Gene-Environment Interactions in Schizophrenia: Review of Epidemiological Findings and Future Directions

Jim van Os¹,²,³, Bart PF Rutten², and Richie Poulton⁴

¹Department of Psychiatry and Neuropsychology, School of Mental Health and Neuroscience, Maastricht University Medical Centre, EURON, SEARCH, PO Box 616 (location DOT 10), Maastricht, 6200 MD, The Netherlands; ²Division of Psychological Medicine, Institute of Psychiatry, London SE5 8AF, UK; ³Dunedin Multidisciplinary Health and Development Research Unit, Department of Preventive and Social Medicine, Dunedin School of Medicine, University of Otago, PO Box 913, Dunedin, New Zealand

Concern is building about high rates of schizophrenia in large cities, and among immigrants, cannabis users, and traumatized individuals, some of which likely reflects the causal influence of environmental exposures. This, in combination with very slow progress in the area of molecular genetics, has generated interest in more complicated models of schizophrenia etiology that explicitly posit gene-environment interactions (EU-GEI. European Network of Schizophrenia Networks for the Study of Gene Environment Interactions. Schizophrenia aetiology: do gene-environment interactions hold the key? [published online ahead of print April 25, 2008] Schizophr Res [254]; S0920-9964(08) 00170–9). Although findings of epidemiological gene-environment interaction (GxE) studies are suggestive of widespread gene-environment interactions in the etiology of schizophrenia, numerous challenges remain. For example, attempts to identify gene-environment interactions cannot be equated with molecular genetic studies with a few putative environmental variables “thrown in”: GxE is a multidisciplinary exercise involving epidemiology, psychology, psychiatry, neuroscience, neuroimaging, pharmacology, biostatistics, and genetics. Epidemiological GxE studies using indirect measures of genetic risk in genetically sensitive designs have the advantage that they are able to model the net, albeit nonspecific, genetic load. In studies using direct molecular measures of genetic variation, a hypothesis-driven approach postulating synergistic effects between genes and environment impacting on a final common pathway, such as “sensitization” of mesolimbic dopamine neurotransmission, while simplistic, may provide initial focus and protection against the numerous false-positive and false-negative results that these investigations engender. Experimental ecogenetic approaches with randomized assignment may help to overcome some of the limitations of observational studies and allow for the additional elucidation of underlying mechanisms using a combination of functional enviromics and functional genomics.

Key words: epidemiology/psychosis/schizophrenia/genetics/gene-environment interaction/gene-environment correlation

Introduction

Attempts to discover genes that relate directly to psychotic disorder (ie, the simple “main effects” approach) have been frustrating and often disappointing, resulting in expression of methodological concerns.⁵ On the other hand, epidemiological research has unveiled high observed rates of schizophrenia in large cities, immigrant populations, traumatized individuals, and cannabis users, at least some of which is thought to be the result of underlying environmental exposures. Exciting findings in other areas of psychiatry have motivated researchers to turn their attention to better understanding the complex ways in which nature interacts with nurture to produce psychosis. This genotype × environmental interaction (hereafter: GxE) approach differs from the linear gene-phenotype approach by positing a causal role not for either genes or environment in isolation but for their synergistic coparticipation in the cause of psychosis where the effect of one is conditional on the other.¹ For example, genes may moderate the psychotogenic effects of dopamine agonist drugs of abuse, or the environment may moderate the level of expression of a gene that is on the causal pathway to psychotic disorder. GxE seems a particularly suitable approach for understanding the development of psychosis because this phenotype is known to be associated with environmentally mediated risks,⁸,⁹ yet people display considerable heterogeneity in their response to those environmental exposures.

The structure of this article is as follows. First, the principles of genetic epidemiology as relevant for the study of...
gene-environment interaction will be reviewed briefly. Second, a brief overview will be given on what “the environment” may consist of in studies of G × E and how environmental mechanisms may be uncovered using “functional enviromics.” Third, the main G × E findings with regard to psychotic disorders will be reviewed, with a particular focus on epidemiological studies that used indirect measures of genetic risk including twin and adoption studies, family studies, and psychometric risk studies. Most of the findings using direct molecular genetic measures of genetic risk will be reviewed elsewhere in this issue. Fourth, considerations will be given to possible underlying mechanisms followed by a discussion of future research and directions.

Ecogenetics

Traditional epidemiology was concerned mainly with environmental risks. Conversely, genetic researchers of complex disorders have mostly focused on molecular genetic approaches in which the environment and interaction between genes and environment were treated as a power-reducing nuisance term. Awareness has been growing, however, that direct or indirect measures of genetic variation can be considered as a conventional epidemiological risk factor in association studies and that epidemiological theory can be readily applied to genetically sensitive datasets. Thus, epidemiologists and human geneticists have been gradually integrating their respective fields of research into a new discipline called genetic epidemiology. Within genetic epidemiology, the term ecogenetics refers to the study of specific gene-environment relationships. Within an ecogenetic framework, several types of gene-environment relationships are relevant for the study of complex disorders, representing different biologically plausible mechanisms by which genes and environment can influence disease outcome.

Ecogenetics in Psychiatry

Until recently, the conventional wisdom within psychiatry and behavioral genetics was that G × E was exceedingly rare and difficult to demonstrate. The revival of interest in G × E derives largely from (1) failures of direct gene-phenotype association studies to uncover genes related to susceptibility for psychiatric disorders and the realization that their multifactorial etiology likely includes many complicated interactive effects requiring more advanced approaches, (2) work demonstrating the operation of G × E in many other branches of medicine; and (3) recent evidence of G × E within psychiatry.

