Identification of Epigenetic Modifications Following Treatment with Olanzapine

By

Jonelle Dickow Villar

This thesis is submitted in partial fulfilment of the requirements for the degree of

Master in Molecular Biology

Department of Biological Sciences

University of Bergen

November 2018

Acknowledgements

I am grateful to the Department of Molecular Biology for the opportunity to realize a dream. It has been a long time coming! Thank you to Rune Male and Lill Knudsen for their support during this adventure.

Thank you to Vidar Steen and my co-supervisor Stéphanie Le Hellard for opening the doors to a very wonderful year of learning at Dr. Einar Martins Research Group in Biological Psychiatry. Not only was I included in lab meetings and seminars, but I also had the opportunity to attend the Bergen Early Psychosis Symposium, the NORMENT Young Researchers Meeting, and the 26th World Congress of Psychiatric Genetics. This year has been full of activity, and I have loved every day that I walked into the lab.

Thank you for the support from SFF NORMENT and the K.B. Jebsen Center for Psychosis Research. All my admiration goes to the broader NORMENT community of researchers for their inspiring work. I also extend my warmest appreciation to all my colleagues at Dr. Einar Martins Research Group. The atmosphere of warm collegiality has really been quite special.

A very special thank you to my main supervisor Anne-Kristin Stavrum for guiding me through the challenging task of using the R-tools required to do this work. I appreciated the careful thought process she showed when evaluating our work, and for encouraging me to do the same. Fortunately, there was good laughter, too.

Thank you to my co-supervisor Kari Ersland for her careful comments to the thesis. I learned to think more clearly about our work while reading her corrections. Her consideration for my questions, especially during the concluding writing phase was greatly appreciated.

Many thanks to Carla P.D. Fernandes for her assistance when preparing the DNA samples to send to Bonn. Her enthusiasm for lab work and for sharing ideas with others is infectious. My thanks to Tatiana Polushina for her encouragement and thoughtful explanations to my statistical questions.

Thank you to Johan Fernø and Martha Haugstøyl with the Hormone Laboratory Research Group. Martha contributed her time and expertise during the cell work and gene analysis.

I am deeply indebted to my dear brother James whose love of science has inspired me for many years. My dearest daughter Liliana is my greatest support and motivator. I look forward to sharing her love of medicine for many years to come.

Jonelle Dickow Villar

November 2018

Abbreviations

Ari	Aripiprazole
AUC	area under the concentration-time curve
BP	biological process
Cmax	maximum drug serum concentration measured following dose 1 and before dose 2
CYP	cytochrome P450
CC	cellular component
DMPs	differentially methylated positions
DMRs	differentially methylated regions
DMSO	Dimethyl sulfoxide
eCells	estimated cell counts
EWAS	epigenome-wide association study
GO	Gene Ontology
GWAS	genome-wide association study
HepG2	human hepatic cell line G2
Limma	linear models for microarray data
MF	molecular function
NT	No treatment
Olz	Olanzapine
SEM	standard error of mean
SNP	single-nucleotide polymorphism
Sv(a)	surrogate variable (analysis)
Que	Quetiapine
Veh	vehicle

Summary

No chronic disease burdens the world more than psychiatric disorders (Collins et al., 2011), with an estimated 40% of the population in 30 European countries affected in any given year (Insel et al., 2012). Current medical treatment for schizophrenia (SCZ), bipolar disorder (BPD) and major depressive disorder (MDD) is based upon well-established antipsychotic drugs and mood stabilizers. Treatment efficacy (30-40%) (P. Lowe et al., 2017) and potentially serious side effects (Leucht et al., 2013) often challenge medication compliance, adding an additional challenge in the path towards wellness. New drugs are required to address the burden of psychotic disorders, and yet a new science is required to address the interplay between the heterogeneous nature of psychotic disorders and drug mechanisms.

Epigenetic mechanisms, particularly alterations of methylation patterns at CpG sites have been shown to alter gene expression in humans, animal models and *in vitro* cell cultures. It is believed that epigenetic modifications induced by antipsychotic drugs plays a role in therapeutic response. Identification of pathways implicated by epigenetically modified genes, including the dopaminergic pathway, for example, has enhanced our understanding of the therapeutic mechanism of the antipsychotic drug olanzapine (Melka et al., 2013).

In the current study, we aimed to identify differentially methylated regions induced by olanzapine. 82 European patients adhering to monotherapy were selected through the TOP Cohort (Thematically Organised Psychosis). Methylation data derived from blood samples was assessed genome-wide using the Illumina 850K EPIC array. The statistical model was corrected for gender and smoking. Following identification of differentially methylated positions (DMPs) in patient blood, we exposed a cultured cell line (HepG2) to verify the modifying effect of olanzapine on DNA methylation levels.

The results of our study provide evidence of differentially methylated positions and regions in the blood of patients adhering to olanzapine monotherapy. A comparison of models adjusting for cell type composition provided evidence of improved *p*-values when cell type adjustment was included in the model. This finding was in concordance with state-of-the-art epigenome-wide-association (EWA) studies. Our results showed concordance between blood and brain for two identified differentially methylated regions, including the Trio and F-actin binding protein (*TRIOBP*) shown to be relevant in schizophrenia (Nicholas J. Bradshaw et al., 2014). The pathways implicated by the differentially methylated genes showed evidence of alterations in immune pathways and the possible mediating effect of olanzapine.

Table of contents

Introduction

1.1 Psychosis and treatment of psychosis

1.1.1 Psychotic disorders

Psychosis is a heterogeneous psychiatric condition that challenges an individual's sense of self and experience with the external environment (Radua et al., 2018). This condition is characterized by two defining features: the occurrence of auditory, sensory, or visual hallucinations; and a system of fixed, false beliefs called delusions (Arciniegas, 2015). Hallucinations may occur with or without the individual's insight into the nature of the hallucination. Delusions may range from ordinary to bizarre, from ideas of persecution or grandiosity to thought control or thought broadcasting. The theme that is apparent in psychosis amongst several psychiatric diagnoses is the individual's lack of insight into the seriousness of their condition (Ibid).

Schizophrenia (SCZ) is the prototypical psychotic disorder (Figure 1.1). The burden of symptoms is divided into positive, negative, and cognitive categories that are somewhat informative of the underlying pathology and/or therapeutic response. Classic positive symptoms include hallucinations, delusions, disorganized thinking and grossly disorganized behavior. Negative symptoms are characterized by passive behavior, social withdrawal, blunted affect and psychomotor slowing. The cognitive symptoms describe a broad group of cognitive dysfunctions (Kahn et al., 2015).

Bipolar Disorder (BP), major depressive disorder (MDD) and schizophrenia share symptoms of debilitating chronicity and recurrent episodes of relapse. Individuals burdened by bipolar disorder and major depression experience episodes of severe depression, with periods of mania or hypomania occurring in bipolar disorder (Hirschfeld, 2014).

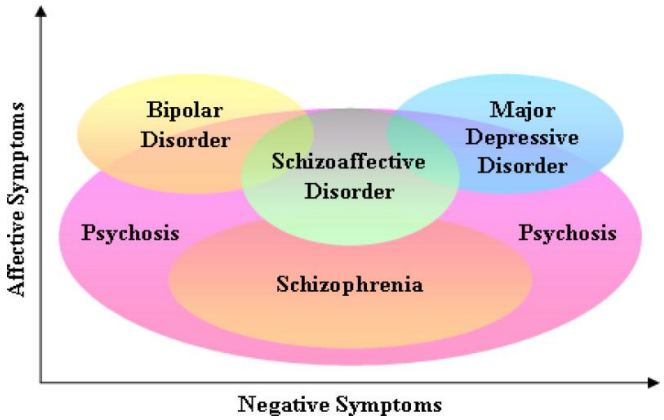


Figure 1.1 Psychotic episodes are experienced on a continuum in addition to negative and affective (mood) symptoms. Psychotic symptoms observed outside these diagnostic categories may be observed in substance abuse and neurodegenerative illness (adapted from (DeRosse & Karlsgodt, 2015).

From epidemiological studies, a number of environmental factors reflect the increased risk for psychosis across the life span: obstetric complications (Kotlicka-Antczak et al., 2017); childhood neglect and trauma (McGrath et al., 2017); poor functioning in adolescence (Fusar-poli et al., 2017); adverse life events including post-traumatic stress disorder (PTSD) (Teicher, 2018; Teicher, Samson, Anderson, & Ohashi, 2016), migration (Cardano, Scarinzi, Costa, & d'Errico, 2018; Norredam, Nellums, Nielsen, Byberg, & Petersen, 2018) and urban living (Kahn et al., 2015). Psychotic experiences are associated with self-harm, suicidal ideation and suicidal attempts (DeVylder et al., 2015; Honings et al., 2016). Sleep disturbances are common indicating a disruption in circadian rhythms (Koyanagi & Stickley, 2004; Oh et al., 2016). Gender and age play a role; with younger males at increased risk (Fusar-poli et al., 2017).

Psychotic disorders are commonly perceived as brain disorders, yet individuals burdened with schizophrenia, bipolar disorder and major depression have higher rates of physical illness, including 15-20 years shorter life expectancy than the non-afflicted (Laursen, Nordentoft, & Mortensen, 2014). Multiple corporal systems may be dysregulated in addition to the central nervous system (CNS) (Pillinger, D'Ambrosio, McCutcheon, & D Howes, 2018). Studies of first episode psychosis have shown dysfunction in cardiovascular (Christoph U Correll et al., 2017), metabolic (Greenhalgh et al., 2017), immune (Delaney et al., 2018), and hypothalamic pituitary adrenal (HPA) systems (Nordholm et al., 2018). The onset of psychosis then may be a predictor of concurrent dysregulation in the body. This evidence challenges the traditional perspective that poor health in psychosis reflects poor life style choices or a consequence of medication side effects (Pillinger et al., 2018).

The clinical boundaries between the disorders are often challenging, in particular in terms of predicting treatment outcomes or improving treatment resistance (P. Lowe et al., 2017; Remington et al., 2017). Genetic analyses have found a high degree of correlation between many of the psychiatric disorders, suggesting that current clinical boundaries are not informative on a genetic level of the underlying processes contributing to disorder onset (Anttila et al., 2018). Currently, contributions from the field of psychiatric epigenetics are providing evidence of the environmental impact, and interplay, with the human genome.

1.1.2 Antipsychotic drug therapy

Effective therapies for individuals suffering from psychotic disorders were not available before the discovery of the first antipsychotic drug in 1952. The development of chlorpromazine and the drugs that followed provided a strategy for discovering a biological basis of schizophrenia (Carpenter & Davis, 2012). Several observations about side effects suggested the mechanism of action was disruption of dopamine transmission. The major dopamine pathways in the brain are involved in motor control, and the predominant side effect initially observed was a Parkinsonian-like tremor or rigidity. Subsequently, drugs that mimicked dopamine were found to induce hallucinations.

Evidence that antipsychotic drugs selectively blocked dopamine receptors occurred in 1974 when an association between dopamine D_2 receptor inhibition and the antipsychotic potencies of the drugs was identified. These early first-generation drugs are commonly referred to as “typical” due to the typical motor “extrapyramidal” side effects. Their efficacy in treating the positive symptoms of psychosis including hallucinations contributed to the hypothesis that dopamine pathways are overactive in schizophrenia (Seeman et al., 1987).

Second-generation drugs were then developed that combined blocking D_2 - receptors with antagonism of serotonin (5-HT) receptors. These drugs were classified as “atypical” due to the reduction of motor side effects, the reduction of negative symptoms, and an improvement in mood and cognitive symptoms (Mauri et al., 2014). These therapeutic effects were associated with receptor binding to multiple serotonin (5-HT) receptors while still achieving antipsychotic effects of D_2 binding (Meltzer, Matsubara, & Lee, 1989).

Nevertheless, the pharmacology of these second-generation drugs is complex due to multiple receptor targeting (Kusumi, Boku, & Takahashi, 2015). In addition to multiple serotonin and dopamine receptors, atypical drugs target alpha-adrenergic, histaminergic, and cholinergic receptors with corresponding side effects. Sedation and weight gain is associated with histaminergic inhibition;

orthostatic hypotension with alpha-adrenergic inhibition; and dry mouth, constipation, blurred vision and tachycardia with cholinergic receptor inhibition (C U Correll, 2010).

Olanzapine, quetiapine, and aripiprazole are examples of widely used second-generation antipsychotics. Olanzapine and quetiapine have similar broad receptor binding profiles, yet quetiapine dissociates faster from the D_2 receptor producing fewer extrapyramidal symptoms (Riedel, Müller, & Strassnig, 2007). Aripiprazole differs from them both by representing a class of *partial* dopamine agonist and serotonin antagonist. It is often referred to in the literature as a third-generation antipsychotic (Tuplin & Holahan, 2017).

The binding affinities of the three antipsychotics with dissociation constant Ki are indicated in Table 1. The strength of binding affinity is associated either with therapeutic effect and/or risk of side effects. Most atypical antipsychotics have more potent 5-HT_{2A} receptor antagonism than D_2 receptor antagonism, resulting in a D_2/5-HT_{2A} ratio below 1 (Kusumi et al., 2015). As seen in Table 1, only aripiprazole has a D_2/ 5-HT_{2A} ratio below 1. Both olanzapine and quetiapine have D_2/ 5-HT_{2A} ratios above 1, with a greater side effect profile than aripiprazole.

Table 1: Receptor binding affinity of atypical second/third-generation antipsychotic drugs.

Receptor Ki (nM)	Olanzapine	Quetiapine	Aripiprazole
D_2 Antagonist	++	+	
D_2 PA			+++
D_3	++	+	+++
$5HT_{1A}$		+*	+++
$5HT_{2A}$	+++	++*	++
$5HT_{2C}$	++	+*	++
$5HT_7$	+	++*	+++
α_1	++	+++	++
M_1	++	++*	
M_3	++	++*	
H_1	+++	+++*	++
Ratio D_2/5-HT_{2A}	8.9	2.6	0.085
+	**Weak binding affinity (100>Ki<1000)**		
++	**Moderate binding affinity (10>Ki<100)**		
+++	**Strong binding affinity (1>Ki<10)**		
PA	**Partial agonist**		
***Binding property due primarily to the metabolite norquetiapine**			

Based on Stahl's Essential Psychopharmacology. 3rd edition and D_2/ 5-HT_{2A} Ki ratios (Kusumi et al., 2015).

The D_2 affinity for each drug is represented by a yellow line in Figure 1.2. Drug effects thought to explain efficacy are represented as blue bars for respective serotonin (5-HT_{1A}, 5-HT_{2A} ,5-HT_{1C}, $5HT_{2C}$)

and dopamine D_2 receptors, while negative side effects are represented as red bars for adrenergic (α1A, α1B), muscarinic (M_2) and histaminergic (H_2) receptors.

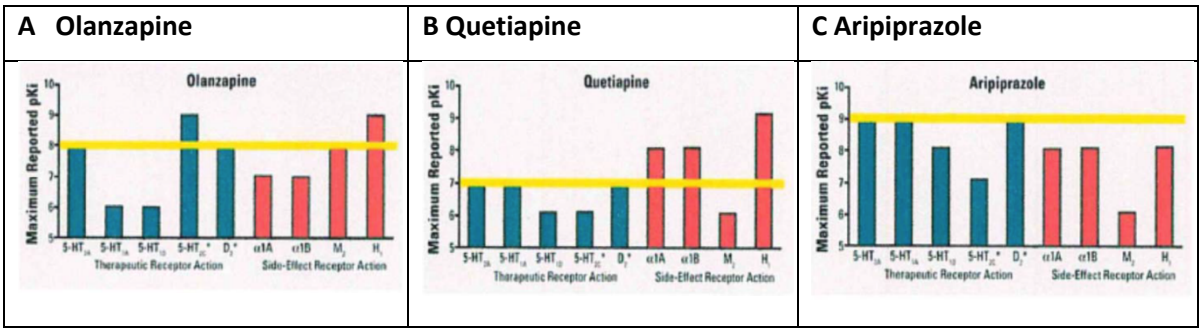


Figure 1.2 Simplified receptor binding affinity profiles: Effects may be potentially therapeutic (5-HT_{1A}, 5-HT_{2A}, 5-HT_{1C}, 5HT_{2C}, D_2) (blue) or potentially give side effects (α1A, α1B M_2 H_2) (red). 5-HT2c (*) and D2 (*) characterize both effects. The yellow line indicates the D2 affinity level for each drug (adapted from (Mauri et al., 2014).

1.1.3 Mechanisms of action of olanzapine

Olanzapine is associated with substantial weight gain, diabetes type II, lipid dysregulation and cardiovascular disease (Lambert et al., 2005). The burden of these physical health-related side effects reduces quality of life, and challenges medication compliance with consequent increased risk of psychotic relapse (Foster, Buckley, Lauriello, Looney, & Schooler, 2017). Olanzapine does show improvement for patients however in the area of cognition. Here cognitive improvement may be enhanced by an increase of prefrontal dopamine release mediated by 5-HT_{2A} receptor antagonism (Castner, Williams, & Goldman-Rakic, 2000).

Therapeutic drug action may also be attributed to the nature of the dopamine D_2 receptor which is a G protein-coupled receptor (GCR). Multiple drug effects may be mediated by intracellular signaling mechanisms and downstream effectors of GCRs including adenylate cyclase, various ion channels, phospholipases, cAMP, cAMP dependent kinase, protein kinase C (PKC), and protein lipase C (PLC) (Fribourg et al., 2011). Alterations in the expression of genes that target neurons induce changes in neuronal plasticity and synaptic remodelling (Horacek et al., 2006).

The primary metabolic pathways for olanzapine include glucuronidation and cytochrome P450 (CYP) mediated oxidation via CYP1A2 and CYP2D6 (Figure 1.3). Glucuronidation occurs mainly in the liver. The rate of oxidation and subsequent available drug serum concentration is affected by factors such as age, gender, obesity, tobacco smoking and medication. Inhibition of CYP1A2 by some antidepressants is known to significantly increase maximum (Cmax) serum concentration of

olanzapine. In addition, concurrent administration of anti-epileptics may induce CYP2D6- activity up to 50% (Soderberg & Dahl, 2013; Urichuk, Prior, Dursun, & Baker, 2008).

Serum concentrations of olanzapine are monitored in patients who smoke due to reduced metabolism of olanzapine by CYP2D. Male smokers show twice the serum concentration (AUC) of olanzapine as female smokers (T. R. Moore, Hill, & Panguluri, 2014). Tiili et al., (2015) reported an association between *CYP2D6* genotype and smoking habits. They found the risk of becoming a heavy smoker was reduced for the poor metabolizer genotype, however the risk was increased for the rapid metabolizers. Additional factors influencing olanzapine-metabolism include co-morbid medical conditions (C U Correll, 2010) and *CYP1A*/*CYP2D6* genotype (Eum, Lee, & Bishop, 2016).

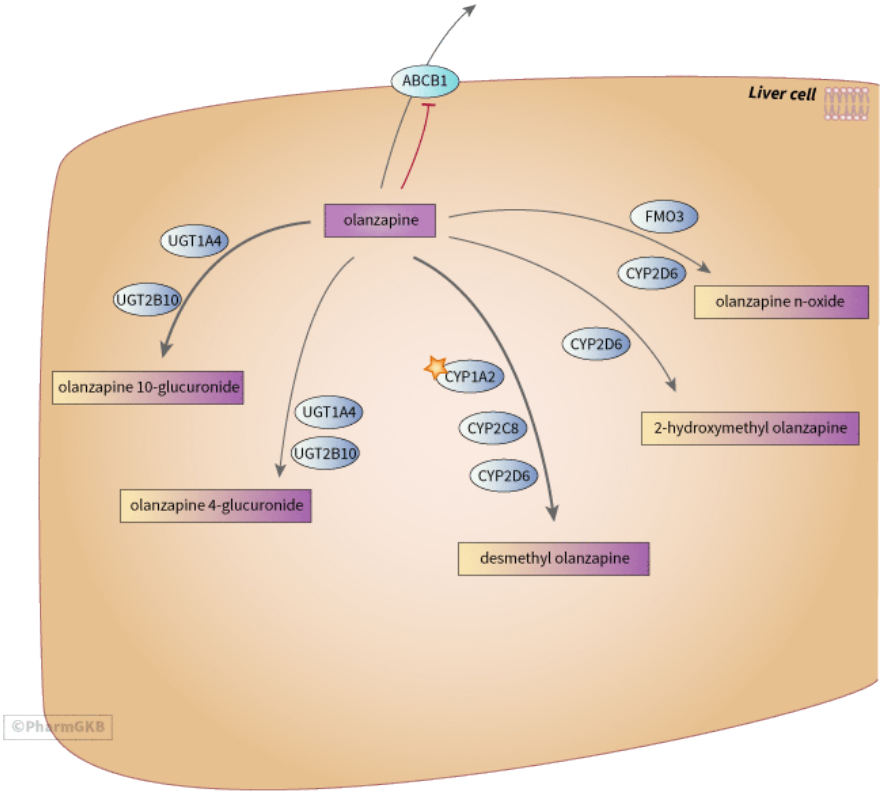


Figure 1.3 Metabolism *of* olanzapine metabolism and transport in the human liver showing candidate genes *CYP1A2* and *CYP2D6*. Metabolism of olanzapine is influenced by age, gender, tobacco smoking, concurrent medications, and illness (Urichuk et al., 2008). (Image credit: PharmaGKB, https://www.pharmgkb.org/pathway/PA166165056)

1.2 Epigenetic mechanisms

1.2.1 Three types of epigenetic modifications

DNA alterations induced by epigenetic modifications are characterized by at least three modes that may interact to regulate transcription: DNA methylation, histone modifications and microRNA (miRNA) silencing (Figure 1.4). Dysregulation of miRNAs can alter the expression of genes associated with neurodevelopment and regulate chromatin structure by targeting epigenetic factors, such as DNA methyltransferases (DNMTs) and histone deacetylases. Histones may be modified by methylation or acetylation of lysine (K), contributing either to compaction of chromatin and gene silencing, or chromatin relaxation and gene expression (Babu Swathy & Banerjee, 2017). Since it is the most accessible, DNA methylation is the most studied epigenetic mechanism and is described in the following chapter.

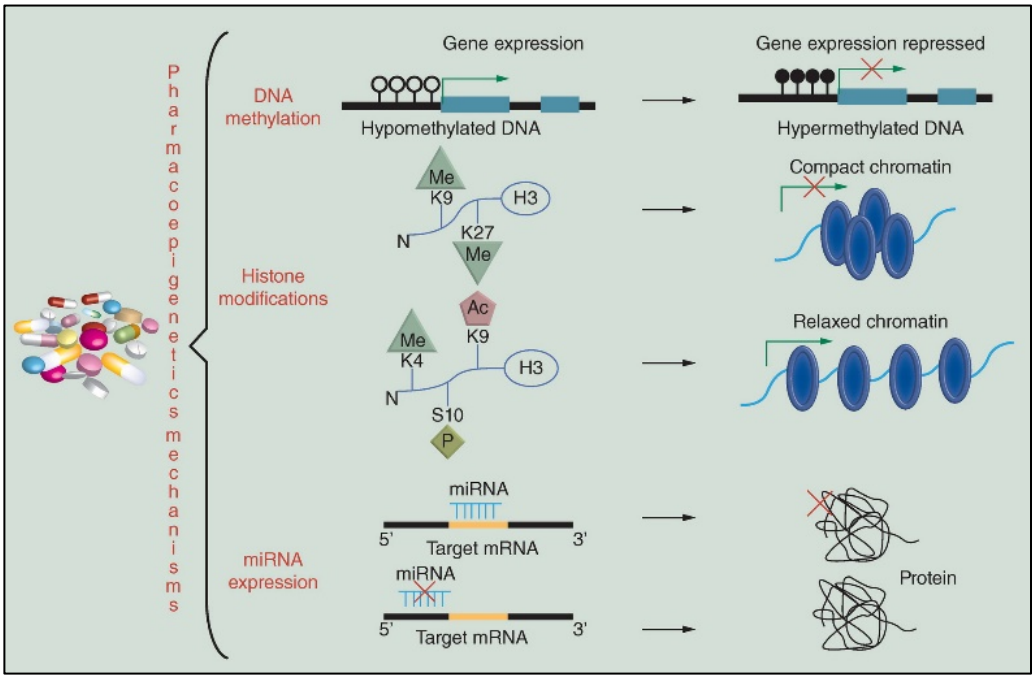


Figure 1.4 Three modes of epigenetic modulation by antipsychotic drugs. Gene expression may be altered by DNA methylation, differential histone modifications, or miRNA expression. (Image credit: (Babu Swathy & Banerjee, 2017).

