The Genetics and Epigenetics of Substance Misuse
An Investigation into Life Stress and Reward Processing in Adolescence

Wong, C Peng

Awarding institution:
King's College London

THE GENETICS AND EPIGENETICS OF SUBSTANCE MISUSE: AN INVESTIGATION INTO LIFE STRESS AND REWARD PROCESSING IN ADOLESCENCE

BY

C PENG WONG

2014

A THESIS SUBMITTED TO KING'S COLLEGE LONDON FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN SOCIAL, GENETIC & DEVELOPMENTAL PSYCHIATRY

Abstract

Substance misuse and addictions impose serious health and socio-economic consequences for both individuals and societies. Substance use during adolescence predicts the severity of addictions in later life, indicating that adolescence is an important milestone for developing addictions. Alcohol and tobacco are the most common form of substance use in adolescents. The heritability estimates of alcohol and tobacco addictions range between 30 - 70%, suggesting that both genetic and environmental factors could contribute to the risks of addictions.

This PhD thesis aimed to identify the genetic and epigenetic factors in alcohol and tobacco misuse in adolescence. The impact of life stress and circadian system on reward sensitivity and substance use was investigated in over 2000 adolescents from the IMAGEN Study. In the first study, both negative life events and ventral striatal activations during reward anticipation were shown to predict the increased alcohol and tobacco use in adolescents. The second study investigated the relationships among reward sensitivity, substance use and the DNA methylation in *PERIOD 1*, a circadian gene that was found to associate with stress-mediated alcohol use in adolescence (Dong et al., 2011). The third study explored the additive genetic and polygenic effects of single nucleotide polymorphisms (SNPs) in the stress and circadian systems on reward sensitivity and substance use. Results from the second and the third studies showed that reward sensitivity, alcohol and tobacco use in adolescents were not associated with the DNA methylation in *PERIOD 1*, and the genetic polymorphisms within the stress and circadian systems. The fourth study examined the genetic influence of stress systems on alcohol misuse. SNP rs1409837 on FYN oncogene related to SRC (*FYN*) gene was found to associate with the reduced drunkenness and binge drinking in adolescents at age 16. When investigating the role of *FYN* rs1409837 on brain functions, *FYN* rs1409837 was found to associate with the reduced amygdala activations during angry face processing. This thesis highlights the genetic influence of life stress on alcohol misuse and provides new approaches that should aid the understanding of genetic underpinnings of substance misuse in adolescence.

Table of Contents

Table of Figures

Table of Tables

Acknowledgements

I would like to thank Dr. Sylvane Desrivières, Prof. Jonathan Mill and Prof. Gunter Schumann for their guidance throughout my PhD. This PhD has been funded by the European Commission FP7 IMAGEN Study under the lead of Prof. Gunter Schumann, hence making all work arising from this thesis possible.

I would like to acknowledge the past and present members of the IMAGEN team, especially Dr. Tianye Jia, Dr. Barbara Ruggeri, Dr. Charlotte Nymberg, Dr. Anna Cattrell, Dr. David Stacey, Dr. Fabiana Carvalho, Dr. Tom Walker-Tilley, Dr. Toni-Kim Clarke, Dr. Eva Loth, Dr. Steven Lubbe, Dr. Anbarasu Lourdusamy, Dr. Alinda Fernandes, Christine Macare, Agnes Kepa, and Poushali Ganguli, who have offered invaluable advice and assistance in navigating the IMAGEN database. I thank the past and present members of the Psychiatric Epigenetics team - Dr. Emma Dempster, Dr. Rebecca Smith, Dr. Chloe Wong, Dr. Ruth Pidsley, Dr. Rachel Kember, Dr. Yogen Patel and Helen Spiers for their technical support and delicious cakes. I thank Dr. Xiaohui Xu, Dr. Conrad Iyegbe and Dr. Karim Malki for their encouragement.

Special thanks to Mum, Dad and Clio, who are always there for me and inspiring me to become the best person I can be. Thank you to my girls Hibba, Lourdes, Nicola, Gigi, Stephanie and Sakina for always believing in me and giving me endless support throughout my PhD.

Declaration of Work

This thesis is based on data from the IMAGEN Study which received funding from the European Commission. The data of the IMAGEN Study was collected and preprocessed by members of the IMAGEN Consortium. In particular, the data used in this thesis had been preprocessed by

Dr. Anne Boland-Auge and team (The Centre National de Génotypage, France), who extracted whole blood DNA and generated the genome-wide genetic data;

Dr. Anbarasu Lourdusamy and Dr. Tianye Jia (Institute of Psychiatry, London), who performed quality control for the genome-wide genetic data;

Dr. Vincent Frouin and team (NeuroSpin, France), who performed SNP imputation, probabilistic mapping and quality control for the functional magnetic resonance images;

Dr. Roman Sasik and team (BIOGEM Laboratory, University California San Diego), who extracted whole blood mRNA and performed quality control procedure for the genome-wide mRNA expression data; and

Dr. Anna Cattrell (Institute of Psychiatry, London), who defined the negative life events from the Life Events Questionnaire.

I was responsible for generating all research questions and conducting all statistical analyses presented in this thesis. I performed the experiments in Chapter 4 and the quality control for the DNA methylation and the preprocessed behavioural, neuroimaging and mRNA expression data. Being part of the IMAGEN Consortium, I was responsible for managing the biological samples during 2010 - 2012. I declare that the work presented in this thesis is my own. This thesis has not been submitted for any other degree at any other university.

Abbreviations

5-HT	5-hydroxytryptamine
3-HPA	3-hydroxypicolinic acid
5-MeC	5-methylcytosine
A	Adenine
AAL	Automated Anatomical Labeling
ADH	Alcohol dehydrogenase
AHRR	Aryl hydrocarbon receptor repressor
ALDH	Aldehyde dehydrogenase
ACTH	Adrenocorticotropic hormone
AUTS2	Autism susceptibility candidate 2 gene
AVP	Arginine vasopressin
BNST	Bed nucleus of the Stria Terminalis
BOLD	Blood oxygen level-dependent
C	Cytosine
CDCV	Common-disease common-variant
CeA	Central extended amygdala
CHRN	Nicotinic acetylcholinergic receptor subunit
CEU	Utah residents with ancestry from northern and western Europe
C.I.	Confidence interval
COMT	Catechol-O-methyl transferase
CRH	Corticotrophin releasing hormone
CRHR1	Corticotropin releasing hormone receptor 1
Cry	Cryptochrome
DA	Dopamine
DAT	Dopamine transporter
Disc1	Disrupted In Schizophrenia 1
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DRD2	Dopamine D2 receptor
DZ	Dizygotic
ENCODE	Encyclopedia of DNA Elements
EPSPs	Excitatory postsynaptic potentials
ESPAD	European School Survey Project on Alcohol and Other Drugs
FDR-BH	Benjamini-Hochberg false discovery rate
fMRI	Functional magnetic resonance imaging
F2RL3	Coagulation factor II (thrombin) receptor-like 3
FYN	FYN oncogene related to SRC
G	Guanine
GABRA2	γ-aminobutyric acid A receptor
GC	Glucocorticoids
GO	Gene Ontology
GR/ NR3C1	Glucocorticoid receptors
GRE	Glucocorticoid response element

GWAs	Genome-wide association studies
HPA	Hypothalamic-pituitary-adrenal
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Locus coeruleus
LD	Linkage disequilibrium
LEQ	Life Events Questionnaire
MAF	Minor allele frequency
MALDI-TOF MS	Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry
MAOA	Monoamine oxidase A
MID	Monetary incentive delay
MNI	Montreal Neurological Institute
mPFC	Medial prefrontal cortex
MR	Magnetic resonance
MZ	Monozygotic
NAc	Nucleus accumbens
NCBI	National Centre for Biotechnology Information
NMDA	N-methyl-D-aspartate
NR3C2/ MR	Mineralocorticoid receptor
PER	Period gene
PDS	Puberty Development Scale
POMC	Pro-opiomelanocortin
PVN	Paraventricular nucleus
Q-Q	Quantile-quantile
RHT	Retinohypothalamic tract
RNA	Ribonucleic acid
ROI	Region-of-interest
SAM	S-adenosyl methionine
SAP	Shrimp alkaline phosphotase
SCN	Suprachiasmatic nucleus
SNAP	SNP Annotation and Proxy Search
SNP	Single nucleotide polymorphism
SPM	Statistical Parametric Mapping
T	Thymine
TF	Transcription factor
TH	Tyrosine hydroxylase
Tm	Melting temperature
U	Uracil
VIP	Vasoactive intestinal polypeptide
VS	Ventral striatum
VTA	Ventral tegmental area
WebGestalt	WEB-based Gene SeT AnaLysis Toolkit
ZT	Zeitgeber time

Chapter 1

General Introduction

1.1 Substance misuse and addictions[1]

The use of psychoactive substances in primates and humans can be traced back to 200 million years ago. Archaeological records indicate the use of nicotine from pituri plants *Duboisia hopwoodii* and *Nicotiana gossel* among the Australian aborigines around 40000 years ago (Saah, 2005). The use of alcohol dates back to 50 million years ago, when the pentailed tree shrews (mammals closely related to the ancestors of modern primates) in West Malaysia consumed nectar from the bertam palm *Eugeissona tristis* that contained about 3.8% concentration of alcohol (Spanagel, 2009).

Nowadays, substance misuse and addictions impose serious health and socio-economic consequences for both individuals and societies. Revealed by the Global Burden of Diseases, Injuries and Risk Factors Study, the contribution of substance use to disease outcomes has increased by 3% - 57% during 1990 to 2010 (Lim et al., 2012). Alcohol and substance addictions account for 86.1 % years of life lost to premature mortality, 20.5 % disability-adjusted life years and 17.3% years lived with disability in worldwide populations (Whiteford et al., 2013). Understanding the aetiology of substance addictions will be crucial in order to prevent substance use related-problems and addictions.

[1] This chapter was adapted from Wong & Schumann (2012). Integration of the circadian and stress systems: influence of neuropeptides and implications for alcohol consumption. *Journal of Neural Transmission, 119,* 1111-1120.

The development of addiction is a multi-stage process characterised by experimentation, repeated use, tolerance, withdrawal and relapse of addictive substances (**Figure 1-1**). Individuals' experience from the occasional, recreational substance use reinforces and shapes their consumption behaviour. Over time, repeated substance use leads to the development of tolerance and loss of control over substance intake, resulting in compulsive substance seeking behaviour as well as difficulties in withdrawal. The addiction cycle can be conceptualised into three main components (Koob and Le Moal, 1997). The Preoccupation/ Anticipation stage describes substance users engage in substance seeking behaviour and craving; followed by the Binge/ Intoxication stage where individuals consume substance excessively and results in substance tolerance and intoxications; and abstinence from substance creates negative affect among substance users at the withdrawal stage. The negative emotion associated with substance withdrawal further encourages craving and substance intake. These three components feed into each other and escalates substance use and addictions.

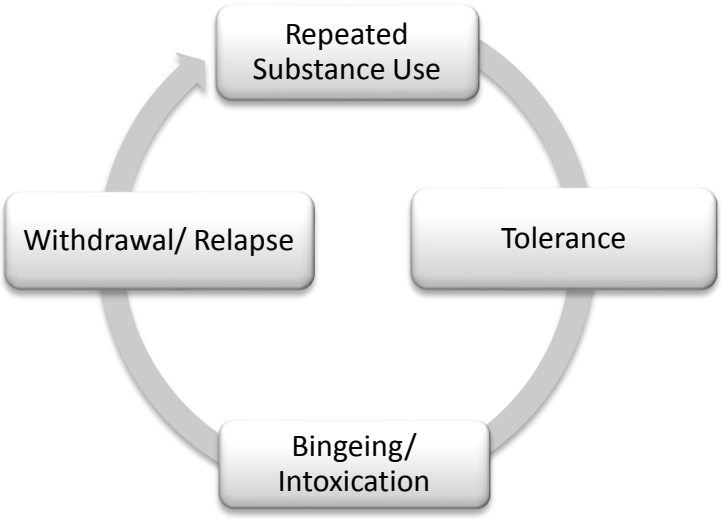


Figure 1-1 The development of addictions. Followed by the experimentation of addictive substance, individuals develop repeated, regular substance use over time. Progressive tolerance to the addictive substance promotes excessive use/ bingeing that are followed by intervals of withdrawal. The negative emotion arises from substance withdrawal causes relapse into substance use, forming a vicious cycle of addiction.

1.2 Substance use in adolescents

Adolescence is a developmental period that occurs between the ages of 12 to 18 (Spear, 2000). During this transitional period, adolescents experience numerous behavioural and physiological changes which will equip individuals with necessary skills for survival and independence in adulthood. These behavioural changes are accompanied by continuous brain maturation (Giedd et al., 1999, Wolf et al., 2013, Blakemore and Mills, 2013, Giedd et al., 1996a). Immature neural connectivity due to the lag between cortical and subcortical brain development may predispose as risks of brain dysfunctions and vulnerability to substance addictions during adolescence (Giedd et al., 1996a). According to the National Comorbidity Survey Replication study in which over 9000 adolescents in the United States were studied during 2001 - 2003, the age of onset for substance use disorders spans from age 14 to adulthood (Kessler et al., 2005). Early onset of substance use predicts comorbid substance use and the severity of addictions in late adolescence and early adulthood (Chambers et al., 2003, Kandel et al., 1992).

Alcohol and tobacco are the two most common forms of substance use (Kendler et al., 2012). According to the National Longitudinal Alcohol Epidemiologic Survey in 1992, over 40% of individuals who started drinking at age 14 or younger had developed alcohol dependence in adulthood (Grant and Dawson, 1997). On the contrary, only 10% of the alcohol dependent individuals started drinking after age 20 (Grant and Dawson, 1997). Likewise, more than 71% of adult smokers reported tobacco consumption before age 18 (Giovino et al., 1995). Therefore, substance use in adolescence is an important milestone for the development of substance misuse and addictions in adulthood.

1.2.1 Heritability of alcohol and tobacco use and addictions

The genetic component of substance use has been demonstrated using twin models. Since monozygotic (MZ) twins are genetically identical and dizygotic (DZ) twins share about 50% of their genetic identity, the genetic component of a disorder or behavioural trait can be confirmed

if higher concordance in the MZ twins compared to the DZ twins is shown. Results from twin studies indicate that the heritability estimates for nicotine dependence range between 30 - 70% (Broms et al., 2007, Kendler et al., 1999), whereas the heritability estimates for alcohol consumption and dependence lies within 30 - 60% (Hansell et al., 2008, Goldman, 1993). Therefore, both genetic and environmental factors can contribute to the development of nicotine and alcohol dependence.

1.2.2 Genetics for alcohol and tobacco use and addictions

The first draft of human genome sequence was completed in 2000 (Lander et al., 2001). The estimated size of human genome is about 3000 mega base pairs, of which 15 - 20% are gene and gene-related sequences. Owing to the advances in genotyping and sequencing techniques, our knowledge in genetic variations and complex disorders has been expanding in the past decade. One of the most studied genetic variations is single nucleotide polymorphism (SNP), which is defined as the variation of deoxyribonucleic acid (DNA) sequence occurring at a single nucleotide at a given position in the genome. Over 38 million validated SNPs across human populations have been documented in the latest SNP database in the National Centre for Biotechnology Information (NCBI dbSNP build 137, released in June 2012).

The genetic vulnerability to substance use and addictions can also be studied in unrelated individuals using genome-wide associations (GWAs) and candidate gene association analyses. Both GWAs and candidate gene association analyses allow the identification of genetic polymorphisms that may associate with disorders or traits of disorders. Conducting GWAs does not require any prior knowledge on the biological basis of phenotypes or disorders, thus providing an unbiased, exploratory approach to screen genetic polymorphisms that may predispose as risk factors for gene functions or disorders. On the other hand, candidate gene association enables the genetic effects of target regions or genes to be investigated.

Various SNPs associated with alcohol and nicotine addictions have been identified in previous studies. The genetic polymorphisms of alcohol metabolising enzymes including alcohol dehydrogenase (*ADH*) and aldehyde dehydrogenase (*ALDH*) were found to associate with alcohol intake and dependence (Frank et al., 2012, Treutlein et al., 2009). Individuals who carried the *ADH1B*2* allele displayed accumulation of acetaldehyde in blood due to the reduced activity of ADH1B enzyme (Rivera-Meza et al., 2010). Acetaldehyde is a toxic intermediate that produces adverse reaction towards alcohol including flushing and nausea, therefore the accumulation of acetaldehyde in blood prevents excessive alcohol intake (Edenberg, 2007). Functional genetic polymorphisms on genes related to receptor signalling and neurotransmission have been found to associate with alcohol addictions. For example, SNP rs6943555 on autism susceptibility candidate 2 gene (*AUTS2*) was associated with alcohol consumption in the European population and the downregulation of *AUTS2* might reduce the sensitivity to alcohol in the *Drosophila* (Schumann et al., 2011). Various SNPs on the alpha-2 subunit of γ-aminobutyric acid A receptor (*GABRA2*) were found to associate with alcohol dependence and neural excitation (Bierut et al., 2010, Edenberg et al., 2004), demonstrating the profound consequences of alcohol on brain functions.

Cigarette smoking and nicotine dependence were found to associate with the genetic polymorphisms of genes coding the nicotinic acetylcholinergic receptor subunits *CHRNB3, CHRNA3, CHRNB4, CHRNA5* (Tobacco and Genetics, 2010, Thorgeirsson et al., 2010, Nees et al., 2013). Nicotine is a major psychoactive components present in tobacco and it targets the nicotinic receptors that are distributed in various brain regions (O'Hara et al., 1998, Markus et al., 2003). It was thought that the genetic polymorphisms on nicotinic receptor genes alter the sensitivity to nicotine and tobacco in the brain (Stevens et al., 2008). Researchers have endeavoured to identify genetic risk variants and narrow the missing heritability in substance use and addictions. Nonetheless, evidence from the current findings supports the polygenic risk of substance addictions.

1.2.3 Epigenetics of substance use and addictions

1.2.3.1 Background

In addition to the genetic polymorphisms, epigenetic mechanisms can play an important role in determining disease susceptibility and progression. The term 'epigenetics' was coined by Conrad H. Waddington in early 1940s to describe the development of phenotypes as a result of gene-environment interaction. In the epigenetic landscape model postulated by Waddington (1942), the developmental events are 'canalised', hence forming various developmental trajectories (Waddington, 1942) (**Figure 1-2**). The valleys and canals of the landscape can be determined by the genetic influence. Imagine a ball rolls down the landscape, the endpoint of the ball will be determined by the paths it has taken. The endpoint of the ball represents the phenotypes that are produced during the development processes in the epigenetic landscape. Nowadays, epigenetics is defined as the study of alterations in phenotypes, chromatin structures and gene transcriptions that are not due to genetic variations. Epigenetic markers are thought to be reprogrammable, reversible and heritable across generations (Lande-Diner and Cedar, 2005, Herb et al., 2012, Reik, 2007).

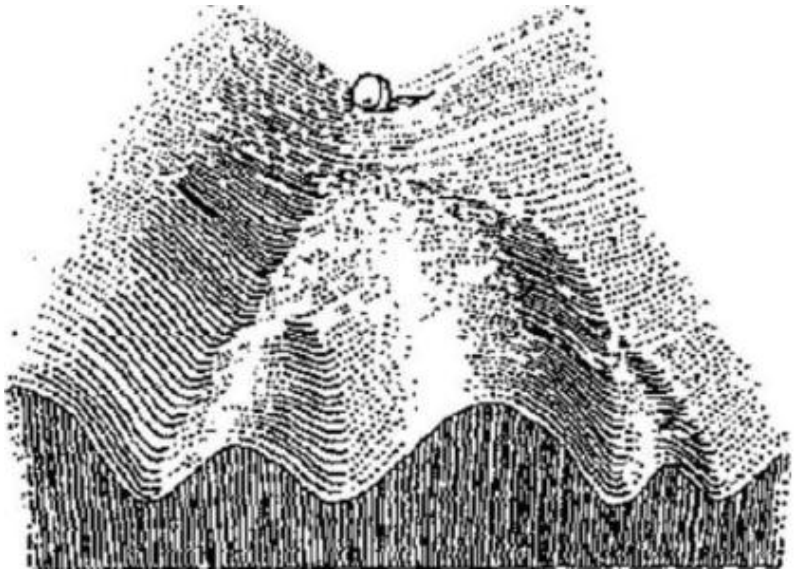
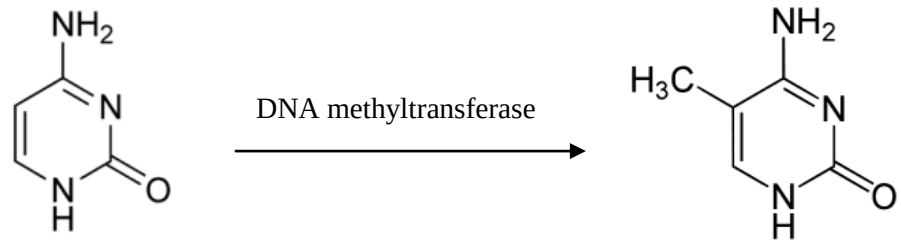


Figure 1-2 Waddington's epigenetic landscape model (1942).

1.2.3.2 DNA methylation in eukaryotes

Various forms of epigenetic modification have been identified in the past decade. Among all identified epigenetic modifications, 5'-cytosine methylation (hereafter DNA methylation) is one of the most commonly studied epigenetic mechanisms in the human genome. It is also the earliest form of DNA modification found in eukaryotic organisms (Bird, 1980, Razin and Riggs, 1980). DNA (5'-cytosine) methylation refers to the addition of methyl group by DNA methyltransferase at the 5th carbon position of cytosine in 5'-CpG-3' dinucleotides (**Figure 1-3**). At least three DNA methyltransferases (DNMT), including DNMT1, DNMT3A and DNMT3B have been identified. DNMT1 preferentially binds to hemimethylated CpG dinucleotide and it is thought to maintain the DNA methylation pattern during replication (Denis et al., 2011, Robertson and Wolffe, 2000). On the other hand, members of the DNMT3 family may be responsible for *de novo* DNA methylation since they target to both hemi- and unmethylated CpG dinucleotides (Denis et al., 2011, Robertson and Wolffe, 2000). All of these DNA methyltransferases use S-adenosyl methionine (SAM) as the methyl donor during the process of DNA methylation (Goll and Bestor, 2005).



Cytosine 5'-Methylcytosine

Figure 1-3 5'-cytosine methylation. The conversion of cytosine to 5'-cytosine methylation (DNA methylation) is catalysed by DNA methyltransferases. All known DNA methyltransferases use S-adenosyl methionine (SAM) as the methyl donor.

In the human genome, regions with over 50% CG content and 0.6 or greater observed-to-expected CpG dinucleotides ratio occurring within a sequence window of at least 200bp are classified as CpG islands (Gardiner-Garden and Frommer, 1987). Over 60% of human genes contain CpG islands are located within gene promoters and the locations of CpG islands are thought to be evolutionarily conserved (Antequera, 2003). DNA methylation at CpG dinucleotides has been linked to transcriptional silencing, which can be crucial to the development and disease progressions (Jaenisch and Bird, 2003). One speculation is that DNA methylation at CpG dinucleotides alter gene transcription efficiency, potentially by reducing the binding affinity of transcriptional factor binding sites (Watt and Molloy, 1988). Other modifications of DNA sequence such as 5'- hydroxymethylation have also been reported after the discovery of 5'-cytosine methylation (Tahiliani et al., 2009).

1.2.3.3 DNA methylation in substance use disorders

The investigation of DNA methylation in complex disorders has come into focus in recent years. The importance of DNA methylation has been implicated in genomic imprinting, embryonic development, silencing of transposable elements and gene expressions. These events may further contribute to the progression and transmission of diseases across generations (Jirtle and Skinner, 2007). Environmental stimuli such as the exposure to radiation and nutritional supplements can also alter the epigenetic landscape as well as the behavioural phenotypes in offsprings (Koturbash et al., 2006, Waterland and Jirtle, 2003).

Since the heritability for substance addictions ranges from 30 - 70% (Broms et al., 2007, Kendler et al., 1999, Hansell et al., 2008, Goldman, 1993), the propensity of alcohol and tobacco use can be modified by various environmental factors. For example, the production of acetyaldehyde is thought to disrupt the production of methyl donor SAM as well as the DNA methylation profiles (Selhub, 1999). Acetyaldehyde inhibits methionine synthase to convert

homocysteine into methionine, which subsequently alter the level of SAM and DNA methylation profile (Selhub, 1999).

Recent studies indicate alterations of DNA methylation landscape in alcohol and tobacco dependent populations. Epigenome-wide analyses revealed differential DNA methylation patterns among patients with alcohol dependence and healthy controls, of which hypomethylation of genes related to alcohol metabolism e.g. alcohol dehydrogenase 1A and *ALDH* were observed in the patients (Zhang et al., 2013). Conversely, hypermethylation and reduced mRNA expression of *DNMT3B* were found in alcohol dependent patients (Bonsch et al., 2006). Several epigenetic loci have been identified to associate smoking behaviour. For example, hypomethylaton of CpG sites on the coagulation factor II (thrombin) receptor-like 3 (*F2RL3*) and aryl hydrocarbon receptor repressor gene (*AHRR*) were consistently observed among smokers (Breitling et al., 2011, Zeilinger et al., 2013, Shenker et al., 2013). Differential protein binding affinity was observed at *AHRR,* suggesting the potential influence of DNA methylation in gene regulations (Zeilinger et al., 2013). Together the results suggest the importance of DNA methylation in predicting substance use.

The propensity of substance use and addictions can be influenced by various factors. In the following sections, I am going to address the relevance of the stress system, circadian system and reward processing in substance use during adolescence. The significance for interactions of these systems on substance use will be discussed.

1.3 The role of stress in substance use

1.3.1 HPA axis and extrahypothalamic stress system

Coping with life stress and negative emotion is essential for the maintenance of psychological well-being in humans (Johnson and Pandina, 1993). In human, the hypothalamic-pituitary-

adrenal (HPA) axis is responsible for monitoring the initiation and feedback of stress responses (Wust, 2004). Exposure to stressful stimuli triggers the production and the release of corticotrophin releasing hormone (CRH) from CRH-neurons in the paraventricular nucleus (PVN) that is transported to the anterior pituitary via the hypophyseal portal system (Antoni, 1986). Upon the binding to CRH receptors, CRH and stress hormones such as arginine vasopressin (AVP) stimulate the synthesis and the release of pro-opiomelanocortin (POMC) and adrenocorticotropic hormone (ACTH) from the anterior pituitary (Simpson and Waterman, 1988). Subsequently ACTH is transported to the adrenal gland and triggers the production of glucocorticoids such as cortisol.

Glucocorticoids regulate stress responses through binding to mineralocorticoid (MR or NR3C2) and glucocorticoid receptors (GR or NR3C1). Both MR and GR are expressed in various tissues and brain regions including the pituitary, hippocampus, amygdala, PVN and hypothalamus (De Kloet et al., 1998). The ligand binding affinity of MR is approximately 5 - 10 fold higher compared to the GR (Reul and de Kloet, 1985). The high affinity of MR maintains receptor activation up to 1 hour (Reul & de Kloet, 1985). GR can only be activated by a large amount of glucocorticoids due to its low ligand-binding affinity (5-10 nM); therefore GR is thought to mediate the delayed inhibition of stress responses (Reul and de Kloet, 1985, de Kloet et al., 2005, Koob, 2003). Upon the binding of glucocorticoids, GR translocates to the nucleus as homodimers or heterodimers with MR. The dimers recruit transcription factors at the glucocorticoid response element (GRE) sites and regulate gene transcriptions (De Kloet et al., 1998). Increase cortisol level promotes additional GR activation, long-term alterations in gene transcriptional activities and neural activations (Joels and de Kloet, 1992).

In addition to the HPA axis, the extrahypothalamic stress system is responsible for monitoring stress responses. The CRH-containing neurons in the PVN projects to the surrounding limbic regions including the central extended amygdala (CeA), hippocampus, bed nucleus of the stria

terminalis (BNST), nucleus accumbens (NAc) and the locus coeruleus (LC) (Herman et al., 2003), allowing the limbic regions receive stress signal from the HPA axis. The activation of the extrahypothalamic stress system can enhance the recruitment of dynorphin, a class of opioid peptide that preferentially binds to κ-opioid receptors and involves in producing aversive dysphoric-effects in mammals (Wee and Koob, 2010, Shippenberg et al., 2007). The neural projections from the extrahypothalamic brain regions to the PVN allow the modulations of both voluntary and reflexive stress responses (Herman et al., 2003).

1.3.2 Stress-induced substance use

The detrimental effect of stress on substance use has been implicated in both healthy and clinical populations. Adolescents were found to consume more alcohol and tobacco when they experienced more stressful life events (Blomeyer et al., 2008, Booker et al., 2008). The elevation of stress level is accompanied by alterations in brain activations. Upon the presentation of stress- or alcohol-related visual cues, patients with alcohol dependence showed increased alcohol craving and blunted brain activations in the prefrontal cortex, anterior cingulate and the amygdala (Seo et al., 2013, Breese et al., 2011). Such observation could be driven by alterations of the HPA axis activity, as higher baseline cortisol level was observed among the smokers and the alcohol dependent patients (Steptoe and Ussher, 2006, Sinha et al., 2009). Family history of substance addictions could be another reason leading to differential HPA axis and stress system reactivity, as children of alcohol-dependent fathers showed higher level of stress-related hormones at baseline compared to children with no family history of alcohol addictions (Zimmermann et al., 2004). At molecular level, the exposure to ethanol or nicotine was found to reduce the mRNA expressions of stress genes including the *CRH* and *c-fos* in the PVN (Richardson et al., 2008, Schmitt et al., 2008). After prolonged ethanol treatment, foetal rats showed reduced mRNA expressions of *POMC* and adenylyl cyclase in neurons expressing β-endorphin (Sarkar et al., 2007). At the same time, the reduction of β-endorphin neurons in the arcuate nuclei and the PVN was also observed (Sarkar et al., 2007).

Genetic polymorphisms and epigenetic modifications of stress genes have been implicated in alcohol and tobacco use. Illustrated by Blomeyer and colleagues (2008), stressful life events can moderate the genetic influence of corticotropin releasing hormone receptor 1 (*CRHR1*) rs1876831 on heavy alcohol use in adolescents (Blomeyer et al., 2008). Nevertheless, DNA methylations of stress genes have been shown to associate with substance use. For example, smoking during pregnancy has been associated with changes in cord blood and placenta methylomes (Joubert et al., 2012, Suter et al., 2011). Some of these candidate CpG loci are located in genes involved in the regulation of oxidative stress (Suter et al., 2011). Differential DNA methylation pattern at *POMC* promoter was observed in patients with alcohol dependence compared to healthy individuals (Muschler et al., 2010). Alterations in the peripheral blood methylome were found among patients with alcohol dependence, in which CpG loci with differential DNA methylation were found to locate in genes involved in stress, signal transduction and immune responses (Zhang et al., 2013). Together the evidence has confirmed the gene-environment interaction of stress-induced substance use in humans.

1.4 Disruption of circadian rhythm as risk factor for substance use

1.4.1 Regulation of the circadian rhythm

Adolescents demonstrate a phase delay for sleep-wake schedule and reduction in sleep duration (Sadeh et al., 2009). The shift in circadian rhythm during adolescence can lead to alterations in arousal, mood, daily activities and brain functions (Spear, 2000, Benca et al., 2009). In mammals, the circadian rhythm is operated by a series of endogenous clocks that oscillate in an approximate 24-hour period (Dunlap, 1999). The circadian clock system is organised in a hierarchical manner, whereby the master clock is located at the suprachiasmatic nucleus (SCN), a highly conserved brain region situated directly above the optic chiasm at the anterior hypothalamus. Brain regions including the arcuate nucleus, medial preoptic area, dorsomedial hypothalamus and the PVN receive input from the SCN (Kalsbeek et al., 2007).

Feedback mechanisms of the circadian cycle

Entrainment of environmental stimuli (e.g. light) activates the SCN and the transcription-translational feedback machinery of the circadian system (Bernard et al., 2007). Upon the activation of the SCN, core circadian components Clock and Bmal1 form heterodimers and bind to the enhancer-box (E-Box) (Gekakis, 1998, Hogenesch et al., 1998), allowing the transcriptions of circadian genes Period (*Per 1*, *Per 2* and *Per 3*) and Cryptochrome (*Cry 1* and *Cry 2*) (Sun et al., 1997, Shearman et al., 1997, Takumi et al., 1998, Miyamoto and Sancar, 1998, Thresher, 1998) (See **Figure 1-4**). The *Per* and *Cry* genes are expressed at different times of the day (Zylka et al., 1998). In mammals, the highest level for *Per1* in the SCN is observed at around 4 hours after the initiation of the circadian cycle (Takumi et al., 1998, Bae et al., 2001). The expressions of *Per2* and *Per3* gradually increased from the beginning of light-dark cycle, with the peak expression levels observed at Zeitgeber time (ZT) 12-16 and ZT4-12 respectively (Takumi et al., 1998, Bae et al., 2001).

Accumulation of circadian genes and proteins due to continuous light exposure brings the circadian cycle to the feedback phase (**Figure 1-4**). Increased expression of orphan nuclear receptor Rev-Erbα inhibits the transcription of *Bmal1* by acting on the Rev-Erbα and ROR response elements (Preitner et al., 2002, Ueda et al., 2002). On the other hand, accumulation of Per and Cry proteins enhances the formation of Per:Cry complex with casein kinases CKIε and CKIδ (Camacho et al., 2001). The resulting Per:Cry:CKIε/ CKIδ complex inhibits *Per* and *Cry* transcriptions by preventing the binding of Clock:Bmal1 complex to E-box, subsequently reduce the transcription of *Rev-Erbα* (also known as nuclear receptor subfamily 1, group D member 1, *Nr1d1*). As a result, the reduction of *Rev-Erbα* activity disinhibits the repression of *Bmal1* transcription and resets the circadian cycle (Preitner et al., 2002). The core circadian component Clock was found to act as histone acetyltransferase, suggesting that epigenetic regulations might be essential for adjusting the circadian rhythm (Doi et al., 2006). Genome-wide gene expression data showed that about 10% of gene transcripts display circadian

oscillations (Akhtar et al., 2002, Duffield et al., 2002). In particular, genes containing circadian motifs such as E-boxes and ROR response elements are rhythmically expressed (Preitner et al., 2002, Ueda et al., 2002).

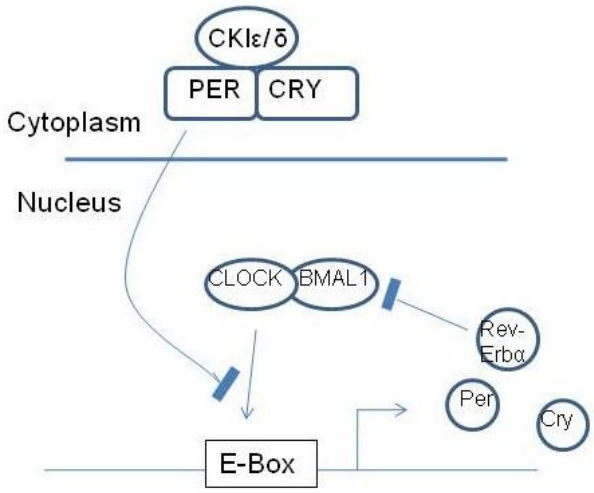


Figure 1-4 Transcriptional-translational feedback mechanism of the circadian system. The circadian cycle is mediated by a series of transcriptional-translational events. The transcriptions of circadian genes including Period (*Per 1*, *Per 2* and *Per 3*), Cryptochrome (*Cry 1* and *Cry 2*) and *Rev-Erbα* are initiated upon the binding of the CLOCK:BMAL1 heterodimer to E-Box. Accumulation of circadian genes and proteins in nucleus brings the circadian cycle to feedback phase, where the formation of Per:Cry:CKIε/ CKIδ protein complex inhibits transcriptional activity by preventing CLOCK:BMAL1 from binding to the E-box. Reduced *Rev-Erbα* level disinhibits the repression of *BMAL1* transcription. Restoring the *BMAL1* transcription resets the circadian clock by enhancing the formation of CLOCK:BMAL1 complex and subsequent transcriptions of circadian genes.

1.4.2 Circadian rhythm and substance use

The role of circadian rhythm in alcohol consumption has been supported by behavioural studies in animals. In a free choice condition where the amount of ethanol intake was not restricted, both rats and hamsters demonstrated alterations in the circadian cycle (Rosenwasser et al.,

2005a, Rosenwasser et al., 2005b, Seggio et al., 2009). Forced ethanol intake significantly shortened the free-running period in mice (Seggio et al., 2009) and these animals were more likely to develop high ethanol preference (Rosenwasser et al., 2005c, Hofstetter et al., 2003).

The chronotype of adolescents has been linked to tobacco smoking. Individuals who exhibited late chronotype were associated with greater tobacco use and poor sleeping quality (Wittmann et al., 2010). Finnish adolescents who exhibited late chronotype were also less likely to quit smoking in a tobacco use intervention programme (Heikkinen et al., 2009). Administration of nicotine to rats was found to disrupt the frequency of neuronal firing in the SCN (Trachsel et al., 1995), which might alter the neural activity in the SCN and circadian rhythm.

The mRNA expressions of circadian genes can be modified by alcohol and tobacco use. Exposure to ethanol was found to reduce *Per2* expression but increase *Per1* and *Bmal1* expressions in the SCN of rats (Farnell et al., 2008); whereas downregulations of *Rev-Erbα* lung mRNA and *c-Fos* protein in the SCN were observed in rats exposed to cigarettes (Vasu et al., 2009, Ferguson et al., 1999). Downregulations of *CLOCK, BMAL1, PER1, PER2, CRY1* and *CRY2* mRNA expressions in the peripheral blood were also observed among patients with alcohol dependence compared to the healthy individuals (Huang et al., 2010). Demonstrated by Spanagel and colleagues (2005), the $Per2^{Brdm1}$ mutant mice showed significantly higher ethanol intake compared to their wild-type counterparts and such observation was accompanied by the reduced expression of glutamate transporter EAAT1 (Spanagel et al., 2005). Verified by a human alcohol drinking population, multiple SNPs on *PER2* were found to associate with alcohol intake in patients with alcohol dependence (Spanagel et al., 2005), confirming the genetic effects of circadian genes on alcohol addictions.

1.5 Sensitivity to reward in substance use

1.5.1 Brain responses in reward processing

Early research indicates electrical stimulation of the septal area monitor the rewarding value towards stimuli in rats (Olds and Milner, 1954). Since then, neural activity during reward processing has been investigated using a variety of neurophysiological techniques. It has been discovered that reward processing is regulated by the mesocorticolimbic dopaminergic pathway originating at the ventral tegmental area (VTA) (Koob, 1992). Dopaminergic neurons at the VTA project to the nucleus accumbens (NAc) and brain regions including the ventral pallidum, substantia nigra, hippocampus, amygdala and prefrontal cortex (Koob, 1992).

In human, brain activations during reward processing have been widely studied using the monetary incentive delay (MID) task. The MID task is a non-invasive, functional neuroimaging tasks that measures brain activations in response to monetary reward (See **Chapter 2 Section 2.3.5** for details). Brain activations measured by the MID task can be regarded as the sensitivity to reward. In young adults, activations of the medial prefrontal cortex (mPFC) and NAc are proportional to the expected amount of reward during the anticipation phase of the MID task, implicating these brain regions are sensitive to reward stimuli (Knutson and Peterson, 2005, Knutson et al., 2005). Since the maturation of frontal cortical networks is slower than that of the limbic regions during adolescence (Giedd et al., 1996a), the imbalance between the mesocortical and mesolimbic dopaminergic pathways may result in hypo-/ hypersensitivity to reward stimuli, therefore modifies reward seeking and risk taking behaviour (Spear, 2000, Forbes and Dahl, 2005).

1.5.2 Reward processing and substance use

Reward processing is one important component that mediates alcohol drinking, smoking and substance use in adolescents (Spear, 2000). Upon the exposure to rewarding stimuli such as alcohol or tobacco, the mesocorticolimbic dopaminergic pathway monitors the emotions

associated with reward stimuli (Spear, 2000). Subsequently the experience of reward can be reinforced and learned, resulting in repetitive reward seeking behaviour (Spear, 2000, Koob, 1992).

Previous literature suggested the brain activity in the VS, which consists of the NAc and the ventral part of the caudate nucleus and putamen, can be a valid predictor for reward processing and substance use. Reduced activations in the VS and larger risk taking bias have been observed in adolescents who consumed alcohol, tobacco and psychoactive drugs (Schneider et al., 2012). Similarly, reduced VS activations have been observed among patients with alcohol dependence and adolescent smokers (Beck et al., 2009, Peters et al., 2011). However, opposite findings have also been suggested, in which higher VS activations during the reward outcome phase of the MID task were found among patients with drug dependence (Bjork et al., 2008b). As illustrated by Wang et al. (2008), smokers who experienced nicotine deprivation demonstrated changes in blood flow in brain regions including the prefrontal cortex, orbitoprefrontal cortex, VS and thalamus. Brain activity during reward processing can be driven by genetic polymorphisms, as the changes in blood flow during nicotine deprivation was shown to associate with the SNPs on the dopamine D2 receptor (*DRD2*) and catechol-O-methyl transferase (*COMT*) (Wang et al., 2008).

1.6 Interactions between the stress, circadian & reward systems in substance use

1.6.1 Interactions between the circadian and stress systems

1.6.1.1 Disruptions of the circadian system alter the production of glucocorticoids

The functions of the circadian genes can be influenced by glucocorticoids and *vice versa*. Rodents with bilateral lesions of the SCN demonstrated impaired circadian activity and stress responses (Buijs et al., 1993). They had higher corticosterone levels compared to their intact counterparts at both baseline and after the exposure to relocation stress (Buijs et al., 1993).

Since the stress and circadian systems share the PVN as the relay centre for neurotransmission, it is possible that the PVN is involved in balancing the signals received from both systems.

Administration of AVP to the PVN was found to reduce corticosterone level in rats with bilateral SCN lesions (Kalsbeek et al., 1992). The actions of AVP during stress responses can be influenced by vasoactive intestinal polypeptide (VIP), as the injection of VIP in the rat PVN significantly increased plasma ACTH and corticosterone in a dose-dependent manner (Alexander and Sander, 1994). Applying VIP antagonist [Lys^{1}, $Pro^{2,5}$, $Arg^{3,4}$, Tyr^{6}-VIP] to the PVN was found to inhibit the stimulatory effect of VIP on the secretions of ACTH and corticosterone (Alexander and Sander, 1994), confirming that the production of glucocorticoids can be regulated by the neuropeptides in the SCN and the PVN (**Figure 1-5**).

Disruptions of circadian genes were also found to impact on the rhythmic production of glucocorticoids. Targeted *Per2* knockout and *Per2/Cry1* double knockouts could lead to arrhythmic secretion of corticosterone in mice (Yang et al., 2009, Oster et al., 2002). The suppression of CRY2 protein has been previously shown to decrease cortisol level in primates (Torres-Farfan et al., 2009), suggesting the expression level of CRY2 can be linked to the production of glucocorticoids.

1.6.1.2 Glucocorticoids alter the expressions of circadian genes

Using animal models, stressors including forced swimming and immobilisation were found to increase *Per1* expression in the CRH-containing neurons in the PVN of mice (Takahashi et al., 2001). Functional analyses indicated the presence of E-boxes and GRE sites in *Per1* and *Per2* (Yamamoto et al., 2005, Yan et al., 2008, So et al., 2007), suggesting that the transcriptional activities of *Per1* and *Per2* can be influenced by signals from the circadian and stress systems. Apart from the GRE motifs, transcription factors including Snail1 can regulate the cortisol-

induced *PER1* mRNA expression in human cell lines (Dong et al., 2011). Nevertheless, adrenalectomy was found to disrupt the rhythmic expression of Per2 protein in the central extended amygdala (CeA) (Lamont et al., 2005). The rhythmic expression of *Per2* mRNA in the CeA and the oval nucleus of the stria terminalis in adrenalectomised rats could be restored by corticosterone replacement (Segall et al., 2006).

GR are highly expressed in the SCN neighbouring brain regions including the BNST, amygdala and hippocampus but absent in the SCN (Amir et al., 2004, Lamont et al., 2005). The projection pathways from these limbic regions to the SCN can potentially adjust the circadian rhythm and the production of neuropeptides in the SCN. Proposed by Balsalobre and colleagues (2000), glucocorticoids can act as a source of entrainment that modifies the circadian cycle (Balsalobre et al., 2000). The absence of GR in the SCN has further strengthened Balsalobre *et al.*'s (2000) proposal of the indirect feedback mechanism.

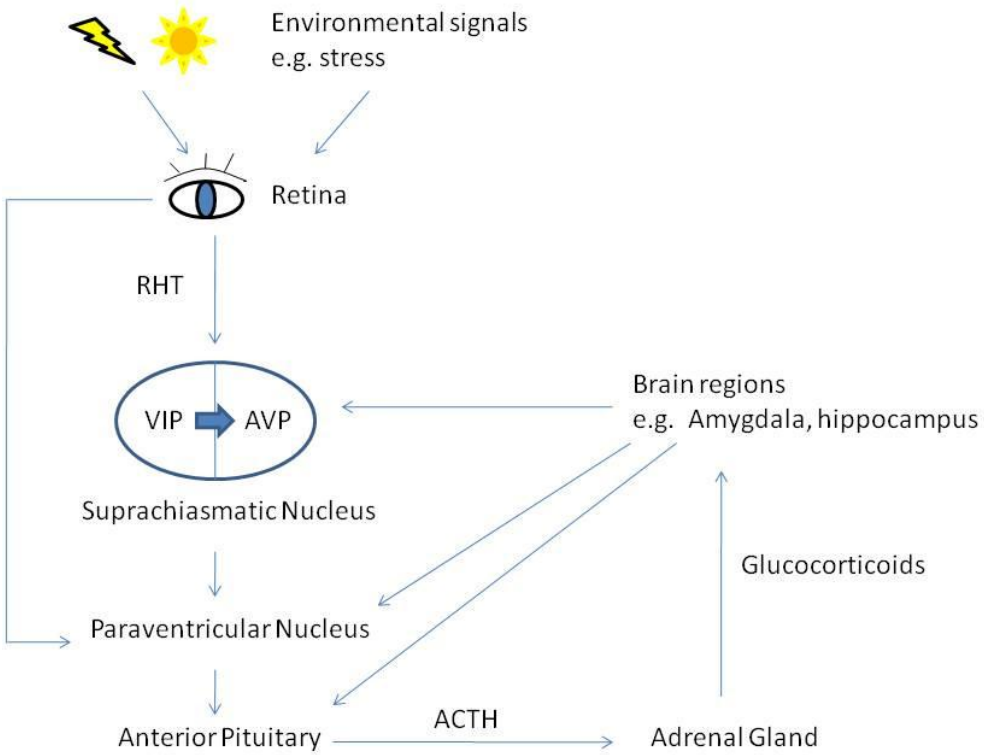


Figure 1-5 Schematic diagram of the interactions between the circadian system and the stress HPA axis. Entrainment signals such as light and stressful stimuli enter the retina. The signal is transmitted to the ventral part of the SCN via the retinohypothalamic tract (RHT). Neurons from the ventral SCN are projected to the dorsal SCN, where the dorsal SCN carries the signals to the paraventricular nucleus (PVN). The PVN contains a cocktail of genes and hormones that are involved in the regulations of the circadian rhythm and stress reactivity. Hormones such as the corticotrophin releasing hormone (CRH) are released into the anterior pituitary via the hypophyseal portal system in the hypothalamus. CRH further stimulates the release of adrenocorticotropic hormone (ACTH) from the anterior pituitary. ACTH is transported to the adrenal gland to stimulate the production of glucocorticoids. Glucocorticoids bind to the glucocorticoid receptors (GR) in target tissues. In mammals, GR is distributed in wide range of tissues and brain areas except the SCN. Signals from other brain areas such as amygdala and hippocampus are required to provide feedback into the SCN, the PVN or the pituitary for modifications of stress responses and circadian rhythm.

1.6.2 Reward processing and circadian rhythm

Reward processing and the dopaminergic reward system have been shown to interact with the circadian system in various animal and human studies. Activations in the mPFC and VS during reward outcome phase were higher among adolescents who delayed their sleeping time in weekends, in which sleep delay was used as an indicator of late chronotype and diurnal preference (Hasler et al., 2012). Components of the dopaminergic system, including the tyrosine hydroxylase (TH, rate-limiting enzyme involved in the synthesis of dopamine (DA)), dopamine receptors and transporter (DAT) can be regulated by the circadian system. For example, disruptions in the rhythmic expressions of DAT and TH have been observed in rats with SCN lesions (Sleipness et al., 2007). Interestingly, monoamine oxidase A (*MAOA*) that is responsible for metabolising DA has been shown to contain E-box binding sites at the promoter (Hampp et al., 2008), indicating the importance of circadian system in monitoring the activity of the dopaminergic reward system.

1.6.3 Influence of stress system on reward processing

Emerging evidence has suggested the role of stress in regulating the dopaminergic system and reward processing. Supplement of reward stimuli to rats exposed to physiological and psychological stress was found to restore locomotor activity and novelty seeking behaviour, suggesting reward might ameliorate the impairments introduced by stress (Dalm et al., 2012). After discovering the presence of GRs in the rat dopaminergic neurons (Harfstrand et al., 1986), it was found that the infusion of corticosterone upregulated the extracellular DA level in the rat NAc during dark phase (Piazza et al., 1996). Mice with Disrupted In Schizophrenia 1 (*Disc1*) knock-out also showed reduced extracellular DA and TH levels in the frontal cortex after the exposure to social isolation stress (Niwa et al., 2013). Such observation was mediated by the gene-environment interaction, since the mice with *Disc1* knock-out or experienced social isolation stress alone did not show any significant alterations in the TH and DA levels (Niwa et al., 2013).

Interestingly, the DNA methylation status of *TH* can be modulated by social isolation stress and glucocorticoids. In Niwa and colleagues (2013) study, the *Disc1* knock-out x social isolation stress mice demonstrated the highest *TH* DNA methylation level in the VTA compared to the mice with either *Disc1* knock-out or experienced social isolation stress (Niwa et al., 2013). The injection of GR antagonist RU38486 in the *Disc1* knock-out x social isolation stress mice not only reduced the DNA methylation level of *TH*, but also restored the locomotor activity and extracellular DA level to baseline, suggesting the role of glucocorticoids in monitoring the dopaminergic system in the brain (Niwa et al., 2013).

1.6.4 Interactions of the stress, circadian and reward systems on alcohol use and addictions

Previous work has illustrated the genetic effect of circadian genes on stress responses and alcohol use and addictions. Dong and colleagues (2011) demonstrated that *PER1* rs3027172 was associated with increased alcohol consumption in adolescents who experienced high level of psychosocial adversity. The risk allele of *PER1* rs3027172 was also associated with alcohol dependence in patients (Dong et al., 2011). Both ethanol injection and increase restraint stress were associated with the upregulation of *Per3* mRNA expression in the hippocampus (Wang et al., 2012), suggesting the role of *Period* gene in mediating stress-induced alcohol intake. Reductions in the *Vip* and *Avp* mRNA levels were observed in mice that received long-term ethanol treatment (Madeira et al., 1997). Combined stimulations of *Vip* and ethanol increased *Pomc* and *ß-endorphin* mRNA expressions more than the application of each substance alone (Poplawski et al., 2005), suggesting the involvement of *Vip* in promoting stress responses during alcohol intake. A risk haplotype (rs3823082 – rs688136) in the *VIP* gene was also shown to associate with alcohol use in a Finnish population (Kovanen et al., 2010).

In addition, interaction between the brain stress and reward systems is thought to influence the substance use behaviour. Suggested by Koob and colleagues, the development of addictions is

mediated by the negative and positive emotions associated with substance use (Koob and Le Moal, 2008, Koob, 2013). Such proposal can be supported by the action of dynorphin under the exposure to ethanol, as the activation of HPA axis enhances the dynorphin-κ-opioid receptors activities and alters the production of DA in the NAc (Wee and Koob, 2010, Shippenberg et al., 2007, Lindholm et al., 2007, Walker et al., 2012). Nevertheless, the core circadian genes also exhibit properties in reward processing, as the genetic polymorphisms of *PER2* rs2304672 and *CLOCK* rs1801260 were found to associate with brain activity during reward outcome and high reward dependence score in adolescents and adults (Forbes et al., 2012, Tsuchimine et al., 2013). Together the evidence demonstrates the role of stress, circadian and reward systems in regulating alcohol use and addictions. The crosstalk among the stress, circadian and reward systems on tobacco use is yet to be determined.

1.7 Thesis aims and objectives

The aim of this thesis is to identify genetic and epigenetic risk factors of alcohol and tobacco use in adolescence. Of particular interest, the impact of life stress and circadian genes on reward processing and substance use will be investigated. All analyses presented in this thesis are performed on the data collected from the IMAGEN Consortium, a multi-centre study of over 2000 adolescents of which the genetic, neuroimaging and behavioural information is available. The IMAGEN study and details of the assessment are described in Chapter 2.

This thesis consists of four empirical chapters:

Chapter 3 identifies the risk factors associated with alcohol and tobacco use in adolescents at age 14 and 16. Based on previous literature and research from the IMAGEN consortium, the consequences of negative life events and the VS activations during reward processing on substance use will be investigated.

Chapter 4 investigates the associations among the DNA methylation of *PER1*, negative life events, reward processing and substance use in adolescents at age 14. The presence of CpG sites at the GRE site and E-Box of *PER1* may indicate potential epigenetic modifications on these transcription regulatory motifs. The putative functions of the selected CpG sites on *PER1* will be explored.

Chapter 5 aims to identify the genetic influence of the stress and circadian systems on reward processing and substance use in adolescents at age 14 and age 16. Using an unbiased gene selection approach, the additive genetic effects of stress and circadian genes on the VS activations during reward anticipation and substance use will be investigated. The genetic influence of the stress and circadian systems at single SNP and polygenic levels will be discussed.

Chapter 6 examines the genetic risk of stress genes on drunkenness in adolescents at age 16. The linkage disequilibrium (LD) structure and functions of the top hit *FYN* rs1409837 will be examined. In the second part of this chapter, the role of *FYN* rs1409837 on negative life events and the amygdala activations during angry face processing will be studied. The potential function of *FYN* rs140837 will be discussed.

In the General Discussion (Chapter 7) I summarised the findings presented in this thesis and highlighted their significance in substance addictions. Limitations of the studies and implications for future research will be presented in this chapter.

Chapter 2

Materials and Methods

This chapter describes the research population, behavioural, neuroimaging and genetic information available in the IMAGEN Study. All analyses presented in this thesis were carried out using the data collected by the IMAGEN Study.

2.1 The IMAGEN Study

2.1.1 The IMAGEN Consortium

IMAGEN is the first multi-centre longitudinal study identifying the genetic and neurobiological basis of reinforcement-related behaviour in adolescents (Schumann et al., 2010). Since December 2007, the IMAGEN Consortium has recruited over 2000 14-year-old adolescents and their parents from 8 cities in Ireland, France, Germany and the UK. Neuroimaging (structural and functional MRI) and behavioural assessments were conducted in the adolescents. Whole blood sample also were collected for genetic and gene expression analyses. Follow-up interviews, online behavioural and clinical assessment were carried out in the adolescents at age 16. The IMAGEN Consortium receives research funding from the European Community's Sixth Framework Programme (LSHM-CT-2007-037286).

Figure 2-1 The IMAGEN Consortium. Over 2000 adolescents were recruited and assessed in eight recruitment centres in London, Nottingham, Dublin, Paris, Berlin, Hamburg, Mannheim and Dresden.

2.1.2 Ethical approval

Ethical approval was obtained from the local ethics committee at each study site (London: Institute of Psychiatry; Nottingham: University of Nottingham; Dublin: Trinity College; Paris: National Institute of Health and medical Research; Berlin: Charité University; Hamburg: University Medical Center Hamburg-Eppendorf; Mannheim: Central Institute of Mental Health; Dresden: Technical University Dresden). A central multi-disciplinary ethic group was established within the IMAGEN Consortium to monitor issues related to consent, confidentiality, data protection, and to resolve issues related to incidental findings that might arise from the genetic, neuroimaging and environmental measurements.

2.1.3 Participants and recruitment

One of the aims of the IMAGEN study is to assess the extent to which the genetic and environmental factors influence behaviours. Therefore, the recruitment procedures were designed and standardised to minimise the ethnic heterogeneity through preferential recruitment

of adolescents with European ethnicity. Researchers visited schools to explain the IMAGEN project and to gain permission to recruit from the schools. After receiving the consent, the IMAGEN team visited the schools again and gave presentation about the project. Adolescents were given information packs to take home and show their parents. Adolescents were asked to send a 'statement of interest' slip to the IMAGEN researchers if they were interested in taking part. Then the IMAGEN team contacted the parent or guardian of the adolescent to discuss the project and answer any queries. Upon the receipt of the informed consent from the adolescents and their parents, appointment was arranged for assessment at the local study centre.

