Heritability Estimation and Risk Prediction in Schizophrenia

Choi Shing Wan

A thesis submitted in partial fulfillment of the
requirements for
the Degree of Doctor of Philosophy

Department of Psychiatry
University of Hong Kong
Hong Kong

October 26, 2015

Declaration

I declare that this thesis represents my own work, except where due acknowledgments is made, and that it has not been previously included in a thesis, dissertation or report submitted to this University or to any other institution for a degree, diploma or other qualification.

Signed.....

Acknowledgements

Abbreviations

CATIE Clinical Antipsychotic Trials of Intervention Effectiveness. 25, 29

CI confidence interval. 11, 27

CNS central nervous system. 20

CNV copy number variation. 14, 15, 17, 19

DSM Diagnostic and Statistical Manual of Mental Disorders. 2

DZ dizygotic. 11

EPS extrapyramidal motor symptoms. 24–26

FGA First Generation Antipsychotic. 23–25, 28

GC Genomic Control. 18

GCTA Genome-wide Complex Trait Analysis. 17, 18

GRM Genetic Relationship Matrix. 17, 18

GWAS Genome Wide Association Study. 14, 17, 18, 20, 27, 29

IL-6 Interleukin-6. 3

IQ intelligence quotient. 6, 7

LD Linkage Disequilibrium. 14, 18–20, 28

LDSC LD SCore. 18, 20, 22

LPS lipopolysaccharide. 3

maf Minor Allele Frequency. 6, 14

MHC major histocompatibility complex. 15

MIA maternal immune activation. 3–5, 15

MZ monozygotic. 11, 12, 26

PANSS Positive and Negative Symptom Scale. 26

PET positron emission tomography. 25

PGC Psychiatric Genomics Consortium. 15, 17

PolyI:C polyriboinosinic-polyribocytidilic acid. 3, 4

SCZ schizophrenia. 5

SE standard error. 19

SGA Second Generation Antipsychotic. 24–26

SNP Single Nucleotide Polymorphism. 13–15, 17–20, 27, 28

TD Tardive dyskinesia. 23, 25, 26

WHO World Health Organization. 1, 2

YLD years lost due to disability. 1, 2

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Schizophrenia

Schizophrenia is a detrimental psychiatric disorder, affecting around $0.3 \sim 0.7\%$ of the population(American Psychiatric Association, 2013). It is characterized by positive symptoms including delusions, hallucinations, disorganized speech and grossly disorganized behavior, and negative symptoms such as the diminished emotional expression(American Psychiatric Association, 2013) with a typical age of onset at late adolescent or late 20s in male and late 20s or early 30s in female(Schultz, North, and C. G. Shields, 2007).

Schizophrenia not only impose long lasting health, social and financial burden not only to the patients, but also to their families(Knapp, Mangalore, and Simon, 2004). Even more so, patients with schizophrenia increased suicide rate (Saha, Chant, and Mcgrath, 2007), leading to a higher mortality. Based on the World Health Organization (WHO) report, schizophrenia is one of the top 20 leading cause of years lost due to disability (YLD) in 2012, ranking 16 among all possible causes (table 1.1), demonstrating the extent of impact from schizophrenia to patients.

Due to the severity of schizophrenia, it has drawn much attention from the research community aiming to delineate the disease mechanics and be able to identify the risk factors. Arguably, the most important first step to any schizophrenia study is to have a robust and reliable disease diagnosis.

Table 1.1: Top 20 leading cause of YLD calculated by WHO in year 2012. Schizophrenia was considered as one of the top 20 leading cause of YLD(World Health Organization, 2013)

Rank	Cause	YLD (000s)	% YLD	YLD per 100k population
0	All Causes	740,545	100	10466
1	Unipolar depressive disorders	76,419	10.3	1080
2	Back and neck pain	53,855	7.3	761
3	Iron-deficiency anaemia	43,615	5.9	616
4	Chronic obstructive pulmonary disease	30,749	4.2	435
5	Alcohol use disorders	27,905	3.8	394
6	Anxiety disorders	27,549	3.7	389
7	Diabetes mellitus	22,492	3	318
8	Other hearing loss	22,076	3	312
9	Falls	20,409	2.8	288
10	Migraine	18,538	2.5	262
11	Osteoarthritis	18,096	2.4	256
12	Skin diseases	15,744	2.1	223
13	Asthma	14,134	1.9	200
14	Road injury	13,902	1.9	196
15	Refractive errors	13,498	1.8	191
16	Schizophrenia	13,408	1.8	189
17	Bipolar disorder	13,271	1.8	188
18	Drug use disorders	10,620	1.4	150
19	Endocrine, blood, immune disorders	10,495	1.4	148
20	Gynecological diseases	10,227	1.4	145

1.2 Diagnosis

Schizophrenia was first named "Dementia Praecox" by Dr. Emil Kraepelin and was later renamed as schizophrenia by Dr. Eugen Bleuler(Jablensky, 2010). Early nosological entity for schizophrenia such as that in Diagnostic and Statistical Manual of Mental Disorders (DSM)-I and DSM-II were vague and unreliable where the inter-rater agreement can be as low as 54%.(Tsuang, Stone, and Faraone, 2000; Harvey et al., 2012)

Later nosologies addressed these problem by introducing structural assessment and clear defined criteria. With these improvements, the inter-rater agreement of DSM-III raised to $\sim 90\%$ (Harvey et al., 2012), suggesting the diagnosis were much more reliable.

Currently DSM is at its 5th edition(American Psychiatric Association, 2013). A patient will be diagnosed with schizophrenia(F20.9) if they suffered from 2 or more of the

following symptoms for a significant portion of time during a 1-month period: 1) delusion; 2) hallucinations; 3) disorganized speech; 4) grossly disorganized or catatonic behaviour; and 5) negative symptoms such as diminished emotional expression, where one of the symptom must be either (1), (2) or (3). Signs of disturbance also need to persist for at least 6-month before the patient can be diagnosed with schizophrenia.

1.3 Risk Factors of Schizophrenia

Considerable effort has been made trying to identify possible risk factors of schizophrenia. It was first observed that there was an increased risk of schizophrenia in individual who were fetuses during the 1957 influenza epidemic(Mednick, 1958). Subsequently, other infectious agents such as HSV-2 and *T.gondii* were also found to increase the risk of schizophrenia if an individual's mother were infected during pregnancy. As different infectious agents all increase the risk of schizophrenia, it leads to the hypothesis of maternal immune activation (MIA) (Brown and Derkits, 2010). It was hypothesized that instead of a particular infectious agents, it was the maternal immune response that disrupt the brain development in the offspring, thus leading to an elevated risk of schizophrenia.

By utilizing the rodent models, it was found when the pregnant rodent was injected with the viral mimic polyriboinosinic-polyribocytidilic acid (PolyI:C) or the bacterial lipopolysaccharide (LPS), the offspring will display neuropathological features similar to those observed in schizophrenia(Urs Meyer, Joram Feldon, and Fatemi, 2009). It was further demonstrated that similar findings can be obtained through the injection of only the Interleukin-6 (IL-6)(Smith et al., 2007), suggesting that it was not the infection, but the maternal immune response that might have disrupted the fetal brain development.

Further studies suggested that MIA might leads to a complex pattern of age-dependent structural abnormalities in the mesoaccumbal and nigrostriatal dopamine systems(Vuillermot et al., 2010). Specifically, MIA induces an early abnormality in specific dopaminergic systems such as those in the striatum and midbrian region(Vuillermot et al., 2010). Based on these observations, U Meyer, Yee, and J Feldon (2007) hypothesize that inflammation in the fetal brain during early gestation not only can disrupt neurodavelopmental processes such as cell proliferation and differentiation, it also predispose the developing nervous system to additional failures in subsequent cell migration, target selection, and synapse maturation(fig. 1.1)(U Meyer, Yee, and J Feldon, 2007).

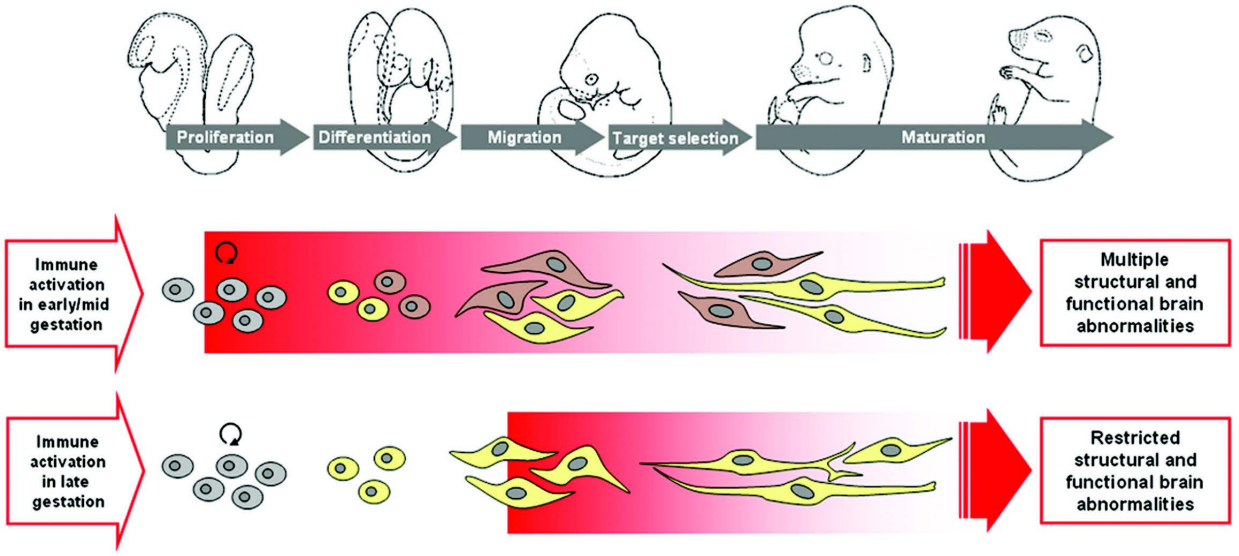


Figure 1.1: Hypothesized model of the impact of prenatal immune challenge on fetal brain development. Maternal infection in early/mid pregnancy may affect early neurodevelopmental events in the fetal brain, thereby influencing the differentiation of neural precursor cells(grey) into particular neuronal phenotype(yellow or brown). This may predispose the developing fetal nervous system to additional failures leading to multiple structural and functional brain abnormalities in later life. Figure used with permission from Journal(U Meyer, Yee, and J Feldon, 2007)

In a separate study by Giovanoli et al. (2013), mice were exposed to low dosage of PolyI:C during early gestation. Offspring born were then left undisturbed or exposed to unpredictable stress during peripubertal development. It was observed that offspring exposed to PolyI:C has an increased level of dopamine in the nucleus accumbens independent to whether if they were exposed to postnatal stress. Whereas serotonin (5-HT) were decreased in the medial prefrontal cortex when exposed to postnatal stress regardless of prenatal exposure. Only when the offspring were exposed to both PolyI:C and postnatal stress will they have an increased dopamine levels in the hippocampus or will sensorimotor gating and psychotomimetic drug sensitivity be affected(Giovanoli et al., 2013). Giovanoli et al. (2013) therefore suggest that the prenatal insult serves as a "disease primer" that increase offspring's vulnerability to subsequent insults.