1.2.2 Focus on DNA Methylation

DNA methylation is catalyzed by a family of DNA methyltransferases (DNMTs) that are essential for establishing DNA patterns in early development and throughout the lifetime. DNA methylation occurs when a methyl group from the *S*-adenyl methionine (SAM) is transferred to the fifth carbon of a cytosine residue to form 5mC. This enzymatic process occurs at genomic sites where a cytosine (C) nucleotide is followed by a guanine (G) nucleotide in a linear sequence in the 5′ → 3′ direction. This newly methylated site is referred to as a CpG and is evaluated in epigenetic studies as a marker of change within an organism (L. D. Moore, Le, & Fan, 2013).

CpG sites are spread throughout the genome and are best studied in areas called CpG islands and CpG shores (Figure 1.5-A/B). CpG islands are highly conserved stretches of DNA approximately 1000 base pairs long (Ibid). They have a higher CpG density than the rest of the genome, although not all CpG islands are methylated (Bird, Taggart, Frommer, Miller, & Macleod, 1985). Instead, the role of these unmethylated CpG islands may be to promote gene expression, enabled by transcription factor binding at GC-rich transcription start sites (Carninci et al., 2006). Areas located up to 2 kb from CpG islands have highly conserved patterns of tissue-specific methylation. These areas are called CpG shores (Irizarry, Wu, & Feinberg, 2009) (Figure 1.5-B).

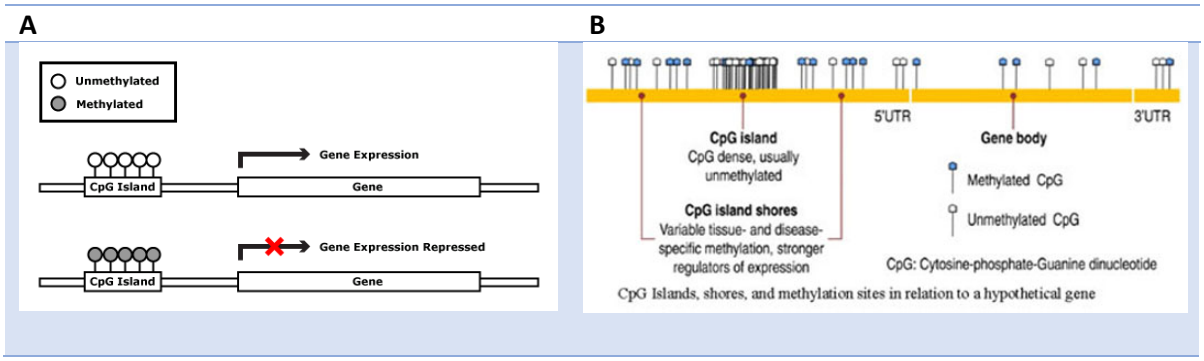


Figure 1.5: Methylation status and location of CpG islands and CpG shores: A) Methylated CpGs upstream of the promoter are associated with repression of gene expression; B) CpG islands, shores and methylation sites in relation to a hypothetical gene body. (Cartoon credits: A) UCSF School of Medicine http://missinglink.ucsf.edu/lm/genes_and_genomes/methylation.html ; B) BioSynthesis: https://www.biosyn.com/tew/inheritance-of-epigenetic-defects.aspx).

1.2.3 Differentially methylated regions

The aim of analyzing the relationship between a phenotype/status and methylation levels, is to identify positions or regions that are differentially methylated between two conditions. This measurement of methylation status may be performed between different tissues, or between cases and controls. Single changes on the CpGs are often referred to as differentially methylated positions (DMPs), while differentially methylated regions (DMRs) have a statistically different DNA methylation pattern between several CpGs in a region (Rakyan, Down, Balding, & Beck, 2011). DMRs may occur throughout the genome, but have been identified particularly around gene promoters and at intergenic regulatory regions (Suzuki & Bird, 2008). DMRs provide more information for biological interpretation due to their size, ranging in length from a few hundred to a few thousand base pairs (Rakyan et al., 2011). Of note, identification of DMRs is more robust and likely to be replicated, given the greater statistical power achieved when a dense region of CpGs is measured as a whole (Robinson et al., 2014; Ziller et al., 2013).

1.3 Epigenetics in psychiatric disorders

A complete understanding of epigenetic mechanisms in the pathogenesis of psychiatric disorders has not yet been detailed, in particular the complex interactions between multiple epigenetic mechanisms, genetic phenotypes, and the environment (Nestler, Peña, Kundakovic, Mitchell, & Akbarian, 2016). A current understanding however provides evidence of the role of neurodevelopmental pathways. The regulation of gene function has a crucial role in neurodevelopment and mediates complex processes involved in brain growth, synaptic plasticity, learning, memory and circadian rhythms (Ovenden, McGregor, Emsley, & Warnich, 2018) (Figure 1.6). Epigenetic mechanisms involved in the dysregulation of genes in these pathways have been shown to be a key determinant in the development of major psychosis (Labrie, Pai, & Petronis, 2012), specifically in bipolar disorder (Fries et al., 2016) and schizophrenia (Mill et al., 2008)

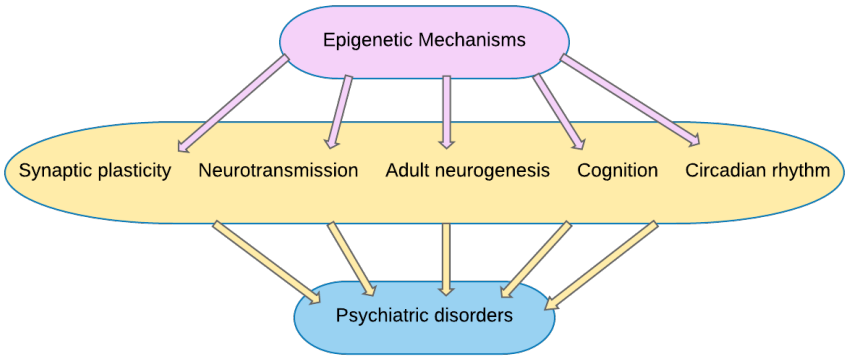


Figure 1.6 Epigenetic mechanisms influence neural mechanisms in psychiatric disorders: DNA methylation, histone modifications and noncoding RNAs are involved in the dysregulation of neural pathways contributing to the pathogenesis of these disorders (adapted from (Kocerha & Aggarwal, 2018).

The involvement of neurotransmitter systems in the psychiatric disorders is well documented, implicating *over*-activity of dopaminergic and glutamatergic systems, as well as *hypo*-function of serotonergic (Garbett, Gal-Chis, Gaszner, Lewis, & Mirnics, 2008) and γ-aminobutyric acid (GABA)-ergic (Orhan et al., 2018) neurotransmitter systems. These systems may be epigenetically modified contributing to the symptoms of psychosis. For instance, the candidate gene catechol-*O*-methyltransferase (*COMT*) has been implicated in the dopamine pathway. *Hypo*-methylation of (*COMT*) and subsequent upregulation of membrane-bound (MB)-COMT is thought to mediate *over*-activity of dopaminergic pathways by enhancing dopamine degradation. This mechanism contributes to the symptom profile in schizophrenia and bipolar disorder .

GABAergic dysfunction has consistently been associated with psychosis and negative symptomology in schizophrenia (Taylor & Tso, 2015). *Hyper*-methylation of reelin (*RELN*) and glutamic acid decarboxylase (*GAD1*) is found in GABAergic neurons (Guidotti, Grayson, & Caruncho, 2016; Huang & Akbarian, 2007). Downregulation of *RELN* and *GAD1* is associated with an increase in DNMT1 expression in the cortex of patients with schizophrenia (Veldic, Guidotti, Maloku, Davis, & Costa, 2005). This may suggest that increased methylation in GABAergic neurons is driven by DNMT1 activity (Dong, Ruzicka, Grayson, & Guidotti, 2015).

Hypo-function of serotonergic systems is thought to be mediated by *hyper*-methylation of serotonin receptor type-1 (HTR1A) (Carrard, Salzmann, Malafosse, & Karege, 2011) and type-2 (*HTR2A*) (Abdolmaleky et al., 2011) mediating multiple signaling pathways involving dopamine transmission (Holloway & González-Maeso, 2015).

1.4 Environmental influences on the epigenome

1.4.1 Antipsychotics and co-medication with psychotrophics

DNA methylation patterns are altered by mood stabilizers prescribed in bipolar disorder such as lithium, while valproic acid indirectly alters DNA methylation by acetylation/deacetylation of histones (Pisanu, 2018). Costa et al., (2002) found that valproic acid enhanced the therapeutic effects of second-generation antipsychotics by indirectly countering the *hyper*-methylation of *GABA* and upregulating expression of *RELN*. Asai et al., (2013) found the therapeutic effects of lithium may be mediated by countering the *hyper*-methylation of solute carrier family 6 member 4 (*SLC6A4*) found in the pre-frontal cortex of patients with bipolar disorder (Sugawara et al., 2011).

1.4.2 Findings with gender

The effects of gender on neural function are important variables and DNA methylation levels are known to be dependent on sex (Eranti, MacCabe, Bundy, & Murray, 2013). The role of differential methylation in sex hormones is well documented, including the variation in tissue-specific distribution of sex hormone receptors (Zouboulis, Chen, Thornton, Qin, & Rosenfield, 2007). Other factors contributing to the genetic differences between the genders includes the presence or absence of genes encoded on the X and Y chromosomes. Zechner et al. (2001) reported on the unusually large number of genes contained on the X chromosome that are involved in the development and function of the nervous system. The masculinizing effect on the brain by the Sex determining region Y (*Sry*) gene located on the Y chromosome was reported by Xu, Burgoyne, & Arnold, (2002).

1.4.3 Findings with smoking behavior

Active tobacco smoking leads to changes in DNA methylation levels (Bauer et al., 2015) which are sensitive biomarkers in the early period of smoking initiation (R. A. Philibert, Beach, Lei, & Brody, 2013). Differentially methylated CpGs associated with smoking can actually predict exposure with high accuracy (Zhang, Florath, Saum, & Brenner, 2016) even years after smoking cessation (Wan et al., 2012). Many smoking-associated CpGs show less persistence and may show normalized methylation levels following smoking cessation. Bauer et al. (2015) reported that active tobacco smoking is also associated with a larger proportion of lymphocytes, suggestive of a protective immune response against the effect of smoking.

1.4.4 Findings with age

Gene expression is dependent on genetic and environmental factors, and undergoes changes during aging (Bryois et al., 2017). It is relatively unknown which aspect of gene regulation is the first to become dysregulated (Booth & Brunet, 2016), nevertheless, several pathways are inter-connected in the aging process and disruption of one pathway by epigenetic dysregulation may lead to dysfunction in others (Kirkpatrick & Kennedy, 2018).

Horvath (2013) proposed the model of an "epigenetic clock" as a chronological age estimator based on the gradual tissue-specific accumulation of differentially methylated genes with age. A comparison of this chronological age predictor to biological aging reveals a pattern of accelerated aging influenced by multiple environmental factors. Degerman et al. (2017) incorporated this "epigenetic clock" into a longitudinal study. Their results showed that individuals with DNA methylation age younger than chronological age, preserved better memory function and cognitive status. In contrast, the impact of negative environmental factors including life style were reflected in accelerated DNA methylation aging and impaired cognitive functions (Ibid).

1.5 DNA methylation assayed in peripheral blood

Peripheral blood is the most widely used tissue for methylation analyses due to ease of access.

1.5.1 Cellular heterogeneity

Whole blood represents a rich collection of diverse cell types, each with a very different DNA methylation profile (Houseman et al., 2012; Jaffe & Irizarry, 2014). The six cells types adjusted for in epigenetic studies are involved in cellular immunity: leukocytes (granulocytes and monocytes), and

lymphocytes (B-cells, CD4+T cells, CD8+T cells, and natural killer cells (NK). Changes in cell type composition, especially in the proportion of lymphocytes may be indicative of an immune response, or a disease phenotype (R. Philibert & Glatt, 2017). It has been shown that antipsychotic drugs alter lymphocyte cell counts and their methylation patterns (Houtepen, van Bergen, Vinkers, & Boks, 2016).

Many lymphocytes, such as B-cells, T- cells and NK cells express similar receptors as neuronal cells including brain-derived neurotrophic factor (BDNF), dopamine, and (GABA) (Gladkevich, Kauffman, & Korf, 2004). Disturbances in main neurotransmitter systems in psychotic disorders are seen concurrently with altered blood lymphocyte function (Ibid). Given that antipsychotic drugs are believed to mediate their therapeutic effect through these same systems, identification of differentially methylated regions mediated by antipsychotics may provide insight into new drug targets.

1.5.2 Peripheral blood

DNA methylation profiles are tissue specific with methylation patterns in the blood differing from other corporal tissues (Hannon, Lunnon, Schalkwyk, & Mill, 2015; R. Lowe, Slodkowicz, Goldman, & Rakyan, 2015). In particular, there are large differences observed between blood and brain (Horvath et al., 2012). However, in a large EWA on schizophrenia, Hannon et al. (2016) introduced a methodological approach that integrated genetic and epigenetic findings. They identified DMPs and DMRs associated with schizophrenia that overlapped with previously identified loci associated with schizophrenia from genome-wide association studies (GWAS). This study was important for EWA studies. Evidence was provided for blood-based DNA methylation analysis to identify differentially methylated candidate genes identified through GWAS (Ibid).

1.6 Methods for interrogating the epigenome

Epigenetic-wide association studies (EWAS) use state-of-the-art methods to evaluate epigenome-wide changes in the methylome. These methods include whole-genome sequencing and microarray technology. Microarray technology offers several advantages that may be preferable in many experimental studies: the method is cost-effective compared to sequencing; it is rapid and reproducible; and it provides a high concordance of methylation values with sequencing methods. Notably, microarray technology facilitates hypothesis-generating studies as the array provides a “fingerprint” of the cellular state (Wright et al., 2016).

1.6.1 Microarray technology for typing DNA methylation

The current microarray of choice is the Illumina Infinium® MethylationEPIC BeadChip. This array has a higher number of probes than previous generations of arrays, thus allowing for a larger proportion of known methylation sites in the genome to be assayed. Currently the EPIC provides single-nucleotide resolution across the genome at over 850 000 methylation sites (Pidsley et al., 2016).

DNA samples are first subjected to sodium bisulfite conversion prior to microarray analysis. During bisulfite conversion, unmodified cytosines (C) are converted to uracils (U). During subsequent amplification, uracils (U) which are copied as thymines (T). Methylated cytosines however are protected from this conversion (Figure 1.7). Probes on the microarray are designed to detect converted thymines or unconverted cytocines at a CG site. The bisulfite converted DNA is then amplified prior to hybridization to the microarray where two different probes with fluorescent colors report on corresponding methylation status.

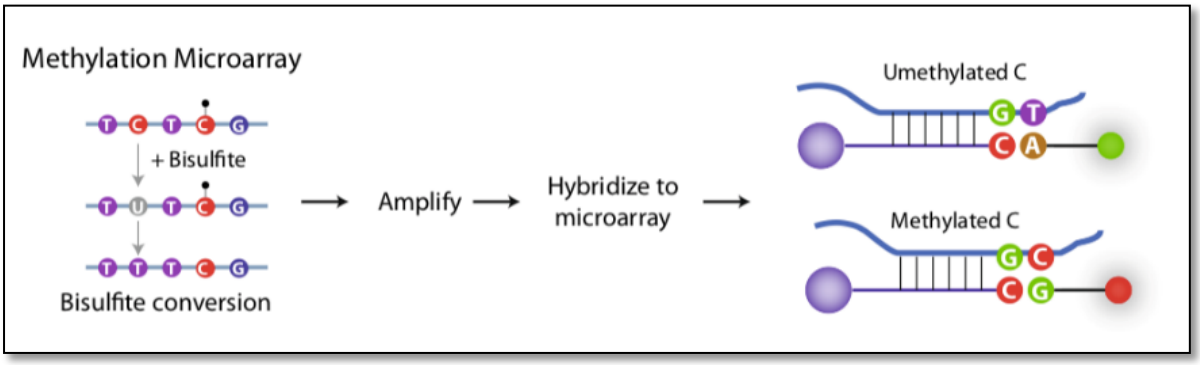


Figure 1.7 Sodium bisulfite conversion of DNA prior to microarray analysis allows for quantitation of methylated CpGs. Unmodified (C)s are converted to (U)s and subsequently copied to (T)s. Probes designed to detect a converted (T) or unconverted (C) have different reporter colors (cartoon adapted from (Masser et al., 2018).

Methylation data assayed from the microarrays will then be preprocessed prior to analysis. The pipeline described in Table 1 provides an overview of the methods used to process, normalize, and analyze the raw methylation data provided by the array. The pipeline provides for quality control (QC), reliability and reproducibility of the reported findings (Wright et al., 2016).

Table 2 Major steps in the EPIC 850K microarray analysis pipeline

Analysis	Motivation for step
Sample filtering	Cohort samples are compared to control probes located on the array to identify samples that fail to adequately detect DNAm. Samples with poor detection are suspected of poor quality and excluded.
Probe filtering	Data screening and quality control of raw data. Probes are removed that fail to meet preset detection values ($p < 0.05$) or are unreliable due to overlap with SNPs which may confound results.
Within-array normalization	Background noise is removed, correction for technical dye-based (red/green) intensity, and probe type (I/II) differences in the array.
Batch effects	Technical or procedural differences may cause hidden variance in the dataset. This non-biological variation may be adjusted for.
Cell-type composition	Each blood sample may contain different proportions of cells types, each with potentially varying DNAm profiles. Statistical methods are incorporated to estimate and correct for cellular heterogeneity.
Differentially methylated positions (DMPs)	Identification of CpG sites; sight-specific DNAm differences at single nucleotide resolution
Differentially methylated regions (DMRs)	DNAm differences identified in CpG-dense regional clusters.
Over-representation and pathway-based enrichment analysis	Genes mapped to DMPs are evaluated in multiple gene databases to discover functional or regulatory enrichment. Pathway analysis provides a snapshot of the relationship of enrichment to cellular mechanisms.
Biological interpretation	Several approaches may be required to interpret the biological relevance of *hypo*- or *hyper*-methylated DMPs, DMRs, and gene expression.

1.7 Biological interpretation of epigenetic modifications

Several interpretation-oriented approaches have been adapted to gain an understanding of the biological and clinical significance of DNA methylation results. The choice of analytical tool and/or publicly available database depends upon the study design. The genomic coordinates of identified differentially methylated regions is easily accessed through publicly available genome browsers

(UCSC). Specialized epigenetic databases detail CpGs, their associated genes, and their concordance between peripheral blood and brain (Edgar, Jones, Meaney, Turecki, & Kobor, 2017). Gene set analysis that leads to implicated pathways is an leading discovery approach (Kamburov et al., 2011). And identification of proteins from these gene sets may be analyzed through their evolutionary relationships (Thomas et al., 2003). Following *in silico* analysis, animal or cell culture models may be used to adapt further studies, or to evaluate the significance of the observations in the clinic (B Swathy, Saradalekshmi, Nair, Nair, & Banerjee, 2018).

Aims

Evidence is provided that antipsychotic medications influence DNA methylation genome-wide and at localized CpG sites of candidate genes. Based on this background, the motivation of the thesis is to address the following issues:

1. To identify the DNA positions and regions associated with altered methylation patterns due to treatment with olanzapine,
    a. To compare models for adjustment for cellular composition,
2. To verify findings in HepG2 cells exposed to olanzapine,
3. To evaluate concordance between blood and brain for CpGs that are differentially methylated in blood,
4. To identify implicated pathways associated with methylation patterns altered by olanzapine.

Materials

3.1 The Thematically Organized Psychosis (TOP) Cohort data set

Data from patients (n = 691) and controls (n = 309) was obtained through the TOP Cohort project affiliated with the Norwegian Centre for Mental Disorders Research (NORMENT). The cohort included patients recruited in the Oslo area diagnosed with psychotic disorders including schizophrenia, bipolar disorder and major depressive disorder. Clinical information relevant to the evaluation of DNA methylation levels and sample selection included ethnicity, gender, age, psychiatric diagnosis and the environmental effects of smoking behavior, antipsychotic and mood stabilizing medications.

3.2 DNA methylation data sets

DNA methylation levels derived from patient blood samples and olanzapine-exposed HepG2 cells were assayed on the Infinium® MethylationEPIC BeadChip at the Institute of Human Genetics, University Hospital of Bonn, Germany.

3.3 Eukaryotic-cell line

Name	Description	Supplier	Catalog no.
HepG2	human hepatoma cultured cell line	American Type Culture Collection (ATCC), Manassas, Virginia, USA	HB-8065

3.4 Cell culture chemicals and reagents

Reagents	Supplier	Catalog no.
Eagle's Minimum Essential Medium (EMEM)	Lonza	12-662F
Penicillin / Streptomycin	Invitrogen	15140-122
Fetal bovine serum (FBS)	Invitrogen	10106-169
L-Glutamine	Lonza	17-605F
Dulbecco's PBS	Sigma-Aldrich	D1408
Trypsin-EDTA	Sigma-Aldrich	T4299
Tris-acetate-EDTA	Thermo Fischer	10123293

3.5 Commercial kits

Name	Components of kit	Supplier	Application	Catalog no.
AllPrep DNA/RNA Mini Kit	Buffer RLT Plus, Buffer RW1, Buffer RPE, Buffer AW1, Buffer AW2, Buffer EB, RNase-Free water	Qiagen	DNA and RNA purification	80204
Amicon®Ultra-0.5 Centrifugal Filter Device	Centrifuge filters, tubes	Merck KGaA	DNA filtration, concentration	UFC500396
High Capacity cDNA Reverse Transcription Kit	MultiScribe Reverse Transcriptase, dNTP, RT buffer, RT Random Primers	Applied Biosystems	cDNA synthesis	4368813
LightCycler® 480 SYBR Green I Master	FastStart Taq DNA Polymerase, SYBR Green I dye, PCR-grade water	Roche	q-PCR	04707516001

3.6 Chemicals

Name	Supplier	Application
Dimethyl sulfoxide (DMSO)	Sigma - Aldrich	Dissolving agent - Olanzapine
Olanzapine	Toronto Research Chemicals Inc., Toronto, Canada	HepG2 treatment
β-mercaptoethanol (β-ME)	Sigma - Aldrich	DNA and RNA purification
Ethanol	AntiBac	DNA and RNA purification
DNA Gel Loading Dye (6X)	Thermo Fischer	Agarose gel electrophoresis
Ethidium bromide stock solution (2.5 mg/ml)	Sigma - Aldrich	Agarose gel electrophoresis
DNA standard ruler	Thermo Fischer	Agarose gel electrophoresis
Agarose	Lonza Sea Kem LE	Agarose gel electrophoresis

3.7 Oligonucleotides

Primers*	Sequence (5'→ 3')
Hs_HPRT_Fwd	tgaccttgatttattttgcatacc
Hs_HPRT_Rev	cgagcaagacgttcagtcct
Hs_ChREBP-α_Fwd	agtgcttgagcctggcctac
Hs_ChREBP-α_Rev	ttgttcaggcggatcttgtc
Hs_ChREBP-β_Fwd	agcggattccaggtgagg
Hs_ChREBP-β_Rev	ttgttcaggcggatcttgtc
Hs_FASN_Fwd	caggcacacacgatggac
Hs_FASN_Rev	cggagtgaatctgggttgat
Hs_SREBF1_Fwd	cgctcctccatcaatgaca
Hs_SREBF1_Rev	tgcgcaagacagcagattta

**All primers were ordered from Sigma-Aldrich*

3.8 Instruments and software

Instrument	Supplier	Application
Countess Automated Cell Counter	Bio-Rad	Cell culture work
NanoDrop ND-100 Spectrophotometer	NanoDrop Technologies	DNA & RNA concentration
GeneAmp®PCR system 9700	Applied Biosystems	cDNA synthesis
LightCycler®480 II	Roche	q-PCR
Hoefer™ HE33 Mini Submarine	Fischer Scientific	Agarose gel electrophoresis
GelDoc ™ XR+	Bio-Rad	Agarose gel image

Methods

4.1 Preprocessing pipeline of methylation data

Preprocessing and quality control (QC) was performed at the Dr. Einar Martins Group for Biological Psychiatry as described in the Appendix. The output from the preprocessing was a data matrix containing M-values. These values are representations of the log2 of the ratio between the methylated and unmethylated probes (Du et al., 2010).

4.2 Analytical pipeline used in methylation analysis

Procedures and standard protocols used in the methylation pipeline were followed using open-source statistical programming R-packages (R-Core Team, 2018) hosted at BioConductor (Huber et al., 2015). BioConductor provides a wide variety of scientific software for the analysis of microarray. The motivation for the procedure and the BioConductor R-packages and their authors are listed in Table 4.1.