The full IMAGEN sample consists of 2248 adolescents (age: 14.55 ± 0.45), of which 51.4% were females and 87.3% were Caucasian. N = 1669 adolescents responded to the follow-up assessment (age: 16.47 ± 1.03), of which 51.2% were females.

2.1.4 Exclusion criteria for recruitment

Participants were excluded prior to assessments if they met any of the following criteria: **a)** born prematurely; **b)** had contraindications for the magnetic resonance imaging, such as braces or other metal implants; **c)** had experienced neurological problems including head trauma or epilepsy; **d)** received medical treatment for disorders related to the central nervous system; and **e) w**ere not able to attend a full assessment day at the local study centres.

2.2 Behavioural assessment

2.2.1 Home assessment

Adolescents completed a 2 hour home assessment through a web-based coordination system *Psytools*, a software that was developed for the purpose of multi-site, multi-lingual assessments (Delosis, London, UK). Participants were provided with instructions for the home assessment, including a unique identification code and an internet link to download the psychometric battery in a computerised format. Context checks were administered at the beginning of each task to

identify whether the data was in a good quality and whether the assessment was undertaken in a confidential environment. At age 14, adolescents were asked to repeat the *Psytools* tasks at the institute assessment if necessary. Follow-up home assessment using *Psytools* was also conducted in adolescents at age 16.

2.2.2 Institute assessment

During the institute assessment, participants completed cognitive, neuropsychological, behavioural and neuroimaging assessment at the study centres. They were asked to give a sample of their blood. The parents or guardians of the adolescents were asked to complete a computerised assessment regarding adolescents' personality and family environment. The institute assessment took approximately eight hours to complete. All assessments were carried out by researchers that were trained by psychologists on a regular basis.

2.2.3 Life Events Questionnaire

2.2.3.1 Background

The Life Events Questionnaire (LEQ) is the modified version of Stressful Life Events Questionnaire developed by (Newcomb et al., 1981) to characterise life events that might introduce stress to adolescents. It contains 39 life events that are allocated into seven dimensions of life stress: Family/Parents, Accident/Illness, Sexuality, Autonomy, Deviance, Relocation and Distress. Individuals were asked to rate their feelings towards each life event on a 5-point rating scale indicating 'very happy', 'happy', 'neutral', 'unhappy' and 'very unhappy'. Dichotomous response on whether each event had occurred in lifetime and in the past year was also recorded. (See Appendix 1: Life Events Questionnaire). The IMAGEN adolescents completed the LEQ assessment at both age 14 and age 16. In the follow-up assessment at age 16, the LEQ was modified such that adolescents were required to report the age during which the life events occurred.

2.2.3.2 Negative life events

Quality control

To ensure the participants had read and understood the questionnaire, only individuals who gave at least two types of feeling ratings to the events were included in the dataset. After the quality control procedure, the LEQ dataset contains N = 2140 adolescents at age 14 and N = 1305 adolescents at age 16.

Defining negative life events

Clustering life events into seven dimensions of stress can be problematic. As stated by the authors, these seven dimensions of stress only explained about 44% variance of their data (Newcomb et al., 1981). It is possible that not all 39 life events listed on the LEQ (Newcomb et al., 1981) are considered as stressful by the adolescents. The ratings were compared among adolescents who had and had not experienced these life events in lifetime. Of all the 39 life events, 20 life events were consistently reported as 'unhappy' and 'very unhappy' by the adolescents. Hence these events were classified as 'negative life events'. Events that were consistently rated as 'happy' and 'very happy' in both populations were classified as 'positive life events'. The classification of life events were performed by Dr. Anna Cattrell at the Institute of Psychiatry. (See **Table 2-1** and **Table 2-2** for negative life events).

Both the frequency and rating towards negative life events were used to quantify adolescents' life stress in this thesis. The frequency of negative life events was measured by adding the number of negative life events that the adolescents had reported experiencing in lifetime. The rating towards negative life events was measured by adding the rating of individual negative life events that the adolescents had reported experiencing in lifetime. The feeling towards all experienced negative life events might indicate the level of stress perceived by the adolescents,

whilst the frequency of negative life events might provide a relatively unbiased way to quantify life stress.

Table 2-1 List of negative life events from the LEQ.

Item no.	***Negative Life Events (20 items)***
1	Parents Divorced
2	Family Accident/Illness
4	In trouble with the law
5	Stole something valuable
8	Death in the family
9	Face broke out with pimples
10	Brother or Sister moved out
16	Thought about suicide
17	Changed Schools
19	Got in trouble at school
20	Got or gave an STD
22	Family had money problems
24	Parents argued or fought
25	Ran away from home
27	Got poor grades in school
30	Broke up with a boyfriend/ girlfriend
31	Family moved
36	Gained a lot of weight
37	Serious accident/illness
39	Parent abused alcohol

Table 2-2 Descriptive statistics of the negative life events.

	Age 14 (N = 2140)		*Age 16 (N = 1305)*	
	Mean	***S.D.***	***Mean***	***S.D.***
Rating	-7.14	4.07	-5.84	3.71
Lifetime frequency	6.65	2.86	5.81	2.64

2.2.4 The European School Survey Project on Alcohol and Other Drugs

2.2.4.1 Background

The European School Survey Project on Alcohol and Other Drugs questionnaire (ESPAD) was designed by the ESPAD consortium to assess substance use behaviour in 15-16 years old teenagers across the Europe (Hibell et al., 1997). The data collection was carried out in every 4 years since 1995. The fifth data collection was carried out across 36 countries in Europe during 2011 (http://www.espad.org/). The ESPAD questionnaire is a self-report questionnaire that provides information about the age of onset and the frequencies of alcohol, tobacco and psychoactive drug use. Specific questions on the expectations of alcohol use and bingeing are asked for those who reported alcohol use. It also contains questions related to substance use-related problems, for example, school performance and truancy. The IMAGEN adolescents completed the ESPAD questionnaire at both age 14 and age 16. (See Appendix 2: ESPAD).

The reliability of ESPAD questionnaire has been tested to ensure data collected in different countries are comparable. Responses related to the recent alcohol use reaches a correlation of over 0.75 across gender in the data collected in 1995 and 1999, whereas the responses related to lifetime drinking experience reaches a correlation of over 0.45 (Lintonen et al., 2004). The test-retest reliability for the data collected in 2004 on drunkenness and the use of cigarettes, cannabis, cocaine and ecstasy among the Italian teenagers has been regarded as 'excellent' (Molinaro et al., 2012). The ESPAD questionnaire was selected to measure the alcohol and tobacco use in the IMAGEN participants due to its reliability and popularity among the research institutes across the Europe. Using the ESPAD questionnaire to investigate substance use in

adolescents will enable comparison and validation of data collected in other European populations in the future.

2.2.4.2 Data quality control

The data collection and quality control of the ESPAD questionnaire were conducted by researchers and assistants who worked with the IMAGEN Study during 2007-2011. Before the individuals started answering the questionnaire, four context questions were asked in order to determine the reliability and quality of responses (see Appendix 4: Quality control questions for ESPAD and PDS). Data of individuals was removed if they were in a hurry (i.e. responded 'yes' to ts_2) and/or indicated 'another person is sitting next to me and watching my responses' in ts_4. To further ensure the reliability of responses of drug use, individuals were asked whether they heard of a fake drug called 'relevin' (question 22) and the number of occasions they had used 'relevin'. Any individuals who had heard of and/ or had taken 'relevin' were excluded. Questionnaires that were not checked or mislabelled by the researchers were excluded from any analyses. After applying the quality control procedure, the ESPAD dataset contains N = 2113 adolescents at age 14 and N = 1491 adolescents at age 16.

2.2.4.3 Alcohol use in adolescents at age 14 and age 16

The number of occasions of alcohol intake in the past 12 months was selected as an indicator of whether a teenager had started drinking alcohol. This measurement can be more reliable than defining drinkers using lifetime alcohol intake in question 8a (Lintonen et al., 2004). It is also a more stable indicator for regular alcohol use compared to the responses from 'drinking in the past 30 days' (Question 8c) since it reduces chances of defining drinkers based on their recent drinking behaviour. Individuals were asked to respond on a 7-point scale indicating the number of alcohol drinking occasions in the past 12 months. An arbitrary cut-off was set such that individuals who responded '0' (i.e. never had any alcohol before) and '1' (i.e. had 1-2 drinking occasions) were regarded as 'non-drinkers', since they were likely to have taken a few sip of

alcohol drink during those occasions; therefore they were classified as 'non-drinkers'. Individuals who responded '2' or above (i.e. had at least 3 alcohol drinking occasions) were classified as 'drinkers' (See **Table 2-3**).

2.2.4.4 Tobacco use in adolescents at age 14 and age 16

The number of occasions of smoking cigarettes in lifetime was selected to indicate whether a teenager had been smoking tobacco. Individuals were classified as 'smokers' or 'non-smokers' according to their responses on a 7-point scale indicating the number of cigarette smoking occasions in lifetime. Similar to the drinker group classification, individuals who responded '0' (i.e. never smoke cigarette before) and '1' (i.e. had 1-2 cigarette smoking occasions) were defined as 'non-smokers' in an arbitrary way. Individuals who responded '2' or above (i.e. had at least 3 cigarette smoking occasions) were classified as 'smokers' (See **Table 2-3**).

2.2.4.5 Alcohol misuse in adolescents at age 16

Since increased alcohol consumption was observed in the adolescents over age, the risk of alcohol misuse in adolescents at age 16 was investigated. Two phenotypes – lifetime drunkenness and bingeing were selected to represent alcohol misuse and intoxications in the adolescents (See **Table 2-3**).

Lifetime drunkenness

Individuals were asked to provide the number of occasions they were drunk from drinking alcoholic beverages in lifetime (Question 19a on the ESPAD) on a six-point scale ranging from zero (never) to six (40 times or more). Individuals who responded zero to the question were allocated to the group ‘never drunk in lifetime’ and individuals who reported at least one occasion of getting drunk were allocated to the group ‘drunk in lifetime’.

Lifetime bingeing

Alcohol bingeing was defined as having five or more alcoholic beverages in one drinking occasion (Wechsler et al., 1994). Adolescents were required to report the number of occasions they had five or more alcoholic drinks in a row in lifetime (Question 17a on the ESPAD) on a five-point scale ranging from zero (never) to five (10 times or more). Individuals who responded zero to the question were allocated to the group 'never binge in lifetime' and individuals who reported at least one occasion of binge drinking were allocated to the group 'binge in lifetime'.

Table 2-3 Alcohol and tobacco use in adolescents.

		Age 14 (N = 2113)	***Age 16 (N = 1491)***
Alcohol use	Non-drinker	1362	280
	Drinker	751	1067
Tobacco use	Non-smoker	1765	939
	Smoker	348	524
Lifetime drunkenness	Never		452
	Ever		814
Lifetime bingeing	Never		418
	Ever		848

2.2.5 Puberty Development Scale

The Puberty Development Scale (PDS) was designed to assess the pubertal status of adolescents at age 14 (Petersen et al., 1988) (See Appendix 3: Puberty Development Scale (PDS)). The PDS contains eight self-report measurement of physical development. Participants answered

questions regarding their growth in stature and pubic hair, as well as menarche in females and voice changes in males. Separate PDS forms were given to male and female adolescents. Five categories of puberty were derived from these measurement **1)** pre-pubertal; **2)** early pubertal; **3)** mid-pubertal; **4)** late pubertal; and **5)** post-pubertal (see **Table 2-4**). High correlations between the adolescent self-rating and physician's rating on puberty development have been observed (Males: *r*=.77 - .84, Females: *r*=.88 - .91; Dorn et al., 1990).

Table 2-4 Puberty development stage of adolescents at age 14.

	Males	***Females***	***Total***
Mean	2.60	2.64	2.62
S.D.	0.53	0.36	0.45
N	1060	1118	2178

2.3 Functional neuroimaging assessment

Neuroimaging assessment was conducted in the IMAGEN adolescents at age 14. This thesis uses data from two functional magnetic resonance imaging (fMRI) tasks, namely the modified monetary incentive delay (MID) task and the face task for analyses.

2.3.1 Functional magnetic resonance imaging

Non-invasive approaches such as functional magnetic resonance imaging (fMRI) are frequently used by neuroscientists to investigate task-dependent activations in human brain. The blood oxygen level-dependent (BOLD) response, which measures the proportion of oxygen-rich and oxygen-poor blood at a certain time point, is commonly used to infer brain activations in the fMRI. When strong magnetic field is applied to the brain, hydrogen atoms from the water spin and align to the magnetic field. The differential magnetic property in the oxygen-rich and oxygen-poor blood allows BOLD response to be calculated.

Blood flow to neurons in specific brain region increases during activation. The oxygen-rich blood displaces oxygen-poor blood around two seconds following the activation of the brain region; with a peak in blood flow that appears 4-6 seconds thereafter. Due to the strong correlation between synaptic activity and BOLD response, the BOLD response can provide a good estimation of neurobiological functions in the brain (Logothetis et al., 2001). Therefore, fMRI has been used to measure task-dependent brain activations in the IMAGEN adolescents.

2.3.2 Image acquisition

The structural and functional magnetic resonance (MR) imaging of the fMRI tasks was performed on 3 Tesla MRI scanners (Siemens, Philips, General Electric & Bruker). High-resolution T1-weighted anatomical images were acquired using the 3D Magnetisation Prepared Rapid Acquisition Gradient Echo (MPRAGE) sequence (Repetition time (TR) = 2300 msec; Echo time (TE) = 2.8 msec; flip angle = 9°; resolution: $1 \times 1 \times 1$ mm^3). The functional T2*-weighted images were acquired using the Gradient-Echo Echo-Planar-Imaging (GE-EPI) sequences (field of view: 22 cm; voxel size: 3.3×3.3 mm^2; slice thickness of 2.4 mm; TE = 30 msec and TR = 2200 msec; flip angle: 75°).

2.3.3 Functional MR image preprocessing, quality control and probabilistic mapping

The acquired EPI images were corrected for motion, realigned and resliced to the first volume. For each participant, the EPI images were coregistered with the T1-weighted anatomical images, rendered and spatially smoothed using a 5mm full-width half maximum Gaussian isotropic kernel. By selecting the haemodynamic response function as the basic function, contrast images of each participant were generated and transformed into a customed template space. Images with distorted magnetic field and structural abnormalities were excluded from any analyses. Estimated movement parameters were added to the design matrix in the form of 18 additional columns (3 translations, 3 rotations, 3 quadratic and 3 cubic translations, and each 3 translations with a shift of ±1 TR). All images were preprocessed and analysed using the

Statistical Parametric Mapping (SPM8, Wellcome Trust Centre for Neuroimaging, U.K.) at the NeuroSpin, France.

2.3.4 Exclusion criteria for the fMRI analyses

In addition to the exclusion criteria presented in **Section 2.1.4**, neuroimaging data of participants were excluded after the preprocessing if they met any of the following criteria: **a)** Reported that they had problems reading the instructions during the functional MRI task investigated or reported falling asleep during the functional MRI assessment; **b)** Moved more than 3 mm or 3 degrees in any direction during scanning; **c)** Showed anatomical abnormalities; and **d)** Showed outlying activation values across voxels in a particular fMRI contrasts under investigation

2.3.5 The Modified Monetary Incentive Delay (MID) fMRI task

2.3.5.1 Task feature

The MID fMRI task measured the blood oxygen level dependent (BOLD) response during reward processing. The task began with a cue that appeared on either the left or right side of the screen for 250 milliseconds (msec). The cue indicated the amount of gain (large win, small win or no win) that the participants could receive on completion of the trial. Participants were required to hit the target that appeared at the same location as the cue after an anticipation phase of 4000 msec. Then a feedback message appeared on the screen for 1450ms to inform the participants whether they had successfully completed the trial and the amount of gain they had received. Using a tracking algorithm, task difficulty (i.e. target duration varied between 250 and 400 msec) was individually adjusted such that each participant successfully responded on about 66% of trials. The MID task consisted of 144 trials. Both the trials and the position of which the cues displayed were randomised. In total 300 volumes of scanned images were acquired in each participant throughout the MID task.

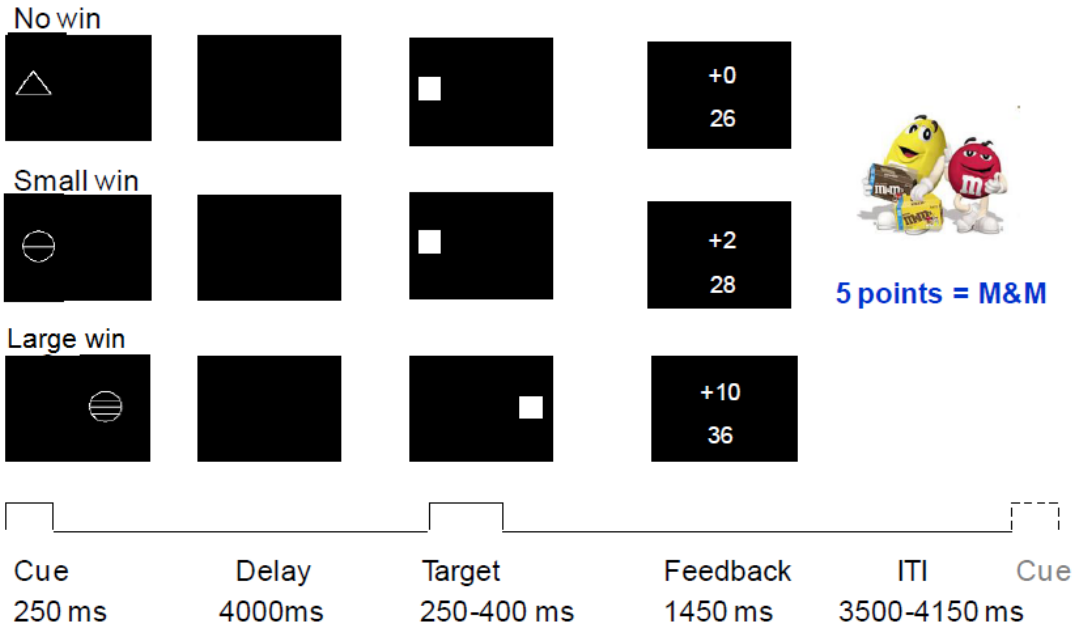


Figure 2-2 The modified monetary incentive delay (MID) task. Participants were asked to hit the target on the screen after one of the three cues was shown. The cues indicated the amount of gain (Triangle (top): no win, Circle with three lines (middle): large win, Circle with one line (bottom): small win) that the participants could receive on completion of the trial. A feedback message was then displayed to indicate the total amount of reward and the amount of reward the participants received after the trial.

2.3.5.2 Ventral striatum BOLD responses during reward anticipation

Reward anticipation 'Large win – No win' contrast

The 'Large win – No win' contrast was created to represent the net brain activation during the large win reward anticipation phase. The brain activation of this contrast was calculated by subtracting the BOLD responses measured in the no win trials from the BOLD responses measured in the large win trials. Only the trials that the individuals hit the targets were included to create the contrast. The brain activation data of 1469 individuals was available at the reward anticipation 'Large win – No win' contrast.

Ventral striatum BOLD responses

The selection of region-of-interest (ROI) was based on Knutson et al's (2005) study on examining neural activations in a monetary incentive delay task in young adults (Knutson et al., 2005). The results showed activations in the bilateral NAc during the anticipation phase were proportional to the amount of gain that the subject could receive after they successfully completed the task. Due to the small volume of NAc and its proximity to the ventral part of the striatum, it would be reasonable to extract BOLD responses within the bilateral ventral striatum (VS) to study reward anticipation. The ROIs for the VS was defined by O'Doherty et al. (2004) at Montreal Neurological Institute (MNI) coordinates x, y, z = ± 15, 9, -9 with a sphere radius of 9 mm (O'Doherty et al., 2004). The average BOLD responses across all voxels within the VS were extracted using the Marsbar toolbox (http://marsbar.sourceforge.net). Additional quality control was performed on the VS BOLD response data to remove outliers (data outside ± 3 S.D.) in the SPSS (Version 20, IBM).

2.3.6 Face task

2.3.6.1 Task feature

The face task involved passive viewing of black-and-white video clips showing angry and emotionally neutral motion faces as well as the non-biological motion control stimuli (Grosbras and Paus, 2006). The non-biological control stimuli were adapted from Beauchamp et al's (2003) study, consisting of concentric circles of various contrasts, expanding and contracting at various speeds that were roughly matched with the size and motion characteristics of the face stimuli (Beauchamp et al., 2003). All video clips were 2-5 seconds long and they were arranged into 19 blocks (18 seconds each). Each block consisted of 7-8 video clips. Five blocks of each biological-motion condition (angry and emotionally neutral faces), and nine blocks of the control condition (circles) were intermixed and presented to the participants through goggles (Nordic Neurolabs, Bergen, Norway) inside the scanner. Participants were instructed to watch

the video clips carefully prior to scanning. In total 444 volumes of scanned images per participant were acquired throughout the face task.

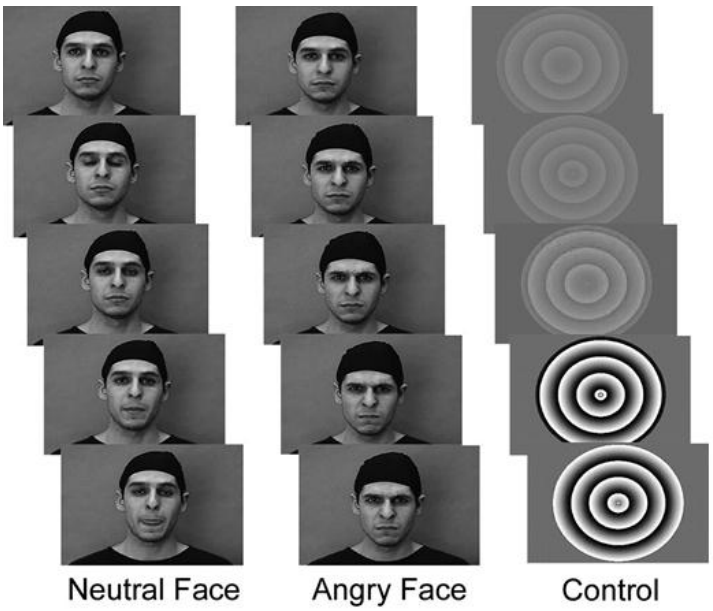


Figure 2-3 Face task motion stimuli. Control stimuli, faces with angry and neutral emotions were organised into 2-5 seconds video clips in the face task.

2.3.6.2 Amygdala BOLD responses in 'Angry face – Control' contrast

Amygdala is one of the most activated limbic regions during emotional processing and response to stressful situations (Phan et al., 2002, Whalen et al., 1998). Emotional processing has been frequently measured by the face task, in which the presentation of fearful or angry faces has been found to increase the amygdala activations (Whalen et al., 1998). To investigate individual's brain activation during emotional processing, the contrast map of 'Angry faces *versus* control stimuli' was selected for analyses in this thesis. The brain activation data of 1635 adolescents was available in the 'Angry face – Control' contrast.

The ROIs for the left and right amygdala were defined by the human Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) and extracted according to the 'angry faces *versus* control' contrast activations using the Marsbar toolbox (http://marsbar.sourceforge.net). The volumes of the resulting amygdala ROI were 1760 mm^3

(left amygdala) and 1984 mm^3 (right amygdala). The average BOLD responses across all voxels within the amygdala ROIs were extracted using the Marsbar toolbox. Additional quality control was performed on the amygdala BOLD responses data to remove outliers (data outside ± 3 S.D.) on the SPSS (Version 20, IBM).

2.4 Genetic data

2.4.1 Biological sample

Ten ml of whole blood was collected in the BD Vacutainer ethylenediaminetetraacetic acid (EDTA) tubes (Becton, Dickinson and Company) from the IMAGEN adolescents at age 14 for deoxyribonucleic acid (DNA) extraction. The whole blood DNA samples were extracted using the Gentra Puregene Blood Kit (QIAGEN) and genotyped at the Centre National de Génotypage, France. For each individual, 200 ng whole blood DNA was hybridised to the BeadChips for genotyping.

2.4.2 Whole-genome genotyping microarray

2.4.2.1 Background

The SNP coverage of whole-genome genotyping microarrays relies on the linkage disequilibrium (LD) structure in human genome. SNPs in that are in high LD (i.e. highly correlated) share similar allele frequencies and they are likely to transmit together through generations. Hence, the SNPs on the genotyping microarrays are designed to 'tag' SNPs that are within the same loci. The tagging nature of SNPs allows researchers to interrogate SNP loci that are associated with phenotypes of disorders and to compensate the reduction in statistical power due to multiple testing and sample size. Using the genetic database of the HapMap project as reference, over 620000 tag SNPs are designed and typed in both Illumina© BeadChips.

Whole-genome genotyping was conducted in three waves. In the first wave, N = 705 individuals were genotyped using the Illumina© Human610Quad BeadChips (Illumina Inc., San

Diego), of which the genetic information on the X chromosome was available. The second and the third waves consisted of N = 993 and N = 393 individuals respectively. Individuals from the second and the third waves were genotyped on the Illumina© Human660w-Quad BeadChips (Illumina Inc., San Diego), of which the genetic information on copy number variations but not the X chromosome was available.

2.4.2.2 Mechanisms for detecting genotypes by probes

The whole-genome genotyping on the Illumina© BeadChips is a multiplexed, single-base primer extension process. The 3' end of single-stranded extension primers (50bp) are designed to anneal adjacent to the SNP site, allowing the incorporation of biotin- or dinitrophenyl-labelled nucleotide at the SNP site. The cytosine and guanine nucleotides are biotin-labelled (generates green signal) and the adenine and thymine are dinitrophenyl-labelled (generates red signal). Then signal intensity of the probes is transformed and normalised in the GenomeStudio software (Illumina Inc., San Diego). Comparing the signal-to-noise intensity of probes will allow researchers to calculate the SNP statistics (e.g. call rate and accuracy) and determine the genotypes of SNPs in each individual.

2.4.2.3 Quality control of the genetic data

Quality control for individuals

The quality control of genetic data was performed by Dr. Anbarasu Lourdusamy and Dr. Tianye Jia at the Institute of Psychiatry, London. They first removed individuals containing ambiguous sex information and excessive missing genotypes (defined by failure rate of over 2%). Then the population stratification were identified using the Structure software (Pritchard et al., 2000). The genome-wide identity-by-descent similarity were used to estimate the relatedness of individuals (Purcell et al., 2007). Closely related individuals were defined as having an outlying heterozygosity larger than 3 standard deviations from the sample mean and with identity-by-descent >0.1875 were removed.

Quality control for SNPs

Next the quality control was performed at the SNP level. SNPs with call rates <95% and minor allele frequency (MAF) <1% were removed. SNPs deviated from the Hardy-Weinberg Equilibrium (p-value $\leq 1 \times 10^{-5}$) were removed from subsequent analyses. Four HapMap populations were used as reference groups to remove individuals with divergent ancestry. The resulting genome-wide genetic data contained 506932 SNPs with an overall genotyping rate >0.99 from N = 1982 adolescents.

2.4.3 SNP imputation

Since only a subset of SNPs is been genotyped in the Illumina© BeadChips, SNP imputation can be performed to maximise SNP coverage and to identify any hidden associations with disorders and traits. Reference populations that contain genetic information at high resolution are required for imputation. Based on the assumption that the reference population shares the same LD structure as the IMAGEN population, the SNP information from the reference population can be used to impute alleles of the untyped SNPs.

Imputation of the IMAGEN genetic data was performed by the IMAGEN biostatisticians led by Dr. Vincent Frouin at the NeuroSpin, France. The genetic database of the 1000 Genomes Project was used as the reference population for SNP imputation. Using the next-generation sequencing technique, the 1000 Genomes Project provides high resolution genetic information of over 14 million SNPs with MAF above 1% in 1000 unrelated individuals across seven populations (Genomes Project et al., 2010). The genotypes of the imputed SNPs are displayed in major allele dosage ranging from 0 to 2, indicating the predicted number of copies of major alleles in the SNPs.

2.5 Whole-genome gene expression data

2.5.1 Biological sample and processing

Ten ml of whole blood was collected in the BD Vacutainer PAXgene tubes (Becton, Dickinson and Company) from N = 683 IMAGEN adolescents at age 14 for total ribonucleic acid (RNA) extraction. Total RNA was extracted using the PAXgene Blood RNA Kit (QIAGEN) at the BIOGEM Laboratory, University of California San Diego (UCSD).

2.5.2 Mechanisms and Procedures

2.5.2.1 Preparation of complementary ribonucleic acid (cRNA) library and hybridisation

Whole-genome gene expression of the IMAGEN adolescents was examined using the Illumina© HumanHT-12 v4 Expression BeadChips (Illumina Inc., San Diego). In order to prepare cRNA for hybridisation, about 50 - 100ng total RNA from each individual underwent *in vitro* transcription amplification using the Illumina© TotalPrep™ RNA Amplification kit (Ambion, Austin). *In vitro* transcription amplification is a multiple-step procedure that incorporates biotin-labelled nucleotides to the cRNA. The cRNA concentration was quantified using the Qubit® 2.0 Fluorometer (Invitrogen). To ensure the quality of the labelled cRNA, fragment size distributions of the labelled-cRNA was determined using the Eukaryotic mRNA Assay with smear analysis on the Bioanalyzer system (Agilent Technologies, Santa Clara).

For each individual, about 750 ng cRNA was hybridised on the Illumina© HumanHT-12 v4 Expression BeadChip (Illumina Inc., San Diego). The Illumina© Expression BeadChip contains 47231 probes to interrogate transcripts and known splice variants across the human transcriptome that are documented in the National Center for Biotechnology Information Reference Sequence Release 38 (NCBI RefSeq, November 2009) and the UniGene.

Each probe consists of a 50bp gene-specific sequence and a 29bp 'address' sequence that are immobilised on the streptavidin-coated beads on the BeadChips for hybridisation. Followed by washing, blocking and streptavadin-Cy3 staining, the intensity of the fluorescence signal emitted by Cy3 can be quantified as gene expression level. Both the labelling and hybridisation were carried out in the BIOGEM Laboratory, UCSD.

2.5.2.2 Quality control

The quality control of the whole-genome gene expression data was performed according to the mloess method developed by Dr. Roman Sasik from the BIOGEM Laboratory, UCSD (Sasik et al., 2002). Any batch effects arise from the RNA extraction and cRNA labelling were corrected during the quality control procedure. Prior to any statistical analyses, the sequences of probes were blast on the UCSC genome browser to ensure binding specificity (i.e. does not contain any putative SNPs in the probe sequence and displays 100% binding to the gene of interest). The gene expression value was $\log_2$ transformed and outliers beyond ±3 S.D. were also removed.

Chapter 3

Negative life events and reward sensitivity as risk factors for substance use

3.1 Introduction

Alcohol and tobacco are the most common forms of substance use in European populations (Kendler et al., 2012). The consumption of these substances in early adolescence can increase risks of developing addictions and other neuropsychiatric disorders in late adolescence and early adulthood (Kessler et al., 2005, Chambers et al., 2003). This chapter explores the extent to which life stress and reward sensitivity are related to alcohol and tobacco use in the IMAGEN adolescents.

3.1.1 Stress-induced substance use

Adolescence is a developmental period during which individuals experience numerous behavioural and physiological changes (Spear, 2000). Overcoming unexpected changes can be stressful, particularly in the case when adolescents fail to cope with these events. The impact of stressful life events during childhood and adolescence on substance use has been investigated. Studies have demonstrated that adverse childhood events such as a violent family environment, parental separation or divorce, are strongly associated with alcohol and tobacco use in adolescents and adults (Dube et al., 2006, Anda et al., 1999). The number of adverse childhood events was found to be proportional to the self-rated severity of alcohol dependence (Pilowsky et al., 2009). The authors found that individuals who had experienced two adverse childhood events were twice as likely to consider themselves as alcoholics compared to individuals who reported only one event (Pilowsky et al., 2009). Increased frequencies of alcohol and tobacco

use were also observed in adolescents who reported more negative life events (Blomeyer et al., 2008, Booker et al., 2008). Results from the Dunedin Longitudinal Study showed that the perceived level of stress due to life events was significantly associated with alcohol and drug use and dependence in 18 - 32 years old adults (Meier et al., 2013). Together these studies suggest that life stress plays an important role in both the development and maintenance of substance use behaviours throughout life.

3.1.2 Reward sensitivity and substance use

The tendency to experiment and develop repeated use of substance in adolescents can also be driven by their sensitivity to reward stimuli (Koob, 1992, Spear, 2000). Individuals' sensitivity to reward has been explored using functional neuroimaging tasks that model different phases of reward processing. One of the most widely used paradigms is the Monetary Incentive Delay (MID) task (Knutson et al., 2005, Knutson and Peterson, 2005). The MID task showed that the BOLD responses of the ventral striatum (VS) and medial prefrontal cortex (mPFC) were proportional to the amount of reward that the participants were exposed during the task (Knutson et al., 2005). In the IMAGEN study, brain activations during reward processing in adolescents were measured by the MID task (See **Chapter 2 Section 2.3.5** for details). The IMAGEN adolescents displayed widespread activations including the striatum, frontal cortex, parietal and occipital lobes in the reward anticipation ('Large win - No win') contrast, with the peak of activation located at the VS (Nees et al., 2012, Nees et al., 2013, Stacey et al., 2012, Peters et al., 2011). Consistent with the previous literature, research from the IMAGEN study has confirmed the involvement of the VS activations during reward anticipation in adolescents.

Two hypotheses on reward processing during adolescence have been emerged from the literature and both hypotheses put emphasis on the activations of the VS. The reward deficiency hypothesis proposes that the reduced VS activation during reward processing will result in the hypoactivation of the reward system, thereby increasing risk-taking behaviour and promoting

substance use in adolescents (Blum et al., 1996); whereas the opposite hypothesis suggests that hyperactivation of the VS and reward system in adolescents will increase risk-taking and reward-seeking behaviour (Galvan et al., 2006, Galvan, 2010). Both hyper- and hypoactivations of the VS during reward processing have been observed in adolescents and among the patients with alcohol dependence (Beck et al., 2009, Peters et al., 2011, Bjork et al., 2008b).

3.1.3 Stress and reward processing

Evidence has demonstrated the influence of stress on sensitivity to reward in animals. Stress reactivity is regulated by the HPA axis and the extrahypothalamic brain stress system (Wust, 2004, Herman et al., 2003). Animal research has suggested that stress may alter the activations of the reward system at molecular level, as the infusion of corticosterone was found to upregulate extracellular DA level in rat NAc during the dark phase (Piazza et al., 1996). Stress introduced by maternal separation was found to increase fearful responses and reduce exploration in rhesus monkeys; such observation was accompanied by the elevations of alcohol consumption and cortisol level (Higley et al., 1991). Similar response had been observed in rodents, in which maternal separation altered place conditioning preference to opioids in rats, suggesting that postnatal life stress can affect the rewarding value of psychoactive drugs (Michaels and Holtzman, 2008). The potential influence of stress on reward sensitivity may also predispose as risk factor for substance use in humans.

3.1.4 Aims and hypotheses

This chapter aims to assess the relationship between known risk factors and substance use behaviour in IMAGEN adolescents at age 14 and age 16. Based on previous literature, it is hypothesised that self-report life stress measured by negative life events, and the ventral striatal BOLD responses measured during the anticipation of a reward, will significantly associate with increased alcohol and tobacco use in adolescents. Whether life stress can alter the VS activations during reward anticipation will be investigated.

3.2 Methods

3.2.1 Participants and phenotypes

At age 14, N = 2025 adolescents completed both the European School Survey Project on Alcohol and Other Drugs questionnaire (ESPAD) (Hibell et al., 1997) and the Life Events Questionnaire (LEQ) (Newcomb et al., 1981) (See

Table 3-1). N = 2030 adolescents completed the MID task at age 14, after applying the quality control procedures, data from N = 1403 adolescents was used for analyses in this chapter (N = 668 Males, N = 735 Females, N = 1228 right-handed, N = 175 left-handed or ambidextrous). In the follow-up assessment of the IMAGEN adolescents at age 16, N = 1169 adolescents completed both the ESPAD and the LEQ (See

Table 3-2).

Alcohol and tobacco use phenotypes in the ESPAD

Information for substance use was measured using the ESPAD questionnaire in adolescents at age 14 and age 16 (Hibell et al., 1997) (See **Chapter 2 Section 2.2.4** for details). Adolescents were classified as either ‘drinkers’ or ‘non-drinkers’, based on the self-report alcohol intake in the past 12 months. Individuals responded on a 7-point scale indicating the number of alcohol drinking occasions in the past 12 months. Adolescents who reported more than two drinking occasions were classified as 'drinkers' and adolescents who reported two drinking occasions or less were classified as 'non-drinkers'. Information on lifetime tobacco use was used to classify adolescents into ‘smoker’ or ‘non-smoker’. Adolescents who had smoked tobacco more than two occasions were classified as 'smoker' and those who smoked tobacco twice or less were classified as 'non-smoker' (See

Table 3-1 and

Table 3-2).

Table 3-1 Alcohol and tobacco use in adolescents at age 14.

	***Total: N* (%)**	***Males***	***Females***
Lifetime tobacco use at age 14			
Non-Smokers	1696 (83.8)	848	848
Smokers	329 (16.2)	160	169
Alcohol use in the past 12 months at age 14			
Non-Drinkers	1306 (64.5)	626	680
Drinkers	719 (35.5)	330	389

Table 3-2 Alcohol and tobacco use in adolescents at age 16.

	***Total: N* (%)**	***Males***	***Females***
Lifetime tobacco use at age 16			
Non-Smokers	731 (62.5)	351	380
Smokers	438 (37.5)	191	247
Alcohol use in the past 12 months at age 16			
Non-Drinkers	245 (21.0)	107	138

Drinkers	924 (79.0)	435	489

Negative life events

Life stress was measured using 20 items from the Life Events Questionnaire (LEQ) (Newcomb et al., 1981) (See **Chapter 2 Section 2.2.3** for details). These items capture life events that might have occurred in the school, family, peers as well as physiological changes and illnesses. Individuals were asked to report whether the life events had occurred and to rate how these events made them feel using a five point scale indicating 'very happy', 'happy', 'neutral', 'unhappy' and 'very unhappy'. Lifetime frequency and ratings towards negative life events were used to characterise life stress experienced by the adolescents at age 14 and age 16.

VS BOLD responses

Functional MRI data were acquired from eight IMAGEN assessment centres, using 3T MRI scanners of different manufacturers described in **Chapter 2 Section 2.3.2**. The VS activations from the reward anticipation ('Large win – No win') contrast of the MID task were selected to indicate reward sensitivity in adolescents (See **Chapter 2 Section 2.3.5**). The VS ROI was extracted based on the coordinates that were previously reported by O'Doherty et al., (2004) using peak coordinates of ± 15, 9 -9 and a sphere radius of 9 mm (O'Doherty et al., 2004). The VS BOLD responses from the reward anticipation ('Large win – No win') contrast were extracted using the Marsbar toolbox (http://marsbar.sourceforge.net). The data was exported for statistical analyses in the SPSS (version 20, IBM).

3.2.2 Statistical analyses

Logistic regressions were performed to investigate whether negative life events and VS BOLD responses could predict alcohol and tobacco use in adolescents. The regression models were in the form

Alcohol use (Drinkers *versus* Non-drinkers)

$$OR = b_0 + b_1 * P + b_2 * \text{Gender} + b_3 * \text{Sites} + \varepsilon,$$

Tobacco use (Smokers *versus* Non-smokers)

Where *P* referred to the predictors **a)** frequency or **b)** ratings towards negative life events measured at age 14 and 16; **c)** left and **d)** right VS BOLD responses). All associations controlled for gender and recruitment sites. For the analyses using VS BOLD responses, handedness information was also included as a covariate.

Linear regressions were performed to investigate whether negative life events could predict VS activations in 14-year-old adolescents. The regression models were in the form

$$\text{VS BOLD responses (Left or Right)} = b_0 + b_1 * P + b_2 * \text{Gender} + b_3 * \text{Sites} + b4 * \text{Handedness} + \varepsilon,$$

Where *P* referred to the frequency or the ratings towards negative life events measured at age 14. The associations were controlled for gender, handedness and recruitment sites. All analyses were performed using the SPSS (version 20, IBM).

3.3 Results

3.3.1 Negative life events promotes substance use

Logistic regression models were applied to investigate whether the frequency and ratings of negative life events could predict substance use in the IMAGEN adolescents. At age 14, both adolescent drinkers and smokers experienced more negative life events compared to the non-drinkers and the non-smokers (Alcohol use: Nagelkerke R^2 = .044; Tobacco use: Nagelkerke R^2 = .093, see **Figure 3-1a** and **b,**

Table 3-3). Adolescent drinkers and smokers also gave lower ratings to negative life events (i.e. perceived greater unhappiness) compared to the non-drinkers and non-smokers (Alcohol use: Nagelkerke R^2 = .017; Tobacco use: Nagelkerke R^2 = .022, see **Figure 3-1c and d,**

Table 3-3).

Similarly, adolescent drinkers and smokers reported significantly more negative life events compared to the non-drinkers and non-smokers at age 16 (Alcohol use: Nagelkerke R^2 = .052; Tobacco use: Nagelkerke R^2 = .114, see **Figure 3-2 a** and **b; Table 3-4**). Adolescent drinkers and smokers also gave lower ratings to negative life events compared to non-drinkers and non-smokers at age 16 (Alcohol use: Nagelkerke R^2 = .027; Tobacco use: Nagelkerke R^2 = .047, see **Figure 3-2 c** and **d**; **Table 3-4**). In addition, there was a significant gender main effect on alcohol use, as female drinkers rated greater unhappiness in the negative life events compared to the male drinkers (Female drinkers: Nagelkerke R^2 = .028, Male drinkers: Nagelkerke R^2 = .026, see **Table 3-4**).

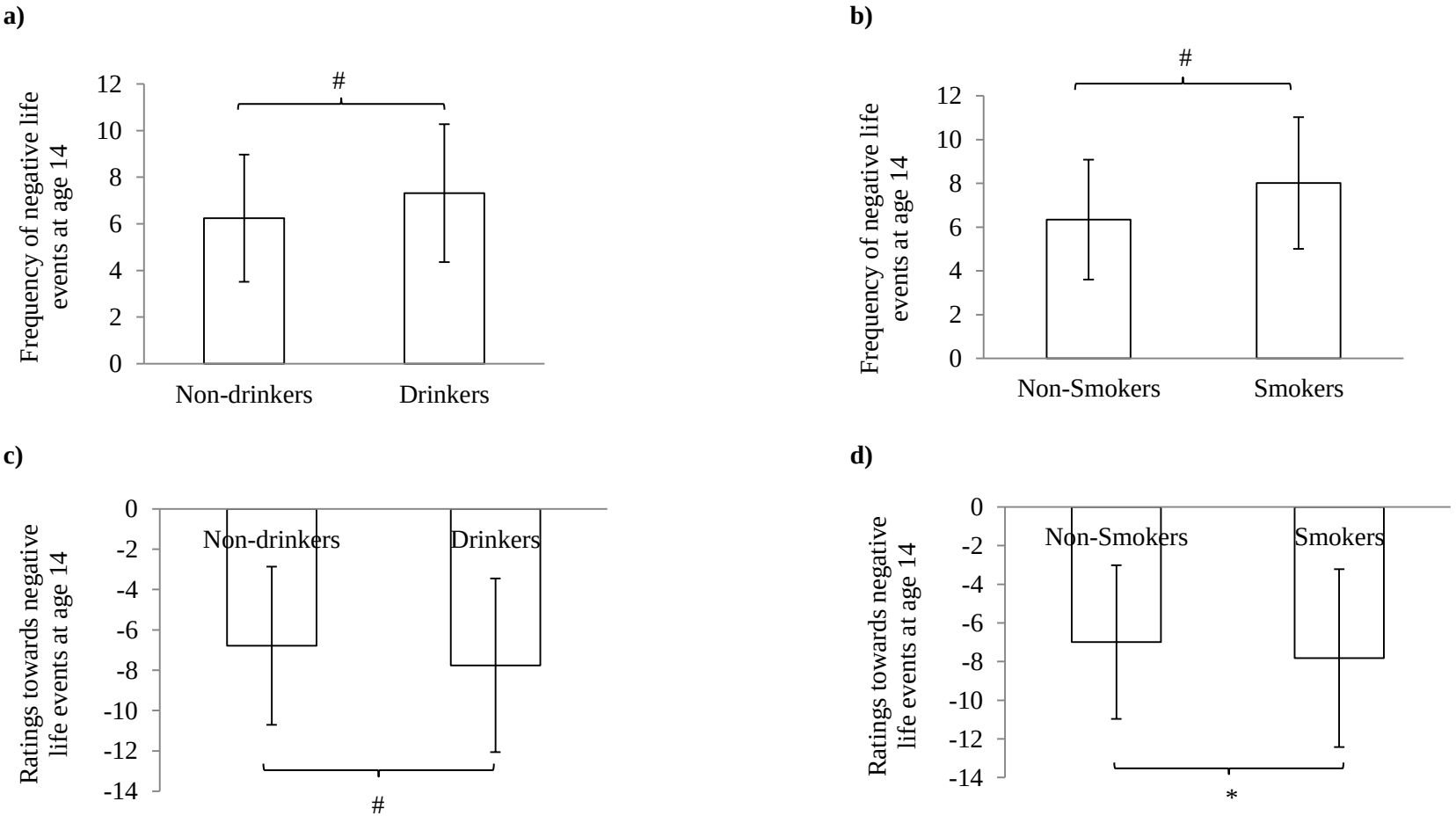


Figure 3-1 Negative life events and substance use in 14-year-old adolescents. Both 14-year-old adolescent drinkers and smokers experienced more negative events and gave lower ratings to the events (i.e. perceived greater unhappiness in the events, Mean ± S.D., * refers to $p < .05$, # refers to $p < .0001$).

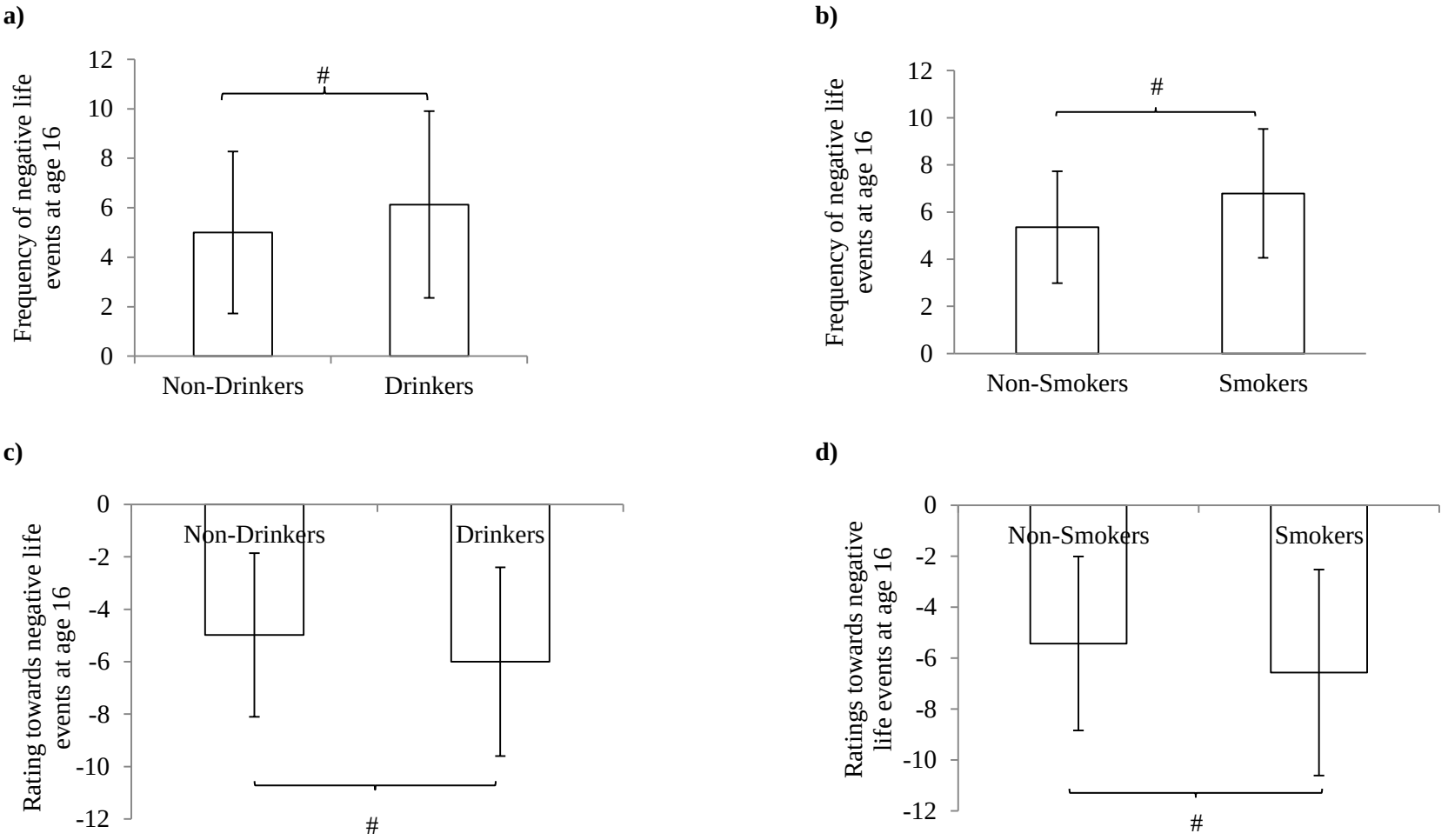


Figure 3-2 Negative life events and substance use in adolescents at age 16. Both 16-year-old adolescent drinkers and smokers experienced more negative events and gave lower ratings to the events (i.e. perceived greater unhappiness in the events, Mean ± S.D., # refers to $p < .0001$).

Table 3-3 Negative life events and substance use in adolescents at age 14

	Negative life events							
	Ratings (Feeling)				*Frequency*			
	β	***95% C.I.***	***p***	***Nagelkerke R^2***	***β***	***95% C.I.***	***p***	***Nagelkerke R^2***
Lifetime tobacco use	-.058	[-.088 - -.028]	1.66×10^{-4}	.022	.232	[.185 - .279]	1.42×10^{-23}	.093
Past 12 months alcohol use	-.055	[-.079 - -.032]	4.79×10^{-6}	.017	.144	[.109 - .179]	1.44×10^{-15}	.044

Table 3-4 Negative life events and substance use in adolescents at age 16

	Negative life events							
	Ratings (Feeling)				*Frequency*			
	β	***95% C.I.***	***p***	***Nagelkerke R^2***	***β***	***95% C.I.***	***p***	***Nagelkerke R^2***
Lifetime tobacco use	-.089	[-.122 - -.056]	1.21×10^{-7}	.047	.228	[.180 - .277]	5.42×10^{-20}	.114
Past 12 months alcohol use								
All	-.092	[-.135 - -.048]	3.80×10^{-5}	.027	.188	[.126 - .249]	2.04×10^{-9}	.052
Male	-.098	[-.175 - -.022]	.012	.026				
Female	-.086	[-.140 - -.032]	.002	.028				

3.3.2 VS activations and substance use

Logistic regressions were applied to investigate whether reward sensitivity, indicated by the VS BOLD responses during reward anticipation, could be a predictor for alcohol and tobacco use in 14-year-old adolescents. The analysis showed that adolescent drinkers displayed significantly greater BOLD response in the bilateral VS during reward anticipation compared to the non-drinkers (Left VS: β = .286, *C.I.* = [.013 - .559], p = .040, Nagelkerke R^2 = .052; Right VS: β = .284, *C.I.* = [.015 - .554], p = .039, Nagelkerke R^2 = .053, see **Figure 3-3**).

Similarly, adolescent smokers displayed significantly greater BOLD responses in the right VS compared to the non-smokers (β = .430, *C.I.* = [.070 - .790], p = .019, Nagelkerke R^2 = .066). However, no significant difference in the left VS BOLD responses was found between the smokers and the non-smokers (β = .274, *C.I.* = [-.089 - .638], p = .139, see **Figure 3-4**).

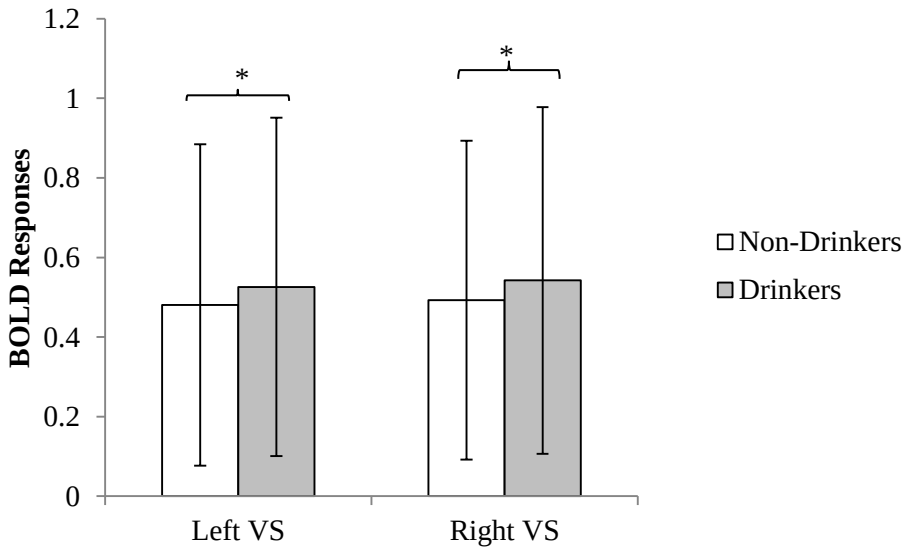


Figure 3-3 VS activations during reward anticipation between the 14-year-old adolescent drinkers and non-drinkers. (Mean ± S.D., * refers to p < .05).

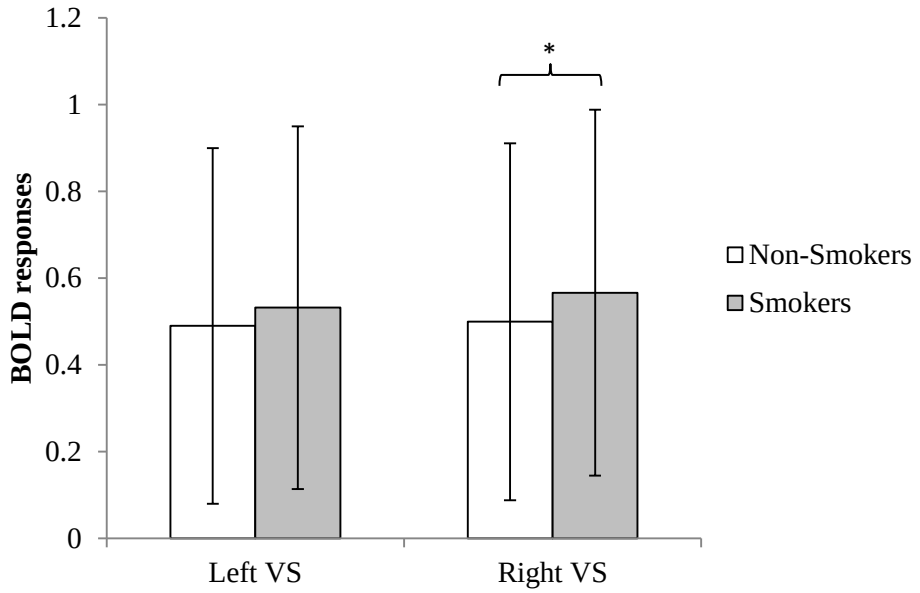


Figure 3-4 VS activations during reward anticipation between the 14-year-old smokers and non-smokers. (Mean ± S.D., * refers to $p < .05$).

3.3.3 Negative life events did not alter the VS activations

Linear regressions were applied to investigate whether the VS activation during reward anticipation could be influenced by life stress or negative life events in the 14-year-old adolescents. Neither the number nor the perceived unhappiness of negative life events the adolescent reported experiencing were associated with the VS activations during reward anticipation (See **Table 3-5**).

Table 3-5 Negative life events and VS activation during reward anticipation

	Negative Life events in 14-year old adolescents					
		Ratings (Feeling)			*Frequency*	
	β	***95% C.I.***	***p***	***β***	***95% C.I.***	***p***
Left VS BOLD responses	.002	[-.004 - .007]	.528	-.005	[-.012 - .003]	.216
Right VS BOLD responses	.001	[-.004 - .007]	.617	-.004	[-.012 - .003]	.270

3.4 Discussion

The use of alcohol and tobacco in adolescents is associated with various risk factors including (and not limited to) stressful environment and reward sensitivity (Blomeyer et al., 2008, Booker et al., 2008, Beck et al., 2009, Peters et al., 2011). This study demonstrated that the number of negative life events an adolescent experienced was strongly associated with alcohol and tobacco use at both age 14 and age 16. At age 16, but not at age 14, female drinkers gave lower ratings to negative life events (i.e. perceived greater unhappiness in the events) compared to the male drinkers. Adolescents who drank or smoked at age 14 also showed greater VS BOLD responses during the anticipation of a large reward compared to their abstinent peers. The VS responses were not associated with the number or ratings of negative life events.

