Patterns of symptoms in major depressive disorder and genetic analysis of the disorder using low-pass sequencing data

Submitted for the Degree of Doctor of Philosophy

Yihan (Jessie) Li

Merton College, University of Oxford

Trinity 2013

Abstract

My thesis aims at identifying both genetic and environmental causes of major depressive disorder (MDD), using a large case-control study: 6,000 Chinese women with recurrent MDD and 6,000 controls. One of the major challenges for conducting genetic research on MDD is disease heterogeneity. The first question addressed is how different MDD is from highly comorbid anxiety disorders. I examine how anxiety disorders predict clinical features of depression and the degree of heterogeneity in their predictive pattern. The second question addressed is whether clinically defined MDD is a single disorder, or whether it consists of multiple subtypes. Results are then compared with and interpreted in the context of Western studies. Furthermore, latent class analysis and factor analysis results are also used in association analysis to explore more genetically homogeneous subtypes. Genetic data were derived using a novel strategy, low pass whole genome sequence analysis. Using genotypes imputed from the sequence data, I show that a cluster of single nucleotide polymorphisms (SNPs) is significantly associated with a binary disease phenotype including only cases with ≥ 4 episodes of MDD, suggesting that recurrence might be an indication of genetic predisposition. The third issue examined is the contribution of rare variants to disease susceptibility. Again using sparse sequence data, I identified exonic sequence variants and performed gene-based analysis by comparing the number of variants between cases and controls in every gene. Furthermore I performed gene enrichment test by combining P values of SNP association tests at different minor allele frequency ranges. Overall, I did not find convincing evidence that rare variants aggregately contribute to disease susceptibility. However, the gene-based analysis resulted in an unexpected finding: cases have an excess of variants in all thirteen-protein coding mitochondrial genes, which was due to copy number differences in the mitochondrial genome. Both human phenotypic data as well as mice experimental data show that the increase in the mitochondrial copy number in cases is due to chronic stress.

Acknowledgements

I would like to express my great appreciation for my supervisor Prof. Jonathan Flint who has provided me with guidance and support throughout my entire DPhil research.

I am also immensely grateful for the scientific advice from Prof. Kenneth S. Kendler and Prof. Richard Mott. I would like to thank the 60 hospitals and nearly 300 interviewers for their assistance with the collection of the CONVERGE data as well as Dong Xie for his technical support. I am particularly grateful for the services provided by Beijing Genomics Institute including sample storage, DNA extraction and whole genome sequencing work.

I would also like to offer my special thanks for my colleagues Leo Goodstadt, Jérôme Nicod, Xiangchao Gan, Warren W. Kretzschmar, Na Cai and Amelie Baud for our stimulating scientific discussions. I would like to thank Andreas Dienst for his encouragement and support throughout the study.

Last but not least, I would like to thank my parents for their selfless decision to send me abroad to pursue further education ten years ago as well as their unconditional support throughout the whole time.

最后，我要感谢我的父母十年前送我出国留学这个无私的决定以及在这个过程中给予我的无条件支持。

Table of Contents

Chapter 1

Introduction

My thesis focuses on understanding the causes of major depressive disorder (MDD). The ultimate aim of the project is to identify genes that increase the risk for MDD by using 6,000 depression patients (cases) and 6,000 healthy controls, all of female Han Chinese origin. The project is a large international collaborative effort named the "China Oxford and Virginia Commonwealth University (VCU) Experimental Research on Genetic Epidemiology" (CONVERGE). The aim of the project is to carry out a genome wide association study (GWAS) of MDD to identify susceptibility loci, as has been achieved for many other common complex diseases. In the introduction I outline the background to the study, starting with a review of the evidence for a genetic cause for MDD and the results of linkage and association studies to date. I then review what is known about the likely non-genetic causes of MDD. Furthermore, I present specific background to each chapter, including the patterns of comorbid anxiety disorders in MDD as well as the symptom diversity in MDD. Moreover, I explain reasons for failed attempt to identify genes that are robustly associated with MDD using genome association and linkage approaches, which has led to the formation of the rare variant hypothesis. I end this chapter with an outline of the aims of my contribution to CONVERGE.

What is known about the genetic basis of MDD?

In this section I present the evidence that MDD has a heritable component, and the estimates of its heritability. I then discuss two important features of the heritability that inform some of the issues I deal with later in my thesis. Both of these features have to do with heterogeneity of genetic effects on depression, in other words evidence that among those with MDD there might be subtypes with differing genetic contributions. The first is the evidence that genetic effects operate differently in men and women. The second feature is the evidence dealing with the presence of more genetically determined subtypes of MDD.

MDD is a highly prevalent and disabling disorder. The lifetime prevalence rate of MDD is 16.2%, reported in the National Comorbidity Survey – Replication (Kessler et al. 2003). It is a disease projected to become the second leading cause of disability worldwide by 2020 (Lopez & Murray 1998). Yet, despite its prevalence, high morbidity and the relative inadequacy of current therapies, we know little about its biological basis. One thing however has been established: MDD has both genetic and environmental causes. Here I review what is known about the genetic contribution to susceptibility to MDD.

There are three ways to detect heritability: family, twin and adoption studies. Family studies report the increased risk to relatives. However this may also be due to environmental causes. Other methods, i.e. adoption or twin studies, are needed to confirm that the relationship is due to genes. Surprisingly there are no adoption studies of MDD, so I present evidence from family and twin studies.

Family study design can be conceptualised as a case control study whereas probands (cases) have MDD and comparison subjects (controls) have no history of

MDD. Cases and controls are usually matched for age and gender. The outcome variable is the prevalence of MDD in relatives, typically first-degree relatives. A meta-analysis of high quality family studies of MDD (Sullivan et al. 2000) reported an odds ratio (OR) for proband versus first degree relative status to be 2.84 (95% confidence interval (CI) = 2.31-3.49). This can be interpreted as the risk for MDD in a first-degree relative of cases to be 2.84 times that for a relative of a matched control. For comparison, family studies on schizophrenia reported an OR of 10 (Kendler & Diehl 1993), whereas the relative increase in risk in siblings of autism cases was between 30 and 100 fold depending on the population base rate (Rutter 2000).

Meta-analysis of twin studies estimated the heritability to be 37% (95% CI = 31%-42%) (Bierut et al. 1999; Kendler et al. 1995; Kendler & Prescott 1999; Lyons et al. 1998; McGuffin et al. 1996). This heritability can be inflated by non-additive genetic effects (i.e. dominance and epistasis), shared familial environment, or gene environment interaction (Visscher et al. 2008; Zuk et al. 2012). The heritability of MDD is comparable with another complex genetic disorder type II diabetes (Almgren et al. 2011) and lower than most psychiatric disorders including >90% for autistic disorders (Freitag 2006), 79% for bipolar and 82%-85% for schizophrenia, schizoaffective disorder and mania (Cardno & Marshall 1999).

One important feature about the genetic susceptibility to depression is that there is a sex difference. First, there is strong evidence that the risk of MDD is greater in women than in men (Kessler et al. 1993; Gater 1998; Weissman et al. 1993). A meta-analysis (Sullivan et al. 2000) did not find substantial differences in heritability between the two sexes. However more recently evidence has emerged that the genetic correlation between men and women is less than one, implying that different genetic loci contribute

to susceptibility to MDD in the two sexes. While there had been indications of this in the literature, the most convincing findings are those from the largest twin study. The Swedish Twin Registry (STR) study using 42,161 twins reported that the heritability of liability to MDD was significantly higher in women (42%) than in men (29%) (Kendler et al. 2006a). This was in concordance with the 44% in women and 24% in men reported in the Australian Twin Registry (ATR) (Bierut et al. 1999) and the 40% in women and 31% in men reported with measurement error in the Virginia Twin Registry (VTR) (Kendler et al. 2001). Furthermore, the genetic risk factors for major depression were moderately correlated (0.55 in VTR and 0.63 in STR) between men and women, which suggested that a substantial proportion of the genetic risk factors were sex-specific (Kendler et al. 2001; Kendler et al. 2006a).

A second important feature about the genetic susceptibility to depression is that there is some evidence indicating the existence of more genetically determined subtypes of the disorder. For many years researchers have been trying to identify a genetically determined subtype of MDD (Winokur 1979); the evidence concerns a subtype that includes early onset. An article reviewed clinical features of depression that most predict familial aggregation for MDD, including early onset, recurrence, impairment, duration, number of symptoms, symptom pattern and comorbidity (Sullivan et al. 2000). Among the seven clinical features, recurrence has predicted familial aggregation in every study investigated (Gershon et al. 1986; Kendler, Gardner, et al. 1999; Kendler et al. 1994; McGuffin et al. 1987; McGuffin et al. 1996; Weissman et al. 1986). The predictive power of age of onset on familial aggregation is problematic to compare due to different definitions in studies and some uncontrolled age effects. Despite the difficulties, early onset of MDD predicted increased familial risk in four (Cadoret et al.

1977; Mendlewicz & Baron 1981; Stancer et al. 1987; Weissman et al. 1986) out of six studies. The groups defined by early onset overlapped considerably with recurrence.

Genetic linkage and association studies

To my knowledge, ten whole genome linkage scans of MDD have been published to date (Abkevich et al. 2003; Breen et al. 2011; Camp et al. 2005, p.2005; Holmans et al. 2004; Holmans et al. 2007; Levinson et al. 2007; McGuffin et al. 2005; Middeldorp et al. 2008, p.2008; Pergadia et al. 2011; Zubenko, Maher, et al. 2003). Linkage analysis reports a logarithm of the odds (LOD) score, which compares the likelihood of obtaining the test data if the two loci are indeed linked, to the likelihood of observing the same data purely by chance. Conventionally, a LOD score greater than 3 is considered evidence for linkage, which can be interpreted as 1000 to 1 odds that the linkage being observed did not occur by chance. There are four regions repeatedly reported on a 5Mb window: chr11 75-80 Mb (Breen et al. 2011; Zubenko, Maher, et al. 2003), chr15 37-42 Mb (Camp et al. 2005, p.2005; Zubenko, Maher, et al. 2003), chr15 87-92 Mb (Breen et al. 2011; Holmans et al. 2004; Holmans et al. 2007; Levinson et al. 2007), and chr3 4-9 Mb (Breen et al. 2011; Middeldorp et al. 2008, p.200). Despite the possibility of false positive findings, linkage studies give evidence for rare and relatively penetrant variants with large effects contributing to the genetic susceptibility to MDD.

To date, there have been eight GWAS conducted for MDD with different sample sizes, phenotypic recruitment criteria and genotyping platforms from different cohorts of different countries with and without replication samples (Kohli et al. 2011; Lewis et

al. 2010; Muglia et al. 2010; Rietschel et al. 2010; Shi et al. 2011; Shyn et al. 2011; Sullivan et al. 2009; Wray et al. 2010). Additionally, there are three meta-analyses (GENDEP Investigators et al. 2013; Shyn et al. 2011; Wray et al. 2010) and one mega-analysis (Ripke et al. 2013) on combined samples to increase statistical power, with the hope of finding single-nucleotide polymorphisms (SNPs) that are above genome-wide significant threshold ($P < 5\times10^{-8}$). Among significant findings: a sample of 604 cases implicated the *HOMER1* gene ($P_{rs7713917} = 1.48E\text{-}6$) (Rietschel et al. 2010); another with 353 cases identified the neuronal transporter gene *SLC6A15* ($P_{rs1545843} = 5.53E\text{-}08$) (Kohli et al. 2011). Both findings lack replication from independent sources.

The largest and the most recent GWAS to date analysed more than 1.2 million SNPs combining 9,240 MDD cases and 9,519 controls of European ancestry in a mega-analysis sourcing samples from nine depression cohorts (Ripke et al. 2013). Results showed that no SNPs achieved genome-wide significance in the MDD discovery phase. The result suggests that common variants of median or large effect do not have main effects in the genetic architecture of MDD.

Environmental risk factors for MDD

MDD arises from the interaction between diverse environments and genetic factors (Winokur 1997). Environmental risk factors of MDD vary in their disease effects, some acting as risk throughout the lifetime (Kendler, Bulik, et al. 2000), others with shorter risk period (Kendler 1998), suggesting environmental factors also operate via more than one pathway. There is also evidence that environmental effects act differently on different subtypes of MDD (Kendler et al. 2007). Here I review what is known about the non-genetic causes of MDD. A number of environmental risk factors, including

stressful life events (SLEs), childhood sexual abuse (CSA), poor parenting have been identified to be associated with MDD.

SLEs are well established to be associated with the risk for MDD (Bebbington et al. 1984; Brown & Harris 2001; Costello 1982; Kendler, Karkowski, et al. 1999; Kendler, Thornton, et al. 2000; Kessler 1997; Paykel 1978; Surtees et al. 1986). A longitudinal study on MDD has examined the causal relationship between SLE and MDD using samples from the VTR (Kendler, Karkowski, et al. 1999). Approximately 75% of the association was causal and the remaining non-causal association was due to the fact that individuals genetically predisposed to MDD choose to put themselves into high-risk environments. Using samples from the same registry (VTR), another longitudinal study empirically proved the "kindling" hypothesis, that is, the role of environmental stressors will progressively diminish with recurrent episodes of MDD (Kendler, Thornton, et al. 2000). They used discrete-time survival, proportional hazards model and piece wide regression analyses to examine the interaction between life events exposure and the number of previous depressive episodes in the prediction of episodes of MDD. Results show that the depressogenic effect of SLEs declined with increasing episode number up to nine, and was not substantially influence by additional episodes.

It is well-established that women who report CSA are at increased risk for developing MDD in adulthood (Fergusson et al. 2008; Hulme 2011; Jumper 1995; Kendler, Bulik, et al. 2000; Neumann et al. 1996; Paolucci et al. 2001; Weiss et al. 1999). A meta-analysis including 37 studies with more than 3 million participants reported a significant association between CSA and MDD with OR 2.66 (95% CIs, 2.14-3.30) (Chen et al. 2010). Assessment of CSA in an early study (Martin et al. 1993) includes questions whether, before the subject was 16 years old, any adult or any other

older person involve the subject in any unwanted incidents, such as (1) inviting or requesting them to do something sexual, (2) kissing or hugging in a sexual way, (3) touching or fondling private parts, (4) showing their sex organs, (5) making them touch the person in a sexual way, or (6) attempting or having sexual intercourse. The responses were used to define different forms of CSA, including non-genital CSA (including sexual invitation, sexual kissing and exposing), genital CSA (including fondling and sexual touching), and attempted or completed intercourse in a study using 1,411 adult female twins from the VTR (Kendler, Bulik, et al. 2000). The study reported that any type of self-report CSA increased the risk for MDD almost two fold (OR = 1.93, 95% CIs = 1.52-2.44), and the most severe type of CSA involving intercourse increased the risk more than three fold (OR = 3.14, 95% CIs = 2.08-4.75) (Kendler, Bulik, et al. 2000). A number of studies demonstrated this "dose-response" relationship between CSA and MDD, that is, risk for MDD increased monotonically with the severity of CSA (Fergusson et al. 1996; Kendler, Bulik, et al. 2000; Mullen et al. 1993). Particular attention has been paid to long term effects of CSA on hypothalamic-pituitary-adrenal axis function, which suggested to be the underlying biological mechanism of earlier stressors causing later MDD in adulthood (Hulme 2011; Weiss et al. 1999).

Poor parenting has been considered an important risk factor for adult psychopathology, in particular MDD (Burbach & Borduin 1986; Gerlsma et al. 1990; Holden & Edwards 1989; Kendler, Myers, et al. 2000; Kendler et al. 1997; Parker 1979; Parker et al. 1995; Parker 1983; Perris et al. 1986; Perris et al. 1994). The Parental bonding instrument (PBI) is widely used to assess different parenting dimensions. The PBI was initially developed by Parker (Parker et al. 1979; Parker 1979; Parker 1989)

and later revised by Kendler to include 16 items which gave rise to three factors: parental warmth, protectiveness and authoritarianism (Kendler 1996). The three dimensions of parenting including coldness (reversely coded from warmth), protectiveness and authoritarianism were used to predict the risk for depression using the VTR samples (Kendler, Myers, et al. 2000). When the three dimensions were examined independently, high levels of coldness and authoritarianism were significantly associated with MDD with ORs less than 1.5, whereas when the three dimensions were examined together in both parents, all the associations were solely with coldness, especially from father (OR = 1.4)

Neuroticism: the role of personality as a risk factor for depression

The personality trait of neuroticism was initially designed to measure emotional stability (Eysenck & Eysenck 1975). This personality trait can be characterised by anxiety, moodiness, worry, envy and jealousy (Thompson 2008). The measurement of neuroticism using Eysenck Personal Questionnaire (EPQ) has been proven to be stable over the past two decades (Wray et al. 2007). Neuroticism is reported to be associated with an increased risk for MDD (Angst & Clayton 1986; Hirschfeld et al. 1989; Kendell & Discipio 1968; Kendler et al. 1993; Wetzel et al. 1980). Twin studies reported that the association between neuroticism and MDD resulted largely from shared genetic risk factors (Fanous et al. 2002; Kendler et al. 2006b).

Broad-sense heritability estimate for neuroticism is reported to be 43% using approximately 9,000 twins (Distel et al. 2009). The largest GWAS of neuroticism to date using a combined sample of 17,375 adults did not report any above genome wide significant threshold SNPs (de Moor et al. 2012). Using GWAS data for neuroticism on

13,835 individuals, a study examined the genetic association between neuroticism and MDD via polygenic score analysis (Middeldorp et al. 2011). Polygenic risk score was calculated by aggregating results from SNP association analysis on neuroticism. This score was then regressed against MDD phenotype. Results showed that polygenic neuroticism scores were significantly positively associated with MDD.

Is depression the same as anxiety?

The currently used criteria for diagnosing an episode of depressive illness arose from a long debate about the nosological status of affective disorders, a debate that concerned the three related problems of diagnostic interface, subtypes and boundaries. In other words, how separate is depression from other disorders (particularly anxiety disorders, and manic depressive psychosis)? Is depression one disorder? Is depression a discrete syndrome, or is it merely diagnostic convention that determines who is ill or who is not? These questions are not merely of nosological interest, since they will inform both the design of experiments, and the interpretation of research findings, including genetic analyses.

MDD frequently overlaps diagnostically with anxiety disorder. A large number of studies have reported the high prevalence rates of comorbid anxiety disorders among MDD patients ranging between 40-70% with different sampling frames and diagnostic criteria in the US and European population (Brown et al. 1996; Brown et al. 2001; Clayton et al. 1991; de Graaf et al. 2003; Fava et al. 2000; Fava et al. 2004; Hasin et al. 2005; Howland et al. 2009; Kessler et al. 1994; Kessler et al. 1996; Kessler et al. 2003; Kessler et al. 2005; Lamers et al. 2011; Melartin et al. 2002; Merikangas et al. 2003; Pini et al. 1997; Sanderson et al. 1990; Zimmerman et al. 2000; Zimmerman &

Chelminski 2003). An article published in 2013 reported the prevalence rate of comorbid anxiety disorder (69.9%) diagnosed with Hamilton Rating Scale for Depression (HRSD)-17 anxiety / somatization factor score ≥ 7 among 375 patients diagnosed with the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV MDD recruited through clinical settings in China (Wu et al. 2013). The paper cited another article (published in Chinese), which reported similar prevalence rate (68.9%) of comorbid DSM-IV anxiety disorder among 508 MDD patients collected through 15 psychiatric facilities in China (Shi et al. 2009). A cross-national epidemiological study has published figures of comorbid panic disorder across eight countries with 5.9% reported in Taiwan (of the most comparable ethnic and cultural background as Han Chinese) which is in the low end compare to the 17% reported in Korea, 16.2% in Germany and 15.3% in Puerto Rico (Weissman et al. 1996). This shows that the comorbidity between MDD and anxiety is universal.

Among different anxiety disorders, generalised anxiety disorder (GAD) is most closely related to MDD. A longitudinal study reported a 37% of cases with GAD preceding MDD, and 32% vice versa (Moffitt et al. 2007). The high prevalence rate and pattern of precedence suggests that the two disorders are very closely related.

Another way of answering the question of whether depression and anxiety are the same disorder is via a genetic approach. Two review articles presented strong evidence for a shared genetic vulnerability between depression and anxiety from twin studies (Cerdá et al. 2010; Middeldorp et al. 2005). Using 5,600 twin pairs in the VTR, a common genetic factor contributing to three internalizing disorders was identified, with MDD and GAD contributing 54% and 53% respectively and phobia contributing 33% (Kendler et al. 2003). Within five internalizing disorders, there were two common

genetic factors identified with the first primarily loaded on MDD and GAD and the second primarily loaded on animal and situational phobia (Kendler et al. 2003). Panic shares more genetic risk with the first common factor. It suggests that depression shares more genetic risk with GAD than panic and phobia, with the latter showing the least genetic correlation. Another study confirmed genetic correlation between depression and anxiety using SNP data to generate genetic risk scores (Demirkan et al. 2011).

One way of helping to answer the question of whether the two disorders are the same is by looking the predictive pattern of each anxiety disorder on the clinical features of depression. It was previously reported that MDD patients with comorbid anxiety disorders generally have a more severe course with greater severity and chronicity, increased suicidal risk (Brown et al. 2001) and poorer treatment response (Brown et al. 2001; Howland et al. 2009), as well as more childhood trauma and higher neuroticism (Lamers et al. 2011). They also have lower social competence (Cerdá et al. 2010) and are more likely to have a family history of MDD (Clayton et al. 1991; Husain et al. 2009; Kendler et al. 1992).

A number of studies have examined the comorbidity between MDD and a single anxiety disorder, including GAD (Gorwood 2004; Hettema 2008; Kendler et al. 1992; Kendler et al. 2007; Kessler et al. 1999) and panic disorder (Cassano et al. 1989; Coryell et al. 1988; Garvey et al. 1987; Grunhaus et al. 1994; Grunhaus 1988; Kessler et al. 1998; Maier et al. 1995; Pini et al. 1994; Reich et al. 1993; Weissman et al. 1993), the conclusion of which generally agree with the theory of shared genetic risk and increased depression severity.

Although, there is evidence for a high degree of shared genetic risk between depression and anxiety, it is possible that some of the genetic susceptibility work for

depression alone and some work on depression plus anxiety. Due to the high comorbidity between depression and anxiety, it is unrealistic to exclude all depression patients with a history of anxiety. Therefore, it is important to know the way in which depression and anxiety are related before we carry out GWAS, which also influences our study design, as it requires building a detailed anxiety symptom profile for each depression patient.

In chapter 3, I will present results of how the seven anxiety disorders predict depression related clinical features and environmental risk factors. Moreover, I will examine the degree of heterogeneity of those seven anxiety disorders in predicting those features.

Is depression one disorder?

The view that depression is one disorder has a long history and has engendered considerable debate. One of the most influential papers in this regard was published in Science in 1973 (Akiskal & McKinney 1973). This paper attempted to unite the diverse views of the causes of depression and show that the phenotype was the outcome of a common final pathway. By conceptualizing the psychoanalytic model of depression as object loss (a proximal cause of depression) as "loss of reinforcement", or loss of control over reinforcement, the authors proposed that experimental investigation in animal models of depression was relevant. They described anatomical, biochemical and pharmacological data as a process occurring in the diencephalic centres of reward (principally the medial forebrain bundle and periventricular system). Depression was seen as a final common pathway, a decrease in the functional capacity of the reward system.

Alternatives to the unitary position have also been proposed, with attempts to differentiate MDD along dimensional, hierarchical, or categorical axes (Eysenck 1970) (Parker 2000). Despite continuing discussion of the value, or otherwise, of identifying subtypes (Cole et al. 2008; Farmer & McGuffin 1989; Grinker et al. 1961; Kendell 1976; Klein DF 1974; Lewis 1934; MacFadyen 1975; Parker 2000), the unitary concept of depression remains the consensus, and primarily the focus of genetic studies. The importance of this debate is the extent to which genetics might resolve it, and potentially guide interpretation of the underlying pathogenic mechanisms.

Here I review evidence that indicates there may be meaningful subtypes of MDD. One of the main research interests has been attempts to identify subtypes of depression based on the coherence of groups of symptoms. Despite limited understanding of the pathophysiology of the disorder, clinicians find that the same symptoms frequently occur together, and this observation is largely responsible for the choice of the nine symptoms that constitute the current DSM-IV nine MDD A criteria, which is widely used as a manual for doctors to diagnose mental disorders in clinical practice.

However the DSM criteria clearly do not capture all depression related symptoms. Hence there are many instruments designed in research to examine a broader spectrum of depression symptoms based on different assumptions - some aim at self-assessed population samples, some require doctor's supervision in clinical setting, some focus on cognitive symptoms, and some include positive symptoms (Shafer 2006). These instruments vary in the number and nature of the symptoms included as well as the types of responses - some collect binary responses and some allow varying degrees of symptom severity.

In order to analyse these instrumental data, statistical methods such as factor analysis (FA) and latent class analysis (LCA) were used to determine the latent symptom dimensions and patient subtypes respectively. There are two types of factor analysis: exploratory and confirmatory, the former operates by reducing the number of items assessed based on the correlations between the items whereas the latter specifies the items loaded on particular factors, forcing the others to have zero loadings on that particular factor. Exploratory FA (EFA) is designed for discovering the latent factor structure; whereas confirmatory FA (CFA) was based on prior knowledge of the potential underlying latent structure of the items. However, it is based on the assumption of multivariate normality to justify a linear model where variables are required to be continuous and normally distributed. When this assumption is violated, such as when feeding it with dichotomous variables, non-informative extraneous factors often occur which possibly due to the various nonlinearities in the model (Magidson & Vermunt 2003). Additionally, FA allows different rotation options, which could result in quite different item loadings making it difficult to interpret (Magidson & Vermunt 2003). LCA, on that other hand, is designed to cluster patients, rather than symptom items, into mutually exclusive and relatively homogeneous groups, based on measure of similarity between individuals across all items examined in the analysis. It was designed to take multivariate categorical data. Although both methods are analogous, I am not aware that they were ever applied to the same psychometric datasets for comparison. Below I summarise what has been discovered for the symptom patterns of depression on different datasets using FA and LCA.

The DSM-IV MDD nine A criteria were aimed at discriminating people with clinical depression from healthy individuals, whose uni-dimensionality was later

supported by empirical evidence from factor analysis using population samples (Aggen et al. 2005; Muthén 1989). A prior analysis from Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD) has detected evidence for heterogeneity within the DSM criteria for MDD between somatic and cognitive criteria (Lux & Kendler 2010). No FA or LCA studies have used the same 14 or 27 items on clinical samples. I will review the literature using different instruments.

A review of ten articles using factor analysis on clinical samples identified between two and four factors, among which four common factors were identified in more than two studies, including the general severe / retarded factor, the somatic factor, cognitive impairment factor, and agitation / anxiety factor (Bech et al. 2011; Beck et al. 1961; E et al. 1967; Friedman et al. 1963; Grinker 1961; Hamilton 1960; Parker 2007; Romera et al. 2008; Shafer 2006; van Loo et al. 2012). A review of 16 LCA articles identified between two and seven clinical classes commonly differed in the levels of severity (Andreasen & Grove 1982; Carragher et al. 2009; Davidson et al. 1989; Davidson et al. 1988; Eaton et al. 1989; Grove et al. 1987; Haslam & Beck 1993; Hybels et al. 2009; Kendler et al. 1996; Lamers et al. 2012; Lamers et al. 2010; Maes et al. 1992; Parker et al. 1999; Schotte et al. 1997; Sullivan et al. 1998; Sullivan et al. 2002). However, the design of these studies differ significantly in a number of ways, including clinical settings and patient recruitment criteria, depression instruments and items included, and sample size. Due to the differences in study design and the sensitivity of the statistical methods to data fed to it, it is impossible to generalise the findings.

In chapter 4, I will present results using FA and LCA on three symptom datasets, with the aim to understand depression symptom dimensions and patient

subtypes. This information together with genetic data will help us to determine whether there are subtype specific SNPs that are buried in GWAS of diseased cases versus controls.

Why GWAS failed

What might be the reason for the failed attempt to identify causal SNPs for MDD in those GWAS studies? Heritability for MDD reported by twin studies is 37%, which is about twice (18%) of what has been reported by SNP-based heritability estimated on case-control scale (Cross-Disorder Group of the Psychiatric Genomics Consortium 2013). Despite the missing heritability explanations (Eichler et al. 2010; Manolio et al. 2009), there is without doubt a genetic component to it. One of the reasons was due to disease heterogeneity, that is, different depression subtypes and environmental risk factors might dilute the power to detect genetic effects. Other reasons lie in the genetic architecture of the disorder and current technological limitations.

GWAS was based on the so called "common disease, common variant" hypothesis, that is common diseases are attributable in part to allelic variant present in more than 1-5% of the population (Manolio et al. 2009). Benefiting from the joint international effort of the International HapMap (Altshuler et al. 2010; Sabeti et al. 2007; Anon 2003) and 1000 Genomes Project (1KGP) (Abecasis et al. 2012), geneticists were able to unravel the genetic structure of human deoxyribonucleic acid (DNA) sequence and construct a haplotype map of sequence variants. Using this valuable knowledge, companies produced genotyping microarray chips to capture common SNPs in order to help scientists test this "common disease common variant" hypothesis. The fact that the so-far-largest GWAS of MDD with more than 18,000

samples and 1.2 million SNPs failed to identify any above genome-wide significance SNPs is disappointing, yet may have taught us a very important lesson.

Power analysis demonstrated that current sample size used for GWAS on MDD does not have enough power to detect SNPs of small effects (with genotypic relative risk (GRR) < 1.2) (Spencer et al. 2009). The largest GWAS was conducted on human height using sample size of 183,727 (Lango Allen et al. 2010). It is claimed that 100,000 schizophrenia case control samples is needed to achieve the same power as the largest published GWAS on height (Visscher et al. 2012) using a method described in a published article (Yang et al. 2010). The fact that a GWAS on MDD using a sample size of over 9,000 cases failed to detect any significant loci makes it very unlikely that common variants with an OR greater than 1.2 contribute to the genetic susceptibility of depression. Furthermore, if it is mostly the common SNPs (SNPs that are present in >1-5% of the population) that are causal for the disorder, with large enough sample size, individual genetic effects of those to-be-detected common SNPs are going to be very small, and most likely to be smaller than the most significant SNP reported in the largest GWAS finding ($OR_{rs1969253}$ = 1.101) (Ripke et al. 2013).

Rare variant hypothesis

So far both genetic linkage (Abkevich et al. 2003; Breen et al. 2011; Camp et al. 2005, p.2005; Holmans et al. 2004; Holmans et al. 2007; Levinson et al. 2007; McGuffin et al. 2005; Middeldorp et al. 2008, p.2008; Pergadia et al. 2011; Zubenko, Maher, et al. 2003) and association approaches (Kohli et al. 2011; Lewis et al. 2010; Muglia et al. 2010; Rietschel et al. 2010; Shi et al. 2011; Shyn et al. 2011; Sullivan et al. 2009; Wray et al. 2010) have failed to identify genes that are robustly associated with depression.

The limitation of both methodological approaches can be interpreted in the context of sample size, effect size, allele frequency and statistical power. As mentioned above, GWAS using sample size in the thousands can detect common variants (minor allele frequency (MAF) > 5%) with modest to large effects (OR > 1.2), but cannot detect common variants with small effects (OR < 1.1) or rare variants (MAF < 5%) with large effects (McCarthy et al. 2008). It is realistic to expect linkage analysis to detect variants with high GRR (> 4) at intermediate allele frequencies (MAF = 5%-50%). However, this approach effectively excludes variants with GRR greater than 2 or rare variants (MAF < 5%) with large effects, due to the unrealistic family sizes needed to detect such variants (Risch 2000; Risch & Merikangas 1996).

Recent sequencing projects show that coding regions contain more low-frequency deleterious mutations than previously anticipated and that common loss of function variants rarely contribute to complex disease (Li et al. 2010; MacArthur et al. 2012; Nelson et al. 2012; Tennessen et al. 2012). Hence there remains the possibility that the individually specific rare variants with large effects aggregately contribute to the genetic cause of MDD.

The presence of such variants has already been documented in other quantitative traits and diseases. For example, sequencing individuals at the extremes of the distributions of quantitative phenotypes has found rare alleles with major phenotypic effects. Sequencing candidate genes in individuals with low (<fifth percentile) and high serum lipids (HDL-C) revealed an enrichment of non-synonymous (NS) variants (Cohen et al. 2004). Sequencing of candidate genes has also found rare variants contributing to susceptibility to high blood pressure (Ji et al. 2008). Sequencing genes that lie within loci identified by GWAS has also found rare mutations with large effects,

as demonstrated by success with type 1 diabetes (Nejentsev et al. 2009) and Crohn's disease (Rivas et al. 2011). The importance of rare, large effect alleles is now well established in psychiatric disease where cases of schizophrenia (Walsh et al. 2008; Xu et al. 2008), autism (Sebat et al. 2007) and bipolar disorder (Malhotra et al. 2011) show enrichment, relative to controls, of structural variants (SVs). More recently, sequencing of patients with severe mental retardation (Vissers et al. 2010), schizophrenia (Girard et al. 2011; Xu et al. 2011) and autism spectrum disorders (O'Roak et al. 2011) has shown that rare deleterious mutations are enriched in cases compared to controls.

Alternative to the study design of deep exome sequencing using sample sizes in the hundreds, which proved to work as mentioned above, I adopted the approach of using low coverage sequence data from ten thousand individuals. The value of low coverage sequence was investigated by simulations, which showed that low coverage sequencing (down to 0.1-0.5X) capture almost as much of the common (MAF > 5%) and low frequency (MAF = 1-5%) variation across the genome as SNP arrays. This makes low coverage sequencing a viable alternative for GWAS (Pasaniuc et al. 2012).

All GWAS on MDD were conducted using genotyping data from designated array chips. If rare variants were the major genetic cause for MDD, then conducting SNP association using genotyping array data would not give us the answer even with larger sample size. Next-generation sequencing technology combined with cutting edge bioinformatics tools gives us a unique opportunity to identify those potential causal variants, however faces analytical challenges and has cost implications.

Thesis overview

In my thesis, I will first describe the nature of the CONVERGE data, including sample recruitment criteria and collection process as well as the bioinformatics pipeline for the GWAS and rare variant analysis, which will be the full body of the method chapter. From chapter 3 to 7, I present results of both phenotypic and genetic analysis using the CONVERGE data. In chapter 3, I describe how the seven comorbid anxiety disorders independently and jointly predict clinical features of depression as well as the degree of heterogeneity within the seven comorbid anxiety disorders in their predictive power to those clinical features. In chapter 4, I identify symptom dimensions and latent patient classes using FA and LCA methods respectively in the case sample. In chapter 5, I present the results from gene-based analysis as well as gene enrichment test based on the rare variant hypothesis focusing on the exome of whole genome sequencing data. In chapter 6, I present the preliminary findings from genome wide association analysis using imputation data on not only disease status but also depression related phenotypes including neuroticism as well as disease status with all controls and subsets of cases partly derived from FA and LCA in chapter 4. The gene-based analysis described in chapter 5 excludes findings from the mitochondrial (MT) genome. In chapter 7, I explain the finding of significantly more variants called for all thirteen protein coding MT genes in the cases compared to controls. Finally, I discuss implications of each of the results chapters in the last chapter of my thesis.

Chapter 2

Methodology

Study design

Data for the present study were drawn from the CONVERGE study of MDD. The project utilizes samples collected for a molecular genetic study of MDD from 58 mental health centres and psychiatric departments of general medical hospitals in 45 cities in 23 provinces. Chapter 3 uses 6,207 cases for comorbid anxiety analysis. Chapter 4 uses 6,008 cases for factor analysis (FA) and latent class analysis (LCA). The discrepancy in sample size between chapter 3 and 4 is due to different sample availability at the time of the analysis. The total number of samples used for the rare variant analysis in chapter 4 is 9,063 (4,287 cases + 4,776 controls). The number of samples used for genome wide association (GWA) analysis on disease status and mitochondria read depth is 11,651 (5,867 cases + 5,784 controls). However, please be aware that 300 samples were excluded in the GWA analysis due to genetic relatedness and minority ethnicity.

All cases and controls were initially identified by Chinese doctors in each centre by a brief interview. Cases were obtained amongst patients attending each of the participating centres. Given evidence that the genetic effects on MDD are different between men and women (Kendler et al. 2001), in order to control for this known genetic heterogeneity, we collected female participants only, with four Han Chinese grandparents. Other detailed case control entry and exclusion criteria are listed in Table 2.1.

Originally, we aimed at obtaining age-matched (+/- 5 years) controls from the non-consanguineous relatives of the cases, which turned out to be unfeasible. Therefore, this particular criterion was relaxed to controls between the age of 40 and 60 that were not related to cases, with the general principle that they were representative of the general population. Psychiatric department in general hospital could recruit surgical or gynecological controls excluding: i) procedures that were related to risk for psychiatric (i.e. all treatments related to traumatic injury or self-induced injury) or ii) substance abuse disorders (i.e. any surgical treatment related to alcoholism or drug abuse), and iii) cancer-related procedures because of high risk for depression. Surgical patients undergone the following procedures could be included as controls: i) surgical removal of organs (especially appendix, gall bladder, uterus, ovaries, colon that is not related to cancer); ii) simple repair operations (i.e. hernia or knee); iii) joint replacements; iv) cardiac procedures including by-pass, stent, valve replacement; v) removal of kidney stones; vi) bladder repair or other urological surgery unrelated to cancer; vii) arthroscopic surgery; or viii) plastic surgery not cancer related.

Eligible individuals were asked for their informed consent to participate in a face-to-face interview to take place at the study centre. Data from the initial interview was entered onto one of the study computers. Any individual or control identified clinically were assigned a study number (patient ID) and asked to attend a detailed phenotypic investigation.

Phenotypic assessment had two components: a diagnostic interview and self-administered questionnaires. The computer assisted personal interview guided the interviewer through the respondent selection process, delivered the questions and directed the interviewer through the questioning sequence. Three assessments were

administered as self-reported paper questionnaires (childhood sexual abuse, neuroticism and Big Five Inventory for controls only). We provided paper versions of questionnaires that could be downloaded and printed out from either the assessment computer or our website. After completion, questionnaire data were entered directly onto the assessment computer. It was also possible to administer these questionnaires as computer assisted interviews. We also provided paper versions of all questionnaires, so that, if necessary, interviewers could use these versions.

All interviewers, who were postgraduate medical students, junior psychiatrists or senior nurses, were trained by the CONVERGE team for a minimum of one week. The interview includes assessment of psychopathology, demographic and personal characteristics, and psychosocial functioning. Interviews were tape-recorded and a proportion of them were listened to by the trained editors, who provided feedback on their quality. The study protocol was approved centrally by the Ethical Review Board of Oxford University and the ethics committee in participating hospitals in China.

	Entry Criteria
Case	Female
Case	Age 30 – 60
Case	Two or more episodes of MD
Case	Age of first MD episode >= 14 and < 50
Case	Four grandparents are Han Chinese
Control	Female
Control	Age 40 – 60
Control	Non blood relative of case or friend from same area
Control	No episode of depression
Control	Four grandparents are Han Chinese
	Exclusion criteria
Case & control	History of bipolar illness
Case & control	History of any type of psychosis
Case	Drug or alcohol dependence with onset before depression
Case & control	Mental retardation

Table 2.1: Entry and exclusion criteria for cases and controls.

Measures and instruments

The diagnoses of depressive disorders (Dysthymia and MDD) and anxiety disorders (Generalized Anxiety Disorder (GAD), Panic Disorder with or without Agoraphobia) were established with the Composite International Diagnostic Interview (CIDI) (WHO lifetime version 2.1; Chinese version), which classifies diagnoses according to the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria. The CIDI interview was supplemented by items from the Structured Clinical Interview for DSM-III-R Diagnosis (SCID) (Spitzer et al. 1987).

Phobias, divided into five subtypes (animal, situational, social and blood-injury, and agoraphobia), were diagnosed using an adaptation of the DSM-IV criterion D (American Psychiatric Association 1994) requiring one or more unreasonable fears, including fears of different animals, social phobia and agoraphobia, that objectively interfered with the respondents' life or resulted in severe anxiety when exposed to the phobic stimulus.

The diagnoses of MDD were based on information collected in the medical records and DSM-IV symptoms during the worst episode. Symptoms information collected for the diagnoses of comorbid disorders of MDD were of lifetime.

The history of lifetime major depression in the parents and siblings was assessed using the Family History Research Diagnostic criteria (Endicott et al. 1975). Neuroticism was measured with the 23-item Eysenck Personality Questionnaire (EPQ) (Eysenck & Eysenck 1975).

Other instruments including premenstrual syndrome, postnatal depression, parent child relationship, stressful life events, social life, childhood sexual abuse were

employed from the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD) (Kendler & Prescott 2006).

The interview was originally translated into Mandarin by a team of psychiatrists in Shanghai Mental Health Centre, with the translation reviewed and modified by members of the CONVERGE team.

All interview sections were fully computerized into a bilingual system of Mandarin and English developed in-house in Oxford, called SysQ. Skip patterns were built into SysQ. Interviews were administered by trained interviewers and entered offline in real time onto SysQ, which was installed in laptops. Once an interview was completed, a backup file containing all the previously entered interview data could be generated with a database compatible format. The backup file together with audio record of the entire interview was uploaded to a designated server, currently maintained in Beijing by a service provider. All the uploaded files in the Beijing server were then transferred to an Oxford server quarterly. The diagnoses of MDD were based on information collected in the medical records and DSM-IV symptoms during the worst episode. The proposed assessment protocol is listed in table 2.2.

Interview sections	Case control	Assessment method
Demographics	Different case control version	Questionnaire
Mania	Case only	CIDI Interview
Psychosis	Case only	CIDI Interview
Major Depression	Different case control version	CIDI Interview
Dysthymia	Case only	CIDI Interview
Generalized Anxiety Disorder	Case only	CIDI Interview
Panic	Case only	CIDI Interview
Phobia	Case only	CIDI Interview
Smoking	Case only	Fagerstrom questionnaire
Family History	Case only	Questionnaire
Stressful Life Events	Both	Questionnaire
Parental Bonding Instrument	Both	Questionnaire
Social Life	Both	Questionnaire
Postnatal Depression	Case only	Questionnaire
Premenstrual Syndrome	Both	Questionnaire
Neuroticism	Both	EPQ paper questionnaire (paper)
Childhood Sexual Abuse	Both	Questionnaire (paper)
Big Five Inventory	Control only	Questionnaire (paper)

Table 2.2: Assessment protocol for interview. EPQ, Eysenck Personality Questionnaire. CIDI, Composite International Diagnostic Interview.

Statistical methods

Generalized linear regression model and ANOVA

In order to test whether comorbid anxiety disorders significantly predict various MDD features, including family history, age of onset, number of episodes, number of DSM-IV diagnostic criteria, suicidal attempt and neuroticism score, generalized linear models (GLMs) with different *link* functions were chosen for modelling different types of dependent variables controlling for age. Data for the age of onset was tested for normality and modelled with linear regression. Suicidal attempt was modelled with logistic regression and results were represented as odds ratio (OR). The family history variable, represented by the total number of first-degree family members among the two parents and all the siblings with a history of MDD, could be considered as a series of Bernoulli trials with a finite number of trials in statistical terms; hence binomial regression was used for modelling the association, weighted against the total number of

family members for each subject. The same reason and choice was applied to variables such as the number of DSM-IV diagnostic criteria and the neuroticism score, which could also be considered as a finite number of Bernoulli trials. Poisson regression was chosen to model the number of episodes data as it is considered to be count data with an upper limit of 96 (denoting 96 or more) being removed. Since binomial and Poisson regressions model the natural logarithms of the count and the event rate respectively, results were represented by count ratio (CR) and rate ratio (RR), both calculated by exponentiating the regression coefficients. The statistical tests were implemented using R (R Development Core Team 2011), which allows the specification of quasibinomial and quasipoisson link functions. They differ from binomial and poisson functions only in that the dispersion parameter is not fixed at one; in other words, over-dispersion is included in the model, which controls for the over-reporting of false positive results.

In addition to calculating the individual effects of the seven comorbid anxieties, the method of Analysis of Deviance for GLM Fits (anova) was used to examine whether the anxiety regression coefficients in the multiple GLMs were homogeneous or heterogeneous in predicting each of the MDD features. Two GLM objects were specified in anova; one with a derived independent variable which is calculated as the sum of all seven anxiety diagnostic status, nested by another containing all the seven comorbid anxiety diagnoses as additive independent variables. The result gives a table with a row for the residual degrees of freedom and deviance for each model. Test statistics could be specified in anova, comparing the reduction in deviance for the row to the residuals. For models with dispersion estimated by moments, such as Gaussian, quasibinomial and quasipoisson fits, F test is considered to be the most appropriate.

Factor analysis

FA was used to examine the latent symptom dimensions for major depression. The total sample was randomly divided into two halves. The first sample was used to perform an exploratory FA (EFA) and the second was used to perform a confirmatory FA (CFA) for validating the EFA symptom structure. Factor analyses were performed using the Mplus program (Muthén & Muthén 1998), with weighted least squared means and variance adjusted (WLSMV) estimator that is designed for ordinal data (Flora & Curran 2004). EFA was performed using a geomin oblique rotation. CFA model fit was evaluated using the Tucker-Lewis Index (TLI) (Tucker & Lewis 1973), the Comparative Fit Index (CFI) (Bentler 1990) and the root mean square error of approximation (RMSEA) (Steiger 1990). For the TLI and CFI, values between 0.90 and 0.95 are considered acceptable, and ≥ 0.95 as good. For the RMSEA, good models have values ≤ 0.05. To confirm that the split-sample procedure produced two random subsamples, we formally tested in Mplus for measurement invariance in our CFA data on the 9 DSM-IV criteria. When we constrained factor loadings to be identical in the two subsamples, the robust χ^2 difference, as expected, was not significant.

Latent class analysis

LCA was performed to examine the topologies of major depression symptoms. Mplus (Muthén & Muthén 1998) was used for running LCA (McCutcheon 1987). A number of criteria were used to guide the decision on the number of classes in mixture modeling, including Akaike Information Criteria (AIC) (Akaike 1987), Bayesian Information Criteria (BIC) (Schwarz 1978), sample-size adjusted BIC, the Vuong-Lo-Mendell-Rubin test (requested using TECH 11 in Mplus), and the bootstrapped parametric

likelihood ratio test (requested using TECH 14 in Mplus). Both tests compared the model of the currently chosen number of classes (K) to a model of K-1 classes. The analysis was performed by fitting a one-class model and gradually increased the number of classes one at a time for model comparison, setting random starting value from 500 to 100 arbitrarily.

Rare variant bioinformatics pipeline

The steps for processing low coverage sequencing data as well as the bioinformatics pipeline for extracting exome data were as follows: 1) DNA was extracted from saliva samples preserved using Oragene kits (Nunes et al. 2012) and sequenced using illumina HiSeq 2000 (http://www.illumina.com/systems/hiseq_2000_1000.ilmn) at Beijing Genomics Institute; 2) sequenced reads were then aligned to the human reference genome (hs37d5.fa released in Jan 2012 by the 1000 Genomes Project) using Stampy (Lunter & Goodson 2011); 3) piling up sequenced reads with filtering options on base quality, mapping quality, and non-pair end reads using SAMtools (Li et al. 2009); 4) aggregating variants in the exonic intervals of each gene; 5) indels filtering and SNP detection; 6) SNP annotation using ANNOVAR (Wang et al. 2010); 7) aggregating different types of SNPs in exonic genes based on different alternative allele counts (proxy for MAF) cut-offs.

The bioinformatics pipeline (Figure 2. 1) was developed using Ruffus (Goodstadt 2010) as a wrapper for all the functional steps from *samtools mpileup* (step 3), which I will explain in more detail. Read alignments were filtered on mapping quality less than 55 (Figure 2. 2a) as well as taking only properly paired reads using *samtools view*. Reads would only be included if the read is paired in sequencing and the

read is mapped in a proper pair. Reads would be excluded if the query sequence itself is unmapped, the mate is unmapped, the alignment is not primary, the reads failed platform / vendor quality checks, or the read is either a polymerase chain reaction (PCR) or an optical duplicate.

