Expression Quantitative Trait Loci as possible Biomarkers on Depression: Candidate Gene and Genome-wide Approaches

Janine Arloth

Dissertation
an der Fakultät für Biologie
der Ludwig–Maximilians–Universität
München

vorgelegt von
Janine Arloth
aus Lutherstadt Wittenberg

München, den 05. August 2014

Diese Dissertation wurde angefertigt
unter der Leitung von Dr. Elisabeth B. Binder am

Erstgutachter: Prof. Dr. John Parsch
Zweitgutachter: PD Dr. Mathias V. Schmidt
Tag der mündlichen Prüfung: 26. November 2014

"We are all feeling our way in the face of the extreme complexity of nature and the daunting task of unraveling her secrets."

-Denis Nobel-

Contents

List of Figures

List of Tables

Zusammenfassung

Depression (Major depressive disorder; MDD) ist eine der häufigsten psychiatrischen Erkrankungen, deren Entstehung durch ein komplexes Zusammenspiel verschiedener Faktoren zustande kommt. So ist das Risiko an einer Depression zu erkranken, durch die Kombination aus genetischer Prädisposition und individuell erlebter Stressbelastung erhöht. Die biologischen Mechanismen wie aus diesen Faktoren eine Depression entsteht, sind bisher weitgehend unbekannt. Eine mögliche molekulare Ursache ist die gestörte Regulierung der Kortisolausschüttung und der Effekt dieses Hormons. Das Kortisol wird als Reaktion des Körpers auf Stress freigesetzt. Durch Bindung des Kortisols an den Glukokortikoid-Rezeptor (GR), steuert es die Aktivität verschiedener Gene und bewirkt dadurch eine stressbedingte Anpassung der Zelle. Der GR wandert, nachdem er das Kortisol in der Zelle gebunden hat, in den Zellkern und bindet dort direkt an die DNS. Eine Störung der Regulation der GR vermittelten Stressantwort wurde bei Depressionserkrankungen nachgewiesen.

In dieser Doktorarbeit wurde zuerst der GR medikamentös aktiviert und die daraus resultierenden Veränderung in der Genaktivität in Blutzellen von depressiven Patienten mit denen von gesunden Kontrollpersonen verglichen. Dazu wurden genomweite Genexpressionsdaten, vor und nach GR-Stimulierung mit 1.5 mg Dexamethason *p.o.*, von 29 männlichen Depressionspatienten und 31 gesunden Kontrollen analysiert.

Durch Dexamethason wurden 2670 Gene von Kontrollpersonen und 1151 Gene von depressiven Patienten reguliert (aktiviert oder unterdrückt). Darunter waren mehrere Gene, die zuvor mit Depression in Verbindung gebracht wurden, wie beispielsweise *FKBP5* und *DUSP1*. Die Genexpressionsprofile dieser GR-stimulierten Gene konnten Patienten und Kontrollpersonen besser unterscheiden (79,2% vs. 41,7% Sensitivität der Klassifikation) als herkömmliche endokrine Tests.

Der zweite Teil dieser Doktorarbeit beschäftigte sich mit der Identifikation von Genvarianten, welche GR-stimulierte Genexpressionsänderungen beeinflussen. Solche Analysen kombinieren Daten aus Genexpressions- mit Genotypisierungs-Microarrays und werden expression quantitative trait locus (eQTL) Analysen genannt. Das Ziel dieser Arbeit war es Einzelnukleotid-Polymorphismen (SNPs) zu identifizieren, die mit Glukokortikoid-vermittelten Genexpressionsänderungen einhergehen (GR-response eQTLs). Dabei konzentrierten wir uns auf Assoziationen mit SNPs innerhalb einer 1-Mb-Region vom 5'- oder 3'-Ende des Transkriptes, den sogenannten *cis*-eQTLs.

Es wurden 3820 GR-response *cis*-eQTLs identifiziert, bei denen SNPs die GR vermittelte Veränderung der Gentranskription beeinflussen. Bei diesen SNPs handelt es sich signifikant häufiger um SNPs, die mit Depression in einer genomweiten Meta-Analyse assoziiert wurden (Psychiatric Genomics Consortium (PGC) Daten mit n >9000 depressiven Patienten

und Kontrollpersonen).

Ähnliche Beobachtungen konnten im Fall von Schizophrenie (SCZ), bipolarer Störung (BPD) und SNPs, die mit Risiko für mehrere psychiatrische Erkrankungen (Crossdisorder Analyse) assoziiert wurden, erzielt werden. Die PGC Crossdisorder Analyse untersuchte das gemeinsame Risiko für fünf psychiatrische Erkrankungen (SZC, MDD, BPD, Autismus und Aufmerksamkeitsdefizit-/ Hyperaktivitätsstörung; n =33000 Patienten und 29000 Kontrollpersonen).

Die 282 gefundenen SNPs, die sowohl eine Assoziation mit GR-regulierter Transkription als auch mit Depression zeigen (GR/MDD SNPs), regulieren die Aktivität von 25 verschiedenen Transkripten. Mithilfe einer Pathway Analyse wurde nachgewiesen, dass diese 25 Transkripte besonders mit molekularen Prozessen, die mit Veränderung synaptischer Plastizität, Immunaktivität oder mit der Pathophysiologie von Depression, zusammenhängen. In relevanten Mausmodellen konnte gezeigt werden, dass über 66% dieser Transkripte eine GR-Regulierung in verschiedenen Gehirnregionen aufweisen. Zusätzlich wurde mit Hilfe des genetischen Risikoprofiles aus GR/MDD SNPs eine veränderte Amygdala Reaktivität in einer unabhängigen Kohorte nachgewiesen.

Zusammenfassend wurde erstens gezeigt, dass ein Genexpressionsprofil aus GR-stimulierten Genen in Blutzellen ein vielversprechender molekularer Biomarker für Depression sein könnte, welcher die Veränderungen in der GR-Funktionalität abbildet, die wiederum ein wichtiger Bestandteil in der zugrundeliegenden molekularen Pathologie bei depressiven Patienten ist. Weiterhin konnte gezeigt werden, dass genetische Varianten, die mit der ersten transkriptionellen Reaktion auf Stress korreliert sind, häufiger mit stress-assoziierten Krankheiten wie Depression in Verbindung gebracht werden. Diese Erkenntnisse verbessern unser Verständnis von psychiatrischen Erkrankungen als Folge der Interaktion von Umwelt- und genetischen Faktoren.

Abstract

The risk for major depressive disorder (MDD) is exacerbated by various genetic factors and stress exposure; however, the underlying biological mechanisms leading to an increase in risk are poorly understood. One putative mechanism implicates the variability in the ability of cortisol, released in response to stress, to trigger a cascade of adaptive genomic and non- genomic processes through glucocorticoid receptor (GR) activation. The GR exerts its main downstream effects via its function as transcription factor.

In the first part of my doctoral thesis a differential gene expression analysis utilizing the dexamethasone challenge test to compare GR-mediated changes in gene expression between depressed patients and healthy controls was conducted. A genome-wide gene expression data set with RNA samples at baseline as well as following GR stimulation with 1.5 mg dexamethasone *p.o.* in peripheral blood cells from 31 male depressed patients and 29 controls was analyzed. I aimed to identify gene expression patterns that would predict MDD disease status from this sample.