3.4.1 Negative life events predict substance use

Both the frequency and the ratings towards negative life events were selected as indicators for life stress experienced by the IMAGEN adolescents. Whilst the frequency of negative life events might provide a relatively unbiased way to quantify stress, the ratings towards negative life events might indicate the level of life stress perceived by the adolescents. Previous research has demonstrated the positive correlation between the frequency of stressful life events and substance use (Pilowsky et al., 2009, Blomeyer et al., 2008, Booker et al., 2008). Consistent

with previous literature, adolescents from the IMAGEN study who experienced more negative life events or gave lower ratings to the events were more likely to consume alcohol and tobacco.

The impact of negative life events on substance use was found in both male and female adolescents. However, significant gender main effect was only observed in the 16-year-old adolescent drinkers. Female drinkers gave lower ratings to the negative life events compared to the male drinkers, indicating that the negative life events could exert a larger influence on alcohol use among the female adolescents. Gender differences in negative life events and stress-related alcohol use and craving were previously observed in adolescents and adults (Newcomb et al., 1981, Chaplin et al., 2008). Due to the self-report nature of the life events data, there has been increasing focus on studying the impact of perceived life stress. A recent longitudinal study showed that the perceived life stress significantly predicted alcohol and cigarette use among the university students (Tavolacci et al., 2013). The perceived life stress continues to influence substance use behaviour in the adulthood (Meier et al., 2013); as higher level of perceived life stress was significantly associated with substance dependence and recovery in adults (Meier et al., 2013). Results from the current study strengthened the role of perceived life stress in substance use during adolescence.

Given that the variance explained of negative life events on substance use increased with age, the data suggested life stress could become increasingly important in regulating substance use throughout adolescence. Reactivity to life stress and stress-induced substance use can be regulated by the brain stress system and the HPA axis in different ways. At the neuroendocrine level, disruptions in diurnal cortisol level were found to associate with elevation of life stress in adolescents and smokers (Essex et al., 2011, Steptoe and Ussher, 2006). Disruptions of the stress system can also be found in children of individuals with alcohol and substance addictions, as they showed blunted cortisol level upon the presentation of psychosocial threat (Evans et al., 2013). Nevertheless, reduced *CRF* and *c-fos* mRNA expressions in the PVN were found in rats

exposed to ethanol or nicotine (Richardson et al., 2008, Schmitt et al., 2008), suggesting alteration of the stress system activity among substance users. Genetic polymorphisms in stress genes, such as *CRHR1* rs1876831 and *PER1* rs3027172 were found to associate with stress-induced alcohol intake in adolescents (Blomeyer et al., 2008, Dong et al., 2011). The vulnerability of substance addictions can be due to the increased interactions between stressful environment with genes and neurotransmitters throughout life. Such interactions may account for the increased variance explained of negative life events on substance use in the IMAGEN adolescents.

3.4.2 Reward sensitivity predicts substance use

In this analysis the increased VS BOLD responses during reward anticipation were found to associate with alcohol and tobacco use in the adolescents at age 14. Consistent with previous studies, the analysis confirmed that the VS activations were responsible for the increased substance use in adolescents (Beck et al., 2009, Peters et al., 2011, Galvan et al., 2006). The findings supported the hyperactivity hypothesis, indicating that the increased activations in the VS might lead to hyperactivation of the reward system, subsequently encourage adolescents to adopt reward-seeking behaviour in the form of substance use (Galvan et al., 2006, Galvan, 2010, Van Leijenhorst et al., 2010).

The right VS showed larger variances explained for alcohol and tobacco use in the IMAGEN adolescents. Such observation could be partly due to the asymmetric activations in the VS, in which greater BOLD responses were observed in the right VS in the adolescents. Findings from the previous studies might also explain the lack of association between the left VS activations and tobacco use in the IMAGEN adolescents. For example, Beck and colleagues showed significant differences were only observed in the right VS activations between patients with alcohol dependence and healthy individuals (Beck et al., 2009). Asymmetric activations were identified in various brain regions including the bilateral VS, orbitofrontal cortex, insula, frontal

gyrus and cerebellum among individuals with nicotine dependence and healthy controls during the presentation of smoking-related cues and craving (Yalachkov et al., 2009, Rose et al., 2007).

3.4.3 Life stress did not alter VS activations

Although negative life events and the VS activation during reward anticipation were associated with increased substance use in the IMAGEN adolescents, there was no significant association between these two risk factors. The role of striatum in emotional regulation was previously demonstrated (Delgado et al., 2008). Despite life stress were shown to alter reward processing in previous studies (e.g. (Michaels and Holtzman, 2008, Higley et al., 1991, Dalm et al., 2012), the types of stress and preference to reward stimuli were manipulated in experimental settings; as opposed to the current analysis of which various forms of negative life events in human adolescents were captured. In addition, the impact of life stress on reward processing can be mediated by neurotransmitters and hormones from the reward and stress systems. For example, psychological stress and glucocorticoids were found to alter the extracellular DA and TH level (Piazza et al., 1996, Niwa et al., 2013). Nevertheless, the DNA methylation level of *TH* in *DISC1* knock-out mice could be influenced by social isolation stress (Niwa et al., 2013), demonstrating the possibility of stress-induced epigenetic modifications in the dopaminergic system. As negative life events could become increasingly important in predicting substance use during adolescence, the interactions between the stress and reward systems at the gene and behavioural levels could become more influential throughout development.

3.4.4 Limitations

Illustrated by a meta-analysis performed on 142 reward-related functional neuroimaging studies, a number of frontal and limbic brain regions including the anterior cingulate cortex, medial orbitofrontal cortex, insula and the thalamus were activated during reward anticipation (Liu et al., 2011). Activations patterns of these additional brain regions may contribute to the vulnerability for substance use in adolescents. In this study, only the VS activations were

selected for analyses since they represented the peak activations during the reward anticipation phase of the MID task.

3.4.5 Conclusions

As demonstrated in this chapter, both negative life events and reward sensitivity were risk factors leading to the increased alcohol and tobacco use among the IMAGEN adolescents. Increase gene-environment interaction might explain the increased variance explained for negative life events on substance use throughout adolescence. Although life stress did not alter reward sensitivity in the adolescents, interactions between the stress and reward systems were observed in the literature at both genetic and epigenetic levels. Together with the findings from previous studies, the analyses presented in this chapter have granted further investigations of the genetics and epigenetics of substance use in the IMAGEN adolescents.

Chapter 4

The role of life stress on *PER1* DNA methylation, reward sensitivity and substance use

4.1 Introduction[2]

4.1.1 Period genes and circadian rhythm

Disruption of the circadian rhythm is a risk factor for substance use and addictions (See Chapter 1 General Introduction). One major component of the circadian system, namely Period genes, is evolutionarily conserved across species. *Per1, Per2* and *Per3* are expressed at different times of the day to sustain the circadian cycle. The mammalian homologue of the *Drosophila* period gene was first reported by Sun et al. (1997) and Tei et al. (1997). It was once named as *Rigui* and then renamed as *Per1* when the other two *Per* homologues, *Per2* and *Per3* were identified by Shearman et al. (1997) and Takumi et al. (1998). Comparison of the *Per* sequences across species indicated that human and mouse *Per1* share 92% homology (Tei et al., 1997). The sequences of *Per1, Per2* and *Per3* were also found to share 44 - 56% similarity with each other (Shearman et al., 1997, Takumi et al., 1998).

4.1.2 The role of *PER1* in substance use

In human, *PER1* is located on chromosome 17 at locus 17p13.1. The promoter region of *PER1* spans about ±1400bp from the transcriptional start site (TSS) (Motzkus et al., 2000). The role of *PER1* in substance use has been demonstrated in both animal and human studies. The genetic

[2] The data of this chapter was presented at The XXth World Congress of Psychiatric Genetics. Abstract title: CP Wong *et al*. (2012). Profiling DNA Methylation in *Period 1*, Negative Life Events and Alcohol Intake in Adolescents.

polymorphism on *PER1* rs3027172 (C/T) was found to associate with increased alcohol consumption in adolescents who experienced high level of psychosocial adversity and among patients with alcohol dependence (Dong et al., 2011). Increased ethanol consumption was also observed in *Per1* mutant mice after the exposure to physiological and psychosocial stressors (Dong et al., 2011). When investigating the molecular mechanism of rs3027172, it was discovered that the risk allele of rs3027172 prevent the induction of *PER1* mRNA expression upon the stimulation with cortisol in human cell lines (Dong et al., 2011). Results from the chromatin immunoprecipitation assay showed that the binding of glucocorticoid receptors to the proximal and distal GRE sites on *Per1* promoter could induce *Per1* mRNA expressions in mice exposed to restraint stress (Yamamoto et al., 2005). The transcriptional activity of *PER1* can also be regulated by the E-Boxes and Snail1 (Yamamoto et al., 2005, Yan et al., 2008, Dong et al., 2011), suggesting that the gene regulatory elements on *PER1* can be crucial in regulating *PER1* mRNA expression and the circadian rhythm during stressful situations.

In addition, *PER1* were also found to associate with reward sensitivity and substance use. In response to an acute cocaine injection, mice with mutant *Per1* showed reduced behavioural sensitisation compared to the non-mutant controls (Abarca et al., 2002). Suggested by Abarca and colleagues, behavioural sensitisation indicated the tendency of cocaine craving and relapse (Abarca et al., 2002, Robinson and Berridge, 1993). Similarly, disruption of *Per1* mRNA expression was found to inhibit the conditional place preference of morphine in mice (Li et al., 2008), indicating the role of *Per1* in mediating reward sensitivity and substance use behaviour.

4.1.3 The epigenetics of substance use and circadian rhythm

Both genes and exposure to environmental risk factors can contribute to the development of substance use and addictions (e.g. (Kendler et al., 1999, Waterland and Jirtle, 2003). One potential mechanism that mediates the gene-environment interactions is via epigenetic modifications. The 5'-cytosine methylation (DNA methylation) is one of the most studied forms of epigenetic modifications, of which the conversion of cytosine to methylated 5’-cytosine is

catalysed by various DNMTs (Denis et al., 2011, Robertson and Wolffe, 2000). Evidence for differential DNA methylation in substance use and addictions has been demonstrated in the literature. For example, variations in DNA methylation of *POMC* and *MAOA* were found among individuals with alcohol and nicotine dependence (Muschler et al., 2010),(Philibert et al., 2010). Apart from the traditional transcriptional-translational machinery, epigenetic modifications can be crucial for adjusting the circadian rhythm in response to environmental input. The major component CLOCK of the circadian system itself is a histone acetyltransferase (Doi et al., 2006), supporting the importance of epigenetic modifications in executing the circadian rhythm.

In addition to the impact of DNA methylation on behaviour and diseases, recent evidence has demonstrated the link between neural activity and DNA methylation in brain. The first epigenome-wide analysis conducted by Guo and colleagues (2011) revealed changes in the DNA methylation landscape of adult mouse brain after electroconvulsive stimulation. The *in vivo* stimulation of dentate granule neurons in hippocampus introduced demethylation and/ or *de novo* methylation in about 1.4% of the assayed CpG sites, of which 31% (N = 945) CpG sites maintained stable DNA methylation status after 24 hours (Guo et al., 2011). In particular, demethylation at CpG sites located in *Crebbp* (CREB binding protein), *Grip1* (glutamate receptor interacting protein 1) and *Per2* in mice were observed (Guo et al., 2011), suggesting that the input of external stimuli might have an impact on neurotransmission and the regulation of circadian rhythm.

4.1.4 DNA methylation in *PER1*

Early investigation of *PER1* DNA methylation was conducted in cancer research, with a particular interest in studying the DNA methylation at the E-boxes due to the presence of CG dinucleotide at the sequence (5'-CACGTG-3'). Evidence suggested that DNA methylation at E-boxes could elevate *PER1* mRNA expression in cervical cancer cell lines (Hsu et al., 2007). Contrary to Hsu et al.'s (2007) study, hypermethylation of *PER1* decreased *PER1* mRNA

expression in the peripheral blood among patients with chronic myeloid leukaemia (Gery et al., 2007). The mixed findings in the literature can be due to small sample size in studies and the fact that the DNA methylation pattern can be specific to both tissue and disease (Eckhardt et al., 2006). Nonetheless, these finding has suggested that methylation in the E-boxes of *PER1* can be crucial for predicting disease status.

4.1.5 Aims and hypotheses

This chapter focuses on identifying the role of epigenetic modifications in *PER1* in response to life stress, and its consequences in reward sensitivity, alcohol and tobacco use in adolescents. It is hypothesised that DNA methylation in *PER1* varies with the level of life stress, which may predict reward sensitivity, alcohol and tobacco use in the adolescents. The DNA methylation at rs3027172 and the regulatory elements in *PER1*, including the GRE site, the first exon and the E-Box, is also expected to vary across individuals due to life stress.

4.2 Methods

4.2.1 Participants and phenotypes

Participants

Whole blood DNA samples from N = 716 adolescents were used for bisulfite conversion and DNA methylation analyses in this chapter.

Alcohol and tobacco use phenotypes

In these 716 individuals, N = 636 completed the European School Survey Project on Alcohol and Other Drugs questionnaire (ESPAD) (Hibell et al., 1997) (See **Chapter 2 Section 2.2.4** for details). Adolescents who had more than two drinking occasions in the past 12 months were regarded as 'drinkers' and those who had drunk alcohol in less than two occasions were regarded as ‘non-drinkers’. Adolescents who had smoked tobacco in more than two occasions in lifetime

were classified as 'smokers' and those who smoked in less than two occasions were classified as ‘non-smokers’ (see **Table 4-1**).

Table 4-1 Substance use information for DNA methylation analyses in 14-year-old adolescents.

	Total: N (%)	***Males***	***Females***
Past 12 months alcohol use at age 14			
Non-Drinkers	390 (61.3)	204	186
Drinkers	246 (38.7)	104	142
Lifetime tobacco use at age 14			
Non-Smokers	528(83.0)	251	277
Smokers	108(17.0)	57	51

Negative life events as an indicator of life stress

N = 652 completed the Life Events Questionnaire (LEQ) to assess the level of life stress in adolescents at age 14. The level of life stress was measured by 20 negative life event items derived from the Life Events Questionnaire (Newcomb et al., 1981) (See **Chapter 2 Section 2.2.3** for details). Individuals were asked to report whether each life events had occurred and rate the events on a five point scale indicating 'very happy', 'happy', 'neutral', 'unhappy' and 'very unhappy'. Both lifetime frequency and ratings towards negative life events were used to characterise life stress experienced by the adolescents.

Reward sensitivity measured by the VS BOLD responses

The VS BOLD responses during reward anticipation ('Large win – No win') contrast in the MID task were used to indicate reward sensitivity of the adolescents. The ROIs for the VS were defined as ± 15, 9 -9, with a sphere radius of 9 mm (O'Doherty et al., 2004). The BOLD responses of the VS ROIs were extracted using the Marsbar toolbox (http://marsbar.sourceforge.net). The MID task were assessed in the adolescents at age 14, of which the VS activation data from N = 438 individuals were used for analyses (N = 195 Males, N = 243 Females, N = 384 right-handed, N = 54 left-handed or ambidextrous, see **Chapter 2 Section 2.3.5** for details).

4.2.2 Detection of DNA 5-cytocine methylation in PER1 through sodium bisulfite treatment and Sequenom MALDI-TOF spectrometry

4.2.2.1 Background of sodium bisulfite treatment

Various methods have been developed to investigate DNA methylation in recent years. Many of them rely on the conversion of genomic DNA into single-stranded DNA using sodium bisulfite (See **Figure 4-1**). During the conversion, sodium bisulfite deaminates cytosine (C) into uracil (U), whereas 5-methylcytosine (5-MeC) remains non-reactive. The change in base-pair after the conversion allows quantification of 5-MeC using different techniques, including PCR amplification, DNA sequencing and microarrays (Laird, 2003). Sodium bisulfite treatment is efficient and its versatile application makes it become the most commonly used approach for studying DNA methylation.

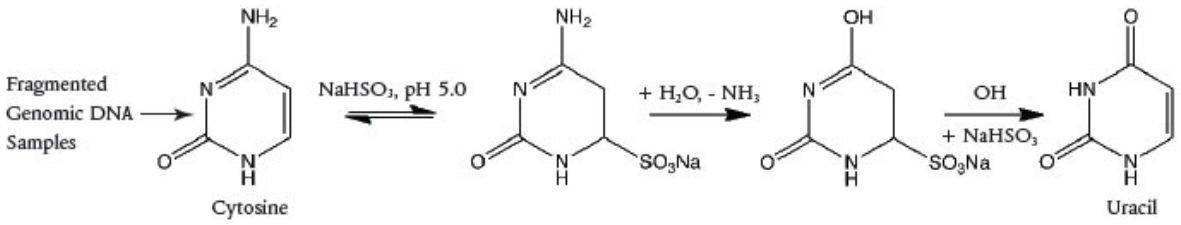


Figure 4-1 Mechanism for sodium bisulfite conversion. Cytosine is susceptible to sodium bisulfite conversion into uracil whereas 5'-methylcytosine remains non-reactive.

(Picture courtesy of The New England Biolabs http://www.neb.com/nebecomm

/tech_reference/ epigenetics/Bisulfite_Conversion.asp#.UE4n3q7LnEU)

4.2.2.2 Procedure for sodium bisulfite conversion of genomic DNA

250ng whole blood genomic DNA from 716 14-year-old adolescents were undergone bisulfite conversion using EZ-96 DNA Methylation™ Kit (Zymo Research, Cat. No. D5003). 250ng 0%, 50% and 100% methylated genomic DNA and double distilled water were included as controls to ensure the quality of bisulfite conversion and subsequent reactions. The 0% and 100% methylated genomic DNA were purchased from Qiagen (EpiTect unmethylated human control DNA, 50 ng/µl, cat. no. 59568) and BioLabs (CpG Methylated HeLa Genomic DNA, 100µg/ml, cat no.N4007S). The 50% methylated DNA was prepared by mixing equal amount of 0% and 100% methylated DNA. All samples and controls were randomly pipetted in the 96-well conversion plates provided in the kit.

The samples were incubated with 5µl M-Dilution Buffer (contains NaOH) and double-distilled water (in a total volume of 50µl) at 37℃ for 15 minutes. After the incubation, 100µl CT Conversion Reagent was added to the denatured DNA and further incubated for 16 hours at 50℃ and then for 10 minutes at 4℃. The CT Conversion Reagent contains sodium bisulfite and it is light sensitive, therefore the plates should be prepared with minimal exposure to light. The CT Conversion Reagent was supplied in powder form in the kit and it was prepared by adding 7.5ml double distilled water and 2.1ml M-Dilution Buffer, followed by frequent vortexing for 10 minutes at room temperature. The CT conversion reagent was prepared in the dark to reduce the exposure to light. It was used immediately following the preparation to enhance the efficiency of conversion.

The samples were transferred to the 96-well Silicon-A^{TM} Binding Plate and mixed with 400µl M-Binding Buffer to allow binding of DNA to the spin columns on the plate. The plates were then centrifuged at 2000 rpm for 5 minutes. After adding 500µl M-Wash Buffer, the plates were centrifuged at 2000 rpm for 5 minutes.

The samples were desulfonated on the column by incubating with 200µl M-Desulphonation Buffer for 18 minutes at room temperature. The samples were centrifuged at 2000 rpm for 5 minutes. Then the samples were washed twice in 500µl M-Wash Buffer and centrifuged at 2000rpm for 5 and 10 minutes respectively. The M-Wash Buffer removes any residue buffers and impurities that may affect sample quality. The samples were then incubated with 25µl M-Elution Buffer for 5 minutes followed by centrifugation at 2000rpm for 3 minutes to elute the DNA. This procedure was repeated again to obtain 50µl DNA.

The bisulfite DNA is fragile and therefore it should be handled with great care. The samples were aliquoted into 5 plates (10µl DNA in each well) to minimise DNA degradation caused by multiple freeze-thaw cycles. All samples were kept at -80℃ as recommended by the protocol.

4.2.3 Amplified regions in *PER1*

Proximal GRE site amplicon

There are two CpG islands located at the promoter region of *PER1* and one of them overlaps with the first exon of *PER1*. One amplicon was designed to cover the GRE site locates near the first exon of *PER1*. Although this GRE site contains a CG dinucleotide which makes it susceptible to DNA methylation, no evidence on epigenetic modifications at GRE sites has been reported in the literature.

First exon amplicon

SNP rs3027172 located at the first exon of *PER1* was shown to influence alcohol intake in adolescents when exposed to stressful environment (Dong et al., 2011). Since the presence of minor allele (C) of rs3027172 will form an additional CpG site with the neighbouring guanine, it is possible that DNA methylation at this CpG will have additional impact on stress-induced alcohol drinking behaviour in adolescents. This amplicon was designed to investigate the DNA methylation pattern in *PER1* first exon and the possible interaction between DNA methylation and rs3027172 on alcohol intake.

Upstream amplicon

The 'upstream' amplicon was designed to cover the region that was previously studied by Hsu et al. (2007) and Gery et al. (2007), in which the mix findings for DNA methylation status and mRNA expression in *PER1* were reported. This amplicon covers one of the E-Box on the *PER1* promoter.

Primer Design

The genomic DNA sequences of the target regions, with 300bp added to both ends of the sequences were extracted from the UCSC genome browser http://genome.ucsc.edu/ (Human genome version 19, University of California, Santa Cruz). SNPs appeared in the extracted sequences were identified using dbSNP132 in the UCSC genome browser. The PCR primers were designed using online software EpiDesigner (http://www.epidesigner.com/start.html, Sequenom). The 'Mass Window' was set to '1500 to 9000' and the 'Product Size' was set to '200-500'bp. Only CpGs in T Reaction were analysed in both strand.

The selection of primer pairs was based on the following criteria:

1. Coverage and number of analysable CpG sites – primers that covered the highest number of analysable CpG sites were selected.
2. Length and melting temperature (*Tm*) of primer pairs – The length of primers should be between 24 - 26nt to ensure specificity. The forward and reverse primers should be of equal or similar length. The melting temperature for forward and reverse primers should be the same or differed by less than 1℃.
3. CpG sites and SNPs – Primers should not contain any CpG sites and any known SNPs to avoid biased amplification and PCR failure.
4. Amplicon size – The amplicon should not exceed 500bp as recommended by Sequenom. However such criteria should be adjusted according to the quality of the genomic and bisulfite DNA.

The reverse primers were tagged with a T7-promoter recognised by RNase A. An 8 base pair insert was attached next to the T7-promoter tag to prevent abortive cycling during the transcription. A 10mer tag was attached to forward primers in order to balance the primer length. To further assist assay design, *in silico* fragmentation analysis of amplicons were generated using 'RSeqMeth' package in R (Version 2.9) (Coolen et al., 2007). The RSeqMeth package predicts the sequence of fragments resulted from T and C cleavage reactions and whether these fragments are analysable in Sequenom (see **Table 4-2** and **Figure 4-2** for information of *PER1* amplicons).

Table 4-2 Details of *PER1* amplified regions and primer sequences. All reverse primers were tagged with T7-promoter whereas the forward primers were 10mer-tagged. The sequences for T7-promoter and 10mer tags were labelled in italics. All primers were purchased from Sigma.

PER1 amplicons	***Location***	***Amplicon size***	***Tm***	***Annealing Temp ℃***	***PCR cycles***
Proximal GRE site					
	-557bp downstream TSS	287bp	60.8	54.0	45
Forward primer: *aggaagagag*GGGAAGGTTGTGGTTAATAGTAGGA	(chr17: 8055196 - 8055483)				
Reverse primer: *cagtaatacgactcactatagggagaaggct*AAAATCTACACTCCTAACCCACTCA			59.5		

First exon and rs302717215					
	-297bp downstream TSS	419bp	60.2	54.0	45
Forward primer: *aggaagagag*GATGAGTGGGTTAGGAGTGTAGATTT	(chr17:8055456 - 8055875)				
Reverse primer: *cagtaatacgactcactatagggagaaggct*ACAACCCTAACCTTAATAAAAACCA			58.5		
Upstream					
	+1276bp upstream TSS	328bp	60.0	56.0	45
Forward primer: *aggaagagag*TTTAGGGAGAGGAGATTGAGAATTT					

(chr17:8057029 -

8057357)

Reverse primer:

*cagtaatacgactcactatagggagaaggct*TCATTATAAAAACACTCCCCTCACT 59.4

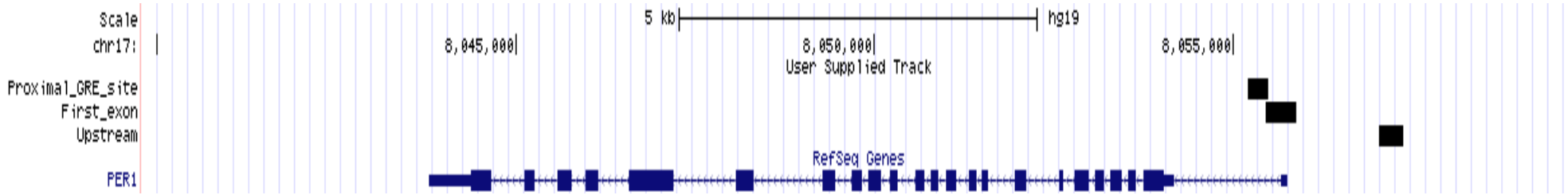


Figure 4-2 Locations of the *PER1* target regions on the UCSC genome browser (Hg19). Sequenom matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS)

Sequenom MALDI-TOF MS enables efficient and reproducible quantification of DNA methylation. This technique relies on successful bisulfite conversion of genomic DNA into single-stranded DNA. After the bisulfite treatment non-methylated cytosine residues (C) are converted into uracil, whilst 5-methyl cytosines (5-MeC) remain unchanged. The converted DNA is selectively-amplified, where the uracils are replaced by thymine (T), or adenine (A) on the complementary strand. The C/T variations at cytosine residues will appear as G/A variation after reverse transcription and base-specific cleavage. The G/A variation at CpG sites results in a mass difference of 16Da which can be detected by the MALDI-TOF MS (See **Figure 4-3**)

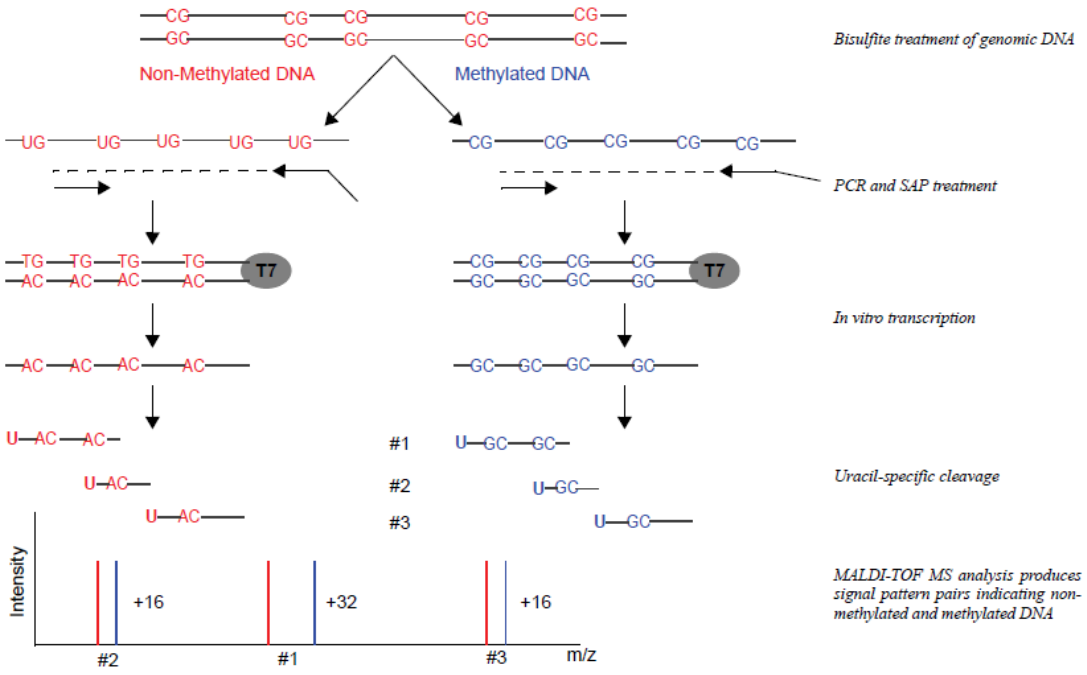


Figure 4-3 Detection of 5-MeC in the Sequenom MALDI-TOF platform. (Courtesy of the Sequenom User Manual).

Bisulfite PCR and clean-up

The PCR was initialised by activating the polymerase at 95℃ for 5 min. The reaction temperature was then reduced to allow the annealing of primers to the single-stranded bisulfite DNA template. The annealing temperature for different primer pairs can be found in **Table 4-2**.

The reaction temperature was increased to 72℃ in order to optimise the ability of polymerase to synthesise complementary DNA strand. The annealing and extension steps were repeated 45 times before the final elongation step, such that any remaining single-stranded DNA were complemented and extended to its full length. The plate was stored at 4℃ in the thermocycler before transferring to the -20℃ freezer for storage.

The PCR amplification of bisulfite DNA and controls was performed on 384-well plates (ABGene Thermo-Fast 384 Plates, Fisher, Cat. no.TF-0384). The PCR cocktail was prepared according to **Table 4-3**, with an overhang of 38% to account for possible pipetting loss. The PCR preparation was performed on ice to reduce DNA degradation and enzyme activity prior the amplifications. The PCR amplifications of samples were performed in duplicates to reduce errors. The duplicates were pooled after the PCR (See **Table 4-4** for PCR thermocycler conditions).

Table 4-3 PCR cocktail for bisulfite DNA amplification. All PCR reagents were purchased from the Sequenom (Cat No. 10132).

Reagent	***Volume for a single reaction (µl)***	***Final concentration for single reaction***	***Volume for a 384-well plate (µl , including 38% overhang: 530x)***
ddH_2O	1.42		752.60
10x Sequenom PCR Buffer (containing 20mM $MgCl_2$)	0.50	1X	265.00
dNTP mix (25mM each)	0.04	200 µM	21.20
Polymerase (5U/µl)	0.04	0.2 unit/ reaction	21.20

Forward primer (1μM)	1.00	0.2 μM	530.00
Reverse primer (1μM)	1.00	0.2 μM	530.00
Total Volume	4.00		2120.00
Bisulfite DNA	1.00		

Table 4-4 Thermocycler condition for the bisulfite PCR.

Temperature	***Time***	
95℃	5 min	
95℃	20 sec	
56℃	30 sec	45 cycles
72℃	1 min	
72℃	3 min	
4℃	∞	

The PCR products were incubated with a cocktail containing shrimp alkaline phosphotase (SAP) to dephosphorylate any unincorporated dNTPs and remove excess primers and salts. Towards the end of the reaction, the samples were heated at 85℃ to inactivate the SAP (See **Table 4-5** and **Table 4-6** for the SAP clean-up conditions).

Table 4-5 Cocktail mix for the SAP clean-up. All reagents were purchased from the Sequenom (Cat. No. 10129).

Reagent	***Volume for a single reaction (µl)***	***Final concentration for single reaction***	***Volume for a 384-well plate (µl , including 38% overhang: 530x)***
RNAase free water	3.40		1802.00
Shrimp alkaline phosphotase (SAP) (1.7 unit/µl)	0.6	1 unit/ reaction	318.00
Total Volume	4.00		2120.00

Table 4-6 Thermocycler condition for the SAP clean-up.

Temperature	***Time***
37℃	40 min
85℃	5 min
4℃	∞

Agarose gel electrophoresis

About 70 samples were selected from each plate to run on 1.5% agarose gel to confirm the quality of the PCR amplification. Gel electrophoresis separates DNA fragments according to their molecular weights, therefore allows the detection of non-specific binding and

contamination arise from the PCR. The addition of ethidium bromide gives fluorescence to the gel to allow visualisation of the DNA fragments.

The 1.5% agarose gel was prepared by dissolving 1.5g agarose to 100ml 1X Tris/Borate/EDTA (TBE) buffer. The mix was heated in a microwave oven for about 1.5 minute to fully dissolve the agarose, 2µl ethidium bromide was added to the mix after it was cooled. The gel was then poured slowly into a gel tank containing combs and was left to set for about 30 minutes. 1µl of PCR product was run in each lane along with 1µl 5X loading buffer and 3µl double distilled water. The HyperLadder V (Bioline, Cat. no. BIO-33031) was used to estimate the size of the PCR products. The ladder produces 12 bands ranging from 25bp to 500bp. The gels were run at 120V for 50 minutes (See **Figure 4-4**).

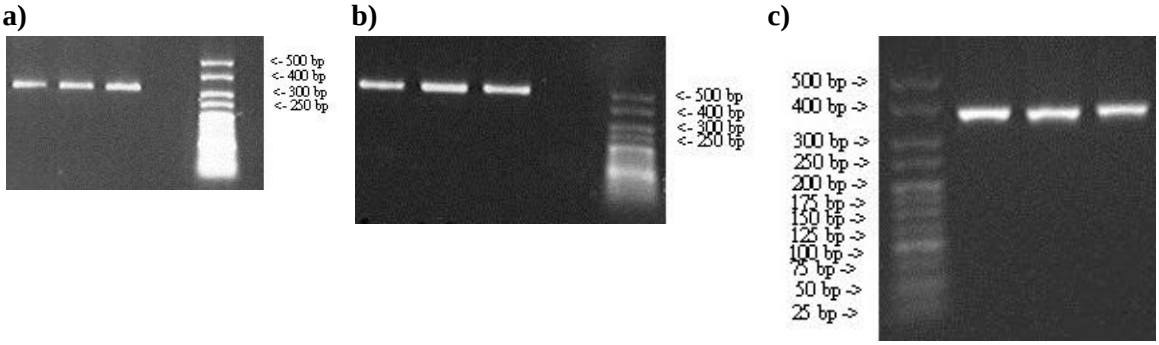


Figure 4-4 Agarose gel electrophoresis for *PER1* amplicons. Agarose gel electrophoresis was performed on a) proximal GRE site b) first exon and c) upstream amplicons in *PER1*. HyperLadder V was used to estimate the size of all amplicons. A single, clear band of PCR product has been observed in all amplifications; indicated that the T-cleavage reaction and the subsequent detection of the DNA methylation were unlikely to be affected by non-specific PCR amplifications.

RNA reverse transcription and base-specific cleavage reaction (T-cleavage reaction)

The PCR product were reverse transcribed into RNA, in which the C/T variation introduced at the bisulfite treatment appeared as G/A variation on the RNA transcript. After the transcription, the RNA was cleaved by the enzyme RNase A. RNase A cleaves the RNA transcript at every C

and U. Using either dCTPs (substitute of rCTPs) or dTTPs (substitute of rTTPs) during the transcription prevents RNase A from producing fragments that are too small to be analysed. Incorporating dCTP makes uracils available for cleavage (T/U-Cleavage) and adding dTTP can prevent the enzyme from cleaving the Ts (i.e. C-Cleavage). T-Cleavage reaction is more informative when analysing regions that have high CG content (See

Table 4-7 and **Table 4-8** for conditions for the T-cleavage reaction).

Table 4-7 Cocktail for T-cleavage reaction. All reagents were purchased from the Sequenom (Cat No. 10132).

T-Cleavage Transcription/ RNase A cocktail	***Volume for a single reaction (µl)***	***Final concentration for single reaction***	***Volume for a 384-well plate (µl, including 38% overhang: 530x)***
RNase free water	3.15		1669.5
5X T7 Polymerase Buffer	0.89	0.64x	471.7
T Cleavage Mix	0.24	NA	127.2
DTT (100mM)	0.22	3.14mM	116.6
T7 RNA & DNA polymerase	0.44	22 units/ reaction	233.2
RNase A	0.06	0.09mg/ml	31.8
PCR/SAP mix	2.00		
Total Volume	7.00		

Table 4-8 Thermocycler condition for T-cleavage reaction. To ensure the quality of samples, the reaction plates were removed from thermocycler upon completion and stored at -20℃ overnight

Temperature	*Time*
37℃	3 hours
4℃	∞

Resin clean-up and chip spotting

Purified water (20 μl) was added to each sample. Afterwards, the samples were conditioned with 6mg clean resin (Sequenom, Cat. No. 10117) to remove excess salt that might interfere with matrix-assisted desorption ionization–time of flight (MALDI-TOF) mass spectrometry. The plate was rotated for 20 minutes and centrifuged at 3800 rpm for 10 minutes.

The plate was transferred to the Sequenom Nanodispenser for spotting. Before spotting the chip, the pins on the Nanodispenser were conditioned with 100% 100 proof ethanol for 30 minutes and washed 10 complete cycles using 50% ethanol prior spotting. The pins were kept in 50% ethanol to prevent the accumulation of analyte when the Nanodispenser was not in use.

The Sequenom SpectroChip (384 SpectroCHIP; Sequenom Cat No. 10117) was placed on the Nanodispenser. SpectroChips contain 3-hydroxypicolinic acid (3-HPA) on the surface, which allows the analyte to crystalise. About 70μl 4-point calibrant (Sequenom, Cat No. 10117) was added to the calibrant reservoir in the Nanodispenser. The 4-point calibrant contains 4 quality-controlled oligos with known spectra mass signals (1479.0, 3004.0, 5044.4 and 8486.6Da). About 12 ± 7nl of samples and calibrant were spotted onto the Sequenom SpectroChip.

Sequenom MALDI-TOF mass spectrometry and the detection of DNA methylation

The Sequenom MALDI-TOF mass spectrometry measures the time-of-flight of the ionised fragments in vacuum (Sequenom, SpectroTyper RT software). The fragmented RNA transcript is ionised by laser and become positively charged. The charge gradient built up inside the vacuum chamber enables ionised fragments to accelerate from the top to the bottom of the chamber where the detector is located. The time-of-flight of an ionised fragment depends on its size and mass, therefore small and less methylated fragments travel quicker. Each methylated CpG site results a shift of 16Da on the mass spectrum. The quality of spectra can be viewed in the MassARRAY© and under 'Analyzer' in EpiTyper© (Sequenom). The DNA methylation status of a CpG site/unit can be calculated from the mass spectrum by comparing the signal intensity of methylated and non-methylated templates within each sample. DNA methylation is reported as a ratio ranging from 0 to 1, corresponding to 0% - 100% methylation.

4.2.4 Quality control for DNA methylation data

Missingness of data

Missing data can be affected by the quality of assays. For example, degradation of the bisulfite DNA can affect the quality of PCR and T-cleavage products. The DNA methylation data was initially examined to find out the percentage of missingness. Individuals with missing DNA methylation information were excluded to achieve ≥ 96% call rates in all *PER1* amplicons (**Table 4-9**).

Table 4-9 Quality control information of *PER1* amplicons. Individuals with missing DNA methylation information for ≥ 2 CpG units (≥ 3 CpG units for the upstream amplicon) were excluded. A call rate of over 96% was achieved after removing individuals with missing data (i.e. more than 2 CpG sites for the promoter amplicons and more than 3 CpG sites the upstream amplicon).

Amplicon	*N*	*Number of CpG analysed*	*Call rate*	*Exclusion criteria*
Proximal GRE site	367	13	97%	≥ 2 CpG units
First exon	349	23	98%	≥ 2 CpG units
Upstream	714	6	96%	≥ 3 CpG units

Detection of outliers

The outlier selection criteria was determined by the sample size (N = ~360) in each batch. By setting an arbitrary overall false positive rate at 5% (i.e. less than 5% chance of having at least one false positive observation), under the assumption of normal distribution of data, the corresponding significance level for each individual was calculated using the following equation:

$$1 - P^{N} = 0.05.$$

The probability (P) for having a true observation within the 5% false discovery rate in about 360 individuals would be

$$1 - P^{360} = 0.05$$

$$P = 0.999938.$$

The Z-score for P = 0.999938 is 3.49. This indicates that given the number of observations in each batch, there is less than 5% chance to observe at least one outlier within the range of +/-3.49SD. Therefore, data points outside +/-3.5SD were removed.

4.2.5 Statistical analyses

Regressions were conducted to study the associations between the DNA methylation of *PER1*, negative life events, VS BOLD responses during reward anticipation and substance use in the 14-year-old adolescents. Gender, experiment batch and recruitment sites were included as covariates. All statistical analyses were conducted in the SPSS (Version 20, IBM).

The regression model for negative life events is displayed below

Y_1 CpG = b_0 + b_1*(Frequency / Rating towards negative life events) + b_2*Gender + b_3*Batch + b_4*Sites + ε,

Where CpG represents the DNA methylation at individual CpG unit.

The regression models for alcohol and tobacco use are shown below.

Y_2 (Alcohol drinking - Drinkers *versus* Non-Drinkers)

OR = b_0 + b_1*CpG + b_2*Gender + b_3*Batch + b_4*Sites + ε.

Y_3 (Tobacco use - Smokers *versus* Non-Smokers)

Individuals' information on handedness was also included as covariate in analyses using VS BOLD responses. The regression models for the VS BOLD responses are shown below.

Y_4 (left VS BOLD responses)

OR $= b_0 + b_1$*CpG + b_2*Gender + b_3*Handedness

Y_5 (right VS BOLD responses) $+ b_4$* Batch + b_5*Sites + ε.

4.3 Results

4.3.1 *PER1* DNA methylation

After filtering the individuals who did not pass the quality control procedure, the DNA methylation data was used for subsequent analyses. DNA methylation is reported as a ratio ranging from 0 to 1, corresponding to 0% - 100% methylation. **Figure 4-5** and **Figure 4-6** displayed the DNA methylation of the amplified *PER1* promoter regions. The CpG units on the promoter CpG island, which was covered by the proximal GRE and the first exon amplicons, displayed 0 - 15% methylation. On the other hand, 18 - 95% methylation was observed in the CpG units in the *PER1* upstream amplicon, suggesting the *PER1* promoter displayed variations in DNA methylation.

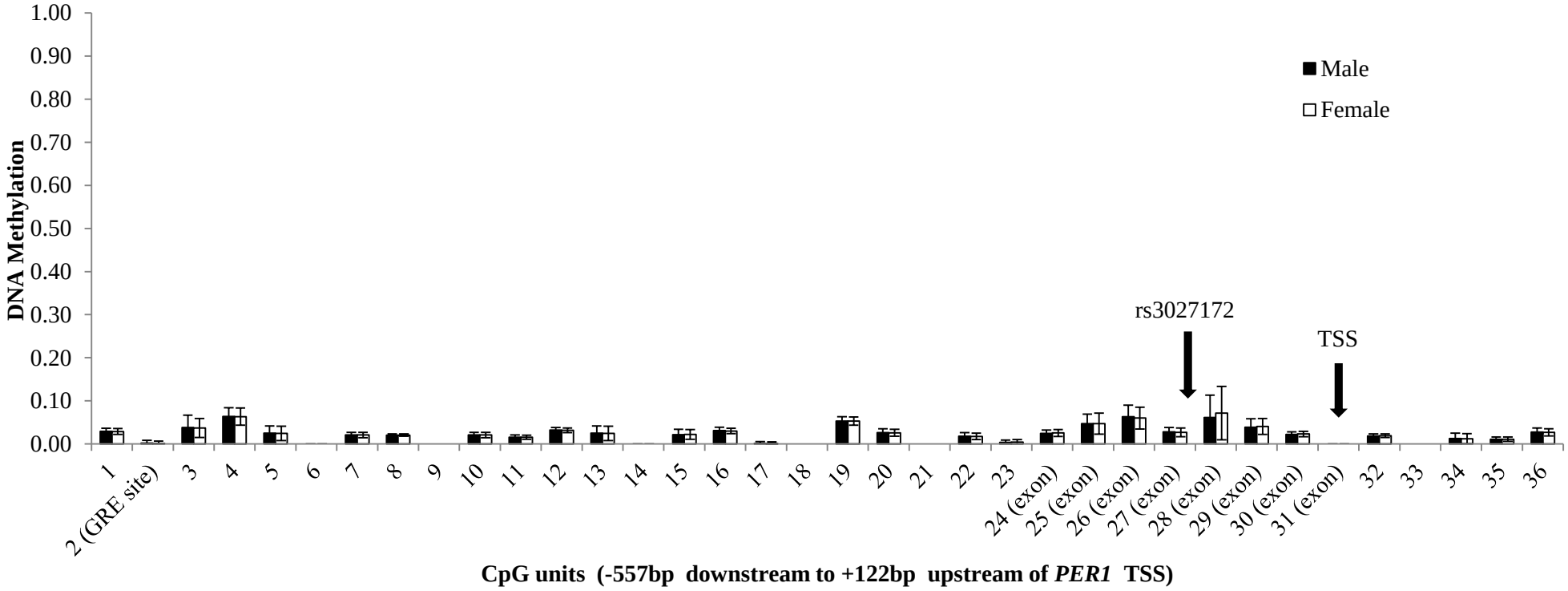


Figure 4-5 DNA methylation in *PER1* promoter CpG island (Mean ± SD). The DNA methylation of 36 CpG units in *PER1* promoter CpG island were analysed by the Sequenom© MALDI-TOF mass spectrometry. CpG units 1 - 15 were covered by the proximal GRE site amplicon (N = 160 in male, N = 197 in female) and CpG units 16-36 were covered by the first exon amplicon (N = 172 in male, N = 167 in female). CpG units 24 -31 are located within the first exon and rs3027172 is located at 30bp downstream of TSS (between CpG units 27 and 28). The DNA methylation at CpG units 9, 18 and 33 were not detectable.

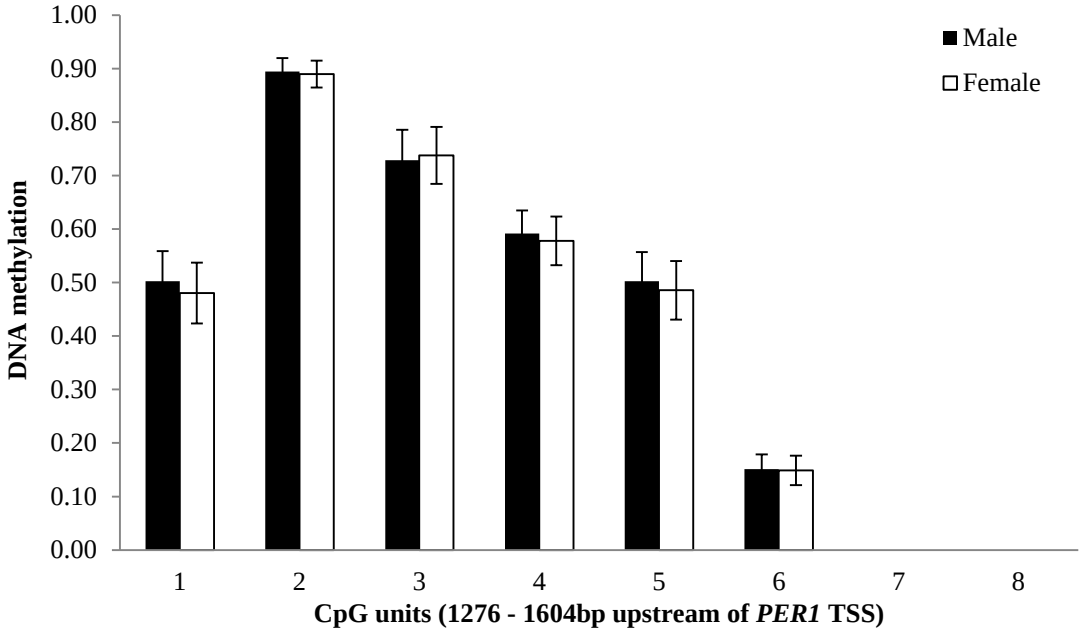


Figure 4-6 DNA methylation in *PER1* upstream (Mean ± SD). The DNA methylation of 8 CpG units located 1276-1604bp upstream of *PER1* TSS was analysed by the Sequenom© MALDI-TOF mass spectrometry (N = 330 in male, N = 355 in female). DNA methylation at CpG units 7 and 8 were not detectable.

DNA methylation at PER1 rs3027172

The presence of the minor allele at *PER1* rs3027172 creates an additional CG dinucleotide. In order to investigate the DNA methylation at *PER1* rs3027172, data obtained from the first exon amplicon was re-analysed by replacing the amplicon sequence with the minor allele of *PER1* rs3027172 (i.e. replacing the major allele T to minor allele C). Results from the EpiTyper showed that the additional CpG dinucleotide created by *PER1* rs3027172 was incorporated into CpG unit 29. The DNA methylation data of CpG unit 29 from 150 individuals were generated. Since CpG unit 29 was composed of 4 CpG dinucleotides including *PER1* rs3027172, the EpiTyper was unable to identify the DNA methylation status of *PER1* rs3027172 alone. Therefore, the additional DNA methylation data of CpG unit 29 is not presented in this chapter.

Of the 42 CpG units assayed in the *PER1* amplicons, only the six CpG units located at the upstream amplicon displayed moderate level of DNA methylation. Inter-individual differences in DNA methylation has also been observed at these CpG units, which might contribute to the variations in reward sensitivity and substance use behaviour in the adolescents. Therefore, all subsequent analyses were carried out on these six upstream CpG units. Since the CpG units in the proximal GRE site and the first exon amplicons were hypomethylated and displayed low inter-individual variations, the associations among these CpG units with the behavioural phenotypes were not analysed.

4.3.2 Putative transcription factor binding sites in the *PER1* upstream amplicon

In order to examine the functions of the six CpG units in *PER1* upstream amplicon, *in silico* analysis for putative transcription factor (TF) binding sites was performed. The putative TF analysis was performed using online software MatInspector http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html. The putative TF binding sites was identified based on the matrix similarity. A matrix similarity of 1 means the input sequence matches with the most conserved nucleotide at each position of the consensus matrix (i.e. the sequence of the TF binding site). According to Cartharius et al. (2005), a good match to the matrix has a matrix similarity of at least 0.80.

Results from the MatInspector predicted that 96 putative TF binding sites were located within the *PER1* upstream amplicon. However, only the putative binding sites for RNA polymerase II transcription factor IIB (consensus sequence 5'-CCGCGCC-3') and E-Box (consensus sequence 5'-CACGTG-3') displayed a matrix similarity of 1.00. The putative binding site for the RNA polymerase II transcription factor IIB was located 1431-1437bp upstream of the TSS and it overlapped with CpG unit 5 at the upstream amplicon. The putative binding site for the E-Box was located 1567 – 1579bp upstream of the TSS and it overlapped with CpG unit 8. In order to

examine the impact of DNA methylation at the *PER1* upstream amplicon, statistical analyses were carried out in each CpG unit.

4.3.3 DNA methylation in *PER1* and negative life events

Regression models were applied to study whether the DNA methylation of *PER1* could be influenced by life stress. Both lifetime occurrence and the ratings towards negative life events were used as the indicators of life stress in adolescents. The results showed that neither the frequency nor the ratings towards negative life events were associated with the DNA methylation in the *PER1* upstream CpG units (**Table 4-10**).

4.3.4 DNA methylation in *PER1* and reward sensitivity

Linear regressions were performed to study the relationship between *PER1* DNA methylation and the bilateral VS activations during reward anticipation. The results showed the DNA methylation in *PER1* upstream CpG units did not predict reward sensitivity in the adolescents

(

Table 4-11).

Table 4-10 DNA methylation at *PER1* upstream CpG units and negative life events. Regression models were applied to study the influence of life stress on the DNA methylation of *PER1* upstream CpG units. The main effects of DNA methylation at CpG units were corrected for gender, experiment batches and recruitment sites.

	Rating towards negative life events			***Frequency***		
	Beta (95% Confidence Interval)	***Wald Chi-Square***	***p***	***Beta (95% Confidence Interval)***	***Wald Chi-Square***	***p***
CpG 1	-.001 (-.002 - .001)	.943	.331	.001 (-.001 - .002)	.494	.482
CpG 2	-.0002 (.0002 - -.001)	.929	.335	.0001 (-.001 - .001)	.239	.625
CpG 3	-.0002 (-.001 - .001)	.135	.713	.001 (-.0003 - .002)	2.13	.145
CpG 4	-.001 (-.001 - .0003)	1.66	.198	.004 (-.001 - .002)	.319	.572
CpG 5	-.001 (-.002 - .001)	.882	.348	.0002 (-.001 - .001)	.078	.780
CpG 6	.0008 (-.0004 - .001)	.092	.762	-.0002 (-.001 - .001)	.306	.580

Table 4-11 DNA methylation in *PER1* upstream CpG units and reward sensitivity. Linear regression models were performed to study the relationship between DNA methylation in *PER1* upstream CpG units and the VS activations during reward anticipation in the MID fMRI task. The main effects of DNA methylation at CpG units on the VS activations were corrected for gender, handedness, experiment batches and recruitment sites.

	Left VS (MNI: -15 9 -9, 9mm)			***Right VS (MNI: 15 9 -9, 9mm)***		
	Beta (95% Confidence Interval)	***Wald Chi-Square***	***p***	***Beta (95% Confidence Interval)***	***Wald Chi-Square***	***p***
CpG 1	.297 (-.393 – .988)	.712	.399	.331 (-.394 - 1.06)	.800	.371
CpG 2	.717 (-.821 – 2.25)	.834	.361	1.00 (-.614 – 2.62)	1.48	.225
CpG 3	.440 (-.461 – 1.34)	.915	.339	.128 (-.821 – 1.08)	.069	.792
CpG 4	.052 (-.797 – .901)	.015	.904	.075 (-.817 – .968)	.027	.868
CpG 5	.207 (-.507 – .920)	.322	.570	.223 (-.527 – .972)	.340	.560
CpG 6	1.33 (-.111 – 2.77)	3.27	.070	1.39 (-.127 – 2.90)	3.22	.073

4.3.5 DNA methylation in *PER1* and substance use

4.3.5.1 Alcohol use

Logistic regressions were performed to study the association between *PER1* DNA methylation and alcohol use in adolescents. Individuals were classified as 'drinkers' or 'non-drinkers' based on the number of alcohol drinking occasions in the past 12 months. The results showed that the DNA methylation status at the upstream CpG units was not associated with alcohol use in the adolescents (**Table 4-12**).

4.3.5.2 Tobacco use

Logistic regressions were performed to study the association between *PER1* DNA methylation and tobacco use in the adolescents. The individuals were classified as 'tobacco users' or 'non-tobacco users' based on their lifetime cigarette smoking occasions. The results showed the DNA methylation at CpG unit 1 was significantly associated with lifetime tobacco use (p = .032) at nominal level. Such association was not driven by the effects of gender, experimental batches or recruitment sites. However, the significance of this association did not survive the correction for multiple testing. The DNA methylation at CpG units 2, 3, 4, 5 and 6 was not associated with tobacco use in these individuals (**Table 4-12**).

Table 4-12 DNA methylation at *PER1* upstream CpG units and substance use. Logistic regression models were performed to study the associations between DNA methylation in *PER1* upstream CpG units and substance use. The main effects of DNA methylation at CpG units on alcohol and tobacco use were corrected for gender, experiment batches and recruitment sites.

	Alcohol Use (Drinkers versus Non-Drinkers)			***Tobacco use (Smokers versus Non-Smokers)***		
	Beta (95% Confidence Interval)	***Wald Chi-Square***	***p***	***Beta (95% Confidence Interval)***	***Wald Chi-Square***	***p***
CpG 1	-1.21 (-4.05 – 1.62)	.702	.402	-3.92 (-7.50 - -.342)	4.61	.032
CpG 2	3.77 (-3.08 – 10.6)	1.16	.281	7.81 (-.840 – 16.5)	3.13	.077
CpG 3	.819 (-3.25 – 4.88)	.156	.693	.347 (-4.93 – 5.62)	.017	.897
CpG 4	-.245 (-4.01 – 3.52)	.016	.898	.047 (-4.82 – 4.92)	3.61×10^{-4}	.985
CpG 5	-2.02 (-5.12 – 1.01)	1.62	.203	-2.57 (-6.57 – 1.44)	1.58	.209
CpG 6	2.78 (-3.43 – 9.00)	.770	.380	1.91 (-6.13 – 9.95)	.216	.642

4.4 Discussion

Alcohol and tobacco use during adolescence is influenced by alterations of the circadian system (e.g. (Spanagel, 2009, Heikkinen et al., 2009, Dong et al., 2011). In this chapter, the DNA methylation landscape of *PER1* promoter region (~1kb) was analysed on the Sequenom™ MALDI-TOF mass spectrometry platform. Whilst the CpG units located upstream of the TSS displayed variations in DNA methylation across individuals, region around the first exon that overlapped with the CpG island was largely hypomethylated. Statistical analyses indicated no significant associations between the DNA methylation at *PER1* upstream with negative life events, reward sensitivity and substance use in adolescents.