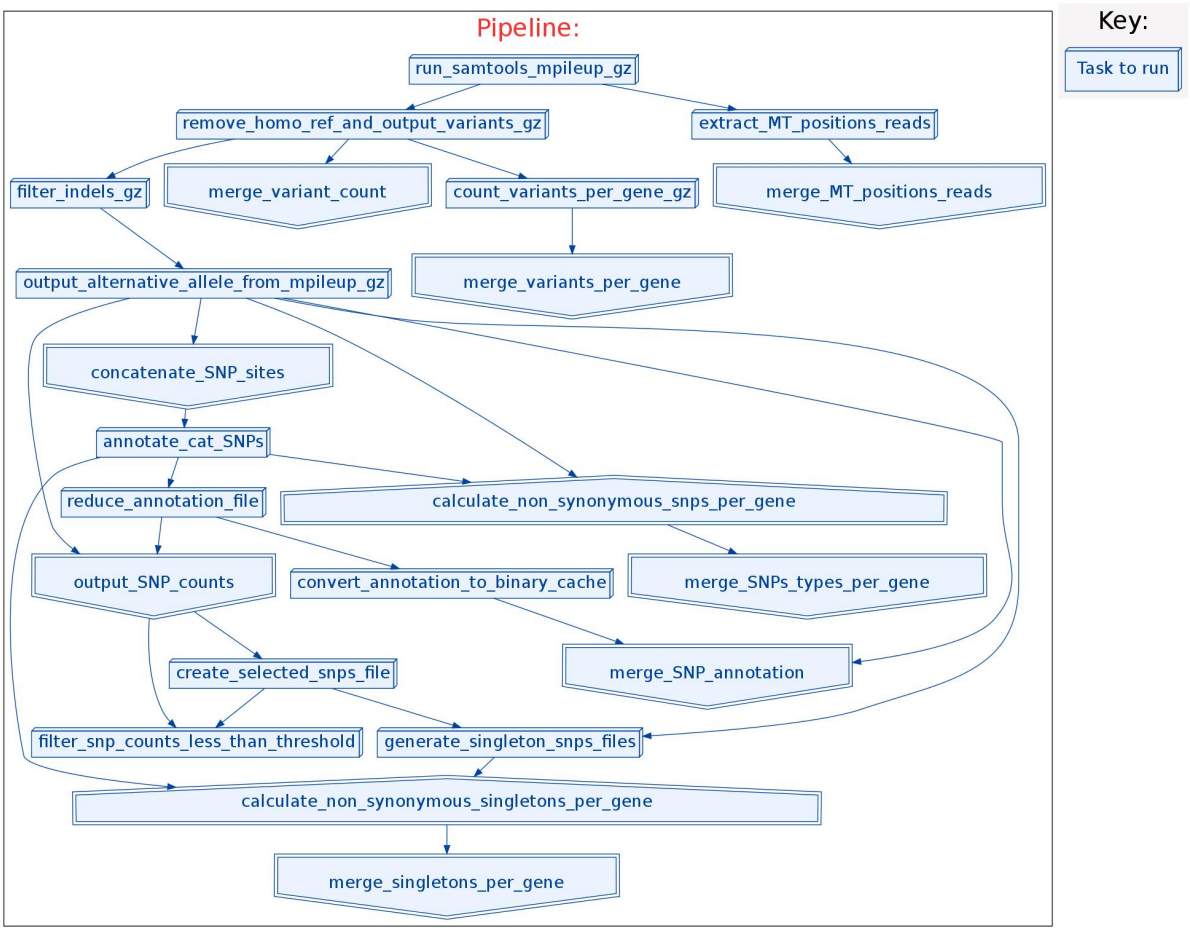


Figure 2. 1: Rare variant bioinformatics pipeline. The pipeline takes sequencing bam files as input and outputs summary statistics and data matrices of the number of exonic variants and different types of annotated SNPs per gene per individual with different alternative allele count thresholds (proxy for minor allele frequency) for subsequent association analysis.

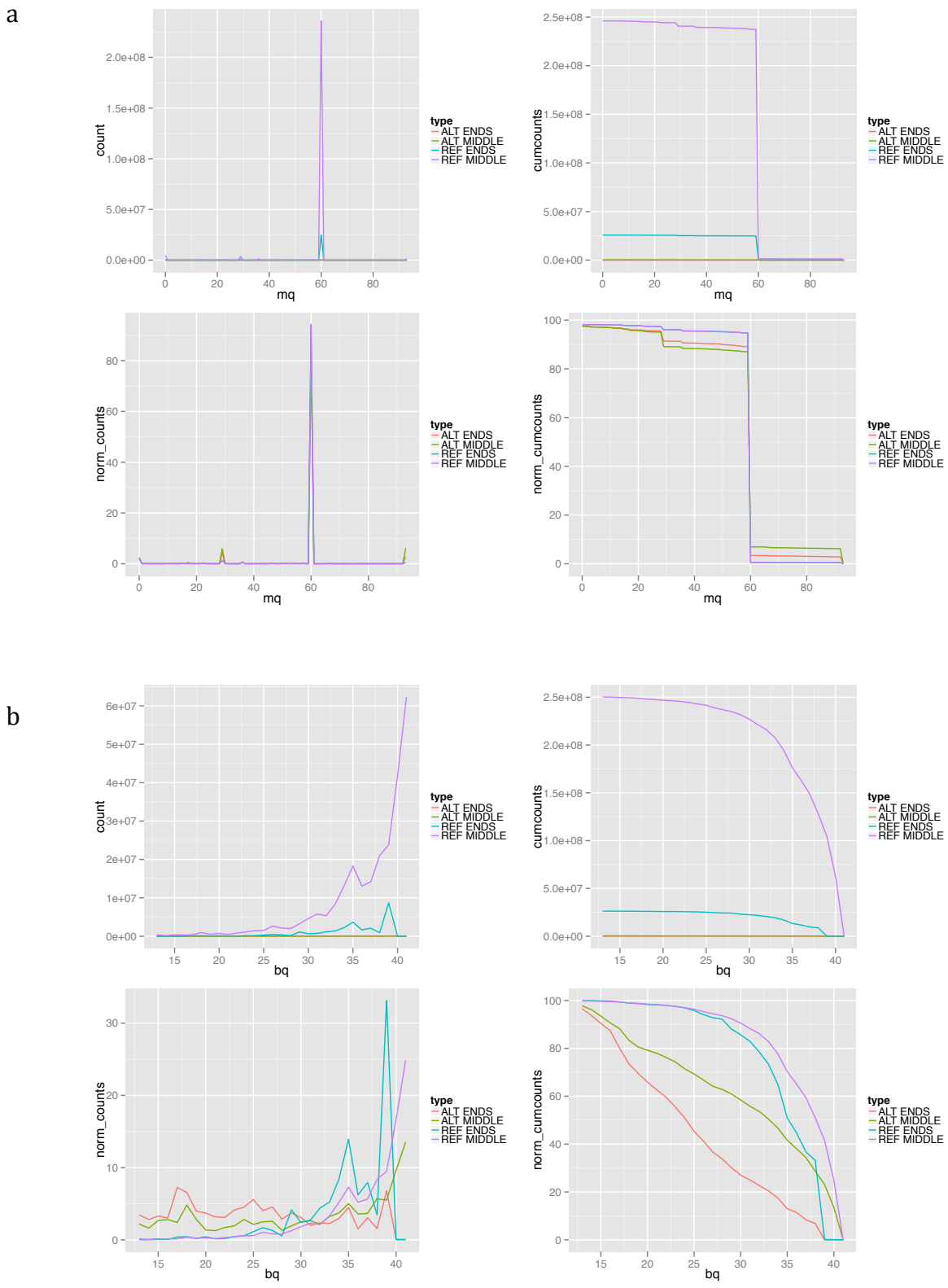


Figure 2. 2: Mapping quality (mq, a) and base quality (bq, b) from 10 randomly selected individuals. ALT, alternative allele; REF, reference allele; ENDS, read positions at the end 5 base pair positions of a read (one read is 83 base pair long); MIDDLE, read positions between 6 and 78 of a read. The x-axis of a and b are mq and bq

respectively. The y-axis for a and b are (cumulative, right) distributions of (normalised, bottom) counts for base pair positions. Cut-offs of 55 and 30 were chosen for mq and bq filtering based on the plots.

During *samtools mpileup*, every individual bam file was piled up against the reference genome in exonic regions across the whole genome with recalibrated base quality cut-off 30 (Figure 2. 2b). The exon coordinates were downloaded from the *ensembl* release 67 Homo sapiens *gtf* file (ftp.ensembl.org/pub/release-67/gtf/homo_sapiens). It contained 96,130,824 base pair positions in 254,986 exons in 21,946 genes. *Samtools mpileup* output all the genotyped sites. I then filtered out all reference sites leaving only the variant sites. A variant would be called as long as there was one alternative allele in the read base column from *samtools mpileup* output. The type of variant could be SNPs, short insertions or deletions (indels). With 1X sequencing coverage, it was not possible to differentiate homozygote from heterozygote at the majority of the SNP sites where there were one read mapped. After that, I aggregated all variants in each exonic gene for every individual and did a case control comparison to see whether there were any genes with an excess of coding variants in the cases compare to the controls.

For the rare SNP analysis, there were a number of additional functional steps. The filtered variant sites were further filtered on indels specifically, and then went through SNP calling one variant file at a time. The SNPs were then concatenated across all individuals and evaluated in an alterative allele frequency spectrum. All exonic SNPs were annotated using ANNOVAR (Wang et al. 2010) and divided into six types, including non-synonymous (NS), synonymous (S), stoploss, stopgain, unknown and untranslated region (UTR). Different allele frequency cut-offs were chosen to compare disease association statistics for different types of SNPs.

Due to an excess of rare variants, I suspected that there would still be many false SNP calls remaining in our dataset, especially rare SNPs down to the singleton level. I then did further filtering on SNPs with at least two reads supporting the alternative allele. This significantly reduced the total number of SNPs called, especially rare SNPs. I then aggregated the SNPs with different allele frequency cutoffs in each gene and did disease association to see whether there was an excess of those SNPs in any gene in the cases compare to the controls. I also did gene enrichment tests to combine the P values of the SNP association tests with different allele frequency cut-offs for different types of SNPs using Fisher's method (Fisher 1970).

GWAS bioinformatics pipeline

All samples were sequenced to a mean depth of 1.2X on the IIlumina Hiseq platform. Sequencing reads for each of the 9300 samples were aligned to GRCh37.p5 and stored in BAM format. The aligner used was Stampy (v1.0.17) (Lunter & Goodson 2011) with BWA (v0.5.9) (Li & Durbin 2009b). A base quality score recalibration table was created for each BAM file using GATK (v2.3.9) (McKenna et al. 2010) and a reference SNP and indel list from dbSNP version 137, excluding all sites after version129. For each BAM file, GATKlite (v2.2.15) (McKenna et al. 2010) was used to remove reads that were not properly mapped and to recalibrate base quality scores. This involved removing between 1–5% of reads per sample. SNPTools (Wang et al. 2013) was used to calculate genotype likelihoods (GLs) at 13,837,179 single nucleotide polymorphisms (SNPs) sites that were found to be polymorphic in the 286 Asian samples in the 1000 Genomes Project (1KGP) Phase 1 samples (Abecasis et al. 2012). BEAGLE (Browning & Browning 2007) was run on the GLs using all 572 1KGP Phase 1 Asian haplotypes

as a reference panel to generate genotype probabilities. Genotype accuracy was assessed by comparing the imputation results to genotypes called from an Illumina Zhonghua-8 900k SNP chip using 16 CONVERGE samples. The squared Pearson correlation coefficient (r^2) between SNP chip genotypes and imputed dose estimates was 92% for SNPs with a minor allele frequency > 5%. Of the 9,300 samples undergone genotyping and imputation, 226 were removed for low quality genotypes. To reduce the effects of population stratification, ancestry principal components (PCs) were constructed using EIGENSOFT 3.0 (Price et al. 2006) and SMARTPCA (Patterson et al. 2006). A final set of 144,929 autosomal SNPs with genotype probability ≥ 0.9 was used to generate 10 PCs with PC1 distinguishing north-south regional differences and the remaining PCs driven by specific chromosomal regions. Association analyses controlling for the 10 PCs were performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/download.shtml).

Chapter 3

Comorbid anxiety disorders in MDD

Introduction

This chapter deals with the relationship between seven anxiety disorders and MDD in our sample of 6,000 cases of recurrent MDD. I address the following questions: (i) is the pattern of comorbidity in China consistent with that reported in the literature, and particularly is there evidence that it is different from that found in Europe and the US? (ii) To what extent do the seven comorbid anxiety disorders independently and jointly predict risk for the different clinical features of MDD? (iii) Is there heterogeneity in the way the seven anxiety disorders predict clinical features?

Results

Here, I present the prevalence of each and any of the comorbid anxiety disorder as well as the pattern of comorbidity. In the final data containing a total of 6,207 cases, the lifetime prevalence rate for any of the seven comorbid anxiety disorders was 50.25%. The prevalence rates for each of the seven comorbid anxiety disorder are listed in Table 3. 1. On average, cases suffered 1.01 (standard deviation (SD) = 1.39) comorbid anxiety disorders. Table 3. 2 lists the prevalence of the 13 commonest combinations of the binary diagnoses of the seven comorbid anxiety disorders. Among half of the cases that suffered at least one of the seven comorbid anxiety disorders, the top 12 (out of 127 possible combinations from 7 binary diagnoses) accounted for 30.7% of all the

combinations. The top 12 combinations contained the seven anxiety disorders exclusively, with generalised anxiety disorder (GAD) 9.08%, animal phobia 5.38%, blood-injury phobia 2.84%, situational phobia 1.93%, social phobia 1.5%, agoraphobia 1.49%, and panic disorder 1.3%.

Anxiety disorders	Negative	Positive	Missing	Positive proportion
Panic disorder	5,687	405	115	6.65%
GAD	4,781	1,328	98	21.74%
Agoraphobia	5,359	715	133	11.77%
Social phobia	5,428	646	133	10.64%
Animal phobia	4,728	1,345	134	22.15%
Situation phobia	5,179	894	134	14.72%
Blood-injury phobia	5,120	953	134	15.69%
Any phobia	3,699	2,377	131	39.12%
Any comorbid anxiety	3,025	3,056	126	50.25%

Table 3. 1: Prevalence rates of comorbid anxiety disorders in patients with major depressive disorder. GAD, generalised anxiety disorder.

Items	Count	Proportion
No Anxiety	3,025	49.97%
GAD only	550	9.08%
Animal phobia only	326	5.38%
Blood-injury phobia only	172	2.84%
Situational phobia only	117	1.93%
GAD and animal phobia	108	1.78%
Social phobia only	91	1.50%
Agoraphobia only	90	1.49%
Animal and situational phobia	84	1.39%
Animal and blood-injury phobia	79	1.30%
Panic only	79	1.30%
GAD and blood-injury phobia	67	1.11%
Animal, situational & blood-injury phobia	59	0.97%
Total (exc. No anxiety)	1,822	30.07%

Table 3. 2: Top 13 combinations among the total of 128 possibilities with seven comorbid anxiety disorders that could either be present or not present in every patient with major depressive disorder. GAD, generalised anxiety disorder.

The pattern illustrates that of the patients who have a comorbid anxiety, more than half of them have at least two different comorbid anxiety disorders. There are more than a hundred combinations of what a patient might or might not have among those

seven comorbid disorders. For instance, one pattern of comorbidity could be due to GAD only; another could be due to three different phobia subtypes combined. Figure 3.1 is a heat map with the colour of each element showing the correlation in response across case individuals with at least one of the seven anxiety disorders. As shown, panic disorder is more closely related to social phobia and agoraphobia. GAD is the furthest away from the rest of the anxiety disorders.

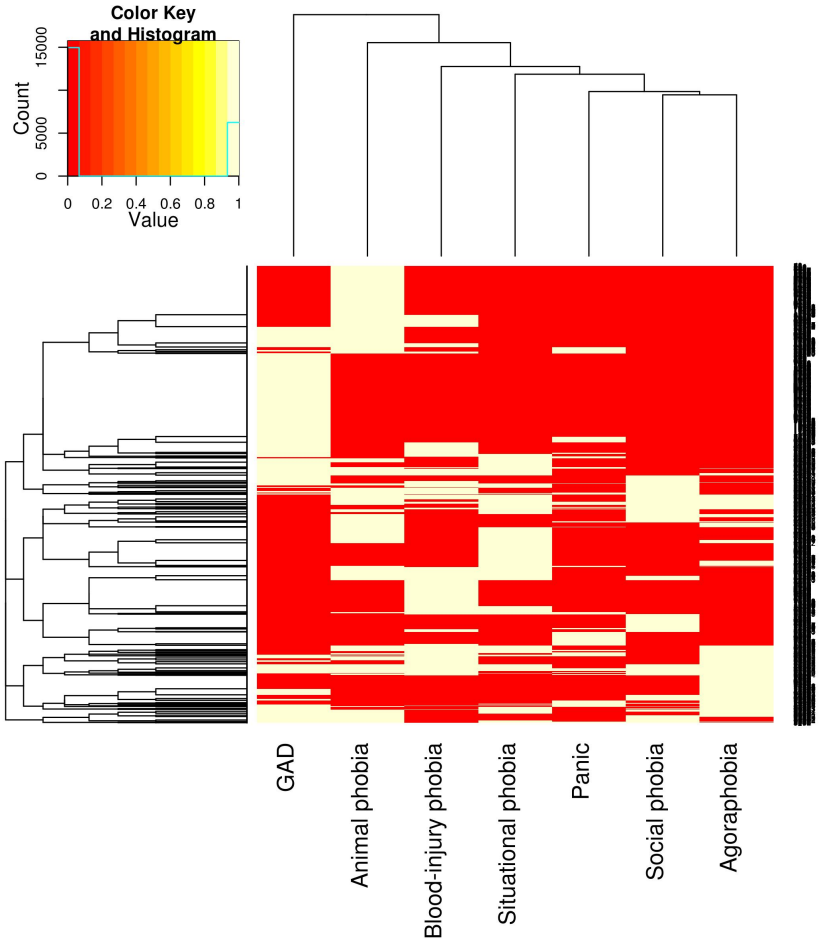


Figure 3.1: Heat map of the seven comorbid anxiety disorder with the colour of each element showing the correlation in response across case individuals with at least one of the seven comorbid anxiety disorders. GAD, generalised anxiety disorder.

In order to tease apart the individual effect of those comorbidities, I use single and multivariate regression to model their predictive pattern in depression related clinical features. Firstly, I regressed each individual comorbid anxiety disorder independently against each depression related features, including binary variable suicidal attempt (22.17% positive response) as well as other numeric variables including age of MDD onset, number of depressive episodes, number of the DSM-IV MDD A criteria, number of depressed family members and neuroticism, whose summary statistics (mean and empirical quantiles) are shown in Table 3.3. Figure 3.2 presents the Forest plots of the effects of the single regression analysis against those clinical features.

Variables	Mean	0%	25%	50%	75%	100%
Number of anxiety disorders	1.01	0	0	0	2	7
Age, years	44.51	30	37	45	52	60
Age of onset, years	34.94	7	27	34	43	60
Number of episodes (excluding 96 or more)	3.93	1	2	3	4	95
Number of DSM-IV diagnostic criteria	8.33	1	8	9	9	9
Number of positively diagnosed family members	0.35	0	0	0	1	6
Number of family members	4.80	2	4	5	6	11
Neuroticism	12.68	0	8	13	18	23

Table 3. 3: Mean and empirical quantiles of depression related clinical features. DSM, Diagnostic and Statistical Manual for Mental Disorders.

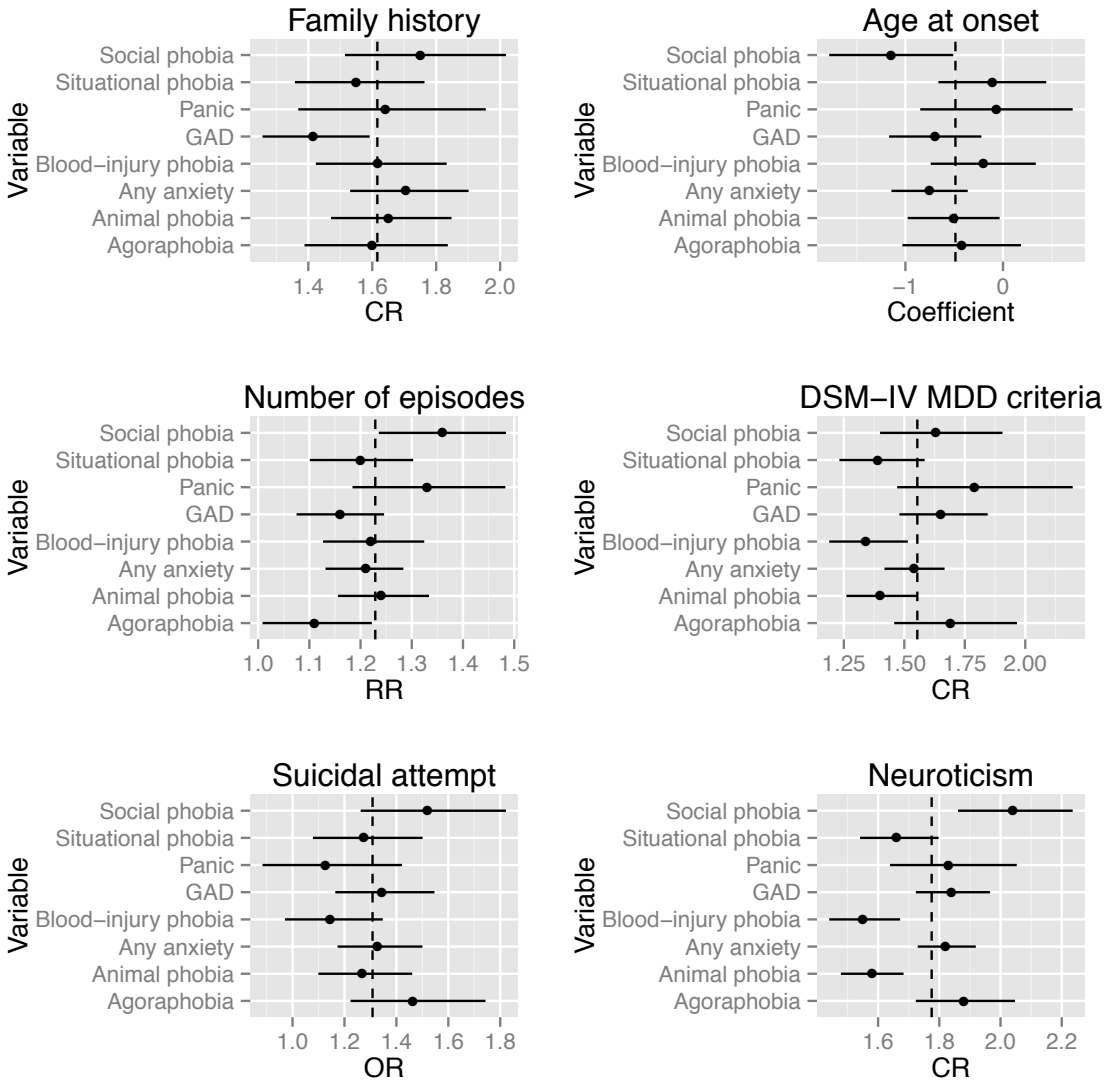


Figure 3.2: Forest plots of single regression analysis between each of the seven comorbid anxiety disorder and clinical features of major depressive disorder (MDD). GAD, generalised anxiety disorder; DSM, Diagnostic and Statistical Manual for Mental Disorders; CR, count ratio; RR, rate ratio; OR, odds ratio. Values used for plotting are for four types of generalised linear regression models including linear regression, logistic regression (OR), quasi-binomial regression (CR), and quasi-Poisson regression (RR). OR, CR and RR were calculated by exponentiating the regression estimates. Linear regression was used to model age at onset; logistic regression was used to model suicidal attempt; quasi-binomial regression was used to model family history, number of DSM-IV MDD A criteria and neuroticism; and quasi-Poisson regression was used to model number of episodes data.

Secondly, I regressed all seven comorbid anxiety disorders additively in one multivariate model controlling for age against those variables of interest as shown in Figure 3.3. Results showed that comorbid anxiety disorders were significantly

associated with clinical features of depression when tested individually and jointly. When adding all anxiety disorders in one regression model controlling for age, significant effects for individual anxiety disorders were generally reduced which gave a clearer picture of specifically which disorders were driving the signals. To control for multiple testing, P values of 0.008 (0.05 / 6 variables) could be used as significance threshold assuming all tests were independent which were not entirely true. Hence I comment on values with $P < 0.01$.

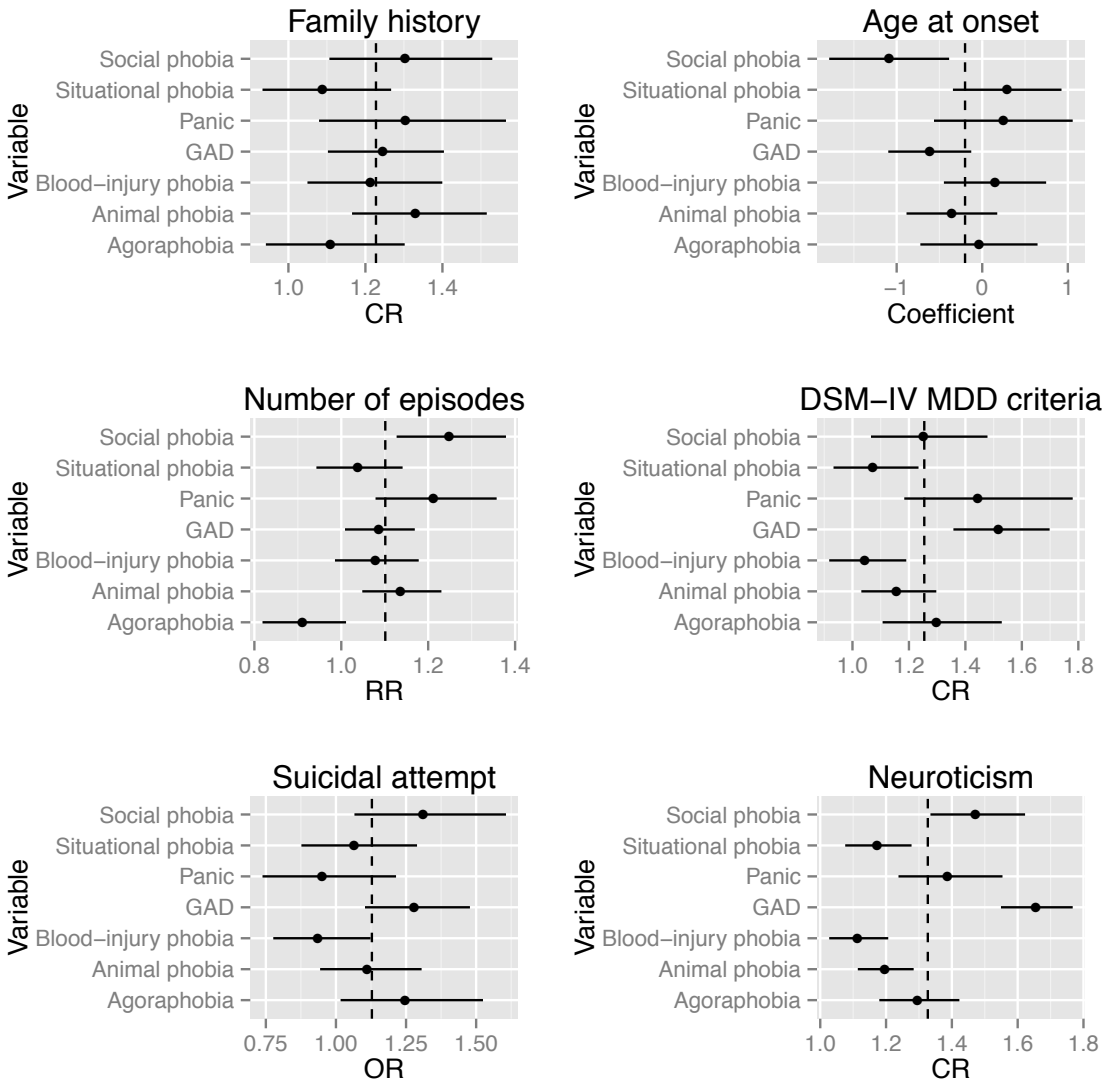


Figure 3.3: Forest plots of multivariate regression analysis for association between seven comorbid anxiety disorders considered altogether and clinical features of major depressive disorder. GAD, generalised anxiety

disorder; DSM, Diagnostic and Statistical Manual for Mental Disorders; CR, count ratio; RR, rate ratio; OR, odds ratio. Values used for plotting are for four types of generalised linear regression models including linear regression, logistic regression (OR), quasi-binomial regression (CR), and quasi-Poisson regression (RR). OR, CR and RR were calculated by exponentiating the regression estimates. Linear regression was used to model age at onset; logistic regression was used to model suicidal attempt; quasi-binomial regression was used to model family history, number of DSM-IV MDD A criteria and neuroticism; and quasi-Poisson regression was used to model number of episodes data.

Multivariate regression analysis resulted in panic, GAD, social, animal and blood-injury phobia all independently predicted risk for family history of MDD with animal phobia showing the highest effect (count ratio (CR) = 1.33). Social phobia predicted earlier onset of depression (BETA = -0.03). Panic, social and animal phobia significantly predicted higher number of depressive episodes with social phobia showing the strongest effect (rate ratio (RR) = 1.25). Panic, GAD, agoraphobia and social phobia were associated with higher number of DSM-IV MDD A criteria, with GAD showing the highest effect (CR = 1.52). GAD and social phobia predicted risk for suicidal attempt with social phobia showing the strongest effect (OR = 1.31). All seven comorbid anxiety disorders were significantly associated with neuroticism with GAD showing the highest effect (CR = 1.65). The predictive pattern of comorbid anxiety disorders in different clinical features of depression seems to be heterogeneous and hard to interpret in the multi-regression framework.

In order to examine the degree of heterogeneity among the seven comorbid anxiety disorders in predicting depression related clinical features, a global heterogeneity test was performed for multiple regression coefficients. Results showed that the seven comorbid anxiety disorders were homogeneous in predicting family history of depression, age at onset and suicidal attempt. They were heterogeneous in predicting the number of episodes (F = 3.24, P = 3.50E-3), DSM-IV MDD A criteria (F = 4.97, P = 4.42E-05) and neuroticism (F = 14.70, P = 9.77E-17). The results were in

concordance with published findings for DSM-IV MDD A criteria (F = 3.48, P = 1.99E-3) and neuroticism (F = 5.72, P = 6.64E-6) with larger degree of heterogeneity. The test on the number of episodes, which was not significant in the published finding (F = 0.59, P = 0.74), resulted in above multiple testing threshold significance when adding the additional two thirds of the samples.

Items	ANOVA F test	ANOVA P value
Family history	0.92	0.48
Age at onset	2.25	0.04
Number of episodes	3.24	3.50E-03
Number of DSM-IV MDD criteria	4.97	4.42E-05
Suicidal attempt	1.97	0.07
Neuroticism	14.70	9.77E-17

Table 3. 4: Heterogeneity test when regressing the seven comorbid anxiety disorders on clinical features of major depressive disorder (MDD). DSM, Diagnostic and Statistical Manual for Mental Disorders.

Summary

To sum up, comorbid anxiety among Chinese women is as prevalent as has been observed in the West. Although it is general consensus that comorbidy increases severity and chronicity, there is conflicting literature on whether and which comorbid anxiety disorder is associated with each clinical feature of depression, such as in the case of panic disorder and suicidal attempt. Our study has provided empirical evidence on specifically which disorder is predictive of which clinical feature of interest when assessed independently and jointly. Overall, the findings presented here support the view that there is more similarity than dissimilarity in the predictive pattern of comorbid anxiety disorders in clinical features of depression between our and Western studies, which encourages us to believe that our findings will be applicable elsewhere. However, the pattern observed can be heterogeneous and hard to interpret in the regression framework, hence we performed global heterogeneity test on the seven comorbid

anxiety disorder in their predictive power over each depression related clinical feature. We found that in the case of neuroticism, number of episodes and diagnostic criteria, there is significant degree of heterogeneity. Hence, when we talk of comorbid anxiety, it is important to distinguish different anxiety disorders rather than generic comorbidity across all. The findings from the study might have implications for the recognition, diagnosis, and possibly treatment of MDD patients with different comorbid anxiety disorders. Discussion of these points is left to the last chapter of the thesis.

Chapter 4

Depression subtypes

Introduction

The aim of this chapter is to identify symptom dimensions and latent patient classes using factor analysis (FA) and latent class analysis (LCA) respectively in a homogeneous sample of female MDD patients recruited through clinical settings. Half of the samples were used to conduct exploratory FA (EFA). The results were then validated in the other half using confirmatory FA (CFA). LCA was performed on the entire case samples. The analysis was based on three sets of items. The first set of items includes the nine DSM-IV A criteria for MDD, with the aim to validate the heterogeneity observed in Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD) using clinical samples. The second set contains 14 items representing the disaggregated DSM-IV criteria (i.e. increased and decreased appetite and weight, insomnia or hypersomnia, and psychomotor agitation or retardation). The primary aim is to examine the atypical somatic subtype of depression. The third set of items consists of detailed assessments of DSM-IV criteria and symptoms of melancholia, anxiety and Beck's cognitive trio (Beck and Alford 2008;Wang et al. 2013). This is to validate the melancholic and anxiety subtypes commonly observed using various instruments. The classes identified using LCA were then validated by examining other depression-related clinical features, comorbidities and environmental risk factors.

Results

For each of the three sets of items, I will describe results from EFA, CFA and LCA in turn. I will justify the decision made on the number of factors from EFA using *eigenvalues* greater than one as well as clinical interpretation of the factors. The factors were named after the items primarily loaded on them. For CFA on the other half of the samples, I will give fit indices as empirical support to validate the EFA results. I will also comment on the factor correlation coefficients for both EFA and CFA. For LCA, the decision made on the number of classes was based on Akaike information criterion (AIC), Bayesian information criterion (BIC) and sample size adjusted BIC, the Vuong-Lo-Mendell-Rubin test and the bootstrapped parametric likelihood ratio test, and the clinical interpretation of the classes. The classes were named after the most distinctive item endorsements across all classes. I will also report Cramer's V as a measure of effect size across classes for every symptom item. Additionally, I will comment on the transition frequencies of patients endorsed in each class at the end of the 14 and 27 items LCAs. Finally, I will describe the external validation results across classes.

FA on 9 items DSM-IV MDD A criteria

Endorsement frequencies for the nine DSM-IV A criteria for MDD were uniformly high and similar in our two sub-samples (Table 4. 1). An EFA in sample one produced two *eigenvalues* exceeding unity. The first factor displayed higher loadings primarily on the somatic symptoms of depression and less prominently on mood symptoms, and reflected a general depressive symptoms factor. The second factor had the highest loadings on suicidal ideation (SI) and worthlessness/guilt, and reflected a cognitive symptoms factor (Table 4. 1). The inter-factor correlation coefficient was + 0.53. The

best fit CFA solution, two factors with all variable loadings ≥ 0.30 from the EFA, had an excellent fit (RMSEA = 0.015, CFI = 0.98, TLI 0.97), which was substantially better than those obtained with one factor (RMSEA = 0.031, CFI = 0.89, TLI 0.86). The inter-factor correlation was +0.78. The loadings for the mood criteria A1 was stronger on the first factor and much weaker on the second factor in the CFA compared to the EFA.

LCA on 9 items DSM-IV MDD A criteria

For a LCA on the same nine DSM-IV criteria for MDD, both BIC and sample-size adjusted BIC reached a low peak for three-class solution (Figure 4. 1). Additionally, the Vuong-Lo-Mendell-Rubin test and the bootstrapped parametric likelihood ratio test both supported a three-class solution. Table 4. 1 depicts the resulting class membership and the item endorsement frequencies for this three-class solution. Class one was the largest class, constituting 85.6% of the cases, and was characterized by high endorsement on all criteria. We called this the severe class. Class two was labeled non-suicidal (8.0%), as it had zero endorsement on SI, and low endorsement of the worthlessness/guilt criterion. Class three was labeled moderate (6.4%), owing to a symptom pattern reflecting lower endorsement frequencies for all nine criteria. We examined the informativeness of the individual criterion for the LCA by examining their Cramer's V statistic (Table 4. 1). Anhedonia, depressed mood and sleep changes were especially uninformative, probably because of their very high overall endorsement in our sample. The three most influential criteria for this LCA were psychomotor changes, SI and worthlessness/guilt.

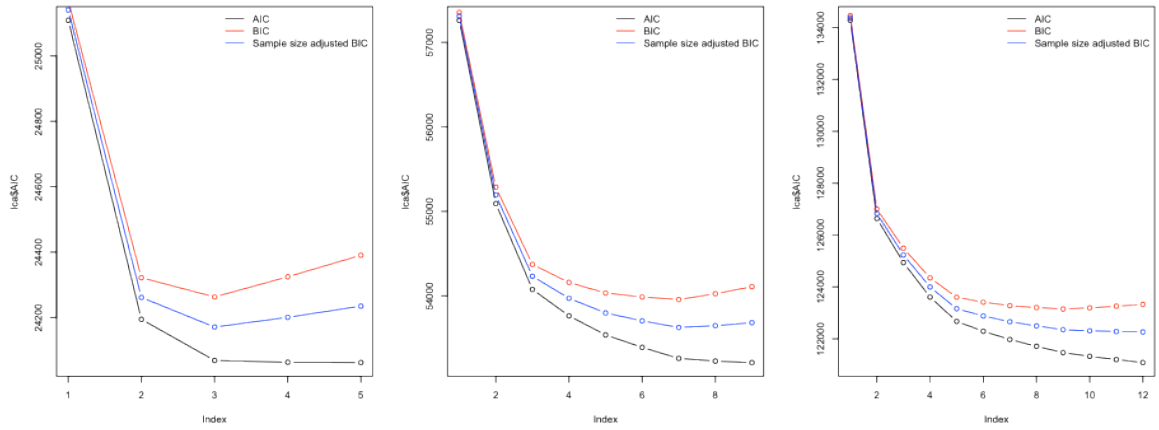


Figure 4. 1: Akaike information criterion (AIC), Bayesian information criterion (BIC) and sample size adjusted BIC for the three datasets for determining the number of classes.

	EFA on sample 1			CFA on sample 2			LCA			
MDD A criteria	**Endorsement**	**F1**	**F2**	**Endorsement**	**F1**	**F2**	**C1**	**C2**	**C3**	**Cramer's V**
A1	99.57%	**0.36**	**0.44**	99.37%	0.68	0.05	1.00	0.99	0.95	0.25
A2	98.87%	**0.27**	0.21	98.84%	0.54		1.00	0.98	0.94	0.23
A3	91.08%	**0.51**	0.00	90.36%	0.40		0.94	0.86	0.68	0.34
A4	95.67%	**0.54**	-0.05	95.58%	0.35		0.97	0.93	0.84	0.25
A5	91.61%	**0.65**	0.02	90.29%	0.71		0.96	0.91	0.50	0.57
A6	93.53%	**0.72**	-0.21	93.51%	0.46		0.96	0.96	0.69	0.38
A7	90.67%	0.11	**0.64**	90.05%		0.77	0.96	**0.63**	0.70	0.54
A8	97.72%	**0.66**	0.07	97.69%	0.73		1.00	0.99	0.81	0.49
A9	76.32%	-0.01	**0.66**	76.39%		0.67	0.89	**0.00**	0.57	0.56

Table 4. 1: Exploratory factor analysis (EFA), confirmatory factor analysis (CFA) and latent class analysis (LCA) on nine Diagnostic and Statistical Manual for Mental Disorders (DSM)-IV major depressive disorder (MDD) A criteria. A1, depressed mood; A2, anhedonia; A3, significant weight or appetite disturbance; A4, sleep disturbance; A5, psychomotor agitation or retardation; A6, loss of energy or fatigue; A7, feeling of worthlessness; A8, diminished ability to think or concentrate; A9, recurrent thoughts of death; F1, general depressive symptoms factor; F2, cognitive symptoms factor; C1, severe class; C2, non-suicidal class; C3, moderate class. For EFA, bold factor loadings indicate that the criterion should load on that particular factor, usually with absolute loading > 0.3, except when all absolute loadings < 0.3 in which case the item will load on the factor with the highest loading as in the case of A2 on F1. This decision affects CFA, as the criterion will be forced loaded as zero on the other factor. For LCA, the values correspond to the endorsement frequencies (percentage of individuals responded positively to that criterion) of the case individuals in that class. Bold endorsement frequencies indicate items that stand out in that particular class compare to all other classes, this can be very high or very low.

FA on the 14 disaggregated DSM-IV MDD A criteria

Next, I examined the 14 items expanded on the three of the nine bi-directional DSM-IV

A criteria representing appetite/weight changes, sleep disturbances, and psychomotor

agitation or retardation. The item endorsement frequencies in the 14 disaggregated DSM-IV A criteria were similar in the two sub-samples and more variable than with the nine criteria. Atypical vegetative symptoms were endorsed by 7-13% of patients. The scree plot for the EFA analyses showed four factors with *eigenvalues* exceeding one and a clear "elbow" at 3 factors. Furthermore, a solution with three factors was considerably more interpretable than one with four and resulted in factors reflecting weight/appetite symptoms, general depressive symptoms and sleep disturbance (Table 4. 2) with modest inter-factor correlations. In our CFA analyses (which produced the most interpretable results when including all EFA loadings ≥ 0.30), the three factor solution had better fit indices (RMSEA 0.032, CFI 0.94, TLI 0.93) than the four factor solution (RMSEA 0.035, CFI 0.93, TLI 0.91) validating our interpretation of the EFA. Loadings from the CFA were comparable to those found in the EFA, identifying three broadly similar factors. Inter-factor correlations were modest except between factors one and three, which equaled -0.44 (Table 4. 3).

	EFA on sample 1				CFA on sample 2				LCA				
Items	**Endorsement**	**F1**	**F2**	**F3**	**Endorsement**	**F1**	**F2**	**F3**	**C1**	**C2**	**C3**	**C4**	**Cramer's V**
A1	99.57%	0.09	**0.68**	-0.20	99.37%		0.69		1.00	0.98	0.98	1.00	0.13
A2	98.87%	-0.08	**0.41**	0.13	98.84%		0.48		1.00	0.97	0.96	0.99	0.16
A3.1	84.23%	**-0.72**	**0.63**	0.00	82.83%	0.67	0.64		0.98	0.31	0.93	**0.57**	0.79
A3.2	60.15%	**-0.62**	**0.40**	0.00	58.52%	0.51	0.43		0.72	0.12	0.72	**0.31**	0.46
A3.3	9.43%	**0.74**	0.01	**0.37**	9.74%	-0.81		0.20	0.03	0.08	0.01	**0.83**	0.71
A3.4	7.11%	**0.84**	0.00	**0.34**	6.33%	-0.88		0.22	0.01	0.03	0.00	**0.73**	0.85
A4.1	92.99%	0.00	**0.47**	**-0.60**	91.87%		0.35	-0.59	0.96	0.83	0.91	0.87	0.20
A4.2	13.12%	0.02	0.00	**0.77**	13.53%			0.95	0.11	0.15	0.08	**0.37**	0.21
A5.1	76.69%	-0.11	**0.48**	0.10	75.18%		0.59		0.86	0.54	0.50	0.80	0.35
A5.2	72.19%	-0.01	**0.46**	-0.13	72.91%		0.45		0.81	0.54	0.54	0.75	0.27
A6	93.53%	-0.10	**0.50**	-0.01	93.51%		0.46		0.98	0.84	0.87	0.92	0.27
A7	90.67%	0.05	**0.60**	0.09	90.05%		0.57		0.98	0.85	0.55	0.94	0.68
A8	97.72%	0.02	**0.72**	-0.06	97.69%		0.73		1.00	0.93	0.91	0.99	0.30
A9	76.32%	-0.07	**0.45**	0.07	76.39%		0.48		0.87	0.64	0.40	0.79	0.37

Table 4. 2: Exploratory factor analysis (EFA), confirmatory factor analysis (CFA) and latent class analysis (LCA) on 14-item dataset. A1, depressed mood; A2, anhedonia; A3.1, loss of appetite; A3.2, loss of weight; A3.3, increase of appetite; A3.4, increase of weight; A4.1, insomnia; A4.2, hypersomnia; A5.1, psychomotor retardation; A5.2, psychomotor agitation; A6, loss of energy or fatigue; A7, feeling of worthlessness; A8, diminished ability to think or concentrate; A9, recurrent thoughts of death. F1, weight / appetite symptoms factor; F2, general depressive symptoms factor; F3, sleep disturbance factor; C1, severe typical class; C2, moderate typical class; C3, non-suicidal class; C4 atypical class. Endorsement here means the percentage of case individuals responded positively to that particular item. For EFA, bold factor loadings indicate that the criterion should load on that particular factor, usually with absolute loading > 0.3, except when all absolute loadings < 0.3 in which case the item will load on the factor with the highest loading. This decision affects CFA, as the criterion will be forced loaded as zero on the other factors. For LCA, the values correspond to the endorsement frequencies (percentage of individuals responded positively to that criterion) of the case individuals in that class. Bold endorsement frequencies indicate items that stand out in that particular class compare to all other classes, this can be very high or very low.

EFA	F1	F2	F3	CFA	F1	F2	F3
F1	1.00			**F1**	1.00		
F2	0.21	1.00		**F2**	-0.12	1.00	
F3	0.18	0.10	1.00	**F3**	-0.44	0.11	1.00

EFA	F1	F2	F3	F4	F5	CFA	F1	F2	F3	F4	F5
F1	1.00					**F1**	1.00				
F2	0.06	1.00				**F2**	0.51	1.00			
F3	-0.08	0.13	1.00			**F3**	-0.47	0.17	1.00		
F4	0.10	0.38	0.12	1.00		**F4**	0.34	0.51	0.14	1.00	
F5	0.00	0.37	0.02	0.22	1.00	**F5**	0.39	0.87	0.18	0.62	1.00

Table 4. 3: Inter-factor correlation coefficients for 14 & 27 items exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) solutions. For 14 items three factor solution, F1, weight / appetite symptoms factor; F2, general depressive symptoms factor; F3, sleep disturbance factor; for 27 items five factor solution, F1, weight/appetite symptoms factor; F2, general regarded depressive symptoms factor; F3, atypical vegetative symptoms factor; F4, suicidal/hopeless factor; F5, agitated depressive symptoms factor.

LCA on 14 disaggregated DSM-IV MDD A criteria

For the LCA on the same expanded 14 items reflecting the disaggregated DSM-IV MDD criteria, the information criteria plot did not show a clear "elbow" (Figure 4. 1). The Vuong-Lo-Mendell-Rubin test and the bootstrapped parametric likelihood ratio tests both suggested a four-class solution, whose class membership and the item endorsement frequencies are depicted in Table 4. 2. By far the most common class was called severe typical (70.4%) and was characterized by highest endorsement on all typical symptoms of depression. The second class called moderate typical (13.6%) had lower endorsement rates for all typical depression symptoms, especially loss of appetite and weight as well as psychomotor changes. Class three was characterized by low endorsement of symptoms of death or SI as well as feelings of worthlessness/guilt, and hence was labeled non-suicidal (8.7%). Class four was labeled atypical (7.4%) owing to its high endorsements on atypical neurovegetative symptoms of increased appetite and weight gain as well as hypersomnia.

The most informative items for these analyses were those describing typical and atypical weight and appetite changes, and feelings of worthlessness/guilt. Although also disaggregated, the psychomotor and sleep items were considerably less influential on class assignment.

Figure 4. 2 shows the proportion classified into the 14-item 4 class solutions coming from the individual classes from the 9-item 3 class solution. 100% of the 14-item severe typical class came from the 9-item severe class, as did 92% of the 14-item atypical class. The two non-suicidal classes were moderately closely related with 52% of the 14-item class coming from the 9-item class. The two moderate classes were even less closely inter-related as 52% of the 14-item moderate class came from the 9-item severe class, and only 27% from the moderate class.