Table 4.1 R-packages used to evaluate the effects of covariates, or unwanted variation on methylation

Package	Motivation for step	Authors
SVA: Surrogate Variable Analysis v.3.28.0	Batch effect and hidden variable identification	(J T Leek, Johnson, Parker, Jaffe, & Storey, 2012)
Limma: Linear Models for Microarray Data, v. 3.38.2	Statistical method of identifying significant DMPs	(Ritchie et al., 2015)
IlluminaHumanMethylationEPICanno.ilm10b2.hg19, v. 0.6.0 **Illumina EPIC annotation library**	Mapping probes to genes and genomic elements	(Hansen, 2016)
DMRcate, v. 1.16.0	Statistical method of identifying significant DMRs	(Peters et al., 2015)
TxDb.Hsapiens.UCSC.hg19.knownGene v. 3.2.2	Genomic annotation database for DMR coordinates	(Carlson, 2015)
GenomicRanges, v. 1.34	Computes annotated ranges for DMRs	(Lawrence et al., 2013)

AnnotationHub, v.2.14.1	Annotated files for DMRs	(Morgan, 2018)
Gviz, v. 1.24.0	Visualization of DMRs	(Hahne & Ivanek, 2016)

4.3 Model selection

In order to evaluate the effects of covariates, or unwanted variation on methylation, linear regression models were compared correcting for variables known to affect methylation, including gender (Eranti et al., 2013), smoking status (Bauer et al., 2015), age (Horvath et al., 2012) and estimated cells counts (eCells) (Jaffe & Irizarry, 2014) as presented in Table 4.2. Each model was evaluated with the inclusion or exclusion of adjustments for eCells.

Table 4.2: Linear regression models evaluated for antipsychotic effect on differential methylation

Model	Description
Mod 1a	y ~ AP1
Mod 1b	y ~ AP1 + eCells
Mod 2a	y ~ AP1 + gender + smoker + age
Mod 2b	y ~ AP1 + gender + smoker + age + eCells
Mod 3a	y ~ AP1 + gender + smoker
Mod 3b	y ~ AP1 + gender + smoker + eCells

Analysis of these models was initiated in R with *Surrogate Variable Analysis* (**sva**) (Table 4.1). The data was formatted as a normalized matrix of methylation (M) expression values with probes in the rows, and barcodes of the samples in the columns. Two model matrices were created: the full model containing the variable of interest (X) and desired covariates (A + B), i.e. y ~ X + A + B; and the null model containing the same covariates as the full model, however without the variable of interest, i.e. y ~ A + B.

The **sva** algorithm then estimated all unmodeled sources of variation directly from the data and returned the surrogates variables (SVs). The SVs were considered for incorporation into the model as additional covariates, i.e. y ~ AP1 + SVs, or for removal from the data set with the **Combat** tool (Table 4.1). As the **sva** tool does not identify the source of the estimated variation, a heatmap plot with the correlation of each SV to a given covariate was evaluated. The number (n) of surrogate variables for each model was recorded for evaluation during the model selection process.

4.4 Analysis of differentially methylated positions

Differential methylation analysis was performed in R using *Linear Models for Microarray Data* **limma** (Table 4.1). **limma** is commonly used to find differences between two groups of samples, even for a large set of features. A typical **limma** pipeline consists of 3 steps: 1) it fits a linear model using e.g. multiple linear regression, which takes into account the covariates to be corrected for in the analysis; 2) it improves the estimation of variance by borrowing information across the different features using Baysian statistics. This method is particularly useful when the sample size is small; 3) lastly, it identifies the features that are different between the two groups under study using a contrast matrix.

4.5 Gene annotation of differentially methylated positions

Differentially methylated probes identified by **limma** were mapped to probes using the **EPIC annotation library** (Table 4.1). This library consists of all genes associated with the probes on the EPIC array including enhancer regions and DNase hypersensitive sites which are important in development and differentiation of blood cells. The output from **limma** provides information on probes that map to multiple genes and overlapping genes.

4.6 Identification of differentially methylated regions

Identification and ranking of the most differentially methylated regions genome-wide was assayed with **DMRcate** and associated R-packages (Table 4.1) (Peters, 2015). The tool relies on functions previously used in **limma** such as designing the linear model and contrast matrix. Following FDR correction (0.3), CpGs from the EPIC array were annotated by regions that grouped from clusters of significant probes within a distance (< *lambda* = 1000) nucleotides to the next probe.

The list of identified DMRs was converted to a genomic range (**GenomicRanges)** that annotates overlapping promoter regions (+/-) 2000 base pairs from the transcription start site (TSS). The ranges were then extracted from the human (hg19) data annotation package **TxDb.Hsapiens.UCSC.hg19.knownGene** and stored with the **AnnotationHub** package. The genome browser hosted at the University of California Santa Cruz (UCSC) (https://genome.ucsc.edu) was used to view the genomic coordinates. The subsequent ranges were plotted using **Gviz** package.

4.7 Validity of CpGs associated with identified differentially methylated regions

Genes containing DMRs identified from Model 3b (FDR 0.3) were evaluated for concordance between blood and brain methylation levels. The names of the associated genes were entered into the online Shiny app for Blood-brain Epigenetic Concordance (BECon) (Table 4.3).

4.8 Gene set over-representation analysis

The ConsensusPathDB from the Max Plank Institute for Molecular Genetics, Berlin, Germany (http://cpdb.molgen.mpg.de/) is a webportal for multiple gene set and pathway-based databases. User defined lists of genes with *p*-values <10E-03 were submitted to the database using HGNC identifiers. A background list of all the genes identified in the analysis was submitted. The publicly available databases used in gene set analysis are seen in Table 4.

Table 4.3 Publicly available databases used to interpret biological relevance

	URL and version release	Reference
BECon: Brain and Blood	https://redgar598.shinyapps.io/BECon/ April, 2017	(Edgar et al., 2017)
ConsensusPathDB	http://cpdb.molgen.mpg.de/ Release 32, January 11, 2017	(Herwig, Hardt, Lienhard, & Kamburov, 2016; Kamburov et al., 2011)
Gene Ontology Consortium	http://geneontology.org/	(Ashburner et al., 2000; Gene Ontology, 2015; The Gene Ontology, 2017)
Pathway Interaction Database (NCI-PID)	hosted at NDEx http://www.ndexbio.org/#/	(Pillich R.T., J., V., D., & D., 2017)
KEGG: Kyoto Encyclopedia of Genes and Genomes	https://www.kegg.jp/ Release 87.1, August 1, 2018	(M Kanehisa, Furumichi, Tanabe, Sato, & Morishima, 2017; Minoru Kanehisa, Sato, Kawashima, Furumichi, & Tanabe, 2016)
Reactome	https://reactome.org/ Release 66, September 27, 2018	(Fabregat et al., 2018; Sidiropoulos et al., 2017)

REVIGO	http://revigo.irb.hr/ Maps to Gene Ontology release Jan 2017	(F Supek, Bosnjak, Skunca, & Smuc, 2011; Fran Supek & Skunca, n.d.)Supek *et al.,* 2017
Wikipathways	http://wikipathways.org/	(Slenter et al., 2018)

4.8.1 Over-representation analysis of gene sets

A list of gene names with *p*-values < 10E-03 and a user defined background list consisting of all the mapped genes assayed from the given model was submitted to the ConsensusPathDB. Gene Ontology (Table 4.3). (GO) categories were searched by selecting gene ontology levels 2 and 3 for biological process (BP), molecular function (MF) and cellular component (CC) with *p*-value cutoff = 0.01. A list of GO IDs and corresponding *p*-values was then submitted to the REVIGO human database (Table 4.3) where redundant gene sets were removed and the BP, MF and CC categories were summarized. Top hits from each category were evaluated for low –log *p*-values, low frequency and high uniqueness of the GO terms.

4.8.2 Over-representation analysis of pathway-based gene sets

Default pathway databases were selected (i.e. Kegg, Biocarta, Wikipathways, Reactome, NCI-PID and Pharmgkb) with *p*-value cutoff = 0.01. The top pathways with *p*-values < 10E-03 were evaluated further in the corresponding pathway databases identified by ConsensusPathDB in Table 4.3.

4.9 Verification of differentially methylated position

4.9.1 HepG2 cell culture exposure to olanzapine

Eukaryotic cell culture experiments were performed under sterile conditions under a Laminar flow hood. HepG2 cells were cultivated in a medium flask (75 cm^2) with Eagle's Minimum Essential Medium (EMEM) supplemented with 10% FBS (v/v), penicillin/streptomycin (final concentration penicillin 100 U/mL, streptomycin 100 µg/mL) and 1% L-Glutamine (v/v) in a 5% CO_2 humidified cell culture incubator at 37 °C. The medium was changed every 48 hours and the cells were washed with 15 ml pre-warmed PBS (37 °C). When the cells reached 80 % confluence, 3 mL trypsin was added (37°C), followed by incubation (37°C) for 3-4 minutes. Trypsination was interrupted by adding 5 mL growth medium (37°C) and resuspending with a pipette. The cells were counted using Countess

Automated Cell Counter (Bio-Rad), and 150 000 cells/well were seeded out in a 12-well plate and placed in the incubator.

Following the 48-hour growth period, the cells were exposed for 72 hours to different concentrations of olanzapine (1 µM, 50 µM and 100 µM) dissolved in 0.25% (v/v) DMSO. No treatment (NT) samples were prepared with EMEM, and Vehicle samples with 0.25% (v/v) DMSO. After the 72-hour exposure period, the medium was removed and the cells were washed with D-PBS (37°C). RLT Lysis buffer (350 µl) (Qiagen) was added to the cells to inhibit RNA degradation from endogenous RNases. The cell lysate suspension was transferred to RNase-free 2 ml Eppendorf tubes and stored at -80°C until extraction.

4.9.2 Total RNA and genomic DNA extraction

Standard protocol for the extraction of gDNA and RNA with the ALLPrep DNA/RNA Mini (Qiagen) was followed. gDNA and RNA concentrations were quantified by NanoDrop 1000 (Thermo Fischer). gDNA samples were then concentrated threefold (60 ng/µl) to laboratory specifications for methylation analysis. This was performed by pipetting 200 µl of the extracted gDNA into an Amicon® Ultra 0.5 mL Centrifugal Filters (Merk KGaA) and centrifuging at 14,000 x g for 20 minutes. The desired concentration of the newly purified gDNA was measured by NanoDrop 1000 (Thermo Fischer) and all samples were diluted to 20 µl.

4.9.3 Analytical agarose gel electrophoresis

An analytical agarose gel was performed to verify the DNA remained double-stranded following the extraction procedure. For nucleic acid detection, a 1.0 % (w/v) agarose gel (100 ml 1x TAE buffer, 1 g agarose, 100 µg/mL ethidium bromide) was used to separate the DNA fragments. The DNA samples were diluted with 6x DNA loading dye, to a final concentration of 1x DNA loading dye. 5 µl of a DNA ladder (Thermo Fischer) was loaded as a standard marker. Electrophoresis was performed in 1x TAE buffer at 120 V for 30 minutes. The electrophoretic mobility of the DNA was visualized with the Bio-Rad Gen Doc EZ Imager System.

4.9.4 cDNA synthesis

Conversion of RNA into complementary DNA (cDNA) was performed using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). A master mix was prepared by combining the components listed in Table 4.4. All components were placed on ice and the work was performed

quickly in sterile conditions under the hood. Each sample of RNA was added as a final component to an aliquot of master mix.

Table 4.4 Components of master mix and RNA for cDNA synthesis

Volume	Components
0.8 µl	25x dNTP mix
2 µl	10x RT buffer
1 µl	MultiScribe Reverse Transcriptase
9.2 µl	Rnase-free water
2 µl	10x RT Random Primers
15 µl	Total Master Mix
5 µl	Diluted RNA (20 ng/µl)
20 µl	Total reaction volume

The thermal cycler settings used in the cDNA thermal reactions are described in Table 4.5.

Table 4.5 Thermal cycler settings

	Step 1	Step 2	Step 3	Step 4
Time (min)	10	120	5	∞
Temperature (°C)	25	37	85	4

The newly synthesized cDNA was diluted 1:3 in PCR-grade water and stored at -20°C until further use.

4.9.5 Quantitative real-time polymerase chain reaction

Quantitative real-time PCR (qPCR) was performed on all samples to measure gene expression in HepG2 cells following olanzapine exposure. The method is accurate and highly sensitive to the amount of mRNA detected in samples. This is done by detecting the level of fluorescent dye that binds to the amplified cDNA in candidate genes. Extracted RNA from HepG2 cells was used as a template for cDNA synthesis and subsequent qPCR-analyses.

Genes associated with lipid biogenesis were selected to validate known drug effects on gene expression from earlier studies. Lipid biogenesis may be induced by activating the sterol regulatory

element binding protein (SREBP), encoded by the sterol regulatory element binding transcription factor 1 (*SFEBF1*) gene (Fernø et al, 2006). Lipid biosynthesis is also induced by the fatty acid synthase (*FASN*) gene, shown to be upregulated by olanzapine treatment (Fernø et al, 2008). And inducing the Carbohydrate-response element binding protein (*ChREBP*) gene upregulates lipid biosynthesis in metabolic tissues and cancer cells (Yu et al, 2014).

The primers for these genes are listed in Table 3.6. The components listed in Table 4.3 were then placed in the LightCycler ® 480 SYBR Green I Master (Roche). During the PCR reaction, single-stranded cDNA is first amplified followed by the formation of double stranded DNA (dsDNA). The fluorescent dye (SYBR green) binds to dsDNA and the detection of an increase of fluorescent signal will be proportional to the amount of dsDNA present. Three series of the reaction were run for statistical analysis.

Table 4.6 Components needed for one reaction

Volume	Components
4 µl	4 µl SYBR green Master I
0.4 µl	Forward primer (20µM)
0.4 µl	Reverse primer (20 µM)
2.2 µl	Rnase-free water
1 µl	cDNA
8 µl	Total reaction volume

4.9.6 Statistical analysis

The q-PCR data was evaluated by testing the relative expression levels of the candidate genes against an endogenous control. The delta-delta Ct method was used for this purpose, followed by a two-sided student *t*-test to determine the statistical significance of the three series. The threshold for statistical significance was set at $p = 0.05$.

4.9.7 Identification of differentially methylated positions in HepG2 cells

Preprocessing and analysis of the methylation data was performed following the methods and pipeline used for blood samples (Appendix) and R-packages (Table 4.1). Differential analysis was performed with **limma**. Four contrast matrices were created: Vehicle was subtracted from each exposure concentration and the control matrix represented Vehicle subtracted from No treatment (NT). The differentially methylated probes identified were mapped to genes using the **EPIC**

annotation library. The top 16 probes were visualized with **ggplot2** to evaluate the variation of average methylation level (M value) for each exposure concentration (1 µM, 50 µM or 100 µM) or condition (NT or Vehicle).

4.9.8 Identification of differentially methylated regions in HepG2 cells

Identification of differentially methylated regions was performed using R-packages (Table 4.1) with standard settings as with patient blood samples: **DMRcate, GenomeRanges, AnnotationHub** and **Gviz.**

4.9.9 Gene set over-representation analysis

A user-defined gene list and a background list (*p*-values < 10E-03) with all mapped genes (n = 25 580) were submitted to ConsensusPathDB. The output provided GO IDs and corresponding *p*-values which were then submitted to the REVIGO human database (Table 4.1). Redundant gene sets were removed and the biological process (BP), molecular function (MF) and cell component (CC) categories were summarized. The top pathways-based gene sets with *p*-value < 10E-03 were evaluated.

4.10 Meta-analysis of HepG2 and Model 3b differentially methylated positions

A meta-analysis was performed for statistical power to identify the enriched genes that ranked highest in both groups.

Results

The main aim of this project was to identify DNA regions that were subject to methylation changes as a result of olanzapine treatment in patients. In our analytical approach we used a data set where DNA methylation patterns have been profiled in blood. The full data set contains data from patients on different antipsychotic drugs. Some patients use multiple drugs, making it difficult to identify changes induced by the individual drugs. A selection of appropriate samples was therefore made. In the following sections the sample selection and results from the identification of differentially methylated regions will and the results from gene set analysis will be described. The results from the verification step using HepG2 cells will be presented and finally, an evaluation of concordance between blood and brain.

5.1 Sample selection and description

DNA methylation data from 158 European samples were selected for analysis: olanzapine (n = 82), quetiapine (n = 48), and aripiprazole (n = 28) (Figure 5.1). Controls were excluded from DNA further analysis.

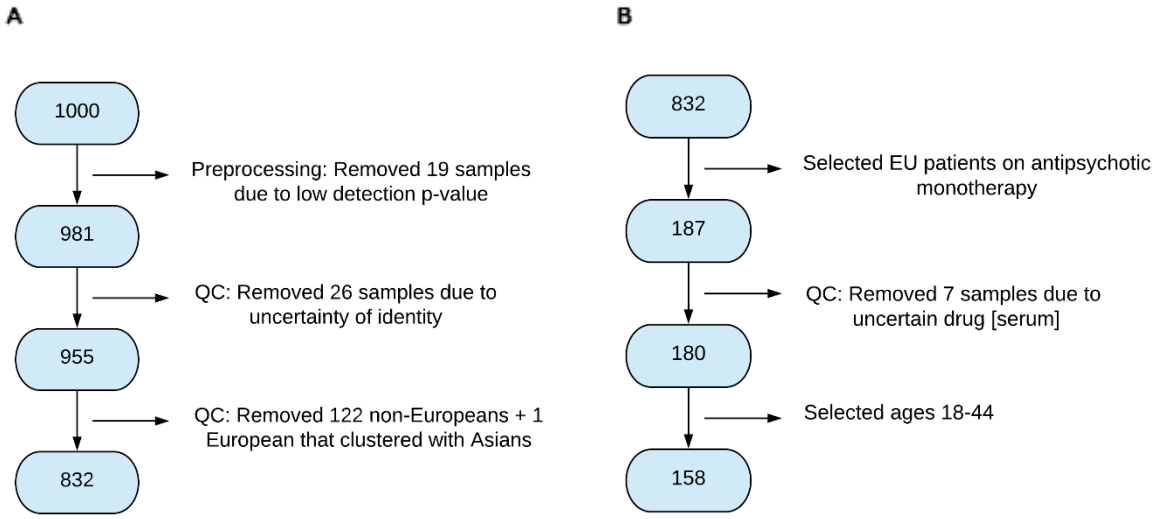


Figure 5.1. TOP cohort sample selection: A) Samples were removed during preprocessing steps as described in the Appendix, and B) samples were selected characterized by European, aged 18-44, and with serum concentration reflecting adherence to antipsychotic monotherapy.

5.1.1 Distribution of gender, age and smoking status

Distribution plots were prepared of the three antipsychotics (Figure 5.2) using the R-package **ggplot2** (see Table 4.1). The distribution of gender (1 = male, 2 = female) in Figure 5.2.A suggests that far more men than women took olanzapine while the gender difference was less with aripiprazole and quetiapine. More patients smoked (0 = no, 1 = yes) while taking olanzapine, far fewer smoked while taking aripiprazole, and the difference was negligible for quetiapine (Figure 5.2.B). Age was evenly distributed for olanzapine while patients tended to be younger while taking either aripiprazole or quetiapine (Figure 5.2.C). The range of serum drug concentrations varied considerably, and values below the recommended range were hardly visible for olanzapine (Figure 5.2.D). In addition, 15 samples total lacked recorded serum concentrations, yet were included due to registered methylation data.

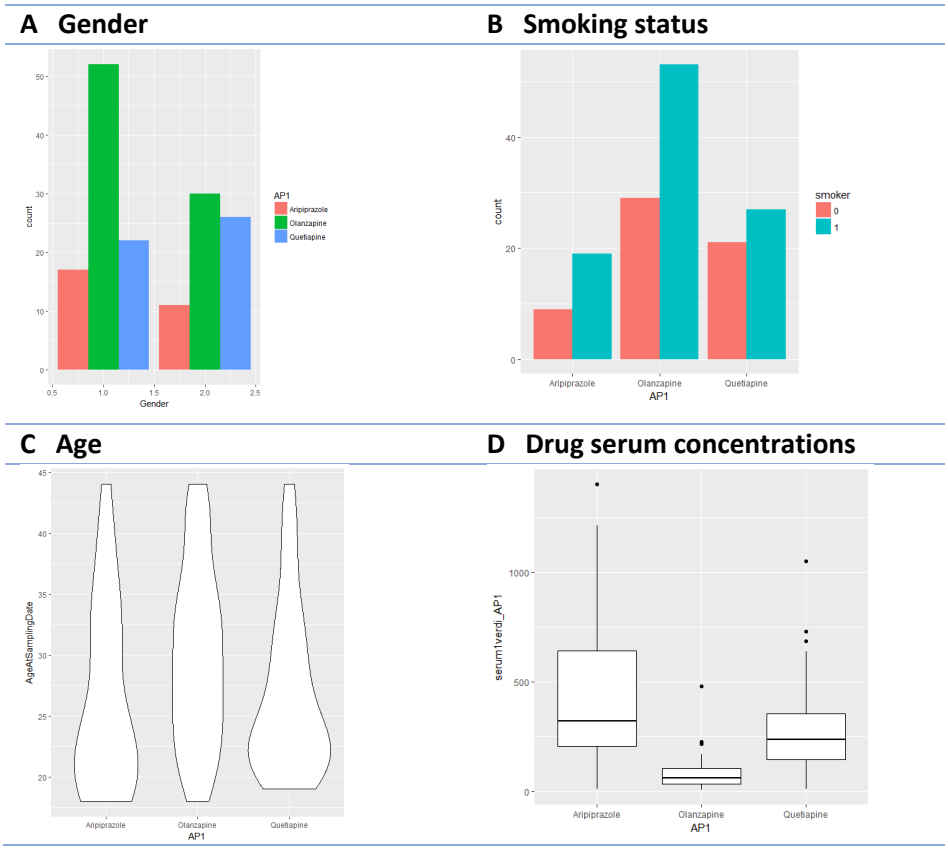


Figure 5.2. Visualization of distributions: A) gender (1 = male, 2 = female), B) smoking status (0 = no, 1 = yes), C) age, and D) drug serum concentration in 158 patients with psychotic disorders taking: olanzapine n= 82, quetiapine n = 48, and aripiprazole n = 28. Plots prepared with **ggplot2**.

The age group of our samples (18-44) was determined by the maximum age for aripiprazole. As age was not evenly distributed for aripiprazole and quetiapine, a cumulative age distribution was plotted. The data suggested a trend towards more patients older than age 35 taking olanzapine, while patients taking aripiprazole and quetiapine would level off after this age (Figure 5.3).

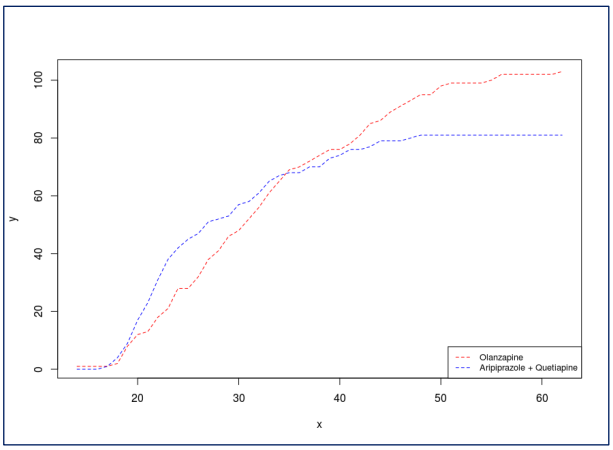


Figure 5.3. Cumulative age distribution of patients adhering to monotherapy: In the TOP Cohort data set, more patients over the age of 35 years take olanzapine (red) than aripiprazole and quetiapine (blue).

5.1.2 Distribution of serum concentrations

Drug serum concentrations from the TOP Cohort data set were compared to the therapeutic reference values provided by Haukeland University Hospital Laboratory (Table 5.1). The number of samples outside the therapeutic ranges is also presented.

Table 5.1: Range of drug serum concentrations recommended for antipsychotic therapeutic effect

Antipsychotic	Therapeutic reference values	Samples below	Samples above	NA serum conc.
Aripiprazole	200 – 1100 nmol/L	6	2	2
Olanzapine	30 – 200 nmol/L	15	4	6
Quetiapine	50 – 700 nmol/L	4	3	7

5.2 Identification of patterns of differential methylation

5.2.1 Model selection

The design of the model was based upon current practice in epigenome-wide association studies (EWAS). These studies correct for environmental factors such as age, gender, smoking status and cell type composition by including the factors as covariates in the model. We included therefore these variables as covariates in our models yet used adjustment for cell type composition as a point of comparison between the models. The aim of this comparison was to determine if cell type

composition itself contributed to differentially methylated patterns associated with an olanzapine-phenotype.

