Oord et al., 2003; Funke et al., 2004; Numakawa et al., 2004; Williams et al., 2005; Li et al., 2005). Few studies have investigated the role of DTNBP1 in bipolar disorder (Breen et al., 2006; Kohn et al., 2004) and it was demonstrated that DTNBP1 could only be involved in cases of bipolar disorder with psychotic features (Raybould et al., 2005). These findings suggest that NRG1 and DTNBP1 may confer susceptibility to a specific clinical phenotype with combined features of psychosis and mania (Craddock et al., 2006; Ileeva et al., 2008). Similarly, it was found that variations in the most promising candidate gene for bipolar disorder, D-amino acid oxidase activator (DAOA) (Detera-Wadleigh & McMahon, 2006), may influence susceptibility to episodes of mood disorder across the traditional definitions of bipolar and schizophrenia (Owen et al., 2007).

Therefore, twin and association studies suggest that psychosis may be conceptualized as a clinical phenotype where hypothetical genes or sets of genes, interacting with environmental factors, may be responsible for vulnerability to psychosis. Depending on additional syndrome-specific genetic determinants and environmental influences, psychosis may coexist with other clinical characteristics; for example, psychosis may be present with affective symptoms or cognitive dysfunction, comprising categorical diagnoses (Ileeva et al., 2008). Similarly, it is possible that while some environmental exposures may play a role in a subgroup of people with psychosis, other forms of psychosis may be due to other risk factors such as genetic ones.

**DISENTANGLING THE ENVIRONMENTAL CONTRIBUTION**

If phenotypic heterogeneity has been invoked as the main reason for the inconsistencies of genetics findings in schizophrenia, another important issue is to establish the role of the environment. It is still unclear what specific roles the environment plays in the pathway which leads to the disorder. Two kinds of epidemiological findings suggest that genes and environment contribute interactively in producing psychosis: the geographic, ethnic and demographic variation in the incidence of schizophrenia (McGrath et al., 2004; Kirkbride et al., 2006) and the marked variability in an individual’s response to the same environmental factor risks.

A recent population-based first-episode study conducted in the UK demonstrated considerable heterogeneity in incidence rates of schizophrenia and other psychoses in terms of sex, age, ethnic group and place of birth. This confirms that environmental effects at the individual, and perhaps neighbourhood level, may interact with genetic factors in the etiology of psychosis (Kirkbride et al., 2006). Moreover, a recent systematic review of studies published in the international literature on the epidemiology of schizophrenia over the last 40 years found a prominent variation in the incidence of schizophrenia between countries in terms of sex, age, ethnic group (McGrath, 2008). This variation is an important tool for understanding and investigating the causes of psychosis. An apparent lack of geographical variation had led to an emphasis on genetic factors, whereas heterogeneity would support environmental causes that most likely interact with the genome.

The higher or lower likelihood of developing schizophrenia in response to a given environmental insult reveals the etiological contribution of gene-environment interactions (GxE). This model postulates that a causal role cannot be found for either genes or environment alone, but in their synergic co-participation. There are a number of environmental exposures that are associated with psychotic disorders and for which a GxE mechanism has been proposed. These include both biological factors, such as malnutrition (Penner & Brown, 2007), obstetric complications (Cannon et al., 2002; Geddes et al., 1999), paternal age (Malaspina et al., 2001; Zammit et al., 2003), maternal infections (Brown et al., 2001) and psychosocial factors such as childhood trauma (Read et al., 2005), stressful life events (Harrison, 2004), migration (Boydell et al., 2001; Cooper et al., 2008), urbanicity (Van Os et al., 2003; 2004) and cannabis use (Henquet et al., 2008).

*Epidemiologia e Psichiatria Sociale, 18, 2, 2009*
Notably, each of these putative causes is actually a class of potential causes. At face value, these data would seem to favour the hypothesis of etiological heterogeneity (Tsuang & Faraone, 1995).

