Molecular genetic gene–environment studies using candidate genes in schizophrenia: A systematic review

Gemma Modinos a,⁎, Conrad Iyegbe a, Diana Prata a, Margarita Rivera b,c, Matthew J. Kempton a, Lucia R. Valmaggia a, Pak C. Sham d, Jim van Os e, Philip McGuire a

⁎ Corresponding author at: Department of Psychosis Studies, PO67, Institute of Psychiatry, King's College London, United Kingdom.

E-mail address: gemma.modinos@kcl.ac.uk (G. Modinos).

1. Introduction

Schizophrenia has a strong genetic component. Classical twin studies estimate its heritability at about 80% (Sullivan et al., 2003), although population-based family studies indicate that the true estimate may be closer to 60% (Lichtenstein et al., 2009), and may be even lower once gene–gene (Zuk et al., 2012) and gene–environment (van Os and Sham, 2003) interactions are taken into account. The substantial heritability associated with schizophrenia naturally led researchers to look for connections with genes. The modern systematic method of genome-wide association studies (GWAS), which allows for thousands of tests of main genetic effects across the genome, has been used in psychosis research in recent years and has produced replicated findings. However, the risk conferred by each common variant is indeed as small as that obtained in the previous hypothesis-based candidate gene studies (typical odds ratio 1.1–1.2) (O’Donovan et al., 2008; Ripke et al., 2011). Furthermore, although copy number variants (CNVs) were shown to have up to a 15 times larger effect on risk (Stefansson et al., 2008), they, on the other hand, are relatively rare, accounting for only a small proportion of the genetic liability to schizophrenia (International Schizophrenia Consortium, 2008). Although from a translational perspective, the impact of GWAS technology is still limited (McCarthy et al., 2010); this GWAS approach has opened new avenues of research in the field, for instance by discovering risk genes that provide new pathophysiological clues to explain the etiology of psychosis.

A separate body of epidemiological research has established that the incidence of schizophrenia is increased in people who have grown up in urban environments, in immigrant groups and their offspring, in individuals who experienced traumatic events in early life, and in regular cannabis users (van Os et al., 2010). While these post-natal environmental risk factors appear to be robust and some are associated with much larger effects than genetic factors (Fearon et al., 2006), the neurobiological mechanisms that underlie their effects on schizophrenia risk are largely unknown. Furthermore, and also in the case of peri-natal risk factors such as obstetric complications during birth, it is plausible that genes moderate the impact of these environmental events.

Gene–environment studies aim to understand how these gene and environment effects interact to increase risk (Moffitt, 2005). In a gene × environment (G × E) interaction, the impact of an environmental risk factor (or vice-versa for a genetic factor) on disease liability is higher when a certain genetic risk factor is also present than when it is not, or when its presence is not taken into account. In psychiatry,
G × E research is still an emerging discipline, and important conceptual and pragmatic questions have been raised on how to conduct and interpret G × E findings from research in psychotic disorders (Zammit et al., 2010b). A number of G × E studies have used quantitative genetic epidemiology methods, namely twin and adoption designs, to study psychosis, and this contribution has been reviewed elsewhere (Van Os et al., 2008). The main effort of this article is to systematically review the current evidence that molecular genetic candidate G × E studies contribute to the etiology of psychosis, to then highlight particular challenges and opportunities of this approach, and to finally articulate a discussion of future research directions.

2. Methods

In order to identify suitable publications for this review, we conducted an online search of the PubMed, Medline and PsychInfo databases using the following search strategy: (“Schizophrenia” OR “Psychosis”) AND “Gene” AND “Environment”. We targeted studies identifying and testing for interaction between an environmental pathogen and a candidate susceptibility gene, and restricted the focus on psychosis as phenotype/outcome (i.e. psychotic symptoms, diagnosis of psychotic disorder). Some of the returned articles did not meet the inclusion criteria upon detailed examination and were subsequently excluded. Studies that did not meet our inclusion criteria included reviews and other articles that did not include original data, studies exploring uniquely genetic (or uniquely environmental) main effects on psychosis, studies examining interactions on other phenotypes (e.g., cognition, white matter volume), articles written in non-English language and studies performed on populations outside of the schizophrenia spectrum. In addition to the online search criteria, in order to ensure that no studies had been omitted, we also conducted a manual search through the bibliographic sections of included articles. This approach identified a total of 21 studies.

3. Results

Due to the considerable variation in susceptibility genes explored among studies, for ease of understanding the articles are described under five sections, based on the environmental factor used: cannabis, seasonality of birth, stress, childhood abuse/trauma, obstetric complications.

Table 1 provides an overview of the 21 selected studies, their sample characteristics, genetic factors, environmental factors, outcome measure and main findings. Supplementary Table 1 gives further details on parameterization of exposure and outcome variables of each study.

3.1. Gene × cannabis

Research into putative gene by cannabis interactions revolves around the notion that genetic liability for psychosis may be expressed as differential sensitivity to the psychotomimetic effect of cannabis (Genetic Risk Outcome in Psychosis Investigators, 2011). Which genes underlie differential sensitivity to cannabis, however, is unclear. To date, nine publications have investigated interactions between a priori underlie differential sensitivity to cannabis, however, is unclear. To date, nine publications have investigated interactions between a priori


determining the relationship between cannabis exposure and schizophrenia risk. A further sub-analysis of cases (n = 493) examined interactions between cannabis and COMT Val158Met and, unlike in the study of Caspi et al. (2005), no significant interaction was found, even when cannabis use was restricted to adolescence. In a later study, Kantrowitz et al. (2009) investigated the interaction of the same polymorphism and adolescent cannabis in African-American and White patients with psychosis and found no significant associations in either group. Finally, Zammit et al. (2011) used a large general population sample to evaluate the risk of psychotic experiences at age 16 as a function of COMT Val158Met and cannabis use at age 14. Although cannabis use increased the risk of later psychotic experiences, it did so irrespectively of the COMT genotype.