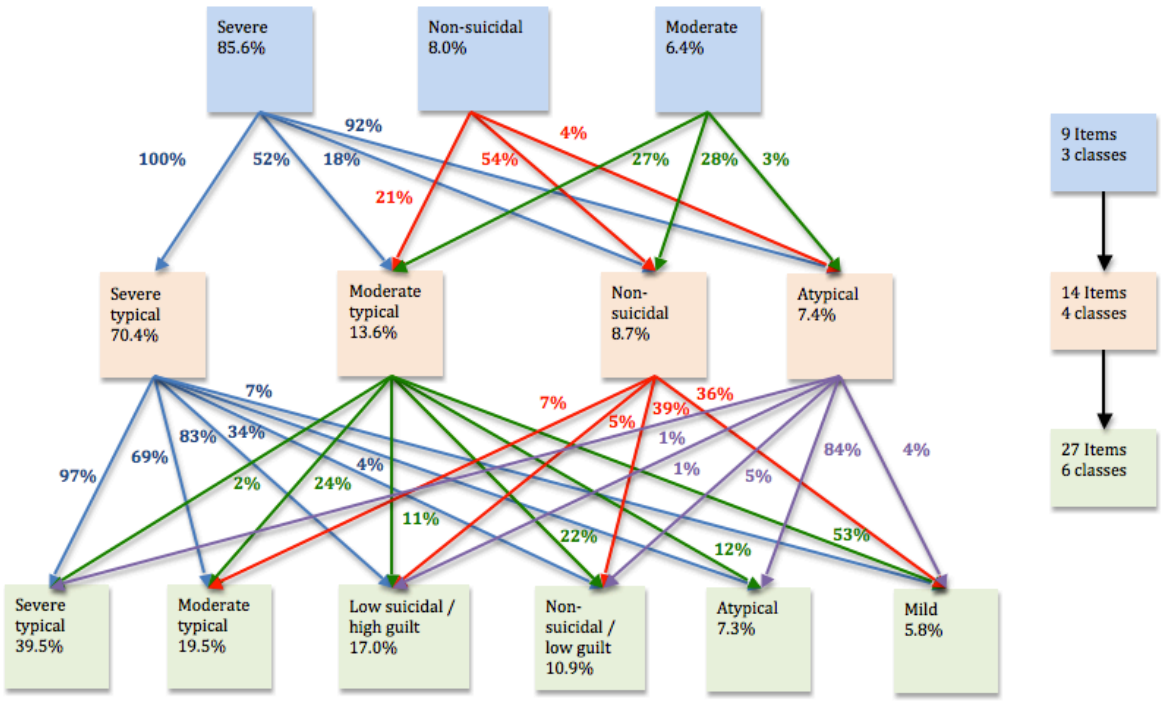


Figure 4. 2: Transition frequencies from 9 to 14 and from 14 to 27 items datasets.

FA on all 27 individual symptoms assessed during worst depressive episode

Next, I examined the 27 depression symptoms including those expanded on the nine DSM-IV MDD A criteria, the symptoms of melancholia and anxiety, as well as Beck's cognitive symptoms. Endorsement rates were variable for the 27 depression-related symptoms individually assessed during the worst lifetime episode and similar in our two sub-samples (Table 4. 4). More than 80% of our sample endorsed both key melancholic symptoms (e.g. lack of mood reactivity and distinct mood quality) and items assessing Beck's cognitive triad (e.g. hopelessness and worthlessness) indicating the clinical severity of the reported depressive episodes.

	EFA on sample 1						CFA on sample 2					
Items	**Endorsement**	**F1**	**F2**	**F3**	**F4**	**F5**	**Endorsement**	**F1**	**F2**	**F3**	**F4**	**F5**
A1	99.57%	-0.06	**0.53**	-0.19	0.13	0.16	99.37%		0.70			
A2	98.87%	0.03	**0.44**	0.05	0.12	-0.12	98.84%		0.52			
A3.1	84.23%	**0.71**	**0.40**	-0.02	-0.02	-0.02	82.83%	0.96	0.02			
A3.2	60.15%	**0.65**	0.18	0.01	-0.04	0.11	58.52%	0.64				
A3.3	9.43%	**-0.65**	-0.01	**0.56**	0.00	0.19	9.74%	-0.17		0.80		
A.3.4	7.11%	**-0.74**	0.09	**0.52**	-0.05	-0.02	6.33%	-0.14		0.90		
A4.1	92.99%	0.06	**0.40**	**-0.49**	-0.03	0.14	91.87%		0.32	-0.44		
A4.2	13.12%	-0.09	-0.03	**0.58**	0.05	-0.01	13.53%			0.50		
A5.1	76.69%	0.04	**0.68**	0.09	0.05	***-0.28***	75.18%		0.96			-0.43
A5.2	72.19%	-0.03	**0.33**	-0.19	0.01	**0.32**	72.91%		0.35			0.15
A.6	93.53%	0.12	**0.51**	0.04	-0.12	0.04	93.51%		0.44			
A7.1	82.39%	-0.02	**0.36**	0.11	**0.44**	0.11	81.07%		0.42		0.44	
A7.2	72.42%	0.02	0.19	0.16	0.17	***0.26***	75.12%					0.65
A8.1	92.15%	-0.06	**0.72**	-0.04	-0.11	0.10	90.89%		0.65			
A8.2	89.30%	-0.18	**0.72**	-0.01	0.02	-0.03	88.38%		0.70			
A8.3	83.86%	-0.10	**0.72**	-0.01	0.03	0.00	82.89%		0.72			
A9.1	71.22%	-0.01	-0.01	-0.07	**0.96**	-0.03	72.17%				0.94	
A9.2	61.72%	-0.07	-0.01	-0.13	**0.97**	-0.01	60.92%				0.95	
ME.A2	87.05%	0.04	**0.41**	0.00	***0.22***	-0.14	86.13%		0.44		0.17	
ME.B1	93.54%	0.10	***0.14***	0.02	0.11	-0.12	92.46%		0.27			
ME.B2	63.70%	0.02	***0.22***	0.06	-0.07	0.08	60.65%		0.29			
MDD1	89.21%	0.02	**0.41**	-0.13	0.04	-0.04	88.71%		0.43			
MDD2	74.59%	-0.07	-0.02	-0.03	0.01	**0.58**	74.81%					0.31
MDD3	80.80%	0.07	***0.21***	0.09	**0.57**	0.20	80.33%		0.29		0.63	
MDD4	67.55%	0.14	-0.09	0.14	0.17	**0.37**	67.34%					0.34
MDD5	89.65%	0.11	0.20	0.19	***0.26***	**0.39**	89.24%				0.22	0.55
MDD6	89.58%	-0.03	0.08	-0.09	0.00	**0.67**	88.87%					0.57

	LCA						
Items	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**Cramer's V**
A1	1.00	1.00	1.00	0.99	1.00	0.94	0.22
A2	1.00	0.99	1.00	0.97	1.00	0.92	0.19
A3.1	0.97	0.81	0.89	0.76	**0.52**	0.55	0.40
A3.2	0.75	0.52	0.62	0.54	**0.27**	0.34	0.32
A3.3	0.04	0.02	0.01	0.07	**0.77**	0.08	0.73
A.3.4	0.01	0.00	0.01	0.05	**0.65**	0.07	0.73
A4.1	0.97	0.90	0.96	0.89	0.85	0.80	0.21
A4.2	0.13	0.12	0.08	0.12	**0.38**	0.11	0.22
A5.1	0.90	0.65	0.82	0.66	0.79	0.25	0.40
A5.2	0.86	0.62	0.77	0.60	0.77	0.34	0.34
A.6	0.99	0.89	0.96	0.93	0.91	0.73	0.29
A7.1	0.98	0.83	**0.85**	**0.35**	0.92	0.41	0.58
A7.2	0.88	0.66	**0.80**	**0.41**	0.82	0.46	0.39
A8.1	0.99	0.85	0.97	0.90	0.97	0.47	0.49
A8.2	0.99	0.81	0.94	0.87	0.97	0.34	0.54
A8.3	0.98	0.72	0.93	0.69	0.89	0.28	0.53
A9.1	0.99	0.96	**0.29**	**0.10**	0.80	0.38	0.78
A9.2	0.94	0.87	**0.05**	**0.03**	0.69	0.23	0.83
ME.A2	0.97	0.85	0.84	0.76	0.88	0.54	0.32
ME.B1	0.96	0.93	0.93	0.90	0.91	0.85	0.12
ME.B2	0.70	0.51	0.68	0.57	0.68	0.42	0.20
MDD1	0.96	0.84	0.92	0.87	0.88	0.63	0.27
MDD2	0.84	0.63	0.77	0.63	0.85	0.66	0.24
MDD3	0.99	0.87	**0.81**	**0.22**	0.89	0.45	0.65
MDD4	0.78	0.62	**0.68**	**0.45**	0.72	0.57	0.25
MDD5	0.99	0.86	**0.98**	**0.61**	0.95	0.67	0.46
MDD6	0.97	0.81	0.95	0.80	0.93	0.70	0.30

Table 4. 4: Exploratory factor analysis (EFA), confirmatory factor analysis (CFA) and latent class analysis (LCA) on 27 items dataset. A1, depressed mood; A2, anhedonia; A3.1, loss of appetite; A3.2, loss of weight; A3.3, increase of appetite; A3.4, increase of weight; A4.1, insomnia; A4.2, hypersomnia; A5.1, psychomotor retardation; A5.2, psychomotor agitation; A6, loss of energy or fatigue; A7, feeling of worthlessness; A8, diminished ability to think or concentrate; A9, recurrent thoughts of death. ME.A2, loss of ability to enjoy good things; ME.B1, depression different from grief or loss; ME.B2, worse mood in the morning; MDD1, loss of sexual drive; MDD2, feel irritable or angry most of the time; MDD3, hopelessness; MDD4, cry a lot; MDD5, helplessness; MDD6, feeling nervous, jittery or anxious. F1, weight/appetite symptoms factor; F2, general regarded depressive symptoms factor; F3, atypical vegetative symptoms factor; F4, suicidal/hopeless factor; F5, agitated depressive symptoms factor; C1, severe typical class; C2, moderate typical class; C3, low suicidal/high guilt class; C4, non-suicidal/low guilt class; C5, atypical class; C6, mild class. For EFA, bold factor loadings indicate that the criterion should load on that particular factor, usually with absolute loading > 0.3, except when all absolute loadings < 0.3 in which case the item will load on the factor with the highest loading. This decision affects CFA, as the criterion will be forced loaded as zero on the other factors. For LCA, the values correspond to the endorsement frequencies (percentage of individuals responded positively to that criterion) of the case individuals in that class. Bold endorsement frequencies indicate items that stand out in that particular class compare to all other classes, this can be very high or very low.

The factor structure of these 27 items was initially examined by EFA (Table 4. 4) resulting in 8 *eigenvalues* exceeding one without a clear "elbow" in the Scree plot. With seven and eight factor solutions, estimated residual variances of some items became negative indicating over-extraction. The five factor EFA model was clinically more sensible than the six factor solution. Similar to the results from the 14 item analysis, the first factor reflected weight/appetite symptoms. The second factor was also similar to that seen with the 14 item solution with a more prominent loading on psychomotor retardation. We called this the general retarded depressive symptom factor. The third factor had highest loadings on the atypical vegetative symptoms of increased appetite and weight, and hypersomnia. The fourth factor had prominent loadings on suicidal symptoms and Beck's cognitive triad of helplessness, hopelessness and worthlessness, and was called a suicidal/hopeless factor. The fifth factor had prominent loadings on agitated depressive symptoms including agitation, nervousness, guilt, irritability, and crying. Inter-factor correlations were generally modest with the highest observed between the general depressive and suicidal/hopeless factors (+0.38).

In the CFA, when the number of factors equalled six or greater, standard errors of model parameter estimates as well as the factor scores could not be computed, which is also an indication of factor over-extraction. Both four factor and five factor solutions (with items of EFA loading cut-off ≥ 0.2) described the data well with the five-factor solution resulting in slightly superior fits (RMSEA 0.030, CFI 0.953, TLI 0.946 vs. RMSEA 0.030, CFI 0.950, TLI 0.943), again congruent with our clinical interpretation. The loadings closely resembled those found in the EFA. The general retarded depressive symptom factor had notable loadings (with the absolute value ≥ 0.2) on by far the most items (n = 15) with higher loadings for typical severe retarded depressive symptoms including several of the

melancholic symptoms. The agitated depressive and suicidal/hopeless factors were next with prominent loadings on six and five symptoms, respectively. The atypical vegetative and weight/appetite symptom factors were the smallest, with prominent loadings on four and two symptoms, respectively.

LCA on all 27 individual symptoms assessed during worst depressive episode

For the LCA on the complete set of 27 items in our interview assessed about the worst lifetime episode of MDD, the information criteria plot also did not show a clear elbow (Figure 4. 1). The Vuong-Lo-Mendell-Rubin test and the bootstrapped parametric likelihood ratio tests both suggested a six-class solution, whose class membership and the item endorsement frequencies are depicted in Table 4. 4. The most common class – class one – was best described as severe typical and was characterized by very high endorsement rates for typical depressive symptoms. Contrary to prior solutions, however, this class now represented only 39.6% of subjects. Class two represented moderate typical depression (19.5%). Classes three and four were both characterized by low rates of SI and thoughts of death. However, compared to class four (10.9%), class three (17.0%) had relatively high rates of guilt, worthlessness and hopelessness. So, we labeled group 3 as low suicidal/high guilt and group 4 as non-suicidal/low guilt. Group 5 is clearly an atypical group (7.3%), and group 6 has consistently lower endorsement rates across most symptoms than any other of the groups and is therefore labeled a mild group.

The most informative items for this LCA reflected SI, classic depressive cognitive symptoms of worthlessness and hopelessness, atypical vegetative symptoms, especially increased appetite and weight gain, and trouble thinking and concentrating. Somewhat surprisingly, the melancholic symptoms of distinct mood quality and mood worse in the

morning were, along with the main items of sad mood and loss of interest/pleasure, the least informative for our class assignments.

Figure 4. 2 presents the relationship between the individuals making up the 14-item 4 class and the 27-item 6 class solutions. Of the many results presented there, the most notable were: i) nearly all subjects (97%) in the 27-item severe typical class were also in the 14 item severe-typical class; ii) most subjects (84%) in the 27-item atypical class were also so classified in the 14-item LCA; iii) the moderate typical class was again moderately stable with a much higher proportion in the 27-item solution (69%) deriving from the 14-item severe typical than moderate typical class (24%); iv) a large proportion of the low SI/high guilt 27 item class (83%) also came from the 14-item severe class; v) the non-suicidal/low guilt class was moderately stable over analyses with 53% of those classified in the 27-item solution also so classified in the 14-item solution; and vi) a majority of the new mild class in the 27-item solution derived from the moderate typical class in the 14-item solution.

An alternative visualization of the FA and LCA results using the three datasets are shown as heat maps in figure 4.3 and 4.4. For LCA, the typical severe classes identified in all three datasets are most closely related to each other, which have the furthest distance from the rest of the classes. Moreover, the two atypical classes identified in the 14 and 27 items data, the low suicidal ideation and cognition classes, and the moderate and mild classes identified in the three datasets are more closely related to each other as expected. For FA, the weight/appetite factors identified in the 14 and 27 items datasets are most closely related to each other and have the furthest distance from the rest of the factors, the atypical factor identified from the 27 items is most closely related to the sleep disturbance factor identified from the 14 items data.

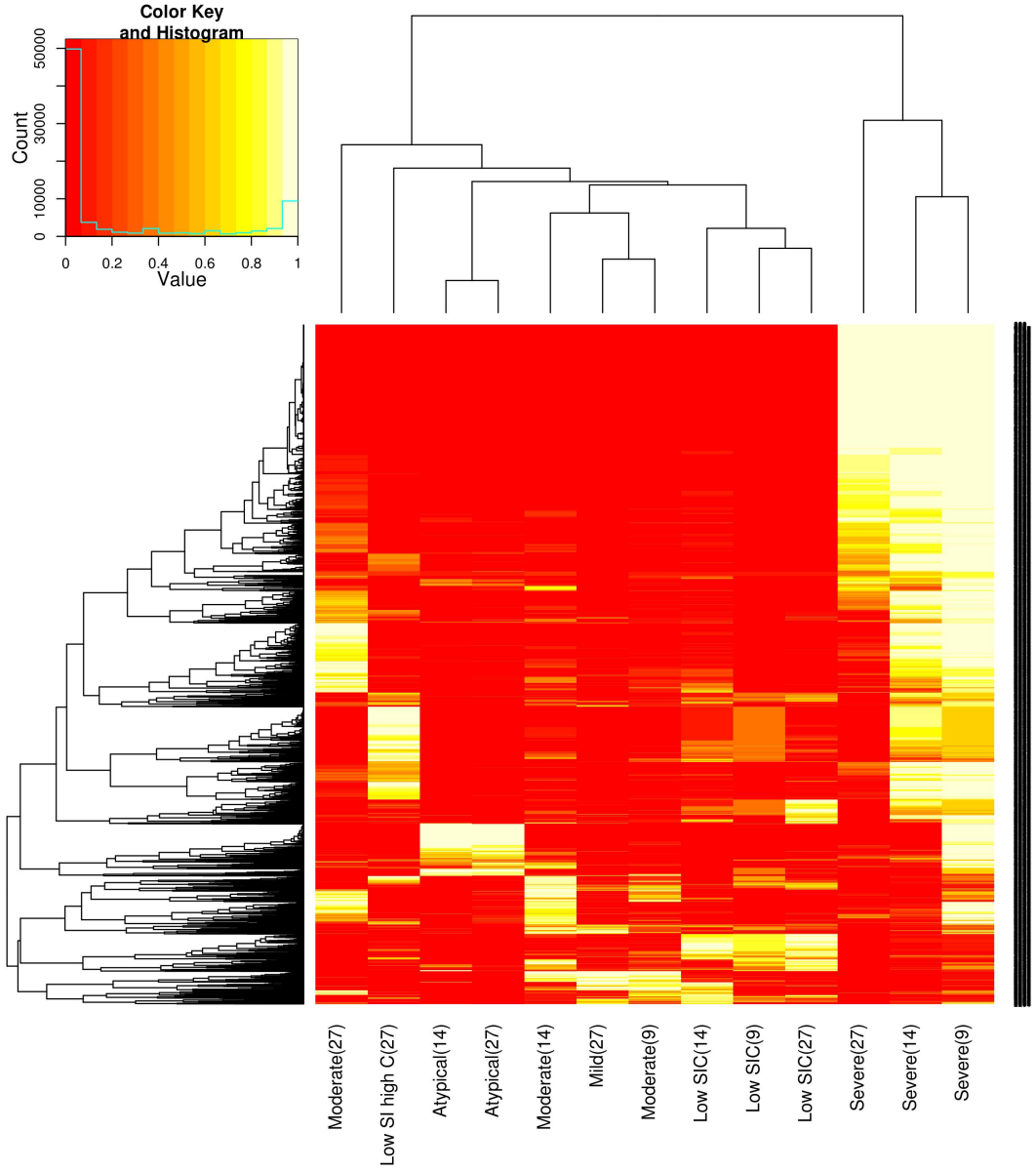


Figure 4.3: Heat map of LCA classes identified from the 9, 14 and 27 items datasets with the colour of each element showing the correlation of class endorsement probability across case individuals. SIC, suicidal ideation & cognition; SI, suicidal ideation; C, cognition.

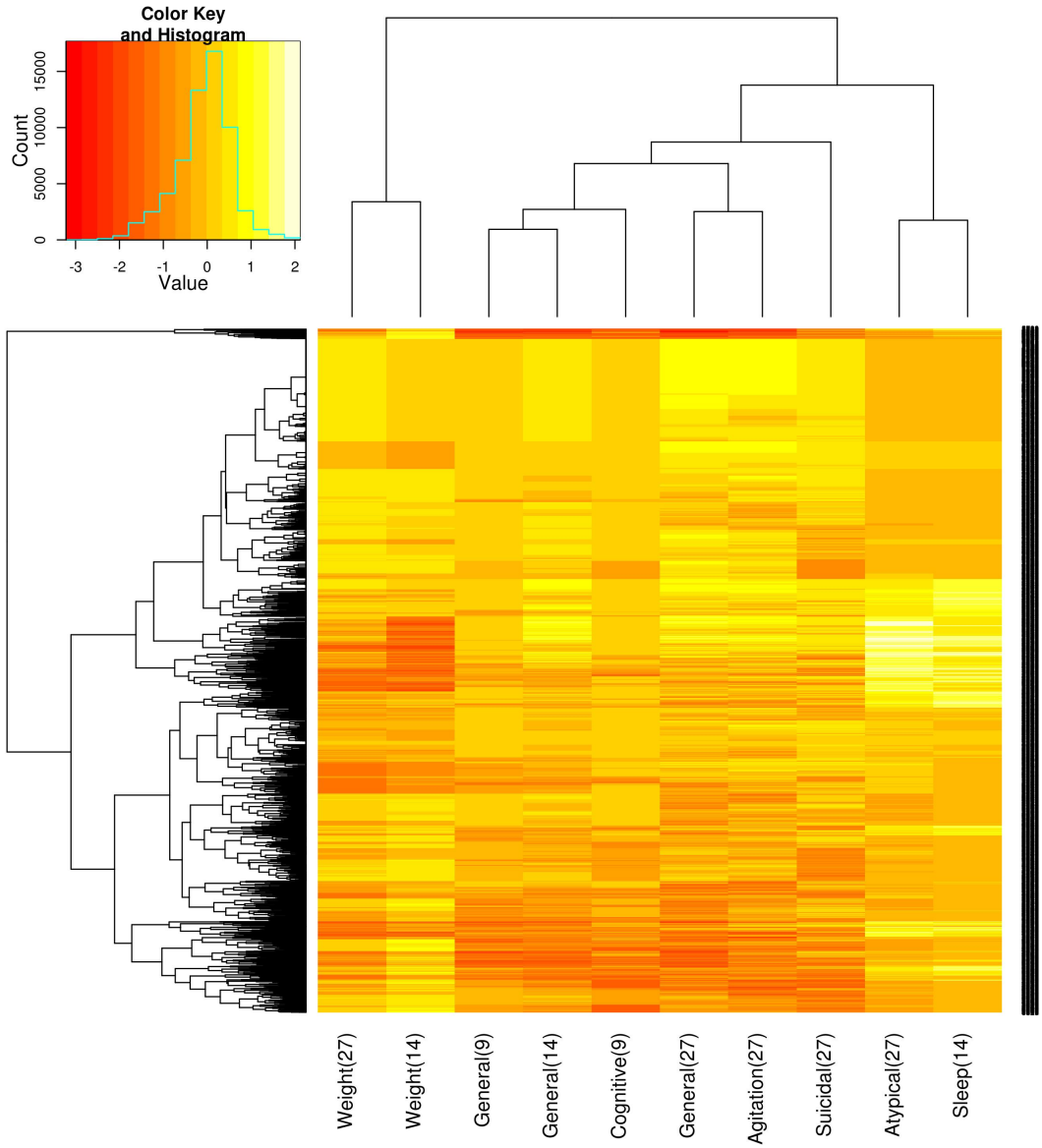


Figure 4.4: Heat map of factor scores generated from the 9, 14 and 27 items factor analysis with the colour of each element showing the correlation of factor scores across case individuals.

External validation

Next we set out to examine how the classes of patients identified differ regarding other depression related features to help us better understand the etiology of the subtypes identified. A number of external validators not used in the LCA were chosen to examine the classes identified by each of the three LCAs, including demographics, comorbidities, clinical

characteristics, and environmental risk factors. For the three classes identified in the nine items solution, the frequencies (for binary variables) as well as the means and standard deviations (for continuous variables) of a number of validators differed significantly across the three classes, mostly driven by the severe class. It was characterized by samples of older age with higher number of episodes and neuroticism scores, as well as higher frequencies of comorbid anxiety disorders including panic, phobia, and generalised anxiety disorder (GAD). Members of the severe class more frequently had a positive family history of depression. Notably, age at onset, body mass index (BMI) and a history of childhood sexual abuse did not discriminate classes.

The pattern of validators differed substantially in the 4 class solution with 14-items. The severe typical class had the highest rates of comorbidity with panic disorder, but the atypical class stood out more with the highest rates of GAD, dysthymia, and phobia, with the highest rates of childhood sexual abuse, family history, and stressful life events, the highest level of neuroticism, the earliest age at onset, most and longest episodes, and highest BMI. The non-suicidal class was characterized by low levels of comorbid anxiety disorder, relatively low neuroticism and the fewest and shortest episodes.

In general, the pattern of validation in the 27-class LCA bore substantial resemblance to that seen with the 14-item solution. The atypical class was most unique with the highest levels of dysthymia, phobia, family history, and childhood sexual abuse, the most frequent and longest duration of episodes, the youngest age and highest BMI, and neuroticism levels. The severe typical class had the highest panic comorbidity and was generally "second" behind the atypical class on other severity markers. The mild and non-suicidal low guilt classes were also the lowest on anxiety disorder comorbidity, neuroticism, childhood sexual abuse, family history, and the number and duration of depressive episodes.

9 items						
	C1	C2	C3	X-squared / F value	P value	Cramer's V / Eta.sq
Panic	7.36%	4.00%	2.38%	20.12	0.0000	0.058
GAD	27.32%	15.61%	15.08%	54.69	0.0000	0.096
Dysthymia	10.21%	6.69%	9.19%	6.30	0.0429	0.033
CSA	10.39%	8.67%	10.19%	1.39	0.4995	0.015
Phobia	41.20%	31.37%	28.31%	39.00	0.0000	0.081
FH	27.45%	24.20%	21.20%	8.57	0.0138	0.039
AAO	34.839 +- 9.779	35.678 +- 9.729	33.901 +- 9.453	0.57	0.4499	0.000
Neuroticism	13.199 +- 5.906	9.911 +- 5.386	9.843 +- 5.661	204.17	0.0000	0.034
SLE	1.588 +- 1.713	1.333 +- 1.524	1.485 +- 1.442	5.17	0.0230	0.001
DLE	48.486 +- 92.427	45.031 +- 78.445	43.07 +- 87.987	1.71	0.1904	0.000
Age	44.557 +- 8.871	44.16 +- 9.003	43.039 +- 8.722	10.34	0.0013	0.002
BMI	22.782 +- 3.411	22.567 +- 3.074	22.523 +- 3.501	3.35	0.0674	0.001
NE	4.733 +- 7.322	3.801 +- 6.083	3.575 +- 4.343	14.83	0.0001	0.002

14 items							
	C1	C2	C3	C4	X-squared / F value	P value	Cramer's V / Eta.sq
Panic	4.10%	6.90%	7.60%	4.13%	19.25	0.0002	0.057
GAD	21.89%	33.41%	27.06%	12.80%	68.20	0.0000	0.107
Dysthymia	11.50%	12.76%	9.74%	5.85%	15.93	0.0012	0.052
CSA	12.88%	13.46%	9.51%	9.18%	13.89	0.0031	0.049
Phobia	33.29%	45.96%	41.22%	30.71%	42.06	0.0000	0.084
FH	24.49%	29.41%	27.23%	24.55%	5.29	0.1515	0.030
AAO	33.312 +- 9.66	32.201 +- 9.703	35.231 +- 9.732	36.387 +- 9.528	55.16	0.0000	0.009
Neuroticism	11.995 +- 5.987	14.537 +- 5.298	13.08 +- 5.951	9.349 +- 5.186	12.95	0.0003	0.002
SLE	1.549 +- 1.65	2.032 +- 2.024	1.541 +- 1.666	1.336 +- 1.484	7.76	0.0054	0.001
DLE	49.357 +- 82.476	70.867 +- 122.683	45.826 +- 88.28	42.46 +- 92.183	7.53	0.0061	0.001
Age	42.955 +- 8.981	42.14 +- 8.556	44.869 +- 8.854	45.116 +- 8.643	47.59	0.0000	0.008
BMI	22.783 +- 3.24	23.822 +- 3.743	22.671 +- 3.389	22.406 +- 3.157	11.40	0.0007	0.002
NE	4.094 +- 5.76	5.421 +- 8.043	4.708 +- 7.366	3.647 +- 5.482	0.02	0.8809	0.000

27 items									
	C1	C2	C3	C4	C5	C6	X-squared / F value	P value	Cramer's V / Eta.sq
Panic	7.64%	9.04%	4.62%	2.34%	3.73%	7.05%	48.01	0.0000	0.090
GAD	33.72%	33.15%	19.79%	12.02%	15.15%	22.47%	180.81	0.0000	0.175
Dysthymia	15.63%	11.14%	9.78%	7.51%	5.83%	7.92%	38.97	0.0000	0.081
CSA	15.89%	9.30%	10.95%	8.77%	8.10%	11.03%	22.40	0.0004	0.062
Phobia	48.26%	46.55%	32.37%	28.07%	31.77%	36.78%	124.28	0.0000	0.145
FH	32.54%	28.69%	24.69%	22.39%	22.24%	26.70%	23.73	0.0002	0.064
AAO	31.957 +- 9.545	35.185 +- 9.744	34.91 +- 9.878	35.497 +- 9.808	36.322 +- 9.422	34.089 +- 9.662	1.65	0.1993	0.000
Neuroticism	15.076 +- 5.063	14.209 +- 5.892	11.621 +- 5.82	9.287 +- 5.44	9.284 +- 5.154	12.893 +- 5.583	206.21	0.0000	0.034
SLE	2.095 +- 2.029	1.671 +- 1.766	1.518 +- 1.642	1.449 +- 1.454	1.195 +- 1.389	1.403 +- 1.545	61.81	0.0000	0.011
DLE	72.134 +- 125.315	47.272 +- 87.939	47.039 +- 88.426	42.4 +- 87.766	36.351 +- 60.631	48.925 +- 99.094	8.59	0.0034	0.001
Age	41.9 +- 8.638	45.146 +- 8.729	44.724 +- 8.909	44.078 +- 8.991	44.843 +- 8.693	43.364 +- 9.087	4.25	0.0393	0.001
BMI	23.712 +- 3.846	22.827 +- 3.382	22.725 +- 3.373	22.67 +- 3.473	22.597 +- 3.377	22.301 +- 3.109	37.79	0.0000	0.006
NE	5.49 +- 8.102	5.086 +- 7.956	4.028 +- 6.052	3.642 +- 4.905	3.872 +- 5.674	4.457 +- 6.903	16.39	0.0001	0.003

Table 4. 5: External validation on the classes identified using latent class analysis (LCA) for nine Diagnostic and Statistical Manual for Mental Disorders (DSM)-IV major depressive disorder (MDD) A criteria three class solution; 14 items four class solution; and 27 items six class solution. GAD, generalised anxiety disorder; CSA, childhood sexual abuse; FH, family history; AAO, age at onset; SLE, stressful life events; DLE, duration of the longest episode; BMI, body mass index; NE, number of episodes. For 9 items LCA, C1, severe class; C2, non-suicidal class; C3, moderate class; for 14 items LCA, C1, severe typical class; C2, moderate typical class; C3, non-suicidal class; C4 atypical class; for 27 items data LCA, C1, severe typical class; C2, moderate typical class; C3, low suicidal/high guilt class; C4, non-suicidal/low guilt class; C5, atypical class; C6, mild class.

Summary

There are three major findings from the study. Firstly, there is significant heterogeneity within the DSM-IV nine MDD A criteria separating the cognitive symptoms of worthless/guilt and suicidal thoughts from the rest. This was further supported by the suicidal/hopeless factor emerged from the 27 items analysis reflecting Beck's cognitive triad and recurrent suicidal thoughts, as well as the non-suicidal classes emerged from all three LCAs. Secondly, when the bi-directional neurovegetative symptoms of appetite/weight changes and sleep disturbances were disaggregated in both the 14 and 27 items datasets, the somatic factors reflecting weight/appetite changes and sleep disturbances as well as the atypical classes, emerged from FA and LCA respectively, support the somatic dimensions and atypical subtype of depression. The external validation on latent classes suggests that patients in the atypical class seem to have the most severe depression. The 14 items FA also separates sleep disturbance from other neurovegetative symptoms, which indicates at least partially independent psychopathology. Thirdly, when including the melancholic and anxiety symptoms in the analysis, it is not evident whether there is a valid melancholic dimension/subtype in our data due to the weak loadings of the melancholic items on the general retarded depressive factor as well as the high invariant endorsements of those melancholic items across all LCA classes. FA suggests an agitation / anxiety dimension.
Discussion of these and other relevant points is left to the last chapter of the thesis.

Chapter 5

Rare variant analysis

Introduction

The focus of this chapter is to identify genes that are enriched with disease causal functional variants. Using whole genome sequence data at one fold (1X) coverage on 9,435 samples, I developed a novel bioinformatics pipeline to examine whether cases might be enriched with variants with and without quality filtering across the whole exome as well as in each gene. Moreover, I applied enrichment tests on functional SNPs with different levels of filtering, and allele frequencies down to 0.1%, to examine whether cases with MDD might be enriched for SNPs affecting gene function. Genes thus identified, could be used as candidates for target re-sequencing.

Results

Sample and variant description

The processing of sequence data was based on a total of 9,435 samples. For gene based analysis as well as gene enrichment test controlling for principal components, a total of 9,063 samples (4,287 cases + 4,776 controls) were actually included for analysis, after filtering out samples with missing phenotype, bad genotype quality based on imputation data, genetically related individuals (identified using identity by state (IBS) on imputation data) as well as ethnic minority samples (identified using principal

component analysis (PCA) on the imputation data). Results for IBS and PCA on the imputation data of complete samples are included in the next chapter.

With an average whole genome read depth of 1.02X across 9,435 individuals, an average of 58% of the 96,130,824 base pair positions in 254,986 exons in 21,946 genes were covered by at least one raw read and 55% filtered on reads with base quality above 30 and mapping quality above 55 (filter 1). The average number of variants called across all 9,435 individuals from raw reads was 79,825 (0.14% of average mapped positions) and 44,690 (0.08% of averaged mapped positions) for filter1 reads.

Table 5. 1 presents data on SNPs after indel filtering, SNP calling and annotation. The table shows the total numbers of synonymous (S) and non-synonymous (NS) SNPs that are singletons, extremely rare (<0.05%) and common (>0.05%) as well as the NS:S ratio in the three minor allele frequency (MAF) ranges. Filter 1 removed 45% (24-13M) NS and 56% (9-5M) S SNPs; filtering on SNPs supported by at least two reads (filter 2) removed 99% (166K) NS and 98% (110K) S SNPs of the filter 1 data.

In filter 1& 2 data, the NS:S ratio increases when MAF decreases, with the trend being less clear in the raw data where the ratio is the highest for common SNPs. Therefore, I plotted the NS:S ratio against alterative allele count (proxy for MAF), which showed that NS SNPs were enriched for low frequency alleles compared to S SNPs in all three levels of filtered datasets (Figure 5. 1), and the high NS:S ratio in common SNPs called from the raw reads was mostly likely due to sequencing errors.

Items	Raw	Filter1	Filter 2	Filter1/ Raw	Filter1 / Filter2
MQ positions (mean)	55,411,735	53,024,292		96%	
MQ sequence coverage (mean)	58%	55%			
Variant count (mean)	79,825	44,690		56%	
Missing (mean)	49,360	28,235	5,461	57%	19%
NS (mean)	14,661	6,158	994	42%	16%
Stopgain (mean)	526	152	6	29%	4%
Stoploss (mean)	27	9	1	34%	9%
S (mean)	8,788	5,479	1,167	62%	21%
Unknown (mean)	182	75	15	41%	20%
All SNPs (mean)	73,543	40,109	7,644	55%	19%
NS.all	24,035,108	13,148,602	166,075	55%	1%
S.all	9,027,174	4,887,137	110,138	54%	2%
NS:S ratio	2.66	2.69	1.51		
NS.singletons	3,328,776	7,554,639	111,506	227%	1%
NS.singletons/all	14%	57%	67%	415%	
S.singletons	1,048,865	2,577,523	63,479	246%	2%
S.singletons/all	12%	53%	58%		
Singletons NS:S ratio	3.17	2.93	1.76		
NS.rare < 0.05%	19,322,980	13,029,445	137,062	67%	1%
NS.rare/all	80%	99%	83%		
S.rare	7,924,177	4,822,992	81,023	61%	2%
S.rare/all	88%	99%	74%		
Rare NS:S ratio	2.44	2.70	1.69		
NS.common > 0.05%	4,712,128	119,157	29,013	3%	24%
NS.common/all	20%	1%	17%		
S.common	1,102,997	64,145	29,115	6%	45%
S.common/all	12%	1%	26%		
Common NS:S ratio	4.27	1.86	1.00		

Table 5. 1: Data summary. MQ. Mapping quality; NS, non-synonymous SNPs; S, synonymous SNPs; raw: variants called without quality filtering on the reads; filter 1: variants called from reads filtered on mapping quality > 55 and base quality > 30 as well as non-pair end reads; filter 2: SNPs called from at least two supportive reads after quality filtering.

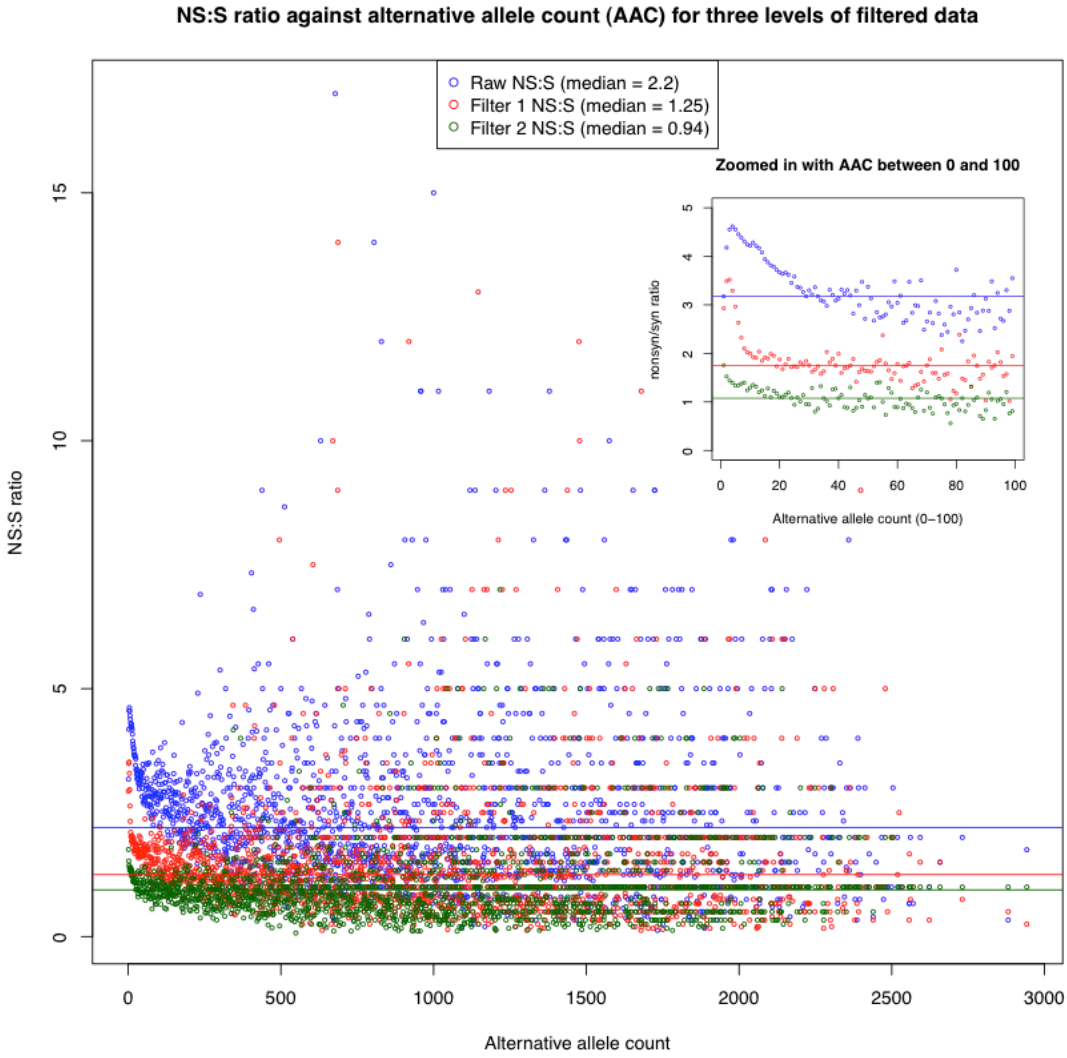


Figure 5. 1: Non-synonymous (NS) : synonymous (S) ratio against alternative allele count (AAC) for three levels of filtered data. Raw: variants called without quality filtering on the reads; filter 1: variants called from reads filtered on mapping quality > 55 and base quality > 30 as well as non-pair end reads; filter 2: SNPs called from at least two supportive reads after quality filtering.

Summary statistics on whole exome data

Table 5. 2 presents the summary statistics for regressing different exome variables against case control disease status controlling for age, the first 10 principal components (PCs) as fixed effects as well as city and sequencing plate as random effects in a mixed model framework. The PCs were generated from principal component analysis (PCA)

on 144K common (MAF > 1%) SNPs pruned for linkage disequilibrium (LD) and randomly selected in 100 SNP windows across the whole genome, which was used to correct for population stratification.

	Raw		Filter 1		Filter 2	
Items	**OR**	**P**	**OR**	**P**	**OR**	**P**
MQ_Positions	1.00	0.086	1.00	0.078		
MQ_Sum	1.00	0.074	1.00	0.152		
MQ_Mean	0.94	0.526	0.98	0.835		
BQ_Positions	1.00	0.079	1.00	0.065		
BQ_Sum	1.00	0.235	1.00	0.254		
BQ_Mean	1.13	0.106	1.29	0.039		
Variant count	1.00	0.200	1.00	0.273		
Reads	1.00	0.111	1.00	0.088		
Missing	1.00	0.455	1.00	0.578	1.00	0.042
NS	1.00	0.635	1.00	0.581	1.00	0.021
Stopgain	1.00	0.328	1.00	0.038	0.99	0.282
Stoploss	1.00	0.560	1.00	0.770	0.95	0.042
S	1.00	0.435	1.00	0.037	1.00	0.008
Unknown	1.00	0.524	1.00	0.244	0.99	0.036
Variants per positions	-	0.319	-	0.473		
Read depth	0.92	0.729	0.93	0.806		
NS per positions	-	0.516	-	0.673		
NS per depth	1.00	0.444	1.00	0.872		
NS+stoploss+stopgain	1.00	0.615	1.00	0.664		
NS+stoploss+stopgain per position	-	0.506	-	0.583		
NS+stoploss+stopgain per depth	1.00	0.428	1.00	0.761		
NS:S ratio	1.31	0.053	2.20	0.029	3.07	0.054

Table 5. 2: Summary statistics regressing variant related items against disease status in a mixed model framework controlling for principal components and age as fixed effects and city and sequencing plate as random effect. MQ, mapping quality; BQ, base quality; OR, odds ratio; NS, non-synonymous SNPs; S, synonymous SNPs. Raw: variants called without quality filtering on the reads; filter 1: variants called from reads filtered on mapping quality > 55 and base quality > 30 as well as non-pair end reads; filter 2: SNPs called from at least two supportive reads after quality filtering.

It has been previously demonstrated that rare variants show stratification that is systematically different from, and typically stronger than common variants (Mathieson & McVean 2012), hence controlling for PCs only might not capture all confounds in our data, especially for SNPs at low MAF. We know rare variants are sensitive to sequencing batch effects and that there is imbalanced sample size from different hospitals and cities; hence we controlled for sequencing plate and city as random effects

and carried out regression in a mixed model framework. Results showed that the number of variants between cases and controls do not significantly differ in raw and filter 1 data (Table 5. 2). The NS:S ratio is significant (with P value threshold < 0.05) in the filter 1 data (P = 0.02, OR = 2.20) as well as S SNPs (P = 0.03). In the filter 2 data, the number of NS SNPs (P = 0.02), S SNPs (P = 0.008) and stoploss (P = 0.04) all independently significantly differ between cases and controls with OR less than but close to 1. Regression analysis on NS:S ratio in filter 2 data resulted in P = 0.05, OR = 3.07. These results suggest that cases are enriched by true NS SNPs than controls.

Gene based analysis using two levels of filtering data

In addition to examining whether cases were enriched by the total number of exonic variants and deleterious SNPs compare to controls, I also examined whether the number of variants and different types of SNPs differed significantly between cases and controls in any of the 21,946 ensembl genes in the human genome. Figure 5. 2 contains QQ plots for all variants, all SNPs, NS and S SNPs (top) as well as those filtered on MQ, BQ and non pair end reads (bottom). There is one autosomal gene *TLR4* ($P_{ENSG00000136869}$ = 5.81, OR = 1.04) on chromosome 9 that exceeded the gene-wide threshold (-log10 P = 5.64) in the raw variant count analysis. However, a QQ plot shows that there is a large degree of inflation in the data; hence the result is likely to be a false positive.

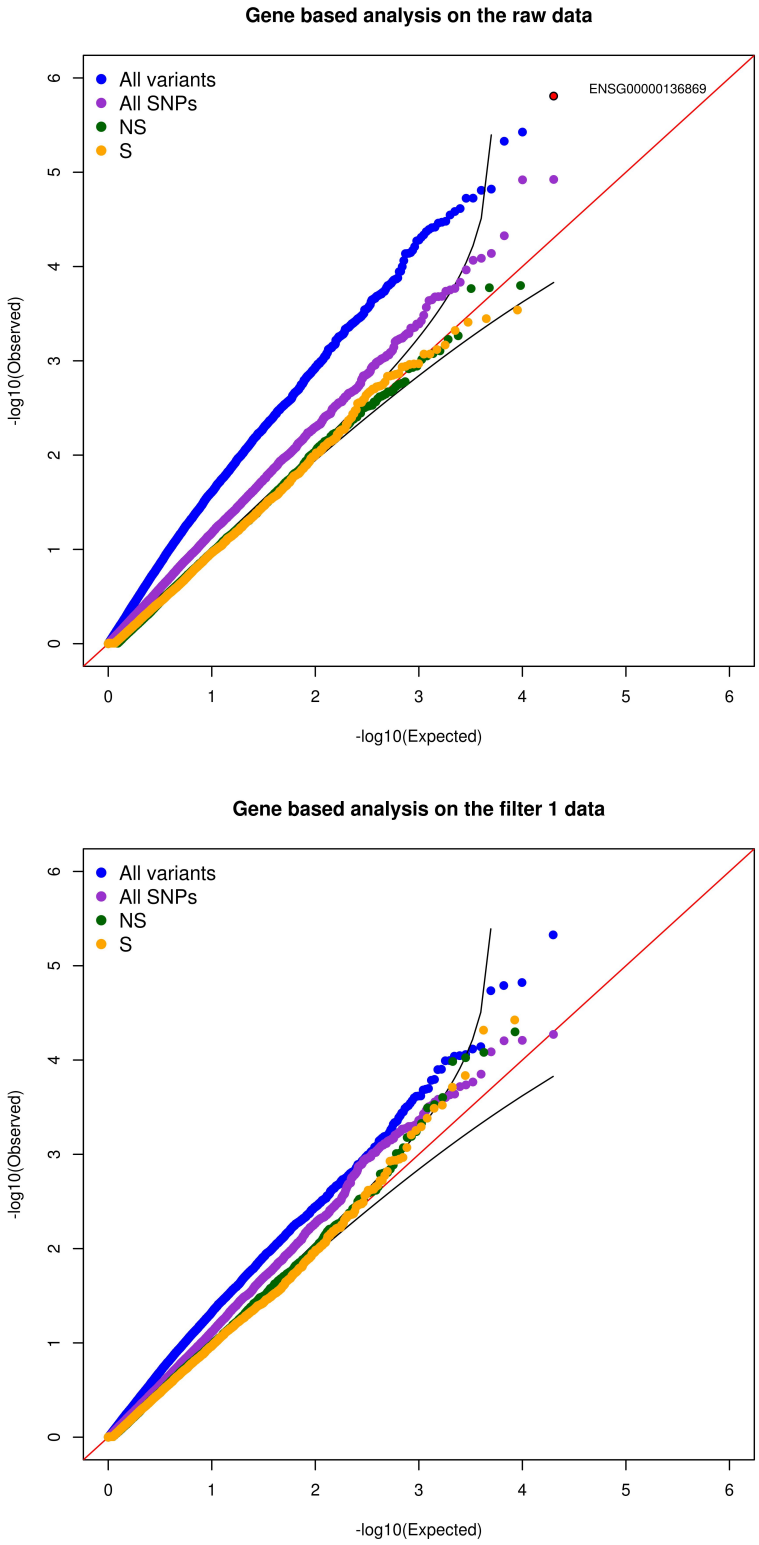


Figure 5. 2: Quantile-quantile plots of P values from regression analysis of all types of variants and SNPs as well as synonymous (s) and non-synonymous (NS) SNP counts per gene against disease status controlling for city as fixed effect removing the 13 protein coding mitochondrial genes. Gene-wide threshold -log10 (0.05/21,946 genes) = 5.64.