Another interesting observation in Giovanoli et al. (2013)'s study was that the combined immune activation and stress led to a 2.5 to 3 fold increase in hippocampal and prefrontal expression of markers characteristic of activated microglia. Considering that microglia is responsible for synaptic pruning during brain development(Paolicelli et al., 2011), perturbation of microglia in the fetal brain might also mediate schizophrenia. Indeed, Onore et al. (2014) demonstrated that MIA has a prolonged effect on macrophages, leading to a

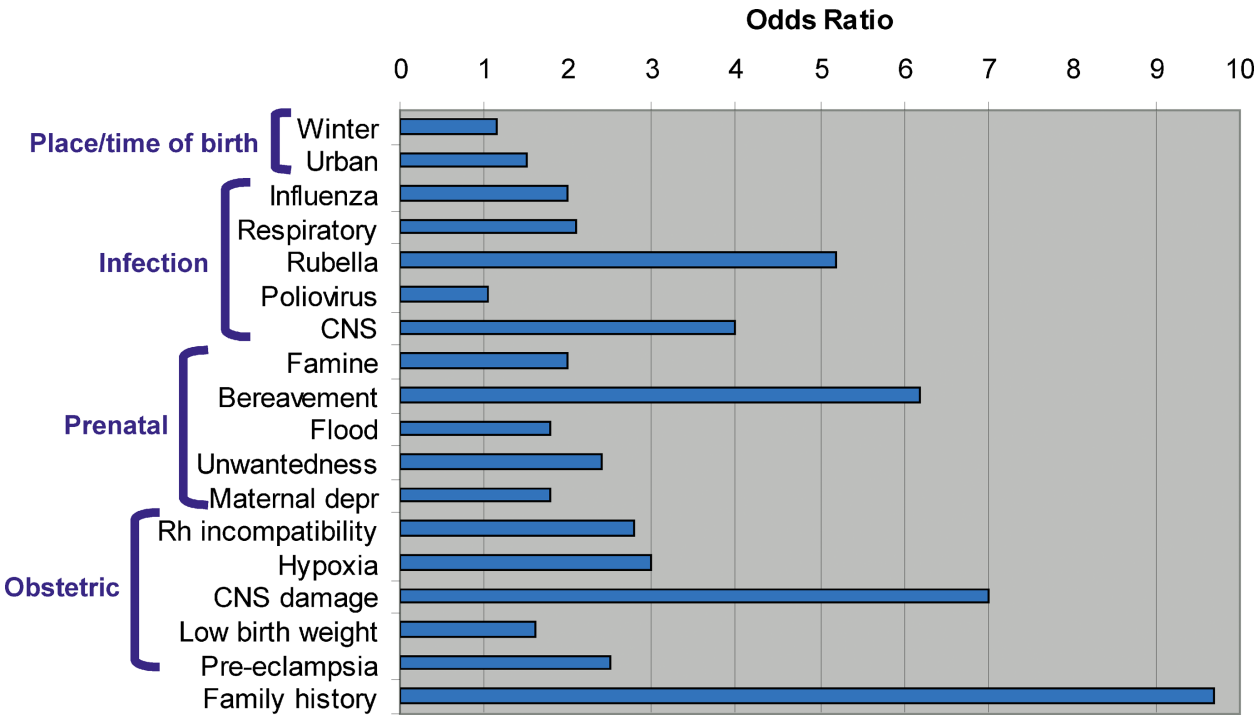


Figure 1.2: Risk factors of schizophrenia. It was observed that family history of schizophrenia was the largest risk factors. Risk of schizophrenia can be more than 9 times higher than the general population for individual with a family history of schizophrenia

shift towards to proinflammatory M1 phenotype. Immune dysfunction in the pathophysiology of schizophrenia has long been speculated(Müller and Schwarz, 2010) and evidence shown that there was a strong influence of the pro- and anti-inflammation cytokines on the glutamatergic neurotransmission(Müller and Schwarz, 2010), suggesting the immune system might played an important role in disease etiology of schizophrenia (SCZ).

Together, these results supports the involvement of MIA in the development of schizophrenia. It was even estimated that one third of all schizophrenia cases could have been prevented shall all infection were prevented from the entire pregnant population(Brown and Derkits, 2010).

Similarly, tobacco consumption (Kelly and McCreadie, 1999), socio economic status and even the area of birth (e.g. urban vs suburb) were also found to be associated with increased risk of schizophrenia(McGrath et al., 2008). However, by and large, the single largest risk factor was family history of schizophrenia(fig. 1.2)(Sullivan, 2005). Studies conducted by Ernst Rüdin, Franz J. Kallmann and Hans Luxenburger, all demonstrated that the relatives of schizophrenia tends to have increased risk of schizophrenia(Irving I Gottesman and James Shields, 1982). The implication of such observation was twofold: as family

members usually shares larger portion of their genetic effects with each other than that of the population, the genetic effects might be the main mediator of schizophrenia; on the other hand, culture, socio-economic status and area of birth usually also transmit within the family, so one cannot separate the environmental factors from the genetic factors.

It was important to study the relative contribution of genetic and environmental influence to individual differences in schizophrenia. If schizophrenia was indeed a genetic disease, we may then focus the resources into study of genetic variations in schizophrenia patients. To quantify the relative contribution of genetic and environmental influence, one will need to estimate the *heritability* of schizophrenia.

1.4 Broad Sense Heritability

A key concept in quantitative genetics is *heritability*, which was defined as *proportion* of total variance of a trait in a population explained by variation of genetic factors in the population. One can partition observed phenotype into a combination of genetic and environmental components(Falconer and Mackay, 1996)

$$\text{Phenotype(P)} = \text{Genotype(G)} + \text{Environment(E)}$$

where the variance of the observed phenotype (σ_P^2) can be expressed as variance of genotype (σ_G^2) and variance of environment (σ_E^2)

$$\sigma_P^2 = \sigma_G^2 + \sigma_E^2$$

The broad sense heritability can then be defined as the ratio between the variance of the observed phenotype and the variance of the genetic effects

$$H^2 = \frac{\sigma_G^2}{\sigma_P^2}$$

One key feature of heritability is that it is a *ratio* of *populational* measurement at a specific time point. As a result of that, the heritability estimation might differ from one population to another due to difference in Minor Allele Frequency (maf) and one might obtain a different heritability estimate if the method or time-point of measurement of the trait differs because of different environmental factors coming into play. A classic example was the study of intelligence quotient (IQ) where the heritability estimation increases with age(Bouchard, 2013). It was hypothesize that the shared environment has a larger effect on

individuals when they were young, and that as they become more independent, the effect of shared environment diminishes, leading to an *increased portion* of variance in IQ explained by the variance in genetic(Bouchard, 2013).

1.5 Narrow Sense Heritability

In reality, the problem of heritability was more complicated for there were different forms of genetic effects. For example, one can partition the genetic variance into variance of additive genetic effects (σ_A^2), variance of dominant genetic effects (σ_D^2) and other epistatic genetic effects (σ_I^2) such that

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2$$

where additive genetic variance was the variance explained by the average effects of all loci involved in the determination of the trait, whereas dominant genetic effects and epistatic genetic effects were the interaction between alleles at the *same* locus or *different* loci respectively.

As individuals only transmit one copy of each allele to their offspring, relatives other than full siblings and identical twins will only share a maximum of one copy of the allele from each other. Considering that dominance and non-additive genetic effects were concerning the interactive effect, which usually involve more than one copy of the alleles, these effects are unlikely to contribute to the resemblance between relatives (Visscher, Hill, and Wray, 2008). On the other hand, the additive genetic effects is usually transmitted from parent to offspring, thus it is usually more useful to consider the narrow sense heritability(h^2) which only consider the additive genetic effects:

$$\begin{aligned} h^2 &= \frac{\sigma_A^2}{\sigma_P^2} \\ h^2 &= \frac{\sigma_A^2}{\sigma_G^2 + \sigma_E^2} \end{aligned} \tag{1.1}$$

To obtain the additive genetic effect, we can first consider the genetic effect of parents to be $G_p = A + D$. As only half of the additive effect were transmitted to their offspring, the child will have a genetic effect of $G_c = \frac{1}{2}A + \frac{1}{2}A' + D'$ where A' is the additive genetic effect obtained from another parent by random and D' is the non-additive genetic

effect in the offspring. If we then consider the parent offspring covariance, we will get

$$\begin{aligned}\text{Cov}_{\text{OP}} &= \sum(\frac{1}{2}A + \frac{1}{2}A' + D')(A + D) \\ &= \frac{1}{2}\sum A^2 + \frac{1}{2}\sum AD + \frac{1}{2}\sum A'(A + D) + D'(A + D) \\ &= \frac{1}{2}V_A + \frac{1}{2}\text{Cov}_{AD} + \frac{1}{2}\text{Cov}_{A'A} + \frac{1}{2}\text{Cov}_{A'D} + \text{Cov}_{D'A} + \text{Cov}_{D'D} \end{aligned} \tag{1.2}$$

Under the assumption of random mating, A' should be independent from A and D. On the other hand, as D' was specific to the child, both of them should be independent from A and D. Moreover, the covariance between the additive genetics and non-additive genetics should be zero(Falconer and Mackay, 1996). Thus, eq. (1.2) becomes

$$\begin{aligned}\text{Cov}_{\text{OP}} &= \frac{1}{2}V_A + \text{Cov}_{AD} \\ &= \frac{1}{2}V_A \end{aligned} \tag{1.3}$$

Now if we assume the variance of phenotype of the parent and offspring were the same, then using eq. (1.3), we can obtain the narrow-sense heritability as

$$h^2 = \frac{1}{2}\frac{V_A}{\sigma_P^2} \tag{1.4}$$

If we consider the simple linear regression equation $Y = X\beta + \epsilon$, its slope can be calculated as

$$\beta_{XY} = \frac{\text{Cov}_{XY}}{\sigma_X Y} \tag{1.5}$$

which resemble eq. (1.4). Therefore, we can calculate the narrow sense heritability as

$$h^2 = 2\beta_{OP} \tag{1.6}$$

where β_{OP} is the slope of the simple linear regression regressing the phenotype of an offspring to the phenotype of *one* of its parents. We can further generalize eq. (1.6) to all possible relativeness

$$h^2 = \frac{\beta_{XY}}{r} \tag{1.7}$$

where r is the relativeness of X and Y.

A key assumption in this calculation was that the relatives does not share anything other than the additive genetic factors. However, this was usually not the case as relatives does tends to be in the same cultural group and might have similar socio-economic status which might all contribute to the variance of the trait. This might therefore lead to bias in

eq. (1.7) and we shall discuss the partitioning of variance in the later sections.

Nonetheless, eq. (1.7) was still useful for the understanding of the calculation of heritability. However, in the case of discontinuous trait (e.g. disease status) the calculation becomes more completed because the variance of the phenotype was dependent on the population prevalence. As eq. (1.7) does not account for the trait prevalence, it cannot be directly applied to discontinuous traits. In order to perform heritability estimation, we will need the concept of liability threshold model popularized by Falconer, 1965.

1.6 Liability Threshold

According the central limit theorem, if a phenotype is determined by a multitude of genetics and environmental factors with relatively small effect, then its distribution will likely follow a normal distribution as is the case of many quantitative traits(Visscher, Hill, and Wray, 2008). The variance of phenotype can therefore be calculated as the variance under the normal distribution. However, such is not the case for disease such as schizophrenia where instead of having a continuous distribution of phenotype, only a dichotomous labeling of "affected" and "normal" were obtained. The variance of these phenotype were therefore more difficult to obtain.

Falconer (1965) proposed the liability threshold model, which suggesting that these discontinuous traits also follow a continuous distribution with an additional parameter called the "liability threshold". Under the liability threshold model, the discontinuous traits were also affected by combination of multitude of genetics and environmental factors, each with a small effects, as in the case of the continuous traits. The main difference was that the phenotype of an individual is determined by whether if the combined effects of these factors("liability") were above a particular threshold ("liability threshold"). So for example, in the case of schizophrenia, only when an individual has a liability above the liability threshold will he/she be affected.

One can then estimate the heritability of the discontinuous by comparing the mean liability of the general population when compared to the relatives of the affected individuals.

For example, if we consider a single threshold model of a dichotomous trait, where

$$T_G = \text{Liability threshold of the general population}$$
$$T_R = \text{Liability threshold of relatives of the index case}$$
$$q_G = \text{Prevalence in the general population}$$
$$q_R = \text{Prevalence in relatives of the index case}$$
$$L_a = \text{Mean Liability of the index case}$$

by assuming both the liability distribution of the general population and that of the relative of the index case both follows the standard normal distribution, we can align the two distribution with respect to T_G and T_R. We can then calculate the mean liability of the index case L_a as $L_a = \frac{z_G}{q_G}$ where z_G is the density of the normal distribution at the liability threshold T_G. Then we can express the regression of relative's liability on the liability of the index case as

$$\beta = \frac{T_G - T_R}{L_a} \tag{1.8}$$

Thus, by applying eq. (1.8) to eq. (1.7), we get

$$h^2 = \frac{T_G - T_R}{L_a r} \tag{1.9}$$

1.7 Twin Studies of Schizophrenia

Now that we can deal with discontinuous traits, we shall come back to the limitation of eq. (1.7). The key limitation of eq. (1.7) was its inability to discriminate the genetic factors from the shared environmental factors. Such problem arise as family not only shared some of their genes, but they also tends to share some of the environmental factors such as diet. In fact, this was the main reason for researchers to discord the argument that schizophrenia was a genetic disorder.