However, Cardno & Farmer (1995) challenge the heterogeneity hypothesis: they argue that differences demonstrated phenotypically by multivariate statistical techniques are quantitative and only compatible with etiological homogeneity resulting from a threshold of disease burden. Variations in phenotype can be placed on this spectrum with non-paranoid, predominantly negative symptoms and/or familial forms of the disorder causing a greater impact than paranoid, predominantly positive symptoms and/or sporadic forms. The problems are complex and far from fully clarified. Perhaps the heterogeneity debate should consider the possibility of rewording the question: ‘Heterogeneity: present or not?’ to ‘Heterogeneity: how much?’. This altered perspective opens up a range of heterogeneity models for empirical testing (Tsuang & Faraone, 1995). In this respect, for the bulk of schizophrenic disorders, etiological heterogeneity is very likely and almost certainly is a hypothesis worth testing (Tsuang & Faraone, 1995).

LINKING GENETICS TO PSYCHOPATHOLOGY AND ENVIRONMENT

Despite the difficulties in measuring and modelling environmental effects (Jones & Cannon, 1998), progress in understanding the etiology of schizophrenia will depend upon the availability of good measures of genetic accountability as well as of relevant environmental exposures.

Driven by clinical observations and research needs, an approach based on phenotypic dissection has emerged to overcome the difficulties that are inherent in research into multifactorial phenotypes (Rietkerk et al., 2008). This approach deconstructs schizophrenia and bipolar disorder into phenotypes based on symptoms, and then, correlates particular phenotypes with genetic variants (Jablensky, 2006). In fact, it has been suggested that the “schizophrenia genes” do not code for schizophrenia per se but for some broader clinical construct such as psychosis (Kendler et al., 1998; Weiser et al., 2005; Craddock & Owen, 2007). Therefore, it is necessary to consider the possibility that there may be no “one” variant of schizophrenia with a definable etiology, and to accept that there are multiple pathways towards psychosis. Perhaps, the category ‘schizophrenia’ may include several diseases whose clinical manifestations are similar.

The prospective of a dimensional, symptom-based approach focused on an individual and sub-syndromal phenotype is attractive because it may provide a model for studying the heterogeneity of schizophrenia, enhancing the ability to identify the underlying pathophysiology of the illness (Carpenter et al., 1993).

To date, relatively limited work has been done to identify genetic variants associated with specific clinical phenomena. Gene-symptom relationships have emerged primarily from follow-up studies of putative schizophrenia risk genes, with only handful of replicated findings (DeRosse et al., 2008). Significant relationships have been thoroughly reported between SNPs in DISC1 gene and the severity of delusions (Hennah et al., 2003; DeRosse et al., 2007) and SNPs in DTNBP1 and negative symptoms (Tosato et al., 2007; DeRosse et al., 2006a; Fanous et al., 2004). In addition, a number of studies have indicated that specific genetic variants may act to modify the clinical presentation of illness without increasing the overall risk for the illness itself. Emblematic is the example of COMT gene: although the evidence in favour of a role of the COMT Val158Met in predisposing to vulnerability to schizophrenia is still controversial (Glatt et al., 2003; Munafò et al., 2005, Allen et al., 2008), the polymorphism may be implicated in determining affective symptomatology in individuals suffering from schizophrenia (McClay et al., 2006; DeRosse et al., 2006b; Herken & Erdal, 2001). Indeed, associations have been reported between negative symptoms and variation in BDNF and DAT1 (Fanous et al., 2004) and between positive symptoms and variations in DRD4 (Serretti et al., 2001) and in DRD2 (Serretti et al., 2000).

