Understanding genetic relationships between circadian function and mood disorders

Amy Christina Ferguson

BSc. Hons

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

School of Life Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

April 2019

Abstract

Mood disorders are amongst the most prevalent and disabling conditions worldwide. There is increasing evidence for the involvement of disrupted circadian rhythms in mood disorders. The mechanism of associations between circadian dysfunction and mood disorders are complex and not fully understood. This thesis explores the influence of genetic variation of circadian function on mood disorder-related phenotypes within two relatively large cohorts, ALSPAC (N=8,100) and UK Biobank (N=500,000). I investigated genetic variants associated with different features of circadian function and how genetic loading for these common variants was associated with risk of mood disorders and related traits. To my knowledge, this is the first application of circadian polygenic risk scores to investigate mood disorder risk.

Both a priori candidate gene profile risk scores (*CACNA1C*) and polygenic risk scores (PRS) were used to investigate the relationship between the genetics of circadian function and mood disorder-related phenotypes. A genome-wide association study (GWAS) was carried out to identify common variants associated with circadian rest/activity rhythmicity and to assess genetic correlation with mood disorders. Mendelian randomisation was used to assess the direction of the relationship between circadian dysfunction and mood disorders.

Chronotype polygenic risk scores (specifically 'eveningness' PRS) were associated with increased risk of bipolar disorder in UK Biobank and with hypomanic features in ALSPAC. The GWAS of low relative amplitude (a measure of circadian rest/activity rhythmicity) identified several associated variants and these variants were used to create a PRS for low relative amplitude. Increased PRS for low relative amplitude was associated with mood instability in UK Biobank.

There are limitations to the population cohorts used in these analyses. They may be under-representative of individuals with clinically-diagnosed mood disorders. Also, the mood phenotypes tested were based on self-report which could be vulnerable to response biases. The polygenic risk scores had small but significant effects on the mood disorder phenotypes investigated.

This work identified associations between genetic variation of circadian function and mood disorder-related phenotypes in both ALSPAC and UK Biobank. With expansion, development and replication, PRS of circadian function could

inform treatment stratification approaches for mood disorders. This thesis also suggests a need for further investigation of the underlying biology of circadian function and how this relates to the pathophysiology of mood disorders.

Strengths to this thesis include the large sample sizes of the cohorts. The actigraph data obtained from UK Biobank allowed for the largest GWAS of rest/activity rhythmicity to date. The extensive self-report and interview-based data available in UK Biobank also provide a breadth of mood disorder-related phenotypes to investigate.

As this is one of the first examples of using circadian polygenic risk scores to investigate the underlying pathophysiology of mood disorders this work requires replication in other population cohorts. It would also be of interest to test these risk scores within clinical populations and assess the extent to which they may support clinical management decisions.

Table of Contents

List of Tables

List of Figures

Acknowledgement

I would like to thank Daniel Smith, Donald Lyall and Jill Pell for all of their guidance and supervision during this PhD thesis. Also the members of our research group for their technical support, advice and encouragement these last few years: thank you Breda, Joey, Keira, Laura, Mark, Nick and Rona.

I would also like to express my gratitude to the Medical Research Council for funding this PhD project, and to all of the participants of ALSPAC and UK Biobank contributing to the cohorts and for making this research possible.

Thank you to all of my family and friends, especially Jenifer, Lucy, Nicola and Rich. I would not have been able to complete this thesis without all of your love and support. I also want to give a special thanks to my mum, dad and brothers, Christina, Kenny, Scott and Ross Ferguson, you have kept me motivated and always been there when I needed you.

This thesis is dedicated to my family.

Author's Declaration

The research reported in this thesis is my own, any contributions from a research group have been acknowledged. This thesis has not been submitted for any other degree.

Amy Ferguson

April 2019

Definitions/Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AIS	Axon Initial Segment
ALSPAC	Avon Longitudinal Study of Parents And Children
BD	Bipolar Disorder
BDI	Bipolar Disorder Type 1
BDII	Bipolar Disorder Type 2
BD-NOS	Bipolar Disorder Non-Specified
CI	Confidence Interval
CIDI-SF	Composite International Diagnostic Interview Short Form
CNS	Central Nervous System
CREB	CAMP Response Element Binding Proteins
CVD	Cardiovascular Disease
DALY	Disability Adjusted Life Year
DHP	Dihydropyridines
DSM	Diagnostic and Statistical Manual of Mental Disorders
EPQ-R-S	Eysenck Personality Questionnaire-Revised Short Form
eQTL	Expression Quantitative Trait Loci
FDR	False Discovery Rate
GAD	Generalised Anxiety Disorder
GPRS	Genetic Profile Risk Score
GTEx	Genotype-Tissue Expression
GWAS	Genome-wide Association Study
h^2_{SNP}	SNP Heritability
HCL-28	Hypomania Checklist- 28
HCL-32	Hypomania Checklist- 32
HPA	Hypothalamic-pituitary-adrenal
HWE	Hardy-Weinberg Equilibrium
ICD	International Classification of Diseases
iPSC	Induced pluripotent Stem Cell
IPSRT	Interpersonal and Social Rhythm Therapy
IVW	Inverse-variance Weighting
LD	Linkage Disequilibrium
LDSR	Linkage Disequilibrium Score Regression
L-VGCC	L-type Voltage-gated Calcium Channel
MAF	Minor Allele Frequency
MAGMA	Multi-marker Analysis of GenoMic Annotation
MDD	Major Depressive Disorder
MHQ	Mental Health Questionnaire
MR	Mendelian Randomisation
O.R	Odds Ratio
pgACC	Perigenual Anterior Cingulate
PGC	Psychiatric Genomics Consortium
PHQ	Patient Health Questionnaire
PHQ-9	Patient Health Questionnaire 9
PRS	Polygenic Risk Score
PTSD	Post-traumatic Stress Disorder
QOL	Quality Of Life
RA	Relative Amplitude
RDoC	Research Domain Criteria

S.E	Standard Error
SCN	Suprachiasmatic Nucleus
SD	Standard Deviation
SMFQ	Short Mood and Feelings Questionnaire
SNP	Single Nucleotide Polymorphism
SR	Self-report
TF	Transcription Factor
UK	United Kingdom
WR	Wald Ratio

Chapter 1 Introduction

Mood disorders are amongst the most common psychiatric conditions, with depression reported as one of the world's leading causes of disability in adults and with increasing prevalence in adolescents (GBD 2016; Fabbri et al. 2018). It is widely accepted that an individual's susceptibility to psychiatric disorders is influenced by a complex mix of genetic and environmental factors (Yoshimizu et al. 2015; Wray et al. 2018).

There have been many studies aiming to elucidate the underlying genetic architecture of mood disorders. This introductory chapter will consider the research to date on the genetic underpinnings of bipolar disorder (BD) and major depressive disorder (MDD) in particular.

1.1 Introduction to Bipolar Disorder and Major Depressive Disorder

BD and MDD are complex, chronic conditions reported as leading causes of disability worldwide; as such these disorders are important public health problems (Palagini et al. 2018; GBD 2016; Fabbri et al. 2018).
BD and MDD are among the most prevalent conditions with lifetime prevalence of up to 4% and 15% in the general population, respectively (Merikangas et al. 2011; Lépine & Briley 2011). As recurring conditions, the impact on public health represents a major concern for global disease burden in the context of disability adjusted life years (DALYs), morbidity and premature mortality (Palagini et al. 2018; Ferrari et al. 2013). Due to the chronic disease course, BD and depression are among the leading causes of DALYs and are responsible for more DALYs lost than all forms of cancer and other major neurological conditions (BD and MDD=79,619, cancer=76,716, major neurological conditions=20,823) (World Health Organization 2002; Merikangas et al. 2011; Ferrari et al. 2013).

1.2 Characteristics of BD and MDD

BD is characterised by recurrent episodes of mania (hypomania) and depression, as well as euthymic phases (McCarthy et al. 2018a; Harrison 2016).
BD is typically divided into subtypes: BD type 1 (BDI) is characterised by manic episodes and high rates of hospitalization (Tharp et al. 2016); BD type 2 (BDII) is characterised by less severe hypomanic episodes; and BD non-specified (BP-NOS) describes individuals with significant bipolar features which fall below the

threshold for BDI and BDII (O'Donovan et al. 2009).

Cognitive impairment is also a common feature of both BD and MDD, and has been shown to persist throughout mood states with patients displaying impaired processing speeds and memory (Cullen et al. 2015; Lépine & Briley 2011). BD patients can present with mixed mood states and chronic mood instability; these symptoms can also persist during periods of remission (Phillips & Kupfer 2013; Harrison et al. 2018; Bauer et al. 2018). Residual symptoms, including cognitive and social impairment, are also often reported during remission by individuals with MDD; these persistent symptoms could influence individuals quality of life (QOL) and increase the risk of relapse (Lépine & Briley 2011; Bauer et al. 2018).

Individuals with mood disorders usually have high comorbidity with other psychiatric and physical health conditions, with a life expectancy approximately 10 years lower than the general population (Jawinski et al. 2015). Bipolar and depressed individuals also have an estimated 20 to 30-fold increased risk of death by suicide (Jawinski et al. 2015; Dong et al. 2018).

An estimated 50% of BD patients, observed in clinical populations, also experience psychotic features which are associated with greater severity of symptoms and long-term morbidity (Neves et al. 2016). Overall, the substantial morbidity seen in BD is due primarily to recurrent depressive episodes (Harrison et al. 2016).

1.3 Diagnostic issues

Diagnosis of mood disorders is dependent upon the presentation of clinical symptoms and the interpretation of those symptoms by the clinician. There are currently no biomarkers which have been identified to aid in the diagnosis of mood disorders (Baryshnikov et al. 2015; Watmuff et al. 2016). The clinical presentations of both BD and MDD are heterogeneous, with BD often referred to as a spectrum disorder (Wray et al. 2018; Smith et al. 2008). BD is thought to be widely under-recognized, often only recognized after a long delay, and is commonly misdiagnosed, as borderline personality disorder or depression (Baryshnikov et al. 2015). The misdiagnosis of BD can have negative effects on the individual and result in the prescription of inappropriate treatments; this

could result in more frequent recurrence of mood episodes, more severe cognitive impairments and an increased risk of suicide (Tseng et al. 2015).

There is shared genetic architecture between psychiatric disorders which suggests that distinct clinical classifications may not be accurate (Cross-Disorder Group of the Psychiatric Genomics Consortium 2013; Forstner et al. 2017; Stahl et al. 2017). By relying solely on clinical classifications of disorders some of the underlying biology influencing disorder-related traits may not be identified (Phillips & Kupfer 2013).

Investigating transdiagnostic components may give a greater understanding of the pathophysiology of psychiatric conditions (Cuthbert & Insel 2013), for example Research Domain Criteria (RDoC) traits, personality traits and circadian measures; the latter (circadian features) will be the main focus of the analyses in this thesis.

1.4 Treatment of BD and MDD

Mood disorders tend to manifest initially during late adolescence or early adulthood; many individuals then experience a chronic illness course which requires lifetime treatment (Watmuff et al. 2016). Currently, the main therapy for both BD and depression is pharmacological (lithium, antipsychotics, antidepressants and anticonvulsants) alongside psychological interventions (Geddes & Miklowitz 2013; Kupfer et al. 2012; Yatham et al. 2018). Many mood disorder patients are continuously symptomatic and experience relapses (Keers et al. 2009; Dunn et al. 2015). Many of the pharmaceutical therapies used result in adverse drug reactions and, therefore, cause negative consequences to an individual's long-term physical health. Antipsychotics have also been reported to influence the cognitive impairment of patients which may lead to non-compliance and the eventual deterioration of the patient's mental health (Keers et al. 2009; Cullen et al. 2015).

This highlights the importance of the development of novel treatments and treatment approaches for BD and MDD. Unfortunately, there has been a lack of suitable therapeutics which have been successfully translated from animal models to patient use (Watmuff et al. 2016). New approaches to develop better therapeutic interventions for these disorders are required and are an important consideration in mood disorder research.

1.5 Current understanding of the genetics of BD and MDD

The underlying genetic architecture of psychiatric disorders is polygenic (to an extent this is explained by the large accumulation of small additive genetic effects) and complex (Sklar et al. 2011). There are several methods by which the genetics of psychiatric disorders have been investigated.

Family, twin and adoption studies have demonstrated the relatively high heritability of mood disorders. Genetic epidemiology estimates were calculated from observational studies of large family pedigrees with mood disorders, hospital and population registry data. Evidence from these studies have reported a high heritability for BD of approximately 70-89% (Craddock & Sklar 2013; Jawinski et al. 2015; Starnawska et al. 2016) and a heritability of 35-40% for MDD (Sullivan et al. 2000; Shih et al. 2004). Linkage studies typically used to identify causative variants and genes in family pedigrees have been unable to robustly identify potential risk variants in BD and MDD (Visscher et al. 2012).

Candidate gene studies have been undertaken in order to investigate the underlying aetiology of BD and MDD. These candidate gene studies are usually hypothesis-driven, based on suspected pathways and processes involved in the pathophysiology of BD and MDD. However, the results of these are usually based on small sample sizes and have not been replicated (Dunn et al. 2015). The overall value of candidate gene studies in this area is considered low.

Another example of how the genetics of psychiatric disorders are studied is the use of genome-wide association studies (GWAS) in large cohorts to identify relatively common genetic variants (minor allele frequency (MAF) 0.01-0.05), referred to as single nucleotide polymorphisms (SNPs), which associate with the phenotypes of interest. There have been several GWAS of BD and MDD which have identified SNPs associated with the disorders; in the case of BD, several variants associated with BD have been replicated in further GWAS (Hou et al. 2016; Stahl et al. 2017; Ferreira et al. 2008; Cross-Disorder Group of the Psychiatric Genomics Consortium 2013; Sklar et al. 2011; Wray et al. 2018).

One example of a replicated finding within BD GWAS is the *CACNA1C* gene; several polymorphisms within this gene have been associated with BD, and

with some evidence for association with MDD; this will be explored in more detail in Chapters 3 and 5.

GWAS has demonstrated that many common variants, each with small effects, influence risk of BD and MDD (Sullivan et al. 2018). Even with the many variants identified, GWAS have been unable to explain the high heritability estimates reported by family and twin studies of BD and MDD (Breuer et al. 2018). It has been suggested that 25% and 30% of phenotypic heritability of MDD and BD, respectively, is attributed to common SNPs; however, most GWAS of mood disorders report much smaller SNP heritability (h^2_{SNP}) (Wray et al. 2018; Bulik-Sullivan et al. 2015; Cross-Disorder Group of the Psychiatric Genomics Consortium et al. 2013; Stahl et al. 2017). It is thought that this 'missing heritability' is due to rare alleles (MAF<0.05), with small or intermediate effect sizes, and structural variation in the genome, such as deletions, insertions, inversions, translocations and copy number variations (repeats of cloned DNA fragments (Feuk et al. 2006)) which are often not included in GWAS (Harrison 2016).

Even though several variants have been identified by GWAS, the true causal variants or genes involved have not yet necessarily been identified (Starnawska et al. 2016; Harrison 2016). Due to linkage disequilibrium (LD) within the genome, it is difficult to reliably identify the causal variant within the loci highlighted by GWAS. It is often not known whether there are coding variants in LD with the risk SNPs which could result in the alteration of the gene product but it is likely that SNPs may influence gene expression (Starnawska et al. 2016).

As mentioned previously, there appears to be shared genetics amongst psychiatric and mood disorders; some cross-disorder studies have identified risk loci which overlap between BD and other major psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium 2013; Forstner et al. 2017; Stahl et al. 2017). There is some overlap of loci with MDD but there is greater overlap with schizophrenia, and it has been suggested that BD and schizophrenia are a continuous spectrum disorder (Harrison 2016). Some variants have also been found to be associated with features of BD, including psychosis and suicidality (Harrison 2016).

1.6 Researching the pathophysiology of BD and MDD

Although many risk variants associated with BD and MDD have been identified by GWAS there has been little progress in using these variants to aid our understanding of the underlying biological mechanisms of the disorder (Wray et al. 2007; Gratten et al. 2014). As BD is heterogeneous, the genetic variants linked to the disorder may influence multiple symptoms by affecting many different pathways (Jawinski et al. 2015). Using a combination of GWAS and gene expression data, there have been a range of pathways reported to have a potential association to BD, including calcium signalling, glutamatergic signalling, second messenger systems, hormone regulation, histone modification and immune pathways (Harrison 2016).

Attempting to understand the effects these genetic variants have on the underlying biological mechanisms of BD and MDD has proved challenging. It is difficult to model the disorder using animals and recent studies have also reported species-specific differences in synaptic biology between animal models and human; the responses seen in humans to specific pathophysiological processes often show a poor correlation with the response displayed by the animal models used (Watmuff et al. 2016).

One method which has been used to investigate the pathophysiological mechanisms of mood disorders is to study reprogrammed induced pluripotent stem cells (iPSCs) obtained from mood disorder patients and controls have been used to investigate differential gene expression and regulation (Madison et al. 2015). iPSCs are cells derived from an individual, transformed by growth factors and signal pathway modifiers to produce a pluripotent cell line which can be induced to differentiate into other cell types – for example, neurons (Soliman et al. 2017). Induced neurons from BD patients displayed significant differences in their neuronal transcriptomes including the upregulation of transcripts for ion channels, membrane-bound receptors and transcripts involved in neuronal cell differentiation compared with controls (Yoshimizu et al. 2015).

At present, the study of MDD patient-derived iPSCs is not as developed as studies using BD patient cells. With the current methods used to generate iPSCs it is not feasible to carry out large scale patient-derived iPSC studies due to the labour intensive and often variable nature of cell line transformation (Soliman et

al. 2017); the findings of these studies are based on small sample sizes (N=20) and require replication (Yoshimizu et al. 2015). However, the small scale of molecular studies based on reprogrammed cells means they are likely to be underpowered to detect effects when investigating single candidate risk variants identified by GWAS. In the future, iPSCs may be a useful tool to model risk variants (highlighted by GWAS or those included in genetic risk scores) and their potential influence on molecular mechanisms. Currently these influences are unclear but with greater information on the genetic architecture of mood disorders and with future advances in iPSC generation, patient-derived cells could become a useful tool to expand on genetic findings and give a further understanding the biology of mood disorders.

1.7 Circadian function in mood disorders

One area of primary interest to this thesis is the potential involvement of dysregulation of circadian rhythmicity in the pathophysiology of BD and MDD. Circadian rhythms are fundamental to homeostasis and are described as variations in physiology and behaviour which occur over an approximate 24 hour period (McClung 2007). These rhythms influence a range of biological and behavioural features, including mood, and are crucial to influences on physical and mental wellbeing (Reppert & Weaver 2001; Merikanto et al. 2017). Circadian rhythmicity is complex and involves the interaction of many different inputs, including core and peripheral genes (described in greater detail in chapter 6) and environmental stimuli. As will be described further in several chapters of this thesis, there is growing evidence, both epidemiological and genetic, supporting the involvement of circadian function in various aspects of mood disorders. For example, both subjective and objective measures of circadian rhythmicity (such as self-reported chronotype and accelerometer-measured activity) have been investigated in mood disorders and disrupted rhythmicity has been associated with increased risk of mood disorders (Burton et al. 2013; Geoffroy et al. 2014; Wulff et al. 2010). However, the majority of investigations into circadian rhythmicity and mood disorders have been based on relatively small clinical populations looking at phenotypic associations. There have been very few genetic studies looking at this relationship with those studies focussed on specific core circadian clock genes (Landgraf et al. 2014; Moon et al. 2016).

The complete genetic architecture of circadian function is unclear and how this influences mood disorder pathophysiology requires further investigation.

1.8 Key gaps in the literature

1. There is currently a lack of understanding of the genetic architecture of mood disorders and mood disorder-related traits. Until very recently the majority of studies looking at mood disorder risk variants have been based on small sample sizes and were often not replicated.

2. A relationship been circadian rhythmicity and mood disorders has been suggested but the genetic architecture of circadian rhythmicity is currently unclear. Also, the influence of circadian genetics on mood phenotypes has not been extensively investigated.

3. Within mood disorders, investigations have used mostly subjective circadian measures to investigate circadian dysregulation in patients.

4. Most studies focus specifically on small patient populations. There are relatively few studies investigating mood disorder genetics in large cohorts which may be more representative of the general population.

1.9 How this thesis will contribute to new knowledge

1. This study uses a relatively large amount of high-quality data to contribute to the current knowledge of mood disorder genetics.

2. It investigates the relationship between disrupted circadian rhythmicity and mood disorders at a genetic level.

3. It will contribute to an understanding of the genetic architecture of the circadian rest-activity cycle using large-scale objective data through GWAS.

4. It uses recent GWAS findings to apply new polygenic risk scores (PRS) to large scale data to investigate associations between circadian measures

(both subjective and objective) and mood disorder-related phenotypes within two separate cohorts (Outline 1.1).

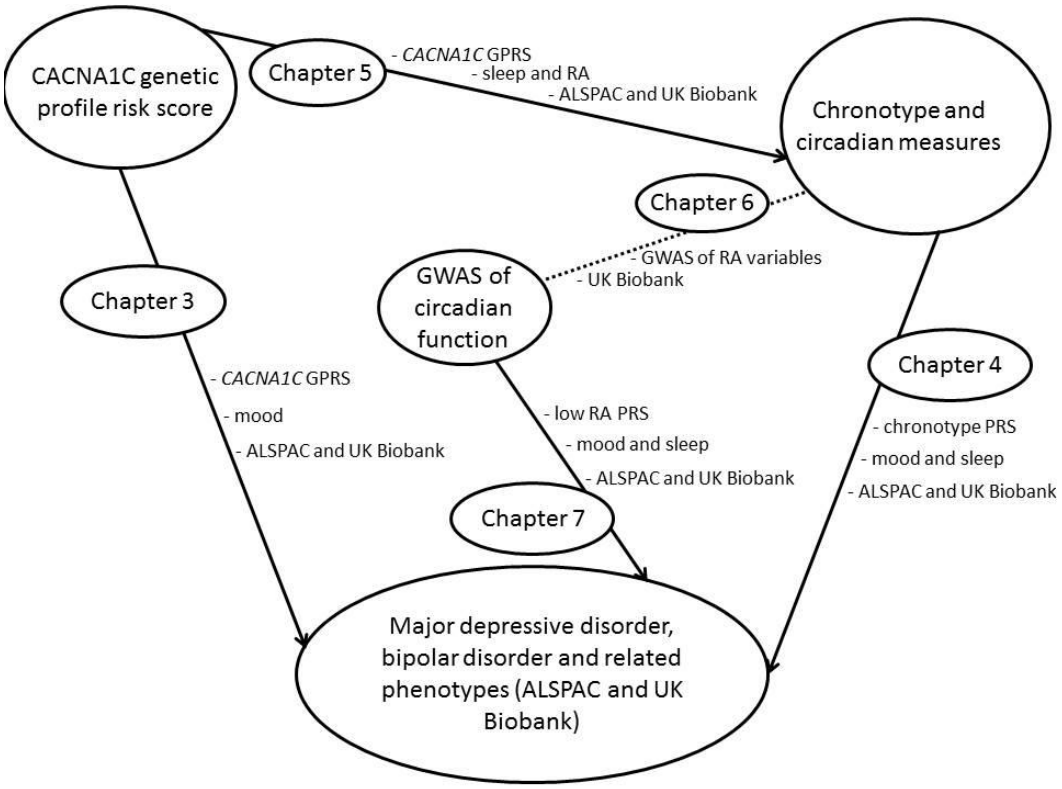
5. This thesis also highlights areas of interest which could inform future clinical considerations (i.e. features of an individual's phenotypic presentation which could be incorporated into a clinical assessment to inform a diagnosis) and potential treatment targets.

1.10 Importance of thesis

A greater understanding of the underlying mechanisms of mood disorders could aid in the identification of biomarkers or treatment targets (Neves et al. 2016). Of specific interest in this thesis is the relationship between circadian rest-activity rhythm function and mood disorders.

As will be described in Chapter 4 and 6, epidemiological evidence has demonstrated associations between disrupted circadian function and mood disorders. This thesis will investigate a possible link between circadian rhythmicity and mood phenotypes at a genetic level. To my knowledge, this is the first example of using PRS for circadian features in the investigation of mood disorder-related traits.

This evidence could inform future research and the development of targeted treatments for mood disorders based on an individual's genotype.



Outline 1.1 Outline of analyses in context of the overall study (Chapter 2 (Methodology) and Chapter 8 (Conclusions) are not included). Diagram highlights the analyses undertaken in each chapter, including which genetic risk score was used, the phenotypes-of-interest and the cohorts investigated. Dotted line represents analyses using the UK Biobank cohort only.

Chapter 2 Dataset description and research methodology

2.1 Avon Longitudinal Study of Parents And Children (ALSPAC)

2.1.1 Participants and ethical approval

ALSPAC is a UK birth cohort recruited from the Avon area of England. Pregnant women with expected delivery dates between April 1991 and December 1992 were recruited to the cohort. All participants provided informed written consent. Data from mothers, partners and children has been collected periodically from September 1990; data includes mother-completed and child-completed questionnaires, interviews, environmental measures, mother and child biological samples, and genetic data.(Golding et al. 2001) The full details of the available data can be found at http://www.bristol.ac.uk/alspac/researchers/access/.

2.1.2 Genotyping and imputation

The Illumina HumanHap550 quad genome-wide SNP genotyping platform was used to genotype DNA samples obtained from 9,912 participants at age 7 (approximately 70% of the sample (Jones et al. 2000)). Participants of ALSPAC who were genotyped were only included in the ALSPAC genotype database after meeting particular quality control criteria. Those individuals found to have >3% individual missingness, evidence of cryptic relatedness (>10% alleles identical by descent) insufficient sample replication and extremes of heterozygosity were excluded from the data. Individuals with gender mismatches and of non-European ancestry were also excluded. Related individuals who passed these quality controls were retained for phasing and imputation, whereupon, further participants were removed due to SNP ID mismatching and violation of Hardy-Weinberg equilibrium (HWE)
(p value < $5x10^{-7}$). Imputation was performed using Impute V2.2.2 and the 1000 genomes reference panel (Phase 1, Version 3) with 2,186 reference haplotypes; SNPs with a quality metric of <0.8, <95% call rate and MAF of <1% were excluded. The genetic data of 8,230 participants with 500,527 SNPs were available (Jones et al. 2016a). Only unrelated individuals were included in analyses in an attempt to prevent shared environmental factors influencing associations (Dudbridge 2016); a list of unrelated individuals (N=8,197; 96% of

the genotyped sample) was provided by ALSPAC. Individuals that were not recorded as Caucasian, or those with missing ethnic background data (N=1,112; 13.08%), were also removed from analyses leaving N=7,390.

2.1.3 Mood disorder phenotypes

2.1.3.1 Hypomania

To assess probable bipolarity in ALSPAC participants, answers given to the 32 item Hypomania checklist-32 (HCL-32) questionnaire were converted into a categorical measure of hypomania. As per the recommendation of Court et al. 2014, based on a Rasch analysis of unidimensionality, only 28 of the 32 items were included to produce Hypomania checklist-28 (HCL-28) and provide a HCL-28 hypomania score (0-28) (Court et al. 2014). Individuals were assessed age 22-23 and those who did not respond, or those with no data, for these 28 items (N=5,700; 69.54% of the whole sample) were coded as missing and were not included in any further analyses. The scores out of 28 were also combined with additional information regarding the duration of "high" states and how often individuals had experienced these "high" states, to produce a categorical measure of hypomanic features, as follows: individuals with a HCL-28 score of greater than 14; a duration of "high" states of "2-3 days" or longer and either "negative consequences" or "negative plus positive consequences" as a response to these "high" states, were designated as "hypomania"(Hayes et al. 2016). Individuals who did not meet these criteria were designated as "no hypomania". Note that this categorical definition of hypomania was the primary outcome measure; the HCL-28 score was then used as a continuous outcome for secondary analyses to further investigate hypomania.

2.1.3.2 Depressive features

Depressive symptoms were assessed in the ALSPAC cohort during late childhood and adolescence using the Short Mood and Feelings Questionnaire (SMFQ) between the ages of 10 and 19. The SMFQ is a 13-item self-reported questionnaire assessing depressive symptoms over a 2 week period (Wiles et al. 2012; Stringaris et al. 2014). Each item is scored 0, 1 or 2 depending on the whether the participant answered "not true", "sometimes true" or "true", respectively. This generated a categorical depression measure (binary

depression) based on a score of greater than 16, as well as an SMFQ score ranging from 0 to 26 (Wiles et al. 2012). As for hypomania, the categorical SMFQ measure was used as the primary outcome measure for depression and the SMFQ score was a secondary outcome.

2.1.4 Sleep phenotypes

Four measures of sleep problems, reported at two different ages, were used for analyses. Mothers responded to the following questions when their child was aged 10: "In the past month child found sleep hard" (kv7034) (referred to as Difficult sleeping-10); "In past month child slept too much" (kv7035) (Too much sleep), both questions were answered with "yes", "no" or "don't know". At child age 13, mothers were again asked "Child had difficulty getting to sleep in past month" (tb7034) (Difficult sleeping-13) answering either "yes", "no" or "don't know". Also at age 13, mothers were asked: "Degree to which child had problems sleeping during last month" (tb5538) (Difficult sleeping-scale) with responses of "not at all"=1, "a little"=2, "yes"=3 and "don't know"=9, and "Frequency worrying interferes with child's sleep" (tb6555) (Worried sleep) answering "not at all"=1, "yes not most days"=2, "yes most days"=3 and "don't know"=9. Those who answered "don't know" were coded as missing and were not included in analyses.

The primary outcome measures for sleep in ALSPAC were the categorical variables "Difficulty sleeping-10" and "Difficulty sleeping-13", which were tested for association with PRS using logistic regression. A secondary outcome measure "Too much sleep" was also tested using logistic regression, the remaining secondary measures ("Difficulty sleeping-scale" and "Worried sleep") were tested using linear regression.

2.2 UK Biobank

2.2.1 Participants and ethical approval

Over 502,000 UK residents aged 37-73 years (most aged 40-70) were recruited to the UK Biobank cohort from 2006-2010. At one of 22 assessment centres across the UK, participants completed a range of lifestyle, demographic, health, mood, cognitive and physical assessments and questionnaires, with DNA samples taken

at baseline assessment (Sudlow et al. 2015). UK Biobank obtained informed consent from all participants and this study was conducted under generic approval from the NHS National Research Ethics Service (approval letter dated 13 May 2016, Ref 16/NW/0274) and under UK Biobank approvals for applications 12761 (PI Cathy Wyse; accelerometer data for use in GWAS) and 6553 (PI Daniel Smith; genetic and phenotypic data).

2.2.2 Genotyping and imputation

UK Biobank released genotypic data for over 500,000 participants using two genotyping arrays specifically designed for UK Biobank with 95% shared marker content (Bycroft et al. 2017). Approximately 10% of these participants were genotyped using Applied Biosystems UK BiLEVE Axiom array by Affymetrix, with the remaining participants being genotyped using Applied Biosystems UK Biobank Axiom Array. Phasing on the autosomes was done using SHAPEIT3 using the 1000 Genomes Phase 3 dataset as a reference panel. Imputation of SNP genotypes was carried out using IMPUTE4; the merged UK10K and 1000 Genomes Phase 3 reference panel, as used for the UK Biobank interim genotype data release. Approximately 850,000 SNPs were directly genotyped with more than 90 million SNPs available after imputation. Stringent quality control was applied to the data, described in an open access document (Bycroft et al. 2017).

2.2.3 Self-reported Bipolar Disorder and Depression

All UK Biobank participants were given the opportunity to provide a self-report of bipolar and depression status (Data-Field 20002). These outcomes are referred to as SR Bipolar Disorder and SR Depression/Recurrent Depression.

2.2.4 Probable Bipolar Disorder and Major Depressive Disorder

Measures of probable BD and MDD were generated for 123,000 UK Biobank participants using questions based on Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), items from the Patient Health Questionnaire (PHQ) and items on help-seeking for mental health. This was obtained during the final two years of recruitment (Smith et al. 2013a). These questions allowed for the distinction of single episode and recurrent depression. Those with single episode of depression were removed from further analyses. From this data, categorical

measures of probable BD (probable BD) and probable depression (probable recurrent depression) were generated, and individuals who did not meet either criteria were classified as controls.

2.2.5 Psychiatric diagnoses, neuroticism and mood instability

2.2.5.1 Bipolar Disorder, Major Depressive Disorder and Generalised Anxiety Disorder

A mental health questionnaire (MHQ) was developed by a UK Biobank mental health research reference group to collect additional mental health phenotype data and was administered during 2016-2017 (Davis et al. 2018). The MHQ was used to obtain information about individuals' lifetime experiences of psychiatric disorders, as well as other risk factors for these disorders, such as anxiety, substance abuse and childhood trauma. The composite questionnaire consisted of 10 sections and was based on a modified Composite International Diagnostic Interview Short Form (CIDI-SF), PHQ-9, Generalised Anxiety Disorder Questionnaire (GAD-7) and questions devised by the mental health research reference group. Lifetime depression (referred to here as 'lifetime MDD'), 'lifetime BD' and lifetime generalised anxiety disorder (referred to as 'lifetime GAD') were evaluated based on answers provided by participants to the online MHQ. Therefore, as with the depression and bipolar disorder phenotypes described above, these assessments represent a likelihood of diagnosis, rather than a confirmed diagnoses (Davis et al. 2018). Individuals who had self-reported BD or MDD were excluded from the control groups. This resulted in variables generated for 157,366 UK Biobank participants.

2.2.5.2 Neuroticism

To define neuroticism a score was taken from the 12-item neuroticism scale of the Eysenck Personality Questionnaire-Revised Short Form (EPQ-R-S) (Eysenck et al. 1985; Smith et al. 2016). Individuals were given a score of 0 or 1 for a "no or yes" answer to each item, with total score from 0 to 12. Higher neuroticism has been associated with higher incidences of psychiatric disorders, therefore, it is linked to greater socioeconomic cost and premature mortality (Lahey 2009; Smith et al. 2016).

2.2.5.3 Mood instability

A “mood instability” phenotype was also obtained from the EPQ-R-S questionnaire. One of these questions was *“Does your mood often go up and down?”* (answer options “yes”, “no”, “don’t know” or “prefer not to answer”) (Eysenck et al. 1985). Individuals who selected “don’t know” or “prefer not to answer” were coded as missing (very few participants); this allowed the generation of a categorical mood instability variable where those who answered “yes” were designated as cases and participants who answered “no” were controls, those answering “don’t know” and “prefer not to answer” were excluded (Ward et al. 2017).

2.2.6 Chronotype phenotype

Chronotype was derived from the participants’ responses to a question from the UK Biobank Morningness-Eveningness questionnaire (Taillard et al. 2003). The question consisted of “Do you consider yourself to be...” with response options “Definitely a “morning” person”, “More a “morning” than “evening” person”, “More an “evening” than “morning” person”, “Definitely an “evening” person”, “Do not know” and “Prefer not to answer”. This chronotype assessment is a widely accepted measure and has previously been reported to explain the greatest variance in individual preference of sleep-wake timings (Taillard et al. 2003). Categorical variables were then generated based upon the responses given, resulting in the generation of four separate chronotype variables (“Definite morning”, “Definite evening”, “Overall morning”, “Overall evening”). The primary outcome measures used for analysis were the “definite morning” and “definite evening” variables. Secondary outcome measures of “overall morning” and “overall evening” were also assessed. Individuals who answered either “don’t know” or “prefer not to answer” were coded as missing for each of the variables and excluded (22.8%) (Figure 2.1).

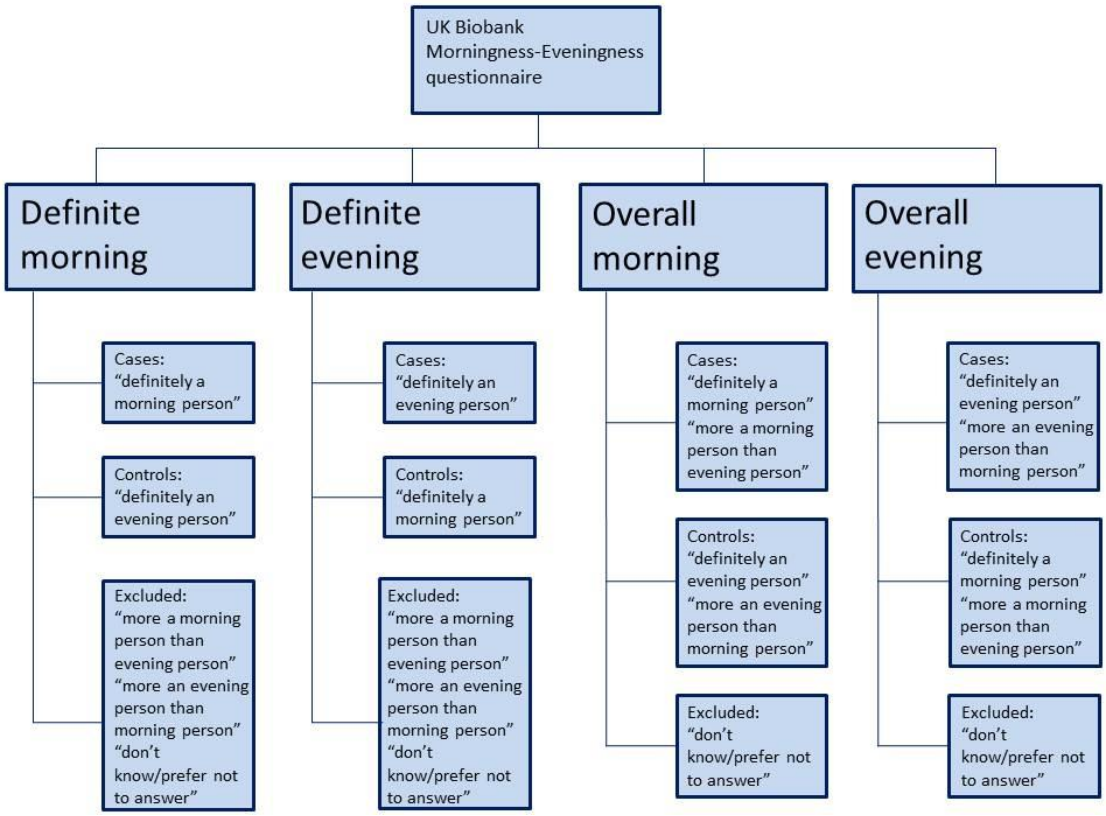


Figure 2.1 Case/control designation from SR Chronotype in UK Biobank

2.2.7 Accelerometry data collection and pre-processing

In 2013, 240,000 UK Biobank participants were invited to wear an accelerometer for seven days as part of a physical activity monitoring investigation (Doherty et al. 2017). Of these, 103,720 (43%) accepted and returned the accelerometer to UK Biobank after use. Participants received a wrist-worn Axivity AX3 triaxial accelerometer in the post and were asked to wear the device on their dominant wrist continuously for seven days, while continuing with their normal activities. At the end of the seven-day period, participants were instructed to return the accelerometer to UK Biobank using a prepaid envelope. Accelerometers were calibrated to local gravity. Devices recorded data at a sampling rate of 94-104Hz, and data were resampled to 100Hz offline. Periods where no data were recorded for >1s were coded as missing, and machine noise was removed using a Butterworth low-pass filter (cut-off 20Hz). Raw activity intensity data were combined into five second epochs. Further details on data pre-processing are available from UK Biobank at http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=131600 (Doherty et al. 2017).

2.2.8 Circadian rest-activity rhythmicity (RA)

From the summary five second epoch data, a measure of relative amplitude (RA) was calculated using Clocklab Version 6 (Actimetrics) by Dr. Cathy Wyse (Wyse et al. 2018, unpublished). This accelerometer-derived activity measure has demonstrated reliability and validity in associating with health measures (Sadeh 2011; Lyall et al. 2018). RA is used commonly as a non-parametric measure of rest-activity rhythm amplitude. It is defined as the relative difference between the most active continuous 10-hour period (M10) and the least active continuous 5-hour period (L5) in an average 24-hour period (Van Someren et al. 1996):

$$RA = \frac{(\mathrm{M10} - \mathrm{L5})}{(\mathrm{M10} + \mathrm{L5})}$$

M10 is the mean activity during the continuous 10 hour period containing maximum activity in each 24 hour recording period (midnight to midnight). L5 is the mean activity for the corresponding 5 hour period containing the minimum activity within the same recording period. For each individual, the RA data point was the mean RA value across all included 24-hour periods (seven days). RA ranges from 0 to 1, with higher values indicating greater distinction between activity levels during the most and least active periods of the day.