Gene based analysis on all SNPs and singletons filtered on two supportive reads in a mixed model framework

In addition to examining whether cases might be enriched with any type of variants and SNPs, I have also examined whether cases might be enriched with all SNPs and singletons supported by at least two reads. Excessive noise in our raw sequencing data would produce a large number of false rare variant calls. Therefore, I examined singletons supported by at least two reads to increase calling accuracy. This harsh filtering removed 98% of the total number of S and NS SNPs across all individuals and an average of 81% per individual.

As shown in Figure 5. 3 (top left), there is inflation for the P value distribution suggesting there are uncontrolled artefacts. I examined two sources of artefacts, including city and sequencing plate as random effects in a mixed model framework. As shown in Figure 5. 3 (bottom left), these resulted in a significantly reduced inflation. The reason for controlling those two sources of artefacts was because samples were neither evenly distributed nor matched between cases and controls in cities and sequencing plates. Figure 5. 3 showed that there are no genes significantly enriched with exonic SNPs and singletons supported by two reads when controlling for the 10 PCs with and without controlling for city and sequencing plate as random effects. In the singleton analysis, adding those random effects caused significant deflation possibly due to model over-fitting.

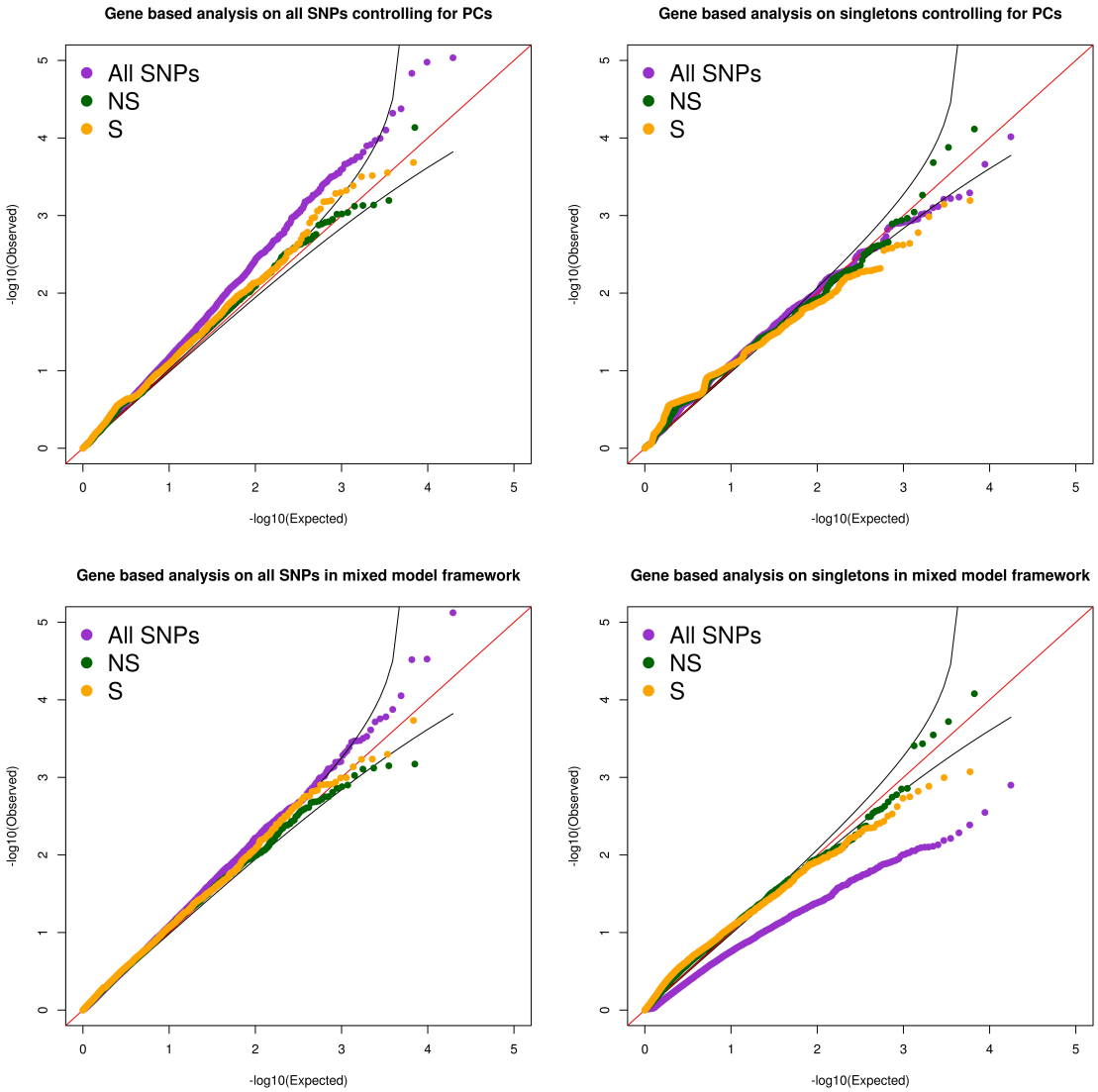


Figure 5. 3: Quantile-quantile plots of P values regressing SNP counts per gene against disease status based on SNPs supported by at least two reads. NS, non-synonymous SNPs; S, synonymous SNPs; the top two graphs are based on regression analysis controlling for the first 10 principal components (PCs), whereas the bottom two graphs are based on regression analysis in a mixed model framework controlling for the PCs as fixed effects, and city and sequencing plate as random effects; the left two graphs are based on all SNPs, whereas the right two graphs are based on singletons only.

Gene enrichment test on common SNPs

After regressing the aggregated number of variants and different types of SNPs per gene against disease status, I carried out a SNP test on the S as well as NS SNPs against disease status with different cut-offs for alternative allele count (as a proxy for MAF).

Subsequently I did a gene enrichment test by combining the P values of the SNP tests to see whether cases had any genes enriched with SNPs. This approach weighted for the significance of the disease causal effect for each SNP, which is an alternative to counting the total number of SNPs. For each gene, the number of (independently significant) NS and S SNPs as well as the combined P values of those SNPs using Fisher's method were calculated.

Using SNP data filtered on MQ, BQ and non pair end reads, controlling for PCs resulted in significant inflation (deviation from the orthogonal line), as shown in a QQ plot (Figure 5.4 top). This suggested uncontrolled artefacts other than population structure. I examined two possible reasons for inflation, including city and sequencing batch effects by permuting disease status within city (Figure 5.4 middle) and sequencing plate (Figure 5.4 bottom) respectively. Gene enrichment results showed that permuting disease status within city removed the QQ plot inflation. This suggested that there were significant batch effects due to where the patients were recruited. Permuting disease status within sequencing plate reduced the degree of inflation and resulted in two significant hits *PABPC3* (-log10 $P_{ENSG00000151846}$ = 11.26) and *CNR2* (-log10 $P_{ENSG00000188822}$ = 9.48), which suggested that there were also sequencing batch effects. Hence, gene enrichment tests on three levels of filtered data were conducted in a mixed model framework by combining P values of SNP regression analysis controlling for PCs as fixed effects, and city and sequencing plate as random effects (Figure 5. 5).

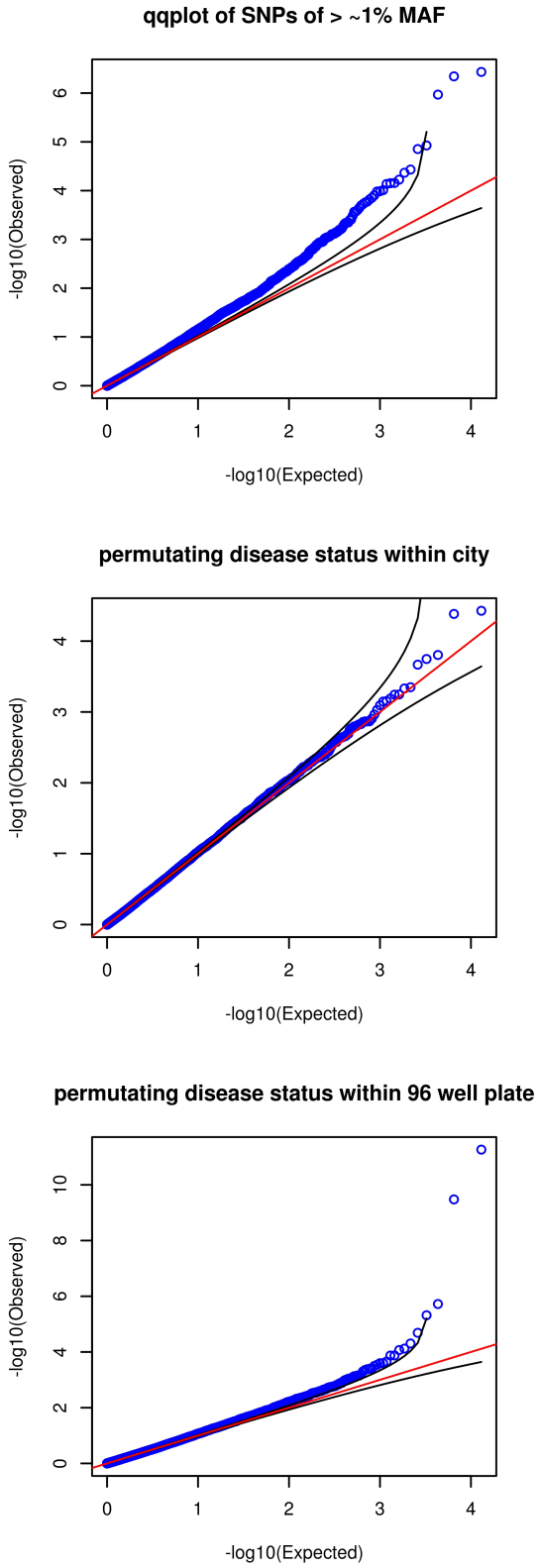
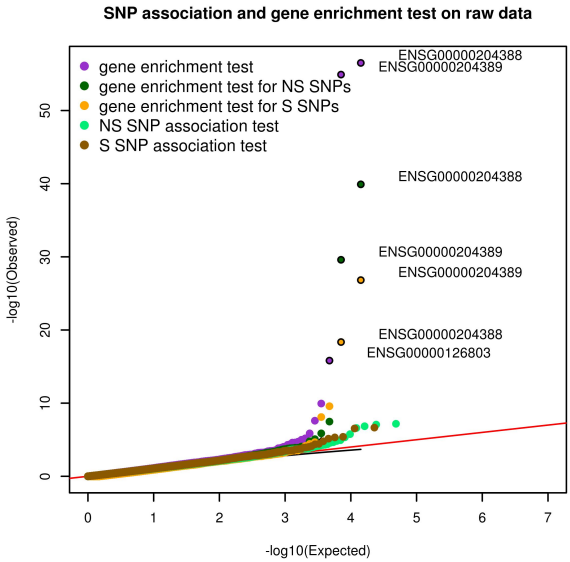
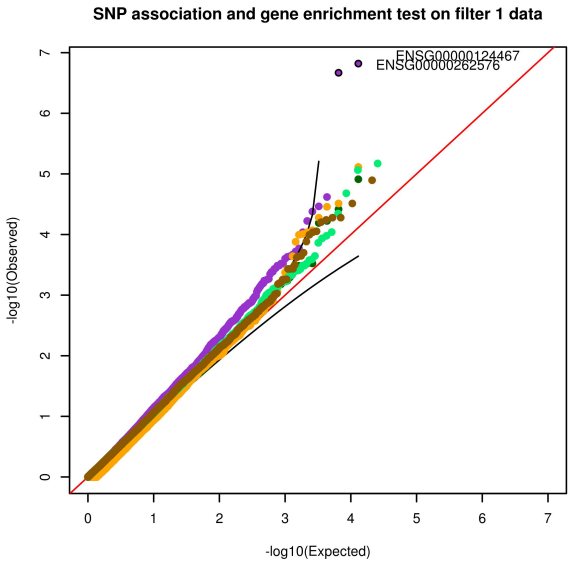


Figure 5. 4: Permutation tests. QQ plots of P values generated from gene based analysis on disease status (top) as well as that permuted within sample origin (middle) and sequencing plate (bottom) controlled for principal components.


SNP association and gene enrichment test on raw data
gene enrichment test
gene enrichment test for NS SNPs
gene enrichment test for S SNPs
NS SNP association test
S SNP association test
ENSG00000204388
ENSG00000204389
ENSG00000204388
ENSG00000204389
ENSG00000204389
ENSG00000204388
ENSG00000126803
-log10(Observed)
-log10(Expected)


SNP association and gene enrichment test on filter 1 data
ENSG00000124467
ENSG00000262576
-log10(Observed)
-log10(Expected)

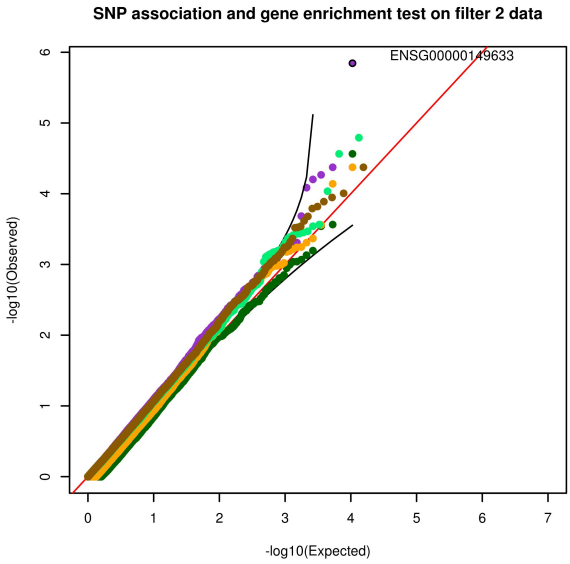

SNP association and gene enrichment test on filter 2 data
ENSG00000149633
-log10(Observed)
-log10(Expected)

Figure 5. 5: QQ plots of P values from common (MAF > ~ 1%) synonymous (S) and non-synonymous (NS) SNP association tests in a mixed model framework controlling for 10 principal components as fixed effects and city and sequencing plates as random effects, as well as QQ plots on combined P values from gene enrichment tests on S and NS SNPs separately and combined. The left graph is based on SNPs called from raw reads with the total alternative allele count (AAC) greater than 95 (MAF proxy > 1%) across all SNP sites; the significant threshold for gene enrichment tests using Fisher's method of combining P values of SNP association tests is –log10 (0.05 / 14,185 genes) = 5.45; the significant threshold for NS SNP association test is –log10 (0.05 / 48,614 NS SNPs) = 5.99; the significant threshold for S SNP association test is –log10 (0.05 / 22,827 S SNPs) = 5.66; the middle graph is based on SNPs called from reads filtered on mapping quality (>55) and base quality (>30) and non pair end reads with the total AAC greater than 90 across all SNP sites; the significant threshold for gene enrichment tests using Fisher's method of combining P values of SNP association tests is –log10 (0.05 / 13,007 genes) = 5.42; the significant threshold for NS SNP association test is –log10 (0.05 / 25,596 NS SNPs) = 5.71; the significant threshold for S SNP association test is –log10 (0.05 / 21,057 S SNPs) = 5.62; the right graph is based on SNPs supported by at least two reads with the total AAC greater than 90 across all SNP sites; the significant threshold for gene enrichment tests using Fisher's method of combining P values of SNP association tests is –log10 (0.05 / 10,590 genes) = 5.33; the significant threshold for NS SNP association test is –log10 (0.05 / 13,264 NS SNPs) = 5.42; the significant threshold for S SNP association test is –log10 (0.05 / 15,546 S SNPs) = 5.49.

I initially looked at SNPs with MAF proxy greater than 1%, with the total number of alternative alleles per exonic site selected to be greater than 95 for the raw data, and 90 on filter 1 and filter 2 data. Among the total of 24M NS and 9M S SNPs in the raw data, SNP count cut-off 95 (MAF ≈ 1%) resulted in 48,614 NS and 22,827 S SNPs selected for a SNP test. This resulted in 4 significant NS SNPs in the *HSPA1B* gene on chromosome 6 with –log10 P ranges between 6.60 and 7.18. S SNP test yielded two significant hits on the same genes, with –log10 P values of 6.54 and 6.66 respectively. A gene enrichment test combining both S and NS SNPs identified 6 significant genes. The top three scoring genes were heat shock 70kDa protein genes (*HSPA1B*, *HSPA1A* and *HSPA2*), with highly significant –log10 P values of 57, 55 and 16 (Table 5. 3). The other above gene-wide threshold (-log10 P = 5.45) hits include *PCDHGA4*, *FLG2* and *ZNF705A*.

Analyses	S.sig	S.n	S.sum. logp	NS.sig	NS.n	NS.sum. logp	S.fisher	NS.fisher	fisher. combined	hgnc_symbol	Chr	Str
Raw (common)	15	44	261.26	26	62	463.21	18.35	39.91	56.51	HSPA1B	6	1
Raw (common)	21	48	330.73	23	58	383.37	26.82	29.60	54.91	HSPA1A	6	1
Raw (common)	7	15	105.32	9	14	88.56	9.58	7.48	15.82	HSPA2	14	1
Raw (common)	**8**	**26**	**131.52**	**10**	**49**	**155.68**	**8.10**	**3.73**	**9.95**	**PCDHGA4**	**5**	**1**
Raw (common)	3	15	63.98	8	34	130.32	3.53	5.08	7.60	FLG2	1	-1
Raw (common)	0	0	0.00	5	12	71.36	0.00	5.87	5.87	ZNF705A	12	1
Filter 1 (common)	1	3	22.52	2	4	36.85	3.01	4.91	6.82	PSG8	19	-1
Filter 1 (common)	**6**	**27**	**111.17**	**7**	**32**	**101.47**	**5.11**	**2.70**	**6.67**	**PCDHGA4**	**5**	**1**
Filter 1 (rare)	2	10	39.07	5	12	64.86	2.18	4.89	6.03	NLRP1	17	-1
Filter 2 (common)	3	4	32.61	1	3	21.12	4.14	2.76	5.84	KIAA1755	20	-1
Filter 2 (rare)	1	1	13.77	2	2	24.18	2.99	4.13	5.94	OR13C9	9	-1

Table 5. 3: Top hits for gene enrichment test combining P values for common SNP analysis in raw data and common and rare SNPs analysis in filter 1&2 data. NS, non-synonymous SNPs; S, synonymous SNPs; S.sig, number of significant S SNPs; S.n, number of S SNPs; NS.sig, number of significant NS SNPs; NS.n, number of NS SNPs. S.sum.logp, the sum of the –log10 P of the S SNP association tests; NS.sum.logp, the sum of the –log10 P of the NS SNP association tests; S.fisher, combining –log10 P values of S SNP association test using Fisher's method; NS.fisher, combining –log10 P values of NS SNP association test using Fisher's method; Chr, chromosome; Str, strand; Raw: variants called without quality filtering on the reads; filter 1: variants called from reads filtered on mapping quality > 55 and base quality > 30 as well as non-pair end reads; filter 2: SNPs called from at least two supportive reads after quality filtering.

Among the 13M NS and 4M S SNPs filtered on BQ, MQ and non pair end reads, 25,596 NS and 21,057 S SNPs remained after applying the 90 SNP count cut-off. No NS or S SNPs exceeded the multiple testing thresholds of 5.71 and 5.62 respectively, determined by using a P value of 0.05 divided by the number of SNP tests. Gene enrichment tests identified two genes exceeding a gene-wide threshold (-log10 P = 5.42): *PSG8* and *PCDHGA4*. The latter was also found from analysis of raw data (Table 5. 3).

The signals from the heat-shock protein genes were removed after level 1 filtering. Using the sequence viewer Integrative Genomics Viewer (IGV) (Figure 5. 6) I found that there was a larger than expected number of reads mapping to the heat shock protein gene region. There was also a high degree of sequence variation, an indication of sequenced reads wrongly mapped to the region. In the last section of the result chapter, I explain what might be the cause of this finding.

Figure 5. 6: Integrative Genomics Viewer (IGV) on *HSPAIB* region on chromosome 6.

When applying the double read filter, no NS or S SNPs exceeded the multiple testing threshold. Gene enrichment test resulted in one borderline significant gene (*KIAA1755* on chromosome 20, -log10 P = 5.84) that exceeded the gene-wide threshold (-log10 P = 5.33). These genes need further validation.

Gene enrichment test on common and rare SNPs in the mixed model framework

As stated above, common SNPs (MAF > 1%) only accounted for a very small proportion of the total exonic SNPs in our population. Therefore, in addition to conducting gene enrichment tests on common SNPs, I also examined SNPs between 1% and 0.1% MAF called from reads filtered on MQ, BQ and non pair end reads (Figure 5. 7 top) as well as SNPs supported by at least two reads (Figure 5. 7 bottom). Gene enrichment test using common SNPs (MAF > 1%) called from reads filtered on MQ, BQ and non pair end reads identified two significant genes *PSG8* and *PCDHGA4*, whereas the same test using rare SNPs gave a different hit: *NLRP1* (-log10 P = 6.03) on chromosome 17. In addition to the *KIAA1755* gene identified by common SNP analysis, gene enrichment tests on rare SNPs identified one significant hit: *OR13C9* (-log10 P = 5.94). Again, without further validation it is unclear whether any of these genes were truly enriched with disease causal variants

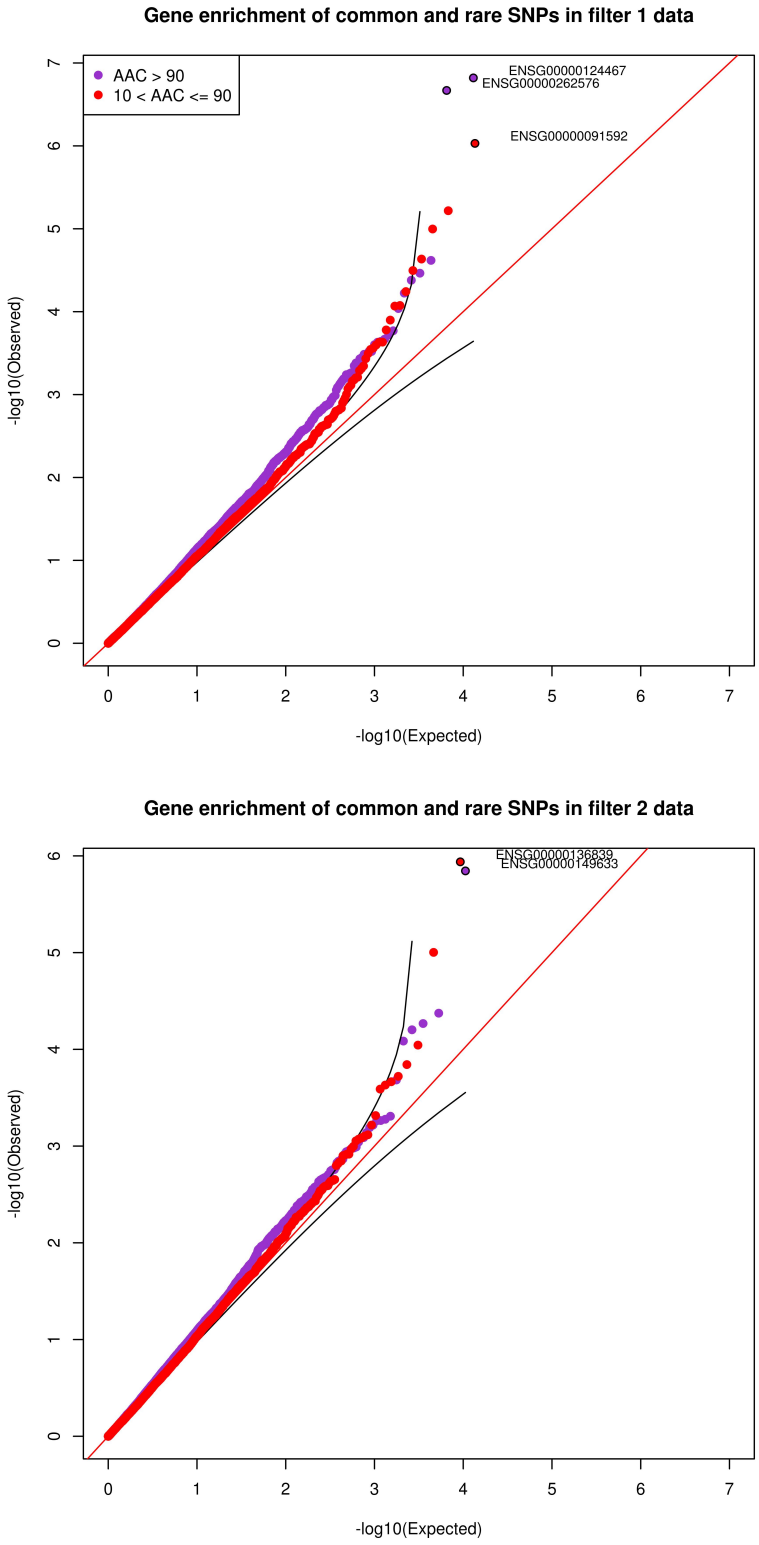


Figure 5. 7: Gene enrichment test on SNPs of two alternative allele count (AAC, proxy for MAF) ranges in filter 1 & 2 data. Filter 1, mapping quality (>55), base quality (>30) and non-pair end reads filter; filter 2, SNPs supported by at least two reads.

Heat shock protein signals are driven by mitochondrial genome read depth

Gene enrichment test combining P values of SNP association test against disease status using SNPs called from the raw reads have resulted in three heat shock protein genes (*HSPA1B, HSPA1A, HSPA2*) that far exceeded the gene wide significant threshold. The fact that the signals disappeared after applying read filters suggested that it was due to wrong mapping of reads. The gene-based analysis excluded the 13 protein coding mitochondrial (MT) genes, which showed highly significant results in the raw data. The explanation for this finding was confirmed to be due to differences in the copy number of the MT genome, evidence for which is was provided in chapter 7. Here I examine whether the heat shock protein signals had anything to do with also highly significant MT copy number difference between cases and controls. Instead of regressing the SNPs called from raw reads against disease status, I regressed them against MT read depth. The P values from SNP association tests were subsequently combined using Fisher's method. Here we show evidence that the heat shock protein signals reported from the gene enrichment analysis were driven by the incorrect mapping of MT reads to the heat shock protein genes in the autosomal human genome (Figure 5. 8).

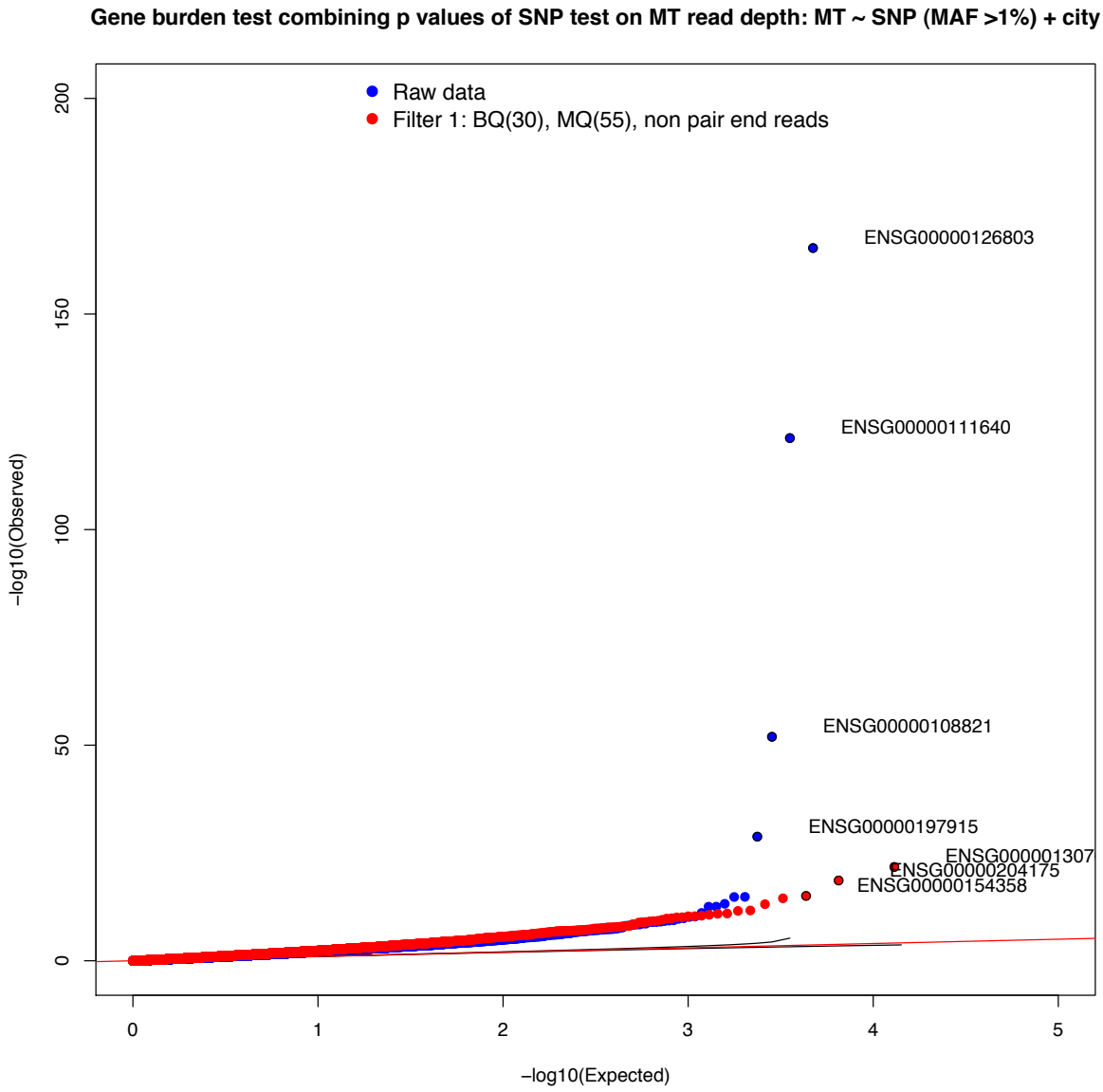


Figure 5. 8: Gene burden test combining P values of SNP association test on MT read depth in raw and filter 1 data.

Summary

In this chapter I have attempted to use low pass sequencing data to identify rare variants that contribute to susceptibility to MDD. The challenge I faced was to identify a signal attributable to a difference in the frequency of true variants, using a sequencing technology that has a high false positive rate. I addressed this problem by assuming that

the error rates in cases and controls would be the same. I tested for differences in the number of variants lying within a gene in cases and in controls. Rather than testing for differences in the frequency of individual alleles between cases and controls, my analyses summed the total contribution of all variants within a gene (thus I am testing for the effect of functional variants). Overall I was unable to find any convincing evidence that functional variants, rare or common, contributed to disease susceptibility. Discussion of these points is left to the last chapter of the thesis.

Chapter 6

GWAS of MDD

Introduction

In this chapter, I present the preliminary results for a genome-wide association study (GWAS) using 11,651 samples. Results in this chapter were originally presented based on 9,300 samples for initial thesis submission, which were updated with the additional samples after the *viva*. Our study differs from currently reported GWAS on MDD, in that instead of using genotype chip data that was designed to capture known common loci, we used imputed genotype probabilities from low pass (~1X) sequencing data, that not only cover common SNPs (minor allele frequency (MAF) > 5%) with high accuracy (imputation vs. genotyping concordance rate >92%) but also capture novel rare variants (MAF = 0.5%-5%) with modest accuracy. Additionally, the way our samples were recruited should reduce known causes of heterogeneity, which should thereby increase power to detect disease susceptible variants. Association analyses were not only performed on depression disease status, but also depression related phenotypes such as neuroticism. I also performed separate analyses with different case subsets: i) endorsed in the *typical severe* LCA class (see chapter 4); ii) with postnatal depression; iii) with low environmental risks; iv) with melancholia; v) earlier age at onset (lower 60%), and vi) high recurrence (≥ 4). I also performed genetic association against eight factor scores from factor analysis on three datasets of different symptom profiles in cases only (see chapter 4).

Results

A total of 11,673 CONVERGE samples (5,867 cases + 5,784 controls + 22 missing phenotypes) were sequenced at an average 1.2X coverage of the nuclear genome. The complete genotypes of each sample were imputed at 13 billion SNPs. There were 29 samples failed imputation, possibly due to quality issues unsolved during upstream processing. This resulted in 11,644 samples with valid imputation data for subsequent identity by state (IBS) and principal component analysis (PCA). These results were added upon examiner's request during the *viva*.

For SNP linkage disequilibrium pruning, I tried two sets of parameters, one with window size=100, shift size=10, r^2=0.7, the other with a smaller window and shift size but more stringent r^2 (50, 5, 0.05) filter. The former resulted in 4,878,535 SNPs and the latter 3,598,384 SNPs. I decided to use the latter SNP set for further MAF-based SNP pruning and PED file generation. The result SNP set contains 136,277 autosomal SNPs .

PCA was performed on the CONVERGE samples augmented with 1000 Genomes data (1,092 samples), as well as on the CONVERGE data alone. To avoid strand problems (i.e. where genotypes on opposing strands were reported in CONVERGE vs. 1000 Genomes), AT/GC sites were taken out when merging with the 1000 Genomes data for calculating PCA. The merged set contains a total of 81,195 autosomal SNPs with the same filtering parameters as the CONVERGE data.

Principal components analysis and identity by state

In the merged dataset, PCA showed that the CONVERGE samples clustered with the CHS (Han Chinese South), CHB (Han Chinese in Beijing) and JPT (Japanese) as shown in figure 6.1.

Figure 6.1: Principal component analysis on CONVERGE + 1000 Genomes SNP data colour coded by ethnic origin. ASW, HapMap African ancestry individuals from SW US; CEU, Centre d'Etude du Polymorphisme Humain (CEPH) individuals; CHB, Han Chinese in Beijing; CHS, Han Chinese South; CLM, Colombian in Medellin, Colombia; FIN, HapMap Finnish individuals from Finland; GBR, British individuals from England and Scotland; IBS, Iberian populations in Spain; JPT, Japanese individuals; LWK, Luhya individuals; MXL, HapMap Mexican individuals from LA California; PUR, Puerto Rican in Puerto Rico; TSI, Tuscan individuals; YRI, Yoruba individuals.

Within the CONVERGE samples only, there are some outliers with unusual ethnicity. With a cut-off of PC1 > 0.0125 (as shown in figure 6.2), there are 46 outliers

with 43 from Shaanxi, one from Xizang (Tibet), one from Sichuan, and one from Zhejiang, based on their self reported origin.

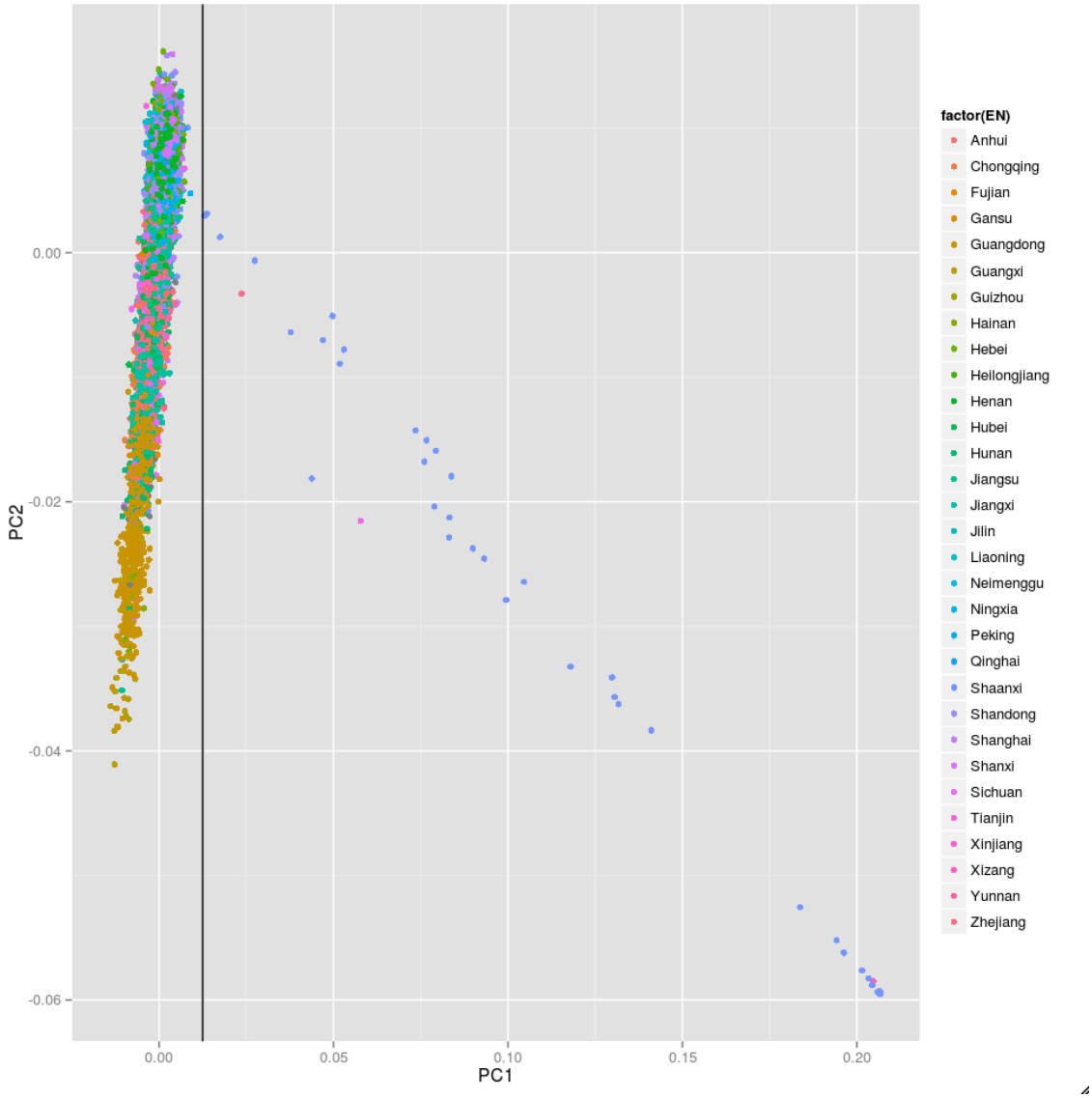


Figure 6.2: Principal component analysis on CONVERGE data only colour coded by the city of origin.

To examine the genetic relatedness among our samples, I performed IBS on all samples and identified 19 genetically identical samples with distinct patient IDs in the phenotypic database. The samples were recruited from the same hospital and clinical data were entered into the same computer. These samples were removed from the subsequent analysis. Samples with other degrees of relatedness (proportion of IBD >

0.3) were also removed from the subsequent analysis indicated as those deviates from the diagonal line in figure 6.3.

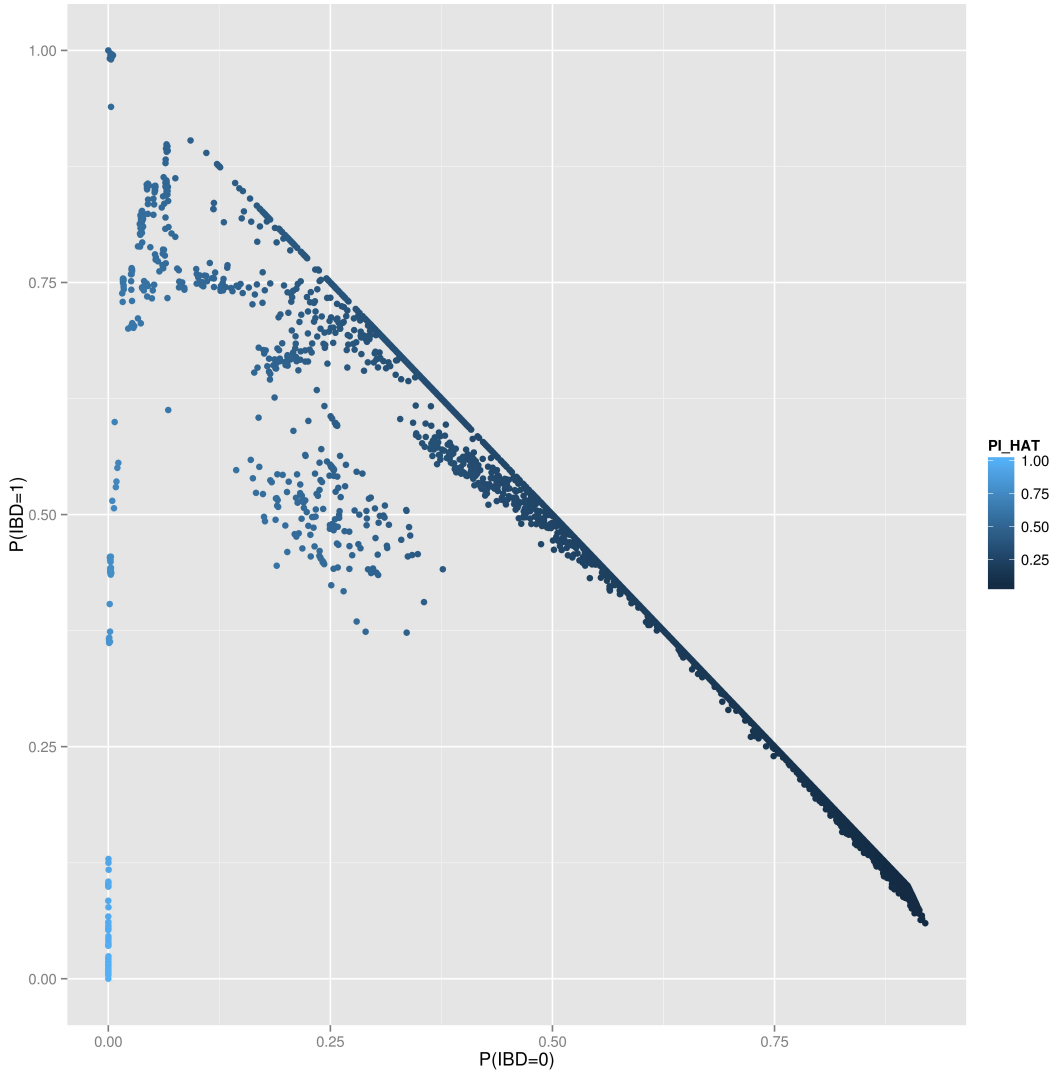


Figure 6.3: Identity by state. PI_HAT, proportion identity by descent (IBD); PI_HAT<0.05 were removed for plotting.

Overall, a total of 300 samples (120 cases + 180 controls) were excluded due to genetic relatedness or ethnicity, resulting in a total of 11,341 samples for the second round of PCA. Figure 6.4 shows a plot of PC1 against PC2, colour coded by latitude derived from city of origin. As can be seen, PC1 captured the transition between North and South. The sample clustering captured by PC2 requires further exploratory work.

Association analysis was performed controlling for either just the first PC or for the first 10 PCs generated from the second round of PCA. Results showed that controlling for the first 10 PCs removed more inflation from the QQ plots, possibly because they capture other sequencing artefacts, hence only these results are presented here (data not shown).

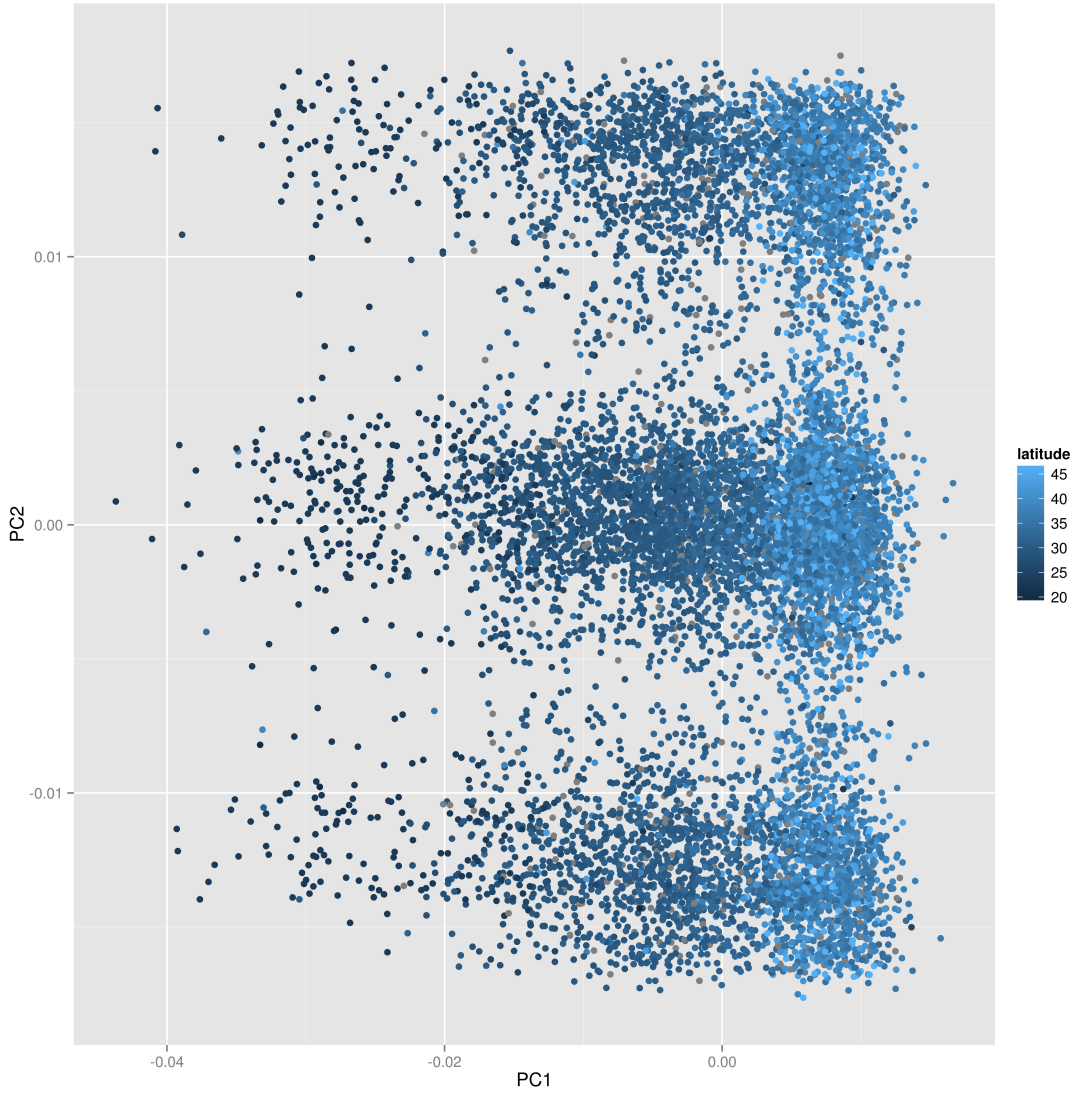


Figure 6.4: Principal component analysis on CONVERGE data only removing genetically related individuals.

Association analysis on height and BMI

Genome wide association (GWA) was carried out on height and body mass index (BMI) controlling for the 10 PCs using 9,417 the samples with non-missing phenotype information. These results were added upon examiner's request during the *viva*. Figure 6.5 presents the Manhattan plots of the P values. Only one SNP (rs182438096) on chromosome 6 exceeded the genome wide significant threshold of 5×10^{-8} for BMI. For the GWA analysis on height, there are two clusters of SNPs exceeding a genome wide significance threshold in a 50 Mb distance. Figure 6.6 and Figure 6.7 contains zoomed-in plots for the top hits of the two regions. Figure 6.8 presents the QQ plots of the P values including the 95% confidence interval. Annotation on the top SNPs is presented in appendix I.