The proposed models (Table 4.2) were then evaluated using *Surrogate Variable Analysis* **sva** (J T Leek et al., 2012, 2010) to uncover hidden variation in the data set. This hidden or latent variation is represented as a surrogate variable (SV) which may be included in the model as a covariate. The simplest model (y ~ AP1) from Table 4.2 was run to evaluate the correlation between surrogate variables, and covariates or variables selected from the data set. A heat map was created (Figure 5.4) to visualize this correlation by –log transforming the *p*-values whereby the resulting white color represented the *p*-values with greatest significance, followed by the colors yellow, orange and red with least significance.

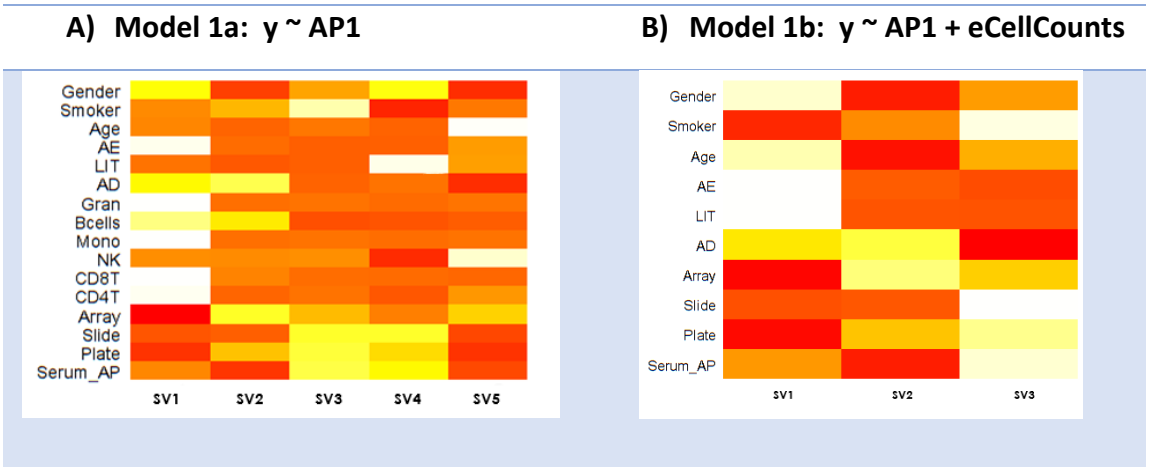


Figure 5.4: Correlation of selected covariates/variables to estimated surrogate variables (SVs): The heatmap represents the correlation between selected variables and the estimated (SVs) identified by **sva.** *P*-values are -log transformed and those with greatest significance are colored white, while red represents variables with the least significance. In Model 1b, SV3 is correlated with serum concentration of antipsychotic drugs.

Five SVs were estimated with **sva** for Model 1a (Figure 5.4.A). Surrogate variable 1 (SV1) showed the most significant correlation with antiepileptic drugs (AE), granulocytes, monocytes, CD8+T and CD4+T cells (Figure 5.4.A). A correlation between AE and these cells may also exist. The serum concentration of antipsychotic drugs (Serum_AP) showed correlation with SV3 and SV4, while lithium (LIT) was significantly correlated only with SV4. Age was correlated with SV5, however some correlation between age and NK cells may also have occurred.

When estimated cell counts values were added to Model 1b (Figure 5.4.B), **sva** estimated only 3 surrogate variables in the data set. The correlation with AE and LIT was most significant for SV1,

followed by gender and age. The most significant correlation for SV2 was with array, while SV3 was correlated with slide, plate, antipsychotic serum drug concentration and smoker.

The **ComBat** tool (Table 4.1) was then used to remove the identified SVs from the models. This resulted in no beneficial reduction of SVAs. The method prescribed then that the SVAs should be added to the model. We selected therefore the models with the least number of SVs (Model 3a and Model 3b) to limit sources of hidden variation during downstream analyses (Table 5.2).

Table 5.2 Models evaluated for differential methylation analysis and their surrogate variables

Models	Description of models	SVAs
Mod 1a	y ~ AP1 + sva	5
Mod 1b	y ~ AP1 + eCells + sva	3
Mod 2a	y ~ AP1 + gender + smoker + age + sva	3
Mod 2b	y ~ AP1 + gender + smoker + age + eCells + sva	3
Mod 3a	y ~ AP1 + gender + smoker + sva	0
Mod 3b	y ~ AP1 + gender + smoker + eCells + sva	0

5.2.2 Differentially methylated positions

Differentially methylated patterns at CpG loci were identified from the data set of methylation values by using **limma** (Table 4.1) to uncover differential expression between the antipsychotic drugs. A contrast matrix was used to essentially extract the effect of olanzapine from quetiapine and aripiprazole. This resulted in a set of differentially methylated probes associated with an olanzapine-effect. Model 3a and Model 3b were run and the identified probes were mapped to 25 580 associated genes using the **EPIC annotation library** (Table 4.1). The number of probes and associated genes for each model are listed in Table 5.3.

Table 5.3 Identification of differentially methylated probes and associated genes

Differentially methylated probes identified	Total N
Probes with *p* < 10E-03	
Model 3a	1 421
Model 3b	1 988
Genes mapped to probes	**25 580**
Model 3a (*p* < 10E-03)	1 178
Model 3b (*p* < 10E-03)	1 581

In addition to the probe names, the output of the **limma** analysis provided metrics for log-fold change (FC), average expression (AveExpr), and *p*-values. The fold change indicates the expression change for a given probe while the AveExpr represents the average log2-intensity values for a particular probe across the arrays. In Tables 5.4 and 5.5, the negative AveExpr is colored pink while positive AveExpr is colored green.

Differentially methylated probes for Model 3a: y ~ AP1 + gender + smoker

The top 18 probes (*p*-values < 10E-05) and associated gene names for Model 3a is shown in Table 5.4. Probes not annotated to a gene in the **EPIC annotation library** may be intergenic and associated with regulatory elements, such as enhancers. Note that probe 10 is located in two overlapping genes.

Table 5.4: Top 18 Differentially methylated probes with *p* < 10E-05 mapped to genes for Model 3a

	Probe	logFC	AveExpr	P-Value	Gene
1	cg05491587	-0.6195	2.1588	8.08E-07	KCNG2
2	cg04695077	-0.2203	-1.8804	1.30E-06	
3	cg25289484	0.3571	3.0323	4.58E-06	SPEN
4	cg13002170	-0.2196	3.8477	4.59E-06	NCOA2
5	cg00136105	0.2108	-3.9058	4.78E-06	GNPTAB
6	cg17256711	0.2261	-1.6588	5.18E-06	OPRM1
7	cg22989843	0.2346	-2.4029	5.66E-06	PAX3
8	cg21706229	0.1762	-2.1700	6.20E-06	
9	cg00553886	0.2030	-1.3150	6.43E-06	TEAD1
10	cg16688303	0.2534	0.6503	6.85E-06	LRRTM2;CTNNA1
11	cg06914048	0.2045	-2.9314	7.28E-06	CABLES1
12	cg08403345	0.2270	-2.3930	7.53E-06	MADCAM1
13	cg18730746	-0.1522	-3.9261	7.67E-06	TEX264
14	cg05646885	-0.1698	0.1278	7.79E-06	
15	cg23855260	-0.1947	3.2232	8.30E-06	
16	cg10027085	-0.2305	-3.8932	8.34E-06	ANGEL2
17	cg23504411	0.1965	-1.1872	8.44E-06	HIST1H3J
18	cg11565785	0.1659	-2.4412	9.73E-06	

The foremost significant gene listed was the potassium voltage-gated channel modifier subfamily G member 2 gene (*KCNG2*); followed spen family transcriptional repressor (*SPEN*); paired box 3 (*PAX3*); and TEA domain transcription factor 1 (*TEAD1*); the nuclear receptor coactivator 2 (*NCOA2*) and opioid mu receptor 1 (*OPRM1*); catenin alpha 1 (*CTNNA1*); mucosal vascular addressin cell adhesion molecule 1 (*MADCAM1*); Cdk5 and Abl enzyme substrate 1 (*CABLES1*); N-acetylglucosamine-1-phosphate transferase subunits alpha and beta (*GNPTAB*); testis expressed 264 (*TEX264*); histone

cluster 1 H3 family member j (*HIST1H3J*); angel homolog 2 (*ANGEL2*); and leucine rich repeat transmembrane neuronal 2 (*LRRTM2*).

Differentially methylated probes for Model 3b (y ~ AP1 + gender + smoker + eCells)

The top 23 probes (*p*-values < 10E-05) and associated gene names for Model 3b is shown in Table 5.5. Probes not annotated to a gene in the **EPIC annotation library** may be intergenic and associated with regulatory elements, such as enhancers. Note that probes number 9 and 19 are located in two overlapping genes.

Table 5.5: Top 23 differentially methylated probes with *p* < 10E-05 mapped to genes for Model 3b

	Probe	logFC	AveExpr	P-Value	Gene
1	cg05491587	-0.6227	2.1588	1.20E-06	KCNG2
2	cg04695077	-0.2133	-1.8804	1.99E-06	
3	cg25289484	0.3696	3.0323	2.46E-06	SPEN
4	cg23296325	0.1823	-4.1634	2.61E-06	NAA30
5	cg05646885	-0.1646	0.1278	4.33E-06	
6	cg14476293	0.2551	2.9945	4.38E-06	MIR548F3
7	cg15986164	-0.1875	3.7964	4.50E-06	GAK
8	cg23504411	0.1967	-1.1872	4.70E-06	HIST1H3J
9	cg16688303	0.2482	0.6503	4.72E-06	LRRTM2;CTNNA1
10	cg18765801	0.1858	3.2942	5.11E-06	MYADM
11	cg23653184	0.1669	0.4235	5.29E-06	
12	cg17256711	0.2258	-1.6588	5.34E-06	OPRM1
13	cg01551350	-0.1850	-1.3218	6.47E-06	
14	cg24004990	-0.1077	-0.2035	7.01E-06	
15	cg13128436	-0.1270	-1.9972	7.50E-06	MECR
16	cg00136105	0.1972	-3.9058	7.55E-06	GNPTAB
17	cg13002170	-0.2043	3.8477	7.71E-06	NCOA2
18	cg19056004	-0.1499	0.4453	7.73E-06	LRRC23;ENO2
19	cg22037687	0.1626	3.3374	8.05E-06	NFATC2
20	cg06647382	0.1562	3.7626	9.10E-06	PDK2
21	cg00553886	0.2007	-1.3150	9.30E-06	TEAD1
22	cg01666716	0.1389	2.5179	9.43E-06	
23	cg22989843	0.2166	-2.4029	9.83E-06	PAX3

Model 3b differs from Model 3a by the identification of N(alpha)-acetyltransferase 30, NatC catalytic subunit (*NAA30*); pyruvate dehydrogenase kinase 2 (*PDK2*) and cyclin G associated kinase (*GAK*); myeloid associated differentiation marker (*MYADM*); mitochondrial trans-2-enoyl-CoA reductase (*MECR*); leucine rich repeat containing 23 (*LRRC23*); enolase 2 (*ENO2*); microRNA 548f-3 (*MIR548F3*); and a nuclear factor of activated T cells 2 (*NFATC2*).

Model 3a and Model 3b coincide with 10 genes mapped to probes with p < 10E-05. They share the top two genes *KCNG2* and *SPEN*, with *KCNG2* slightly more significant in Model 3a. *P*-values and fold-change values differ slightly for *HIST1H3J*, *LRRTM2*, *CTNNA1*, *OPRM1*, *GNPTAB*, *NCOA2*, *TEAD1* and *PAX3.* The average of probe intensities is the same for both models (Table 5.4 and Table 5.5).

Shared genes between Model 3a and Model 3b

An analysis was made of all genes shared between the models (*p*-value < 10E-03) to evaluate any differences after adjusting for estimated cell counts in Model 3b. The total number of shared genes (n = 959) represented 53.3% of the total number of genes. These genes were visualized in a Venn diagram (Figure 5.5).

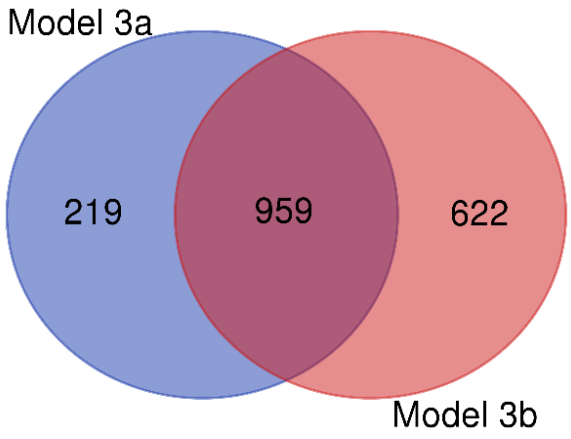


Figure 5.5. Unique genes shared between Model 3a and Model 3b annotated by the EPIC annotation library (*p*-value < 10E-03)**:** Model 3a (n = 1178), Model 3b (n = 1581). Genes shared between the two models (n = 959) represented 53.3% of the total number of genes (n = 1800) for both models. Venn diagram generated with webtool created by the Van de Peer lab of Bioinformatics and Evolutionary Genomics (http://bioinformatics.psb.ugent.be/beg/tools/venn-diagrams).

***Hypo* and *hyper*-methylated genes shared between the models**

An analysis was then made of the shared *hypo*- and *hyper*- methylated genes to evaluate if there was a general methylation pattern that might be associated with olanzapine (Figure 5.6). A methylation score was calculated as the inverse of the *p*-value for each gene. Genes with negative scores indicated *hypo*-methylated status, while genes with positive scores indicated *hyper*-methylated status. The number of genes from each state and model is presented in Table 5.6 and represented in Venn diagrams in Figure 5.6.

Table 5.6 *Hypo*- and *hyper*-methylated genes coincide between Model 3a and Model 3b

State	Model 3a	Model 3b	Shared	Unique
hypo-methylated	435	683	352	766
hyper-methylated	743	898	590	1051

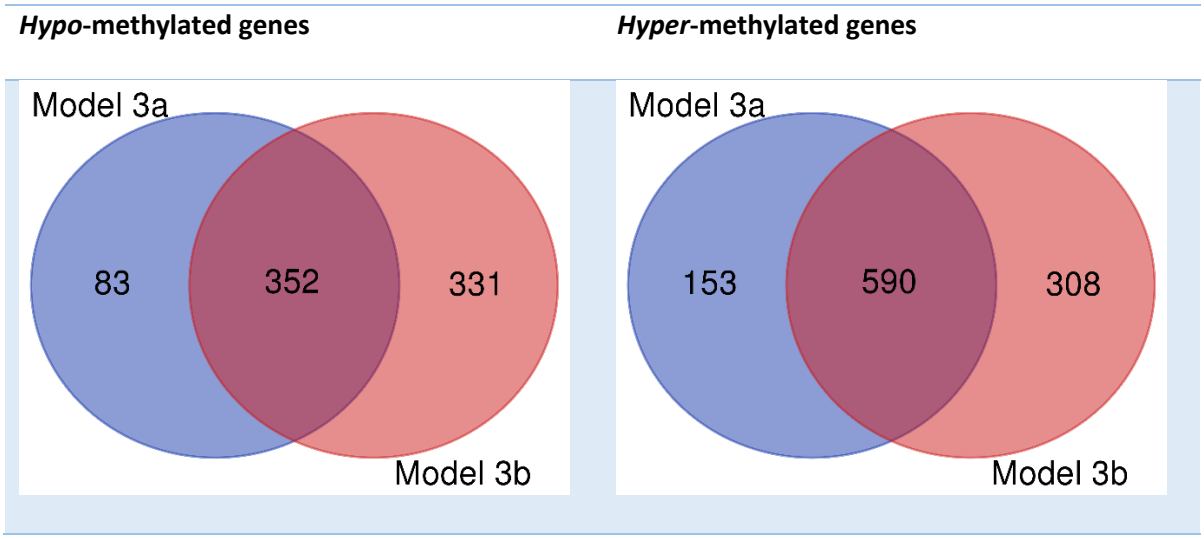


Figure 5.6. Shared *hypo*- and *hyper*-methylated genes as measured by inverse *p*-value < 10E-03. Model 3a and Model 3b share 766 *hypo*-methylated genes and 1051 *hyper*-methylated genes. Venn diagram generated with webtool created by the Van de Peer lab of Bioinformatics and Evolutionary Genomics (http://bioinformatics.psb.ugent.be/beg/tools/venn-diagrams).

The data in Table 5.6 indicate that a greater number of genes that coincide between the models are *hyper*-methylated (63%) compared to *hypo*-methylated (37%). A complete understanding of the role of methylation status on gene expression is still under investigation. Nevertheless, identification of the genomic coordinates of richly methylated regions may provide greater insight into the gene regulatory mechanisms at these loci.

5.2.3 Differentially methylated regions

The aim of EWAS is to identify differentially methylated regions that are significantly different in methylation status between tissues or individuals. These densely methylated regions have more statistical power than individual CpGs because they represent persistent change across a region.

Biological interpretation is therefore enhanced by the identification of a correlation between differentially methylated regions and phenotype.

R-packages for identifying differentially methylated regions included **DMRcate,** computation of genomic ranges with **GenomeRanges,** annotation of files in **AnnotationHub,** and subsequent visualization in **Gviz** (Table 4.1).

Model 3a (y ~ AP1 + gender + smoker): Three differentially methylated regions (DMRs) were identified with a false discovery rate (FDR 0.4). As there were no findings with FDR 0.3, further evaluation was discontinued with Model 3a.

Model 3b (y ~ AP1 + gender + smoker + eCells): Fifteen DMRs were identified with false discovery rate (FDR 0.3). Two DMRs were evaluated for further analysis: one *hyper*-methylated DMR at the Trio and F-actin binding protein (*TRIOBP*) on Chr. 22q13 (Figure 5.4), and one *hypo*-methylated DMR at the transcription factor Sry-box 30 (*SOX30*) on Chr. 5q33.3 (Figure 5.5).

In Figure 5.7, methylation levels for *TRIOBP* CpGs (pink) were higher in patients taking olanzapine, than methylation levels in patients taking aripiprazole or quetiapine (blue).

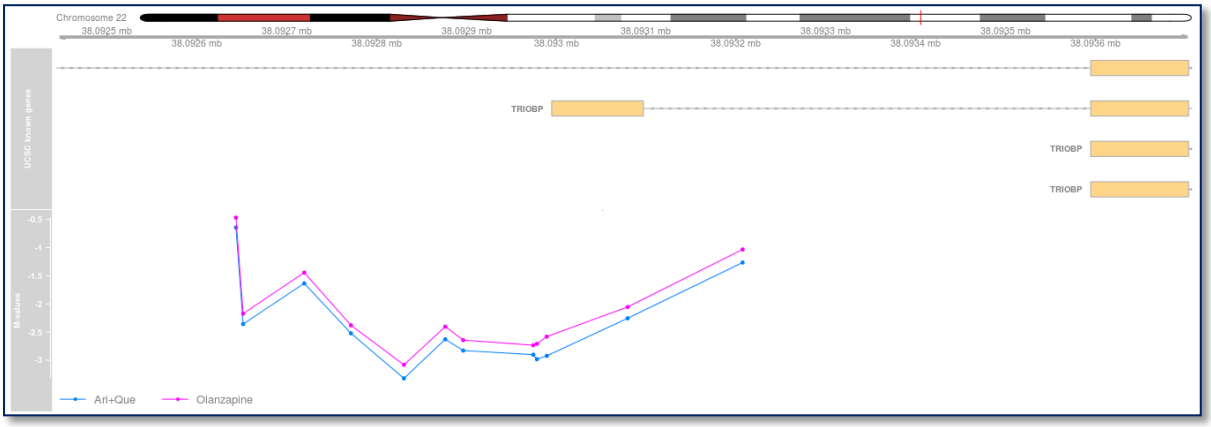


Figure 5.7. DMR identified (FDR 0.3) on Chr. 22q13: *Hyper*-methylation of *TRIOBP* at most CpGs was identified in patients taking olanzapine (pink) compared to patients taking aripiprazole or quetiapine (blue).

In contrast, methylation levels for most CpGs at the *SOX30* gene were lower in patients taking olanzapine (pink) than in patients taking aripiprazole or quetiapine (blue) (Figure 5.8).

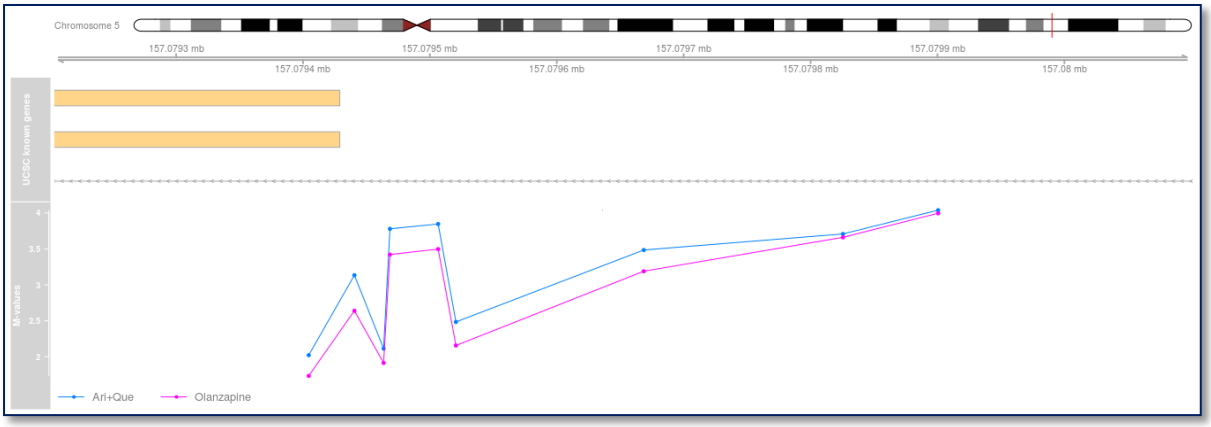


Figure 5.8. DMR identified (FDR 0.3) on Chr. 5.q33.3: *Hypo*-methylation of *SOX30* identified at most CpGs in patients taking olanzapine (pink) compared to patients taking aripiprazole or quetiapine (blue).

The biological interpretation of these DMRs may be evaluated in relation to their role in cellular or metabolic pathways. Identification of significant pathways may suggest the mediating effect of olanzapine.

5.3 Gene set over-representation analysis

Gene set over-representation analysis enhances the biological interpretation of epigenetically modified genes by using statistical methods to map gene sets in curated databases to our user-defined gene list. Over-representation analysis determines whether a set of identified gene categories are present more than would be expected by chance in our set of differentially methylated genes. Gene Ontology (GO) categories are used for this purpose. The approach is then extended to identify pathways containing over-represented gene sets that map to our list. ConsensusPathDB accesses both GO and publicly available pathway databases for this purpose (Table 5.3).

Gene lists with *p*-value < 10E-03 prepared from Model 3a (n = 1178) and Model 3b (n= 1581) were submitted to ConsensusPathDB with default settings (level 2 and 3 GO categories with *p*-value cut-off = 0.01).

Gene set over-representation analysis: The output of the database search in ConsensusPathDB provided a list of Gene Ontology (GO) categories and term IDs which classify different aspects of gene function. The top 5 gene sets identified in each category: Biological process (Table 5.6), Molecular function (Table 5.7) and Cellular component (Table 5.8) were organized by corresponding log10 *p*-value for each model.

Terms relating to **biological process** in GO describe cellular processes that require multiple molecular activities. Gene sets characterized by cellular adhesion, development and structure are present in the top 5 hits for both models. The adjustment for cell type composition in Model 3b may have resulted in slightly more significant *p*-values than in Model 3a (Table 5.6).

Table 5.6 Comparison of over-represented gene sets classified by GO terms for biological process

Biological process					
Model 3a			**Model 3b**		
Term ID	**Description**	**log10 *p*-value**	**Term ID**	**Description**	**log10 *p*-value**
GO:0007155	Cell adhesion	-7.502	**GO:0048856**	Anatomical structure development	-8.083
GO:0048468	Cell development	-6.519	**GO:0048468**	Cell development	-7.863
GO:0048731	System development	-6.491	**GO:0007155**	Cell adhesion	-7.498
GO:0007275	Multicellular organism development	-6.490	**GO:0007275**	Multicellular organism development	-7.010
GO:0048856	Anatomical structure development	-6.176	**GO:0048731**	System development	-6.848

Terms relating to **molecular function** in GO are characterized by the activities of gene products that require multiple molecular activities performed on the molecular level. Gene sets characterized by the binding of proteins and enzymes are represented in both models, although Model3a has 2 terms that are more specific for DNA binding (Chromatin binding and RNA polymerase II transcription factor activity, sequence-specific DNA binding (Table 5.7).