Participants who provided accelerometer data for less than 72 hours or who did not provide data for all one-hour periods within the 24-hour cycle were excluded from analyses. Over 10,000 participants were also excluded because their data was identified by UK Biobank as having poor calibration, poor wear compliance, or flagged by UK Biobank as unreliable (unexpectedly small or large size) and where participants whose wear-time overlapped with a daylight savings clock change (Lyall et al. 2018).

2.3 Polygenic risk scores

Weighted PRS were generated for all individuals with genotype data available in ALSPAC and UK Biobank using Plink 1.9 (Purcell et al. 2007).
Only SNPs with a MAF of greater than 0.01 – here the minor allele is designated A1 – within the cohorts were considered for inclusion in the PRS calculation (generally most SNPs included). All SNPs were tested for HWE in both cohorts; SNPs with HWE p value of less than 0.001 were excluded from further analysis.

All statistical analyses (i.e. logistic and linear regressions) were carried out using the statistical software package Stata Student Edition 13 for Windows; the standard nominal p value significance cut-off value of $p<0.05$ was used for all analyses. Changes in the weighted PRS scores are small due to the use of the log of O.R; to provide a more comprehensive interpretation of the analyses, the weighted scores were standardised to their z values.

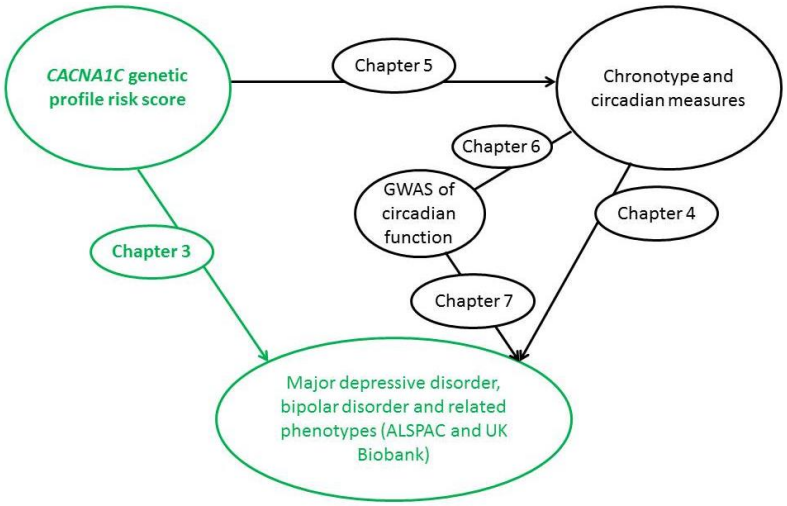
Most of the association analyses used logistic regression assuming an additive allelic (0/1/2) effect model. For ALSPAC, potential confounders, sex and socioeconomic status were included in the models. For UK Biobank, the adjusted analysis included age, sex, socioeconomic status (measured using the Townsend deprivation score), 8 UK Biobank genetic principal components, assessment centre, genotyping chip and batch as covariates.

Methodology specific to each analyses is detailed in the appropriate chapters.

2.3.1 Power calculation

As the effect sizes reported for polygenic risk scores are often relatively small a large sample size was required for each analysis (Dudbridge 2013). To determine the sample sizes required to detect the effects of genetic risk scores on the various phenotypes in ALSPAC and UK Biobank to significance post-hoc power calculations were undertaken using GPower software (Version 3.1). The estimated effect size and desired power (95% power) were inputted and the required minimum sample size was calculated based on a logistical regression model (Faul et al. 2009; Faul et al. 2007). As described in the limitations of each analysis, some investigations were underpowered to detect effects to significance due to the relatively small number of cases available in the categorical mood-related phenotypes.

Chapter 3 Investigating the effect of Calcium Voltage-Gated Channel Subunit Alpha 1 C (*CACNA1C*) single nucleotide polymorphisms on mood disorder phenotypes within two population cohorts



Outline 3.1 Chapter in context of overall study

3.1 Introduction

This chapter will investigate the potential associations of *CACNA1C* SNPs with mood disorder-related traits in ALSPAC and UK Biobank using a GPRS approach.

In GWAS, several SNPs within the Calcium Voltage-Gated Channel Subunit Alpha 1 C (*CACNA1C*) gene have been associated with neuropsychiatric disorders, including BD, schizophrenia and MDD (Bigos et al. 2010; Sklar et al. 2011; Smoller 2013; Heilbronner et al. 2015; Stahl et al. 2017). This association appears to be strongest for BD, the genome-wide significant allelic associations localize within intron 3 of *CACNA1C* ($p<2x10^{-8}$) (Fiorentino et al. 2014).

Most genetic variants associated with common neuropsychiatric disorders are relatively common (typically MAF>5%) in the general population (Bigos et al. 2010). For example, the *CACNA1C* minor allele rs1006737 SNP – which is a replicated finding and has been of considerable interest in previous studies (Green et al. 2013; Ferreira et al. 2008; Bigos et al. 2010) – is found in approximately one third of individuals of European ancestry and is a common variant in other populations (see Table 3.2) (NCBI 2018). These genetic variants may also highlight pathways which could lead to the development of more targeted pharmaceutical therapies (Dunn et al. 2015; Harrison et al. 2018).

3.1.1 The *CACNA1C* gene

The *CACNA1C* gene encodes the alpha subunit of the L-type voltage-gated calcium channel (L-VGCC) $Ca_v1.2$ (Erk et al. 2014a). The *CACNA1C* locus is

located on the short arm of chromosome 12, spanning a genomic region of approximately 740kb (Soldatov 1994). The gene contains approximately 55 exons and several splice variants have been detected, however the exact range of transcripts and protein isoforms produced is unknown (Soldatov 1994; Harrison 2016). The majority of risk variants associated with psychiatric disorders, including BD, identified through GWAS are found within introns and 5' and 3' untranslated regions (Fiorentino et al. 2014; Kabir et al. 2017). A particular area of interest is *CACNA1C* intron 3; this intronic region is highly conserved in mammals, suggesting a fundamental functional importance for the region (Fiorentino et al. 2014). Variation in this region may have an effect on regulatory elements within the large intronic region with several possible outcomes, including altered expression of *CACNA1C* transcripts, altered expression of *CACNA1C* splice variants and differential expression during development (Yoshimizu et al. 2015).

3.1.2 Potential pathways of pathophysiology

Functional studies have suggested that risk SNPs lie within regions of tight transcriptional control: gene expression may become altered through the differential binding of regulatory nuclear proteins or by altering interactions between intronic enhancers and promoters (Kabir et al. 2017). One study, undertaken in BD patients, suggests that risk variants affect the genome architecture and, therefore, influences interactions with transcription start sites and altering gene expression (Starnawska et al. 2016). Several studies have reported expression changes of the *CACNA1C* gene associated with risk SNPs in various tissues using both post-mortem samples and induced neurons derived from BD patients (Harrison 2016).

DNA methylation is also a potential mechanism by which non-coding variants can result in the phenotype variability seen in BD. Altered methylation is linked to abnormal gene expression, differential splicing and the use of alternate gene promoters. Overall, hypermethylation has been reported in BD patients (N=582) and the rs1006737 *CACNA1C* risk SNP has been found to be associated with the hypermethylation of CpG island 3 (found within intron 3) of the *CACNA1C* gene in BD subjects compared to controls (Starnawska et al. 2016). It is therefore possible that the *CACNA1C* risk SNPs influence the underlying mechanisms of BD via altered DNA methylation patterns (Starnawska et al.

2016).

It is also possible that non-coding risk variants influence the pathophysiology of BD by affecting non-coding RNAs or antisense transcripts influencing other regulatory elements or genes (Harrison 2016).

3.1.3 The $Ca_v1.2$ channel

The protein encoded by *CACNA1C*, the alpha subunit of L-VGCC $Ca_v1.2$, is the only component of the channel which forms the transmembrane pore vital for allowing action potentials on the cell membrane to be converted to calcium influx (Erk et al. 2014a; Catterall 2011). This transduction of electric excitability is critical to the physiological roles of many organs (Liao & Soong 2010). $Ca_v1.2$ channels are expressed by a variety of cell types, including neurons, lymphocytes, smooth muscle cells, pancreatic beta cells, and cardiomyocytes (Moosmang et al. 2005; Cabral et al. 2010; Liao & Soong 2010; Bidaud & Lory 2011), and are involved in many processes, summarised in Figure 3.1, such as excitation-contraction coupling, hormone regulation and secretion, integration of synaptic input, cell survival and neurotransmitter release (Tabuchi et al. 2000; Gomez-Ospina et al. 2006; Kolarow et al. 2007; Liao & Soong 2010; Bidaud & Lory 2011; Catterall 2011; Erk et al. 2014b; VT et al. 2015).

The influx of calcium ions via $Ca_v1.2$ channels activates several pathways within neurons (Soeiro-De-Souza et al. 2017) including the subsequent activation of a series of calmodulin-dependent protein kinases leading to the transduction of molecular cascades and gene expression, via cAMP response element binding proteins (CREB) (Kabir et al. 2017). CREB-activated genes are critically involved in synaptic and neuronal plasticity (Kabir et al. 2017). It has been reported that *CACNA1C* risk SNPs are associated with higher intracellular calcium ion concentrations in neurons of BD individuals compared to healthy controls; these risk SNPs may therefore influence BD through greater activity of calcium-dependent cascades. It is of note this study was undertaken using a small case sample (N=50) (Soeiro-De-Souza et al. 2017).

3.1.4 Possible influence of *CACNA1C* in the brain

Several studies have reported that carriers of *CACNA1C* risk variants show altered activity and connectivity in various regions of the brain, and some

evidence of structural differences in areas such as the hippocampus, perigenual anterior cingulate (pgACC) and amygdala (Erk et al. 2010; Tesli et al. 2013; Erk et al. 2014b; Jakobsson et al. 2015; Kamali et al. 2015). However, as mentioned previously, these studies were based on relatively small sample sizes and are likely to be underpowered. *CACNA1C* may influence brain structure via its roles in neuronal plasticity and dendritic retraction (Soeiro-De-Souza et al. 2017).

Evidence of *CACNA1C* involvement in brain activity has been demonstrated in both BD patients and unaffected first degree relatives who carry the risk variants. Examples of this include reduced activity in the pgACC during stress and mood-related responses (Erk et al. 2014b), altered activity during emotional processing (Heyes et al. 2015), reduced activity in the hippocampus, and dysfunction in the amygdala (Tesli et al. 2013; Erk et al. 2014b) (see Figure 3.1). The *CACNA1C* variant rs1006737 has been previously associated with increased anxiety and depression scores in individuals carrying the SNP relative to those who do not; this has been demonstrated in both individuals with BD and unaffected first degree relatives (heterozygous risk variant carriers: N=119, mean 0.16 vs homozygous wild-type: N=141, mean 0.09, p 0.02) (Erk et al. 2014a).

3.1.5 Potential as a therapeutic target for mood disorders

L-VGCCs are known to be highly sensitive to calcium channel antagonists (Bidaud & Lory 2011). The alpha subunit of $Ca_v1.2$, encoded by *CACNA1C*, is the main binding site for antagonists such as dihydropyridines (DHPs), benzothiazepines and phenylalkylamines (Bidaud & Lory 2011). Mice exposed to L-VGCC agonists displayed severe neurobehavioral symptoms and prolonged depressive symptoms; these symptoms were reversed after treatment with a DHP drug (Keers et al. 2009). The calcium channel blockers verapamil and nimodipine may have efficacy in the mood stabilisation of BD, making $Ca_v1.2$ a possible target for therapeutic treatment (Keers et al. 2009; Bidaud & Lory 2011; Erk et al. 2014b).

Overall, it is reasonable to conclude that common variants at the *CACNA1C* locus may have an important role in the pathophysiology of mood disorders, particularly BD, and that *CACNA1C* potential as a therapeutic target.

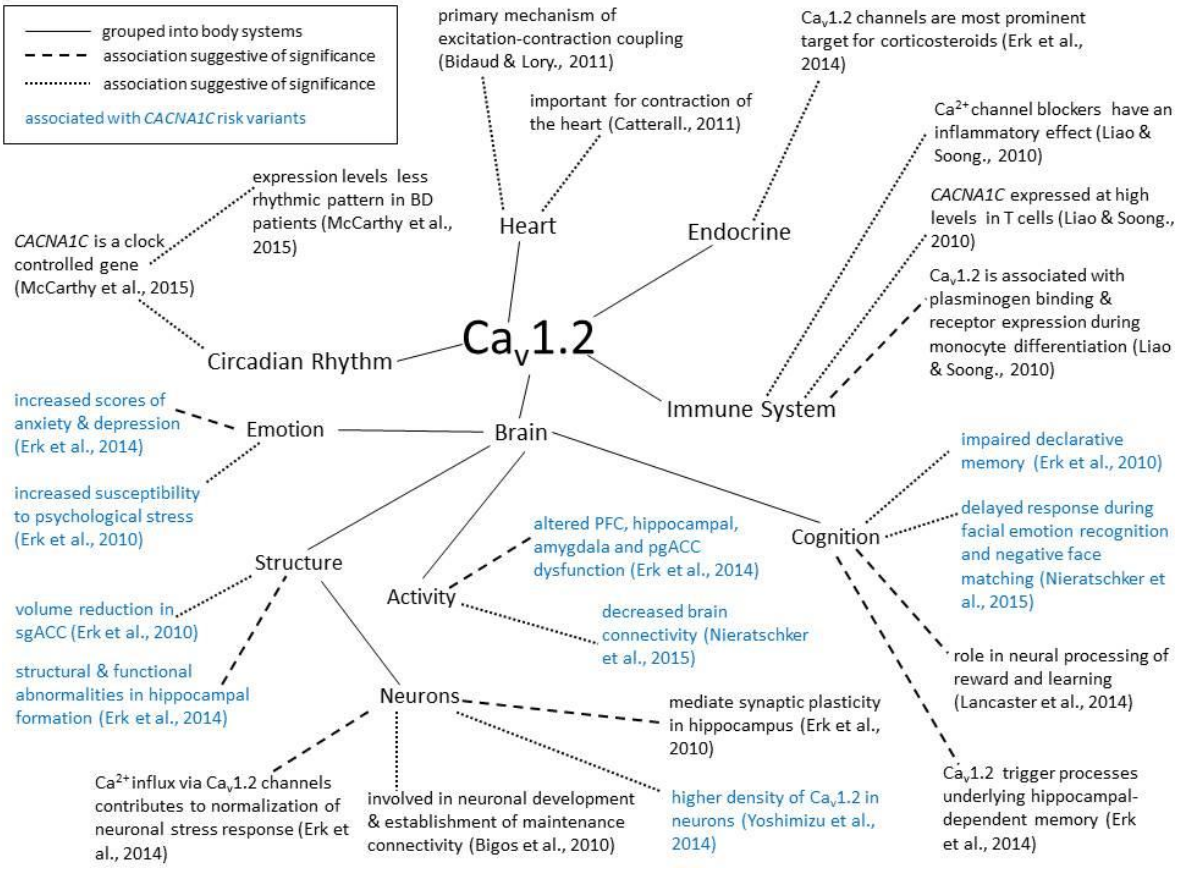


Figure 3.1 Various $Ca_v1.2$ channel functions and potential pathophysiological influences

3.2 Rationale

Previous investigations into the effects of *CACNA1C* SNPs have focussed only on single variants in relatively small samples and are likely to be underpowered (Bigos et al. 2010; Heilbronner et al. 2015; Dao et al. 2010; Strohmaier et al. 2013). Generating a weighted score using several risk SNPs may allow for the detection of small effects within our cohorts. As this risk score is calculated using SNPs from only one gene locus this risk score is referred to as the genetic profile risk score (GPRS) for the *CACNA1C* gene, as opposed to a PRS which uses variants from many gene loci. The preliminary analyses make use of very large samples from the first release of genetic data from UK Biobank sample (N=152,000), and all available genetic data in ALSPAC (N=8,365) and, therefore, are likely to be sufficiently powered to find low effect sizes (greater than 90% power to detect low effects).

Previous studies investigating the influence of *CACNA1C* in the context of psychiatric disorders have suggested that risk alleles may have a greater effect on one sex compared to the other (Dao et al. 2010; Strohmaier et al. 2013;

Heilbronner et al. 2015; Starnawska et al. 2016). The analyses were also undertaken upon separating the samples by sex.

The subsequent release of genetic data for the remaining UK Biobank sample and further phenotype data allows the UK Biobank to act as a replication cohort. These analyses were carried out to test the robustness of the findings from the primary analyses.

3.3 Hypotheses to be tested

In this chapter, the primary hypothesis being tested was: is there an association between greater genetic loading of *CACNA1C* variants and the clinical expression of several mood disorder phenotypes within two population cohorts (ALSPAC and UK Biobank)? These phenotypes included: hypomania and depressive features within ALSPAC; and mood instability, neuroticism, BD status and MDD status within the UK Biobank cohort.

3.4 Methods

3.4.1 GPRS analysis

A weighted GPRS was generated for all individuals with genotype data available in ALSPAC and UK Biobank using Plink 1.9 (Purcell et al. 2009). To provide weighting for the *CACNA1C* SNPs, the log of the odds ratios provided by BD GWAS and meta-analyses were used (Kloiber et al. 2012; Green et al. 2013; Fiorentino et al. 2014; Heilbronner et al. 2015). Only SNPs with a MAF of greater than 0.01 – here the minor allele is designated A1 – within the cohorts were considered for inclusion in the GPRS calculation. All SNPs were tested in both cohorts for LD and HWE; SNPs with a HWE p value of less than 0.001 were excluded from further analysis (see Table 3.2).

For the analyses using the ALSPAC cohort only, individuals genotyped at all 19 SNPs-of-interest were included. In the case of UK Biobank, analyses were run using individuals genotyped for all 15 SNPs found to be in HWE.

Changes in the weighted GPR scores are small due to the use of the log of O.R; to provide a more comprehensive interpretation of the analyses, the weighted scores were standardised to z values (i.e. per standard deviation).

All O.R and coefficients reported are per SD of GPRS.

3.4.2 ALPSAC

Details of recruitment to the cohort and genotyping are described in Chapter 2. Only unrelated individuals were included in these analyses to prevent shared environmental factors influencing associations (Dudbridge 2016); a list of unrelated individuals (N=8,197; 98% of the whole sample) was provided by ALSPAC. Individuals not recorded as Caucasian, or those with missing ethnic background data (N=973; 11.87%), were also removed from analyses leaving N=7,224.

3.4.2.1 Hypomania

To test associations between *CACNA1C* GPRS and features of BD in ALSPAC, categorical and continuous measures of hypomania were used. The full details of how these measures are generated using HCL-28 can be found in Chapter 2. HCL-28 score is used as the continuous measure of hypomania and represents an

individual's score out of 28 using the checklist. The categorical measure of hypomania (binary hypomania) combines the HCL-28 score with additional information regarding the duration of "high" states and how often individuals had experienced these "high" states. The categorical definition of hypomania was the primary outcome measure and the HCL-28 score was a secondary outcome.

3.4.2.2 Depressive features

Depressive symptoms were assessed in the ALSPAC cohort during late childhood and adolescence using the SMFQ. This generates an SMFQ score ranging from 0 to 26 (Wiles et al. 2012), as well as a categorical depression measure (binary depression) based on a score of greater than 16. The categorical SMFQ measure was used as the primary outcome measure and the SMFQ score was the secondary outcome.

These primary outcome measures of hypomania and depression were tested for association with *CACNA1C* GPRS using logistic regression assuming an additive allelic effect model. The dimensional secondary outcome measures (HCL-28 score and SMFQ score) were analysed using linear regression.

3.4.3 UK Biobank cohort

Individuals were filtered from the initial cohort of N=152,000 based upon several quality control criteria, including relatedness, ancestry (Non-Caucasian individuals), gender mismatch and quality control failure in the UK BiLEVE study. This left N=119,953 (78.9% of the cohort). Individuals missing genotype information for any of the 15 chosen SNPs were also excluded from analyses, leaving N=95,073.

3.4.3.1 BD and MDD

Individuals with BD and recurrent MDD were identified according to the criteria previously used by Smith and colleagues (Smith et al. 2013a) and as described within Chapter 2. These data were only available on approximately one third of the UK Biobank cohort because specific questions on manic features were only introduced towards the end of the recruitment period. BD and MDD defined in

this way were the primary outcomes of interest.
However, at initial recruitment all UK Biobank participants were given the opportunity to provide a self-report of bipolar and depression status. These were used as secondary outcomes in the current analyses and are referred to as SR Bipolar Disorder and SR Recurrent Depression.

3.4.3.2 Neuroticism

To define neuroticism a score taken from the 12 item neuroticism scale of the Eysenck Personality Questionnaire-Revised Short Form (EPQ-R-S) (Eysenck et al. 1985). Individuals were given a score of 0 or 1 for a "no or yes" answer, respectively, for each item and given a total neuroticism score ranging from 0 to 12. As described in Chapter 2.

3.4.3.3 Mood instability

As described in Chapter 2, a "mood instability" outcome measure was also obtained from the EPQ-R-S questionnaire: participants were asked *"Does your mood often go up and down?"* and given the option to answer "yes", "no", "don't know" or "prefer not to answer". Individuals who selected "don't know" or "prefer not to answer" were coded as missing (<5%); this allowed the generation of a categorical mood instability variable where those who answered "yes" were designated as cases and participants who answered "no" were included as controls.

3.4.3.4 Mental Health Questionnaire phenotypes

A MHQ was developed by a UK Biobank mental health research reference group to collect additional mental health phenotype data and was administered during 2016-2017. Lifetime BD and lifetime MDD variables were generated for 157,366 UK Biobank participants. Further details for these variables have been described in Chapter 2.

Most of the association analysis carried out used logistic regression assuming an additive allelic effect model (the association analysis using neuroticism score used a linear regression). The adjusted analysis included age, sex, socioeconomic status (assessed using the Townsend deprivation score), 10 UK

Biobank genetic principal components, assessment centre, batch and array as covariates. The association analyses were also performed separately for males and females, both unadjusted and adjusted, with age, deprivation index, 8 UK Biobank genetic principal components, assessment centre, batch and array included as covariates.

3.5 Results

Table 3.1 Demographics of cohorts

	ALSPAC N total = 8,197	**UK Biobank** N total = 119,953
Sex		
Female, N (%)	3,525 (48.8)	63,088 (52.6)
Age		
Mean (SD)	24.497 (0.5)	56.867 (7.93)
Deprivation		
Mean (SD)	2.831 (1.33)	-1.466 (2.99)
Range	3	-2.278

As the two population cohorts capture participants of different age groups, and parents were excluded from ALSPAC analysis (the ALSPAC table above refers to child participants only), it is unlikely that there is any overlap between the samples (Table 3.1). It is of note, in ALSPAC, DNA was obtained at age 7 with the hypomanic and depressive data used below collected between ages 22-23 and 10-19, respectively. UK Biobank DNA samples and self-report mood phenotypes were obtained at baseline assessment. The deprivation data in ALSPAC was obtained from the maternal socioeconomic status and Townsend score was used in UK Biobank.

As described in the methodology above, many *CACNA1C* SNPs-of-interest (listed in Table 3.2) were selected to generate GPR scores for genotyped individuals in ALSPAC and UK Biobank (LD structure of SNPs displayed in Supplemental Figure 3.1).

Table 3.2 SNP-of-interest information (ALSPAC/UK Biobank)

SNP ID	Position	A1	A2	MAF	O.R (for BD)	HWE p value
rs2007044	2344960	G	**A**	0.393/0.395	1.198	0.371/0.002
rs1006737	2345295	**A**	G	0.337/0.332	1.198	0.147/0.069
rs2159100	2346393	**T**	C	0.337/0.333	1.198	0.134/0.074
rs1024582*	2402246	**A**	G	0.348/0.338	1.19	0.286/0.000
rs4765913	2419896	**A**	T	0.216/0.217	1.14	0.622/0.833
rs4765914	2420377	**T**	C	0.212/0.206	1.14	0.646/0.495
rs3819532	2436837	**T**	C	0.395/0.401	1.32	0.963/0.721
rs3819534	2436868	**A**	G	0.395/0.394	1.32	1/0.825
rs2238065	2442631	**A**	G	0.264/0.268	1.06	0.497/0.252
rs2238066	2445399	**G**	A	0.293/0.294	1.02	0.651/0.161
rs2283302	2452619	**A**	G	0.289/0.289	1.35	1/0.53
rs2238070	2456115	**G**	T	0.437/0.448	1.33	0.874/0.005
rs2238071*	2456416	**A**	G	0.427/0.431	1.34	0.839/0.000
rs2239073	2538500	C	**T**	0.433/0.431	0.83	0.369/0.509
rs4765681	2557196	T	**C**	0.43/0.434	0.82	0.835/0.919
rs4765937	2570535	C	**T**	0.419/0.432	0.92	0.856/0.004
rs16929470	2601742	**T**	C	0.039/0.04	0.65	0.558/0.007
rs11062247*	2616128	**G**	A	0.169/0.167	0.79	$0.781/1.62 \times 10^{-7}$
rs11062248*	2616188	**T**	A	0.169/0.168	0.78	$0.812/2.23 \times 10^{-7}$

The effect alleles used to calculate the risk scores are indicated in bold. Those excluded from analysis in UK Biobank are indicated by *. The MAF given is for Allele 1(A1). All SNPs found within intron 3 of *CACNA1C*.

3.5.1 ALSPAC cohort

The primary outcome used to investigate hypomanic features in ALSPAC was the categorical hypomania measure (based on the answers given to HCL-28 at age 22-23). An unadjusted logistic regression was conducted to establish the potential for association between increasing GPRS and a designation of "hypomania". Although genotypic data was available for 7,224 Caucasian individuals in ALSPAC, not all individuals genotyped for the 19 risk SNPs-of-interest have HCL or SMFQ scores.

Table 3.3 Primary hypomania and depression outcome measures

ALSPAC	Cases	Controls
Hypomania, N (%)	181 (7.25)	2,316 (92.75)
Depression, N (%)	33(0.58)	5,689 (99.42)

A logistic regression was carried out on the 2,072 individuals genotyped for all 19 SNPs-of-interest available, using the primary outcome measure – hypomania status – as the response variable and GPR score as the explanatory variable: no association was seen (Table 3.4). A linear regression was also carried out using the secondary outcome measure (Supplemental Figure 3.2) and GPR score. Again, no association between hypomania and *CACNA1C* risk scores was apparent (Table 3.4).

Table 3.4 Regressions of primary and secondary outcomes and GPRS

Phenotype	O.R	S.E	p value	95% CI	r^2
Binary Hypomania	0.983	0.085	0.843	0.831/1.164	0.0000
Binary Depression	1.311	0.247	0.151	0.906/1.897	0.0065
	Coefficient				
HCL score	-0.163	0.132	0.217	-0.422/0.096	0.0008
SMFQ score	0.057	0.05	0.254	-0.041/0.156	0.0003

Unadjusted model

A similar strategy was used to investigate depressive features: the influence of *CACNA1C* GPR scores on the primary depression outcome (binary depression, generated from SMFQ score which was assessed between the ages of 10 and 19) was investigated using an unadjusted logistic regression (4,638 observations). The largest effect seen was for binary depression (O.R 1.311), however, this was not significant ($p<0.05$). When linear regression was applied to the secondary outcome measure, SMFQ score, the association was not significant.

Table 3.5 Primary hypomania and depression outcomes separated by sex

ALSPAC	Female	
	Cases	*Controls*
Hypomania, N (%)	99(54.7)	1,500(64.8)
Depression, N (%)	13(39.4)	2,823(49.6)

In order to investigate the potential of a sex effect, the sample was stratified. The above analyses were then repeated separately for males and females. Splitting the sample according to sex resulted in an uneven distribution of observations: only 746 males had both HCL-28 scores and were genotyped for all 19 SNPs-of-interest, whereas, 1,326 females were included in the hypomania analyses.

Table 3.6 Regressions of primary and secondary outcomes and GPRS separated by sex

Phenotype	**O.R**	**S.E**	**p value**	**95% CI**	**r^2**
Hypomania					
Female	0.862	0.102	0.21	0.683/1.088	0.0027
Male	1.135	0.143	0.317	0.886/1.453	0.0023
Depression					
Female	1.737	0.554	0.083	0.93/3.245	0.0003
Male	1.13	0.269	0.607	0.709/1.801	0.0013
HCL score	**Coefficient**				
Female	-0.332	0.167	**0.046**	-0.659/-0.005	0.0032
Male	0.088	0.215	0.681	-0.333/0.51	0.0002
SMFQ score					
Female	0.147	0.069	**0.033**	0.012/0.283	0.0022
Male	-0.036	0.073	0.622	-0.18/0.108	0.0001

Unadjusted model

The primary outcome measures for both hypomania and depression had no significant association to an increasing GPRS in either sex groups. However, increasing HCL score had a negative association to GPRS (coefficient -0.332) that was nominally significant (p 0.046). It is however worth noting that no association was found for the primary hypomania variable. A separate association was also seen in females for SMFQ score (coefficient 0.147, p 0.033) (Table 3.6). Again, no association was identified for the primary depressive outcome measure. In males, there were no significant associations for the primary or secondary outcome measures.

3.5.2 UK Biobank

To investigate the effect of *CACNA1C* GPRS on the mood disorder phenotypes logistic regression assuming an additive allelic model was used. The primary outcomes of interest were BD and recurrent MDD (Table 3.7). A logistic regression was carried out using the GPRS of individuals who were genotyped for all 15 SNPs-of-interest (which passed HWE testing).

Table 3.7 Mood phenotypes in UK Biobank

UK Biobank	*Cases*	*Controls*
Probable BD, N (%)	354(1.12)	31,386(98.88)
Probable Recurrent Depression, N (%)	5,687(15.79)	30,325(84.21)
SR Bipolar Disorder, N (%)	332(0.28)	119,621(99.72)
SR Recurrent Depression, N (%)	7,384(6.16)	112,569(93.84)
Neuroticism, N (%)	49,692(50.84)	48,045(49.16)
Mood Instability, N (%)	53,271(45.47)	63,898(54.53)

As was the case for ALSPAC, not all individuals with mood disorder data were genotyped for the chosen SNPs, resulting in 25,187 available observations for BD and 28,536 for Recurrent MDD. The logistic regression yielded no significant effects for the standardised GPRS on BD or Recurrent MDD (Table 3.8). In secondary analyses, the outcome measures of SR Bipolar Disorder and SR Recurrent MDD were assessed. However, no significant effects were found. Both unadjusted and adjusted logistic regressions were conducted. The adjusted analyses included age, sex and deprivation as well as the principal components and the quality control steps described above as potential confounders. Again, no significant associations between *CACNA1C* GPRS and BD or depression were identified.

Table 3.8 Regressions of mood phenotypes and GPRS

Phenotype	O.R	S.E	p value	95% CI	r^2
Probable BD	1.086 (1.09)	0.063 (0.063)	0.153 (0.138)	0.97/1.217 (0.973/1.222)	0.0007 (0.0228)
Probable Recurrent Depression	1.008 (1.01)	0.016 (0.016)	0.611 (0.543)	0.977/1.04 (0.979/1.042)	0.0000 (0.0218)
SR Bipolar Disorder	1.01 (1.014)	0.06 (0.061)	0.874 (0.821)	0.898/1.135 (0.901/1.14)	0.0000 (0.0254)
SR Recurrent Depression	0.984 (0.986)	0.013 (0.013)	0.216 (0.303)	0.959/1.01 (0.961/1.012)	0.0000 (0.0216)
Mood Instability	1.003 (1.005)	0.006 (0.007)	0.613 (0.425)	0.991/1.016 (0.992/1.018)	0.0000 (0.0185)
	Coefficient				
Neuroticism score	0.011 (0.011)	0.011 (0.011)	0.349 (0.307)	-0.012/0.033 (-0.011/0.33)	0.0000 (0.0411)

Unadjusted model (adjusted model)

Analyses were then carried out on the other mood disorder phenotypes of interest. Associations between the binary "mood instability" outcome variable and the GPRS were tested using logistic regression. As above, an unadjusted regression was first completed and then repeated with adjustment for potential confounders. There were no significant associations found between mood instability and *CACNA1C* risk score for either the unadjusted or adjusted regressions.

A standardized linear regression was used to test the continuous neuroticism score, also adjusted for potential confounders, but no evidence of association was found.

Table 3.9 Mood phenotypes in UK Biobank separated by sex

UK Biobank	Female	
	Cases	*Controls*
Probable BD, N(%)	163(46.04)	15,548(49.54)
Probable Recurrent Depression, N(%)	3,598(63.27)	14,826(41.17)
SR Bipolar Disorder, N(%)	179(53.91)	62,909(52.44)
SR Recurrent Depression, N(%)	4,780(64.73)	58,308(51.8)
Mood Instability, N(%)	29,528(55.43)	32,127(50.29)

The sample was then stratified to test for a potential sex effect of *CACNA1C* variants. Logistic regressions were used to investigate the effect of the *CACNA1C* GPR scores on both the primary and secondary BD/Recurrent Depression outcome measures and mood instability. As above, a linear regression was used to test the association between risk scores and the continuous neuroticism measure.

Table 3.10 Regressions of mood phenotypes and GPRS separated by sex.

Phenotype	O.R	S.E	p value	95% CI	r^2
Prob BD					
Female	1.194	0.098	**0.031** (**0.027**)	1.021/1.409	0.0031
	(1.199)	(0.099)		(1.015/1.406)	(0.0347)
Male	0.99 (0.99)	0.082	0.901 (0.904)	0.842/1.16	0.0000
		(0.082)		(0.841/1.165)	(0.0198)
Prob Rec Depression					
Female	0.996	0.02	0.841 (0.982)	0.957/1.037	0.0000
	(0.999)	(0.021)		(0.96/1.041)	(0.0108)
Male	1.025	0.026	0.34 (0.32)	0.975/1.077	0.0001
	(1.026)	(0.026)		(0.976/1.079)	(0.0084)
SR BD					
Female	1.053	0.085	0.525 (0.503)	0.899/1.233	0.0002
	(1.056)	(0.085)		(0.901/1.237)	(0.0219)
Male	0.961	0.085	0.65 (0.695)	0.808/1.143	0.0001
	(0.966)	(0.086)		(0.811/1.15)	(0.0383)
SR Rec Depression					
Female	0.973	0.016	0.103 (0.187)	0.942/1.005	0.0001
	(0.978)	(0.016)		(0.947/1.011)	(0.0141)
Male	0.999	0.022	0.949 (0.95)	0.957/1.042	0.0000
	(1.001)	(0.022)		(0.959/1.046)	(0.0144)
Mood Instability					
Female	1.003	0.009	0.757 (0.45)	0.986/1.02	0.0000
	(1.007)	(0.009)		(0.989/1.025)	(0.0193)
Male	1.002	0.009	0.773 (0.694)	0.984/1.021	0.0000
	(1.004)	(0.01)		(0.985/1.023)	(0.0151)
Neuroticism score	**Coefficient**				
Female	0.011	0.016	0.499 (0.31)	-0.02/0.042	0.0000
	(0.016)	(0.016)		(-0.015/0.047)	(0.0152)
Male	0.006	0.016	0.717 (0.69)	-0.026/0.374	0.0000
	(0.006)	(0.016)		(-0.249/0.038)	(0.0222)

Unadjusted model (adjusted model)

For males, there was no association between any of the outcome measures and *CACNA1C* GPR score (Table 3.10). However, for females, a small effect was

found for the primary BD outcome measure: both the unadjusted (p 0.031) and the adjusted (p 0.027) regressions suggested that *CACNA1C* GPRS may have a (small) positive effect on the risk of BD. However, this association was not found for the secondary BD outcome measure (SR Bipolar Disorder) and no other associations were seen in females for the other mood disorder phenotypes.

3.5.3 Replication in UK Biobank

The preliminary analyses above were undertaken using the first release of genetic data from UK Biobank (N=152,000). The more recent collection of data using the MHQ is potentially a more reliable assessment of BD and MDD status. The *CACNA1C* GPRS were then tested for association with lifetime BD and MDD status. However, there were no associations found using these MHQ-defined phenotypes (Table 3.11).

The associations between *CACNA1C* GPRS and neuroticism, and mood instability were also tested in the larger UK Biobank cohort but no associations were identified (Table 3.11).

Table 3.11 Regressions of MHQ mood phenotypes and GPRS

Phenotype	O.R	S.E	p value	95% CI	r^2
Lifetime BD	0.994	0.026	0.834	0.945/1.047	0.000
	(0.994)	(0.026)	(0.815)	(0.994/1.046)	(0.018)
Lifetime MDD	1.005	0.007	0.422	0.993/1.018	0.000
	(1.004)	(0.007)	(0.532)	(0.944/1.046)	(0.043)
Mood	1.001	0.003	0.72	0.995/1.007	0.000
Instability	(1.001)	(0.003)	(0.66)	(0.995/1.008)	(0.017)
	Coefficient				
Neuroticism	0.003	0.005	0.55	-0.007/0.014	0.000
score	(0.003)	(0.005)	(0.552)	(-0.007/0.014)	(0.02)

Unadjusted model (adjusted model)

Consistent with the preliminary analyses above, UK Biobank participants with MHQ data were stratified by sex to assess potential associations between *CACNA1C* GPRS and the MHQ mood disorder measures, neuroticism, and mood instability. However, no significant associations between the GPRS and these mood phenotypes were found in males or females (Table 3.12).

Table 3.12 Regressions of mood phenotypes and GPRS separated by sex

Phenotype	O.R	S.E	p value	95% CI	r^2
Lifetime BD					
Female	0.962	0.003	0.209 (0.198)	0.906/1.022	0.0031
	(0.961)	(0.029)		(0.905/1.021)	(0.0347)
Male	1.004	0.003	0.896 (0.957)	0.942/1.071	0.0000
	(1.002)	(0.003)		(0.939/1.068)	(0.0198)
Lifetime MDD					
Female	0.998 (1)	0.008	0.787 (0.968)	0.982/1.014	0.0000
		(0.008)		(0.984/1.016)	(0.0108)
Male	1.012	0.011	0.282 (0.411)	0.99/1.034	0.0001
	(1.009)	(0.011)		(0.987/1.031)	(0.0084)
Mood Instability					
Female	0.997	0.004	0.48 (0.7)	0.989/1.005	0.0000
	(0.998)	(0.004)		(0.99/1.007)	(0.0193)
Male	1.005	0.005	0.26 (0.27)	0.996/1.014	0.0000
	(1.005)	(0.008)		(0.996/1.014)	(0.0151)
Neuroticism score	**Coefficient**				
Female	-0.001	0.007	0.846 (0.31)	-0.015/0.014	0.0000
	(0.001)	(0.007)		(-0.013/0.016)	(0.0152)
Male	0.005	0.008	0.52 (0.52)	-0.01/0.021	0.0000
	(0.005)	(0.008)		(-0.01/0.02)	(0.0222)

Unadjusted model (adjusted model)

3.6 Discussion

Overall, within the initial preliminary investigation using both cohorts described above, there were no significant associations between *CACNA1C* risk scores and the mood disorder-related traits when assessing the samples as a whole. The samples were then stratified by sex based on a priori rationale, and there was evidence of a small association between the GPRS and risk of BD for females (but not for males).

Within ALSPAC, higher *CACNA1C* GPR scores were associated with a lower HCL-28 score in females. This finding is in a direction which is opposite to what was anticipated from the existing literature and is difficult to account for (Witt et al. 2014; Court et al. 2014; Smith et al. 2015; Hayes et al. 2016). It contrasts with the finding in UK Biobank that an increasing risk score was associated with greater risk of bipolar status in females (O.R 1.199, 95% CI 1.015/1.406). One possibility is that this is related to the different age ranges of the ALSPAC and UK Biobank cohorts: perhaps *CACNA1C* variants have a differential impact on mood symptoms across the life course, the variants may be protective at one stage and deleterious at another as a result of evolutionary pressures ("antagonistic pleiotropy") (Carter & Nguyen 2011).