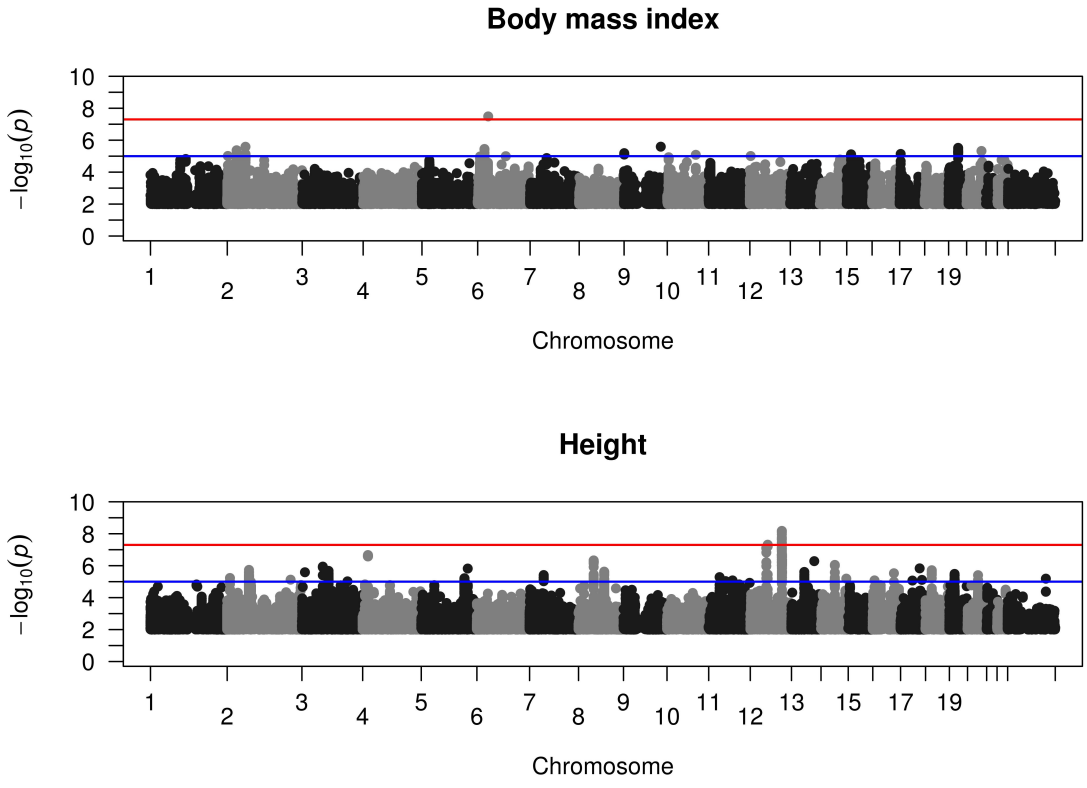


Figure 6.5: Manhattan plots of the association analysis on body mass index and height.

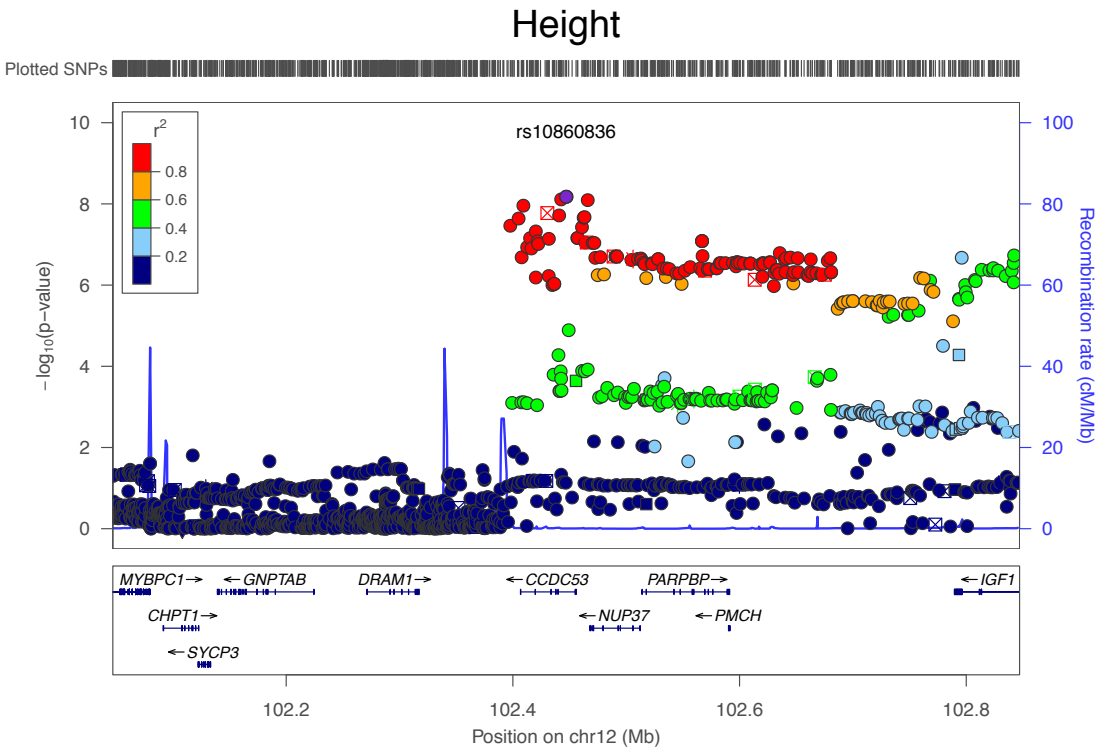


Figure 6.6: Zoom in plot for top hits region for association analysis on height. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

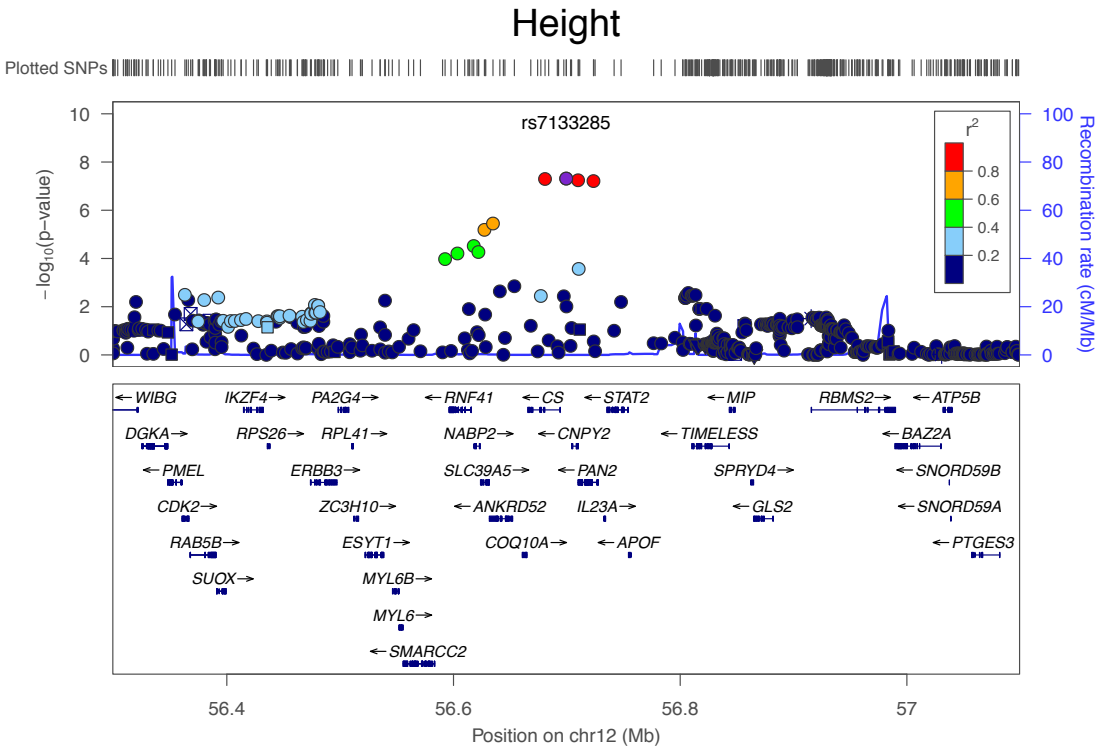


Figure 6.7: Zoom in plot for top hits region for association analysis on height. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

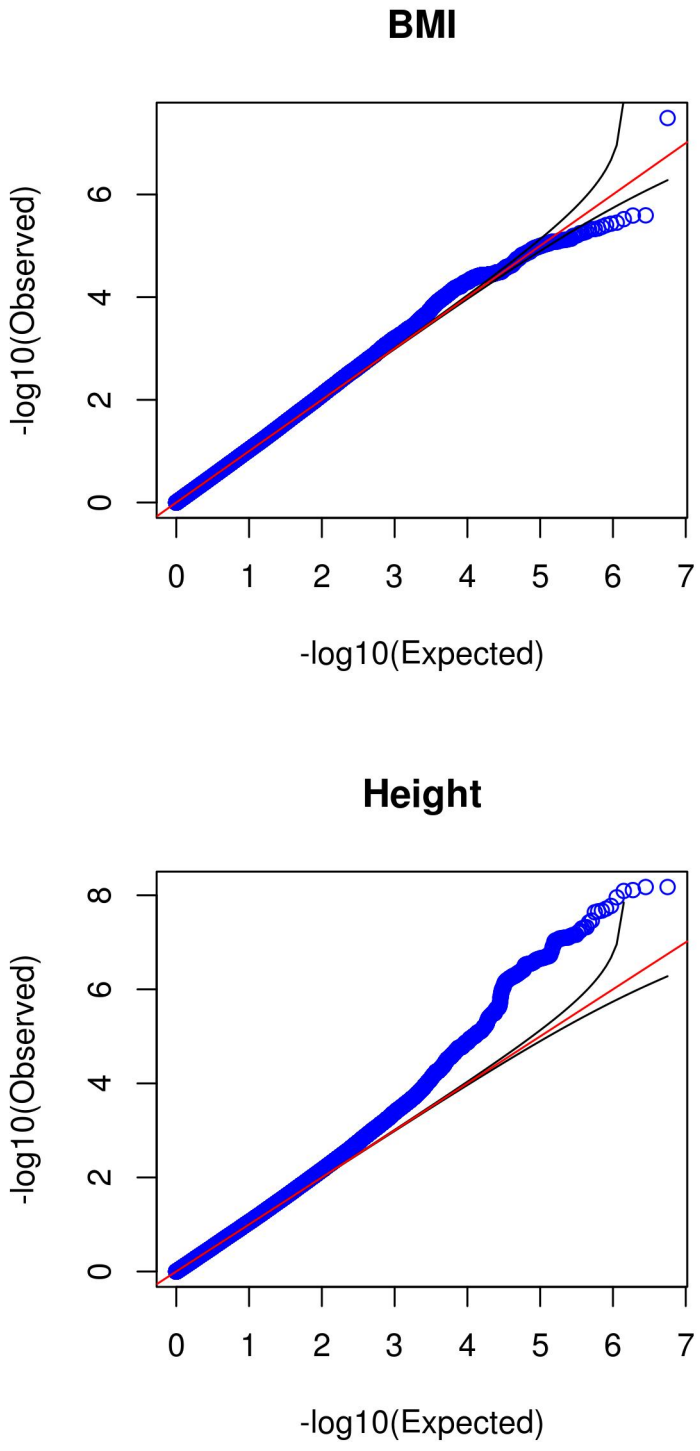


Figure 6.8: QQ plots of body mass index (BMI and height). Black lines indicated 95% confidence interval.

Association analysis on MDD case (subsets) and control status

GWA analysis was carried out on depression disease status controlling for 10 PCs to correct for population structure. GWA analysis was also carried out on neuroticism and a number of derived binary phenotypes. These phenotypes consist of all controls (scored 0) and different subsets of cases (scored 1), which would potentially increase sample homogeneity at the cost of reduced sample size. The subsets of cases include only cases that belong to the severe typical class from the latent class analysis described in chapter 4, meet diagnostic criteria for melancholia, postnatal depression, with low environmental risk, early age of onset or high recurrence. Binary diagnoses of melancholia and postnatal depression were based on DSM-IV criteria. Information on postnatal depression was assessed using an adaptation of the Edinburgh Scale. The stressful life event section assessed 16 traumatic events and the age of their occurrence. Cases with low environmental risk include all those who reported no history of childhood sexual abuse (involving genital contact or intercourse), rape and the lower 80% of the number of stressful life events (minus the rape item). Cases with early age at onset included only those in the lower 60%, whereas cases with high recurrence include only those that reported equal to or higher than four episodes. Table 6.1 presents the number of samples and the proportions of cases included in each derived binary phenotype.

Phenotypes	Case	Case subset proportion	Control	Total
Disease status	5867	100.00%	5784	11651
Postnatal depression	975	16.62%	5783	6758
Melancholia	4993	85.10%	5783	10776
Low environmental risk	3971	67.68%	5783	9754
Early onset	3669	62.54%	5783	9452
High recurrence	1779	30.32%	5783	7562
LCA severe (14 items)	3929	66.97%	5783	9712

Table 6. 1: Number and proportions of case subsets selected for genome wide association analysis. LCA, latent class analysis.

I performed a GWA on eight phenotypes. Therefore the genome-wide significance threshold ($P = 5\times10^{-8}$) should be lowered to account for multiple testing ($P = 6.25\times10^{-9}$). Figure 6.9 presents Manhattan plots regressing SNP dosages against each of those eight phenotypes, removing SNPs with $-\log_{10}(P) < 2$ and MAF < 5% (due to probable imputation inaccuracy). For GWA on disease status, as well as a number of binary phenotypes with different case selection criteria such as early age at onset, severe LCA class, low environmental risks and melancholia, there is one SNP (rs7257142) on chromosome 19 that exceeded the genome wide significance threshold. Figure 6.10 shows the zoomed-in plot for that particular SNP on disease status. Detailed information including chromosomal position, MAF, BETA, P value and SNP annotation are included in appendix I. Furthermore, there is a cluster of SNPs on chromosome 10 that exceeded genome wide significance threshold for cases selected with high recurrence, which overlapped with the top SNPs for melancholia. Figure 6.11 and Figure 6.12 show the zoomed in plots on the same region for those two phenotypes. Figure 6.13 presents the QQ plots of the P values of the association analysis of the eight phenotypes. For GWA on high recurrence, the cluster of significant SNPs clearly deviates from the 95% confidence interval.

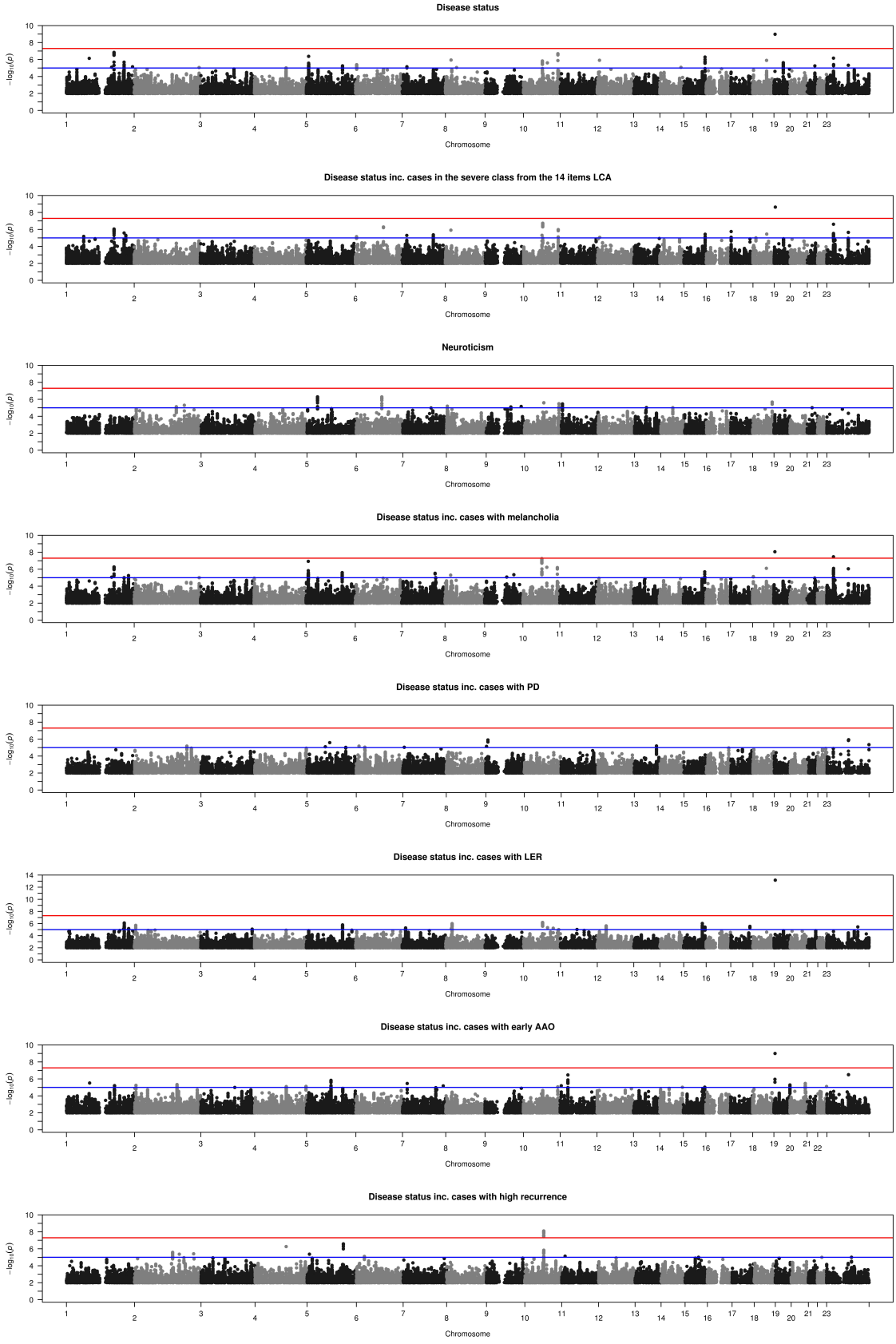


Figure 6.9: Manhattan plots on eight depression related phenotypes. LCA, latent class analysis; PD, postnatal depression; LER, low environmental risks; AAO, age at onset.

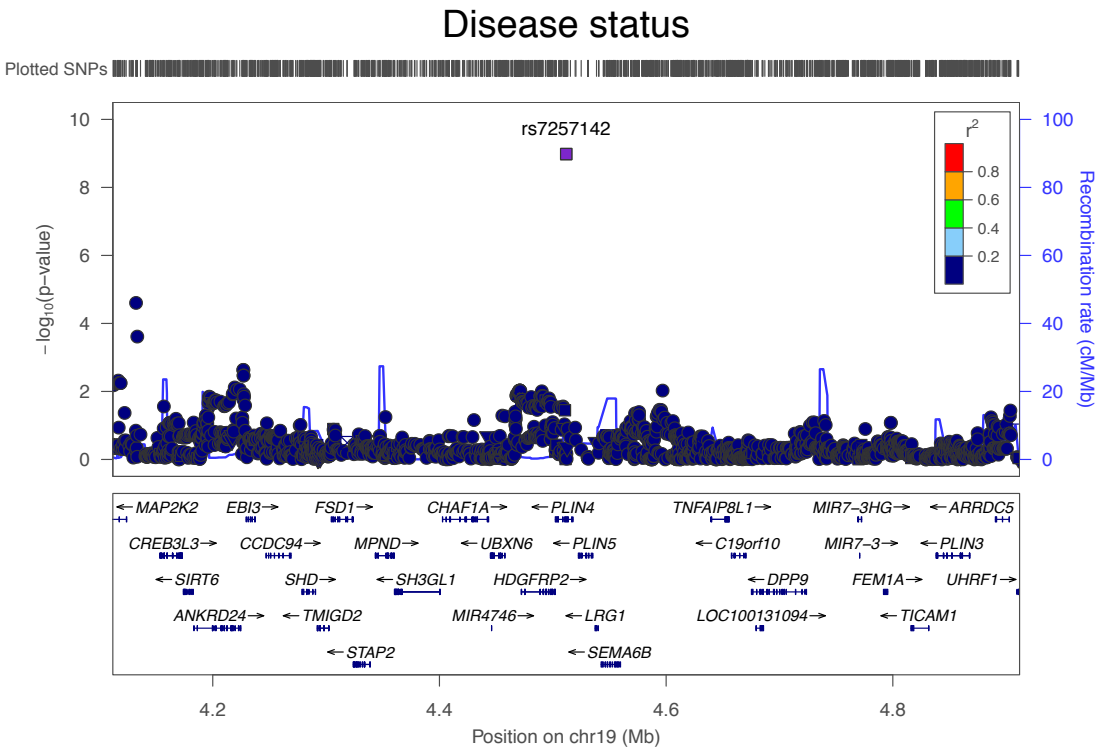


Figure 6.10: Zoom in plot for top hit for association analysis on disease status. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

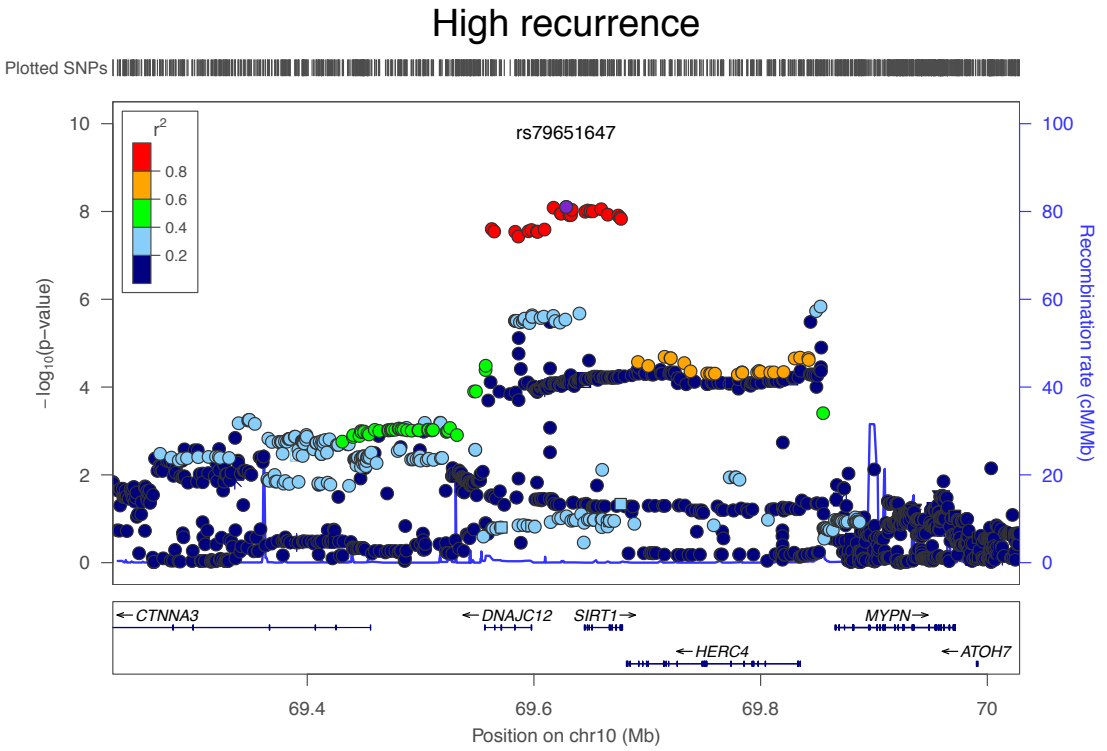


Figure 6.11: Zoom in plot for top hit for association analysis on disease status including only cases with >= 4 episodes. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

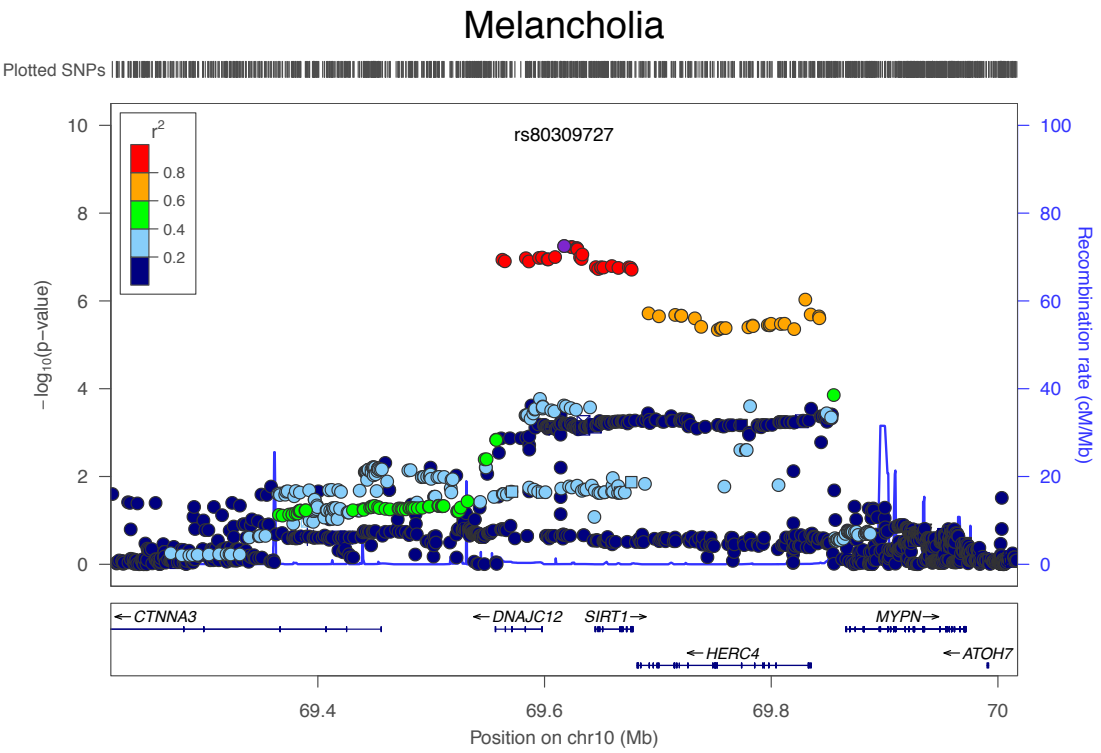


Figure 6.12: Zoom in plot for top hits region for association analysis on controls and case subsets including those with high recurrence (top) and melancholia (bottom). Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

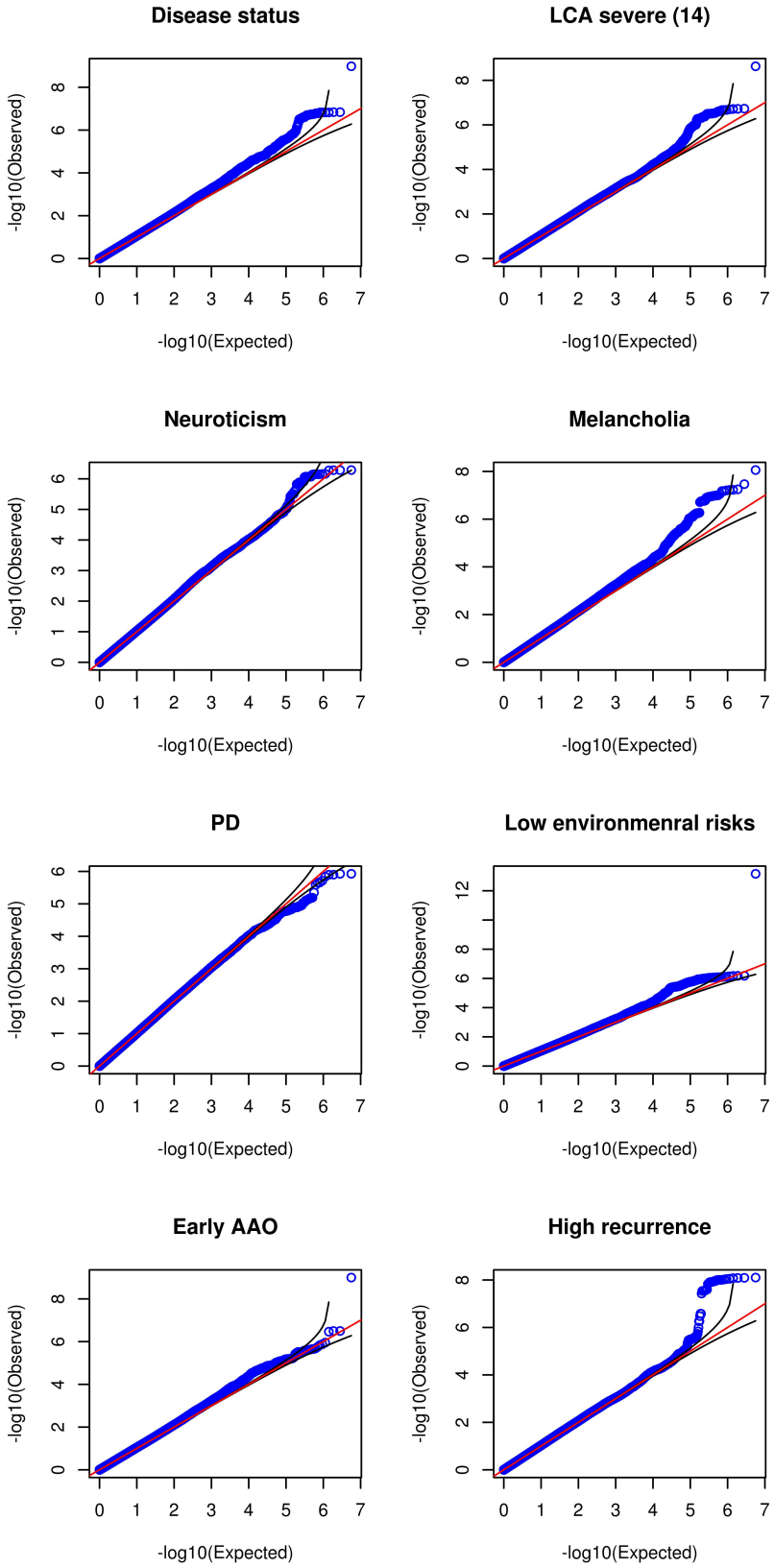


Figure 6.13: QQ plots on the eight depression related phenotypes. LCA, latent class analysis, PD, postnatal depression; AAO, age at onset.

Association analysis on factor scores generated from factor analysis on cases

In addition to carrying out association analysis on cases and controls with different case subsets, I also performed association analysis on continuous measures of factor scores generated from factor analysis using case samples only. In chapter 4, I described the factors generated from each of the three datasets with different symptom profiles. I performed GWA analysis on two factor scores generated from factors primarily loaded on general depressive symptoms and symptoms of worthlessness and suicidal ideation using nine DSM-IV diagnostic criteria. I carried out GWA analysis for three factor scores generated from the 14 items dataset expanded on the bidirectional neurovegetative symptoms. The first factor primarily loaded on symptoms reflecting typical severe depression; the second on appetite and weight changes; and the third on sleep disturbance as well as atypical appetite and weight items. For the 27 items dataset, I also selected three factors for conducting GWA, with the first primarily loaded on symptoms reflecting general regarded depressive symptoms, the second on typical depressive symptoms, and the third on symptoms of worthlessness, hopelessness and helplessness as well as suicidal ideation. Figure 6.14 presents eight Manhattan plots of association analysis regressing SNP dosages against these factor scores. No SNPs exceeded the genome wide significance threshold. Figure 6.15 presents the QQ plots of the P values of the association analysis on those eight factor scores, which further confirmed the non-significant association results.

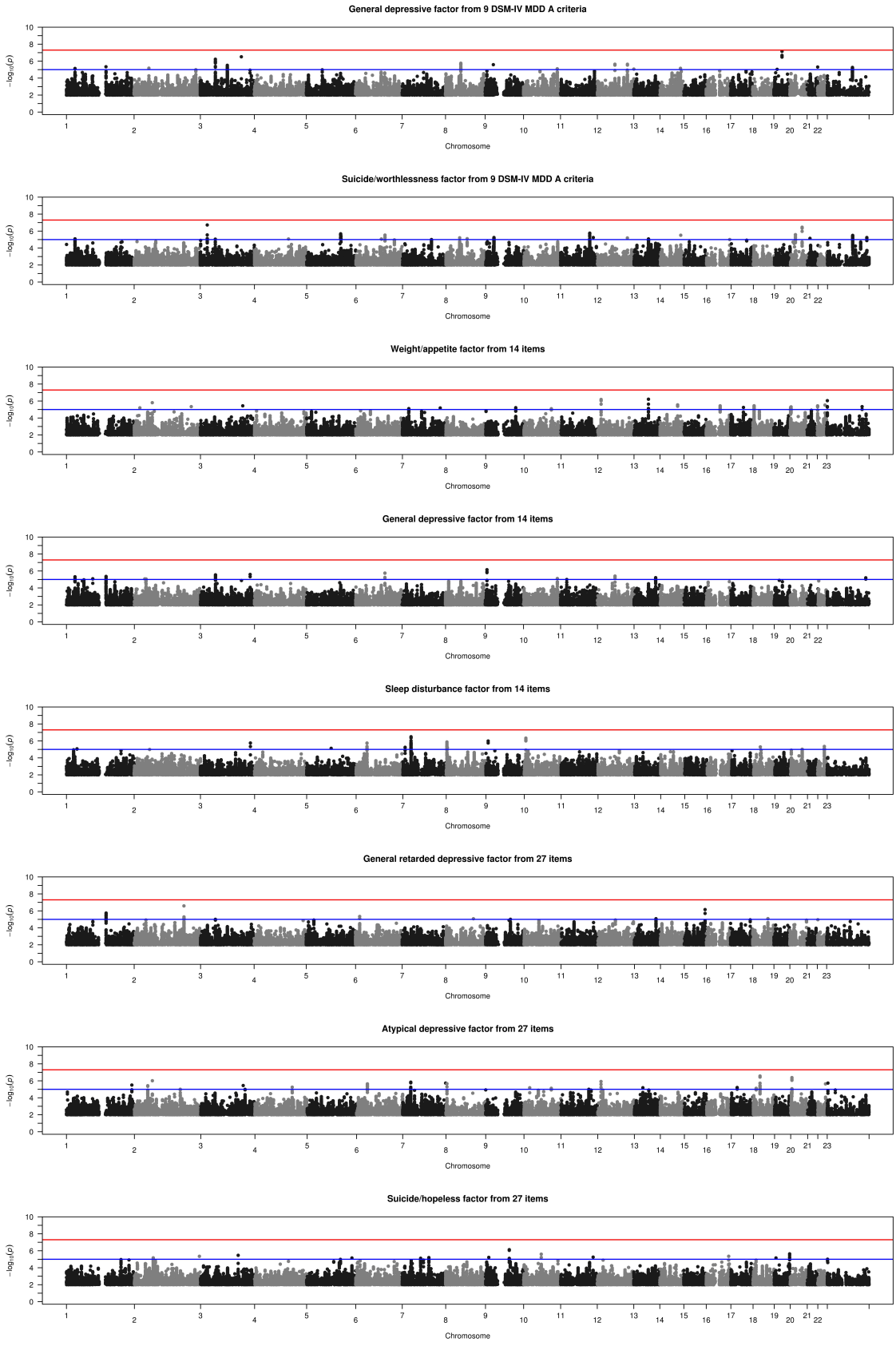


Figure 6.14: Manhattan plots on 8 phenotypic factor scores in cases only controlling for 10 principal components.

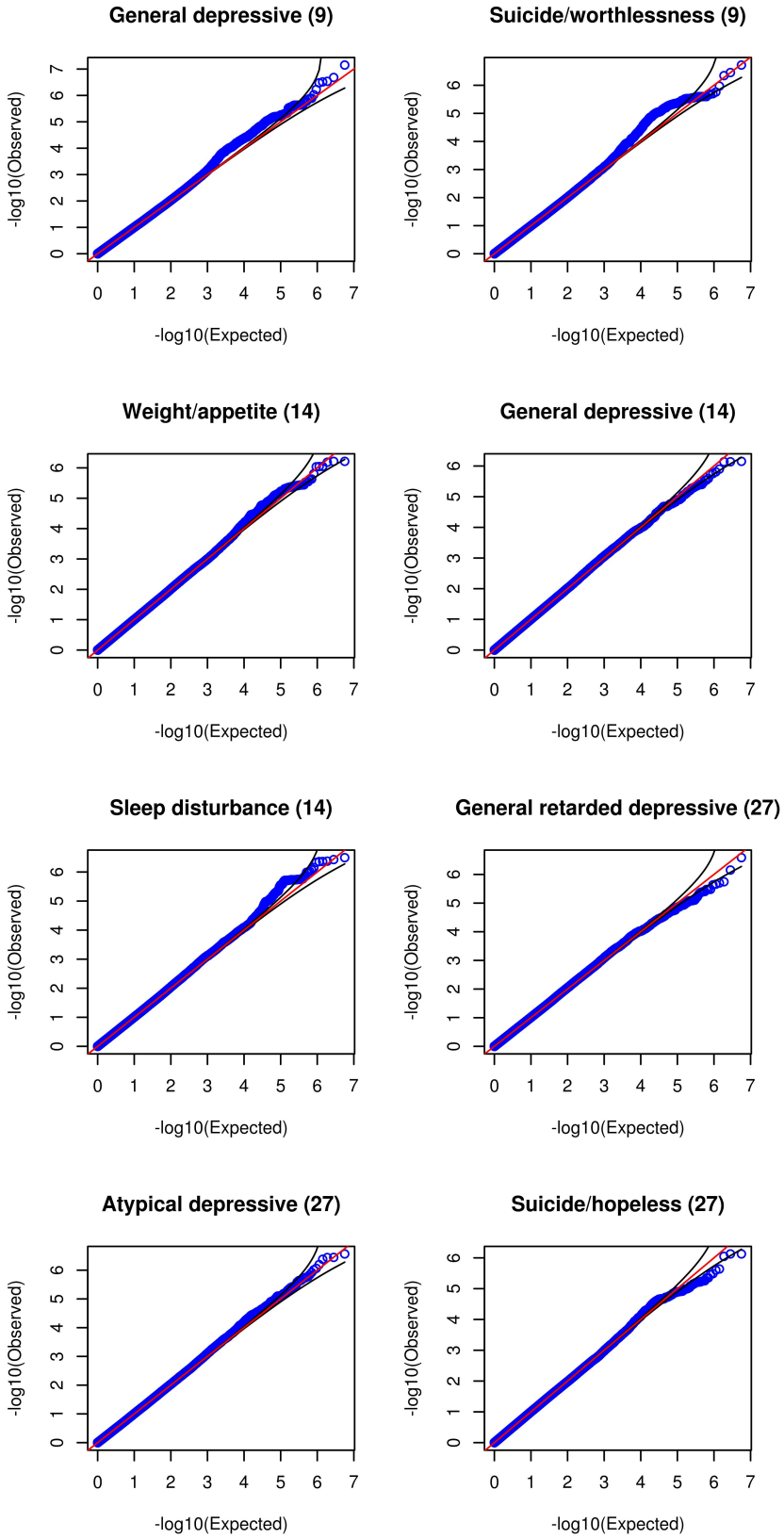


Figure 6.15: QQ plots on 8 case only factor scores. The black lines are 95% confidence interval.

Summary

In this chapter, I have presented the preliminary results for a GWA analysis using genotype dose estimates imputed from low pass (1X coverage) sequencing data. Firstly, GWA analysis was performed on BMI and height. The clusters of SNPs that exceeded the genome wide significance threshold were not previously reported to be associated with BMI and height, suggesting false positive findings. Among the 11,673 samples, 1,957 controls have missing height information due to the fact that the information was not collected for controls at the beginning of the study. However, PCA was performed on all the samples. Therefore, if samples with missing height information have population bias compare to the ones included in the analysis, this might result in a false positive finding due to uncontrolled population structure. GWA on MDD as well as a number of binary phenotypes with different case selection criteria correcting for population structure resulted in one SNP exceeding genome wide significance threshold. However, it does not cluster with other SNPs and is therefore likely to be a false positive. GWA on the binary disease phenotype of cases with more than four episodes gave us a cluster of SNPs on chromosome 10 exceeding a genome wide significance threshold, which overlapped with the top SNPs for melancholia. Similarly to height and BMI, it is important to note that selecting cases with certain phenotypic criteria without recalculating PCs might result in population structure not being properly controlled for. Further work needs to be done by performing PCA on only the selected samples. Finally, GWA on the scores from factor analysis using symptom data did not result in any SNPs exceeding a genome wide significant threshold. Discussion of these points is left to the last chapter of the thesis.

Chapter 7

Mitochondrial genome copy number

Introduction

In chapter 5, I performed the analysis to compare the number of variants between cases and controls in each of the 21,946 *ensembl* gene based on the assumption that there was equal amount of noise and sequencing error in the data. The result from 9,435 samples showed that the number of variants called from the raw reads were significantly higher in all 13-protein coding mitochondrial (MT) genes in cases compare to the controls (Table 7. 1), which raised suspicion. In this chapter, I consider possible explanations for this finding.

Chr	Gene	OR	-log10 p
MT	ENSG00000198888	1.007	13.26
MT	ENSG00000198763	1.007	14.79
MT	ENSG00000198804	1.004	12.50
MT	ENSG00000198712	1.009	14.47
MT	ENSG00000228253	1.038	12.62
MT	ENSG00000198938	1.009	13.90
MT	ENSG00000198899	1.008	13.59
MT	ENSG00000198840	1.020	14.72
MT	ENSG00000212907	1.022	16.09
MT	ENSG00000198886	1.005	16.28
MT	ENSG00000198786	1.004	15.99
MT	ENSG00000198695	1.013	14.62
MT	ENSG00000198727	1.005	15.77

Table 7. 1: Summary statistics of the variant count comparison between cases and controls in all 13-protein coding mitochondrial genes. Chr, chromosome; OR, odds ratio.

Results

Cases have more reads mapping to mitochondria than controls

I first tested whether the extra variants I had seen might be due to the presence of more reads mapping to the mitochondria. I extracted the mean MT read depth on raw reads from whole genome low coverage (1X) sequencing data averaged by whole exome read depth, which was then regressed against disease status controlling for the 10 principal components (PCs) and age as fixed effects and city and sequencing plate as random effects. The mathematical notation of the mixed effect logistic model is shown below:

$$\Pr(case = 1) = \frac{e^{\mu_i}}{1 + e^{\mu_i}}$$

$$\mu_i = \mu + d_i\delta + a_i\alpha + \sum_{m=1}^{10} z_{mi}w_m + \beta_{b_i}$$

where μ_i is linear predictor; m is the grand mean (intercept), d_i is MT divided by exome read depth, a_i is age, z_{mi} (m=1..10) are the 10 PCs; m, d, a are fixed effects to be estimated. b_i is a factor indicating sequencing batch 1…k; β_k is the random effect due to batch k. $\beta_k \sim N(0,\delta_B^2)$. The model was fitted using the R function *glmer()* and models compared using a likelihood ratio test implemented in the R function *anova()*.

The regression analysis showed that cases do indeed have significantly more raw reads mapped to the MT genome than controls (P = 9.95E-29, OR = 1.0116, 95% confidence interval (CI) = 1.0096 - 1.0137). This confirmed that the variant signals in the 13 MT genes were purely driven by higher read depth instead of true genetic variation. Figure 7. 1 and Figure 7. 2 present the kernel density plot and boxplot of MT read depth divided by whole exome read depth in a log scale in cases and controls respectively. Although the P value of the regression analysis is highly significant, the

effect size is very small. The odds ratio (OR), as a measure of effect size, can be interpreted as the following: with an increase of 1X coverage of the MT genome, there is a 1% (95% CI = 0.96% - 1.37%) increase in the odds of developing depression.

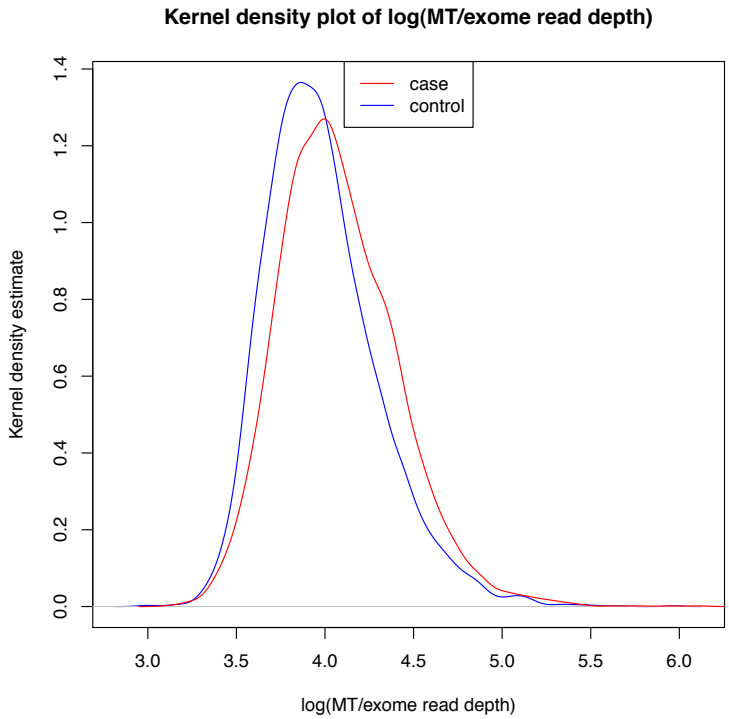


Figure 7. 1: Kernel density plot of mitochondrial read depth divided by whole exome read depth in a log scale in cases and controls.

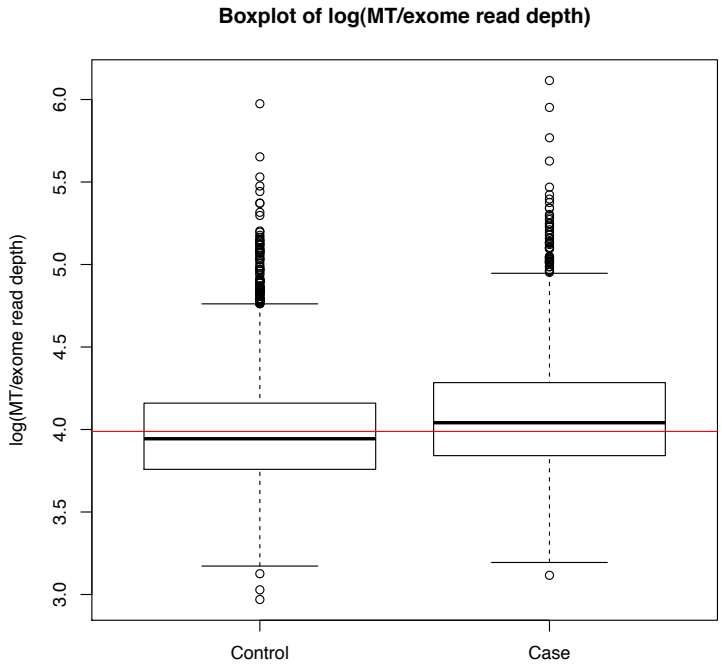


Figure 7. 2: Boxplot of mitochondrial read depth divided by exome read depth in a log scale in cases and controls

Age was controlled for as a covariate in the mixed model. In our study, the ages of cases lay between 30 and 60 years at the time of recruitment, whereas control ages were between 40 and 60. Association analyses show that MT read depth and age were negatively correlated (R^2= -0.05, P = 0.003) in cases and positively correlated (R^2=0.04, P = 0.006) in controls. These results were based on 4,287 cases and 4,776 controls. When sample sizes increased to 5,867 cases and 5,784 controls, the positive effect for controls was still significant (R^2= 0.05, P = 0.0005); however, the negative effect for cases became insignificant (R^2=-0.02, P = 0.08) suggesting that the previously significant finding in cases might be a false positive.

The highly significant association between MT read depth and depression does not explain the origin of the additional reads in cases. Reads might, for example, be

derived from an exogenous source of DNA. In fact, I found that there is a potential contaminating source of DNA. While there is no significant difference between cases and controls on the whole genome sequence coverage (P = 0.40), mean read depth (P = 0.49) and rate of clean reads (P = 0.48), cases do have significantly lower stampy mapping rates (P = 5.2e-17) than controls.

The likely source of the unmapped reads is bacterial DNA. The DNA from the CONVERGE project was extracted from saliva, which, unlike blood, may also contain large quantities of bacteria. I took the raw reads that were mapped to the human MT genome and converted them to *fastq* format. The *fastq* files were then aligned to the whole bacteria assemblies from NCBI using default BWA parameters (Li & Durbin 2009a). Result shows that there is significantly more MT reads mapped to bacteria genome in cases compared to controls (P = 1.23e-06, OR = 1.03).

If the additional MT reads were due to the wrong mapping of reads (for example to bacterial genomes), the MT read depth difference between cases and controls would be weakened after quality filtering on the reads (as it would remove the wrongly mapped reads). Results showed that filtering on read quality did not reduce the significance of the result (P = 8.05E-29, OR = 1.02). This indicates that the signal was most likely not driven by exogenous source of DNA.