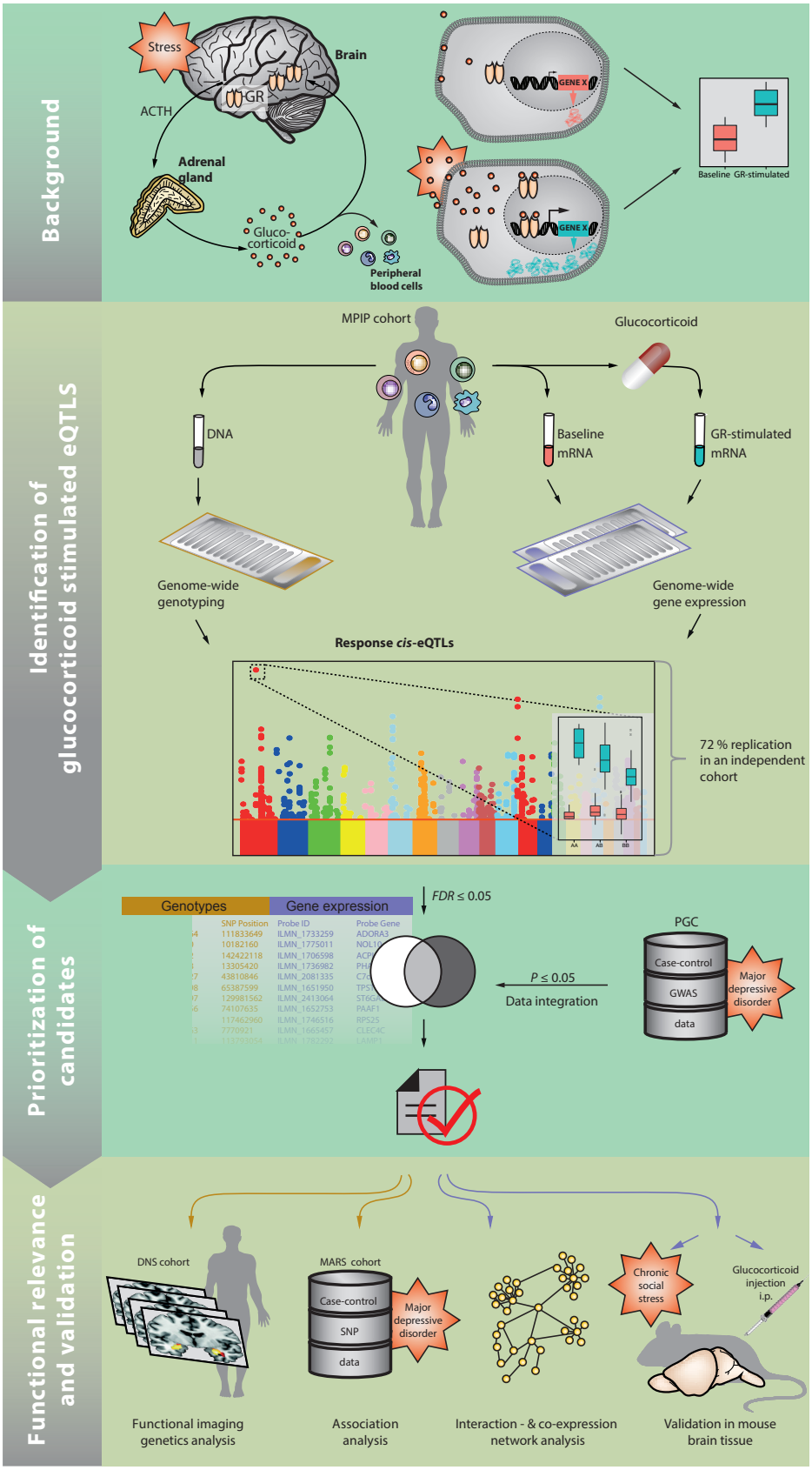
The dexamethasone intake led to a reproducible regulation of 2,670 transcripts in controls and 1,151 regulated transcripts in depressed patients, including several genes previously associated with the pathophysiology of MDD, e.g. *FKBP5* and *DUSP1*. Furthermore, using a machine learning algorithm I showed that a gene expression profile of GR-stimulated transcripts outperforms baseline gene expression as a classifier for MDD disease status with sensitivity of 79.2% vs. 41.7%, respectively.

The second part of my doctoral thesis presents a novel approach based on the analysis of GR-response expression quantitative trait loci (eQTLs). I investigated on a genome-wide level, whether variants that alter the immediate transcriptional response to GR activation may alter the risk to suffer from stress-related disorders, like MDD. The eQTL analysis was performed on imputed single nucleotide polymorphism (SNP) data in a *cis*-window of ±1 Mb and the differences in gene expression between GR-stimulated and baseline samples from peripheral blood cells of 160 male individuals (see summary figure below for illustration of the sequence of experiments and analyses investigated in this study).

We identified 3,820 GR-response *cis*-eQTLs with SNPs modulating the GR induction of gene transcription. These SNPs were highly enriched among variants associated with MDD, as identified in a meta-analysis for MDD using the PGC data with an n of over 9,000 MDD cases and controls. Furthermore, there was also evidence for significant enrichment of these GR-response eSNPs with schizophrenia (SCZ), bipolar disorder (BPD) and variants conferring psychiatric risk for cross disorders. The PGC cross disorder analysis measures the shared risk on five major psychiatric disorders (SCZ, BPD, MDD, attention deficit hyperactivity disorder and autism spectrum disorder; n =33,000 cases and 29,000 controls).

The 282 SNPs showing both an association with GR-mediated transcription and MDD (GR/MDD SNPs) regulate 25 distinct transcripts. Pathway analysis suggests an involvement of these 25 transcripts in pathways associated with ubiquitination and proteasome degradation and the inflammatory response- systems that have been implicated in the pathophysiology of MDD and in stress-related changes in synaptic plasticity. Additionally, in corresponding mouse models, we found over 66% of these 25 transcripts to be regulated following GR agonist stimulation in hippocampus, prefrontal cortex or amygdala. In addition, the genetic risk profile of the GR/MDD SNPs was associated with altered centromedial amygdala reactivity to threat-related cues.

In summary, it was first shown that in vivo stimulated gene expression in peripheral blood cells could be a promising molecular marker of altered GR functioning, an important component of the underlying pathology, in patients suffering from depressive episodes. Secondly, our data suggests that genetic variants that modulate the first transcriptional response to stress are more likely to be associated with stress-related disorders. This strongly supports the importance of molecular gene by environment interactions for the understanding of the pathophysiology of MDD and related disorders.


Background
Stress
Brain
GR
ACTH
Adrenal gland
Gluco-corticoid
Peripheral blood cells
GENE X
Baseline
GR-stimulated
Identification of glucocorticoid stimulated eQTLS
MPIP cohort
Glucocorticoid
DNA
Baseline mRNA
GR-stimulated mRNA
Genome-wide genotyping
Genome-wide gene expression
Response cis-eQTLs
72 % replication in an independent cohort
Prioritization of candidates
Genotypes
Gene expression
FDR ≤ 0.05
P ≤ 0.05
Data integration
PGC
Case-control
GWAS
data
Major depressive disorder
Functional relevance and validation
DNS cohort
MARS cohort
Case-control
SNP
data
Major depressive disorder
Chronic social stress
Glucocorticoid injection i.p.
Functional imaging genetics analysis
Association analysis
Interaction - & co-expression network analysis
Validation in mouse brain tissue

1. Introduction

1.1. Background of major depressive disorder

1.1.1. Epidemiology and clinical features

Major depressive disorder (MDD) is the most common psychiatric disorder with a reported lifetime prevalence up to 17% [108]. In Europe, MDD is the third leading cause of disability [62] and rated as one of the disorders with the highest global burden of disease according to the world health organisation (WHO) [218]. From 1990 to 2010, the global burden of MDD increased by 37% [156]. Although several studies reported MDD to be equally heritable in man and women [105], there is some evidence suggesting that women are more likely to experience an episode of depression during their lifetime [106, 3]. Probably there are genes acting differently on the risk for MDD in both sexes [105]. The average age-of-onset is estimated to be in the early to mid twenties [5]. Young adults (18-29 years) are more likely to suffer from depression compared to older adults (>60 years) [107]. Epidemiological studies have shown the high comorbidity of MDD with other psychiatric disorders, especially with anxiety disorders [107]. Around 90% of the patients with an anxiety disorder experience MDD at some point in their life [182].

Despite its high prevalence and impact, the pathophysiological mechanisms underlying MDD are not sufficiently understood, resulting in non-optimal treatments with high rates of recurrence and treatment resistance [232]. Treatment options include pharmacologic therapy, electroconvulsive therapy, psychotherapy or a combination of some or all of these therapies.

The most common symptoms of MDD include affective abnormalities, like depressed mood and loss of interest or pleasure. Other symptoms are sleep distruptions, poor concentration, recurring thoughts of death and significant weight loss without dieting or weight gain [34]. These symptoms are often recurrent and can become chronic.

Currently, depression diagnosis is based on a set of such signs or symptoms, defined by the diagnostic and statistical manual of mental disorders forth edition, text revision (DSM-IV-TR [6]; see table 1.1) or international classification of diseases (ICD-10). Various other measurements have been used to screen for depressive symptoms including the self-report rating scales, e.g. Hamilton rating scale for depression (HAM-D) and center for epidemiological studies depression scale (CES-D) as well as the interview-based rating scales, e.g. Beck depression invertury (BDI) and patient health questionnaire (PHQ-9). The BDI is the most widely used assessment scale for depression. It is a 21-item questionnaire (<9 no depression, 10-15 mild depression, 16-23 moderate depression and >24 severe depression)

with questions rated on a four-point scale. The HAM-D is a 17-item questionnaire (<7 normal depression, 8-13 mild depression, 14-18 moderate depression, 19-22 severe depression, >22 very severe depression) and questions are rated on a five-point scale [35]. Both measurements were used in this thesis (see subsection 2.1).

At present, the optimal diagnostic method or screening tool is unknown. No biological criteria or measures are established for routine diagnosis, thus, finding better biomarkers is one challenge we aim to answer in this thesis.

A. At least five of the following symptoms have been present during the same 2-week period and represent a change from previous functioning: at least one of the symptoms is either 1) depressed mood or 2) loss of interest or pleasure.
(1) Depressed mood most of the day, nearly every day, as indicated either by subjective report or observation made by others. (2) Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day. (3) Significant weight loss when not dieting or weight gain, or decrease or increase in appetite nearly every day (4) Insomnia or hypersomnia nearly every day. (5) Psychomotor agitation or retardation nearly every day. (6) Fatigue or loss of energy nearly every day. (7) Feelings of worthlessness or excessive or inappropriate guilt nearly every day. (8) Diminished ability to think or concentrate, or indecisiveness, nearly every day. (9) Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or specific plan for committing suicide.
B. The symptoms do not meet criteria for a mixed episode.
C. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
D. The symptoms are not due to the direct physiological effects of a substance or a general medical condition.
F. The symptoms are not better accounted for by bereavement, i.e., after the loss of a loved one, the symptoms persist for longer than 2 months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation.