Table 5.7 Comparison of over-represented gene sets classified by GO terms for molecular function

Molecular function					
Model 3a			**Model 3b**		
Term ID	**Description**	**log10 *p*-value**	**Term ID**	**Description**	**log10 *p*-value**
GO:0005515	Protein binding	-6.731	**GO:0005515**	Protein binding	-6.445
GO:0003682	Chromatin binding	-4.213	**GO:0019899**	Enzyme binding	-6.100
GO:0019899	Enzyme binding	-4.091	**GO:0005088**	Ras guanyl-nucleotide exchange factor activity	-4.791
GO:0000981	RNA pol. II TF activity, sequence-specific DNA binding	-3.969	**GO:0019904**	Protein domain specific binding	-3.852
GO:0044877	Macromolecular complex binding	-3.734	**GO:0008092**	Cytoskeletal protein binding	-3.309

- transcription factor (TF)

Terms relating to **cellular component** in GO are characterized by the locations where a gene product performs a function in relation to the cellular structure. In Model 3b, all five gene sets are characterized by the neuron, and include specifically the cell body, dendrites, synapses and postsynaptic components such as neurotransmitter receptors. Model 3a shares 3 of these gene sets however the log10 *p*-values are lower than the values in Model 3b (Table 5.8).

Table 5.8 Comparison of over-represented gene sets classified by GO terms for cellular component

Cellular component					
Model 3a			**Model 3b**		
Term ID	**Description**	**log10 *p*-value**	**Term ID**	**Description**	**log10 *p*-value**
GO:0098794	Postsynapse	-5.420	**GO:0097458**	Neuron part	-7.333
GO:0097458	Neuron part	-4.854	**GO:0036477**	Somatodendritic compartment	-5.979
GO:0036477	Somatodendritic compartment	-4.295	**GO:0098794**	Postsynapse	-5.886
GO:0043296	Apical junction complex	-4.270	**GO:0014069**	Postsynaptic density	-5.492
GO:0019897	Extrinsic component of plasma membrane	-4.001	**GO:0099572**	Postsynaptic specialization	-5.398

The over-represented gene sets characterized by GO terms identified categories of gene functions where the modifying effects of olanzapine may be located. These gene sets indicated activity localized to the neuron in synaptic receptors. Gene sets involved in the process of cell adhesion and structural development were also over-represented. Pathway-based gene sets were then evaluated to see which pathways were implicated our differentially methylated gene sets.

Pathway-based over-representation analysis: The top pathways identified by the gene lists submitted to ConsensusPathDB provided additional information for comparison of the two models as seen in Table 5.9 (Model 3a) and Table 5.10 (Model 3b).

Table 5.9 Model 3a: 11 pathway-enriched gene sets identified with *p* < 10E-03

Pathway	*p*-value	q-value	Database	External_id
Hippo signaling pathway	9.85E-06	9.72E-03	KEGG	path:hsa04390
Posttranslational regulation of adherens junction stability and dissassembly	1.08E-05	9.72E-03	PID	ajdiss_2pathway
White fat cell differentiation	1.54E-04	6.37E-02	Wikipathways	WP4149
Syndecan-2-mediated signaling events	1.77E-04	6.37E-02	PID	syndecan_2
Cardiac Progenitor Differentiation	2.94E-04	8.82E-02	Wikipathways	WP2406
Mitochondrial tRNA aminoacylation	4.81E-04	1.19E-01	Reactome	R-HSA-379726
Wnt Signaling Pathway and Pluripotency	5.83E-04	1.19E-01	Wikipathways	WP399
TCR signaling in naive CD4+ T cells	5.97E-04	1.19E-01	PID	tcr_pathway
Transcriptional activation by NRF2	6.63E-04	1.19E-01	Wikipathways	WP3
MAPK Signaling Pathway	7.86E-04	1.22E-01	Wikipathways	WP382
Pathways in cancer - Homo sapiens	8.16E-04	1.22E-01	KEGG	path:hsa05200

Pathways associated with signaling and cell differentiation were identified in Model 3a. The MAPK and Wnt signaling pathways are involved in cell proliferation, differentiation and cell migration. T cell receptor (TCR) signaling mediates diverse functional outcomes in naïve versus memory CD4 T cells.

Table 5.10 Model 3b: 9 pathway-enriched gene sets identified with *p* < 10E-03

Pathway	*p*-value	q-value	Source	External_id
Hippo signaling pathway	1.44E-06	2.86E-03	KEGG	path:hsa04390
RUNX1 regulates transcription of genes involved in WNT signaling	1.41E-04	1.40E-01	Reactome	R-HSA-8939256
Neuronal System	2.66E-04	1.76E-01	Reactome	R-HSA-112316
RUNX3 regulates p14-ARF	4.06E-04	2.00E-01	Reactome	R-HSA-8951936
Protein-protein interactions at synapses	5.21E-04	2.00E-01	Reactome	R-HSA-6794362
TP53 regulates transcription of additional cell cycle genes whose exact role in the p53 pathway remain uncertain	7.59E-04	2.00E-01	Reactome	R-HSA-6804115
RUNX3 regulates YAP1-mediated transcription	8.73E-04	2.00E-01	Reactome	R-HSA-8951671
G alpha (12/13) signaling events	9.00E-04	2.00E-01	Reactome	R-HSA-416482
FCERI mediated NF-kB activation	9.88E-04	2.00E-01	Reactome	R-HSA-2871837

The top pathway (*p*-value < 10E-05) for both Model 3a and Model 3b was the Hippo signaling pathway. This pathway is involved in the regulation of cell proliferation and apoptosis. The difference in nominal *p*-values and q-values between the models was less than expected considering the adjustment for cell type composition in Model 3b.

Pathways identified in Model 3b related to the neuronal system and synaptic mediated neurotransmitter activity. Intercellular communication and cell-cell adhesion also occur at synapses mediated by protein-protein interactions. G(12/13) signaling pathway in involved in the regulation of cellular proliferation and NF-kB activation by FCERI is necessary for proinflammatory cytokine production and allergic inflammatory disease process (Fabregat et al. 2018).

In order to verify the olanzapine-mediated changes in methylation patterns found in the blood, a cell culture trial was performed with hepatic HepG2 cells. The human hepatic HepG2 cell line is widely used in drug metabolism and hepatoxicity studies. These cells offer a stable phenotype that does not depend on donor characteristics (Donato, Tolosa, & Gómez-Lechón, 2015).

5.4 Verification of epigenetic modifications in blood using olanzapine-exposed HepG2 cells

Three different analytical-approaches were performed with the hepatic cell line (HepG2) to identify drug-induced effects following a 72-hour exposure to olanzapine. Gene expression was evaluated with genes known to alter fatty acid synthesis. DNA isolated from the exposed cells was purified and methylation status was assayed on the EPIC microarray. The methylation data was then evaluated in *silico* using the same pipeline as with blood samples (see Appendix). Prior to methylation analysis, the quality of the isolated DNA was assessed for purity, concentration and integrity.

5.4.1 Analytical agarose gel electrophoresis

Agarose gel electrophoresis is a technique used for separating linear DNA molecules by size in an electric field. Negatively charged DNA molecules migrate towards the positive electrode and smaller fragments will migrate faster within the gel. A standard analytical agarose gel was performed to confirm the integrity of the DNA prior to sending the samples for methylation analysis. In Figure 5.9, all samples corresponded to circa 6 000 base pairs without visible smaller DNA fragments. The clarity of the gel indicated that the double-stranded DNA remained intact following the purification procedure.

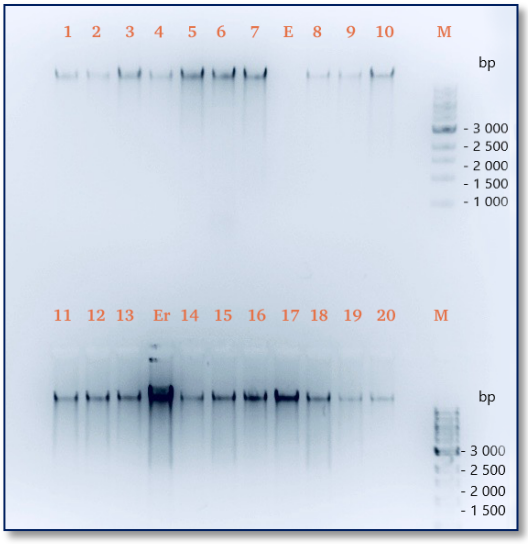


Figure 5.9. Analytical agarose gel (1% w/v) electrophoresis confirms integrity of DNA samples. Lanes 1 – 4 (No treatment), lanes 5 – 8 (Vehicle – DMSO), lanes 9 – 12 (1 µM), lanes 13 – 16 (50 µM) and lanes 17 – 20 (100 µM). Lane M: 6 000 bp DNA Marker (25 ng). Lane E is empty and lane Er is a highly concentrated sample.

5.4.2 Validation of gene response to olanzapine following 72-hour exposure

Quantitative real-time PCR (qPCR) was then performed to measure the fold change in gene expression. Genes known to induce alterations in gene expression following exposure to second-generation drugs were selected due to evidence of their role in fatty acid synthesis. Lipid biogenesis may be induced by activating the sterol regulatory element binding protein (SREBP), encoded by sterol regulatory element binding transcription factor 1 (*SFEBF1*) (Fernø, Skrede, Vik-Mo, Håvik, & Steen, 2006), as well as by the fatty acid synthase (*FASN*) gene, shown to be upregulated by olanzapine treatment (Vik-Mo, Ferno, Skrede, & Steen, 2009). A third gene known to induce lipid biogenesis is the carbohydrate-response element binding protein (*ChREBP*) in response to glucose, although it has been shown to be induced by a different antipsychotic drug and not olanzapine (Liu, 2017).

Expression of *SREBF1* (Figure 5.10-A) shows a clear dose response (50 µM olanzapine average fold change ± standard error of mean (SEM), 1,52 ± 0,26, *p*-value:0,03), 100 µM olanzapine (3,26 ± 0,76, *p*-value: 0,02). A dose trend was seen also with 1 µM olanzapine (*p*-value: 0,07) however it did not pass the significance level. The dose response for *FASN* is noticeable only between 50 µM and 100 µM however was not significant (data not shown). *ChREBPα* expression showed greater variability amongst all samples and was not significant (data not shown).

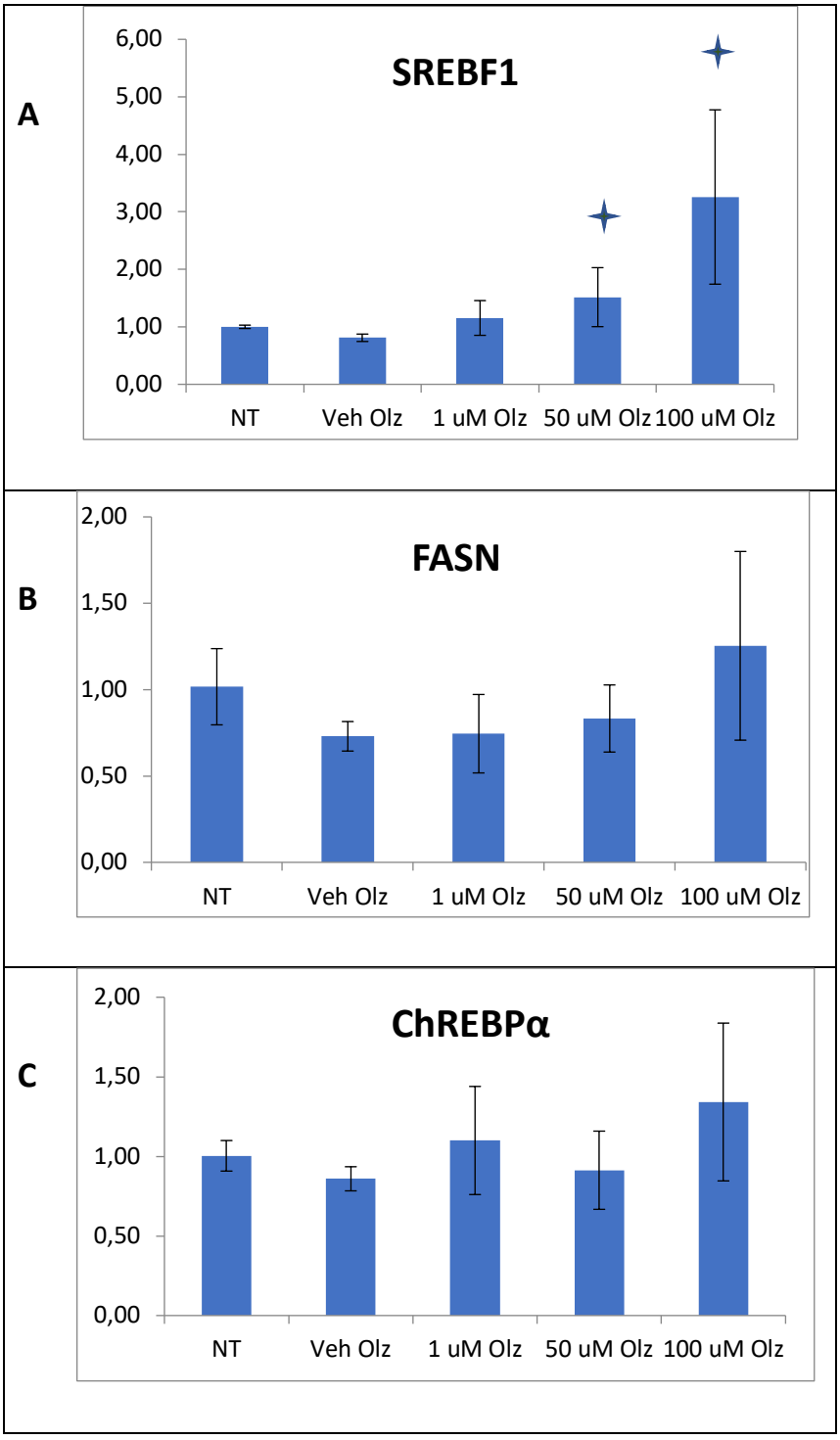


Figure 5.10 Gene expression following 72 hours of Olanzapine (Olz) exposure: A) Expression of *SREBF1* followed a dose response trend response (50 µM olanzapine average fold change ± SEM, 1,52 ± 0,26, *p*-value: 0,03), 100 µM olanzapine (3,26 ± 0,76, *p*-value: 0,02) although not significant at 1 µM Olz; B) The dose response for *FASN* is noticeable only between 50 µM and 100 µM however was not significant (data not shown); C) Expression of *ChREBPα* showed greatest variability amongst all samples however not significant. NT: ENEM medium; Vehicle: DMSO.

5.4.4 Differentially methylated regions after olanzapine exposure in HepG2

Nine DMRs were identified in HepG2 cells (FDR 0.3) containing between 2 - 7 CpGs. The DMR shown in Figure 5.12 consists of 7 CpGs and overlaps with the ALKBH3 antisense RNA1 (*ALKBH3-AS1*) located on the minus strand of chromosome 11p11.2. The average methylation levels do not differ for the 100 µM samples (blue), except for one CpG which is *hypo*-methylated in comparison to Vehicle-DMSO (pink).

Overlapping with this gene, but on the opposite strand is ALKBH3, which belongs to a family of proteins that mediates DNA damage repair: Alkylated DNA Repair Protein AlkB Homolog 3 (*ALKBH3*). The protein protects against cytotoxicity of methylating agents by demethylating DNA and RNA containing 1- methyladenosine (m1A) (Duncan et al., 2002).

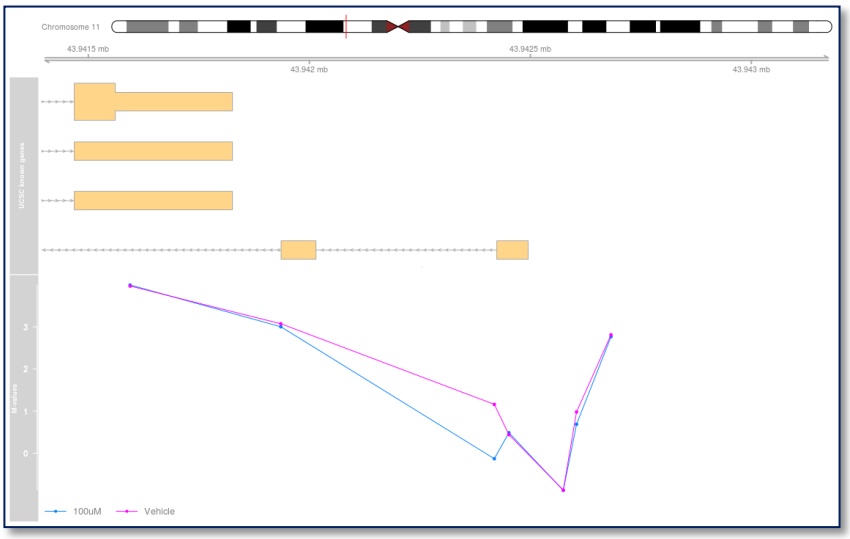


Figure 5.12. DMR identified in 100 µM samples (FDR 0.3) on Chr. 11.p11.2: The DMR identified for *ALKBH3-AS1* spans 7 CpG sites where only one CPG is markedly *hypo*-methylated in 100 µM samples (blue) compared to Vehicle-DMSO (pink). The genomic coordinates: chr11: 43941594-43942682.

5.4.3 Differentially methylated probes identified following olanzapine-exposure

The methylation data was for HepG2 cells was then assayed using **limma** to identify differentially methylated probes following 72-hour exposure to olanzapine. The genome-wide statistical analysis for the 100 µM concentration samples was selected for downstream work as a greater effect response was anticipated at this concentration. The top 21 differentially methylated probes (p-value < 10E-07) and their associated genes, mapped using the **Epic annotation library** are presented in Table 5.9.

These genes included: Rho associated coiled-coil containing protein kinase 2 (*ROCK2*); protein activator of interferon induced protein kinase EIF2AK2 (*PRKRA*); tight junction protein 2 (*TJP2*); F-box

protein 31 (*FBXO31*); EMAP like 1 (*EML1*); TOX high mobility group box family member 3 (*TOX3*); ATPase phospholipid transporting 11A (*ATP11A*); ring finger protein 145 (*RNF145*); CTD small phosphatase2 (*CTDSP2*); nucleoporin 93 (*NUP93*); zinc finger AN1-type containing 4 (*ZFAND4*); zinc finger and BTB domain containing 10 (*ZBTB10*); zinc finger protein 557 (*ZNF557*); coiled-coil alpha-helical rod protein 1 (*CCHCR1*); long intergenic non-protein coding RNA 1947 (*CTB-7E3.1*); microRNA 548n (*MIR548N*); transcription factor 19 (*TCF19*); BCL6, transcription repressor (*BCL6*); atonal bHLH transcription factor 8 (*ATOH8*); activating transcription factor 7 (*ATF7*); and core-binding factor subunit beta (*CBFB*).

Table 5.9 Top 21 probes from 100 µM samples with nominal *p*-value < 10E-07 and associated genes. Negative average expression of the probes is colored light pink while positive expression is colored light green. Genes were mapped using overlap at probes 5 and 9.

	Probe	logFC	AveExpr	P-Value	Gene
1	cg12394201	-1.2887	0.9224	8.99E-13	
2	cg12771518	-0.6895	-0.6113	1.08E-09	ROCK2
3	cg09211893	1.3579	-3.5961	2.45E-09	EML1
4	cg11389179	1.3891	-4.1691	1.30E-08	ZFAND4
5	cg13151361	-0.6477	1.0152	1.61E-08	PRKRA;MIR548N
6	cg01571169	-0.6319	2.6449	1.62E-08	FBXO31
7	cg03876705	0.9820	-2.0613	1.74E-08	TOX3
8	cg19742870	-0.6083	1.0162	2.09E-08	TJP2
9	cg00080118	0.7573	-3.1322	3.67E-08	CCHCR1;TCF19
10	cg01744729	-0.8273	3.2903	4.50E-08	ATP11A
11	cg03271007	-0.5595	1.2194	5.06E-08	RNF145
12	cg23087770	-1.1663	-0.3865	6.42E-08	
13	cg10091025	-0.5340	0.7108	6.52E-08	BCL6
14	cg02464123	0.5491	-2.5533	6.84E-08	CTDSP2
15	cg19956166	0.8559	-3.1523	7.25E-08	ATOH8
16	cg03319894	-0.9717	1.2080	7.77E-08	ZBTB10
17	cg23415810	-0.7549	1.2225	8.75E-08	ATF7
18	cg12270595	0.7824	-0.5593	8.90E-08	NUP93
19	cg04508379	0.5708	-3.6236	8.95E-08	ZNF557
20	cg09000178	1.7868	-4.2566	9.44E-08	CBFB
21	cg09733789	-0.8264	1.0838	9.86E-08	CTB-7E3.1

Box plots were also made for the top 16 probes using **ggplot2** as shown in Figure 5.11. The aim of this figure is to present the methylation values (M-values) as box plots in order to compare each exposure condition to 100 µM samples. The average methylation level of the top 16 probes for 100 µM generally tended to be lower than for 50 µM probes, except for 6 probes with associated genes: *EML1*, *ZFAND4*, *TOX3*, *CCHCR1*, *CTDSP2*, and *ATOH8*. The 50 µM probes had lower methylation levels in 4

probes with associated genes: *NA*, *ROCK2*, *ZFAND4 and BCL6,* compared to 1 µM. There was generally little variation with 1 µM probes compared to Vehicle or NT. These probes tended to share a similar pattern of expression.

Figure 5.11 The top 16 significant DMPs identified in HepG2 cells following 100 µM 72 hours with either no treatment (NT), Vehicle, or concentration of exposure to olanzapine. The average expression (AveExpr) of the probes are represented as box plots with NT (orange), Vehicle (olive), 1 µM (green), 50 µM (blue), or 100 µM (pink) olanzapine on the X-axis and the M-value on the Y-axis (note values are different for each boxplot). The gene names associated with each probe are listed beside the probe name.

Despite the lower methylation levels in 100 µM probes compared to 50 µM probes, the difference in methylation levels between 100 µM and Vehicle was larger than the difference between 50 µM and Vehicle. The downstream analysis was continued with the 100 µM samples.

5.4.5 Meta-analysis Model 3b and HepG2

To see how the results obtained from the analysis in the patient blood samples using Mod3b compared to the results obtained from the olanzapine-exposed HepG2 cells, a meta-analysis was performed between the results from the two analyses. The meta-analysis will enrich for differentially

methylated positions that are ranked high on the lists of either analysis, and where the direction of change goes in the same direction. The top 34 probes *p*-value < 10E-05 is shown in Table 5.10. The direction of methylation pattern between Mod3b and HepG2 is indicated as (++) for *hyper*-methylated and (- -) for *hypo*-methylated.

Table 5.10 Top 34 probes *p*-value < 10E-05 from meta-analysis with direction of methylation pattern

No.	Probe	P.value	Direction	Gene
1	cg12698713	1.27E-07	?-	NA
2	cg04695077	1.54E-07	--	
3	cg01551350	9.46E-07	--	
4	cg23504411	1.06E-06	++	HIST1H3J
5	cg08525481	1.36E-06	--	OGFR
6	cg12412751	1.57E-06	++	SBNO2
7	cg09593028	1.70E-06	++	ZNF518A
8	cg23724384	1.96E-06	++	
9	cg18730746	2.12E-06	--	TEX264
10	cg06914048	2.45E-06	++	CABLES1
11	cg05646885	3.01E-06	--	
12	cg25289484	3.26E-06	++	SPEN
13	cg23653184	3.27E-06	++	
14	cg24004990	3.33E-06	--	
15	cg27408541	3.86E-06	++	TANC1
16	cg01666716	3.86E-06	++	
17	cg06647382	3.88E-06	++	PDK2
18	cg11068289	3.99E-06	--	
19	cg17256711	4.11E-06	++	OPRM1
20	cg13756796	4.67E-06	++	LOC102724421
21	cg22989843	4.78E-06	++	PAX3
22	cg00799842	4.98E-06	?+	NA
23	cg22678739	6.19E-06	--	
24	cg00947599	6.19E-06	--	GNA12
25	cg19786998	6.46E-06	++	PRTG
26	cg02474628	6.48E-06	++	WDFY1
27	cg21908248	6.52E-06	++	PPP1R15B
28	cg18541417	6.74E-06	++	RASGRP1
29	cg00236832	7.00E-06	++	RARA
30	cg10266418	7.32E-06	--	CMTM3
31	cg23149016	8.15E-06	--	SP7
32	cg05008948	8.48E-06	--	ZNF655
33	cg23265096	9.37E-06	--	CTSZ
34	cg18765801	9.66E-06	++	MYADM

Despite the differences in olanzapine concentration between patient blood samples and 100 µM-exposed HepG2 cells, the pattern of methylation in the identified genes is concordant between the two groups in the top 34 probes. The genes identified as *hypo*-methylated included: opioid growth factor receptor (*OGFR*); testis expressed 264 (*TEX264*); G protein subunit alpha *12* (*GNA12*); CKLF like MARVEL transmembrane domain containing 3 (*CMTM3*); Sp7 transcription factor (*SP7*); zinc finger protein 655 (*ZNF655*); and cathepsin Z (*CTSZ*).