The effects of cannabis on psychosis have also been investigated with other candidate genes. van Winkel and the Genetic Risk Outcome Psychosis Group (2011) included patients with schizophrenia, siblings, and controls, and applied at-risk, case-control, case–sibling and case–only approaches, with an initial pool of 152 a priori candidate SNPs within 42 genes. In the sibling group (at-risk approach), three SNPs survived Bonferroni correction for multiple testing of their main effects on the presence of schizotypal experiences after recent cannabis exposure (as determined by urinalysis), within the protein kinase B gene (AKT1 rs2494732 and rs1130233) and LRRMT1 (rs673871). These three SNPs then became the subject of case-only, case–sibling and case–control analyses, to examine the effect of genetic moderation of lifetime cannabis exposure (as determined by the Composite International Diagnostic Interview, CDI; Robins et al., 1988) on adult risk of psychosis. The interaction between AKT1 rs2494732 and cannabis use was the only finding supported in all three sub-analyses. In subjects with early-life experience of cannabis, the C/C genotype increased the risk of being diagnosed with psychosis approximately two-fold. There was no evidence for an interaction with COMT Val158Met. Interestingly, Di Forti et al. (2012) recently reported a direct replication of the study of van Winkel and the Genetic Risk Outcome Psychosis Group (2011), showing an interaction between variation at AKT1 rs2494732 and lifetime history of cannabis use on increasing psychosis risk. C homozygotes with a history of cannabis use showed a greater than two-fold increased likelihood of psychosis when compared to those who were T homozygotes. In addition, the interaction between AKT1 rs2494732 and lifetime frequency of use was also significant, with C homozygotes who were daily cannabis users showing a seven-fold increase in the odds of psychosis compared to T homozygotes. Di Forti et al. found no main effects of AKT1 or cannabis use alone on the risk of psychosis, and no significant change in risk associated with AKT1 among subjects who had never used cannabis.

Bhattacharyya et al. (2012) reported an effect of the dopamine transporter gene (DAT1) 3′ UTR VNTR and a trend for an effect of the AKT1 rs1130233 genotype on the increase in psychotic symptoms...
Table 1
Summary of studies examining interactions between molecular candidate genetic (G) and environmental risk factors (E) for psychosis.