It is also possible that *CACNA1C* may specifically influence the depressive aspect of BD because a significant effect on SMFQ scores was seen in females within the ALSPAC cohort. In this analysis there seemed to be an increased risk of greater SMFQ scores with an increasing *CACNA1C* GPR score. It may be that *CACNA1C* has a greater influence on the depressive features of BD relative to the hypomanic/manic features; however, it is worth noting that it is unclear why this would be the case mechanistically, and no significant associations were identified in females (or males) for any of the depression outcome measures in UK Biobank. The findings within ALSPAC also contradict the hypothesis that *CACNA1C* variants are implicated in the manic phases of BD. Interestingly, one study has found that treatment-resistant bipolar patients administered L-VGCC antagonists display improvement of manic symptoms, although there was no improvement of depressive symptoms (Kabir et al. 2017).

It is important to note that by dividing the ALSPAC sample by sex there were fewer observations available for testing which may have resulted in the

analyses becoming underpowered to detect true effects (type 2 error), or contributed to spurious findings (type 1 error). To detect small effects at 95% power would require a sample size of greater than 1,800; when stratifying by sex in ALSPAC a sample of less than 1,600 females remained (Faul et al. 2007). Nonetheless, some effects were found for females within both cohorts during the preliminary analyses, perhaps suggesting that the *CACNA1C* locus may have a potential sex effect in the context of mood disorders, and particularly for BD. Previous studies have also identified a potential sex effect of *CACNA1C* variants where variants were more strongly associated with the phenotypes-of-interest in females versus males, although these studies tended to focus solely on single variants tested separately (Dao et al. 2010; Strohmaier et al. 2013; Heilbronner et al. 2015; Starnawska et al. 2016). Overall, the effects seen in both the ALSPAC and UK Biobank cohorts might be considered to provide some preliminary evidence for a potential differential effect of *CACNA1C* variants in females.

However, it is important to note that these findings were in fact contradictory between the two cohorts (effects were in opposite directions) and were based on self-reported data. Clearly before any firm conclusions can be drawn these findings require replication. Upon release of the remaining UK Biobank genetic data and release of MHQ data for over 157,000 participants, replication was possible. Using the lifetime mood disorder variables derived from the more extensive MHQ, there was no evidence for associations between *CACNA1C* GPRS and mood disorder outcomes to replicate the preliminary findings in UK Biobank.

It is of note that there currently are no further, up-to-date, mood-related traits in ALSPAC to replicate the significant associations between *CACNA1C* GPRS and hypomania, and depressive state in females within ALSPAC.

It has been reported that calcium channel antagonists may have some benefit in the mood stabilisation of BD (Bidaud & Lory 2011; Erk et al. 2014b), and with weak evidence of a potential sex effect of *CACNA1C* it is theoretically possible that female patients may benefit more from the inclusion of calcium channel antagonists than male patients. This is of course a speculative suggestion that requires much more detailed future investigation.

Furthermore, although some significant effects were found in females,

the *CACNA1C* GPRS explained very little of the variance observed within the cohorts (maximum r^2 of 0.0347, i.e., 3%).

Despite the lack of evidence for associations between *CACNA1C* risk score and mood disorder-related traits, there are strengths to this investigation. To my knowledge, this is the first application of using a GPRS – generated using variation in a single BD candidate risk gene – to investigate the potential influence of *CACNA1C* on mood disorder phenotypes.

This chapter also demonstrates the evaluation of *CACNA1C* on relevant mood phenotypes in larger samples than existing *CACNA1C* literature. A strength of this chapter is the application of these analyses in two separate cohorts of two different age groups.

3.7 Limitations

In testing the hypotheses in this way there are limitations. One is the use of self-reported outcome measures for all analyses carried out; the outcome measure may be subject to reporting bias (Ganna & Ingelsson 2015). However, the measures of BD and recurrent depression within UK Biobank have proved to be useful and reasonably consistent with expected associations within previous studies (Smith et al. 2013a). Similarly, for the ALSPAC data, the HCL-32 has been widely used to define hypomania (Smith et al. 2015) and the SMFQ has been used to measure depressive symptoms (Stochl et al. 2015).

A limitation of the investigations in ALSPAC is the lack of statistical power. There were a much smaller number of observations in ALSPAC relative to UK Biobank for the desired mood phenotypes tested. The ALSPAC analyses may therefore be underpowered to detect true associations.

When stratifying the samples for sex in ALSPAC, there were an uneven number of males and females for the analyses, which could have influenced the findings as the female sample could have been better powered to detect effects, although the adjusted analysis controlled for this to some extent within the main analysis in ALSPAC by adjusting for sex.

The analyses within ALPSAC are also unadjusted for potential confounding factors such as deprivation and require further correction. It is important to note, also, that the results displayed above (both from the ALSPAC and UK

Biobank cohorts) have not been corrected for multiple testing and as such can only be considered preliminary exploratory findings at best (the observed observations did not survive false discovery rate (FDR) correction).

Also of note is that not all of the SNPs-of-interest were independent of each other; some of the chosen SNPs are in high LD with each other (rs1006737 is in high LD with rs2007044, rs2159100 and rs1024582, rs4765913 is in high LD with rs4765914, rs3819532 is high LD with rs3819534, rs2238070 and rs2238071 with the two latter SNPs in high LD with each other, rs2238065 is in high LD with rs2238066 and rs2283302, and rs11062247 and rs11062248 (Supplemental Figure 3.1)) (Dudbridge 2016) which could be confounding the effects reported. As to be expected, the LD patterns were consistent between the two cohorts.

Another limitation of this work is the difficulty in drawing comparisons between the two cohorts chosen. UK Biobank is a longitudinal cohort of an older sample from different areas of the UK, whereas, ALSPAC is a birth cohort with individuals currently now in early adulthood, all of whom were from a specific area of the UK. In addition to this, the mood phenotypes tested here have been recorded differently between the two cohorts, making it difficult to compare findings directly. However, one novel and potentially useful strength of this work is that it examined different age groups and as such covers a wide age-range of adults.

Clearly, not all individuals who carry these variants go on to develop neuropsychiatric disorders (Yoshimizu et al. 2015). It is also possible that the SNPs associated with greater susceptibility to BD are not causative but rather are in linkage disequilibrium with the true risk variants (Heyes et al. 2015). Furthermore, the risk SNPs may be incompletely penetrant (Yoshimizu et al. 2015), consistent with the prevailing view that neuropsychiatric disorders are complex polygenic conditions influenced by a diverse combination of genetic, epigenetic and environmental factors (Yoshimizu et al. 2015).
Identifying these causal variants would provide greater understanding of the underlying pathophysiology of psychiatric conditions.

Finally, the findings above cannot explain the underlying molecular mechanisms by which variants in the *CACNA1C* gene may influence mood phenotypes and, hence, provide no insight into the underlying biology of mood disorders.

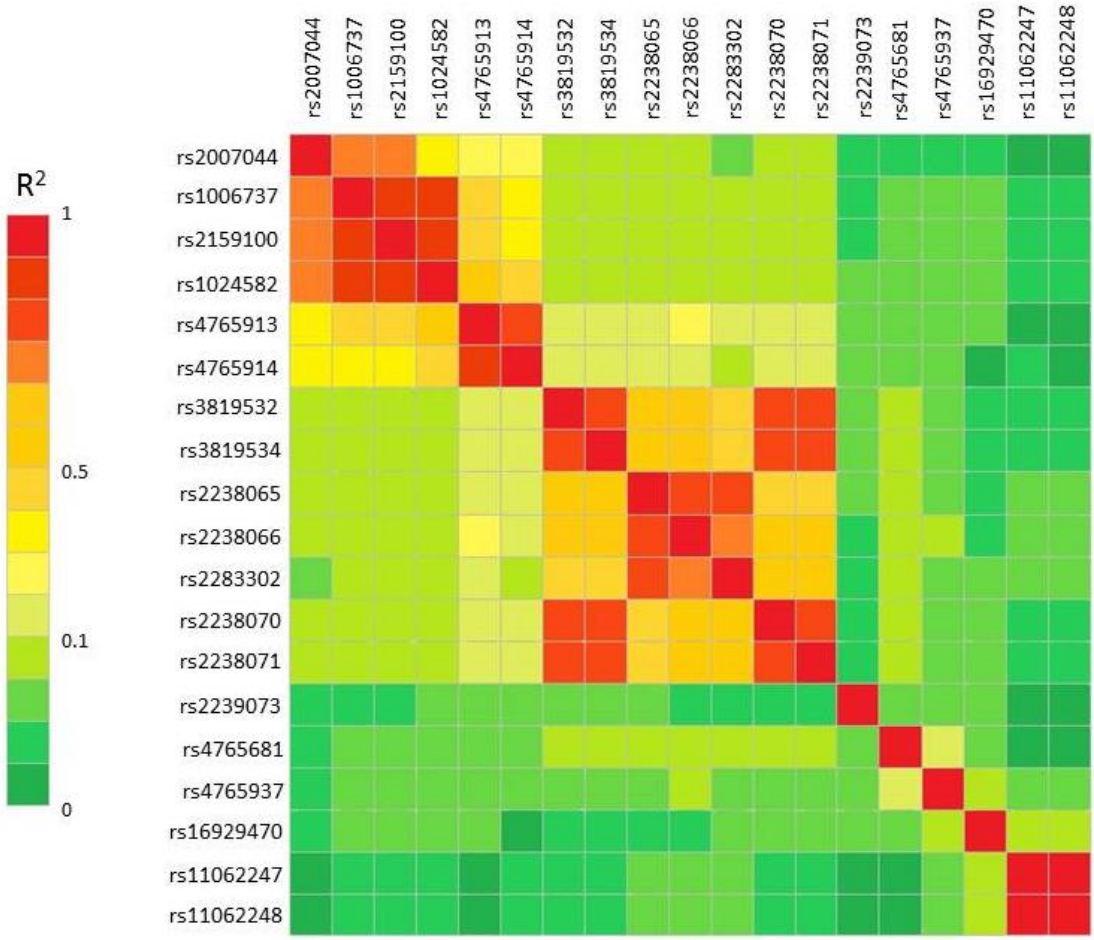
3.8 Future Work

As noted above, the analyses in this chapter are to some extent exploratory. This investigation focussed on the *CACNA1C* risk score in the context of mood disorders and mood disorder-related phenotypes. As there was some weak evidence of the potential influence of *CACNA1C* GPRS on these phenotypes, it may be of interest to investigate the *CACNA1C* GPRS in the context of other traits known to be associated with mood disorders, such as disrupted circadian rhythms. Circadian rhythm disruptions are a common feature of BD (Hayashi et al. 2015; Steinan et al. 2015) and the *CACNA1C* gene is known to have a role in circadian rhythm physiology (Shi et al. 2008; Schmutz et al. 2015; McCarthy et al. 2016).

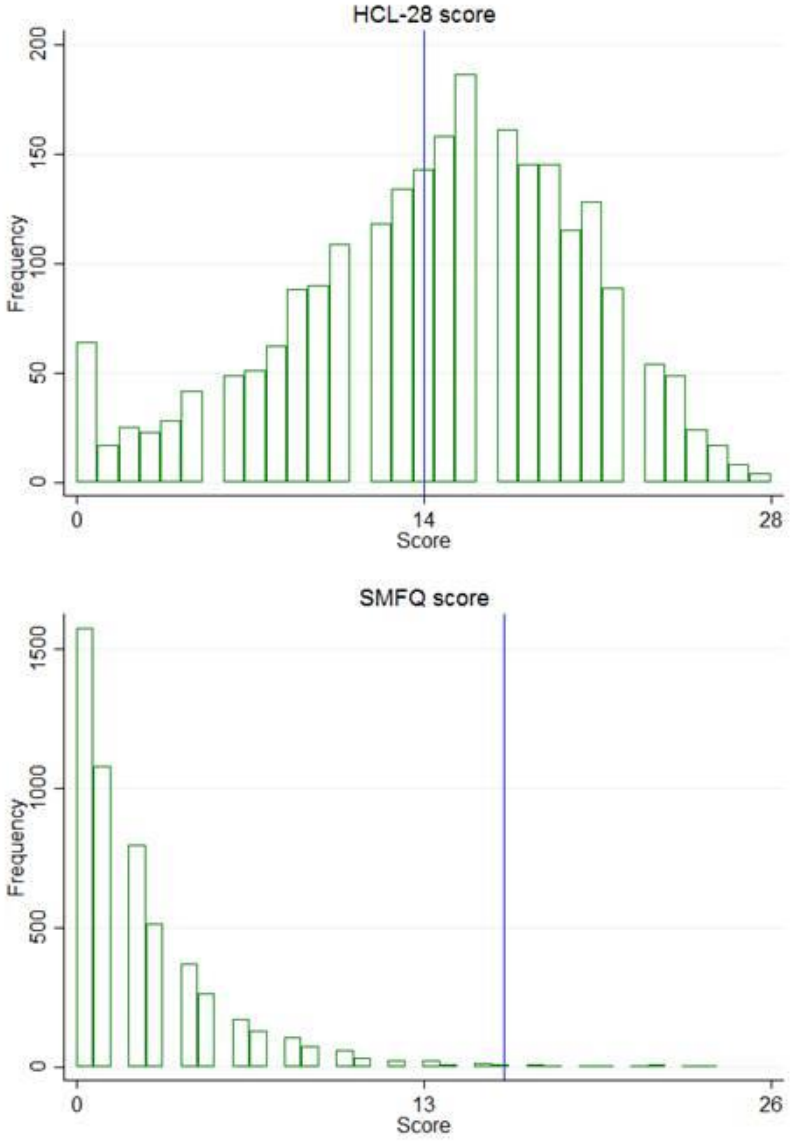
3.9 Conclusions

The exact impact of *CACNA1C* genetic variants on mood-related phenotypes in generally healthy populations is unclear. However, the potential for an increasing *CACNA1C* risk score influencing BD and/or MDD pathophysiology and for sex-specific effects (as seen in other studies) cannot be dismissed; further study is required in independent large cohorts.

3.10 Supplementary Figures

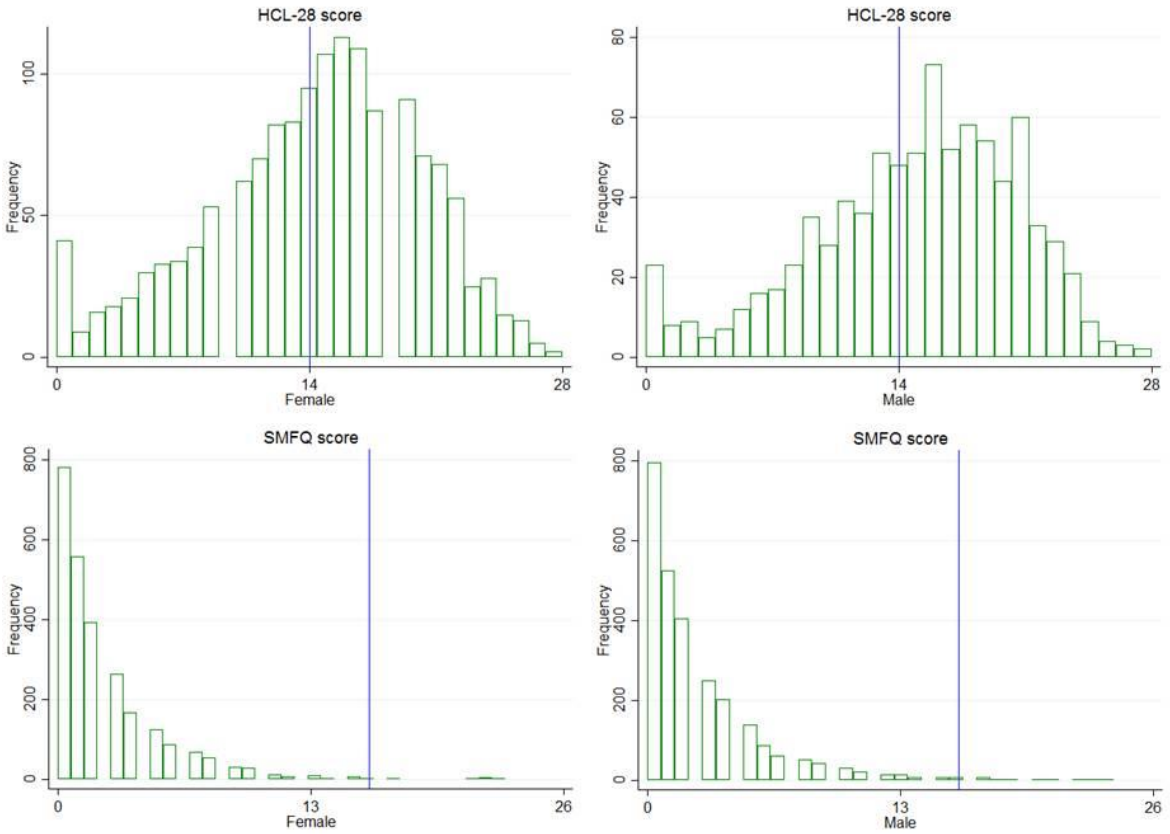


Supplemental Figure 3.1 LD heat map of selected *CACNA1C* SNPs-of-interest



Supplemental Figure 3.2 Histogram of HCL-28 and SMFQ scores

Blue line on HCL-28 score graph indicates threshold at which individuals could be designated as hypomanic, however, three additional questions are required to assign individual as hypomanic (detailed in chapter 2).
Blue line on SMFQ score graph shows threshold for designating individual as a depressive case for the primary depression outcome in ALSPAC (score of 16 or greater).



Supplemental Figure 3.3 Histogram of HCL-28 and SMFQ scores separated by sex

Blue line on HCL-28 score graph indicates threshold at which individuals could be designated as hypomanic, however, three additional questions are required to assign individual as hypomanic (detailed in chapter 2).
Blue line on SMFQ score graph shows threshold for designating individual as a depressive case for the primary depression outcome in ALSPAC (score of 16 or greater).

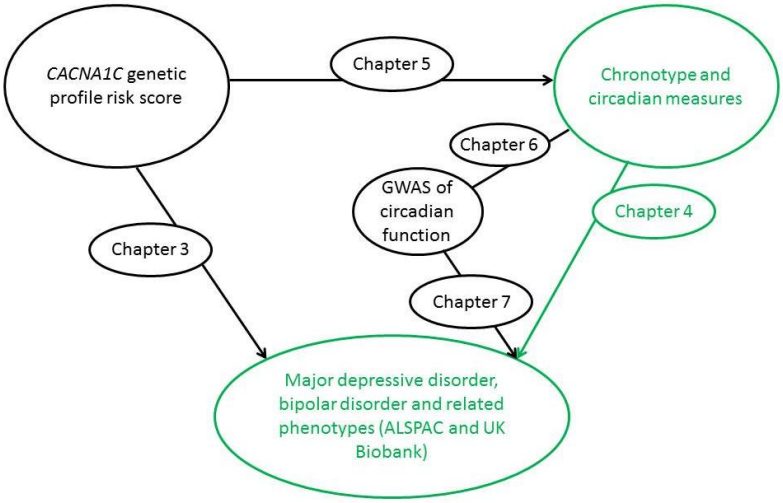
Supplemental Table 3.1 Summary of HCL-28 and SMFQ scores

ALSPAC	Mean	Median	Min	Max	N
HCL score	14.548	15	0	28	2326
SMFQ score	2.315	1	0	26	4879

Supplemental Table 3.2 Summary of HCL-28 and SMFQ scores separated by sex

ALSPAC	Mean	Median	Min	Max	N
HCL score					
Female	14.315	15	0	28	1476
Male	14.952	16	0	28	850
SMFQ score					
Female	2.331	1	0	26	2447
Male	2.299	1	0	26	2432

Chapter 4 Investigating the effect of genetic variants for chronotype preference on mood disorder and sleep phenotypes in two population cohorts



Outline 4.1 Chapter in context of overall study

4.1 Introduction

The *CACNA1C* gene locus has been shown to have an association with mood disorders including BD and MDD (Bigos et al. 2010; Sklar et al. 2011; Smoller 2013; Heilbronner et al. 2015; Kabir et al. 2017) and there is increasing evidence that altered calcium signalling leading to disruption of second messenger systems, possibly as a result of variation at this locus, influences the pathophysiology of BD (Harrison 2016).

Analyses in Chapter 3 did not demonstrate a clear association between *CACNA1C* GPRS and mood disorder phenotypes. It is of note that *CACNA1C* has been robustly associated with BD and related traits, with many of these associations being replicated; but the underlying mechanism by which *CACNA1C* variants influence the full range of bipolar phenotypes remains unclear.

CACNA1C gene expression is known to both be influenced by circadian rhythms and this expression in turn influences circadian rhythmicity (Schmutz et al. 2015; McCarthy et al. 2016). There is potential that *CACNA1C* is involved in the pathophysiology of BD through altering circadian rhythms. However, the exact influence of circadian rhythmicity genetics on mood disorder phenotypes in ALSPAC and UK Biobank is unclear. This chapter aims to investigate potential associations between mood disorder-related phenotypes and features of circadian rhythmicity, in particular chronotype.

4.1.1 Circadian rhythmicity in mood disorders

Mood disorders are often associated with sleep disturbances (Pagani et al. 2016; Dmitrzak-Węglarz et al. 2016; Asaad et al. 2016). Sleep-wake cycle abnormalities are observed in all phases of BD, as well as between episodes (Baek et al. 2016; Pagani et al. 2016). When compared to healthy controls, individuals with BD display a variety of sleep problems including longer sleep latency, higher sleep fragmentation and greater sleep disturbances (Geoffroy et al. 2014). BD patients appear to be susceptible to sleep-wake cycle disturbances and tend to easily shift their circadian phase when sleep-wake cycles are disrupted or exposed to inappropriate artificial light; this high sensitivity to circadian rhythm phase shifts could be a potential marker of bipolarity (Moon et al. 2016). Individuals at risk of developing BD also display sleep and circadian rhythm irregularities; these abnormalities can be seen before the onset of the disorder and may be considered a modifier of disease course (Bellivier et al. 2015).

Several aspects of sleep profiles and circadian rhythms have been investigated in the context of mood disorders, including circadian phase preference (chronotype), rest and activity measures and melatonin peak time via self-report questionnaires, actigraphy and salivary secretions of melatonin and cortisol (Bellivier et al. 2015; Baek et al. 2016; Dmitrzak-Węglarz et al. 2016). BD patients demonstrate greater daytime dysfunction relative to healthy controls (Geoffroy et al. 2014). These circadian rhythm abnormalities are associated with disrupted brain function, including impaired cognition and emotional processing (Wulff et al. 2010), and have also been suggested to lead to deterioration in the mental health of otherwise healthy individuals (Landgraf et al. 2014). This may be indicative of bidirectional relationship between an individual's circadian rhythm and their mood and overall cognitive functioning (Landgraf et al. 2014). The misalignment between endogenous circadian rhythms and an individual's environment is a common feature for individuals with BD and has been associated with both acute manic and depressive relapses (Castro et al. 2015; Moon et al. 2016).

4.1.2 Chronotype

Chronotype, or diurnal preference, is defined as an individual's preference for wakefulness and activity at a particular time of day (Alloy et al. 2017). Individuals are usually classified into three broad chronotypes: morning ("larks"), evening ("owls") and intermediate (a combination of morning and evening) (Berdynaj et al. 2016). Chronotype is considered to be a combination of genetics, biological processes and psychosocial processes (Etain et al. 2014; Dmitrzak-Węglarz et al. 2016). Diurnal preference is a physiological trait with a clear biological basis and has been associated with other endogenous phase markers including melatonin secretion, cortisol-awakening response and the circadian shift in body temperature (Bellivier et al. 2015). Differences in these endogenous phase markers have been reported in individuals with different chronotype preferences; for example, evening-types demonstrate a phase delay in peak body temperature (N=14) and cortisol levels (N=125), and reduced night time melatonin peak levels relative to morning-types (N=170) (Randler & Schaal 2010; Kerkhof & Van Dongen 1996; Burgess & Fogg 2008) and the diurnal peak in physiological functions (including core temperature and melatonin levels) occurs earlier in morning than in evening-types (Merikanto et al. 2016; Desanctis 2017).

There is evidence that an individual's timing preference for daily activities, their chronotype, may be associated to greater risk of adverse health conditions: evening chronotypes have an increased risk of sleep problems, hypertension and type 2 diabetes, compared to individuals with a morning chronotype (Merikanto et al. 2013; Merikanto et al. 2016).

4.1.3 Potential influence of chronotype on mood

Chronotype has previously been associated with several mood disorder-related phenotypes, including anxiety and depression (Corruble et al. 2014; Dmitrzak-Węglarz et al. 2016). Certain chronotypes may predispose some individuals to mood disorders (Merikanto et al. 2013). Evening-types have been seen to associate with depressive symptoms in both depressed patients and healthy controls; depressed patients were more likely to be evening-types while healthy individuals with a late chronotype were found to have higher depression scores and had an increased likelihood of presenting moderate or severe depression

symptoms (Chan et al. 2014; Berdynaj et al. 2016; Levandovski et al. 2011; Hidalgo et al. 2009). A study of depressed patients demonstrated that individuals reporting an evening chronotype had higher non-remission rates than those reporting a morning chronotype upon follow-up (O.R 3.36, 95%CI 1.35/8.34). It is of note, this study is based on a relatively small sample size of 253 patients with a high proportion of females (82.6%) (Chan et al. 2014).

There is increasing evidence to suggest that chronotype has an involvement in the pathophysiology of depression as evening chronotypes reportedly display similar negative biases in emotional processing often seen in depressed patients even in participants with no history of depression (Berdynaj et al. 2016). Individuals with BD also more commonly report an evening chronotype preference compared to healthy controls (Baek et al. 2016; Alloy et al. 2017); plus eveningness has also been associated with rapid mood cycling and earlier age of onset in BD (Bellivier et al. 2015).

4.1.4 Genome wide association studies of chronotype

Previous GWAS of BD and MDD have identified several clock genes which have core functions in maintaining circadian rhythms (Landgraf et al. 2014; Maclukiewicz et al. 2014) and some studies have suggested that variation within the clock genes are associated with specific BD clinical subtypes (Moon et al. 2016). Research involving animal models has demonstrated the link between circadian clock genes and brain functions associated with psychiatric illness through the manipulation of light-dark cycles and gene knockout experiments including deletion of *PER1*, *PER2* and *CLOCK* genes resulting in hyperactivity, greater reward-seeking behaviour, depression-like and mania-like behaviour (Landgraf et al. 2014).

At the time of this analysis, three GWAS were undertaken in different sample populations to identify variants which associate with an individual's likelihood of reporting a particular chronotype preference (Jones et al. 2016b; Lane et al. 2016; Hu et al. 2016). Two of the studies used self-reported chronotype of over 100,000 individuals from UK Biobank (Lane et al. 2016; Jones et al. 2016b), whereas the third used self-reported chronotype responses from 89,000 participants of 23andMe (Hu et al. 2016). Across the three genome-wide studies,

a total of 37 independent SNPs were reported to associate with self-reported chronotype. One study exclusively used the 23andMe cohort to investigate self-reported morningness in over 89,000 individuals and found 15 SNPs associated with identifying as "a morning person" (Hu et al. 2016); a further 13 SNPs were found to associate with self-reported morningness in UK Biobank in approximately 128,000 participants and were replicated within the 23andMe cohort (Jones et al. 2016b).

Self-reported eveningness was also investigated in UK Biobank (100,000 participants) and identified 11 SNPs (Lane et al. 2016); the alternate alleles of two of those SNPs (rs2050122 and rs10157197) were also found to associate with self-reported morningness (Jones et al. 2016b). Both Hu et al. and Lane et al. treated chronotype as a binary trait, whereas, Jones et al. used a linear chronotype measure for their GWAS which could have provided greater statistical power to detect variants.

Although both Lane et al. and Jones et al. used UK Biobank participants, the exclusion criteria between the two studies were different. Jones et al. excluded individuals with any reports of diabetes to allow for downstream investigations, whereas, Lane et al. excluded shift workers and those on medication for sleep problems which could have potentially confounding effects on circadian rhythmicity (Ferguson et al. 2018).

Each of the studies reported SNP loci which implicate known circadian genes; many other pathways were also implicated including, energy metabolism, immune response, nucleotide metabolism, gene expression and light detection. The self-reported chronotype was also found to genetically correlate with schizophrenia in UK Biobank (Lane et al. 2016; Jones et al. 2016b).

4.2 Rationale

Approximately 40% of the population have a distinct morning or evening chronotype (Alloy et al. 2017); chronotype preference is reported to correlate with many endogenous circadian phase markers including measures of rest and activity(Duffy et al. 2001; Etain et al. 2014). As noted above, evening chronotypes have been linked to increased risk of mood disorders (often demonstrated by significantly lower composite scale of morning scores in mood disorder patients ($p<0.0001$) (Alloy et al. 2017; Baek et al. 2016; Wood et al.

2009) and several studies have demonstrated BD patients are more likely to be evening-types (Etain et al. 2014). Circadian preference is thought to be genetically determined (Etain et al. 2014); however, as yet, genetic variants associated with chronotype have not been tested for any association to different features of mood disorders. An evening chronotype has been hypothesised to be a pre-existing factor for BD (Pagani et al. 2016; Alloy et al. 2017).

If genetic loading for a particular chronotype associates with features of mood disorder, it may help strengthen evidence of this hypothesis and support the suggested requirement for the stabilisation of rest-activity rhythms to be included in the management of mood disorders. It is often suggested that the treatment of disrupted sleep and circadian rhythm disruptions should be combined with typical pharmacotherapy to create a better treatment approach for BD and depression (Bellivier et al. 2015); a clear clinical assessment of both a range of mood and sleep symptoms may be required in order to more effectively manage and treat mood disorders (Chan et al. 2014; Geoffroy et al. 2014). Manipulation of sleep has been reported to influence mood symptoms (Baek et al. 2016) and individuals with depressive symptoms have described some improvement of symptoms after chronotherapeutic interventions (Li et al. 2013).

4.3 Hypothesis to be tested

The aim of this investigation is to establish whether, on average, higher polygenic loading for both morningness and (separately) eveningness were associated with specific mood disorder related phenotypes within two population cohorts (ALSPAC and UK Biobank). These phenotypes included: hypomania and depressive features (within ALSPAC); and mood instability, neuroticism, lifetime BD status, lifetime MDD status and lifetime Generalised Anxiety Disorder (GAD) status (within the UK Biobank). The key question was whether individuals who were more genetically predisposed to a morning chronotype preference were less likely to display mood disorder-related phenotypes? Conversely, were individuals with a greater polygenic loading for an evening chronotype preference more likely to display these mood phenotypes?

4.4 Methods

4.4.1 ALSPAC

Participants of ALSPAC who were genotyped were only included in the ALSPAC genotype database after meeting particular quality control criteria. Details of quality control measures, imputation and phasing described in Chapter 2. Only unrelated individuals were included in these analyses in an attempt to prevent shared environmental factors influencing associations (Dudbridge 2016); a list of unrelated individuals (N=8,197; 96% of the genotyped sample) was provided by ALSPAC. Individuals that were not recorded as Caucasian, or those with missing ethnic background data (N=1,112; 13.08%), were also removed from analyses leaving N=7,390.

4.4.1.1 Sleep phenotypes

There is no measure of chronotype within ALSPAC and so the PRS were tested against two measures of sleep problems reported at two different ages (age 10 and age 13). This resulted in four sleep-related variables; these variables were generated from mother-answered questionnaires regarding child sleep habits. These phenotypes are detailed in Chapter 2.

As a proxy for disrupted circadian rhythm (and due to the lack of chronotype data in ALSPAC), the primary outcome measures for sleep in ALSPAC were the categorical variables “Difficulty sleeping-10” and “Difficulty sleeping-13” which were tested for their association to chronotype PRS using logistic regression. It is of note that due to the small numbers of observations available for individuals with both mood outcome and sleep phenotype data for use in phenotypic association analysis (Supplemental Table 4.1) individuals that were not of white European ancestry were included in this phenotypic analysis.

4.4.1.2 Hypomania

To test associations between chronotype PRS and features of BD in ALSPAC categorical and continuous measures of hypomania were used. The full details of how these measures are generated using HCL-28 can be found in Chapter 2. The

categorical definition of hypomania was the primary outcome measure and the continuous HCL-28 score was a secondary outcome.

4.4.1.3 Depressive features

Depressive symptoms were assessed in the ALSPAC cohort during late childhood and adolescence using the SMFQ. This generates an SMFQ score ranging from 0 to 26 (Wiles et al. 2012), as well as a categorical depression measure based on a score of greater than 16. As for hypomania, the categorical SMFQ measure was used as the primary outcome measure and the SMFQ score was the secondary outcome. Detailed in Chapter 2.

The primary outcome measures of categorical hypomania and categorical depression were tested for association with the chronotype PRS using logistic regression assuming an additive risk allelic effect model. The dimensional secondary outcome measures (HCL-28 score and SMFQ score) were analysed using linear regression.

4.4.2 UK Biobank cohort

4.4.2.1 Chronotype phenotype

Chronotype was derived from the participants' responses to a question obtained from the Morningness-Eveningness questionnaire (Taillard et al. 2003). This question is an accepted measure of chronotype and has been reported to explain the greatest variance in preference of sleep-wake timings (Taillard et al. 2003). Categorical variables were then generated based upon the responses given, resulting in the generation of four separate chronotype variables. The primary outcome measures used for analysis were the "definite morning" and "definite evening" variables with "overall morning" and "overall evening" providing secondary outcome measures. The method by which these variables were generated is detailed in Chapter 2.

4.4.2.2 Bipolar Disorder, Depression and Generalised Anxiety Disorder

A MHQ was developed by a mental health research reference group to collect additional mental health phenotype data in UK Biobank and was administered during 2016-2017. Lifetime BD, lifetime MDD and lifetime GAD variables were

generated for 157,366 UK Biobank participants. Further detail for these variables contained in Chapter 2.

4.4.2.3 Neuroticism

To define neuroticism a score taken from the 12 item neuroticism scale of the EPQ-R-S (Smith et al. 2013a). Individuals were given a score of 0 or 1 for a "no or yes" answer, respectively, for each item and given a total neuroticism score ranging from 0 to 12. As described in Chapter 2.

4.4.2.4 Mood instability

A "mood instability" outcome measure was also obtained from the EPQ-R-S questionnaire: participants were asked *"Does your mood often go up and down?"* and given the option to answer "yes", "no", "don't know" or "prefer not to answer". Individuals who selected "don't know" or "prefer not to answer" were coded as missing; this allowed the generation of a categorical mood instability variable where those who answered "yes" were designated as cases and participants who answered "no" were included as controls. As detailed in Chapter 2.

4.4.2.5 Association analyses

Most of the association analysis carried out used logistic regression assuming an additive allelic effect model (the association analysis using neuroticism score used a linear regression). The adjusted analysis included age, sex, socioeconomic status (assessed using the Townsend deprivation score), 8 UK Biobank genetic principal components, assessment centre and batch as covariates. Chronotype preference has been reported to associate with gender (Hu et al. 2016) and differences in self-reported chronotype can be seen in the sample of the UK Biobank cohort used in these analyses. The association analyses were also performed separately for males and females, both unadjusted and adjusted, with age, deprivation index, 8 UK Biobank genetic principal components, assessment centre and batch included as covariates (Ward et al. 2017).

4.4.2.6 Generating polygenic risk scores

Weighted PRS were generated for all individuals with genotype data available in ALSPAC and UK Biobank using Plink 1.9 (Purcell et al. 2007). To provide weighting for the SNPs the log of the odds ratios provided by the chronotype GWAS literature were used (15 SNPs from Hu et al., 14 SNPs from Jones et al. and 9 SNPs Lane et al.) (Hu et al. 2016; Jones et al. 2016b; Lane et al. 2016). PRS methodology is detailed in Chapter 2. Two separate polygenic risk scores were generated for each individual; one composed of the genome-wide significant SNPs associated with being a "morning person" compared to being an "evening person" and a second composed of the genome-wide significant variants associated with being an "evening person" compared to being a "morning person".

Of the 27 SNPs reported to associate with self-reported morningness by two of the GWAS papers, only 25 SNPs were included for analyses as some SNPs were not genotyped within ALSPAC or UK Biobank. For previous investigations only observations for individuals genotyped at all SNPs-of-interest were used for testing, however, as there were approximately only 2500 individuals in ALSPAC genotyped for these SNPs-of-interest who also have hypomania data, only individuals missing more than 3 of the SNPs-of-interest were excluded. For the evening PRS, 11 SNPs were investigated but not all SNPs were found to be genotyped in ALSPAC and UK Biobank. In ALSPAC, only 9 SNPs associated with self-reported eveningness had been genotyped and only individuals missing 2 or more SNPs were excluded.

As UK Biobank was used as the initial discovery sample, individuals included in the chronotype GWAS were removed from further analyses. Only individuals genotyped for all 25 SNPs-of-interest and the 10 eveningness SNPs genotyped in UK Biobank were included for the association analyses using the morningness PRS and eveningness PRS, respectively.

Statistical analyses used are detailed in Chapter 2. Changes in the weighted PRS scores are small due to the use of the log of O.R; to provide a more comprehensive interpretation of the analyses, the weighted scores were standardised to their z values (i.e. per SD). The standard nominal p value significance cut-off value of $p<0.05$ was used for all analyses.

4.4.3 Two-Sample Mendelian Randomisation (MR)

Genetic instruments were selected for morningness and depression using two non-overlapping samples were obtained from the summary statistics of Jones et al. (2016), PLoS Genetics (Jones et al. 2016b), and Wray et al. (2018), Nature Genetics (Wray et al. 2018). Overlapping SNPs between genome-wide significant SNPs associated with MDD in Psychiatric Genomics Consortium (PGC) (excluding UK Biobank and 23andMe) and all SNPs associated with morningness in UK Biobank were identified and tested for overall genetic effect on morningness. Overlapping SNPs between genome-wide significant SNPs associated with morningess and all SNPs associated with MDD were identified and tested for overall genetic effect on MDD. Overall genetic effects were obtained via meta-analysis of SNP WR, with causal relationships investigated using MREgger and IVW. MREgger controls for potential pleiotropy, however, is a more conservative method than IVW (Bowden et al. 2016).

4.5 Results

Table 4.1 Demographics of cohorts

	ALSPAC N total = 8,197	**UK Biobank** N total = 119,953
Sex		
Female, N (%)	3,525 (48.8)	63,088 (52.6)
Age		
Mean (SD)	24.497 (0.5)	56.867 (7.93)
Deprivation (Townsend score)		
Mean (SD)	2.831 (1.33)	-1.466 (2.99)
Range	3	-2.278

As demonstrated in a previous chapter, the two population cohorts capture participants of different age groups and so it is unlikely that there is any overlap between the samples. The deprivation data in ALSPAC was obtained from the maternal socioeconomic status and Townsend score was used in UK Biobank.

4.5.1 ALSPAC cohort

The two PRS contain SNPs which were identified to have an association with self-reported chronotype in UK Biobank; however, it is unclear as to whether these SNPs are also associated with sleep phenotypes in ALSPAC. Analyses were carried out to explore the potential association between the chronotype risk scores and available sleep variables in ALSPAC (Table 4.2).

Table 4.2 Logistic regressions of sleep outcome measures and chronotype polygenic risk scores

Phenotype	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Difficult sleeping-10	0.955 (0.954)	0.069 (0.069)	0.521 (0.513)	0.828/1.1 (0.827/1.099)	0.000 (0.004)
Difficult sleeping-13	0.926 (0.929)	0.069 (0.068)	0.298 (0.316)	0.802/1.07 (0.804/1.073)	0.001 (0.003)
Eveningness PRS					
Difficult sleeping-10	1.074 (1.074)	0.074 (0.074)	0.302 (0.299)	0.938/1.229 (0.938/1.23)	0.001 (0.002)
Difficult sleeping-13	0.916 (0.91)	0.063 (0.063)	0.201 (0.171)	0.801/1.048 (0.796/1.041)	0.002 (0.005)

Unadjusted model (adjusted model)

There were no significant associations when testing the morningness and eveningness PRS against several measures of sleep problems during childhood (Difficult sleeping-10, assessed age 10) and adolescence in ALSPAC (Difficult sleeping-13, assessed age 13) (Table 4.2).