Table 1.1.: DSM-IV-TR diagnosis criteria for MDD. The table is based on [6].

1.1.2. Etiology and candidate systems

The etiology of MDD includes psychological, physiological as well as environmental factors, resulting in a very heterogeneous disorder [82].

1.1.2.1. Stress hormone system

The stress hormone system, or hypothalamic-pituitary-adrenal (HPA) axis, is an important mediator in depression [71]. A stress response, e.g. after exposure to adverse life events, elicits the activation of this system by releasing corticotrophin-releasing factor/hormone (CRF/CRH) from the hypothalamus. This hormone is mediated by the corticotropin releasing hormone receptor 1 (*CRHR1*), leading to secretion of adrenocorticotrophic hormone (ACTH) from the anterior pituitary. ACTH then triggers the release of glucocorticoids (cortisol, a stress hormone, in humans and corticosterone in rodents) from the cortex of the adrenal glands (see figure 1.1a). Glucocorticoids interact with their receptors to inhibit further secretion of ACTH and corticotrophin releasing hormone (CRH) in a negative-feedback loop [43].

Hyperactivity of the HPA axis in depressed patients is the most common finding in psychiatry. Central mechanisms for this hyperactivity are an increased neurotransmission of CRH and an impaired negative feedback of the HPA axis [11, 158].

Two different types of nuclear hormone receptors mediate the action of glucocorticoids: the mineralocorticoid receptor (MR) and the GR. The MR is selectively expressed in the limbic system and shows a high affinity for glucocorticoids, i.e. a ten-fold higher affinity for cortisol than the GR [43]. It is already activated at basal glucocorticoid levels (low stress levels), such as the cortisol concentration in the afternoon and night [205]. The second receptor is the GR, present in the pituitary, hypothalamus area and prefrontal cortex (PFC). The GR will only be occupied during stress response (high stress levels), but always after the complete saturation of the MRs [205, 42].

1.1.2.2. Glucocorticoid receptor

The GR is a prime candidate for associations with susceptibility for MDD and a target of psychiatric therapy. For example, Modell et al. [154] indicated that people with MDD have a disturbed function or decreased expression of the GR [42].

Glucocorticoid binding allows the GR to translocate from cytoplasm to nucleus, where it binds to specific sequences of the deoxyribonucleic acid (DNA) known as glucocorticoid response elements (GREs) and regulates the expression of target genes (see figure 1.1b). The GR is able to stimulate or repress transcription, and interact with other transcription factors (TF)s such as the activating protein-1 (AP1) and the nuclear factor kappa B (NFκB) [145]. The main function of GR activation is to promote proper negative feedback of the HPA axis to terminate the stress response.

The GR is encoded by the nuclear receptor subfamily 3, group C, member 1 (*NR3C1*) gene, located on chromosome 5.

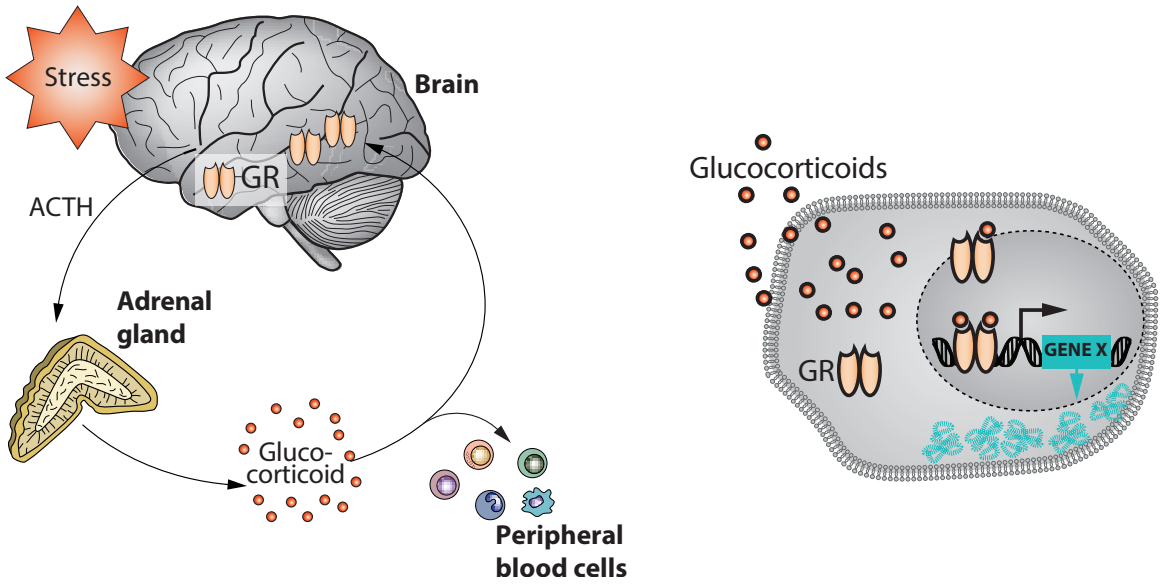
1.1.2.3. Dexamethasone

A way to evaluate the reactivity on the HPA axis is provided by the dexamethasone suppression test (DST) [127]. Since endogenous glucocorticoids could serve as prime candidates for stress-related disorders, synthetic analogs of glucocorticoids, i.e. dexamethasone, serve

as potent GR agonists [191]. By stimulating the GR, dexamethasone activates the negative feedback loop reducing the activity of the HPA axis and leading to a decrease in the production and release of ACTH and cortisol, which in turn is measured by the DST [127]. Thus, the DST is can be interpreted as a measurement of GR sensitivity. To potentially characterize MDD patients and uncover alteration in the stress hormone system, the DST received considerable attention in psychiatric research [27].

Besides, by altering cortisol secretion, dexamethasone leads to extensive and reproducible gene expression changes (for example see section 3.1) that can be used as molecular markers for GR sensitivity. Genetic variants (see subsection 1.2.1) that influence the transcriptional effects of the GR activation are interesting candidate polymorphisms (see subsection 1.2.1) for MDD. The genetics of variation in gene expression (see subsection 1.2.2 and section 3.2) has gained much attention in the last decade, resulting in meaningful studies to characterize the genetic architecture of transcriptional regulation [150, 224, 155].

Genes which are componets of the HPA axis and known to be associated with MDD are listed in table 1.2.



(a) The stress hormone system.

(b) Circulation of the GR between cytoplasm and nucleus.

Figure 1.1.: Schematic representation of the stress hormone system and its main effector, the GR, which is know to have a disrupted function in MDD. (a) Stressful events activate the HPA axis by including the release of CRH from the hypothalamus, which promotes the secretion of ACTH from the pituitary. ACTH in turn stimulates the adrenal gland to release glucocorticoids into blood stream. Normalization of CRH after stress exposure is archived via negative feedback mechanism, whereby glucocorticoids activate the GR, which terminates the stress reaction [42]. (b) The GR is a nuclear hormone receptor and upon activation, it translocates from the cytoplasm to the nucleus, where it binds to GREs and regulates gene expression.

Endophenotype	Gene	Function	Chr.	Reference
HPA axis	*NR3C1*	GR	5	[221]
	CRHR1	CRH/CRF receptor	17	[19, 126]
Monoamines	*TPH2*	involved in the biosynthesis of serotonin	12	[246]
	MAOA	degrades amine neurotransmitters	X	[50]
	SLC6A4	serotonin transporter	17	[195]
	COMT	breakdown of dopamine in brain	22	[141]
Glutamate	*GAD1*	production of GABA	2	[91]
	GRIA3	glutamate receptor	X	[114]
	P2RX7	ATP binding, ion channel activity	12	[128]
Neuroplasticity	*DISC-1*	neurodevelopment and neural signaling	1	[81]
	BDNF	plasticity	11	[192]
	NTRK2	receptor for brain-derived neurotrophic factor (*BDNF*)	9	[49]
	CREB1	transcription factor	2	[49]

Table 1.2.: A list of candidate MDD genes. The table is based on [219].

1.1.2.4. Neurotransmitters

Neurotransmitters are chemical messengers that transmit signals across nerve cells. Alterations in their function in the brain have been implicated in the pathophysiology of MDD before [219].

Major types of neurotransmitters are amino acids (glutamate), neuropeptides (CRH) and monoamines. After introducing the monoamine hypothesis in 1965 [188], greatest attention in MDD research was given by the monoamine system, including serotonin or catecholamines (adrenaline, noradrenaline and dopamine). It proposes that "the underlying biological or neuroanatomical basis for depression is a deficiency of central noradrenergic and/or serotonergic systems and that targeting this neuronal lesion with an antidepressant would tend to restore normal function in depressed patients" [92]. As decreased serotonin levels and noradrenergic and dopaminergic dysfunction has been associated with depression and popular antidepressants prevent serotonin reuptake (selective serotonin re-uptake inhibitor (SSRI)s) candidate studies investigated genes regulating the monoaminergic pathway. The main candidates are summarized in table 1.2 [82].