The recent G × E findings in psychiatry suggest that genes are likely to influence disorder mostly indirectly, via their impact upon physiological pathways, and work by increasing (or decreasing) the likelihood of developing a psychiatric disorder, rather than as direct causes of disorder per se. Thus, the notion of “a gene for ...” is misleading and diverts attention from more important issues. Further, some theorists now suggest that (1) additive, noninteractive genetic effects may be less common than previously assumed; (2) studying genes in isolation from known environmental risks may fail to detect important genetic influences; and (3) traditional notions of multiplicative interaction are probably not appropriate for “real-world” interactions, particularly given the ubiquity of some environmental exposures. Thus, biological synergism (coparticipation of causes to some outcome) between environmental exposure and background genetic vulnerability is thought to be common in multifactorial disorders such as psychosis. The classic problem, however, is how coparticipation between causes in nature (biological synergism) can be inferred from statistical manipulations with research data (statistical interaction), in particular with regard to the choice of additive (change in risk occurs by adding a quantity) or multiplicative (change in risk occurs by multiplying with a quantity) models. It has been shown that the true degree of biological synergism can be better estimated from—but is not the same as—the additive statistical interaction rather than the much more often used multiplicative interaction.

Genetic Moderation of Sensitivity to Environment. According to the concept of genetic moderation of sensitivity to the environment, differences in genetic endowment explain why people respond differently to the same environment (figure 1). Most evidence for this type of G × E in psychosis has come indirectly from twin and adoption studies and a variety of naturalistic designs in which nonspecific genetic contributions have been assessed. More recently, researchers have obtained information about how variation in specific measured genes interacts with specific measured environments. Genetic moderation of environmental sensitivity gives rise to synergism, or interaction, because the biological effects of G and E are dependent on each

![Figure 1: Gene × Environment Interaction: Genes Controlling Environmental Sensitivity.](image-url)
other in such a way that exposure to neither or either one alone does not result in the outcome in question, whereas exposure to both does. For example, a well-known example of gene-environment interaction is the observation that among Orientals, alcohol sensitivity is strongly regulated by genetic polymorphism of the aldehyde dehydrogenase (ALDH2) gene. Similarly, there is strong evidence that some polymorphisms may be involved in psychiatric disorders. For example, the gene encoding the serotonin transporter (5-HTT) contains a regulatory variation (5-HTTLPR), the short (“s”) allele of which is associated with lower transcriptional efficiency of the promoter compared with the long (“l”) allele. Data from animal and human research indicate that 5-HTTLPR may interact with environmental adversity to cause depression, reflecting underlying developmental mechanisms that affect the structural connectivity and, as a consequence, functional interactions, within a neural circuit involved in the regulation of emotional reactivity and extinction of fear.27,28,29,30,31

Although gene-environment synergism is likely prevalent, other models of disease causation, including models that imply that there is no synergism (synergism is zero), may also apply, although likely to a lesser degree. For example, an individual may get schizophrenia only if in possession of a certain type of vulnerability conferred by either genetic or environmental factors. An environmental factor could disrupt early brain development in the same fashion as a genetic mutation. In this model, synergism is zero, and the effect of genes and environment is said to be additive.

Environmental Impact on DNA Sequence and Methylation. Apart from genes impacting on sensitivity for environmental risk factors, G × E in psychotic disorder may also take the form of environmental factors impacting on either the DNA sequence (causing de novo mutations) or DNA methylation (causing altered gene expression through epimutations). The most suggestive epidemiological evidence for such mechanisms in psychosis comes from studies linking advanced paternal age to the risk of schizophrenia in the offspring.32,33,34,35 Paternal age varies as a function of the sociocultural environment.36 The observed paternal age effect on schizophrenia may consist of mutagenesis, causing de novo spontaneous mutations that would then propagate, and accumulate in successive generations of sperm-producing cells. Alternatively, the mechanism underlying the paternal age effect may be genomic imprinting.37 Genomic imprinting is the phenomenon whereby a small subset of all the genes in the genome is expressed according to their parent of origin. Some imprinted genes are expressed from a maternally inherited chromosome and silenced on the paternal chromosome, while other imprinted genes show the opposite expression pattern and are only expressed from a paternally inherited chromosome.38 One of the mechanisms for gene silencing is DNA methylation. The inherited methylation pattern is maintained in somatic cells but is erased and reestablished late in spermatogenesis for paternally imprinted genes, a process that could become impaired as age advances.

Although research on DNA methylation as an “epigenetic” mechanism underlying G × E in psychiatry is in an early phase, this field appears promising. For example, early maternal behavior in animals can affect offspring stress sensitivity through altered DNA methylation of key neuronal receptor genes involved in the stress response.39,40 Environmentally induced epigenetic mechanisms may explain a range of epidemiological findings including typical

---

**Fig. 2.** Promoter Activity of 5-HTT Gene Is Modified by Sequence Elements Within the Proximal Regulatory Region; the Short (“s”) Allele Is Associated With Lower Transcriptional Efficiency of the Promoter as Compared With the Long (“l”) Allele: Converging Evidence for Gene × Environment Interaction in Depression.
Gene-Environment Correlation

In contrast to G × E, gene-environment correlation (hereafter rGE) refers to how differences in an individual’s genotype can “drive” differential environmental exposure (figure 3). In rGE, exposure to environmental events is not a random phenomenon but rather stems (at least partly) from differences in genetic makeup. rGEs come in 3 main forms: passive rGE refers to environmental influences linked to genetic effects external to the person. For example, parents create the early child-rearing environment, as well as providing genetic material to their offspring. Passive rGE occurs when parental behavior, which is partly under genetic control, influences the nature of the early child-rearing environment. Thus, parental genes can exert an influence upon the child via the environment, but whose effects are independent of the child itself. In contrast, active rGE (eg, selection of specific environments or “niche picking”) and evocative rGE arise largely as a result of genetic factors nested within the individual. Evocative rGE refers to the impact of the child’s behavior on their social environment, in particular the responses they elicit from people around them. One person’s preference for sporting activities over another person’s penchant for artistic endeavors, thus selecting themselves into different environments, is an example of active rGE, while the different responses elicited from the social environment by gregarious vs shy individuals exemplifies evocative rGE. Combining examples of rGE and G × E in one illustrative situation: rGE might manifest as arguments and disagreements preceding marital dissolution, yet G × E may determine who becomes depressed as result of that relationship breakdown.