<table>
<thead>
<tr>
<th>Study reference (year)</th>
<th>Sample</th>
<th>Candidate G</th>
<th>Candidate E</th>
<th>Outcome variable</th>
<th>Results</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Di Forti et al. (2012)</td>
<td>489 first-episode psychosis patients, 278 HC</td>
<td>AKT1 rs2494732</td>
<td>Cannabis (lifetime history of use)</td>
<td>Psychotic disorder</td>
<td>+</td>
<td>AKT1 rs2494732 × lifetime history of cannabis use interaction (likelihood ratio test = 8.54; P = .014). C homozygotes with history of cannabis use showed greater than twofold increased odds of having psychosis (OR = 2.18; 95% CI: 1.10–4.31) compared with T homozygotes. AKT1 rs2494732 × lifetime frequency of cannabis use interaction (likelihood ratio test = 13.39; P = .010). C homozygotes who were daily cannabis users showed a greater probability of having psychosis compared with T homozygotes (OR = 7.23; 95% CI: 1.37–38.12).</td>
</tr>
<tr>
<td>Bhattacharyya et al. (2012)</td>
<td>35 HC</td>
<td>DAT1 3′ UTR VNTR, AKT1 rs1130233</td>
<td>Cannabis (delta-9-THC)</td>
<td>Delta-9-THC-induced psychotic experiences</td>
<td>+</td>
<td>Under delta-9-THC, AKT1 rs1130233 G homozygotes that were also 9-repeat DAT1 3′ UTR VNTR carriers had an increase in psychotic experiences relative to 10-repeat homozygotes (P &lt; .001).</td>
</tr>
<tr>
<td>Zammit et al. (2011)</td>
<td>750 SZ, 688 HC</td>
<td>COMT Val158Met (rs4680), rs4818, rs6269, rs737865, rs2097603, rs165599</td>
<td>Cannabis</td>
<td>Psychotic experiences</td>
<td>−</td>
<td>No interactions (P-values from .304 to .981).</td>
</tr>
<tr>
<td>Kantrowitz et al. (2009)</td>
<td>92 SZ patients (33 White, 46 African-American)</td>
<td>COMT Val158Met (rs4680)</td>
<td>Cannabis</td>
<td>Genotype associations with cannabis use in Caucasians and African-Americans</td>
<td>−</td>
<td>No differences in COMT Val158Met associations with adolescent cannabis use and in African American (χ²(2) = 29; P = .23) or White (χ²(2) = 1.45; P = .49) patients.</td>
</tr>
<tr>
<td>Henquet et al. (2009)</td>
<td>31 psychotic disorder, 25 HC</td>
<td>COMT Val158Met (rs4680)</td>
<td>Cannabis</td>
<td>Psychotic experiences (ESM)</td>
<td>+</td>
<td>Cannabis use increased hallucinatory experiences only in COMT Val carriers with high scores on psychometric psychosis (Val/Val: β = .78, 95% CI: .36–1.19, P &lt; .001; Val/Met: β = .21, 95% CI: .01–.44, P = .063; Met/Met: β = .21, 95% CI: .13–.55, P = .22).</td>
</tr>
<tr>
<td>Zammit et al. (2007)</td>
<td>750 SZ, 688 HC</td>
<td>CRN1 rs1049353</td>
<td>Cannabis</td>
<td>Psychotic disorder</td>
<td>−</td>
<td>No CRN1 rs1049353 × cannabis interaction (OR = .83, 95% CI = .65–1.05). No association between COMT Val158Met and cannabis use (OR = .98, 95% CI = .76–1.27, P = .89) in patients with schizophrenia, regardless of age at start of cannabis use.</td>
</tr>
<tr>
<td>Henquet et al. (2006)</td>
<td>30 psychotic disorder, 12 relatives, 32 HC</td>
<td>COMT Val158Met (rs4680)</td>
<td>Cannabis (delta-9-THC)</td>
<td>Delta-9-THC-induced psychotic experiences</td>
<td>+</td>
<td>No interactions (OR = .90, 95% CI = .47–1.70; P = .744).</td>
</tr>
<tr>
<td>Caspi et al. (2005)</td>
<td>803 HC</td>
<td>COMT Val158Met (rs4680)</td>
<td>Cannabis</td>
<td>Schizophreniform disorder</td>
<td>−</td>
<td>Frequency of the 5-HTTLPR allele showed one-cyclic season of birth variation in men with SZ (P = .01).</td>
</tr>
<tr>
<td>Alemany et al. (2011)</td>
<td>533 HC</td>
<td>BDNF Val66Met (rs6265)</td>
<td>Childhood abuse</td>
<td>Psychotic experiences</td>
<td>+</td>
<td>BDNF Met carriers with childhood abuse reported more positive psychotic-like experiences than Val homozygotes (β = .27, 95% CI: -.2–.54; Val/Met: OR = .5, 95% CI: .76–.82; but not in Met homozygotes (OR = 1.1, 95% CI: 21–54).</td>
</tr>
<tr>
<td>Munjewerff et al. (2011)</td>
<td>742 SZ, 884 HC</td>
<td>MTHFR C677T (rs1801133)</td>
<td>Childhood neglect</td>
<td>Psychotic experiences</td>
<td>−</td>
<td>No interactions (OR = .90, 95% CI = .47–1.70; P = .744).</td>
</tr>
<tr>
<td>Chotai et al. (2003)</td>
<td>954 UPAD, BPAD, and SZ</td>
<td>TPH1 A218C (rs1800532)</td>
<td>Seasonality of birth</td>
<td>Season of birth variations in UPAD, BPAD, and SZ</td>
<td>+</td>
<td>Frequency of the TPH1 allele A showed season of birth variations with one cycle per year in HC women (P = .05) and men with BPAD (P = .05).</td>
</tr>
<tr>
<td></td>
<td>395 HC</td>
<td>5-HTTLPR</td>
<td>Seasonality of birth</td>
<td>Season of birth variations in UPAD, BPAD, and SZ</td>
<td>+</td>
<td>Frequency of the 5-HTTLPR allele s showed one-cyclic season of birth variation in men with UPAD (P = .01).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DRD4</td>
<td>Seasonality of birth</td>
<td>Season of birth variations in UPAD, BPAD, and SZ</td>
<td>+</td>
<td>The frequency of the DRD4 7-repeat allele showed one-cyclic season of birth variation in women with SZ (P = .01).</td>
</tr>
<tr>
<td>Study Authors</td>
<td>Sample Size</td>
<td>Genotype/SNP</td>
<td>Mode of Analysis</td>
<td>Outcome/Interaction</td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>------------------</td>
<td>------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Tochigi et al. (2002)</td>
<td>110 SZ, 493 HC</td>
<td>HLA (HLA-A24, HLA-A26)</td>
<td>Seasonality of birth</td>
<td>Association between HLA-A and winter birth in SZ</td>
<td>No association between HLA-A24 or A26 and winter birth (December-March) in patients with schizophrenia. Frequencies of winter birth were 3% in patients with HLA-A24. 25% in those without HLA-A24 (P = .06), and 35% in patients with HLA-A26 vs. 44% in those without HLA-A26 (P = .4). Increased incidence of winter births (February-March) in patients with HLA-A24 than in patients without (P = .003).</td>
<td></td>
</tr>
<tr>
<td>Narita et al. (2000)</td>
<td>60 SZ + HLA-DR1, 307 SZ no HLA-DR1</td>
<td>HLA (HLA-DR1)</td>
<td>Seasonality of birth</td>
<td>Association between HLA-DR1 and winter birth in SZ</td>
<td>+ AKTI: Proband with OCs were more likely to have the more allele at rs1820510 (OR = 3.39, 95% Cl = .93–11.3; P = .011). AKTI: Proband with OCs were more likely to have the more allele at rs1820510 (OR = 3.39, 95% Cl = .93–11.3; P = .011).</td>
<td></td>
</tr>
<tr>
<td>Nicodemus et al. (2008)</td>
<td>116 SZ spectrum disorders, 134 HC</td>
<td>AKTI, BDNF, CAPON, CHRNA7, COMT, DNTNIP1, GADD1, GRM3, NOTCH4, NRG1, PRODH, RCS4, TNFα</td>
<td>Obstetric complications</td>
<td>SZ</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Keri et al. (2009)</td>
<td>200 SZ</td>
<td>NRG1 SNP98NRG243177 (rs10954867), rs70052888 (the latter two as control SNPs)</td>
<td>Psychosocial stress (neutral or conflict-related family interactions)</td>
<td>Unusual thoughts</td>
<td>No differences between NRG1 T/T and C/C patients (P &lt; .001). There were no differences in effect size between Val/Val carriers and Val/Met carriers (P &lt; .01). No differences among patients with different NRG1 genotypes during neutral family interactions (P &gt; .5).</td>
<td></td>
</tr>
<tr>
<td>Simons et al. (2009)</td>
<td>579 young adult female twins (general population)</td>
<td>COMT Val158Met (rs4680)</td>
<td>Stress (ESM) (&quot;event&quot; and &quot;social&quot; stress)</td>
<td>Feelings of paranoia (ESM)</td>
<td>No effect size between Val vs. Met homozygotes was in the highest categories of unpleasant (P = .03) and very unpleasant appraisals (P = .001). No COMT Val158Met × &quot;social stress&quot; interaction: Val/Met carriers showed more feelings of paranoia than Val/Val carriers (P = .001). Differences in effect size between the Val/Val and the Val/Met variants were significant for unpleasant (P = .04) and very unpleasant (P = .004) appraisals.</td>
<td></td>
</tr>
<tr>
<td>van Winkel et al. (2008a)</td>
<td>31 psychotic disorder + cannabis, 25 HC + cannabis</td>
<td>COMT Val158Met (rs4680)</td>
<td>Stress (ESM)</td>
<td>Psychotic experiences (ESM)</td>
<td></td>
<td>No differences in effect size between Val/Val and Val/Met carriers for Delusions (P = .03), and no differences in effect size between Val/Val and Val/Met carriers for Delusions (P = .03). Differences in effect size between Val/Val and Val/Met carriers for Delusions (P = .03) and no differences in effect size between Val/Val and Val/Met carriers for Delusions (P = .03). Differences in effect size between Val/Val and Val/Met carriers for Delusions (P = .03).</td>
</tr>
<tr>
<td>Stefanis et al. (2007)</td>
<td>306 HC</td>
<td>COMT Val158Met (rs4680)</td>
<td>Stress</td>
<td>Psychotic symptoms</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