Glutamate is the most abundant free amino acid in the brain and has excitatory effects on nerve cells [245]. Dysregulation of proteins involved in glutamatergic signaling are implicated in alterations in animal models for depression. Drugs influencing the glutamate receptor tend to have an antidepressant like effect in these models. Additionally, significantly higher levels of glutamate are present in patients with MDD [153]. Studies, which focused on genes involved in the regulation of the glutamatergic neurotransmission, are listed in table 1.2.

1.1.2.5. Neuroplasticity

Plasticity refers to the capacity of cells or organs to change their phenotype in response to alterations in their environment [198]. Especially the term neuroplasticity denotes environmental adaptability through modification of the connectivity between neurons and neuronal circuits, i.e. adding new nerve cells, strengthening or weakening nerve connections (synapses) [140]. In the last few years, the view has gained ground that impairment of neuroplasticity may play an important role in psychiatric diseases, like MDD, and the major goal is to identify the specific transmitter systems involved in those diseases and design appropriate interventions [25].

Patients with MDD display structural brain changes, such as a reduction in hippocampal volume, which may be due to glutamate neurotoxicity-induced reduction in neurogenesis [197]. Moreover, an increased density of hippocampal neurons and glia cells (synaptic plasticity) was observed in MDD patients [209]. In contrast the PFC is associated with decreasing density of neurons and glia cells [180]. Decreased glia density has also been found in amygdala (AM) and cingulate cortex (Cg25) [191, 18]. Furthermore, hippocampal strength is changed by long-term potentiation and depression.

Exposure to chronic stress disrupting hippocampus dependent memory functions and deficits in memory formation are observed in MDD [174]. Synaptic function that underlie memory and learning were recently associated with an orchestrated function of protein synthesis and degradation. Genes involved in the etiology of depression and also in the cellular mechanisms of neuroplasticity are listed in table 1.2.

1.1.2.6. Neuroimaging

Neuroimaging techniques can be used to further study brain function and structure, and help to better understand the relationship between certain brain regions and specific mental function, especially in the light of stress-related disorders. Common techniques include hemodynamic (blood flow or circulation) techniques- such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), as well as electro-magnetic techniques-like electroencephalography (EEG) and magnetoencephalography (MEG) [39]. In this thesis we used blood oxygenation level-dependent (BOLD) fMRI to analyze the importance of functional expression SNP (eSNP)s for MDD (see section 3.2).

An MRI scanner works with a magnetic field inside the scanner and magnetic susceptibility effects of deoxygenated hemoglobin (deoxygenated hemoglobin is much more magnetic than oxygen) were utilized for the fMRI technique. Changes in blood oxygenation and flow, which occur in response to neural activity, can be detected by this technique. When the brain is activated by a specific task, the magnetic resonance signal intensity is increased in the thereby activated regions. This is due to greater uptake of oxygen, resulting in an increased blood flow in this area [78]. An fMRI scan produces activation maps visualizing which parts of the brain are involved in a particular mental process [65].

Human neuroimaging studies have examined alterations in the activation of specific brain regions in MDD patients relative to controls. Brain regions involved in emotion experience,

like PFC and Cg25, as well as hippocampus (HC) and AM, which are part of the emotional memory formation and retrieval process, are characterized by dysregulated neural activity in common psychiatric disorders [182].

Structural neuroimaging studies provide strong evidence of decreased volume of brain regions- such as AM, PFC and the HC [65, 131, 220]- that control emotion, mood and cognition, in MDD patients compared to healthy controls.

Furthermore, genetic variants in the serotonergic system (*5-HTTLPR*, serotonin-transporter-linked polymorphic region in solute carrier family 6 (neurotransmitter transporter), member 4 (*SLC6A4*)) are associated with increased amygdala activation in patients with MDD [65]. Gene by environment interactions indicating risk-allele carriers to be associated with low grey matter volume and processing of negative affect have been found in this genetic variants [182, 171].

Other studies report a decrease of the connectivity between HC and PFC as well as other brain regions, although there are contradicting reports of increased connectivity of some regions, indicating a more complex disruption of reciprocal connections [53].

1.1.3. Genetic and environmental factors

MDD is characterized by both genetic as well as environmental influences. The primary known environmental risk factors for MDD are life stress events including sexual, physical or emotional abuse, childhood neglect, loss of a parent or living with mentally ill parent [85].

Case-control family studies showing that MDD aggregates within families date back to the first few decades of the 20^{th} century [62]. A meta-analysis of high-quality family studies found that the prevalence of MDD in the first-degree relatives was 2.84-fold higher in relatives of affected subjects compared to relatives of unaffected subjects [213, 62]. This meta-analysis estimated heritability for MDD to be 37% (95% confidence interval (CI) of 31-42%) with minimal shared environmental effects between siblings (0% with 95% CI of 0-5%), but large individual specific environmental effects (63% with 95% CI of 58-67%). The absence of shared family environmental influences points to aggregation within MDD families due to genetics effects. Consistent heritability estimated are found in the Swedish national twin study, comprising more than 15,000 twin-pairs, which estimated a 38% heritability for MDD [104].

Numerous genes have been associated with risk of depression in various studies, the most common candidate genes are listed before in table 1.2 [17]. One of the candidate genes is the most widely studied serotonin transporter gene (*SLC6A4*) on chromosome 17, a therapeutic target for the SSRI class of drugs [72]. A common polymorphism located in the promoter of this gene is 5-hydroxytriptamine-transporter-linked polymorphic region (*5-HTTLPR*), a 44base pairs (bp) insertion/deletion short/long polymorphism. The short allele (deletion) correlates with reduced serotonin transporter messenger RNA (mRNA) transcription [84, 86, 72]. Mixed results of the relationship of this polymorphism to depression have been reported [90], including positive and negative associations. More consistent results have been obtained by studying gene by environment interactions between *5-HTTLPR* and

stressful life events. Caspi et al. [30] reported that the *5-HTTLPR* short allele carriers were more susceptible to depression if they experience early life stress.

Abnormal HPA axis regulation is a key neurobiological characteristic of MDD and GR function has been shown to be disturbed in MDD patient, as described above. Thus polymorphisms altering the transcriptional effects of GR activation might be interesting candidates for this disorder. One important regulator of GR sensitivity is the FK506 binding protein 5 (*FKBP5*) gene on chromosome six. Binder et al. [11, 12] showed that depression is associated with GR supersensitivity in *FKBP5*. Variants in *FKBP5* are associated with antidepressant response. These variants were also correlated with increased intracellular *FKBP5* protein expression, which activates alterations in GR and thereby HPA axis regulation. A less HPA axis hyperactivity during depressive episodes was observed in individuals carrying the correlated genotypes of these variants.

Despite this estimated substantial genetic contribution, candidate genes and genome-wide association studies (GWASs) for MDD have failed to identify robust genetic associations. Flint and Kendler [62] recently found over 1,500 articles reporting genetic associations for MDD. Only few groups agree with each other, which is reflected by the conflicting resolution of meta-analysis results for MDD. Table 1.3 summarizes data of 26 candidate genes analyzed by a meta-analysis of which only seven yield a significant nominal p-value ($P < 0.05$) [62]. Furthermore, the Psychiatric Genomics Consortium (PGC) adopted a large mega-analysis within over 9,000 cases and the same number of controls and was not able to identify a maker reaching the genome-wide significance threshold ($P < 5 \times 10^{-8}$) [133].

Thus, one aim of this thesis is to identify SNPs associated with glucocorticoid-induced gene expression changes and to functionally characterize their relevance for psychiatric disorders, especially MDD.

1.1.4. Animal models of stress-related disorders

The underlying disease process of depression is not fully understood and recreating the disease in animal models is not possible. At present, the models exhibit a depression-like behavior in simulating parts of the human symptoms. But, not all symptoms can be reproduced in animals, for example subjective feeling and appetite change cannot be used for modeling. Examples for symptoms assessable in animals include physiological, endocrinological and neuroanatomical alterations as well as behavioral traits. Such models can be used to predict variability to detect accurate treatments that are useful for the clinic. Depending on the stressor (physical/systemic vs. cognitive/psychological), different neurological circuits are activated [52].

To assess depression and antidepressant-like behavior the forced swim test (FST) [129] and tail suspension test (TST) [208] were used. An animal model of MDD incorporating disease etiology and predisposition is the learned helplessness (LH) paradigm [194]. All three models rely on relatively short-term aversive stress exposure [45].