4.4.1 DNA methylation at *PER1* promoter

4.4.1.1 Hypomethylation at the first exon and the promoter CpG island

As illustrated in this study, the *PER1* promoter displayed wide range of DNA methylation. It might be possible that the hypomethylation observed at the first exon and the promoter CpG island of *PER1* did not happen by chance but instead, could be of evolutionarily significance. In mammals, over 60% genes contain at least one CpG islands overlapping with the promoter region and transcription start site (Antequera, 2003). Despite the abundance of CpGs in the promoter CpG islands, most of them are hypomethylated (Weber et al., 2007, Suzuki and Bird, 2008). Computational analyses revealed that the DNA methylation status at CpG islands could be dependent on their locations in the genome (Cocozza et al., 2011). Cocozza et al (2011) studied the DNA methylation data of over 28,000 CpG islands obtained by the Encyclopedia of DNA Elements (ENCODE) Consortium and found CpG islands at the promoters displayed the lowest DNA methylation compared to those located in the 3'-untranslated and intragenic regions. Hypomethylation at the *PER1* promoter CpG islands might be important for regulating gene functions (Antequera, 2003, Cocozza et al., 2011), possibly by protecting transcriptional activities from disruptions caused by DNA methylation.

4.4.1.2 DNA methylation at *PER1* upstream amplicon

Of all the CpG units examined, only the six CpG units located at *PER1* upstream amplicon demonstrated inter-individual differences in DNA methylation. The selection of target region was based on the previous studies, in which inter-individual differences in DNA methylation at the same region was found to associate with the disease status of cancers as well as the mRNA expression of *PER1* (Hsu et al., 2007, Gery et al., 2007). Differential DNA methylation was frequently observed in CpG shores in the brain, liver and spleen tissues, in which the CpG shores were defined as the genomic regions stretch up to 2kb from both sides of the CpG islands (Irizarry et al., 2009). Consistent with Irizarry et al's (2009) findings, the *PER1* upstream amplicon (about 1kb upstream of the *PER1* promoter CpG island) displayed medium to high DNA methylation, compared to the low DNA methylation observed in the first exon/ promoter CpG island in the IMAGEN adolescents.

4.4.2 DNA methylation in *PER1*, life stress, reward sensitivity and substance use

4.4.2.1 Hypomethylation at the *PER1* proximal GRE site

The proximal GRE site located outside the first exon was not methylated among the adolescents. Such observation could be due to various reasons. Firstly, since the GRE site was found to be highly conserved across species (So et al., 2007, Yamamoto et al., 2005), the absence of DNA methylation at the *PER1* proximal GRE site could be of evolutionarily and functional significance. Based on the speculative view of which the DNA methylation alters the binding affinity of transcriptional factor binding sites (Watt and Molloy, 1988), the unmethylated *PER1* proximal GRE site potentially may allow the binding of GR to initiate stress responses during stressful situations.

Secondly, as pointed out by Yamamoto et al (2005), the proximal GRE site on the *mPER1* promoter might not be functional, based on the observation that the mutated proximal GRE site did not completely abolish the induction of *Per1* mRNA expression upon dexamethasone

treatment in mice. Therefore DNA methylation at the proximal GRE site in *PER1* might not be susceptible to stress. The abundance of putative TF binding sites at the *PER1* promoter further suggested the complex interactions among the gene regulatory elements and DNA methylation along the promoter. Such interaction might explain the mixed findings in Hsu et al (2007) and Gery et al's (2007) studies, where DNA methylation was found to associate with both upregulation and downregulation of *PER1* mRNA expression.

4.4.2.2 *PER1* DNA methylation was not influenced by life stress and did not predict reward sensitivity and substance use

Neither the frequency nor the ratings towards negative life events was found to influence DNA methylation in the six *PER1* upstream CpG units. The results suggested that the DNA methylation at these CpG units were not altered by life stress. Nevertheless, the DNA methylation at these CpG units had no effect on reward sensitivity or substance use in the adolescents. Such findings could be constrained by the selection of target region as well as the potential tissue-specific variations in DNA methylation. As demonstrated by previous studies, the expression of *Per1* was shown to influence reward sensitivity (Abarca et al., 2002, Li et al., 2008). Recent evidence also suggested alterations in DNA methylation at the VTA in response to reward stimuli, indicating the role of DNA methylation in the activations of dopaminergic neurons and reward-related learning behaviour (Day et al., 2013).

Due to the increase influence in life stress on substance use in the IMAGEN adolescents (see Chapter 3), the possibility that stress-induced DNA methylation at *PER1* can predict reward sensitivity and substance use at later stage of adolescence or adulthood cannot be excluded. It is also possible that the DNA methylation in *PER1* may have a larger impact on reward sensitivity, alcohol drinking and smoking in clinical populations compared to adolescents who had just developed alcohol drinking and smoking behaviour. Since the associations of life stress, reward sensitivity and substance use was demonstrated in the IMAGEN population (see

Chapter 3), interactions among these factors may be observable at genetic and neurotransmission levels as well as in alternative forms of epigenetic modifications.

4.4.3 Additional factors influencing the DNA methylation in *PER1*

4.4.3.1 Tissue- and gender-specific DNA methylation

Emerging evidence has indicated differential DNA methylation pattern in the brain between patients and healthy individuals in human brain. When comparing the epigenetic and whole transcriptome gene expression profiles of basolateral amygdala, central nucleus of amygdala and superior frontal cortex from 17 individuals suffered from chronic alcohol abuse and 15 matched control cases, Ponomarev and colleagues (2012) discovered that alterations of gene transcripts could be dependent on the GC-content of gene modules. Alcohol abuse might result in the up-regulations of genes from GC-rich modules and long-terminal repeat retrotransposons, but downregulation of genes from the GC-poor modules (Ponomarev et al., 2012), confirming that gene expressions in human brain could be regulated in response to environmental input such as heavy alcohol use or abuse. Analyses also revealed differential gene expression profiles across cell types, with downregulation of transcripts from neuronal modules and upregulation of transcripts in microglia modules observed among the alcoholics (Ponomarev et al., 2012). In contrast, Manzardo and colleagues (2012) investigated the global DNA methylation in the frontal cortex of 10 alcoholics and 10 age and gender-matched controls. Their results showed no significance difference in global DNA methylation; with over 80% genes displayed similar peak score values between two groups (Manzardo et al., 2012).

It is unclear to what extent the observed DNA methylation in peripheral blood reflects the DNA methylation pattern in brain tissues. This might explain the lack of association between *PER1* DNA methylation and the VS activation during reward anticipation in this study. As illustrated by the bisulfite DNA sequencing data from Eckhardt and colleagues (Eckhardt et al., 2006), differences in DNA methylation of 12 human tissue types ranged from 5% to 20%. Although

the comparison between the DNA methylation in brain and whole blood in the IMAGEN participants could not be made, the use of publicly available database such as the BrainCloudMethyl (Numata et al., 2012) might enhance researchers' understanding to the DNA methylation in human brain. Demonstrated by Iwamoto et al.'s (2011) study on prefrontal cortex in human, neuronal cells exhibited larger inter-individual differences and lower global DNA methylation compared to non-neuronal cells. Despite the DNA methylation pattern in neurons and non-neurons achieved a correlation of over 0.86, the methylated regions in neurons and non-neurons were associated with different functions and contained different sets of transcriptional factor binding sites (Iwamoto et al., 2011). Based on the current literature, it is expected that differential DNA methylation can occur across cell types in both blood and brain tissues. As a result, one major limitation of the current study is the absence of correction of cell heterogeneity in whole blood DNA samples. Such limitation also applies to the majority of published studies, as the DNA methylation status across cell types were not corrected prior to analyses.

In addition, larger differences in DNA methylation were found across tissue types compared to gender and age (Eckhardt et al., 2006). Confirmed by this study, the *PER1* DNA methylation data of IMAGEN participants displayed very little gender differences, with the maximum difference of 2.23% observed at CpG unit 1 in the *PER1* upstream amplicon.

4.4.3.2 Histone modifications

The consequences of DNA methylation on *PER1* can be more complex than previously thought. Interestingly, DNA methylation at promoter CpG islands appears to be modified by the structure of chromatin. In human, the core component of chromatin is the nucleosome, which comprises of 147bp DNA wrapped around histones (H3, H4, H2A and H2B). Dimethylated H3K4 (H3K4me2) was found in hypomethylated promoter regions and CpG islands (Weber et al., 2007). The reduction of H3K4me2 was accompanied by the methylation of CpG islands and

promoters (Weber et al., 2007). Similarly, the level of trimethylated H3K4 was negatively correlated with the level of DNA methylation at the intragenic CpG islands (Maunakea et al., 2010). It can be envisaged that the DNA methylation in *PER1* covaries with the extent to which the histones are modified. As previously mentioned, *CLOCK* protein is a histone acetyltransferase (Doi et al., 2006). This has raised the question of whether histone modifications of *CLOCK* and other circadian components can influence the downstream targets such as *PER1*. Future investigation on the interactions between histone modifications and DNA methylation will enable better understanding of epigenetic modifications in *PER1*.

4.4.4 Statistical and technical concerns

Similar to other univariate analyses such as genome-wide association of common variants, having a large sample size is crucial for reducing chances of false discovery arise from multiple testing. Due to various constraints, most epigenetic studies were conducted with sample sizes of about 50-100 individuals. The large sample adopted in this study can be advantageous for studying DNA methylation since it increases the sensitivity and power for analyses. To ensure the quality of the DNA methylation data in *PER1*, a careful quality procedure was applied in this study. The selection of outliers was determined by the observed data range and sample size. The advantage of selecting outliers in a data-driven manner is that it ensures only the data outside the 5% arbitrary false discovery rate are removed, hence minimising chances of excluding any true observations. Nonetheless, this study is one of the large-scale projects investigating the DNA methylation in humans.

The Sequenom MALDI-TOF MS is one of the best available methods to detect DNA methylation of *PER1*. The Sequenom platform enables high-throughput detection of DNA methylation in an accurate and economical way (Coolen et al., 2007). Due to the limitation of the technology, each assayed region has to be restricted to ~500bp. Therefore the Sequenom platform is suitable for candidate gene studies with specific target regions. However, restricting

the size of target regions can introduce biases and limit researchers' understanding to the functional significance of DNA methylation. The importance for the DNA methylation in promoters, CpG islands, CpG shores, gene bodies and intragenic regions in regulating transcriptional activities and behaviour has been suggested (Irizarry et al., 2009, Maunakea et al., 2010). Platforms for studying epigenome-wide DNA methylation have become increasingly available in recent years and they will provide researchers a better understanding to the functional significance of CpG loci.

The resolution of the Sequenom platform can be an issue, as illustrated in the *PER1* data presented in this chapter. The detection of DNA methylation in the Sequenom platform heavily relies on the size of fragments produced during the T-cleavage reaction. Fragments that are out of the detection range or having the same mass cannot be identified in the mass spectra. This could explain the missing data at the E-Box (CpG unit 8) and CpG unit 7 at the *PER1* upstream amplicon since these fragments shared the same mass. It is also unlikely to detect the DNA methylation at a single CpG dinucleotide when a cleaved fragment contains multiple CpG dinucleotides that are adjacent to each other. Due to the high CG content at the *PER1* promoter CpG island, it is difficult to identify DNA methylation at individual CpG dinucleotide. Since rs3027172 is located in close proximity to three CpG dinucleotides, only the overall DNA methylation across CpG dinucleotides was reported by the Sequenom. Therefore, the Sequenom platform may not be ideal to investigate the DNA methylation at rs3027172. Alternative approach was adopted to further analyse the DNA methylation at the E-Box and rs3027172. Using the single nucleotide extension technology available at the Applied Biosystems SNaPshot platform, the methylated cytosine in bisulfite converted DNA produces a mixed signal of C and T. Hence the DNA methylation at the cytosine can be calculated from the C/T ratio. However, due to the difficulties in optimising the SnaPshot primers, I failed to complete the DNA methylation analysis at the E-box and rs3027172.

Other types of DNA methylation such as 5-hydroxymethylcytosine (5hmeC) cannot be distinguished from the signals of 5meC in the Sequenom platform. This problem is not limited to the Sequenom but also found in other DNA methylation platforms that require amplifications or sequencing of bisulfite DNA. Interestingly, 5meC can be converted into 5hmeC by enzymes from the ten-eleven translocation family (Tahiliani et al., 2009). The conversion of 5meC to 5hmeC further promotes the demethylation of cytosine (Liutkeviciute et al., 2009), suggesting that the possibility that 5meC and 5hmeC may not share the same role in gene regulations. Despite all the limitations, the current technology has enabled researchers to discover the epigenome. Techniques with improved resolution will enhance researchers' understanding to human epigenome in the future.

4.4.5 Conclusions

In this chapter the DNA methylation profile of *PER1* was shown to vary within the promoter and across 714 individuals. Due to the limitation of the Sequenom platform, the selection of target regions on *PER1* might introduce biases in studying the impact of DNA methylation at the CpG units on substance use. Therefore, the lack of associations between *PER1* DNA methylation in the selected region with life stress, reward processing and substance use may not necessarily reflect the functional significance of the DNA methylation at *PER1*. In searching the functional significance of DNA methylation in *PER1* in substance use, this study has put forward more questions in identifying functional epigenetic loci in substance use for future investigations. Increasing knowledge in the interactions between DNA methylation markers and other epigenetic modifications such as 5-hydroxymethylation and histone modifications has aided the discovery of functional epigenetic markers. This study is one of the pioneer studies investigating the influence of DNA methylation on substance use using a large sample. The research of epigenetic modifications in circadian rhythm is still at its early stage, adopting large scale or genome-wide approaches will be beneficial in answering our research question in a comprehensive way.

Chapter 5

Genetics of the stress and circadian systems on reward sensitivity, alcohol and tobacco use in adolescents

5.1 The genetics of alcohol and tobacco addictions

Recent advances in technology have aided the investigation of genetic epidemiology and molecular genetics of complex diseases. Several genome-wide association studies (GWAs) have successfully identified risk loci for alcohol and nicotine addictions (Schumann et al., 2011, Bierut et al., 2010, Tobacco and Genetics, 2010, Thorgeirsson et al., 2010). Although the heritability estimates of alcohol and nicotine addictions range from 30 - 70% across various populations (Broms et al., 2007, Kendler et al., 1999, Hansell et al., 2008, Goldman, 1993), the genetic risk variants have remained sparse in the published literature and they did not fully account for the heritability of the addictions.

One possible cause of missing heritability is the genotyping platform used for studying the genetics of addictions. According to the common-disease common-variant (CDCV) hypothesis, most common diseases are caused by common risk variants that individually has small effect on the diseases. Early whole-genome genotyping platforms were designed to detect associations between common SNP markers and complex traits in unrelated individuals. The minor allele frequency (MAF) of the common SNPs is defined as $\geq$ 1% and the SNPs are tagged through linkage disequilibrium (LD) (Reich et al., 2001, Sullivan et al., 2012). Hence other forms of genetic variations such as SNPs with low MAFs, copy number variations and gene-gene interactions cannot be captured in the GWAs.

Another reason that can explain the missing heritability is the lack of statistical power for detecting genetic risk variants. The detection of risk variants in GWAs depends on both the effect sizes (i.e. penetrance) and the MAFs of SNPs. Illustrated by Wang et al (2005), a larger sample size will be required to detect risk variants with lower MAFs and smaller effect sizes. For example, at least 20000 individuals are required to detect a risk variant with MAF below 0.1 and an odd ratio of 1.3. However, only 5000 individuals are required to detect a risk variant with an odd ratio of 2 and MAF below 0.1 (Wang et al., 2005). The extremely large sample size required for GWAs prevents the identification of risk variants in complex disorders.

An alternative way to circumvent the low phenotypic variance explained by individual SNPs is by investigating the aggregate genetic effect of risk variants. The genetic contribution of complex traits, including alcohol and tobacco use, is thought to be additive and polygenic (Sullivan et al., 2012, Balding, 2006). Additive genetic effect states that the contribution of a SNP to a complex trait increases with number of copies of the risk allele in a linear manner. Therefore, increasing the number of risk alleles will proportionally increase the risk of developing complex disorders and traits (Plomin et al., 2009). On this basis, the polygenic model developed by the International Schizophrenia Consortium (2009) successfully improved the predictive heritability of schizophrenia. By applying an arbitrary threshold to identify the genetic risks associated with schizophrenia, the risk alleles *en masse* explained about 3% variance in the target populations (International Schizophrenia et al., 2009). The same approach was adopted by Kos et al (2013), in which the common SNPs could explain 0.73% and 2.14% variance in the African American and European American alcohol dependent populations.

There has been increasing interest to understand the functions of genes in networks. It can be envisaged that genes sharing similar functions converge to common pathways and lead to the development of complex disorders (Hardy and Singleton, 2009). In contrast to the GWAs where no *a priori* hypothesis is required, targeting SNPs on genes with known biological functions can

reduce the number of independent tests and therefore increase chances to identify risk variants that are associated with the complex disorders and traits. The selection of genes, however, can be challenging and prone to biases. This applies to the research in psychiatric genetics in which many risk variants are located in genes involved in neurotransmissions. Several association studies were designed to reduce gene selection bias whilst integrating the existing knowledge of psychiatric disorders at the same time. Sun et al (2009) prioritised 4062 candidate genes associated with schizophrenia using data from various source including literature search, GWAs, linkage analyses, gene expression microarray analyses. Using a weighting approach, 33 genes were identified as core genes associated with schizophrenia and 502 genes were prioritised for follow-up analyses (Sun et al., 2009). By assigning genes according to the pre-defined biological functions listed in the Kyoto Encyclopedia of Genes and Genomes (KEGG), gene-set analysis revealed several pathways such as 'synthesis and degradation of ketone bodies (KEGG ID:72)' and 'neuroactive ligand-receptor interaction (KEGG ID: 4080)' were associated with alcohol dependence (Biernacka et al., 2013). Although the gene-set analysis approach did not enable testing of individual SNPs located in the pathways, single SNP associations were then conducted to identify top SNPs associated with alcohol dependence (Biernacka et al., 2013).

As discussed in Chapter 1, the regulations of stress and circadian rhythm rely on the input from multiple genes. Several genetic polymorphisms in the circadian and stress genes, such as *PER1*, *PER2, VIP,* and *CRHR1* have been identified to associate with risks of substance dependence and addictions (Dong et al., 2011, Spanagel et al., 2005, Kovanen et al., 2010, Blomeyer et al., 2008). Nevertheless, the genetic polymorphisms on *PER2* and *CLOCK* were found to associate with the sensitivity to reward stimuli (Forbes et al., 2012, Tsuchimine et al., 2013). Since the initiation of substance use can be predicted by sensitivity to reward in adolescents (see Chapter 3), the genetic polymorphisms within stress and circadian systems can be the risk factors for substance use and addictions.

5.1.1 Aims and hypotheses

This chapter aims to identify the additive genetic effects of stress and circadian genes on substance use in adolescents using the single SNP association and polygenic score analyses. By adopting the mass candidate gene approach, it is hypothesised that common SNPs of genes from **a)** the stress system; **b)** the circadian system and **c)** both the stress and circadian systems are associated with alcohol and tobacco use in adolescents at age 14 and age 16. The genetic influence of the stress and circadian systems on reward sensitivity as risk factor of substance use will be studied. The second part of this chapter aims to estimate the phenotypic variances for reward sensitivity and substance use that are attributable to the polygenic effects of the stress and circadian systems. Due to the additive effect of SNPs, the polygenic influences from the stress and circadian systems are expected to account for a larger proportion of phenotypic variances than an individual SNP.

5.2 Methods

5.2.1 Participants and phenotypes

The genome-wide genetic data from N = 1982 individuals were used for genetic association analyses (see **Chapter 2 Section 2.4** for details). Of these 1982 individuals, N = 1857 individuals responded to the past 12 month alcohol use question in the European School Survey Project on Alcohol and Other Drugs questionnaire (ESPAD, See **Chapter 2 Section 2.2.4**) (Hibell et al., 1997) at age 14; N = 1250 individuals responded to the past 12 month alcohol use question and N = 1350 individuals responded to lifetime tobacco use question in the ESPAD at age 16 (see **Table 5-1**). After applying the quality control procedures, the Monetary Incentive Delay (MID, See **Chapter 2 Section 2.3.5**) fMRI data from N = 1282 adolescents was used for analyses in this chapter (N = 613 Males, N = 669 Females, N = 1117 right-handed, N = 165 left-handed or ambidextrous).

Table 5-1 ESPAD information used for genetic association analyses.

	Total: N (%)	*Males*	*Females*
Past 12 months alcohol use at age 14			
Non-Drinkers	1183 (63.7)	595	588
Drinkers	674 (36.3)	314	360
Past 12 months alcohol use at age 16			
Non-Drinkers	254 (20.3)	120	134
Drinkers	996 (79.7)	471	525
Lifetime tobacco use at age 16			
Non-Smokers	868 (64.3)	430	438
Smokers	482 (35.7)	211	271

5.2.2 Identification of candidate genes

A systematic search of candidate genes related to regulations of stress and circadian rhythm was performed on the National Centre for Biotechnology Information (NCBI) Gene database in April 2012 (http://www.ncbi.nlm.nih.gov/gene).

Stress gene list (N = 2348)

The NCBI Gene database was screened on search terms 'stress', 'psychosocial stress', 'stress HPA axis', 'HPA axis' and 'life events stress' to create a list of candidate genes involved in the regulation of stress and psychosocial stress. Based on all published literature and sequence information available until April 2012, 55493 genes in 'stress', 25 genes in 'psychological stress', 53 genes in 'stress HPA axis', 66 genes in 'HPA axis' and 165 genes in 'life events

stress' were identified. N = 2348 human stress genes were obtained by filtering all non-mammalian vertebrate genes and duplicates (See **Appendix 5: Stress genes**).

Circadian gene list (N = 335)

To identify candidate genes involved in the regulation of circadian rhythm, keywords including 'circadian rhythm, 'clock', 'circadian clock' and 'circadian cycle' were chosen to screen the NCBI Gene database. The NCBI database found 1763 genes in 'circadian rhythm', 1533 genes in 'clock', 1331 genes in 'circadian clock', 396 genes in 'circadian cycle. A total of 335 human circadian genes were obtained by filtering all non-mammalian vertebrate genes and duplicates (See **Appendix 6: Circadian genes**).

Common/ 'Overlap' genes in the stress and circadian gene lists (N =186)

To identify genes that might involve in both the regulations of stress and circadian rhythm, genes appeared in both the stress and circadian gene lists were extracted. The resulting gene list contained 186 genes that were appeared in both lists (See **Appendix 7: Overlap stress x circadian genes**).

Neurotransmitter genes in the 'overlap' stress and circadian gene list (N =18)

To further distinguish genes coding for neurotransmitters in the gene lists, a gene list consisted of all neurotransmitters in human was created by entering keyword 'neurotransmitters' in the NCBI Gene database. The resulting human neurotransmitter gene list contained 1236 genes, of which 18 neurotransmitter genes were appeared in the overlapping stress and circadian gene list (See **Figure 5-1** for the summary of gene lists and **Appendix 8: Neurotransmitter stress x circadian genes**).

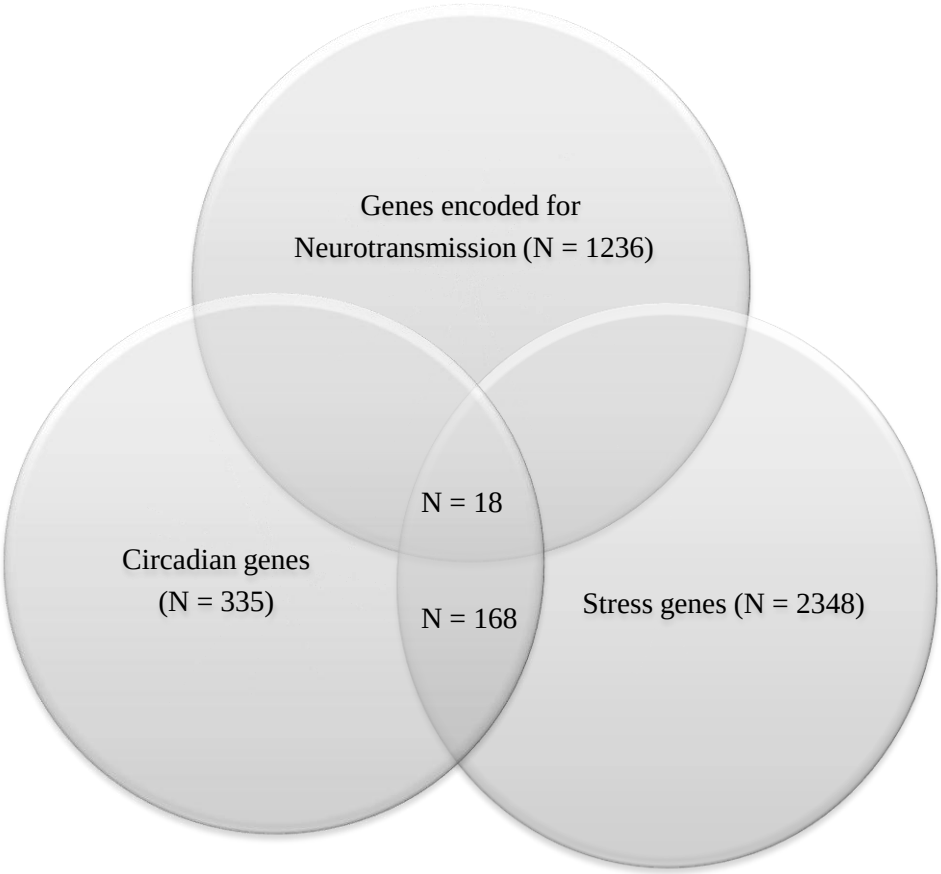


Figure 5-1 Summary for the stress and circadian gene lists. The NCBI Gene database was screened in April 2012 to identify autosomal candidate genes involved in the regulations of stress and circadian rhythm. After filtering the non-mammalian vertebrate genes and duplicates, the resulting stress and circadian gene lists contained 2348 and 335 genes respectively. N = 186 genes were appeared in both the stress and circadian gene lists, of which N = 18 were genes coding for neurotransmissions.

5.2.3 Functional annotation of stress and circadian genes

In order to characterise the biological functions of the identified candidate genes, both the stress and circadian gene lists were uploaded to the WebGestalt (WEB-based Gene SeT AnaLysis Toolkit, http://genereg.ornl.gov/webgestalt/). WebGestalt is a web-based data mining system that allows exploration and enrichment of gene sets according to the functional features shared by the genes (Zhang et al., 2005). The Gene Ontology (GO) terms were selected to define the functional feature of genes. The GO is a collection of controlled vocabulary of terms describing

the biological functions of gene products. The GO terms are organised into three independent domains, namely the biological process, cellular component and molecular function (Ashburner et al., 2000).

5.2.4 Single SNP association analyses and quality control

In order to maximise the coverage of gene regulatory region, SNPs located within +/- 10kb of the autosomal genes were selected for association analyses. For genes that exist in multiple versions, the longest versions of genes were used for SNP extraction. The genetic information of SNPs in the four gene lists (stress, circadian, genes common or 'overlap' in the stress and circadian lists, and neurotransmitters in the overlap gene list) were extracted from the genome-wide genetic data using PLINK v1.07 (Purcell et al., 2007). The command '--missing' was used to assess the quality of the genetic information in each gene list. The overall genotyping rates were > 99% in all gene lists.

Linear regressions were applied to study the additive genetic effects of SNPs in the VS activations during reward anticipation, past 12 months alcohol use (age 14 and 16) and lifetime tobacco use at age 16 in the adolescents. Performing genetic association analyses using linear models would allow greater tolerance of the residual terms and account for the additive genetic effect of SNPs at the same time. The same strategy was adopted by Dudbridge (2013) and Wray et al (2010) to analyse the binary phenotypes. All associations were controlled for gender and sites of recruitment. Handedness was included as an additional covariate in the association analyses using the VS activations. The linear regression models of substance use phenotypes were in the form

$$Y_n = b_0 + b_1*\text{G} + b_2*\text{Gender} + b_3*\text{ Sites} + \varepsilon,$$

of which Y_n refers to the substance use phenotypes and G refers to '0', '1' or '2'copies of the minor allele of each SNP. Similarly, the linear regression models of the VS activations during reward anticipation were in the form

$$\text{VS activations} = b_0 + b_1\text{*G} + b_2\text{*Gender} + b_3\text{* Handedness} + b_4\text{* Sites} + \varepsilon.$$

For each genetic association analysis, the Benjamini-Hochberg false discovery rate (FDR-BH) control was applied to correct for multiple testing. After performing the analyses, quantile-quantile (Q-Q) *p*-value plots were produced in R to examine the quality of the association statistics. The Q-Q *p*-value plots display the distribution of the observed –log10 *p*-values against the distribution of the expected (null) –log10 *p*-values. Any systematic deviation of the observed –log10 *p*-values from the null distribution implicates population stratification and hidden individual relatedness in the genetic data (Balding, 2006). The genomic inflation factor λ was calculated in PLINK and it quantified the systematic deviation of associations due to population structure.

5.2.5 Polygenic score analysis

The polygenic score analysis aimed to identify the additive genetic effect of common SNPs *en masse* of the complex disease traits. The analysis involved selecting the genetic risk variants in the discovery sample, followed by calculating the weighted score of the associated variants in each subject in an independent target sample. Association between the weighted scores and a trait implies significant genetic basis for the disease trait. To facilitate the polygenic score analyses and subsequent validations, individuals' genetic information was organised into two lists (N = 712 individuals in List 1 and N = 1446 individuals in List 2).

SNP pruning

The SNPs from the four gene lists described in **Section 5.2.2** were pruned based on a pairwise r^2 threshold of 0.5 and within a 5-SNP sliding window in a window size of 50 SNPs. Pruning

removes SNPs that are in strong linkage disequilibrium (LD, $r^2 > 0.5$) and retains relatively independent SNPs for calculating the polygenic scores of individuals See **Table 5-2** for information of the resulting SNP set in the gene lists.

Table 5-2 Coverage of SNPs in the gene lists. The SNPs in each gene list were pairwise-pruned to obtain an independent SNP set.

Gene Lists	***Unpruned SNPs (N)***	***Pruned SNPs (N)***	***Pruned/ Unpruned SNP ratio***
Stress (N = 2348)	41307	21708	0.53
Circadian (N = 335)	5318	2856	0.54
Overlap in stress x circadian genes (N = 186)	2953	1660	0.56
Neurotransmitters in stress x circadian genes (N = 18)	439	225	0.51

Calculating polygenic scores in the discovery sample

Linear regressions were conducted in the discovery samples, using gender and recruitment sites as covariates. Information on handedness was also included as a covariate when analysing the VS activations. Using the results of the association, SNPs were selected based on the liberal *p*-value threshold $p < 0.5$ to calculate the polygenic scores. The polygenic score of each individual in the target sample was obtained by averaging the weighted score of the selected risk alleles, in which the weighted score of each SNP was calculated by multiplying copies of the minor allele to the corresponding β value. The scoring was performed using the '--score' command in PLINK.

Estimation of the phenotypic variance explained by polygenic scores in the target sample

Semi-partial correlation was applied to estimate the phenotypic variance explained by the polygenic scores of the target sample. The phenotype of individuals from the target sample was controlled for gender and recruitment sites; subsequently it was used to correlate with the polygenic scores. One-tailed correlation was used since positive correlation between the adjusted phenotype and the polygenic scores was expected. The squared correlation coefficient (R^2) was used to indicate the proportion of phenotypic variance explained by the polygenic scores.

Validation of the results

The calculation of polygenic score was repeated twice by using both individual lists as target samples. Individuals in List 1 were treated as the discovery sample to calculate the polygenic scores of individuals in List 2 and *vice versa*.

5.3 Results

5.3.1 Functional annotations of candidate genes

Using all identified human genes as the reference gene list, the stress and circadian gene lists were enriched with Gene Ontology (GO) terms using hypergeometric statistical test in the WebGestalt. The results showed 'GO: 0006950 response to stress' was the most enriched GO term in the stress gene list, with a ratio of enrichment = 2.90 (Adjusted $p = 1.17 \times 10^{-212}$, N = 761 genes, see **Figure 5-2**). In the circadian gene list, the GO term 'GO:0007623 circadian rhythm' was the most enriched with a ratio of enrichment = 43.54 (Adjusted $p = 6.71 \times 10^{-131}$, N = 103 genes, see **Figure 5-3**).

Figure 5-2 Functional annotation of the stress genes using WebGestalt. The directed acyclic graph (DAG) displayed the predicted GO functions under biological process and molecular function of the stress gene list.

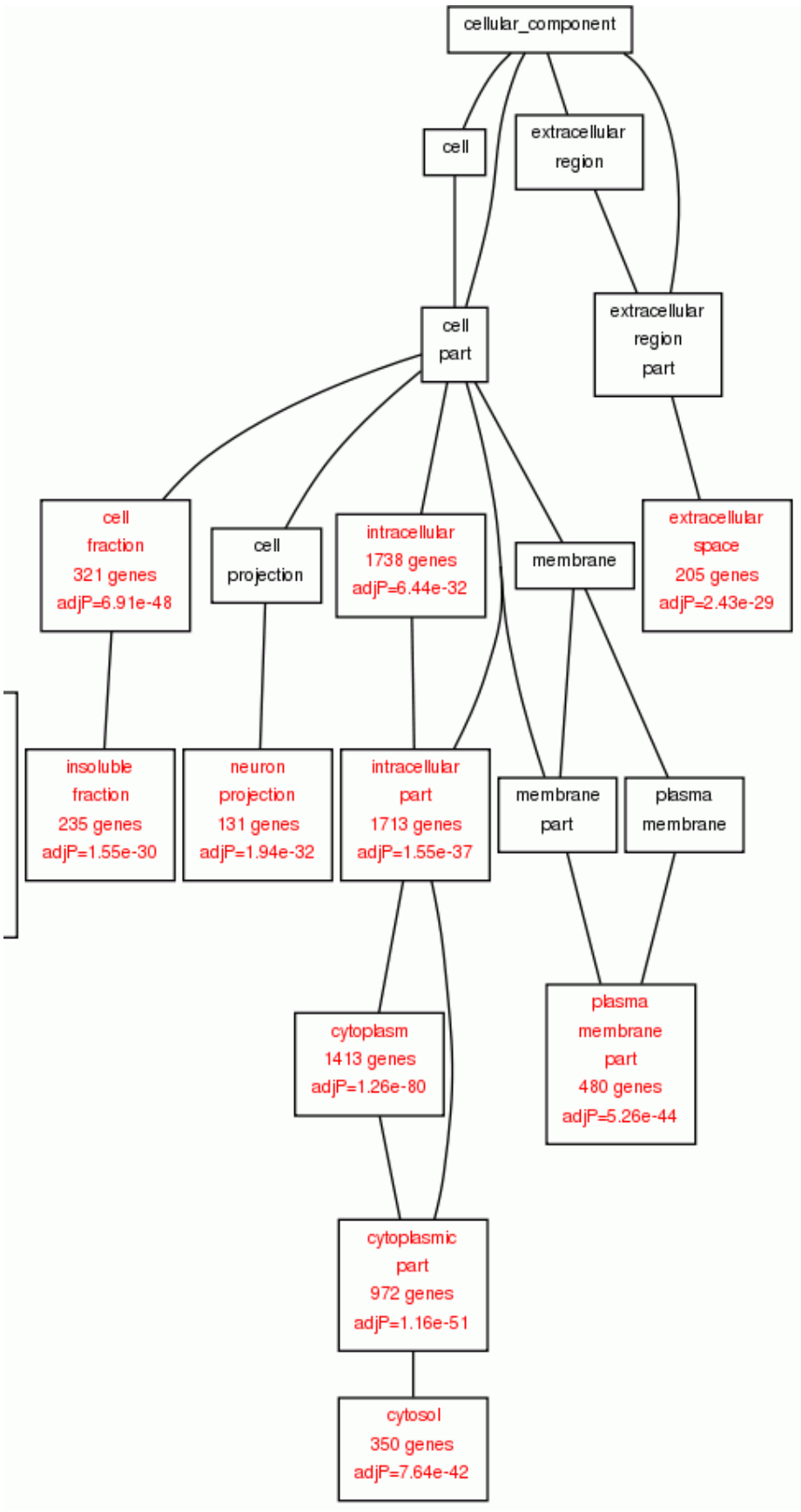


Figure 5-2 (continued) Functional annotation of the stress genes using WebGestalt. The directed acyclic graph displayed the predicted GO functions under 'cellular component' of the stress gene list.

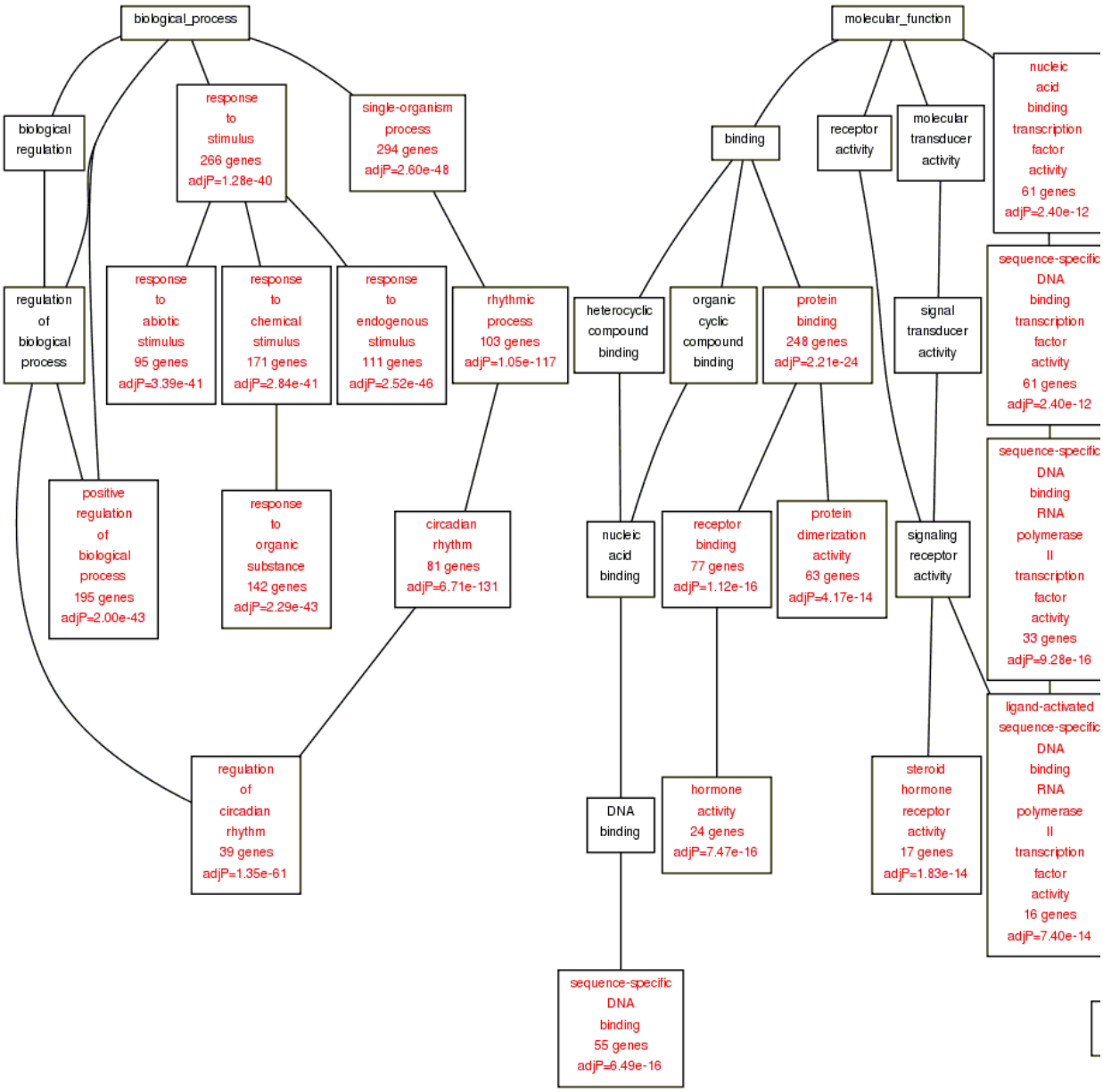


Figure 5-3 Functional annotation of the circadian genes using WebGestalt. The directed acyclic graph (DAG) displayed the predicted GO functions under biological process and molecular function of the circadian gene list.

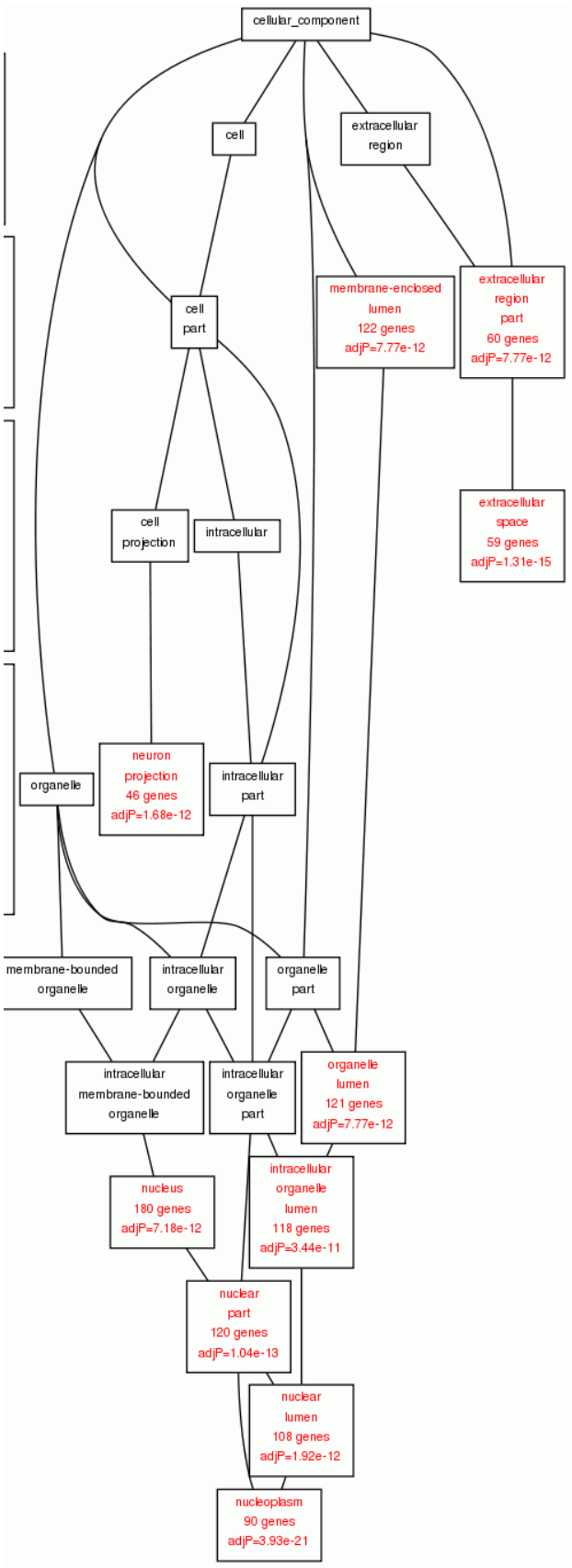


Figure 5-3 (continued) Functional annotation of the circadian genes using WebGestalt. The directed acyclic graph (DAG) displayed the predicted GO functions under cellular component of the circadian gene list.

5.3.2 Q-Q plots of the single SNP associations

The Q-Q plots for the single SNP association analyses are displayed below (see **Figure 5-4, Figure 5-5, Figure 5-6, Figure 5-7** and **Figure 5-8**). The distribution of the observed –log *p*-value followed the null distribution, suggesting the results of the associations were unlikely to be affected by any hidden population structure. The genomic inflation factors (λ) ranged from 1 to 1.3 across all the analyses, with higher λ observed in the associations using the neurotransmitter stress x circadian gene list. The increased genomic inflation factors in the neurotransmitter stress x circadian gene list could be due to the strong relatedness of SNPs as they were located on small number of genes.

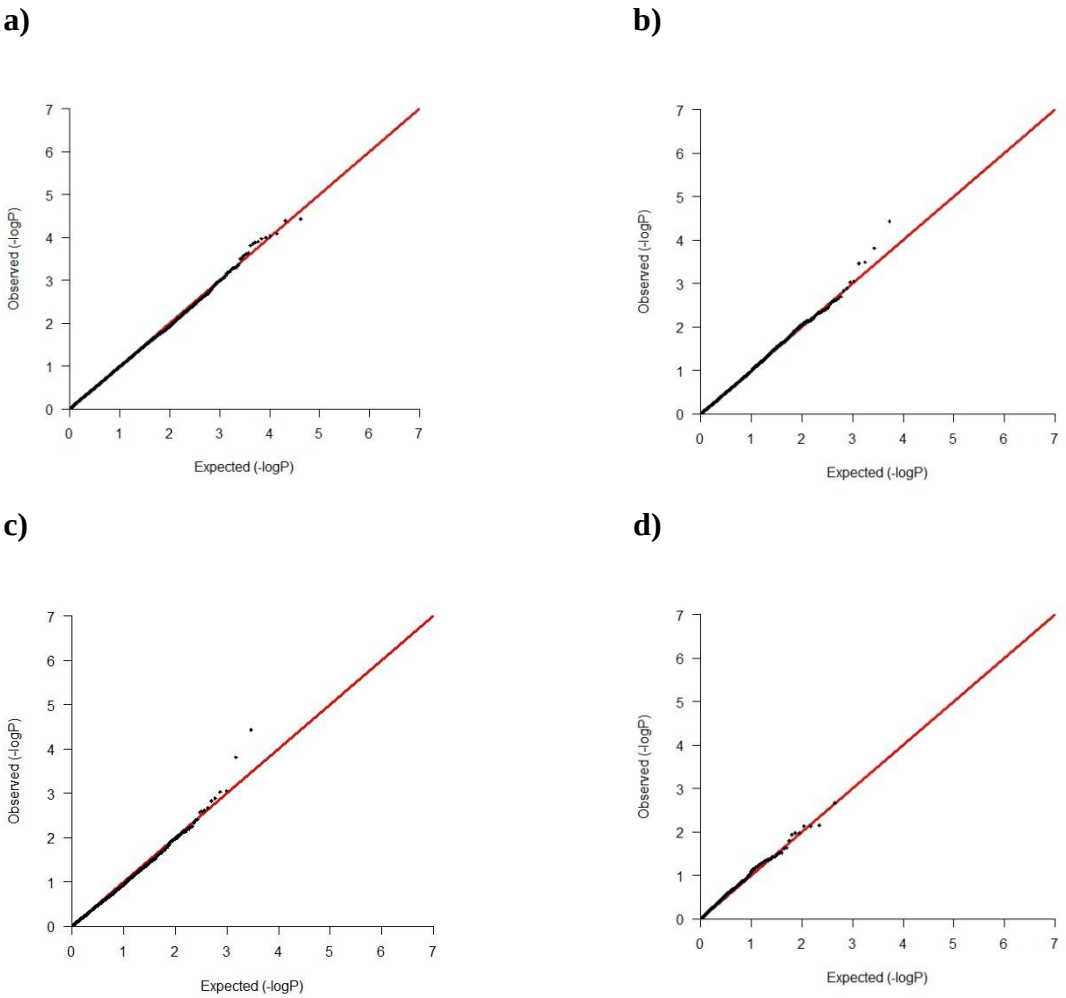


Figure 5-4 Q-Q plots of genetic associations of the left VS activation during reward anticipation with a) stress genes; b) circadian genes; c) overlap stress x circadian genes; and d) neurotransmitter stress x circadian genes. The genomic inflation factors (λ) were 1.00

in all gene lists except the neurotransmitter stress x circadian gene list (λ = 1.15). All red lines indicate x = y.

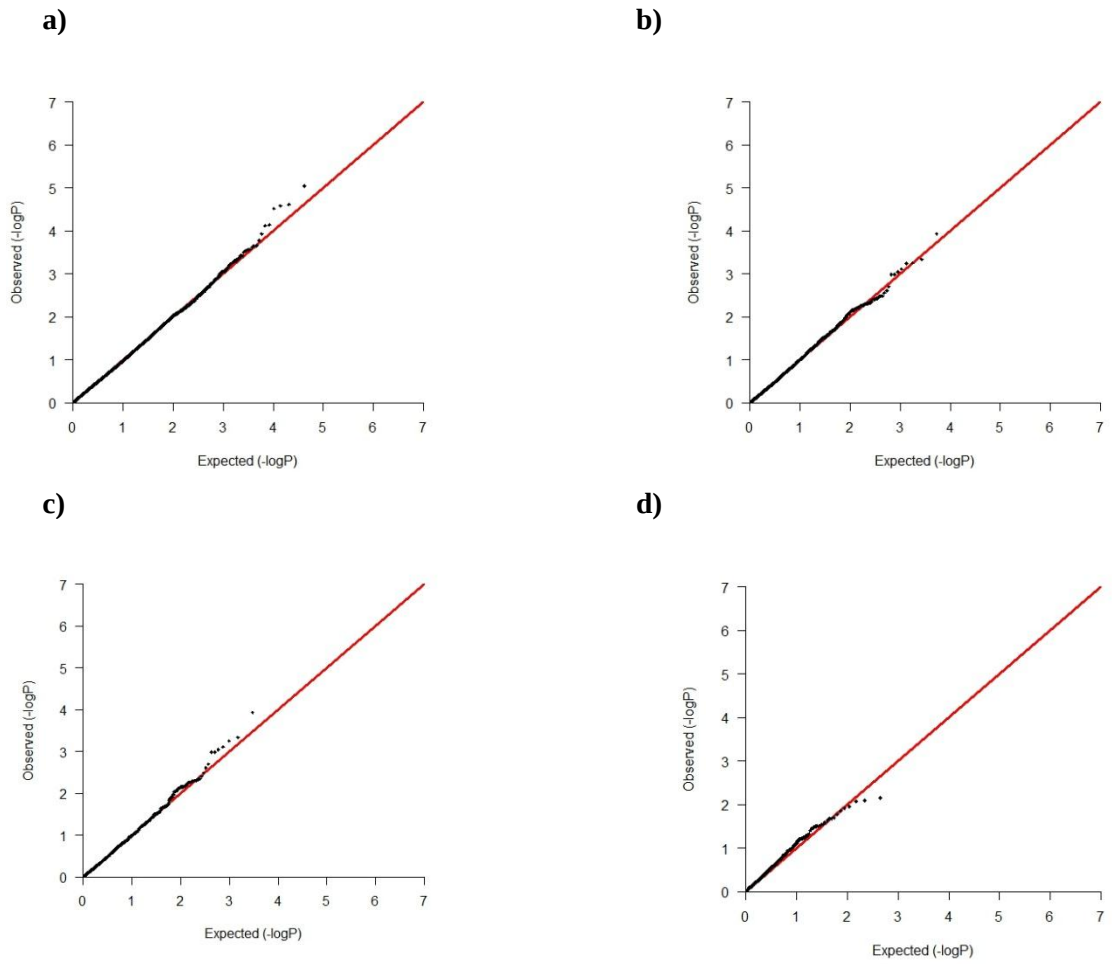


Figure 5-5 Q-Q *p*-value plots of genetic associations of the right VS activation during reward anticipation with a) stress genes; b) circadian genes; c) overlap stress x circadian genes; and d) neurotransmitter stress x circadian genes. The genomic inflation factors (λ) were 1.00 in all gene lists except the neurotransmitter stress x circadian gene list (λ = 1.05). All red lines indicate x = y.

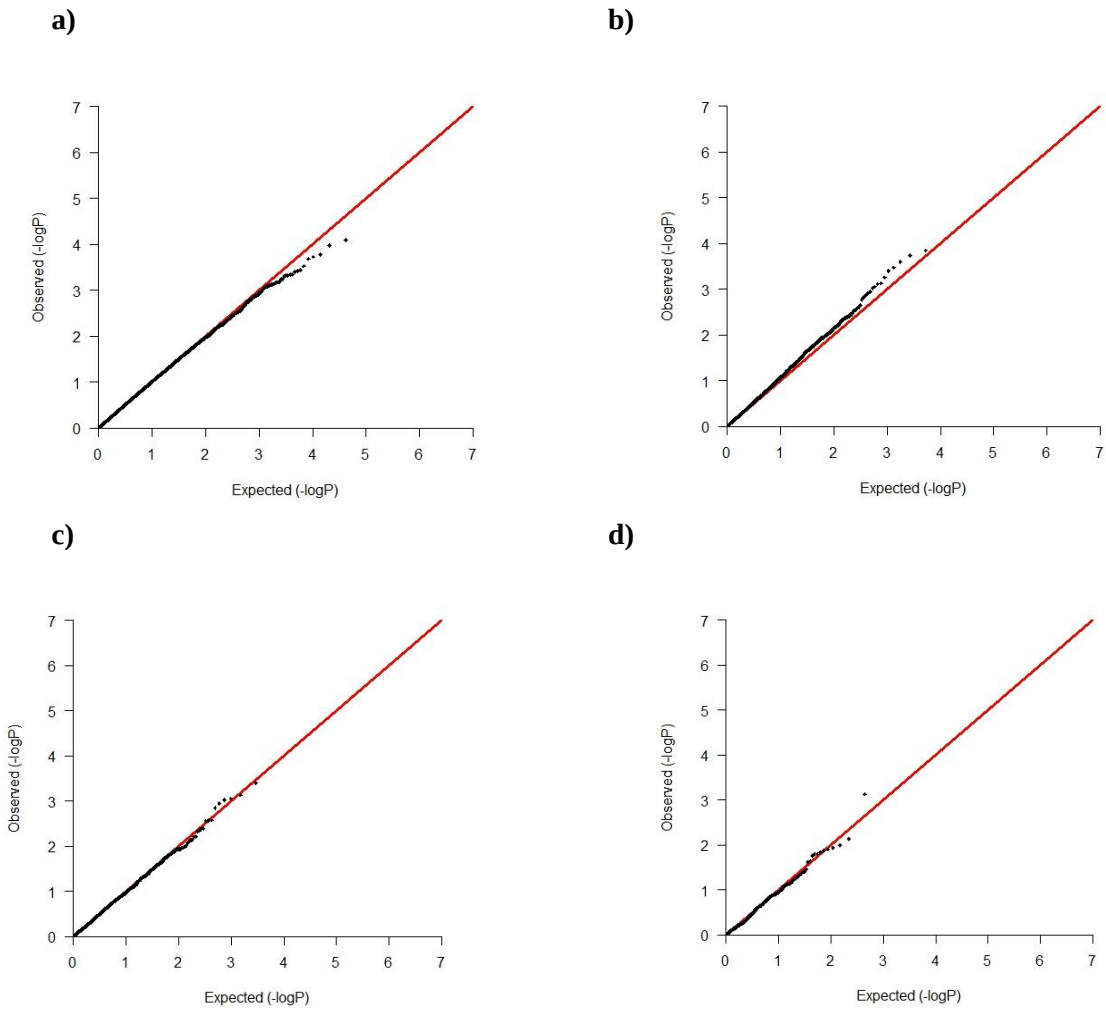


Figure 5-6 Q-Q *p*-value plots of genetic associations of past 12 month alcohol use in 14-year- old adolescents with a) stress genes; b) circadian genes; c) overlap stress x circadian genes; and d) neurotransmitter stress x circadian genes. The genomic inflation factors (λ) were 1.00 in the neurotransmitter and overlap stress x circadian gene lists. λ = 1.02 in the stress gene list and λ = 1.05 in the circadian gene list. All red lines indicate x = y.