SNP = single nucleotide polymorphism; SZ = patients with schizophrenia; ESM = experience sampling method; GADD1 = glutamate decarboxylase 1; GRM3 = metabotropic glutamate receptor 3; HC = healthy controls; HLA = human leukocyte antigen; MTHFR = methylenetetrahydrofolate reductase; NOTCH4 = neuralized locus notch homolog protein 4; NRG1 = neuregulin; OR = odds ratio; PRODH = proline dehydrogenase; RCS4 = regulator of G protein signaling 4; SNP = single nucleotide polymorphism; SZ = patients with schizophrenia; TNF-α = tumor necrosis factor; TPH = tryptophan hydroxylase gene; UPAD = unipolar affective disorder.
induced by acute D-9-THC administration. There was an interaction between the effects of the DAT1 and AKT1 genotypes on the effect of D-9-THC, with a greater increase in symptoms in G homozygotes for AKT1 who were also DAT1 9-repeat carriers relative to those who were not. Although this would support van Winkel et al.’s suggestion of AKT1 as a modulator of the effects of cannabinoids, the effect in the van Winkel et al. study was of a history of early cannabis use on the risk of psychotic disorder across a sample including patients, relatives and healthy volunteers, as opposed to the acute effect of a single dose of D-9-THC on psychotic symptoms in healthy volunteers. In addition, the van Winkel et al. study reported effects for rs2494732, while Bhattacharyya et al. did so for rs1130233. Finally, an important issue that represents an important caveat in interpretation with regard to cannabis studies but also across all the reviewed studies is that sample sizes differed greatly between studies (see Table 1 for sample characteristics of each study). In this particular case, van Winkel et al. investigated a sample of nearly 2000 individuals, whereas Bhattacharyya et al. included 35 healthy controls, therefore their results should be contrasted with caution.

3.2. Gene × seasonality of birth

A number of studies have been published reporting associations between schizophrenia and season of birth. In particular, for individuals born in the Northern Hemisphere, birth in late winter or early spring has been associated with a 5–10% greater likelihood of developing schizophrenia (Torrey et al., 1997; Davies et al., 2003; Bembenek and Kociuba, 2005; Mino and Oshima, 2006). More recently, it was proposed that seasonally varying environmental factors might influence the risk of schizophrenia in concert with candidate genes; to date, four studies have sought to address this issue.

Narita et al. (2000) reported an association between incidence of births in winter and the Human Leukocyte Antigen (HLA)-DR1 in patients with schizophrenia, suggesting an interaction of these two risk factors. However, another Japanese study (Tochigi et al., 2002) failed to replicate this, reporting no significant association between HLA-A24 or HLA-A26 and season of birth in schizophrenia. Chotai et al. (2003) examined the risk of psychosis as a function of three polymorphisms within the tryptophan hydroxylase (TPH), the serotonin transporter (5-HTTLPR) and the dopamine D4 receptor (DRD4) genes and season of birth. Interactions were sought in patients with unipolar affective disorder (UPAD), bipolar affective disorder (BAD), schizophrenia, and healthy controls. The authors observed that the effects of variations in these genes dependent on season of birth were different for the different psychiatric disorders (see Table 1 for further details).