Stress exposure is the main environmental risk factor for MDD; therefore, the majority of animal models of MDD are based on the exposure to various types of acute or chronic

Relevance	# of studies	# of cases	# of controls	P value
5-HTR2A	7 - 11	768 - 1,491	959 - 2,937	≥ 0.12
5-HT-6R	4	701	2,422	0.41
5HTTLPR/SLC6A4	4 - 39	275 - 6,836	739 - 14,903	$\geq$ **0.007**
ACE	4 - 15	586 - 2,479	5,169 - 7,744	> 0.1
BDNF	2 - 23	285 - 4,173	688 - 12,747	> 0.1
CLOCK	6	930	2,305	0.47
COMT	6	NA	NA	NS
DRD3	4	541	606	NS
DRD4	5	318	814	**0.003**
GABRA3	NA	NA	NA	NS
GNB3	3	375	492	$<$ **0.05**
HTR1A	4 - 13	1,658 - 3,199	2,046 - 4,380	$\geq$ **0.006**
HTR1B	NA	NA	NA	NS
HTR2A	4 - 11	768 - 1,491	959 - 2,937	> 0.1
HTR2C	2	NA	NA	NS
HTR6	4	701	2,422	0.406
MAOA	4	NA	NA	NS
MTHFR	4 - 17	291 - 3,341	835 - 13,840	$\geq$**0.003**
NET/SLC6A2	3 - 6	1,673 - 1,681	1,410 - 2,938	0.78
DAT/SLC6A3	3	151	272	$<$ **0.05**
TPH1	10	1,812	2,223	> 0.1

Table 1.3.: Candidate gene polymorphisms implicated in meta-analyses of genetic association studies of MDD. NA refers to not available, NS nonsignificant and # to number of. The number of samples and studies included in the meta-analysis is given as range of size (samllest to largest). Significant meta-analysis p-values are highlighted in bold. The table is based on [62].

stressors. Example for models of MDD that include a stress component in adulthood are: the chronic mild stress (CMS) paradigm, aiming to model a chronic depressive-like state that develops gradually over time in response to a mild stressor (isolation or crowed housing, food or water deprivation, disruption in dark-light cycle, etc.) [52]. The CMS has been linked to result in long lasting behavioral and neuroendocrinological changes, which resemble dysfunctions in MDD patients [45]. Other models of adult stress are social conflict animal models. Stress can be a chronic and a recurring factor occurring in lives of all higher animal species. Humans experiencing social defeat show increased MDD symptoms [13]. Different paradigm for the social stress model have been established, like dyadic, i.e. animal is exposed to a dominant and aggressive other animal, and group social stress [45]. In this thesis we used a paradigm developed by Schmidt et al. [189], utilizing chronic social stress as a key pathogenic factor during adolescence. An unstable social environment for a prolonged period of time is created. Animals are exposed to a continuous, stressful situation with no possibility to escape and adapt to. Briefly, animals

were exposed to seven weeks of chronic social stress, e.g. the group composition in each cage is changed twice per week, so that each time four different mice are put together in a new clean cage. After the seven weeks stress procedure, all animals were single housed for five weeks. After this recovery phase, susceptible animals exhibit depression-like behavioral and endocrine phenotypes, while this is not the case for resilient animals, which show strong stress-related disturbance immediately after the seven weeks of stress but recovery to the level of unstressed control animals [190, 189, 110].

Contrary to the adult stress models, the early life stress models, like the maternal separation paradigm, have been investigated.

Further evidence is provided by genetic animal models in which components of the HPA axis were modified by mutagenesis (GR or *CRHR1* knockout mice) or models which assess the functionality of the HPA axis by challenge tests, such as the DST or the combined dexamethasone-CRH test [42]. The dexamethasone treated mice experiment was part of this thesis (see section 3.2 and subsection 2.1).

1.2. Technological background to identify complex traits

1.2.1. Genome-wide association studies

The genome-wide association study (GWAS) methods screen the whole genome for associations between common genetic variants (single nucleotide polymorphism (SNP)s) and a phenotype without any prior selection for specific regions, genes or variants of interest. Therefore, the GWAS approach is also called hypothesis-free approach. This method utilizes high-throughput genotyping arrays (see section 1.2.1), which capture a remarkable proportion of common variation in the genome [109]. GWAS can be used to detect case-control associations. Thereby, a set of cases and matched controls is used to assess the difference in SNP frequencies between both sets [22]. The basic design of a GWAS includes two steps: the first one is statistical testing of the correlation between SNPs and phenotype and the second one follows up the best hits in an independent sample for statistical validation [226].

GWASs are an important advance for the identification of genetic variants influencing common human diseases, but there are several limitations. Firstly, there is a huge gap between statistical association and identifying the underlying functional basis between a genomic interval and a given complex trait. Secondly, associations identified in one population are often not reproducible and cannot be replicated in other populations and furthermore a large number of identified loci are cell and tissue specific. Thirdly, the enormous number of loci identified by GWASs have only been able to account for a very small proportion of the heritability of the complex traits [64]. This phenomenon has become known as the "case of missing heritability" [143, 136].

Regardless of the drawbacks, these studies mainly represent a valuable discovery tool for examining genomic function and elucidate pathophysiologic mechanisms [168].

1.2.1.1. Analysis of GWASs

In the following section, several methodologies, which are crucial for GWASs, are described. The outcome highly relies on the correct genotype data, which otherwise reduces the power to fine map the trait of interest.

SNP genotyping

The increased interest in SNPs has been reflected by the rapid development of diverse and dense SNP genotyping methods. This constitutes the process of determining the genotype of an individual by examining its DNA sequence with the use of biological assays. A higher throughput was provided through the implementation of oligonucleotide microarrays. Hundreds of thousands of fixed sequences can be arranged in a small area, enabling very high-throughput data generation. The typical feature of a microarray based genotyping platform is the large number of SNPs that can be genotyped from one or more samples at the same time.

SNP genotyping methods hold two components: at first a method for discrimination between alternative alleles and second a method for reporting the presence of the allele/s in a given DNA sample. Determination methods include primer extension, enzyme cleavage and allele-specific hybridization. The detection is based on light signals emitted at specific spots on these chips [216].

Several companies such as Applied Biosystems, Luminex, Fluidigm, Affymetrix or Illumina provide high-throughput genotyping arrays commonly used in GWASs.

In this thesis, all individuals of the Max-Planck Institute of Psychiatry (MPIP) cohort were independently genotyped utilizing the Illumina Human610-Quad and Illumina Human660W-Quad Genotyping BeadChips, using the llumina Sentix Human-1, HumanHap300, Illumina Human610-Quad and HumanOmniExpress Genotyping BeadChips for the Munich antidepressant response signature (MARS) cohort and Illumina HumanOmniExpress BeadChips for the Duke Neurogenetics Study (DNS) cohort (see section 3.2).

Hardy-Weinberg equilibrium, genotypic- and allelic frequencies

Allele frequency defines the rate of a single allele in a population and is calculated by dividing the number of times the allele of interest is observed in a population by the total number of copies of the alleles at a particular locus[1]. The genotype frequency refers to the frequency of the different combinations of those alleles in the population. One of these alleles will appear less frequently than the other, which is then defined as the minor allele. Typically, SNPs with a low minor allele frequency (MAF)(<5%) are excluded in GWASs to avoid misclassification bias, since variants with a low MAF do not show much variation across the population and their detection becomes unlikely [136].

The Hardy-Weinberg equilibrium (HWE) is a principle describing that the genetic variation in a population in the absence of disturbing factors will remain constant over generations. Moreover, when mating is random in a large population and disturbing factors,

[1] http://www.nature.com/scitable/definition/hardy-weinberg-equilibrium-122

e.g. natural selection, mutations and inbreeding, are excluded the law predicts that both genotypes and allele frequencies will remain constant because they are in equilibrium (Eq. 1.1). For instance, a mutation can destroy the equilibrium by introducing new alleles into a population or nonrandom mating and natural selection by changing gene frequencies[1]. Deviation from HWE within genotyped SNPs can provide wrong evidence for association. Therefore testing for HWE is part of the normal quality control of GWASs and markers are typically exclude if the P value of the HWE test is less than 10^{-5}.

The Hardy-Weinberg principle can be illustrated with the following equation:

$$\begin{aligned} p^2 &= f(AA) \\ 2pq &= f(Aa) \\ q^2 &= f(aa) \\ p^2 + 2pq + q^2 &= 1 \end{aligned} \tag{1.1}$$

Linkage disequilibrium, haplotypes and tagging SNPs

The pattern of association between SNPs in the genome can be derived from haplotypes and linkage disequilibrium (LD).

A haplotype is a combination of a set of alleles at a number of closely spaced sites on a chromosome [74].

In a GWAS, the genetic phenomenon of a non-random association of alleles at two or more loci is important. Resulting genetic markers in proximity of a disease-causing variant will be more often co-inherited with this disease-causing variant than expected under independent conditions. In fact, the closer two genes are on a chromosome the higher their chances of being inherited together. Contrary, for more distant genes the likelihood of separation during recombination is greater[2]. This lack of independence among different genetic variants is termed LD. Some studies reported that physical distance does not always explain the level of linkage and small distance does not ensure a high level of LD [207, 228, 185, 199]. For example, Abecasis et al. [1] showed that only 45% oft their observed variation in disequilibrium measures could be explained by physical distance. Additional factors, such as allele frequency, type of polymorphism, and genomic location must be taken into consideration.