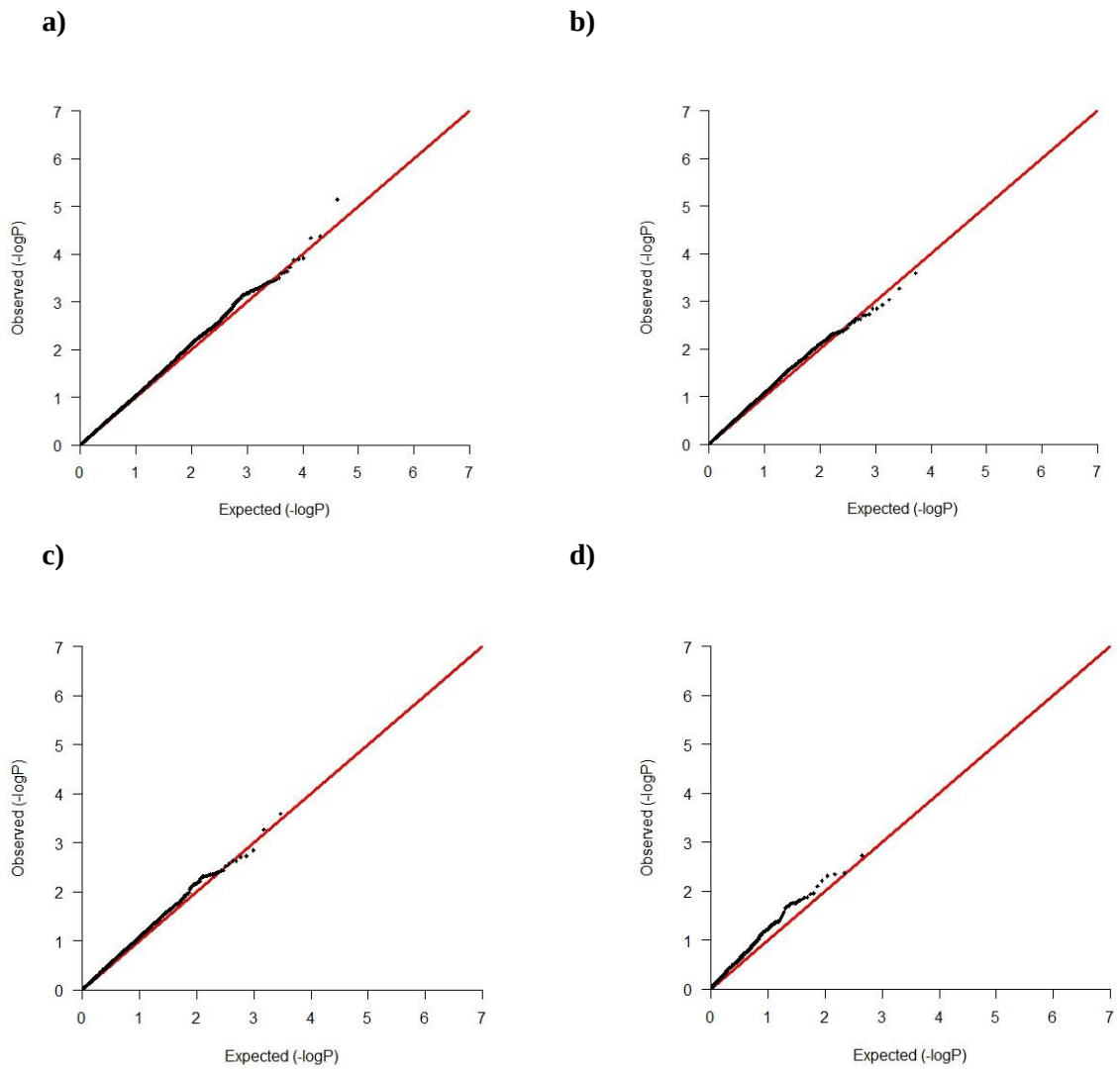


Figure 5-7 Q-Q *p*-value plots of genetic associations of past 12 month alcohol use in 16-year- old adolescents with a) stress genes; b) circadian genes; c) overlap stress x circadian genes; and d) neurotransmitter stress x circadian genes. $\lambda = 1.05$ in the stress gene list; $\lambda = 1.16$ in the circadian gene list; $\lambda = 1.15$ in the overlap stress x circadian gene list; $\lambda = 1.3$ in the neurotransmitter stress x circadian gene list. All red lines indicate x = y.

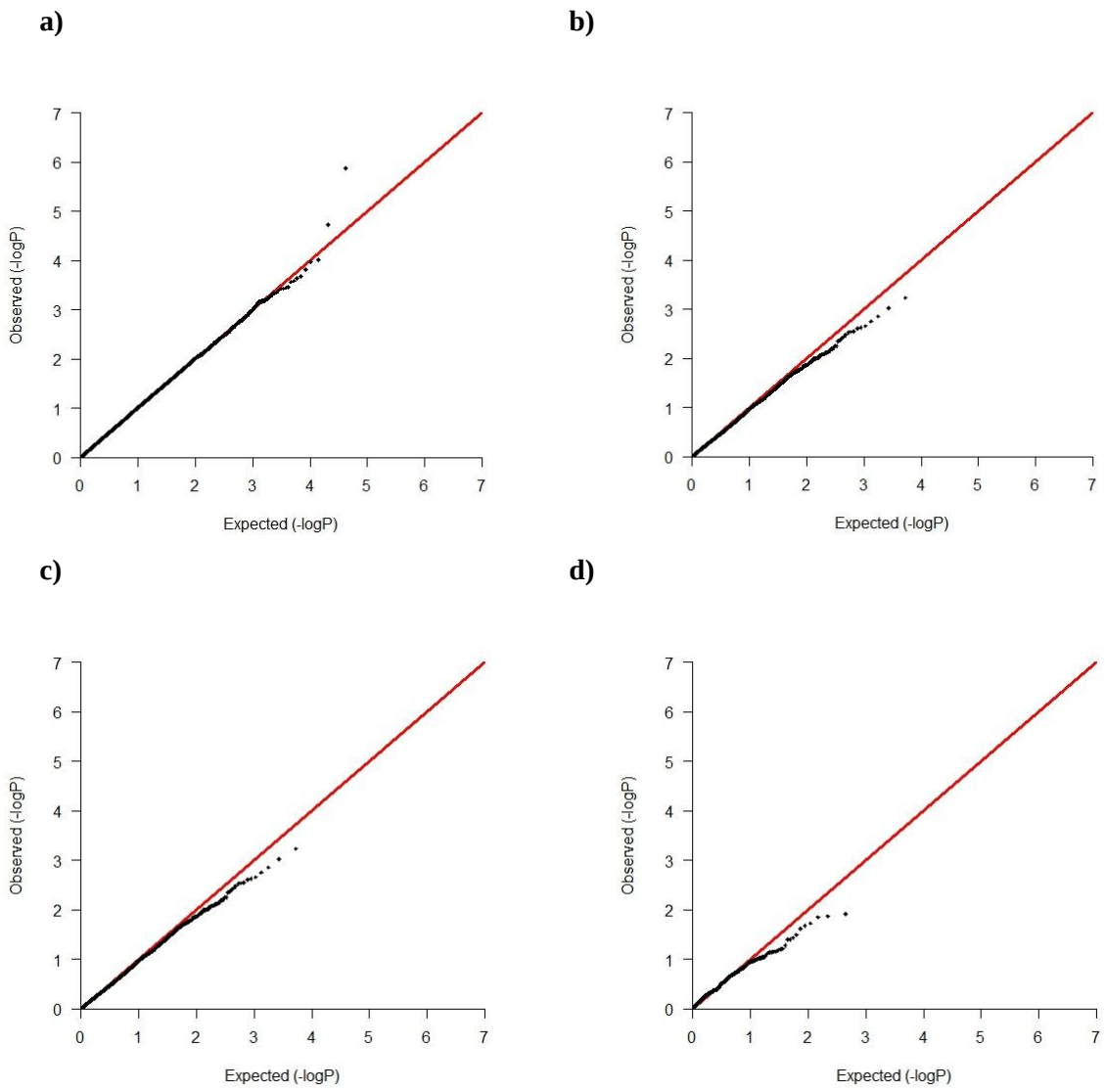


Figure 5-8 Q-Q *p*-value plots of genetic associations of lifetime tobacco use in 16-year-old adolescents with a) stress genes; b) circadian genes; c) overlap stress x circadian genes; and d) neurotransmitter stress x circadian genes. λ = 1.01 in the stress gene list; λ = 1 in the circadian gene list; λ = 1 in the overlap stress x circadian gene list; λ = 1.12 in the neurotransmitter stress x circadian gene list. All red lines indicate x = y.

5.3.3 Associations with VS activation during reward anticipation

Genetic association analyses were performed to examine the associations between the VS activations during reward anticipation with the **a)** stress genes; **b)** circadian genes; **c)** overlap stress x circadian genes; and **d)** neurotransmitter stress x circadian genes. No significant associations were found after the corrections of multiple testing using the FDR-BH procedure (

Table 5-3 for associations with the left VS activations and **Table 5-4** for associations with the right VS activations).

Table 5-3 Associations with the left VS activation during reward anticipation. The analyses showed SNPs from the stress and circadian genes were not associated with the left VS activation during reward anticipation. For each analysis five SNPs with the lowest *p*-values are reported below.

Chromosome	*Gene Symbol*	*Gene*	*SNP*	*Location*	*Beta*	*p*	p_{FDR-BH}
Stress genes							
19	*PPP5C*	Protein phosphatase 5	rs741231	51586003	0.095	3.68E-05	0.642
6	*AGER*	Advanced glycosylation end product-specific receptor	rs2070600	32259421	-0.161	4.09E-05	0.642
1	*RYR2*	Ryanodine receptor 2	rs11583646	235873442	0.101	8.27E-05	0.642
8	*FBXO32*	F-box protein 32	rs4436105	124617617	0.065	9.34E-05	0.642
21	*ABCG1*	ATP-binding cassette, sub-family G, member 1	rs225406	42572299	-0.107	0.000101	0.642
Circadian genes							
19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs741231	51586003	0.095	3.68E-05	0.196

19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs917948	51548855	0.099	0.000156	0.414
17	*RAI1*	Retinoic acid induced 1	rs11649804	17637480	-0.063	0.000328	0.460
20	*GHRH*	Growth hormone releasing hormone	rs6032470	35320941	-0.085	0.000346	0.460
3	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	rs3774262	188054508	-0.086	0.000880	0.827

Overlap stress x circadian genes

19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs741231	51586003	0.095	3.68E-05	0.109
19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs917948	51548855	0.099	0.000156	0.230
3	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	rs3774262	188054508	-0.086	0.000880	0.689
19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs759290	51583951	0.059	0.000933	0.689
19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs11083822	51570135	0.067	0.00131	0.745

Neurotransmitter, stress x circadian genes

11	*CCKBR*	Cholecystokinin B receptor	rs2947029	6244858	0.049	0.00214	0.734
7	*DDC*	Dopa decarboxylase (aromatic L-amino acid decarboxylase)	rs3735275	50484765	0.057	0.00702	0.734
11	*CCKBR*	Cholecystokinin B receptor	rs906895	6236824	0.043	0.00733	0.734
10	*ADRB1*	Adrenergic, beta-1, receptor	rs17875474	115791055	0.099	0.00748	0.734
19	*CACNA1A*	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	rs10411276	13296829	0.057	0.0106	0.734

Table 5-4 Associations with the right VS activation during reward anticipation. The analyses showed SNPs from the stress and circadian genes were not associated with the right VS activation during reward anticipation. For each analysis five SNPs with the lowest *p*-values are reported below.

Chromosome	*Gene Symbol*	*Gene*	*SNP*	*Location*	*Beta*	*p*	$p_{FDR\text{-}BH}$
Stress genes							
8	*FGF17*	Fibroblast growth factor 17	rs3176292	21959262	-0.088	9.09E-06	0.315
8	*FGF17*	Fibroblast growth factor 17	rs1078363	21960158	-0.078	2.43E-05	0.315
21	*ABCG1*	ATP-binding cassette, sub-family G, member 1	rs225406	42572299	-0.118	2.63E-05	0.315
1	*CHI3L1*	Chitinase 3-like 1	rs12123883	201431465	-0.129	3.05E-05	0.315
1	*RYR2*	Ryanodine receptor 2	rs11583646	235873442	0.104	7.25E-05	0.524

Circadian genes

19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs741231	51586003	0.090	0.000119	0.633
3	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	rs3774262	188054508	-0.092	0.000466	0.698
19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs917948	51548855	0.092	0.000566	0.698
20	*GHRH*	Growth hormone releasing hormone	rs6032470	35320941	-0.084	0.000573	0.698
15	*RORA*	RAR-related orphan receptor A	rs1437547	59221873	0.095	0.000777	0.698

Overlap stress x circadian genes

19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs741231	51586003	0.090	0.000119	0.351
3	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	rs3774262	188054508	-0.092	0.000466	0.443
19	*PPP5C*	Protein phosphatase 5, catalytic subunit	rs917948	51548855	0.092	0.000566	0.443
15	*RORA*	RAR-related orphan receptor A	rs1437547	59221873	0.095	0.000777	0.443

16	*CDH1*	Cadherin 1, type 1, E-cadherin (epithelial)	rs2276329	67420815	0.106	0.000908	0.443
Neurotransmitter, stress x circadian genes							
7	*DDC*	Dopa decarboxylase (aromatic L-amino acid decarboxylase)	rs4947510	50492914	0.050	0.00699	0.698
7	*DDC*	Dopa decarboxylase (aromatic L-amino acid decarboxylase)	rs2329340	50587723	0.045	0.00805	0.698
7	*DDC*	Dopa decarboxylase (aromatic L-amino acid decarboxylase)	rs3735275	50484765	0.056	0.00844	0.698
7	*DDC*	Dopa decarboxylase (aromatic L-amino acid decarboxylase)	rs1966839	50594595	0.043	0.0111	0.698
11	*CCKBR*	Cholecystokinin B receptor	rs2947029	6244858	0.041	0.0121	0.698

5.3.4 Associations with alcohol use in adolescents at age 14 and 16

Genetic association analyses were performed to examine the associations of **a)** stress genes; **b)** circadian genes; **c)** overlap stress x circadian genes; and **d)** neurotransmitter stress x circadian genes and the past 12 months alcohol use in adolescents at age 14 and age 16. No significant associations were found after the corrections of multiple testing using the FDR-BH procedure (**Table 5-5**

and

Table 5-6).

Table 5-5 Associations with alcohol use in adolescents at age 14. The analyses showed SNPs from the stress and circadian genes were not associated with the alcohol use in the adolescents at age 14. For each analysis five SNPs with the lowest *p*-values are reported below.

Chromosome	***Gene Symbol***	***Gene***	***SNP***	***Location***	***Beta***	***p***	$p_{FDR\text{-}BH}$
Stress genes							
14	*AHSA1*	Activator of heat shock 90kDa ATPase homolog 1	rs3742729	77012331	0.222	8.11E-05	0.977
6	*AKAP12*	A kinase (PRKA) anchor protein 12	rs2626356	151695962	-0.224	0.000106	0.977
14	*FOXA1*	Forkhead box A1	rs3742595	37126014	-0.375	0.000167	0.977
9	*IKBKAP*	Inhibitor of kappa light polypeptide gene enhancer in B-cells	rs13299328	110743382	0.197	0.000188	0.977
14	*AHSA1*	Activator of heat shock 90kDa ATPase homolog 1	rs17751319	77000631	0.241	0.000207	0.977
Circadian genes							

2	*NCKAP5*	NCK-associated protein 5	rs6735119	133645739	0.231	0.000142	0.423
2	*NCKAP5*	NCK-associated protein 5	rs6715483	133641172	0.176	0.000185	0.423
2	*NCKAP5*	NCK-associated protein 5	rs10803536	133645695	0.173	0.000254	0.423
2	*NCKAP5*	NCK-associated protein 5	rs6749758	133645716	0.169	0.000338	0.423
15	*RORA*	RAR-related orphan receptor A	rs893286	59168370	-0.153	0.000398	0.423

Overlap stress x circadian genes

15	*RORA*	RAR-related orphan receptor A	rs893286	59168370	-0.153	0.000398	0.667
2	*IL1B*	Interleukin 1, beta	rs1143634	113306861	0.124	0.000741	0.667
4	*HNRNPD*	Heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)	rs3944	83493823	0.126	0.000885	0.667
10	*CDK1*	Cyclin-dependent kinase 1	rs744383	62206784	0.121	0.000938	0.667

10	*CDK1*	Cyclin-dependent kinase 1	rs2448358	62227266	0.117	0.00113	0.667
Neurotransmitter, stress x circadian genes							
2	*IL1B*	Interleukin 1, beta	rs1143634	113306861	0.124	0.000741	0.325
11	*TPH1*	Tryptophan hydroxylase 1	rs591556	18017976	-0.112	0.00737	0.766
12	*NOS1*	Nitric oxide synthase 1 (neuronal)	rs2293048	116149208	-0.120	0.0102	0.766
17	*SLC6A4*	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	rs1906451	25539605	0.078	0.0116	0.766
17	*SLC6A4*	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	rs12449783	25551779	0.077	0.0125	0.766

Table 5-6 Associations with alcohol use in adolescents at age 16. The analyses showed SNPs from the stress and circadian genes were not associated with the alcohol use in adolescents at age 16. For each analysis five SNPs with the lowest *p*-values are reported below.

Chromosome	*Gene Symbol*	*Gene*	*SNP*	*Location*	*Beta*	*p*	$p_{FDR\text{-}BH}$
Stress genes							
19	*WTIP*	Wilms tumour 1 interacting protein	rs2965282	39676632	0.145	7.24E-06	0.299
7	*STEAP1*	Six transmembrane epithelial antigen of the prostate 1	rs259135	89634060	-0.160	4.18E-05	0.605
15	*INO80*	Chromatin-remodelling ATPase INO90 homolog	rs2928147	39171941	-0.513	4.58E-05	0.605
19	*WTIP*	Wilms tumour 1 interacting protein	rs11671685	39677306	0.124	0.000123	0.605
9	*CTSL1*	Cathepsin L1	rs3128511	89537270	0.122	0.000129	0.605
Circadian genes							
3	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	rs6444175	188062438	-0.133	0.000258	0.746

1	*ADORA1*	Adenosine A1 receptor	rs17511192	201386114	0.107	0.000539	0.746
1	*ELF3*	E74-like factor 3 (ets domain transcription factor, epithelial-specific)	rs2819360	200243877	0.103	0.000926	0.746
2	*NCKAP5*	NCK-associated protein 5	rs7595365	133228231	-0.103	0.00119	0.746
1	*ADORA1*	Adenosine A1 receptor	rs17465037	201397294	0.101	0.00144	0.746

Overlap stress x circadian genes

3	*ADIPOQ*	Adiponectin, C1Q and collagen domain containing	rs6444175	188062438	-0.133	0.000258	0.626
1	*ADORA1*	Adenosine A1 receptor	rs17511192	201386114	0.107	0.000539	0.626
1	*ADORA1*	Adenosine A1 receptor	rs17465037	201397294	0.101	0.00144	0.626
17	*SLC6A4*	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	rs2020936	25574940	-0.124	0.00187	0.626
1	*ADORA1*	Adenosine A1 receptor	rs34639219	201389739	0.098	0.00195	0.626

Neurotransmitter, stress x circadian genes

17	*SLC6A4*	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	rs2020936	25574940	-0.124	0.00187	0.437
12	*TPH2*	Tryptophan hydroxylase 2	rs11179022	70657013	-0.252	0.00425	0.437
6	*VIP*	Vasoactive intestinal peptide	rs633596	153107180	-0.121	0.00458	0.437
17	*SLC6A4*	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	rs8071667	25576899	-0.115	0.00496	0.437
17	*SLC6A4*	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	rs6354	25574024	-0.110	0.00612	0.437

5.3.5 Associations with lifetime tobacco use in adolescents at age 16

Genetic association analyses were performed to examine the associations of **a)** stress genes; **b)** circadian genes; **c)** overlap stress x circadian genes; and **d)** neurotransmitter stress x circadian genes and the lifetime tobacco use in adolescents at age 16. No significant associations were found after the corrections of multiple testing using the FDR-BH procedure (**Table 5-7**).

Table 5-7 Associations with tobacco use in adolescents at age 16. The analyses showed SNPs from the stress and circadian genes were not associated with the tobacco use in adolescents at age 16. For each analysis five SNPs with the lowest *p*-values are reported below.

Chromosome	***Gene Symbol***	***Gene***	***SNP***	***Location***	***Beta***	***p***	$p_{FDR\text{-}BH}$
Stress genes							
20	*PLCB1*	Phospholipase C, beta 1	rs727684	8183815	-0.186	1.32E-06	0.055
20	*KCNB1*	Potassium voltage gated channel, Shab-related subfamily, member 1	rs6095526	47502279	0.572	1.90E-05	0.392
20	*PLCB1*	Phospholipase C, beta 1	rs6133566	8202666	-0.151	9.57E-05	0.900

1	*CACNA1E*	Calcium channel, voltage-dependent, R type, alpha 1E subunit	rs16858051	179980807	0.204	0.000107	0.900
5	*GRIA1*	Glutamate receptor, ionotropic, AMPA 1	rs4958671	153024649	-0.146	0.000152	0.900
Circadian genes							
6	*EYS*	Eyes shut homolog (Drosophila)	rs9363329	65822684	-0.226	0.000585	0.999
11	*CRY2*	Cryptochrome 2 (photolyase-like)	rs1554338	45863406	0.258	0.000958	0.999
15	*UBE3A*	Ubiquitin protein ligase E3A	rs2719881	23157820	0.332	0.00140	0.999
15	*RORA*	RAR-related orphan receptor A	rs8032023	59199434	-0.114	0.00177	0.999
10	*KCNMA1*	Potassium large conductance calcium-activated channel, subfamily M, alpha member 1	rs2719995	78992108	0.112	0.00219	0.999
Overlap stress x circadian genes							

15	*RORA*	RAR-related orphan receptor A	rs8032023	59199434	-0.114	0.00177	0.991
10	*KCNMA1*	Potassium large conductance calcium-activated channel, subfamily M, alpha member 1	rs2719995	78992108	0.112	0.00219	0.991
6	*TNF*	Tumor necrosis factor	rs2256965	31663109	0.109	0.00285	0.991
15	*RORA*	RAR-related orphan receptor A	rs782931	59195995	-0.113	0.00288	0.991
10	*CXCL12*	Chemokine (C-X-C motif) ligand 12	rs2839692	44194873	0.131	0.00330	0.991

Neurotransmitter, stress x circadian genes

19	*CACNA1A*	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	rs16042	13202037	-0.137	0.0124	0.905
12	*PMCH*	Pro-melanin-concentrating hormone	rs10860845	101121832	0.090	0.0137	0.905
19	*CACNA1A*	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	rs7250783	13237536	0.101	0.0143	0.905

19	*CACNA1A*	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	rs2112460	13451412	0.085	0.0189	0.905
6	*GRIK2*	Glutamate receptor, ionotropic, kainate 2	rs1417182	102440034	-0.084	0.0213	0.905

5.3.6 Polygenic score analyses for the VS activations during reward anticipation

Polygenic score analyses were performed to identify the proportion of variance in the VS activations during reward anticipation explained by the stress and circadian genes *en masse*. Each analysis was repeated using two discovery samples. As illustrated in **Table 5-8**, the stress and circadian genes *en masse* did not account for the VS activations during reward anticipation.

Table 5-8 Polygenic score analyses for the VS BOLD responses during reward anticipation.

		Left VS		***Right VS***	
Gene List	***Discovery sample***	***r***	***p (one-tailed)***	***r***	***p (one-tailed)***
Stress genes	List 1	0.009	0.394	-0.049	0.078
	List 2	0.019	0.349	-0.048	0.158
Circadian genes	List 1	-0.011	0.373	0.002	0.474
	List 2	-0.003	0.476	0.013	0.391
Overlap stress x circadian genes	List 1	-0.074	0.016	-0.033	0.167
	List 2	-0.092	0.029	-0.020	0.336

Neurotransmitter, stress x circadian genes	List 1	-0.028	0.207	0.030	0.191
	List 2	-0.049	0.153	0.037	0.219

5.3.7 Polygenic score analyses for alcohol use

Polygenic score analyses were performed to study the past 12 months alcohol use in adolescents at age 14 and age 16. The results showed that the additive genetic effects of stress and circadian genes *en masse* did not explain alcohol use in the adolescents at age 14. The polygenic influence of stress genes on alcohol use in adolescents at age 16 had reached nominal significance but they did not survive the correction of multiple testing. Inconsistent findings were found in adolescents' alcohol use at age 16, in which significant polygenic effects of the circadian and the overlap stress x circadian genes were observed when using List 2 as discovery sample (**Table 5-9**).

Table 5-9 Polygenic score analyses for alcohol use in adolescents measured at age 14 and 16

		Age 14		*Age 16*	
Gene List	***Discovery sample***	***r***	***p (one-tailed)***	***r***	***p (one-tailed)***
Stress genes	List 1	-0.008	0.389	0.065	0.033
	List 2	0.002	0.477	0.111	0.013
Circadian genes	List 1	0.014	0.316	0.039	0.133
	List 2	0.048	0.118	0.090	0.036
Overlap stress x circadian genes	List 1	-0.012	0.343	0.044	0.105
	List 2	0.017	0.337	0.089	0.038
Neurotransmitter, stress x circadian genes	List 1	0.01	0.368	0.035	0.160
	List 2	0.004	0.466	0.052	0.151

5.3.8 Polygenic score analysis for lifetime tobacco use

Polygenic score analyses were performed to study the genetic effects of the stress and circadian genes on tobacco use in 16-year-old adolescents. The results showed that the tobacco use in adolescents was not explained by the additive genetic effects of stress and circadian genes *en masse* (**Table 5-10**).

Table 5-10 Polygenic score analyses for lifetime tobacco use in 16-year-old adolescents

Gene List	*Discovery sample*	*r*	*p (one-tailed)*
Stress genes	List 1	-0.026	0.219
	List 2	-0.031	0.261
Circadian genes	List 1	-0.052	0.065
	List 2	-0.041	0.193
Overlap stress x circadian genes	List 1	-0.033	0.167
	List 2	-0.012	0.398
Neurotransmitter, stress x circadian genes	List 1	0.026	0.225

	List 2	0.033	0.243

5.4 Discussion

Evidence for the genetic influence on substance use has been mentioned in the literature. In order to investigate the additive genetic effects of the stress and circadian systems on substance use during adolescence, large-scale candidate gene association analyses were performed on stress and circadian genes that were systematically selected from the NCBI Gene database. After performing the FDR-BH control, no common SNPs from the identified stress and circadian genes were associated with the VS activations during reward anticipation, alcohol and tobacco use in adolescents. The polygenic effect of the stress and circadian systems did not explain the VS activations and tobacco use in the adolescents either. Inconsistent results were found in alcohol use among the adolescents at age 16, of which the significant polygenic effect of the stress genes *en masse* did not survive the correction of multiple testing. The polygenic influence from the circadian system and stress x circadian genes were inconsistent across the validation analyses.

5.4.1 Confirming the stress and circadian properties of the gene lists

Selecting stress and circadian genes for association analysis enabled their genetic influences on substance use in adolescents to be investigated. To minimise the selection biases, the stress and circadian genes were identified based on the information available in the NCBI gene and literature database. The WebGestalt functional annotation analyses confirmed the biological functions represented by the stress and circadian gene lists, in which the lists were enriched with the GO terms related to the regulations of stress and circadian systems. Large number of genes was classified under functions related to cellular signaling and these functions could be important for regulating the stress and circadian signals.

5.4.2 Stress and circadian systems on substance use in adolescents

Previous studies suggested the associations between the disruptions of the stress system and circadian rhythm with increased substance use (Chandra et al., 2011, Seggio et al., 2009,

Rosenwasser et al., 2005a, Steptoe and Ussher, 2006, Richardson et al., 2008, Schmitt et al., 2008, Sarkar et al., 2007). The genetic effects of the stress and circadian systems in alcohol use were also suggested (Dong et al., 2011, Blomeyer et al., 2008, Wang et al., 2012, Kovanen et al., 2010). However the results presented in this chapter failed to replicate previous findings due to various reasons.

Firstly, the genetic architecture of substance use and addictions varies across populations. Due to the discrepancies in allele frequencies and LD structure in different populations, the genetic effect of any previously identified stress or circadian-related risk variants in substance use might not be detectable in the IMAGEN participants. The results were consistent with the CDCV hypothesis, indicating that individual SNPs only displayed small genetic effect on complex traits (Sullivan et al., 2012). Corrections of multiple testing in the single SNP association analyses might also hinder the detection of risk variants on substance use.

Secondly, the significant polygenic risks for the stress and circadian systems on alcohol use in adolescents at age 16 might support the increased gene-environment interactions on alcohol use throughout development. Gene-environment interactions might be less detectable in adolescents at age 14; thus the polygenic effects of the stress and circadian systems on alcohol use were not observed. During adolescence, individuals are exposed to numerous life events and physiological changes such as alterations in circadian rhythm and stress reactivity (reviewed by Spear (2000)). These developmental changes could become more prominent and impose a larger impact on alcohol use in the 16-year-old adolescents. Although the polygenic risk of stress genes on alcohol use did not survive multiple testing corrections, the results might suggest the genetic basis of stress-related alcohol use during adolescence. The inconsistent findings from the circadian and the overlap stress x circadian genes might implicate the contribution of the circadian genes and the potential interaction between the stress and circadian systems on the increased alcohol use in adolescents at age 16.

Nevertheless, detecting risk variants and polygenic effects is dependent on the prevalence of risk phenotypes. The proportion of adolescents who had consumed alcohol increased from ~36% at age 14 to ~80% at age 16. The increased alcohol consumption in the adolescents at age 16 could increase the power for genetic associations and this might account for the nominally significant polygenic effects of the stress and circadian systems. On the other hand, only ~36% adolescents had smoked tobacco at age 16. Larger sample size might be required to detect the polygenic influence of the stress and circadian systems on tobacco use.

The sensitivity of phenotypes is also important for detecting the risk variants. In the current study, the substance use phenotypes were selected based on the regular alcohol and tobacco use in adolescents. The dichotomous nature of the substance use phenotypes did not distinguish adolescents with heavy alcohol or tobacco consumption from the rest of the population. It could be possible that the genetic influence of the stress and circadian systems might impose a larger impact on heavy substance use or abuse populations and account for the non-significant findings in the single SNP and polygenic score analyses.

5.4.3 Impact of stress and circadian systems on reward sensitivity

As suggested by the previous literature and the results presented in Chapter 3, the tendency of substance use in adolescents could be predicted by reward sensitivity, indicated by the VS activations during reward processing (Bjork et al., 2008a, Schneider et al., 2012). Despite the VS was one of the most strongly activated brain region during reward anticipation in the IMAGEN adolescents, the VS activations were not explained by the additive genetic influence of the stress and circadian genes.

As reported in Chapter 3, life stress did not alter the VS activations during reward anticipation. Such observation might account for the lack of genetic effect from the stress genes and the VS

activations. Previous studies demonstrated the associations between reward processing and the genetic polymorphisms of *PER2* and *CLOCK* (Forbes et al., 2012, Tsuchimine et al., 2013), and this has supported the motivation for studying the genetic influences of circadian rhythm on reward processing in this chapter. Since there were no circadian phenotypes available in the IMAGEN Study, the relationship between circadian rhythm and reward sensitivity in adolescents had remained speculative.

It is also possible that the genetic polymorphisms of stress and circadian genes alter brain activations in multiple regions during reward processing. Indicated by the meta-analysis performed by Liu and colleagues (2011), a number of frontal and limbic brain regions including the anterior cingulate cortex, medial orbitofrontal cortex, insula and the thalamus were activated during reward anticipation. The VS activations during reward anticipation did not fully account for the reward sensitivity in the adolescents. It has been proposed that brain regions connecting to the NAc, such as the extended amygdala that consists of the central nucleus of the amygdala (CeA) and the bed nucleus of the stria terminalis (BNST) in rats, can be responsible for integrating signals from the stress and reward systems. Evidence suggested that the CRF-containing neurons in the extended amygdala could mediate the negative consequences of stress and cocaine withdrawal (Koob and Kreek, 2007). Upon the exposure of morphine, the adrenalectomised rats showed alterations of FosB protein expression in the NAc, BNST, CeA and PVN (Garcia-Perez et al., 2012), supporting the interactions between the stress and reward systems in multiple brain regions. As discussed in the General Introduction, the PVN is responsible for integrating stress and circadian signals (Wong and Schumann, 2012), hence suggesting the input of the circadian system in stress-related reward processing. Another point to consider is that the stress and circadian systems might impact on different aspects of reward processing such as reward-related outcomes. Brain activations during reward outcome could not be captured by the reward anticipation contrast in the MID task. Additional fMRI contrasts will be required to model brain activations in response to reward stimuli.

5.4.4 Methodological issues

5.4.4.1 Single SNP association analyses

Focusing on genes in the stress and circadian systems for genetic association analyses could be beneficial due to the amount of the genetic data available (N = 1982 individuals) in the IMAGEN Study. Gene prioritisation could reduce the number of SNPs tested, compensate the statistical power and reduce false positives. Selecting genes with similar functions might increase chances of producing spurious Q-Q plots due to the increased relatedness of SNPs; therefore producing large genomic inflations in the analyses using the neurotransmitter stress x circadian gene list. Despite the single SNP associations failed to generate any significant findings, it was observed that several top associated SNPs were located on the same genes. Such observation could be explained by the additive genetic influence of the candidate genes (Sullivan et al., 2012). Alternative approach will be required to resolve the statistical power and multiple testing issues in the single SNP association analyses.

5.4.4.2 Polygenic score analyses

The polygenic score analysis has been applied to study various complex disease traits (International Schizophrenia et al., 2009, Kos et al., 2013). As demonstrated by the literature and the results presented in this chapter, the phenotypic variance explained by polygenic risk scores could be influenced by several factors.

The selection of risk variants was based on the liberal thresholds of LD and *p*-values of SNPs in the discovery samples. Pruning SNPs based on the LD structure could enhance the independence of SNPs and reduce chances of overestimating the phenotype in target samples; whereas including SNPs below the arbitrary *p*-value threshold ($p < 0.5$) could remove SNPs that were less relevant to the phenotypes. Since the SNP scoring procedure involved averaging the SNP x effect size weightings over the number of SNPs tested, filtering the less associated SNPs prior scoring could reduce the dilution of genetic effects on the predicted phenotypic outcomes.

One possible explanation for the negative correlations observed in the VS BOLD responses, alcohol use in adolescents at age 14 and lifetime tobacco use in adolescents at age 16 was including SNPs that were less associated with the VS activations and substance use phenotypes. This could be due to the arbitrary selection of SNPs at *p*-value below 0.5. As illustrated by the polygenic score analyses on schizophrenia conducted by The International Schizophrenia Consortium (2009), the maximum phenotypic variance explained was observed at the most stringent *p*-value thresholds ($p < 0.01$). It might be possible that only a small set of SNPs in the stress and circadian genes were responsible for reward sensitivity and increased substance use in the adolescents. Including more SNPs that were less associated with the phenotypes when calculating the polygenic scores might contribute to the observed negative correlations.

The inconsistent findings in the circadian system and stress x circadian genes on alcohol use in 16-year-old adolescents have put the reliability of polygenic score analyses into question. As demonstrated by The International Schizophrenia Consortium (2009), the estimated polygenic influence of schizophrenia was found to increase with the discovery sample size. The polygenic risk score was found to account for ~ 20% phenotypic variance in schizophrenia when using 20000 case-control pairs in the discovery sample, compared to about 3% variance explained when 6900 case-control pairs were used (International Schizophrenia et al., 2009). This might also explain the significant findings in the circadian and stress x circadian genes when using List 2 (N = 1446) as the discovery sample. The small sample size in List 1 (N = 712) might reduce the power for predicting the polygenic influence of alcohol use. The findings from this chapter have confirmed that the estimated phenotypic variances could be influenced by the discovery sample size, thus highlighting the importance of results validation.

In addition, the estimated phenotypic variances can be limited by other factors such as the properties of SNPs. Yang et al. (2011) demonstrated that about 45% genetic variation of human height could be explained by all common SNPs. It was later regarded as the 'chip heritability' or

the maximum amount of additive genetic information captured by the genotyping platforms. Nevertheless, the estimated phenotypic variances by polygenic risk score could be affected by the genetic architecture of complex traits and the allele frequencies of SNPs. Given the same number of SNPs, SNPs with lower MAF (MAF = 0.05) provided larger variance explained for schizophrenia than SNPs with MAF = 0.5 (International Schizophrenia et al., 2009). Despite the limitations and statistical pitfalls, the polygenic score analysis has provided an alternative approach to predict genetic risks of complex traits.

5.4.4.3 Genome-wide genetic data

The main advantage of using genome-wide genetic data for large-scale candidate gene association analyses is to obtain high quality of genetic data. Many quality control procedures can only be performed in genome-wide scale. For example, identifying population relatedness and stratification requires a large amount of genetic information. Genetic markers with low call-rate, significant deviation from the Hardy-Weinberg equilibrium and in high LD with other markers can also be assessed accurately and in a cost-effective way.

5.4.4.4 Limitations

One major concern in statistical genetics research is the lack of replication of findings. Ideally the validation analyses should be carried out in the independent populations. Due to the amount of genetic information available, only internal replication could be performed in the polygenic score analyses. Future analyses include repeating the polygenic score analyses on alcohol use in other adolescent population can be considered.

In contrast to traditional approaches, this thesis has made a novel attempt to incorporate gene functions in conducting genetic association analyses. By employing WebGestalt and NCBI Gene database, the functions of stress and circadian genes were further classified according to

the GO terms. Similar to previous studies, the gene prioritisation approach adopted in this thesis did not free from selection biases.

As mentioned in the General Introduction, about 10% genes transcripts are found to be rhythmically expressed (Akhtar et al., 2002; Duffield et al., 2002). Ideally more circadian genes should be identified through the NCBI gene database. In contrast, large number of stress genes was classified under the cellular component GO terms, such as 'cell fraction', 'cytoplasm' and 'plasma membrane'. These genes might not be directly related to the regulation of stress in addictions or any neuropsychiatric disorders. Therefore the stress gene list might include genes that were less relevant to stress regulations in neuropsychiatric disorders and false positives. Since there were only limited number of genes associated with addictions and neuropsychiatric disorders in the current literature, it would be reasonable to develop a more stringent system to filter stress genes according to their associated GO terms. For future analysis, genes that were solely classified under the cellular component GO terms 'cell fraction' and 'cytoplasm' might be excluded.

5.4.5 Conclusions

This chapter adopted both single SNP association and polygenic score analyses in attempt to investigate the additive genetic influence from the stress and circadian systems on reward sensitivity and substance use in adolescents. The lack of significant findings could be due to the selection of phenotypes as well as a number of statistical issues. Identifying genetic variants for disorders has remained an important challenge in psychiatric genetics research. Future research should focus on developing statistical models to account for the gene-gene/ environment interactions and reducing false positives by applying rigorous validation procedures.

Chapter 6

The role of stress system and *FYN* rs1409837 on alcohol misuse

6.1 Introduction[3]

6.1.1 Alcohol misuse, drunkenness and bingeing in adolescents

Alcohol misuse or excessive alcohol consumption during adolescence has been considered as the gateway for developing alcohol addiction (Windle and Zucker, 2010, Koob and Le Moal, 1997). Alcohol-related outcomes such as drunkenness can be important for mediating alcohol use during adolescence. Drunkenness refers to the perceived level of inebriation and it indicates the level of alcohol intoxication in the individuals (Ahlström and Österberg, 2005, Windle et al., 2005). Report from the Monitoring the Future National Survey of the National Institutes of Health indicated that 19 - 42% adolescents had reported drunkenness, and 10 - 22% adolescents had experienced binge drinking, of which binge drinking was defined as the consumption of five or more alcoholic beverages on a single occasion (Johnston et al., 2006, Wechsler et al., 1994). The prevalence of drunkenness and binge drinking in adolescents may increase the risk of alcohol addiction in adulthood (Windle et al., 2005, Windle and Zucker, 2010).

Adolescents are less sensitive to the sedative effect of alcohol compared to adults (Doremus et al., 2003, Varlinskaya and Spear, 2010). When given the same dose of ethanol, adult mice showed reduced social interaction to greater extent compared to adolescent mice (Doremus et al., 2003, Varlinskaya and Spear, 2010). Adolescent rats also showed significantly higher level of locomotor activity and less anxious behaviour compared to adult rats upon acute and chronic

[3] Results from this chapter are reported in the manuscript 'Protective effect of *FYN* rs1409837 on drunkenness and amygdala BOLD responses in adolescents' (Wong *et al*., in preparation).

alcohol withdrawal (Doremus et al., 2003), suggesting that the adolescents might be less sensitive to alcohol. Such proposal was confirmed by the results from animal models, showing that the adolescent rats had higher voluntary ethanol intake than adults (Vetter et al., 2007, Doremus et al., 2005). The reduced sensitivity to alcohol in adolescents can impact on their alcohol consumption behaviour and level of drunkenness.

6.1.2 Life stress, amygdala activations and alcohol misuse

Subject to a series of developmental and life events, adolescents who experience high level of life stress are more vulnerable to alcohol addictions as they are more likely to engage in repeated alcohol use (Spear, 2000, Windle et al., 2005). Similarly, increased alcohol consumption has been reported in adolescent and young adult populations who experienced high level of life stress (Dube et al., 2006, Meier et al., 2013). Confirmed by the analyses presented in Chapter 3, IMAGEN adolescents who experienced higher level of life stress were more likely to consume alcohol. Increased contribution of life stress to alcohol use has been observed among the adolescents, with the variance explained increased from 1.7 – 4.4% at age 14 to 2.7 – 5.2% at age 16 (see Chapter 3).

The amygdala is one of the most activated brain regions during emotional processing and the exposure to stressful situations (Phan et al., 2002, Whalen et al., 1998). Emotional processing has been frequently measured by the faces task, in which the presentation of angry faces has been found to increase the activations of the amygdala, indicating that the amygdala is sensitive to emotional or stressful stimuli (Whalen et al., 1998). Alteration in the amygdala activations might predict the level of alcohol consumption, as heavy alcohol drinkers showed attenuation of the amygdala activations upon the presentation of angry faces (Gorka et al., 2013). However, opposite pattern has also been observed in heavy drinkers as they demonstrated increased amygdala activations when alcohol-related cues were presented (Dager et al., 2013).

Alterations in the amygdala activations and alcohol misuse have been found to associate with the genetic polymorphisms of the stress genes. For example, genetic polymorphisms of the serotonin transporter gene was found to associate with the amygdala activations during angry face processing in adults, as well as alcohol intoxication and increased alcohol consumption in primates that were exposed to life stress (Hariri et al., 2002, Barr et al., 2003). Genetic polymorphisms from stress receptor-coding genes *NR3C1* and *CRHR1* were found to associate with drunkenness among the adolescents and alcohol dependence populations (Desrivieres et al., 2011, Treutlein et al., 2006), suggesting that the amygdala activations during emotional processing might mediate the gene-environment interactions of alcohol misuse under stressful environment.

6.1.3 Aims and hypotheses

This chapter aims to dissect the genetic influence of the stress system on alcohol misuse and intoxication in the 16-year-old IMAGEN adolescents. Alcohol misuse and intoxication were characterised by drunkenness and binge drinking reported by the adolescents. It is hypothesised that the activations of the amygdala can mediate the genetic influence of stress genes and alcohol misuse in adolescents.

6.2 Methods

6.2.1 Materials

Genetic information and stress gene list

The genome-wide genetic data from N = 1982 individuals were used for the genetic association analyses (see **Chapter 2 Section 2.4.2** for details). To generate the stress gene list, systematic search of stress-related genes was performed in the NCBI Gene database. The procedure for gene selection was described in **Chapter 5 Section 5.2.2**). N = 41307 SNPs located ±10kb of 2348 stress genes were extracted from the genome-wide genetic data for association analyses.

Drunkenness and binge drinking phenotypes

Adolescents' information on lifetime drunkenness (Question 19a) and binge drinking (Question 17a) was measured in the ESPAD at age 16 (Hibell et al., 1997) (See **Chapter 2 Section 2.2.4).** At this age majority of adolescents had already consumed alcohol. Measuring drunkenness allows researchers to study the effect of alcohol intoxication in the adolescents, whereas the bingeing allows researchers to study the relevance of drunkenness on excessive alcohol consumption. For the lifetime drunkenness phenotype, individuals were asked to report the number of occasions they were drunk from drinking alcoholic beverages in lifetime on a six-point scale ranging from zero (never) to six (40 times or more). Individuals' responses were dichotomised into 'never been drunk in lifetime' (N = 452)/ 'had been drunk in lifetime' (N = 814). For the lifetime binge drinking phenotype, adolescents were asked to report the number of occasions in lifetime they had five or more alcoholic drinks in a row on a five-point scale ranging from zero (never) to five (10 times or more). Alcohol bingeing was defined as having five or more alcoholic beverages in one drinking occasion (Wechsler et al., 1994). Individuals' responses were dichotomised into those who had never binged in their lifetime (N = 418) versus those that had binged in lifetime (N = 848).

Amygdala BOLD responses

Individuals' amygdala BOLD responses were extracted from the faces task 'angry faces *versus* control' contrast (See **Chapter 2 Section 2.3.6** for face task). The ROIs of the left and right amygdala were defined by the human Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) and extracted to match the 'angry faces *versus* control' contrast using the Marsbar toolbox (http://marsbar.sourceforge.net). The average amygdala BOLD responses from 1484 adolescents were used.

Negative life events

Negative life events were measured using the LEQ in N = 852 16-year-old adolescents (Newcomb et al., 1981) (see **Chapter 2 Section 2.2.3** for details). Individuals were asked to report whether the life events had occurred and rate how the event made them feel using a five point scale indicating 'very happy', 'happy', 'neutral', 'unhappy' and 'very unhappy'. Both lifetime frequency and ratings towards negative life events were used to characterise life stress experienced by the adolescents.

6.2.2 Association analyses

Genotyped and imputed SNPs association analyses

Linear regressions were applied to study the additive genetic effect of stress genes on drunkenness in PLINK v1.07 (Purcell et al., 2007). The regression models of drunkenness were in the form

$$Drunkenness = b_0 + b_1*\mathrm{G} + b_2*\mathrm{Gender} + b_3*\ \mathrm{Sites} + \varepsilon,$$

where G referred to the copies of minor allele of each SNP. All associations were controlled for gender and sites of recruitment. Benjamini-Hochberg false discovery rate (FDR-BH) control was applied to correct for the number of SNPs tested.

The same model was applied to analyse the imputed SNP data, of which individuals' drunkenness information was regressed against the major allele dosage of the imputed SNPs. Both gender and sites were included as the covariate. The analyses of imputed SNPs were performed on the SPSS (IBM, version 20).

Regional association analysis

Regional association analysis was conducted on all genotyped SNPs ±500 kbp of the strongest associated (i.e. target) SNPs. The additive genetic effect of SNPs on drunkenness was analysed using PLINK v1.07 (Purcell et al., 2007). All associations were controlled for gender and recruitment sites. The results of the regional association analysis were displayed on the SNAP (SNP Annotation and Proxy Search, http://www.broad.mit.edu/mpg/snap).

mRNA expression analysis

The mRNA expression of target gene was measured by probes on the Illumina© gene expression microarray and normalised on the $\log_2$ scale (See **Chapter 2 Section 2.5** for details). The sequences of the probes were blasted on the UCSC genome browser prior any analyses. Probes that contained SNPs or displayed non-specific binding to the target gene were excluded from the analyses. Outlier detection was performed such that gene expression data outside ±3 S.D. were removed prior analyses. The $\log_2$ mRNA expressions were regressed against the copies of minor allele of the target SNP. Gender, recruitment sites and puberty development scores measured by the Puberty Development Scale (PDS, See **Chapter 2 Section 2.2.5**) (Petersen et al., 1988) were included as covariates. The associations were controlled for puberty development scores in order to account for differences in mRNA expression due to differences in physical development across individuals (Ojeda et al., 2010). The mRNA expression data from N = 628 adolescents were used in this chapter.

Neuroimaging analyses

The additive genetic effect of the target SNP on bilateral amygdala BOLD responses was studied in the following linear regression model

Amygdala BOLD responses (Left or Right) = $b_0 + b_1$*G + b_2*Gender + b_3* Handedness + b_4* Sites + ε,

where G referred to the copies of minor allele of each SNP. The bilateral amygdala BOLD responses were used to predict life stress in the linear regression model displayed below

Lifetime negative life events (rating or frequency) = b_0 + b_1* Amygdala BOLD response (Left or Right) + b_2*Gender + b_3* Sites + ε.

6.2.3 Bioinformatic analyses

Identify SNPs in 1000 Genomes database

Due to the tagging nature of SNPs, SNPs in high linkage disequilibrium (LD) with the target SNPs were identified through the 1000 Genomes SNP data set (Utah residents with ancestry from northern and western Europe, CEU population) using the Proxy Search function on SNAP (http://www.*broad*.mit.edu/mpg/*snap)*. The 1000 Genomes SNP data set was used because it provided better coverage of the SNP information across human genome. The distance limit was set to ±500kb of the target SNP.

The genetic information of SNPs in high LD with the target SNP was extracted from the IMAGEN imputed database, of which the SNPs were imputed using the 1000 Genomes genetic data as reference (See **Chapter 2 Section 2.4.3**). The genotypes of the imputed SNPs were displayed in the major allele dosage ranging from 0 to 2, indicating the predicted copies of the major allele.

Identify putative transcription factor binding sites at SNPs

Alterations of putative transcription factor (TF) binding sites by SNPs were studied using the Genomatix© (http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html). The DNA sequence ±50bp of SNPs were extracted and entered into the Genomatix©. The putative TF binding sites was identified based on matrix similarity. A matrix similarity of 1 means the input sequence matches with the most conserved nucleotide at each position of the consensus matrix (i.e. the sequence of the TF binding site). Any putative TF binding sites with a matrix similarity over 0.75 were reported.

6.3 Results

6.3.1 Genetic association analyses with drunkenness in adolescents

To identify the influence of stress-related genes on drunkenness, linear regression analyses were performed on 41307 tag SNPs from 2348 stress genes.

Table 6-1 displayed the five most associated SNPs from the genetic association analyses. The most significantly associated SNP, rs1409837, was located at 956 bp upstream of the FYN oncogene related to SRC (*FYN*) gene on chromosome 6 (β = -.354, $p_{uncorrected}$ = 1.31 x 10^{-7}, p_{FDR} = .005 see **Figure 6-1** for Q-Q plot). The minor allele of rs1409837 (C/T, MAF C = 0.088) was associated with reduced drunkenness in the IMAGEN adolescents at age 16. As illustrated by the Q-Q plot, the distribution of the observed –log *p*-value followed the null distribution (Genomic inflation factor λ = 1.03), suggesting the association were unlikely to be affected by any hidden population structure. To investigate the significance of the association between rs1409837 and drunkenness on alcohol consumption, genetic association analysis was performed. The results showed that the C allele of rs1409837 was associated with reduced binge drinking in the adolescents at age 16 (β = -.159, p = .016).

Table 6-1 Association of stress genes and lifetime drunkenness. Genetic association of stress genes and adolescents' information on lifetime drunkenness were performed in PLINK. After the correction of multiple testing, rs1409837 located 956bp upstream of *FYN* remained significantly associated with drunkenness in adolescents. Results of the five most associated SNPs are displayed below.

Chromosome	*Gene Symbol*	*Gene*	*SNP*	*Location*	*Beta*	*p*	*FDR BH*
6	*FYN*	FYN oncogene related to SRC	rs1409837	112302304	-0.354	1.31E-07	0.005
6	*FYN*	FYN oncogene related to SRC	rs9372316	112309728	-0.193	1.52E-05	0.314
6	*CSNK2B*	Casein Kinase II Beta Subunit	rs3130617	31735502	0.173	6.50E-05	0.604
6	*POU5F1*	POU class 5 homeobox 1	rs3094187	31234923	0.154	7.12E-05	0.604
8	*NAT1*	N-acetyltransferase 1	rs2188023	18083511	0.173	7.31E-05	0.604

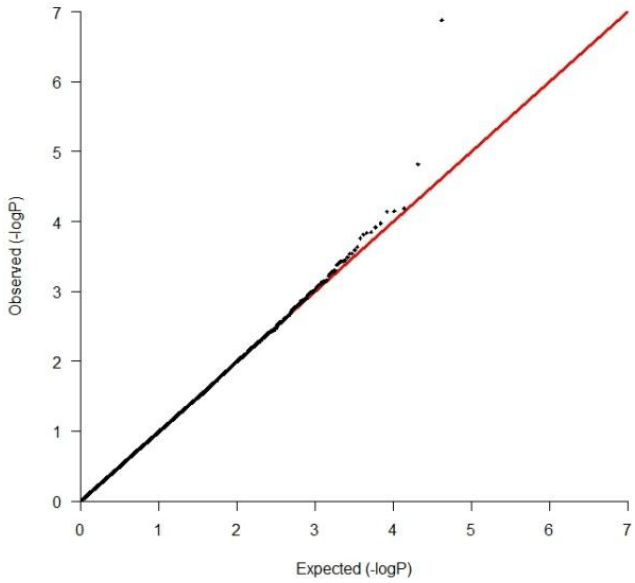


Figure 6-1 Q-Q plot for single SNP association of stress genes and lifetime drunkenness. The genomic inflation factor (λ) of the associations among 2348 stress genes and drunkenness was 1.03, indicating that the association was unlikely to be influenced by any hidden population structure.

6.3.2 Regional association analysis of rs1409837

In order to study the functional significance of rs1409837 in respect to its location on the genome, regional association analysis was conducted in 223 genotyped SNPs ±500 kbp of rs1409837. SNP rs1409837 showed the most significant association with drunkenness among all the surrounding SNPs (**Table 6-2**). The regional association analysis plot was created based on the associations ($-\log_{10}$ *p*-values) of SNPs with drunkenness (**Figure 6-2**). SNPs localised at the 5' region of *FYN* showed stronger association (i.e. larger $-\log_{10}$ *p*-values) with drunkenness compared to SNPs that were further away from the 5'upstream region of *FYN*. The results supported the genetic influence of *FYN* on the association between rs1409837 and drunkenness.

Table 6-2 Regional association analysis of 223 genotyped SNPs ±500 kbp of rs1409837 and drunkenness. Regional association analysis was conducted in 223 SNPs ±500 kbp of rs1409837 to examine the functional significance of rs1409837 in respect to its location on the genome. Results indicated that rs1409837 remained the most significant SNP associated with drunkenness. Results of the five most associated SNPs are displayed below.

Chromosome	*SNP*	*Location*	*Beta*	*p*
6	rs1409837	112302304	-0.354	1.31E-07
6	rs1998933	112324860	-0.1969	9.47E-06
6	rs9372316	112309728	-0.1925	1.52E-05
6	rs9320379	112291314	-0.1512	0.002421
6	rs2024832	112296405	-0.1465	0.003274

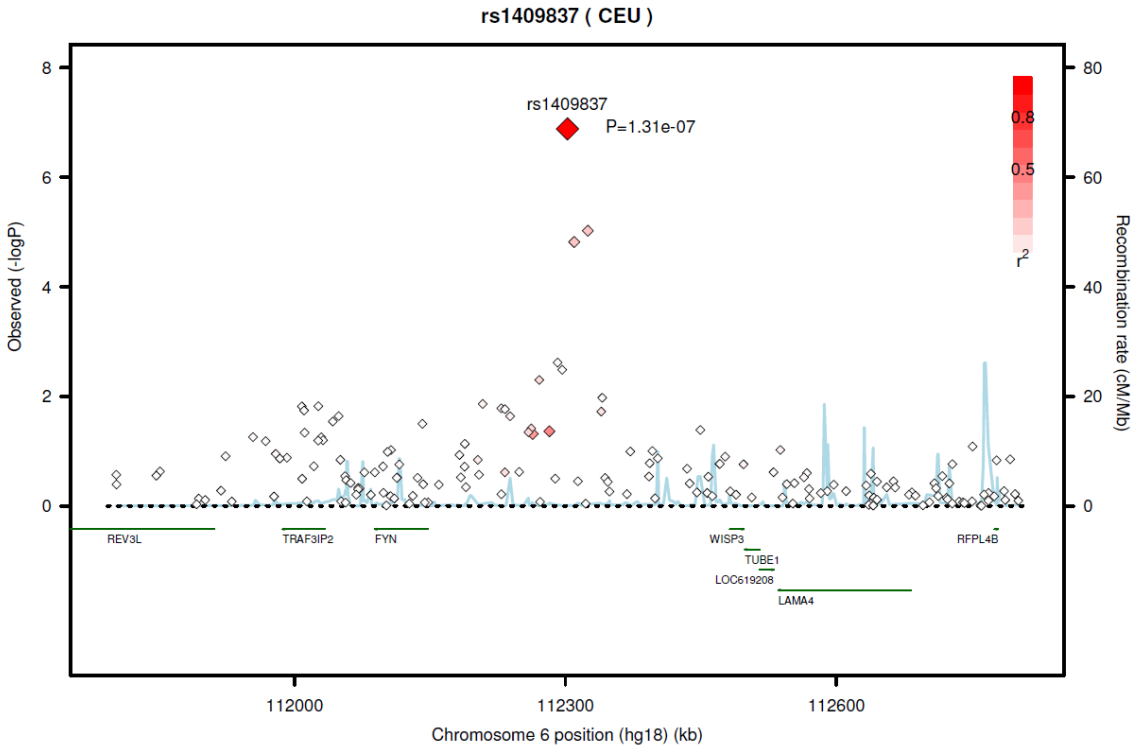


Figure 6-2 Regional association plot for rs1409837 ±500 kbp. Regional association analysis was conducted on 223 genotyped SNPs ±500 kbp from rs1409837 and drunkenness. The SNP

rs1409837 was located at the 5' upstream region of *FYN*. Using the HapMap CEU population as reference panel, the $-\log_{10} p$-values of SNPs were plotted against the chromosomal position (kb) on SNAP (http://www.broad.mit.edu/mpg/snap). Pairwise LD (r^2) of rs1409837 and SNPs were indicated by the colour shading. The recombination rate was indicated by the light-blue lines. There are three versions *FYN* gene and only the shortest version of *FYN* is displayed in this figure (see **Figure 6-3** for other versions of *FYN* genes). Results from the regional association analysis supported the genetic influence from *FYN* on the association between rs1409837 and drunkenness.