Finally, Muntjewerff et al. (2011) examined the interaction between the folate-regulating gene Methylene tetrahydrofolate reductase (MTHFR) C677T (rs1801133) and winter birth on schizophrenia risk using a case–control design, and reported no significant main effects or interactions.

3.3. Gene × stress

Evidence from general population studies suggests that individuals with certain genetic polymorphisms are more sensitive to stress, and stress is a trait which has long been associated with the development of psychosis (see van Winkel et al., 2008b for a review). To date, there have been six studies specifically testing candidate gene by stress interactions in psychosis.

Stefanis et al. (2007) reported, in a sample of healthy men assessed at recruitment for military service and after 18 months of training, that COMT Val carriers were more susceptible to have a psychotic outcome under the effect of stress than Met homozygotes. van Winkel et al. (2008a) showed, in patients with psychosis, that COMT Met homozygotes had the largest increase in psychotic experiences in response to stress. This is opposite to the previous report by Stefanis et al. (2007), who had found that it was COMT Val carriers who showed increased risk of psychosis. However, a difference between studies is that van Winkel et al. used patients with psychosis, whilst Stefanis et al. investigated healthy individuals. Later, Collip et al. (2011) replicated the van Winkel et al. finding of COMT Val158Met by stress interaction in risk of psychosis outcome patients, showing that COMT Val158Met moderated the association between stress and psychotic experiences, with Met homozygotes showing the largest psychotic reactivity to stress.

Simons et al. (2009) reported that healthy female carriers of the COMT Val allele displayed more feelings of paranoia in response to event stress (stress related to a recent event) than Met carriers. On the other hand, carriers of the BDNF Met allele showed more feelings of paranoia in response to social stress (stress related to the person they were with) than Val homozygotes. Using a similar methodology, Peerbooms et al. (2012) found that MTHFR C677T moderated the interaction between COMT Val158Met and stress on increasing psychotic symptoms in patients. In particular, COMT Met/Met patients with the MTHFR T allele displayed the largest increases in psychotic symptoms in reaction to stress, whereas for patients who were MTHFR C homozygotes there was no symptomatological difference. Finally, Keri et al. (2009) investigated the interaction between the Neuregulin gene (NRG1) rs6994992 and psychosocial stress on unusual thoughts in patients with schizophrenia. Patients who were NRG1 T homozygotes expressed more unusual thoughts than C-carriers during conflict-related interactions, but not during neutral interactions. Two control polymorphisms of the NRG1 gene, which are not related to schizophrenia and do not affect gene expression (rs10854867 and rs7005288), showed no such effect.

3.4. Gene × childhood abuse/trauma

Childhood adversity has been shown to be a risk factor for the development of psychotic symptoms in clinical and non-clinical samples (see Schafer and Fisher, 2011 for a review). Despite this established relationship, it is necessary to consider the type and severity of the environmental exposure, together with the specific individual genetic background of risk, in order to understand the development of psychosis in adulthood. To date, only one study sought to address the issue of gene by childhood trauma interactions in psychosis. Alemany et al. (2011) examined interactions between childhood adversity and the BDNF Val66Met polymorphism on lifetime prevalence of positive and negative psychotic experiences in adulthood. They observed that Met carriers had higher scores on adult positive psychotic-like experiences when childhood abuse was present, as compared with Val homozygotes.

3.5. Gene × obstetric complications (OCs)

OCs, occurring during pregnancy, delivery, and the neonatal period, have been documented as risk factors for schizophrenia (Geddes et al., 1999; Cannon et al., 2002). In terms of G × E, severe OCs are proposed to interact with genetic risk factors to increase risk for schizophrenia (see Mittal et al. 2008 for a review). Nicodemus et al. (2008) tested this hypothesis by including genes thought to be influenced by hypoxia or involved in brain vascular function in a family-based study of probands with a range of schizophrenia spectrum disorders. Of 13 genes examined (with a total number of 290 SNPs), four showed significant interaction with OCs in schizophrenia risk: AKT1 (rs2494735, rs3803300, rs1130233), BDNF (rs2049046, ss76882600), DTNBP1 (rs875462) and GRM3 (rs7808623).

4. Discussion

4.1. Challenges and opportunities of the molecular genetic candidate G × E approach

Psychosis research is not immune from the general pitfalls afflicting G × E research in psychiatry. Limitations in sample size and study...
design have been features of research in this field (Uher and McGuffin, 2008; Caspi et al., 2010), such that studies reporting negative findings need to have a six-fold difference in their sample size for acceptance into the public domain, compared to positive studies (Duncan and Keller, 2011). Several more technical themes have also conspired to thwart the success of replication studies in G × E research, such as measurement error and power, scaling of effects, distribution of risk exposures, and schizophrenia-specific factors.

4.1.1. Measurement error and power

Environmental factors are typically subject to much greater misclassification than genetic factors. A higher level of measurement error is innate to large-scale studies of candidate-based G × E (Caspi et al., 2010), but it is suggested that strategies that optimize the specificity and accuracy with which exposures are measured can offset the deficits in power incurred through under-sampling. Simulations of measurement error by Wong et al. (2004) help to qualify this point and provide empirical support to it. They suggest that an increase in correlation of the measured values with the true values of “Environment” from .4 to .7 can equate to as much as a 20-fold gain in sample size. This would indicate that the problem of a small sample could, to an extent, be overcome by maximizing the precision of environmental measures. This could be achieved by, for example, the use of mobile health (M-Health) measures in real life using variations of Experience Sampling in combination with ambulatory physiological parameters. In addition, environmental measures can be improved by the use of experimental assessments (at the behavioral, fMRI, PET, or cognitive levels) of the impact of, for instance, drug use, social defeat, or stress.