Various statistical measures are used to quantify LD between alleles of two loci. D' and r^2 are most widely used and depend on linkage coefficient, D (Eq. 1.2), which is defined for a specific pair of alleles, A and B, and does not depend on how many other alleles are at the two loci- each pair has its own D.

$$\begin{aligned} D &= f_{AB} - f_A f_B \\ f &= frequency \end{aligned} \tag{1.2}$$

The smallest possible value, D_{min}, is the less negative value of $-f_A f_B$ and $-1(1-f_A)(1-f_B)$. The largest possible value, D_{max}, is the smallest of $f_A(1-f_B)$ and $f_B(1-f_A)$. D'

[2] http://www.nature.com/scitable/definition/linkage-51

is defined as the ratio of D to its maximum possible absolute values of D (Eq. 1.3), given the allele frequencies [199].

$$D' = \frac{D}{D_{min/max}} \tag{1.3}$$

The second measure to quantify LD is:

$$r^2 = \frac{D^2}{f_A(1-f_A)f_B(1-f_B)} \tag{1.4}$$

D' and r^2 range from zero (independence; no LD) to one (complete LD), but their interpretation is slightly different. D' equal to one if just two or three of the possible haplotypes are present, and it is less than one if all four possible haplotypes are present. An r^2 equal to one can just be reached if only two haplotypes are present. In association mapping the r^2 is the LD measure of choice, since there is a inverse relationship between r^2 and the sample size required to detect association between susceptibility loci and SNPs (for more details please see [247]) [228].

The human genome is split into blocks of high LD regions, which are known as haplotype/LD blocks. The length of a LD block varies across different ethnics. The LD in Europeans and European-Americans extends larger distances than in Africans and African-Americans, which might reflect a population bottleneck at the time when modern humans first left Africa[228].

The most widely studied region in humans is the major histocompatibility complex (MHC)- also referred as human leukocyte antigen (HLA) complex, which is located on chromosome 6 and spans 3.6mega base pairs (Mb) (extended MHC spans 7.6Mb [93]). This regions is known for its high degree of LD. It is arranged into conserved extended haplotypes of variable size, which makes portioning into LD blocks highly complex [15].

The LD pattern can further be used to create non-redundant sets of SNPs (LD bins/ SNP bins), this process is called "tagging". SNPs, capturing other SNPs on the basis of LD patterns are defined as tagging SNPs. If a tagging SNP is correlated with a trait of interest the markers in high LD should exhibit association to this trait as well. A sufficient selection of tagging SNPs can provide enough information to predict the information about the other variants in LD [101].

Population stratification

One of the key challenges in GWASs is to avoid spurious associations. Such misleading associations can occur due to confounding factors, e.g. inconsistency in data collection methods or differences in allele frequencies within subpopulations (for more details see section 1.2.2.2). Especially in GWASs with a case-control design, differences in allele frequency among cases and controls unrelated to the outcome of interest can cause spurious associations between phenotype and genotype. Population stratification refers to those

ancestral differences. Large samples comprising multiple populations can interfere the LD structure or lead to deviation from the HWE [176]. Methods dealing with correction for population stratification include the genomic control approach, structured association approach, multidimensional scaling (MDS), principal component analysis (PCA) or linear mixed models. The genomic control [46] corrects stratification by adjusting association statistics at each marker by a common factor for all SNPs (λ_{GC}). The structured association correction [177] assigns the sample to discrete subpopulation clusters and then accumulates evidence of association within each cluster. The most widely used method is the PCA approach. The EIGENSTRAT method [176], uses PCA to identify several top principal components (PC)s and includes them as covariates in the association analysis. A PC is defined as the product of a weight vector and a genotype vector, with weights reflecting the marginal information about ancestry [229]. Another method correcting for stratification, which is equivalent to PCA for certain similarity matrices, is the MDS approach.

Missing genotypic values and imputation

Missing genotype information is a frequent problem within GWASs. Since most of the analysis tools cannot handle missing values, they have to be removed prior to their application, resulting in a considerable loss of information.

Imputation methods address this problem by using the LD information within a region to predict missing genotypes at typed SNPs or genotypes that are not genotyped at all [80, 69]. Based on a reference panel of samples from identical or similar populations that was produced by whole-genome sequencing, imputation methods infer genotypes at markers that were not directly typed in a study. Samples from the HapMap project [96] and/or 1,000 Genomes Project (1KGP) [225] are used as reference panels. Those reference panels contain a much larger number of SNPs, because they were produced by genome sequencing. The 1KGP reference panel allows a deeper analysis of the contribution of genetic variation than the HapMap data. It benefited from whole-genome sequencing technology, which increased resolution significantly. The obvious advantage of imputation is a considerable gain of information at no or low additional costs. Before the rise of GWASs, imputation methods were successfully applied to association studies [83]. Unfortunately, most of them do not work for genome-wide approaches due to excessive computational costs of the algorithms. However, other algorithms especially developed for the genome-wide task proved themselves to be particularly useful because of the dense and numerous marker sets and background information made available by these kinds of studies.

Computationally intensive tools used for genotype imputation include IMPUTE [139], MACH [123], BEAGLE [24] and fastPHASE [187]. They mainly differ in the approach to choose the ancestral haplotypes.

In this thesis we utilized IMPUTE version 2 (see section 2.3) to estimate the genotypes of incomplete or untyped SNPs applying combined reference data of HapMap and 1KGP. Briefly, IMPUTE first aims to identify shared haplotypes of the individuals from a study panel and the haplotypes in the reference sets. Individuals in the study and reference panels share a degree of common ancestry. Therefore, different parts of the study data will

be more closely related to different individuals of the reference panel, thus, the haplotypes of a given individual could be modeled as a mosaic of haplotypes of the related individuals. Missing genotypes in the study sample were then imputed on the basis of the matched haplotypes of the reference set. The output for each imputed SNP is the probability of the distributions over the genotypes 0, 1 or 2, which was used to estimate the missing genotype [139, 138].

Analyzing the imputed SNPs can lead to more significant associations and a more detailed view of associated regions.

1.2.2. Gene expression as molecular phenotype

The expression of genes is an intermediate molecular phenotype, which can help to identify genetic variation responsible for psychiatric disorders. Gene expression likely reflects both, state and trait dependent disease-related influences and these have been shown to be highly heritable [211].

1.2.2.1. Analysis of gene expression profiles

Gene expression can be detected using sequencing-based (RNA sequencing (RNA-seq)) or hybridization-based approaches (microarray). In this thesis, gene expression microarrays (see section 2.2) provided the measurements on gene expression. Microarrays utilize the principle of complementary hybridization between nucleic acids and the advantage of the knowledge of the genome. High-density microarrays harbore probe sequences complementary to thousands of genes, each immobilized at a specific coordinate on the surface of the array. To measure gene abundance from certain cells of tissues the ribonucleic acid (RNA) is extracted and labeled fluorescently or radioactively. The tagged RNA hybridizes specifically to complementary DNA (cDNA) sequences on the array and the signal is proportional to the abundance of RNA in the sample [147]. The emitted light signal is detected using autoradiography, chemiluminescence, or fluorescent scanning. Quantitative signal intensity scanning allows gene expression levels to be measured by their positions on the microarray and level of hybridization [224].

Microarrays allow researchers to measure the expression of thousands of genes simultaneously with relatively low costs as compared to the sequencing methods. Nevertheless, microarrays have some limitations. They depend on the quality of the available genome annotations, there is cross-hybridization between similar array probe sequences and they are only able to detect known transcripts- since they interrogate a fixed content.

Furthermore, microarray technologies are very sensitive to create batch effects. Due to practical reasons, the number of samples that can be hybridized or amplified at the same time is limited, resulting in different runs that might be several days or weeks apart. Differences in lab conditions and preparation methods can further contribute to the variation within the gene expression data [118].

Differential gene expression

Examining the differences in gene expression levels across two or more experimental groups is referred to as "differential expression analysis". This can be done by targeting a specific gene or by utilizing genome-wide gene expression levels [41]. In this thesis, we identified significant differences in gene expression levels between cases and controls as well as differences in gene expression levels after dexamethasone between cases and controls (see section 3.1). Further we showed that genetic variants altering the transcriptional response to GR activation are relevant for psychiatric disorders (see section 3.2).

Functional genomic analyses of gene expression

Transcriptional regulation of gene expression

The aim of functional genomics is to elucidate the functions of genes by identifying the locations of all their regulatory elements. A regulatory element, including TFs, DNA methylation patterns, promoters and enhancers, refers to a discrete region in the genome that encodes a defined product or a reproducible biochemical signature such as transcription or specific chromatin structure [57]. Since, the regulation of gene expression mainly occurs at the transcriptional level, the identifiction of regulatory elements plays an important role in understanding regulation of gene transcription [231]. The formation of regions of open chromatin is a key factor elucidating functional regulatory activity. The structure of the DNA is organized by nucleosome packing, acting as a regulator of transcription by enabling or restricting protein binding, and therefore influencing the activity of a gene [132]. Regions of open chromatin are indicated by deoxyribonuclease I hypersensitive sites (DHSs) and deoxyribonuclease I (DNaseI) mappings which have been instrumental in the discovery of regulatory elements. Chromatin structure can be profiled with DNase-seq [201], a combination of DNaseI digestion and high-throughput sequencing, as well as FAIRE-seq (formaldehyde-assisted isolation of regulatory elements followed by sequencing) [70]. Histone modifications can be also assessed by chromatin immunoprecipitation (ChIP) followed by high-throughput sequencing (ChIP-seq) [67]. The same technique can be used to map the genomic location for TF binding sites. The binding of TFs can modulate transcription levels and influence the activity of specific genes. Further, chromatin contact between specific regulatory elements highlights an important feature in gene expression regulation. The characterization of these physical interactions can be carried out by chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) [242]. This technique converts chromatin structure into millions of short tag sequences and by combining ChIP with chromatin conformation capture (3C) technology and high-throughput sequencing. Thereby, higher order chromatin interactions at genome-wide resolution can be assessed.