A classical adoption study carried out by Heston (1966) in 1966 set off to discriminate whether if the increased risk of schizophrenia in relatives of schizophrenia was caused by the shared environmental factors or the shared genetic factors. An advantages of adoption studies was that if the child was separated from their family early after birth, then the shared environmental factors should be minimized, thus any resemblance between the parent and child should be driven mainly by the shared genetic factors. Heston (1966) collected data of

47 individuals born from a schizophrenic mother during the period from 1915 to 1947. They were separated from their mother within three day of birth and were sent to a foster family. 50 matched control were also recruited to the study. It was observed that there was an increased risk of schizophrenia in individual born to schizophrenic mother when compared to the control group even-though they were brought up in a different environment as that of their mother. This result suggested that schizophrenia was likely driven by the shared genetic factors instead of the shared environmental factors.

Despite the usefulness of adoption studies in delineating the effect of shared environment from the genetic factors, collection of adoption data were difficult. Moreover, any prenatal influence such as alcohol abuse during pregnancy might confound the results. Therefore, an alternative way would be the twin studies using the relationship between the monozygotic (MZ) and dizygotic (DZ) twins.

Theoretically, MZ twins should share all their genetic components (both additive(A) and non-additive(D) genetic factors) and also their common environmental factors(C) where the only difference between a twin pair would be the non-shared environmental factors(E). As for the DZ twins, they should also share the same common environmental factors yet they only share $\frac{1}{2}$ of their additive genetic factors and $\frac{1}{4}$ of their non-additive genetic factors. The non-shared environmental was also by definition not shared among the twins(Rijsdijk and Pak C Sham, 2002). Based on these assumptions, Falconer and Mackay, 1996 derived the heritability as

$$h^2 = 2(\rho_{MZ} - \rho_{DZ}) \tag{1.10}$$

where ρ_{MZ} and ρ_{DZ} were the phenotype correlation between the MZ twins and DZ twins respectively.

By combining Falconer's formula and the concept of liability threshold model, I I Gottesman and J. Shields (1967) estimated that the heritability of schizophrenia to be $> 60\%$ based on previously collected twin data, strongly suggesting schizophrenia as a genetic disorder. The result was further supported by one of the landmark meta-analysis study conducted by Sullivan, Kendler, and Neale, 2003. Based on data obtained from 12 published schizophrenia twin studies, the authors found that although there was a non-zero contribution of environmental influence on liability of schizophrenia (11%,confidence interval (CI)=3% − 19%), there was a much larger contribution from genetics (81%, CI=73% − 90%), further supporting that schizophrenia was largely mediated by the genetic factors.

Such findings were not limited to twin-studies but were also reported in large scale population based studies. A recent large scale population based study in Sweden popu-

lation(Lichtenstein et al., 2009) also found that there was a large genetic contribution in schizophrenia (64%). Although the estimated heritability(64%(Lichtenstein et al., 2009) vs 81%(Sullivan, Kendler, and Neale, 2003)) differs between the two studies, they, there is no doubt that schizophrenia is highly heritable, leading to the initiative of genetic research in schizophrenia.

1.8 Genetic Analysis of Schizophrenia

1.8.1 Genetic Architecture of Schizophrenia

Studies on estimation of heritability of schizophrenia strongly support schizophrenia as a genetic disorder. However, little was known about the mechanism of schizophrenia nor the genetic architecture of the disorder. All data from adoption studies, twin studies and family studies shown that schizophrenia does not follow the Mendelian framework(I I Gottesman and J. Shields, 1967; Irving I Gottesman and James Shields, 1982). Specifically, shall schizophrenia be a Mendelian disorder, then we would expect all MZ siblings of the proband to also suffer from schizophrenia. However, the life time morbid risk of monozyogitc twins were only 48%(fig. 1.3)(I. Gottesman, 1991), making it unlikely for schizophrenia to follow a Mendelian pattern.

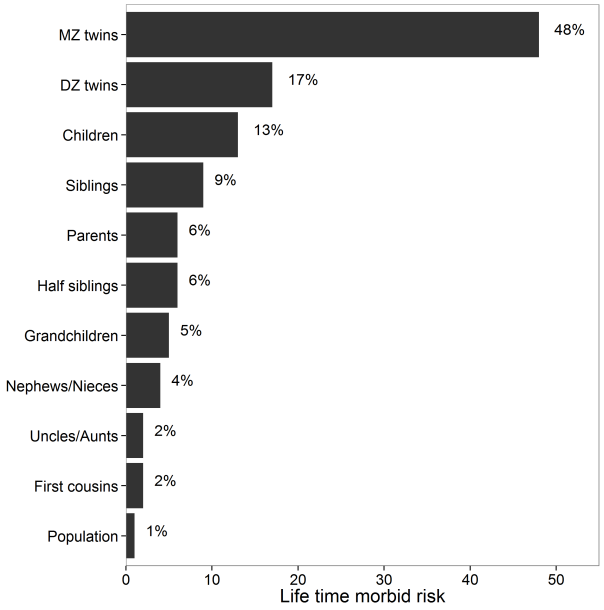


Figure 1.3: Lifetime morbid risks of schizophrenia in various classes of relatives of a proband. It was noted that the morbid risk of monozygotic (MZ) twins were only 48%, much lower than one would expect if schizophrenia follows a Mendelian pattern. Reproduced with permission from journal(Riley and Kendler, 2006).

Based on these observations, I. Gottesman and J. Shields, 1967 proposed that schizophrenia follows a polygenic model where disease phenotype were determined by the additive effects from multiple genes. Thus, schizophrenia is a complex genetic disorder with complicated pattern of inheritance. Their hypothesis was supported by the calculation of Risch, 1990a by taking into account of dif-

ferent inheritance model and the life time morbid risk observed in relatives of affected individuals.

Another interesting conclusion from the calculation of Risch (1990a) was the effect size of individual locus. By comparing the observed life time morbid risk and the calculated risk from different models, Risch suggested that genetic models with a single locus with risk of 3.0 and with all other loci of small effect or models with two or three loci with risk of 2.0 were most consistent with the observed life time morbid risk of schizophrenia. (Risch, 1990b).

Risch's calculation provided an explanation for the early inconsistent findings of linkage studies in schizophrenia(Harrison and Weinberger, 2005). As linkage studies were aimed to identify genetic variation of large effect size they failed to capture genetic loci with small effect size. It was therefore tempting to suggest that schizophrenia only follows the "common disease-common variant" model, which stated that schizophrenia should be mediated by large amount of common variants such as Single Nucleotide Polymorphism, each carries a small effect size.

However, another possible hypothesis was that the variation mediating schizophrenia were rare, therefore require a large sample size to detect. The inconsistent results of the early linkage studies might be due to the inadequate sample size. This lead to some researchers suggesting the "common disease-rare variant" hypothesis, which propose that schizophrenia was mediated by a small amount of rare variants, each with a large effect size(McClellan, Susser, and King, 2007).

Nevertheless, success in genetic research of schizophrenia remains limited. Only until the initiation of Human Genome Project and the technological advance resulted from it that does genetic research of schizophrenia entered an era of success.

1.8.2 The Human Genome Project and HapMap Project

In 1990, the Human genome project was initiated, aiming at constructing the first physical map of the human genome at per nucleotide resolution(Lander et al., 2001). The completion of the human genome project has opened up a new era of genetic research, allowing researchers to identify Single Nucleotide Polymorphisms (SNPs) on the human genome, which is one of the major source of genetic variation.

Soon after the completion of the human genome project, the HapMap Project was

initiated(Consortium, 2005), aiming to provide a genome-wide database of common human sequence variation such as SNPs with maf ≥ 0.05. More importantly was that the HapMap Project also provided a detailed Linkage Disequilibrium (LD) map of the human genome.

LD was of particular importance to genetic research for it was the non-random correlation of genotypes between 2 genetic loci. SNPs in high LD were usually observed together in the human genome. When a large amount of SNPs were in high LD together, they form what was known as a LD block. By performing association testing on SNPs representing a LD block("tagging"), one can avoid the need of performing association on the whole genome, therefore reducing the cost of the experiment. This was the fundamental concept of Genome Wide Association Study (GWAS) which was now extensively used in the genetic research.

1.8.3 Genome Wide Association Study

In GWAS, genome-wide genotyping array were commonly used to systematically detect genetic variants such as SNP and copy number variation (CNV). For quantitative traits, the association between the trait and frequency of the variants were calculated using methods such as linear regression. On the other hand, for dichotomous traits such as schizophrenia, the frequency of the variants were compared between the case and control samples using methods such as chi-square test or logistic regression. Because of the problem of multiple testing, only variants with a p-value passing a genome wide threshold (p-value $\leq 5 \times 10^{-8}$) were considered significant. Another possible method to decide the significant threshold was to consider the "effective number" of tests(Li et al., 2011) taking into consideration of LD as not all tests in a GWAS were independent of each other. The power of the GWAS were determined by the magnitude of effect, sample size, and required level of statistical significance(the false-positive, or type I, error rate)(S. Purcell, Cherny, and P C Sham, 2003).

Single Nucleotide Polymorphism

Despite the great promise from GWAS, early GWAS in schizophrenia remain largely disappointing and were unable to identify any robust genetic markers associated with schizophrenia. The failure of early GWAS in schizophrenia were mainly due to the relative small sample size of the studies, which result in low detection power.

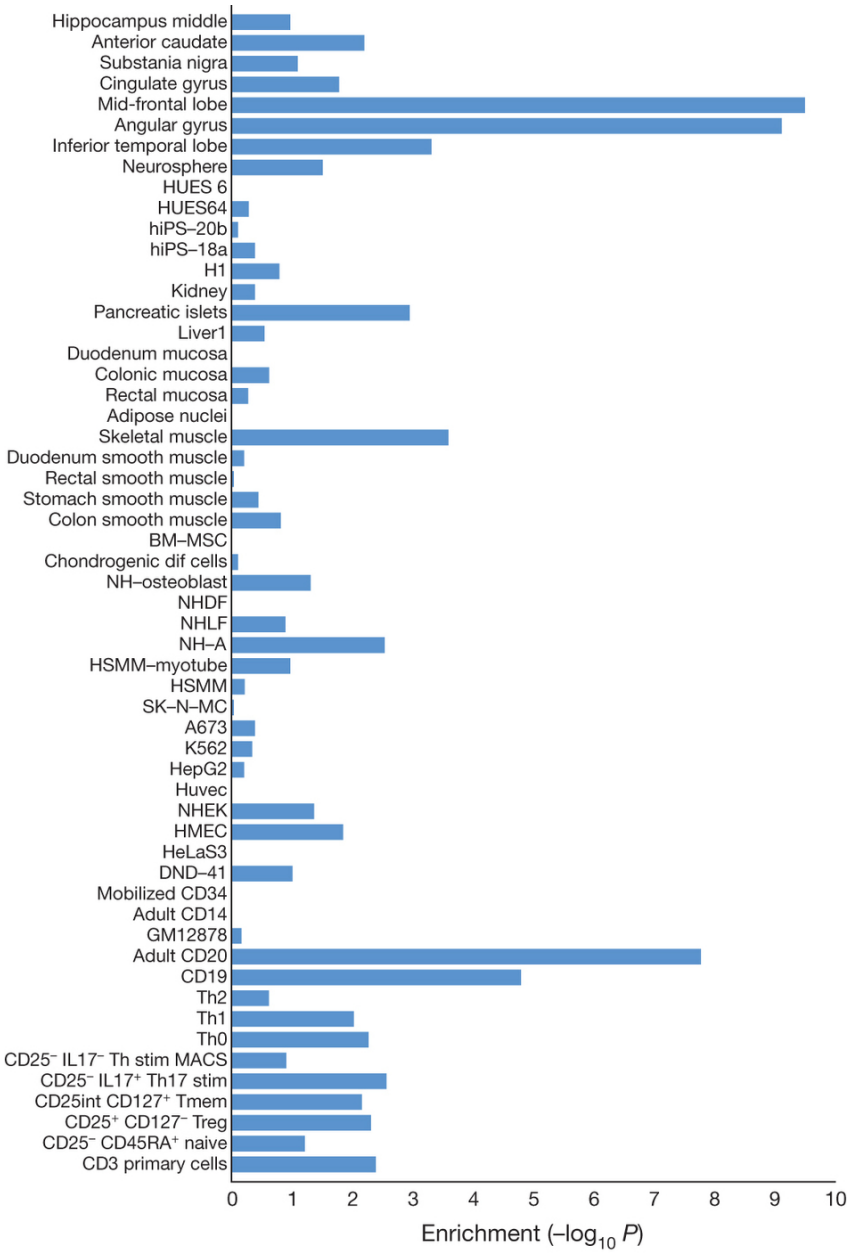
To overcome the problem of small sample size, large consortium were formed such that data from different research groups from different countries were combined, essentially providing a large sample size for the analysis. By 2014, the Schizophrenia Working group of the Psychiatric Genomics Consortium (PGC) has collected 34,241 schizophrenia samples and 45,604 controls(Stephan Ripke et al., 2014). By combining the samples with those obtained by deCODE genetics, a total of 36,989 schizophrenia samples and 113,075 controls were used for the largest meta-analysis of schizophrenia. In their study(Stephan Ripke et al., 2014), 128 linkage-disequilibrium-independent SNPs were found to exceeded the genome-wide significance(p-value $\leq 5 \times 10^{-8}$), corresponding to 108 genetic loci. 75% of these loci contain protein coding genes and a further 8% of these loci were within 20kb of a gene. It was found that genes involved in glutamatergic neurotransmission (e.g. *GRM3*, *GRIN2A* and *GRIA1*), synaptic plasticity and genes encoding the voltage-gated calcium channel subunits (e.g. *CACNA1C*, *CACNB2* and *CACNA1I*) were among the genes associated within these loci. Importantly, *DRD2*, the target of all effective anti-psychotic drug were also associated with schizophrenia. This result converges with existing knowledge of *DRD2* being involved in the pathology of schizophrenia, supported by multiple lines of research(Talkowski et al., 2007). It was further demonstrated that schizophrenia association were significantly enriched at enhancers active in brain and enriched at enhancers active in tissues with important immune functions(fig. 1.4)(Stephan Ripke et al., 2014).