The retrospective nature of exposure assessment in many studies is also problematic, especially in patients with schizophrenia in whom recall and interpretation may be biased by cognitive impairments and delusional beliefs. Studies that involve the prospective assessment of environmental factors at the time they are active are thus preferable, but arelogically demanding. Nevertheless, such studies are ongoing. For example, one study within the European Network of National Schizophrenia Networks studying the Gene–Environment Interactions Program (EU-GEI; European Network of Schizophrenia Networks for the Study of Gene–Environment Interactions, 2008) is assessing the exposure of subjects at ultra-high risk for psychosis to established environmental risk factors for psychosis (e.g., stress, cannabis use), collecting genetic data, and then following a large cohort over several years in order to determine which individuals will subsequently develop psychosis (http://www.eu-gei.eu/).

4.1.2. Scaling of effects

As for any type of statistical analysis, a typical G × E analysis requires large samples to facilitate the detection of small interaction effects. A wider debate surrounds how these interactions should be scaled. In order to determine the presence of an interaction, a product term is added to the regression model. In linear regression, the regression coefficient of the product term defines interaction as departure from additivity of the effects on the dependent variable. However, in logistic regression interaction is defined as a departure from multiplicity in the risk of disease (Knol et al., 2007). While an additive model is often thought to best epitomize the concept of biological interaction (Darroch, 1997; Kendler and Gardner, 2010), logistic regression is the standard statistical tool for the analysis of epidemiological studies.

At the heart of the issue is the question of biological validity, as understanding the biological basis of an interaction is the ultimate goal for G × E research, particularly from a translational perspective. Biological interactions need not give any statistical clues to their existence. This is demonstrated by the example of phenylketonuria (PKU), which results from a combination of homozygous “loss-of-function” mutations in the gene encoding the phenylalanine hydroxylase enzyme, and dietary exposure to phenylalanine. Under natural conditions, any statistical trace of this biological interaction is obscured by the ubiquitous nature of phenylalanine in the human diet. Biological validity remains a gold standard for all G × E research because the concept of a biological interaction is easy to understand and forms the basis for designing interventions (assuming an interaction is large enough to merit this course of action). In contrast, inferring a mechanistic relationship out of a statistical effect relies on conditions and assumptions (VanderWeele et al., 2010) that may not necessarily hold true for schizophrenia (Zammit et al., 2010a). The difference between these two definitions (of biological versus statistical interaction) can be problematic, as there remains plenty of scope for conflict between the two. A statistical interaction may still have great predictive value nonetheless. In some cases discrepancies between the two may be artifactual. For example, logarithmic transformation of variables that exhibit multiplicative effects can cause bona fide interactions to disappear; while in other scenarios transformation may induce interactions spuriously (Kendler and Gardner, 2010).

These issues have fueled a debate about the most appropriate way to scale interaction effects (e.g., Kendler and Gardner, 2010; Zammit et al., 2010a). A key step to obtaining a definitive answer to this question will be the introduction of more systematic approaches to G × E discovery. Such approaches, as epitomized by GWAS (Engelman et al., 2009; Thomas, 2010), may unveil consistent patterns of G × E for multiple inter-related genes within critical pathways. In the future this will be further complemented by the genome sequencing projects now underway in schizophrenia (Bickeboller et al., 2011).

4.1.3. Distribution of risk exposures

Statistical power could be potentially compromised if the distribution of genotypes and exposure for a given sample are unfavorable. Power to detect interactions is at an optimum for “balanced” designs, when both minor allele frequencies and exposure rates are 50%. However, this is unlikely under case–control designs, or indeed even for prospective studies when multiple genetic polymorphisms are to be tested. A more feasible way to guarantee greater power to detect G × E is the use of selective sampling (Boks et al., 2007), that is, selecting subjects with extremely high and low environmental exposure. A limitation to the approach is that the lack of information on subjects with intermediate levels of environmental exposure may obscure the exact pattern of G × E.

The issue of replication is further complicated by the fact that, depending on the frequency of the exposure, the same G × E construct may exhibit: (i) no effect when the prevalence of environmental exposure is very low, (ii) statistical interaction when the prevalence is moderate or (iii) a main effect when the prevalence is very high (Caspi et al., 2010). Variation in genotypic frequency has the same impact on the observed form of an underlying G × E. Replication is therefore only valid when the exposure and genotypic profiles of the replication sample mimic the same characteristics of the discovery sample.

4.1.4. Statistical model

One of the major problems with replication studies in G × E research is the lack of testing for G × E directly as performed in the original study. For instance, as will be described in more detail in the next section, while the original study of Caspi et al. (2005) tested for a direct G × E interaction, the studies of Henquet et al. (2006, 2009) tested for three-way interactions, and Peerbooms et al. (2012) tested for a four-way interaction. Consistency in the choice of a statistical model is an essential requirement for candidate G × E studies aiming at replication, therefore it is important to consider the patterns of interaction reported when attempting to conclude in favor of replication.