The encyclopedia of DNA elements (ENCODE) project [58] proposes to build a comprehensive map of all functional elements in the human genome. Parts of the ENCODE data were used in this thesis to evaluate whether the long-range regulation of GR-response expression quantitative trait locus (eQTL)s may also be associated with long-range physical chromatin interactions, (see section 3.2.2). Therefore, our eQTL data (see section 1.2.2.2) were compared to the ENCODE DHSs and ChIA-PET data. Furthermore, we investigated

whether specific TF binding motifs are enriched in the set of GR-response eSNPs.

Of course, other mechanisms to modulate gene expression exist, containing e.g. microRNA (miRNA)s, non-coding RNAs or DNA methylation on the epigenetic level.

Gene network analysis

Gene networks provide a straightforward representation of the relationship between genes. Different network-based approaches can be used, including gene co-expression, protein-protein interaction as well as cell-cell interaction [241].

In general, gene networks are used to identify higher-level features of gene-gene relationships based on graph theoretic consideration like node degree or clustering coefficient. Genes functionally related to each other, often show correlated gene expression profiles. Therefore, co-expression can be utilized to identify clusters of genes, which share a biological function. Similarity can be measured by Pearson correlation coefficient, Kendall's τ correlation or Spearman's rank correlation. The similarity measure is applied to all possible gene pairs generating a symmetric matrix of correlation values [151]. The associations established from a co-expression analysis can be illustrated in a network, depicting a gene expression profiles as a node and shared edges between nodes indicate a significant pairwise expression profile association [241]. Evaluating the broader structure of networks allows detecting groups of even higher co-regulation (modules). The genes in a network can be characterized by the number of connections they have. Highly connected genes are major players in a network and are referred to as hubs [152].

In this thesis we used the Pearson similarity measure to test if a set of gene products identified in the GR-response eQTL analysis tends to be more co-regulated on a transcriptional level than random eQTL genes (see section 3.2.6).

Furthermore, these genes were analyzed for physical interactions, regulatory interactions and association with psychiatric disorders based on manually curated relationships extracted from the scientific literature (see section 3.2.6).

Classification

The ability to distinguish between classes of samples, like patients and healthy controls, is especially important in psychiatric research, since the diagnosis is solely based on verbal information. Gene expression data can be used for classification and help to identify biomarkers for psychiatric disorders. The primary classification methods include k-nearest neighbor classifiers, discriminant analysis, neural networks, logistic regression, support vector machines and classification trees.

The classification of microarray gene expression data is challenging due to the large number of genes relative to the small number of samples. Typically, these data sets contain around 12,000 genes for less than 100 samples [51], including a large number of irrelevant or redundant genes uninformative for classification. Therefore, it is crucial to reduce the number of genes in order to achieve good classification accuracy. Feature selection and dimension reduction are the main approaches for this purpose. Feature selection methods are: entropy-based, t-statistics, correlation-based, χ^2 and signal-to-noise statistic [2].

Creating a classifier is a multi-step approach. Firstly, the set of samples is divided randomly into training and test set, using the hold-out, k-fold cross validation or leave-one-out approach. The model is fit on the training set and then used to predict the responses for observations in the test set. The performance of this model can be represented in a confusion matrix that reports the number of true (TP) and false plosives (FP) as well as true (TN) and false negatives (FN). From these values the true positive rate (TPR) (sensitivity; see Eq. 1.5), true negative rate (TNR) (specificity; see Eq. 1.6) and accuracy (ACC; see Eq. 1.7) can be calculated.

$$TRP = \frac{TP}{(TP + FN)} = \frac{TP}{P} \tag{1.5}$$

$$TNR = \frac{TN}{(FP + TN)} = \frac{TN}{N} \tag{1.6}$$

$$ACC = \frac{TP + TN}{(P + N)} \tag{1.7}$$

In this thesis, we performed classification with the random forest (RF) classification technique [20]. It gains some advantages compared to other classifiers: usage of bagging on samples, random subset of variables and majority vote schema [2]. The RF uses the training set to create an in-bag partition to construct the tree and the test set, which is not used in the construction of the tree, is used as an out-of-bag (OOB) partition to test the constructed tree and to evaluate the performance. All trees vote to determine the prediction results and an OOB estimate of the error rate is calculated. The RF provides the mean decrease Gini as an importance measure that calculates the quality of a split for each node of a tree by means of the Gini index. Each time a node is split on a variable, the Gini index for both descendent nodes is less than the ancestor node. A higher mean decrease Gini represents a higher variable importance [20, 161]. In this thesis, the Gini index was used as importance measure to run a feature section before starting to create a classification model (see section 3.1).

1.2.2.2. Analysis of expression quantitative trait loci

The study of the genetic source of variation in gene expression is known as expression quantitative trait locus (eQTL) study. It combines established quantitative trait locus (QTL) mapping with genome-wide gene expression data (see figure 1.2 for illustration). These studies take advantage of technical developments such as microarrays, which allow the parallel measurement of expression levels and genotypes of thousands of genes in large numbers of samples [32].

The mapping of eQTLs is based on the assumption that gene expression levels can be incorporated with genotype data in the same way as any other phenotype, such as age or body weight. Typical data sets for eQTL studies consist of a set of two data matrices. The first $n \times G$ matrix contains gene expression values for diverse individuals (n) and the measured transcripts (G; plus several phenotypes). The second $n \times S$ matrix

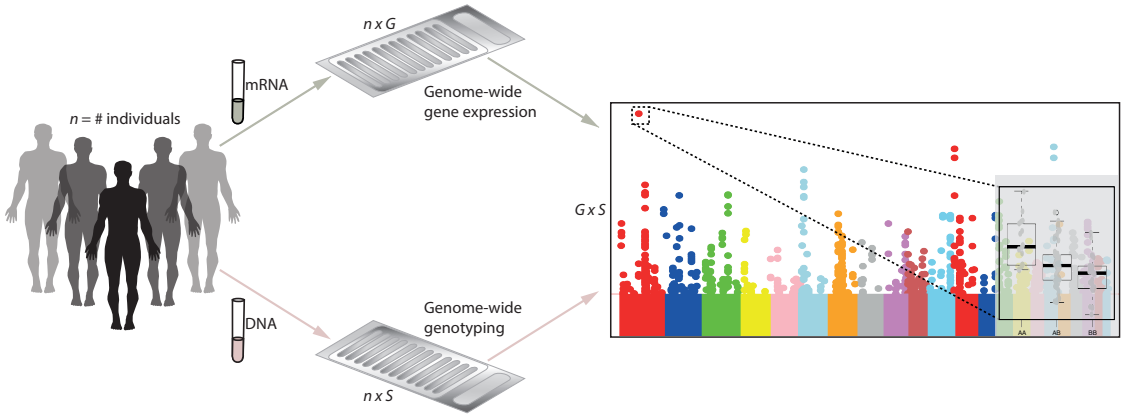


Figure 1.2.: Genetic association mapping of genome-wide gene expression data. Genome-wide genotyping (SNP) and gene expression data (treated as quantitative phenotype) were combined to identify genetic loci that control quantitative variation in gene expression (eQTLs).

similarly contains all participants (n) and collected information on genomic markers (S) [33]. Thereby eQTL studies connect variation in DNA sequence level to that at the RNA level. The significance of these eQTL studies is further enhanced by the possibility to scan for regulators without the need of prior knowledge of the mechanisms and by simultaneously investigating many gene expression probes enabling the identification of co-regulation, i.e. co-expression networks [32]. The understanding of eQTL data can be further increased by integrating additional biological information e.g. epigenetic or regulatory factors [36]. eQTLs are influenced not only by genetic polymorphisms but also by other biological effects, such as heritability, and *cis* and *trans* effects, which are further characterized in the following paragraphs.

Application of eQTL studies are deeper understanding of the genetic basis of complex disorders and identification of gene networks involved in disease pathogenesis.