Disturbed sleep is often associated with mood disorders; however, the association between sleep phenotypes and mood-disorder related phenotypes in ALPSAC has not yet been investigated. For the depression outcome (age 10-19), the categorical sleep phenotypes tested appeared to be significantly associated with increased depression risk (Table 4.3). There appears to be no significant associations between the primary sleep phenotypes and categorical hypomania (age 22-23).

Table 4.3 Logistic regressions of sleep measures and primary mood outcome measures in ALSPAC

Phenotype	O.R	S.E	p value	95% CI	r^2
Categorical hypomania					
Difficult sleeping-10	1.124	0.478	0.783	0.488/2.587	<0.001
	(1.114)	(0.475)	(0.801)	(0.483/2.569)	(0.003)
Difficult sleeping-13	0.682	0.252	0.3	0.33/1.407	0.002
	(0.336)	(0.294)	(0.212)	(0.061/1.863)	(0.042)
Categorical depression					
Difficult sleeping-10	5.461	3.546	**0.009**	1.529/19.499	0.008
	(5.698)	(3.71)	**(0.008)**	(1.59/20.416)	(0.01)
Difficult sleeping-13	13.284	14.037	**0.014**	1.675/105.377	0.01
	(13.462)	(14.242)	**(0.014)**	(1.692/107.067)	(0.012)

Unadjusted model (adjusted model)

Logistic regressions were carried out to test for association between an increasing chronotype PRS and a designation of "hypomania". Similar regressions were carried out to test for association between chronotype PRS and an individual's depression status. The regressions revealed no significant association (at $p<0.05$) between increased morningness PRS and the primary mood phenotype outcome measures (Table 4.4). However, an increasing eveningness PRS was found to be associated with a decreased risk of hypomania and this remained significant after adjustment for age and sex (Table 4.4). There appeared to be no significant association between depression and eveningness PRS (Table 4.4).

Table 4.4 Regressions of primary mood outcome measures and chronotype PRS in ALSPAC

Phenotype	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Categorical	1.10	0.091	0.234	0.939/1.296	0.001
hypomania	(1.113)	(0.092)	(0.196)	(0.946/1.309)	(0.005)
Categorical	0.982	0.194	0.925	0.666/1.447	0.000
depression	(0.99)	(0.195)	(0.959)	(0.672/1.458)	(0.012)
Eveningness PRS					
Categorical	0.842	0.066	**0.029**	0.721/0.982	0.004
hypomania	(0.839)	(0.066)	**(0.027)**	(0.718/0.98)	(0.009)
Categorical	0.869	0.157	0.439	0.61/1.239	0.002
depression	(0.871)	(0.159)	(0.449)	(0.609/1.245)	(0.014)

Unadjusted model (adjusted model)

Linear regressions were also carried out on the secondary outcome measures; this was used to investigate the association between the chronotype risk scores and continuous hypomanic and depressive scores obtained from the HCL-32 and SMFQ questionnaires, respectively. There were no significant associations with the chronotype PRS identified using either of the mood phenotype outcome measures (Supplemental Table 4.2).

Chronotype preference has been reported to associate with sex; with morningness shown to be more prevalent in females (Hu et al. 2016); the potential association between the chronotype PRS and mood phenotypes was, therefore, investigated splitting the sample by sex (Table 4.5).

Table 4.5 Regressions of primary mood outcome measures and chronotype PRS in ALSPAC split by sex

Female	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Categorical	1.101	0.118	0.369	0.893/1.359	0.001
hypomania	(1.101)	(0.118)	(0.369)	(0.892/1.359)	(0.001)
Categorical	0.918	0.309	0.8	0.475/1.776	0.001
depression	(0.915)	(0.31)	(0.793)	(0.47/1.779)	(0.009)
Eveningness PRS					
Categorical	0.928	0.092	0.448	0.764/1.126	0.001
hypomania	(0.923)	(0.092)	(0.421)	(0.76/1.122)	(0.003)
Categorical	0.957	0.255	0.868	0.567/1.615	0.000
depression	(0.96)	(0.261)	(0.88)	(0.563/1.636)	(0.015)

Male	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Categorical	1.13	0.147	0.348	0.875/1.458	0.002
hypomania	(1.131)	(0.147)	(0.344)	(0.877/1.46)	(0.002)
Categorical	1.024	0.25	0.921	0.635/1.651	0.000
depression	(1.036)	(0.256)	(0.886)	(0.639/1.68)	(0.04)
Eveningness PRS					
Categorical	0.857	0.093	0.157	0.692/1.061	0.003
hypomania	(0.852)	(0.094)	(0.145)	(0.687/1.057)	(0.007)
Categorical	0.857	0.193	0.493	0.551/1.333	0.002
depression	(0.857)	(0.197)	(0.504)	(0.547/1.345)	(0.033)

Unadjusted model (adjusted model)

There were no significant associations between morningness or eveningness PRS and the primary mood phenotypes in males or females (Table 4.5). Using the continuous hypomania and depression scores as a secondary analysis, a nominally significant association was found between greater polygenic loading for morningness and greater hypomania score (coefficient 0.309 (standardized), p 0.049, Supplemental Table 4.3) which is the reverse of what would be expected. No significant associations were identified in males.

4.5.2 UK Biobank cohort

As mentioned previously, over 100,000 participants from the UK Biobank were included in the self-report chronotype GWA studies. The individuals included in these studies were removed before any analysis using the morningness PRS; this resulted in approximately 330,000 observations available for use in further analysis (Table 4.6).

Table 4.6 Summary of mood phenotypes in UK Biobank (N=157,366)

UK Biobank	*Cases*	*Controls*
Lifetime BD, N (%)	1,366 (1.43)	94,234 (98.57)
Lifetime MDD, N (%)	22,671 (28.07)	58,108 (71.93)
Lifetime GAD, N(%)	6,722 (10.15)	59,473 (89.85)
Neuroticism, N (%)	121,919 (51.16)	116,380 (48.84)
Mood Instability, N (%)	129,267 (45.07)	157,524 (54.93)

Within the UK Biobank, individuals who reported a morning chronotype were found to have decreased risk of mood disorder-related phenotypes. Whereas, self-reported evening chronotype was associated with increased risk of all mood disorder phenotypes (Table 4.7). There is clearly a significant association between chronotype and mood disorder-related phenotypes in UK Biobank, however, it is unclear as to whether this association has a genetic underpinning. The derivation of chronotype risk scores can be used to investigate this association in the UK Biobank participants not included in the original GWAS investigation.

Table 4.7 Logistic regressions of mood phenotypes and self-reported chronotype

Phenotype	O.R	S.E	p value	95% CI	r^2
Morning chronotype (vs. evening)					
Lifetime BD	0.742	0.065	**$6.6x10^{-4}$**	0.625/0.881	0.023
Lifetime MDD	0.63	0.019	**$1.6x10^{-54}$**	0.594/0.667	0.046
Lifetime GAD	0.649	0.03	**$2.3x10^{-21}$**	0.594/0.71	0.037
Mood Instability	0.739	0.011	**$3.2x10^{-95}$**	0.718/0.761	0.018
	Coefficient				
Neuroticism	-0.709	0.023	**$6x10^{-207}$**	-0.754/-0.664	0.04
Evening chronotype (vs. morning)					
Lifetime BD	1.348	0.118	**$6.6x10^{-4}$**	1.135/1.6	0.023
Lifetime MDD	1.588	0.047	**$1.6x10^{-54}$**	1.498/1.684	0.046
Lifetime GAD	1.54	0.07	**$2.3x10^{-21}$**	1.408/1.683	0.037
Mood Instability	1.353	0.02	**$3.2x10^{-95}$**	1.315/1.392	0.018
	Coefficient				
Neuroticism	0.709	0.023	**$6x10^{-207}$**	0.664/0.754	0.04

Adjusted model

There was a clear phenotypic association between self-reported chronotype preference and mood disorder-related phenotypes.

Before testing the chronotype PRS for associations to mood phenotypes, regressions – adjusted for age, sex and deprivation – were used to investigate the potential association between the chrontype PRS and self-reported chronotype preference. Many of the SNPs incorporated in the chronotype PRS were discovered by investigating self-chronotype in the UK Biobank sample. Here, the chronotype PRS were tested against the self-reported chronotypes of the remaining Biobank cohort – after removal of those involved in the GWA studies – to ensure these SNPs do indeed associate with self-reported chronotype preference.

The morningness PRS was found to be associated with self-reported morningness; both greater definite morningness and overall morningness were significantly associated with an increased PRS (Table 4.8). The morningness PRS was also found to negatively associate with self-reported eveningness (Supplemental Table 4.4). The eveningness PRS appears to be positively associated with self-

reported eveningness significantly, however, the effect size is relatively small (O.R 1.069) (Table 4.8).

Table 4.8 Logistic regressions of self-reported chronotype and polygenic risk scores

Phenotype	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Definite morningness	1.162	0.009	**$3.8x10^{-79}$**	1.144/1.180	0.02
Overall morningness	1.10	0.005	**$2x10^{-117}$**	1.091/1.109	0.008
Eveningness PRS					
Definite eveningness	1.078	0.009	**$4.4x10^{-21}$**	1.061/1.095	0.018
Overall eveningness	1.049	0.004	**$5.5x10^{-31}$**	1.04/1.057	0.007

Adjusted model

The risk scores were then used to investigate the association between genetic loading for chronotype and mood-disorder related phenotypes. Regressions were carried out to test the morningness and eveningness PRS for associations to lifetime BD, lifetime MDD, lifetime GAD, neuroticism and mood instability. The logistic regressions were also adjusted for age, sex and townsend score.

Table 4.9 Logistic regressions of mood phenotypes and chronotype PRS

Phenotype	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Lifetime BD	0.947	0.031	0.102	0.887/1.011	0.000
	(0.946)	(0.032)	(0.097)	(0.886/1.01)	(0.018)
Lifetime MDD	0.988	0.009	0.218	0.97/1.007	0.000
	(0.988)	(0.01)	(0.2)	(0.949/1.007)	(0.043)
Lifetime GAD	0.984	0.015	0.287	0.954/1.014	0.000
	(0.985)	(0.015)	(0.325)	(0.955/1.015)	(0.031)
Mood	0.997	0.004	0.504	0.988/1.006	0.000
Instability	(0.998)	(0.005)	(0.68)	(0.989/1.007)	(0.017)
	Coefficient				
Neuroticism	-0.007	0.006	0.273	-0.019/0.005	0.000
	(-0.001)	(0.007)	(0.859)	(-0.014/0.012)	(0.038)
Eveningness PRS					
Lifetime BD	1.069	0.03	**0.019**	1.011/1.13	0.000
	(1.069)	(0.03)	**(0.02)**	(1.011/1.13)	(0.016)
Lifetime MDD	1.000	0.008	0.987	0.985/1.016	0.000
	(0.999)	(0.008)	(0.943)	(0.983/1.016)	(0.044)
Lifetime GAD	0.983	0.013	0.182	0.958/1.008	0.000
	(0.982)	(0.013)	(0.159)	(0.957/1.007)	(0.033)
Mood	1.000	0.004	0.938	0.993/1.008	0.000
Instability	(1.00)	(0.004)	(0.979)	(0.992/1.007)	(0.016)
	Coefficient				
Neuroticism	-0.003	0.006	0.598	-0.016/0.009	0.000
	(-0.003)	(0.006)	(0.671)	(-0.016/0.01)	(0.38)

Unadjusted model (adjusted model)

The morningness PRS showed no significant association with the mood phenotypes tested ($p>0.05$), however, the eveningness PRS was found to be significantly associated with an increased risk of lifetime BD both with and without adjustment (O.R 1.069, p 0.019). It is of note that a relatively small effect size is seen, therefore, the increase in risk is relatively low (Table 4.9).

As mentioned previously, associations have been found between chronotype preference and sex, with slight differences in self-reported chronotype also seen

in the sample of the UK Biobank cohort used in these analyses (56% females report a definite morning chronotype). As with the analyses done in ALSPAC, the relationship between chronotype PRS and mood phenotypes was investigated separately in females and males due to a priori rationale.

Table 4.10 Logistic regressions of mood phenotypes and chronotype PRS separated by sex

Female	**O.R**	**S.E**	**p value**	**95% CI**	**r^2**
Morningness PRS					
Lifetime BD	0.965	0.044	0.428	0.882/1.054	0.000
	(0.961)	(0.044)	(0.386)	(0.879)	(0.016)
Lifetime MDD	0.984	0.012	0.166	0.961/1.007	0.000
	(0.981)	(0.012)	(0.116)	(0.958/1.005)	(0.018)
Lifetime GAD	0.98	0.019	0.294	0.943/1.018	0.000
	(0.979	(0.019)	(0.273)	(0.942/1.017)	(0.018)
Mood	1.00	0.004	0.848	0.992/1.01	0.000
Instability	(0.993)	(0.005)	(0.157)	(0.983/1.003)	(0.018)
	Coefficient				
Neuroticism	-0.018	0.009	**0.035**	-0.035/-0.001	0.000
	(-0.012)	(0.009)	(0.177)	(-0.03/0.006)	(0.013)
Eveningness PRS					
Lifetime BD	1.086	0.042	**0.034**	1.006/1.172	0.000
	(1.087)	(0.042)	**(0.031)**	(1.008/1.174)	(0.015)
Lifetime MDD	0.997	0.01	0.766	0.977/1.017	0.000
	(0.997)	(0.01)	(0.752)	(0.958/1.017)	(0.019)
Lifetime GAD	0.975	0.016	0.118	0.945/1.006	0.000
	(0.974)	(0.016)	(0.105)	(0.944/1.005)	(0.02)
Mood	1.009	0.005	0.065	0.999/1.018	0.000
Instability	(1.005)	(0.005)	(0.326)	(0.995/1.015)	(0.018)
	Coefficient				
Neuroticism	0.014	0.008	0.109	-0.003/0.03	0.000
	(0.012)	(0.009)	(0.19)	(-0.006/0.029)	(0.013)

Unadjusted model (adjusted model)

Male	O.R	S.E	p value	95% CI	r^2
Morningness PRS					
Lifetime BD	0.928	0.045	0.122	0.843/1.02	0.001
	(0.929)	(0.0450	(0.133)	(0.844/1.023)	(0.019)
Lifetime MDD	0.999	0.016	0.947	0.968/1.031	0.000
	(1.00)	(0.016)	(0.988)	(0.968/1.032)	(0.026)
Lifetime GAD	0.992	0.026	0.748	0.942/1.044	0.000
	(0.996)	(0.026)	(0.882)	(0.946/1.049)	(0.021)
Mood	1.001	0.005	0.78	0.991/1.012	0.000
Instability	(0.995)	(0.006)	(0.344)	(0.984/1.006)	(0.013)
	Coefficient				
Neuroticism	0.009	0.009	0.328	-0.009/0.027	0.000
	(0.012)	(0.01)	(0.217)	(-0.007/0.031)	(0.02)
Eveningness PRS					
Lifetime BD	1.049	0.044	0.253	0.967/1.138	0.000
	(1.047)	(0.044)	(0.267)	(0.965/1.136)	(0.017)
Lifetime MDD	1.005	0.014	0.722	0.978/1.032	0.000
	(1.004)	(0.014)	(0.76)	(0.977/1.032)	(0.025)
Lifetime GAD	0.998	0.022	0.927	0.955/1.043	0.000
	(0.996)	(0.023)	(0.873)	(0.953/1.041)	(0.022)
Mood	0.995	0.005	0.351	0.985/1.005	0.000
Instability	(0.993)	(0.006)	(0.227)	(0.982/1.004)	(0.013)
	Coefficient				
Neuroticism	-0.02	0.009	**0.029**	-0.038/-0.002	0.000
	(-0.02)	(0.01)	**(0.037)**	(-0.038/-0.001)	(0.019)

Unadjusted model (adjusted model)

When investigating chronotype PRS and mood phenotypes in males there appeared to be no significant associations found with the exception of neuroticism score. Contrary to the phenotypic associations, greater polygenic loading for eveningness was associated with lower neuroticism score in males. In females, eveningness PRS was found to be significantly associated with an increased risk of lifetime BD and this finding remained significant after adjustment. As with the significant association between eveningness PRS and BD seen in the whole sample, the effect size seen in females was relatively small (namely O.R 0.96-1.09) (Table 4.10). An association was also found between

morningness PRS and lower neuroticism score in females, however, this association did not remain significant after adjustment for confounders.

As there appear to be differences in the PRS findings in males and females, the analyses were repeated to establish whether there was an interaction between sex and chronotype PRS. When testing for an interaction between sex and both the morningness and eveningness PRS, there appeared to be no significant associations with BD, MDD, GAD or mood instability. However, there was a significant association with decreased neuroticism score when investigating the interaction between sex and eveningness PRS (coefficient -0.032, p 0.016, Supplemental Table 4.5) which lines up with the finding of decreased neuroticism score and evening PRS in males.

As the primary sex stratified analyses in ALSPAC showed no significant results, and was likely to be underpowered, an interaction test was also likely to be underpowered and was not undertaken in ALSPAC.

4.5.3 Morningness and Depression MR

Table 4.11 Two sample Mendelian Randomisation

Two sample MR	MREgger/ IVW Beta	MREgger/ IVW S.E	p value	95% CI	igx^2
Depression on morningness					
Slope	-0.24	0.029	**7.31x10^{-10}**	-0.3/-0.18	0.969
Intercept	0.002	0.001	0.31	-0.001/0.005	0.969
Morningness on depression					
Slope	0.031	0.013	**0.024**	0.06/2.27	0.966
Intercept	-0.002	0.0003	2.14x10^{-8}	-0.003/-0.001	0.966

To investigate potential causal relationships between morningness and depression, a two sample MR was performed. As relationships between circadian rhythms and mood disorders have been hypothesised to be bidirectional (Hidalgo et al. 2009), MR was used to test the influence of depression on morningness and the effect of morningness on depression (Table 4.11). There was evidence of depression conferring a significantly lower propensity for morningness (IVW beta 0.24, IVW p 7.31x10^{-10}). However, when investigating the influence of morningness on depression there was evidence of morningness increasing the risk of depression (IVW beta 0.03, IVW p 0.024) which is contradictory to the

phenotypic analyses above (Table 4.7) and findings from other studies which report morningness conferring a reduced risk of depression (Jones et al. 2019). It is of note that due to the significant intercept (p $2.14x10^{-8}$) this finding may be due to pleiotropy in this direction. There may also be bias introduced here by weak instruments, the overall genetic effect of morningness on depression using meta-analysis of SNP WR was not significant (p 0.71, Supplemental Table 4.6) suggesting weak instruments bias in this analysis.

4.6 Discussion

An evening chronotype has been suggested as a pre-existing risk factor for BD (Alloy et al. 2017); therefore, it seems logical to hypothesise that polygenic loading for morningness may be protective against mood disorders. The aim of this study was to investigate potential associations between an individual's chronotype PR scores and mood disorder-related phenotypes within ALSPAC and UK Biobank.

Within ALSPAC, the chronotype PRS were not associated with sleep phenotypes. It is important to note that these associations were tested with relatively small sample sizes (N=978) and were likely underpowered to detect small effects (Supplemental Table 4.1). The sleep phenotypes tested in ALSPAC were associated with the categorical measure of depression with relatively large effect sizes; however, the effect sizes seen for the associations between sleep phenotypes and mood were relatively small and did not meet nominal statistical significance.

When investigating the association between chronotype PRS and mood the only significant finding was the negative association between eveningness PRS and binary hypomania (as PRS increased; risk of hypomania decreased). This effect is in the opposite direction as to what was expected and as to what was found in UK Biobank. This raises the possibility that an increasing eveningness PRS is protective against hypomania in adolescence and early adulthood; although, as the regressions were potentially underpowered to detect small effects, no definitive conclusions can be drawn on this point.

In the UK Biobank, the chronotype PRS were significantly associated with self-reported chronotype preference; increasing morningness PRS was positively associated with morning preference and negatively associated with evening preference. The opposite was true for an increasing eveningness PRS (in the expected direction).

Overall, chronotype preference was also significantly associated with all of the mood phenotypes of interest; as expected being a "morning person" decreased the risk of several mood phenotypes, whereas, being an "evening person" was associated with an increased risk of all mood disorder phenotypes. As the genetic loading for chronotype was significantly associated with self-reported

chronotype preference and the chronotype preference was significantly associated with mood disorder phenotypes, there was the potential that the chronotype PRS was also associated with these mood phenotypes. However, there appeared to be no association to these mood phenotypes when investigating the morningness PRS. An increasing eveningness PRS was however found to be associated with a small but significant increase in the risk of lifetime BD in the UK Biobank sample.

Investigating the association between chronotype PRS and mood in ALSPAC showed only a significant finding was the negative association between eveningness PRS and binary hypomania. This effect is in the opposite direction as to what was expected and as to what was found in UK Biobank. This raises the intriguing possibility that an increasing eveningness PRS is protective against hypomania in adolescence and early adulthood; although, as the regressions were underpowered to detect small effects, no definitive conclusions can be drawn on this point.

Upon stratifying this sample by sex, this association between eveningness PRS and lifetime BD was seen in females did not reach statistical significance in males. The genetic loading for evening preference appears to influence the risk of BD and has an effect in females and not males, however, the effect sizes seen are small and eveningness PRS seems to have a relatively small effect on the risk of BD. The variants associated with eveningness may have some small direct or indirect influence on the pathophysiology of BD. When investigated this in ALSPAC, a reverse association was seen in females: greater morningness PRS was associated with greater hypomania score in females. Again, there is potential that chronotype PRS has different effects throughout an individual's life course. Also, as eveningness PRS appears to be protective against neuroticism in males and there is evidence of an interaction between sex and eveningness PRS; this further suggests a relationship between sex and the underlying genetics of chronotype preference. However, these findings are preliminary and in need of replication.

It is still difficult to determine causality; however, there does appear to be a significant causal link between chronotype and mood disorders (in this case, specifically morningness and depression). Due to the conflicting evidence

between the findings of the above MR and recent studies of morningness and mood disorders the direction of causality is unclear (Jones et al. 2019).

4.7 Limitations

In testing the above hypotheses there are several limitations.

1. The use of self-reported outcome measures for all analyses carried out; the outcome measure may be subject to reporting bias. However, the measures of HCL-32 to investigate hypomania (Smith et al. 2015) and the SMFQ to measure depressive symptoms (Stochl et al. 2015) in ALSPAC have been implemented in previous studies. The use of the lifetime mood disorder variables derived from the MHQ in UK Biobank should reduce this self-report bias.

2. A limitation of the investigations in ALSPAC is a lack of statistical power. There were a smaller number of observations in ALSPAC relative to UK Biobank for the desired mood and sleep phenotypes. Particularly in the case of the sleep phenotype data available in ALSPAC there was a very small sample available for analysis.

3. The UK Biobank sample may also be underpowered to detect small effects as individuals with psychiatric and mood disorders appear to be underrepresented in this cohort (Fry et al. 2017).

4. There are also limitations to the MR analysis: there may be biases introduced by weak instruments and pleiotropy. These findings require replication. Strengthening the genetic instruments used in the analysis using more variants associated with depression and chronotype identified through GWAS may help to give more evidence of the causal relationships between these traits.

5. It is of note, PRS analyses show only small effects on the traits-of-interest and explain very little of the trait variance. A more effective PRS is needed to explain a greater proportion of the variance seen.

4.8 Future Work

Upon completion of this analysis, there were more GWAS of chronotype preference released which could be used to expand the chronotype risk scores used here. The extension of the chronotype PRS may allow the estimations of effect on mood disorder to explain more of the phenotypic variance. This may also make the PRS more accurate for distinguishing affected individuals in a clinical setting. A more extensive PRS could be used to further investigate the potential interaction chronotype preference has with sex.

However, as many of the GWAS which have been conducted to identify variants associated with chronotype include UK Biobank participants it may result in confounding of the findings. For future investigations using such risk scores would require a separate sample cohort.

The collection of more participant-reported and objective sleep phenotypes measures in ALSPAC could provide a greater understanding of the influence these PRS have in young adulthood.

Also, there has been strong evidence for the involvement of circadian rhythm in mood disorders, such as BD and MDD, and it may be of benefit to investigate a variety of circadian phenotypes or sleep symptoms in the context of mood disorder related phenotypes to help elucidate a greater understanding of the pathophysiology of BD.

4.9 Conclusions

This is the first use of a chronotype PRS investigating mood disorders. The relationship between mood disorders and circadian behaviours has been widely suggested (Merikanto et al. 2013), however, the influence of genetics on this relationship has been unclear. These findings strengthen the evidence for the relationship between circadian rhythms and mood disorders, and related traits. This may help highlight direct or indirect pathways which influence the development and pathophysiology of mood disorders. However, these analyses are based on a subjective circadian trait and further investigations are required to understand the relationships between other circadian rhythm traits and mood disorders.

4.10 Supplementary Data

Supplemental Table 4.1 Cross tabulation of observations of ALSPAC mood outcomes and sleep phenotypes

ALSPAC	**Binary hypomania**		**Binary depression**	
Difficult sleeping-10	*Yes*	*No*	*Yes*	*No*
Yes	11	134	12	345
No	13	178	3	471
Difficult sleeping-13	*Yes*	*No*	*Yes**	*No**
Yes	12	146	9	313
No	24	199	1	462

* excluding individuals of non-white ethnic background resulted in no observations for the desired variables. Individuals were not excluded based on ancestry for this phenotypic association test.

Supplemental Table 4.2 Linear regression of chronotype PRS and continuous hypomania and depression scores

Phenotype	**Coefficient**	**S.E**	**p value**	**95% CI**	**r2**
Morningness PRS					
HCL score	0.069	0.128	0.589	-0.182/0.32	0.000
	(0.084)	(0.128)	(0.51)	(-0.166/0.335)	(0.003)
SMFQ score	-0.001	0.042	0.975	-0.084/0.081	0.000
	(-0.0003)	(0.042)	(0.995)	(-0.083/0.082)	(0.001)
Eveningness PRS					
HCL score	0.013	0.121	0.915	-0.224/0.25	0.000
	(0.01)	(0.121)	(0.931)	(-0.226/0.247)	(0.003)
SMFQ score	-0.046	0.041	0.265	-0.127/0.035	0.002
	(-0.046)	(0.041)	(0.269)	(-0.127/0.035)	(0.014)

Unadjusted model (adjusted model)

Supplemental Table 4.3 Linear regression of chronotype PRS and continuous hypomania and depression scores split by sex

Female	Coefficient	S.E	p value	95% CI	r2
Morningness PRS					
HCL score	0.309	0.157	**0.049**	0.001/0.616	0.003
	(0.309)	(0.157)	**(0.049)**	(0.001/0.616)	(0.003)
SMFQ score	0.003	0.058	0.962	-0.111/0.116	0.000
	(0.002)	(0.058)	(0.97)	(-0.111/0.116)	(0.001)
Eveningness PRS					
HCL score	0.006	0.144	0.967	-0.276/0.288	0.000
	(0.003)	(0.144)	(0.982)	(-0.279/0.285)	(0.000)
SMFQ score	-0.035	0.052	0.501	-0.138/0.068	0.000
	(-0.035)	(0.052)	(0.51)	(-0.137/0.068)	(0.001)

Male	Coefficient	S.E	p value	95% CI	r2
Morningness PRS					
HCL score	-0.355	0.22	0.107	-0.787/0.077	0.003
	(-0.355)	(0.22)	(0.108)	(-0.788/0.078)	(0.003)
SMFQ score	-0.002	0.061	0.974	-0.121/0.117	0.000
	(-0.002)	(0.061)	(0.971)	(-0.122/0.117)	(0.000)
Eveningness PRS					
HCL score	-0.099	0.176	0.573	-0.444/0.246	0.000
	(-0.107)	(0.176)	(0.543)	(-0.453/0.0.238)	(0.002)
SMFQ score	-0.069	0.055	0.215	-0.177/0.04	0.001
	(-0.071)	(0.055)	(0.2)	(-0.18/0.038)	(0.001)

Unadjusted model (adjusted model)

Supplemental Table 4.4 Logistic regressions of self-reported eveningness and morningness PRS in UK Biobank

Phenotype	O.R	S.E	p value	95% CI	r2
Definite eveningness	0.922	0.007	**1.6x10-24**	0.908/0.937	0.018
Overall eveningness	0.949	0.004	**4.1x10-37**	0.941/0.956	0.071

Adjusted model

Supplemental Table 4.5 Logistic regression investigating sex interaction with PRS in UK Biobank

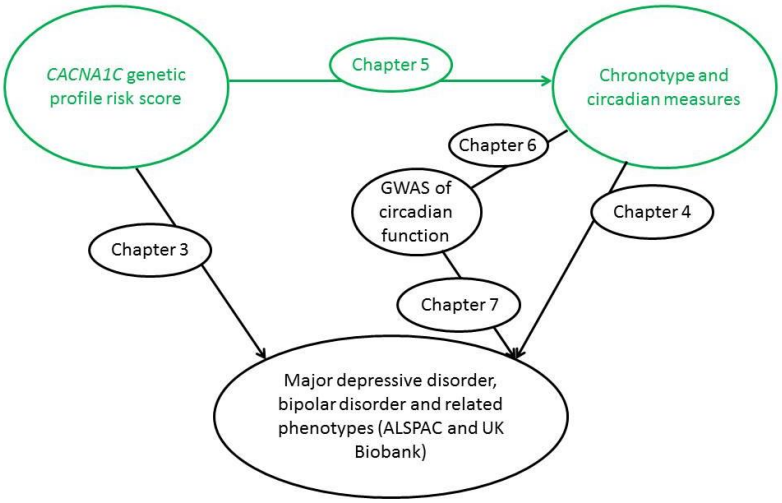
	O.R	S.E	p value	95% CI	r2
Morningness PRS					
Lifetime BD	1.006	0.056	0.92	0.901/1.122	0.016
Lifetime MDD	1.007	0.028	0.673	0.974/1.041	0.043
Lifetime GAD	1.005	0.028	0.861	0.952/1.061	0.033
Mood Instability	1.002	0.008	0.8	0.987/1.017	0.017
	Coefficient				
Neuroticism	0.024	0.013	0.07	-0.002/0.05	0.038
Eveningness PRS					
Lifetime BD	0.96	0.053	0.466	0.861/1.071	0.016
Lifetime MDD	1.003	0.017	0.881	0.97/1.036	0.043
Lifetime GAD	1.016	0.028	0.567	0.963/1.071	0.033
Mood Instability	0.988	0.008	0.112	0.973/1.003	0.017
	Coefficient				
Neuroticism	-0.032	0.013	**0.016**	-0.057/-0.006	0.038

Adjusted model

Supplemental Table 4.6 Two sample MR overall genetic effects using meta-analysed SNP WR

Overall genetic effects	Log odds	S.E	p value	p chi
Depression to morningness	-0.224	0.014	**$3.01x10^{-54}$**	**$1.1x10^{-4}$**
Morningness to depression	-0.009	0.025	0.709	1

Chapter 5 The potential influence of *CACNA1C* GPRS on circadian rhythm phenotypes



Outline 5.1 Chapter in context of overall study

5.1 Introduction

As described in Chapter 3, there have been multiple SNPs within the *CACNA1C* loci found to have an association with several psychiatric conditions including BD, schizophrenia and MDD. *CACNA1C* has been identified as a risk locus by both GWAS and subsequent meta-analyses (Bigos et al. 2010; Sklar et al. 2011; Smoller 2013; Heilbronner et al. 2015; Kabir et al. 2017). However, it is still unclear as to the exact mechanism by which *CACNA1C* genetic variants influence the pathophysiology of BD.

The *CACNA1C* gene is known to have a role in circadian rhythm (Schmutz et al. 2015; McCarthy et al. 2016) and disruptions of circadian rhythm are frequently reported in BD (Hayashi et al. 2015; Steinan et al. 2015). In Chapter 3, *CACNA1C* GPRS were used to investigate the potential associations between *CACNA1C* risk variants and mood disorders phenotypes. Upon replication, there was little evidence for an association between the GPRS and mood disorders, particularly in UK Biobank. However, it is possible that *CACNA1C* may be exacting influence on mood disorders through disrupted circadian rhythms. As demonstrated in Chapter 4, there is phenotypic association between chronotype (a subjective measure of circadian rhythmicity and rest-activity preference) and mood disorder-related traits in both ALSPAC and UK Biobank. The aim of this chapter is to determine whether there is an association between *CACNA1C* GPRS and circadian rhythm phenotypes, specifically chronotype preference and relative amplitude (an objective measure of rest-activity cycles).

5.1.1 *CACNA1C* in circadian rhythms

The *CACNA1C* gene function is influenced by the circadian clock, as well as having a role in rhythmic calcium signalling (Schmutz et al. 2015; McCarthy et al. 2016). The $Ca_V1.2$ protein is expressed rhythmically, with its peak occurring late at night, and contributes to the induction of clock genes. Knockout of the $Ca_V1.2$ channels in mouse models has shown a role in circadian clock resetting. The influx of calcium ions via the $Ca_V1.2$ channels activate second messenger systems leading to the transcription of circadian clock genes (Schmutz et al. 2015). Greater detail on the functions of the $Ca_V1.2$ channel and influence of the *CACNA1C* gene is reported in Chapter 3.

CACNA1C polymorphic variants associated with psychiatric and mood disorders have been reported to associate with increased sleep latency in both children and adults; sleep latency refers to the time taken to transition from wakefulness to sleep (Kantojarvi et al. 2017); and it has been suggested that *CACNA1C* partially modulates the electrophysiology of sleep (Kantojarvi et al. 2017).

5.1.2 Circadian rhythms in mood disorders

As described in chapters 1 and 4, disrupted circadian rhythmicity are commonly observed in mood disorders, including higher sleep disturbances, longer circadian phase and greater evening chronotype preference compared to healthy individuals (Pagani et al. 2016; Dmitrzak-Węglarz et al. 2016; Asaad et al. 2016). Individuals with, or at risk of, mood disorders (particularly BD) appear to be more susceptible to disruptions of circadian rhythmicity and sleep-wake cycle disturbances when compared to healthy controls (Geoffroy et al. 2014; Moon et al. 2016). Chronotype has been reported to associate with mood disorders and mood-related traits, in a previous chapter and in several other studies, with evening chronotype reported more frequently than in healthy controls (Corruble et al. 2014; Dmitrzak-Węglarz et al. 2016; Pagani et al. 2016; Chan et al. 2014; Berdynaj et al. 2016; Hidalgo et al. 2009; Baek et al. 2016; Alloy et al. 2017). In a recent study, within the UK Biobank, objective measures of rest-activity rhythmicity from accelerometer data were derived (mainly RA) (Lyall et al. 2018). It was found that low RA, a measure showing little average difference

between an individual's rest and activity over the course of a week, was associated with several mood disorder phenotypes (Lyall et al. 2018).

5.2 Rationale

As disruptions of sleep homeostasis and circadian rhythm systems are frequently associated with mood disorders, it is logical to hypothesise that *CACNA1C* may influence BD and MDD through its circadian effects. There are subjective measures of sleep phenotypes available in ALPSAC which could be used to investigate this. Also, chronotype and RA are measures of sleep-wake preference and rest-activity, respectively, in UK Biobank and may reflect altered sleep homeostasis and circadian rhythmicity. If *CACNA1C* risk scores show associations with these phenotypes, which have already been linked to mood disorders (Corruble et al. 2014; Dmitrzak-Węglarz et al. 2016; Pagani et al. 2016; Chan et al. 2014; Berdynaj et al. 2016; Hidalgo et al. 2009; Baek et al. 2016; Alloy et al. 2017; Lyall et al. 2018), it may be beneficial to consider an individual's *CACNA1C* genotype when developing treatment interventions. Breast and ovarian cancer risk and treatment is often informed by *BRCA1*/*BRCA2* genotyping (Tung & Garber 2018), a more extensive genotyping of *CACNA1C* could eventually be used in a similar way to inform mood disorder treatment.

5.3 Hypothesis to be tested

The aim of this chapter is to investigate whether a higher genetic loading for *CACNA1C* risk variants is associated with both subjective and objective measures of circadian rhythmicity within ALSPAC and UK Biobank. These findings could, potentially, strengthen the evidence of *CACNA1C*'s involvement in sleep homeostasis and circadian rhythmicity; as well as highlight a possible pathway by which variation in the *CACNA1C* gene could be influencing mood disorder pathophysiology.

5.4 Methods

A weighted GPRS was generated for all individuals with genotype data available in ALSPAC and UK Biobank using Plink 1.9 (Purcell et al. 2007); *CACNA1C* SNP weightings were obtained from several BD GWAS and meta-analyses (chapter 3)(Kloiber et al. 2012; Green et al. 2013; Fiorentino et al. 2014; Heilbronner et al. 2015). Chapters 2 and 3 describe *CACNA1C* GPRS/PRS methodology. Weighted scores were standardised to z values (i.e. per SD). Statistical analysis described in Chapter 2.

5.4.1 ALSPAC

Participants of ALSPAC who were genotyped were only included in the ALSPAC genotype database after meeting particular quality control criteria. Details of quality control measures, imputation and phasing described in Chapter 2. Only unrelated individuals were included in these analyses in an attempt to prevent shared environmental factors influencing associations (Dudbridge 2016); a list of unrelated individuals (N=8,197; 96% of the genotyped sample) was provided by ALSPAC. Individuals that were not recorded as Caucasian, or those with missing ethnic background data (N=1,112; 13.08%), were also removed from analyses leaving N=7,390.

5.4.1.1 Sleep phenotypes

There is no measure of chronotype within ALSPAC and so the *CACNA1C* GPRS was tested against four measures of sleep problems reported at two different ages. This resulted in four sleep-related variables; these variables were generated from mother-answered questionnaires regarding child sleep habits. These phenotypes are detailed in Chapter 2.

The primary outcome measures for sleep in ALSPAC were the categorical variables “Difficulty sleeping-10” and “Difficulty sleeping-13” (ALSPAC variable kv7034 and tb7034, respectively) which were tested for their association to *CACNA1C* GPRS using logistic regression. Secondary outcome measures were also tested (“Difficulty sleeping-scale” (tb5538) and “Worried sleep” (tb6555)) using linear regression.

5.4.2 UK Biobank cohort

Before any analyses were undertaken with UK Biobank data individuals were removed from the initial cohort of N=152,000 based upon several quality control criteria, including relatedness, ancestry (Non-Caucasian individuals), gender mismatch and quality control failure in the UK BiLEVE study. Chapter 2 describes the quality control and exclusion criteria. This left N=119,953 (78.9% of the cohort). As with the analyses undertaken in Chapter 3, individuals missing genotype information for any of the 15 chosen SNPs were also excluded from analyses leaving N=95,073.

5.4.2.1 Chronotype phenotype

As described in Chapter 2, chronotype was derived from the participants' responses to a question obtained from the Morningness-Eveningness questionnaire (Taillard et al. 2003). This question is an accepted measure of chronotype and has been reported to explain the greatest variance in preference of sleep-wake timings (Taillard et al. 2003). Categorical variables were then generated based upon the responses given, resulting in the generation of four separate chronotype variables. The primary outcome measures used for analysis were the "definite morning" and "definite evening" variables with "overall morning" and "overall evening" providing secondary outcome measures.

5.4.2.2 Relative amplitude phenotype

Accelerometer data were obtained from 103,720 UK Biobank participants. This data underwent extensive pre-processing and exclusion criteria in order to derive a RA variable. RA is defined as the relative difference between the most active continuous 10-hour period and the least active continuous 5-hour period in an average 24-hour period; or each individual, the RA data point was the mean RA value across seven days. RA ranges from 0 to 1, with higher values indicating greater distinction between activity levels during the most and least active periods of the day.

Chapters 2 and 6 detail the pre-processing, exclusions and generation of the RA variable.

5.5 Results

5.5.1 ALSPAC

It is unclear as to whether *CACNA1C* has any influence on sleep features. Logistic regression was used to test the associations between *CACNA1C* GPRS and subject sleep phenotypes in ALSPAC.

Table 5.1 Logistic regressions of sleep outcome measures and *CACNA1C* GPRS

Phenotype	O.R	S.E	p value	95% CI	r^2
Difficult sleeping-	1.055	0.07	0.423	0.926/1.202	0.001
10	(1.054)	(0.07)	(0.428)	(0.925/1.201)	(0.002)
Difficult sleeping-	0.997	0.066	0.964	0.876/1.135	0.000
13	(0.999)	(0.066)	(0.984)	(0.877/1.137)	(0.001)

Unadjusted model (adjusted model)

There were no associations found between the GPRS and difficulty sleeping (Table 5.1). Compared to the GPRS analyses undertaken in Chapter 3 there are far fewer observations available for this analyses (N= 978, maximum sample size in above analysis) and is therefore likely to be underpowered.