The enrichment of immune related enhancers remains significant even after the removal of major histocompatibility complex (MHC) region from the analysis, provided further genetic support of the involvement of the immune system in the etiology of schizophrenia. Because of its role in neural development(Zhao and Schwartz, 1998; Deverman and Patterson, 2009), it is likely that the perturbation in the immune system might disrupt the brain development, therefore increasing the risk of schizophrenia. Indeed, studies on MIA has demonstrated that cytokine imbalance might predispose individual to schizophrenia(U Meyer, J Feldon, and Yee, 2009).

Copy Number Variation

Another important arm of genetic research in schizophrenia was to identify copy number variation (CNV) associated with schizophrenia. CNV were classified as segment of DNA that is 1kb or larger and that is present at a different copy number when compared to the reference genome, usually in the form of insertion, deletion or duplication(Feuk, Carson, and Scherer, 2006). Due to the length of these variants, the CNV might contain the entire

Figure 1.4: Enrichment of enhancers of SNPs associated with schizophrenia. It was observed that the largest enrichment were in cell lines related to the brain and in tissues with important immune functions. Graphs reproduced with permission from the journal.(Stephan Ripke et al., 2014)



genes and their regulatory regions which might in turn contribute to significant phenotypic differences(Feuk, Carson, and Scherer, 2006).

To identify robust association between CNV and schizophrenia, Szatkiewicz et al., 2014 conducted a GWAS for CNV association with schizophrenia used the Swedish national sample (4,719 schizophrenia samples and 5,917 controls). In their study, they were able to association between schizophrenia and CNV such as 16p11.2 duplications, 22q11.2 deletions, 3q29 deletions and 17q12 duplications were identified. Through the gene set association analysis, calcium channel signaling and binding partners of the fragile X mental retardation protein were found to be associated with these CNV(Szatkiewicz et al., 2014). Interestingly, the calcium channel signaling were also enriched in the PGC GWAS on SNP association, suggesting that the variants were converging on similar set of pathway or gene sets.

Unlike the result form the GWAS on SNP data, the CNV identified were rare(≤ 12 in 4,719 samples) and has a relative large effect (e.g. 22q11 deletion has an odd ratio of 16.32(Szatkiewicz et al., 2014)). The results from the SNP GWAS supports the "common disease-common variant" model whereas the GWAS on CNV supports the "common disease-rare variant" model, illustrating the complex genetic model behind the etiology of schizophrenia.

Although the GWAS in schizophrenia seems to return a lot of interesting results, the question remains: How much of the known genetic risk factors associated explain the disease risk of schizophrenia? To answer these question, we need to estimate the heritability based on the GWAS data. However, in order to obtain the large volume of data, most of the samples were not relatives. How can one estimate the heritability based only on the genetic data of the general population instead of family or twin data?

1.8.4 Genome-wide Complex Trait Analysis

Unlike family based data, the relationship between the samples were unknown. Yet in a typical GWAS, the genotype of each individuals were known. The "genetic distance" between two individual will provide an estimate of their relationship, thus allowing the calculation of heritability. J Yang et al. (2011) use the concept of genetic distance to calculate the Genetic Relationship Matrix (GRM) to represent the relationship between individuals. The GRM were then used in the restricted maximum likelihood analysis(REML) to estimate the heritability of the trait(J Yang et al., 2011). This was implemented in Genome-wide Complex Trait Analysis (GCTA) and were now wildly used in the estimation of heritability

on GWAS data.

The problem with GCTA was that it require the genotype data to estimate the heritability. However, for complex disease like schizophrenia, the data were usually obtained from multiple data source. Because of privacy issues, usually only the test statistic were shared among the research groups and only meta-analysis were performed. Given there was no raw genotype data, it is impossible to calculate the GRM, thus making the use of GCTA impossible.

1.8.5 LD SCore

Sometimes, in a GWAS study, one can observe a general inflation of test statistics. It was usually considered to be contributed to the presence of confounding factors such as population stratification under the assumption that most of the SNPs should have no association to the disease. It was therefore a common practice for one to perform the Genomic Control (GC) on the GWAS results(Zheng, Freidlin, and Gastwirth, 2006).

The problem of the GC was that the basic assumption of a small number of causal SNPs might not be true. Through careful simulation, Jian Yang et al. (2011) demonstrated that in the absence of population stratification and other form of technical artifacts, the presence of polygenic inheritance can also inflate the test statistic(Jian Yang et al., 2011). More importantly, they observed that the magnitude of inflation was determined by the *heritability*, the LD structure, sample size and the number of causal SNPs of the trait.

Following on this observation, B. K. Bulik-Sullivan et al. (2015) developed the LD SCore (LDSC). The fundamental concept of LDSC was that the more genetic variant a SNP tag, the more likely for it to be able to tag a causal variant; whereas population stratification and cryptic relatedness should not be associated with LD. The number of genetic variants tagged by a SNP_j (l_j)(LD score) was then defined as the sum of r^2 of the k SNPs within a 1cM window of SNP_j:

$$l_j = \sum_k r_{jk}^2 \tag{1.11}$$

The expected χ^2 of SNP_j was then defined as a function of the LD score (l_j), the number of samples (N), the number of SNPs in the analysis(M), the contribution of confounding factors (a) and most importantly, the heritability (h^2):

$$\text{E}[\chi_j^2|l_j] = \frac{Nl_jh^2}{M} + Na + 1 \tag{1.12}$$

If one express the LD score and the χ^2 as vectors $\boldsymbol{L}$ and $\boldsymbol{\chi^2}$ respectively, eq. (1.12) becomes a regression of the χ^2 against the LD score:

$$\boldsymbol{\chi^2} = \frac{N}{M}\boldsymbol{L}h^2 + Na + 1 \tag{1.13}$$

As a result of that, the heritability h^2 will be the slope of the regression and the intercept minus one will represent the mean contribution of the confounding bias such as those of population stratification. Thus, eq. (1.13) can be used for the estimation of heritability given only the test statistics and the population LD were provided.

Using data from Stephan Ripke et al. (2014), and applying the liability threshold adjustment, B. K. Bulik-Sullivan et al. (2015) estimated the heritability of schizophrenia should be 0.555 with standard error (SE) of 0.008. The estimated heritability was lower than what was previously estimated from population based study(64%(Lichtenstein et al., 2009)) and twin studies(81%(Sullivan, Kendler, and Neale, 2003)). Possible reasons of such discrepancies might be that in Stephan Ripke et al. (2014)'s study, only SNPs data were collected. From Szatkiewicz et al. (2014), it was clearly demonstrated that other than SNPs, CNVs were also associated with schizophrenia. By ignoring CNVs in the estimation of heritability, the estimation of B. K. Bulik-Sullivan et al. (2015) would only provide a lower bound of heritability estimated. Another possibility of the"missing" heritability can be due to interaction between the genetic and environmental factors. Although previous studies(I I Gottesman and J. Shields, 1967) suggested that the non-additive genetic factors were unlikely to contribute to schizophrenia, the possibility of involvement of gene-environmental interaction $G \times E$ were not ruled out. Indeed, in the adoption study conducted by Tienari et al. (2004), it was found that individuals with higher genetic risk were significantly more sensitive to "adverse" vs "healthy" rearing patterns in adoptive families than are adoptees at low genetic risk(Tienari et al., 2004), providing support to a possible interaction between genetic and environmental factors. Therefore, in order to account for the "missing" heritability, one might need to consider genetic variations other than SNPs and might need to take into consideration of the $G \times E$ interaction.

Nonetheless, the heritability estimation from Stephan Ripke et al. (2014) were still encouraging, as for the first time in genetic research of schizophrenia, a large portion of heritability of schizophrenia were finally identified. This permit the genetic research of schizophrenia to move beyond statistical association and focus on the functional basis of the genetic susceptibility locus of schizophrenia.

1.8.6 Partitioning of Heritability of Schizophrenia

Traditionally, functional enrichment analysis in GWAS only take into account of SNPs that passed the genome wide significance threshold. However, for complex traits such as that of schizophrenia, much of the heritability might lies in SNPs that do not reach genome wide significance threshold at the current sample size. For example, in 2013, only 13 risk loci were detected using 13,833 schizophrenia samples and 18,310 controls (S Ripke et al., 2013). When the sample size increased to 34,241 schizophrenia samples and 45,604 controls in 2014, 108 risk loci were identified(Stephan Ripke et al., 2014). Thus, if one only consider the significant loci, risk loci that have not reach genome wide significance threshold might be ignored from the analysis, decreasing the power of the functional enrichment analysis.

Unlike traditional functional enrichment analysis, LDSC uses information from all SNPs and taking into account of the LD structure to partition heritability into different functional categories. Thus should be more powerful when compared to traditional analysis and should help to provide useful insight into the disease etiology of schizophrenia.

Finucane et al. (2015) used data from Stephan Ripke et al. (2014) and functional categories derived from the ENCODE annotation(ENCODE Project Consortium, 2012), the NIH Roadmap Epigenomics Mapping Consortium annotation(Bernstein et al., 2010) and other studies(Finucane et al., 2015), it was found that the brain cell types were most enriched in schizophrenia, especially those related to the central nervous system (CNS). Of all the functional categories, the most enriched category in schizophrenia was the H3K4me3 mark in the fetal brain(table 1.2). As H3K4me3 was mostly linked to active promoters, it was likely for genes that were active in fetal brain (e.g. genes related to brain development) to be associated with schizophrenia, supporting the idea of schizophrenia as a neuro-developmental disorder.

Moreover, it was also observed that the second most enriched cell types were those related to immunity. Undoubtedly, the CNS and the immune system have an important role in the disease etiology of schizophrenia.