Confounding of G × E by rGE. In studies aimed at detecting G × E, rGE is noise and must be ruled out. In other words, the “E” in G × E must be shown to be a true-environmentally mediated effect rather then a genetic epiphenomenon. For example, does the genetic liability for schizophrenia increase the psychotogenic effect of cannabis or does schizophrenia genetic liability increase the likelihood of using cannabis? Experimental paradigms (see below) are able to deal effectively with this problem by randomly assigning participants to the exposed and unexposed conditions. In observational designs, however, confounding by rGE is difficult to rule out but can be tested separately. An interesting example concerns urbanicity and schizophrenia. As discussed below, 4 independent studies have suggested that the urban environment may contribute to the onset of psychotic disorder in individuals at genetic risk (ie, evidence for G × E). An alternative explanation, however, is that the genetic liability for schizophrenia increases the likelihood of moving to the big city, ie, there may be rGE. A priori this is unlikely, given the fact that the effect of urbanicity on schizophrenia is restricted to the window of childhood and adolescence: children do not make the family decision to move to the big city, regardless of whether they are genetically inclined to do so or not. Two twin studies from Australia and The Netherlands on urban mobility support this notion. The Australian study showed more evidence for influence of genetic factors on urban mobility than the Dutch study. However, genetic influence in the Australian study was mostly apparent in older individuals who were well past the age at risk for onset of schizophrenia; environmental factors accounted for most of the variation in younger individuals. The reason for the discrepancy in genetic contribution to urban mobility between the Australian and the Dutch study is likely related to contextual factors. Just as the heritability of alcoholism has been shown to differ as a function of societal availability (severe restriction resulting in alcohol use only by those who are genetically most predisposed), so was the genetic influence on urban mobility shown to vary as a function of base rate of the urban outcome that was only 10% in Australia vs around 30%–50% (very heavy and heavy urbanization) in The Netherlands. More evidence of genetic influence in Australia therefore may in part be the result of the lower base rate of urbanicity.
Thus, the conclusion from the Australian and Dutch twin studies is that there are likely only very few human characteristics beyond any genetic influence, including urban mobility. However, in young adulthood, the age range during which psychotic disorder typically declares itself, environmental more than genetic factors may influence exposure to the risk environment that urbanicity represents, making rGE unlikely.

Another important issue in rGE is that genetic effects on the outcome can be direct or indirect (figure 4). For example, genes may have an effect on both the outcome and the environmental exposure, while the environment has no effect on the outcome. In this case, the observed association between the environment and the outcome is genetically confounded (figure 4A). On the other hand, genes may have an effect on the environment, but no direct effect on the outcome because only the environment has a causal effect (figure 4B). This is the situation where the environment is on the causal pathway between genes and environment, a situation that can help in providing evidence for a true causal contribution of an environmental factor to disease (referred to sometimes as “Mendelian randomization”). For example, evidence in the situation of figure 4B of an association between the gene and the outcome can only be explained if there is a true causal relationship between the environmental risk factor and the outcome. Given random assortment of genes from parents to offspring during gamete formation and conception, gene-outcome associations representing gene-causal exposure associations are not generally susceptible to the reverse causation or confounding that may plague conventional observational studies.

The Environment, Experimental Ecogenetics, and Functional Enviromics

The Environment and Psychosis

Here, we refer to the environment broadly as all nongenetic influences that are associated with at least 2 exposure states. Sometimes a distinction is made between “biological” and “social” environmental exposures, but such a distinction may not be helpful as long as the underlying mechanisms, which are likely overlapping, are not elucidated. There are a number of environmental exposures that are associated with psychotic disorders and symptoms and for which a mechanism of gene-environment interaction has been proposed. These environmental exposures are summarized in box 1, together with an indication to what degree the evidence for an association with schizophrenia is supported by meta-analytic estimates from systematic reviews. The most solid evidence for an association with schizophrenia and related psychosis outcomes is for paternal age, migration, urbanicity, and cannabis use, the latter 2 particularly in the case of exposure during development.

Environmental Measurement and Experimental Ecogenetics

There are legitimate concerns about how to accurately capture the environmental risk exposure history of participants. This task is particularly challenging when measuring psychosocial risk factors whose negative effects may act cumulatively across long periods of the life course. Equally challenging are the inherent difficulties in precisely measuring “unit exposure” for illicit substances such as cannabis that can be ingested in different forms, with different tetrahydrocannabinol levels, using different methods. Measuring tobacco intake is comparatively straightforward, but even this presents problems with accuracy of recall over long periods.

Henquet et al have introduced the term of “experimental ecogenetics” in human psychosis research to refer to some obvious advantages: (1) randomization precludes confounding by not only known but, critically, also unknown confounders; (2) rGE is not an issue if “G” is randomly allocated to “E,” and (3) it is relatively easy to make the sample size match the required power. In figure 5, an example is given of how the association between migration and schizophrenia, and possible genetic moderation thereof, can be examined in the context of an experimental ecogenetic design by reducing migration to an experimental exposure of “social hostility” and the psychosis outcome to an experimental outcome of “abnormal salience attribution” and testing the association between exposure and outcome in a genetically sensitive test design. The advent of controlled experiments with virtual-reality environments may similarly represent an important asset for the study of environmental exposures.