Consistent with the analyses undertaken in Chapter 3, the sample was stratified by sex in order to investigate a potential sex effect between *CACNA1C* GPRS and sleep phenotypes. The above regressions were repeated separately in females and males.

Table 5.2 Logistic regressions of sleep outcome measures and *CACNA1C* GPRS

Phenotype	O.R	S.E	p value	95% CI	r^2
Female					
Difficult sleeping-	0.996	0.086	0.964	0.84/1.181	0.000
10	(0.997)	(0.086)	(0.971)	(0.841/1.182)	(0.000)
Difficult sleeping-	0.983	0.084	0.837	0.832/1.161	0.000
13	(0.981)	(0.083)	(0.818)	(0.83/1.159)	(0.001)
Male					
Difficult sleeping-	1.139	0.119	0.21	0.929/1.397	0.003
10	(1.143)	(0.119)	(0.201)	(0.932/1.402)	(0.004)
Difficult sleeping-	1.019	0.107	0.855	0.83/1.251	0.000
13	(1.027)	(0.108)	(0.798)	(0.836/1.263)	(0.001)

Unadjusted model (adjusted model)

No evidence of associations was identified in females or males. However, the analyses including only males appears to show a greater (still small) effect on difficultly sleeping at age 10 and explain a slightly higher proportion of variance than in females (Table 5.2). This association is not significant, although it is of note that only 428 males had both genotype and sleep phenotype data; it is likely this analysis is underpowered and there could be a potential link between *CACNA1C* GPRS and sleep phenotypes in young males. Of course this is speculation based on these preliminary findings.

5.5.2 UK Biobank

To investigate the effect of *CACNA1C* GPRS on chronotype preference logistic regression assuming an additive allelic model was used. The analyses were also adjusted for confounders for age, sex, deprivation, principal component and other quality control measures. “Definite morning” and “definite evening” were the primary outcome measures for this analysis. The regressions yielded no significant effects on the primary or secondary measurement outcomes using the standardized GPRS (Table 5.3).

Table 5.3 *CACNA1C* GPRS vs chronotype

Outcome	O.R	S.E	p value	95% CI	r^2
Definite	1.005	0.013	0.695	0.98/1.03	<0.0001
morning	(1.003)	(0.013)	(0.797)	(0.978/1.029)	(0.0205)
Definite	0.995	0.013	0.695	0.97/1.02	<0.0001
evening	(0.997)	(0.129)	(0.797)	(0.972/1.022)	(0.0205)
Overall	1.005	0.007	0.474	0.991/1.019	<0.0001
morning	(1.004)	(0.007)	(0.533)	(0.991/1.018)	(0.008)
Overall	0.995	0.007	0.474	0.982/1.009	<0.0001
evening	(0.996)	(0.007)	(0.533)	(0.982/1.009)	(0.008)

Unadjusted model (adjusted model)

As above, the regressions were also conducted in females and males separately.

Table 5.4 *CACNA1C* GPRS vs chronotype after splitting sample by sex

Outcome	O.R	S.E	p value	95% CI	r^2
Female					
Definite	0.989	0.017	0.527	0.956/1.024	<0.0001
morning	(0.989)	(0.018)	(0.532)	(0.955/1.024)	(0.0185)
Definite	1.011	0.018	0.527	0.977/1.047	<0.0001
evening	(1.011)	(0.018)	(0.532)	(0.977/1.047)	(0.0185)
Overall	1.002	0.009	0.859	0.983/1.02	<0.0001
morning	(1.002)	(0.009)	(0.838)	(0.984/1.021)	(0.007)
Overall	0.998	0.009	0.859	0.98/1.017	<0.0001
evening	(0.998)	(0.009)	(0.838)	(0.98/1.017)	(0.007)
Male					
Definite	1.021	0.019	0.256	0.985/1.059	0.0001
morning	(1.02)	(0.019)	(0.289)	(0.983/1.059)	(0.0233)
Definite	0.979	0.018	0.256	0.944/1.015	0.0001
evening	(0.98)	(0.018)	(0.289)	(0.945/1.017)	(0.0233)
Overall	1.008	0.01	0.404	0.989/1.029	<0.0001
morning	(1.007)	(0.01)	(0.469)	(0.987/1.028)	(0.0097)
Overall	0.992	0.01	0.404	0.972/1.011	<0.0001
evening	(0.993)	(0.01)	(0.469)	(0.973/1.013)	(0.0097)

Unadjusted model (adjusted model)

When stratifying the sample by sex there were no associations found between an increasing GPRS and a specific chronotype preference. There appeared to be a greater effect in males who reported a definite chronotype, however, this finding was not significant (Table 5.4).

The above analyses are based on self-reported and subjective circadian measures. As described previously, objective measures of rest-activity cycles has been derived from UK Biobank accelerometer data. Linear regressions were conducted to determine whether there is an association between *CACNA1C* GPRS and the objective rest-activity measure RA.

Table 5.5 Linear regression of RA and *CACNA1C* GPRS

Outcome	Beta	S.E	p value	95% CI	r^2
RA	-0.001	0.018	0.973	-0.036/0.035	0.000
	(-0.001)	(0.018)	(0.978)	(-0.036/0.035)	(0.000)

Unadjusted model (adjusted model)

There was no association found between *CACNA1C* GPRS and RA (Table 5.5).

5.6 Discussion

There are multiple lines of evidence to support the hypothesis of altered calcium signalling and the *CACNA1C* genotype in BD (Harrison 2016). However, it is still unclear as to the exact mechanism by which *CACNA1C* genetic variants influence the pathophysiology of BD. It has been suggested that *CACNA1C* may directly or indirectly affect the underlying mechanisms of BD by disrupting circadian rhythms (Kantojarvi et al. 2017; Schmutz et al. 2015; McCarthy et al. 2016). The relationship between the *CACNA1C* gene and peripheral features of the circadian clock, such as chronotype and RA, are currently not understood. However, the above preliminary analyses did not find evidence of associations between *CACNA1C* GPRS, subjective, and objective measures of circadian rhythmicity.

Therefore, further investigation is required to understand the potential mechanisms of *CACNA1C* in BD in the context of circadian rhythmicity.

5.7 Limitations

There appears to be no studies which have combined *CACNA1C* SNPs in this way; therefore, the usefulness of the GPRS generated here has not been validated and could be lacking by limiting the collection of SNPs to only those reported as genome-wide significant. Theoretically, a more extensive *CACNA1C* GPRS may be more accurate for detecting potential associations with circadian measures. Due to the small sample size of individuals with both genotype and sleep phenotype data in ALSPAC the analyses was underpowered. As described in Chapter 3 and 4, compared to UK Biobank there are much smaller sample sizes available in ALSPAC for the outcomes-of-interest.

A limitation of this study was the subjective circadian measures used. In the case of ALSPAC, the sleep phenotypes available were based on mother-reported questionnaires and may not align with the child's experience. Also, chronotype is based upon the self-reported responses to a single question in the UK Biobank (Ganna & Ingelsson 2015). However, this is an accepted measure of chronotype and has been reported to explain the greatest variance of sleep-wake preference (Taillard et al. 2003).

It is important to note, also, that the results displayed above have not been corrected for multiple testing and after applying multiple testing, again, no associations were found.
As described in Chapter 3, the *CACNA1C* GPRS did not consist of completely independent SNPs which could have affected the findings (Williams & Haines 2011).

5.8 Future work

With the release of more recent psychiatric GWAS there is potential to generate a more extensive *CACNA1C* GPRS which could be more effective for investigating the link between this candidate gene and mood disorders. A more robust GPRS may also be of use in determining whether there is an association between *CACNA1C* and circadian rhythmicity.

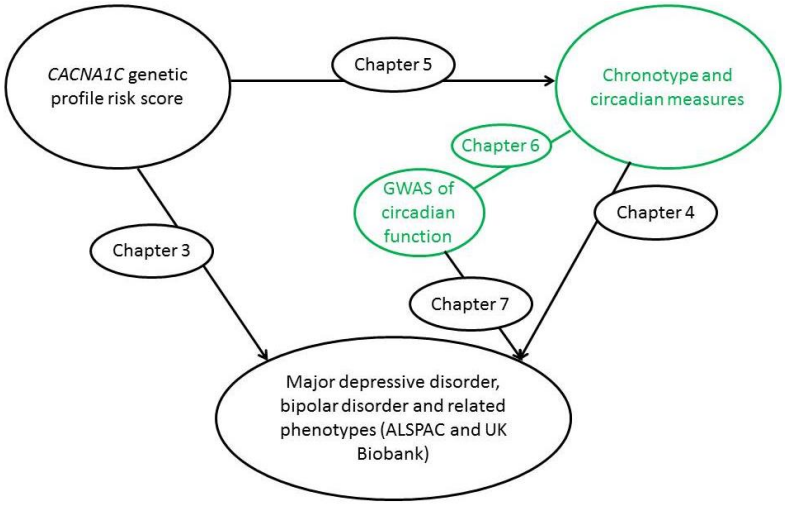
Chapter 4 reported significant associations between chronotype preference and mood disorders. As mentioned in the previous chapter these associations were based on subjective measures of circadian rhythmicity, further study is required to understand the underlying genetic architecture of objective circadian traits, of particular interest is RA.

5.9 Conclusion

Overall, there was no evidence of associations between *CACNA1C* GPRS and measures of circadian rhythmicity. The influence of *CACNA1C* on the underlying pathophysiology of mood disorders currently remains unclear.

Chapter 6 Genome-wide association study of circadian rest-activity rhythmicity in over 77,000 UK Biobank participants

This chapter appears in a published format at Ferguson et al. (2018), EBioMedicine, doi: https://doi.org/10.1016/j.ebiom.2018.08.004 (Ferguson et al. 2018)



Outline 6.1 Chapter in context of overall study

6.1 Introduction

In chapter 4, the influences of chronotype PRS were investigated for associations with several mood disorder-related phenotypes of hypomania, depressive features, lifetime BD, lifetime MDD, mood instability and neuroticism. Increased PRS for evening chronotype positively associated with BD, MDD, mood instability and neuroticism. As chronotype is a subjective measure, investigating a PRS of an objectively measured circadian rhythmicity parameter is preferable. To date there are no large GWAS of objective circadian rhythmicity measures in humans, the largest GWAS using objective measures of sleep parameters obtained from actigraph data consisted of 956 participants from the LIFE Adult Study (Spada et al. 2016). Therefore, this chapter focusses on the first large-scale GWAS of relative amplitude, an objective measure of rest-activity rhythm, in over 77,000 individuals from UK Biobank.

6.1.1 Circadian rhythms

Circadian rhythms are variations in physiology and behaviour that recur approximately every 24 hours (McClung 2007). They include rhythms of body temperature, hormone release, activity, concentration, mood, eating and sleeping.

Circadian rhythmicity plays a fundamental role in homeostasis and in the

maintenance of physical and mental wellbeing (Reppert & Weaver 2001; Merikanto et al. 2017). Circadian disruption is associated with a range of adverse health outcomes, including cardiovascular disease, obesity, diabetes and some cancers (Reutrakul & Knutson 2015; Wulff et al. 2010; Sigurdardottir 2012), as well as increased risk for MDD and BD (Burton et al. 2013; Bullock & Murray 2014; Ng et al. 2015).

Circadian rhythmicity is co-ordinated centrally by the suprachiasmatic nucleus (SCN) in the anterior hypothalamus (Reppert & Weaver 2001) and is regulated by both exogenous environmental stimuli ("zeitgebers") and by genetic factors (Charrier et al. 2017). The circadian clock can be further classified into the central clock and peripheral clocks (Mohawk et al. 2012).

The central clock refers to the autoregulatory transcription/translational feedback loops that maintain cell-cycle function (Koike et al. 2012). The SCN acts as the main regulator of central circadian oscillators which rhythmically alter gene expression to sustain many biological processes, involving several core circadian clock genes (Mohawk et al. 2012). However, circadian oscillators are expressed differentially and independently in different tissues referred to as the peripheral clock (Mohawk et al. 2012; Albrecht 2012). As well as being subject to central oscillators of the circadian clock, cells also contain their own intrinsic clock with peripheral oscillators (Mohawk et al. 2012). There is little overlap in the expression of genes under circadian control between tissues (Mohawk et al. 2012; Albrecht 2012), suggesting a need for specific spatial and temporal controls to function efficiently (Mohawk et al. 2012; Albrecht 2012).

As noted above, several core genes involved in regulating the central circadian clock have been identified, however, communication between the SCN, central and peripheral circadian clocks is not yet well understood (Albrecht 2012). Given the complexity required to synchronise the central and peripheral clocks, and to regulate the mechanisms required to create differential spatial and temporal gene expression, the control of circadian rhythms is likely to be polygenic, with many regulatory genes and pathways still to be identified (Zhang et al. 2009; Merikanto et al. 2017; Albrecht 2012; Mohawk et al. 2012).

To date, the most commonly used measure of circadian phenotypes in humans has been subjectively-reported chronotype, defined as an individual's preference for morning or evening wakefulness and activity (Alloy et al. 2017).

As noted above, evening chronotype is more likely to be associated with adverse health outcomes (Corruble et al. 2014; Dmitrzak-Węglarz et al. 2016; Merikanto et al. 2017; Goel et al. 2014). Recently, GWAS of chronotype, self-reported sleep duration, and accelerometer-derived sleep traits have identified several independent genetic loci previously implicated in the regulation of circadian function (including *PER2*, *PER3*, *RSG16*, *AK5*, *FBXL13*), in addition to novel associated genetic loci (Hu et al. 2016; Jones et al. 2016b; Lane et al. 2016; Jones et al. 2018; Dashti et al. 2018; Jones et al. 2019).

6.1.2 Circadian rhythmicity in mood disorders

Several aspects of circadian rhythms and sleep profiles have been investigated within mood disorders, including chronotype and rest/activity measures (Burton et al. 2013). Disruption of circadian rhythmicity, assessed from both subjective questionnaires and actigraph data, is associated with mood disorders, as well as daytime dysfunction, and impaired cognitive and emotional processing (Geoffroy et al. 2014; Wulff et al. 2010). Previous chapters have explored sleep disturbances, rest-activity circadian rhythmicity disruption and chronotype in relation to mood disorder-related traits, both phenotypically and genetically. Most of these analyses were based on variables generated from subjectively-reported data (chronotype).

6.2 Rationale

Chronotype, as a subjective measure, is vulnerable to response bias. It may also have inconsistent associations with more objective measures of circadian rhythmicity (Taillard et al. 2003). In a recent study within the UK Biobank cohort, objective measures of rest-activity rhythmicity from accelerometer data were derived and it was found that low RA, an objective measure of an individual's rest-activity rhythmicity, was associated with several mood disorder phenotypes (Lyall et al. 2018). This chapter will build on that study by conducting a GWAS of the rest-activity measure low RA.

6.3 Hypothesis to be tested

The aim of this investigation was to conduct a GWAS of low RA in UK Biobank to provide data regarding genetic architecture underlying disrupted rest-activity circadian rhythmicity. A secondary GWAS of continuous RA was also undertaken to provide a more complete understanding of the genetics of rest-activity cycles. Further, the findings of the primary GWAS were used to assess the degree of genetic correlation between low RA and several psychiatric phenotypes, including attention deficit hyperactivity disorder (ADHD), BD, MDD, mood instability, post-traumatic stress disorder (PTSD), schizophrenia, anxiety and insomnia.

6.4 Methods

6.4.1 Participants and ethical approval

Information on participants' recruitment to UK Biobank is detailed in Chapter 2. Here, the data from 91,448 participants who provided genetic and accelerometer data that passed quality control (details below) and who had available data on all covariates included within fully and/or partially adjusted models were used in the GWAS.

6.4.2 Accelerometry data collection and pre-processing

In 2013, 240 000 UK Biobank participants were invited to wear an accelerometer for seven days as part of a physical activity monitoring investigation; with 103,720 (43%) accepted and returned the accelerometer to UK Biobank after use (Doherty et al. 2017). Raw activity intensity data were combined into five second epochs (Supplemental Figure 6.1). Further details on data pre-processing are found in Chapter 2 and are available from UK Biobank at http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=131600 (Doherty et al. 2017).

6.4.3 Circadian rest-activity rhythmicity (relative amplitude, RA)

From the summary five second epoch data, a measure of RA was calculated using Clocklab Version 6 (Actimetrics) by Dr. Cathy Wyse (Wyse et al. 2018, unpublished). This accelerometer-derived activity measure has demonstrated reliability and validity (Sadeh 2011). RA is used commonly as a non-parametric measure of rest-activity rhythm amplitude. It is defined as the relative difference between the most active continuous 10-hour period (M10) and the least active continuous 5-hour period (L5) in an average 24-hour period (Van Someren et al. 1996):

$$RA = \frac{(\mathrm{M10} - \mathrm{L5})}{(\mathrm{M10} + \mathrm{L5})}$$

M10 is the mean activity during the continuous 10 hour period containing maximum activity in each 24 hour recording period (midnight to midnight). L5 is the mean activity for the corresponding 5 hour period containing the minimum activity within the same recording period. For each individual, the RA data point

was the mean RA value across all included 24-hour periods (seven days). RA ranges from 0 to 1, with higher values indicating greater distinction between activity levels during the most and least active periods of the day.

Exclusion criteria detailed in Chapter 2. In the current sample (N=91,870), mean RA was 0.87 (Standard deviation (SD) 0.06; range 0.121-0.997), similar to previously reported values in healthy populations (Bullock & Murray 2014), however, the distribution of RA values was negatively skewed (Supplemental Figure 6.2).

6.4.4 Genotyping and imputation

Full details of the genotyping and imputation methods used by UK Biobank have been described in Chapter 2.

6.4.5 Primary GWAS of low RA

Our primary GWAS was a study of cases of low RA defined as a "pathological tail" of mean RA greater than two standard deviations below the overall mean RA, with the remaining participants classified as controls (Supplemental Figure 6.2) (Lyall et al. 2018). Before proceeding with genetic analyses, further exclusions were applied to the data. Individuals were removed according to UK Biobank genomic analysis exclusions, failure of quality control, gender mismatch, sex chromosome aneuploidy, ethnicity (not Caucasian), lack of accelerometry data, plus the other accelerometry exclusions as noted above. For related individuals (first cousins or closer), a single individual was randomly selected from each pair of related individuals for inclusion in the analysis. After these (and the accelerometry-based) exclusions, 71,500 individuals were available for GWAS. Data were further refined by removing SNPs with an imputation score of <0.8, minor allele frequency of <0.01 and Hardy-Weinberg equilibrium $p<1x10^{-6}$, resulting in 7,969,123 variants.

The primary association analysis was conducted using logistic regression in PLINK (Purcell et al. 2007); an additive allelic effects model was used with sex, age, genotyping array, and the first 8 genetic principal components as covariates. For the GWAS, genome-wide significance was set at $p<5x10^{-8}$.

6.4.6 Secondary GWAS of RA as a continuous trait

The BOLT-LMM method allows the inclusion of related individuals within GWAS. This method requires less restrictive assumptions compared with standard GWAS, as used above, by using a mixture of two Gaussian priors and is a generalisation of a standard mixed model. This mixed model accounts for both relatedness and population stratification, resulting in greater power compared to principal component analysis (Loh et al. 2015). After exclusions, 77,440 individuals were available for this GWAS. As above, genome-wide significance was less than $p<5x10^{-8}$. Note that due to the imbalance between cases and controls available for low RA as a dichotomous measure, it was not appropriate to use a BOLT-LMM approach for the primary GWAS (Loh et al. 2015).

6.4.7 eQTL analysis

The lead SNP from each locus, identified by GWAS, was assessed for the possibility of eQTLs. This genotype-specific gene expression was assessed using the GTEx portal (GTEx Consortium 2013). The portal was also used to investigate tissue-specific expression for the implicated genes (GTEx Consortium 2013).

6.4.8 Gene-based analysis

The summary statistics from both the primary and secondary GWAS analyses were uploaded to the FUMA web application for gene-based analyses (Watanabe et al. 2017). Gene-based analyses were carried out based on the MAGMA method using all genetic associations within the summary statistics (Watanabe et al. 2017; de Leeuw et al. 2015). For these analyses genome-wide significance was set at $p<5x10^{-5}$.

6.4.9 Genetic correlations between low RA and psychiatric phenotypes

LDSR was applied to the GWAS summary statistics to provide an estimate of SNP heritability (h^2_{SNP}) (Bulik-Sullivan et al. 2015; Zheng et al. 2017). LDSR was also used to investigate genetic correlations between low RA and anxiety, ADHD, BD, MDD, mood instability, PTSD, schizophrenia and insomnia. The LD scores for these disorders were obtained using the summary statistics from the Psychiatric

Genomics Consortium, the CNCR-Complex Trait Genomics group, and from UK Biobank. ADHD, BD, MDD, schizophrenia and insomnia were analysed using LD Hub (Zheng et al. 2017).

Methods also detailed in Ferguson et al., 2018.

6.5 Results

6.5.1 GWAS of low RA

The primary analysis was a case-control GWAS of low RA. The GWAS data showed only a slight deviation in test statistics compared to the null (λ_{GC} 1.016, Figure 6.1); this deviation may be due to the polygenic architecture of low relative amplitude. The h^2_{SNP} accounted for less than 1% of the population variance in low RA (h^2_{SNP} 0.0067, S.E 0.0054). The Manhattan plot for low RA GWAS is presented in Figure 6.1. Two independent genomic loci on chromosomes 1 and 22 were associated with low RA at genome-wide significance, including three SNPs (Supplemental Table 6.2 Genome-wide significant loci associated with low RA). These SNPs highlighted two candidate gene loci: Neurofascin (*NFASC*) on chromosome 1 and Solute Carrier Family 25 Member 17 (*SLC25A17*) on chromosome 22 (Figure 6.2). As each of these SNPs is an intronic variant, the exact effect of each polymorphism is unclear.

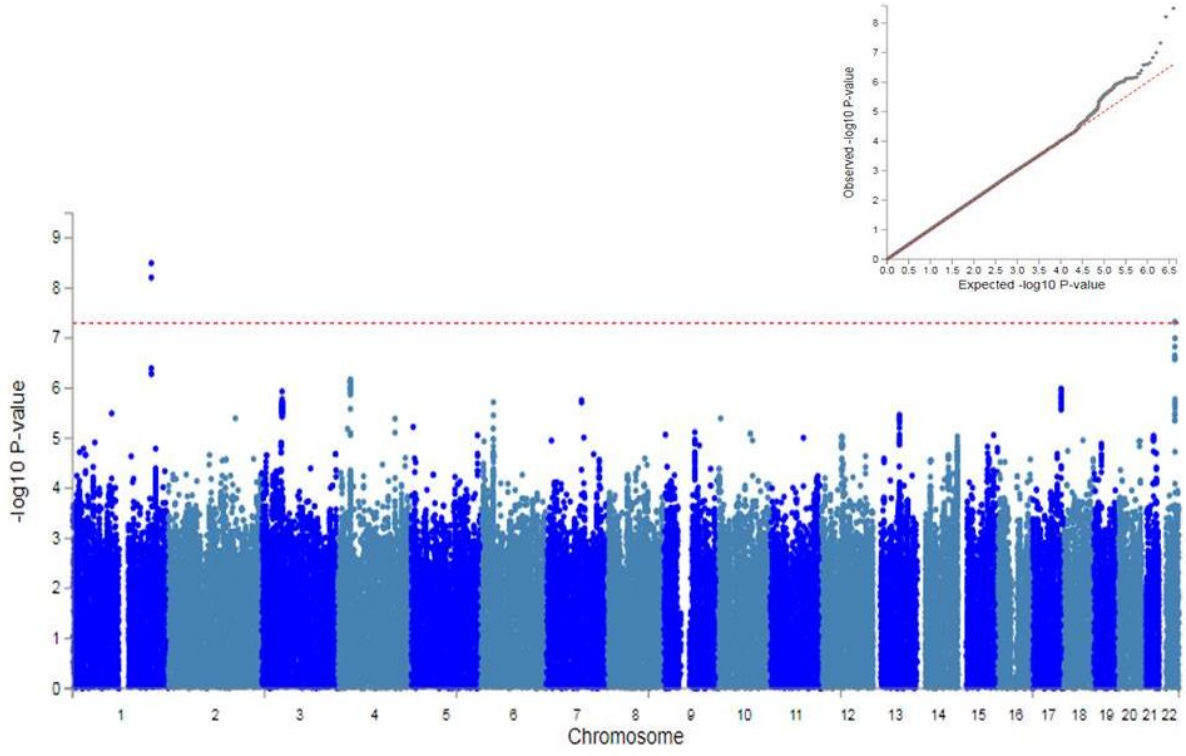


Figure 6.1 SNP Manhattan plot and QQ plot (inset) of low RA GWAS (N=2700 cases verses N=68,300 controls)

Red line of Manhattan plot represents genome-wide significance ($p<5\times10^{-8}$).

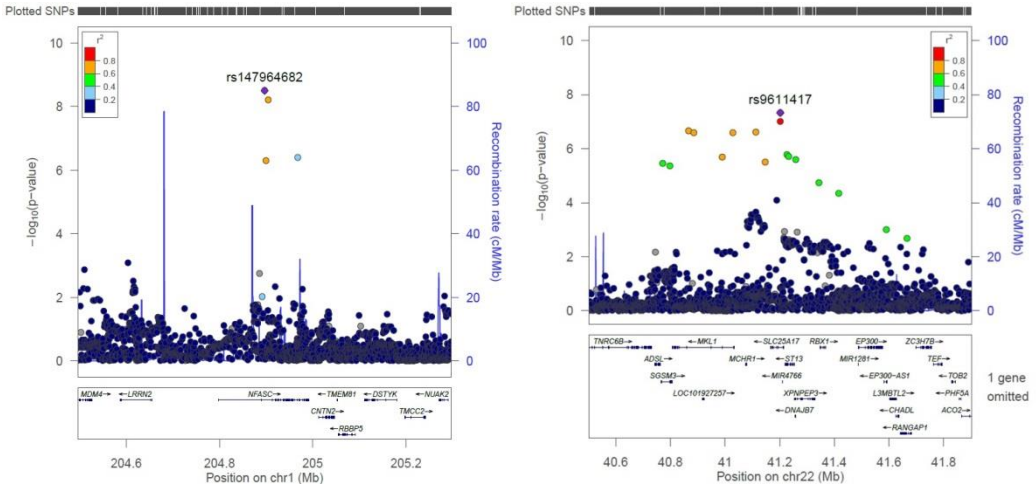


Figure 6.2 Regional plots of *NFASC* and *SLC25A17*

Regional plots of SNPs produced by FUMA (Watanabe et al. 2017).

6.5.2 GWAS of continuous RA

As a secondary analysis, a GWAS of a continuous measure of RA using a BOLT-LMM model was performed. The BOLT-LMM GWAS showed a slight deviation in the test statistics compared to the null (λ_{GC} 1.054, Figure 6.3), again consistent with a polygenic architecture for RA. The h^2_{SNP} for RA as a continuous measure accounted for greater than 8% of the population variance (h^2_{SNP} 0.085, S.E 0.00035). This estimate is much higher compared to that found for low RA (RA=8.5%, low RA=0.67%) which may be a result of the less restrictive BOLT-LMM method used for the continuous RA GWAS. Five SNPs, all localised to one locus on chromosome 2, were associated with continuous RA at genome-wide significance (Supplemental Table 6.3 Genome-wide significant loci associated with continuous RA using BOLT-LMM). These SNPs highlight the Meis Homeobox 1 (*MEIS1*) gene. Again, as noted above, these are intronic SNPs and their exact effects are not known.

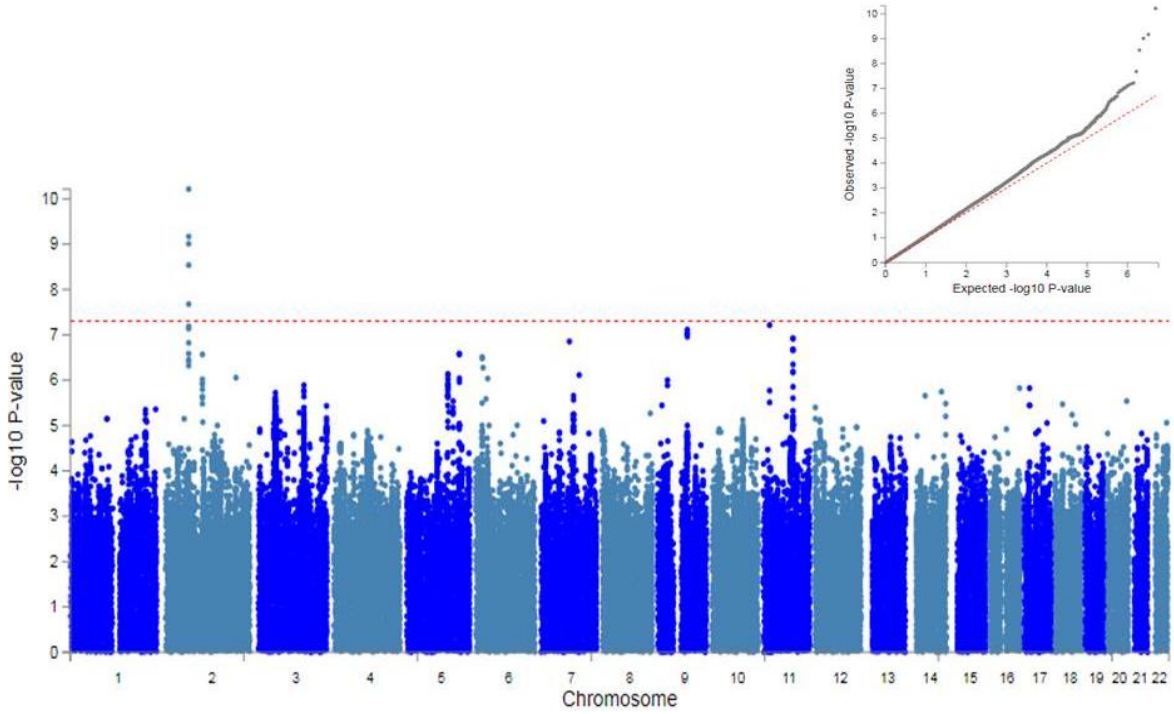


Figure 6.3 SNP Manhattan plot and QQ plot (inset) of continuous RA GWAS (N=77,440)

Red line of Manhattan plot represents genome-wide significance ($p<5\times10^{-8}$).

6.5.3 Gene-based analysis of RA

Gene-based analyses of both low RA and continuous RA were undertaken. The gene-based analysis of low RA identified two genes significantly associated with low RA: Forkhead Box J1 (*FOXJ1*) on chromosome 17 and Zinc Finger FYVE-type Containing 21 (*ZFYVE21*) on chromosome 14 (Figure 6.4). The gene set analysis of continuous RA identified three genes: Copine 4 (*CPNE4*) and Chromosome 3 open reading frame 62 (*C3orf62*) on chromosome 3, and Renalase (*RNLS*) on chromosome 10 (Figure 6.5).

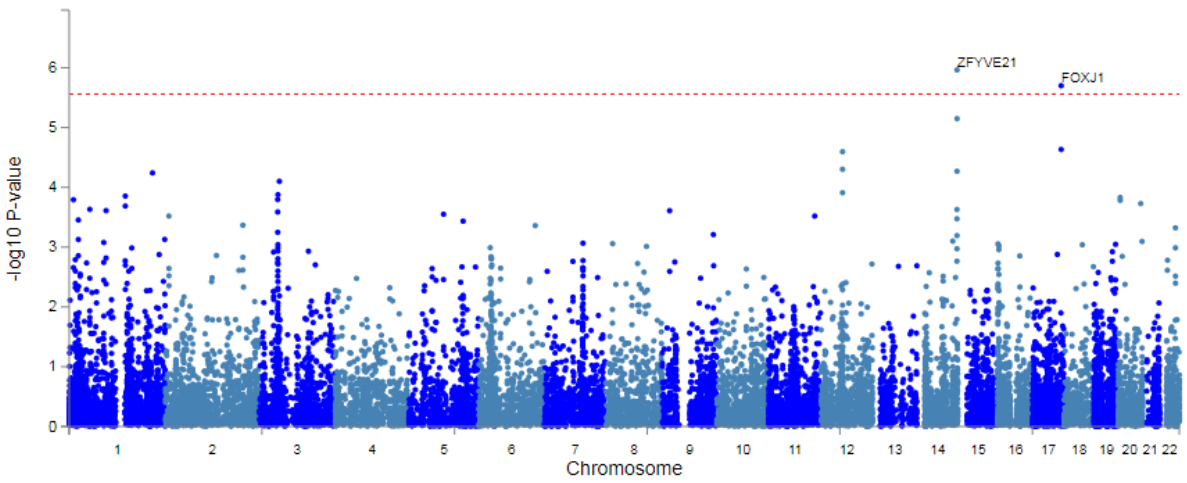


Figure 6.4 Low RA gene-based analysis

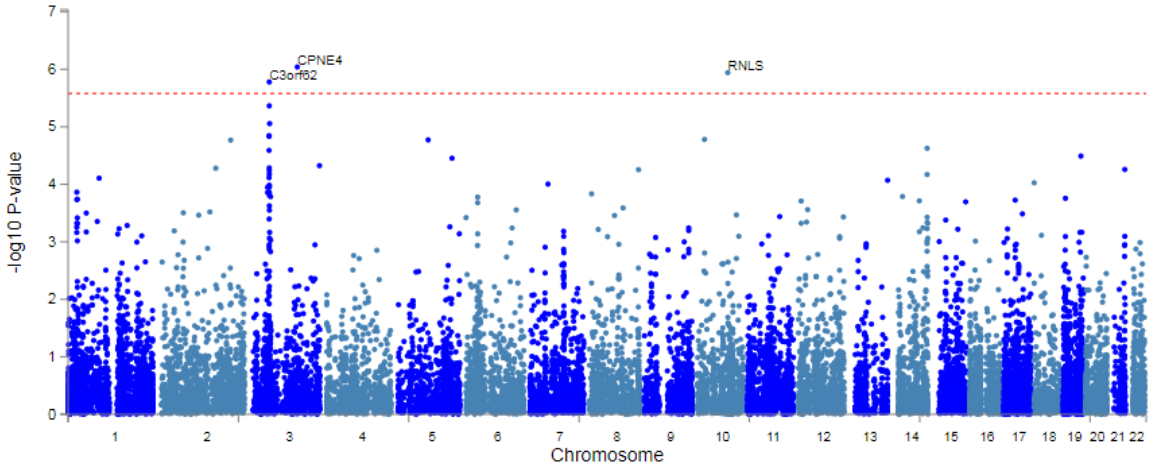


Figure 6.5 Continuous RA gene-based analysis

Red line represents genome-wide significance ($p<5x10^{-8}$).

6.5.4 eQTL analysis

The lead SNPs from both GWAS were assessed for potential eQTLs. Only the lead SNP found on chromosome 22 (rs9611417) was associated with the expression of a nearby gene. Being heterozygous at rs9611417 was associated with lower expression of *RANGAP1* gene in oesophageal mucosa in comparison to rs9611417 C allele homozygotes (beta -0.43, p $7.2x10^{-5}$, Figure 6.6). Information on the influence of G homozygotes was not available.

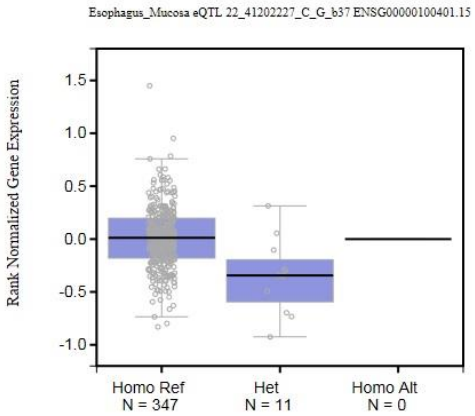


Figure 6.6 eQTLs of rs9611417 box plot

Homo Ref: rs9611417 CC, Het: rs9611417 CG, Homo Alt: rs9611417 GG. Obtained from GTex portal.(GTEx Consortium 2013)

6.5.5 Genetic correlation between low RA and psychiatric phenotypes

There were no significant genetic correlations identified between low RA and ADHD, anxiety, BD, MDD, mood instability, PTSD and schizophrenia. There was weak evidence of genetic correlation between low RA and insomnia (r_g 0.90, S.E 0.42, p 0.033), suggesting that the biology underlying low RA may be associated with the regulation of sleep (Table 6.1) but this finding did not survive FDR correction.

Table 6.1 Genetic correlations between low relative amplitude and ADHD, anxiety, BD, MDD, mood instability, PTSD, schizophrenia and insomnia

Phenotype	r_g	S.E	z	p value	p FDR corrected
ADHD	0.35	0.45	0.77	0.44	0.93
Anxiety	-0.004	0.42	-0.01	0.99	1.00
BD	-0.06	0.18	-0.34	0.73	1.00
MDD	0.005	0.25	0.02	0.99	1.00
Mood Instability	-0.16	0.27	-0.58	0.56	0.94
PTSD	0.74	0.62	1.19	0.23	0.71
Schizophrenia	0.15	0.14	1.14	0.25	0.71
Insomnia	0.90	0.42	2.13	**0.03**	0.28

ADHD: meta-analysis of Children's Hospital of Philadelphia, International Multicenter ADHD Genetics Project (phase I and II) and Pfizer funded study from University of California, Los Angeles, Washington University and Massachusetts General Hospital (N=5,414) (Neale et al. 2010).
Anxiety: Anxiety NeuroGenetics Study Consortium (N=18,000) (Otowa et al. 2016).
BD: Psychiatric Genomics Consortium – BD working group (N=63,766) (Sklar et al. 2011).
MDD: Meta-analysis of Psychiatric Genomics Consortium 29, deCODE, Generation Scotland, Genetic Epidemiology Research on Aging, iPSYCH, UK Biobank and 23andMe (N=480,359) (Wray et al. 2018).
Mood Instability: UK Biobank (N=113,968) (Ward et al. 2017).
PTSD: Psychiatric Genomics Consortium – PTSD working group (N=20,070) (Duncan et al. 2018).
Schizophrenia: Psychiatric Genomics Consortium – Schizophrenia working group (N=150,064) (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al. 2014).

6.6 Discussion

The primary GWAS of low RA identified three genome-wide significant SNPs within two independent loci; the two *NFASC* SNPs highlighted were found to be in relatively high LD (r^2 0.66-1.00) in many populations (Ensembl 2018c) and are potentially tagging a single underlying functional variant influencing low RA. The secondary GWAS of RA as a continuous measure identified five genome-wide significant SNPs at a single locus, within the *MEIS1* gene on chromosome 2. Again, the SNPs identified are in medium to high LD with each other (r^2 0.39-1.00) in many populations, including UK Biobank (Supplemental Figure 6.4) (Ensembl 2018b).

6.6.1 Genes of interest

One of the genes highlighted by GWAS was *NFASC* (Figure 6.2), a gene encoding the neurofascin protein (Taylor et al. 2017). Neurofascin is an L1 family immunoglobulin cell adhesion molecule that interacts with several proteins to anchor voltage-gated Na^+ channels to the intracellular skeleton in neurons. It is involved in neurite outgrowth and organization of axon initial segments (AIS) during early development (Taylor et al. 2017). These AIS complexes (Figure 6.7, comprising neurofascin, ankyrin G (encoded by *ANK3*), gliomedin and betaIV spectrin) are important for the generation of action potentials and for the maintenance of neuronal integrity (Thaxton et al. 2010; Leterrier et al. 2015).