Cell type	cell-type group	Mark	P-value
Fetal brain**	CNS	H3K4me3	3.09×10^{-19}
Mid frontal lobe**	CNS	H3K4me3	3.63×10^{-15}
Germinal matrix**	CNS	H3K4me3	2.09×10^{-13}
Mid frontal lobe**	CNS	H3K9ac	5.37×10^{-12}
Angular gyrus**	CNS	H3K4me3	1.29×10^{-11}
Inferior temporal lobe**	CNS	H3K4me3	1.70×10^{-11}

Cingulate gyrus**	CNS	H3K9ac	5.37×10^{-11}
Fetal brain**	CNS	H3K9ac	5.75×10^{-11}
Anterior caudate**	CNS	H3K4me3	2.19×10^{-10}
Cingulate gyrus**	CNS	H3K4me3	4.57×10^{-10}
Pancreatic islets**	Adrenal/Pancreas	H3K4me3	2.24×10^{-09}
Anterior caudate**	CNS	H3K9ac	3.16×10^{-9}
Angular gyrus**	CNS	H3K9ac	4.68×10^{-9}
Mid frontal lobe**	CNS	H3K27ac	7.94×10^{-9}
Anterior caudate**	CNS	H3K4me1	1.20×10^{-8}
Inferior temporal lobe**	CNS	H3K4me1	3.72×10^{-8}
Psoas muscle**	Skeletal Muscle	H3K4me3	4.17×10^{-8}
Fetal brain**	CNS	H3K4me1	6.17×10^{-8}
Inferior temporal lobe**	CNS	H3K9ac	9.33×10^{-8}
Hippocampus middle**	CNS	H3K9ac	9.33×10^{-7}
Pancreatic islets**	Adrenal/Pancreas	H3K9ac	1.62×10^{-6}
Penis foreskin melanocyte primary**	Other	H3K4me3	2.09×10^{-6}
Angular gyrus**	CNS	H3K27ac	2.34×10^{-6}
Cingulate gyrus**	CNS	H3K4me1	2.82×10^{-6}
Hippocampus middle**	CNS	H3K4me3	2.82×10^{-6}
CD34 primary**	Immune	H3K4me3	4.68×10^{-6}
Sigmoid colon**	GI	H3K4me3	5.01×10^{-6}
Fetal adrenal**	Adrenal/Pancreas	H3K4me3	6.31×10^{-6}
Inferior temporal lobe**	CNS	H3K27ac	8.32×10^{-6}
Peripheralblood mononuclear primary**	Immune	H3K4me3	9.33×10^{-6}
Gastric**	GI	H3K4me3	1.17×10^{-5}
Substantia nigra*	CNS	H3K4me3	1.95×10^{-5}
Fetal brain*	CNS	H3K4me3	2.63×10^{-5}
Hippocampus middle*	CNS	H3K4me1	3.31×10^{-5}
Ovary*	Other	H3K4me3	6.46×10^{-5}
CD19 primary (UW)*	Immune	H3K4me3	7.08×10^{-5}
Small intestine*	GI	H3K4me3	8.51×10^{-5}
Lung*	Cardiovascular	H3K4me3	1.17×10^{-4}
Fetal stomach*	GI	H3K4me3	1.29×10^{-4}
Fetal leg muscle*	Skeletal Muscle	H3K4me3	1.51×10^{-4}
Spleen*	Immune	H3K4me3	1.70×10^{-4}
Breast fibroblast primary*	Connective/Bone	H3K4me3	2.04×10^{-4}
Right ventricle*	Cardiovascular	H3K4me3	2.14×10^{-4}
CD4+ CD25- Th primary*	Immune	H3K4me3	2.19×10^{-4}
CD4+ CD25- IL17- PMA Ionomycin stim MACS Th sprimary*	Immune	H3K4me1	2.19×10^{-4}
CD8 naive primary (UCSF-UBC)*	Immune	H3K4me3	2.24×10^{-4}
Pancreas*	Adrenal/Pancreas	H3K4me3	2.34×10^{-4}

CD4+ CD25- Th primary*	Immune	H3K4me1	2.75×10^{-4}
CD4+ CD25- CD45RA+ naive primary*	Immune	H3K4me1	2.75×10^{-4}
Colonic mucosa*	GI	H3K4me3	3.24×10^{-4}
Right atrium*	Cardiovascular	H3K4me3	3.31×10^{-4}
Fetal trunk muscle*	Skeletal Muscle	H3K4me3	3.39×10^{-4}
CD4+ CD25int CD127+ Tmem primary*	Immune	H3K4me3	3.47×10^{-4}
Substantia nigra*	CNS	H3K9ac	3.63×10^{-4}
Placenta amnion*	Other	H3K4me3	4.17×10^{-4}
Breast myoepithelial*	Other	H3K9ac	5.50×10^{-4}
CD8 naive primary (BI)*	Immune	H3K4me1	5.75×10^{-4}
Substantia nigra*	CNS	H3K4me1	6.61×10^{-4}
Cingulate gyrus*	CNS	H3K27ac	7.94×10^{-4}
CD4+ CD25- CD45RA+ naive primary*	Immune	H3K4me3	8.71×10^{-4}

Table 1.2: Enrichment of Top Cell type of Schizophrenia. * = significant at False Discovery Rate < 0.05. ** = significant at p < 0.05 after correcting for multiple hypothesis. Reproduce with permission from Journal.(Finucane et al., 2015)

1.8.7 Genetic Correlation

Another very important application of LDSC is that it allow one to identify the genetic correlation between traits(B. Bulik-Sullivan et al., 2015). The genetic correlation can be used as a genetic analogue to co-morbidity, thus allowing deeper understanding to the etiology of the traits. Above all, genetic correlation was important in studying the treatment response. It has been observed that there was an increased prevalence of anxiety, depression and substance abuse in schizophrenia(Buckley et al., 2009). These co-morbidity were generally associated with more severe psychopathology and with poorer outcome(Buckley et al., 2009). A deeper understanding of possible co-morbidity between different traits and schizophrenia might provide insight not only to the disease etiology of schizophrenia, it might even provide important information in possible treatment options for schizophrenia. Using breast cancer as an example, it was found that patients with comorbidity had poorer survival than those without comorbidity(Søgaard et al., 2013) and it was suggested that by treating the comorbid diseases, one might be able to delay mortality in breast cancer patients(Ording et al., 2013).

By applying their method to 25 different phenotypes, B. Bulik-Sullivan et al. (2015) shown that schizophrenia has significant genetic correlation with bipolar disorder, major depression and more surprisingly, anorexia nervosa. Previous studies have always suggest

there to be a co-morbidity between schizophrenia and bipolar disorder (Lichtenstein et al., 2009; S. M. Purcell et al., 2009; Buckley et al., 2009). Similarly, it was not uncommon for schizophrenia to display depressive symptoms(Buckley et al., 2009). It was even observed that individuals at high risk and ultrahigh risk for developing schizophrenia have generally demonstrated a significant degree of depressive symptoms prior to and during the emergence of psychotic symptoms, suggesting a close relationship between schizophrenia and depression.

On the other hand, the genetic correlation between schizophrenia and anorexia nervosa were slightly unexpected for there has been a lack of study in the co-morbidity between eating disorder and schizophrenia. Nonetheless, this finding raises the possibility of similarity between anorexia and nervosa.

1.9 Antipsychotics

Despite the success in the genetic research of schizophrenia, an effective cure of schizophrenia was yet to be found. Currently, the main treatment method for schizophrenia was the use of antipsychotic drugs to reduce symptoms and prevent relapse. However, there was a large variability between individuals in their response to treatment, some might even suffer from adverse side effects such as agranulocytosis and Tardive dyskinesia (TD). Thus, it is vital to administrate the antipsychotics according to individual conditions. Unfortunately, there was a lack of understanding of the factors influencing the drug response, forcing clinicians to administrate antipsychotics on a trial and error process. There is a therefore a pressing need for better understanding treatment response in schizophrenia such that an optimal treatment can be provided for the patients.

1.9.1 History of Antipsychotic

Early research in treatment of schizophrenia largely follows a random trial and error process where methods such as prolonged sleep treatment, insulin coma therapy and pharmacoconvulsive treatment were proposed(Lehmann and Ban, 1997). The first antipsychotic drug Chlorpromazine, a phenothiazine, were developed in early 1950s. Subsequently within a period of less than 10 years, 20 other antipsychotic phenothiazine were in development. Collectively, they were considered as the First Generation Antipsychotics (FGAs).

FGAs were found to be extremely effective in reducing the positive symptoms of

schizophrenia such as delusions, hallucinations and disorganized thinking. However, the FGAs were found to be ineffective against negative and cognitive symptoms, and might even cause acute extrapyramidal motor symptoms (EPS) such as parkinsonism, dysphoria and tardive dyskinesia, making them unpopular among patients(Tandon, 2007).

In 1966, a new drug, name Clozapine was introduced(Lehmann and Ban, 1997). Clozapine has been shown to be more effective when compared to FGAs and was less likely to cause EPS and tardive dyskinesia. Moreover, it was shown to reduce suicidality and was more effective in reducing negative and cognitive symptoms(Lehmann and Ban, 1997; Tandon, 2007). Despite the superior performance of Clozapine, it found to be associated with the severe and potentially lethal adverse side effect, agranulocytosis(Alvir et al., 1993), limiting its use as a first line treatment of schizophrenia(Remington et al., 2013). Subsequently, a number of antipsychotic were developed in hope of a "safe clozapine" which have the same level of effectiveness as clozapine and not having the adverse side effects. These were considered as the Second Generation Antipsychotics (SGAs) which includes risperidone, olanzapine and quetiapin. Although the SGAs tends to have lower risk for EPS, they tends to be associated with significant metabolic side effects such as weight gain, diabetes mellitus and hyperlipidemia(Üçok and Gaebel, 2008).

1.9.2 Mechanism of Action of Antipsychotic

The difference between the FGAs and SGAs provides valuable information on possible mechanisms associated with treatment response and adverse side effects such as EPS. It was first demonstrated on 1963 that FGAs tends to block the dopamine receptors(Lehmann and Ban, 1997) and it was hypothesized that the binding of dopamine receptors, especially the D_2 receptors were required for reduction of positive symptoms(M J Arranz and Leon, 2007). Indeed, it was found that dopamine receptor blockade was not unique to FGAs but was also required for SGAs and there has yet been any successful antipsychotic drugs that works without dopamine D_2 blockade(Jian-Ping Zhang and Anil K Malhotra, 2011). However, it was observed that there were significant differences between FGAs and SGAs affinities.

When compared to FGAs, SGAs have a lower affinity for and occupancy at the D_2 receptors and tends to have a more diverse receptor binding profiles. For example, the ratio between affinity of serotonin receptor (5-HT_2) to that of the D_2 receptor were significantly greater (15.8 times) for SGAs when compared to FGAs(Meltzer, 1991). These leads to two competing hypothesis of antipsychotic action: the serotonin-dopamine hypothesis, which

stated that the ratio of serotonin 5-HT_2 to dopamine D_2 affinity was the main mechanism accounting for the superior performance of SGAs; and the dopamine hypothesis which stated that the modulation of the dopamine D_2 receptor was the single most important factor affecting the performance of the antipsychotic(Shitij Kapur and Mamo, 2003).

One common characteristics for most SGAs except amisulpride was their affinity to the serotonin receptors such as 5-HT_2. It was therefore suggested that the reduction of negative symptoms were resulting from the serotonin blockade and the serotonin-dopamine interactions were important to the antipsychotic drug actions(Meltzer, 1999). However, amisulpride serves as a counter example to the serotonin-dopamine hypothesis. Amisulpride is a SGA that does not have any affinity for serotonin receptors yet have comparable performance in reduction of negative response when compared to olanzapine(Kumar and Chaudhury, 2014). Thus, serotonin receptor blockade might not be required for the reduction of negative symptoms.

Moreover, positron emission tomography (PET) studies have shown that a minimum occupancy of 60%-65% of striatal D_2 like receptors is required to obtain clinical response whereas D_2 occupancy of above 80% is considered as the main cause of EPS(M J Arranz and Leon, 2007; Shitij Kapur and Mamo, 2003). Upon further investigation, it was found that clozapine preferentially target the mesolimbic dopamine system while sparing the nigrostriatal dopamine system(Gardner, Walker, and Paredes, 1993). This raise the possibility that the main difference between FGAs and SGAs was the preferential blockade of cortical dopamine D_2 receptors compared with striatal dopamine D_2 receptors(Shitij Kapur and Mamo, 2003). Based on these observation, it was now hypothesized that schizophrenia was a result of both "hypodopaminergia" in the prefrontal cortex and "hyperdopaminergia" in the straitum, with a possible involvement of the glutamate system(Howes and S Kapur, 2009).