The genes identified as *hyper*-methylated included: histone cluster 1 H3 family member j (*HIST1H3J*); strawberry notch homolog 2 (*SBNO2*); zinc finger protein 518A (*ZNF518A*); Cdk5 and Abl enzyme substrate 1 (*CABLES1*); spen family transcriptional repressor (*SPEN*); paired box 3 (*PAX3*); tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 1 (*TANC1*); pyruvate dehydrogenase kinase 2 (*PDK2*); opioid mu receptor 1 (*OPRM1*); protogenin (*PRTG*); WD repeat and FYVE domain containing 1 (*WDFY1*); protein phosphatase 1 regulatory subunit 15B (*PPP1R15B*); RAS guanyl releasing protein 1 (*RASGRP1*); retinoic acid receptor alpha (*RARA*); myeloid associated differentiation marker (*MYADM*); and uncharacterized (*LOC102724421*).

Pathway-based gene set over-representation of meta-analysis: Model 3b and HepG2 *p < 10E-03*

An analysis was made on the pathway level of the genes enriched in the meta-analysis between Model 3b and HepG2. A list of 2158 probes and their associated genes were submitted to ConsensusPathDB with *p*-value < 10E-03. The top four pathways are shown in Table 5.11.

Table 5.11 Validation of over-represented pathway-based genes in meta-analysis: Model 3b and HepG2

Pathway	*p*-value	q-value	Source
Extracellular matrix organization	1.49E-05	3.12E-02	Reactome
Hippo signaling pathway - Homo sapiens	4.39E-05	4.61E-02	KEGG
L1CAM interactions	1.54E-04	5.36E-02	Reactome
DNA Damage Response (only ATM dependent)	1.67E-04	5.36E-02	Wikipathways

The extracellular matrix performs many dynamic functions and is able to influence cell behaviours including proliferation, adhesion and migration (Hynes, 2014). The *L1CAM* family cell adhesion molecules have been implicated in processes integral to nervous system development, including inter-neuronal adhesion. They are predominately expressed by neuronal cells during development

(Yu, Yang, Fu, & Jin, 2016). In disease, the ATM dependent DNA damage response is involved in signal transduction pathway deregulation (Berger, Stanley, Moore, & Goodarzi, 2017).

5.5 Concordance of identified differentially methylated regions in brain

The blood samples from patients adhering to olanzapine treatment were used as a surrogate for brain tissue to identify epigenome-wide modifications to the DNA. Although DNA methylation is tissue-specific, many CpGs show concordance between blood and brain tissue. Differentially methylated regions in *TRIOBP* and *SOX30* genes were therefore evaluated for their concordance using the Blood-brain Epigenetic Concordance (**BECon**) method (Edgar et al., 2017). This method evaluates the concordance (measured as variability) of CpGs between blood and Brodmann Area 10 (anterior prefrontal cortex), Brodmann Area 20 (temporal cortex), and Broadmann Area 7 (parietal cortex).

5.51 Trio and F-actin Binding Protein (*TRIOBP*)

Twenty-five of the 27 CpGs identified for *TRIOBP* by Edgar et al. (2017) were present in our data set, while only two of the CpGs had *p*-value < 10E-03 (cg03360992, *p*-value 2.57E-05 and cg12379720, *p*-value 8.36E-04). The top 12 CpGs showed inter-individual variability between three brain regions (Brodmann Areas 10, 20 and 7) and peripheral blood samples (Figure 5.13).

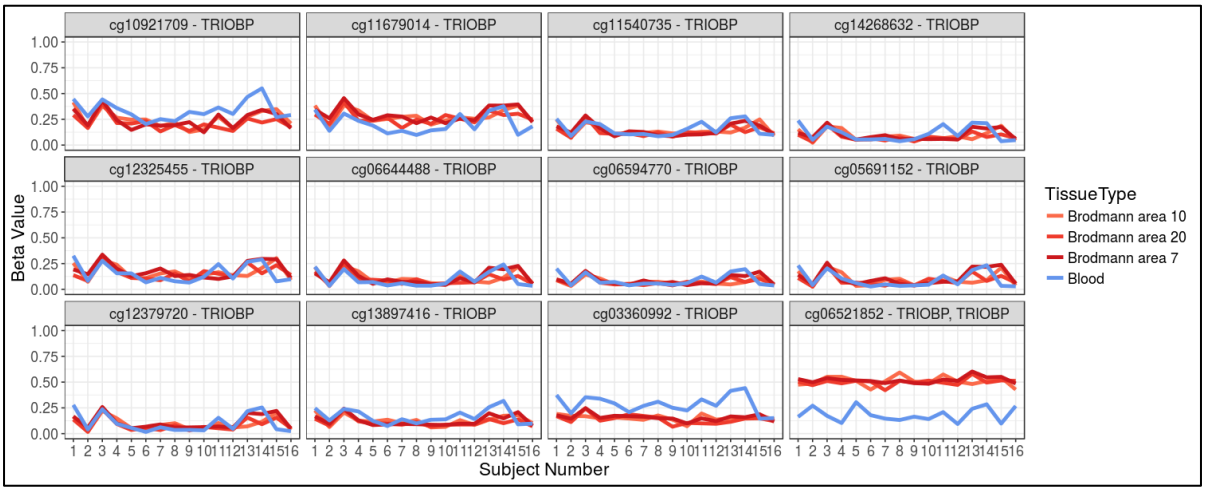


Figure 5.13 CpGs associated with *TRIOBP* show inter-individual variability between brain and blood at BA10 (orange), BA20 (pink), BA7 (red) and blood (blue) as evaluated by the BECon method (Edgar et al., 2017). 27 concordant CpGs were identified by submitting the *TRIOBP* gene name to the Shiny app https://redgar598.shinyapps.io/BECon/. Note probes cg03360992 *p*-value 2.57E-05 and cg12379720 *p*-value 8.36E-04). The x-axis represents the inter-individual variability while the y-axis represents the Beta values of the methylation data.

The top 12 CpGs ranked by variability, correlation and cell composition in brain and blood are presented in Figure 5.14. Variability and correlation metrics are calculated from cell composition adjusted data. Variability (green) is the reference range of methylation beta values between the 10th – 90th percentile of all samples. Correlation is represented as Spearman correlation values of methylation between blood and the brain area, with darker gray representing lower correlations and darker orange higher correlations. The cell composition metrics represent how much the beta values change on average at a CpG with the cell composition adjustment.

The top 10 CpGs showed 90% positive correlation values (0.51-0.81) in BA10, BA20, and BA7. The beta values changed at a given CpG on average 4% in 75-90% of blood cells, and on average 2% in 50-90% of brain cells (Figure 5.14).

					Variability				Correlation			Cell Composition	
	Chr	Coor	Gene(s)	Gene Region(s)	BA10	BA20	BA7	Blood	BA10	BA20	BA7	Blood	Brain
cg10921709	22	38092643	TRIOBP	promoter	0.21	0.14	0.19	0.22	0.59	0.45	0.6	0.05	0.01
cg11679014	22	38092719	TRIOBP	promoter	0.17	0.12	0.17	0.23	0.52	0.54	0.6	0.05	0.02
cg11540735	22	38092771	TRIOBP	promoter	0.14	0.1	0.13	0.16	0.69	0.81	0.72	0.04	0.02
cg14268632	22	38092830	TRIOBP	promoter	0.13	0.08	0.12	0.17	0.44	0.65	0.66	0.04	0.02
cg12325455	22	38092876	TRIOBP	promoter	0.17	0.17	0.18	0.21	0.66	0.66	0.67	0.04	0.03
cg06644488	22	38092896	TRIOBP	promoter	0.16	0.11	0.17	0.17	0.61	0.74	0.68	0.03	0.01
cg06594770	22	38092974	TRIOBP	promoter	0.07	0.07	0.11	0.15	0.41	0.56	0.61	0.02	0.01
cg05691152	22	38092978	TRIOBP	promoter	0.16	0.12	0.18	0.19	0.61	0.7	0.7	0.03	0.02
cg12379720	22	38092989	TRIOBP	promoter	0.14	0.12	0.15	0.22	0.67	0.59	0.71	0.04	0.02
cg13897416	22	38093079	TRIOBP	promoter	0.13	0.07	0.12	0.16	0.51	0.6	0.61	0.04	0.01
cg03360992	22	38093207	TRIOBP	promoter	0.05	0.08	0.06	0.22	0.39	0.15	0.34	0.04	0
cg06521852	22	38141419	TRIOBP, TRIOBP	intragenic, promoter	0.11	0.04	0.06	0.18	0.2	0.18	0.37	0.05	0.02
cg01497487	22	38141814	TRIOBP, TRIOBP	intragenic, promoter	0.05	0.03	0.04	0.04	−0.27	−0.04	−0.26	0.01	0
cg05797418	22	38141834	TRIOBP, TRIOBP	intragenic, promoter	0.02	0.01	0.01	0.02	−0.27	−0.04	−0.16	0.01	0
cg17904489	22	38142133	TRIOBP, TRIOBP	intragenic, promoter	0.01	0.01	0.01	0.01	0.18	0.29	0.17	0	0
cg09192333	22	38142150	TRIOBP, TRIOBP	intragenic, promoter	0.02	0.03	0.03	0.03	0.26	0.03	0.15	0	0
cg22596557	22	38142152	TRIOBP, TRIOBP	intragenic, promoter	0.02	0.02	0.02	0.02	−0.2	0.19	−0.44	0	0
cg05463782	22	38142159	TRIOBP, TRIOBP	intragenic, promoter	0.02	0.02	0.03	0.03	0.01	0.37	0.4	0	0
cg03313666	22	38142162	TRIOBP, TRIOBP	intragenic, promoter	0.02	0.02	0.02	0.03	−0.14	0.15	−0.12	0	0
cg12354986	22	38142174	TRIOBP, TRIOBP	intragenic, promoter	0.02	0.03	0.03	0.03	0.29	0.34	0.34	0	0
cg16761131	22	38142377	TRIOBP, TRIOBP	intragenic, promoter	0.01	0.01	0.01	0.01	−0.1	−0.35	0.12	0	0
cg12361663	22	38142561	TRIOBP	intragenic	0.01	0.01	0.01	0.01	0.45	−0.27	−0.36	0	0
cg24232236	22	38142998	TRIOBP	intragenic	0.15	0.07	0.09	0.01	0.32	0.14	0.14	0	0.02
cg25404088	22	38143168	TRIOBP	intragenic	0.13	0.09	0.08	0.11	0.2	0.19	0.47	0.01	0.03
cg00219303	22	38146969	TRIOBP	intragenic	0.04	0.05	0.02	0.14	−0.31	−0.65	−0.13	0.05	0.01
cg22615910	22	38155575	TRIOBP	intragenic	0.02	0.01	0.02	0.02	−0.29	−0.62	0.23	0	0
cg26028571	22	38172403	TRIOBP	three_plus	0.15	0.08	0.07	0.05	−0.1	−0.29	0.1	0.01	0.03

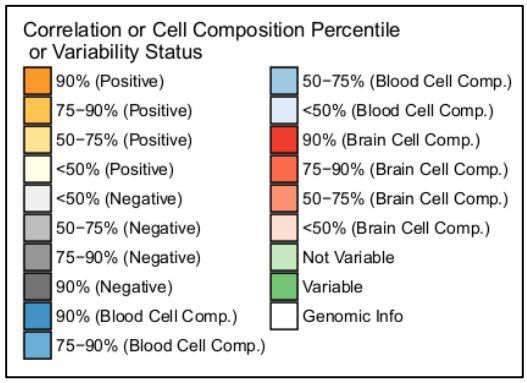


Figure 5.14 27 CPGs associated with *TRIOBP* show concordance between brain and blood. The genomic coordinates are provided including the gene region(s). The top 12 CpGs show variability between Brodmann Areas 10, 20, 7 and blood. Spearman correlation values between brain regions (0.41-0.81). The beta values changed at a given CpG on average 2-5% in the blood and 1-3% in the brain.

5.5.2 SRY-box 30 (*SOX30*)

All 13 CpGs identified for *SOX30* by Edgar et al. (2017) were present in our data set, while only one CpG had *p*-value < 10E-03 (cg06200339 *p*-value 7.99E-04). These CpGs showed inter-individual variability between Brodmann Areas 20, 10 and 7 and peripheral blood samples (Figure 5.15).

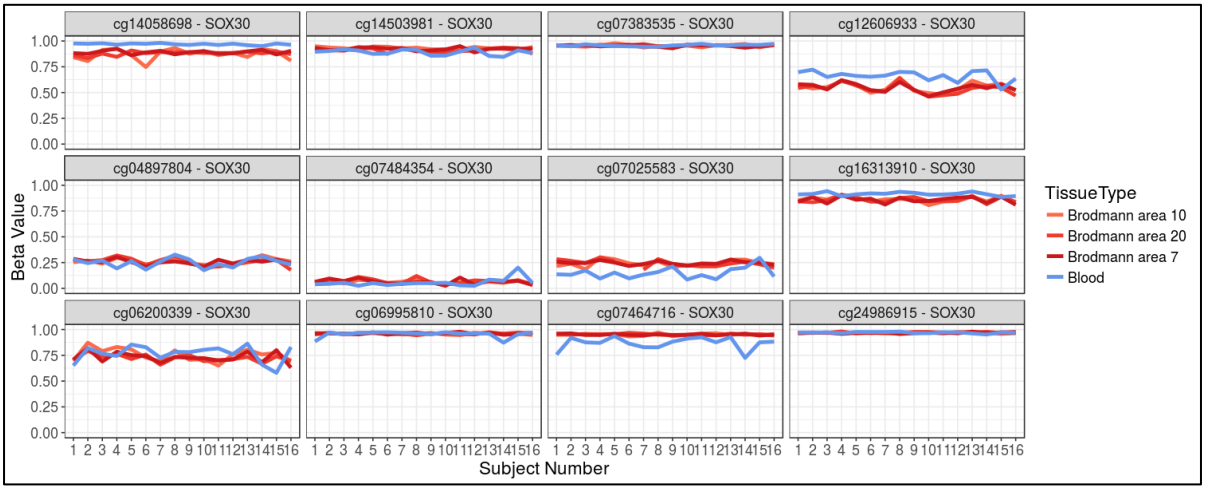
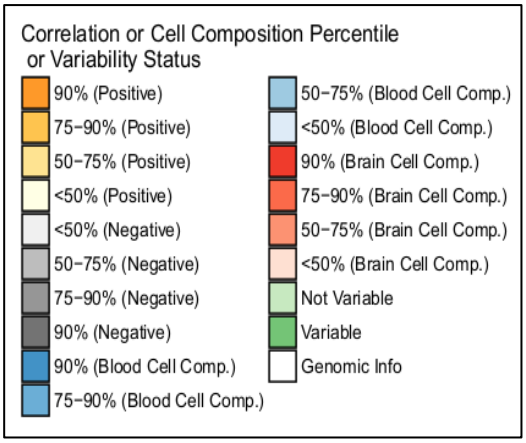


Figure 5.15 CpGs associated with *SOX30* show inter-variability between brain and blood at BA10 (orange), BA20 (pink), BA7 (red) and blood (blue) as evaluated by the BECon method (Edgar et al., 2017). 13 concordant CpGs were identified by submitting the *SOX30* gene name to the Shiny app https://redgar598.shinyapps.io/BECon/. Note probe cg06200339 with *p*-value 7.99E-04. The x-axis represents the inter-individual variability while the y-axis represents the Beta values of the methylation data.

Six of the 13 CpGs showed higher variability between brain and blood, although the correlation values in BA10, BA20, and BA7 for these 6 CpGs were low (-0.11 – 0.55): cg12606933, cg04897804, cg07484354, cg07025583, cg16313910, cg06200339. The beta values changed at a given CpG on average 2% in < 90% of blood cells, and on average 0-1% in < 75% of brain cells (Figure 5.16).

	Chr	Coor	Gene(s)	Gene Region(s)	Variability				Correlation			Cell Composition	
					BA10	BA20	BA7	Blood	BA10	BA20	BA7	Blood	Brain
cg14058698	5	157053392	SOX30	intragenic	0.12	0.04	0.04	0.02	0.13	−0.26	0.12	0	0.02
cg14503981	5	157076297	SOX30	intragenic	0.03	0.05	0.03	0.06	0.03	0.02	−0.14	0.02	0.01
cg07383535	5	157078082	SOX30	intragenic	0.03	0.01	0.02	0.02	−0.04	0.04	0.27	0	0
cg12606933	5	157078573	SOX30	intragenic	0.12	0.12	0.09	0.11	0.34	0.12	0.31	0.03	0.01
cg04897804	5	157078629	SOX30	intragenic	0.06	0.1	0.06	0.11	0.55	0.17	0.46	0.02	0
cg07484354	5	157078822	SOX30	intragenic	0.04	0.06	0.06	0.05	0.15	0.01	0.44	0.02	0
cg07025583	5	157078881	SOX30	intragenic	0.07	0.08	0.05	0.12	0.06	0.29	0.47	0.02	0
cg16313910	5	157079312	SOX30	promoter	0.06	0.06	0.07	0.04	0.1	−0.11	0.1	0.01	0
cg06200339	5	157079404	SOX30	promoter	0.13	0.08	0.11	0.19	0.26	0.45	0.17	0.03	0
cg06995810	5	157079468	SOX30	promoter	0.02	0.02	0.02	0.05	−0.48	0.32	0.21	0.01	0
cg07464716	5	157079520	SOX30	promoter	0.03	0.02	0.01	0.13	0.14	0.1	0.14	0.02	0
cg24986915	5	157079825	SOX30	promoter	0.02	0.01	0.01	0.01	−0.03	−0.13	0.14	0.01	0
cg03853593	5	157080900	SOX30	promoter	0.04	0.05	0.03	0.04	−0.03	0	−0.05	0.01	0.01

Figure 5.16 13 CPGs in the *SOX30* gene show concordance between brain and blood. The genomic coordinates are provided including the gene region(s). Only 6 CpGs show consistent variability between Brodmann Areas 10, 20, 7 and blood. Spearman correlation values for these 6 CpGs between brain regions (-0.11-0.55). The beta values changed at a given CpG on average 2% in < 75% of blood cells and on average 0-1% in < 75% of brain cells.



6. Discussion

Epigenome-wide association studies are relevant for the study of psychosis and the drugs used to treat them. This is apparent by the increasing number of EWAS in psychiatric disorders in the last decade. While these studies have identified genes implicated in the etiology of psychotic disorders, much still remains unknown regarding the underlying molecular mechanisms. A complete understanding of antipsychotic drug mechanisms is lacking. It is important to know if genes involved in the induction of side effects are also implicated in the therapeutic drug response.

The motivation of this study then was to identify the DNA positions and regions associated with altered methylation patterns in patients adhering to monotherapy with the antipsychotic drug olanzapine. We identified two differentially methylated regions from patient blood samples and the concordance of these regions between blood and brain showed that these blood samples could be used as a surrogate for brain tissue, especially for the *TRIOBP* DMR.

In addition, we identified differentially methylated positions in cultured HepG2 cells exposed to a high concentration of olanzapine. The positions from HepG2 cells and patient blood were subjected to a meta-analysis, which showed concordance in the top 34 probes, with a predominance of *hyper*-methylated genes. Finally, pathway analysis indicated the Hippo signaling pathway and Extracellular matrix organization pathway as implicated by the differentially methylated genes.

6.1 Identification of DNA positions and regions associated with altered methylation patterns following treatment with olanzapine

Differentially methylated positions were identified from methylation data assayed on the state of the art Illumina Infinium® MethylationEPIC BeadChip. The study design was developed in accordance with EWAS whereby models were designed that addressed the influence of environmental factors on methylation levels. We incorporated gender and smoking as covariates into the model, however chose to remove the age factor despite the reported influence of age on the methylome (Booth & Brunet, 2016). This decision was based upon **sva** reporting 0 surrogate variables for the model when age was excluded (Model 3b).

Before deciding on a model for analyzing differentially methylated positions, we checked the hidden variance in the Top Cohort data set by using the statistical method, *Surrogate Variable Analysis* (Jeffrey T Leek, Johnson, Parker, Jaffe, & Storey, 2012). Using this method we were able to estimate

sources of variation in the data set that were correlated with the environmental factors under consideration. Jaffe and Irizarry (2014) amongst others have reported on the correlation between age and cell type composition. Using **sva** we uncovered one surrogate variable that was correlated with age. This exercise facilitated our choice of model.

Influence of cell type composition between Model 3a and Model 3b. One additional consideration was evaluated before selecting the model for downstream analysis. We wanted to evaluate the impact of cell type composition on methylation levels identified with our model. Most EWA studies today correct for cellular heterogeneity due to different methylation profiles in different cell types (Jaffe & Irizarry, 2014). While changes in leukocyte composition may reflect an underlying phenotype related to a disease process (Houseman et al., 2012), an alternative perspective suggests that these changes may of themselves represent a biomarker of interest (R. Philibert & Glatt, 2017). Philibert and Glatt (2017) recommended therefore that methylation results be reported with and without adjustments for cellular heterogeneity. We employed this perspective in our study design to determine if there was any significant difference between the models.

We used therefore adjustment for cell type composition as a point of comparison to evaluate possible differences in the models associated with olanzapine-treatment. Our results indicated that while Model 3b (adjusted) identified over 550 more probes and over 400 more genes than Model 3a (unadjusted), the top significant probe for both models mapped to the potassium voltage-gated channel modifier subfamily G member 2 (*KCNG2*) gene. The role of potassium ion channel pathology is recognized in the development of psychiatric disorders including Bipolar Disorder. Multiple neurobiological mechanisms are mediated by potassium ion channels, including regulation of dopaminergic pathways, synaptic plasticity, and myelination (Balaraman, Lahiri, & Nurnberger, 2015). Pathway analyses from larger studies within the Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) have shown enrichment of genes related to potassium ion response (2018).

Next we looked at the number of differentially methylated genes that coincided in both models (n = 959). These shared genes represented approximately 53 % of the total number of genes (n = 1800). The higher proportion of *hyper*-methylated genes (63%) compared to *hypo*-methylated genes (37%), may be suggestive of a modifying effect of olanzapine. Similar findings were reported in a rat model where olanzapine–induced *hyper*-methylation levels were localized in hippocampal, cerebellum and liver tissues (Melka et al., 2014).

Comparison of the models through gene set over-representation analysis. Gene lists from both models submitted to ConsensusPathDB returned Gene Ontology (GO) terms for Biological process relating to cellular adhesion, development and structure; Molecular function terms characterizing binding activities; and Cellular component terms corresponding to the neuron and the synapse in particular. The models differed however. The unadjusted model (Model 3a) showed 2 GO terms specific to DNA binding that the adjusted model (Model 3b) did not have. Nevertheless, Model 3b had more terms specific to the synapse, in addition to lower *p*-values. Genes with less significant *p*-values may be representative of greater variation and therefore may be mapped less specifically to gene sets in pathway-based over-representation analysis. Further investigation therefore of Model 3a was discontinued.

Identification of differentially methylated regions. Differentially methylated regions are more informative than individual CpGs given their size (< 1 kb up to 1 MB) (Rakyan et al., 2011). We identified 15 CpGs using Model 3b (FDR 0.3) and presented two examples for further study: The Trio and F-actin binding protein (*TRIOBP*) gene was *hyper*-methylated at most CpGs in individuals taking olanzapine compared to aripiprazole or quetiapine; and the transcription factor Sry (sex determining region Y)-box 30 (*SOX30*) was *hypo*-methylated at most CpGs in individuals taking olanzapine compared to aripiprazole or quetiapine.

Three major variants have been identified for *TRIOBP* (*TRIOBP-1* and *TRIOBP-4/5*). There is no genetic overlap between *TRIOBP-1* and *TRIOBP-4* although both are associated with the regulation of actin cytoskeletal reorganization. TRIOBP-1 has been found to accumulate in the brain of individuals with schizophrenia (N J Bradshaw et al., 2014). Accumulation of TRIOBP-1 at adherens junctions (Park et al., 2018) has implications for disruption of endothelial cells in the blood-brain barrier and risk of CNS inflammation (Stamatovic, Keep, & Andjelkovic, 2008). There is evidence that these aggregations are mediated by the disruption of protein degradation; a process affecting cell morphology and cell development (N J Bradshaw et al., 2014).

SOX30 encodes a member of the SOX (SRY-related HMG-box) family of transcription factors involved in the regulation of embryonic development and in the determination of cell fate (Osaki et al., 1999). Yet when it forms a protein complex with another protein, SOX30 acts as an interactor protein. Together with Four and a half LIM domains protein 2 (FHL2), the complex locates to the nucleoplasm and interacts in the lipid metabolism pathway (REACTOME, EMBI: interaction_id:EBI-19031624).