6.3.3 LD structure and *in-silico* functional characterisation of *FYN* rs1409837

6.3.3.1 Identify SNPs in LD with *FYN* rs1409837

The significant association between *FYN* rs1409837 and drunkenness could be due to various reasons. *FYN* rs1409837 could be the causal variant that reduced the likelihood of drunkenness in adolescents. Alternatively, the protective effect of *FYN* rs1409837 on drunkenness might represent an indirect association of a SNP that is in high LD with *FYN* rs1409837, or the significant results of *FYN* rs1409837 could be a false-positive that was caused by chance or any hidden population structure. As illustrated in **Section 6.3.1** and the Q-Q plot (**Figure 6-1**), the observed association of *FYN* rs1409837 was unlikely to be influenced by any hidden population structure.

Due to the tagging nature of the genotyped SNPs, SNPs that were in high LD with *FYN* rs1409837 might also contribute to the significant association between *FYN* rs1409837 and drunkenness. To dissect the nature of the association between *FYN* rs1409837 and drunkenness, *in-silico* analysis was carried out to identify the LD structure of *FYN* rs1409837. As revealed by the regional association plot (**Figure 6-2**), the genotyped SNPs ±500 kbp from *FYN* rs1409837 showed weak to modest LD ($r^2 < 0.5$) with *FYN* rs1409837.

Additional genetic information from the 1000 Genomes database was used to achieve a better SNP coverage. Using the 1000 Genomes (CEU population) genetic information available on SNAP (http://www.*broad*.mit.edu/mpg/*snap)*, six SNPs (rs72944247, rs72944248, rs17073026, rs72944257, rs72944258 and rs72944244) that were in strong LD (r^2 > 0.8) with rs1409837 were identified (**Table 6-3**).

6.3.3.2 Associations of drunkenness and imputed SNPs

The genetic associations of drunkenness and SNPs in strong LD with *FYN* rs1409837 were investigated. The drunkenness information of adolescents was regressed against the imputed major allele dosage of the six SNPs in strong LD with *FYN* rs1409837. Among all the imputed SNPs, rs72944247 displayed the strongest association with drunkenness, indicating that the reduced major allele dosage of rs72944247 was associated with the reduced drunkenness in the adolescents (**Table 6-3**). These results was consistent with the association between *FYN* rs1409837 and drunkenness reported in **Section 6.3.1**, as the minor allele of *FYN* rs1409837 was associated with reduced drunkenness in adolescents. The results from the imputed SNPs further strengthened the association between *FYN* rs1409837 and drunkenness was unlikely to be due to false positives.

Table 6-3 Associations of imputed SNPs and drunkenness. SNPs in LD (r^2 > 0.8) with *FYN* rs1409837 were identified and imputed using the 1000 Genomes database. Adolescents' information on drunkenness was regressed against the major allele dosage of the imputed SNPs. Gender and recruitment sites were included as covariates.

Imputed SNPs	*Distance away from rs1409837 (bp)*	r^2	*Allele*	*MAF*	*Beta*	*p*
rs72944244	-2764	0.85	C/T	C = 0.066	.304	5.86E-06
rs72944247	+2741	1.00	A/G	A = 0.066	.310	3.35E-06
rs72944248	+4059	1.00	A/G	A = 0.067	.310	3.50E-06
rs17073026	+9036	1.00	A/G	A = 0.071	.307	4.20E-06
rs72944257	+12931	1.00	A/G	G = 0.051	.304	8.32E-06
rs72944258	+13023	1.00	G/T	G = 0.051	.304	8.32E-06

6.3.3.3 Putative transcription factor (TF) binding sites in rs1409837 and imputed SNPs

Since rs1409837 is located at the 5' upstream region of *FYN*, gene regulatory elements e.g. TF binding sites located at rs1409837 and the SNPs in strong LD with rs1409837 might underlie mechanisms contributing to the association with drunkenness. *In silico* analysis was carried out to identify the putative TF sites at *FYN* rs1409837 and the six imputed SNPs. Using the Genomatix© software, alterations of putative TF binding sites due to the presence of the minor alleles at *FYN* rs1409837, as well as the imputed SNPs rs72944247, rs72944248, rs17073026, rs72944257, rs72944258 and rs72944244 were investigated.

Table 6-4 listed the predicted alterations in TF binding sites at *FYN* rs1409837 and the six SNPs in strong LD with *FYN* rs1409837. The minor allele of *FYN* rs1409837 was found to replace the putative TF binding site for SRY-related high mobility group box (SOX) factors by ETS1 factors. Stronger candidates were observed in the imputed SNPs, in which the minor allele of rs72944247 could create a putative TF binding site for forkhead domain factors (matrix similarity = 1.000). The minor allele of rs72944258 might remove the putative TF binding site for two-handed zinc finger homeodomain (matrix similarity = 0.981) and create an additional binding site for MAF/ AP1 (matrix similarity = 0.999). The alterations in putative TF binding sites at *FYN* 1409837 and the imputed SNPs might indicate the functional significance of these SNPs on drunkenness at molecular level.

Table 6-4 Alterations of putative TF binding sites by *FYN* rs1409837 and imputed SNPs. *In silico* analysis was conducted in the Genomatix© to investigate the putative TF binding sites at *FYN* rs1409837 and the imputed SNPs that were in strong LD with *FYN* rs1409837. The putative TF sites and their binding sequences created by the minor alleles were listed under 'New TF binding sites'. Putative TF sites that could be removed due to the presence of the minor alleles were listed under 'TF binding sites removed'. The putative TF binding sites was identified using matrix similarity. A matrix similarity of 1 means the input sequence matches with the most conserved nucleotide at each position of the consensus matrix. Only putative TF sites with matrix similarity (MS) >0.75 were reported.

	New TF binding sites			***TF binding sites removed***		
SNPs	***TF family***	***MS***	***Details***	***TF family***	***MS***	***Details***
						SOX/SRY-sex/testis determining and related high mobility group box factors
rs1409837	V$ETSF	0.794	Human & murine ETS1 factors (5'-AGGAAG-3')	V$SORY	0.771	(5’-CCAAAGT-3’)
rs72944244	V$GABF	0.812	GA-boxes (5’-GAGAG-3’)	V$GATA	0.962	GATA binding factors (5’-GATAA-3’)
	V$ABDB	0.853	Abdominal-B type homeodomain transcription factors (5’-CTCTCAA-3’)	V$HNF1	0.805	Hepatic Nuclear Factor 1 (5’-GTTA-3’)

	V$LEFF	0.867	LEF1/TCF (5’-ATCAAA-3’)			
rs72944247	V$FKHD	1.000	Forkhead domain factors (5’-AACA-3’)			
						Krueppel like transcription factors
rs72944248	V$HOXC	0.888	HOX - PBX complexes (5’-GATAT-3’)	V$KLFS	0.970	(5’-AGGGT-3’)
						Nucleoside diphosphate kinase
				V$NDPK	0.895	(5’-GGGAGGG-3’)
			SOX/SRY-sex/testis determining and related			
rs17073026	V$SORY	0.941	HMG box factors (5’-ATT-3’)			
			Distal-less homeodomain transcription factors			
	V$DLXF	0.914	(5’-TAATTT-3’)			
	V$HBOX	0.952	Homeobox transcription factors (5-TAAT-3’)			

rs72944257	V$PARF	0.853	PAR/bZIP family (5’-TAA-3’)	V$ABDB	0.924	Abdominal-B type homeodomain transcription factors (5’-ATAAAA-3’)
	V$PRDF	0.819	Positive regulatory domain I binding factor (5’-AGTAAAAGT-3’)	V$CDXF	0.887	Vertebrate caudal related homeodomain protein (5’-ACTTTTAT-3’)
				V$FKHD	0.865	Fork head domain factors (5’-AATAAAA-3’)
rs72944258	V$AP1R	0.999	MAF and AP1 related factors (5’-TGCTGA-3’)	V$ZFHX	0.981	Two-handed zinc finger homeodomain transcription factors (5’-GTTT-3’)

6.3.4 *FYN* rs1409837 and mRNA expressions

Alterations of TF binding sites at *FYN* rs1409837 and at the SNPs in strong LD with *FYN* rs1409837 might potentially disrupt the transcriptional processes and the mRNA expressions. To test this hypothesis, the association between rs1409837 and *FYN* mRNA levels was investigated.

The mRNA expressions of *FYN* in the peripheral blood were measured by four probes on the Illumina© expression microarray in 628 IMAGEN adolescents (**Figure 6-3**). Two probes located at the 3' untranslated region (probe IDs: ILMN_1686555 and ILMN_2380801) were excluded from the analyses due to the presence of SNPs and non-specific binding. The sequences of probes ILMN_2249920 and ILMN_1781207 showed 100% binding to the *FYN* gene and did not contain any SNPs. Therefore, the *FYN* mRNA expressions measured by ILMN_2249920 and ILMN_1781207 were used for analyses. Due to the high correlations of *FYN* rs1409837 and the imputed SNPs, only the associations between *FYN* rs1409837 and *FYN* mRNA levels were analysed. As illustrated in **Table 6-5**, the additive genetic effect of *FYN* rs409837 was not associated with the mRNA expressions of *FYN* (ILMN_2249920: β = -.018, *C.I.* = [-.097 – .060], p = .649; ILMN_1781207: β = -.005, *C.I.* = [-.060 – .050], p = .868).

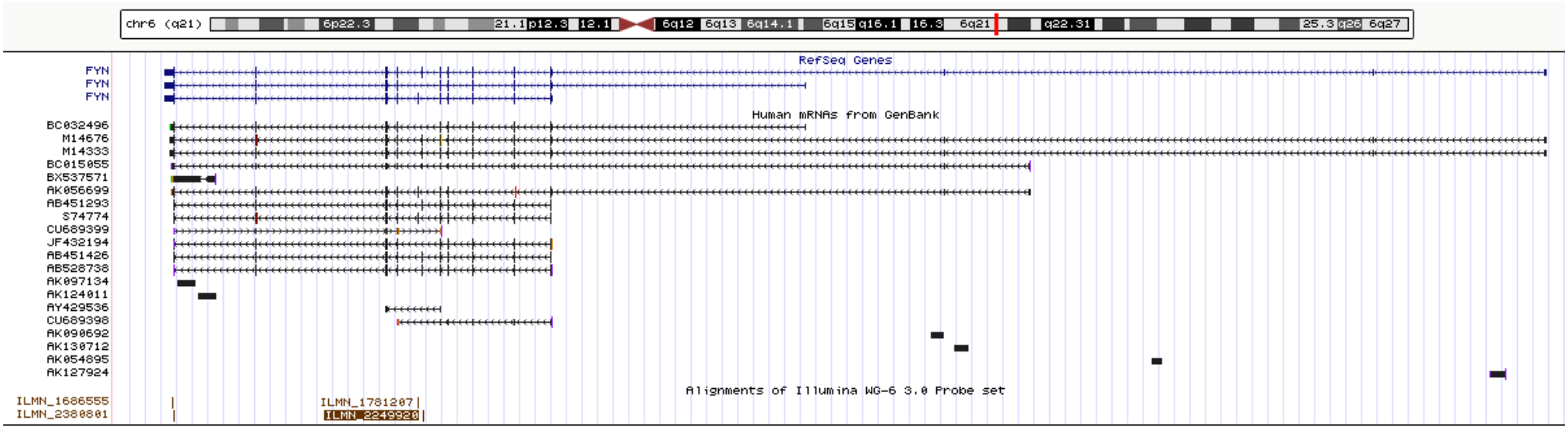


Figure 6-3 Locations of the *FYN* mRNA expression probes. The locations of four *FYN* mRNA expressions probes were indicated on the UCSC genome browser (Hg18). There were three versions of *FYN* RefSeq genes. Various forms of *FYN* mRNA isoforms in human were also displayed in this figure.

Table 6-5 Associations of rs1409837 and *FYN* mRNA expressions. The *FYN* mRNA expressions in peripheral blood were measured by two probes on the Illumina© expression microarray.

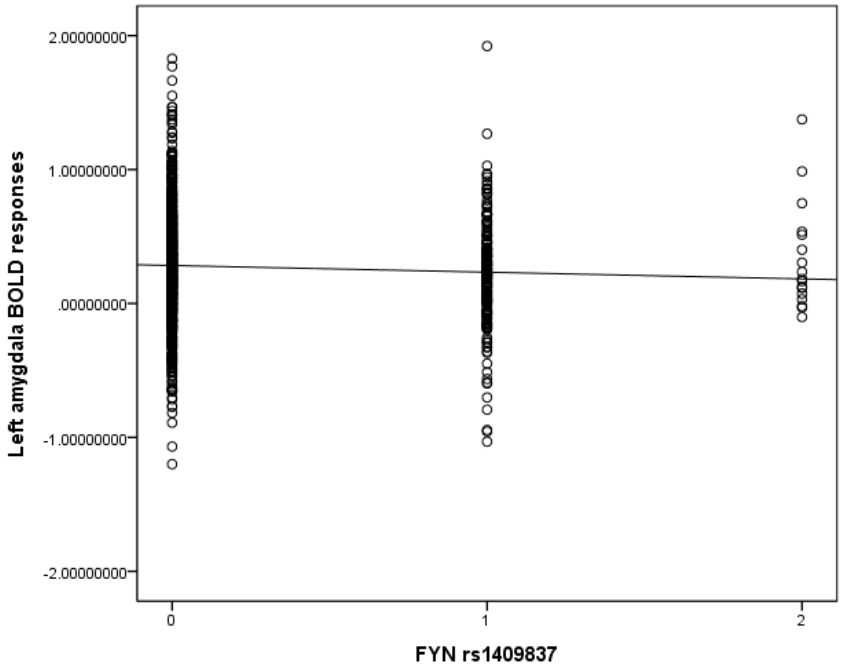
FYN mRNA probes		***FYN rs1409837***	
(Location)	***Beta***	***95% C.I.***	***p***
ILMN_2249920	-.018	[-.097 – .060]	.649
(Chr6: 112128094 – 112128143)			
ILMN_1781207	-.005	[-.060 – .050]	.868
(Chr6: 112127414 – 112127463)			

6.3.5 *FYN* rs1409837, amygdala BOLD responses and life stress

6.3.5.1 *FYN* rs1409837 predicts amygdala BOLD responses but not life stress

In order to understand the relationship between *FYN* rs1409837 and stress processing in the brain, region-of-interest analyses were conducted to examine the association between *FYN* rs1409837 and the amygdala activations during angry face processing. The amygdala is one of the most activated brain regions during emotional processing and the exposure to stressful situations (Phan et al., 2002, Whalen et al., 1998). Emotional processing has been frequently measured by the face task, in which the presentation of fearful or angry faces has been found to alter the amygdala activations (Whalen et al., 1998). The minor allele of *FYN* rs1409837 was significantly associated with the reduced amygdala BOLD responses during angry face processing (Left amygdala: β = -.053, *C.I.* = [-.097 to -.008], p = .020; Right amygdala: β = -.045, *C.I.* = [-.086 to -.004], p = .033; **Figure 6-4**). The results suggested that *FYN* rs1409837 might predict the sensitivity towards emotional stimuli in adolescents.

a)



b)

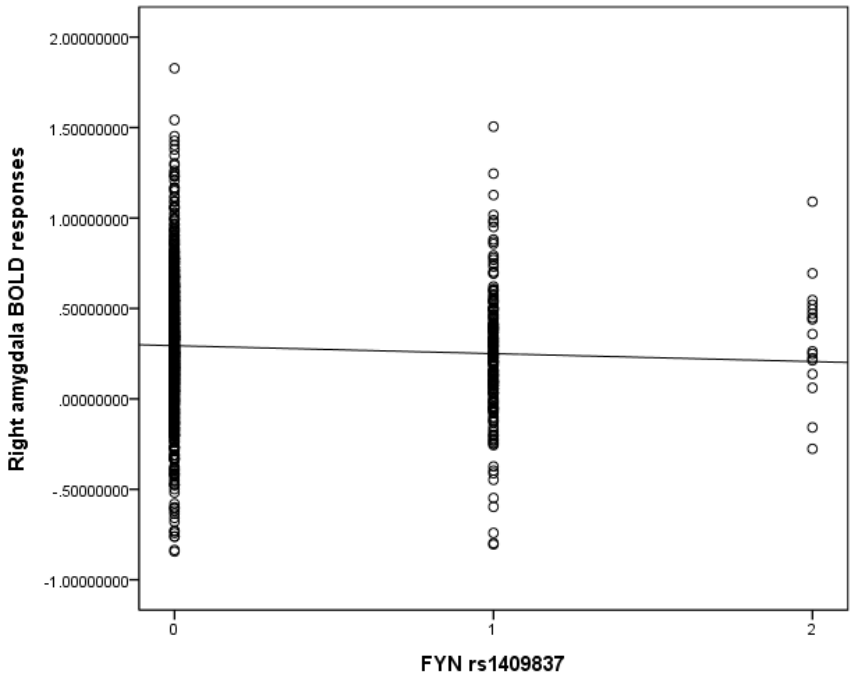


Figure 6-4 *FYN* rs1409837 and the amygdala BOLD responses in during angry face processing. The BOLD responses of **a)** the left amygdala and **b)** the right amygdala in the 'angry faces *versus* control' contrast were regressed against the copies of minor allele of *FYN* rs1409837.

Next, the relationship between *FYN* rs1409837 and self-report life stress was investigated. Life stress was measured by both lifetime frequency and rating towards the negative life events in adolescents at age 16. Whilst the rating towards negative life events could indicate the perceived level of life stress in the adolescents, the frequency of negative life events might quantify life stress in a less subjective way. The results showed that *FYN* rs1409837 was not associated with neither the frequency nor the ratings towards negative life events (Frequency: β = .171, *C.I.* = [-.252 to .593], p = .429; Ratings: β = -.215, *C.I.* = [-.823 to .393], p = .488). The results suggested that *FYN* rs1409837 did not have any direct effect on life stress.

6.3.5.2 Amygdala BOLD responses predict the ratings but not the frequency of negative life events

The relationship between the amygdala activations during angry face processing and life stress was studied. Regression models were applied to investigate whether the amygdala BOLD responses could predict life stress reported by the adolescents. Adolescents who gave lower ratings to negative life events (i.e. perceived greater unhappiness) had shown higher amygdala BOLD responses (Left amygdala: β = -.794, *C.I.* = [-1.46 to -.127], p = .020; Right amygdala: β = -.772, *C.I.* = [-1.50 to -.045], p = .038), indicating that the activations of the amygdala increased when adolescents experienced higher level of perceived life stress. However the amygdala BOLD responses did not predict life stress measured by the lifetime frequency of negative life events (Left amygdala: β = .210, *C.I.* = [-.267 to .686], p = .389; Right amygdala: β = .307, *C.I.* = [-.209 to .823], p = .244). Together the evidence suggested that the amygdala activation could be important in monitoring the perception towards life events and stress in the adolescents.

6.3.5.3 Life stress increases the likelihood of drunkenness

To investigate the relationship between life stress and drunkenness, the associations between the amygdala activations during angry face processing and negative life events on drunkenness in

adolescents were performed. The results showed that the BOLD responses of the amygdala did not associate with drunkenness (Left amygdala: β = .117, *C.I.* = [-.048 to .282], p = .166; Right amygdala: β = .102, *C.I.* = [-.077 to .281], p = .263). On the other hand, both the frequency and ratings towards negative life events were significantly associated with increased drunkenness (Frequency: β = .160, *C.I.* = [.102 to .218], $p = 6.56 \times 10^{-8}$; Ratings: β = -.071, *C.I.* = [-.113 to -.028], p = .001). Together the results indicated that adolescents who had higher level of life stress were more likely to experience drunkenness, yet individuals' sensitivity towards emotional stimuli did not directly predict drunkenness.

6.4 Discussion

Summary of findings

In this study, the genetic influence of stress-related genes and alcohol misuse in adolescents was investigated. The minor allele of *FYN* rs1409837 was found to associate with reduced drunkenness and bingeing in the 16-year-old IMAGEN adolescents. In order to understand the relevance of *FYN* rs1409837 and stress, further analyses on *FYN* rs1409837, amygdala activations during angry face processing and negative life events were performed. The results indicated that *FYN* rs1409837 was significantly associated with the reduced bilateral amygdala activations. Although individuals with high amygdala activation were more likely to give lower ratings to negative life events and recognise the events as more stressful, *FYN* rs1409837 did not directly associated with the negative life events. Both the frequency and ratings towards negative life events could predict the increased drunkenness in the adolescents, confirming the role of life stress in alcohol misuse.

6.4.1 Protective effect of *FYN* rs1409837 on drunkenness

Results from this study demonstrated the protective role of *FYN* rs1409837 on drunkenness in the IMAGEN adolescents. The minor allele of *FYN* rs1409837 was found in 8.8% of the IMAGEN adolescents and in 5.3% of the HapMap CEU population. The low MAF of *FYN*

rs1409837 suggested that its protective effect on drunkenness could only be observed in a small number of individuals. Defining the LD of SNPs using r^2 had taken the MAFs into account, therefore low MAFs were also observed in the six SNPs showing high LD with *FYN* rs1409837. Consistent with the *FYN* rs1409837 and drunkenness association, increase major allele dosages of the imputed SNPs were associated with increased drunkenness in the adolescents. Similar beta and *p* values were observed in the associations of the imputed SNPs and drunkenness; the results further supported the genetic influence of *FYN* rs1409837 and its tagged region on drunkenness.

The functional significance of *FYN* rs1409837 and the imputed SNPs on gene transcriptions and mRNA expressions was examined. *In-silico* analyses showed that the minor allele of *FYN* rs1409837 might replace the putative TF binding sites for SOX family by the ETS1 family. However, these putative TF binding sites displayed low matrix similarities (below 0.80), implying that the prediction for TF binding sites was not highly confident. Alterations in TF families among SNPs in high LD with *FYN* rs1409837 were also suggested. For example, the putative TF binding sites for forkhead domain family and MAF/AP1 related factors were created in the minor allele carriers of rs72944247 and rs72944258. Both of these putative TF binding sites showed high matrix similarities. Previous studies indicated the involvement of forkhead domain family and MAF/AP1 related factors in stress regulations. Members of the forkhead domain were previously found to regulate stress responses in the pituitary (Belgardt et al., 2008), whereas members of the MAF/AP1 family were involved in regulating cellular and oxidative stress responses (reviewed by Suzuki et al. (2005)). Based on these results, it could be envisaged that *FYN* rs1409837 and the imputed SNPs might influence the molecular mechanisms of drunkenness via alterations of TF binding sites. As some of these putative TF binding sites might be responsible for stress regulations, suggesting the association between the tagged region of *FYN* rs1409837 and drunkenness could also be mediated by stress-related regulatory mechanisms. Further analysis will be required to confirm the role of TF binding sites on *FYN* and drunkenness.

Due to the availability of gene expression data, the association between rs1409837 and *FYN* mRNA expressions in the peripheral blood was only investigated in a subset of participants. The lack of associations suggested no cis-acting effect of rs1409837 on *FYN* mRNA expressions. As shown by the UCSC genome database, the *FYN* mRNA could exist in several isoforms. However, the *FYN* probes were designed to cover multiple isoforms; thus hinder the detection of any cis-acting effects of *FYN* rs1409837 on individual mRNA isoforms. Any post-transcriptional activities such as alternative splicing of precursor mRNA could also impact on the detection of *FYN* mRNA expressions. Nevertheless, gene expressions could be specific to gender, stages of puberty and tissues (Ojeda et al., 2010, Cheung and Spielman, 2009). Therefore the peripheral blood mRNA expression might not fully reflect the mRNA expressions in other tissues such as the brain and the liver. Essentially, the *FYN* mRNA expressions were measured in the adolescents at age 14. Since the association between *FYN* rs1409837 and drunkenness was observed in the adolescents at age 16, any genetic effect of *FYN* rs1409837 on mRNA expressions might not be detectable either. Further analysis will be required to confirm the cis-acting effect of *FYN* rs1409837 on mRNA expressions.

6.4.2 FYN, NMDA receptor and drunkenness

FYN is a non-receptor protein tyrosine kinase that belongs to the Src family. It is widely distributed in the central nervous system throughout development (Umemori et al., 1992). FYN is known to mediate the activity of N-methyl-D-aspartate (NMDA) receptor. Upon the activation of the cAMP-dependent PKA signalling pathway, the release of FYN increases phosphorylation of the NMDAR2B subunit (equivalent to the NMDARε2 subunit in mice) of the NMDA receptor (Yaka et al., 2003). Phosphorylation of NMDA receptor by FYN enhances channel activity, and subsequently promotes calcium ion influx and downstream NMDA receptor signalling (Yaka et al., 2003). These changes are essential for modulating long-term potentiation and glutamate transmission in the brain (Yaka et al., 2003, Woodward et al., 2006).

Ethanol is known as a powerful inhibitor of NMDA receptors. Acute exposure to ethanol was found to reduce excitatory postsynaptic potentials (EPSPs) in the NMDA receptors (Lovinger et al., 1990). Upon the injection of ethanol, *Fyn*-deficient mice showed reduced NMDARε2 phosphorylation and EPSPs (Miyakawa, 1997). The results illustrated the involvement of FYN in mediating the sensitivity to ethanol through NMDA receptors. The genetic influence of *FYN* on risks of alcohol dependence and schizophrenia was demonstrated in previous studies (Schumann et al., 2003, Ishiguro et al., 2000, Pastor et al., 2009). Consistent with previous research, results from the current study supported the protective role of *FYN* on alcohol misuse and intoxication in adolescents.

6.4.3 FYN, amygdala activations and negative life events stress

The relationship between FYN and stress was demonstrated at cell signalling and neurotransmission levels in the previous literature. For example, glucocorticoids were found to dissociate the FYN/ T-cell receptor complex and inhibit downstream T-cell receptor signalling (Lowenberg et al., 2006). Apart from the glucocorticoids, additional candidates such as serotonin receptors are responsible for monitoring the activity of the stress system and stress-related alcohol consumption (reviewed by Koob (2008)). A link between serotonin receptors and *Fyn* has also been found, in which the activation of the serotonin (5-hydroxytryptamin, 5-HT_6) receptors was found to enhance autophosphorylation of Fyn and upregulate the surface expression of 5-HT_6 receptors in the rat brain (Yun et al., 2007). FYN is involved in regulating oxidative and chemical stress, as well as altering the immunological response in rats suffered from traumatic stress (Niture et al., 2011, Goebel-Goody et al., 2012, Xiao et al., 2009). The evidence for *FYN* in stress-related neuropsychiatric disorders has remained sparse in the current literature, implying that the selection of stress genes from the NCBI Gene Database has captured genes that are less studied in the psychiatric research.

In addition to the genetic association with drunkenness, *FYN* rs1409837 was also associated with the reduced BOLD responses in the bilateral amygdala during angry face processing. The additive genetic effect of *FYN* rs1409837 was specific to the amygdala activations and *FYN* rs1409837 had no direct effect on life stress measured by the frequency and ratings towards negative life events in the adolescents. Consistent with Whalen et al.'s (1998) findings, reduced amygdala activations were associated with the reduced fearfulness towards emotional stimuli. In the current study, adolescents who demonstrated lower amygdala BOLD responses also perceived negative life events as less unhappy or less stressful. During adolescence, the maturation of limbic brain regions including the amygdala and the hippocampus precedes the cortical regions (Gogtay et al., 2006, Giedd et al., 1996b). *FYN* rs1409837 might be involved in stabilising the NMDA receptor activity in the amygdala and maintained the NMDA receptor functions.

Life stress characterised by frequency and ratings towards negative life events was found to increase drunkenness in the adolescents. The results were consistent with previous studies, in which increased alcohol craving and relapse were linked to emotional stress among adolescent drinkers and patients with alcohol dependence (Pilowsky et al., 2009, Blomeyer et al., 2008, Sinha et al., 2009). Since the amygdala activations could also predict the subjective perception of life stress, such association might indicate the impact of the amygdala activations on stress-related alcohol consumption and drunkenness in the adolescents. Further investigation will be required to understand the gene-environment interaction of *FYN* rs1409837 on drunkenness during stressful environment.

6.4.4 Limitations of the study

In attempt to understand the functions of *FYN* rs1409837 on drunkenness, additional bioinformatic resources could be explored to provide a more comprehensive view. For example, the interaction between *FYN* rs1409837 and TF binding sites could also be studied using data

from the ENCODE project (The ENCODE Project Consortium, 2012). The ENCODE project provides open resources for functional elements encoded in the human genome. It allows the enrichment of *FYN* rs1409837 with functional elements residing within the ENCODE-defined regions and complements the predicted TF alterations using MatInspector.

The face task is a classical fMRI paradigm for measuring emotional processing. However, brain activity measured in the face task might not fully reflect individuals' emotional sensitivity towards stressful situations. Firstly, depending on individuals' attention and fatigue level, the recorded brain activations could be subject to the habituation of stimuli over time. Counterbalancing the order of stimuli presented in the face task might reduce chances of habituation. Secondly, although the amygdala has been regarded as the key brain region of emotional processing, the associations between *FYN* rs1409837 and amygdala activation could also be driven by face recognition. Further analyses using the amygdala activations during happy face processing or non-biological threatening stimuli could be informative in confirming the involvement of *FYN* rs1409837 in emotional processing. It should be noted that the fMRI assessment was conducted in the IMAGEN adolescents at age 14; whereas the drunkenness and negative life events measurements in adolescents at age 16 were used for analyses in this study. Despite the limitations, the analyses using the amygdala activations in the face task would allow researchers to understand the relationship between *FYN* rs1409837 and the neurobiology of emotional processing and stress reactivity.

Similarly, the *FYN* mRNA expressions were only measured in a subset of adolescents at age 14. To improve the accuracy of the mRNA expressions, techniques including the real-time PCR should be applied to measure the mRNA expressions of *FYN* isoforms. Ideally, the mRNA expressions of *FYN* should be measured in the adolescents at age 16. Validation analysis can be carried out in additional adolescent populations where alcohol consumption, genetic and gene expression data are available. Additional analysis should be carried out to examine the

association between *FYN* rs1409837 and brain mRNA expressions. Using published resources such as the BrainCloud, the *cis*-acting effects of *FYN* rs1409837 in the brain can be investigated in age and/ or gender-specific manner (Colantuoni et al., 2011). At present, the BrainCloud database only provides mRNA expressions in the prefrontal cortex. Therefore, further investigation will be required to study the significance of *FYN* rs1409837 on the mRNA expression in the amygdala.

Additional work should be also carried out to validate the associations of *FYN* rs1409837, drunkenness and the amygdala activations. Despite the high accuracy of genotyping and imputed SNP data (Huang et al., 2009, Halperin and Stephan, 2009), further analyses including sequencing and functional characterisation will be required to confirm the nature of association and the functions of *FYN* rs1409837. Rigorous replications should be performed in other adolescent populations to confirm the associations.

6.4.5 Conclusions

By focusing on genes in the stress system, this study has successfully identified the protective effect of *FYN* rs1409837 on drunkenness and bingeing in adolescents. One possible mechanisms of *FYN* rs1409837 is to reduce the amygdala activations during emotional processing and stressful situations, which may further influence the perceived level of life stress. This study has presented the first evidence for the genetic effect of *FYN* on both alcohol drunkenness and brain activations in human. The functional mechanisms of *FYN* rs1409387 are yet to be determined.

Chapter 7

General Discussion

7.1 Overview

My thesis aimed to identify the genetic and epigenetic risk factors for repeated substance use and misuse in adolescents. Substance use in adolescence can be served as an important milestone for developing substance addictions (Grant and Dawson, 1997, Giovino et al., 1995). The importance of using multidisciplinary approaches to dissect the genetic underpinnings of neuropsychiatric disorders has been emphasised in recent years (Munoz et al., 2009). On this basis, I adopted various approaches to integrate the genetic, epigenetic, neuroimaging and behavioural data in order to understand the substance use behaviour in the IMAGEN adolescents. In Chapters 3, 4 and 5, I investigated the role of life stress, circadian genes and reward sensitivity in alcohol and tobacco use in adolescents. In Chapter 6, I focused on identifying the gene-neuroimaging relationship of *FYN* rs1409837 on drunkenness and the amygdala activations in the adolescents.

Table 7-1 displays the summary of hypotheses and findings in this thesis. To summarise, the results presented in this thesis have supported the relevance of life stress and reward sensitivity in predicting the increased alcohol and tobacco use in adolescents. Evidence from Chapter 6 partially supported the role of stress system on drunkenness and binge drinking behaviour. However, the interactions between stress and circadian systems did not influence reward sensitivity and substance use in adolescents.

Table 7-1 Summary of hypotheses and results.

Research questions and hypotheses		*Specific Results*
Research question: Negative life events and disruptions of the stress system increase substance use in adolescents		
Hypotheses:		
Increase life stress promotes alcohol and tobacco use in adolescents (Chapter 3)	✓	Adolescent drinkers and smokers displayed higher level of life stress at age 14 and age 16. Both drinkers and smokers experienced more negative life events and gave lower ratings to the events compared to non-drinkers and non-smokers.
Genetic polymorphisms of the stress system predict alcohol and tobacco use in adolescents (Chapter 5)	x	Stress genes were not associated with alcohol and tobacco use in adolescents at single SNP and polygenic levels.
Genetic polymorphisms of the stress system predict alcohol misuse in adolescents (Chapter 6)	✓	*FYN* rs1409837 from the stress gene list was associated with reduced drunkenness and bingeing in adolescents at age 16.

	✗	No cis-acting effect of *FYN* rs1409837 on mRNA expressions.
Life stress increases risks of alcohol misuse (Chapter 6)	✓	*FYN* rs1409837 was significantly associated with the reduced bilateral amygdala activations during angry face processing.
	✗	*FYN* rs1409837 was not associated with the frequency and ratings towards negative life events.
	✓	Adolescents who had higher level of life stress were more likely to experience drunkenness.
Research question: Disruptions of the circadian system promote substance use in adolescents		
Hypotheses:		
Inter-individual variations in *PER1* DNA methylation predict alcohol	✗	No associations between the DNA methylation at *PER1* upstream CpG units with

and tobacco use in adolescents (Chapter 4)		alcohol and tobacco use in adolescents.
Genetic polymorphisms of the circadian system predict alcohol and tobacco use in adolescents (Chapter 5)	x	Circadian genes were not associated with alcohol and tobacco use in adolescents at both single SNP and polygenic levels.
Research question: Disruptions of both stress and circadian systems promote substance use in adolescents		
Hypotheses:		
Life stress alters DNA methylation in *PER1* (Chapter 4)	x	Frequency and ratings toward negative life events did not predict the DNA methylation at *PER1* upstream CpG units.
DNA methylation at the *PER1* gene regulatory elements predict alcohol and tobacco use in adolescents (Chapter 4)	x	Lack of inter-individual variations in DNA methylation at the *PER1* proximal GRE site.

	✗	The DNA methylation at *PER1* rs3027172 and the E-box could not be detected using the Sequenom platform.
Genetic polymorphisms of stress x circadian genes predict alcohol and tobacco use in adolescents (Chapter 5)	✗	Stress x circadian genes were not associated with alcohol and tobacco use in adolescents at both single SNP and polygenic levels.
Research question: Sensitivity to reward predicts substance use in adolescents		
Hypothesis:		
VS activations during reward anticipation predicts alcohol and tobacco use in adolescents (Chapter 3)	✓	Adolescent drinkers showed significantly higher bilateral VS activations compared to non-drinkers at age 14.
	✓	Adolescent smokers demonstrated significantly higher right VS activations compared to non-smokers at age 14.

	x	No difference in left VS activations was found between smokers and non-smokers.
Research question: Disruptions in the stress and circadian systems alter reward sensitivity in adolescents at age 14		
Hypotheses:		
Increase life stress alters reward sensitivity in adolescents (Chapter 3)	x	Neither the frequency nor ratings towards negative life events predicted the bilateral VS activations during reward anticipation in the 14-year-old adolescents.
Inter-individual variations in *PER1* DNA methylation reward sensitivity in adolescents (Chapter 4)	x	Lack of significant associations between the DNA methylation at *PER1* upstream CpG units and the bilateral VS activations during reward anticipation in adolescents.
Genetic polymorphisms of stress genes predict reward sensitivity in	x	Stress genes were not associated with the bilateral VS activations during reward

adolescents (Chapter 5)		anticipation at both single SNP and polygenic levels.
Genetic polymorphisms of circadian genes predict reward sensitivity in adolescents (Chapter 5)	x	Circadian genes were not associated with the bilateral VS activations during reward anticipation at single SNP and polygenic levels.
Genetic polymorphisms of stress x circadian genes predict reward sensitivity in adolescents Chapter 5)	x	Stress x circadian genes were not associated with the bilateral VS activations during reward anticipation at single SNP and polygenic levels.

7.2 Life stress and reward sensitivity as risk factor for substance use and misuse

7.2.1 Life stress promotes substance use and misuse in adolescents

The exposure to environmental stress has been implicated in various stages of addiction, from the initiation to repeated substance use, bingeing and withdrawal (Blomeyer et al., 2008, Booker et al., 2008, Sinha et al., 2009, Vendruscolo et al., 2012, Feltenstein et al., 2012). Part of this thesis investigated the impact of life stress, characterised by both the frequency and ratings towards negative life events, on alcohol and tobacco use in adolescents. Consistent with the previous studies, Chapter 3 illustrated that both alcohol and tobacco use in the IMAGEN adolescents were driven by the negative life events. The variance explained for negative life events ranged between 1.7 - 9.3% at age 14 and between 2.7 - 11.4% at age 16 (see Chapter 3), demonstrating that life stress could become increasingly important in predicting substance use throughout adolescence. The increased variance explained of negative life events has supported the addictions cycle proposed by Koob and Le Moal (1997) and the development of substance misuse and addictions can be maintained by negative emotions. Negative emotions due to stressful environment encourage adolescents to consume alcohol and tobacco in order to alleviate stress. The spiralling effect of stress-induced substance use may contribute to the strong association between negative life events and the increased substance use in the adolescents at age 16, which may have profound consequence in substance use behaviour in later life.

7.2.2 Reward sensitivity predicts substance use

In addition to life stress, this thesis investigated the impact of reward sensitivity on substance use in the IMAGEN adolescents. Proposed by Koob and Le Moal's (2007) model of addiction cycle, the sensations and experience associated with substance use could be rewarding and therefore encouraged repetitive substance use (Koob and Le Moal, 1997). Analyses from Chapter 3 supported the hyperactivation hypothesis of reward processing (Galvan, 2010, Galvan

et al., 2006). As a result, adolescent drinkers and smokers who demonstrated larger sensitivity to reward might engage in sensation and reward seeking behaviour more frequently; therefore they were more likely to consume alcohol and tobacco. Results from this thesis supported the use of neuroimaging phenotypes in dissecting the neural basis of reward processing.

7.3 Stress genes as the risk factor for alcohol misuse in adolescents

Based on the associations of negative life events and substance use presented in Chapter 3, I examined the additive genetic and polygenic effects of stress genes on alcohol and tobacco use in adolescents. Contrary to the findings from previous studies (Blomeyer et al., 2008, Tobacco and Genetics, 2010), the results reported in Chapter 5 suggested lack of associations between stress genes and substance use in the IMAGEN adolescents. The lack of significant associations could be due to various reasons, e.g. the detection of risk variants was dependent on the MAFs as well as the penetrance of SNPs (Reich and Lander, 2001, Sullivan et al., 2012). Based on the CDCV hypothesis, the genetic effects of common variants on substance use phenotypes are likely to be very small. Applying stringent *p*-value threshold to reduce false positives further prevented the detection of genetic risk variants within the stress system. Many published findings were based on *a priori* hypotheses in which genes related to stress regulations were selected for genetic association analyses. The correction of multiple testing for 41307 SNPs might hinder the detection of any significant associations between stress genes and substance use.

Nevertheless, it is expected that gene-environment interaction can influence the detection of genetic risk variants from the stress system on substance use. This hypothesis was partly supported by the increased variance explained of life stress on substance use in the 16-year-old adolescents (Chapter 3), as well as the increased variance explained of stress genes *en masse* on alcohol use in the 16-year-old adolescents (Chapter 5). Since life stress is a crucial factor for

propagating the addiction cycle, it can be expected that the genetic influence from the stress system will be of larger significance among heavy alcohol and tobacco users.

Therefore, Chapter 6 focused on identifying the associations between stress genes and drunkenness as an indicator of alcohol intoxication. The minor allele of *FYN* rs1409837 was found to reduce the likelihood of drunkenness and bingeing in the 16-year-old adolescents. In attempt to assess the stress property of *FYN* rs1409837, it was discovered that *FYN* rs1409837 was associated with the reduced bilateral amygdala activations in angry face processing. The likelihood of drunkenness was also influenced by increased life stress. Together with Chapter 3, the results presented in Chapter 6 have confirmed the involvement of life stress at different stages of addiction. In addition to the previous studies where the genetic influence of *FYN* on alcohol addictions has been demonstrated (Schumann et al., 2003, Pastor et al., 2009), Chapter 6 has presented the first genetic association of *FYN* rs1409837 and drunkenness in healthy adolescents; of which the findings can be important for reducing the missing heritability of alcohol addictions.

As FYN is known for regulating the activations of NMDA receptors via phosphorylation (Miyakawa, 1997, Yaka et al., 2003), the protective effect of *FYN* rs1409837 on drunkenness may involve in maintaining the NMDA receptor activity in the amygdala and the brain stress system. Illustrated by Vendruscolo and colleagues (2012), binge drinking and alcohol withdrawal could lead to differential *GR* mRNA expression in limbic regions that might alter stress reactivity of the brain stress system (Vendruscolo et al., 2012). Further confirmed by the associations between the amygdala activations and negative life events, increased bilateral amygdala activations were found to elevate the perceived life stress. Whilst the role of *FYN* rs1409837 on stress reactivity requires further investigation, my analyses have demonstrated the potential influence of *FYN* rs1409837 in alcohol misuse and addictions.

7.4 Stress did not alter reward sensitivity

Proposed by Koob and colleague's research, the development of repeated substance use and addictions is mediated by positive reinforcement associated with the pleasurable experience and sensation of substance use, as well as the negative reinforcement which promotes the maintenance of substance use to alleviate the negative emotions (Koob and Le Moal, 1997, Koob, 2013). Since both negative life events and the VS activations during reward anticipation were found to predict alcohol and tobacco use during adolescence, I was motivated to investigate the effect of life stress on reward processing. The lack of significant influence of stress genes on reward sensitivity could be due to the selection of ROIs. Although the VS was the most activated region in the reward anticipation contrast, other brain regions could be more relevant in regulating emotion-related reward processing (Liu et al., 2011). Therefore performing genetic analyses on brain ROIs might limit the chances of detecting risk variants (See Chapter 5).

Koob and colleagues' proposal may provide an alternative explanation for the lack of associations among life stress, stress genes and reward sensitivity in the IMAGEN adolescents. The interactions between the stress and reward systems may become more prominent throughout the development of addictions and impose a larger impact among the heavy substance users. Demonstrated by Feltenstein and colleagues, the exposure to stress was found to enhance nicotine-seeking behaviour in mice suffered from nicotine withdrawal (Feltenstein et al., 2012). The impact of stress system on reward processing might be less observable among adolescents who had begun to consume alcohol and tobacco.

7.5 Circadian system did not influence substance use and reward sensitivity

As discussed in the General Introduction, alterations in circadian rhythm were linked to the substance use and addictions in adolescents and clinical populations. Therefore one main part of

this thesis aimed to identify the relationship between circadian genes and substance use in adolescents.

I first investigated whether the DNA methylation of *PER*1 could act as a potential mechanism for stress-induced substance use in Chapter 4. Subject to the limitations of the Sequenom platform, only 44 CpG units on the *PER1* promoter were examined. The results presented in Chapter 4 showed that negative life events did not alter the DNA methylation of *PER1* and that the inter-individual differences in *PER1* DNA methylation did not predict the VS activations during reward anticipation, alcohol and tobacco use in adolescents. As a result, it can be possible that DNA methylation outside the assayed region may predict the stress-induced substance use and reward sensitivity in adolescents. Recent epigenome-wide studies indicated that the differences in DNA methylation could be due to substance use and alterations in brain activations (Zhang et al., 2013, Breitling et al., 2011, Zeilinger et al., 2013, Guo et al., 2011). Therefore, it is also possible that substance use and differences in the VS activations might influence the *PER1* DNA methylation level in the adolescents. The analyses reported in this chapter are largely inconclusive, yet it has informed researchers the impact of DNA methylation as a potential mechanism on stress-induced substance use.

Evidence from various genetic association studies suggested that alcohol and tobacco addictions are polygenic disorders triggered by gene-environment interactions (Schumann et al., 2011, Tobacco and Genetics, 2010, Dong et al., 2011, Thorgeirsson et al., 2010). As indicated by the association analyses presented in Chapter 5, genetic variations of circadian genes did not significantly predict substance use in adolescents at both age 14 and age 16. The analyses failed to replicate results from previous studies owing to the differences in MAFs and substance use phenotypes across populations. The inconsistent findings of polygenic score analyses has highlighted issues including statistical methodology that may mask the genetic effects of circadian system on substance use.

The analyses from Chapter 5 also suggested the lack of genetic effects of circadian genes on reward sensitivity in adolescents. Again, the lack of significant associations could be due to the statistical methodology as well as the sensitivity of substance use phenotypes. Such findings can also be supported by a previous study, demonstrating the lack of associations between circadian genes and reward processing (Zghoul et al., 2007). The *Per1*^{Brdm1} mutant mice did not show significant differences in ethanol seeking, reinforcement and relapse compared to their healthy counterparts (Zghoul et al., 2007). It is also possible that the impact of circadian system on reward processing is specific to different stages of addictions. For example, blunted circadian rhythm and reduced locomotor activity were observed in mice suffered from long-term opiate withdrawal (Li et al., 2010). Moreover, disruptions in the mRNA expression of *Per1, Per2* and *Per3* in the VTA, NAc and the amygdala had been observed, implying the impaired functioning of mesocorticolimic dopaminergic system at the withdrawal phase in these mice (Li et al., 2010). The interplay between the circadian and reward systems can be expected towards the later stage of addiction cycle.

To summarise, the analyses presented in this thesis suggested that the circadian system had no significant impact on reward sensitivity and substance use in adolescents. The lack of circadian rhythm information in the IMAGEN adolescents did not allow us to identify the relationships of circadian rhythm with negative life events, reward sensitivity and substance use. It is therefore unclear the extent to which alterations in circadian rhythm can influence substance use in the adolescents.

7.6 Thesis limitations

The IMAGEN study has provided a large sample that is essential for genetic and neuroimaging analyses. Nonetheless, this thesis is subject to various limitations and methodological issues. These issues should be taken into consideration when interpreting the results. Whilst specific

limitations pertaining to different aspects of this thesis are discussed in each chapter, the main methodological issues are listed below.

Community sample

Adolescents of the IMAGEN study were recruited from local communities. Symptoms of alcohol and tobacco misuse might not be observed in these adolescents. Additional analyses will be required to validate the significance of findings in clinical populations.

Behavioural phenotypes

Negative life events

The level of life stress experienced by adolescents is restricted to the 20 negative life events items measured in the Life Events Questionnaire (Newcomb et al., 1981). In particular, the frequency of each occurred life event was reported on a dichotomous scale. Therefore events that might occur more than once (e.g. family moved, face broke out with pimples) was not taken into consideration. Thus the negative life events did not fully capture the consequences of life stress on substance use and reward processing in adolescents.

Substance use phenotypes in the ESPAD

Similar to many self-report questionnaires, the reliability of substance use occasions measured in the ESPAD is dependent on individuals' ability to recall the information. Therefore, the classifications of substance use and abstinence groups can be subject to self-report bias. Additional measurements from other questionnaires can be used to cross-validate the reliability of substance use phenotypes.

Brain activations in fMRI tasks

The BOLD response measured in the fMRI tasks has provided an indirect measurement of brain activity. Despite all efforts have been made to standardise the scanning procedure, changes in the BOLD responses could be subject to head movements, task habituation, as well as the anxiety and arousal states of the adolescents. Performing ROI analyses ignores the connectivity of brain regions which can have an impact on the ROI BOLD responses.

MID and face fMRI contrasts

The BOLD responses of the VS from the 'Large win- No win' reward anticipation contrast of the MID task were analysed in this thesis. Other aspects of reward processing that may indicate reward sensitivity, for example reward feedback and the omission of an expected reward, were modelled in other contrasts of the MID task. Similarly, the amygdala BOLD responses of the 'angry face – control' contrast in the face task did not fully represent individual's neural response when processing emotional stimuli. Other aspects of emotional processing, including the processing of positive emotions, can be characterised by including happy faces in the face task.

Genome-wide genetic data

Since SNPs with MAF below 1% were excluded during the quality control procedure, any rare variants that might contribute to the risk phenotypes could not be examined. Due to the tagging nature of the genotyped SNPs, any target SNPs should undergo further analyses to confirm the nature of the association. This will involve exploring the LD structure and performing SNP imputation of the target region. If necessary, additional genotyping, sequencing and functional studies will be required to determine the functions of the target SNPs.

Imputed genetic data

Imputation of SNPs was performed based on the genetic information of individuals available from the 1000 Genomes project. The allele dosage of the imputed SNPs only reflects the estimated copies of major allele expressed in the IMAGEN adolescents. Therefore, genotyping the imputed SNPs will be required to confirm the genetic variations and the MAFs within the IMAGEN population.

Gene expression microarray data

The design of probes has limited the accuracy for the detection of mRNA expressions. As illustrated by the analyses presented in Chapter 6, the *FYN* mRNA expression data of two probes could not be used due to the presence of SNPs and non-specific binding. Although the probes were designed to capture the expressions of all *FYN* mRNA isoforms, the data would not allow researchers to distinguish the mRNA expression of individual isoform. Technique such as real-time PCR should be used to validate the mRNA expression of target genes and isoforms.

7.7 Implications for future research and conclusions

This thesis has investigated the risks of alcohol and tobacco misuse in adolescents at behavioural, genetic and epigenetic levels. Given the availability of the cross-sectional behavioural and neuroimaging data in the IMAGEN Study, future research should focus on investigating the development trajectories of substance addictions. Based on the work established in this thesis, I propose the following projects:

Project 1: Epigenetic regulation of life stress and risks of substance misuse in 14 year old adolescents

The genome-wide DNA methylation data of ~700 IMAGEN adolescents was made available in summer 2013. Based on the data from Chapter 4, analysis can be carried out to explore the

DNA methylation of additional CpG loci on *PER1* in associate with alcohol and tobacco misuse in adolescents. Alterations in the genome-wide DNA methylation pattern in response to negative life events can also be studied. Upon the identification of candidate CpG loci, their functional significance on substance misuse can be investigated by performing bioinformatic analyses and associations with the mRNA expressions. This project will allow researchers to identify the consequences of life stress on gene functions and substance use behaviour in the adolescents.

Project 2: Identify the impact of life stress on reward sensitivity in heavy substance users

The impact of stress on reward processing can be less prominent in adolescents who had just started consuming alcohol and tobacco. In addition to the ventral striatum, the dorsal part of striatum as well as other brain regions may be more relevant in integrating stress signals during the anticipation of reward. Hence future work should be focused on identifying brain ROIs associated with reward sensitivity and life stress. Analyses can be performed to identify genetic risk variants that are associated with brain activations during reward anticipation among heavy substance users. As the development of substance addictions can be seen as a process of continuous positive and negative reinforcements, this project will allow researchers to capture the neural basis of reward sensitivity in response to stress and how the interaction between the stress and reward systems can contribute to the escalation of substance use.

Project 3: Establish the gene-environment relationship of *FYN* rs1409837, life stress and drunkenness

Based on the results from Chapter 6, future work should aim to characterise mechanisms underlying the protective effect of *FYN* rs1409837 on drunkenness. Additional analyses should be carried out to establish the interaction effect of *FYN* rs1409837, amygdala activation/ negative life events and drunkenness in adolescents at age 16. To improve the power of associations, alternative approaches such as using haplotype as a unit for analysis should be considered. As the follow-up assessment in adolescents at age 18 is underway, the genetic effect of *FYN* rs1409837 on alcohol misuse can also be tested in the participants at later life. In

addition, functional analysis can be conducted at cellular level to confirm the role of *FYN* rs1409837 in mediating stress reactivity.

Replication of findings

Finally, all results should subject to rigorous replications in order to minimise chances of detecting false positives. Performing internal replication within the IMAGEN population maximises the homogeneity across samples and data, thus increases the confidence of findings. Whereas performing external replication in populations that resembles the IMAGEN population may allow generalisation of findings to a wider population.

Conclusions

In the past decades, extensive research has been dedicated to investigate the genetic basis of substance misuse and addictions. However, it has remained a challenge to understand the functions of genetic variants related to the development of addictions. In this thesis, I examined the stress and circadian systems in mediating reward sensitivity and risks of alcohol and tobacco misuse in over 2000 adolescents. The data of my thesis has supported the role of life stress in substance use and drunkenness during adolescence at the behavioural, neuroimaging and genetic levels. The consequences of reward sensitivity on substance use have also been confirmed. By adopting the gene-neuroimaging approach, I investigated the potential mechanisms of *FYN* rs409837 on drunkenness in attempt to resolve the missing heritability problem of alcohol addictions. Results obtained in this thesis provide new insights and approaches that should support the understanding of the genetic underpinnings of substance misuse in adolescence. A comprehensive understanding to the genetics and neurobiology of substance addictions will have profound consequences, from designing therapeutic targets to policy making and implementation of substance use intervention programs in the future.

References

ABARCA, C., ALBRECHT, U. & SPANAGEL, R. 2002. Cocaine sensitization and reward are under the influence of circadian genes and rhythm. *Proc Natl Acad Sci U S A,* 99**,** 9026-30.

AHLSTR M, S. K. & ÖSTERBERG, E. L. 2005. International perspectives on adolescent and young adult drinking. *Alcohol Research & Health,* 28**,** 259-268.

AKHTAR, R. A., REDDY, A. B., MAYWOOD, E. S., CLAYTON, J. D., KING, V. M., SMITH, A. G., GANT, T. W., HASTINGS, M. H. & KYRIACOU, C. P. 2002. Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. *Curr Biol,* 12**,** 540-50.

ALEXANDER, L. D. & SANDER, L. D. 1994. Vasoactive intestinal peptide stimulates ACTH and corticosterone release after injection into the PVN. *Regul Pept,* 51**,** 221-7.

AMIR, S., LAMONT, E. W., ROBINSON, B. & STEWART, J. 2004. A circadian rhythm in the expression of PERIOD2 protein reveals a novel SCN-controlled oscillator in the oval nucleus of the bed nucleus of the stria terminalis. *J Neurosci,* 24**,** 781-90.

ANDA, R. F., CROFT, J. B., FELITTI, V. J., NORDENBERG, D., GILES, W. H., WILLIAMSON, D. F. & GIOVINO, G. A. 1999. Adverse childhood experiences and smoking during adolescence and adulthood. *JAMA,* 282**,** 1652-8.

ANTEQUERA, F. 2003. Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci,* 60**,** 1647-58.

ANTONI, F. A. 1986. Hypothalamic control of adrenocorticotropin secretion: advances since the discovery of 41-residue corticotropin-releasing factor. *Endocr Rev,* 7**,** 351-78.

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet,* 25**,** 25-9.

BAE, K., JIN, X., MAYWOOD, E. S., HASTINGS, M. H., REPPERT, S. M. & WEAVER, D. R. 2001. Differential functions of mPer1, mPer2, and mPer3 in the SCN circadian clock. *Neuron,* 30**,** 525-36.

BALDING, D. J. 2006. A tutorial on statistical methods for population association studies. *Nat Rev Genet,* 7**,** 781-91.

BALSALOBRE, A., BROWN, S. A., MARCACCI, L., TRONCHE, F., KELLENDONK, C., REICHARDT, H. M., SCHUTZ, G. & SCHIBLER, U. 2000. Resetting of circadian time in peripheral tissues by glucocorticoid signaling. *Science,* 289**,** 2344-7.

BARR, C. S., NEWMAN, T. K., BECKER, M. L., CHAMPOUX, M., LESCH, K. P., SUOMI, S. J., GOLDMAN, D. & HIGLEY, J. D. 2003. Serotonin transporter gene variation is associated with alcohol sensitivity in rhesus macaques exposed to early-life stress. *Alcohol Clin Exp Res,* 27**,** 812-7.

BEAUCHAMP, M. S., LEE, K. E., HAXBY, J. V. & MARTIN, A. 2003. FMRI responses to video and point-light displays of moving humans and manipulable objects. *J Cogn Neurosci,* 15**,** 991-1001.

BECK, A., SCHLAGENHAUF, F., WUSTENBERG, T., HEIN, J., KIENAST, T., KAHNT, T., SCHMACK, K., HAGELE, C., KNUTSON, B., HEINZ, A. & WRASE, J. 2009. Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. *Biol Psychiatry,* 66**,** 734-42.

BELGARDT, B. F., HUSCH, A., ROTHER, E., ERNST, M. B., WUNDERLICH, F. T., HAMPEL, B., KLOCKENER, T., ALESSI, D., KLOPPENBURG, P. & BRUNING, J. C. 2008. PDK1 deficiency in POMC-expressing cells reveals FOXO1-dependent and -independent pathways in control of energy homeostasis and stress response. *Cell Metab,* 7**,** 291-301.