It was worth noting that most clinical studies of SGAs were sponsored by industry, leading to questions of their validity. Two government lead clinical trial, Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE)(Lieberman et al., 2005) and CUtLASS(Jones et al., 2006), were therefore performed to provide unbiased comparison between FGAs and SGAs. Unfortunately, the superior performance of SGAs over FGAs were not observed nor were the SGAs associated with better cognitive or social outcomes. It therefore seems like the only advantages of SGAs over FGAs were the reduced risk of adverse side effects such as EPS and TD.

1.9.3 Antipsychotic Response

Although the government lead studies does not support SGAs's role in reducing negative and cognitive symptoms of schizophrenia, there is without doubt that SGAs were better in terms of reduced risk of EPS and TD. There is no question that a better treatment is required yet it is just as important to learn how to better utilize currently available antipsychotics. Simply a better understanding of factors behind the variation in individual responses to different antipsychotic drugs will be extremely beneficial. It will allow researchers to categorize people by their personal profile and provide the most optimal antipsychotic drug for their treatment.

It is worth noting that the antipsychotic drug response is a multidimensional problem which not only focus in the reduction of symptoms, but the instance of adverse drug effect is also an important research focus. However, due to limited scope of the current thesis, we will focus only on studies on the reduction of symptoms.

Positive and Negative Symptom Scale (PANSS)

In order to study the response of antipsychotic, it is important to have an objective scale to quantify the reduction of symptoms. The Positive and Negative Symptom Scale (PANSS)(Kay, Fiszbein, and Opler, 1987) were among one of the most commonly scale used to measure the core symptoms of schizophrenia and is composed of 3 subscales: positive, negative and general psychopathology. There were a total of 30 different symptoms included in PANSS and each symptoms were rated from 1 to 7, thus the minimal score for PANSS is 30. To calculate the percentage reduction of PANSS, which represent a reduction in severity of symptoms, the reduction of PANSS will then be divided by the original PANSS minus 30:

$$\%\text{improvement} = \frac{\text{PANSS}_{after} - \text{PANSS}_{before}}{\text{PANSS}_{before} - 30} \times 100\%$$

Factors Associated with Antipsychotic Responses

Factors such as diet, smoking and concomitant medications were known to significantly affect metabolic enzyme activity rates, thus have an impact to antipsychotic treatment response(Maria J Arranz and Munro, 2011). On the other hand, clinical features such as treatment adherence and duration of illness; individual variation such as gender and ethnicity all influence the treatment efficacy(Maria J Arranz and Munro, 2011).

Considering the heritability of schizophrenia were up to 80%, genetic variations

can explain much of the variation in schizophrenia. Therefore, people hypothesize that the genetic variations might also be able to explain much of the variation in antipsychotic drug response. However, although there were incidence report of concordance of response in MZ twin data(Vojvoda et al., 1996; Mata et al., 2001), the sample size were not enough for heritability estimation(studies usually consist of only one pair of twins). Nonetheless, these studies shades lights on the possibility that variation in antipsychotic response might be able to be explained by genetic variations of individuals.

1.9.4 Pharmacogenetics and Pharmacogenomics

Given that genetic variations might be able to explain the variation in antipsychotic drug response, it was therefore compelling to study the association between genetic variations and antipsychotic drug response. The terms "pharmacogenetics" and "pharmacogenomics" were introduced to define study of variability in drug response due to genetic variations and can be used interchangeably(Pirmohamed, 2001). Before the popularity of GWAS, pharmacogenetic studies were mainly conducted based on the candidate gene approach. Genes targeted by antipsychotic drugs such as genes coding for dopamine receptors and serotonin receptors were among the major target of research. Similarly, genes involve in the metabolizing the antipsychotic drugs such as the P450 family of enzymes were extensively studied.

Dopamine Receptors

The dopamine D_2 receptor plays a critical role in antipsychotic drug action, with D_2 receptor antagonism considered to be necessary and sufficient for antipsychotic drug efficacy(Shitij Kapur and Mamo, 2003). As such, polymorphisms on the *DRD2* gene, which codes for the D_2 receptor, were extensively studied. The SNP(rs1799732) representing a deletion at position -141, which was located in the 5' promoter region of *DRD2* were found to be able to influence the density of D_2 receptor density in the striatum in healthy samples unexposed to antipsychotic drug treatment(Arinami et al., 1997). A significant difference in response rate between deletion carrier and patients with homozygous insertion genotype were observed (odds ratio = 0.65, 95% CI: 0.43-0.97), indicating patients who carry one or two deletion allele were more likely to have less favorable antipsychotic drug responses.

Other than the D_2 receptor, most antipsychotics also shown similar affinity for the dopamine D_3 receptor(Sokoloff et al., 2006), leading to pharmacogenetic studies of variants on the *DRD3* gene, which codes for the D_3 receptors. Much of the research were focused

on the SNP(rs6280) coding for the serine to glycine substitution at amino acid position 9 in the N-terminal extracellular domain of the D_3 receptor protein. It was suggested that the dopamine has 4-5 times higher affinity to the glycine-9 variant when compared to the serine-9 variant(Jeanneteau et al., 2006), thus it was hypothesized that the serine to glycine substitution might modulate the antipsychotic drug response. Interestingly, it was found that the serine allele was associated with better response to FGAs but was associated with non-response to clozapine treatment(Jian-Ping Zhang and Anil K Malhotra, 2011). However, this finding was not replicated and there was yet any consistent evidence of the association of the serine to glycine substitution with antipsychotic response(Jian-Ping Zhang and Anil K Malhotra, 2011).

Serontonin Receptors

Serontonin receptors, especially the 5-HT_{2A} receptors first gain attention because of it critical involvement in the pathophysiology of hallucinations(Aghajanian and Marek, 1999), leading to speculation of its role in the etiology of schizophrenia where hallucinations is one of the main symptoms. Although there are debates on the importance of serontonin receptors in antipsychotic drug responses(Shitij Kapur and Mamo, 2003), pharmacogenetic studies on serotonin receptors such as the 5-HT_{2A} receptors remains popular.

Polymorphisms on *HTR2A* gene, which codes for the 5-HT_{2A} receptors, were extensively studied. The synonymous SNP at codon 10(T102C,rs6313) and the A-1438G SNP(rs6311) in the promoter region of *HTR2A* are in complete LD. It was found that the C allele of the T102C SNP, together with the G allele of the A-1438G SNP might cause lower promoter activities of *HTR2A* and may decrease the 5-HT_{2A} densities in some brain areas(Jian-Ping Zhang and Anil K Malhotra, 2011). However, results from studies on the association of T102C and A-1438G have not reach an agreement(Jian-Ping Zhang and Anil K Malhotra, 2011).

Cytochrome P450 enzymes

There are many other factors that might affect the antipsychotic drug response. For example, the time course of the absorption, the bioavailability, the distribution of the drug in the body, the excretion of the drugs and the metabolism of the drugs all influences the efficacy of antipsychotics. Genetic variants in enzymes mediating these factors are therefore interesting target for pharmacogenetic studies.

The Cytochrome P450 enzyme family, including CYP1A1, CYP2A6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A5 and many others, in the liver is one of the major target of pharmacogenetic studies because of its role in metabolizing many of the antipsychotic drugs(Cacabelos, Hashimoto, and Takeda, 2011). Around 40% of antipsychotics are major substrate for CYP2D6(Cacabelos, Hashimoto, and Takeda, 2011) making it an ideal target to study. There are more than 100 genetic variations observed on the *CYP2D6* gene and by combining different alleles, the CYP2D enzyme can be categorized based on the degrees of the enzymatic activities: poor metabolizer, intermediate metabolizer, extensive metabolizer(normal) and ultra-rapid metabolizer(Jian-Ping Zhang and Anil K Malhotra, 2011). It was hypothesized that individuals' CYP2D enzymatic activities can affect the level of drugs in their blood. For example, people with *CYP2D* alleles from the poor metabolizer categories were expected to have a higher drug levels in the blood when compared to people with *CYP2D* alleles from the ultra-rapid metabolizer categories.

Although there were data suggesting the association of poor metabolizer with higher rate of drug induced side effects(Ravyn et al., 2013), most studies to date have been unable to provide sufficient evidence to support the use of Cytochrome 450 genotype testing to improve therapeutic efficacy in the use of antipsychotic medications(Ravyn et al., 2013). However, Ravyn et al. (2013) do agree that the use of cytochrome 450 genotype testing might help to prevent adverse side effects in patients receiving some antipsychotics such as Risperidone and Aripiprazole.

Genome Wide Association of Antipsychotic Drug response

Despite the usefulness of the candidate gene approach, it was restricted by our limited knowledge regarding the mechanism behind antipsychotic response. With the popularization of GWAS and advancement of sequencing technology, we now have the ability to perform association on variants across the whole human genome, allowing a hypothesis free approach.

The CATIE project conducted a total of four GWAS, on phenotype such as antipsychotic treatment response(McClay et al., 2011), antipsychotic-induced Parkinsonism(Alkelai et al., 2009), movement related adverse antipsychotic effect(Aberg et al., 2010) and metabolic side effects(Adkins et al., 2011). For the study of antipsychotic treatment response(McClay et al., 2011), a total of 738 subjects from the CATIE project, each from different ethnic background, were genotyped. Principle component analysis (PCA) were performed to control for subtle and extensive variation due to both genomic and experimental features. Based on Oord et al. (2009), it was assumed that it takes on average about 30 days for treatment to

exert an effect. Therefore the total PANSS score change within a 30 days period, along with change of the five scale PANSS, including positive symptoms, negative symptoms, disorganization symptoms, excitement and emotional distress within a 30 days period were used to represent the treatment effect. Because of variation in efficacy for different antipsychotic drugs, the treatment effect of the five antipsychotic used(olanzapine, quetiapine, isperidone, ziprasidone and perphenazine) were estimated independently. In total, there were 30 PANSS outcome measured (5 drugs $\times$ 6 PANSS scales) and were associated with the genotypes.

Unfortunately, none of the SNPs passed through the genome wide significance threshold(p-value$\leq 5 \times 10^{-8}$). When considering the false discovery rate (FDR) instead, rs17390445 was found to be significantly associated with change in positive symptoms score when Ziprasidone were administrated(q-value= 0.049). The rs17390445 is located in the intergenic region of chromosome 4p15 and does not associate with any genes. On the other hand, SNP in the Ankrin Repeat and Sterile Alpha Motif Domain-Containing Protein 1B gene(*ANKS1B*) was found to be associated in change in negative symptoms when Olanzapine was administrated and SNP in the Contactin-Associated Protein-Like 5 gene(*CNTNAP5*) was found to be associated with change in negative symptoms when Risperidone were administrated.

Despite being the largest GWAS on antipsychotic treatment response, the sample size per drug group were relatively small ($\sim$ 150 samples per drug group) compared to other GWAS on psychiatric phenotypes. Similar to schizophrenia, it was hypothesized that the antipsychotic treatment response is more likely to be affected by rare variants with large effect or common variants with modest effect(Jorgensen and Williamson, 2008). As such, a large number of samples will be required to provide adequate power in genetic association study on antipsychotic treatment outcome. Given the current sample size of CATIE and only by assuming all antipsychotic drugs efficacy were influenced by the same genetic variant, one can at best detect a common (maf$\leq$ 0.2) genetic variant with odd ratio of 2 or above(Jorgensen and Williamson, 2008). However, the calculation in Jorgensen and Williamson (2008) did not take into account of LD and the calculated power was likely to be over-estimated. Therefore, the CATIE GWAS is likely to be under-powered and might contain large amount of false negative results.

Another problem faced by the CATIE GWAS was the large non-adherence rate. 74% of patients discontinued the study medication before the 18 months period ends(Lieberman et al., 2005) with almost 30% stopped medication because of 'patient's decision'. It was estimated that with a sample size of 400 and non-adherence rate of 50%, the power of the study will be less than 0.7 and the power might further drop to below 0.4 when the non-

adherence rate reaches 70%(A K Malhotra, J-P Zhang, and Lencz, 2012). Considering that the non-adherence rate in schizophrenia ranges from 20 to 70%(A K Malhotra, J-P Zhang, and Lencz, 2012), it was more than likely that majority of the samples were not adhered to their medication, thus decreasing the power of the study.