6.2 Validation of findings in HepG2 cells exposed to olanzapine

In order to further analyze the effects of olanzapine on DNA methylation observed in patients, we performed experiments using cultured HepG2 cells. As a control for the model system, gene expression of selected genes was analyzed prior to methylation status.

6.2.1 Gene expression analysis

Olanzapine is known to affect gene expression of SREBF1 (Le Hellard et al., 2008; Steen et al., 2017) and FASN (Vik-Mo et al., 2008) but not in ChREBPα (Liu et al., 2017) in cultured HepG2 cells. These genes were selected therefore as controls for the effect of olanzapine in HepG2 cells.

A clear dose response was apparent for SREBF1, displaying significant upregulated expression in response to 50 µM and 100 µM. Of note, the standard error of the mean (SEM) was larger for the 100 µM samples, indicating more variation between the samples at higher doses. The possible source of the variation could not be attributed to the solvent DMSO as we prepared the samples with 0.25% (w/v) DMSO, well within the widely accepted concentration range of (0.1% to 1%, w/v) in cell culture research (Tunçer et al., 2018).

The results of the gene expression study indicate a measure of cytotoxicity, possibly induced by the 100 µM concentration. At this point a cytotoxicity assay would have been beneficial to clarify this point. We assumed that the 100 µM concentration would be sufficient exposure to affect methylation levels while the 50 µM and 1 µM samples showed lower gene expression levels. The decision to continue working with the 100 µM samples may be considered a limitation of the study.

Overall, alterations in gene expression were observed for SREBF1 and FASN, with both displaying upregulated expression patterns in response to olanzapine, as previously described (Fernø et al., 2006). These results show that the cells indeed have an effect of the drug at the selected candidate gene level. This could imply that alterations at the DNA methylation level are also present.

6.2.2 DNA Methylation analysis and meta-analysis

The top 21 probes and associated genes for 100 µM samples were identified and boxplots of the top 16 probes from each exposure group were presented for comparison. Generally, the top 16 probes for 100 µM samples showed a lower trend in average methylation level than the 50 µM samples. Examining the genes associated with the top differentially methylated probes shows genes involved in cellular activities pertaining to a cytotoxic environment. This was also supported by the analysis of

differentially methylated regions where one of the seven DMRs found is associated with *human homolog 3* ALKBH*(ALKBH3-AS1*).

ALKBH3 functions as a demethylase enzyme to repair DNA, especially targeting single-stranded DNA that has been methylated by metabolic or environmental insults. *ALKBH3* is upregulated otherwise in cancer cells and is necessary for cellular proliferation (Liefke et al., 2015). *ALKBH3-AS1* is an antisense RNA located on the opposite strand on Chromosome 11 and can regulate gene expression by silencing *AKLBH3* (Magistri, Faghihi, St Laurent 3rd, & Wahlestedt, 2012). *ALKBH3-AS1* is an antisense RNA located on the opposite strand on Chromosome 11 and can regulate gene expression by silencing *AKLBH3* (Magistri et al., 2012).

We were hesitant therefore to use these samples in a pathway-based over-representation analysis. The suspicion remained that these top genes for the 100 µM samples were associated with the cytotoxic event and therefore pathways implicated would reflect a bias towards diseased pathways. Instead, we performed a meta-analysis between Model 3b and HepG2 in order to identify the differentially methylated positions that ranked high on both lists.

6.2.3 Blood-brain concordance

DNA methylation levels vary between subjects and the question arises regarding the extent to which this variability can be measured in inaccessible tissues in the brain. Hannon et al., (2015) addressed this question by first identifying 4 regions in the brain highly correlated with each other. Next, they developed a method to use the variation of DNA methylation measured in the blood as a predictive metric of inter-individual variability identified in the brain. They found that while blood may give limited information relating to underlying pathological processes, it may be useful in the identification of disease biomarkers already present in the brain (Ibid).

In our study we used **BECon** (https://redgar598.shinyapps.io/BECon/) based on Hannon et al.'s (2015) work to evaluate blood-brain concordance between the differentially methylated regions identified for *TRIOBP* and *SOX30*. The top 12 CpGs for *TRIOBP* showed concordance between blood and Brodmann Area 10 (anterior prefrontal cortex), Brodmann Area 20 (inferior temporal cortex), and Brodmann Area 7 (parietal cortex). 10 CPGs for *TRIOBP* showed 90% positive correlation values (0.51-0.81) in these three brain regions. All three regions are associated with psychosis (Woodward & Heckers, 2016).

In contrast, the top 12 CpGs for *SOX30* showed less concordance compared to *TRIOBP* amongst all probes. Six of the probes with greatest variability amongst Brodmann Area 10, Brodmann Area 20,

Brodmann Area 7 and blood, showed lower correlation values between brain regions (-0.11-0.55). Lower correction value may reflect different functions for the transcription factor SOX30 in these three brain regions.

6.4 Pathways implicated by differentially methylated genes identified

We used ConsensusPathDB to map lists of differentially methylated genes to pathway-based gene sets in the database. We compared the top pathways for blood alone (Model 3b) and the meta-analysis. Amongst the top pathways for both groups was the Hippo signaling pathway and the WNT signaling pathway. The Hippo signaling pathway functions as part of the immune response, modulating the development and function of leukocytes (Taha, Janse van Rensburg, & Yang, 2018). Altered methylation levels in leukocytes have been shown in schizophrenics (Melas et al., 2012), nevertheless olanzapine has been shown to have antioxidant activity in neutrophils, reducing free radical-induced damage and mediating a neuroprotective effect (Brinholi et al., 2016).

The WNT signaling pathway is critical for neurodevelopment and has been shown to be associated with psychosis (Mill et al., 2008; Wesseling, Gottschalk, & Bahn, 2015). A recent study using the TOP Cohort dataset showed abnormal gene expression in the WNT signaling pathway in patients with Schizophrenia and Bipolar Disorder (Hoseth et al., 2018). They showed upregulation the Nuclear factor of activated T cells 3 (*NFATC3*) in the patient blood of a mixed-antipsychotic drug group. Our study identified *hyper*-methylation of (*NFATC2*) in patients treated with olanzapine mono-therapy, indicating altered gene transcription during an immune response. These results indicate a role for antipsychotic drug targeting in the WNT pathway in the treatment of psychosis (Hoseth et al., 2018).

6.5 Future directions

Several additional analyses may be performed with the methylation data obtained in this project. An analysis of genotyping data for *CYPD26* from the TOP Cohort could be undertaken to determine if patients with recorded low serum levels of olanzapine were high metabolizers, or unambiguously non-compliant with their medication.

Increased sample size may provide the statistical power to evaluate the difference in methylation patterns between patients who are taking additional medications. Mood stabilizers such as lithium (Pisanu, Papadima, Del Zompo, & Squassina, 2018) and valproic acid (Babu Swathy & Banerjee, 2017) are known to alter methylation patterns, often in combination with other drugs.

Gene expression data from the TOP Cohort overlaps with methylation data for samples in this project. The gene expression data therefore could be evaluated with regard to the DMRs identified. The DMR for *TRIOBP* was *hyper*-methylated, a state generally associated with reduced gene expression. The effect of altered methylation levels on gene expression however, is complex. Since TRIOBP-1 was found to aggregate in the brain of patients with schizophrenia (N J Bradshaw et al., 2014), it would be interesting to ascertain if reduced gene expression may be associated with a therapeutic effect of olanzapine. The DMR for SOX30 was *hypo*-methylated, suggestive of increased gene expression. Further study will be required to evaluate the role of this transcription factor in embryonic development and cell fate.

Finally, the biological interpretation of DNA methylation in patients taking olanzapine may be enhanced by incorporating clinical data. Quantifying the relationship between differential methylation and clinical response to the drug may aid in the development of new guidelines for personalized treatment in psychiatry.

EWA studies are still in their infancy although promising work has already been done in psychiatric epigenetics. There is good reason to believe these studies will continue to contribute important work towards unraveling the complexity between multiple environmental factors, and the underlying genetic vulnerabilities in individuals with psychotic disorders.

7 Conclusion

Differentially methylated positions and regions were identified in the blood of patients adhering to olanzapine monotherapy. A comparison of models adjusting for cell type composition provided evidence of improved *p*-values when cell type adjustment was included in the model. Our results showed concordance between blood and brain for two identified differentially methylated regions. The pathways implicated by the differentially methylated genes showed evidence of alterations in immune pathways and the possible mediating effect of olanzapine.

In conclusion, epigenetic modifications induced by the antipsychotic drug olanzapine may provide insight into mechanisms that inform treatment response. Given the burden of psychosis worldwide and the challenge to successfully treat psychotic disorders, the epigenetics of drug treatment response may contribute to new drug targets, improved treatment outcomes and personalized care.

8 References

Abdolmaleky, H. M., Yaqubi, S., Papageorgis, P., Lambert, A. W., Ozturk, S., Sivaraman, V., & Thiagalingam, S. (2011). Epigenetic dysregulation of HTR2A in the brain of patients with schizophrenia and bipolar disorder. *Schizophrenia Research*, *129*(2–3), 183–190. https://doi.org/10.1016/j.schres.2011.04.007

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., … Neale, B. M. (2018). Analysis of shared heritability in common disorders of the brain. *Science*, *360*(6395). https://doi.org/10.1126/science.aap8757

Arciniegas, D. B. (2015). Psychosis. *Continuum (Minneapolis, Minn.)*, *21*(3 Behavioral Neurology and Neuropsychiatry), 715–736. https://doi.org/10.1212/01.CON.0000466662.89908.e7

Asai, T., Bundo, M., Sugawara, H., Sunaga, F., Ueda, J., Tanaka, G., … Iwamoto, K. (2013). Effect of mood stabilizers on DNA methylation in human neuroblastoma cells. *Int J Neuropsychopharmacol*, *16*(10), 2285–2294. https://doi.org/10.1017/S1461145713000710

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, *25*(1), 25–29. https://doi.org/10.1038/75556

Balaraman, Y., Lahiri, D. K., & Nurnberger, J. I. (2015). Variants in Ion Channel Genes Link Phenotypic Features of Bipolar Illness to Specific Neurobiological Process Domains. *Mol Neuropsychiatry*, *1*(1), 23–35. https://doi.org/10.1159/000371886

Bauer, M., Fink, B., Thurmann, L., Eszlinger, M., Herberth, G., & Lehmann, I. (2015). Tobacco smoking differently influences cell types of the innate and adaptive immune system-indications from CpG site methylation. *Clin Epigenetics*, *7*, 83. https://doi.org/10.1186/s13148-016-0249-7

Berger, N. D., Stanley, F. K. T., Moore, S., & Goodarzi, A. A. (2017). ATM-dependent pathways of chromatin remodelling and oxidative DNA damage responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1731). https://doi.org/10.1098/rstb.2016.0283

Bird, A., Taggart, M., Frommer, M., Miller, O. J., & Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell*, *40*(1), 91–99.

Booth, L. N., & Brunet, A. (2016). The Aging Epigenome. *Mol Cell*, *62*(5), 728–744. https://doi.org/10.1016/j.molcel.2016.05.013

Bradshaw, N. J., Bader, V., Prikulis, I., Lueking, A., Mullner, S., & Korth, C. (2014). Aggregation of the protein TRIOBP-1 and its potential relevance to schizophrenia. *PLoS One*, *9*(10), e111196. https://doi.org/10.1371/journal.pone.0111196

Bradshaw, N. J., Bader, V., Prikulis, I., Lueking, A., Müllner, S., & Korth, C. (2014). Aggregation of the Protein TRIOBP-1 and its potential relevance to schizophrenia. *PLoS ONE*, *9*(10), 1–11. https://doi.org/10.1371/journal.pone.0111196

Brinholi, F. F., Farias, C. C. de, Bonifácio, K. L., Higachi, L., Casagrande, R., Moreira, E. G., & Barbosa, D. S. (2016). Clozapine and olanzapine are better antioxidants than haloperidol, Quetiapine, Risperidone and ziprasidone in in vitro models. *Biomedicine and Pharmacotherapy*, *81*, 411–415. https://doi.org/10.1016/j.biopha.2016.02.047

Bryois, J., Buil, A., Ferreira, P. G., Panousis, N. I., Brown, A. A., Viñuela, A., … Dermitzakis, E. T. (2017). Time-dependent genetic effects on gene expression implicate aging processes. *Genome Research*, *27*(4), 545–552. https://doi.org/10.1101/gr.207688.116

Cardano, M., Scarinzi, C., Costa, G., & d'Errico, A. (2018). Internal migration and mental health of the second generation. The case of Turin in the age of the Italian economic miracle. *Social Science & Medicine (1982)*, *208*, 142–149. https://doi.org/10.1016/j.socscimed.2018.04.055

Carlson, M. (2015). TxDb.Hsapiens.UCSC.hg19.knownGene, v. 3.2.2.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., … Hayashizaki, Y. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genetics*, *38*(6), 626–635. https://doi.org/10.1038/ng1789

Carpenter, W. T. J., & Davis, J. M. (2012). Another view of the history of antipsychotic drug discovery and development. *Molecular Psychiatry*, *17*(12), 1168–1173. https://doi.org/10.1038/mp.2012.121

Carrard, A., Salzmann, A., Malafosse, A., & Karege, F. (2011). Increased DNA methylation status of the serotonin receptor 5HTR1A gene promoter in schizophrenia and bipolar disorder. *Journal of Affective Disorders*, *132*(3), 450–453. https://doi.org/10.1016/j.jad.2011.03.018

Castner, S. A., Williams, G. V, & Goldman-Rakic, P. S. (2000). Reversal of antipsychotic-induced

working memory deficits by short-term dopamine D1 receptor stimulation. *Science (New York, N.Y.)*, *287*(5460), 2020–2022.

Collins, P. Y., Patel, V., Joestl, S. S., March, D., Insel, T. R., Daar, A. S., … Stein, D. J. (2011). Grand challenges in global mental health. *Nature*, *475*(7354), 27–30. https://doi.org/10.1038/475027a

Correll, C. U. (2010). From receptor pharmacology to improved outcomes : individualising the selection , dosing , and switching of antipsychotics. *European Psychiatry*, *25*, S12–S21. https://doi.org/10.1016/S0924-9338(10)71701-6

Correll, C. U., Solmi, M., Veronese, N., Bortolato, B., Rosson, S., Santonastaso, P., … Stubbs, B. (2017). Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *16*(2), 163–180. https://doi.org/10.1002/wps.20420

Costa, E., Chen, Y., Davis, J., Dong, E., Noh, J. S., Tremolizzo, L., … Guidotti, A. (2002). REELIN and schizophrenia: a disease at the interface of the genome and the epigenome. *Molecular Interventions*, *2*(1), 47–57. https://doi.org/10.1124/mi.2.1.47

Delaney, S., Fallon, B., Alaedini, A., Yolken, R., Indart, A., Feng, T., … Javitt, D. (2018). Inflammatory biomarkers in psychosis and clinical high risk populations. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2018.10.017

DeRosse, P., & Karlsgodt, K. H. (2015). Examining the Psychosis Continuum. *Current Behavioral Neuroscience Reports*, *2*(2), 80–89. https://doi.org/10.1007/s40473-015-0040-7

DeVylder, J. E., Jahn, D. R., Doherty, T., Wilson, C. S., Wilcox, H. C., Schiffman, J., & Hilimire, M. R. (2015). Social and psychological contributions to the co-occurrence of sub-threshold psychotic experiences and suicidal behavior. *Social Psychiatry and Psychiatric Epidemiology*, *50*(12), 1819–1830. https://doi.org/10.1007/s00127-015-1139-6

Donato, M. T., Tolosa, L., & Gómez-Lechón, M. J. (2015). Culture and Functional Characterization of Human Hepatoma HepG2 Cells. In M. Vinken & V. Rogiers (Eds.), *Protocols in In Vitro Hepatocyte Research* (pp. 77–93). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4939-2074-7_5

Dong, E., Ruzicka, W. B., Grayson, D. R., & Guidotti, A. (2015). DNA-methyltransferase1 (DNMT1) binding to CpG rich GABAergic and BDNF promoters is increased in the brain of schizophrenia and bipolar disorder patients. *Schizophr Res*, *167*(1–3), 35–41. https://doi.org/10.1016/j.schres.2014.10.030

Du, P., Zhang, X., Huang, C. C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, *11*, 587. https://doi.org/10.1186/1471-2105-11-587

Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T., & Sedgwick, B. (2002). Reversal of DNA alkylation damage by two human dioxygenases. *Proceedings of the National Academy of Sciences*, *99*(26), 16660–16665. https://doi.org/10.1073/pnas.262589799

Edgar, R. D., Jones, M. J., Meaney, M. J., Turecki, G., & Kobor, M. S. (2017). BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry*, *7*(8), e1187. https://doi.org/10.1038/tp.2017.171

Eranti, S. V, MacCabe, J. H., Bundy, H., & Murray, R. M. (2013). Gender difference in age at onset of schizophrenia: a meta-analysis. *Psychological Medicine*, *43*(1), 155–167. https://doi.org/10.1017/S003329171200089X

Eum, S., Lee, A. M., & Bishop, J. R. (2016). Pharmacogenetic tests for antipsychotic medications: clinical implications and considerations. *Dialogues in Clinical Neuroscience*, *18*(3), 323–337.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., … D'Eustachio, P. (2018). The Reactome Pathway Knowledgebase. *Nucleic Acids Res*, *46*(D1), D649–D655. https://doi.org/10.1093/nar/gkx1132

Fernø, J., Skrede, S., Vik-Mo, A. O., Håvik, B., & Steen, V. M. (2006). Drug-induced activation of SREBP-controlled lipogenic gene expression in CNS-related cell lines: marked differences between various antipsychotic drugs. *BMC Neuroscience*, *7*, 69. https://doi.org/10.1186/1471-2202-7-69

Foster, A., Buckley, P., Lauriello, J., Looney, S., & Schooler, N. (2017). Combination Antipsychotic Therapies: An Analysis From a Longitudinal Pragmatic Trial. *Journal of Clinical Psychopharmacology*, *37*(5), 595–599. https://doi.org/10.1097/JCP.0000000000000766

Fribourg, M., Moreno, J. L., Holloway, T., Provasi, D., Baki, L., Mahajan, R., … Logothetis, D. E. (2011).

Decoding the signaling of a GPCR heteromeric complex reveals a unifying mechanism of action of antipsychotic drugs. *Cell*, *147*(5), 1011–1023. https://doi.org/10.1016/j.cell.2011.09.055

Fries, G. R., Li, Q., McAlpin, B., Rein, T., Walss-Bass, C., Soares, J. C., & Quevedo, J. (2016). The role of DNA methylation in the pathophysiology and treatment of bipolar disorder. *Neurosci Biobehav Rev*, *68*, 474–488. https://doi.org/10.1016/j.neubiorev.2016.06.010

Fusar-poli, P., Tantardini, M., Simone, S. De, Ramella-cravaro, V., Oliver, D., Kingdon, J., … Mcguire, P. (2017). Deconstructing vulnerability for psychosis : Meta-analysis of environmental risk factors for psychosis in subjects at ultra high-risk. *European Psychiatry*, *40*, 65–75. https://doi.org/10.1016/j.eurpsy.2016.09.003

Garbett, K., Gal-Chis, R., Gaszner, G., Lewis, D. A., & Mirnics, K. (2008). Transcriptome alterations in the prefrontal cortex of subjects with schizophrenia who committed suicide. *Neuropsychopharmacologia Hungarica : A Magyar Pszichofarmakologiai Egyesulet Lapja = Official Journal of the Hungarian Association of Psychopharmacology*, *10*(1), 9–14.

Gene Ontology, C. (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res*, *43*(Database issue), D1049-56. https://doi.org/10.1093/nar/gku1179

Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. (2018). *Cell*, *173*(7), 1705–1715.e16. https://doi.org/10.1016/j.cell.2018.05.046

Gladkevich, A., Kauffman, H. F., & Korf, J. (2004). Lymphocytes as a neural probe: Potential for studying psychiatric disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *28*(3), 559–576. https://doi.org/10.1016/j.pnpbp.2004.01.009

Greenhalgh, A. M., Gonzalez-Blanco, L., Garcia-Rizo, C., Fernandez-Egea, E., Miller, B., Arroyo, M. B., & Kirkpatrick, B. (2017). Meta-analysis of glucose tolerance, insulin, and insulin resistance in antipsychotic-naïve patients with nonaffective psychosis. *Schizophrenia Research*, *179*, 57–63. https://doi.org/10.1016/j.schres.2016.09.026

Guidotti, A., Grayson, D. R., & Caruncho, H. J. (2016). Epigenetic RELN Dysfunction in Schizophrenia and Related Neuropsychiatric Disorders. *Frontiers in Cellular Neuroscience*, *10*, 89. https://doi.org/10.3389/fncel.2016.00089

Hahne, F., & Ivanek, R. (2016). Visualizing genomic data using Gviz and bioconductor. In *Methods in Molecular Biology* (Vol. 1418, pp. 335–351). Humana Press, New York, NY.

https://doi.org/10.1007/978-1-4939-3578-9_16

Hannon, E., Dempster, E., Viana, J., Burrage, J., Smith, A. R., Macdonald, R., … Schalkwyk, L. (2016). An integrated genetic-epigenetic analysis of schizophrenia : evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biology*, 1–16. https://doi.org/10.1186/s13059-016-1041-x

Hannon, E., Lunnon, K., Schalkwyk, L., & Mill, J. (2015). Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, *10*(11), 1024–1032. https://doi.org/10.1080/15592294.2015.1100786

Hansen, K. D. (2016). IlluminaHumanMethylationEPICanno.ilm10b2.hg19: Annotation for Illumina's EPIC methylation arrays.

Herwig, R., Hardt, C., Lienhard, M., & Kamburov, A. (2016). Analyzing and interpreting genome data at the network level with ConsensusPathDB. *Nature Protocols*, *11*, 1889. https://doi.org/10.1038/nprot.2016.117https://www.nature.com/articles/nprot.2016.117#supplementary-information

Hirschfeld, R. M. (2014). Differential diagnosis of bipolar disorder and major depressive disorder. *Journal of Affective Disorders*, *169 Suppl*, S12-6. https://doi.org/10.1016/S0165-0327(14)70004-7

Holloway, T., & González-Maeso, J. (2015). Epigenetic Mechanisms of Serotonin Signaling. *ACS Chemical Neuroscience*, *6*(7), 1099–1109. https://doi.org/10.1021/acschemneuro.5b00033

Honings, S., Drukker, M., van Nierop, M., van Winkel, R., Wittchen, H.-U., Lieb, R., … van Os, J. (2016). Psychotic experiences and incident suicidal ideation and behaviour: Disentangling the longitudinal associations from connected psychopathology. *Psychiatry Research*, *245*, 267–275. https://doi.org/10.1016/j.psychres.2016.08.002

Horacek, J., Bubenikova-Valesova, V., Kopecek, M., Palenicek, T., Dockery, C., Mohr, P., & H??schl, C. (2006). Mechanism of Action of Atypical Antipsychotic Drugs and the Neurobiology of Schizophrenia. *CNS Drugs*, *20*(5), 389–409. https://doi.org/10.2165/00023210-200620050-00004

Horvath, S., Zhang, Y., Langfelder, P., Kahn, R. S., Boks, M. P., van Eijk, K., … Ophoff, R. A. (2012).

Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol*, *13*(10), R97. https://doi.org/10.1186/gb-2012-13-10-r97

Hoseth, E. Z., Krull, F., Dieset, I., Mørch, R. H., Hope, S., Gardsjord, E. S., … Ueland, T. (2018). Exploring the Wnt signaling pathway in schizophrenia and bipolar disorder. *Translational Psychiatry*, *8*(1), 55. https://doi.org/10.1038/s41398-018-0102-1

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., … Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, *13*(1), 86. https://doi.org/doi:10.1186/1471-2105-13-86

Houtepen, L. C., van Bergen, A. H., Vinkers, C. H., & Boks, M. P. M. (2016). DNA methylation signatures of mood stabilizers and antipsychotics in bipolar disorder. *Epigenomics*, *8*(2), 197–208. https://doi.org/10.2217/epi.15.98

Huang, H.-S., & Akbarian, S. (2007). GAD1 mRNA expression and DNA methylation in prefrontal cortex of subjects with schizophrenia. *PloS One*, *2*(8), e809. https://doi.org/10.1371/journal.pone.0000809

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., … Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, *12*(2), 115–121. https://doi.org/10.1038/nmeth.3252

Hynes, R. O. (2014, December). Stretching the boundaries of extracellular matrix research. *Nature Reviews. Molecular Cell Biology*. England. https://doi.org/10.1038/nrm3908

Insel, T. R., Sahakian, B. J., Voon, V., Nye, J., Brown, V. J., Altevogt, B. M., … Williams, J. H. (2012). Drug research: A plan for mental illness. *Nature*, *483*(7389), 483269a. https://doi.org/10.1038/483269a

Irizarry, R. A., Wu, H., & Feinberg, A. P. (2009). A species-generalized probabilistic model-based definition of CpG islands. *Mammalian Genome : Official Journal of the International Mammalian Genome Society*, *20*(9–10), 674–680. https://doi.org/10.1007/s00335-009-9222-5

Jaffe, A. E., & Irizarry, R. A. (2014). Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biology*, *15*(2), R31. https://doi.org/10.1186/gb-2014-15-2-r31

Kahn, R. S., Sommer, I. E., Murray, R. M., Meyer-lindenberg, A., Cannon, T. D., Correll, C. U., … Insel,

T. R. (2015). Schizophrenia, *1*(November). https://doi.org/10.1038/nrdp.2015.67

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H., & Herwig, R. (2011). ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res*, *39*(Database issue), D712-7. https://doi.org/10.1093/nar/gkq1156

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*, *45*(D1), D353–D361. https://doi.org/10.1093/nar/gkw1092

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation, *44*(October 2015), 457–462. https://doi.org/10.1093/nar/gkv1070

Kirkpatrick, B., & Kennedy, B. K. (2018). Accelerated aging in schizophrenia and related disorders: Future research, *196*. https://doi.org/10.1016/j.schres.2017.06.034

Kocerha, J., & Aggarwal, N. (2018). Epigenetics in Neurobehavioral Disease. In *Epigenetics in Human Disease* (2nd ed., pp. 251–267). Academic Press. https://doi.org/10.1016/b978-0-12-812215-0.00008-x

Kotlicka-Antczak, M., Pawelczyk, A., Pawelczyk, T., Strzelecki, D., Zurner, N., & Karbownik, M. S. (2017). A history of obstetric complications is associated with the risk of progression from an at risk mental state to psychosis. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2017.10.039

Koyanagi, A., & Stickley, A. (2004). The Association between Sleep Problems and Psychotic Symptoms in the General Population : A Global Perspective.

Kusumi, I., Boku, S., & Takahashi, Y. (2015). Psychopharmacology of atypical antipsychotic drugs: From the receptor binding profile to neuroprotection and neurogenesis. *Psychiatry Clin Neurosci*, *69*(5), 243–258. https://doi.org/10.1111/pcn.12242

Labrie, V., Pai, S., & Petronis, A. (2012). Epigenetics of major psychosis: progress, problems and perspectives. *Trends Genet*, *28*(9), 427–435. https://doi.org/10.1016/j.tig.2012.04.002

Lambert, M., Haro, J. M., Novick, D., Edgell, E. T., Kennedy, L., Ratcliffe, M., & Naber, D. (2005). Olanzapine vs. other antipsychotics in actual out-patient settings: six months tolerability results from the European Schizophrenia Out-patient Health Outcomes study. *Acta*

Psychiatrica Scandinavica, *111*(3), 232–243. https://doi.org/10.1111/j.1600-0447.2004.00451.x

Laursen, T. M., Nordentoft, M., & Mortensen, P. B. (2014). Excess early mortality in schizophrenia. *Annual Review of Clinical Psychology*, *10*, 425–448. https://doi.org/10.1146/annurev-clinpsy-032813-153657

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., … Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLoS Computational Biology*, *9*(8), 1–10. https://doi.org/10.1371/journal.pcbi.1003118

Le Hellard, S., Mühleisen, T. W., Djurovic, S., Fernø, J., Ouriaghi, Z., Mattheisen, M., … Steen, V. M. (2008). Polymorphisms in SREBF1 and SREBF2, two antipsychotic-activated transcription factors controlling cellular lipogenesis, are associated with schizophrenia in German and Scandinavian samples. *Mol. Psychiatry*, *15*(5), 463. https://doi.org/10.1038/mp.2008.110

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, *28*(6), 882–883. https://doi.org/10.1093/bioinformatics/bts034

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments, *28*(6), 882–883. https://doi.org/10.1093/bioinformatics/bts034

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., … Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet*, *11*(10). https://doi.org/10.1038/nrg2825

Leucht, S., Cipriani, A., Spineli, L., Mavridis, D., Örey, D., Richter, F., … Davis, J. M. (2013). Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *The Lancet*, *382*(9896), 951–962. https://doi.org/10.1016/s0140-6736(13)60733-3

Liefke, R., Windhof-Jaidhauser, I. M., Gaedcke, J., Salinas-Riester, G., Wu, F., Ghadimi, M., & Dango, S. (2015). The oxidative demethylase ALKBH3 marks hyperactive gene promoters in human cancer cells. *Genome Medicine*, *7*(1), 66. https://doi.org/10.1186/s13073-015-0180-0

Liu, X., Wu, Z., Lian, J., Hu, C. H., Huang, X. F., & Deng, C. (2017). Time-dependent changes and

potential mechanisms of glucose-lipid metabolic disorders associated with chronic clozapine or olanzapine treatment in rats. *Scientific Reports*, *7*(1), 1–13. https://doi.org/10.1038/s41598-017-02884-w

Lowe, P., Krivoy, A., Porffy, L., Henriksdottir, E., Eromona, W., & Shergill, S. S. (2017). When the drugs don't work: treatment-resistant schizophrenia, serotonin and serendipity. *Therapeutic Advances in Psychopharmacology*, 204512531773700. https://doi.org/10.1177/2045125317737003

Lowe, R., Slodkowicz, G., Goldman, N., & Rakyan, V. K. (2015). The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. *Epigenetics*, *10*(4), 274–281. https://doi.org/10.1080/15592294.2014.1003744

Magistri, M., Faghihi, M. A., St Laurent 3rd, G., & Wahlestedt, C. (2012). Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends in Genetics : TIG*, *28*(8), 389–396. https://doi.org/10.1016/j.tig.2012.03.013

Masser, D. R., Hadad, N., Porter, H., Stout, M. B., Unnikrishnan, A., Stanford, D. R., & Freeman, W. M. (2018). Analysis of DNA modifications in aging research. *GeroScience*, *40*(1), 11–29. https://doi.org/10.1007/s11357-018-0005-3

Mauri, M. C., Paletta, S., Maffini, M., Colasanti, A., Dragogna, F., Di Pace, C., & Altamura, A. C. (2014). Clinical pharmacology of atypical antipsychotics: an update. *EXCLI Journal*, *13*, 1163–1191.

McGrath, J. J., Saha, S., Lim, C. C. W., Aguilar-Gaxiola, S., Alonso, J., Andrade, L. H., … Kessler, R. C. (2017). Trauma and psychotic experiences: transnational data from the World Mental Health Survey. *The British Journal of Psychiatry : The Journal of Mental Science*, *211*(6), 373–380. https://doi.org/10.1192/bjp.bp.117.205955

Melas, P. A., Rogdaki, M., Osby, U., Schalling, M., Lavebratt, C., & Ekstrom, T. J. (2012). Epigenetic aberrations in leukocytes of patients with schizophrenia: association of global DNA methylation with antipsychotic drug treatment and disease onset. *FASEB J*, *26*(6), 2712–2718. https://doi.org/10.1096/fj.11-202069

Melka, M. G., Castellani, C. A., Laufer, B. I., Rajakumar, N., O'Reilly, R., & Singh, S. M. (2013). Olanzapine induced DNA methylation changes support the dopamine hypothesis of psychosis. *Journal of Molecular Psychiatry*, *1*(1), 19. https://doi.org/10.1186/2049-9256-1-19

Melka, M. G., Laufer, B. I., McDonald, P., Castellani, C. A., Rajakumar, N., O'Reilly, R., & Singh, S. M. (2014). The effects of olanzapine on genome-wide DNA methylation in the hippocampus and cerebellum. *Clinical Epigenetics*, *6*(1), 1–12. https://doi.org/10.1186/1868-7083-6-1

Meltzer, H. Y., Matsubara, S., & Lee, J. C. (1989). Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin2 pKi values. *The Journal of Pharmacology and Experimental Therapeutics*, *251*(1), 238–246.

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., … Petronis, A. (2008). Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet*, *82*(3), 696–711. https://doi.org/10.1016/j.ajhg.2008.01.008

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *38*(1), 23–38. https://doi.org/10.1038/npp.2012.112

Moore, T. R., Hill, A. M., & Panguluri, S. K. (2014). Pharmacogenomics in psychiatry: implications for practice. *Recent Patents on Biotechnology*, *8*(2), 152–159.

Morgan, M. (2018). AnnotationHub: Client to access AnnotationHub resources.

Nestler, E. J., Peña, C. J., Kundakovic, M., Mitchell, A., & Akbarian, S. (2016). Epigenetic Basis of Mental Illness. https://doi.org/10.1177/1073858415608147

Nordholm, D., Rostrup, E., Mondelli, V., Randers, L., Nielsen, M. O., Wulff, S., … Glenthoj, B. (2018). Multiple measures of HPA axis function in ultra high risk and first-episode schizophrenia patients. *Psychoneuroendocrinology*, *92*, 72–80. https://doi.org/10.1016/j.psyneuen.2018.03.015

Norredam, M., Nellums, L., Nielsen, R. S., Byberg, S., & Petersen, J. H. (2018). Incidence of psychiatric disorders among accompanied and unaccompanied asylum-seeking children in Denmark: a nation-wide register-based cohort study. *European Child & Adolescent Psychiatry*, *27*(4), 439–446. https://doi.org/10.1007/s00787-018-1122-3

Oh, H. Y., Singh, F., Koyanagi, A., Jameson, N., Schiffman, J., & Devylder, J. (2016). Sleep disturbances are associated with psychotic experiences : Findings from the National Comorbidity Survey Replication. *Schizophrenia Research*, *171*(1–3), 74–78. https://doi.org/10.1016/j.schres.2016.01.018

Orhan, F., Fatouros-Bergman, H., Goiny, M., Malmqvist, A., Piehl, F., Cervenka, S., … Engberg, G. (2018). CSF GABA is reduced in first-episode psychosis and associates to symptom severity. *Molecular Psychiatry*, *23*(5), 1244–1250. https://doi.org/10.1038/mp.2017.25

Osaki, E., Nishina, Y., Inazawa, J., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., … Semba, K. (1999). Identification of a novel Sry-related gene and its germ cell-specific expression. *Nucleic Acids Research*, *27*(12), 2503–2510. https://doi.org/10.1093/nar/27.12.2503

Ovenden, E. S., McGregor, N. W., Emsley, R. A., & Warnich, L. (2018). DNA methylation and antipsychotic treatment mechanisms in schizophrenia: Progress and future directions. *Prog Neuropsychopharmacol Biol Psychiatry*, *81*, 38–49. https://doi.org/10.1016/j.pnpbp.2017.10.004

Park, S., Lee, H., Kim, M., Park, J., Kim, S. H., & Park, J. (2018). Emerging roles of TRIO and F-actin-binding protein in human diseases. *Cell Commun Signal*, *16*(1), 29. https://doi.org/10.1186/s12964-018-0237-y

Peters, T. J. (2015). De Novo Identification of differentially methylated regions in the human genome.

Peters, T. J., Buckley, M. J., Statham, A. L., Pidsley, R., Samaras, K., R, V. L., … Molloy, P. L. (2015). De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin*, *8*, 6. https://doi.org/10.1186/1756-8935-8-6

Philibert, R. A., Beach, S. R. H., Lei, M.-K., & Brody, G. H. (2013). Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. *Clinical Epigenetics*, *5*(1), 19. https://doi.org/10.1186/1868-7083-5-19

Philibert, R., & Glatt, S. J. (2017). Optimizing the chances of success in the search for epigenetic biomarkers: Embracing genetic variation. *Am J Med Genet B Neuropsychiatr Genet*, *174*(6), 589–594. https://doi.org/10.1002/ajmg.b.32569

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., … Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, *17*(1). https://doi.org/10.1186/s13059-016-1066-1

Pillich R.T., J., C., V., R., D., W., & D., P. (2017). NDEx: A Community Resource for Sharing and

Publishing of Biological Networks. In W. C., A. C., & R. K. (eds) (Eds.), *Protein Bioinformatics. Methods in Molecular Biology* (pp. 271–301). New York, NY USA. https://doi.org/10.1007/978-1-4939-6783-4_13

Pillinger, T., D'Ambrosio, E., McCutcheon, R., & D Howes, O. (2018). Is psychosis a multisystem disorder? A meta-review of central nervous system, immune, cardiometabolic, and endocrine alterations in first-episode psychosis and perspective on potential models. *Molecular Psychiatry*, 1–19. https://doi.org/10.1038/s41380-018-0058-9

Pisanu, C., Papadima, E. M., Del Zompo, M., & Squassina, A. (2018). Understanding the molecular mechanisms underlying mood stabilizer treatments in bipolar disorder: Potential involvement of epigenetics. *Neurosci Lett*, *669*, 24–31. https://doi.org/10.1016/j.neulet.2016.06.045

R-Core Team. (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Radua, J., Ramella-Cravaro, V., Ioannidis, J. P. A., Reichenberg, A., Phiphopthatsanee, N., Amir, T., … Fusar-Poli, P. (2018). What causes psychosis? An umbrella review of risk and protective factors. *World Psychiatry*, *17*(1), 49–66. https://doi.org/10.1002/wps.20490

Rakyan, V. K., Down, T. A., Balding, D. J., & Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nat Rev Genet*, *12*(8), 529–541. https://doi.org/10.1038/nrg3000

Remington, G., Addington, D., Honer, W., Ismail, Z., Raedler, T., & Teehan, M. (2017). Guidelines for the Pharmacotherapy of Schizophrenia in Adults. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, *62*(9), 604–616. https://doi.org/10.1177/0706743717720448

Riedel, M., Müller, N., & Strassnig, M. (2007). Quetiapine in the treatment of schizophrenia and related disorders, *3*(2), 219–235.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Robinson, M. D., Institute of Molecular Life Sciences Zurich, Switzerland, U. of Z., SIB Swiss Institute of Bioinformatics Zurich, Switzerland, U. of Z., mark.robinson@imls.uzh.ch, Kahraman, A., Institute of Molecular Life Sciences Zurich, Switzerland, U. of Z., … SIB Swiss Institute of Bioinformatics Zurich, Switzerland, U. of Z. (2014). Statistical methods for detecting

differentially methylated loci and regions. *Front Genet*, *5*. https://doi.org/10.3389/fgene.2014.00324

Seeman, P., Bzowej, N. H., Guan, H. C., Bergeron, C., Reynolds, G. P., Bird, E. D., … Tourtellotte, W. W. (1987). Human brain D1 and D2 dopamine receptors in schizophrenia, Alzheimer's, Parkinson's, and Huntington's diseases. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *1*(1), 5–15.

Sidiropoulos, K., Viteri, G., Sevilla, C., Jupe, S., Webber, M., Orlic-milacic, M., … Fabregat, A. (2017). Reactome enhanced pathway visualization, *33*(July), 3461–3467. https://doi.org/10.1093/bioinformatics/btx441

Slenter, D. N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., … Willighagen, E. L. (2018). WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Research*, *46*(D1), D661–D667. https://doi.org/10.1093/nar/gkx1064

Soderberg, M. M., & Dahl, M.-L. (2013). Pharmacogenetics of olanzapine metabolism. *Pharmacogenomics*, *14*(11), 1319–1336. https://doi.org/10.2217/pgs.13.120

Stamatovic, S., Keep, R., & Andjelkovic, A. (2008). Brain Endothelial Cell-Cell Junctions: How to “Open” the Blood Brain Barrier. *Current Neuropharmacology*, *6*(3), 179–192. https://doi.org/10.2174/157015908785777210

Steen, V. M., Skrede, S., Polushina, T., López, M., Andreassen, O. A., Fernø, J., & Le, S. (2017). Genetic evidence for a role of the SREBP transcription system and lipid biosynthesis in schizophrenia and antipsychotic treatment. *European Neuropsychopharmacology*, *27*(6), 589–598. https://doi.org/10.1016/j.euroneuro.2016.07.011

Sugawara, H., Iwamoto, K., Bundo, M., Ueda, J., Miyauchi, T., Komori, A., … Kato, T. (2011). Hypermethylation of serotonin transporter gene in bipolar disorder detected by epigenome analysis of discordant monozygotic twins. *Translational Psychiatry*, *1*, e24. https://doi.org/10.1038/tp.2011.26

Supek, F., Bosnjak, M., Skunca, N., & Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*, *6*(7), e21800. https://doi.org/10.1371/journal.pone.0021800

Supek, F., & Skunca, N. (n.d.). Visualizing GO annotations, (1).

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews. Genetics*, *9*(6), 465–476. https://doi.org/10.1038/nrg2341

Swathy, B., & Banerjee, M. (2017). Understanding epigenetics of schizophrenia in the backdrop of its antipsychotic drug therapy. *Epigenomics*, *9*(5), 721–736. https://doi.org/10.2217/epi-2016-0106

Swathy, B., Saradalekshmi, K. R., Nair, I. V, Nair, C., & Banerjee, M. (2018). Understanding the influence of antipsychotic drugs on global methylation events and its relevance in treatment response. *Epigenomics*, *10*(3), 233–247. https://doi.org/10.2217/epi-2017-0086

Taha, Z., Janse van Rensburg, H. J., & Yang, X. (2018). The hippo pathway: Immunity and cancer. *Cancers*, *10*(4), 1–20. https://doi.org/10.3390/cancers10040094

Taylor, S. F., & Tso, I. F. (2015). GABA abnormalities in schizophrenia: a methodological review of in vivo studies. *Schizophrenia Research*, *167*(1–3), 84–90. https://doi.org/10.1016/j.schres.2014.10.011

Teicher, M. H. (2018). Childhood trauma and the enduring consequences of forcibly separating children from parents at the United States border. *BMC Medicine*, *16*(1), 4–6. https://doi.org/10.1186/s12916-018-1147-y

Teicher, M. H., Samson, J. A., Anderson, C. M., & Ohashi, K. (2016). The effects of childhood maltreatment on brain structure, function and connectivity. *Nature Reviews. Neuroscience*, *17*(10), 652–666. https://doi.org/10.1038/nrn.2016.111

The Gene Ontology, C. (2017). Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res*, *45*(D1), D331–D338. https://doi.org/10.1093/nar/gkw1108

Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., … Narechania, A. (2003). PANTHER: A library of protein families and subfamilies indexed by function. *Genome Research*, *13*(9), 2129–2141. https://doi.org/10.1101/gr.772403

Tunçer, S., Gurbanov, R., Sheraj, I., Solel, E., Esenturk, O., & Banerjee, S. (2018). Low dose dimethyl sulfoxide driven gross molecular changes have the potential to interfere with various cellular processes. *Scientific Reports*, *8*(1), 14828. https://doi.org/10.1038/s41598-018-33234-z

Tuplin, E. W., & Holahan, M. R. (2017). Aripiprazole, A Drug that Displays Partial Agonism and Functional Selectivity. *Curr Neuropharmacol*, *15*(8), 1192–1207.

https://doi.org/10.2174/1570159X15666170413115754

Urichuk, L., Prior, T. I., Dursun, S., & Baker, G. (2008). Metabolism of atypical antipsychotics: involvement of cytochrome p450 enzymes and relevance for drug-drug interactions. *Current Drug Metabolism*, *9*(5), 410–418.

Veldic, M., Guidotti, A., Maloku, E., Davis, J. M., & Costa, E. (2005). In psychosis, cortical interneurons overexpress DNA-methyltransferase 1. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(6), 2152–2157. https://doi.org/10.1073/pnas.0409665102

Vik-mo, A. O., Birkenaes, A. B., Fernø, J., Jonsdottir, H., Andreassen, O. A., & Steen, V. M. (2008). Increased expression of lipid biosynthesis genes in peripheral blood cells of olanzapine-treated patients, 679–684. https://doi.org/10.1017/S1461145708008468

Vik-Mo, A. O., Ferno, J., Skrede, S., & Steen, V. M. (2009). Psychotropic drugs up-regulate the expression of cholesterol transport proteins including ApoE in cultured human CNS- and liver cells. *BMC Pharmacol*, *9*, 10. https://doi.org/10.1186/1471-2210-9-10

Wan, E. S., Qiu, W., Baccarelli, A., Carey, V. J., Bacherman, H., Rennard, S. I., … Demeo, D. L. (2012). Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Human Molecular Genetics*, *21*(13), 3073–3082. https://doi.org/10.1093/hmg/dds135

Wesseling, H., Gottschalk, M. G., & Bahn, S. (2015). Targeted multiplexed selected reaction monitoring analysis evaluates protein expression changes of molecular risk factors for major psychiatric disorders. *International Journal of Neuropsychopharmacology*, *18*(1), 1–13. https://doi.org/10.1093/ijnp/pyu015

Woodward, N. D., & Heckers, S. (2016). Mapping Thalamocortical Functional Connectivity in Chronic and Early Stages of Psychotic Disorders. *Biological Psychiatry*, *79*(12), 1016–1025. https://doi.org/10.1016/j.biopsych.2015.06.026

Wright, M. L., Dozmorov, M. G., Wolen, A. R., Jackson-Cook, C., Starkweather, A. R., Lyon, D. E., & York, T. P. (2016). Establishing an analytic pipeline for genome-wide DNA methylation. In *Clin Epigenetics* (Vol. 8). https://doi.org/10.1186/s13148-016-0212-7

Xu, J., Burgoyne, P. S., & Arnold, A. P. (2002). Sex differences in sex chromosome gene expression in

mouse brain. *Human Molecular Genetics*, *11*(12), 1409–1419.

Yu, X., Yang, F., Fu, D.-L., & Jin, C. (2016). L1 cell adhesion molecule as a therapeutic target in cancer. *Expert Review of Anticancer Therapy*, *16*(3), 359–371. https://doi.org/10.1586/14737140.2016.1143363

Zechner, U., Wilda, M., Kehrer-Sawatzki, H., Vogel, W., Fundele, R., & Hameister, H. (2001). A high density of X-linked genes for general cognitive ability: a run-away process shaping human evolution? *Trends in Genetics : TIG*, *17*(12), 697–701.

Zhang, Y., Florath, I., Saum, K.-U., & Brenner, H. (2016). Self-reported smoking, serum cotinine, and blood DNA methylation. *Environmental Research*, *146*, 395–403. https://doi.org/10.1016/j.envres.2016.01.026

Ziller, M. J., Gu, H., Müller, F., Donaghey, J., Tsai, L. T.-Y., Kohlbacher, O., … Meissner, A. (2013). Charting a dynamic DNA methylation landscape of the human genome. *Nature*, *500*, 477. Retrieved from https://doi.org/10.1038/nature12433

Zouboulis, C. C., Chen, W.-C., Thornton, M. J., Qin, K., & Rosenfield, R. (2007). Sexual hormones in human skin. *Hormone and Metabolic Research = Hormon- Und Stoffwechselforschung = Hormones et Metabolisme*, *39*(2), 85–95. https://doi.org/10.1055/s-2007-961807

9. Appendix

Genome-wide quantification of DNA methylation

Data for 1000 samples obtained using the Illumina EPIC platform was imported to R and preprocessed for data analysis, using BioConductor packages minfi (Aryee et al., 2014) and watermelon (Pidsley et al., 2013), the following way: 1) 19 samples having 1% of sites with a detection p-value greater than 0.05 were removed, as well as 1762 sites with a bead count less than 3 in 5% of the samples and 6614 sites having 1% of samples with a detection p-value greater than 0.05 were removed.

2) The dataset was normalised using Dasen (Pidsley et al., 2013).

3) Removed probes on the X and Y chromosomes.

4) Removed probes flagged by (Zhou et al., 2017) using their supplementary file EPIC.manifest.pop.rda and MASK.general.EUR column. Here probes that may have issues with cross-hybridisation or contain snps close to the target CPG are marked for removal.

5) Compared SNPs from the EPIC array to genotype data from the same samples and removed samples that did not match.

6) Calculated pairwise correlation between all and removed samples with a correlation to another sample in the dataset that was higher than 0.9. The dataset passing quality control had 957 samples and 776023 probes.

Prediction of cell type composition

Cell type composition was predicted using the estimateCellCounts function in the minfi package.

Preprocessing of the HepG2 data

Preprocessing of the HepG2 data was done using a similar pipeline as for the blood samples:

1) No samples were removed due to high detection p-value, 6769 sites were removed.

2) The dataset was normalised using Dasen (Pidsley et al., 2013).

3) Removed probes on the X and Y chromosomes.

4) Removed probes flagged by (Zhou et al., 2017) using their supplementary file EPIC.manifest.pop.rda and MASK.general.EUR column.

References

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*(10), 1363–1369.

https://doi.org/10.1093/bioinformatics/btu049

Pidsley, R., Y Wong, C. C., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, *14*, 293. https://doi.org/10.1186/1471-2164-14-293

Zhou, W., Laird, P. W., & Shen, H. (2017). Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Research*, *45*(4), e22. https://doi.org/10.1093/nar/gkw967