We carried out a direct test of the hypothesis that cases have more MT reads than controls by amplifying mitochondrial DNA by quantitative polymerase chain reaction (qPCR). For this experiment, we used MT DNA in 125 pilot samples including 88 cases and 37 controls from China. Results initially showed that cases had higher MT copy number compared to controls. However, our lab assistant repeated the qPCR experiment and the initial borderline significant finding was no longer present. There

were three explanations for the failed experiment, including the inconsistent nature of qPCR results, DNA degradation and small sample size. Further experiments are requested to confirm this finding.

Antidepressant drug response

We then examined whether the MT copy number difference between cases and controls might be a result of antidepressant drug intake. Although patients' drug history is notoriously imprecise, we did include one question asking whether or not patient has ever taken any types of psychotropic drugs, including antipsychotics, mood stabilizers, antidepressants, other psychotropic medicine or never took any psychotropic medicine. Of the 4,287 cases, less than 5% report taking anti-psychotics or mood stabilizers, 83.5% have taken antidepressants, 45.9% have taken other psychotropic medicine, and 12.7% said they had never taken any psychotropic medicine. I treated each drug response as a binary trait, for example categorizing the cases into those who do and those who do not report taking antidepressants, and looked for a relationship with MT read depth. I found no significant association with any of the classifications I used, notably none were seen for those reporting the use of any drug, compared to those who were drug free. I then compared the MT read depth between cases with no drug history and controls. This showed a highly significant effect (OR = 1.01, P = 6.06e-08, n = 5,264). I also compared MT read depth between cases with antidepressants and controls, which again showed a highly significant result (OR = 1.01, P = 7.79e-24, n = 6,444), but notably the effect size was the same. Finally I examined the MT read depth between patients with and without antidepressant drug history in cases only (excluding cases who have taken other psychotropic drugs). The result is insignificant (P = 0.37).

Altogether these analyses indicate that the difference I observed in the MT read depth between cases and controls is not due to antidepressant drug intake.

We then carried out a direct test of the hypothesis that antidepressant increases MT copy number by designing a mouse experiment with 6 cases and 6 controls between 5-6 weeks old. Cases were given fluoxetine (the most commonly used antidepressant in the CONVERGE study) dissolved in their drinking water. We then took saliva and blood samples on a weekly basis over a period of 50 days. Real time qPCR was conducted on saliva and blood samples at 4 and 6 time points respectively. Figure 7. 3 shows the MT copy number as well as fitted residuals after controlling for weight. Results indicate that antidepressant does not increase MT copy number over a period of 50 days in mice.

If the MT copy number increase in the cases was not a result of drugs, it is unclear at this stage what might be the possible explanation. Our result was obtained from sequencing DNA extracted from saliva. Given that depression is thought to be a disease of the brain, a finding from saliva DNA was unexpected. Perhaps therefore our result was not related to the cause of depression, as it surely must be in the brain. Our next hypothesis was to test whether the MT copy number increase in cases might have a purely environmental origin. If so, variation in MT number should not be heritable.

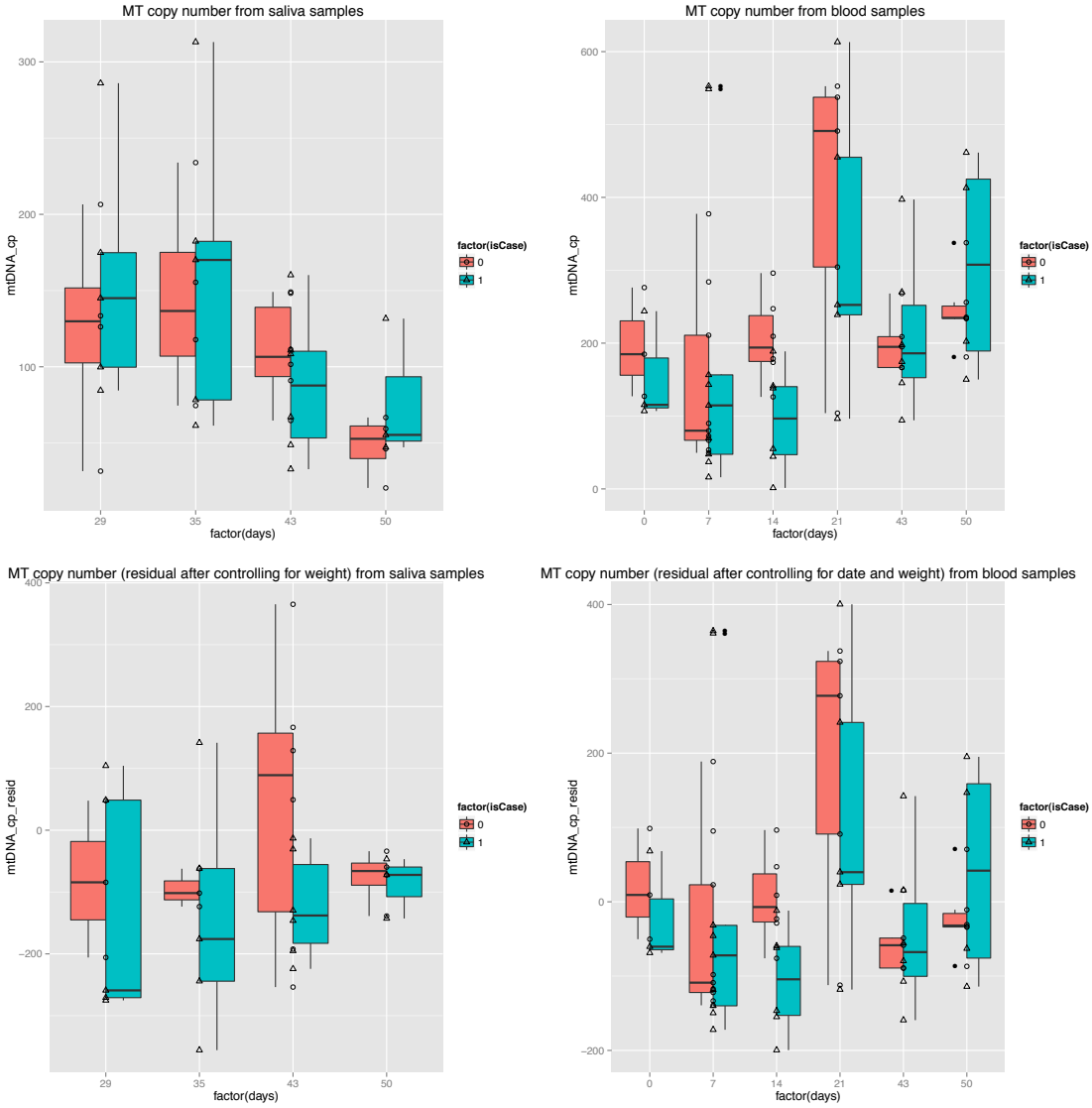


Figure 7. 3: Boxplots and scatter plots of mitochondrial copy number quantified as saliva and blood qPCR expression data for the mice antidepressant drug experiment.

Genetic basis for MT read depth

To determine whether genetic variation contributed to MT copy number I performed a whole genome scan on the MT read depth using 11,651 sequencing samples. The results were initially presented with 9,300 samples for initial thesis submission, which were updated with the additional samples for the final submission. Figure 7. 4 and 7.5 show the Manhattan plots and QQ plots on MT read depth for all samples as well as cases and

controls separately. The top hits with –log10 P greater than 7 for the genome scans are listed in appendix I. The multiple testing burden for GWAS was taken into account when determining the genome wide significance threshold (Pe'er et al. 2008). For case control combined samples, there are two clusters of above genome wide significance SNPs, one of which lies in the *TFAM, CISD1, UBE2D1, BICC1* and *IPMK* genes on chromosome 10, the other lies in the *RNU6-10P, FAM133B* and *CDK6* genes on chromosome 7 respectively. The most significant SNPs ($-\log10_{rs1937} = 17.25$; $-\log10_{rs2282989} = 8.73$) for each of these gene clusters were shown in the zoomed-in plots in figure 7.6 and 7.7. The *TFAM* gene, which showed above genome wide significance is known to be responsible for MT replication and repair (Shi et al. 2012). The mapping results established a genetic basis for the MT copy number. There are three sporadic single SNPs on chromosome 4, 8 and 18, that have significant P values. The SNP on chromosome 4, which has 10% MAF, is also significant on control only analysis. It is unclear what the reason was for these sporadic single significant SNPs.

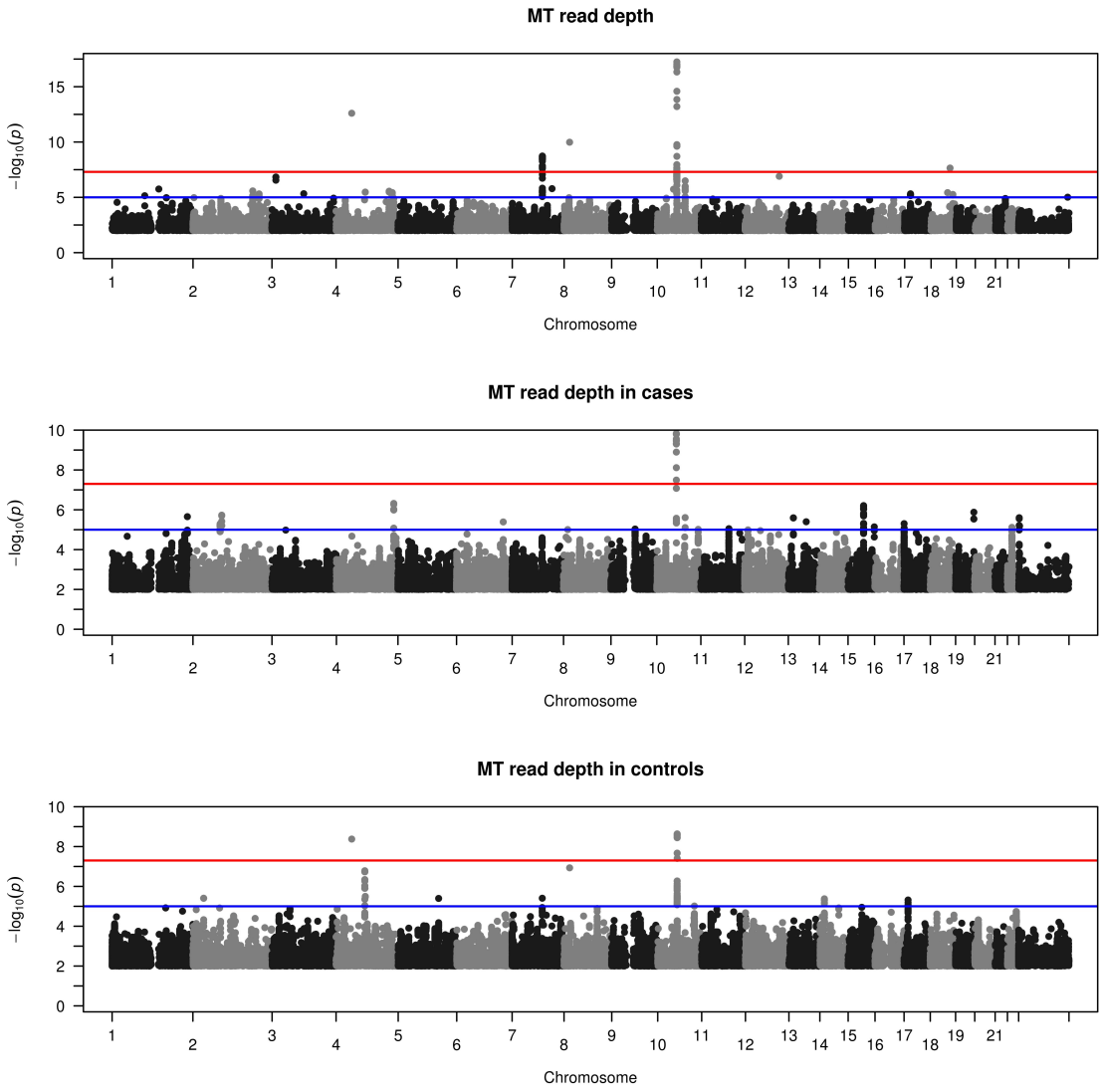


Figure 7. 4: Manhattan plots for genome wide association on MT read depth in combined sample as well as in cases and controls separately.

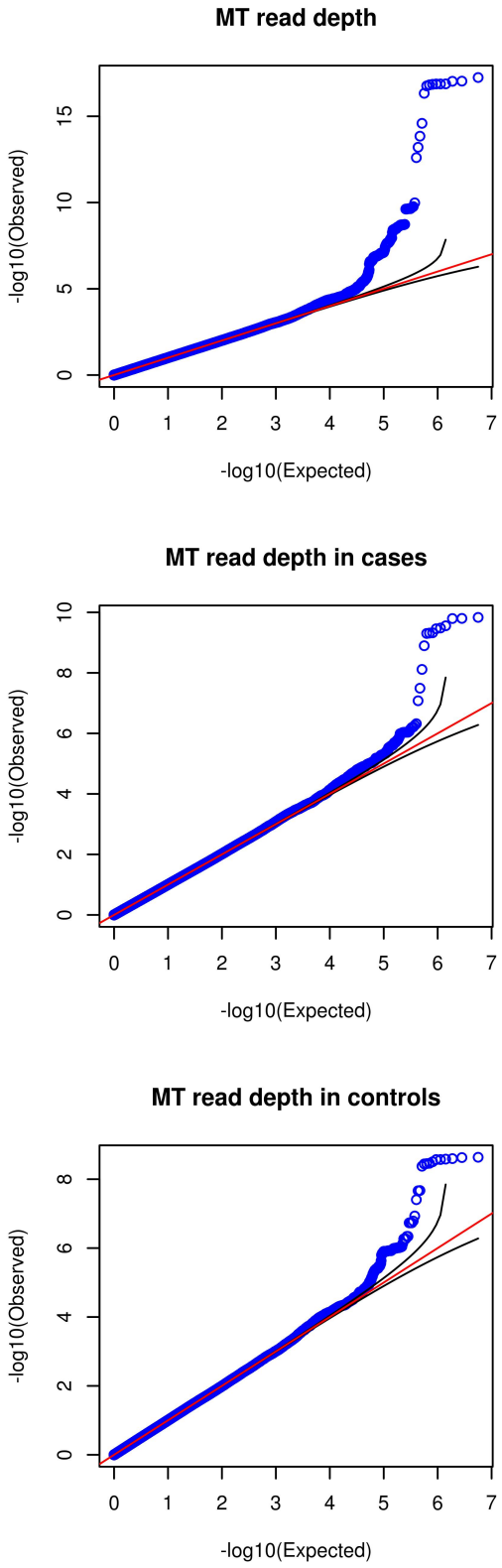


Figure 7. 5: QQ plots for P values from genome wide association analysis on MT read depth in combined sample as well as in cases and controls separately.

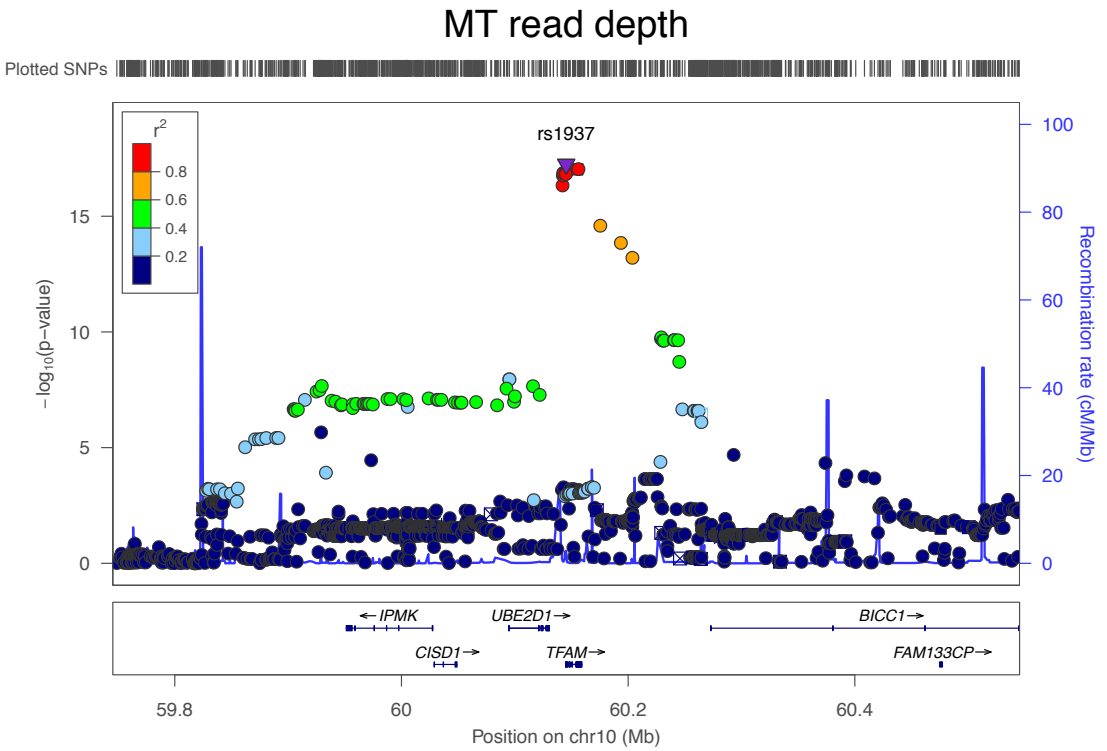


Figure 7. 6: Zoomed in on the top hit for genome scan on MT read depth using combined sample. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

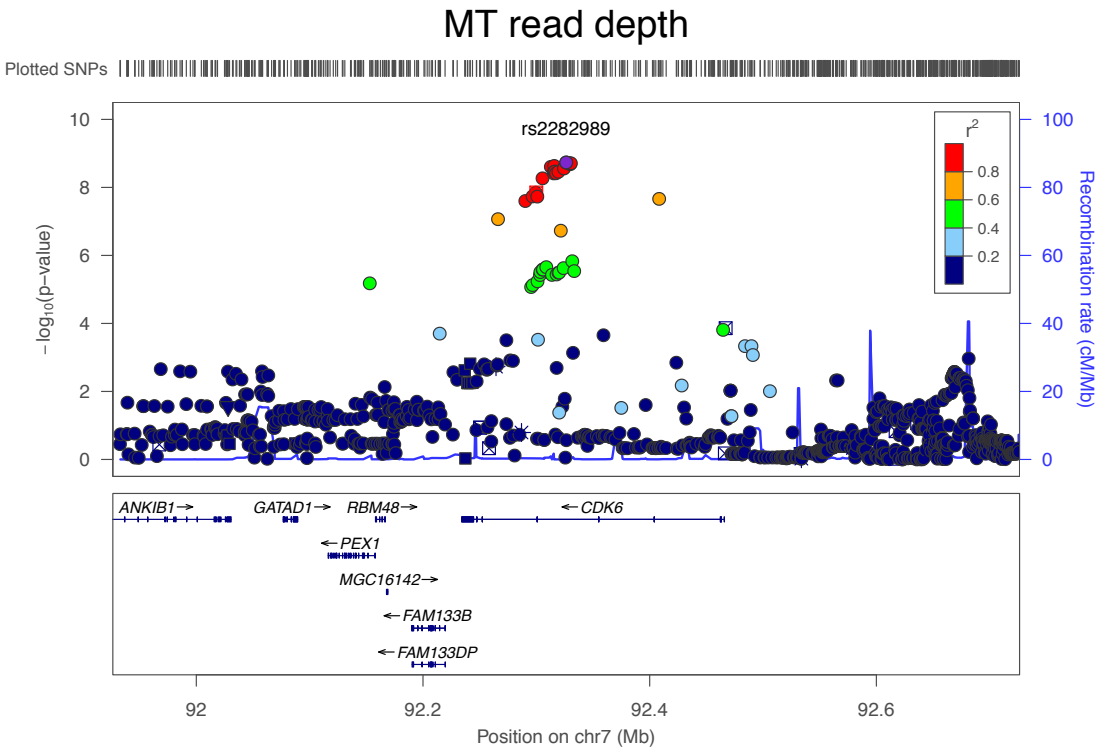


Figure 7. 7: Zoomed in plot for the top hit of the second cluster of above genome wide significance SNPs for combined sample. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

I also mapped MT copy number in cases and controls separately. Potentially, variation in MT number might only have a genetic basis in cases (the larger variance in the measure for example might be due to the additional contribution of genetic variation to the measure). I performed F test to compare the variance of the MT read depth between cases and controls. The result is highly significant (F = 1.296, P < 2.2e-16).

For case and control separately, the signals from the combined samples were significantly weakened. However, the above genome wide significance hits from the cluster of SNPs in the *TFAM* gene on chromosome 10 overlapped with that of the cases (P_{rs1937} = 9.83) (figure 7.8) and controls ($P_{rs11006126}$ = 8.63) (figure 7.9).

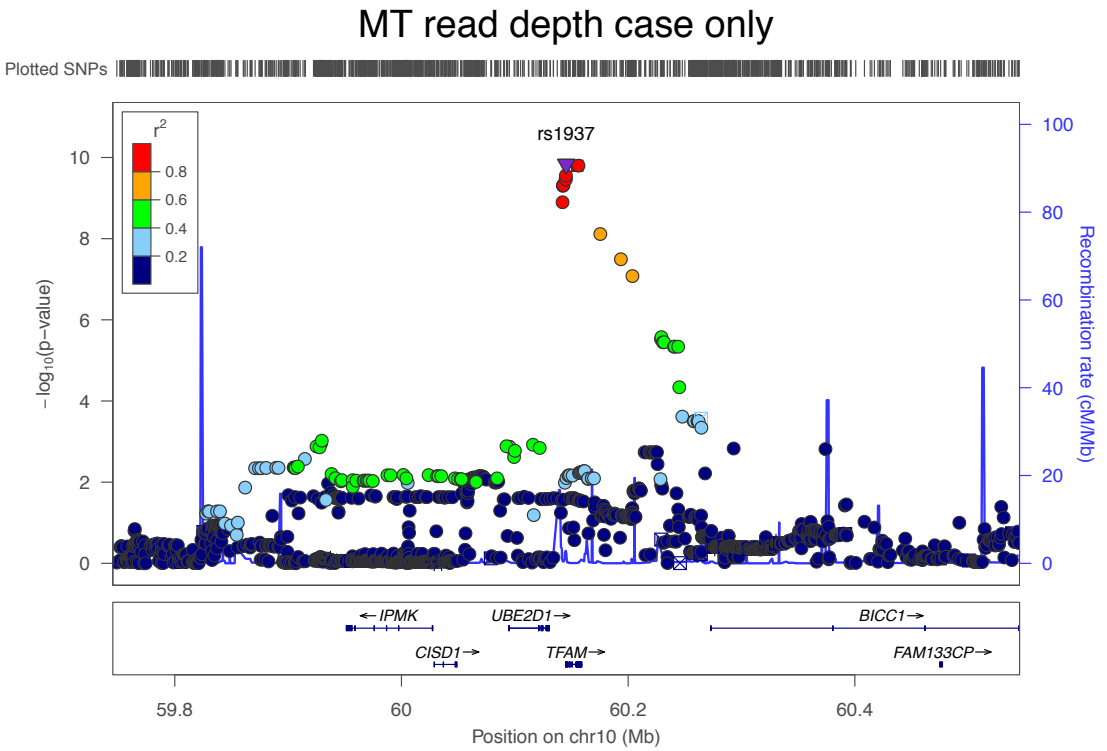


Figure 7. 8: Zoomed in plot for the top hit from genome scan on mitochondrial read depth in cases only. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

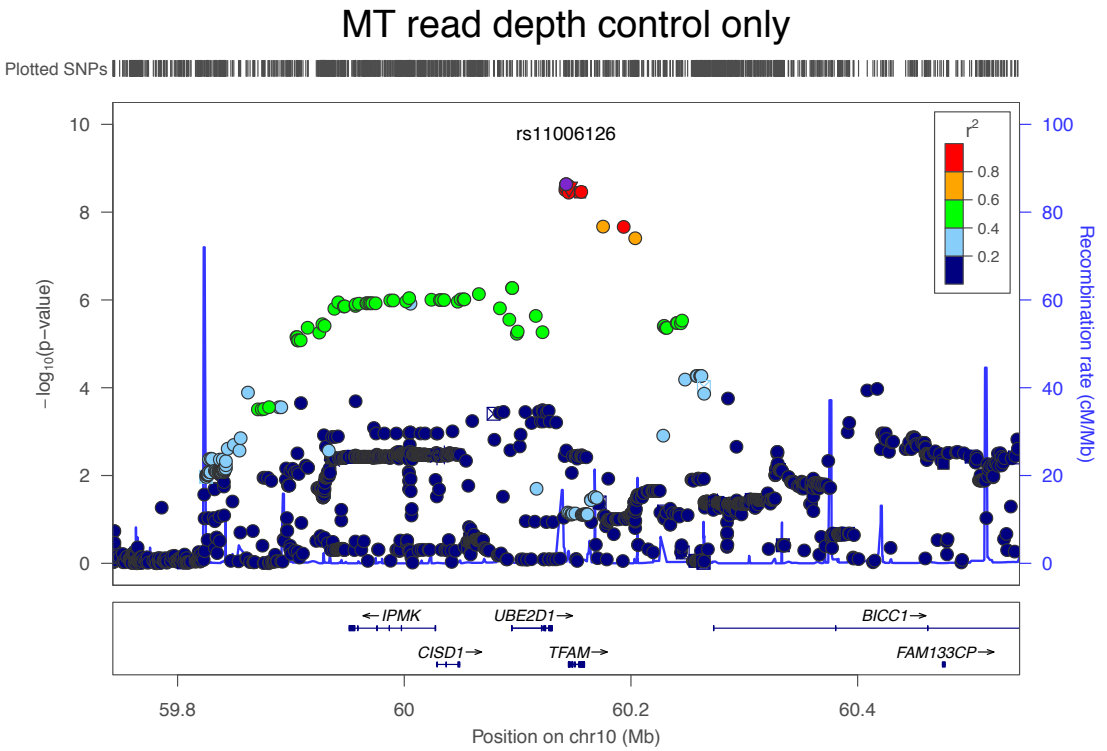


Figure 7. 9: Zoomed in plot for the top hit from genome scan on mitochondrial read depth in controls only. Functional annotation of all 1000 Genomes and HapMap SNPs includes the following categories with a description of its graphical representation in brackets: framestop (triangle) splice, (triangle), non-synonymous (inverted triangle), synonymous (square), UTR (square), TFBScons (star), MCS44 Placental (square with diagonal lines) and none of the above (filled circle).

I have also compared the effects of the top SNPs in *TFAM* (27 SNPs) and *CDK6* (23 SNPs) genes between cases and controls. Figure 7.10 contains Forest plots of the effects of the top SNPs in those two genes in cases and controls separately for comparison. The mean effects (presented as the dashed vertical line) of the top SNPs on the *TFAM* and *CDK6* genes are larger in cases ($Coef_{TFAM} = 2.42$, $Coef_{CDK6} = 1.52$) than in controls ($Coef_{TFAM} = 2.31$, $Coef_{CDK6} = 1.47$).

Mapping results in cases and controls and the failure to see an increased signal in the combined sample at some loci indicated that the genetic basis of MT variation differed between cases and controls. This suggests that there is a gene by environment interaction involved, a hypothesis that will need additional analyses to confirm.

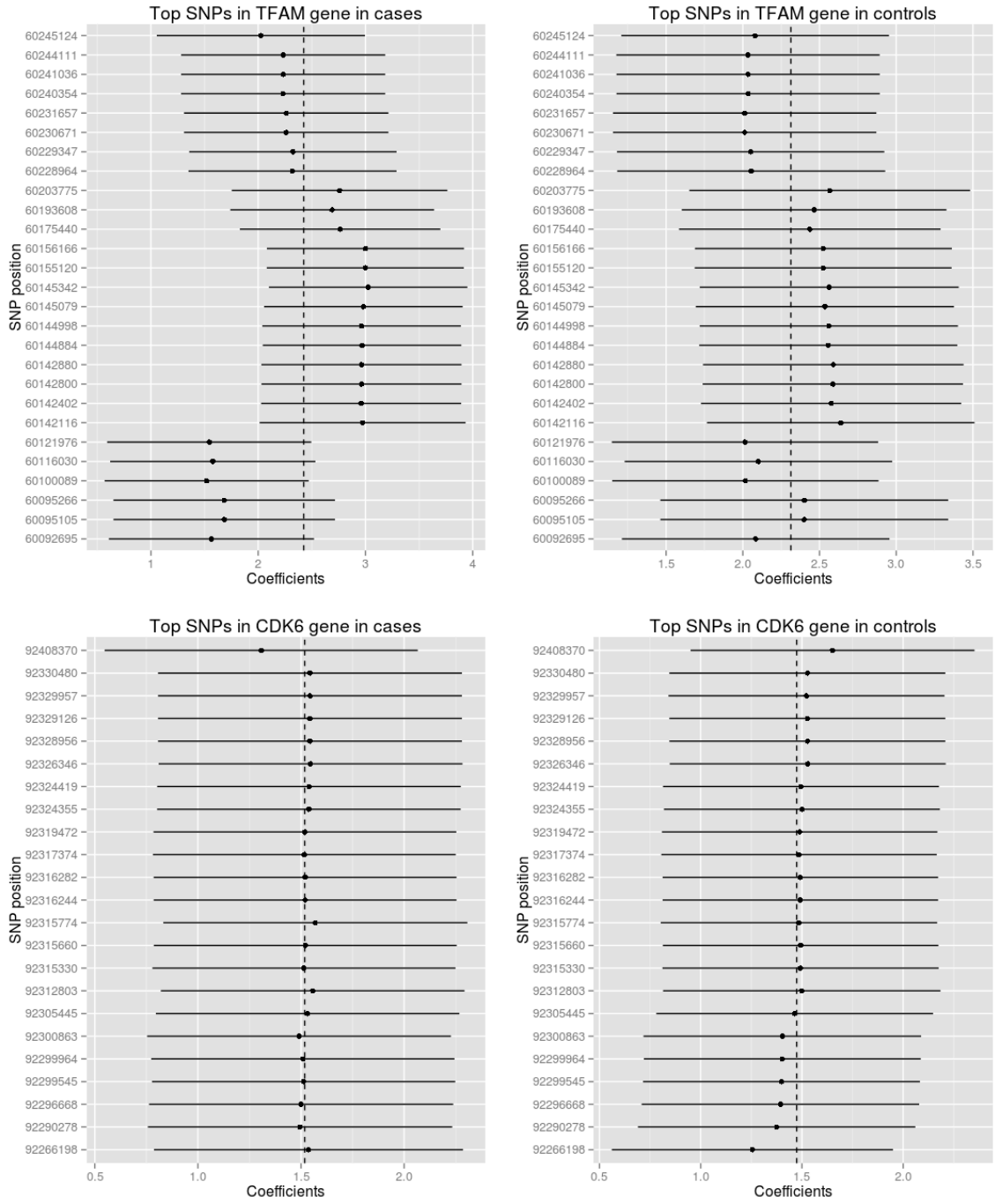


Figure 7.10: Forest plot of top SNPs in the *TFAM* (chr10) and *CDK6* (chr7) genes respectively. The mean of the regression coefficient is presented as the dashed vertical line.

Common genetic basis between mitochondria and MDD

I next asked if the genetic effects that contribute to variation in mitochondria might also contribute to the variation in MDD. In other words I set out to establish if there is a genetic correlation between the two phenotypes. I performed a polygenic risk analysis (Dudbridge 2013) by testing whether the SNPs that predict MT read depth jointly predict risk for MDD in 9,300 samples.

The polygenic risk scores were calculated by taking the sum of the scores for the selected SNPs with different P value thresholds. The score for each SNP for each individual is calculated as the product of the effect size (BETA coefficient) and genotype dose estimate. The Nagelkerke R squared was used to evaluate the goodness of fit of the logistic regression model, in other words the variance in disease status explained by the MT polygenic scores. Table 7. 2 showed the predictive power of the polygenic score calculated from MT scan against disease status in the total 9,300 samples with different P value thresholds. When taking 82,381 SNPs with –log10 P threshold greater than two, the polygenic score calculated from BETA generated from MT scan was highly predictive of disease status (P = 1.26E-21), which suggested that those selected SNPs for predicting MT read depth were jointly predictive of disease status.

MT scan - log10 P value threshold	Number of SNPs	MDD polygenic P value	NagelkerkeR2 fit for PCs only	NagelkerkeR2 fit for PCs and polygenic score
7	40	0.477	0.005	0.005
6	64	0.805	0.005	0.005
5	147	0.608	0.005	0.005
4	1,356	0.015	0.005	0.006
3	10,002	8.6E-09	0.005	0.010
2	82,381	1.26E-21	0.005	0.019

Table 7. 2: Polygenic risk analysis. Polygenic score was calculated for each individual by combining a product of genotype dose estimate and effect size (BETA) from SNP association analysis on MT read depth. The polygenic risk score was then regressed against disease status.

I then calculated the predictive power of the polygenic score from the BETA values generated by MT scan. I divided the sample into two halves, and used one half to create a polygenic score and the other to predict disease status from the score. If the polygenic score in one half significantly predicts disease status in the other half, there is evidence for a shared genetic basis (Table 7. 3). The predictive power of the MT SNPs was not significant in the other half with -log10 P threshold reduced to 0.3 (P = 0.5) including 3.6M SNPs (P = 0.38). This result shows that there is no evidence of a common genetic basis for MT read depth and MDD. Furthermore, it also suggests that the MT copy number variation could not be a pre-determining factor for depression. Instead it is more likely to be a consequence of depression.

		Sample 2 (4,535)			Sample 1 (4,538)		
MT scan - log10 P value threshold for randomly selected half samples	**Number of SNPs**	**MDD polygenic P value on the second half samples**	**NagelkerkeR2 fit for PCs only**	**NagelkerkeR2 fit for PCs and polygenic score**	**MDD polygenic P value on the same half samples**	**NagelkerkeR2 fit for PCs only**	**NagelkerkeR2 fit for PCs and polygenic score**
7	23	0.709	0.007	0.007	0.456	0.006	0.007
6	269	0.361	0.007	0.007	0.551	0.006	0.007
5	639	0.304	0.007	0.007	0.506	0.006	0.007
4	1,870	0.469	0.007	0.007	0.204	0.006	0.007
3	9,518	0.778	0.007	0.007	0.002	0.006	0.009
2	77,795	0.558	0.007	0.007	2.73E-15	0.006	0.026
1.3	373,934	0.180	0.007	0.007	2.41E-18	0.006	0.030
0.3	3,664,976	0.380	0.007	0.007	3.64E-07	0.006	0.014

Table 7. 3: Polygenic risk analysis on a random split half samples

Phenotypic association on MT read depth

In order to test the hypothesis that MT copy number increase in cases might be triggered by environmental factors, I performed linear mixed model on all the available phenotypes against MT read depth controlling for the 10 PCs and age as fixed effect, as

well as sequencing batch and sample origin as random effects. The P value was calculated by anova Chi-squared test on two nested models one with only the covariates, the other with the phenotype of interest as an additional fixed effect. With approximately 1,000 raw and derived phenotypes, a significant threshold of $P = 5\times10^{-5}$ was applied. Table 7.4 presents the most significant results for the mixed model analysis. The phenotypic variables including the number of DSM-IV MDD A criteria based on the information collected from the medical records (which most resembles disease status) and neuroticism are most significantly associated with MT copy number, with a –log10 P of 31 and 23 respectively. The next most significant results included premenstrual syndrome, childhood sexual abuse with different levels of severity, the number of stressful life events, marital status, an item from the parental bonding instrument (talk things over with father) and occupation. The phenotypic association results suggested that the highly significant result between MT read depth and disease status is a result of chronic stress.

Description	Sample size	Estimate	SE	T value	Chisq P value	$-\log_{10}$ (P)
Number of DSM-IV MDD A criteria based on MR	11028	0.62	0.05	11.72	1.45E-31	30.84
Neuroticism	10531	0.27	0.03	10.03	1.46E-23	22.84
PMS	10563	-0.73	0.12	-5.90	3.59E-09	8.44
CSA severity (0-3)	10842	2.05	0.35	5.79	7.44E-09	8.13
Intercourse (exc.)	10303	7.95	1.39	5.73	1.04E-08	7.98
Intercourse (inc.)	10842	7.64	1.40	5.45	5.23E-08	7.28
Non genital (inc.)	10842	4.00	0.78	5.14	2.77E-07	6.56
Any CSA (binary)	10842	3.68	0.72	5.10	3.48E-07	6.46
SLE	10274	0.55	0.12	4.60	4.50E-06	5.35
Marital status (0-5)	11026	0.89	0.20	4.42	9.67E-06	5.01
Talking thing over (father)	9164	0.66	0.15	4.41	1.04E-05	4.98
Occupation	9216	0.36	0.08	4.27	1.89E-05	4.72
Marital status (binary)	10834	0.58	0.14	4.11	3.96E-05	4.40
Genital CSA	10842	3.71	0.93	3.99	6.63E-05	4.18

Table 7. 4: Regressing phenotypic variables against MT read depth. SE, standard error; DSM, diagnostic and statistical manual for mental disorders; MR, medical record; PMS, premenstrual syndrome; CSA, childhood sexual abuse; inc., inclusive of CSA with higher severity; exc., exclusive of CSA with higher severity; SLE, number of stressful life events;

We further tested the hypothesis that stress might lead to an increase in the number of MT genomes by inducing chronic stress in mice, and measuring their weekly MT copy number using qPCR. The experiment was conducted using eight female cases and controls over four weeks period. Stress was induced by repeating a sequence of the following approaches: tail suspension on the first day (10mins with 5 mins rest x 3), forced swimming on the second day (10mins with 10mins rest x 2), foot shock (0.75 mA 10 sec with 10 sec rest x 3) on the third day, and restraint for 3hrs and sleep deprivation for 24hrs on the fourth day.

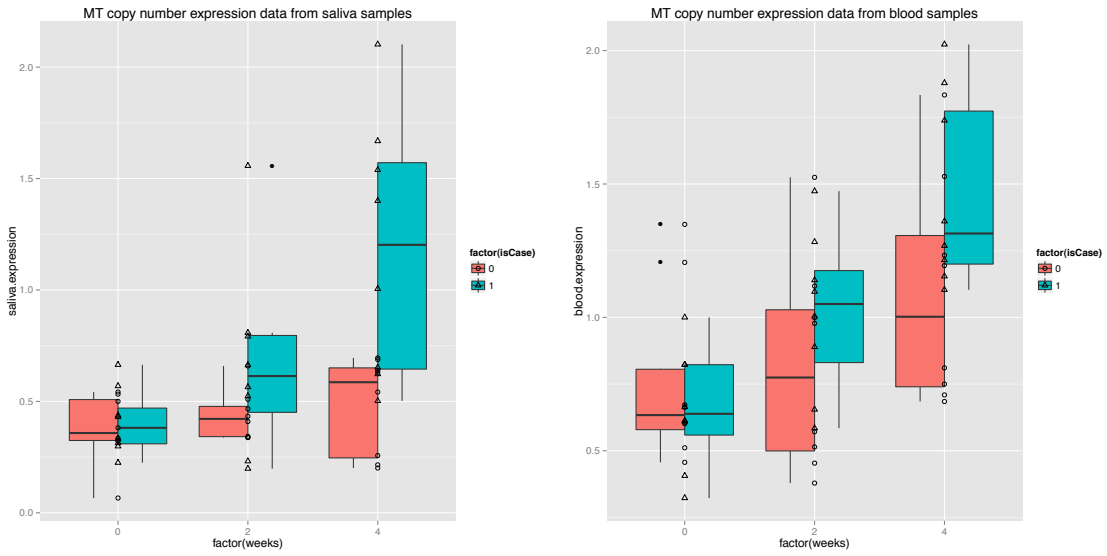


Figure 7. 11: Mice chronic stress experiment.

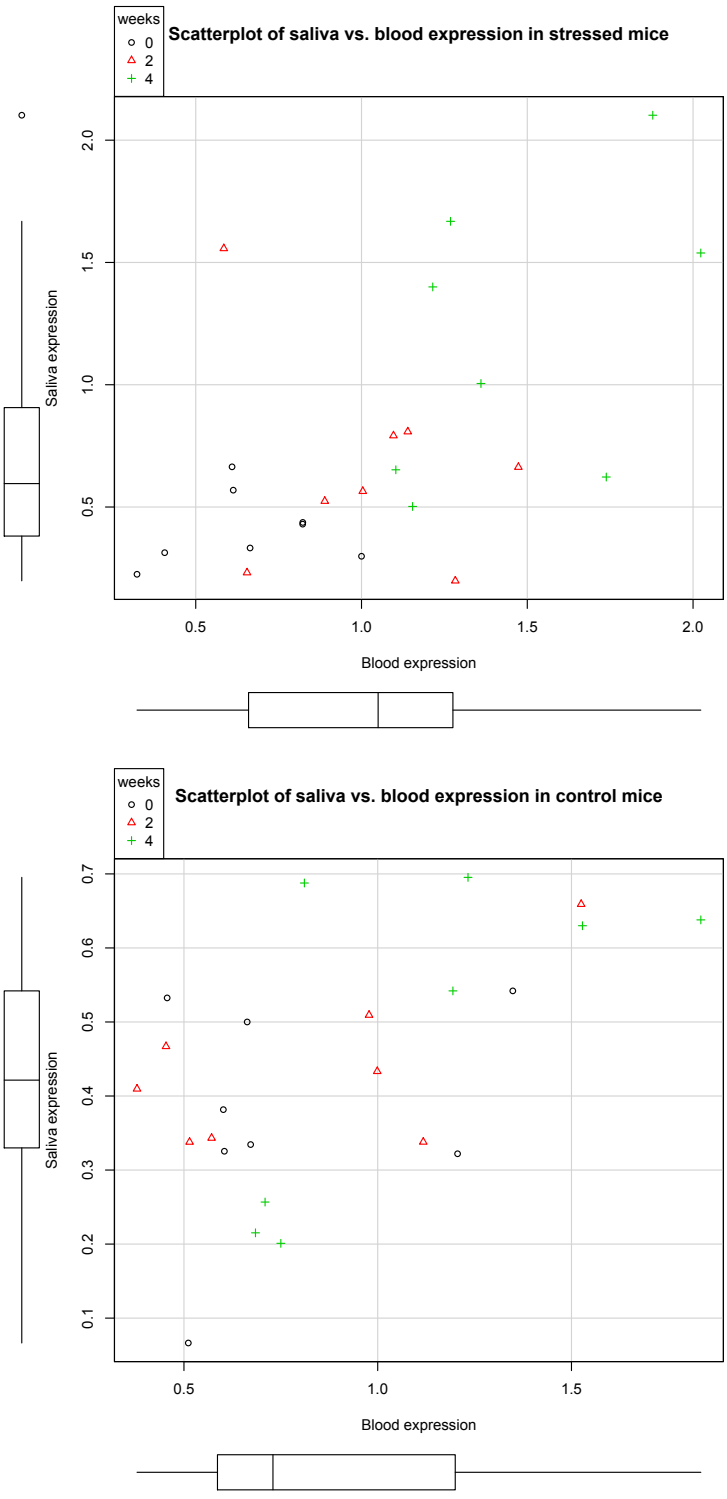


Figure 7.12: Scatter plots of saliva versus blood qPCR data in stressed (top) and control (bottom) mice coloured by the three time points collected over a four-week period.

Figure 7. 11 clearly shows that the MT copy number quantified using qPCR increases more in cases compare to controls over time. Figure 7.12 presents the scatter plots of saliva against blood qPCR data in stressed and control mice coloured by the three time points over four weeks period. We analysed the data by ANOVA on mixed models on case control status. The first one treated week as a fixed effect and individual mouse as random effect. The second added an interactive term of MT copy number and week. Result showed that MT copy number is highly significant (P = 7.988e-05) in predicting chronic stress when controlling for effect of week and individual mouse. The mice data provides good evidence that the MT copy number difference between cases and controls is a result of chronic stress.

Summary

In this chapter, I have examined the possible reasons for the significantly higher MT copy number in cases compared to controls. The initial finding came from the observation that cases have significantly more variants in all 13 protein coding MT genes in cases compare to controls using the gene based approach described in chapter 5. This finding was due to the presence of significantly more reads mapping to the MT genome in cases. The results of genome scans demonstrate that there is a genetic basis for the MT read depth. I identified a genome-wide significant hit in the *TFAM* gene, a gene known to be involved in MT replication and repair (Shi et al. 2012). Polygenic risk analysis however rejects the hypothesis that there is common genetic basis between MT copy number and MDD. By regressing the polygenic score calculated in a randomly selected half against disease status in the other half I was unable to find a significant prediction of disease status. Human phenotypic analysis indicates that the MT copy

number difference between cases and controls might be a result of chronic stress. The hypothesis was validated by inducing chronic stress in mice and comparing their MT copy number changes over time. Results indicate that mice who have undergone chronic stress show a greater increase in the MT copy number than controls. Both human and mice data indicate that chronic stress might be the cause for the increase in MT copy number in cases. Discussion of these points is left to the last chapter of the thesis.

Chapter 8

Discussion

My thesis presents findings from the CONVERGE study, the largest case-control collection of subjects with recurrent MDD yet conducted. I set out to use the data from CONVERGE to address questions about the phenotypic structure of MDD and its genetic substrate. To do so, I took advantage of a novel strategy for genome analysis, namely the collection of low coverage (1X) whole genome sequence data. The size of the data set, the complexity of the phenotypes, and the unknown nature of the biology of depression present formidable challenges.

In this discussion chapter I review the principal findings from each chapter and discuss their implications. I begin with summaries of the phenotypic analyses (my consideration of the importance and structure of comorbid anxiety disorders, and the attempts to recognize subtypes of MDD from factor and latent class analyses), then progress to the genetic analyses. Finally I turn to the most intriguing result, the finding that patients with MDD have more mitochondrial (MT) DNA than controls.

For the genetic analysis, I developed a novel bioinformatics approach to examine the rare variant disease hypothesis by analysing the exome of whole genome low coverage sequencing data on a large sample based on the assumption that there is equal amount of sequencing errors in cases and controls. Overall, I was unable to find convincing evidence that functional variants, common or rare, contributed to disease susceptibility.

I then took a more conventional approach to test the "common disease common SNP" hypothesis by using imputed genotype dose estimates. In addition to presenting the preliminary finding of the first pass genome wide association analysis on MDD, I also did SNP association analysis on subtypes of MDD to see whether increasing disease homogeneity at the cost of reduced sample size would increase the power to detect susceptible SNPs for specific subtypes. Furthermore, I carried out SNP association analyses on factor scores generated from case sample to examine whether there is a genetic basis for phenotypically determined subtypes of depression.

Finally, I presented explanation for a novel observation in the MT genome using the rare variant analytical approach. Further experiments were carried out to test several hypotheses, including bacteria contamination and antidepressant drug effect. Finally, I was able to present evidence supporting that the MT copy number increase in cases was a result of chronic stress.

Comorbid anxiety in MDD

The first phenotypic analysis I carried out was to examine the predictive patterns of comorbid anxiety disorders on MDD related clinical features. This analysis shed light on the question raised in the introduction about whether depression and anxiety are the same disorder or not. Despite the inconsistent reporting in the single association analyses, the prevalence rate and the predictive patterns agree with what have been reported in the literature. The seven comorbid anxiety disorders are significantly heterogeneous when predicting the risk for neuroticism, the number of MDD episodes and criteria.

The results presented in Chapter 3 made three points. First, the prevalence of anxiety disorders among patients with MDD is high, and is commensurate with that reported elsewhere in the literature. Specifically, the 50.25% prevalence rate for any of the seven comorbid anxiety disorders in our sample is within the 95% confidence interval of published figures using Western samples of comparable recruitment and diagnostic criteria. It is lower than the ~70% reported in the two Chinese studies (Shi et al. 2009; Wu et al. 2013), possibly due to the fact that our cases were all recurrently depressed which might have different patterns of comorbid anxiety from samples including single episode cases.

Second, the association between clinical features of MDD and comorbid anxiety disorders is similar to that reported in the West. Despite the sampling differences, single and multivariate regression analysis using the entire case sample confirmed the published literature that comorbid anxiety disorders predict earlier depression onset (Fava et al. 2000; Fava et al. 2004; Howland et al. 2009; Lamers et al. 2011; Penninx et al. 2011), family history (Clayton et al. 1991; Gorwood 2004; Weissman et al. 1986) and neuroticism (Lamers et al. 2011). There are conflicting findings on whether comorbid anxiety disorder predicts increased number of depression episodes, where two studies reported negative results (Fava et al. 2000; Fava et al. 2004) and one positive (Howland et al. 2009). Our result is in line with the significant finding although it is important to note that we exclude single episode cases, which presumably should reduce the degree of variation in the data.