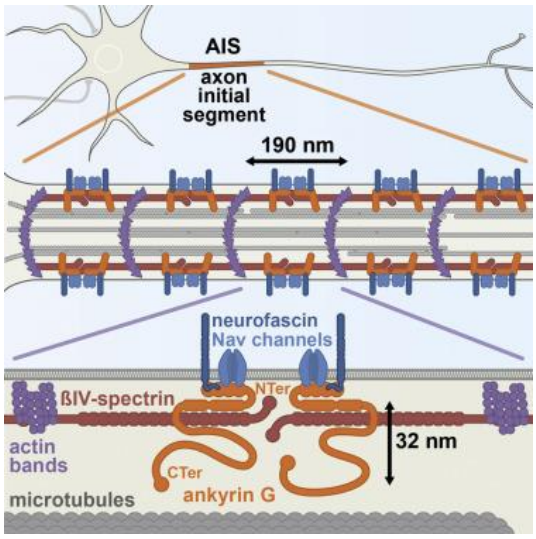


Figure 6.7 Diagram of AIS displaying neurofascin protein interacting with Na+ channels (Leterrier et al. 2015)

Polymorphisms in *ANK3* have been found to be associated with BD through several case-control GWAS (Ferreira et al. 2008; Sklar et al. 2011). The direct binding of neurofascin to ankyrin G at the AIS therefore represents a potentially important biological link between circadian rhythmicity and BD. Interestingly, variants in the *NFASC* gene have previously been found to have a suggestive significant association ($p<5x10^{-5}$) with increased daytime sleep (napping) in UK Biobank (Lane et al. 2017). The *NFASC* SNPs identified by the low RA GWAS were intronic variants and the *NFASC* transcript undergoes extensive alternative splicing with not all variants being functionally categorised (Taylor et al. 2017). The precise influence these variants have on the *NFASC* gene is currently unclear.

One of the genome-wide significant SNPs from the primary GWAS is located within *SLC25A17*. This gene encodes a peroxisomal solute carrier membrane protein that transports several cofactors from the cytosol to the peroxisomal matrix (Agrimi et al. 2012). Variants of this gene have previously been associated with autism spectrum disorder, schizophrenia (The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium 2017), and more recently with BD (Stahl et al. 2017) and morning chronotype preference to suggestive significance (Lane et al. 2016; Jones et al. 2016b). The *SLC25A17* gene is also involved in adrenomyeloneuropathy (Van Veldhoven 2010), an inherited condition in which long chain fatty acids accumulate in the CNS disrupting several brain functions (Moser et al. 1999).
The SNP identified in the GWAS of low RA (rs9611417) was associated with lower expression of *RANGAP1*, a GTPase activator protein involved in nuclear transport (Bischoff et al. 1995). *RANGAP1* is 439kb downstream of rs9611417 and shows relatively high expression in the brain (Supplemental Figure 6.3); this gene has been found to be suggestively associated ($p<5x10^{-5}$) with several sleep traits in UK Biobank, including short sleep duration, frequent insomnia symptoms and excess daytime sleepiness (Lane et al. 2017).

The GWAS of RA as a continuous measure highlighted SNPs within the *MEIS1* gene. This gene encodes a homeobox transcription factor (TF) protein crucial for the normal development of several tissues, including the CNS (Wang et al. 2014; Xiong et al. 2009). The *MEIS1* gene has been reported to have significant associations with insomnia and suggestive significant associations with sleep

duration and excessive daytime sleepiness in UK Biobank (Lane et al. 2017). This gene is also associated with myeloid leukaemia (Wang et al. 2014) and restless leg syndrome 7 (RLS 7) (Wang et al. 2014; Xiong et al. 2009). Restless leg syndromes are neurological sleep-wake disorders characterised by involuntary muscle movement during rest periods, resulting in sleep deprivation and insomnia (Xiong et al. 2009). This is perhaps further evidence of the potential role of the *MEIS1* gene in rest-activity cycles.

6.6.2 Gene-based analyses

The gene-based analysis of low RA identified two genes: *FOXJ1* on chromosome 17, and *ZFYVE21* on chromosome 14. *FOXJ1* encodes a forkhead TF protein which has a role in cell differentiation in respiratory, reproductive, immune, and CNS tissues. It is required for the formation of cilia (Jacquet et al. 2011). In mouse models, *FOXJ1* was reported to be important for neurogenesis within the forebrain and olfactory bulb (Jacquet et al. 2011).

The *ZFYVE21* gene encodes the zinc-finger FYVE-type containing 21 protein, involved in cell migration and adhesion (Nagano et al. 2011). There is relatively little characterisation of the function of this protein and gene such that the potential involvement of this gene in the brain and circadian function is unclear.

For the continuous measure of RA, three genes were identified in the gene-based analysis: *CPNE4* and *C3orf62* (chromosome 3); and *RNLS* on (chromosome 10). The *CPNE4* gene encodes a calcium-dependent phospholipid binding protein involved in membrane trafficking and may be involved in intracellular calcium-mediated processes (Tomsig & Creutz 2002). Deletion of this gene has been associated with earlier age-of-onset of Alzheimer's disease (Szigeti et al. 2013), and has been associated at genome-wide significance with sleep traits in UK Biobank, including frequent insomnia and daytime sleep (Lane et al. 2017). Currently, the *C3orf62* gene is an uncharacterised protein coding gene that has not been functionally annotated (Ensembl 2018a).

RNLS encodes a flavin adenine dinucleotide-dependent amine oxidase, known as Renalase, an enzyme hormone secreted from the kidney into the bloodstream (Xu et al. 2005). Renalase is involved in mediating cardiac function and blood

pressure by influencing heart rate and has been associated with hypertension, chronic kidney failure and type 1 diabetes prediction (Xu et al. 2005; Lv et al. 2016; Frohnert et al. 2018; Desir 2008). It is perhaps worth noting that disrupted circadian rhythmicity has been associated with both diabetes and hypertension in several studies (Merikanto et al. 2013; Merikanto et al. 2016; Reutrakul & Knutson 2015), and that *RNLS* has been associated with both treatment outcome and episode recurrence in BD, as well as showing suggestive significant associations with frequent insomnia symptoms (Fabbri & Serretti 2016; Lane et al. 2017).

There was no evidence of genetic correlation between low RA and any of the psychiatric disorders or mood traits tested. A suggestive genetic correlation was identified between low RA and sleep traits, specifically insomnia; however, this did not remain significant after correcting for multiple testing. The GWAS of low RA is still preliminary and is limited by the small case sample size; it could have been underpowered to detect a large proportion of variants influencing low RA. More extensive low RA GWAS may be required to effectively assess genetic correlations to psychiatric disorders and mood traits.

6.7 Limitations

There are several limitations to the design used to test the above hypothesis:

1. UK Biobank may not be representative of the general UK population (Fry et al. 2017), with a possible under-representation of individuals with disrupted rest-activity cycles (i.e. those with low RA) resulting in a relatively small sample size of low RA cases. It is likely that the case-control GWAS of low RA was underpowered to detect variants which have only small or moderate effects on RA. Much larger case sample sizes are required to detect variants (Dunn et al. 2015).

2. The use of the LDSC package for estimating SNP heritability in low RA cases (1.2 million SNPs rather than 9 million SNPs assessed by the BOLT-LMM approach) could have resulted in an underestimation of trait heritability. Also, with the relatively small sample size of low RA cases, our analyses may be underpowered resulting in a lower estimate of SNP

heritability than expected (for example, compared to chronotype) (Jones et al. 2016b). Previous investigations of accelerometer-based phenotypes using selected clinical samples (such as BD) reported higher heritability estimates (h^2>0.30) (Pagani et al. 2016). The use of a non-clinical, general population cohort with potential under-representation of psychiatric disorders may have resulted in a SNP heritability estimate that is lower than might be expected.

3. Circadian rhythms are subject to influence from both biological factors and environmental stimuli. It is a potential limitation of this investigation that the analyses were not adjusted for potential non-genetic confounders operating during the accelerometry data collection period, such as medical illness, medication status, chronic pain, transmeridian air travel, obesity and irregular work patterns. The accelerometers were worn for 7 days which may not fully represent rhythmicity, particularly in working participants where weekend rhythms might differ substantially from weekday rhythms. Further studies are required to investigate the possible interaction between RA and environmental factors; future investigations may also benefit from the inclusion of other measures of RA variability to adjust for intra- and inter-individual activity levels. There are limitations to accelerometers also; they are not valid for measuring some physical activities, including cycling and resistance force activities. It is also difficult to extrapolate accelerometry data as individuals typically perform activities for a shorter time period than when not wearing the accelerometer (de Almeida Mendes et al. 2018).

6.8 Future work

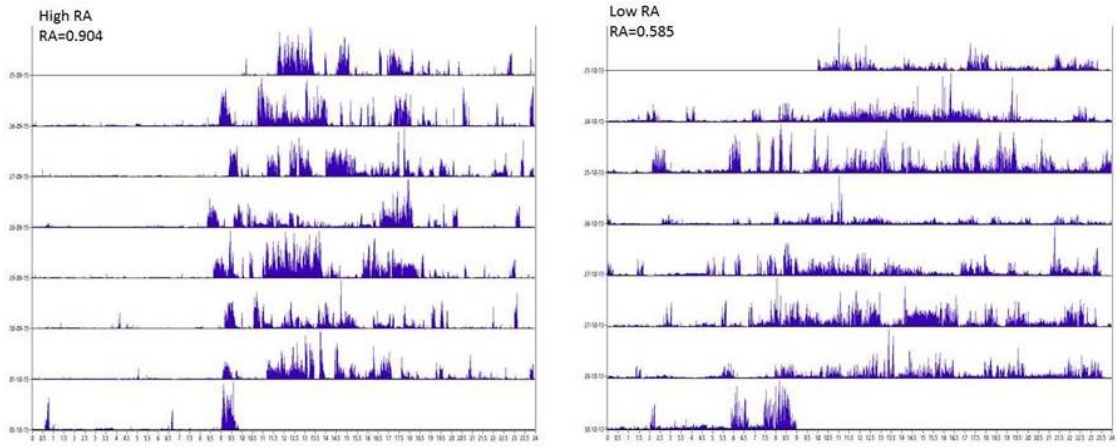
The direct influence that genes-of-interest highlighted here have on circadian rhythmicity and mood disorders is unclear. Further studies are required to elucidate their function with respect to depression and BD. The identification of genes with putative biological links to, or associated with, mood disorders does however suggest a link between circadian rhythmicity and mood disorders. This merits further investigation, for example, with polygenic risk scores. The following chapter will investigate the associations between low RA polygenic risk scores and mood disorder phenotypes.

The results of this investigation are preliminary and require replication. Accelerometery data from large cohorts such as 23andMe may help to strengthen the findings of the above GWAS. If accelerometery data was collected in a cohort such as ALSPAC, this would provide data on rest-activity data and, by extension, information on individuals' circadian rhythmicity at an earlier age which may highlight variants and pathways involved in the development of mood disorders and related traits.

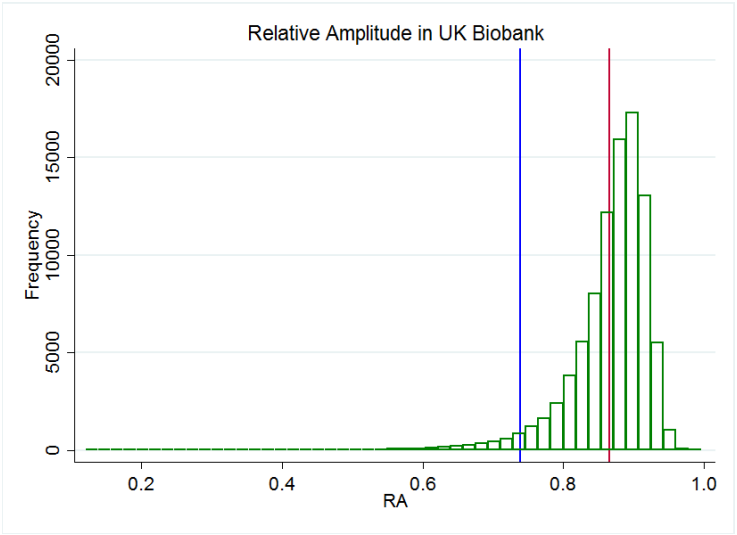
6.9 Conclusions

Overall, these findings contribute new knowledge on the genetic architecture of circadian rhythmicity. Several of the genetic variants are located within or close to genes which may have a role in the pathophysiology of mood disorders. These findings provide novel genetic variants to investigate in the context of mood disorders and potentially strengthen the evidence for the relationship between disrupted circadian rhythmicity and psychiatric traits and conditions.

6.10 Supplemental

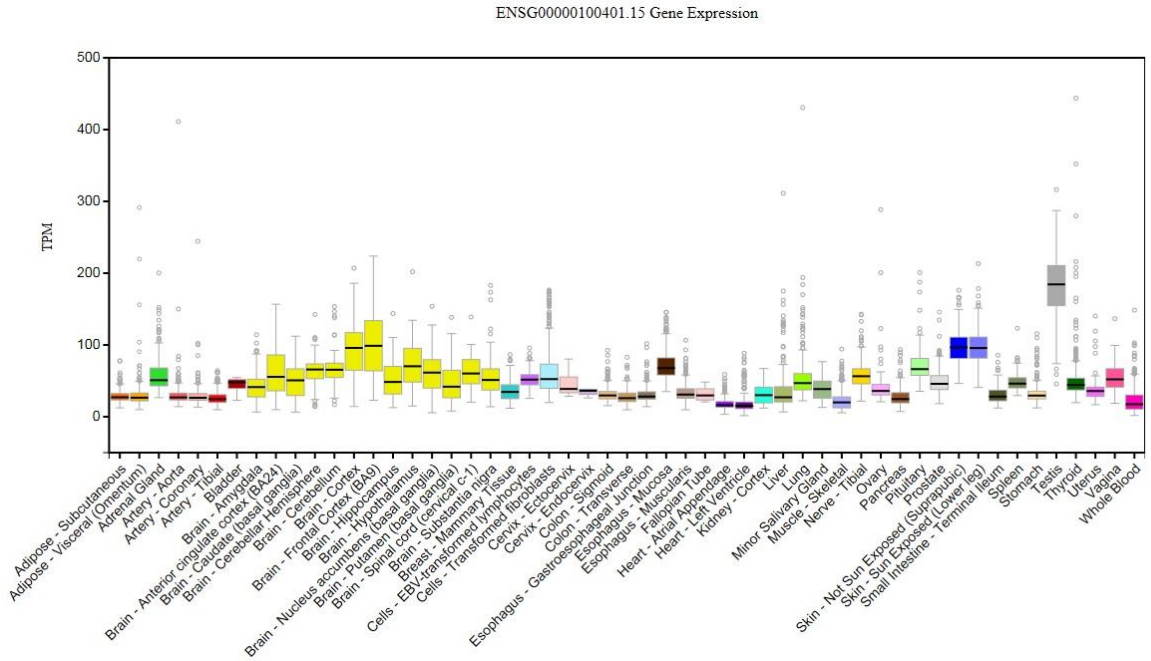


Supplemental Figure 6.1 Representative actograms of high RA and low RA



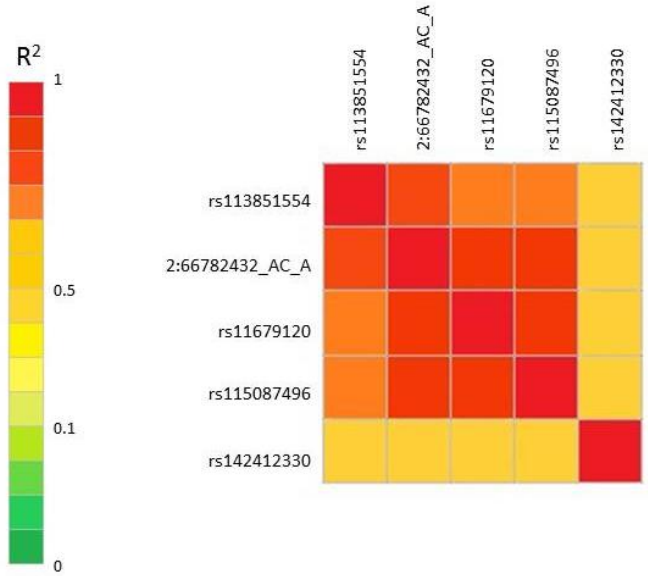
Supplemental Figure 6.2 Relative amplitude histogram indicating (N=2,987 Cases and N=80,352 Controls)

Red line represents mean value of RA. Blue line is two standard deviations from the mean, designating cases for use in the primary GWAS. These numbers indicate all (Caucasian) individuals available for GWAS before genetic exclusions and QC were applied.



Supplemental Figure 6.3 Tissue specific expression of *RANGAP1*

Obtained from GTex portal.(GTEx Consortium 2013)



Supplemental Figure 6.4 LD heat map of continuous RA GWAS significant SNPs (within *MEIS1* gene)

Supplemental Table 6.1 Demographics of UK Biobank participants included in accelerometer data subset

	Low RA cases	Controls	Test statistic	p value
Age at baseline, years			3.15	<0.0001
Mean(SD)	55.94(0.15)	56.39(0.03)		
Median(IQR)	57(50-64)	58(52-65)		
Age at accelerometry, years			2.92	<0.0001
Mean(SD)	61.69(0.15)	62.11(0.03)		
Median(IQR)	62(56-69)	63(57-69)		
Sex			-18.18	<0.0001
Female	1,187(39.78)	45,405(56.57)		
Male	1,797(60.22)	34,857(43.43)		
Townsend deprivation score			-21.1	<0.0001
Mean(SD)	-0.84(0.06)	-1.91(0.01)		

Supplemental Table 6.2 Genome-wide significant loci associated with low RA

SNP	Chr	Position	A1/A2	MAF	Beta	p value	Nearby gene
rs147964682	1	204,896,804	G/C	0.012	0.584	**3.179x10^{-9}**	*NFASC*
rs146042826	1	204,904,528	A/G	0.013	0.568	**6.17x10^{-9}**	*NFASC*
rs9611417	22	41,202,227	G/C	0.012	0.5616	**4.753x10^{-8}**	*SLC25A17*

A1: reference allele

Supplemental Table 6.3 Genome-wide significant loci associated with continuous RA using BOLT-LMM

SNP	Chr	Position	A1/A2	Info	Beta	p value	Nearby gene
rs113851554	2	66,750,564	G/T	0.93	0.0048	**6.10x10^{-11}**	*MEIS1*
2:66782432_AC_A	2	66,782,432	AC/A	0.92	0.0048	**9.80x10^{-10}**	*MEIS1*
rs11679120	2	66,785,180	G/A	0.93	0.0045	**2.10x10^{-8}**	*MEIS1*
rs115087496	2	66,793,725	G/C	0.92	0.0047	**2.90x10^{-9}**	*MEIS1*
rs142412330	2	66,802,493	T/TCTC	0.93	-0.0036	**6.80x10^{-10}**	*MEIS1*

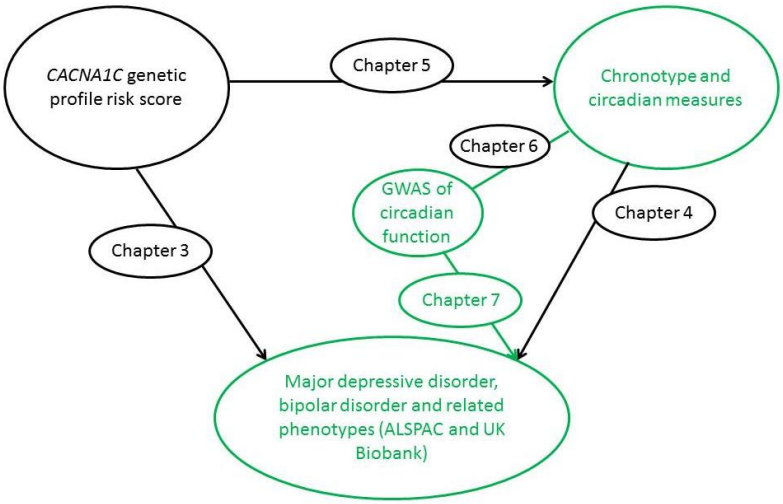
Supplemental Table 6.4 Suggestive significant loci associated with low RA

SNP	Chr	Position	A1/A2	Info	Beta	p value	Nearby gene
rs11715894	3	53,040,459	A/T	0.421	0.138	$1.16x10^{-6}$	*SFMBT1*
rs560770640	3	53,044,907	G/T	0.42	0.136	$1.52x10^{-6}$	*SFMBT1*
rs578200280	3	53,068,709	C/T	0.421	0.136	$1.67x10^{-6}$	*SFMBT1*
rs151079563	3	53,194,321	T/C	0.447	0.145	$4.01x10^{-7}$	*PRKCD*
rs190135744	4	32,918,842	C/T	0.054	0.285	$2.71x10^{-7}$	*intergenic*
rs9277979	6	33,294,098	T/C	0.177	0.167	$1.91x10^{-6}$	*DAXX*
rs546882114	6	33,308,438	T/C	0.181	0.169	$1.95x10^{-6}$	*intergenic*
rs372171356	6	33,295,111	G/A	0.177	0.164	$3.10x10^{-6}$	*DAXX*
rs76775274	7	91,713,047	C/T	0.011	0.521	$1.75x10^{-6}$	*AKAP9*
rs117704951	7	91,876,485	T/A	0.011	0.514	$1.92x10^{-6}$	*ANKIB1*
rs146314842	7	97,662,316	T/G	0.03	-0.459	$9.71x10^{-6}$	*intergenic*
rs289055	13	68,464,763	G/A	0.364	0.134	$3.44x10^{-6}$	*intergenic*
rs289056	13	68,476,397	T/C	0.366	0.132	$4.74x10^{-6}$	*OR7E111P*
rs2094932	13	68,482,537	G/A	0.363	0.13	$8.42x10^{-6}$	*intergenic*
rs556389482	17	74,166,151	A/G	0.268	0.16	$4.57x10^{-7}$	*RNF157*
rs562449594	17	74,171,132	G/A	0.264	0.152	$9.40x10^{-7}$	*RNF157*
rs754706	17	74,150,113	C/T	0.264	0.151	$1.03x10^{-6}$	*RNF157*
rs547968601	17	74,171,356	A/G	0.264	0.151	$1.05x10^{-6}$	*RNF157*
rs1868822	17	74,149,524	C/T	0.264	0.151	$1.06x10^{-6}$	*RNF157*
rs9277979	6	33,294,098	T/C	0.177	0.167	$1.91x10^{-6}$	*DAXX*
rs546882114	6	33,308,438	T/C	0.181	0.169	$1.95x10^{-6}$	*intergenic*
rs372171356	6	33,295,111	G/A	0.177	0.164	$3.10x10^{-6}$	*DAXX*

Supplemental Table 6.5 Suggestive significant loci associated with continuous RA

SNP	Chr	Position	A1/A2	Info	Beta	p value	Nearby gene
rs10194961	2	106,304,263	T/A	0.338	0.002	$2.80x10^{-7}$	*intergenic*
rs115291000	2	149,409,639	G/A	0.958	0.004	$2.90x10^{-6}$	*EPC2*
rs11686221	2	107,263,240	C/T	0.988	0.006	$4.70x10^{-6}$	*intergenic*
rs139169199	2	141,262,927	C/CA	0.96	0.004	$1.10x10^{-5}$	*LRP1B*
rs11693221	2	66,799,986	C/T	0.953	0.004	$1.00x10^{-7}$	*MEIS1*
rs142704867	5	151,700,577	T/C	0.964	0.004	$4.40x10^{-7}$	*intergenic*
rs79593753	5	151,777,074	C/T	0.974	0.005	$3.20x10^{-6}$	*NMUR2*
rs11538104	5	133,727,052	T/G	0.989	0.007	$3.30x10^{-6}$	*UBE2B*
rs36072161	5	133,729,345	T/A	0.989	0.007	$4.10x10^{-6}$	*intergenic*
rs3842139	6	34,665,678	C/CAA	0.264	-0.002	$4.10x10^{-7}$	*intergenic*
rs12194155	6	18,557,377	G/T	0.94	-0.003	$2.00x10^{-6}$	*MIR548A1HG*
rs12215669	6	18,561,460	C/T	0.94	-0.003	$2.10x10^{-6}$	*MIR548A1HG*
rs115595252	7	110,014,522	T/A	0.991	0.008	$3.90x10^{-6}$	*intergenic*
rs7809370	7	93,708,083	A/G	0.63	0.001	$4.80x10^{-6}$	*intergenic*
rs11139851	9	85,493,040	C/G	0.852	0.002	$6.00x10^{-7}$	*intergenic*
rs142398474	11	18,669,128	C/T	0.982	0.006	$1.90x10^{-7}$	*intergenic*
rs7951433	11	86,084,878	C/A	0.986	0.007	$6.10x10^{-7}$	*intergenic*

Chapter 7 Polygenic risk score for low relative amplitude and the association with mood disorder-related traits, and sleep-related traits in UK Biobank and ALSPAC



Outline 7.1 Chapter in context of overall study

7.1 Introduction

In a previous chapter, a GWAS of RA, an objective measure of rest-activity cycles and circadian health, was conducted (Ferguson et al. 2018). This chapter builds on the findings of this work by assessing, within both UK Biobank and ALSPAC, how PRS for low RA associates with both mood disorders and sleep phenotypes.

7.1.1 Circadian rhythms in mood disorders

Chapters 4-6 have highlighted that disrupted circadian rhythmicity is associated with a range of adverse health outcomes, including cardiovascular disease, obesity, diabetes and some cancers (Reutrakul & Knutson 2015; Wulff et al. 2010; Sigurdardottir 2012), as well as increased risk for MDD and BD (Burton et al. 2013; Bullock & Murray 2014; Ng et al. 2015). Many studies have found an association between mood disorders, psychiatric traits and disruption of circadian rhythmicity (Baek et al. 2016; Pagani et al. 2016; Geoffroy et al. 2014; Bellivier et al. 2015). In UK Biobank, low RA (reflective of greater circadian disruption) was associated with increased risk of BD, MDD, mood instability and neuroticism, as well as, subjective measures of wellbeing. Individuals with low RA were also more likely to have lower health satisfaction scores, lower subjective happiness and greater subjective loneliness (Lyall et al. 2018).

As described in Chapter 6, there are layers of complexity within the circadian clock; many biological processes under circadian control which require both synchronization and spatial-temporal control. Disrupted rest-activity cycles, i.e., low RA, could be influenced by alterations in any one of these multiple layers of

complexity within the circadian control system. Although there have been major advances in understanding the molecular substrates of the circadian clock, there is still much to be discovered in this area.

In Chapter 6, a primary GWAS of low RA in UK Biobank identified three SNPs of genome-wide significance ($p<5x10^{-8}$) as well as many SNPs of suggestive significance ($p<5x10^{-5}$) but there was little to no genetic correlation found between low RA, sleep traits and psychiatric phenotypes (as reported in Chapter 6) (Ferguson et al. 2018). However, as genetic correlations can be limited by the design, power and results of GWAS investigations (in the sense that some GWAS methodologies are more restrictive and underestimate heritability or may be underpowered due to smaller sample size) some associations may have been missed (Loh et al. 2015). The epidemiological data suggests that low RA may be genetically correlated with mood disorders and related traits, as well as with subjective sleep features.

Understanding how low RA affects these traits should provide a greater understanding of the underlying pathophysiology of mood disorders and could provide a novel target for treatment or management of the disorders. In this chapter, these potential associations will be tested in both UK Biobank and ALSPAC using low RA PRS. This approach allows for the assessment of the possible link between genetic loading for low RA and these traits-of-interest in a wide age-range of adults, as there may be a differential affect at younger versus older ages.

7.2 Rationale

Low RA has previously been, phenotypically, associated with mood disorder-related outcomes in UK Biobank (Lyall et al. 2018). However, it was unclear whether genetic loading for low RA associates with mood disorders, mood-related traits and sleep-related traits in UK Biobank and ALSPAC. The advantage of investigating the potential associations in both UK Biobank and ALSPAC is the possibility of replicating associations. However, it is reasonable to suggest that the differences in factors such as age, lifestyles, life experience and socioeconomic status between the two cohorts could highlight potential external influences on the effects driven by low RA PRS.

There appear to be no studies using PRS of objective circadian measures in this way. The following analyses may provide a greater understanding of the underlying mechanisms of mood-disorders and related traits.

7.3 Hypothesis to be tested

The aim of this investigation was to determine whether higher genetic loading (based on a minimum of 21 SNPs) for low RA was associated with mood disorder-related outcomes and features of disrupted sleep.

Within a group of UK Biobank participants who were not part of the primary GWAS study (141,000 individuals), the association between increased PRS for low RA and mood trait phenotypes (specifically BD, MDD, GAD, mood instability and neuroticism) was tested. I also tested the association between low RA PRS and self-reported chronotype.

Within ALSPAC, low RA PRS was tested for association with hypomania, depression and subjective measures of disturbed sleep.

7.4 Methods

7.4.1 Genotyping and imputation

Descriptions of the genotyping, imputation and phasing methods used in UK Biobank and ALSPAC are detailed in Chapter 2.

7.4.2 Bipolar Disorder, Depression and Generalised Anxiety Disorder

A MHQ was developed by a UK Biobank mental health research reference group to collect additional mental health phenotype data and was administered during 2016-2017. Lifetime BD, lifetime MDD and lifetime GAD variables were generated for 157,366 UK Biobank participants. Further details for these variables have been described in Chapter 2.

7.4.3 Neuroticism

To define neuroticism a score taken from the 12 item neuroticism scale of the EPQ-R-S (Eysenck et al. 1985). Individuals were given a score of 0 or 1 for a "no or yes" answer, respectively, for each item and given a total neuroticism score ranging from 0 to 12. As described in Chapter 2.

7.4.4 Mood instability

As described in Chapter 2, a "mood instability" outcome measure was also obtained from the EPQ-R-S questionnaire: participants were asked *"Does your mood often go up and down?"* and given the option to answer "yes", "no", "don't know" or "prefer not to answer". Individuals who selected "don't know" or "prefer not to answer" were coded as missing; this allowed the generation of a categorical mood instability variable where those who answered "yes" were designated as cases and participants who answered "no" were included as controls.

7.4.5 Chronotype phenotype in UK Biobank

Chronotype was derived from the participants' responses to a question obtained from the Morningness-Eveningness questionnaire. Categorical variables were

then generated based upon the responses given, resulting in the generation of four separate chronotype variables. The primary outcome measures used for analysis were the “definite morning” and “definite evening” variables with “overall morning” and “overall evening” providing secondary outcome measures. The method by which these variables were generated is detailed in Chapters 2 and 4.

7.4.6 Hypomania

To test associations between low RA PRS and features of BD in ALSPAC categorical and continuous measures of hypomania were used. The full details of how these measures are generated using HCL-28 can be found in Chapter 2. The categorical definition of hypomania was the primary outcome measure and the continuous HCL-28 score was a secondary outcome.

7.4.7 Depressive features

Depressive symptoms were assessed in the ALSPAC cohort during late childhood and adolescence using the SMFQ. This generates an SMFQ score ranging from 0 to 26 (Wiles et al. 2012), as well as a categorical depression measure (binary depression) based on a score of greater than 16. The categorical SMFQ measure was used as the primary outcome measure and the SMFQ score was the secondary outcome.

7.4.8 Sleep phenotypes in ALSPAC

There is no measure of chronotype within ALSPAC therefore the low RA PRS was tested against four measures of sleep problems reported at two different ages (age 10 and age 13). This resulted in four sleep-related variables; these variables were generated from mother-answered questionnaires regarding child sleep habits. These phenotypes are detailed in Chapter 2 and Chapter 4.

The primary outcome measures for sleep in ALSPAC were the categorical variables “Difficulty sleeping-10” and “Difficulty sleeping-13” which were tested for their association to low RA PRS using logistic regression. Secondary outcome measures were also tested (“Difficulty sleeping-scale” and “Worried sleep”) using linear regression.

7.4.9 Association between PRS for low RA and mood disorder phenotypes in UK Biobank

Associations between higher PRS for low RA and psychiatric diagnoses were examined in up to 76,018 Caucasian individuals who had completed the MHQ and who were not included in the primary low RA GWAS, described in Chapter 6. Similarly, associations between low RA PRS and mood instability/neuroticism were examined in between 91,248 and 140,504 individuals (depending on the dependent variable) not included in the low RA GWAS. PRS including SNPs at 6 different significance thresholds ($p<5x10^{-8}$, $p<5x10^{-5}$, p<0·01, p<0·05, p<0·1, p<0·5) were divided into quartiles, with the exception of $p<5x10^{-8}$ which was divided into tertiles as there were only three scores generated for participants at this threshold. The top and bottom quantiles were compared in logistic regression models that were adjusted for age, sex, Townsend deprivation index (Townsend 1987), genotype array and the first 8 genetic principal components. FDR correction was applied which is less conservative than Bonferroni correction (Pike 2011; Benjamini & Hochberg 1995).

7.4.10 Association between PRS for low RA and mood disorder phenotypes in ALSPAC

Associations between higher PRS for hypomania and depression were examined in up to 2,500 Caucasian individuals. PRS were generated at 5 SNP significance thresholds, using variants identified by the low RA GWAS undertaken in UK Biobank, and divided into quantiles, as above. Due to the small number of SNPs used to generate a PRS for genome-wide significant SNPs, there was insufficient genotype data available in ALSPAC to generate scores for this threshold. The top vs. bottom quantiles were contrasted in regression models that were adjusted for sex and deprivation. As in previous chapters, ALSPAC analysis was not adjusted for age due to participants being of similar age (25-27 years old).

7.5 Results

7.5.1 Association between PRS for low RA and mood disorder phenotypes in UK Biobank

The findings of analyses assessing the association between low RA PRS and several mood disorder-related phenotypes are presented in Table 7.1. Positive associations were identified between increased PRS and (increased risk of) mood instability at all PRS thresholds (these associations met significance with $p<0.05$), with the exception of genome-wide significance threshold ($p<5x10^{-8}$). For MDD, small positive associations were found for the low RA PRS at the top three significance thresholds (O.R 1·02-1·03), which remained significant after FDR correction (p 0·025-0·05). A positive association with neuroticism was found for the highest threshold (p 0·004, FDR adjusted p 0·021). However, other associations between the remaining PRS thresholds and neuroticism score were not significant (Ferguson et al. 2018).

Table 7.1 Associations between low RA PRS and mood disorder phenotypes

PRS p threshold	Outcome(Cases/Controls)	O.R (95% CI)	p uncorrected	p FDR corrected
$p<5x10^{-8}$	BD	0·99 (0·92,1·06)	0·748	0·785
$p<5x10^{-5}$	(406/37,699)	1·05 (0·98, 1·12)	0·206	0·746
p<0·01		1·03 (0·96, 1·11)	0·355	0·746
p<0·05		1·04 (0·96, 1·12)	0·309	0·746
p<0·1		1·02 (0·94, 1·10)	0·617	0·785
p<0·5		1·02 (0·93, 1·11)	0·754	0·785
$p<5x10^{-8}$	MDD	1·00 (0·99, 1·02)	0·812	0·805
$p<5x10^{-5}$	(9,543/24,317)	1·00 (0·98, 1·02)	0·966	0·805
p<0·01		1·01 (0·99, 1·03)	0·395	0·494
p<0·05		1·02 (1·00, 1·04)	**0·03**	**0·050**
p<0·1		1·03 (1·01, 1·05)	**0·005**	**0·025**
p<0·5		1·03 (1·00, 1·05)	**0·021**	**0·050**
$p<5x10^{-8}$	GAD	0·97 (0·95, 1·00)	0·092	0·300
$p<5x10^{-5}$	(2,587/23,564)	0·98 (0·96, 1·01)	0·274	0·548
p<0·01		0·99 (0·97, 1·02)	0·729	0·729
p<0·05		1·01 (0·98, 1·04)	0·475	0·713
p<0·1		1·03 (0·99, 1·06)	0·1	0·300
p<0·5		1·01 (0·97, 1·04)	0·699	0·729
$p<5x10^{-8}$	Mood Instability	1·00 (0·99, 1·01)	0·913	0·940
$p<5x10^{-5}$	(78,710/91,248)	1·01 (1·00, 1·02)	**0·019**	**0·0096**
p<0·01		1·01 (1·01, 1·02)	**$9·5x10^{-5}$**	**$2·2x10^{-4}$**
p<0·05		1·02 (1·01, 1·02)	**$3·6x10^{-6}$**	**$9·6x10^{-5}$**
p<0·1		1·01 (1·01, 1·02)	**$8·3x10^{-5}$**	**$5·9x10^{-4}$**
p<0·5		1·02 (1·01, 1·03)	**$1·2x10^{-6}$**	**$1·5x10^{-5}$**
PRS p threshold	**Outcome (N)**	**Beta (95% CI)**	**p uncorrected**	**p FDR corrected**
$p<5x10^{-8}$	Neuroticism score	-0·004 (-0·02, 0·007)	0·456	0·399
$p<5x10^{-5}$	(140,504)	0·01 (-0·003, 0·02)	0·124	0·134
p<0·01		0·01 (-0·002, 0·02)	0·098	0·134
p<0·05		0·01 (-0·005, 0·03)	0·059	0·134
p<0·1		0·01 (-0·003, 0·02)	0·128	0·134
p<0·5		0·02 (0·007, 0·04)	**0·004**	**0·021**

7.5.2 Association between PRS for low RA and mood disorder phenotypes in ALSPAC

The PRS for low RA was tested for any association with hypomanic and depressive symptoms in ALSPAC (Table 7.2). There were no significant associations found between low RA and the primary mood-related trait measures in ALSPAC. There was a nominally significant association between low RA and SMFQ score at two significance thresholds, however, these associations did not survive correction.

Table 7.2 Associations between low RA PRS and mood disorder-related traits in ALSPAC

PRS p threshold	Outcome	O.R (95% CI)	p uncorrected	p FDR corrected
$p<5x10^{-5}$	Binary Hypomania	0·98 (0·84,1·14)	0·793	0.991
p<0·01	(age 22-23)	0·99 (0·81,1·22)	0·970	0.991
p<0·05		0.98 (0·81, 1·20)	0·872	0.991
p<0·1		1.08 (0·89, 1·32)	0·424	0.991
p<0·5		1·00 (0·82, 1·22)	0·991	0.991
$p<5x10^{-5}$	Binary Depression	1·16 (0·94, 1·43)	0·173	0.286
p<0·01	(age 10-19)	1·22 (0·94, 1·59)	0·129	0.286
p<0·05		1·19 (0·45, 1·57)	0·229	0.286
p<0·1		1·27 (0.97, 1·68)	0·087	0.286
p<0·5		1·14 (0.86, 1·51)	0·355	0.355
PRS p threshold	**Outcome**	**Beta (95% CI)**	**p uncorrected**	**p FDR corrected**
$p<5x10^{-5}$	HCL score	0·05 (-0·21, 0·31)	0·720	0.720
p<0·01	(age 22-23)	0·20 (-0·16, 0·57)	0·274	0.567
p<0·05		0·17 (-0·18, 0·53)	0·340	0.567
p<0·1		0·31 (-0·05, 0·66)	0·089	0.445
p<0·5		0·13 (-0·22, 0·49)	0·46	0.575
$p<5x10^{-5}$	SMFQ score	0·02 (-0·02, 0·007)	0·647	0.647
p<0·01	(age 10-19)	0·09 (-0·003, 0·02)	0·079	0.132
p<0·05		0·11 (-0·002, 0·02)	**0·031**	0.108
p<0·1		0·10 (-0·005, 0·03)	**0·043**	0.108
p<0·5		0·03 (-0·003, 0·02)	0·512	0.640

7.5.3 Association between PRS for low RA and chronotype in UK Biobank

It was unclear as to whether the polygenic loading for low RA would show any association to more subjective circadian/sleep phenotypes. These analyses could be a partial indication of how low RA aligns with other circadian traits. In chapter 4, the influences of different chronotypes on both physical and mental health were discussed in depth; the earlier chapter reported the association between evening chronotype and mood disorder-related phenotypes.
As expected, the genetic loading for low RA was found to be associated with decreased self-reported morningness and increased eveningness at several PRS significance thresholds (Table 7.3).