BENCA, R., DUNCAN, M. J., FRANK, E., MCCLUNG, C., NELSON, R. J. & VICENTIC, A. 2009. Biological rhythms, higher brain function, and behavior: Gaps, opportunities, and challenges. *Brain Res Rev,* 62**,** 57-70.

BERNARD, S., GONZE, D., CAJAVEC, B., HERZEL, H. & KRAMER, A. 2007. Synchronization-induced rhythmicity of circadian oscillators in the suprachiasmatic nucleus. *PLoS Comput Biol,* 3**,** e68.

BIERNACKA, J. M., GESKE, J., JENKINS, G. D., COLBY, C., RIDER, D. N., KARPYAK, V. M., CHOI, D. S. & FRIDLEY, B. L. 2013. Genome-wide gene-set analysis for identification of pathways associated with alcohol dependence. *Int J Neuropsychopharmacol,* 16**,** 271-8.

BIERUT, L. J., AGRAWAL, A., BUCHOLZ, K. K., DOHENY, K. F., LAURIE, C., PUGH, E., FISHER, S., FOX, L., HOWELLS, W., BERTELSEN, S., HINRICHS, A. L., ALMASY, L., BRESLAU, N., CULVERHOUSE, R. C., DICK, D. M., EDENBERG, H. J., FOROUD, T., GRUCZA, R. A., HATSUKAMI, D., HESSELBROCK, V., JOHNSON, E. O., KRAMER, J., KRUEGER, R. F., KUPERMAN, S., LYNSKEY, M., MANN, K., NEUMAN, R. J., NOTHEN, M. M., NURNBERGER, J. I., JR., PORJESZ, B., RIDINGER, M., SACCONE, N. L., SACCONE, S. F., SCHUCKIT, M. A., TISCHFIELD, J. A., WANG, J. C., RIETSCHEL, M., GOATE, A. M., RICE, J. P. & GENE, E. A. S. C. 2010. A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci U S A,* 107**,** 5082-7.

BIRD, A. 1980. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res,* 8**,** 1499-1504.

BJORK, J. M., KNUTSON, B. & HOMMER, D. W. 2008a. Incentive-elicited striatal activation in adolescent children of alcoholics. *Addiction,* 103**,** 1308-19.

BJORK, J. M., SMITH, A. R. & HOMMER, D. W. 2008b. Striatal sensitivity to reward deliveries and omissions in substance dependent patients. *Neuroimage,* 42**,** 1609-21.

BLAKEMORE, S. J. & MILLS, K. L. 2013. Is Adolescence a Sensitive Period for Sociocultural Processing? *Annu Rev Psychol*.

BLOMEYER, D., TREUTLEIN, J., ESSER, G., SCHMIDT, M. H., SCHUMANN, G. & LAUCHT, M. 2008. Interaction between CRHR1 gene and stressful life events predicts adolescent heavy alcohol use. *Biol Psychiatry,* 63**,** 146-51.

BLUM, K., SHERIDAN, P. J., WOOD, R. C., BRAVERMAN, E. R., CHEN, T. J., CULL, J. G. & COMINGS, D. E. 1996. The D2 dopamine receptor gene as a determinant of reward deficiency syndrome. *J R Soc Med,* 89**,** 396-400.

BONSCH, D., LENZ, B., FISZER, R., FRIELING, H., KORNHUBER, J. & BLEICH, S. 2006. Lowered DNA methyltransferase (DNMT-3b) mRNA expression is associated with genomic DNA hypermethylation in patients with chronic alcoholism. *J Neural Transm,* 113**,** 1299-304.

BOOKER, C. L., UNGER, J. B., AZEN, S. P., BAEZCONDE-GARBANATI, L., LICKEL, B. & JOHNSON, C. A. 2008. A longitudinal analysis of stressful life events, smoking behaviors, and gender differences in a multicultural sample of adolescents. *Subst Use Misuse,* 43**,** 1521-43.

BREESE, G. R., SINHA, R. & HEILIG, M. 2011. Chronic alcohol neuroadaptation and stress contribute to susceptibility for alcohol craving and relapse. *Pharmacol Ther,* 129**,** 149-71.

BREITLING, L. P., YANG, R., KORN, B., BURWINKEL, B. & BRENNER, H. 2011. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet,* 88**,** 450-7.

BROMS, U., MADDEN, P. A., HEATH, A. C., PERGADIA, M. L., SHIFFMAN, S. & KAPRIO, J. 2007. The Nicotine Dependence Syndrome Scale in Finnish smokers. *Drug Alcohol Depend,* 89**,** 42-51.

BUIJS, R. M., MARKMAN, M., NUNES-CARDOSO, B., HOU, Y. X. & SHINN, S. 1993. Projections of the suprachiasmatic nucleus to stress-related areas in the rat hypothalamus: a light and electron microscopic study. *J Comp Neurol,* 335**,** 42-54.

CAMACHO, F., CILIO, M., GUO, Y., VIRSHUP, D. M., PATEL, K., KHORKOVA, O., STYREN, S., MORSE, B., YAO, Z. & KEESLER, G. A. 2001. Human casein kinase

Idelta phosphorylation of human circadian clock proteins period 1 and 2. *FEBS Lett,* 489**,** 159-65.

CARTHARIUS, K., FRECH, K., GROTE, K., KLOCKE, B., HALTMEIER, M., KLINGENHOFF, A., FRISCH, M., BAYERLEIN, M. & WERNER, T. 2005. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics,* 21**,** 2933-42.

CHAMBERS, R. A., TAYLOR, J. R. & POTENZA, M. N. 2003. Developmental Neurocircuitry of Motivation in Adolescence: A Critical Period of Addiction Vulnerability. *Am J Psychiatry,* 160**,** 1041-1052.

CHANDRA, S., SCHARF, D. & SHIFFMAN, S. 2011. Within-day temporal patterns of smoking, withdrawal symptoms, and craving. *Drug Alcohol Depend,* 117**,** 118-25.

CHAPLIN, T. M., HONG, K., BERGQUIST, K. & SINHA, R. 2008. Gender differences in response to emotional stress: an assessment across subjective, behavioral, and physiological domains and relations to alcohol craving. *Alcohol Clin Exp Res,* 32**,** 1242-50.

CHEUNG, V. G. & SPIELMAN, R. S. 2009. Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nat Rev Genet,* 10**,** 595-604.

COCOZZA, S., AKHTAR, M. M., MIELE, G. & MONTICELLI, A. 2011. CpG islands undermethylation in human genomic regions under selective pressure. *PLoS One,* 6**,** e23156.

COLANTUONI, C., LIPSKA, B. K. YE, T., HYDE, T. M.,TAO, R., LEEK, J. T., COLANTUONI, E. A., ELKAHLOUN, A. G., HERMAN, M. M., WEINBERGER, D. R. & KLEINMAN, J. E. 2011. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature,* 478, 519-23.

COOLEN, M. W., STATHAM, A. L., GARDINER-GARDEN, M. & CLARK, S. J. 2007. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. *Nucleic Acids Res,* 35**,** e119.

DAGER, A. D., ANDERSON, B. M., STEVENS, M. C., PULIDO, C., ROSEN, R., JIANTONIO-KELLY, R. E., SISANTE, J. F., RASKIN, S. A., TENNEN, H., AUSTAD, C. S., WOOD, R. M., FALLAHI, C. R. & PEARLSON, G. D. 2013. Influence of alcohol use and family history of alcoholism on neural response to alcohol cues in college drinkers. *Alcohol Clin Exp Res,* 37 Suppl 1**,** E161-71.

DALM, S., DE KLOET, E. R. & OITZL, M. S. 2012. Post-training reward partially restores chronic stress induced effects in mice. *PLoS One,* 7**,** e39033.

DAY, J. J., CHILDS, D., GUZMAN-KARLSSON, M. C., KIBE, M., MOULDEN, J., SONG, E., TAHIR, A. & SWEATT, J. D. 2013. DNA methylation regulates associative reward learning. *Nat Neurosci,* 16**,** 1445-52.

DE KLOET, E. R., JOELS, M. & HOLSBOER, F. 2005. Stress and the brain: from adaptation to disease. *Nat Rev Neurosci,* 6**,** 463-75.

DE KLOET, E. R., VREUGDENHIL, E., OITZL, M. S. & JOELS, M. 1998. Brain corticosteroid receptor balance in health and disease. *Endocr Rev,* 19**,** 269-301.

DELGADO, M. R., GILLIS, M. M. & PHELPS, E. A. 2008. Regulating the expectation of reward via cognitive strategies. *Nat Neurosci,* 11**,** 880-1.

DENIS, H., NDLOVU, M. N. & FUKS, F. 2011. Regulation of mammalian DNA methyltransferases: a route to new mechanisms. *EMBO Rep,* 12**,** 647-56.

DESRIVIERES, S., LOURDUSAMY, A., MULLER, C., DUCCI, F., WONG, C. P., KAAKINEN, M., POUTA, A., HARTIKAINEN, A. L., ISOHANNI, M., CHAROEN, P., PELTONEN, L., FREIMER, N., ELLIOTT, P., JARVELIN, M. R. & SCHUMANN, G. 2011. Glucocorticoid receptor (NR3C1) gene polymorphisms and onset of alcohol abuse in adolescents. *Addict Biol,* 16**,** 510-3.

DOI, M., HIRAYAMA, J. & SASSONE-CORSI, P. 2006. Circadian regulator CLOCK is a histone acetyltransferase. *Cell,* 125**,** 497-508.

DONG, L., BILBAO, A., LAUCHT, M., HENRIKSSON, R., YAKOVLEVA, T., RIDINGER, M., DESRIVIERES, S., CLARKE, T. K., LOURDUSAMY, A., SMOLKA, M. N., CICHON, S., BLOMEYER, D., TREUTLEIN, J., PERREAU-LENZ, S., WITT, S., LEONARDI-ESSMANN, F., WODARZ, N., ZILL, P., SOYKA, M., ALBRECHT, U.,

RIETSCHEL, M., LATHROP, M., BAKALKIN, G., SPANAGEL, R. & SCHUMANN, G. 2011. Effects of the circadian rhythm gene period 1 (per1) on psychosocial stress-induced alcohol drinking. *Am J Psychiatry,* 168**,** 1090-8.

DOREMUS, T. L., BRUNELL, S. C., RAJENDRAN, P. & SPEAR, L. P. 2005. Factors Influencing Elevated Ethanol Consumption in Adolescent Relative to Adult Rats. *Alcoholism: Clinical & Experimental Research,* 29**,** 1796-1808.

DOREMUS, T. L., BRUNELL, S. C., VARLINSKAYA, E. I. & SPEAR, L. P. 2003. Anxiogenic effects during withdrawal from acute ethanol in adolescent and adult rats. *Pharmacology Biochemistry and Behavior,* 75**,** 411-418.

DUBE, S. R., MILLER, J. W., BROWN, D. W., GILES, W. H., FELITTI, V. J., DONG, M. & ANDA, R. F. 2006. Adverse childhood experiences and the association with ever using alcohol and initiating alcohol use during adolescence. *J Adolesc Health,* 38**,** 444 e1-10.

DUDBRIDGE, F. 2013. Power and predictive accuracy of polygenic risk scores. *PLoS Genet,* 9**,** e1003348.

DUFFIELD, G. E., BEST, J. D., MEURERS, B. H., BITTNER, A., LOROS, J. J. & DUNLAP, J. C. 2002. Circadian programs of transcriptional activation, signaling, and protein turnover revealed by microarray analysis of mammalian cells. *Curr Biol,* 12**,** 551-7.

DUNLAP, J. C. 1999. Molecular bases for circadian clocks. *Cell,* 96**,** 271-90.

ECKHARDT, F., LEWIN, J., CORTESE, R., RAKYAN, V. K., ATTWOOD, J., BURGER, M., BURTON, J., COX, T. V., DAVIES, R., DOWN, T. A., HAEFLIGER, C., HORTON, R., HOWE, K., JACKSON, D. K., KUNDE, J., KOENIG, C., LIDDLE, J., NIBLETT, D., OTTO, T., PETTETT, R., SEEMANN, S., THOMPSON, C., WEST, T., ROGERS, J., OLEK, A., BERLIN, K. & BECK, S. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet,* 38**,** 1378-85.

EDENBERG, H. J. 2007. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res Health,* 30**,** 5-13.

EDENBERG, H. J., DICK, D. M., XUEI, X., TIAN, H., ALMASY, L., BAUER, L. O., CROWE, R. R., GOATE, A., HESSELBROCK, V., JONES, K., KWON, J., LI, T. K., NURNBERGER, J. I., JR., O'CONNOR, S. J., REICH, T., RICE, J., SCHUCKIT, M. A., PORJESZ, B., FOROUD, T. & BEGLEITER, H. 2004. Variations in GABRA2, encoding the alpha 2 subunit of the GABA(A) receptor, are associated with alcohol dependence and with brain oscillations. *Am J Hum Genet,* 74**,** 705-14.

ENCODE PROJECT CONSORTIUM 2014. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature,* 447, 799-816.

ESSEX, M. J., SHIRTCLIFF, E. A., BURK, L. R., RUTTLE, P. L., KLEIN, M. H., SLATTERY, M. J., KALIN, N. H. & ARMSTRONG, J. M. 2011. Influence of early life stress on later hypothalamic-pituitary-adrenal axis functioning and its covariation with mental health symptoms: a study of the allostatic process from childhood into adolescence. *Dev Psychopathol,* 23**,** 1039-58.

EVANS, B. E., GREAVES-LORD, K., EUSER, A. S., FRANKEN, I. H. & HUIZINK, A. C. 2013. Cortisol levels in children of parents with a substance use disorder. *Psychoneuroendocrinology,* 38**,** 2109-20.

FARNELL, Y. Z., ALLEN, G. C., NAHM, S. S., NEUENDORFF, N., WEST, J. R., CHEN, W. J. & EARNEST, D. J. 2008. Neonatal alcohol exposure differentially alters clock gene oscillations within the suprachiasmatic nucleus, cerebellum, and liver of adult rats. *Alcohol Clin Exp Res,* 32**,** 544-52.

FELTENSTEIN, M. W., GHEE, S. M. & SEE, R. E. 2012. Nicotine self-administration and reinstatement of nicotine-seeking in male and female rats. *Drug Alcohol Depend,* 121**,** 240-6.

FERGUSON, S. A., KENNAWAY, D. J. & MOYER, R. W. 1999. Nicotine phase shifts the 6-sulphatoxymelatonin rhythm and induces c-Fos in the SCN of rats. *Brain Res Bull,* 48**,** 527-38.

FORBES, E. E. & DAHL, R. E. 2005. Neural systems of positive affect: relevance to understanding child and adolescent depression? *Dev Psychopathol,* 17**,** 827-50.

FORBES, E. E., DAHL, R. E., ALMEIDA, J. R., FERRELL, R. E., NIMGAONKAR, V. L., MANSOUR, H., SCIARRILLO, S. R., HOLM, S. M., RODRIGUEZ, E. E. &

PHILLIPS, M. L. 2012. PER2 rs2304672 polymorphism moderates circadian-relevant reward circuitry activity in adolescents. *Biol Psychiatry,* 71**,** 451-7.

FRANK, J., CICHON, S., TREUTLEIN, J., RIDINGER, M., MATTHEISEN, M., HOFFMANN, P., HERMS, S., WODARZ, N., SOYKA, M., ZILL, P., MAIER, W., MOSSNER, R., GAEBEL, W., DAHMEN, N., SCHERBAUM, N., SCHMAL, C., STEFFENS, M., LUCAE, S., ISING, M., MULLER-MYHSOK, B., NOTHEN, M. M., MANN, K., KIEFER, F. & RIETSCHEL, M. 2012. Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. *Addict Biol,* 17**,** 171-80.

GALVAN, A. 2010. Adolescent development of the reward system. *Front Hum Neurosci,* 4**,** 6.

GALVAN, A., HARE, T. A., PARRA, C. E., PENN, J., VOSS, H., GLOVER, G. & CASEY, B. J. 2006. Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *J Neurosci,* 26**,** 6885-92.

GARCIA-PEREZ, D., LAORDEN, M. L., MILANES, M. V. & NUNEZ, C. 2012. Glucocorticoids regulation of FosB/DeltaFosB expression induced by chronic opiate exposure in the brain stress system. *PLoS One,* 7**,** e50264.

GARDINER-GARDEN, M. & FROMMER, M. 1987. CpG islands in vertebrate genomes. *J Mol Biol,* 196**,** 261-82.

GEKAKIS, N. 1998. Role of the CLOCK Protein in the Mammalian Circadian Mechanism. *Science,* 280**,** 1564-1569.

GENOMES PROJECT, C., ABECASIS, G. R., ALTSHULER, D., AUTON, A., BROOKS, L. D., DURBIN, R. M., GIBBS, R. A., HURLES, M. E. & MCVEAN, G. A. 2010. A map of human genome variation from population-scale sequencing. *Nature,* 467**,** 1061-73.

GERY, S., KOMATSU, N., KAWAMATA, N., MILLER, C. W., DESMOND, J., VIRK, R. K., MARCHEVSKY, A., MCKENNA, R., TAGUCHI, H. & KOEFFLER, H. P. 2007. Epigenetic silencing of the candidate tumor suppressor gene Per1 in non-small cell lung cancer. *Clin Cancer Res,* 13**,** 1399-404.

GIEDD, J., BLUMENTHAL, J., JEFFRIES, N. O., CASTELLANOS, F. X., EVANS, A. C. & RAPOPORT, J. L. 1999. Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience,* 2**,** 861-863.

GIEDD, J. N., SNELL, J. W., LANGE, N., RAJAPAKSE, J. C., CASEY, B. J., KOZUCH, P. L., VAITUZIS, A. C., VAUSS, Y. C., HAMBURGER, S. D., KAYSEN, D. & RAPOPORT, J. L. 1996a. Quantitative magnetic resonance imaging of human brain development: ages 4-18. *Cereb Cortex,* 6**,** 551-60.

GIEDD, J. N., VAITUZIS, A. C., HAMBURGER, S. D., LANGE, N., RAJAPAKSE, J. C., KAYSEN, D., VAUSS, Y. C. & RAPOPORT, J. L. 1996b. Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: ages 4-18 years. *J Comp Neurol,* 366**,** 223-30.

GIOVINO, G. A., HENNINGFIELD, J. E., TOMAR, S. L., ESCOBEDO, L. G. & SLADE, S. 1995. Epidemiology of tobacco use and dependence. *Epidemiologic Reviews,* 17**,** 48-65.

GOEBEL-GOODY, S. M., BAUM, M., PASPALAS, C. D., FERNANDEZ, S. M., CARTY, N. C., KURUP, P. & LOMBROSO, P. J. 2012. Therapeutic implications for striatal-enriched protein tyrosine phosphatase (STEP) in neuropsychiatric disorders. *Pharmacol Rev,* 64**,** 65-87.

GOGTAY, N., NUGENT, T. F., 3RD, HERMAN, D. H., ORDONEZ, A., GREENSTEIN, D., HAYASHI, K. M., CLASEN, L., TOGA, A. W., GIEDD, J. N., RAPOPORT, J. L. & THOMPSON, P. M. 2006. Dynamic mapping of normal human hippocampal development. *Hippocampus,* 16**,** 664-72.

GOLDMAN, D. 1993. Recent developments in alcoholism:genetic transmission. *Recent Dev Alcohol,* 11**,** 231-48.

GOLL, M. G. & BESTOR, T. H. 2005. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem,* 74**,** 481-514.

GORKA, S. M., FITZGERALD, D. A., KING, A. C. & PHAN, K. L. 2013. Alcohol attenuates amygdala-frontal connectivity during processing social signals in heavy social drinkers: a preliminary pharmaco-fMRI study. *Psychopharmacology (Berl),* 229**,** 141-54.

GRANT, B. F. & DAWSON, D. A. 1997. Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse,* 9**,** 103-10.

GROSBRAS, M. H. & PAUS, T. 2006. Brain networks involved in viewing angry hands or faces. *Cereb Cortex,* 16**,** 1087-96.

GUO, J. U., MA, D. K., MO, H., BALL, M. P., JANG, M. H., BONAGUIDI, M. A., BALAZER, J. A., EAVES, H. L., XIE, B., FORD, E., ZHANG, K., MING, G. L., GAO, Y. & SONG, H. 2011. Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nat Neurosci,* 14**,** 1345-51.

HALPERIN, E. & STEPHAN, D. A. 2009. SNP imputation in association studies. *Nat Biotechnol,* 27**,** 349-51.

HAMPP, G., RIPPERGER, J. A., HOUBEN, T., SCHMUTZ, I., BLEX, C., PERREAU-LENZ, S., BRUNK, I., SPANAGEL, R., AHNERT-HILGER, G., MEIJER, J. H. & ALBRECHT, U. 2008. Regulation of monoamine oxidase A by circadian-clock components implies clock influence on mood. *Curr Biol,* 18**,** 678-83.

HANSELL, N. K., AGRAWAL, A., WHITFIELD, J. B., MORLEY, K. I., ZHU, G., LIND, A., PERGADIA, M. L., MADDEN, P. A. F., TODD, R. D., HEATH, A. C. & MARTIN, N. G. 2008. Long-Term Stability and Heritability of Telephone Interview Measures of Alcohol Consumption and Dependence. *Twin Research and Human Genetics,* 11**,** 287-305.

HARDY, J. & SINGLETON, A. 2009. Genomewide Association Studies and Human Disease. *New England Journal of Medicine,* 360**,** 1759-68.

HARFSTRAND, A., FUXE, K., CINTRA, A., AGNATI, L. F., ZINI, I., WIKSTROM, A. C., OKRET, S., YU, Z. Y., GOLDSTEIN, M., STEINBUSCH, H. & ET AL. 1986. Glucocorticoid receptor immunoreactivity in monoaminergic neurons of rat brain. *Proc Natl Acad Sci U S A,* 83**,** 9779-83.

HARIRI, A. R., MATTAY, V. S., TESSITORE, A., KOLACHANA, B., FERA, F., GOLDMAN, D., EGAN, M. F. & WEINBERGER, D. R. 2002. Serotonin transporter genetic variation and the response of the human amygdala. *Science,* 297**,** 400-3.

HASLER, B. P., DAHL, R. E., HOLM, S. M., JAKUBCAK, J. L., RYAN, N. D., SILK, J. S., PHILLIPS, M. L. & FORBES, E. E. 2012. Weekend-weekday advances in sleep timing are associated with altered reward-related brain function in healthy adolescents. *Biol Psychol,* 91**,** 334-41.

HEIKKINEN, A. M., BROMS, U., PITKANIEMI, J., KOSKENVUO, M. & MEURMAN, J. 2009. Key factors in smoking cessation intervention among 15-16-year-olds. *Behav Med,* 35**,** 93-9.

HERB, B. R., WOLSCHIN, F., HANSEN, K. D., ARYEE, M. J., LANGMEAD, B., IRIZARRY, R., AMDAM, G. V. & FEINBERG, A. P. 2012. Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nat Neurosci,* 15**,** 1371-3.

HERMAN, J. P., FIGUEIREDO, H., MUELLER, N. K., ULRICH-LAI, Y., OSTRANDER, M. M., CHOI, D. C. & CULLINAN, W. E. 2003. Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo–pituitary–adrenocortical responsiveness. *Frontiers in Neuroendocrinology,* 24**,** 151-180.

HIBELL, B., ANDERSSON, B., BJARNASON, T., KOKKEVI, A., MORGAN, M. & NARUSK, A. 1997. The 1995 ESPAD report : alcohol and other drug use among students in 26 European countries. *Swedish Council for Information on Alcohol and Other Drugs*.

HIGLEY, J. D., HASERT, M. F., SUOMI, S. J. & LINNOILA, M. 1991. Nonhuman primate model of alcohol abuse: effects of early experience, personality, and stress on alcohol consumption. *Proc Natl Acad Sci U S A,* 88**,** 7261-5.

HOFSTETTER, J. R., GRAHAME, N. J. & MAYEDA, A. R. 2003. Circadian activity rhythms in high-alcohol-preferring and low-alcohol-preferring mice. *Alcohol,* 30**,** 81-85.

HOGENESCH, J. B., GU, Y. Z., JAIN, S. & BRADFIELD, C. A. 1998. The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proc Natl Acad Sci U S A,* 95**,** 5474-9.

HSU, M. C., HUANG, C. C., CHOO, K. B. & HUANG, C. J. 2007. Uncoupling of promoter methylation and expression of Period1 in cervical cancer cells. *Biochem Biophys Res Commun,* 360**,** 257-62.

HUANG, L., WANG, C. & ROSENBERG, N. A. 2009. The relationship between imputation error and statistical power in genetic association studies in diverse populations. *Am J Hum Genet,* 85**,** 692-8.

HUANG, M. C., HO, C. W., CHEN, C. H., LIU, S. C., CHEN, C. C. & LEU, S. J. 2010. Reduced expression of circadian clock genes in male alcoholic patients. *Alcohol Clin Exp Res,* 34**,** 1899-904.

INTERNATIONAL SCHIZOPHRENIA, C., PURCELL, S. M., WRAY, N. R., STONE, J. L., VISSCHER, P. M., O'DONOVAN, M. C., SULLIVAN, P. F. & SKLAR, P. 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature,* 460**,** 748-52.

IRIZARRY, R. A., LADD-ACOSTA, C., WEN, B., WU, Z., MONTANO, C., ONYANGO, P., CUI, H., GABO, K., RONGIONE, M., WEBSTER, M., JI, H., POTASH, J. B., SABUNCIYAN, S. & FEINBERG, A. P. 2009. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet,* 41**,** 178-86.

ISHIGURO, H., SAITO, T., SHIBUYA, H., TORU, M. & ARINAMI, T. 2000. Mutation and association analysis of the Fyn kinase gene with alcoholism and schizophrenia. *Am J Med Genet,* 96**,** 716-20.

IWAMOTO, K., BUNDO, M., UEDA, J., OLDHAM, M.C., UKAI, W., HASHIMOTO, E., SAITO, T., GESCHWIND, D.H. & KATO, T. 2011. Neurons show distinctive DNA methylation profile and higher interindividual variations compared with non-neurons. *Genome Res,* 21, 688-96.

JAENISCH, R. & BIRD, A. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet,* 33 Suppl**,** 245-54.

JIRTLE, R. L. & SKINNER, M. K. 2007. Environmental epigenomics and disease susceptibility. *Nat Rev Genet,* 8**,** 253-62.

JOELS, M. & DE KLOET, E. R. 1992. Control of neuronal excitability by corticosteroid hormones. *Trends Neurosci,* 15**,** 25-30.

JOHNSON, V. & PANDINA, R. J. 1993. A Longitudinal Examination of the Relationships Among Stress, Coping Strategies, and Problems Associated with Alcohol Use. *Alcohol Clin Exp Res,* 17**,** 696-702.

JOHNSTON, L. D., O'MALLEY, P. M., BACHMAN, J. G. & SCHULENBERG, J. E. 2006. Monitoring the Future national results on adolescent drug use: Overview of key findings. *(NIH Publication No. 06-5882). Bethesda, MD: National Institute on Drug Abuse***,** 67.

JOUBERT, B. R., HABERG, S. E., NILSEN, R. M., WANG, X., VOLLSET, S. E., MURPHY, S. K., HUANG, Z., HOYO, C., MIDTTUN, O., CUPUL-UICAB, L. A., UELAND, P. M., WU, M. C., NYSTAD, W., BELL, D. A., PEDDADA, S. D. & LONDON, S. J. 2012. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect,* 120**,** 1425-31.

KALSBEEK, A., BUIJS, R. M., VAN HEERIKHUIZE, J. J., ARTS, M. & VAN DER WOUDE, T. P. 1992. Vasopressin-containing neurons of the suprachiasmatic nuclei inhibit corticosterone release. *Brain Res,* 580**,** 62-7.

KALSBEEK, A., KREIER, F., FLIERS, E., SAUERWEIN, H. P., ROMIJN, J. A. & BUIJS, R. M. 2007. Minireview: Circadian control of metabolism by the suprachiasmatic nuclei. *Endocrinology,* 148**,** 5635-9.

KANDEL, D. B., YAMAGUCHI, K. & CHEN, K. 1992. Stages of Progression in Drug Involvement from Adolescence to Adulthood: Further Evidence for the Gateway Theory. *J. Stud. Alcohol,* 53**,** 447-457.

KENDLER, K. S., CHEN, X., DICK, D., MAES, H., GILLESPIE, N., NEALE, M. C. & RILEY, B. 2012. Recent advances in the genetic epidemiology and molecular genetics of substance use disorders. *Nat Neurosci,* 15**,** 181-9.

KENDLER, K. S., NEALE, M. C., SULLIVAN, P., COREY, L. A., GARDNER, C. O. & PRESCOTT, C. A. 1999. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychological Medicine,* 29**,** 299-308.

KESSLER, R. C., BERGLUND, P., DEMLER, O., JIN, R., MERIKANGAS, K. R. & WALTERS, E. E. 2005. Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry,* 62**,** 593-602.

KNUTSON, B. & PETERSON, R. 2005. Neurally reconstructing expected utility. *Games and Economic Behavior,* 52**,** 305-315.

KNUTSON, B., TAYLOR, J., KAUFMAN, M., PETERSON, R. & GLOVER, G. 2005. Distributed neural representation of expected value. *J Neurosci,* 25**,** 4806-12.

KOOB, G. 1992. drug of abuse: anatomy, pharmacology and function of reward pathways. *TiPs,* 13**,** 177 - 184.

KOOB, G. & KREEK, M. J. 2007. Stress, dysregulation of drug reward pathways, and the transition to drug dependence. *Am J Psychiatry,* 164**,** 1149-59.

KOOB, G. F. 2003. Alcoholism: allostasis and beyond. *Alcohol Clin Exp Res,* 27**,** 232-43.

KOOB, G. F. 2008. A role for brain stress systems in addiction. *Neuron,* 59**,** 11-34.

KOOB, G. F. 2013. Addiction is a Reward Deficit and Stress Surfeit Disorder. *Front Psychiatry,* 4**,** 72.

KOOB, G. F. & LE MOAL, M. 1997. Drug Abuse: Hedonic Homeostatic Dysregulation. *Science,* 278**,** 52-58.

KOOB, G. F. & LE MOAL, M. 2008. Addiction and the brain antireward system. *Annu Rev Psychol,* 59**,** 29-53.

KOS, M. Z., YAN, J., DICK, D. M., AGRAWAL, A., BUCHOLZ, K. K., RICE, J. P., JOHNSON, E. O., SCHUCKIT, M., KUPERMAN, S., KRAMER, J., GOATE, A. M., TISCHFIELD, J. A., FOROUD, T., NURNBERGER, J., JR., HESSELBROCK, V., PORJESZ, B., BIERUT, L. J., EDENBERG, H. J. & ALMASY, L. 2013. Common biological networks underlie genetic risk for alcoholism in African- and European-American populations. *Genes Brain Behav,* 12**,** 532-42.

KOTURBASH, I., BAKER, M., LOREE, J., KUTANZI, K., HUDSON, D., POGRIBNY, I., SEDELNIKOVA, O., BONNER, W. & KOVALCHUK, O. 2006. Epigenetic dysregulation underlies radiation-induced transgenerational genome instability in vivo. *Int J Radiat Oncol Biol Phys,* 66**,** 327-30.

KOVANEN, L., SAARIKOSKI, S. T., HAUKKA, J., PIRKOLA, S., AROMAA, A., LONNQVIST, J. & PARTONEN, T. 2010. Circadian clock gene polymorphisms in alcohol use disorders and alcohol consumption. *Alcohol Alcohol,* 45**,** 303-11.

LAIRD, P. W. 2003. The power and the promise of DNA methylation markers. *Nat Rev Cancer,* 3**,** 253-66.

LAMONT, E. W., ROBINSON, B., STEWART, J. & AMIR, S. 2005. The central and basolateral nuclei of the amygdala exhibit opposite diurnal rhythms of expression of the clock protein Period2. *Proc Natl Acad Sci U S A,* 102**,** 4180-4.

LANDE-DINER, L. & CEDAR, H. 2005. Silence of the genes--mechanisms of long-term repression. *Nat Rev Genet,* 6**,** 648-54.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L.

W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature*, 409**,** 860-921.

LI, S. X., LIU, L. J., JIANG, W. G., SUN, L. L., ZHOU, S. J., LE FOLL, B., ZHANG, X. Y., KOSTEN, T. R. & LU, L. 2010. Circadian alteration in neurobiology during protracted opiate withdrawal in rats. *J Neurochem*, 115**,** 353-62.

LI, S. X., WANG, Z. R., LI, J., PENG, Z. G., ZHOU, W., ZHOU, M. & LU, L. 2008. Inhibition of Period1 gene attenuates the morphine-induced ERK-CREB activation in frontal cortex, hippocampus, and striatum in mice. *Am J Drug Alcohol Abuse*, 34**,** 673-82.

LIM, S. S., VOS, T., FLAXMAN, A. D., DANAEI, G., SHIBUYA, K. & ADAIR-ROHANI, H. A. M., ANDERSON HR, ANDREWS KG, ARYEE M, ATKINSON C, BACCHUS LJ, BAHALIM AN, BALAKRISHNAN K, BALMES J, BARKER-COLLO S, BAXTER A, BELL ML, BLORE JD, BLYTH F, BONNER C, BORGES G, BOURNE R, BOUSSINESQ M, BRAUER M, BROOKS P, BRUCE NG, BRUNEKREEF B, BRYAN-HANCOCK C, BUCELLO C, BUCHBINDER R, BULL F, BURNETT RT, BYERS TE, CALABRIA B, CARAPETIS J, CARNAHAN E, CHAFE Z, CHARLSON F, CHEN H, CHEN JS, CHENG AT, CHILD JC, COHEN A, COLSON KE, COWIE BC, DARBY S, DARLING S, DAVIS A, DEGENHARDT L, DENTENER F, DES JARLAIS DC, DEVRIES K, DHERANI M, DING EL, DORSEY ER, DRISCOLL T, EDMOND K, ALI SE, ENGELL RE, ERWIN PJ, FAHIMI S, FALDER G, FARZADFAR F, FERRARI A, FINUCANE MM, FLAXMAN S, FOWKES FG, FREEDMAN G, FREEMAN MK, GAKIDOU E, GHOSH S, GIOVANNUCCI E, GMEL G, GRAHAM K, GRAINGER R, GRANT B, GUNNELL D, GUTIERREZ HR, HALL W, HOEK HW, HOGAN A, HOSGOOD HD 3RD, HOY D, HU H, HUBBELL BJ, HUTCHINGS SJ, IBEANUSI SE, JACKLYN GL, JASRASARIA R, JONAS JB, KAN H, KANIS JA, KASSEBAUM N, KAWAKAMI N, KHANG YH, KHATIBZADEH S, KHOO JP, KOK C, LADEN F, LALLOO R, LAN Q, LATHLEAN T, LEASHER JL, LEIGH J, LI Y, LIN JK, LIPSHULTZ SE, LONDON S, LOZANO R, LU Y, MAK J, MALEKZADEH R, MALLINGER L, MARCENES W, MARCH L, MARKS R, MARTIN R, MCGALE P, MCGRATH J, MEHTA S, MENSAH GA, MERRIMAN TR, MICHA R, MICHAUD C, MISHRA V, MOHD HANAFIAH K, MOKDAD AA, MORAWSKA L, MOZAFFARIAN D, MURPHY T, NAGHAVI M, NEAL B, NELSON PK, NOLLA JM, NORMAN R, OLIVES C, OMER SB, ORCHARD J, OSBORNE R, OSTRO B, PAGE A, PANDEY KD, PARRY CD, PASSMORE E, PATRA J, PEARCE N, PELIZZARI PM, PETZOLD M, PHILLIPS MR, POPE D, POPE CA 3RD, POWLES J, RAO M, RAZAVI H, REHFUESS EA, REHM JT, RITZ B, RIVARA FP, ROBERTS T, ROBINSON C, RODRIGUEZ-PORTALES JA, ROMIEU I, ROOM R, ROSENFELD LC, ROY A, RUSHTON L, SALOMON JA, SAMPSON U, SANCHEZ-RIERA L, SANMAN E, SAPKOTA A, SEEDAT S, SHI P, SHIELD K, SHIVAKOTI R, SINGH GM, SLEET DA, SMITH E, SMITH KR, STAPELBERG NJ, STEENLAND K, ST CKL H, STOVNER LJ, STRAIF K, STRANEY L, THURSTON GD, TRAN JH, VAN DINGENEN R, VAN DONKELAAR A, VEERMAN JL, VIJAYAKUMAR L, WEINTRAUB R, WEISSMAN MM, WHITE RA, WHITEFORD H, WIERSMA ST, WILKINSON JD, WILLIAMS HC, WILLIAMS W, WILSON N, WOOLF AD, YIP P, ZIELINSKI JM, LOPEZ AD, MURRAY CJ, EZZATI M, ALMAZROA MA, MEMISH ZA. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380**,** 2224-60.

LINDHOLM, S., ROSIN, A., DAHLIN, I., GEORGIEVA, J. & FRANCK, J. 2007. Ethanol alters the effect of kappa receptor ligands on dopamine release in the nucleus accumbens. *Physiol Behav*, 92**,** 167-71.

LINTONEN, T., AHLSTROM, S. & METSO, L. 2004. The reliability of self-reported drinking in adolescence. *Alcohol Alcohol,* 39**,** 362-8.

LIU, X., HAIRSTON, J., SCHRIER, M. & FAN, J. 2011. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev,* 35**,** 1219-36.

LIUTKEVICIUTE, Z., LUKINAVICIUS, G., MASEVICIUS, V., DAUJOTYTE, D. & KLIMASAUSKAS, S. 2009. Cytosine-5-methyltransferases add aldehydes to DNA. *Nat Chem Biol,* 5**,** 400-2.

LOGOTHETIS, N. K., PAULS, J., AUGATH, M., TRINATH, T. & OELTERMANN, A. 2001. Neurophysiological investigation of the basis of the fMRI signal. *Nature,* 412**,** 150-7.

LOVINGER, D. M., WHITE, G. & WEIGHT, F. F. 1990. NMDA receptor-mediated synaptic excitation selectively inhibited by ethanol in hippocampal slice from adult rat. *J Neurosci,* 10**,** 1372-9.

LOWENBERG, M., VERHAAR, A. P., BILDERBEEK, J., MARLE, J., BUTTGEREIT, F., PEPPELENBOSCH, M. P., VAN DEVENTER, S. J. & HOMMES, D. W. 2006. Glucocorticoids cause rapid dissociation of a T-cell-receptor-associated protein complex containing LCK and FYN. *EMBO Rep,* 7**,** 1023-9.

MADEIRA, M. D., ANDRADE, J. P., LIEBERMAN, A. R., SOUSA, N., ALMEIDA, O. F. & PAULA-BARBOSA, M. M. 1997. Chronic alcohol consumption and withdrawal do not induce cell death in the suprachiasmatic nucleus, but lead to irreversible depression of peptide immunoreactivity and mRNA levels. *J Neurosci,* 17**,** 1302-19.

MANZARDO, A. M., HENKHAUS, R. S. & BUTLER, M. G.2012. Global DNA promoter methylation in frontal cortex of alcoholics and controls. *Gene,* 498, 5-12.

MARKUS, R. P., SANTOS, J. M., ZAGO, W. & RENO, L. A. 2003. Melatonin nocturnal surge modulates nicotinic receptors and nicotine-induced [3H]glutamate release in rat cerebellum slices. *J Pharmacol Exp Ther,* 305**,** 525-30.

MAUNAKEA, A. K., NAGARAJAN, R. P., BILENKY, M., BALLINGER, T. J., D'SOUZA, C., FOUSE, S. D., JOHNSON, B. E., HONG, C., NIELSEN, C., ZHAO, Y., TURECKI, G., DELANEY, A., VARHOL, R., THIESSEN, N., SHCHORS, K., HEINE, V. M., ROWITCH, D. H., XING, X., FIORE, C., SCHILLEBEECKX, M., JONES, S. J., HAUSSLER, D., MARRA, M. A., HIRST, M., WANG, T. & COSTELLO, J. F. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature,* 466**,** 253-7.

MEIER, M. H., CASPI, A., HOUTS, R., SLUTSKE, W. S., HARRINGTON, H., JACKSON, K. M., BELSKY, D. W., POULTON, R. & MOFFITT, T. E. 2013. Prospective developmental subtypes of alcohol dependence from age 18 to 32 years: implications for nosology, etiology, and intervention. *Dev Psychopathol,* 25**,** 785-800.

MICHAELS, C. C. & HOLTZMAN, S. G. 2008. Early postnatal stress alters place conditioning to both mu- and kappa-opioid agonists. *J Pharmacol Exp Ther,* 325**,** 313-8.

MIYAKAWA, T. 1997. Fyn-Kinase as a Determinant of Ethanol Sensitivity: Relation to NMDA-Receptor Function. *Science,* 278**,** 698-701.

MIYAMOTO, Y. & SANCAR, A. 1998. Vitamin B2-based blue-light photoreceptors in the retinohypothalamic tract as the photoactive pigments for setting the circadian clock in mammals. *Proc Natl Acad Sci U S A,* 95**,** 6097-102.

MOLINARO, S., SICILIANO, V., CURZIO, O., DENOTH, F. & MARIANI, F. 2012. Concordance and consistency of answers to the self-delivered ESPAD questionnaire on use of psychoactive substances. *Int J Methods Psychiatr Res,* 21**,** 158-68.

MOTZKUS, D., MARONDE, E., GRUNENBERG, U., LEE, C. C., FORSSMANN, W. & ALBRECHT, U. 2000. The human PER1 gene is transcriptionally regulated by multiple signaling pathways. *FEBS Lett,* 486**,** 315-9.

MUNOZ, K. E., HYDE, L. W. & HARIRI, A. R. 2009. Imaging genetics. *J Am Acad Child Adolesc Psychiatry,* 48**,** 356-61.

MUSCHLER, M. A., HILLEMACHER, T., KRAUS, C., KORNHUBER, J., BLEICH, S. & FRIELING, H. 2010. DNA methylation of the POMC gene promoter is associated with craving in alcohol dependence. *J Neural Transm,* 117**,** 513-9.

NEES, F., VOLLSTADT-KLEIN, S., FAUTH-BUHLER, M., STEINER, S., MANN, K., POUSTKA, L., BANASCHEWSKI, T., BUCHEL, C., CONROD, P. J., GARAVAN,

H., HEINZ, A., ITTERMANN, B., ARTIGES, E., PAUS, T., PAUSOVA, Z., RIETSCHEL, M., SMOLKA, M. N., STRUVE, M., LOTH, E., SCHUMANN, G., FLOR, H. & CONSORTIUM, I. 2012. A target sample of adolescents and reward processing: same neural and behavioral correlates engaged in common paradigms? *Exp Brain Res,* 223**,** 429-39.

NEES, F., WITT, S. H., LOURDUSAMY, A., VOLLSTADT-KLEIN, S., STEINER, S., POUSTKA, L., BANASCHEWSKI, T., BARKER, G. J., BUCHEL, C., CONROD, P. J., FRANK, J., GALLINAT, J., GARAVAN, H., HEINZ, A., ITTERMANN, B., LOTH, E., MANN, K., ARTIGES, E., PAUS, T., PAUSOVA, Z., SMOLKA, M. N., STRUVE, M., SCHUMANN, G., RIETSCHEL, M., FLOR, H. & CONSORTIUM, I. 2013. Genetic risk for nicotine dependence in the cholinergic system and activation of the brain reward system in healthy adolescents. *Neuropsychopharmacology,* 38**,** 2081-9.

NEWCOMB, M. D., HUBA, G. J. & BENTLER, P. M. 1981. A Multidimensional Assessment of Stressful Life Events among Adolescents: Derivation and Correlates. *Journal of Health and Social Behavior,* 22**,** 400-415.

NITURE, S. K., JAIN, A. K., SHELTON, P. M. & JAISWAL, A. K. 2011. Src subfamily kinases regulate nuclear export and degradation of transcription factor Nrf2 to switch off Nrf2-mediated antioxidant activation of cytoprotective gene expression. *J Biol Chem,* 286**,** 28821-32.

NIWA, M., JAARO-PELED, H., TANKOU, S., SESHADRI, S., HIKIDA, T., MATSUMOTO, Y., CASCELLA, N. G., KANO, S., OZAKI, N., NABESHIMA, T. & SAWA, A. 2013. Adolescent stress-induced epigenetic control of dopaminergic neurons via glucocorticoids. *Science,* 339**,** 335-9.

NUMATA, S., YE, T., HYDE, T. M., GUITART-NAVARRO, X., TAO, R., WININGER, M., COLANTUONI, C., WEINBERGER, D. R., KLEINMAN, J. E. & LIPSKA, B. K. 2012. DNA methylation signatures in development and aging of the human prefrontal cortex. *Am J Hum Genet,* 90**,** 260-72.

O'DOHERTY, J., DAYAN, P., SCHULTZ, J., DEICHMANN, R., FRISTON, K. & DOLAN, R. J. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science,* 304**,** 452-4.

O'HARA, B. F., EDGAR, D. M., CAO, V. H., WILER, S. W., HELLER, H. C., KILDUFF, T. S. & MILLER, J. D. 1998. Nicotine and nicotinic receptors in the circadian system. *Psychoneuroendocrinology,* 23**,** 161-73.

OJEDA, S. R., DUBAY, C., LOMNICZI, A., KAIDAR, G., MATAGNE, V., SANDAU, U. S. & DISSEN, G. A. 2010. Gene networks and the neuroendocrine regulation of puberty. *Mol Cell Endocrinol,* 324**,** 3-11.

OLDS, J. & MILNER, P. 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J Comp Physiol Psychol,* 47**,** 419-27.

OSTER, H., YASUI, A., VAN DER HORST, G. T. & ALBRECHT, U. 2002. Disruption of mCry2 restores circadian rhythmicity in mPer2 mutant mice. *Genes Dev,* 16**,** 2633-8.

PASTOR, I. J., LASO, F. J., INES, S., MARCOS, M. & GONZALEZ-SARMIENTO, R. 2009. Genetic association between -93A/G polymorphism in the Fyn kinase gene and alcohol dependence in Spanish men. *Eur Psychiatry,* 24**,** 191-4.

PETERS, J., BROMBERG, U., SCHNEIDER, S., BRASSEN, S., MENZ, M., BANASCHEWSKI, T., CONROD, P. J., FLOR, H., GALLINAT, J., GARAVAN, H., HEINZ, A., ITTERMAN, B., LATHROP, M., MARTINOT, J. L., PAUS, T., POLINE, J. B., ROBBINS, T. W., RIETSCHEL, M., SMOLKA, M., STROHLE, A., STRUVE, M., LOTH, E., SCHUMANN, G., BUCHEL, C. & CONSORTIUM, I. 2011. Lower ventral striatal activation during reward anticipation in adolescent smokers. *Am J Psychiatry,* 168**,** 540-9.

PETERSEN, A., CROCKETT, L., RICHARDS, M. & BOXER, A. 1988. A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence,* 17(2)**,** 117-133.

PHAN, K. L., WAGER, T., TAYLOR, S. F. & LIBERZON, I. 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage,* 16**,** 331-48.

PHILIBERT, R. A., BEACH, S. R., GUNTER, T. D., BRODY, G. H., MADAN, A. & GERRARD, M. 2010. The effect of smoking on MAOA promoter methylation in DNA prepared from lymphoblasts and whole blood. *Am J Med Genet B Neuropsychiatr Genet,* 153B**,** 619-28.

PIAZZA, P. V., ROUGE-PONT, F., DEROCHE, V., MACCARI, S., SIMON, H. & LE MOAL, M. 1996. Glucocorticoids have state-dependent stimulant effects on the mesencephalic dopaminergic transmission. *Proc Natl Acad Sci U S A,* 93**,** 8716-20.

PILOWSKY, D. J., KEYES, K. M. & HASIN, D. S. 2009. Adverse childhood events and lifetime alcohol dependence. *Am J Public Health,* 99**,** 258-63.

PLOMIN, R., HAWORTH, C. M. & DAVIS, O. S. 2009. Common disorders are quantitative traits. *Nat Rev Genet,* 10**,** 872-8.

PONOMAREV, I., WANG, S., ZHANG, L., HARRIS, R. A. & MAYFIELD, R. D. 2012. Gene coexpression networks in human brain identify epigenetic modifications in alcohol dependence. *J Neurosci,* 32, 1884-97.

POPLAWSKI, M. M., BOYADJIEVA, N. & SARKAR, D. K. 2005. Vasoactive Intestinal Peptide and Corticotropin-Releasing Hormone Increase ??-Endorphin Release and Proopiomelanocortin Messenger RNA Levels in Primary Cultures of Hypothalamic Cells: Effects of Acute and Chronic Ethanol Treatment. *Alcoholism: Clinical & Experimental Research,* 29**,** 648-655.

PREITNER, N., DAMIOLA, F., LOPEZ-MOLINA, L., ZAKANY, J., DUBOULE, D., ALBRECHT, U. & SCHIBLER, U. 2002. The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell,* 110**,** 251-60.

PRICE, A. L., PATTERSON, N. J., PLENGE, R. M., WEINBLATT, M. E., SHADICK, N. A. & REICH, D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet,* 38**,** 904-9.

PRITCHARD, J. K., STEPHENS, M., ROSENBERG, N. A. & DONNELLY, P. 2000. Association mapping in structured populations. *Am J Hum Genet,* 67**,** 170-81.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I., DALY, M. J. & SHAM, P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet,* 81**,** 559-75.

RAZIN, A. & RIGGS, A. D. 1980. DNA methylation and gene function. *Science,* 210**,** 604-10.

REICH, D. E., CARGILL, M., BOLK, S., IRELAND, J., SABETI, P. C., RICHTER, D. J., LAVERY, T., KOUYOUMJIAN, R., FARHADIAN, S. F., WARD, R. & LANDER, E. S. 2001. Linkage disequilibrium in the human genome. *Nature,* 411**,** 199-204.

REICH, D. E. & LANDER, E. S. 2001. On the allelic spectrum of human disease. *Trends Genet,* 17**,** 502-10.

REIK, W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature,* 447**,** 425-32.

REUL, J. M. & DE KLOET, E. R. 1985. Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology,* 117**,** 2505-11.

RICHARDSON, H. N., LEE, S. Y., O'DELL, L. E., KOOB, G. F. & RIVIER, C. L. 2008. Alcohol self-administration acutely stimulates the hypothalamic-pituitary-adrenal axis, but alcohol dependence leads to a dampened neuroendocrine state. *European Journal of Neuroscience,* 28**,** 1641-1653.

RIVERA-MEZA, M., QUINTANILLA, M. E., TAMPIER, L., MURA, C. V., SAPAG, A. & ISRAEL, Y. 2010. Mechanism of protection against alcoholism by an alcohol dehydrogenase polymorphism: development of an animal model. *FASEB J,* 24**,** 266-74.

ROBERTSON, K. D. & WOLFFE, A. P. 2000. DNA methylation in health and disease. *Nat Rev Genet,* 1**,** 11-9.

ROBINSON, B. E. & BERRIDGE, K. C. 1993. The neural basis of drug craving: an incentive sensitisation theory of addiction. *Brain Research Reviews,* 18**,** 247-291.

ROSE, J. E., BEHM, F. M., SALLEY, A. N., BATES, J. E., COLEMAN, R. E., HAWK, T. C. & TURKINGTON, T. G. 2007. Regional brain activity correlates of nicotine dependence. *Neuropsychopharmacology,* 32**,** 2441-52.

ROSENWASSER, A., LOGAN, R. & FECTEAU, M. 2005a. Chronic Ethanol Intake Alters Circadian Period-Responses to Brief Light Pulses in Rats. *Chronobiology International,* 22**,** 227-236.

ROSENWASSER, A. M., FECTEAU, M. E. & LOGAN, R. W. 2005b. Effects of ethanol intake and ethanol withdrawal on free-running circadian activity rhythms in rats. *Physiol Behav,* 84**,** 537-42.

ROSENWASSER, A. M., FECTEAU, M. E., LOGAN, R. W., REED, J. D., COTTER, S. J. & SEGGIO, J. A. 2005c. Circadian activity rhythms in selectively bred ethanol-preferring and nonpreferring rats. *Alcohol,* 36**,** 69-81.

SAAH, T. 2005. The evolutionary origins and significance of drug addiction. *Harm Reduct J,* 2**,** 8.

SADEH, A., DAHL, R. E., SHAHAR, G. & ROSENBLAT-STEIN, S. 2009. Sleep and the transition to adolescence: a longitudinal study. *Sleep,* 32**,** 1602-9.

SARKAR, D. K., KUHN, P., MARANO, J., CHEN, C. & BOYADJIEVA, N. 2007. Alcohol exposure during the developmental period induces beta-endorphin neuronal death and causes alteration in the opioid control of stress axis function. *Endocrinology,* 148**,** 2828-34.

SASIK, R., CALVO, E. & CORBEIL, J. 2002. Statistical analysis of high-density oligonucleotide arrays: a multiplicative noise model. *Bioinformatics,* 18**,** 1633-40.

SCHMITT, H. F., HUANG, L. Z., SON, J. H., PINZON-GUZMAN, C., SLATON, G. S. & WINZER-SERHAN, U. H. 2008. Acute nicotine activates c-fos and activity-regulated cytoskeletal associated protein mRNA expression in limbic brain areas involved in the central stress-response in rat pups during a period of hypo-responsiveness to stress. *Neuroscience,* 157**,** 349-59.

SCHNEIDER, S., PETERS, J., BROMBERG, U., BRASSEN, S., MIEDL, S. F., BANASCHEWSKI, T., BARKER, G. J., CONROD, P., FLOR, H., GARAVAN, H., HEINZ, A., ITTERMANN, B., LATHROP, M., LOTH, E., MANN, K., MARTINOT, J. L., NEES, F., PAUS, T., RIETSCHEL, M., ROBBINS, T. W., SMOLKA, M. N., SPANAGEL, R., STROHLE, A., STRUVE, M., SCHUMANN, G., BUCHEL, C. & CONSORTIUM, I. 2012. Risk taking and the adolescent reward system: a potential common link to substance abuse. *Am J Psychiatry,* 169**,** 39-46.

SCHUMANN, G., COIN, L. J., LOURDUSAMY, A., CHAROEN, P., BERGER, K. H., STACEY, D., DESRIVIERES, S., ALIEV, F. A., KHAN, A. A., AMIN, N., AULCHENKO, Y. S., BAKALKIN, G., BAKKER, S. J., BALKAU, B., BEULENS, J. W., BILBAO, A., DE BOER, R. A., BEURY, D., BOTS, M. L., BREETVELT, E. J., CAUCHI, S., CAVALCANTI-PROENCA, C., CHAMBERS, J. C., CLARKE, T. K., DAHMEN, N., DE GEUS, E. J., DICK, D., DUCCI, F., EASTON, A., EDENBERG, H. J., ESKO, T., FERNANDEZ-MEDARDE, A., FOROUD, T., FREIMER, N. B., GIRAULT, J. A., GROBBEE, D. E., GUARRERA, S., GUDBJARTSSON, D. F., HARTIKAINEN, A. L., HEATH, A. C., HESSELBROCK, V., HOFMAN, A., HOTTENGA, J. J., ISOHANNI, M. K., KAPRIO, J., KHAW, K. T., KUEHNEL, B., LAITINEN, J., LOBBENS, S., LUAN, J., MANGINO, M., MAROTEAUX, M., MATULLO, G., MCCARTHY, M. I., MUELLER, C., NAVIS, G., NUMANS, M. E., NUNEZ, A., NYHOLT, D. R., ONLAND-MORET, C. N., OOSTRA, B. A., O'REILLY, P. F., PALKOVITS, M., PENNINX, B. W., POLIDORO, S., POUTA, A., PROKOPENKO, I., RICCERI, F., SANTOS, E., SMIT, J. H., SORANZO, N., SONG, K., SOVIO, U., STUMVOLL, M., SURAKK, I., THORGEIRSSON, T. E., THORSTEINSDOTTIR, U., TROAKES, C., TYRFINGSSON, T., TONJES, A., UITERWAAL, C. S., UITTERLINDEN, A. G., VAN DER HARST, P., VAN DER SCHOUW, Y. T., STAEHLIN, O., VOGELZANGS, N., VOLLENWEIDER, P., WAEBER, G., WAREHAM, N. J., WATERWORTH, D. M., WHITFIELD, J. B., WICHMANN, E. H., WILLEMSEN, G., WITTEMAN, J. C., YUAN, X., ZHAI, G., ZHAO, J. H., ZHANG, W., MARTIN, N. G., METSPALU, A., et al. 2011. Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. *Proc Natl Acad Sci U S A,* 108**,** 7119-24.