A further issue is that the environment can be conceptualized at many levels that may all be relevant to behavioral phenotypes associated with schizophrenia, varying from minor stressors in the flow of daily life as assessed...
published environmental exposures for psychosis for which G × E has been suggested (M+: at least one positive meta-analytic estimate; M+/-: inconclusive meta-analytic estimate; M-: no meta-analytic estimate available)

Environmental variables with likely impact in fetal life:

1. M+: Maternal pregnancy complications, in particular fetal hypoxia and proxies for fetal folate deficiency
2. M+/-: Prenatal maternal infection, prenatal maternal stress, prenatal maternal folate deficiency
3. M+: Paternal age
4. M-: Prenatal exposure to chemical agents (eg, lead)

Environmental variables with likely impact in early life:

5. M-: Quality of early rearing environment (institutional care, school, parents)
6. M+/-: Childhood trauma (abuse or neglect)

Environmental variables with likely impact in middle childhood/adolescence:

7. M+: Urban environment during development: a variable indicating the level of population density, or size of a city within a country, of the place where the individual was growing up (between the ages of 5 and 15 years)
8. M+: Cannabis use
9. M+: Migration
10. M+/-: Stressful life events
11. M-: Traumatic brain injury

Measures of the wider social environment:

12. M-: Neighborhood measures of social fragmentation, social capital, and social deprivation

Measures of the microenvironment in the flow of daily life:

13. M-: Small daily life stressors, assessed using momentary assessment technology, subtly impacting on affect, salience, and reward

by momentary assessment technologies to contextual effects of the wider social environment such as neighborhood-type or ethnic density. Finally, some environmental risks such as “urbanicity” and “ethnicity” are proxies for as yet unidentified environmental or possibly even partly genetic factors.

Functional enviroimics

Functional envoironics, or the study of the mechanisms underlying environmental impact on the individual to increase the risk for psychopathology is still in its infancy, with many hypotheses yet to be tested. These include effects of the environment on (1) developmental programming and adult functional circuits of the brain, (2) neuroendocrine and neurotransmitter functioning, (3) patterns of interpersonal interactions that may shape risk for later psychopathology, and (4) affective and cognitive processing. Conversely, hypotheses need to be tested about the neural mechanism by which genetic variation may increase susceptibility to environmental stressors. These mechanisms and their underlying pathophysiological pathways need to be clarified in order to develop a priori gene-environment interaction research paradigms. For example, it has been suggested that there may be synergistic effects of genes and environment in bringing about a “sensitization” of mesolimbic dopamine neurotransmission. This hypothesis is supported by (1) evidence quantifying the impact of stress and dopamine agonist drugs on mesolimbic dopamine release and subsequent sensitization as well as stress-dopamine agonist cross-sensitisation; (2) evidence indicating that genetic risk for schizophrenia is associated with underlying alterations in the dopamine system, including increased dopamine synaptic availability, increased striatal dopamine synthesis, and increased dopamine reactivity to stress, and (3) human and animal evidence that effects of environmental risk factors associated with schizophrenia have lasting effects on dopamine neurotransmission including developmental trauma, defeat stress associated with ethnic minority group, prenatal hypoxia, and prenatal maternal immune activation.

Thus, although there is evidence to suggest that many other neurotransmitter systems can also be targeted, a case can be made, as an example of functional envoironics, for investigating genetic variation affecting dopamine neurotransmission in interaction with environmental risk factors such as stress and dopamine agonist drugs. Molecular genetic and functional genomic studies focusing on genes associated with dopamine neurotransmission suggest that this gene group may be useful for G × E studies. For example, a recent large study focusing on gene-gene interaction (epistasis) and functional effects suggested that a network of interacting dopaminergic polymorphisms may increase risk for schizophrenia. Evidence for epistasis between genes impacting on dopamine signaling can be validated using a neural systems-level intermediate phenotype approach in humans. Recent work of this type, using a prefrontal function fMRI phenotype, similarly suggests epistasis between polymorphisms in genes that control dopamine signaling. More specifically, there is evidence that schizophrenia may be characterized by a combination
of prefrontal cortical dysfunction and subcortical dopaminergic disinhibition. Research has shown that the valine allele carriers of a functional polymorphism in the catechol-O-methyltransferase gene (COMT Val158Met), an important enzyme regulating prefrontal dopamine turnover, predicted increased dopamine synthesis in the midbrain, suggesting that this allele may increase the risk for schizophrenia in interaction with, e.g., stress and dopamine agonist drugs. Several studies suggest that valine-allele carriers may indeed be more sensitive to the psychotogenic effects of drugs of abuse or stress.

There are examples of many other avenues that may be explored in functional enviromics. Thus, a recent systematic review suggested that more than 50% of genes potentially associated with schizophrenia, particularly AKT1, BDNF, CAPON, CCKAR, CHRNA7, CNR1, COMT, DNTBP1, GAD1, GRM3, IL10, MLC1, NOTCH4, NR4A2/NURR1, PRODH, RELN, RGS4, RTN4/NOGO, and TNF, are subject to regulation by hypoxia and/or are expressed in the vasculature. Thus, future studies of genes proposed as candidates for susceptibility to schizophrenia should include their possible regulation by physiological or pathological hypoxia during development as well as their potential role in gene-environment interactions involving events inducing hypoxia during early development.

**Epidemiological G × E Studies in Psychosis**

**Epidemiological Findings**

Two robust epidemiological findings suggest that “genes” and “environments” operate interactively to produce schizophrenia. First, there is widespread geographic, temporal, ethnic, and other demographic variation in the incidence of schizophrenia, reinforcing the etiological role played by environmental factors. Second, there is marked variability in people’s responses to these environmental risk factors, ranging from obvious vulnerability to extreme resilience. This well-recognized heterogeneity in response points to the operation of G × E. A number of studies have examined G × E using indirect measures of genetic risk, such as being a relative, a twin or adopted away offspring of a person with schizophrenia, or the level of psychometric psychosis proneness in a person as an expression of distributed genetic risk for psychotic disorder (see below). The advantage of these studies is that the measure of genetic risk, while nonspecific and therefore not able to capture gene-environment interactions with very specific mechanisms, is nevertheless (1) well validated and (2) represents the complete net genetic load including all gene-gene interactions. While newer studies using direct molecular genetic measures of genetic risk have the advantage of using specific measures, they are also prone to false-positive findings, given the enormous amount of molecular genetic variation that can be used for G × E modeling, and the absence of all other factors influencing genetic risk in the model of G × E using a small contribution to genetic variation in the form of a single-nucleotide polymorphism (SNP). Therefore, epidemiological studies using indirect measures of genetic risk remain useful and may point the way to G × E studies using direct measures of genetic risk; to date, they remain the most informative. A review of these findings is presented here.