Table 7.3 Associations between low RA PRS and self-reported chronotype in UK Biobank

PRS p threshold	Outcome	O.R (95% CI)	p uncorrected	p FDR corrected
$p<5x10^{-8}$	Definite morning	0·99 (0·98, 1.01)	0·610	0.610
$p<5x10^{-5}$		0.98 (0·96, 1.00)	0·057	0.068
p<0·01		0.96 (0·94, 0.97)	**$1.1x10^{-4}$**	**$1.65x10^{-4}$**
p<0·05		0.96 (0·94, 0.97)	**$6.5x10^{-7}$**	**$1.76x10^{-6}$**
p<0·1		0.95 (0·94, 0.97)	**$2.2x10^{-8}$**	**$1.32x10^{-7}$**
p<0·5		0.96 (0·94, 0.97)	**$8.8x10^{-7}$**	**$1.76x10^{-6}$**
$p<5x10^{-8}$	Definite evening	1·01 (0·99, 1·03)	0·610	0.610
$p<5x10^{-5}$		1·02 (1.00, 1·04)	0·057	0.068
p<0·01		1·04 (1.02, 1·06)	**$1.1x10^{-4}$**	**$1.65x10^{-4}$**
p<0·05		1·05 (1·03, 1·06)	**$6.5x10^{-7}$**	**$1.76x10^{-6}$**
p<0·1		1·05 (1·03, 1·07)	**$2.2x10^{-8}$**	**$1.32x10^{-7}$**
p<0·5		1·05 (1·03, 1·06)	**$8.8x10^{-7}$**	**$1.76x10^{-6}$**
$p<5x10^{-8}$	Overall morning	1.00 (0·99, 1·00)	0·700	0.840
$p<5x10^{-5}$		1.00 (0·99, 1·01)	0·860	0.860
p<0·01		0·99 (0·98, 1·00)	**0·0054**	**0.0081**
p<0·05		0.98 (0·97, 0.99)	**$3.4x10^{-5}$**	**$6.8x10^{-5}$**
p<0·1		0.98 (0·97, 0.99)	**$3x10^{-6}$**	**$9x10^{-6}$**
p<0·5		0.98 (0·97, 0.99)	**$8.9x10^{-7}$**	**$5.34x10^{-6}$**
$p<5x10^{-8}$	Overall evening	1·00 (0·99, 1·01)	0·700	0.840
$p<5x10^{-5}$		1·00 (0·99, 1·01)	0·860	0.860
p<0·01		1·01 (1·01, 1·02)	**0·0054**	**0.0081**
p<0·05		1·02 (1·01, 1·03)	**$3.4x10^{-5}$**	**$6.8x10^{-5}$**
p<0·1		1·02 (1·01, 1·03)	**$3x10^{-6}$**	**$9x10^{-6}$**
p<0·5		1·02 (1·01, 1·03)	**$8.9x10^{-7}$**	**$5.34x10^{-6}$**

7.5.4 Association between PRS for low RA and sleep phenotypes in ALSPAC

As there was a positive finding between low RA and chronotype in UK Biobank, it was of interest to assess for an association between low RA and sleep phenotypes in ALSPAC. However, there were no significant findings using the primary or secondary sleep phenotype data in ALSPAC (Table 7.4). There were relatively few individuals with both sufficient genotype data for low RA PRS and

sleep phenotype data. Therefore, the analysis may have been underpowered to detect effects of low RA PRS in the ALSPAC cohort.

Table 7.4 Associations between low RA PRS and mother-reported sleep traits in ALSPAC

PRS p threshold	**Outcome**	**O.R (95% CI)**	**p uncorrected**	**p FDR corrected**
$p<5x10^{-5}$	Difficulty sleeping-10	0·91 (0·79,1·05)	0·202	0.960
p<0·01		0·94 (0·78,1·13)	0·506	0.960
p<0·05		0.97 (0·80, 1·17)	0·757	0.960
p<0·1		1.00 (0·82, 1·21)	0·960	0.960
p<0·5		1·03 (0·85, 1·24)	0·770	0.960
$p<5x10^{-5}$	Difficulty sleeping-13	0.88 (0·76, 1·02)	0·093	0.465
p<0·01		0.89 (0·74, 1·08)	0·237	0.505
p<0·05		1·01 (0·83, 1·23)	0·906	0.906
p<0·1		1·02 (0.83, 1·25)	0·860	0.906
p<0·5		0.90 (0.76, 1·10)	0·303	0.505
PRS p threshold	**Outcome**	**Beta (95% CI)**	**p uncorrected**	**p FDR corrected**
$p<5x10^{-5}$	Difficulty sleeping-score	-0·03 (-0·06, 0·01)	0·134	0.335
p<0·01		-0·04 (-0.09, 0·01)	0·072	0.335
p<0·05		-0·03 (-0·08, 0·02)	0·273	0.455
p<0·1		-0·01 (-0·06, 0·04)	0·623	0.740
p<0·5		-0·01 (-0·05, 0·04)	0·740	0.740
$p<5x10^{-5}$	Worried sleep	0·02 (-0·04, 0·01)	0·149	0.745
p<0·01		0·09 (-0·04, 0·02)	0·560	0.970
p<0·05		0·11 (-0·02, 0·03)	0·756	0.970
p<0·1		0·10 (-0·03, 0·03)	0·961	0.970
p<0·5		0·03 (-0·03, 0·03)	0·970	0.970

7.6 Discussion

As noted previously (Chapter 6), there was little evidence of genetic correlation between low RA and psychiatric phenotypes (Ferguson et al. 2018). This is perhaps surprising given the cross-sectional, observational associations in the literature on circadian rhythmicity and mood disorders, although a lack of genetic correlation in the case of exposure/outcome associations could imply direct causality – rather than shared genetic architecture. Many features of circadian rhythmicity have been associated with BD and depression, including low RA and circadian phase preference (chronotype) (Lyall et al. 2018; Baek et al. 2016; Dmitrzak-Węglarz et al. 2016). Also, core circadian clock genes have been associated with both BD and depression, with altered circadian biology suggested to be a vulnerability marker for mood disorders (Partonen 2012; Etain et al. 2011; Liberman et al. 2018; Geoffroy et al. 2014; McCarthy et al. 2013; Ferguson et al. 2018).

However, as described in Chapter 6 and above, genetic correlations may be limited by the design of both the preliminary GWAS of low RA and GWAS of traits-of-interest. There is also potential that the genetic overlap between circadian rhythms and mood disorders are due to the variants underlying circadian rhythmicity traits other than RA. The GWAS were also unable to detect rare variants (occurring in less than 1% of the population); some of the genetic overlap could be due to rare variants with moderate effects (Manolio et al. 2009).

It has been suggested that the treatment of disrupted circadian rhythmicity could be used in combination with current pharmaceutical therapies to develop more effective treatments for mood disorders (Bellivier et al. 2015). Therefore, a more complete understanding of circadian rhythmicity in the context of mood disorders is important clinically (Ferguson et al. 2018).

Within the UK Biobank cohort lower RA has been associated with prevalent adverse mental health (Lyall et al. 2018). In the current study, some evidence was found for an association between greater polygenic risk for low RA and both MDD and neuroticism (in independent sub-samples of the cohort). Across several PRS thresholds, there was a more robust association between increasing PRS for low RA and the phenotype of mood instability. Mood instability is a common

symptom that cuts across several psychiatric disorders and, as potential RDoC trait, may be a more useful phenotype than categorical diagnoses for understanding underlying biology (Broome et al. 2015; Insel 2014). This also demonstrates a possible direct link between genetic loading for circadian disruption and the subjective experience of dysregulated or unstable mood. For this reason, the observed association is of interest and merits further investigation (Ferguson et al. 2018).

As relapse occurs often in mood disorders (Hofmann et al. 2012) it may in future be useful to monitor rest-activity cycles, for example, of individuals with known higher PRS, to potentially pre-emptively treat or manage incipient mood disorder relapses. Again, more complete investigations into these relationships are required, as the mood instability phenotype used here was based on a self-reported subjective measure that may be influenced by response bias (Ferguson et al. 2018).

The few significant associations found between PRS and lifetime BD and MDD could be partially due to the possible under-representation of individuals with psychiatric disorders in UK Biobank; UK Biobank may not be fully representative of the general UK population (Fry et al. 2017).

The association between low RA PRS and chronotype also appears to strengthen the evidence for the involvement of genetic variants associated with low RA in circadian rhythms. It should be noted however that the risk scores showed relatively small effects on the traits of interest and overall explain only a small proportion of the variance within the traits.

When testing the PRS for low RA in ALSPAC there were no significant associations identified with most of the mood trait phenotypes tested with the exception of weak associations between low RA PRS and depression score. These analyses may have been underpowered because of the relatively small sample size and there is some uncertainty about the validity of self-reported mood phenotypes, such that a potential influence of genetic loading for low RA on hypomania and depression cannot be completely dismissed. The small sample size also means that outlying or unrepresentative participants are more prominent, leading to potential type-1 error and spurious findings.

There were no associations found between the low RA PRS and sleep outcomes in ALSPAC and, as above, the small sample size available is an obvious limitation. Another limitation of ALSPAC is the reliance on self-report and mother-reported variables which could be influenced by response biases. The collection of objective measures of sleep and circadian function in ALSPAC, similar to those used in UK Biobank, would be useful in the longer-term.

It is possible that genetic variants affecting circadian rhythmicity may have an unknown developmental-specific effect on how mood disorder-related traits manifest in adolescence and adulthood; although as there is currently no way to verify this using data which is currently available. There are also limitations to this PRS analysis, as only 3 genome-wide significant SNPs were included for the most stringent PRS threshold the analysis is likely to be underpowered; the lack of associations found between mood phenotypes and genome-wide PRS threshold could be in part due to this lack of statistical power.

These preliminary results add some support to disrupted circadian rhythmicity as a potential endophenotype of mood disorders; which could eventually be a target for treatment. Several studies have demonstrated clear differences in motor activity between mood disorder patients and healthy controls (Scott et al. 2017). Investigations of actigraph data have reported differences distinguishing MDD and BD patients however it is of note that these studies were limited by relatively small sample sizes (Scott et al. 2017). Nevertheless, the potential of activity monitoring as a diagnostic tool or for the observation of treatment response could be beneficial to patients with greater genetic loading for circadian disruption.

7.7 Future work

The current analyses are not adjusted for environmental factors which may influence circadian rest-activity cycles, such as, medication status, irregular work patterns and chronic pain status. Further study is required to look at the potential interaction between genetic variants (risk scores) and environmental factors, and how they associate with psychiatric phenotypes.

As described in Chapter 6, further study is required to gain a better understanding of the genetic architecture of circadian rhythmicity. Larger GWAS could provide many more variants for inclusion in PRS which then may be more

effective at assessing the associations between circadian rhythmicity and mood disorders.

The above PRS analyses require replication in cohorts, such as 23andMe and All of Us, which may be more representative of the general population.

7.8 Conclusions

Overall, the findings above contribute new knowledge on how the underlying genetics of circadian rhythmicity (namely, the rest-activity cycle measure of relative amplitude) overlaps with mood disorder phenotypes, particularly mood instability.

Chapter 8 Conclusions

8.1 Overview of main findings

Mood disorders are highly prevalent and are a leading cause of disability. They comprise many different features and how these features impact on pathophysiology is currently not well-understood. Specifically, disrupted circadian rhythmicity (for example, disrupted daily activity and altered sleep patterns) is a core feature of mood disorders but the relationship between circadian function and mood disorders is unclear. This thesis aimed to contribute new knowledge on the genetic architecture of mood disorders and circadian rhythms.

This thesis has attempted to understand this relationship from a genetic perspective, with a focus on polygenic risk scores.

In Chapter 3, as an exploratory analysis, GPRS was used to investigate the influence of the BD candidate risk gene *CACNA1C* on mood disorder-related traits in UK Biobank and in ALSPAC. There were no clear associations identified in either cohort. However, when investigating the samples separately by sex there appeared to be a weak association between *CACNA1C* and BD traits in females within both cohorts (although according to the MHQ-defined outcomes available for some UK Biobank participants there were no associations).

In Chapter 4, I investigated the relationship between circadian function and mood disorders more directly. Associations were found between chronotype PRS and mood disorder-related traits suggesting a potential link between diurnal preference and mood disorders. A limitation of using PRS for chronotype is the fact that it is based on a self-reported, subjective measure which can be subject to reporting bias.

In Chapter 4, an exploratory MR analysis was used to investigate the potential causal relationship between chronotype and depression. Depression was found to be causative for eveningness (as expected) however morningness was found to have a causative effect on depression. This was contradictory to the phenotypic study undertaken in UK Biobank which reported that greater morningness PRS was associated with decreased risk of depression and may suggest a degree of reverse causality in the observational correlations. This finding also did not align

with another MR study of chronotype and depression which reported that morningness SNPs were associated with decreased depression risk (Jones et al. 2019). The findings of both that study and my analysis may indicate a bidirectional relationship (Landgraf et al. 2014). Clearly, further investigation is required to establish the true directions of this relationship.

As *CACNA1C* has been reported to be involved in circadian function (Schmutz et al. 2015; McCarthy et al. 2016), this GPRS was also tested for associations with both chronotype preference (morningness and eveningness) and RA (a marker of circadian rhythmicity). It is possible that previously-identified associations between *CACNA1C* and mood disorders were occurring via the gene's influence of circadian function. There were no associations between the *CACNA1C* GPRS and both subjective and objective measures of circadian rhythmicity within ALSPAC and UK Biobank, as described within Chapter 5.

At the time of the chronotype-focussed study there were no large-scale GWAS of objective circadian parameters available from which to generate PRS of circadian dysfunction but with the release of the UK Biobank accelerometer data I was able to conduct a GWAS using RA, a derived objective rest-activity measure that is less likely to be affected by response bias compared to subjective methods. The case-control GWAS of low RA identified three variants associated with low RA and a secondary GWAS of continuous RA identified several variants associated with RA. As detailed in Chapter 6, the variants associated with low RA highlighted a potential biological link between rest-activity cycles and a replicated BD candidate risk gene (*ANK3*) and *SLC25A17* was associated with BD in a recent GWAS (Stahl et al. 2017).

When investigating low RA PRS in Chapter 7, a significant association was only found between greater PRS for disrupted rest-activity cycles (low RA) and mood instability. However, the chronotype PRS has greater statistical power to detect associations as the variants were identified in GWAS using larger sample sizes with a better balance of case and controls (evening chronotype vs morning chronotype) relative to the GWAS of low RA. This highlights the requirement of a larger sample with accelerometer data as well as a greater balance of individuals with low RA and higher RA than the GWAS undertaken in Chapter 6; with this type of sample cohort more variants associated with disrupted rest-

activity cycles could be identified and improved PRS generated to better explore the link between mood disorders and disrupted circadian rhythmicity.

8.2 Contribution to the literature

1. The lack of evidence for an association between *CACNA1C* GPRS and mood traits in ALSPAC and UK Biobank suggests that focussing on single candidate risk gene is ineffective for evaluating mood disorder risk in the general population.

2. With the emergence of large population cohorts, including UK Biobank and 23andMe, which have information on self-reported subjective circadian measures, several GWAS of chronotype preference have now been published (Lane et al. 2016; Jones et al. 2016b; Hu et al. 2016). This allowed for one of the first assessments using morningness and eveningness PRS to study the relationship between chronotype preference and mood disorders at a genetic level. This analysis identified an association between greater eveningness PRS and BD.

3. The underlying genetic architecture of circadian rhythmicity is complex and not well understood. However, as mentioned above, with the release of UK Biobank accelerometer data the largest GWAS of an objective circadian measure (RA) was undertaken. This GWAS highlighted a variant within the *NFASC* gene; as described in chapter 6, neurofascin (the protein product of the *NFASC* gene) physically interacts with ankyrin G. The gene which encodes this product (*ANK3*) is a replicated BD candidate risk gene (Ferreira et al. 2008; Sklar et al. 2011). The *NFASC* had not been implicated in mood disorders previously, however, this finding suggests its potential involvement in the pathophysiology of BD. The *SLC25A17* gene which was also identified in the GWAS of low RA has reported to be associated with BD in a recent GWAS (Stahl et al. 2017).

The associations identified in this thesis provide new knowledge of the potential biology underlying the pathophysiology of mood disorders and provide some support for a biological relationship between disrupted circadian function and mood disorders.

8.3 Implications

As reported in Chapters 4 and 7, PRS of circadian features were associated with features of mood disorders including mood instability, hypomania and BD. In the future, using circadian PRS alongside clinical data may help to provide a better understanding of the specific aspects of an individual's mood disorder and potentially predict their likely clinical presentations.

If PRS can be further developed to explain a greater proportion of the variance seen it may be more directly clinically applicable (Dudbridge 2013). Using a circadian PRS could potentially indicate how patients may respond to different treatments or management strategies and therefore treat BD and MDD with greater efficiency than is seen with current treatments. For example, schizophrenia PRS has been shown to predict psychotic symptoms in BD (Hamshere et al. 2011); applying circadian PRS in this way may be able to predict specific features of BD or MDD.

In the future PRS may be clinically useful; more research is required to establish how these PRS associate with mood traits in clinical populations. Further work is also required to understand the potential influence PRS could have on treatment response, for example because the most effective dose of pharmaceutical treatments could be dependent on the patient's genotype. Genes found to associate with depression have shown to be potential targets for existing drugs, hence, genotyping could be useful to develop the most appropriate treatment for an individual (Howard et al. 2019).

Potentially developing tailored genotyping arrays to identify an individual's polygenic loading for risk variants and calculating several risk scores for example, BD, MDD, low RA and chronotype PRS could provide clinicians with useful information. A larger risk score containing multiple PRS may have better ability to predict an individual's risk (discussed in more detail in section 8.5.4 below) (Krapohl et al. 2018).

Understanding a patient's specific risk scores may allow for the design of personalised treatment plans. Using an individual's genotype to predict risk and provide the appropriate treatment has been successfully included in clinical practice for treating cancer patients, particularly breast and ovarian cancer (Tung & Garber 2018).

There is increasing evidence of the involvement of circadian dysfunction in the pathophysiology of mood disorders. A composite assessment of circadian rhythm, including polygenic risk scores, rest-activity measures, sleep quality and chronotype measures may be beneficial to inform more individual-based treatment and management strategies which could help to maintain remission and reduce relapse in mood disorders (Dong et al. 2019).

One of the main considerations in psychiatric research is how findings could eventually be applied to inform faster and more efficient diagnoses and provide information for the development of effective treatments.

An example is the identification of objective biomarkers which could not only identify individuals with psychiatric conditions but also aid in distinguishing disorders. These biomarkers could also be potential biological targets for treatments (Phillips & Kupfer 2013; Ruderfer et al. 2014).

The variants included in PRS which associate with mood disorders could highlight genes whose products could be viable targets for future treatments. A recent study demonstrated it may be possible to identify new treatment targets by investigating the genes implicated by variants associated with MDD. The study used findings from the latest PGC GWAS to integrate gene expression information and drug-target networks to indicate which genes are likely to be affected by a chosen treatment (Gaspar et al. 2019). Although this study highlighted several potentially targetable genes which were enriched in MDD the results require further investigation and replication. In the future this could be a useful method for identifying novel treatment targets from GWAS data (Gaspar et al. 2019).

Integration of many different diagnostic tools which encompass lifestyle assessments, clinical presentation and biological measures could be a better representation of an individual's specific pathophysiology (Phillips & Kupfer 2013; Ruderfer et al. 2014); for example, accelerometer, genotype and chronotype data could be integrated with other clinical assessments to improve diagnoses and allow for more personalised treatment.

8.4 Limitations

8.4.1 Limitations: polygenic scores

A strength of this thesis is the use of novel PRS to investigate the relationship between circadian genetics and mood disorders. However, this is also a potential limitation. As this is the first use of these PRS they have not been externally validated (although the PRS found associations in both UK Biobank and ALSPAC).

It is of note that the PRS used had only a small effect on the mood phenotypes tested. Replication of the findings of this study and a PRS which has a greater effect on mood phenotypes may be required before any definitive conclusions can be drawn. Currently, PRS may be more useful for associations tests, as they have been used in this thesis, than for predicting disease (Dudbridge 2013). Using a less conservative PRS could improve the accuracy of predicting individuals at risk, however the PRS should not be so extensive as to include uninformative variants that are identified by GWAS at higher p-values (Stocker et al. 2018). For example, studies have found that including SNPs with GWAS p value of <0.01 in PRS demonstrated better distinction between healthy controls and Alzheimer patients than more conservative p value thresholds (Stocker et al. 2018). Current GWAS do not have large enough case sample sizes to identify more risk variants, therefore PRS could also be further expanded by undertaking GWAS using increased case sample sizes (Dudbridge 2013). With replication, further findings of GWAS and less conservative p value thresholds PRS could be expanded to increase predictive accuracy and clinical utility (Stocker et al. 2018; Dudbridge 2013).

As detailed in Chapters 4-7, there are limitations to both the subjective and objective circadian measures used in the chronotype and low RA PRS analyses, and the GWAS of low RA. The derived RA measure – although exposed to various QC measures – was not adjusted for some potentially confounding variables (described in Chapter 6). However, it should also be noted that over-adjusting for confounding factors could result in false negatives and variants influencing RA may have been missed (Aschard et al. 2015).

8.4.2 Limitations: cohorts used

The mood phenotypes investigated in ALSPAC and UK Biobank are based on self-report or structured interviews which could be sensitive to responder bias and are a limitation to this study as the outcomes could be under- or over-representative of the true incidences of mood disorders (Ganna & Ingelsson 2015). Being able to make use of formal diagnoses of BD and MDD to investigate the relationships between the PRS and mood disorder would have been desirable.

In the case of ALSPAC, primary and secondary care data has been collected on a subset of participants; however these data are still being processed by ALSPAC and were not available for use by researchers during the course of this study (http://www.bristol.ac.uk/alspac/researchers/our-data/linkage/).

There is currently no primary care data in UK Biobank available to researchers, however, standardised primary care data with information on participants diagnoses, treatments etc. should be available in the near future (https://www.ukbiobank.ac.uk/wp-content/uploads/2018/12/Primary-Care-Data.pdf). Unfortunately, this standardised primary health care data was not available during the time frame of this thesis.

With the available data it is not possible to reliably separate individuals in UK Biobank with different clinical subtypes of BD. Even though there are potential biases and limitations for clinically-defined subtypes of BD (as described in Chapter 1), discrepancies in clinical presentations may indicate differences in the biological underpinnings between the subtypes. There is evidence to support genetic differences between BDI and BDII, BDI has a relatively strong genetic correlation with schizophrenia, whereas BDII is more strongly correlated with MDD (Stahl et al. 2017). The BD outcomes used for the analyses in this thesis could not be separated into different BD classifications. It is possible more subtle associations between PRS and specific BD subtypes could be lost.

As the mental health data in both ALSPAC and UK Biobank are cross-sectional, there are limitations to the conclusions that can be drawn. Cross-sectional data is taken from certain time points and so does not account for possible changes over time: the data may not be representative of the true incidences of the

phenotypes-of-interest. Cross-sectional data also make causal inferences difficult (Bowen & Wiersema 1999). The cohorts represent samples of two different age groups (adolescent/young adult and older adult/elderly) and the results obtained in this thesis have sometimes been in opposing directions when comparing ALSPAC and UK Biobank. It is possible there are some differential effects of the PRS at different developmental stages. However, as there was little longitudinal data it is difficult to investigate how the influence of PRS may change over an individual's life course.

UK Biobank is not representative of the general UK population. Of the approximate 9.2 million UK residents invited to join the cohort there was only a 5.5% response rate; those who volunteered to join were more likely to be female, older, better educated, living in less socio-economically deprived areas and with fewer health conditions compared to the general population (Fry et al. 2017).

Also, as the ALSPAC cohort participants were recruited from a relatively small area of the UK it may not be representative of the wider population. Therefore, the findings from UK Biobank and ALSPAC may not be generalisable.

8.5 Possible future work

8.5.1 Replication of findings

The GWAS of RA and low RA detailed in Chapter 6 were the first large-scale investigations of the genetics of rest-activity rhythmicity. These results therefore must be considered preliminary and require replication. As there was an uneven proportion of cases compared to controls in the low RA GWAS, a replication GWAS could be undertaken in UK Biobank if more participants could be recruited for accelerometer data collection in the future. If a better balance of cases and controls could be obtained, a less conservative GWAS method could be used (such as BOLT-LMM which was used for the GWAS of RA); also, a greater number of cases would increase power to detect variants of low effect.

A replication GWAS of low RA able to identify a greater number of associated variants would also allow for an investigation of the potential causal relationship (and the direction of the causality) between low RA and mood disorders. This could serve as a preliminary study into the causal relationship

between circadian disruption and mood disorders.

Also, with the release of more up-to-date GWAS of BD and other mood disorders this causal relationship could be investigated in more disorders; depression was used to investigate causal relationships as it has relatively current GWAS findings compared to BD.

As mentioned above, there is currently no primary care data available in ALSPAC or UK Biobank, however these data are reportedly available for use soon. It would be of interest to test the associations between the PRS used in the previous chapters and primary-care diagnosed cases of BD and MDD in these cohorts.

There are several cohorts that contain both genetic data and a variety of phenotypic information, including self-report and clinical psychiatric measures, for example 23andMe and Generation Scotland (Eriksson et al. 2010; Smith et al. 2013b). A replication study could be undertaken to investigate the association between the risk scores generated in this thesis and the self-reported mood phenotypes available in these cohorts.

Some population cohorts, such as deCODE (Balkau 2000), also contain both genetic data and phenotypic information obtained from electronic health records. The benefit of these cohorts is the linked data and variety of phenotypic information available which could provide replication cohorts for testing the PRS used in this thesis.

In order to investigate how the PRS may influence the development of mood disorders, consistent mental health data collected at different stages of life is required. An example of a longitudinal cohort with both mental health and genotype data is Add Health (National Longitudinal Study of Adolescent to Adult Health); PRS could be tested for their associations to self-report mood outcomes at various ages to investigate the potential developmental influences of PRS (Evans & Erickson 2019).

However, these cohorts do not include the extensive phenotyping of objective circadian measures required to replicate the analyses detailed in Chapters 6 and 7.

Also, with the use of growing emergence of “mobile health data”, there is potential to apply information on both the subjective and objective measures generated by these accessible technologies and software to large cohorts to

further explore the relationship between circadian function and mood disorders (Merikangas et al. 2019). Developing cohorts, for example All of Us (https://allofus.nih.gov/), advertise a wide variety of data which links basic lifestyle/demographics, genotype data, self-reported health, electronic health records and mobile health data. It would be of interest to investigate circadian measures derived from this mobile health data in the context of psychiatric and mood disorders.

8.5.2 Investigating within clinical samples

The analyses undertaken in this thesis focussed on birth and population cohorts which could be under-representative of individuals with psychiatric conditions. Future analyses could be applied to clinical populations to investigate the applicability of circadian PRS for predicting mood disorder risk. Hypothetically, applying the chronotype or circadian PRS to clinical subsets may be useful for investigating treatment responses. There is recent evidence of chronotype preference and circadian preference associating with lithium response in BD patients on lithium monotherapy (N=193) (McCarthy et al. 2018b). Non-responding patients were more likely to report an evening chronotype compared to patients responding to lithium. A subset of patients included in this study donated fibroblasts which allowed for the investigation internal circadian rhythm measures. Lithium non-responders also displayed greater circadian dysregulation, such as longer circadian period, than lithium responders (McCarthy et al. 2018b). There is potential that using circadian PRS may predict patient response to lithium before subjecting patients to the treatment. Providing appropriate treatment to patients could improve patient QOL and reduce morbidity (Phillips & Kupfer 2013; Bauer et al. 2018).

PRS may also be useful in determining which patients may benefit most from psychosocial or lifestyle interventions. Individuals with high polygenic loading for circadian dysfunction could potentially benefit from treatment interventions focussed on establishing better circadian rhythmicity. Interpersonal and social rhythm therapy (IPSRT) was designed specifically for BD patients based on the hypothesis of circadian disruption and abnormal sleep-wake cycles are involved in the symptomology of BD (Frank et al. 2000). IPSRT is a psychotherapy, to be implemented in conjunction with medication, using various behavioural

techniques to aid patients in dealing with interpersonal changes and regiment their daily routines. There is some epidemiological evidence that mood disorder patients benefit from strict, rhythmic schedules; however, it is worth noting that this evidence is based on relatively small-scale studies (Frank et al. 2000).

By investigating circadian PRS within a clinical population it may also be possible to focus on specific characteristics of mood disorders to establish whether the PRS would be useful in predicting a patient's clinical presentation. For example, does greater polygenic loading for circadian disruption associate with greater risk of patient relapse or poorer prognosis?

8.5.3 Gene-environment interactions

Mood disorders are known to have both genetic and environmental inputs (Yoshimizu et al. 2015; Wray et al. 2018), however, the relationships between these risk factors are currently unclear. As this thesis describes one of the few incidences of using circadian measure PRS it is also unknown at present how the potential interactions between these risk scores and environmental risk factors influence the risk of mood disorders. Some environmental exposures may be more detrimental to patients with greater polygenic loading for circadian dysfunction. Therefore, patients with both high risk scores and exposure to specific environments or life events could have more severe presentations than patients with high risk scores but no exposure to the environmental risks.

8.5.4 Combining PRS

The associations identified between the PRS and mood phenotypes have relatively small effects on the phenotypes and more work is needed to further develop the PRS in order to explain a greater proportion of the trait variance. With development the PRS could have a more effective clinical application than as they stand currently (Dudbridge 2013). It would be of interest to generate a combined risk score which contains both variants associated with BD or MDD and variants associated with circadian disruptions.

A recent study has demonstrated that using a risk score composed of multiple PRS is able to give better prediction of phenotypic variation compared to a single PRS. A multiple PRS approach could be useful in providing individual-

specific estimates of risk (Krapohl et al. 2018). This method suggests that even risk scores of traits which are only slightly genetically related could be combined to increase predictive power. The combined power of multiple PRS obtained from many discovery GWAS could explain a greater proportion of trait variance (Krapohl et al. 2018). An expanded risk score may explain more of the phenotypic variance seen in complex disorders and could be more efficient at predicting individuals at risk of developing mood disorders.

It is theoretically possible that using several risk scores (i.e. using RA, chronotype, BD and MDD PRS in conjunction) to determine an individual's susceptibility to mood disorders, potential disease prognosis or treatment response could improve patient outcomes (Hamshere et al. 2011).

8.6 Conclusions

The literature on mood disorder genetics has historically suffered from low quality, small sample size data. This thesis made use of relatively large high-quality datasets and identified several associations which strengthen the theory of a relationship between circadian function and the risk of mood disorders. This study was able to use new PRS to identify links between both subjective and objective circadian measures, and mood disorder-related phenotypes, and also identified genetic variation associated with disrupted rest-activity cycles in the first large-scale GWAS of an objective circadian measure. However, these findings require replication.

In conclusion, this thesis has provided new and supporting evidence of the genetic relationship between circadian function and mood disorders and argues for further research to develop a greater understanding of how these relationships influence the development of mood disorders and how they may inform the development of novel treatment targets.

Appendix

During the course of this PhD thesis I was also involved in producing the following publications, each of these publications was important for the progress of this thesis:

- Strawbridge. R.J. et al. 2019. Identification of novel genome-wide associations for suicidality in UK Biobank, genetic correlation with psychiatric disorders and polygenic association with completed suicide. *EBioMedicine*, 41, pp.517-525.

- Ward. J. et al. 2018. Polygenic risk scores for major depressive disorder and neuroticism as predictors of antidepressant response: Meta-analysis of three treatment cohorts. *PLoS One,* 13(9).

- Strawbridge. R.J. et al. 2018. Genetics of self-reported risk-taking behaviour, trans-ethnic consistency and relevance to brain gene expression. *Translational Psychiatry*, 8(1), pp.178.

- Ferguson. A. et al. 2018. Genome-wide association study of circadian rhythmicity in 71,500 UK Biobank participants and polygenic association with mood instability. *EBioMedicine,* 35, pp.279-287

- Lyall. L.M. et al. 2018, Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: a cross-sectional study of 91,105 participants from the UK Biobank. *Lancet Psychiatry*, 5(6), pp.507-514.

- Strawbridge. R.J. et al. 2018. Genome-wide analysis of self-reported risk-taking behaviour and cross-disorder genetic correlations in the UK Biobank cohort. *Translational Psychiatry*, 8(1), pp.39.

- Lyall. L.M. et al. 2018. Seasonality of depressive symptoms in women but not men: A cross-sectional study in the UK Biobank cohort. *Journal of Affective Disorders*, 229, pp.296-305.

- Ward. J. et al. 2017, Genome-wide analysis in UK Biobank identifies four loci associated with mood instability and genetic correlation with major depressive disorder, anxiety disorder and schizophrenia. Translational Psychiatry, 7(11), pp.1264.

List of References

Agrimi, G. et al., 2012. The human gene *SLC25A17* encodes a peroxisomal transporter of coenzyme A, FAD and NAD $^{+}$. *Biochemical Journal*, 443(1), pp.241–247. Available at: http://biochemj.org/lookup/doi/10.1042/BJ20111420.

Albrecht, U., 2012. Timing to Perfection: The Biology of Central and Peripheral Circadian Clocks. *Neuron*, 74(2), pp.246–260. Available at: https://www.sciencedirect.com/science/article/pii/S0896627312003327.

Alloy, L.B. et al., 2017. Circadian Rhythm Dysregulation in Bipolar Spectrum Disorders. *Current Psychiatry Reports*.

Asaad, T. et al., 2016. Polysomnographic characteristics of bipolar hypomanic patients: Comparison with unipolar depressed patients. *Journal of Affective Disorders*, 191, pp.274–279. Available at: http://dx.doi.org/10.1016/j.jad.2015.12.001.

Aschard, H. et al., 2015. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *American journal of human genetics*, 96(2), pp.329–39. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25640676.

Baek, J.H. et al., 2016. Lifetime Characteristics of Evening-Preference and Irregular Bed-Rise Time Are Associated With Lifetime Seasonal Variation of Mood and Behavior: Comparison Between Individuals With Bipolar Disorder and Healthy Controls. *Behavioral Sleep Medicine*, 14(2), pp.155–168.

Balkau, B., 2000. The DECODE study. Diabetes epidemiology: collaborative analysis of diagnostic criteria in Europe. *Diabetes & metabolism*, 26(4), pp.282–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11011220.

Baryshnikov, I. et al., 2015. Differences and overlap in self-reported symptoms of bipolar disorder and borderline personality disorder. *European psychiatry : the journal of the Association of European Psychiatrists*, 30(8), pp.914–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26647866.

Bauer, M. et al., 2018. Areas of uncertainties and unmet needs in bipolar disorders: clinical and research perspectives. *The Lancet Psychiatry*, 5(11), pp.930–939. Available at: https://www.sciencedirect.com/science/article/pii/S2215036618302530?via%3Dihub.

Bellivier, F. et al., 2015. Sleep- and circadian rhythm–associated pathways as therapeutic targets in bipolar disorder. *Expert Opinion on Therapeutic Targets*, 19(6), pp.747–763. Available at: http://www.tandfonline.com/doi/full/10.1517/14728222.2015.1018822.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B*, 57(1), pp.289–300. Available at: http://www.stat.purdue.edu/~doerge/BIOINFORM.D/FALL06/Benjamini and Y FDR.pdf%5Cnhttp://engr.case.edu/ray_soumya/mlrg/controlling_fdr_benjamini95.pdf.

Berdynaj, D. et al., 2016. Effect of chronotype on emotional processing and risk taking.

Chronobiology International, 33(4), pp.406–418. Available at: http://dx.doi.org/10.3109/07420528.2016.1146739.

Bidaud, I. & Lory, P., 2011. Hallmarks of the channelopathies associated with L-type calcium channels: A focus on the Timothy mutations in Cav1.2 channels. *Biochimie*, 93(12), pp.2080–2086. Available at: http://dx.doi.org/10.1016/j.biochi.2011.05.015.

Bigos, K.L. et al., 2010. Genetic Variation in CACNA1C Affects Brain Circuitries Related to Mental Illness. *Archives of general psychiatry*, 67(9), pp.939–945.

Bischoff, F.R. et al., 1995. Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport. *Proceedings of the National Academy of Sciences of the United States of America*, 92(5), pp.1749–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7878053.

Bowden, J. et al., 2016. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *International journal of epidemiology*, 45(6), pp.1961–1974. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27616674.

Bowen, H.P. & Wiersema, M.F., 1999. Matching method to paradigm in strategy research: limitations of cross-sectional analysis and some methodological alternatives. *Strategic Management Journal*, 20(7), pp.625–636.

Breuer, R. et al., 2018. Detecting significant genotype–phenotype association rules in bipolar disorder: market research meets complex genetics. *International Journal of Bipolar Disorders*, 6(1), p.24. Available at: https://journalbipolardisorders.springeropen.com/articles/10.1186/s40345-018-0132-x.

Broome, M.R. et al., 2015. Mood instability: Significance, definition and measurement. *British Journal of Psychiatry*, 207(4), pp.283–285.

Bulik-sullivan, B. et al., 2015. An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, 47(11), pp.1236–1241. Available at: http://dx.doi.org/10.1038/ng.3406.

Bullock, B. & Murray, G., 2014. Reduced amplitude of the 24 hour activity rhythm: A biomarker of vulnerability to bipolar disorder? *Clinical Psychological Science*, 2(1), pp.86–96.

Burgess, H.J. & Fogg, L.F., 2008. Individual Differences in the Amount and Timing of Salivary Melatonin Secretion S. Yamazaki, ed. *PLoS ONE*, 3(8), p.e3055. Available at: http://dx.plos.org/10.1371/journal.pone.0003055.

Burton, C. et al., 2013. Activity monitoring in patients with depression: A systematic review. *Journal of Affective Disorders*, 145(1), pp.21–28.

Bycroft, C. et al., 2017. Genome-wide genetic data on ~ 500 , 000 UK Biobank participants.

Cabral, M.D. et al., 2010. Knocking down Cav1 calcium channels implicated in Th2 cell

activation prevents experimental asthma. *American Journal of Respiratory and Critical Care Medicine*, 181(12), pp.1310–1317.

Carter, A.J. & Nguyen, A.Q., 2011. Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. *BMC Medical Genetics*, 12(1), p.160. Available at: http://bmcmedgenet.biomedcentral.com/articles/10.1186/1471-2350-12-160.

Castro, J. et al., 2015. Circadian rest-activity rhythm in individuals at risk for psychosis and bipolar disorder. *Schizophrenia Research*, 168(1–2), pp.50–55.

Catterall, W.A., 2011. Voltage-Gated Calcium Channels. *Cold Spring Harbor Perspectives in Biology* , 3(8), pp.1–24. Available at: http://cshperspectives.cshlp.org/content/3/8/a003947.abstract.

Chan, J.W.Y. et al., 2014. Eveningness and Insomnia: Independent Risk Factors of Nonremission in Major Depressive Disorder. *Sleep*, 37(5), pp.911–917. Available at: https://academic.oup.com/sleep/article-lookup/doi/10.5665/sleep.3658.

Charrier, A. et al., 2017. Clock Genes and Altered Sleep – Wake Rhythms : Their Role in the Development of Psychiatric Disorders. *International Journal of Molecular Sciences*, 18(938), pp.1–22.

Corruble, E. et al., 2014. Morningness-eveningness and treatment response in major depressive disorder. *Chronobiology International*, 31(2), pp.283–289.

Court, H. et al., 2014. Improving the psychometric utility of the hypomania checklist (HCL-32): A Rasch analysis approach. *Journal of Affective Disorders*, 152–154(1), pp.448–453. Available at: http://dx.doi.org/10.1016/j.jad.2013.10.014.

Craddock, N. & Sklar, P., 2013. Genetics of bipolar disorder. *The Lancet*, 381(9878), pp.1654–1662. Available at: http://dx.doi.org/10.1016/S0140-6736(13)60855-7.

Cross-Disorder Group of the Psychiatric Genomics Consortium, C.-D.G. of the P.G. et al., 2013. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics*, 45(9), pp.984–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23933821.

Cross-Disorder Group of the Psychiatric Genomics Consortium, C.-D.G. of the P.G., 2013. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*, 381(9875), pp.1371–1379. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23453885.

Cullen, B. et al., 2015. Cognitive function and lifetime features of depression and bipolar disorder in a large population sample: Cross-sectional study of 143,828 UK Biobank participants. *European Psychiatry*, 30(8), pp.950–958. Available at: http://www.sciencedirect.com/science/article/pii/S0924933815001522.

Cuthbert, B.N. & Insel, T.R., 2013. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, 11(1), p.126. Available at: http://bmcmedicine.biomedcentral.com/articles/10.1186/1741-7015-11-126.