The direct empirical evidence for a plausible biological mechanism linking environment to schizophrenia is, so far, very limited. The most notable exceptions are a new generation of birth cohort studies which prospectively assesses the impact of a given exposure on genotype over a long period of time; a brilliant example of this approach is represented by the Dunedin birth cohort study. In one of a series of papers from these researchers, it was found that individuals who smoked cannabis before the age of 15 years and who carry the COMT Valine allele had an increased risk of developing a schizophreniform disorder. They exhibit more psychotic symptoms than individuals carrying the Methionine allele (Caspì et al., 2005). This study, notable for its longitudinal design, also documents legitimate concerns about how to accurately assess the environmental risk exposure of participants: it is in fact difficult to measure the amount of active drug that is ingested in different forms with different tetrahydrocannabinol levels during recreational cannabis use over

Epidemiologia e Psichiatria Sociale, 18, 2, 2009
85
many years. In another study conducted within the framework of the Dunedin cohort it was found that individuals with one or two copies of the “short” serotonin transporter allele exhibited more depressive episodes and suicidality following adverse life events than individuals with two copies of the “long” allele (Caspi et al., 2003). This study shows how it is challenging to measure the frequency, timing and extent of the trauma caused by adverse psychosocial events whose negative effects may act cumulatively across long periods of an individual’s lifetime (Van Os et al., 2008; Caspi & Moffitt, 2006).

CONCLUSION

The difficulty in gaining a consistent and clear-cut picture of the genetics of schizophrenia mirrors the marked clinical and neurobiological heterogeneity of the disorder (Tamminga & Holcomb, 2005). A comprehensive global model to understand clinical heterogeneity in schizophrenia is still lacking. As long as we are not able to disentangle the question of heterogeneity at the clinical level, it is not likely that heterogeneity at the etiological and pathophysiological levels will be solved. Acknowledging that schizophrenia is neither an entity nor a unity, we should analyze its psychopathology through its different elements and search for the factors which determine each of them. It can therefore be expected that research on symptoms or longitudinally consistent psychopathological dimensions is better suited for etio-pathogenetic investigations than research based on diagnoses.

It will be necessary in the future to conduct studies that define persistent aspects of the schizophrenic profile which are more likely to represent an underlying biological pathogenesis as opposed to fluctuating, possibly environmentally mediated symptoms. The definition of a particular clinical phenotype based on symptoms, social functioning and prognosis may be achieved using a multiwave longitudinal design. The correlation of genetic risk factors with this new clinical phenotype will not only permit an assessment of the clinical heterogeneity of the disease, but will also improve the classification of mental disorders, and potentially enable the identification of useful biological markers. Moreover, the identification of genetic factors will be important for clinical prognostication and planning, treatment outcome and social adaptation (Rosenman et al., 2003; Allardyc et al., 2007). The identification of genetic factors will facilitate the search for independent environmental factors and enable investigation into the mechanism of interaction between genes and the environment. Again, the cohort approach may be a helpful design because it permits the collection of prospective longitudinal histories of participants’ environmental exposure, of the outcome of mental disorders and of the characterization of genotypes. In fact, sample size requirements can be reduced with high-quality measurements of environmental risk factors, especially when measurements are repeated over time (Wong et al., 2003).

In the next few years, the task of epidemiologists researching the genetics of schizophrenia will be:

1) the provision of convincing evidence that allows a better understanding of the mechanisms that underlie the hypothesized GxE interactions;
2) the development of models of control factors to rule out alternative explanations, reinforcing confidence in the GxE interactions;
3) to define whether such interactions account for a non-trivial proportion of the disorder in the human population (Caspi & Moffitt, 2006).

At the same time, progress in molecular genetics will produce a great amount of data about the genome. A multidisciplinary approach which carefully takes into account both environmental factors and genetics, and which is capable of producing hypothesis-driven research strategies on final biological pathways, may represent the most productive and fruitful approach.

REFERENCES


Epidemiologia e Psichiatria Sociale, 18, 2, 2009
The contribution of epidemiology to defining the most appropriate approach to genetic research on schizophrenia


Epidemiologia e Psicopatologia, 18, 2 2009


Epidemiologia e Psichiatria Sociale, 18, 2, 2009


Welcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661-678.