4.1.5. Schizophrenia-specific factors

Schizophrenia research is complicated by the variety of environmental etiologies linked to the disorder, and the failure to conceptualize these into a smaller number of coherent domains when assessing the level of environmental risk exposure. Approaches to the study of
G × E in depression acknowledge stress as a major stimulus. This is a main reason why the parameterization of stress into a single cumulative score (of stressful life events) has been successful for this disorder. In contrast, efforts to understand the genetic basis of G × E in schizophrenia have been diluted by the heavy diversification of etiological models proposed. Known environmental factors include urbanicity, migration, ethnicity, social capital, discrimination, social defeat, childhood trauma, stress, substance use, and stressful life events. There is a need to find a common mechanism tying these risk factors together. For example, social defeat (the negative experience of being excluded from the majority group) is thought to represent the specific “exposure” underlying the effects of different environmental risks, namely urban upbringing, migration, childhood trauma, low intelligence and drug abuse (Selten and Cantor-Graae, 2005, 2007). The parameterization of social defeat into a cumulative score can be successful for approaches to the study of G × E in schizophrenia, using specific social defeat designs rather than weak environmental indicators of social defeat (Selten et al., 2013).

4.2. Future prospects

In view of the current state of the art in G × E research, with the available evidence encouraging further research while also highlighting inconsistencies and methodological concerns, a number of future prospects may aid in taking this field forward.

4.2.1. Agnostic molecular genetic G × E studies using GWAS: GEWIS

The GWAS approach has undergone a series of modifications that allow it to be used to conduct genomic screens for environmentally responsive variants, known as Genome–Environment–Wide Interaction Studies (GEWIS) (Khoury and Wacholder, 2009). Potential advantages include the possibility to improve the targeting of interventions and treatment as well as providing new leads for understanding the wider biological context (Caspi and Moffitt, 2006). Amongst the current challenges are the questions of (i) how best to maximize sensitivity to detect true signals while minimizing the statistical penalties of a liberal approach to multiple testing (Caspi and Moffitt, 2006), and (ii) how to interpret G × E findings (Zammit et al., 2010a). To date, GEWIS have been used to study the etiology of a number of neurodevelopmental and neurological phenotypes. For example, it has been used to examine the effect of genetic moderators of the effect of coffee drinking on Parkinson’s Disease (Hamza et al., 2011). Although this innovative approach is currently one of the many long-term aspirations for policymakers in the psychiatric genetics community (Psychiatric GWAS Consortium Steering Committee, 2009), it has yet to be applied to schizophrenia research. Although the big issue for GEWIS is statistical, there is evidence that environmental risk factors cluster (Rodgers et al., 2004), hence attenuating statistical penalties on GEWIS. Furthermore, the quantification of environmental risk factors could be optimized through the use of a new generation of instruments (questionnaires) and devices that enable information on pathological exposures to be captured with greater sensitivity. An emphasis on post-hoc explorations of candidate pathways, genes and variants may help these approaches become more substantive and mechanism-revealing. Finally, a main advantage is that the biological impact of an environmental risk factor can often be studied in detail in animal models, allowing for more focused and hypothesis-driven GEWIS enquiries. Finally, a main advantage is that the biological impact of an environmental risk factor can often be studied in detail in animal models, allowing for more focused and hypothesis-driven GEWIS enquiries. Animal studies are useful in the sense that the impact of an environmental risk factor on the organism can be studied at the molecular level both in animals and humans, as well as the study of epigenetic mediation and ecogenetic moderation. "Model animals" can be a very productive strategy for G × E research when, as suggested by Insel (2010), an animal is “implanted” with a hypothesized biological factor (e.g., a CNV), then exposed to a variety of high-risk environments, and subsequently monitored for developmental, structural and functional consequences.

Thus far, the many lines of derivative research resulting from GWAS systematic findings in schizophrenia collectively demonstrate how both systematic and hypothesis-based molecular candidate approaches can work in tandem (Psychiatric GWAS Consortium Steering Committee, 2009; Thomas, 2010). GWAS has and will continue to deliver new common and rare risk loci for schizophrenia via the Psychiatric GWAS Consortium (Ripke et al., 2011). In terms of a rare variation, CNV discovery has been very important and this is likely to extend with new sequencing technologies (Xu et al., 2012). Output from these studies can then be integrated in hypothesis-based G × E models so that the effect of the genetic variations found can be measured in specific environmental risk exposures. Taking into account the environmental context may lead to a more precise estimation of the magnitude (effect size) of the risk originating from a particular variation. In fact, in the cases of “hidden” qualitative or extreme interactions, G × E will actually provide gains in power to detect these genetic effects (Zammit et al., 2010b). The effect of a genetic variation may go undetected even in large GWAS if environmental effects are not taken into account, if, for example the effect of the risk gene variation is reversed or only present in a subgroup of the sample exposed to a specific environment.

4.2.2. Biobanks

One way to address the challenge of balancing sample size and measurement error for optimal statistical benefit is to apply greater epidemiological rigor to the collection, storage and power of genetic datasets. The rapid proliferation of biobanks in biomedical research is accompanied by the expectation that this will have a positive impact on the quality of G × E research in schizophrenia and also the success rate for translation of new findings into clinical practice (Iyegbe et al., 2012). Their main functions are the processing and storage of biological samples and the collection of phenotype and other data to facilitate statistical analysis. Biobanks thus provide an infrastructure for high-quality population data that would be ideal for G × E studies.