On the other hand, chronic schizophrenia patients were recruited for the CATIE study, which were associated with an increased duration of psychotic symptoms, increased likelihood of substance abuse, and functional/social disabilities that may influence drug response rates and confound the result of association(Jian-Ping Zhang and Anil K Malhotra, 2013). Previous treatment of antipsychotic might also confound the results for a better dose can be given to patients based on previous treatments.

Nonetheless, although no genome-wide significant SNP was identified in the CATIE GWAS, a number of SNP were marginally significant. By increasing the study power, we might start to identify some of the genetic variants that are associated with antipsychotic treatment response. With the increased knowledge in schizophrenia with the success of PGC, we might soon be welcoming a better clinical application of the genetic data in treatment of schizophrenia, helping the schizophrenia patients to have a better quality of life.

Chapter 5

Conclusion

Bibliography

Aberg, Karolina et al. (2010). "Genomewide association study of movement-related adverse antipsychotic effects." eng. In: *Biological psychiatry* 67.3, pp. 279–282. DOI: 10.1016/j.biopsych.2009.08.036 (cit. on p. 29).

Adkins, D E et al. (2011). "Genomewide pharmacogenomic study of metabolic side effects to antipsychotic drugs." eng. In: *Molecular psychiatry* 16.3, pp. 321–332. DOI: 10.1038/mp.2010.14 (cit. on p. 29).

Aghajanian, G K and G J Marek (1999). "Serotonin and hallucinogens." eng. In: *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* 21.2 Suppl, 16S–23S. DOI: 10.1016/S0893-133X(98)00135-3 (cit. on p. 28).

Alkelai, Ana et al. (2009). "Genome-wide association study of antipsychotic-induced parkinsonism severity among schizophrenia patients." eng. In: *Psychopharmacology* 206.3, pp. 491–499. DOI: 10.1007/s00213-009-1627-z (cit. on p. 29).

Alvir, J M et al. (1993). "Clozapine-induced agranulocytosis. Incidence and risk factors in the United States." eng. In: *The New England journal of medicine* 329.3, pp. 162–167. DOI: 10.1056/NEJM199307153290303 (cit. on p. 24).

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, p. 991. DOI: 10.1176/appi.books.9780890425596.744053 (cit. on pp. 1, 2).

Arinami, T et al. (1997). "A functional polymorphism in the promoter region of the dopamine D2 receptor gene is associated with schizophrenia." eng. In: *Human molecular genetics* 6.4, pp. 577–582 (cit. on p. 27).

Arranz, M J and J de Leon (2007). "Pharmacogenetics and pharmacogenomics of schizophrenia: a review of last decade of research". In: *Mol Psychiatry* 12.8, pp. 707–747 (cit. on pp. 24, 25).

Arranz, Maria J and Janet C Munro (2011). "Toward understanding genetic risk for differential antipsychotic response in individuals with schizophrenia". In: *Expert Review of Clinical Pharmacology* 4.3, pp. 389–405. DOI: 10.1586/ecp.11.16 (cit. on p. 26).

Bernstein, Bradley E et al. (2010). "The NIH Roadmap Epigenomics Mapping Consortium." eng. In: *Nature biotechnology* 28.10, pp. 1045–1048. DOI: 10.1038/nbt1010-1045 (cit. on p. 20).

Bouchard, Thomas J (2013). "The Wilson Effect: the increase in heritability of IQ with age." In: *Twin research and human genetics : the official journal of the International Society for Twin Studies* 16.5, pp. 923–30. DOI: 10.1017/thg.2013.54 (cit. on pp. 6, 7).

Brown, A S and E J Derkits (2010). "Prenatal infection and schizophrenia: a review of epidemiologic and translational studies". eng. In: *Am J Psychiatry* 167.3, pp. 261–280. DOI: appi.ajp.2009.09030361[pii]10.1176/appi.ajp.2009.09030361 (cit. on pp. 3, 5).

Buckley, Peter F et al. (2009). "Psychiatric Comorbidities and Schizophrenia". In: *Schizophrenia Bulletin* 35.2, pp. 383–402. DOI: 10.1093/schbul/sbn135 (cit. on pp. 22, 23).

Bulik-Sullivan, Brendan K et al. (2015). "LD Score regression distinguishes confounding from polygenicity in genome-wide association studies". In: *Nature Genetics* 47.3, pp. 291–295. DOI: 10.1038/ng.3211 (cit. on pp. 18, 19).

Bulik-Sullivan, Brendan et al. (2015). "An atlas of genetic correlations across human diseases and traits". In: *Nat Genet* advance online publication (cit. on p. 22).

Cacabelos, Ramon, Ryota Hashimoto, and Masatoshi Takeda (2011). "Pharmacogenomics of antipsychotics efficacy for schizophrenia." eng. In: *Psychiatry and clinical neurosciences* 65.1, pp. 3–19. DOI: 10.1111/j.1440-1819.2010.02168.x (cit. on p. 29).

Consortium, The International HapMap (2005). "A haplotype map of the human genome". In: *Nature* 437, pp. 1299–1320. DOI: 10.1038/nature04226 (cit. on p. 14).

Deverman, B E and P H Patterson (2009). "Cytokines and CNS development". eng. In: *Neuron* 64.1, pp. 61–78. DOI: 10.1016/j.neuron.2009.09.002S0896-6273(09)00680-1[pii] (cit. on p. 15).

ENCODE Project Consortium (2012). "An integrated encyclopedia of DNA elements in the human genome". In: *Nature* 489.7414, pp. 57–74 (cit. on p. 20).

Falconer, Douglas S (1965). "The inheritance of liability to certain diseases, estimated from the incidence among relatives". In: *Annals of Human Genetics* 29.1, pp. 51–76. DOI: 10.1111/j.1469-1809.1965.tb00500.x (cit. on p. 9).

Falconer, Douglas S and Trudy F C Mackay (1996). *Introduction to Quantitative Genetics (4th Edition)*. Vol. 12, p. 464 (cit. on pp. 6, 8, 11).

Feuk, Lars, Andrew R Carson, and Stephen W Scherer (2006). "Structural variation in the human genome". In: *Nat Rev Genet* 7.2, pp. 85–97 (cit. on pp. 15, 17).

Finucane, Hilary K et al. (2015). "Partitioning heritability by functional annotation using genome-wide association summary statistics". In: *Nat Genet* advance online publication (cit. on pp. 20, 22).

Gardner, E L, L S Walker, and W Paredes (1993). "Clozapine's functional mesolimbic selectivity is not duplicated by the addition of anticholinergic action to haloperidol: a brain stimulation study in the rat." eng. In: *Psychopharmacology* 110.1-2, pp. 119–124 (cit. on p. 25).

Giovanoli, S. et al. (2013). "Stress in puberty unmasks latent neuropathological consequences of prenatal immune activation in mice". eng. In: *Science* 339.6123, pp. 1095–1099. DOI: 10.1126/science.1228261339/6123/1095[pii] (cit. on p. 4).

Gottesman, I I and J Shields (1967). "A polygenic theory of schizophrenia". In: *Proceedings of the National Academy of Sciences* 58.1, pp. 199–205 (cit. on pp. 11, 12, 19).

Gottesman, II (1991). *Schizophrenia genesis: The origins of madness.* WH Freeman/Times Books/Henry Holt & Co (cit. on p. 12).

Gottesman, II and J Shields (1967). "A polygenic theory of schizophrenia". In: *Proceedings of the National Academy of Sciences* 58.1, pp. 199–205 (cit. on p. 12).

Gottesman, Irving I and James Shields (1982). *Schizophrenia: The Epigenetic Puzzle*. Cambridge University Press (cit. on pp. 5, 12).

Harrison, P J and D R Weinberger (2005). "Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence." In: *Molecular psychiatry* 10.1, 40–68, image 5. DOI: 10.1038/sj.mp.4001686 (cit. on p. 13).

Harvey, Philip D. et al. (2012). "Diagnosis of schizophrenia: Consistency across information sources and stability of the condition". In: *Schizophrenia Research* 140.1-3, pp. 9–14. DOI: 10.1016/j.schres.2012.03.026 (cit. on p. 2).

Heston, Leonard L (1966). "Psychiatric Disorders in Foster Home Reared Children of Schizophrenic Mothers". In: *The British Journal of Psychiatry* 112.489, pp. 819–825 (cit. on p. 10).

Howes, O D and S Kapur (2009). "The dopamine hypothesis of schizophrenia: version III–the final common pathway". eng. In: *Schizophr Bull* 35.3, pp. 549–562. DOI: 10.1093/schbul/sbp006sbp006[pii] (cit. on p. 25).

Jablensky, Assen (2010). "The diagnostic concept of schizophrenia: its history, evolution, and future prospects." In: *Dialogues in clinical neuroscience* 12.3, pp. 271–87 (cit. on p. 2).

Jeanneteau, Freddy et al. (2006). "A functional variant of the dopamine D3 receptor is associated with risk and age-at-onset of essential tremor." eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 103.28, pp. 10753–10758. DOI: 10.1073/pnas.0508189103 (cit. on p. 28).

Jones, Peter B et al. (2006). "Randomized controlled trial of the effect on Quality of Life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1)." eng. In: *Archives of general psychiatry* 63.10, pp. 1079–1087. DOI: 10.1001/archpsyc.63.10.1079 (cit. on p. 25).

Jorgensen, Andrea L and Paula R Williamson (2008). "Methodological quality of pharmacogenetic studies: Issues of concern". In: *Statistics in Medicine* 27.30, pp. 6547–6569. DOI: 10.1002/sim.3420 (cit. on p. 30).

Kapur, Shitij and David Mamo (2003). "Half a century of antipsychotics and still a central role for dopamine D2 receptors." eng. In: *Progress in neuro-psychopharmacology & biological psychiatry* 27.7, pp. 1081–1090. DOI: 10.1016/j.pnpbp.2003.09.004 (cit. on pp. 25, 27, 28).

Kay, Stanley R, Abraham Fiszbein, and Lewis A Opler (1987). "The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia". In: *Schizophrenia Bulletin* 13.2, pp. 261–276. DOI: 10.1093/schbul/13.2.261 (cit. on p. 26).

Kelly, C and R G McCreadie (1999). "Smoking habits, current symptoms, and premorbid characteristics of schizophrenic patients in Nithsdale, Scotland." eng. In: *The American journal of psychiatry* 156.11, pp. 1751–1757 (cit. on p. 5).

Knapp, Martin, Roshni Mangalore, and Judit Simon (2004). "The global costs of schizophrenia." In: *Schizophrenia bulletin* 30.2, pp. 279–293 (cit. on p. 1).

Kumar, Subodh and Suprakash Chaudhury (2014). "Efficacy of amisulpride and olanzapine for negative symptoms and cognitive impairments: An open-label clinical study". In: *Industrial Psychiatry Journal* 23.1, pp. 27–35. DOI: 10.4103/0972-6748.144953 (cit. on p. 25).

Lander, E S et al. (2001). "Initial sequencing and analysis of the human genome." eng. In: *Nature* 409.6822, pp. 860–921. DOI: 10.1038/35057062 (cit. on p. 13).

Lehmann, H E and T A Ban (1997). "The history of the psychopharmacology of schizophrenia." eng. In: *Canadian journal of psychiatry. Revue canadienne de psychiatrie* 42.2, pp. 152–162 (cit. on pp. 23, 24).

Li, Miao-Xin Xin et al. (2011). "Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets". In: *Human Genetics* 131.5, pp. 747–756. DOI: 10.1007/s00439-011-1118-2 (cit. on p. 14).

Lichtenstein, Paul et al. (2009). "Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study". In: *The Lancet* 373.9659, pp. 234–239. DOI: 10.1016/S0140-6736(09)60072-6 (cit. on pp. 12, 19, 23).

Lieberman, Jeffrey A et al. (2005). "Effectiveness of Antipsychotic Drugs in Patients with Chronic Schizophrenia". In: *New England Journal of Medicine* 353.12, pp. 1209–1223. DOI: 10.1056/NEJMoa051688 (cit. on pp. 25, 30).