**Findings From Twin, Adoption, and Family Studies**

Twin and adoption studies provide strong but nonspecific evidence for the involvement of both genes and environmental factors in the etiology of schizophrenia. Both have shown moderate to high heritability for
were similar (8.4% and 8.9%, respectively). Findings from several adoption studies are consistent with G × E in the development of psychotic disorders. For example, Carter et al. compared, in a 25-year longitudinal study, 212 children of schizophrenic mothers with 99 children of normal parents in terms of exposure to environmental risk (ie, institutional care and family instability). Very few cases of psychosis were identified in those families without a history of schizophrenia but, among those with a family history, strong environmental effects were observed. Consistent with this, Tienari et al. compared adopted-away offspring (N = 145) of mothers with a history of psychotic illness vs those without illness (N = 158). Measures of the rearing environment in the adoptive home were obtained (measures on scales of “critical/conflictual,” “constricted,” and “boundary problems”) and revealed strong effects for those with a biological predisposition (odds ratio around 10) that were absent in those with low genetic risk (odds ratio around 1).

Findings in support of G × E also come from migration designs which, eg, have demonstrated a higher risk of psychosis among Caribbean immigrants to the United Kingdom compared with the majority population in the United Kingdom. Further, family studies of UK-born Afro-Caribbeans have demonstrated a particularly high risk of schizophrenia among the siblings of young, Afro-Caribbean patients (15.9% compared with 1.8% in siblings of white patients), whereas the rates of schizophrenia among the white and Afro-Caribbean parents were similar (8.4% and 8.9%, respectively).

**Studies Using a Psychometric Psychosis Liability Approach**

Subtle subclinical expression of psychosis can be measured in the general population. There is evidence that this phenotype of “psychometric psychosis proneness” represents in part the distributed genetic risk for psychotic disorder, suggesting that it could be used as a proxy to represent the factor “G” in studies of G × E, although to the degree that environmental factors contribute to the psychometric psychosis proneness measure these cannot be excluded as a source of confounding. Thus, Vollema et al. reported that scores on the positive dimension of a schizotypy questionnaire administered to relatives of patients with psychotic disorders corresponded to their genetic risk of psychosis. Fanous et al. demonstrated that interview-based positive and negative symptoms in schizophrenia predicted their equivalent subclinical symptom dimensions in nonpsychotic relatives, implying an etiological continuum between the subclinical and the clinical psychosis phenotypes. Kendler and Hewitt studied twins from the general population and concluded that the variance in most self-report schizotypy scales, except for perceptual aberration, involved substantial genetic contributions. MacDonald et al. found in their general population-based twin study only one common schizotypy factor, mainly explained by perceptual aberration, magical ideation, schizotypal cognitions, and to a lesser extent social anhedonia. The common schizotypy factor was influenced by shared environmental, nonshared environmental, and possibly genetic effects. Recently, a general population female twin study by Linney et al. showed that additive genetic and unique environmental effects influenced self-reported psychotic experiences. The multivariate structural equation model generated 2 independent latent factors, namely a positive (ie, cognitive disorganization, unusual experiences, and delusional ideation) and a negative dimension (ie, cognitive disorganization and introvertive anhedonia), suggesting different etiological mechanisms for the various scales of the subclinical psychosis phenotype. In a recent, general population study using both self-report and interview-based measures of positive and negative dimensions of psychotic experiences in 257 subjects belonging to 82 families, significant family-specific variation for both positive and negative subclinical psychosis dimensions were demonstrated, with between-family proportions of total variance between 10% and 40%. Thus, both the positive and the negative dimensions of subclinical psychosis show familial clustering in samples unsel ected for psychiatric disease. Operationalizing the genetic effect “G” along these lines, Henquet et al. showed that a psychometric measure of psychosis proneness interacted with cannabis use to predict the likelihood of developing psychotic symptoms. In this study, rGE was unlikely to have been a confounder because no association between baseline psychosis proneness and subsequent use of cannabis was observed. Nonetheless, confounding cannot be ruled out entirely because the proxy genetic measure of psychometric psychosis proneness will also be influenced by environmental factors. As a complement to the observational designs described above, Verdoux et al. used a quasi-experimental “experience sampling” method and obtained similar findings showing that psychosis liability moderated the effect of cannabis in terms of “switching on” psychotic symptoms in the flow of daily life. For more details on possible gene × cannabis interactions, we refer to the article by Henquet and colleagues in this issue. Other studies using psychometric psychosis liability as a proxy measure for genetic risk were able to demonstrate G × E with childhood urbanicity (see below for more details) and childhood trauma.