A large number of international bodies have been created to regulate the collection, storage and power of genetic datasets, and many have overlapping functions. PHOEBE (Promoting Harmonisation Of Epidemiological Biobanks in Europe), ENGAGE (European Network of Genomic and Genetic Epidemiology) and P3G (Public Population Project in Genomics) are three examples of large biobanking consortia operating within Europe; a more comprehensive list of international organizations responsible for forging an integrated biobanking superstructure can be found in Harris et al. (2012). The use of these biobanks for G × E studies will rely on the quality of the environmental data collected. Accordingly, one of the more immediate by-products of this unprecedented cooperation between international biobanking agencies has been a new international consensus on the generation, sharing, pooling and analysis of data and samples (Guerin et al., 2010; Yuille et al., 2010; Harris et al., 2012).

4.2.3. Electronic medical records (EMR)

Clinical databases are following the footsteps of epidemiological biobanks. It will soon become much easier to harvest valuable clinical data derived from routine patient contact with clinical services, given that a switchover to EMR is underway in many geographical regions. The integrative blueprint for the digital clinical age allows a full, personalized profile of clinical, molecular (including DNA sequencing) and environmental exposure data to be compiled for each patient. The front-end portal for this serves as a personal record that can follow the individual around as they move between different mental health services. Back-end access to such data (for research purposes) is possible and necessarily anonymized (Stewart et al., 2009). Stewart et al. (2009) describe the formation of the British South London and Maudsley Biomedical Research Centre (SLAM BRC) psychiatric Case Register. The development of this case register paid careful attention
to privacy and security issues, and followed the principle of “consent or anonymize”, by which consent is not required if the data is used for other purposes than those for which it was created, as long as the data are anonymous. In fact, under UK law consent is not required for use of anonymized information. However, the Case Register allows anyone who objects to their data being used for research purposes to opt out, so that their data would not be searchable. These ethical issues, which may vary across countries, need to be considered by any project intending to use EMR, including G × E research. The true potential of the EMR model will however only be fully unlocked once high-dimensional genetic and molecular profiling becomes economically feasible, making it possible to combine clinical data with diagnostic/prognostic genetics.

4.2.4. Organizational initiatives
Collaborative efforts that bring together expertise in statistics, genetics, epidemiology, experimental psychiatry, brain imaging, and clinical psychiatry will be required to succeed in the challenging task of developing etiological models of psychosis that integrate genetic risk with environmental factors associated with the disorder. Some examples have already been devised for schizophrenia, sponsored by various national and supra-national policy-making agencies to enhance the efficiency, quality and authority of G × E studies. More specifically, the large multi-center EU-GEI study has been set out within the framework of a EU commitment to harnessing the potential of G × E research to devise and update mental health policy across the continent. Of particular importance is a work package entitled “Functional enviroomics”, which aims to reduce the current ambiguity surrounding social–environmental risk exposures in schizophrenia. More specifically, the objectives of this work package are: (1) To develop and apply methods for the detailed assessment of candidate environmental exposures, at both individual and area levels, by using an optimum, family-based, case-control design, in a diverse range of settings across Europe; (2) To investigate the impact of hypothesized environmental exposures, measured at individual and area levels, on (a) risk of schizophrenia spectrum disorders, and (b) high rates of disorder in urban centers and in migrant and ethnic minority groups; (3) To examine evidence for (a) hypothesized G × E and (b) hypothesized environment × environment interactions across the life course; and (4) To develop a translational risk assessment chart indexing environmental load.

The European Network of Schizophrenia Networks for the Study of Gene–Environment Interactions has argued that systematic attempts to identify gene–environment interactions cannot simply be equated with traditional molecular genetic studies with a number of putative environmental variables thrown in (European Network of Schizophrenia Networks for the Study of Gene–Environment Interactions, 2008), Therefore, the challenge in the years to come is to bring many disciplines together to work on the identification of gene–environment interactions.

5. Conclusion
This systematic review has identified a number of inconsistencies and methodological issues that may have limited the meaningful impact of molecular genetic candidate G × E interaction studies in the field of psychosis research. Nevertheless, some common findings across studies have emerged, especially with regard to genetic interactions with cannabis and stress. Replication studies in psychiatry are currently only rare (Duncan and Keller, 2011; Decoster et al., 2012), therefore strong and consistent replication is warranted before molecular genetic candidate G × E reports are accepted as a true reflection of how exposures and genetic background combine to alter disease risk. It is important to realize that epidemiological and molecular genetic candidate G × E research is rarely informative with regard to molecular mechanisms because it is not the final, but rather the first step that needs to be followed by targeted experimental follow-up animal and human research aimed at identifying biological mechanisms (van Os et al., 2010). In order to take this field forward, the research perspective needs to be multi-disciplinary, combining different paradigms, such as GWAS, animal studies, imaging genetics, epigenetic approaches, neuroscience, and G × E interactions.

Role of funding source
Financial support: This work was supported by the European Community’s Seventh Framework Program under grant agreement No. HEALTH-F2-2009-241909 (Project EU-GEI), DP was supported by a UK National Institute for Health Research (NIHR) fellowship (PC3022A).

Contributors
All authors were extensively involved in the conception and planning of the focus and content of this review manuscript, and in directing acquisition, analysis, and interpretation of the published literature included for review. Authors were also involved in the drafting and critical revision of the manuscript for important intellectual content and provided approval of the final submitted version of the manuscript.

Conflict of interest
All authors declare that they have no conflicts of interest.

Acknowledgements
The authors would like to thank Dr. Stanley Zammit for his advice and stimulating discussions during the preparation of the manuscript.

Appendix A. Supplementary data
Supplementary data to this article can be found online at dx.doi.org/10.1016/j.schres.2013.09.010.

References