Dao, D.T. et al., 2010. Mood disorder susceptibility gene CACNA1C modifies mood-related behaviors in mice and interacts with sex to influence behavior in mice and diagnosis in humans. *Biological Psychiatry*, 68(9), pp.801–810. Available at: http://dx.doi.org/10.1016/j.biopsych.2010.06.019.

Dashti, H. et al., 2018. GWAS in 446,118 European adults identifies 78 genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *bioRxiv*. Available at: https://www.biorxiv.org/content/early/2018/04/19/274977.full.pdf+html.

Davis, K.A.S. et al., 2018. Mental health in UK Biobank : development , implementation and results from an online questionnaire completed by 157 366 participants. *BJPsych Open*, 4, pp.83–90.

de Almeida Mendes, M. et al., 2018. Calibration of raw accelerometer data to measure physical activity: A systematic review. *Gait & Posture*, 61, pp.98–110. Available at: https://www.sciencedirect.com/science/article/pii/S0966636217310573.

de Leeuw, C.A. et al., 2015. MAGMA: Generalized Gene-Set Analysis of GWAS Data H. Tang, ed. *PLOS Computational Biology*, 11(4), p.e1004219. Available at: http://dx.plos.org/10.1371/journal.pcbi.1004219.

Desanctis, M. V, 2017. Circadian principles : Behavioral health implications. *Journal of Applied Behaviour Research*, 24(5), pp.1–17.

Desir, G. V, 2008. Renalase deficiency in chronic kidney disease, and its contribution to hypertension and cardiovascular disease. *Current Opinion in Nephrology and Hypertension*, 17(2), pp.181–185. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18277152.

Dmitrzak-Węglarz, M. et al., 2016. Chronotype and sleep quality as a subphenotype in association studies of clock genes in mood disorders. *Acta Neurobiologiae Experimentalis*, 76(1), pp.32–42.

Doherty, A. et al., 2017. Large scale population assessment of physical activity using wrist worn accelerometers: The UK biobank study. *PLoS ONE*, 12(2), pp.1–14.

Dong, L. et al., 2019. A composite measure of sleep health predicts concurrent mental and physical health outcomes in adolescents prone to eveningness. *Sleep Health*. Available at: https://www.sciencedirect.com/science/article/pii/S2352721818302328.

Dong, M. et al., 2018. Prevalence of suicide attempt in individuals with major depressive disorder: a meta-analysis of observational surveys. *Psychological Medicine*, 4, pp.4–14. Available at: https://doi.org/10.1017/S0033291718002301.

Dudbridge, F., 2016. Polygenic Epidemiology. *Genetic Epidemiology*, 40(4), pp.268–272.

Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. *PLoS genetics*, 9(3), p.e1003348. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23555274.

Duffy, J.F., Rimmer, D.W. & Czeisler, C.A., 2001. Association of intrinsic circadian period

with morningness-eveningness, usual wake time, and circadian phase. *Behavioral neuroscience*, 115(4), pp.895–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11508728.

Duncan, L.E. et al., 2018. Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Molecular Psychiatry*, 23(3), pp.666–673. Available at: http://www.nature.com/articles/mp201777.

Dunn, E.C. et al., 2015. Genetic determinants of depression: recent findings and future directions. *Harvard review of psychiatry*, 23(1), pp.1–18. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25563565.

Ensembl, 2018a. Gene: C3orf62 (ENSG00000188315) - Summary - Homo sapiens - Ensembl genome browser 92. Available at: http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000188315;r=3:49268602-49277909.

Ensembl, 2018b. rs113851554 (SNP) - Linkage disequilibrium - Homo sapiens - Ensembl genome browser 92. Available at: http://www.ensembl.org/Homo_sapiens/Variation/HighLD?db=core;r=2:66522932-66523932;v=rs113851554;vdb=variation;vf=21204591;second_variant_name=rs142412330.

Ensembl, 2018c. rs147964682 (SNP) - Linkage disequilibrium - Homo sapiens - Ensembl genome browser 92. Available at: http://www.ensembl.org/Homo_sapiens/Variation/HighLD?db=core;r=1:204927176-204928176;v=rs147964682;vdb=variation;vf=32409503.

Eriksson, N. et al., 2010. Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits G. Gibson, ed. *PLoS Genetics*, 6(6), p.e1000993. Available at: https://dx.plos.org/10.1371/journal.pgen.1000993.

Erk, S. et al., 2010. Brain Function in Carriers of a Genome-wide Supported Bipolar Disorder Variant. *Archives of General Psychiatry*, 67(8), p.803. Available at: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archgenpsychiatry.2010.94.

Erk, S., Meyer-Lindenberg, A., Schmierer, P., et al., 2014b. Hippocampal and frontolimbic function as intermediate phenotype for psychosis: Evidence from healthy relatives and a common risk variant in cacna1c. *Biological Psychiatry*, 76(6), pp.466–475. Available at: http://dx.doi.org/10.1016/j.biopsych.2013.11.025.

Erk, S., Meyer-Lindenberg, A., Linden, D.E.J., et al., 2014a. Replication of brain function effects of a genome-wide supported psychiatric risk variant in the CACNA1C gene and new multi-locus effects. *NeuroImage*, 94, pp.147–154. Available at: http://dx.doi.org/10.1016/j.neuroimage.2014.03.007.

Etain, B. et al., 2014. Association between circadian genes, bipolar disorders and chronotypes. *Chronobiology International*, 31(7), pp.807–814.

Etain, B. et al., 2011. Genetics of circadian rhythms and mood spectrum disorders. *European Neuropsychopharmacology*, 21, pp.S676–S682. Available at: https://www.sciencedirect.com/science/article/pii/S0924977X11001611?via%3Dihub.

Evans, C.R. & Erickson, N., 2019. Intersectionality and depression in adolescence and early adulthood: A MAIHDA analysis of the national longitudinal study of adolescent to adult health, 1995–2008. *Social Science & Medicine*, 220, pp.1–11. Available at: https://www.sciencedirect.com/science/article/pii/S0277953618306099?via%3Dihub.

Eysenck, S.B.G., Eysenck, H.J. & Barrett, P., 1985. A revised version of the psychoticism scale. *Personality and Individual Differences*, 6(1), pp.21–29.

Fabbri, C. et al., 2018. Pleiotropic genes in psychiatry: Calcium channels and the stress-related FKBP5 gene in antidepressant resistance. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 81, pp.203–210. Available at: http://dx.doi.org/10.1016/j.pnpbp.2017.10.005.

Fabbri, C. & Serretti, A., 2016. Genetics of long-term treatment outcome in bipolar disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 65, pp.17–24. Available at: https://www.sciencedirect.com/science/article/pii/S0278584615300282?via%3Dihub.

Faul, F. et al., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), pp.175–191. Available at: http://www.springerlink.com/index/10.3758/BF03193146.

Faul, F. et al., 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behaviour Research Methods*, 41(4), pp.1149–1160. Available at: http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPower31-BRM-Paper.pdf.

Ferguson, A. et al., 2018. Genome-Wide Association Study of Circadian Rhythmicity in 71,500 UK Biobank Participants and Polygenic Association with Mood Instability. *EBioMedicine*, 35, pp.279–287. Available at: http://www.ncbi.nlm.nih.gov/pubmed/30120083.

Ferrari, A.J. et al., 2013. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS medicine*, 10(11), p.e1001547. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24223526.

Ferreira, M.A.R. et al., 2008. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nature genetics*, 40(9), pp.1056–8. Available at: http://dx.doi.org/10.1038/ng.209.

Feuk, L., Carson, A.R. & Scherer, S.W., 2006. Structural variation in the human genome. *Nature Reviews Genetics*, 7(2), pp.85–97. Available at: http://www.nature.com/articles/nrg1767.

Fiorentino, A. et al., 2014. Analysis of ANK3 and CACNA1C variants identified in bipolar disorder whole genome sequence data. *Bipolar Disorders*, (7), pp.583–591.

Forstner, A.J. et al., 2017. Identification of shared risk loci and pathways for bipolar disorder and schizophrenia. *PLoS ONE*, 12(2), pp.1–14.

Frank, E., Swartz, H.A. & Kupfer, D.J., 2000. Interpersonal and social rhythm therapy:

managing the chaos of bipolar disorder. *Biological Psychiatry*, 48(6), pp.593–604. Available at: https://www.sciencedirect.com/science/article/pii/S0006322300009690?via%3Dihub.

Frohnert, B.I. et al., 2018. Prediction of type 1 diabetes using a genetic risk model in the Diabetes Autoimmunity Study in the Young. *Pediatric Diabetes*, 19(2), pp.277–283. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28695611.

Fry, A. et al., 2017. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. *American Journal of Epidemiology*, 186(9), pp.1026–1034. Available at: https://academic.oup.com/aje/article/186/9/1026/3883629.

Ganna, A. & Ingelsson, E., 2015. 5 year mortality predictors in 498,103 UK Biobank participants: a prospective population-based study. *Lancet (London, England)*, 386(9993), pp.533–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26049253.

Gaspar, H.A. et al., 2019. Using genetic drug-target networks to develop new drug hypotheses for major depressive disorder. *Translational Psychiatry*, 9(1), p.117. Available at: http://www.nature.com/articles/s41398-019-0451-4.

GBD, 2016. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*, 388(10053), pp.1545–1602. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27733282.

Geddes, J.R. & Miklowitz, D.J., 2013. Treatment of bipolar disorder. *The Lancet*, 381(9878), pp.1672–1682. Available at: http://dx.doi.org/10.1016/S0140-6736(13)60857-0.

Geoffroy, P.A. et al., 2014. Sleep in remitted bipolar disorder: A naturalistic case-control study using actigraphy. *Journal of Affective Disorders*, 158, pp.1–7. Available at: http://dx.doi.org/10.1016/j.jad.2014.01.012.

Goel, N. et al., 2014. Circadian Rhythms, Sleep Deprivation, and Human Performance. *Prog Mol Biol Transl Sci.*, 119, pp.155–190. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3963479&tool=pmcentrez&rendertype=abstract.

Golding, J. et al., 2001. ALSPAC-The Avon Longitudinal Study of Parents and Children. *Paediatric and Perinatal Epidemiology*, 15(1), pp.74–87. Available at: http://doi.wiley.com/10.1046/j.1365-3016.2001.00325.x.

Gomez-Ospina, N. et al., 2006. The C Terminus of the L-Type Voltage-Gated Calcium Channel CaV1.2 Encodes a Transcription Factor. *Cell*, 127(3), pp.591–606.

Gratten, J. et al., 2014. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nature neuroscience*, 17(6), pp.782–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24866044.

Green, E.K. et al., 2013. Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Molecular psychiatry*, 18(12), pp.1302–7. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3971368&tool=pmcentrez&rendertype=abstract.

GTEx Consortium, T.Gte., 2013. The Genotype-Tissue Expression (GTEx) project. *Nature genetics*, 45(6), pp.580–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23715323.

Hamshere, M.L. et al., 2011. Polygenic dissection of the bipolar phenotype. *The British journal of psychiatry : the journal of mental science*, 198(4), pp.284–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21972277.

Harrison, P.J. et al., 2016. Innovative approaches to bipolar disorder and its treatment. *Annals of the New York Academy of Sciences*, 1366(1), pp.76–89. Available at: http://doi.wiley.com/10.1111/nyas.13048.

Harrison, P.J., 2016. Molecular neurobiological clues to the pathogenesis of bipolar disorder. *Current Opinion in Neurobiology*, 36, pp.1–6. Available at: http://dx.doi.org/10.1016/j.conb.2015.07.002.

Harrison, P.J., Geddes, J.R. & Tunbridge, E.M., 2018. The Emerging Neurobiology of Bipolar Disorder. *Trends in Neurosciences*, 41(1), pp.18–30. Available at: http://creativecommons.org/licenses/by/4.0/.

Hayashi, A. et al., 2015. Calcium-dependent intracellular signal pathways in primary cultured adipocytes and ANK3 gene variation in patients with bipolar disorder and healthy controls. *Molecular Psychiatry*, 20(February), pp.931–940. Available at: http://www.nature.com/doifinder/10.1038/mp.2014.104.

Hayes, J.F. et al., 2016. Childhood interleukin-6 , C-reactive protein and atopic disorders as risk factors for hypomanic symptoms in young adulthood : a longitudinal birth cohort study. *Psychological Medicine*, pp.1–11.

Heilbronner, U. et al., 2015. A common risk variant in CACNA1C supports a sex-dependent effect on longitudinal functioning and functional recovery from episodes of schizophrenia-spectrum but not bipolar disorder. *European Neuropsychopharmacology*, 25(12), pp.2262–2270. Available at: http://dx.doi.org/10.1016/j.euroneuro.2015.09.012.

Heyes, S. et al., 2015. Genetic disruption of voltage-gated calcium channels in psychiatric and neurological disorders. *Progress in Neurobiology*, 134, pp.36–54. Available at: http://dx.doi.org/10.1016/j.pneurobio.2015.09.002.

Hidalgo, M.P. et al., 2009. Relationship between depressive mood and chronotype in healthy subjects. *Psychiatry and Clinical Neurosciences*, 63(3), pp.283–290.

Hofmann, S.G. et al., 2012. EMOTION DYSREGULATION MODEL OF MOOD AND ANXIETY DISORDERS. *Depression and Anxiety*, 29(5), pp.409–416. Available at: http://doi.wiley.com/10.1002/da.21888.

Hou, L. et al., 2016. Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *The Lancet*, 387(10023), pp.1085–1093. Available at: https://www.sciencedirect.com/science/article/pii/S0140673616001434?via%3Dihub.

Howard, D.M. et al., 2019. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, 22(3), pp.343–352. Available at: http://www.nature.com/articles/s41593-018-0326-7.

Hu, Y. et al., 2016. GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nature Communications*, 7, pp.1–9. Available at: http://dx.doi.org/10.1038/ncomms10448.

Insel, T.R., 2014. The NIMH Research Domain Criteria (RDoC) Project: Precision Medicine for Psychiatry. *American Journal of Psychiatry*, 171(4), pp.395–397. Available at: http://psychiatryonline.org/doi/abs/10.1176/appi.ajp.2014.14020138.

Jacquet, B. V et al., 2011. Specification of a Foxj1-dependent lineage in the forebrain is required for embryonic-to-postnatal transition of neurogenesis in the olfactory bulb. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(25), pp.9368–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21697387.

Jakobsson, J. et al., 2015. CACNA1C polymorphism and altered phosphorylation of tau in bipolar disorder. *The British Journal of Psychiatry*. Available at: http://bjp.rcpsych.org/cgi/doi/10.1192/bjp.bp.114.159806.

Jawinski, P. et al., 2015. Brain Arousal Regulation in Carriers of Bipolar Disorder Risk Alleles. *Neuropsychobiology*, 72(2), pp.65–73.

Jones, H.J. et al., 2016a. Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry*. Available at: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2015.3058.

Jones, R.W. et al., 2000. *A new human genetic resource: a DNA bank established as part of the Avon Longitudinal Study of Pregnancy and Childhood (ALSPAC)*, Available at: www.nature.com/ejhg.

Jones, S.E. et al., 2018. Genetic studies of accelerometer-based sleep measures in 85,670 individuals yield new insights into human sleep behaviour. *bioRxiv*. Available at: https://www.biorxiv.org/content/early/2018/04/19/303925.full.pdf+html.

Jones, S.E. et al., 2016b. Genome-Wide Association Analyses in 128 , 266 Individuals Identifies New Morningness and Sleep Duration Loci. *PLoS Genetics*, 12(8), pp.1–19.

Jones, S.E. et al., 2019. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nature Communications*, 10(1), p.343. Available at: http://www.nature.com/articles/s41467-018-08259-7.

Kabir, Z.D., Martínez-Rivera, A. & Rajadhyaksha, A.M., 2017. From Gene to Behavior: L-Type Calcium Channel Mechanisms Underlying Neuropsychiatric Symptoms. *Neurotherapeutics*, 14(3), pp.588–613.

Kamali, A. et al., 2015. Mapping the trajectory of the stria terminalis of the human limbic system using high spatial resolution diffusion tensor tractography. *Neuroscience Letters*, 608, pp.45–50. Available at: http://dx.doi.org/10.1016/j.neulet.2015.09.035.

Kantojarvi, K. et al., 2017. Variants in calcium voltage-gated channel subunit Alpha1 C-gene (CACNA1C) are associated with sleep latency in infants. *PLoS ONE*, pp.1–12.

Keers, R., Farmer, A.E. & Aitchison, K.J., 2009. Extracting a needle from a haystack: Reanalysis of whole genome data reveals a readily translatable finding. *Psychological Medicine*, 39(8), pp.1231–1235.

Kerkhof, G.A. & Van Dongen, H.P.A., 1996. Morning-type and evening-type individuals differ in the phase position of their endogenous circadian oscillator. *Neuroscience Letters*, 218(3), pp.153–156. Available at: https://www.sciencedirect.com/science/article/pii/S0304394096131402?via%3Dihub.

Kloiber, S. et al., 2012. ANK3 and CACNA1C - Missing genetic link for bipolar disorder and major depressive disorder in two German case-control samples. *Journal of Psychiatric Research*, 46(8), pp.973–979.

Koike, N. et al., 2012. Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals Nobuya. *Science*, 338(6105), pp.349–354.

Kolarow, R., Brigadski, T. & Lessmann, V., 2007. Postsynaptic secretion of BDNF and NT-3 from hippocampal neurons depends on calcium calmodulin kinase II signaling and proceeds via delayed fusion pore opening. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(39), pp.10350–10364.

Krapohl, E. et al., 2018. Multi-polygenic score approach to trait prediction. *Molecular Psychiatry*, 23(5), pp.1368–1374. Available at: http://www.nature.com/doifinder/10.1038/mp.2017.163.

Kupfer, D.J., Frank, E. & Phillips, M.L., 2012. Major depressive disorder: new clinical, neurobiological, and treatment perspectives. *The Lancet*, 379(9820), pp.1045–1055. Available at: https://www.sciencedirect.com/science/article/pii/S0140673611606028.

Lahey, B.B., 2009. Public health significance of neuroticism. *The American psychologist*, 64(4), pp.241–56. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19449983.

Landgraf, D., McCarthy, M.J. & Welsh, D.K., 2014. Circadian Clock and Stress Interactions in the Molecular Biology of Psychiatric Disorders. *Current Psychiatry Reports*, 16(10).

Lane, J.M. et al., 2017. Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. *Nature genetics*, 49(2), pp.274–281. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27992416.

Lane, J.M. et al., 2016. Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the UK Biobank. *Nature Communications*, 7, pp.1–10. Available at: http://dx.doi.org/10.1038/ncomms10889.

Lépine, J.-P. & Briley, M., 2011. The increasing burden of depression. *Neuropsychiatric disease and treatment*, 7(Suppl 1), pp.3–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21750622.

Leterrier, C. et al., 2015. Nanoscale Architecture of the Axon Initial Segment Reveals an Organized and Robust Scaffold. *Cell Reports*, 13(12), pp.2781–2793. Available at: https://www.sciencedirect.com/science/article/pii/S2211124715013820.

Levandovski, R. et al., 2011. Depression Scores Associate With Chronotype and Social Jetlag in a Rural Population. *Chronobiology International*, 28(9), pp.771–778. Available at: http://www.tandfonline.com/doi/full/10.3109/07420528.2011.602445.

Li, J.Z. et al., 2013. Circadian patterns of gene expression in the human brain and disruption in major depressive disorder. *Pnas*, 110(24), pp.9950–9955. Available at: http://www.pnas.org/content/110/24/9950.full.

Liao, P. & Soong, T.W., 2010. CaV1.2 channelopathies: From arrhythmias to autism, bipolar disorder, and immunodeficiency. *Pflugers Archiv European Journal of Physiology*, 460(2), pp.353–359.

Liberman, A.R. et al., 2018. Modeling Strengthens Molecular Link between Circadian Polymorphisms and Major Mood Disorders. *Journal of Biological Rhythms*, 33(3), pp.318–336. Available at: http://journals.sagepub.com/doi/10.1177/0748730418764540.

Loh, P.-R. et al., 2015. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature genetics*, 47(3), pp.284–90. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25642633.

Lv, Y.-B. et al., 2016. Association of Renalase SNPs rs2296545 and rs2576178 with the Risk of Hypertension: A Meta-Analysis A. T. DeWan, ed. *PLOS ONE*, 11(7), p.e0158880. Available at: http://dx.plos.org/10.1371/journal.pone.0158880.

Lyall, L.M. et al., 2018. Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: a cross-sectional study of 91 105 participants from the UK Biobank. *The Lancet Psychiatry*, 5(6), pp.507–514. Available at: http://linkinghub.elsevier.com/retrieve/pii/S2215036618301391.

MacIukiewicz, M. et al., 2014. Analysis of genetic association and epistasis interactions between circadian clock genes and symptom dimensions of bipolar affective disorder. *Chronobiology International*, 31(6), pp.770–778.

Madison, J.M. et al., 2015. Characterization of bipolar disorder patient-specific induced pluripotent stem cells from a family reveals neurodevelopmental and mRNA expression abnormalities. *Molecular psychiatry*, 20(6), pp.703–17. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25733313.

Manolio, T.A. et al., 2009. Finding the missing heritability of complex diseases. *Nature*, 461(7265), pp.747–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19812666.

McCarthy, M.J., Korelova, O., et al., 2018a. A common genetic variant in CACNA1C predicts heart rate in patients with bipolar disorder. *Psychiatry Research*, pp.1–5. Available at: https://doi.org/10.1016/j.psychres.2018.01.032.

McCarthy, M.J. et al., 2016. Calcium channel genes associated with bipolar disorder modulate lithium's amplification of circadian rhythms. *Neuropharmacology*, 101, pp.439–448. Available at: http://dx.doi.org/10.1016/j.neuropharm.2015.10.017.

McCarthy, M.J., Wei, H., et al., 2018b. Chronotype and cellular circadian rhythms predict the clinical response to lithium maintenance treatment in patients with bipolar disorder. *Neuropsychopharmacology*, p.1. Available at: http://www.nature.com/articles/s41386-018-0273-8.

McCarthy, M.J. et al., 2013. Genetic and clinical factors predict lithium's effects on PER2 gene expression rhythms in cells from bipolar disorder patients. *Translational Psychiatry*, 3(10), pp.e318-8. Available at: http://dx.doi.org/10.1038/tp.2013.90.

McClung, C.A., 2007. Circadian genes, rhythms and the biology of mood disorders. *Pharmacology & therapeutics*, 114(2), pp.222–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17395264.

Merikangas, K.R. et al., 2011. Prevalence and Correlates of Bipolar Spectrum Disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry*, 68(3), p.241. Available at: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archgenpsychiatry.2011.12.

Merikangas, K.R. et al., 2019. Real-time Mobile Monitoring of the Dynamic Associations Among Motor Activity, Energy, Mood, and Sleep in Adults With Bipolar Disorder. *JAMA Psychiatry*, 76(2), p.190. Available at: http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2018.3546.

Merikanto, I. et al., 2017. Circadian preference towards morningness is associated with lower slow sleep spindle amplitude and intensity in adolescents. *Scientific Reports*, 7(1), pp.1–12.

Merikanto, I. et al., 2013. Evening types are prone to depression. *Chronobiology International*, 30(5), pp.719–725.

Merikanto, I. et al., 2016. Eveningness relates to burnout and seasonal sleep and mood problems among young adults. *Nordic Journal of Psychiatry*, 70(1), pp.72–80.

Mohawk, J.A., Green, C.B. & Takahashi, J.S., 2012. Central and peripheral circadian clocks in mammals. *Annual review of neuroscience*, 35, pp.445–462. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22483041 [Accessed October 18, 2018].

Moon, J.H. et al., 2016. Advanced Circadian Phase in Mania and Delayed Circadian Phase in Mixed Mania and Depression Returned to Normal after Treatment of Bipolar Disorder. *EBioMedicine*, 11, pp.285–295. Available at: http://dx.doi.org/10.1016/j.ebiom.2016.08.019.

Moosmang, S. et al., 2005. Mouse models to study L-type calcium channel function. *Pharmacology & therapeutics*, 106(3), pp.347–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15922017.

Moser, A.B. et al., 1999. Plasma very long chain fatty acids in 3,000 peroxisome disease patients and 29,000 controls. *Annals of Neurology*, 45(1), pp.100–110.

Nagano, M. et al., 2011. ZF21 protein, a regulator of the disassembly of focal adhesions and cancer metastasis, contains a novel noncanonical pleckstrin homology domain.

The Journal of biological chemistry, 286(36), pp.31598–609. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21768110.

NCBI, 2018. Reference SNP (refSNP) Cluster Report: rs1006737 **Clinical Channel**. *dbSNP*. Available at: https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?searchType=adhoc_search&type=rs&rs=rs1006737.

Neale, B.M. et al., 2010. Meta-analysis of genome-wide association studies of attention deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(9), p.884. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20732625.

Neves, M. de C. et al., 2016. Neural correlates of hallucinations in bipolar disorder. *Revista Brasileira de Psiquiatria*. Available at: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-44462016005016102&lng=en&nrm=iso&tlng=en.

Ng, T.H. et al., 2015. Sleep-wake disturbance in interepisode bipolar disorder and high-risk individuals: A systematic review and meta-analysis. *Sleep Medicine Reviews*, 20, pp.46–58. Available at: http://dx.doi.org/10.1016/j.smrv.2014.06.006.

O'Donovan, M.C., Craddock, N.J. & Owen, M.J., 2009. Genetics of psychosis; insights from views across the genome. *Human Genetics*, 126(1), pp.3–12.

Otowa, T. et al., 2016. Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular Psychiatry*, 21(10), pp.1391–1399. Available at: http://www.nature.com/articles/mp2015197.

Pagani, L. et al., 2016. Genetic contributions to circadian activity rhythm and sleep pattern phenotypes in pedigrees segregating for severe bipolar disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 113(6), pp.E754-61. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26712028.

Palagini, L. et al., 2018. Developmental pathways towards mood disorders in adult life: is there a role for sleep disturbances? *Journal of Affective Disorders*. Available at: https://www.sciencedirect.com/science/article/pii/S0165032718308954.

Partonen, T., 2012. Clock gene variants in mood and anxiety disorders. *Journal of Neural Transmission*, 119(10), pp.1133–1145. Available at: http://link.springer.com/10.1007/s00702-012-0810-2.

Phillips, M.L. & Kupfer, D.J., 2013. Bipolar disorder diagnosis: challenges and future directions. *Lancet*, 381(9878), pp.1663–1671. Available at: http://dx.doi.org/10.1016/S0140-6736(13)60989-7.

Pike, N., 2011. Using false discovery rates for multiple comparisons in ecology and evolution. *Methods in Ecology and Evolution*, 2, pp.278–282.

Purcell, S. et al., 2007. PLINK : A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(September), pp.559–575.

Purcell, S.M. et al., 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 10(AuGuST), pp.8192–8192.

Randler, C. & Schaal, S., 2010. Morningness–eveningness, habitual sleep-wake variables and cortisol level. *Biological Psychology*, 85(1), pp.14–18. Available at: https://www.sciencedirect.com/science/article/pii/S030105111000116X?via%3Dihub.

Reppert, S.M. & Weaver, D.R., 2001. Molecular analysis of mammalian circadian rhythms. *Annual review of physiology*, 63, pp.647–676.

Reutrakul, S. & Knutson, K.L., 2015. Consequences of circadian disruption on cardiometabolic health. *Sleep Medicine Clinics*, 10(4), pp.455–468.

Ruderfer, D.M. et al., 2014. Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Molecular psychiatry*, 19(9), pp.1017–1024. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24280982.

Sadeh, A., 2011. The role and validity of actigraphy in sleep medicine: An update. *Sleep Medicine Reviews*, 15(4), pp.259–267. Available at: http://dx.doi.org/10.1016/j.smrv.2010.10.001.

Schizophrenia Working Group of the Psychiatric Genomics Consortium, S.W.G. of the P.G. et al., 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510), pp.421–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25056061.

Schmutz, I. et al., 2015. A Specific Role for the REV-ERB α controlled L-type Voltage Gated Calcium Channel Ca V 1 . 2 in Resetting the Circadian Clock in the Late Night. *Journal of Biological Rhythms*, 29(4), pp.288–298.

Scott, J. et al., 2017. A pilot study to determine whether combinations of objectively measured activity parameters can be used to differentiate between mixed states, mania, and bipolar depression. *International journal of bipolar disorders*, 5(5). Available at: http://www.ncbi.nlm.nih.gov/pubmed/28155205.

Shi, J. et al., 2008. Clock Genes may Influence Bipolar Disorder Susceptibility and Dysfunctional Circadian Rhythm. *American Journal Of Medical Genetics Part B Neuropsychiatric Genetics*, (147B), pp.1047–1055.

Shih, R.A., Belmonte, P.L. & Zandi, P.P., 2004. A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *International Review of Psychiatry*, 16(4), pp.260–283. Available at: http://www.tandfonline.com/doi/full/10.1080/09540260400014401.

Sigurdardottir, L.E.A., 2012. Circadian Disription, Sleep Loss and Prostate Cancer risk: A Systematic Review of Epidemiological Studies. *Cancer Epidemiol Biomarkers Prev*, 21(7), pp.1002–1011.

Sklar, P. et al., 2011. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics*, 43(10), pp.977–983. Available at: http://www.nature.com/doifinder/10.1038/ng.943%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/21926972%5Cnfile:///Users/Claustrum/Dropbox/Papers2/Articles/2011/Sklar/S

klar_Nat_Genet_2011.pdf%5Cnhttp://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=219.

Smith, B.H. et al., 2013b. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology*, 42(3), pp.689–700. Available at: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dys084.

Smith, D., Ghaemi, S. & Craddock, N., 2008. The broad clinical spectrum of bipolar disorder: implications for research and practice. *Journal of Psychopharmacology*, 22(4), pp.397–400. Available at: http://journals.sagepub.com/doi/10.1177/0269881108089585.

Smith, D.J. et al., 2015. Childhood IQ and risk of bipolar disorder in adulthood: prospective birth cohort study. *BJPsych Open*, 1, pp.74–80. Available at: https://www.rcpsych.ac.uk/pdf/childhood IQ and bipolar link.pdf.

Smith, D.J. et al., 2016. Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Molecular Psychiatry*, 21(6), pp.749–757. Available at: http://www.nature.com/articles/mp201649.

Smith, D.J. et al., 2013a. Prevalence and characteristics of probable major depression and bipolar disorder within UK Biobank: Cross-sectional study of 172,751 participants. *PLoS ONE*, 8(11), pp.1–7.

Smoller, J.W., 2013. Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. *The Lancet*, 381(9875), pp.1371–1379.

Soeiro-De-Souza, M.G. et al., 2017. The CACNA1C risk allele rs1006737 is associated with age-related prefrontal cortical thinning in bipolar i disorder. *Translational Psychiatry*, 7(4), pp.2–7. Available at: http://dx.doi.org/10.1038/tp.2017.57.

Soldatov, N.M., 1994. Genomic Structure of Human L-Type Ca2+ Channel. *Genomics*, 22, pp.77–87.

Soliman, M.A. et al., 2017. Pluripotent stem cells in neuropsychiatric disorders. *Molecular psychiatry*, 22(9), pp.1241–1249. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28322279.

Spada, J. et al., 2016. Genome-wide association analysis of actigraphic sleep phenotypes in the LIFE Adult Study. *Journal of Sleep Research*, 25(6), pp.690–701. Available at: http://doi.wiley.com/10.1111/jsr.12421.

Stahl, E. et al., 2017. Genome-wide association study identifies 30 loci associated with bipolar disorder. *bioRxiv*, p.173062. Available at: https://www.biorxiv.org/content/early/2017/08/08/173062.1.full.pdf+html.

Starnawska, A. et al., 2016. CACNA1C hypermethylation is associated with bipolar disorder. *Translational Psychiatry*, 6(6), pp.1–7.

Steinan, M.K. et al., 2015. Sleep problems in bipolar disorders: More than just insomnia. *Acta Psychiatrica Scandinavica*, pp.1–10.

Stochl, J. et al., 2015. Mood, anxiety and psychotic phenomena measure a common psychopathological factor. *Psychological Medicine*, 45(7), pp.1483–1493.

Stocker, H. et al., 2018. The genetic risk of Alzheimer's disease beyond APOE ε4: systematic review of Alzheimer's genetic risk scores. *Translational Psychiatry*, 8(1), p.166. Available at: http://www.nature.com/articles/s41398-018-0221-8.

Stringaris, A., Lewis, G. & Maughan, B., 2014. Developmental pathways from childhood conduct problems to early adult depression: Findings from the ALSPAC cohort. *British Journal of Psychiatry*, 205(1), pp.17–23.

Strohmaier, J. et al., 2013. The psychiatric vulnerability gene CACNA1C and its sex-specific relationship with personality traits, resilience factors and depressive symptoms in the general population. *Molecular Psychiatry*, 18(5), pp.607–613.

Sudlow, C. et al., 2015. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3), p.e1001779. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25826379.

Sullivan, P.F. et al., 2018. Psychiatric Genomics : An Update and an Agenda. *Amercian Journal of Psychiatry*, 175(1), pp.15–27.

Sullivan, P.F., Neale, M.C. & Kendler, K.S., 2000. Genetic Epidemiology of Major Depression: Review and Meta-Analysis. *American Journal of Psychiatry*, 157(10), pp.1552–1562. Available at: http://psychiatryonline.org/doi/abs/10.1176/appi.ajp.157.10.1552.

Szigeti, K. et al., 2013. Genome-Wide Scan for Copy Number Variation Association with Age at Onset of Alzheimer's Disease. *Journal of Alzheimer's Disease*, 33(2), pp.517–523. Available at: https://content.iospress.com/articles/journal-of-alzheimers-disease/jad121285 [Accessed May 30, 2018].

Tabuchi, A. et al., 2000. Differential activation of brain-derived neurotrophic factor gene promoters I and III by Ca2+ signals evoked via L-type voltage-dependent and N-methyl-D-aspartate receptor Ca2+ channels. *Journal of Biological Chemistry*, 275(23), pp.17269–17275.

Taillard, J. et al., 2003. The circadian and homeostatic modulation of sleep pressure during wakefulness differs between morning and evening chronotypes. *Journal of Sleep Research*, 12(4), pp.275–282. Available at: http://doi.wiley.com/10.1046/j.0962-1105.2003.00369.x.

Taylor, A.M., Saifetiarova, J. & Bhat, M.A., 2017. Postnatal Loss of Neuronal and Glial Neurofascins Differentially Affects Node of Ranvier Maintenance and Myelinated Axon Function. *Frontiers in Cellular Neuroscience*, 11(February). Available at: http://journal.frontiersin.org/article/10.3389/fncel.2017.00011/full.

Tesli, M. et al., 2013. CACNA1C Risk Variant and Amygdala Activity in Bipolar Disorder, Schizophrenia and Healthy Controls. *PLoS ONE*, 8(2), pp.6–11.

Tharp, J.A. et al., 2016. Goals in bipolar I disorder: Big dreams predict more mania. *Motivation and Emotion*, 40(2), pp.290–299. Available at: http://link.springer.com/10.1007/s11031-015-9519-5.

Thaxton, C. et al., 2010. In Vivo Deletion of Immunoglobulin Domains 5 and 6 in Neurofascin (Nfasc) Reveals Domain-Specific Requirements in Myelinated Axons. *Journal of Neuroscience*, 30(14), pp.4868–4876. Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5951-09.2010.

The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, 2017. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular autism*, 8, p.21.

Tomsig, J.L. & Creutz, C.E., 2002. Copines: a ubiquitous family of Ca(2+)-dependent phospholipid-binding proteins. *Cellular and molecular life sciences : CMLS*, 59(9), pp.1467–1477. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12440769.

Townsend, P., 1987. Deprivation. *Journal of Social Policy*, 16(2), pp.125–146. Available at: http://www.stat.purdue.edu/~doerge/BIOINFORM.D/FALL06/Benjamini and Y FDR.pdf%5Cnhttp://engr.case.edu/ray_soumya/mlrg/controlling_fdr_benjamini95.pdf.

Tseng, M.-C.M. et al., 2015. *Prevalence and correlates of bipolar disorders in patients with eating disorders.*, Elsevier. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26583349.

Tung, N.M. & Garber, J.E., 2018. BRCA1/2 testing: therapeutic implications for breast cancer management. *British Journal of Cancer*, 119(2), pp.141–152. Available at: http://www.nature.com/articles/s41416-018-0127-5.

Van Someren, E.J.W. et al., 1996. Circadian rest-activity rhythm disturbances in Alzheimer's disease. *Biological Psychiatry*, 40(4), pp.259–270.

Van Veldhoven, P.P., 2010. Biochemistry and genetics of inherited disorders of peroxisomal fatty acid metabolism. *Journal of lipid research*, 51(10), pp.2863–2895. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20558530.

Visscher, P.M. et al., 2012. Five years of GWAS discovery. *American journal of human genetics*, 90(1), pp.7–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22243964.

VT, C. et al., 2015. Golli Myelin Basic Proteins Modulate Voltage-Operated Ca++ Influx and Development in Cortical and Hippocampal Neurons. *Molecular Neurobiology*.

Wang, Q. et al., 2014. Regulation of MEIS1 by distal enhancer elements in acute leukemia. *Leukemia*, 28(1), pp.138–146. Available at: http://www.nature.com/articles/leu2013260.

Ward, J. et al., 2017. Genome-wide analysis in UK Biobank identifies four loci associated with mood instability and genetic correlation with major depressive disorder, anxiety disorder and schizophrenia. *Translational Psychiatry*, 7(11), p.1264. Available at: http://www.nature.com/articles/s41398-017-0012-7.

Watanabe, K. et al., 2017. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, 8(1826). Available at: http://www.nature.com/articles/s41467-017-01261-5.

Watmuff, B. et al., 2016. Disease signatures for schizophrenia and bipolar disorder using patient-derived induced pluripotent stem cells. *Molecular and Cellular Neuroscience*. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1044743116300033.

Wiles, N.J. et al., 2012. Physical activity and depression in adolescents: Cross-sectional findings from the ALSPAC cohort. *Social Psychiatry and Psychiatric Epidemiology*, 47(7), pp.1023–1033.

Williams, S.M. & Haines, J.L., 2011. Correcting Away the Hidden Heritability. *Annals of human genetics*, 75, pp.348–350.

Witt, S.H. et al., 2014. Investigation of manic and euthymic episodes identifies state- and trait-specific gene expression and STAB1 as a new candidate gene for bipolar disorder. *Translational Psychiatry*, 4(8), p.e426. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4150244&tool=pmcentrez&rendertype=abstract.

Wood, J. et al., 2009. Replicable differences in preferred circadian phase between bipolar disorder patients and control individuals. *Psychiatry research*, 166(2–3), pp.201–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19278733.

World Health Organization, 2002. *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*, Geneva, Switzerland. Available at: https://www.who.int/whr/2002/en/whr02_en.pdf?ua=1.

Wray, N.R. et al., 2018. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), pp.668–681. Available at: http://www.nature.com/articles/s41588-018-0090-3.

Wray, N.R., Goddard, M.E. & Visscher, P.M., 2007. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome research*, 17(10), pp.1520–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17785532.

Wulff, K. et al., 2010. Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nature Reviews Neuroscience*, 11(8), pp.589–599.

Wyse, C. et al., 2018. Variation in Circadian Rhythms by Sex , Age and Chronotype in 88 , 098 UK Biobank Participants. , pp.1–24.

Xiong, L. et al., 2009. MEIS1 intronic risk haplotype associated with restless legs syndrome affects its mRNA and protein expression levels. *Human molecular genetics*, 18(6), pp.1065–1074. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19126776.

Xu, J. et al., 2005. Renalase is a novel, soluble monoamine oxidase that regulates cardiac function and blood pressure. *The Journal of clinical investigation*, 115(5), pp.1275–1280. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15841207.

Yatham, L.N. et al., 2018. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar disorders*, 20(2), pp.97–170. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29536616.

Yoshimizu, T. et al., 2015. Functional implications of a psychiatric risk variant within CACNA1C in induced human neurons. *Molecular psychiatry*, 20(2), pp.162–9. Available at: http://dx.doi.org/10.1038/mp.2014.143.

Zhang, E.E. et al., 2009. A Genome-Wide RNAi Screen for Modifiers of the Circadian Clock in Human Cells. *Cell*, 139(1), pp.199–210.

Zheng, J. et al., 2017. LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*, 33(2), pp.272–279.