**Summary of Epidemiological G × E Studies To Date**

In Table 1, the different epidemiological G × E studies are summarized. For each study, the proxy genetic factor, the proxy environmental factor, and the main findings as well as main limitations are summarized. Environmental
### Table 1. First-Generation Studies of Proxy Gene-Environment Interaction in Psychotic Disorder

<table>
<thead>
<tr>
<th>Proxy Genetic Variable</th>
<th>Proxy Environmental Variable</th>
<th>Findings</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive FH</td>
<td>Ethnic group</td>
<td>Familial morbid risk for psychotic disorder higher in siblings of African-Caribbean probands than in siblings of white probands.(^{96,108})</td>
<td>May be informative however preferably environmental exposure status and clinical status is measured in both cases and all first-degree relatives and analyses are adjusted for age, sex, and number of relatives.</td>
</tr>
<tr>
<td>Urban birth</td>
<td>No association between urban birth and a positive FH for psychotic disorder(^ {109}), however, interaction was tested on multiplicative rather than additive scale (see below)</td>
<td>Level of misclassification may be high because many unaffected relatives may carry the high-risk genotype.</td>
<td></td>
</tr>
<tr>
<td>Obstetric complications</td>
<td>Mostly inconclusive findings with regard to FH.(^ {110,112,113})</td>
<td>Absence of an association between positive FH and environmental exposure does not rule out gene-environment interaction, and presence of an association does not rule out lack of gene-environment interaction.(^ {109}) Evidence can be considered stronger if replicated (eg, urbanicity findings).</td>
<td></td>
</tr>
<tr>
<td>Birth in winter/spring</td>
<td>Positive, negative, and inconclusive associations with FH.(^ {114,115,116,117})</td>
<td>Testing for interaction on additive scale likely more informative.(^ {25}) Suggests that environmental factor is necessary for expression of high-risk genotype in affected twin or inhibition of protective genotype in unaffected twin.</td>
<td></td>
</tr>
<tr>
<td>Stressful life events</td>
<td>Positive association with FH.(^ {118})</td>
<td>Risk of psychotic disorder spectrum disorder 3% in absence of environmental risk, and 62% in presence of environmental risk. This difference seems extremely high.</td>
<td></td>
</tr>
<tr>
<td>Urbanicity</td>
<td>Evidence for synergism between urban environment (proxy environmental risk) and FH (proxy genetic risk) when tested on additive scale.(^ {119})</td>
<td>Children destined to develop psychotic disorder may have contributed to dysfunctional family environment rather than the other way round.</td>
<td></td>
</tr>
<tr>
<td>Having an identical twin with psychotic disorder</td>
<td>Being discordant for psychotic disorder</td>
<td>Children of both affected and nonaffected twin in discordant pair have higher rate of psychotic disorder.(^ {120,121,122})</td>
<td>Case-control comparison difficult because many other factors may be involved.</td>
</tr>
<tr>
<td>Biological parent with psychotic disorder</td>
<td>Growing up in dysfunctional adoptive family environment</td>
<td>Risk of psychotic disorder spectrum disorder or psychotic disorder–associated thought disorder higher in high-risk adoptees who had been brought up in dysfunctional adoptive family environment.(^ {94,123,124,125})</td>
<td></td>
</tr>
<tr>
<td>Institutional care and family instability</td>
<td>Very few cases of psychosis were identified in those families without a history of psychotic disorder, but among those with a FH strong environmental effects were observed.(^ {93})</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having positive relationships with father and mother</td>
<td>High-risk children with positive parental relationships had lower risk for developing psychotic disorder.(^ {126})</td>
<td>May suggest a negative G × E.</td>
<td></td>
</tr>
<tr>
<td>Having neither, 1 or 2 parents with psychotic disorder</td>
<td>Obstetric complications</td>
<td>The greater the proxy genetic risk, the greater the effect of obstetric complications, in particular fetal hypoxia, on ventricular enlargement (the psychotic disorder endophenotype).(^ {127,128})</td>
<td>Genetic risk may increase risk of obstetric complication (rGE). Genetic risk may increase risk of heavy alcohol consumption or head injury resulting in greater OC effect sizes.</td>
</tr>
<tr>
<td>Proxy Genetic Variable</td>
<td>Proxy Environmental Variable</td>
<td>Findings</td>
<td>Remarks</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Having a parent with psychotic disorder and additionally having an electrodermal abnormality as a child</td>
<td>Paternal absence</td>
<td>Higher rate of paternal absence in children who subsequently developed psychotic disorder.</td>
<td>Status of electrodermal abnormality as a marker of genetic risk for psychotic disorder unclear.</td>
</tr>
<tr>
<td>Having an MZ twin with psychotic disorder</td>
<td>Sharing the same chorion with the co-twin</td>
<td>Concordance rate was higher for MZ twins whose marker suggested that they were monochorionic than those whose marker indicated that they were dichorionic.</td>
<td>These results are compatible with an environmental factor in the prenatal environment facilitating expression of genetic risk for psychotic disorder.</td>
</tr>
<tr>
<td>Having expression of genetically influenced psychometric psychosis liability</td>
<td>Early trauma</td>
<td>Evidence that trauma and psychometric psychosis liability synergistically increase the risk for psychosis persistence over time.</td>
<td>Difficult to disentangle rGE from G x E.</td>
</tr>
<tr>
<td></td>
<td>Cannabis use</td>
<td>Evidence that cannabis and psychometric psychosis liability synergistically increase the risk for psychosis persistence over time, see also Verdoux et al.</td>
<td>Psychometric psychosis liability very indirect measure of genetic risk.</td>
</tr>
<tr>
<td></td>
<td>Growing up in urban environment</td>
<td>Evidence that urbanicity and psychometric psychosis liability synergistically increase the risk for psychosis persistence over time.</td>
<td></td>
</tr>
<tr>
<td>Being a member of a schizophrenia pedigree</td>
<td>Traumatic brain injury</td>
<td>Within the schizophrenia pedigrees but not bipolar pedigrees, traumatic brain injury was associated with a greater risk of schizophrenia, consistent with synergistic effects between genetic vulnerability for schizophrenia and traumatic brain injury.</td>
<td>Similar comments as for positive FH.</td>
</tr>
<tr>
<td>None</td>
<td>Having an older father</td>
<td>Having an older father is associated with an increased risk of schizophrenia in the offspring.</td>
<td>The underlying mechanism of this association may represent a special case of gene-environment interaction whereby the environment impacts on DNA sequence (de novo mutation) or DNA methylation (affecting gene expression). Thus, age of the father is a variable that is partly under control from the sociocultural environment, and older age may have an effect on DNA methylation in the sex cells (the inherited methylation pattern in humans is maintained in somatic cells but is erased and reestablished late in spermatogenesis for paternally imprinted genes, a process that could become impaired as age advances). Alternatively, advanced paternal age may lead to an increased rate of de novo mutations in gametes.</td>
</tr>
</tbody>
</table>
Exposures used in G × E studies include migration, urbanicity, obstetric complications, cannabis, stress, developmental trauma, and others. In most studies, the effect of genes and environment alone was rather small, and the bulk of their effect mediated through gene-environment interactions (Table 1).