Malhotra, A K, J-P Zhang, and T Lencz (2012). "Pharmacogenetics in psychiatry: translating research into clinical practice". In: *Mol Psychiatry* 17.8, pp. 760–769 (cit. on p. 31).

Mata, I et al. (2001). "Olanzapine: concordant response in monozygotic twins with schizophrenia". In: *The British Journal of Psychiatry* 178.1, p. 86 (cit. on p. 27).

McClay, J L et al. (2011). "Genome-wide pharmacogenomic analysis of response to treatment with antipsychotics." eng. In: *Molecular psychiatry* 16.1, pp. 76–85. DOI: 10.1038/mp.2009.89 (cit. on p. 29).

McClellan, Jon M, Ezra Susser, and Mary-Claire King (2007). "Schizophrenia: a common disease caused by multiple rare alleles". In: *The British Journal of Psychiatry* 190.3, pp. 194–199 (cit. on p. 13).

McGrath, John et al. (2008). "Schizophrenia: A Concise Overview of Incidence, Prevalence, and Mortality". In: *Epidemiologic Reviews* 30.1, pp. 67–76. DOI: 10.1093/epirev/mxn001 (cit. on p. 5).

Mednick (1958). "Schizophrenia Following Prenatal Exposure to an Influenza Epidemic". In: 1111.1 (cit. on p. 3).

Meltzer, Herbert Y (1991). "The Mechanism of Action of Novel Antipsychotic Drugs". In: *Schizophrenia Bulletin* 17.2, pp. 263–287. DOI: 10.1093/schbul/17.2.263 (cit. on p. 24).

— (1999). "The Role of Serotonin in Antipsychotic Drug Action". In: *Neuropsychopharmacology* 21.S1, 106S–115S (cit. on p. 25).

Meyer, U, J Feldon, and B K Yee (2009). "A review of the fetal brain cytokine imbalance hypothesis of schizophrenia". eng. In: *Schizophr Bull* 35.5, pp. 959–972. DOI: 10.1093/schbul/sbn022sbn022[pii] (cit. on p. 15).

Meyer, U, B K Yee, and J Feldon (2007). "The neurodevelopmental impact of prenatal infections at different times of pregnancy: the earlier the worse?" eng. In: *Neuroscientist* 13.3, pp. 241–256. DOI: 13/3/241[pii]10.1177/1073858406296401 (cit. on pp. 3, 4).

Meyer, Urs, Joram Feldon, and S Hossein Fatemi (2009). "In-vivo rodent models for the experimental investigation of prenatal immune activation effects in neurodevelopmental brain disorders". In: *Neuroscience & Biobehavioral Reviews* 33.7, pp. 1061–1079. DOI: http://dx.doi.org/10.1016/j.neubiorev.2009.05.001 (cit. on p. 3).

Müller, Norbert and Markus J Schwarz (2010). "Immune System and Schizophrenia". In: *Current immunology reviews* 6.3, pp. 213–220 (cit. on p. 5).

Onore, C E et al. (2014). "Maternal immune activation leads to activated inflammatory macrophages in offspring". eng. In: *Brain Behav Immun* 38, pp. 220–226. DOI: 10.1016/j.bbi.2014.02.007S0889-1591(14)00053-1[pii] (cit. on p. 4).

Oord, Edwin J C G van den et al. (2009). "A systematic method for estimating individual responses to treatment with antipsychotics in CATIE". In: *Schizophrenia Research* 107.1, pp. 13–21. DOI: http://dx.doi.org/10.1016/j.schres.2008.09.009 (cit. on p. 29).

Ording, Anne Gulbech et al. (2013). "Comorbid Diseases Interact with Breast Cancer to Affect Mortality in the First Year after Diagnosis—A Danish Nationwide Matched Cohort Study". In: *PLoS ONE* 8.10, e76013 (cit. on p. 22).

Paolicelli, R C et al. (2011). "Synaptic pruning by microglia is necessary for normal brain development". eng. In: *Science* 333.6048, pp. 1456–1458. DOI: 10.1126/science.1202529science.1202529[pii] (cit. on p. 4).

Pirmohamed, Munir (2001). "Pharmacogenetics and pharmacogenomics". In: *British Journal of Clinical Pharmacology* 52.4, pp. 345–347. DOI: 10.1046/j.0306-5251.2001.01498.x (cit. on p. 27).

Purcell, S, S S Cherny, and P C Sham (2003). "Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits". en. In: *Bioinformatics* 19, pp. 149–150. DOI: 10.1093/bioinformatics/19.1.149 (cit. on p. 14).

Purcell, Shaun M et al. (2009). "Common polygenic variation contributes to risk of schizophrenia and bipolar disorder." eng. In: *Nature* 460.7256, pp. 748–752. DOI: 10.1038/nature08185 (cit. on p. 23).

Ravyn, Dana et al. (2013). "CYP450 Pharmacogenetic treatment strategies for antipsychotics: A review of the evidence". In: *Schizophrenia Research* 149.1–3, pp. 1–14. DOI: http://dx.doi.org/10.1016/j.schres.2013.06.035 (cit. on p. 29).

Remington, Gary et al. (2013). "Clozapine's role in the treatment of first-episode schizophrenia." eng. In: *The American journal of psychiatry* 170.2, pp. 146–151. DOI: 10.1176/appi.ajp.2012.12060778 (cit. on p. 24).

Rijsdijk, Fruhling V and Pak C Sham (2002). "Analytic approaches to twin data using structural equation models." eng. In: *Briefings in bioinformatics* 3.2, pp. 119–133 (cit. on p. 11).

Riley, Brien and Kenneth S Kendler (2006). "Molecular genetic studies of schizophrenia." In: *European journal of human genetics : EJHG* 14.6, pp. 669–680. DOI: 10.1038/sj.ejhg.5201571 (cit. on p. 12).

Ripke, Stephan et al. (2014). "Biological insights from 108 schizophrenia-associated genetic loci". In: *Nature* 511, pp. 421–427. DOI: 10.1038/nature13595 (cit. on pp. 15, 16, 19, 20).

Ripke, S et al. (2013). "Genome-wide association analysis identifies 13 new risk loci for schizophrenia". eng. In: *Nat Genet* 45.10, pp. 1150–1159. DOI: 10.1038/ng.2742ng.2742[pii] (cit. on p. 20).

Risch, N (1990a). "Linkage strategies for genetically complex traits. I. Multilocus models." In: *American Journal of Human Genetics* 46.2, pp. 222–228 (cit. on pp. 12, 13).

— (1990b). "Linkage strategies for genetically complex traits. II. The power of affected relative pairs." In: *American Journal of Human Genetics* 46.2, pp. 229–241 (cit. on p. 13).

Saha, Sukanta, David Chant, and John Mcgrath (2007). "A Systematic Review of Mortality in Schizophrenia". In: *Archives of general psychiatry* 64.10, pp. 1123–1131. DOI: 10.1001/archpsyc.64.10.1123 (cit. on p. 1).

Schultz, Stephen H., Stephen W. North, and Cleveland G. Shields (2007). "Schizophrenia: A review". In: *American Family Physician* 75.12, pp. 1821–1829 (cit. on p. 1).

Smith, S E et al. (2007). "Maternal immune activation alters fetal brain development through interleukin-6". eng. In: *J Neurosci* 27.40, pp. 10695–10702. DOI: 27/40/10695[pii]10.1523/JNEUROSCI.2178-07.2007 (cit. on p. 3).

Søgaard, Mette et al. (2013). "The impact of comorbidity on cancer survival: a review". In: *Clinical Epidemiology* 5.Suppl 1, pp. 3–29. DOI: 10.2147/CLEP.S47150 (cit. on p. 22).

Sokoloff, P et al. (2006). "The dopamine D3 receptor: a therapeutic target for the treatment of neuropsychiatric disorders." eng. In: *CNS & neurological disorders drug targets* 5.1, pp. 25–43 (cit. on p. 27).

Sullivan, Patrick F (2005). "The Genetics of Schizophrenia". In: *PLoS Med* 2.7, e212 (cit. on p. 5).

Sullivan, Patrick F, Kenneth S Kendler, and Michael C Neale (2003). "Schizophrenia as a Complex Trait". In: *Archives of general psychiatry* 60, pp. 1187–1192. DOI: 10.1001/archpsyc.60.12.1187 (cit. on pp. 11, 12, 19).

Szatkiewicz, J P et al. (2014). "Copy number variation in schizophrenia in Sweden". In: *Mol Psychiatry* 19.7, pp. 762–773 (cit. on pp. 17, 19).

Talkowski, Michael E et al. (2007). "Dopamine Genes and Schizophrenia: Case Closed or Evidence Pending?" In: *Schizophrenia Bulletin* 33.5, pp. 1071–1081 (cit. on p. 15).

Tandon, Rajiv (2007). *Antipsychotic treatment of schizophrenia: two steps forward, one step back*. eng (cit. on p. 24).

Tienari, Pekka et al. (2004). "Genotype-environment interaction in schizophrenia-spectrum disorder". In: *The British Journal of Psychiatry* 184.3, pp. 216–222 (cit. on p. 19).

Tsuang, Ming T., William S. Stone, and Stephen V. Faraone (2000). "Toward reformulating the diagnosis of schizophrenia". In: *American Journal of Psychiatry* 157.7, pp. 1041–1050. DOI: 10.1176/appi.ajp.157.7.1041 (cit. on p. 2).

Üçok, A L P and Wolfgang Gaebel (2008). "Side effects of atypical antipsychotics: a brief overview". In: *World Psychiatry* 7.1, pp. 58–62 (cit. on p. 24).

Visscher, Peter M, William G Hill, and Naomi R Wray (2008). "Heritability in the genomics era [mdash] concepts and misconceptions". In: *Nat Rev Genet* 9.4, pp. 255–266 (cit. on pp. 7, 9).

Vojvoda, Dolores et al. (1996). "Monozygotic twins concordant for response to clozapine". In: *The Lancet* 347.8993, p. 61. DOI: http://dx.doi.org/10.1016/S0140-6736(96)91594-9 (cit. on p. 27).

Vuillermot, Stéphanie et al. (2010). "A longitudinal examination of the neurodevelopmental impact of prenatal immune activation in mice reveals primary defects in dopaminergic development relevant to schizophrenia". eng. In: *J Neurosci* 30.4, pp. 1270–1287. DOI: 10.1523/JNEUROSCI.5408-09.201030/4/1270[pii] (cit. on p. 3).

World Health Organization (2013). *WHO methods and data sources for global burden of disease estimates*. Tech. rep. Geneva (cit. on p. 2).

Yang, Jian et al. (2011). "Genomic inflation factors under polygenic inheritance". In: *Eur J Hum Genet* 19.7, pp. 807–812 (cit. on p. 18).

Yang, J et al. (2011). "GCTA: a tool for genome-wide complex trait analysis". eng. In: *Am J Hum Genet* 88.1, pp. 76–82. DOI: 10.1016/j.ajhg.2010.11.011S0002-9297(10)00598-7[pii] (cit. on p. 17).

Zhang, Jian-Ping and Anil K Malhotra (2011). "Pharmacogenetics and antipsychotics: therapeutic efficacy and side effects prediction." eng. In: *Expert opinion on drug metabolism & toxicology* 7.1, pp. 9–37. DOI: 10.1517/17425255.2011.532787 (cit. on pp. 24, 28, 29).

— (2013). "Pharmacogenetics of antipsychotics: recent progress and methodological issues." eng. In: *Expert opinion on drug metabolism & toxicology* 9.2, pp. 183–191. DOI: 10.1517/17425255.2013.736964 (cit. on p. 31).

Zhao, B and J P Schwartz (1998). "Involvement of cytokines in normal CNS development and neurological diseases: recent progress and perspectives". eng. In: *J Neurosci Res* 52.1, pp. 7–16. DOI: 10.1002/(SICI)1097-4547(19980401)52:1<7::AID-JNR2>3.0.CO;2-I[pii] (cit. on p. 15).

Zheng, Gang, Boris Freidlin, and Joseph L Gastwirth (2006). "Robust genomic control for association studies." eng. In: *American journal of human genetics* 78.2, pp. 350–356. DOI: 10.1086/500054 (cit. on p. 18).

Supplementary Materials

Appendix