SCHUMANN, G., LOTH, E., BANASCHEWSKI, T., BARBOT, A., BARKER, G., BUCHEL, C., CONROD, P. J., DALLEY, J. W., FLOR, H., GALLINAT, J., GARAVAN, H., HEINZ, A., ITTERMAN, B., LATHROP, M., MALLIK, C., MANN, K., MARTINOT, J. L., PAUS, T., POLINE, J. B., ROBBINS, T. W., RIETSCHEL, M., REED, L., SMOLKA, M., SPANAGEL, R., SPEISER, C., STEPHENS, D. N., STROHLE, A., STRUVE, M. & CONSORTIUM, I. 2010. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol Psychiatry,* 15**,** 1128-39.

SCHUMANN, G., RUJESCU, D., KISSLING, C., SOYKA, M., DAHMEN, N., PREUSS, U. W., WIEMAN, S., DEPNER, M., WELLEK, S., LASCORZ, J., BONDY, B., GIEGLING, I., ANGHELESCU, I., COWEN, M. S., POUSTKA, A., SPANAGEL, R., MANN, K., HENN, F. A. & SZEGEDI, A. 2003. Analysis of genetic variations of protein tyrosine kinase fyn and their association with alcohol dependence in two independent cohorts. *Biological Psychiatry,* 54**,** 1422-1426.

SEGALL, L. A., PERRIN, J. S., WALKER, C. D., STEWART, J. & AMIR, S. 2006. Glucocorticoid rhythms control the rhythm of expression of the clock protein, Period2, in oval nucleus of the bed nucleus of the stria terminalis and central nucleus of the amygdala in rats. *Neuroscience,* 140**,** 753-7.

SEGGIO, J. A., FIXARIS, M. C., REED, J. D., LOGAN, R. W. & ROSENWASSER, A. M. 2009. Chronic ethanol intake alters circadian phase shifting and free-running period in mice. *J Biol Rhythms,* 24**,** 304-12.

SELHUB, J. 1999. Homocysteine metabolism. *Annu Rev Nutr,* 19**,** 217-46.

SEO, D., LACADIE, C. M., TUIT, K., HONG, K. I., CONSTABLE, R. T. & SINHA, R. 2013. Disrupted ventromedial prefrontal function, alcohol craving, and subsequent relapse risk. *JAMA Psychiatry,* 70**,** 727-39.

SHEARMAN, L. P., ZYLKA, M. J., WEAVER, D. R., KOLAKOWSKI, L. F., JR. & REPPERT, S. M. 1997. Two period homologs: circadian expression and photic regulation in the suprachiasmatic nuclei. *Neuron,* 19**,** 1261-9.

SHENKER, N. S., POLIDORO, S., VAN VELDHOVEN, K., SACERDOTE, C., RICCERI, F., BIRRELL, M. A., BELVISI, M. G., BROWN, R., VINEIS, P. & FLANAGAN, J. M. 2013. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet,* 22**,** 843-51.

SHIPPENBERG, T. S., ZAPATA, A. & CHEFER, V. I. 2007. Dynorphin and the pathophysiology of drug addiction. *Pharmacol Ther,* 116**,** 306-21.

SIMPSON, E. R. & WATERMAN, M. R. 1988. Regulation of the synthesis of steroidogenic enzymes in adrenal cortical cells by ACTH. *Annu Rev Physiol,* 50**,** 427-40.

SINHA, R., FOX, H. C., HONG, K. A., BERGQUIST, K., BHAGWAGAR, Z. & SIEDLARZ, K. M. 2009. Enhanced negative emotion and alcohol craving, and altered physiological responses following stress and cue exposure in alcohol dependent individuals. *Neuropsychopharmacology,* 34**,** 1198-208.

SLEIPNESS, E. P., SORG, B. A. & JANSEN, H. T. 2007. Diurnal differences in dopamine transporter and tyrosine hydroxylase levels in rat brain: dependence on the suprachiasmatic nucleus. *Brain Res,* 1129**,** 34-42.

SO, A. Y., CHAIVORAPOL, C., BOLTON, E. C., LI, H. & YAMAMOTO, K. R. 2007. Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS Genet,* 3**,** e94.

SPANAGEL, R. 2009. Alcoholism: A systematic approach from molecular physiology to addictive behavior. *Physiol Rev,* 89**,** 649-705.

SPANAGEL, R., PENDYALA, G., ABARCA, C., ZGHOUL, T., SANCHIS-SEGURA, C., MAGNONE, M. C., LASCORZ, J., DEPNER, M., HOLZBERG, D., SOYKA, M., SCHREIBER, S., MATSUDA, F., LATHROP, M., SCHUMANN, G. & ALBRECHT, U. 2005. The clock gene Per2 influences the glutamatergic system and modulates alcohol consumption. *Nat Med,* 11**,** 35-42.

SPEAR, L. P. 2000. The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioural Reviews,* 24**,** 417-463.

STACEY, D., BILBAO, A., MAROTEAUX, M., JIA, T., EASTON, A. C., LONGUEVILLE, S., NYMBERG, C., BANASCHEWSKI, T., BARKER, G. J., BUCHEL, C.,

CARVALHO, F., CONROD, P. J., DESRIVIERES, S., FAUTH-BUHLER, M., FERNANDEZ-MEDARDE, A., FLOR, H., GALLINAT, J., GARAVAN, H., BOKDE, A. L., HEINZ, A., ITTERMANN, B., LATHROP, M., LAWRENCE, C., LOTH, E., LOURDUSAMY, A., MANN, K. F., MARTINOT, J. L., NEES, F., PALKOVITS, M., PAUS, T., PAUSOVA, Z., RIETSCHEL, M., RUGGERI, B., SANTOS, E., SMOLKA, M. N., STAEHLIN, O., JARVELIN, M. R., ELLIOTT, P., SOMMER, W. H., MAMELI, M., MULLER, C. P., SPANAGEL, R., GIRAULT, J. A., SCHUMANN, G. & CONSORTIUM, I. 2012. RASGRF2 regulates alcohol-induced reinforcement by influencing mesolimbic dopamine neuron activity and dopamine release. *Proc Natl Acad Sci U S A,* 109**,** 21128-33.

STEPTOE, A. & USSHER, M. 2006. Smoking, cortisol and nicotine. *Int J Psychophysiol,* 59**,** 228-35.

STEVENS, V. L., BIERUT, L. J., TALBOT, J. T., WANG, J. C., SUN, J., HINRICHS, A. L., THUN, M. J., GOATE, A. & CALLE, E. E. 2008. Nicotinic receptor gene variants influence susceptibility to heavy smoking. *Cancer Epidemiol Biomarkers Prev,* 17**,** 3517-25.

SULLIVAN, P. F., DALY, M. J. & O'DONOVAN, M. 2012. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet,* 13**,** 537-51.

SUN, J., JIA, P., FANOUS, A. H., WEBB, B. T., VAN DEN OORD, E. J., CHEN, X., BUKSZAR, J., KENDLER, K. S. & ZHAO, Z. 2009. A multi-dimensional evidence-based candidate gene prioritization approach for complex diseases-schizophrenia as a case. *Bioinformatics,* 25**,** 2595-6602.

SUN, Z. S., ALBRECHT, U., ZHUCHENKO, O., BAILEY, J., EICHELE, G. & LEE, C. C. 1997. RIGUI, a putative mammalian ortholog of the Drosophila period gene. *Cell,* 90**,** 1003-11.

SUTER, M., MA, J., HARRIS, A., PATTERSON, L., BROWN, K. A., SHOPE, C., SHOWALTER, L., ABRAMOVICI, A. & AAGAARD-TILLERY, K. M. 2011. Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. *Epigenetics,* 6**,** 1284-94.

SUZUKI, A., IIDA, S., KATO-URANISHI, M., TAJIMA, E., ZHAN, F., HANAMURA, I., HUANG, Y., OGURA, T., TAKAHASHI, S., UEDA, R., BARLOGIE, B., SHAUGHNESSY, J., JR. & ESUMI, H. 2005. ARK5 is transcriptionally regulated by the Large-MAF family and mediates IGF-1-induced cell invasion in multiple myeloma: ARK5 as a new molecular determinant of malignant multiple myeloma. *Oncogene,* 24**,** 6936-44.

SUZUKI, M. M. & BIRD, A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet,* 9**,** 465-76.

TAHILIANI, M., KOH, K. P., SHEN, Y., PASTOR, W. A., BANDUKWALA, H., BRUDNO, Y., AGARWAL, S., IYER, L. M., LIU, D. R., ARAVIND, L. & RAO, A. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**,** 930-5.

TAKAHASHI, S. 2001. Physical and Inflammatory Stressors Elevate Circadian Clock Gene mPer1 mRNA Levels in the Paraventricular Nucleus of the Mouse. *Endocrinology,* 142**,** 4910-4917.

TAKUMI, T., TAGUCHI, K., MIYAKE, S., SAKAKIDA, Y., TAKASHIMA, N., MATSUBARA, C., MAEBAYASHI, Y., OKUMURA, K., TAKEKIDA, S., YAMAMOTO, S., YAGITA, K., YAN, L., YOUNG, M. W. & OKAMURA, H. 1998. A light-independent oscillatory gene mPer3 in mouse SCN and OVLT. *EMBO J,* 17**,** 4753-9.

TAVOLACCI, M. P., LADNER, J., GRIGIONI, S., RICHARD, L., VILLET, H. & DECHELOTTE, P. 2013. Prevalence and association of perceived stress, substance use and behavioral addictions: a cross-sectional study among university students in France, 2009-2011. *BMC Public Health,* 13**,** 724.

TEI, H., OKAMURA, H., SHIGEYOSHI, Y., FUKUHARA, C., OZAWA, R., HIROSE, M. & SAKAKI, Y. 1997. Circadian oscillation of a mammalian homologue of the Drosophila period gene. *Nature,* 389**,** 512-6.

THORGEIRSSON, T. E., GUDBJARTSSON, D. F., SURAKKA, I., VINK, J. M., AMIN, N., GELLER, F., SULEM, P., RAFNAR, T., ESKO, T., WALTER, S., GIEGER, C., RAWAL, R., MANGINO, M., PROKOPENKO, I., MAGI, R., KESKITALO, K., GUDJONSDOTTIR, I. H., GRETARSDOTTIR, S., STEFANSSON, H., THOMPSON, J. R., AULCHENKO, Y. S., NELIS, M., ABEN, K. K., DEN HEIJER, M., DIRKSEN, A., ASHRAF, H., SORANZO, N., VALDES, A. M., STEVES, C., UITTERLINDEN, A. G., HOFMAN, A., TONJES, A., KOVACS, P., HOTTENGA, J. J., WILLEMSEN, G., VOGELZANGS, N., DORING, A., DAHMEN, N., NITZ, B., PERGADIA, M. L., SAEZ, B., DE DIEGO, V., LEZCANO, V., GARCIA-PRATS, M. D., RIPATTI, S., PEROLA, M., KETTUNEN, J., HARTIKAINEN, A. L., POUTA, A., LAITINEN, J., ISOHANNI, M., HUEI-YI, S., ALLEN, M., KRESTYANINOVA, M., HALL, A. S., JONES, G. T., VAN RIJ, A. M., MUELLER, T., DIEPLINGER, B., HALTMAYER, M., JONSSON, S., MATTHIASSON, S. E., OSKARSSON, H., TYRFINGSSON, T., KIEMENEY, L. A., MAYORDOMO, J. I., LINDHOLT, J. S., PEDERSEN, J. H., FRANKLIN, W. A., WOLF, H., MONTGOMERY, G. W., HEATH, A. C., MARTIN, N. G., MADDEN, P. A., GIEGLING, I., RUJESCU, D., JARVELIN, M. R., SALOMAA, V., STUMVOLL, M., SPECTOR, T. D., WICHMANN, H. E., METSPALU, A., SAMANI, N. J., PENNINX, B. W., OOSTRA, B. A., BOOMSMA, D. I., TIEMEIER, H., VAN DUIJN, C. M., KAPRIO, J., GULCHER, J. R., CONSORTIUM, E., MCCARTHY, M. I., PELTONEN, L., THORSTEINSDOTTIR, U. & STEFANSSON, K. 2010. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet*, 42**,** 448-53.

THRESHER, R. J. 1998. Role of Mouse Cryptochrome Blue-Light Photoreceptor in Circadian Photoresponses. *Science*, 282**,** 1490-1494.

TOBACCO & GENETICS, C. 2010. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet*, 42**,** 441-7.

TORRES-FARFAN, C., ABARZUA-CATALAN, L., VALENZUELA, F. J., MENDEZ, N., RICHTER, H. G., VALENZUELA, G. J. & SERON-FERRE, M. 2009. Cryptochrome 2 expression level is critical for adrenocorticotropin stimulation of cortisol production in the capuchin monkey adrenal. *Endocrinology*, 150**,** 2717-22.

TRACHSEL, L., HELLER, H. C. & MILLER, J. D. 1995. Nicotine phase-advances the circadian neuronal activity rhythm in rat suprachiasmatic nuclei explants. *Neuroscience*, 65**,** 797-803.

TREUTLEIN, J., CICHON, S., RIDINGER, M., WODARZ, N., SOYKA, M., ZILL, P., MAIER, W., MOESSNER, R., GAEBEL, W., DAHMEN, N., FEHR, C., SCHERBAUM, N., STEFFENS, M., LUDWIG, K. U., FRANK, J., WICHMANN, H. E., SCHREIBER, S., DRAGANO, N., SOMMER, W. H., LEONARDI-ESSMANN, F., LOURDUSAMY, A., GEBICKE-HAERTER, P., WIENKER, T. F., SULLIVAN, P. F., N THEN, M. M., KIEFER, F., SPANAGEL, R., MANN, K. & RIETSCHEL, M. 2009. Genome-wide association study of alcohol dependence. *Arch Gen Psychiatry*, 66**,** 773-784.

TREUTLEIN, J., KISSLING, C., FRANK, J., WIEMANN, S., DONG, L., DEPNER, M., SAAM, C., LASCORZ, J., SOYKA, M., PREUSS, U. W., RUJESCU, D., SKOWRONEK, M. H., RIETSCHEL, M., SPANAGEL, R., HEINZ, A., LAUCHT, M., MANN, K. & SCHUMANN, G. 2006. Genetic association of the human corticotropin releasing hormone receptor 1 (CRHR1) with binge drinking and alcohol intake patterns in two independent samples. *Mol Psychiatry*, 11**,** 594-602.

TSUCHIMINE, S., YASUI-FURUKORI, N., KANEDA, A. & KANEKO, S. 2013. The CLOCK C3111T polymorphism is associated with reward dependence in healthy Japanese subjects. *Neuropsychobiology*, 67**,** 1-5.

TZOURIO-MAZOYER, N., LANDEAU, B., PAPATHANASSIOU, D., CRIVELLO, F., ETARD, O., DELCROIX, N., MAZOYER, B. & JOLIOT, M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15**,** 273-89.

UEDA, H. R., CHEN, W., ADACHI, A., WAKAMATSU, H., HAYASHI, S., TAKASUGI, T., NAGANO, M., NAKAHAMA, K., SUZUKI, Y., SUGANO, S., IINO, M.,

SHIGEYOSHI, Y. & HASHIMOTO, S. 2002. A transcription factor response element for gene expression during circadian night. *Nature,* 418**,** 534-9.

UMEMORI, H., WANAKA, A., KATO, H., TAKEUCHI, M., TOHYAMA, M. & YAMAMOTO, T. 1992. Specific expressions of Fyn and Lyn, lymphocyte antigen receptor-associated tyrosine kinases, in the central nervous system. *Brain Res Mol Brain Res,* 16**,** 303-10.

VAN LEIJENHORST, L., ZANOLIE, K., VAN MEEL, C. S., WESTENBERG, P. M., ROMBOUTS, S. A. & CRONE, E. A. 2010. What motivates the adolescent? Brain regions mediating reward sensitivity across adolescence. *Cereb Cortex,* 20**,** 61-9.

VARLINSKAYA, E. I. & SPEAR, L. P. 2010. Sensitization to social anxiolytic effects of ethanol in adolescent and adult Sprague-Dawley rats after repeated ethanol exposure. *Alcohol,* 44**,** 99-110.

VASU, V. T., CROSS, C. E. & GOHIL, K. 2009. Nr1d1, an important circadian pathway regulatory gene, is suppressed by cigarette smoke in murine lungs. *Integr Cancer Ther,* 8**,** 321-8.

VENDRUSCOLO, L. F., BARBIER, E., SCHLOSBURG, J. E., MISRA, K. K., WHITFIELD, T. W., JR., LOGRIP, M. L., RIVIER, C., REPUNTE-CANONIGO, V., ZORRILLA, E. P., SANNA, P. P., HEILIG, M. & KOOB, G. F. 2012. Corticosteroid-dependent plasticity mediates compulsive alcohol drinking in rats. *J Neurosci,* 32**,** 7563-71.

VETTER, C. S., DOREMUS-FITZWATER, T. L. & SPEAR, L. P. 2007. Time course of elevated ethanol intake in adolescent relative to adult rats under continuous, voluntary-access conditions. *Alcohol Clin Exp Res,* 31**,** 1159-68.

WADDINGTON, C. H. 1942. Canalisation of development and the inheritance of acquired characters. *Nature,* 150**,** 563-565.

WALKER, B. M., VALDEZ, G. R., MCLAUGHLIN, J. P. & BAKALKIN, G. 2012. Targeting dynorphin/kappa opioid receptor systems to treat alcohol abuse and dependence. *Alcohol,* 46**,** 359-70.

WANG, W. Y., BARRATT, B. J., CLAYTON, D. G. & TODD, J. A. 2005. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet,* 6**,** 109-18.

WANG, X., MOZHUI, K., LI, Z., MULLIGAN, M. K., INGELS, J. F., ZHOU, X., HORI, R. T., CHEN, H., COOK, M. N., WILLIAMS, R. W. & LU, L. 2012. A promoter polymorphism in the Per3 gene is associated with alcohol and stress response. *Transl Psychiatry,* 2**,** e73.

WANG, Z., RAY, R., FAITH, M., TANG, K., WILEYTO, E. P., DETRE, J. A. & LERMAN, C. 2008. Nicotine abstinence-induced cerebral blood flow changes by genotype. *Neurosci Lett,* 438**,** 275-80.

WATERLAND, R. A. & JIRTLE, R. L. 2003. Transposable Elements: Targets for Early Nutritional Effects on Epigenetic Gene Regulation. *Molecular and Cellular Biology,* 23**,** 5293-5300.

WATT, F. & MOLLOY, P. L. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & Development,* 2**,** 1136-1143.

WEBER, M., HELLMANN, I., STADLER, M. B., RAMOS, L., PAABO, S., REBHAN, M. & SCHUBELER, D. 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet,* 39**,** 457-66.

WECHSLER, H., DAVENPORT, A., DOWDALL, G., MOEYKENS, B. & CASTILLO, S. 1994. Health and behavioral consequences of binge drinking in college. A national survey of students at 140 campuses. *JAMA,* 272**,** 1672-7.

WEE, S. & KOOB, G. F. 2010. The role of the dynorphin-kappa opioid system in the reinforcing effects of drugs of abuse. *Psychopharmacology (Berl),* 210**,** 121-35.

WHALEN, P. J., RAUCH, S. L., ETCOFF, N. L., MCINERNEY, S. C., LEE, M. B. & JENIKE, M. A. 1998. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J Neurosci,* 18**,** 411-8.

WHITEFORD, H. A., DEGENHARDT, L., REHM, J., BAXTER, A. J., FERRARI, A. J., ERSKINE, H. E., CHARLSON, F. J., NORMAN, R. E., FLAXMAN, A. D., JOHNS, N., BURSTEIN, R., MURRAY, C. J. L. & VOS, T. 2013. Global burden of disease

attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet*.

WINDLE, M., MUN, E. Y. & WINDLE, R. C. 2005. Adolescent-to-young adulthood heavy drinking trajectories and their prospective predictors. *J Stud Alcohol,* 66**,** 313-22.

WINDLE, M. & ZUCKER, R. A. 2010. Reducing underage and young adult drinking: how to address critical drinking problems during this developmental period. *Alcohol Res Health,* 33**,** 29-44.

WITTMANN, M., PAULUS, M. & ROENNEBERG, T. 2010. Decreased psychological well-being in late 'chronotypes' is mediated by smoking and alcohol consumption. *Subst Use Misuse,* 45**,** 15-30.

WOLF, L. K., WRIGHT, N. D., KILFORD, E. J., DOLAN, R. J. & BLAKEMORE, S. J. 2013. Developmental changes in effects of risk and valence on adolescent decision-making. *Cogn Dev,* 28**,** 290-299.

WONG, C. C. & SCHUMANN, G. 2012. Integration of the circadian and stress systems: influence of neuropeptides and implications for alcohol consumption. *J Neural Transm,* 119**,** 1111-20.

WOODWARD, J. J., RON, D., WINDER, D. & ROBERTO, M. 2006. From blue states to up states: a regional view of NMDA-ethanol interactions. *Alcohol Clin Exp Res,* 30**,** 359-67.

WRAY, N. R., YANG, J., GODDARD, M. E. & VISSCHER, P. M. 2010. The genetic interpretation of area under the ROC curve in genomic profiling. *PLoS Genet,* 6**,** e1000864.

WUST, S. 2004. Common Polymorphisms in the Glucocorticoid Receptor Gene Are Associated with Adrenocortical Responses to Psychosocial Stress. *Journal of Clinical Endocrinology & Metabolism,* 89**,** 565-573.

XIAO, S., WANG, J., JIANG, J., CAO, X., WU, G. & ZHAO, H. 2009. Characterization of Fyn signaling on the age-dependent immuno-modulation on traumatic rats. *Brain Res,* 1255**,** 162-9.

YAKA, R., PHAMLUONG, K. & RON, D. 2003. Scaffolding of Fyn kinase to the NMDA receptor determines brain region sensitivity to ethanol. *J Neurosci,* 23**,** 3623-32.

YALACHKOV, Y., KAISER, J. & NAUMER, M. J. 2009. Brain regions related to tool use and action knowledge reflect nicotine dependence. *J Neurosci,* 29**,** 4922-9.

YAMAMOTO, T., NAKAHATA, Y., TANAKA, M., YOSHIDA, M., SOMA, H., SHINOHARA, K., YASUDA, A., MAMINE, T. & TAKUMI, T. 2005. Acute physical stress elevates mouse period1 mRNA expression in mouse peripheral tissues via a glucocorticoid-responsive element. *J Biol Chem,* 280**,** 42036-43.

YAN, J., WANG, H., LIU, Y. & SHAO, C. 2008. Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput Biol,* 4**,** e1000193.

YANG, J., MANOLIO, T. A., PASQUALE, L. R., BOERWINKLE, E., CAPORASO, N., CUNNINGHAM, J. M., DE ANDRADE, M., FEENSTRA, B., FEINGOLD, E., HAYES, M. G., HILL, W. G., LANDI, M. T., ALONSO, A., LETTRE, G., LIN, P., LING, H., LOWE, W., MATHIAS, R. A., MELBYE, M., PUGH, E., CORNELIS, M. C., WEIR, B. S., GODDARD, M. E. & VISSCHER, P. M. 2011. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet,* 43**,** 519-25.

YANG, S., LIU, A., WEIDENHAMMER, A., COOKSEY, R. C., MCCLAIN, D., KIM, M. K., AGUILERA, G., ABEL, E. D. & CHUNG, J. H. 2009. The role of mPer2 clock gene in glucocorticoid and feeding rhythms. *Endocrinology,* 150**,** 2153-60.

YUN, H. M., KIM, S., KIM, H. J., KOSTENIS, E., KIM, J. I., SEONG, J. Y., BAIK, J. H. & RHIM, H. 2007. The novel cellular mechanism of human 5-HT6 receptor through an interaction with Fyn. *J Biol Chem,* 282**,** 5496-505.

ZEILINGER, S., KUHNEL, B., KLOPP, N., BAURECHT, H., KLEINSCHMIDT, A., GIEGER, C., WEIDINGER, S., LATTKA, E., ADAMSKI, J., PETERS, A., STRAUCH, K., WALDENBERGER, M. & ILLIG, T. 2013. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One,* 8**,** e63812.

ZGHOUL, T., ABARCA, C., SANCHIS-SEGURA, C., ALBRECHT, U., SCHUMANN, G. & SPANAGEL, R. 2007. Ethanol self-administration and reinstatement of ethanol-seeking behavior in Per1(Brdm1) mutant mice. *Psychopharmacology (Berl),* 190**,** 13-9.

ZHANG, B., KIROV, S. & SNODDY, J. 2005. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res,* 33**,** W741-8.

ZHANG, R., MIAO, Q., WANG, C., ZHAO, R., LI, W., HAILE, C. N., HAO, W. & ZHANG, X. Y. 2013. Genome-wide DNA methylation analysis in alcohol dependence. *Addict Biol,* 18**,** 392-403.

ZIMMERMANN, U., SPRING, K., KUNZ-EBRECHT, S. R., UHR, M., WITTCHEN, H. U. & HOLSBOER, F. 2004. Effect of ethanol on hypothalamic-pituitary-adrenal system response to psychosocial stress in sons of alcohol-dependent fathers. *Neuropsychopharmacology,* 29**,** 1156-65.

ZYLKA, M. J., SHEARMAN, L. P., WEAVER, D. R. & REPPERT, S. M. 1998. Three period homologs in mammals: differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. *Neuron,* 20**,** 1103-10.

Appendices

Appendix 1: Life Events Questionnaire

Questions for each life event in the baseline assessment conducted in adolescents at age 14:

1) How would this event make you feel?

2) Has this event EVER occurred?

3) Has this event occurred IN THE PAST YEAR?

Questions for each life event in the follow-up assessment conducted in adolescents at age 16:

1) Has this event EVER occurred?

2) How old were you when this happened?

3) How did this event make you feel?

Life events items

leq_01 'Parents divorced'.

leq_02 'Family accident or illness'.

leq_03 'Found a new group of friends'.

leq_04 'Got in trouble with the law'.

leq_05 'Stole something valuable'.

leq_06 'Given medication by physician'.

leq_07_year 'Fell in love'.

leq_08_year 'Death in family'.

leq_09_year 'Face broke out with pimples'.

leq_10_year 'Brother or sister moved out'.

leq_11_feel 'Started seeing a therapist'.

leq_12_year 'Parent changed jobs'.

leq_13_year 'Began a time-consuming hobby'.

leq_14_year 'Got or made pregnant'.

leq_15_year 'Decided about college / university'.

leq_16_year 'Thought about suicide'.

leq_17_year 'Changed schools'.

leq_18_year 'Joined a club or group'.

leq_19_year 'Got in trouble at school'.

leq_20_year 'Got or gave sexually transmitted disease'.

leq_21_year 'Met a teacher I liked a lot'.

leq_22_year 'Family had money problems'.

leq_23_year 'Got own TV or computer'.

leq_24_year 'Parents argued or fought'.

leq_25_year 'Ran away from home'.

leq_26_year 'Started going out with a girlfriend/boyfriend'.

leq_27_year 'Got poor grades in school'.

leq_28_year 'Went on holiday without parents'.

leq_29_year 'Started driving a motor vehicle'.

leq_30_year 'Broke up with boy/ girl-friend'.

leq_31_year 'Family moved'.

leq_32_year 'Started making own money'.

leq_33_year 'Found religion'.

leq_34_year 'Parent remarried'.

leq_35_year 'Had a gay experience'.

leq_36_year 'Gained a lot of weight'.

leq_37_year 'Serious accident or illness'.

leq_38_year 'Lost virginity'.

leq_39_year 'Parent abused alcohol'.

Appendix 2: ESPAD

General information about « missings » or jump rules

CHILD ESPAD

6 – If(0) 8a

8a – If(0) 22a

8b – If(0) 29beer

8c – If(0) 29beer

17a – If(0) prev31

17b – If(0) prev31

19a – If(0) 22a

19b – If(0) 29d

24a – If(0) 25a

24b – If(0) 29i

24c – If(0) 29i

25a – If(0) 27tranq

25b – If(0) 29r

25c – If(0) 29r

27* - If(NO) JUMP 4

1) Variable labels

- TRUANCY

VARIABLE LABELS C.4a 'During the LAST 30 DAYS how many whole days of school have you missed because of illness?'.

VARIABLE LABELS C.4b 'During the LAST 30 DAYS how many whole days of school have you missed because you skipped or ”cut”?'.

VARIABLE LABELS C.4c 'During the LAST 30 DAYS how many whole days of school have you missed for other reasons?'.

- SCHOOL PERFORMANCE

VARIABLE LABELS C.5 'Which of the following best describes your average grade in the end of the last term?'.

- SMOKING (FTND)

VARIABLE LABELS C.6 'On how many occasions during your lifetime have you smoked cigarettes?'.

VARIABLE LABELS C.29e 'When did you smoke your FIRST cigarette?'.

VARIABLE LABELS C.29f 'When did you first smoke cigarettes ON A DAILY BASIS?'.

VARIABLE LABELS C.7 'How frequently have you smoked cigarettes during the LAST 30 DAYS?'.

VARIABLE LABELS C.ftnd1 'How soon after you wake up do you smoke your first cigarette?'.

VARIABLE LABELS C.ftnd2 'Do you find it difficult to refrain from smoking in places where it is forbidden, e.g. at the mosque (church), at the bus?' .

VARIABLE LABELS C.ftnd3 'Which cigarette would you hate most to give up?'.

VARIABLE LABELS C.ftnd4 'How many cigarettes/day do you smoke?'.

VARIABLE LABELS C.ftnd5 'Do you smoke more frequently during the first hours after waking than the rest of the day?'.

VARIABLE LABELS C.ftnd6 'Do you smoke when you are so ill that you are in bed most of the day?'.

- ALCOHOL USE ONSET + FREQUENCY, TYPE OF ALCOHOL

VARIABLE LABELS C.8a 'On how many occasions IN YOUR WHOLE LIFETIME have you had any alcoholic beverage to drink?'.

VARIABLE LABELS C.8b 'On how many occasions OVER THE LAST 12 MONTHS have you had any alcoholic beverage to drink?'.

VARIABLE LABELS C.8c 'On how many occasions OVER THE LAST 30 DAYS have you had any alcoholic beverage to drink?'.

VARIABLE LABELS C.9a 'On how many occasions OVER THE LAST 30 DAYS have you had beer to drink? (do not include low alcohol beer)'.

VARIABLE LABELS C.29beer 'When did you FIRST drink beer (at least one glass)?'.

VARIABLE LABELS C.9b 'On how many occasions OVER THE LAST 30 DAYS have you had wine to drink?'.

VARIABLE LABELS C.29wine 'When did you FIRST drink wine (at least one glass)?'.

VARIABLE LABELS C.9c 'On how many occasions OVER THE LAST 30 DAYS have you had an alcopop to drink? (e.g. Barcardi Breezer)'.

VARIABLE LABELS C.29alcopop 'When did you FIRST drink alcopops (at least one bottle)?'.

VARIABLE LABELS C.9d 'On how many occasions OVER THE LAST 30 DAYS have you had spirits to drink? (whisky, cognac, shot drinks etc) (also include spirits mixed with soft drinks, except alcopops)'.

VARIABLE LABELS C.29spirits 'When did you FIRST drink spirits (at least one glass)?'.

VARIABLE LABELS C.17a 'How many times IN YOUR WHOLE LIFETIME have you had five or more drinks in a row?'.

VARIABLE LABELS C.17b 'How many times OVER THE LAST 12 MONTHS have you had five or more drinks in a row?'.

VARIABLE LABELS C.17c 'How many times OVER THE LAST 30 DAYS have you had five or more drinks in a row?'.

- ALCOHOL USE EXPECTANCIES (only in those who report alcohol use): “How likely is it that each of the following things would happen to you personally, if you drink alcohol?”

VARIABLE LABELS C.18a 'Feel relaxed'.

VARIABLE LABELS C.18b 'Get into trouble with police'.

VARIABLE LABELS C.18c 'Harm my health'.

VARIABLE LABELS C.18d 'Feel happy'.

VARIABLE LABELS C.18e 'Forget my problems'.

VARIABLE LABELS C.18f 'Not be able to stop drinking'.

VARIABLE LABELS C.18g 'Get a hangover'.

VARIABLE LABELS C.18h 'Feel more friendly and outgoing'.

VARIABLE LABELS C.18i 'Do something I would regret'.

VARIABLE LABELS C.18j 'Have a lot of fun'.

VARIABLE LABELS C.18k 'Feel sick'.

- BINGEING

VARIABLE LABELS C.prev31 'How many drinks containing alcohol do you have on a TYPICAL DAY when you are drinking?'.

VARIABLE LABELS C.19a 'On how many occasions IN YOUR WHOLE LIFETIME have you been drunk from drinking alcoholic beverages?'.

VARIABLE LABELS C.19b 'On how many occasions OVER THE LAST 12 MONTHS have you been drunk from drinking alcoholic beverages?'.

VARIABLE LABELS C.19c 'On how many occasions OVER THE LAST 30 DAYS have you been drunk from drinking alcoholic beverages?'.

VARIABLE LABELS C.29d 'When did you FIRST get drunk from drinking alcoholic beverages?'.

VARIABLE LABELS C.20 'Please indicate on this scale from 1 to 10 how drunk you would say you were the last time you were drunk.'.

VARIABLE LABELS C.21 'How many drinks do you usually need to get drunk?'.

- DRUG USE: “Have you ever heard of…”

VARIABLE LABELS C.22_hash 'Marijuana (weed, grass, skunk, pot, ganja) or hashish (hash, hash oil)'.

VARIABLE LABELS C.22_inhalants 'Inhalants like glue or aerosols'.

VARIABLE LABELS C.22_tranq 'Tranquillisers or sedatives (e.g. benzodiazepines such as valium or xanax; barbiturates, barbs or downers such as amytal, or seconal).'.

VARIABLE LABELS C.22_amphet 'Amphetamines (speed, uppers, crystal meth) or desoxyn.'.

VARIABLE LABELS C.22_lsd 'LSD (acid)'.

VARIABLE LABELS C.22_mushrooms 'Magic mushrooms or other hallucinogens (excluding LSD)'.

VARIABLE LABELS C.22_crack 'Crack'.

VARIABLE LABELS C.22_coke 'Cocaine (coke)'.

VARIABLE LABELS C.22_relevin 'Relevin'.

VARIABLE LABELS C.22_heroin 'Heroin'.

VARIABLE LABELS C.22_narcotic 'Narcotics (e.g. opium, morphine, codeine)'.

VARIABLE LABELS C.22_mdma 'Ecstasy (MDMA)'.

VARIABLE LABELS C.22_ketamine 'Ketamine (Ket, K) or Phencyclidine (PCP, or angel dust)'.

VARIABLE LABELS C.22_ghb 'GHB or liquid ecstasy'.

VARIABLE LABELS C.22_anabolic 'Anabolic steroids'.

VARIABLE LABELS C.23 'Have you ever wanted to try any of the drugs mentioned in the previous questions?'.

- DRUG USE ONSET + FREQUENCY for marijuana, inhalants, tranquilisers, amphetamines, LSD, magic mushrooms, crack, cocaine, relevin, heroin, narcotics, MDMA, ketamine, GHB, anabolic steroids

VARIABLE LABELS C.life_hash 'On how many occasions IN YOUR WHOLE LIFETIME have you used marijuana (grass, pot) or hashish (hash, hash oil)?'.

VARIABLE LABELS C.year_hash 'On how many occasions OVER THE LAST 12 MONTHS have you used marijuana (grass, pot) or hashish (hash, hash oil)?'.

VARIABLE LABELS C.month_hash 'On how many occasions OVER THE LAST 30 DAYS have you used marijuana (grass, pot) or hashish (hash, hash oil)?'.

VARIABLE LABELS C.week_hash 'On how many occasions OVER THE LAST WEEK have you used marijuana (grass, pot) or hashish (hash, hash oil)?'.

VARIABLE LABELS C.first_hash 'When did you first try marijuana (grass, pot) or hashish (hash, hash oil)?'.

VARIABLE LABELS C.life_glue 'On how many occasions IN YOUR WHOLE LIFETIME have you sniffed a substance (glue, aerosols etc) to get high??'.

VARIABLE LABELS C.year_glue 'On how many occasions OVER THE LAST 12 MONTHS have you sniffed a substance (glue, aerosols etc) to get high??'.

VARIABLE LABELS C.month_glue 'On how many occasions OVER THE LAST 30 DAYS have you sniffed a substance (glue, aerosols etc) to get high??'.

VARIABLE LABELS C.week_glue 'On how many occasions OVER THE LAST WEEK have you sniffed a substance (glue, aerosols etc) to get high??'.

VARIABLE LABELS C.first_glue 'When did you first use Inhalants (glue, etc) to get high?'.

VARIABLE LABELS C.life_tranq 'On how many occasions IN YOUR WHOLE LIFETIME have you used tranquillisers or sedatives (without a doctors prescription) (e.g. benzodiazepines such as valium or xanax; barbiturates, barbs or downers such as amytal, or seconal)?'.

VARIABLE LABELS C.year_tranq 'On how many occasions OVER THE LAST 12 MONTHS have you used tranquillisers or sedatives (without a doctors prescription)

(e.g. benzodiazepines such as valium or xanax; barbiturates, barbs or downers such as amytal, or seconal)?'.

VARIABLE LABELS C.month_tranq 'On how many occasions OVER THE LAST 30 DAYS have you used tranquillisers or sedatives (without a doctors prescription) (e.g. benzodiazepines such as valium or xanax; barbiturates, barbs or downers such as amytal, or seconal)?'.

VARIABLE LABELS C.week_tranq 'On how many occasions OVER THE LAST WEEK have you used tranquillisers or sedatives (without a doctors prescription) (e.g. benzodiazepines such as valium or xanax; barbiturates, barbs or downers such as amytal, or seconal)?'.

VARIABLE LABELS C.first_tranq 'When did you first try tranquillisers or sedatives (without a doctors prescription) (e.g. benzodiazepines such as valium or xanax; barbiturates, barbs or downers such as amytal, or seconal)?'.

VARIABLE LABELS C.life_amphet 'On how many occasions IN YOUR WHOLE LIFETIME have you used amphetamines (speed), methamphetamine (crystal meth) or desoxyn?'.

VARIABLE LABELS C.year_amphet 'On how many occasions OVER THE LAST 12 MONTHS have you used amphetamines (speed), methamphetamine (crystal meth) or desoxyn?'.

VARIABLE LABELS C.month_amphet 'On how many occasions OVER THE LAST 30 DAYS have you used amphetamines (speed), methamphetamine (crystal meth) or desoxyn?'.

VARIABLE LABELS C.week_amphet 'On how many occasions OVER THE LAST WEEK have you used amphetamines (speed), methamphetamine (crystal meth) or desoxyn?'.

VARIABLE LABELS C.first_amphet 'When did you first try amphetamines (speed), methamphetamine (crystal meth) or desoxyn?'.

VARIABLE LABELS C.life_lsd 'On how many occasions IN YOUR WHOLE LIFETIME have you used LSD (acid)?'.

VARIABLE LABELS C.year_lsd 'On how many occasions OVER THE LAST 12 MONTHS have you used LSD (acid)?'.

VARIABLE LABELS C.month_lsd 'On how many occasions OVER THE LAST 30 DAYS have you used LSD (acid)?'.

VARIABLE LABELS C.week_lsd 'On how many occasions OVER THE LAST WEEK have you used LSD (acid)?'.

VARIABLE LABELS C.first_lsd 'When did you first try LSD (acid)?'.

VARIABLE LABELS C.life_mushrooms 'On how many occasions IN YOUR WHOLE LIFETIME have you used magic mushrooms or other hallucinogens (excluding LSD)?'.

VARIABLE LABELS C.year_mushrooms 'On how many occasions OVER THE LAST 12 MONTHS have you used magic mushrooms or other hallucinogens (excluding LSD)?'.

VARIABLE LABELS C.month_mushrooms 'On how many occasions OVER THE LAST 30 DAYS have you used magic mushrooms or other hallucinogens (excluding LSD)?'.

VARIABLE LABELS C.week_mushrooms 'On how many occasions OVER THE LAST WEEK have you used magic mushrooms or other hallucinogens (excluding LSD)?'.

VARIABLE LABELS C.first_mushrooms 'When did you first try magic mushrooms or other hallucinogens (excluding LSD)?'.

VARIABLE LABELS C.life_crack 'On how many occasions IN YOUR WHOLE LIFETIME have you used crack?'.

VARIABLE LABELS C.year_crack 'On how many occasions OVER THE LAST 12 MONTHS have you used crack?'.

VARIABLE LABELS C.month_crack 'On how many occasions OVER THE LAST 30 DAYS have you used crack?'.

VARIABLE LABELS C.week_crack 'On how many occasions OVER THE LAST WEEK have you used crack?'.

VARIABLE LABELS C.first_crack 'When did you first try crack?'.

VARIABLE LABELS C.life_coke 'On how many occasions IN YOUR WHOLE LIFETIME have you used cocaine (coke)?'.

VARIABLE LABELS C.year_coke 'On how many occasions OVER THE LAST 12 MONTHS have you used cocaine (coke)?'.

VARIABLE LABELS C.month_coke 'On how many occasions OVER THE LAST 30 DAYS have you used cocaine (coke)?'.

VARIABLE LABELS C.week_coke 'On how many occasions OVER THE LAST WEEK have you used cocaine (coke)?'.

VARIABLE LABELS C.first_coke 'When did you first try cocaine (coke)?'.

VARIABLE LABELS C.life_relevin 'On how many occasions IN YOUR WHOLE LIFETIME have you used relevin?'.

VARIABLE LABELS C.year_relevin 'On how many occasions OVER THE LAST 12 MONTHS have you used relevin?'.

VARIABLE LABELS C.month_relevin 'On how many occasions OVER THE LAST 30 DAYS have you used relevin?'.

VARIABLE LABELS C.week_relevin 'On how many occasions OVER THE LAST WEEK have you used relevin?'.

VARIABLE LABELS C.first_relevin 'When did you first try relevin?'.

VARIABLE LABELS C.life_heroin 'On how many occasions IN YOUR WHOLE LIFETIME have you used heroin?'.

VARIABLE LABELS C.year_heroin 'On how many occasions OVER THE LAST 12 MONTHS have you used heroin?'.

VARIABLE LABELS C.month_heroin 'On how many occasions OVER THE LAST 30 DAYS have you used heroin?'.

VARIABLE LABELS C.week_heroin 'On how many occasions OVER THE LAST WEEK have you used heroin?'.

VARIABLE LABELS C.first_heroin 'When did you first try heroin?'.

VARIABLE LABELS C.life_narcotic 'On how many occasions IN YOUR WHOLE LIFETIME have you used narcotics (e.g. opium, morphine, codeine)?'.

VARIABLE LABELS C.year_narcotic 'On how many occasions OVER THE LAST 12 MONTHS have you used narcotics (e.g. opium, morphine, codeine)?'.

VARIABLE LABELS C.month_narcotic 'On how many occasions OVER THE LAST 30 DAYS have you used narcotics (e.g. opium, morphine, codeine)?'.

VARIABLE LABELS C.week_narcotic 'On how many occasions OVER THE LAST WEEK have you used narcotics (e.g. opium, morphine, codeine)?'.

VARIABLE LABELS C.first_narcotic 'When did you first try narcotics (e.g. opium, morphine, codeine)?'.

VARIABLE LABELS C.life_mdma 'On how many occasions IN YOUR WHOLE LIFETIME have you used ecstasy (MDMA)?'.

VARIABLE LABELS C.year_mdma 'On how many occasions OVER THE LAST 12 MONTHS have you used ecstasy (MDMA)?'.

VARIABLE LABELS C.month_mdma 'On how many occasions OVER THE LAST 30 DAYS have you used ecstasy (MDMA)?'.

VARIABLE LABELS C.week_mdma 'On how many occasions OVER THE LAST WEEK have you used ecstasy (MDMA)?'.

VARIABLE LABELS C.first_mdma 'When did you first try ecstasy (MDMA)?'.

VARIABLE LABELS C.life_ketamine 'On how many occasions IN YOUR WHOLE LIFETIME have you used ketamine (Ket, K) or Phencyclidine (PCP, or angel dust)?'.

VARIABLE LABELS C.year_ketamine 'On how many occasions OVER THE LAST 12 MONTHS have you used ketamine (Ket, K) or Phencyclidine (PCP, or angel dust)?'.

VARIABLE LABELS C.month_ketamine 'On how many occasions OVER THE LAST 30 DAYS have you used ketamine (Ket, K) or Phencyclidine (PCP, or angel dust)?'.

VARIABLE LABELS C.week_ketamine 'On how many occasions OVER THE LAST WEEK have you used ketamine (Ket, K) or Phencyclidine (PCP, or angel dust)?'.

VARIABLE LABELS C.first_ketamine 'When did you first try ketamine (Ket, K) or Phencyclidine (PCP, or angel dust)?'.

VARIABLE LABELS C.life_ghb 'On how many occasions IN YOUR WHOLE LIFETIME have you used GHB or liquid ecstasy?'.

VARIABLE LABELS C.year_ghb 'On how many occasions OVER THE LAST 12 MONTHS have you used GHB or liquid ecstasy?'.

VARIABLE LABELS C.month_ghb 'On how many occasions OVER THE LAST 30 DAYS have you used GHB or liquid ecstasy?'.

VARIABLE LABELS C.week_ghb 'On how many occasions OVER THE LAST WEEK have you used GHB or liquid ecstasy?'.

VARIABLE LABELS C.first_ghb 'When did you first try GHB or liquid ecstasy?'.

VARIABLE LABELS C.life_anabolic 'On how many occasions IN YOUR WHOLE LIFETIME have you used anabolic steroids?'.

VARIABLE LABELS C.year_anabolic 'On how many occasions OVER THE LAST 12 MONTHS have you used anabolic steroids?'.

VARIABLE LABELS C.month_anabolic 'On how many occasions OVER THE LAST 30 DAYS have you used anabolic steroids?'.

VARIABLE LABELS C.week_anabolic 'On how many occasions OVER THE LAST WEEK have you used anabolic steroids?'.

VARIABLE LABELS C.first_anabolic 'When did you first try anabolic steroids?'.

2) Value labels

VALUE LABELS C.4a 1 'None' 2 '1 day' 3 '2 days' 4 '3-4 days' 5 '5-6 days' 6 '7 days or more' .

VALUE LABELS C.4b 1 'None' 2 '1 day' 3 '2 days' 4 '3-4 days' 5 '5-6 days' 6 '7 days or more' .

VALUE LABELS C.4c 1 'None' 2 '1 day' 3 '2 days' 4 '3-4 days' 5 '5-6 days' 6 '7 days or more' .

VALUE LABELS C.5 1 'A (93-100)' 2 'A- (90-92)' 3 'B+ (87-89)' 4 'B (83-86)' 5 'B- (80-82)' 6 'C+ (77-79)' 7 'C (73-76)' 8 'C- (70-72)' .

VALUE LABELS C.6 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.29e 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old'.

VALUE LABELS C.29f 0 'Never' 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.7 0 'Not at all' 1 'Less than 1 cigarette per week' 2 'Less than 1 cigarette per day' 3 '1-5 cigarettes per day' 4 '6-10 cigarettes per day' 5 '11-20 cigarettes per day' 6 'More than 20 cigarettes per day' 999 'non smoker' .

VALUE LABELS C.ftnd1 3 'Within 5 minutes' 2 '6-30 minutes' 1 '31-60 minutes' 0 'after 60 minutes' 999 'non smoker' .

VALUE LABELS C.ftnd2 1 'Yes' 0 'No' 999 'non smoker' .

VALUE LABELS C.ftnd3 1 'The first one in the morning' 0 'All others' 999 'non smoker' .

VALUE LABELS C.ftnd4 0 '10 or less' 1 '11-20' 2 '21-30' 3 '31 or more' 999 'non smoker' .

VALUE LABELS C.ftnd5 1 'Yes' 0 'No' 999 'non smoker' .

VALUE LABELS C.ftnd6 1 'Yes' 0 'No' 999 'non smoker' .

VALUE LABELS C.8a 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.8b 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.8c 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.9a 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.29beer 0 'Never' 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.9b 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.29wine 0 'Never' 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.9c 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.29alcopop 0 'Never' 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.9d 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.29spirits 0 'Never' 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.17a 0 '0' 1 '1' 2 '2' 3 '3-5' 4 '6-9' 5 '10 or more times' .

VALUE LABELS C.17b 0 '0' 1 '1' 2 '2' 3 '3-5' 4 '6-9' 5 '10 or more times' .

VALUE LABELS C.17c 0 '0' 1 '1' 2 '2' 3 '3-5' 4 '6-9' 5 '10 or more times' .

VALUE LABELS C.prev31 1 '1 or 2' 2 '3 or 4' 3 '5 or 6' 4 '7 to 9' 5 '10 or more' .

VALUE LABELS C.18a 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18b 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18c 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18d 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18e 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18f 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18g 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18h 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18i 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18j 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.18k 1 'Very likely' 2 'Likely' 3 'Unsure' 4 'Unlikely' 5 'Very unlikely' .

VALUE LABELS C.19a 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.19b 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.19c 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.29d 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.20 1 '1 Somewhat merry only' 2 '2' 3 '3' 4 '4' 5 '5' 6 '6' 7 '7' 8 '8' 9 '9' 10 '10 Heavily intoxicated to the point of being unable to stand on my feet.' .

VALUE LABELS C.21 1 '1-2 drinks' 2 '3-4 drinks' 3 '5-6 drinks' 4 '7-8 drinks' 5 '9-10 drinks' 6 '11-12 drinks' 7 '13 drinks or more' .

VALUE LABELS C.22_hash 1 'Yes' 0 'No' .

VALUE LABELS C.22_inhalants 1 'Yes' 0 'No' .

VALUE LABELS C.22_tranq 1 'Yes' 0 'No' .

VALUE LABELS C.22_amphet 1 'Yes' 0 'No' .

VALUE LABELS C.22_lsd 1 'Yes' 0 'No' .

VALUE LABELS C.22_mushrooms 1 'Yes' 0 'No' .

VALUE LABELS C.22_crack 1 'Yes' 0 'No' .

VALUE LABELS C.22_coke 1 'Yes' 0 'No' .

VALUE LABELS C.22_relevin 1 'Yes' 0 'No' .

VALUE LABELS C.22_heroin 1 'Yes' 0 'No' .

VALUE LABELS C.22_narcotic 1 'Yes' 0 'No' .

VALUE LABELS C.22_mdma 1 'Yes' 0 'No' .

VALUE LABELS C.22_ketamine 1 'Yes' 0 'No' .

VALUE LABELS C.22_ghb 1 'Yes' 0 'No' .

VALUE LABELS C.22_anabolic 1 'Yes' 0 'No' .

VALUE LABELS C.23 1 'Yes' 0 'No' .

VALUE LABELS C.life_hash 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_hash 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_hash 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_hash 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_hash 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_glue 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_glue 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_glue 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_glue 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_glue 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_tranq 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_tranq 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_tranq 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_tranq 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_tranq 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_amphet 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_amphet 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_amphet 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_amphet 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_amphet 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_lsd 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_lsd 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_lsd 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_lsd 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_lsd 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_mushrooms 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_mushrooms 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_mushrooms 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_mushrooms 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_mushrooms 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_crack 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_crack 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_crack 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_crack 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_crack 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_coke 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_coke 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_coke 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_coke 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_coke 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_relevin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_relevin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_relevin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_relevin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_relevin 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_heroin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_heroin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_heroin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_heroin 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_heroin 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_narcotic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_narcotic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_narcotic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_narcotic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_narcotic 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_mdma 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_mdma 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_mdma 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_mdma 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_mdma 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_ketamine 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_ketamine 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_ketamine 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_ketamine 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_ketamine 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_ghb 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_ghb 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_ghb 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_ghb 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_ghb 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

VALUE LABELS C.life_anabolic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.year_anabolic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.month_anabolic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.week_anabolic 0 '0' 1 '1-2' 2 '3-5' 3 '6-9' 4 '10-19' 5 '20-39' 6 '40 or more' .

VALUE LABELS C.first_anabolic 11 '11 years old or less' 12 '12 years old' 13 '13 years old' 14 '14 years old' 15 '15 years old' 16 '16 years old' .

Appendix 3: Puberty Development Scale (PDS)

1) Variable labels

For Female adolescents:

VARIABLE LABELS a8_f 'Would you say that your growth in height: …?'.

VARIABLE LABELS a9_f 'And how about the growth of body hair (body hair means underarm and pubic hair), would you say that your body hair has:…?'.

VARIABLE LABELS a10_f 'Have you noticed any skin changes, especially pimples?'.

VARIABLE LABELS a11_f 'Have your breasts begun to grow?'.

VARIABLE LABELS a12a_f 'Have you begun to menstruate?'.

VARIABLE LABELS a12b_f 'How old were you when you had your first period?'.

VARIABLE LABELS a13_f 'Do you think your development is any earlier or later than most other girls your age?'.

For Male adolescents:

VARIABLE LABELS a8_m 'Would you say that your growth in height: …?'.

VARIABLE LABELS a9_m 'And how about the growth of body hair (body hair means underarm and pubic hair), would you say that your body hair has:…?'.

VARIABLE LABELS a10_m 'Have you noticed any skin changes, especially pimples?'.

VARIABLE LABELS a11_m 'Have you noticed a deepening of your voice?'.

VARIABLE LABELS a12_m 'Have you begun to grow hair on your face?'.

VARIABLE LABELS a13_m 'Do you think your development is any earlier or later than most other boys your age?'.

2) Value labels

For Female adolescents:

VALUE LABELS a8_f 1 'Has not yet begun to spurt (spurt means more growth than usual)' 2 'Has barely started' 3 'Is definitely underway' 4 'Seems completed' .

VALUE LABELS a9_f 1 'Not yet started growing' 2 'Has barely started growing' 3 'Is definitely underway' 4 'Seems completed' .

VALUE LABELS a10_f 1 'Not yet started showing changes' 2 'Have barely started showing changes' 3 'Skin changes are definitely underway' 4 'Skin changes seem completed' .

VALUE LABELS a11_f 1 'Not yet started growing' 2 'Has barely started changing' 3 'Breast growth is definitely underway' 4 'Breast growth seems completed' .

VALUE LABELS a12a_f 1 'Yes' 0 'No' .

VALUE LABELS a12b_f 10 '10 years or younger' 11 '11' 12 '12' 13 '13' 14 '14' .

VALUE LABELS a13_f 5 'Much earlier' 4 'Somewhat earlier' 3 'About the same' 2 'Somewhat later' 1 'Much later' .

For Male adolescents:

VALUE LABELS a8_m 1 'Has not yet begun to spurt (spurt means more growth than usual)' 2 'Has barely started' 3 'Is definitely underway' 4 'Seems completed' .

VALUE LABELS a9_m 1 'Not yet started growing' 2 'Has barely started growing' 3 'Is definitely underway' 4 'Seems completed' .

VALUE LABELS a10_m 1 'Not yet started showing changes' 2 'Have barely started showing changes' 3 'Skin changes are definitely underway' 4 'Skin changes seem completed' .

VALUE LABELS a11_m 1 'Not yet started changing' 2 'Has barely started changing' 3 'Voice change is definitely underway' 4 'Voice change seems completed' .

VALUE LABELS a12_m 1 'Not yet started growing hair' 2 'Has barely started growing hair' 3 'Facial hair growth is definitely underway' 4 'Facial hair growth seems completed' .

VALUE LABELS a13_m 5 'Much earlier' 4 'Somewhat earlier' 3 'About the same' 2 'Somewhat later' 1 'Much later' .

Appendix 4: Quality control questions for ESPAD and PDS

Variable name	Q/A
ts_1	What computer are you using for this session?
	1 computer at home
	2 computer at the research institute
	3 computer at school
	4 computer at a friend’s home
	5 computer at a library

6 computer in an internet café

7 computer at a parent's office

8 computer elsewhere

ts_2 Are you in a hurry, e.g. are you rushing off to another activity very soon?

1 yes

2 no

ts_3 How do you feel right now?

1 relaxed

2 tense

3 annoyed

4 happy

5 angry

6 curious

ts_4 Who is with you in the room?

1 I am alone

2 Others are nearby

3 Another person is sitting next to me and watching my responses

ts_5 Please tell us about the noise level in the room:

1 The room is quiet, there is no noise

2 There is some noise (e.g. tv, music, talking) but it does not disturb me

3 There is noise that disturbs me (e.g. tv, music, talking) but I can lower it

4 There is a lot of noise and it is quite distracting

5 I listen to music because I think this might help me focusing on the tasks

ts_6 How tired are you at the moment?

1 I am full of energy

2 I am at a normal level of energy

3 I am a bit tired

4 I am quite tired

5 I feel like I am falling asleep

Psytools Valid Flag	0 not valid 1 valid

The Psytools Valid flag is shown as ‘invalid’ if at least one of the following cases apply:

1) Automated flags for context questions:
ts_2 = "1", ts_4 = "3", ts_5 ="4" or ts_5 = "5"

1a) The questions ts_1 – ts_4 are asked before each of the following Psytools questionnaires:

ESPAD_CHILD
PDS

Appendix 5: Stress genes

Enclosed in CD-ROM

Appendix 6: Circadian genes

Enclosed in CD-ROM

Appendix 7: Overlap stress x circadian genes

Enclosed in CD-ROM

Appendix 8: Neurotransmitter stress x circadian genes

Enclosed in CD-ROM