**Epidemiological Replications of Gene-Urbanicity Interaction.** The finding that the rate of psychotic disorder is higher in children and adolescents growing up in an urban environment is well replicated and unlikely to be confounded entirely by rGE due to selective drift to urban areas in those at genetic risk for psychosis, although rGE may operate to some degree as it will in the case of many environmental risks. "Urbanicity" is a proxy for an as yet unidentified environmental factor(s) prevalent in urban areas and, if causal, may contribute to up to 20%-30% of the incidence of psychotic disorder in some countries. For this reason, urbanicity is an interesting factor to study in the context of G × E. Four studies in The Netherlands, Germany, Israel, and Denmark have attempted to examine gene-urbanicity interactions using epidemiological designs and indirect measures of genetic risk. All studies found evidence for gene-urbanicity interaction and are summarized in table 2. Clearly, the possibility of interaction between an environmental exposure in urban areas and genetic risk is in need of further study, focusing on (1) the precise nature of the urban exposure, eg, growing up in an area lacking in trust and cohesion, (2) the psychological and neurobiological mechanism of the environmental exposure in order to develop rational hypotheses about gene-environment interaction, (3) the nature of the genetic variation involved, and ultimately (4) the mechanism of the gene-environment interactions.

**Future Prospects**

To date, the study of gene-environment interactions has largely been epidemiological, where genotype, risk exposure, and disorder are studied as they occur in the population. A key contribution of a robust G × E comes from knowing that 3 apparently unconnected factors (gene, environmental risk factor, and disorder) are in fact causally linked. However, there are a number of methodological concerns that continue to challenge genetic-epidemiological research mainly because observational methods struggle to achieve the degree of control that is possible using experimental designs. Concerns are listed below.

**The Ideal Sample Size for G × E Research**

Clearly the optimal sample size required to detect G × E will vary according to the design used. For example,

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Measure Genetic Risk</th>
<th>Measure Urbanicity</th>
<th>Psychotic Disorder</th>
<th>Psychotic Symptom</th>
<th>Psychotic Incidence Associated with Urbanicity</th>
<th>Rate Unexposed</th>
<th>Rate G</th>
<th>Rate GE</th>
<th>Rate G + E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Os et al</td>
<td>Netherlands</td>
<td>Family history psychosis</td>
<td>Population density—dichotomous</td>
<td>Psychotic disorder</td>
<td>Psychotic symptoms</td>
<td>Summary increase in incidence of psychotic disorder</td>
<td>0.85%</td>
<td>3.01%</td>
<td>9.72%</td>
<td>1.59%</td>
</tr>
<tr>
<td>Van Os et al</td>
<td>Denmark</td>
<td>Family history psychosis</td>
<td>Five categories from capital city to rural area—5 levels</td>
<td>Psychotic disorder</td>
<td>Psychotic symptoms</td>
<td>Summary increase in incidence of psychotic disorder</td>
<td>14.9%</td>
<td>14.9%</td>
<td>14.9%</td>
<td>29%</td>
</tr>
<tr>
<td>Spauwen et al</td>
<td>Germany</td>
<td>Psychometric psychosis liability</td>
<td>City of Munich vs surrounding villages—dichotomous</td>
<td>Psychotic disorder</td>
<td>Psychotic symptoms</td>
<td>Summary increase in incidence of psychotic disorder</td>
<td>0.054%</td>
<td>0.054%</td>
<td>0.054%</td>
<td>0.054%</td>
</tr>
<tr>
<td>Weiser et al</td>
<td>Israel</td>
<td>Cognitive impairment endophenotype</td>
<td>Population density—5 levels</td>
<td>Psychotic disorder</td>
<td>Psychotic symptoms</td>
<td>Summary increase in incidence of psychotic disorder</td>
<td>0.011%</td>
<td>0.011%</td>
<td>0.011%</td>
<td>0.011%</td>
</tr>
</tbody>
</table>

aRate unexposed refers to those exposed to neither urbanicity nor genetic risk.
bRate G refers to those exposed to genetic risk only.
cRate GE refers to those exposed to genetic risk and urbanicity.
dRate G + E refers to those exposed to both genetic risk and urbanicity.

11
case-control studies will generally require very large sample sizes simply because the genetic effects are expected to be small. However, even with prospective cohort studies, large sample sizes may be required when the environmental risk factor(s) and/or disorder of interest occur at low frequencies. However, large sample sizes are not always necessary or desirable given the costs of amassing large samples. Indeed, sample size requirements can be substantially reduced with high-quality measurement of environmental risk factors, especially when measures are repeated over time; in particular the use of momentary assessment technologies with many repeated measures holds promise for the detection of subtle gene-environment interactions. Other methods to reduce sample size, based on selection of extreme exposure groups, may also apply.