However there are some features that, in the literature, are not consistently reported to be associated. Thus there are conflicting reports as to whether comorbid anxiety disorder increases suicidal risk among depression patients (Bolton et al. 2008;

Fava et al. 2004; Placidi et al. 2000). For comorbid panic disorder specifically, one study reported positive increased risk for suicidal attempt in women (Bolton et al. 2008) and four other studies gave negative findings (Claassen et al. 2007; King et al. 1995; Perroud et al. 2010; Placidi et al. 2000). Our result is in line with the finding that comorbid panic disorder in patients with MDD does not increase risk for suicidal attempt. Furthermore, two depression studies GENDEP (Perroud et al. 2010) and NESARC (Bolton et al. 2008) have reported insignificant findings when regressing comorbid GAD, agoraphobia and social phobia independently against suicidal attempt. Another study (STAR*D) (Claassen et al. 2007) reported positive finding for GAD and social phobia unadjusted for covariates, which is most comparable to our results. When all the comorbid anxiety disorders were assessed jointly in one multivariate regression model, GAD and social phobia are significant in predicting risk for suicidal attempt.

The third conclusion from Chapter 3 is that there is considerable degree of heterogeneity among the seven anxiety disorders in predicting risk for certain depression related clinical features. Comorbid anxiety disorders are significantly heterogeneous when predicting risk for neuroticism and the number of DSM-IV MDD A criteria. In other words, there are significant individual effects for each comorbid anxiety disorder in predicting risk for neuroticism and the number of DSM-IV criteria. A point worth noting is the nature of those two variables. The number of DSM-IV MDD A criteria was calculated by counting the number of positive responses for the nine criteria, each of which were independent questions regarding symptoms. Similarly, neuroticism was calculated by counting the number of positive responses for the 23 independent questions in Eysenck Personal Questionnaire (EPQ). Hence the same number could represent a different combination of positive responses to those questions.

Heterogeneity could also be a result of multiple comorbid anxieties operating in opposite effects, which was supported by the significant results in both single and multivariate regression for GAD (OR = 1.38) and blood-injury phobia (OR = 0.66) in the published finding (Li et al. 2011). Although insignificant, the multivariate regression analysis presented here resulted in panic and blood injury phobia both having protective effects. It is unclear whether the previous significant finding was false positive which disappeared when sample size increased, or stringent editing resulted in less qualified diagnoses for anxiety is having a reduction in its predictive power for suicidal attempt.

Family history of depression resulted in all seven comorbid anxiety disorders having significant predictive effects when assessed independently and four anxiety disorders showing significant effects when assessed jointly. Heterogeneity test resulted in the least significant result (largest P value) across all clinical features of interest, which suggest that many if not all seven comorbid anxiety disorders are significant in predicting risk for depression family history and their effects are homogeneous. This leads to discussion about the shared genetic propensity to depression and anxiety, which partly explains the comorbid phenomenon (Cerdá et al. 2010; Kendler et al. 2003). A study, modelled using 5,600 population twins on seven psychiatric syndromes, resulted in one internalising dimension with relatively large additive genetic effects (MDD 0.54, GAD 0.53 and phobia 0.33) and modest unique or non-shared environmental effects (0.16-0.47), which suggested that genetic factors were largely responsible for the pattern of comorbidity among internalising disorders (Kendler et al. 2003). When including five internalising disorders by expanding phobia into animal and situational phobia and adding panic disorder in the model, two factors were identified with one

primarily loaded on MDD and GAD, the other on animal and situational phobia, both with strong additive genetic effects on the disorders, which suggested that genetic factors were also largely responsible for the pattern of comorbidity that resulted in the division of internalising disorders into two dimensions reflecting anxious-misery and fear respectively (Kendler et al. 2003). A review article supported the view that genetic factors play an important role in explaining the comorbidity between MDD and GAD, and suggested that non-shared environments explain almost all the covariance between MDD and agoraphobia and more than 40% between MDD and GAD for female twins (Cerdá et al. 2010).

Factor and latent class analyses of MDD

The second phenotypic analysis chapter aimed at answering the question of whether or not depression is one disorder. I examined symptom dimensions and patient subtypes within depression using different symptom profiles. The statistical methods adopted are factor analysis (FA) and latent class analysis (LCA), which provided complementary results, suggesting a number of clinically significant subtypes including the suicidal cognitive, the somatic atypical and anxiety subtypes. Using FA and LCA on three levels of symptom data, I was able to identify three both statistically and clinically significant subtypes including the suicidal cognitive subtype, the atypical somatic subtype and the anxiety / agitation subtype. I will discuss each of them in turn.

The suicidal cognitive subtype

There are two factors from FA on the 9 DSM-IV MDD A criteria, separating cognitive symptoms of worthlessness/guilt and suicidal ideation from the other criteria. Studies using population samples support a coherent dimension in the nine DSM-IV MDD A criteria (Aggen et al. 2005; Muthén 1989), whereas a study using DSM MDD samples detected significant heterogeneity (Lux & Kendler 2010). Our results are convergent with prior analysis from Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD) that there is significant heterogeneity within the DSM-IV A criteria for MDD, particularly between somatic and cognitive criteria (Lux & Kendler 2010).

With more detailed assessment of DSM-IV criteria (symptoms of melancholia, anxiety and Beck's cognitive trio were all included in the analysis), a more informative picture emerged. The suicide/hopeless factor had prominent loadings on suicidal symptoms and Beck's triad of hopelessness, worthlessness and helplessness (somewhat more weakly), which support the Beck's cognitive theory reflecting distorted views of the self, the world and the future (Beck & Alford 2009; Beck et al. 1961; Beck & Young 1979). This, together with the cognitive factor emerged from the nine DSM-IV MDD A criteria suggest that the A7 criterion of feelings of worthlessness and guilt together with A9 criterion of suicidal thought do not add to a unitary psychopathological construct as they were loaded on different depression symptom factors.

The non-suicidal class identified in the nine DSM-IV criteria, characterised by low endorsement on guilt/worthlessness and suicide is the first time it has been identified using LCA. Our non-suicidal class with zero endorsement on suicidal

thoughts and low on feelings of worthlessness and guilt somewhat complimented the cognitive factor primarily loaded on those items. LCA on 27 fully expanded dataset resulted in two non-suicidal classes differed primarily on endorsements on worthlessness, guilt/remorse, hopelessness and helplessness. The non-suicidal/low guilt class has low endorsements on cognitive symptoms indicative of Beck's cognitive model of depression, the result of which complimented the interpretation of the suicidal/hopeless factor from the FA (Beck 1967; Beck 1963).

The epidemiology of suicide in Chinese women differs from Western studies in a number of ways, which might give partial explanation for the non-suicidal class from all three LCAs. First, suicidal rate is two to three times higher in China compare to global average (Harvard School of Public Health et al. 1996). Second, China is one of the few countries where suicidal rate is higher in women than in men (3:1) (Ji et al. 2001; Anon n.d.), and is especially high in rural areas (3:1) (Phillips et al. 2002; Phillips et al. 2007; Qin & Mortensen 2001). Third, the prevalence rates of Axis-I mental illnessness including MDD is much lower (30%) among the suicides than reported in the West (Law & Liu 2008). Law and Liu described extensively about how culture might be interacting with the changing role of women in modern China and hypothesised that the high number of suicides among young women in rural China might be due to acute interpersonal or financial crisis leading to impulsive act of lethal pesticides intake and the subsequent lack of sufficient medical care.

The somatic atypical subtype

When the bi-directional neurovegetative symptoms of appetite/weight changes and sleep disturbances were expanded from the nine DSM-IV criteria, two atypical

symptom dimensions emerged from FA which is consistent with several prior findings (Bech et al. 2011; Kendler et al. 1996; Matza et al. 2003). This gives evidence to atypical depression being a valid independent diagnostic entity (Stewart et al. 1993). In contrary to most prior findings, the atypical feature of sleep disturbance was moderately independent from appetite/weight symptoms, possibly due to our sample size and statistical power. This is however in congruent with a meta-analysis which resulted in separate factors for sleep problems and weight loss (Shafer 2006). Sleep disturbances in severe MDD may form a psychopathological dimension, which is at least partially independent of other neurovegetative changes (Kupfer 1995).

We have identified five LCA studies that included the disaggregated DSM-IV criteria, all of which supported at least one atypical class (Kendler et al. 1996; Lamers et al. 2010; Lamers et al. 2012; Sullivan et al. 1998; Sullivan et al. 2002). However, the endorsement frequencies of the atypical class in the most comparable Western studies: 15.6% in NCS-R (Lamers et al. 2012) and 24.6% in NESDA (Lamers et al. 2010), were greater than in our Han Chinese sample (~7.4%). The atypical class can be interpreted in the context of the two factors primarily loaded weight/appetite changes and sleeping disturbances respectively.

The melancholia and anxiety subtypes

When the melancholic symptoms of loss of ability to enjoy good things, depression different from grief and worse mood in the morning were included in FA, they were all weakly loaded on the general retarded depressive factor which contained prominent loadings on traditional mood changes, psychomotor retardation symptoms. The endorsement frequencies of those three items do not differ significantly across all LCA

classes, which compliments the FA findings that the melancholic subtype of depression is not apparent in our data. This is most likely to be due to the fact cases recruited in our study suffered from the most severe depression. The prevalence rate of cases meeting DSM-IV melancholia is 85.57%, which empirically supports this view.

Epidemiological studies have found that rates of MDD are lower in East Asia than in most countries (Parker et al. 2001). There has been supporting evidence that the Chinese tend to deny depression or express it somatically (Kleinman 1982; Parker et al. 2001). The leading theory is cultural stoicism and stigmatization. A recent epidemiological study in Taiwan (of similar ethnic origin to the Han Chinese) reported that individuals with MDD were much more impaired than in parallel US studies and were only one third as likely to seek professional help (Liao et al. 2012). This might explain the clinical severity of our sample that was ascertained entirely through psychiatric treatment centres.

The second largest factor, which was newly emerged from the 27 items FA, reflected agitated and anxious depressive symptoms. It most closely resemble those describing agitated depression (Koukopoulos & Koukopoulos 1999), which was recognised as a subtype in the Research Diagnostic Criteria (Spitzer et al. 1978). We know depression is highly comorbid with anxiety disorder, which has been previously reported in our study using preliminary data (Li et al. 2011). The prevalence rate of any of the seven comorbid anxiety disorders assessed in our study using the entire case sample of size 6,207 is 50.25%. There are both unique as well as shared genetic effects between the two disorders, whose uniqueness give evidence to the agitation / anxiety dimension from our data.

Consistency of depression subtypes

Due to the difference in sample characteristics and items assessed between our data and those reported using the same statistical methods, it is hard to compare results. However, the most commonly identified subtypes of depression using Western samples including cognitive, somatic and anxiety dimensions were confirmed in our study. Furthermore, the observation that latent classes identified using Western samples most commonly differ in levels of severity, is also supported in our data. This evidence suggests that there is a high degree of similarity in the symptom patterns of depression between China and the West. Furthermore, I was able to identify two classes that do not differ in severity but in nature, including the atypical class and the non-suicidal class. The atypical subtype has been previously reported. However, my study is the first to identify the non-suicidal class, which most likely to be due to cultural reasons described above. The fact that I was able to identify subtypes of depression among severely depressed patients in a homogeneous group which coincides with previously reported subtypes using different instruments in different populations suggests that the symptom patterns observed in the data are applicable elsewhere. The subtypes identified will be used for association with genotype data, with the goal of identifying loci that are not only susceptible to MDD, but also to subtypes of MDD.

Rare variant analysis

In chapter 5, I presented results from gene based analysis as well as gene enrichment test on exonic SNPs. Although I did not find convincing evidence that functional variants, rare or common, contributed to disease susceptibility, I did find suggestive evidence that cases might be enriched for the total number of non-synonymous (NS)

SNPs. This is supported by results from a regression analysis of NS : synonymous (S) ratio on disease status. The *TLR4* gene exceeded gene-wide significant threshold in the variant count analysis, but inflation in the QQ plot should caution that the finding might be a false positive. I was unable to account for this inflation. Fisher's method of combining P values of SNP association tests identified one gene, *PCDHGA4* that exceeded gene-wide significance. The other borderline significant hits resulted from gene enrichment tests on both common and rare SNPs included *PSG8*, *NLRP1*, *KIAA1755*, *OR13C9*. All these genes will need further validation and replication.

My analysis uncovered a number of difficulties for interpretation of the findings. Perhaps most importantly, I found that mismapping of DNA reads in the cases could lead to false positives. This was most evident for the results obtained from raw reads covering in six genes that scored above threshold: *HSPA1B, HSPA1A, HSPA2, PCDHGA4, FLG2, ZNF705A*. The three heat shock protein genes had remarkably high scores (i.e. $-\log_{10}(P_{HSPA1B}) > 50$), raising suspicion that there might be some artefacts. Visual inspection of the results, using the viewer Integrative Genomics Viewer (IGV), indicated that the strong signals are mostly likely to be driven by incorrect mapping of the sequence reads. The signals disappeared when analysed SNPs called from filtered reads, confirming that the signals were driven by poor mapping of bad quality reads. It is not clear why cases should have more bad quality reads mapped to the heat shock protein regions than controls. In fact as I showed, some reads that had been mapped to the heat shock proteins should have been mapped to the human mitochondrial genome.

A second problem for my approach is that sequencing error likely contaminates my findings. In this respect it is important to note that the vast majority of variants identified were rare. This observation might be attributable to the poor sequencing

quality, in other words most of what I found could simply be sequencing error. Even with stringent filtering I still observed that almost 80% of the total variant sites have five or fewer alternative alleles and 63% are singletons. If the variants were primarily due to sequencing error, then most would be classified as non-synonymous. In our data this is true: overall, the NS:S ratio was about 2.66 for raw data. But when I applied increasingly stringent filters, the ratio dropped, indicating that the filters were behaving as expected in enriching for true variants.

My finding is consistent with recent papers that have explored allele frequency spectra in exomic data. Deep (111X) exome sequencing on more than 15K genes on 2,440 individuals demonstrated that the majority of single nucleotide variants are rare (86% with MAF<0.5%) and previously unknown with 57% singletons and 72% with three or fewer minor alleles (Tennessen et al. 2012). Another sequencing project (with median read depth 27) on fewer (202) genes but more (14K) individuals revealed that more than 95% variants were rare (MAF <=0.5%) with approximately 90% previously unknown (Nelson et al. 2012). The 1000 Genome Exon Pilot project sequenced roughly 1000 genes for 700 samples and detected SNPs that are 70% novel and 74% below 1% MAF (Marth et al. 2011). Re-sequencing of 200 exomes at 12X has identified 80% SNPs below 5% MAF and 44% previously unreported (Li et al. 2010). These sequencing projects, though vary in sample size, number of target genes and sequencing depth, resulted in an important and common finding, which is that the vast majority of human functional variants are rare and previously unknown. The proportions of rare variants that were defined by different MAF threshold (0.5%, 1% or 5%) are relative to the sample size.

Given that the addition of a double read filter did enrich for true positives, it might be expected that the findings using the SNPs identified after applying this filter would be more robust. The enrichment P values obtained from the analysis of these data were disappointing, suggesting that there are no rare functional variants. As Table 5. 1 (Chapter 5) shows, the number of variants that passed this filter was extremely small. One possibility is that I did not have sufficient power to detect an effect. With an average of 55% whole exome coverage from reads filtered on 1X sequencing data, calling only SNPs supported by two reads would remove most SNPs covered by only one read. Data show that the harsh filtering has removed 98% of the concatenated SNP sites compare to filter 1 data (this is the most likely reason for the removal of the *PCDHGA4* signals for example). The borderline significant signals resulted from gene enrichment test on both common and rare SNP tests with double read filtered data were driven by very few SNPs, unlike in the cases of raw and filter 1 data.

I can assess the loss of power from using 1X sequence data by comparing the number of SNPs detected in my sample with other sequencing data, based on the findings discussed above. In our data, the total number of SNPs supported by at least two reads in the exome (targeting 96Mb) of whole genome sequencing data of 1X coverage resulted in an average of 7,644 exonic SNPs per individual. In comparison, whole exome sequencing (targeting 28Mb) from 438 samples and whole genome sequencing on 37 samples focusing only on the same genomic regions as the exome data resulted in 18K unfiltered and 16K filtered SNPs (Kiezun et al. 2012). Their numbers are more than twice of what we discovered with a target region three times less using 5% of our sample size.

Despite the loss of power, it is still worth considering whether the genes I have identified are likely to be involved in MDD. One simple requirement is that the genes should be expressed in the central nervous system (CNS). The most promising finding from this study was the *PCDHGA4* gene, which showed significant effect from gene enrichment test on SNPs called from both the raw reads as well as the filtered reads. This gene is a member of the protocadherin gamma gene cluster, which has a immunoglobulin-like organization. A large number of cadherin superfamily genes have been identified to express in the CNS, as synaptic components and neuronal circuitry (Yagi & Takeichi 2000). No previous publication has linked this gene to depression, but a number of cadherin superfamily genes have been linked to schizophrenia (Barden & Morissette 1999; Crowe & Vieland 1999; Riley et al. 1998; Schwab et al. 1997; Van Broeckhoven & Verheyen 1999). The only way to verify the finding is by deep sequencing of this and other candidate genes using cost effective approach such as the molecular inversion probes technique (Hiatt et al. 2013; O'Roak et al. 2012).

GWAS of MDD

In chapter 6, I presented preliminary results for a genome wide association (GWA) analysis using genotype dose estimates imputed from low pass (1X coverage) sequencing data. GWA analysis on MDD correcting for population structure resulted in one common SNP exceeding a genome wide significant threshold. None of the eight genes near the SNP have been reported to be associated with depression. However, the *SEMA6B* gene encodes a member of the semaphorin family, which might be involved in both peripheral and central nervous system development. For a GWA analysis on cases with high recurrence, the top SNPs lie in the *SIRT1* gene, which has been reported to be

associated with major depressive disorder in the Japanese population (Kishi et al. 2010). Another study demonstrated that the *SIRT1* influenced anxiety and exploratory behaviour in mice (Feliciano 2012).

Our finding is consistent with conclusion from previous GWAS on MDD using thousands of samples of European ancestry that common SNPs with large effects do not contribute to the genetic susceptibility for depression. The replicability of GWAS results across populations has been assessed in a study which presented evidence that most disease loci identified using GWAS have been replicated using people of European and East Asian ancestry with strong correlation of odds ratios indicating that the underlying common causal variants are shared between the two ancestries (Marigorta & Navarro 2013).

However, GWA analysis on disease status with controls and case subsets (i.e. severe class from LCA, melancholia, postnatal depression, low environmental risks) showed significant inflation of P values, which suggested that there might be significant heterogeneity within our case sample that dilutes the effects of disease susceptible loci. Therefore, taking a subset of cases with higher degree of severity and homogeneity might give us higher power to detect disease susceptible loci for that specific subtype. Other explanations for the inflation are uncontrolled artefacts, such as city and sequencing plate.

The SNP with lowest P value was found when the analysis included only cases with low environmental risk ($-\log_{10}(P_{rs7257142}) = 13$), which is the same as the top SNP for a GWA on disease status. It has been previously reported that there is significant heterogeneity in the environmental impact on depression, (Foley et al. 1996; Kendler et al. 1993; Kendler et al. 2004). That is to say, there are potentially a number of different

ways in which a stressful life event might impact on a person's mental health; the pathways leading from environmental insult to depression onset are varied. Furthermore, the outcome of a life event will likely depend, at least in part, on the genetic makeup of the individual. In other words there is likely to be an effect of gene by environmental interaction. Our significant result shows that including only cases with low environmental risks, although reduced sample size, increased the power of detecting disease susceptible loci.

Another source of heterogeneity in a single diagnosis of depression lies in the multiple dimensions of genetic liability, previously shown in a twin study (Kendler et al. 2013) and replicated here (see chapter 4). Therefore the effect of susceptibility loci that are unique to each dimension will be diluted when they are aggregated within a single diagnostic category. In order to test the genetic heterogeneity for specific subtypes of depression, I carried out association analyses on eight factor scores representing symptom dimensions of interest generated from factor analysis on three datasets (see chapter 4). Of the association analyses on the eight factor scores, none exceeded a genome wide significance threshold. One explanation for the negative finding is that there is not enough statistical power with the samples available to detect genetically significant subtypes within this highly homogeneous case population. The other explanation is that although symptoms might manifest differently in depressed individuals, there is actually one common genetic pathway.

One of the major issues in research in depression is the question of disease heterogeneity. This is particularly relevant to genetic mapping since it impacts on two related issues: the question of power to detect an effect and the interpretation of any findings that emerge from a GWAS. Clearly, if we were able to define genetically

homogenous subgroups we might be better placed to find significant hits, since we would expect the genetic effects on the subtype to be relatively larger than in the entire sample. (In the subgroup, the genetic effects are not diluted by the confounding presence of other subtypes, with differing genetic bases). Second, genetic loci identified within a single homogenous sub-group are expected to reflect the biological basis of depression in that subgroup and potentially therefore to have more biological coherence, belonging for example to the same biochemical pathway.

My results add to this debate, which can be interpreted in the following ways. The fact that some case subtypes gave more significant results could be due to chance, or uncontrolled population structure in that particular subtype population as principal component analysis (PCA) was done on the entire sample. Another interpretation of the result is that there is true genetic heterogeneity in our depression case sample. Subdividing case subtypes based on symptom profiles does indeed give more power to detect disease or subtype specific SNP. Supporting evidence is the cluster of associated SNPs observed when selecting cases with high recurrence.

In the thesis, I performed LCA and FA on depression symptoms and tested for the genetic basis of each subtype. Another way would be to calculate the genetic correlations maximised between individuals using genotype data, and then test the genetic correlations between LCA classes as well as between dysthymia and melancholia. There remains a rich area of research to seek evidence for the common and individual genetic bases of depression subtypes.

Mitochondrial coverage in MDD

Chapter 7 follows up from the unexpected finding that cases have significantly more variants in all 13 protein coding mitochondrial (MT) genes using the gene-based analytical approach described in Chapter 5. The finding was shown to be due to MT copy number difference between cases and controls. Further experimental results rejected the hypotheses that the excess of reads mapping to the MT genome could be due to bacterial contamination or drug effects. Both human and mice experimental data suggested that the finding was a result of chronic stress.

There are a number of points that warrant discussion. The relationship between MT copy number and depression has not, to my knowledge, been reported before. MT is essential for neurotransmission, neuronal plasticity, cellular resilience to stress and behavioural adaption (Chen et al. 2010; Gleichmann & Mattson 2011; Jonas 2009; Jonas 2006; Kato 2008; MacAskill et al. 2010; Mattson et al. 2008; Quiroz et al. 2008). Mitochondria dysfunction is reported to play an important role in neurodegenerative disorders, including Parkinson's disease (Henchcliffe & Beal 2008), Huntington's disease (Turner & Schapira 2010), and Alzheimer's disease (Galindo et al. 2010). Depression or mood related features are common to all these neurodegenerative disorders (Bassiony et al. 2002; Paulsen et al. 2001; Reijnders et al. 2008; Zubenko, Zubenko, et al. 2003) and could be early signs of later onset of Parkinson's (Ishihara & Brayne 2006) and Alzheimer's diseases (Green et al. 2003). Moreover, a review article presented pharmacological, genetic, cellular and post-mortem evidence for the role of mitochondrial dysfunction in mood disorders (Manji et al. 2012). A number of point mutations in the MT genome have been reported to be associated with psychiatric

disorders including schizophrenia, bipolar and MDD (Rollins et al. 2009; Sequeira et al. 2012).

I found one study reporting a significant association between leukocyte MT copy number and depression in community dwelling elderly women in Korea. Interestingly this study reported a significantly *lower* MT copy number in cases compared to controls (Kim et al. 2011). This result might be explained by the presence of covariates modulating the relationship between depression and MT copy number. For example we see a complex effect of age, in that while age is related to an increase in copy number in controls, copy number decreases with age in cases. Our results are limited to one sex, and it is possible that there is a sex effect.

Two features of the MT copy number deserve comment: the size of the effect and the fact that we find this in saliva. The effect is relatively small, and its biological interpretation not clear, since a single mitochondrion can contain a variable number of MT genomes. Our OR was calculated for the increase in the average MT read depth, which is not easily translated into the number of mitochondria; however taking an overall measure we see that the mean MT read depth increases from 57 in controls to 63 in cases with the raw reads and 43 to 47 with the filtered reads, which suggests that the increase in meaningful.

Finding the MT increase in saliva indicates that the effect is likely to be systemic. We found increase in mitochondria in the blood of stressed mice and we presume, but do not currently know, that the same is true in humans. It will be important to determine whether, as we suspect, all tissues show this change. We do not yet know the functional consequences of these changes.

The cause of the mitochondria increase is unknown, but on the assumption that the finding occurs in all tissues, then a systemic origin is likely. One possibility is an increase in corticosteroids. The association between chronic stress and corticosteroids level is well known (de Kloet et al. 2005). Corticosteroids promote metabolic activity, and this may drive the MT biogenesis. In human and laboratory animals, stress increases levels of glucocorticoids (Schmidt et al. 2011). Evidence shows that glucocorticoids bound to glucocorticoid receptors regulates mitochondria gene expression, which is mediated by B-cell lymphoma protein 2 (Du et al. 2009). These data support the MT link between chronic stress and psychiatric disorders. Further experiments are needed to establish the association between corticosteroids level and MT copy number.

Finally it is worth stressing that the MT copy number increase is likely to be an important biomarker for depression. However its utility remains to be established. For example, it will be important to determine whether there is an association between MT copy number and antidepressant drug response. In current clinical practice, half of the patients respond positively on antidepressants. It takes at least three months for patients to show signs of remission. An association between MT copy number and drug response would help with treatment decisions.

Bibliography

Abecasis, G.R. et al., 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422), pp.56–65.

Abkevich, V. et al., 2003. Predisposition locus for major depression at chromosome 12q22-12q23.2. *American journal of human genetics*, 73(6), pp.1271–1281.

Aggen, S.H., Neale, M.C. & Kendler, K.S., 2005. DSM criteria for major depression: evaluating symptom patterns using latent-trait item response models. *Psychological medicine*, 35(4), pp.475–487.

Akaike, H., 1987. Factor analysis and AIC. *Psychometrika*, 52(3), pp.317–332.

Akiskal, H.S. & McKinney, W.T., 1973. Depressive Disorders: Toward a Unified Hypothesis Clinical, experimental, genetic, biochemical, and neurophysiological data are integrated. *Science*, 182(4107), pp.20–29.

Almgren, P. et al., 2011. Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia*, 54(11), pp.2811–2819.

Altshuler, D.M. et al., 2010. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311), pp.52–58.

American Psychiatric Association, 1994. *Diagnostic and Statistical Manual of Mental Disorders* Fourth Edition., Washington DC: American Psychiatric Association.

Andreasen, N.C. & Grove, W.M., 1982. The classification of depression: traditional versus mathematical approaches. *The American journal of psychiatry*, 139(1), pp.45–52.

Angst, J. & Clayton, P., 1986. Premorbid personality of depressive, bipolar, and schizophrenic patients with special reference to suicidal issues. *Comprehensive psychiatry*, 27(6), pp.511–532.

Anon, 2003. The International HapMap Project. *Nature*, 426(6968), pp.789–796.

Anon, WHO | Suicide prevention. *WHO*. Available at: http://www.who.int/mental_health/prevention/en/ [Accessed June 17, 2013].

Barden, N. & Morissette, J., 1999. Chromosome 13 workshop report. *American journal of medical genetics*, 88(3), pp.260–262.

Bassiony, M.M. et al., 2002. The relationship between delusions and depression in Alzheimer's disease. *International journal of geriatric psychiatry*, 17(6), pp.549–556.

Bebbington, P.E. et al., 1984. Misfortune and resilience: a community study of women. *Psychological medicine*, 14(2), pp.347–363.

Bech, P. et al., 2011. Factor structure and dimensionality of the two depression scales in STAR*D using level 1 datasets. *Journal of Affective Disorders*, 132(3), pp.396–400.

Beck, A.T. et al., 1961. An inventory for measuring depression. *Archives of general psychiatry*, 4, pp.561–571.

Beck, A.T., 1967. *Depression: clinical, experimental, and theoretical aspects*, Hoeber Medical Division, Harper & Row.

Beck, A.T., 1963. THINKING AND DEPRESSION. I. IDIOSYNCRATIC CONTENT AND COGNITIVE DISTORTIONS. *Archives of General Psychiatry*, 9, pp.324–333.

Beck, A.T. & Alford, B.A., 2009. *Depression: causes and treatment*, Philadelphia: University of Pennsylvania Press.

Beck, A.T. & Young, J.E., 1979. *Cognitive therapy of depression*, New York: Guilford Press.

Bentler, P.M., 1990. Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), pp.238–246.

Bierut, L.J. et al., 1999. Major depressive disorder in a community-based twin sample: are there different genetic and environmental contributions for men and women? *Archives of general psychiatry*, 56(6), pp.557–563.

Bolton, J.M. et al., 2008. Exploring the correlates of suicide attempts among individuals with major depressive disorder: findings from the national epidemiologic survey on alcohol and related conditions. *The Journal of Clinical Psychiatry*, 69(7), pp.1139–1149.

Breen, G. et al., 2011. A genome-wide significant linkage for severe depression on chromosome 3: the depression network study. *The American journal of psychiatry*, 168(8), pp.840–847.

Van Broeckhoven, C. & Verheyen, G., 1999. Report of the chromosome 18 workshop. *American journal of medical genetics*, 88(3), pp.263–270.

Brown, C. et al., 1996. Treatment outcomes for primary care patients with major depression and lifetime anxiety disorders. *The American Journal of Psychiatry*, 153(10), pp.1293–1300.

Brown, G.W. & Harris, T., 2001. *Social origins of depression: a study of psychiatric disorder in women*, London: Routledge.

Brown, T.A. et al., 2001. Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, 110(4), pp.585–599.

Browning, S.R. & Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American journal of human genetics*, 81(5), pp.1084–1097.

Burbach, D.J. & Borduin, C.M., 1986. Parent-child relations and the etiology of depression: A review of methods and findings. *Clinical Psychology Review*, 6(2), pp.133–153.

Cadoret, R.J., Woolson, R. & Winokur, G., 1977. The relationship of age of onset in unipolar affective disorder to risk of alcoholism and depression in parents. *Journal of psychiatric research*, 13(3), pp.137–142.

Camp, N.J. et al., 2005. Genome-wide linkage analyses of extended Utah pedigrees identifies loci that influence recurrent, early-onset major depression and anxiety disorders. *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 135B(1), pp.85–93.

Cardno, A. & Marshall, E., 1999. Heritability estimates for psychotic disorders: The maudsley twin psychosis series. *Archives of General Psychiatry*, 56(2), pp.162–168.

Carragher, N. et al., 2009. Subtypes of depression in a nationally representative sample. *Journal of affective disorders*, 113(1-2), pp.88–99.

Cassano, G.B. et al., 1989. The nature of depression presenting concomitantly with panic disorder. *Comprehensive Psychiatry*, 30(6), pp.473–482.

Cerdá, M. et al., 2010. Genetic and environmental influences on psychiatric comorbidity: a systematic review. *Journal of Affective Disorders*, 126(1-2), pp.14–38.

Chen, G., Henter, I.D. & Manji, H.K., 2010. Translational research in bipolar disorder: emerging insights from genetically based models. *Molecular psychiatry*, 15(9), pp.883–895.

Chen, L.P. et al., 2010. Sexual abuse and lifetime diagnosis of psychiatric disorders: systematic review and meta-analysis. *Mayo Clinic proceedings. Mayo Clinic*, 85(7), pp.618–629.

Claassen, C.A. et al., 2007. Clinical differences among depressed patients with and without a history of suicide attempts: findings from the STAR*D trial. *Journal of Affective Disorders*, 97(1-3), pp.77–84.

Clayton, P.J. et al., 1991. Follow-up and family study of anxious depression. *The American Journal of Psychiatry*, 148(11), pp.1512–1517.

Cohen, J.C. et al., 2004. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science (New York, N.Y.)*, 305(5685), pp.869–872.

Cole, J., McGuffin, P. & Farmer, A.E., 2008. The classification of depression: are we still confused? *The British journal of psychiatry: the journal of mental science*, 192(2), pp.83–85.

Coryell, W. et al., 1988. Depression and panic attacks: the significance of overlap as reflected in follow-up and family study data. *The American Journal of Psychiatry*, 145(3), pp.293–300.

Costello, C.G., 1982. Social factors associated with depression: a retrospective community study. *Psychological medicine*, 12(2), pp.329–339.

Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, 45(9), pp.984–994.

Crowe, R.R. & Vieland, V., 1999. Report of the Chromosome 5 Workshop of the Sixth World Congress on Psychiatric Genetics. *American journal of medical genetics*, 88(3), pp.229–232.

Davidson, J. et al., 1988. A study of depressive typologies using grade of membership analysis. *Psychological medicine*, 18(1), pp.179–189.

Davidson, J.R. et al., 1989. Classification of depression by grade of membership: a confirmation study. *Psychological medicine*, 19(4), pp.987–998.

Demirkan, A. et al., 2011. Genetic risk profiles for depression and anxiety in adult and elderly cohorts. *Molecular psychiatry*, 16(7), pp.773–783.

Distel, M.A. et al., 2009. The five-factor model of personality and borderline personality disorder: a genetic analysis of comorbidity. *Biological psychiatry*, 66(12), pp.1131–1138.

Du, J. et al., 2009. Dynamic regulation of mitochondrial function by glucocorticoids. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), pp.3543–3548.

Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. *PLoS genetics*, 9(3), p.e1003348.

E, T., Muir, W. & J, A., 1967. A factor analysis of the Beck Inventory of Depression. *Journal of Consulting Psychology*, 31(1), pp.23–28.

Eaton, W.W. et al., 1989. DSM-III major depressive disorder in the community. A latent class analysis of data from the NIMH epidemiologic catchment area programme. *The British journal of psychiatry: the journal of mental science*, 155, pp.48–54.

Eichler, E.E. et al., 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nature reviews. Genetics*, 11(6), pp.446–450.

Endicott, J., Andreasen, N. & Spitzer, R.L., 1975. *Family History-Research Diagnostic Criteria*, New York: Biometrics Research, New York State Psychiatric Institute.

Eysenck, H.J., 1970. The Classification of Depressive Illnesses. *The British Journal of Psychiatry*, 117(538), pp.241–250.

Eysenck, H.J. & Eysenck, S.B.G., 1975. *Manual of the Eysenck Personality Questionnaire*, London: Hodder and Stoughton.

Fanous, A. et al., 2002. Neuroticism, major depression and gender: a population-based twin study. *Psychological medicine*, 32(4), pp.719–728.

Farmer, A. & McGuffin, P., 1989. The classification of the depressions. Contemporary confusion revisited. *The British journal of psychiatry: the journal of mental science*, 155, pp.437–443.

Fava, M. et al., 2000. Anxiety disorders in major depression. *Comprehensive Psychiatry*, 41(2), pp.97–102.

Fava, M. et al., 2004. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR*D. *Psychological Medicine*, 34(7), pp.1299–1308.

Feliciano, P., 2012. SIRT1 and anxiety. *Nature Genetics*, 44(2), pp.120–120.

Fergusson, D.M., Boden, J.M. & Horwood, L.J., 2008. Exposure to childhood sexual and physical abuse and adjustment in early adulthood. *Child abuse & neglect*, 32(6), pp.607–619.

Fergusson, D.M., Horwood, L.J. & Lynskey, M.T., 1996. Childhood sexual abuse and psychiatric disorder in young adulthood: II. Psychiatric outcomes of childhood sexual abuse. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35(10), pp.1365–1374.

Fisher, R.A., 1970. *Statistical methods for research workers* 14th ed., revised and enlarged., Edinburgh: Oliver and Boyd.

Flora, D.B. & Curran, P.J., 2004. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), pp.466–491.

Foley, D.L., Neale, M.C. & Kendler, K.S., 1996. A longitudinal study of stressful life events assessed at interview with an epidemiological sample of adult twins: the basis of individual variation in event exposure. *Psychological medicine*, 26(6), pp.1239–1252.

Freitag, C.M., 2006. The genetics of autistic disorders and its clinical relevance: a review of the literature. *Molecular Psychiatry*, 12(1), pp.2–22.

Friedman, A.S. et al., 1963. SYNDROMES AND THEMES OF PSYCHOTIC DEPRESSION: A FACTOR ANALYSIS. *Archives of general psychiatry*, 9, pp.504–509.

Galindo, M.F. et al., 2010. Mitochondrial biology in Alzheimer's disease pathogenesis. *Journal of neurochemistry*, 114(4), pp.933–945.

Garvey, M.J., Tollefson, G.D. & Tuason, V.B., 1987. Comparison of major depressions with and without panic attacks. *Comprehensive Psychiatry*, 28(1), pp.65–67.

Gater R, T.M., 1998. Sex differences in the prevalence and detection of depressive and anxiety disorders in general health care settings: Report from the world health organization collaborative study on psychological problems in general health care. *Archives of General Psychiatry*, 55(5), pp.405–413.

GENDEP Investigators, MARS Investigators & STAR*D Investigators, 2013. Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. *The American journal of psychiatry*, 170(2), pp.207–217.

Gerlsma, C., Emmelkamp, P.M.G. & Arrindell, W.A., 1990. Anxiety, depression, and perception of early parenting: a meta-analysis. *Clinical Psychology Review*, 10(3), pp.251–277.

Gershon, E.S. et al., 1986. Validation of criteria for major depression through controlled family study. *Journal of affective disorders*, 11(2), pp.125–131.

Girard, S.L. et al., 2011. Increased exonic de novo mutation rate in individuals with schizophrenia. *Nature genetics*, 43(9), pp.860–863.

Gleichmann, M. & Mattson, M.P., 2011. Neuronal calcium homeostasis and dysregulation. *Antioxidants & redox signaling*, 14(7), pp.1261–1273.

Goodstadt, L., 2010. Ruffus: A Lightweight Python Library for Computational Pipelines. *Bioinformatics*, p.btq524.

Gorwood, P., 2004. Generalized anxiety disorder and major depressive disorder comorbidity: an example of genetic pleiotropy? *European Psychiatry: The Journal of the Association of European Psychiatrists*, 19(1), pp.27–33.

De Graaf, R. et al., 2003. Temporal sequencing of lifetime mood disorders in relation to comorbid anxiety and substance use disorders--findings from the Netherlands Mental Health Survey and Incidence Study. *Social Psychiatry and Psychiatric Epidemiology*, 38(1), pp.1–11.

Green, R.C. et al., 2003. Depression as a risk factor for Alzheimer disease: the MIRAGE Study. *Archives of neurology*, 60(5), pp.753–759.

Grinker, R.R. et al., 1961. *The phenomena of depressions*, Oxford, England: Hoeber Med. Div., Harper & Row.

Grinker, R.R., 1961. *The Phenomena of Depressions*, Paul B. Hoeber.

Grove, W.M. et al., 1987. Isolation and characterization of a nuclear depressive syndrome. *Psychological medicine*, 17(2), pp.471–484.

Grunhaus, L., 1988. Clinical and psychobiological characteristics of simultaneous panic disorder and major depression. *The American Journal of Psychiatry*, 145(10), pp.1214–1221.

Grunhaus, L. et al., 1994. Clinical characteristics of patients with concurrent major depressive disorder and panic disorder. *The American Journal of Psychiatry*, 151(4), pp.541–546.

Hamilton, M., 1960. A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*, 23, pp.56–62.

Harvard School of Public Health, World Health Organization & World Bank, 1996. *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*, Cambridge, MA: Published by the Harvard School of Public Health on behalf of the World Health Organization and the World Bank ; Distributed by Harvard University Press.

Hasin, D.S. et al., 2005. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Archives of General Psychiatry*, 62(10), pp.1097–1106.

Haslam, N. & Beck, A.T., 1993. Categorization of major depression in an outpatient sample. *The Journal of nervous and mental disease*, 181(12), pp.725–731.

Henchcliffe, C. & Beal, M.F., 2008. Mitochondrial biology and oxidative stress in Parkinson disease pathogenesis. *Nature clinical practice. Neurology*, 4(11), pp.600–609.

Hettema, J.M., 2008. The nosologic relationship between generalized anxiety disorder and major depression. *Depression and Anxiety*, 25(4), pp.300–316.

Hiatt, J.B. et al., 2013. Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. *Genome Research*, 23(5), pp.843–854.

Hirschfeld, R.M. et al., 1989. Premorbid personality assessments of first onset of major depression. *Archives of general psychiatry*, 46(4), pp.345–350.

Holden, G.W. & Edwards, L.A., 1989. Parental attitudes toward child rearing: Instruments, issues, and implications. *Psychological Bulletin*, 106(1), pp.29–58.

Holmans, P. et al., 2007. Genetics of recurrent early-onset major depression (GenRED): final genome scan report. *The American journal of psychiatry*, 164(2), pp.248–258.

Holmans, P. et al., 2004. Genomewide significant linkage to recurrent, early-onset major depressive disorder on chromosome 15q. *American journal of human genetics*, 74(6), pp.1154–1167.

Howland, R.H. et al., 2009. Concurrent anxiety and substance use disorders among outpatients with major depression: clinical features and effect on treatment outcome. *Drug and Alcohol Dependence*, 99(1-3), pp.248–260.

Hulme, P.A., 2011. Childhood sexual abuse, HPA axis regulation, and mental health: an integrative review. *Western journal of nursing research*, 33(8), pp.1069–1097.

Husain, M.M. et al., 2009. Family history of depression and therapeutic outcome: findings from STAR*D. *The Journal of Clinical Psychiatry*, 70(2), pp.185–195.

Hybels, C.F. et al., 2009. Profiles of depressive symptoms in older adults diagnosed with major depression: latent cluster analysis. *The American journal of geriatric psychiatry: official journal of the American Association for Geriatric Psychiatry*, 17(5), pp.387–396.

Ishihara, L. & Brayne, C., 2006. A systematic review of depression and mental illness preceding Parkinson's disease. *Acta neurologica Scandinavica*, 113(4), pp.211–220.

Ji, J., Kleinman, A. & Becker, A.E., 2001. Suicide in contemporary China: a review of China's distinctive suicide demographics in their sociocultural context. *Harvard review of psychiatry*, 9(1), pp.1–12.

Ji, W. et al., 2008. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nature genetics*, 40(5), pp.592–599.

Jonas, E., 2006. BCL-xL regulates synaptic plasticity. *Molecular interventions*, 6(4), pp.208–222.

Jonas, E.A., 2009. Molecular participants in mitochondrial cell death channel formation during neuronal ischemia. *Experimental neurology*, 218(2), pp.203–212.

Jumper, S.A., 1995. A meta-analysis of the relationship of child sexual abuse to adult psychological adjustment. *Child abuse & neglect*, 19(6), pp.715–728.

Kato, T., 2008. Role of mitochondrial DNA in calcium signaling abnormality in bipolar disorder. *Cell calcium*, 44(1), pp.92–102.

Kendell, R.E., 1976. The classification of depressions: a review of contemporary confusion. *The British journal of psychiatry: the journal of mental science*, 129, pp.15–28.

Kendell, R.E. & Discipio, W.J., 1968. Eysenck Personality Inventory Scores of Patients with Depressive Illnesses. *The British Journal of Psychiatry*, 114(511), pp.767–770.

Kendler, K.S., Neale, M.C., et al., 1993. A longitudinal twin study of personality and major depression in women. *Archives of general psychiatry*, 50(11), pp.853–862.

Kendler, K.S. et al., 1995. A pilot Swedish twin study of affective illness including hospital- and population-ascertained subsamples: results of model fitting. *Behavior genetics*, 25(3), pp.217–232.

Kendler, K.S. et al., 2006a. A Swedish national twin study of lifetime major depression. *The American Journal of Psychiatry*, 163(1), pp.109–114.

Kendler, K.S., Neale, M., et al., 1993. A twin study of recent life events and difficulties. *Archives of general psychiatry*, 50(10), pp.789–796.

Kendler, K.S., 1998. Anna-Monika-Prize paper. Major depression and the environment: a psychiatric genetic perspective. *Pharmacopsychiatry*, 31(1), pp.5–9.

Kendler, K.S., Bulik, C.M., et al., 2000. Childhood sexual abuse and adult psychiatric and substance use disorders in women: an epidemiological and cotwin control analysis. *Archives of General Psychiatry*, 57(10), pp.953–959.

Kendler, K.S. et al., 2007. Clinical indices of familial depression in the Swedish Twin Registry. *Acta Psychiatrica Scandinavica*, 115(3), pp.214–220.

Kendler, K.S. et al., 2001. Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? *Psychological Medicine*, 31(4), pp.605–616.

Kendler, K.S. et al., 1992. Major depression and generalized anxiety disorder. Same genes, (partly) different environments? *Archives of General Psychiatry*, 49(9), pp.716–722.

Kendler, K.S., 1996. Parenting: a genetic-epidemiologic perspective. *The American Journal of Psychiatry*, 153(1), pp.11–20.

Kendler, K.S. et al., 2006b. Personality and major depression: a Swedish longitudinal, population-based twin study. *Archives of general psychiatry*, 63(10), pp.1113–1120.

Kendler, K.S. et al., 1994. The clinical characteristics of major depression as indices of the familial risk to illness. *The British Journal of Psychiatry: The Journal of Mental Science*, 165(2), pp.66–72.

Kendler, K.S. et al., 1996. The identification and validation of distinct depressive syndromes in a population-based sample of female twins. *Archives of General Psychiatry*, 53(5), pp.391–399.

Kendler, K.S. et al., 2007. The sources of co-morbidity between major depression and generalized anxiety disorder in a Swedish national twin sample. *Psychological Medicine*, 37(3), pp.453–462.

Kendler, K.S. et al., 2003. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, 60(9), pp.929–937.

Kendler, K.S., Aggen, S.H. & Neale, M.C., 2013. Evidence for Multiple Genetic Factors Underlying DSM-IV Criteria for Major Depression. *JAMA psychiatry (Chicago, Ill.)*, 70(6), pp.599–607.

Kendler, K.S. & Diehl, S.R., 1993. The genetics of schizophrenia: a current, genetic-epidemiologic perspective. *Schizophrenia bulletin*, 19(2), pp.261–285.

Kendler, K.S., Gardner, C.O. & Prescott, C.A., 1999. Clinical characteristics of major depression that predict risk of depression in relatives. *Archives of general psychiatry*, 56(4), pp.322–327.

Kendler, K.S., Karkowski, L.M. & Prescott, C.A., 1999. Causal relationship between stressful life events and the onset of major depression. *The American journal of psychiatry*, 156(6), pp.837–841.

Kendler, K.S., Kuhn, J.W. & Prescott, C.A., 2004. Childhood sexual abuse, stressful life events and risk for major depression in women. *Psychological medicine*, 34(8), pp.1475–1482.

Kendler, K.S., Myers, J. & Prescott, C.A., 2000. Parenting and adult mood, anxiety and substance use disorders in female twins: an epidemiological, multi-informant, retrospective study. *Psychological Medicine*, 30(2), pp.281–294.

Kendler, K.S. & Prescott, C.A., 1999. A population-based twin study of lifetime major depression in men and women. *Archives of general psychiatry*, 56(1), pp.39–44.

Kendler, K.S. & Prescott, C.A., 2006. *Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders* 1st ed., New York: Guilford Press.

Kendler, K.S., Sham, P.C. & MacLEAN, C.J., 1997. The determinants of parenting: an epidemiological, multi-informant, retrospective study. *Psychological Medicine*, 27(3), pp.549–563.

Kendler, K.S., Thornton, L.M. & Gardner, C.O., 2000. Stressful life events and previous episodes in the etiology of major depression in women: an evaluation of the "kindling" hypothesis. *The American Journal of Psychiatry*, 157(8), pp.1243–1251.

Kessler, R.C. et al., 1996. Comorbidity of DSM-III-R major depressive disorder in the general population: results from the US National Comorbidity Survey. *The British Journal of Psychiatry. Supplement*, (30), pp.17–30.

Kessler, R.C. et al., 1999. Impairment in pure and comorbid generalized anxiety disorder and major depression at 12 months in two national surveys. *The American Journal of Psychiatry*, 156(12), pp.1915–1923.

Kessler, R.C. et al., 1994. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry*, 51(1), pp.8–19.