Biostatistics

It is likely that mass genome-wide molecular genetic approaches, “enriched” with a few measures of “environmental” exposures will create invalid and confusing findings, largely because of the extent of multiple testing and the opportunities for post hoc analyses afforded by such studies. It is of paramount importance to consider the study of G x E as a separate discipline, requiring a highly specialized and multidisciplinary approach taking both environment and genes seriously. A hypothesis-driven strategy focusing on final common pathways in which biological synergism between genetic and environmental mechanisms take place, fed by information from functional enviromics and functional genomics pointing to promising neural systems and processes may constitute the most productive approach. In combination, this will enable a translational approach for systematically studying the effect of environmental manipulations on neural systems linked to genetic risk for schizophrenia. However, even a hypothesis-driven approach is likely to face major challenges in the area of biostatistics. Even allowing for, as discussed earlier, the major problem of how to bridge the gap between statistical interaction (statistical manipulations of data) and biological synergism (biological processes in nature), which currently cannot be estimated directly, solutions to, eg, modeling multiple ambiguous haplotype x environment interactions need to be developed. Fortunately, software allowing for modeling complicated interactions is currently being incorporated in several statistical programs.

Which Endophenotypes To Study?

In order to elucidate converging pathways that are the site of biological synergism between genes and environments, a wide range of approaches employing intermediate (or endo-) phenotypes may be used. For example, one may focus on the domain of neural systems-level intermediate phenotypes, cognition, neuroanatomy, salience attribution, treatment response, measures of course and outcome, subclinical psychosis expression, neurotic symptoms, and dynamic cerebral phenotypes in early-onset groups. The appeal of studying endophenotypes is obvious in that, compared with clinical diagnoses that are often characterized by substantial heterogeneity, endophenotypes appear to be cleaner, simpler constituents of psychopathology and (maybe falsely) promise improved chances of detecting true gene effects. Nonetheless, questions remain about which endophenotypes, for which disorder, are most worthy of study in a G x E framework. One argument against the use of endophenotypes is their apparently lower heritability estimates than the clinical phenotype. Although at first glance this may seem a valid argument, lower heritability estimates are only to be expected if endophenotypes reflect the “pure” contribution of genes and the clinical phenotype additionally represents the contribution of gene-environment interactions. The reason for this is that heritability estimates are derived from genetic epidemiological studies that estimate simple genetic and simple environmental contributions to schizophrenia liability. Unfortunately, these studies do not model the contribution of gene-environment interactions (G x E) because researchers tend to not include direct measures of the environment in such studies, thus precluding the quantification of gene-environment interactions. Therefore, the heritability of schizophrenia may be 80%, but simulations show that gene-environment interactions may make up the bulk of this proportion. Thus, endophenotypes may be more suitable measures of “pure” genetic risk because heritability estimates of the clinical disorder may be inflated by gene-environment interactions. Further research on this issue is needed.

Multiple Tests

As mentioned earlier, there are legitimate concerns about low prior probability testing for associations between a large number of polymorphisms (eg, via SNP chips) and specific disorders in the absence of some guiding theory that will allow researchers to sort true from false-positive associations. Guarding against “fishing trips” is important if we are to advance our understanding of how G x E operates in the development of schizophrenia.

Conclusion

Not only is there meta-analytic support for environmental effects on schizophrenia risk, evidence is now accumulating that environmental exposures are impacting on the risk for psychotic disorder in coparticipation with genetic factors and that effects of genes and environment in isolation are likely small or nonexistent.
Embracing a G × E approach has implications for gene discovery. That is, selecting and/or stratifying samples based on documented environmental risk exposure may not only help in the quest to identify new susceptibility genes for psychotic disorders but also in unraveling the pathway(s) to the onset of first-episode psychosis. For molecular genetic research, this means that the strategy of “brute force,” used to compensate for loss of power due to underlying G × E by inclusion of huge samples of many thousands of patients and hundreds of thousands of markers along the genome, may be complemented by imaginative approaches based on environmental stratification. Genetic odds ratios of 1.1 in nonstratified samples may be considerably higher in exposed samples. In addition, distal tiny genetic contributions by themselves explain little if more proximal interactions with environmental component causes explain the underlying pathophysiology.

It is obvious that more funding needs to be directed to G × E research—after nearly 1500 inconclusive molecular genetic investigations in schizophrenia complementary approaches no longer need to be excluded. The European Network of Schizophrenia Networks for the Study of Gene-Environment Interactions (EU-GEI)\textsuperscript{1} has suggested that part of the funding may be necessary to bring together the multitude of disciplines, currently working in isolation of each other, which is necessary for the study of gene-environment interactions.

Future research needs to better integrate epidemiological and experimental paradigms focusing on functional envrionics and functional genomics.\textsuperscript{1,58} This is desirable because neither traditional genetic epidemiology nor epidemiologic studies on isolated environmental factors can tell us much about the biological mechanisms involved in a G × E. These approaches are complementary, with each informing the other, and ideally should be used in unison for best effect. Many (but by no means all) of the challenges confronting genetic epidemiology listed above can be addressed using experimental designs with their advantages of greater experimental control and precision. However, these benefits have to be balanced against the loss of ecological validity that can sometimes result.

Epidemiologists should be encouraged to incorporate more physiological (ie, mechanistic) measures in their studies, and to move beyond 2-way interactions to models involving multiple genes and environments as well as gene-gene and environment-environment interactions.

Acknowledgments


References

5. Sullivan PF. The dice are rolling for schizophrenia genetics. [published online ahead of print June 4, 2008]. Psychol Med. doi:10.1017/S0144686X0800534X.
7. Crow TJ. The emperors of the schizophrenia polygene have no clothes. [published online ahead of print April 4, 2008]. Psychol Med. doi:10.1017/S0033291708003395.
J. van Os et al.


45. Pedersen CB, Mortensen PB. Are the cause(s) responsible for urban-rural differences in schizophrenia risk rooted in families or in individuals? Am J Epidemiol. 2006;163:971–978.


62. Collip D, Myin-Germeys I, Van Os J. Does the concept of “sensitization” provide a plausible mechanism for the putative link between the environment and schizophrenia? Schizophr Bull. 2008;34:220–225.