Kessler, R.C. et al., 1998. Lifetime panic-depression comorbidity in the National Comorbidity Survey. *Archives of General Psychiatry*, 55(9), pp.801–808.

Kessler, R.C. et al., 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), pp.617–627.

Kessler, R.C. et al., 1993. Sex and depression in the National Comorbidity Survey I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders*, 29(2–3), pp.85–96.

Kessler, R.C., 1997. The effects of stressful life events on depression. *Annual Review of Psychology*, 48, pp.191–214.

Kessler, R.C. et al., 2003. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA: The Journal of the American Medical Association*, 289(23), pp.3095–3105.

Kiezun, A. et al., 2012. Exome sequencing and the genetic basis of complex traits. *Nature Genetics*, 44(6), pp.623–630.

Kim, M.-Y. et al., 2011. Leukocyte mitochondrial DNA (mtDNA) content is associated with depression in old women. *Archives of gerontology and geriatrics*, 53(2), pp.e218–221.

King, M.K. et al., 1995. Suicide attempt history in depressed patients with and without a history of panic attacks. *Comprehensive Psychiatry*, 36(1), pp.25–30.

Kishi, T. et al., 2010. SIRT1 gene is associated with major depressive disorder in the Japanese population. *Journal of affective disorders*, 126(1-2), pp.167–173.

Klein DF, 1974. Endogenomorphic depression: A conceptual and terminological revision. *Archives of General Psychiatry*, 31(4), pp.447–454.

Kleinman, A., 1982. Neurasthenia and depression: a study of somatization and culture in China. *Culture, medicine and psychiatry*, 6(2), pp.117–190.

De Kloet, E.R., Joëls, M. & Holsboer, F., 2005. Stress and the brain: from adaptation to disease. *Nature reviews. Neuroscience*, 6(6), pp.463–475.

Kohli, M.A. et al., 2011. The neuronal transporter gene SLC6A15 confers risk to major depression. *Neuron*, 70(2), pp.252–265.

Koukopoulos, A. & Koukopoulos, A., 1999. Agitated depression as a mixed state and the problem of melancholia. *The Psychiatric clinics of North America*, 22(3), pp.547–564.

Kupfer, D.J., 1995. Sleep research in depressive illness: clinical implications--a tasting menu. *Biological psychiatry*, 38(6), pp.391–403.

Lamers, F. et al., 2011. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, 72(3), pp.341–348.

Lamers, F. et al., 2010. Identifying depressive subtypes in a large cohort study: results from the Netherlands Study of Depression and Anxiety (NESDA). *The Journal of clinical psychiatry*, 71(12), pp.1582–1589.

Lamers, F. et al., 2012. Structure of major depressive disorder in adolescents and adults in the US general population. *The British journal of psychiatry: the journal of mental science*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22700082 [Accessed June 22, 2012].

Lango Allen, H. et al., 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467(7317), pp.832–838.

Law, S. & Liu, P., 2008. Suicide in China: unique demographic patterns and relationship to depressive disorder. *Current psychiatry reports*, 10(1), pp.80–86.

Levinson, D.F. et al., 2007. Genetics of recurrent early-onset major depression (GenRED): significant linkage on chromosome 15q25-q26 after fine mapping with single nucleotide polymorphism markers. *The American journal of psychiatry*, 164(2), pp.259–264.

Lewis, A.J., 1934. Melancholia: A Clinical Survey of Depressive States. *The British Journal of Psychiatry*, 80(329), pp.277–378.

Lewis, C.M. et al., 2010. Genome-wide association study of major recurrent depression in the U.K. population. *The American Journal of Psychiatry*, 167(8), pp.949–957.

Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, 25(16), pp.2078–2079.

Li, H. & Durbin, R., 2009a. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 25(14), pp.1754–1760.

Li, H. & Durbin, R., 2009b. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), pp.1754–1760.

Li, Y. et al., 2011. Patterns of co-morbidity with anxiety disorders in Chinese women with recurrent major depression. *Psychological Medicine*, pp.1–9.

Li, Y. et al., 2010. Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. *Nature genetics*, 42(11), pp.969–972.

Liao, S.-C. et al., 2012. Low prevalence of major depressive disorder in Taiwanese adults: possible explanations and implications. *Psychological medicine*, 42(6), pp.1227–1237.

Van Loo, H.M. et al., 2012. Data-driven subtypes of major depressive disorder: a systematic review. *BMC medicine*, 10, p.156.

Lopez, A.D. & Murray, C.C., 1998. The global burden of disease, 1990-2020. *Nature Medicine*, 4(11), pp.1241–1243.

Lunter, G. & Goodson, M., 2011. Stampy: A statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, 21(6), pp.936–939.

Lux, V. & Kendler, K.S., 2010. Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychological medicine*, 40(10), pp.1679–1690.

Lyons, M.J. et al., 1998. A registry-based twin study of depression in men. *Archives of general psychiatry*, 55(5), pp.468–472.

MacArthur, D.G. et al., 2012. A systematic survey of loss-of-function variants in human protein-coding genes. *Science (New York, N.Y.)*, 335(6070), pp.823–828.

MacAskill, A.F., Atkin, T.A. & Kittler, J.T., 2010. Mitochondrial trafficking and the provision of energy and calcium buffering at excitatory synapses. *The European journal of neuroscience*, 32(2), pp.231–240.

MacFadyen, H.W., 1975. The classification of depressive disorders. *Journal of Clinical Psychology*, 31(3), pp.380–394.

Maes, M. et al., 1992. A clinical and biological validation of the DSM-III melancholia diagnosis in men: results of pattern recognition methods. *Journal of psychiatric research*, 26(3), pp.183–196.

Magidson, J. & Vermunt, J.K., 2003. Comparing Latent Class Factor Analysis with the Traditional Approach in Data Mining. In *Statistical Data Mining and Knowledge Discovery*. CRC Press, pp. 373–383.

Maier, W., Minges, J. & Lichtermann, D., 1995. The familial relationship between panic disorder and unipolar depression. *Journal of Psychiatric Research*, 29(5), pp.375–388.

Malhotra, D. et al., 2011. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron*, 72(6), pp.951–963.

Manji, H. et al., 2012. Impaired mitochondrial function in psychiatric disorders. *Nature Reviews Neuroscience*. Available at: http://www.nature.com/nrn/journal/v13/n5/box/nrn3229_BX1.html [Accessed August 30, 2013].

Manolio, T.A. et al., 2009. Finding the missing heritability of complex diseases. *Nature*, 461(7265), pp.747–753.

Marigorta, U.M. & Navarro, A., 2013. High trans-ethnic replicability of GWAS results implies common causal variants. *PLoS genetics*, 9(6), p.e1003566.

Marth, G.T. et al., 2011. The functional spectrum of low-frequency coding variation. *Genome biology*, 12(9), p.R84.

Martin, J. et al., 1993. Asking about child sexual abuse: methodological implications of a two stage survey. *Child abuse & neglect*, 17(3), pp.383–392.

Mathieson, I. & McVean, G., 2012. Differential confounding of rare and common variants in spatially structured populations. *Nature genetics*, 44(3), pp.243–246.

Mattson, M.P., Gleichmann, M. & Cheng, A., 2008. Mitochondria in neuroplasticity and neurological disorders. *Neuron*, 60(5), pp.748–766.

Matza, L.S. et al., 2003. Depression with atypical features in the National Comorbidity Survey: classification, description, and consequences. *Archives of general psychiatry*, 60(8), pp.817–826.

McCarthy, M.I. et al., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9(5), pp.356–369.

McCutcheon, A.L., 1987. *Latent Class Analysis*, SAGE.

McGuffin, P. et al., 1996. A hospital-based twin register of the heritability of DSM-IV unipolar depression. *Archives of general psychiatry*, 53(2), pp.129–136.

McGuffin, P. et al., 2005. Whole genome linkage scan of recurrent depressive disorder from the depression network study. *Human molecular genetics*, 14(22), pp.3337–3345.

McGuffin, P., Katz, R. & Bebbington, P., 1987. Hazard, heredity and depression. A family study. *Journal of psychiatric research*, 21(4), pp.365–375.

McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9), pp.1297–1303.

Melartin, T.K. et al., 2002. Current comorbidity of psychiatric disorders among DSM-IV major depressive disorder patients in psychiatric care in the Vantaa Depression Study. *The Journal of Clinical Psychiatry*, 63(2), pp.126–134.

Mendlewicz, J. & Baron, M., 1981. Morbidity risks in subtypes of unipolar depressive illness: differences between early and late onset forms. *The British journal of psychiatry: the journal of mental science*, 139, pp.463–466.

Merikangas, K.R. et al., 2003. Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich Cohort Study. *Archives of General Psychiatry*, 60(10), pp.993–1000.

Middeldorp, C.M. et al., 2008. Linkage on chromosome 14 in a genome-wide linkage study of a broad anxiety phenotype. *Molecular psychiatry*, 13(1), pp.84–89.

Middeldorp, C.M. et al., 2005. The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. *Psychological Medicine*, 35(5), pp.611–624.

Middeldorp, C.M. et al., 2011. The genetic association between personality and major depression or bipolar disorder. A polygenic score analysis using genome-wide association data. *Translational psychiatry*, 1, p.e50.

Moffitt, T.E. et al., 2007. Depression and generalized anxiety disorder: cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Archives of general psychiatry*, 64(6), pp.651–660.

De Moor, M.H.M. et al., 2012. Meta-analysis of genome-wide association studies for personality. *Molecular psychiatry*, 17(3), pp.337–349.

Muglia, P. et al., 2010. Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. *Molecular psychiatry*, 15(6), pp.589–601.

Mullen, P.E. et al., 1993. Childhood sexual abuse and mental health in adult life. *The British journal of psychiatry: the journal of mental science*, 163, pp.721–732.

Muthén, B.O., 1989. Dichotomous Factor Analysis of Symptom Data. *Sociological Methods & Research*, 18(1), pp.19–65.

Muthén, L.K. & Muthén, B.O., 1998. *Mplus User's Guide. Sixth Edition*, Los Angeles: CA: Muthén & Muthén.

Nejentsev, S. et al., 2009. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science (New York, N.Y.)*, 324(5925), pp.387–389.

Nelson, M.R. et al., 2012. An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science (New York, N.Y.)*, 337(6090), pp.100–104.

Neumann, D.A. et al., 1996. The Long-Term Sequelae of Childhood Sexual Abuse in Women: A Meta-Analytic Review. *Child Maltreatment*, 1(1), pp.6–16.

Nunes, A.P. et al., 2012. Quality of DNA extracted from saliva samples collected with the Oragene™ DNA self-collection kit. *BMC medical research methodology*, 12, p.65.

O'Roak, B.J. et al., 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nature genetics*, 43(6), pp.585–589.

O'Roak, B.J. et al., 2012. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science (New York, N.Y.)*, 338(6114), pp.1619–1622.

Paolucci, E.O., Genuis, M.L. & Violato, C., 2001. A meta-analysis of the published research on the effects of child sexual abuse. *The Journal of psychology*, 135(1), pp.17–36.

Parker, G., 2000. Classifying Depression: Should Paradigms Lost Be Regained? *American Journal of Psychiatry*, 157(8), pp.1195–1203.

Parker, G., 2007. Defining melancholia: the primacy of psychomotor disturbance. *Acta psychiatrica Scandinavica. Supplementum*, (433), pp.21–30.

Parker, G. et al., 1995. Low parental care as a risk factor to lifetime depression in a community sample. *Journal of Affective Disorders*, 33(3), pp.173–180.

Parker, G., 1979. Parental characteristics in relation to depressive disorders. *The British Journal of Psychiatry*, 134, pp.138–147.

Parker, G., 1983. *Parental overprotection: a risk factor in psychosocial development*, New York: Grune & Stratton.

Parker, G. et al., 1999. Subtyping depression: testing algorithms and identification of a tiered model. *The Journal of nervous and mental disease*, 187(10), pp.610–617.

Parker, G., 1989. The Parental Bonding Instrument: psychometric properties reviewed. *Psychiatric developments*, 7(4), pp.317–335.

Parker, G., Gladstone, G. & Chee, K.T., 2001. Depression in the planet's largest ethnic group: the Chinese. *The American Journal of Psychiatry*, 158(6), pp.857–864.

Parker, G., Tupling, H. & Brown, L.B., 1979. A parental bonding instrument. , 52, pp.1–10.

Pasaniuc, B. et al., 2012. Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. *Nature genetics*, 44(6), pp.631–635.

Patterson, N., Price, A.L. & Reich, D., 2006. Population Structure and Eigenanalysis. *PLoS Genet*, 2(12), p.e190.

Paulsen, J. et al., 2001. Neuropsychiatric aspects of Huntington's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 71(3), pp.310–314.

Paykel, E.S., 1978. Contribution of life events to causation of psychiatric illness. *Psychological medicine*, 8(2), pp.245–253.

Pe'er, I. et al., 2008. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic epidemiology*, 32(4), pp.381–385.

Penninx, B.W.J.H. et al., 2011. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Affective Disorders*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21496929 [Accessed June 16, 2011].

Pergadia, M.L. et al., 2011. A 3p26-3p25 genetic linkage finding for DSM-IV major depression in heavy smoking families. *The American journal of psychiatry*, 168(8), pp.848–852.

Perris, C. et al., 1986. Perceived depriving parental rearing and depression. *The British Journal of Psychiatry*, 148, pp.170–175.

Perris, C., Arrindell, W.A. & Eisemann, M., 1994. *Parenting and psychopathology*, J. Wiley.

Perroud, N. et al., 2010. History of suicide attempts among patients with depression in the GENDEP project. *Journal of Affective Disorders*, 123(1-3), pp.131–137.

Phillips, M.R. et al., 2007. Assessing depressive symptoms in persons who die of suicide in mainland China. *Journal of affective disorders*, 98(1-2), pp.73–82.

Phillips, M.R. et al., 2002. Risk factors for suicide in China: a national case-control psychological autopsy study. *Lancet*, 360(9347), pp.1728–1736.

Pini, S. et al., 1994. Phenomenology of panic disorder and major depression in a family study. *Journal of Affective Disorders*, 30(4), pp.257–272.

Pini, S. et al., 1997. Prevalence of anxiety disorders comorbidity in bipolar depression, unipolar depression and dysthymia. *Journal of Affective Disorders*, 42(2-3), pp.145–153.

Placidi, G.P. et al., 2000. Anxiety in major depression: relationship to suicide attempts. *The American Journal of Psychiatry*, 157(10), pp.1614–1618.

Price, A.L. et al., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), pp.904–909.

Qin, P. & Mortensen, P.B., 2001. Specific characteristics of suicide in China. *Acta psychiatrica Scandinavica*, 103(2), pp.117–121.

Quiroz, J.A. et al., 2008. Mitochondrially mediated plasticity in the pathophysiology and treatment of bipolar disorder. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 33(11), pp.2551–2565.

R Development Core Team, 2011. *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.r-project.org/ [Accessed July 29, 2011].

Reich, J. et al., 1993. Comorbidity of panic and major depressive disorder. *Journal of Psychiatric Research*, 27 Suppl 1, pp.23–33.

Reijnders, J.S.A.M. et al., 2008. A systematic review of prevalence studies of depression in Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society*, 23(2), pp.183–189; quiz 313.

Rietschel, M. et al., 2010. Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. *Biological Psychiatry*, 68(6), pp.578–585.

Riley, B.P. et al., 1998. Failure to exclude a possible schizophrenia susceptibility locus on chromosome 13q14.1-q32 in southern African Bantu-speaking families. *Psychiatric genetics*, 8(3), pp.155–162.

Ripke, S. et al., 2013. A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular psychiatry*, 18(4), pp.497–511.

Risch, N. & Merikangas, K., 1996. The future of genetic studies of complex human diseases. *Science (New York, N.Y.)*, 273(5281), pp.1516–1517.

Risch, N.J., 2000. Searching for genetic determinants in the new millennium. *Nature*, 405(6788), pp.847–856.

Rivas, M.A. et al., 2011. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nature genetics*, 43(11), pp.1066–1073.

Rollins, B. et al., 2009. Mitochondrial Variants in Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. *PLoS ONE*, 4(3), p.e4913.

Romera, I. et al., 2008. Factor analysis of the Zung self-rating depression scale in a large sample of patients with major depressive disorder in primary care. *BMC Psychiatry*, 8, p.4.

Rutter, M., 2000. Genetic Studies of Autism: From the 1970s into the Millennium. *Journal of Abnormal Child Psychology*, 28(1), pp.3–14.

Sabeti, P.C. et al., 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164), pp.913–918.

Sanderson, W.C., Beck, A.T. & Beck, J., 1990. Syndrome comorbidity in patients with major depression or dysthymia: prevalence and temporal relationships. *The American Journal of Psychiatry*, 147(8), pp.1025–1028.

Schmidt, H.D., Shelton, R.C. & Duman, R.S., 2011. Functional biomarkers of depression: diagnosis, treatment, and pathophysiology. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 36(12), pp.2375–2394.

Schotte, C.K. et al., 1997. Cluster analytic validation of the DSM melancholic depression. The threshold model: integration of quantitative and qualitative distinctions between unipolar depressive subtypes. *Psychiatry research*, 71(3), pp.181–195.

Schwab, S.G. et al., 1997. Evidence suggestive of a locus on chromosome 5q31 contributing to susceptibility for schizophrenia in German and Israeli families by multipoint affected sib-pair linkage analysis. *Molecular psychiatry*, 2(2), pp.156–160.

Schwarz, G., 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), pp.461–464.

Sebat, J. et al., 2007. Strong association of de novo copy number mutations with autism. *Science (New York, N.Y.)*, 316(5823), pp.445–449.

Sequeira, A. et al., 2012. Mitochondrial mutations and polymorphisms in psychiatric disorders. *Frontiers in genetics*, 3, p.103.

Shafer, A.B., 2006. Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology*, 62(1), pp.123–146.

Shi, J. et al., 2011. Genome-wide association study of recurrent early-onset major depressive disorder. *Molecular psychiatry*, 16(2), pp.193–201.

Shi, S. et al., 2009. Multi-center study of the clinical features in depression comorbidity with anxiety disorders. *Shanghai Archives of Psychiatry*, 21, pp.198–202.

Shi, Y. et al., 2012. Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41), pp.16510–16515.

Shyn, S.I. et al., 2011. Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. *Molecular Psychiatry*, 16(2), pp.202–215.

Spencer, C.C.A. et al., 2009. Designing Genome-Wide Association Studies: Sample Size, Power, Imputation, and the Choice of Genotyping Chip. *PLoS Genet*, 5(5), p.e1000477.

Spitzer, R.L. et al., 1978. *Research Diagnostic Criteria (RDC) for a Selected Group of Functional Disorders*, New York State Psychiatric Institute.

Spitzer, R.L., Williams, J.B. & Gibbon, J., 1987. *Structured Clinical Interview for DSM-III-R - Patient Version (SCID-P, 4/1/87)*, New York State Psychiatric Institute.

Stancer, H.C. et al., 1987. Evidence for homogeneity of major depression and bipolar affective disorder. *Journal of psychiatric research*, 21(1), pp.37–53.

Steiger, J., 1990. Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research*, 25(2), pp.173–180.

Stewart, J.W. et al., 1993. Atypical depression. A valid clinical entity? *The Psychiatric clinics of North America*, 16(3), pp.479–495.

Sullivan, P.F. et al., 2009. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Molecular psychiatry*, 14(4), pp.359–375.

Sullivan, P.F., Kessler, R.C. & Kendler, K.S., 1998. Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. *The American Journal of Psychiatry*, 155(10), pp.1398–1406.

Sullivan, P.F., Neale, M.C. & Kendler, K.S., 2000. Genetic epidemiology of major depression: review and meta-analysis. *The American Journal of Psychiatry*, 157(10), pp.1552–1562.

Sullivan, P.F., Prescott, C.A. & Kendler, K.S., 2002. The subtypes of major depression in a twin registry. *Journal of Affective Disorders*, 68(2-3), pp.273–284.

Surtees, P.G. et al., 1986. Life events and the onset of affective disorder. A longitudinal general population study. *Journal of affective disorders*, 10(1), pp.37–50.

Tennessen, J.A. et al., 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science (New York, N.Y.)*, 337(6090), pp.64–69.

Thompson, E.R., 2008. Development and Validation of an International English Big-Five Mini-Markers. *Personality and Individual Differences*, 45(6), pp.542–548.

Tucker, L. & Lewis, C., 1973. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), pp.1–10.

Turner, C. & Schapira, A.H.V., 2010. Mitochondrial matters of the brain: the role in Huntington's disease. *Journal of bioenergetics and biomembranes*, 42(3), pp.193–198.

Visscher, P.M. et al., 2012. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Molecular psychiatry*, 17(5), pp.474–485.

Visscher, P.M., Hill, W.G. & Wray, N.R., 2008. Heritability in the genomics era--concepts and misconceptions. *Nature reviews. Genetics*, 9(4), pp.255–266.

Vissers, L.E.L.M. et al., 2010. A de novo paradigm for mental retardation. *Nature genetics*, 42(12), pp.1109–1112.

Walsh, T. et al., 2008. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science (New York, N.Y.)*, 320(5875), pp.539–543.

Wang, K., Li, M. & Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), p.e164.

Wang, Y. et al., 2013. An integrative variant analysis pipeline for accurate genotype/haplotype inference in population NGS data. *Genome research*, 23(5), pp.833–842.

Weiss, E.L., Longhurst, J.G. & Mazure, C.M., 1999. Childhood sexual abuse as a risk factor for depression in women: psychosocial and neurobiological correlates. *The American journal of psychiatry*, 156(6), pp.816–828.

Weissman, M.M. et al., 1996. Cross-national epidemiology of major depression and bipolar disorder. *JAMA: the journal of the American Medical Association*, 276(4), pp.293–299.

Weissman, M.M. et al., 1993. Sex differences in rates of depression: cross-national perspectives. *Journal of Affective Disorders*, 29(2–3), pp.77–84.

Weissman, M.M. et al., 1993. The relationship between panic disorder and major depression. A new family study. *Archives of General Psychiatry*, 50(10), pp.767–780.

Weissman, M.M. et al., 1986. Understanding the clinical heterogeneity of major depression using family data. *Archives of General Psychiatry*, 43(5), pp.430–434.

Wetzel, R.D. et al., 1980. Personality as a subclinical expression of the affective disorders. *Comprehensive psychiatry*, 21(3), pp.197–205.

Winokur, G., 1997. All roads lead to depression: clinically homogeneous, etiologically heterogeneous. *Journal of Affective Disorders*, 45(1-2), pp.97–108.

Winokur, G., 1979. Unipolar depression: is it divisible into autonomous subtypes? *Archives of general psychiatry*, 36(1), pp.47–52.

Wray, N.R. et al., 2007. Genetic and phenotypic stability of measures of neuroticism over 22 years. *Twin research and human genetics: the official journal of the International Society for Twin Studies*, 10(5), pp.695–702.

Wray, N.R. et al., 2010. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Molecular Psychiatry*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21042317 [Accessed August 30, 2011].

Wu, Z. et al., 2013. Difference in remission in a Chinese population with anxious versus nonanxious treatment-resistant depression: A report of OPERATION study. *Journal of affective disorders*.

Xu, B. et al., 2011. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nature genetics*, 43(9), pp.864–868.

Xu, B. et al., 2008. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nature genetics*, 40(7), pp.880–885.

Yagi, T. & Takeichi, M., 2000. Cadherin superfamily genes: functions, genomic organization, and neurologic diversity. *Genes & development*, 14(10), pp.1169–1180.

Yang, J., Wray, N.R. & Visscher, P.M., 2010. Comparing apples and oranges: equating the power of case-control and quantitative trait association studies. *Genetic epidemiology*, 34(3), pp.254–257.

Zimmerman, M. & Chelminski, I., 2003. Clinician recognition of anxiety disorders in depressed outpatients. *Journal of Psychiatric Research*, 37(4), pp.325–333.

Zimmerman, M., McDermut, W. & Mattia, J.I., 2000. Frequency of anxiety disorders in psychiatric outpatients with major depressive disorder. *The American Journal of Psychiatry*, 157(8), pp.1337–1340.

Zubenko, G.S., Zubenko, W.N., et al., 2003. A collaborative study of the emergence and clinical features of the major depressive syndrome of Alzheimer's disease. *The American journal of psychiatry*, 160(5), pp.857–866.

Zubenko, G.S., Maher, B., et al., 2003. Genome-wide linkage survey for genetic loci that influence the development of depressive disorders in families with recurrent, early-onset, major depression. *American journal of medical genetics. Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*, 123B(1), pp.1–18.

Zuk, O. et al., 2012. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4), pp.1193–1198.

Appendix

SNP annotations

Phenotype	rsID	CHR	SNP	FreqA	BETA	P	logp	N.genesymbol
Disease status	rs7257142	19	4511830	0.54	-0.076	1.05E-09	8.98	CHAF1A,MIR4746,SEMA6B,LRG1,PLIN5,PLIN4,UBXN6,HDGFRP2
Early AAO	rs7257142	19	4511830	0.54	-0.0825	1.01E-09	9.00	CHAF1A,MIR4746,SEMA6B,LRG1,PLIN5,PLIN4,UBXN6,HDGFRP2
LCA severe	rs7257142	19	4511830	0.54	-0.0796	2.34E-09	8.63	CHAF1A,MIR4746,SEMA6B,LRG1,PLIN5,PLIN4,UBXN6,HDGFRP2
Low environmental risks	rs7257142	19	4511830	0.54	-0.1006	6.88E-14	13.16	CHAF1A,MIR4746,SEMA6B,LRG1,PLIN5,PLIN4,UBXN6,HDGFRP2
Melancholia	rs7257142	19	4511830	0.54	-0.0742	8.75E-09	8.06	CHAF1A,MIR4746,SEMA6B,LRG1,PLIN5,PLIN4,UBXN6,HDGFRP2
Melancholia	rs6628027	23	24655626	0.42	-0.0452	3.41E-08	7.47	LOC100288233,LOC101928444,PDK3,POLA1,PCYT1B
Melancholia	rs80309727	10	69617347	0.62	0.0378	5.63E-08	7.25	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,DNAJC12,RPL12P8
Melancholia	rs12415454	10	69623737	0.62	0.0377	5.90E-08	7.23	DNAJC19P1,RPL21P92,TRNAS21,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8

Melancholia	rs12415800	10	69624180	0.62	0.0377	5.98E-08	7.22	DNAJC19P1,RPL21P92,TRNAS21,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
Melancholia	rs79651647	10	69628475	0.63	0.0377	6.22E-08	7.21	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
Melancholia	rs12572973	10	69628936	0.63	0.0378	6.57E-08	7.18	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
Melancholia	rs76164521	10	69629036	0.63	0.0376	6.68E-08	7.18	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
Melancholia	rs75255354	10	69633348	0.63	0.0373	8.76E-08	7.06	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
Melancholia	rs75818614	10	69609323	0.62	0.0371	1.00E-07	7.00	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs79651647	10	69628475	0.63	0.0411	7.90E-09	8.10	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs12572973	10	69628936	0.63	0.0412	8.11E-09	8.09	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs80309727	10	69617347	0.62	0.041	8.20E-09	8.09	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs76164521	10	69629036	0.63	0.041	8.36E-09	8.08	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8

High recurrence	rs33992752	10	69659402	0.63	0.0409	8.89E-09	8.05	DNAJC19P1,RPS3AP38,SIRT1,HERC4, DNAJC12,RPL12P8
High recurrence	rs75255354	10	69633348	0.63	0.041	9.13E-09	8.04	DNAJC19P1,RPL21P92,RPS3AP38,SIRT 1,HERC4,DNAJC12,RPL12P8
High recurrence	rs79269275	10	69647397	0.63	0.0409	9.50E-09	8.02	DNAJC19P1,RPL21P92,RPS3AP38,SIRT 1,HERC4,DNAJC12,RPL12P8
High recurrence	rs34294800	10	69649904	0.62	0.0409	9.82E-09	8.01	DNAJC19P1,RPL21P92,RPS3AP38,SIRT 1,HERC4,DNAJC12,RPL12P8
High recurrence	rs35480510	10	69649902	0.63	0.0408	9.90E-09	8.00	DNAJC19P1,RPL21P92,RPS3AP38,SIRT 1,HERC4,DNAJC12,RPL12P8
High recurrence	rs12570807	10	69652003	0.63	0.0408	9.90E-09	8.00	DNAJC19P1,RPL21P92,RPS3AP38,SIRT 1,HERC4,DNAJC12,RPL12P8
High recurrence	rs12571880	10	69645280	0.63	0.0408	1.01E-08	7.99	DNAJC19P1,RPL21P92,RPS3AP38,SIRT 1,HERC4,DNAJC12,RPL12P8
High recurrence	rs12415454	10	69623737	0.62	0.0407	1.11E-08	7.96	DNAJC19P1,RPL21P92,TRNAS21,RPS3 AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs12415800	10	69624180	0.62	0.0407	1.12E-08	7.95	DNAJC19P1,RPL21P92,TRNAS21,RPS3 AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs35416150	10	69665189	0.63	0.0406	1.18E-08	7.93	DNAJC19P1,RPS3AP38,SIRT1,HERC4, DNAJC12,RPL12P8

High recurrence	rs12412886	10	69632695	0.63	0.0406	1.21E-08	7.92	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs145356089	10	69631362	0.63	0.0406	1.22E-08	7.91	DNAJC19P1,RPL21P92,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs16924945	10	69674347	0.63	0.0406	1.24E-08	7.91	POU5F1P5,DNAJC19P1,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs74790878	10	69675935	0.63	0.0404	1.37E-08	7.86	POU5F1P5,DNAJC19P1,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs4746720	10	69676830	0.63	0.0404	1.47E-08	7.83	POU5F1P5,DNAJC19P1,RPS3AP38,SIRT1,HERC4,DNAJC12,RPL12P8
High recurrence	rs76931017	10	69562826	0.63	0.0401	2.51E-08	7.60	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs75818614	10	69609323	0.62	0.0397	2.58E-08	7.59	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs2273772	10	69597824	0.62	0.0396	2.64E-08	7.58	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs12416174	10	69595187	0.63	0.0396	2.82E-08	7.55	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8

High recurrence	rs74719503	10	69564969	0.62	0.0397	2.86E-08	7.54	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs79651769	10	69602616	0.62	0.0395	2.86E-08	7.54	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs78784886	10	69583513	0.63	0.0399	2.90E-08	7.54	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs77546299	10	69603288	0.62	0.0395	2.91E-08	7.54	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
High recurrence	rs74642861	10	69586239	0.64	0.0402	3.72E-08	7.43	DNAJC19P1,RPL21P92,TRNAS21,SIRT1,HERC4,AKR1B10P1,DNAJC12,RPL12P8
i9.f1	rs12609923	19	30063144	0.20	0.0754	7.06E-08	7.15	POP4,LOC284395,VSTM2B,PLEKHF1
BMI	rs182438096	6	34754144	0.91	2.7289	3.24E-08	7.49	LOC101929243,UHRF1BP1,RPS10P13,C6orf106,SNRPC,TAF11
Height	rs10860836	12	102447060	0.70	-0.4476	6.67E-09	8.18	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs10860837	12	102447422	0.70	-0.4476	6.67E-09	8.18	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs10860835	12	102442631	0.70	-0.4455	7.74E-09	8.11	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs11111149	12	102466079	0.70	-0.4441	8.08E-09	8.09	CCDC53,PARPBP,NUP37

Height	rs11111132	12	102409283	0.70	-0.4402	1.10E-08	7.96	LOC100129880,RPL9P24,CCDC53,DRAM1,NUP37
Height	rs17041470	12	102430430	0.70	-0.4342	1.68E-08	7.77	LOC100129880,RPL9P24,CCDC53,PARPBP,NUP37
Height	rs10860834	12	102440836	0.70	-0.4326	1.93E-08	7.72	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs2373569	12	102463242	0.70	-0.4327	2.12E-08	7.67	CCDC53,PARPBP,NUP37
Height	rs11111147	12	102462468	0.70	-0.4305	2.19E-08	7.66	CCDC53,PARPBP,NUP37
Height	rs11111127	12	102405082	0.70	-0.43	2.29E-08	7.64	LOC100129880,RPL9P24,CCDC53,DRAM1,NUP37
Height	rs11111121	12	102397730	0.69	-0.4275	3.45E-08	7.46	LOC100129880,RPL9P24,CCDC53,DRAM1,NUP37
Height	rs10860839	12	102460991	0.69	-0.4349	3.75E-08	7.43	CCDC53,PARPBP,NUP37
Height	rs11111136	12	102420265	0.72	-0.4445	4.76E-08	7.32	LOC100129880,RPL9P24,CCDC53,PARPBP,NUP37
Height	rs7133285	12	56699429	0.79	-0.4691	4.85E-08	7.31	LOC100128676,RNU7-40P,LOC100419033,RNF41,CNPY2,CS,ANKRD52,SLC39A5,APOF,IL23A,STAT2,NABP2,COQ10A,PAN2
Height	rs3816804	12	56680745	0.79	-0.468	5.04E-08	7.30	LOC100128676,RNU7-40P,TRNAS11,LOC100419033,RNF41,CNPY2,CS,ANKRD52,SLC39A5,APOF,IL23A,SMARCC2,STAT2,NABP2,COQ10A,PAN2

Height	rs3809128	12	56709919	0.79	-0.4685	5.70E-08	7.24	LOC100128676,RNU7-40P,LOC100419033,RNF41,CNPY2,CS,ANKRD52,SLC39A5,APOF,IL23A,STAT2,NABP2,COQ10A,PAN2
Height	rs146426492	12	56723494	0.80	-0.4811	6.16E-08	7.21	LOC100128676,RNU7-40P,LOC100419033,CNPY2,CS,ANKRD52,SLC39A5,APOF,IL23A,STAT2,NABP2,TIMELESS,COQ10A,PAN2
Height	rs79077890	12	102457301	0.71	-0.4199	6.90E-08	7.16	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs75423089	12	102456819	0.71	-0.4199	6.93E-08	7.16	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs17032226	12	102415211	0.72	-0.4324	6.99E-08	7.16	LOC100129880,RPL9P24,CCDC53,PARPBP,DRAM1,NUP37
Height	rs11111142	12	102431759	0.71	-0.431	7.25E-08	7.14	LOC100129880,CCDC53,PARPBP,NUP37
Height	rs78080209	12	52247711	0.83	0.5102	7.43E-08	7.13	ANKRD33,FIGNL2,SCN8A,LOC728503,ACVR1B,ACVRL1
Height	rs117596976	12	52251496	0.83	0.5094	7.90E-08	7.10	ANKRD33,FIGNL2,SCN8A,LOC728503,ACVR1B,ACVRL1
Height	rs78653647	12	52253105	0.83	0.5096	7.93E-08	7.10	ANKRD33,FIGNL2,SCN8A,LOC728503,ACVR1B,ACVRL1
Height	rs118091512	12	52239411	0.84	0.5107	7.93E-08	7.10	ANKRD33,FIGNL2,SCN8A,LOC728503,ACVRL1
Height	rs141157575	12	52252572	0.83	0.5096	7.94E-08	7.10	ANKRD33,FIGNL2,SCN8A,LOC728503,ACVR1B,ACVRL1

Height	rs138195537	12	52251642	0.83	0.5091	8.03E-08	7.10	ANKRD33,FIGNL2,SCN8A,LOC728503, ACVR1B,ACVRL1
Height	rs12424451	12	102566908	0.71	-0.4322	8.23E-08	7.08	LOC100418882,PMCH,PARPBP,NUP37
Height	rs12427090	12	102566906	0.71	-0.4322	8.23E-08	7.08	LOC100418882,PMCH,PARPBP,NUP37
Height	rs79835345	12	102566915	0.71	-0.4322	8.23E-08	7.08	LOC100418882,PMCH,PARPBP,NUP37
Height	rs17032237	12	102421888	0.71	-0.4172	8.38E-08	7.08	LOC100129880,RPL9P24,CCDC53,PAR PBP,NUP37
Height	rs1962472	12	102463961	0.71	-0.4155	8.93E-08	7.05	CCDC53,PARPBP,NUP37
Height	rs76734806	12	102465188	0.71	-0.4155	8.94E-08	7.05	CCDC53,PARPBP,NUP37
Height	rs78471714	12	102470434	0.71	-0.4151	9.15E-08	7.04	CCDC53,PARPBP,NUP37
Height	rs139377474	12	102471860	0.71	-0.4151	9.15E-08	7.04	CCDC53,PARPBP,NUP37
Height	rs66500550	12	102422359	0.71	-0.4151	9.72E-08	7.01	LOC100129880,RPL9P24,CCDC53,PAR PBP,NUP37
MT	rs1937	10	60145342	0.83	2.7889	5.67E-18	17.25	CISD1,TFAM,UBE2D1
MT	rs11006132	10	60156166	0.83	2.7508	9.18E-18	17.04	TFAM,UBE2D1
MT	rs1049432	10	60155120	0.83	2.7495	9.39E-18	17.03	TFAM,UBE2D1
MT	rs2279339	10	60144998	0.83	2.7544	1.32E-17	16.88	CISD1,TFAM,UBE2D1
MT	rs11006126	10	60142880	0.83	2.7767	1.34E-17	16.87	CISD1,TFAM,UBE2D1
MT	rs2279340	10	60144884	0.83	2.7533	1.34E-17	16.87	CISD1,TFAM,UBE2D1
MT	rs9971104	10	60142800	0.83	2.7735	1.40E-17	16.85	CISD1,TFAM,UBE2D1
MT	rs12247015	10	60145079	0.82	2.7491	1.51E-17	16.82	CISD1,TFAM,UBE2D1
MT	rs11006125	10	60142402	0.83	2.7649	1.73E-17	16.76	CISD1,TFAM,UBE2D1
MT	rs9971282	10	60142116	0.83	2.8063	4.66E-17	16.33	CISD1,TFAM,UBE2D1
MT	rs61190999	10	60175440	0.82	2.579	2.57E-15	14.59	TFAM,UBE2D1,BICC1
MT	rs76455223	10	60193608	0.82	2.5444	1.42E-14	13.85	TFAM,UBE2D1,BICC1
MT	rs150887692	10	60203775	0.84	2.6282	6.24E-14	13.21	TFAM,UBE2D1,BICC1

MT	rs182839028	4	48661495	0.90	6.3087	2.50E-13	12.60	RNU5E-3P,FRYL
MT	rs145351516	8	17321297	0.28	-4.2093	1.04E-10	9.98	ADAM24P,SLC7A2,MTMR7
MT	rs1427215	10	60229347	0.86	2.1389	1.72E-10	9.76	TFAM,UBE2D1,BICC1
MT	rs61875518	10	60228964	0.86	2.1385	1.96E-10	9.71	TFAM,UBE2D1,BICC1
MT	rs17644676	10	60240354	0.85	2.092	2.27E-10	9.64	TFAM,BICC1
MT	rs35258735	10	60241036	0.85	2.092	2.27E-10	9.64	TFAM,BICC1
MT	rs61875523	10	60244111	0.85	2.0918	2.29E-10	9.64	TFAM,BICC1
MT	rs61875520	10	60230671	0.85	2.0914	2.39E-10	9.62	TFAM,BICC1
MT	rs7093118	10	60231657	0.85	2.0914	2.39E-10	9.62	TFAM,BICC1
MT	rs2282989	7	92326346	0.68	1.5548	1.85E-09	8.73	RNU6-10P,CDK6
MT	rs1004051	7	92328956	0.68	1.5524	1.95E-09	8.71	RNU6-10P,CDK6
MT	rs3731318	7	92329126	0.68	1.5524	1.95E-09	8.71	RNU6-10P,CDK6
MT	rs3802079	7	92330480	0.68	1.5524	1.95E-09	8.71	RNU6-10P,CDK6
MT	rs77192138	10	60245124	0.86	2.0151	1.97E-09	8.71	TFAM,BICC1
MT	rs1004052	7	92329957	0.68	1.55	2.07E-09	8.68	RNU6-10P,CDK6
MT	rs11981374	7	92315774	0.68	1.5467	2.33E-09	8.63	RNU6-10P,CDK6,FAM133B
MT	rs2282987	7	92312803	0.68	1.5455	2.49E-09	8.60	RNU6-10P,CDK6,FAM133B
MT	rs10225660	7	92324355	0.68	1.5391	2.61E-09	8.58	RNU6-10P,CDK6
MT	rs17164721	7	92324419	0.68	1.5368	2.77E-09	8.56	RNU6-10P,CDK6
MT	rs11981340	7	92315660	0.68	1.5261	3.32E-09	8.48	RNU6-10P,CDK6,FAM133B
MT	rs13437843	7	92316244	0.68	1.5246	3.44E-09	8.46	RNU6-10P,CDK6,FAM133B
MT	rs60726864	7	92316282	0.68	1.5246	3.44E-09	8.46	RNU6-10P,CDK6,FAM133B
MT	rs11533993	7	92319472	0.68	1.5238	3.48E-09	8.46	RNU6-10P,CDK6,FAM133B

MT	rs2282988	7	92317374	0.68	1.5197	3.83E-09	8.42	RNU6-10P,CDK6,FAM133B
MT	rs6964803	7	92315330	0.68	1.5221	3.86E-09	8.41	RNU6-10P,CDK6,FAM133B
MT	rs75882441	7	92305445	0.68	1.5101	5.40E-09	8.27	RNU6-10P,CDK6,FAM133B
MT	rs112660736	10	60095105	0.87	2.0523	1.12E-08	7.95	IPMK,CISD1,TFAM,UBE2D1
MT	rs16912153	10	60095266	0.87	2.0522	1.12E-08	7.95	IPMK,CISD1,TFAM,UBE2D1
MT	rs2106135	7	92299964	0.69	1.4677	1.41E-08	7.85	RNU6-10P,CDK6,FAM133B
MT	rs2282986	7	92299545	0.69	1.4674	1.42E-08	7.85	RNU6-10P,CDK6,FAM133B
MT	rs56353205	7	92296668	0.69	1.4598	1.84E-08	7.73	RNU6-10P,CDK6,FAM133B
MT	rs2301557	7	92300863	0.69	1.4588	1.86E-08	7.73	RNU6-10P,CDK6,FAM133B
MT	rs77599079	10	59929700	0.84	1.8876	2.17E-08	7.66	TPT1P10,IPMK,CISD1
MT	rs445	7	92408370	0.67	1.4908	2.17E-08	7.66	RNU6-10P,AC002454.1,CDK6
MT	rs12255735	10	60116030	0.84	1.865	2.21E-08	7.66	IPMK,CISD1,TFAM,UBE2D1
MT	rs9958410	18	58122704	0.32	-2.5268	2.25E-08	7.65	LOC100421385,MRPS5P4,MC4R
MT	rs41459146	7	92290278	0.69	1.446	2.51E-08	7.60	RNU6-10P,CDK6,FAM133B
MT	rs61873953	10	60092695	0.84	1.8507	2.81E-08	7.55	IPMK,CISD1,TFAM,UBE2D1
MT	rs57851212	10	59927787	0.84	1.8638	3.28E-08	7.48	TPT1P10,IPMK
MT	rs7072206	10	59925048	0.84	1.8577	3.76E-08	7.43	TPT1P10,IPMK

MT	rs11006120	10	60121976	0.84	1.8055	5.24E-08	7.28	IPMK,CISD1,TFAM,UBE2D1
MT	rs4948291	10	60100089	0.84	1.7957	6.12E-08	7.21	IPMK,CISD1,TFAM,UBE2D1
MT	rs61873887	10	60023959	0.85	1.7977	7.47E-08	7.13	TPT1P10,IPMK,CISD1,UBE2D1
MT	rs11006087	10	59990325	0.85	1.7939	7.95E-08	7.10	TPT1P10,IPMK,CISD1
MT	rs12572236	10	59987662	0.85	1.7939	7.95E-08	7.10	TPT1P10,IPMK,CISD1
MT	rs7900114	10	60001675	0.85	1.7926	8.04E-08	7.09	TPT1P10,IPMK,CISD1,UBE2D1
MT	rs12572520	10	60031460	0.85	1.7888	8.57E-08	7.07	TPT1P10,IPMK,CISD1,UBE2D1
MT	rs2275442	10	60035152	0.85	1.7888	8.57E-08	7.07	TPT1P10,IPMK,CISD1,UBE2D1
MT	rs2790176	10	60032636	0.85	1.7888	8.57E-08	7.07	TPT1P10,IPMK,CISD1,UBE2D1
MT	rs147826719	7	92266198	0.70	1.4096	8.57E-08	7.07	RNU6-10P,CDK6,FAM133B,RBM48,MGC16142
MT	rs12571364	10	59914787	0.84	1.8161	8.62E-08	7.06	TPT1P10,IPMK
MT	rs139600710	10	60004329	0.85	1.7914	8.89E-08	7.05	TPT1P10,IPMK,CISD1,UBE2D1
MT	rs12570088	10	59938336	0.85	1.7904	9.39E-08	7.03	TPT1P10,IPMK,CISD1
MT_case	rs1937	10	60145342	0.83	3.0305	1.47E-10	9.83	CISD1,TFAM,UBE2D1
MT_case	rs11006132	10	60156166	0.83	3.0051	1.59E-10	9.80	TFAM,UBE2D1
MT_case	rs1049432	10	60155120	0.83	3.0035	1.60E-10	9.80	TFAM,UBE2D1
MT_case	rs12247015	10	60145079	0.82	2.9854	2.78E-10	9.56	CISD1,TFAM,UBE2D1
MT_case	rs2279340	10	60144884	0.83	2.9732	3.29E-10	9.48	CISD1,TFAM,UBE2D1
MT_case	rs2279339	10	60144998	0.83	2.9691	3.49E-10	9.46	CISD1,TFAM,UBE2D1

MT_case	rs9971104	10	60142800	0.83	2.9679	4.74E-10	9.32	CISD1,TFAM,UBE2D1
MT_case	rs11006126	10	60142880	0.83	2.9677	4.84E-10	9.32	CISD1,TFAM,UBE2D1
MT_case	rs11006125	10	60142402	0.83	2.9649	4.98E-10	9.30	CISD1,TFAM,UBE2D1
MT_case	rs9971282	10	60142116	0.83	2.9791	1.27E-09	8.90	CISD1,TFAM,UBE2D1
MT_case	rs61190999	10	60175440	0.82	2.7581	7.68E-09	8.11	TFAM,UBE2D1,BICC1
MT_case	rs76455223	10	60193608	0.82	2.6844	3.22E-08	7.49	TFAM,UBE2D1,BICC1
MT_case	rs150887692	10	60203775	0.84	2.7525	8.35E-08	7.08	TFAM,UBE2D1,BICC1
MT_control	rs11006126	10	60142880	0.83	2.5921	2.31E-09	8.64	CISD1,TFAM,UBE2D1
MT_control	rs9971104	10	60142800	0.83	2.589	2.36E-09	8.63	CISD1,TFAM,UBE2D1
MT_control	rs2279339	10	60144998	0.83	2.563	2.53E-09	8.60	CISD1,TFAM,UBE2D1
MT_control	rs2279340	10	60144884	0.83	2.5595	2.61E-09	8.58	CISD1,TFAM,UBE2D1
MT_control	rs11006125	10	60142402	0.83	2.5779	2.69E-09	8.57	CISD1,TFAM,UBE2D1
MT_control	rs1937	10	60145342	0.83	2.5658	2.69E-09	8.57	CISD1,TFAM,UBE2D1
MT_control	rs9971282	10	60142116	0.83	2.6402	3.13E-09	8.50	CISD1,TFAM,UBE2D1
MT_control	rs11006132	10	60156166	0.83	2.5277	3.45E-09	8.46	TFAM,UBE2D1
MT_control	rs1049432	10	60155120	0.83	2.5261	3.53E-09	8.45	TFAM,UBE2D1
MT_control	rs12247015	10	60145079	0.82	2.5375	3.60E-09	8.44	CISD1,TFAM,UBE2D1
MT_control	rs182839028	4	48661495	0.90	7.1986	4.20E-09	8.38	RNU5E-3P,FRYL
MT_control	rs61190999	10	60175440	0.82	2.4395	2.13E-08	7.67	TFAM,UBE2D1,BICC1
MT_control	rs76455223	10	60193608	0.82	2.4675	2.17E-08	7.66	TFAM,UBE2D1,BICC1
MT_control	rs150887692	10	60203775	0.84	2.5687	3.93E-08	7.41	TFAM,UBE2D1,BICC1

Appendix I: Annotation of SNPs that exceeded a –log10 P of 7 in GWA analysis. AAO, age at onset; LCA, latent class analysis; MT, mitochondria /chr20 read depth.