Heritability of eQTLs

eQTL studies take advantage of variation in gene expression to detect the underlying genetic cause. A key question is to what extend the phenotypic variation in a trait can be attributed to genetic factors, i.e. heritability [236]. Heritability, attributable to additive genetic factors, is defined as narrow-sense heritability. It can be inflated by non-additive genetic effects, such as epistasis, referring to broad-sense heritability [136]. It is inherently difficult to restrict the contribution of non-genetic factors in humans. An inference of the genetic contribution to polymorphic variation on the level of gene expression can be assessed by estimating heritability of genes by familial aggregation studies. Evidence for familial aggregation was observed by comparing variation among unrelated individuals, among siblings within families and between monozygotic twins [31].

Previous studies [240, 56, 73, 48, 31] have shown that levels of gene expression are highly

heritable using a variety of cells and tissues. For example, Dixon et al. [48] identified 15,084 eQTL probes (out of 54,675 probes) representing 6,660 genes with narrow-sense heritabilities > 0.3 in lymphoblastoid cell lines (LCLs). Lately, in the Multiple Tissue Human Expression Resource (MuTHER) study [76] gene expression across multiple tissues in 856 twins was analyzed. For expressed genes they reported mean heritability values of 0.16 (skin), 0.21 (LCLs) and 0.26 (adipose), with higher estimates for significant *cis*-eQTLs (> 0.3). Most recently, Wright and colleagues [237] analyzed gene expression of 18,392 genes from peripheral blood in 2,752 twins and reported low heritability estimates. Interestingly, they detected that a significantly positive association between heritability and GWAS genes and heritability is strongly associated with expression mean and variance.

Cis- and trans-effects of eQTLs

When interpreting eQTL data one must consider that identified loci can influence the gene expression levels either in *cis* (local; see figure 1.3) or in *trans* (distal; see figure 1.4). To distinguish both effects a search radius (ϵ) has to be selected and for each expression probe p a *cis* -window is termed by Eq. 1.8.

$$N_p = N_{p,\epsilon} = (a_p - \epsilon, b_p + \epsilon) \tag{1.8}$$

,where a_p refers to the coordinated of the 5' end and b_p to the 3' end of the expression probe p. $|N_p|$ denotes the number of SNPs in the defined search space. In this thesis ϵ was set to 1Mb to define *cis*-eQTLs (see section 2.5.2). *Cis*-acting eQTLs include SNPs in the vicinity of the regulated transcripts. To date, eQTL studies have found more *cis* than *trans* effects. This is probably due to the diverse thresholds used for statistical significance and sample sizes [32]. *Cis*- associations have larger effects, are more stable across statistical methodologies and are more common [41], resulting in an easier detectability than *trans* associations. Several studies have indicated that $> 90\%$ of the *cis*-eQTLs in humans are located 100kilo base pairs (kb) upstream and downstream of the transcript [48]. The *cis*-eQTLs cluster symmetrically around the transcription start or end sites and reflect polymorphisms located directly upstream of the transcript [211].

When the genomic distance between gene expression probe and genetic variant is large (usually greater than 1Mb), the eQTLs are defined as *trans* [211]. Because *trans*-acting variants can be anywhere in the genome relative to the target gene and stringent statistical correction for multiple testing has to be applied, it is more difficult to detect these effects. Furthermore, their effects are usually weaker than *cis* effects. This is possibly due to the combined influence on a gene by multiple *trans*-acting regulators, whereas usually one or only very few *cis*-acting regulators [32] have sufficient influence. However, *trans* associations are of interest since they are mainly mediated through transcription factors and can possibly reveal master regulators (hubs) of transcription [41, 36].

Of note, the declaration as *cis* or *trans*-regulator only implies the distance of the genetic signal to the target gene [32] but carries no functional significance.

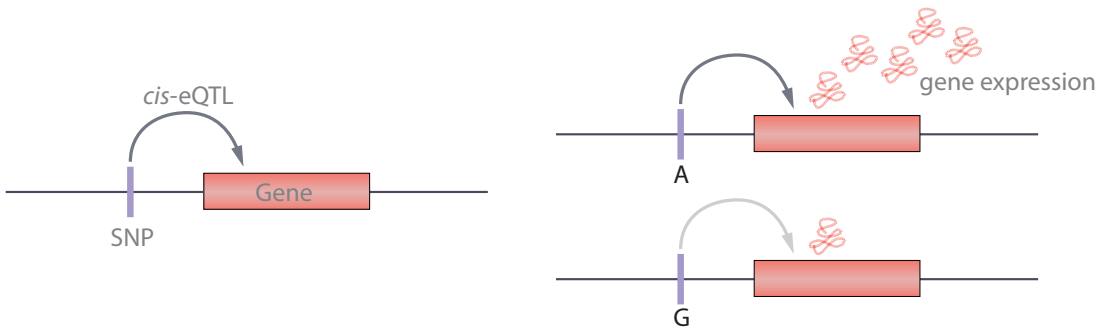


Figure 1.3.: Effect of *cis*-acting variants on expression levels of genes. *Cis*- acting variants are found to be in the vicinity to the target genes. Different allelic forms of the genetic variant show different influences on gene expression. Here, individuals with the A-allele have higher expression levels of the target gene than individuals with the G-allele.

Tissue specificity of eQTLs

One of the first questions in designing an eQTL study regards the type of cells to utilize in the study [32]. Ideally, RNA for eQTL analysis should be obtained from a wide range of tissues, to downstream analyze disease-associated SNPs in the disease-relevant tissue/s. For many diseases, it is very difficult to obtain disease-relevant tissue from living humans, this is especially true for brain tissue needed in psychiatric research. However, the Genotype-Tissue Expression (GTEx) project achieved to collect gene expression data in over 60 tissues across approximately 1,400 individuals [77].

Studies in humans are primarily investigated in blood and subcutaneous adipose tissues [56], cells from liver samples [186], lymphocytes [73] and cells from tissue from brain banks [157]. These studies indicated that eQTLs are common [77]. Several studies have reported tissue specific effects of eQTLs [76, 244]. The simplest approach to determine eQTLs in different tissues is by mapping eQTLs separately for each tissue and subsequently examining the overlap (tissue-by-tissue analysis) [63]. Another approach is the joint-analysis on all tissues, while simultaneously allowing for differences among eQTLs present in each tissue. This method maximizes the power to detect eQTLs that are actually shared among tissues. It is superior to the tissue-by tissue analysis, which, due to incomplete power, leads to high numbers of falsely interpreted tissue-specific eQTLs. For example, Flutre and colleagues [63] re-analyzed *cis*-eQTLs in samples of LCLs, T-cells and fibroblasts from Dimas and colleagues [47] by applying both, the joint and tissue-by tissue approach. They showed an increased power to identify eQTLs with the joint-analysis, and discovered additional 63% genes with eQTLs. Moreover, they concluded only 8% of the eQTLs to be tissue-specific and 88% to be common between all three tissues. Originally, using the tissue-by-tissue approach 69-80% of the eQTL were thought to operate in tissue-specific manner. Another approach to investigate eQTLs shared among different tissues is to study cross-heritability of gene expression. The method is based on polygenic models and attempts to estimate

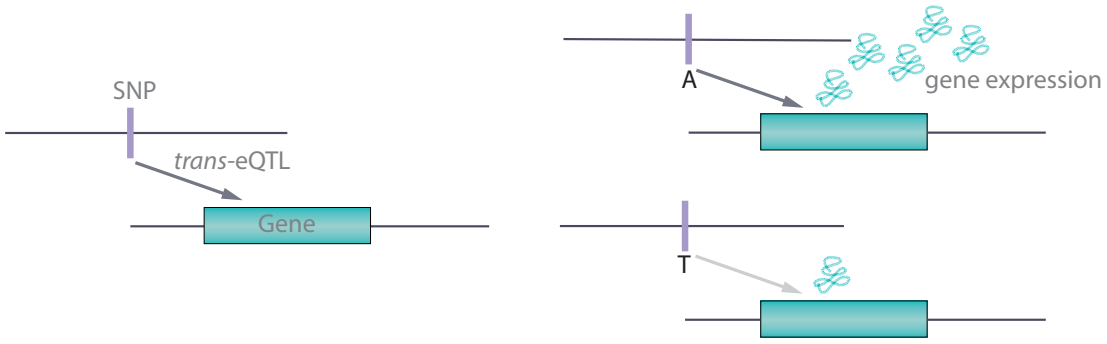


Figure 1.4.: Effect of *trans*-acting variants on gene expression levels of genes. *Trans*-acting variants are distally located from the target gene, often on other chromosomes. Different allelic forms of the genetic variant exhibit different influences on gene expression. Here, individuals with the A-allele have a higher expression level of the target gene than those with the T-allele.

the combined influence of all shared eQTLs [63].

An attractive choice of tissue in psychiatric studies is blood since it is the most accessible tissue [32] in the clinic. Many studies show the utility of blood as a surrogate for brain tissue. This has been supported mainly through indirect investigations of expression profiles in relation to specific neurological and psychiatric disorders [26]. However, the question remains, are the resulting eQTLs meaningful? In this thesis, we show that it is indeed possible to differentiate between MDD cases and controls utilizing GR-stimulated gene expression patterns obtained from peripheral blood samples (see section 3.1). Furthermore, we compared gene expression profiles of human GR-response eQTL genes from peripheral blood cells to mice assays from blood and brain tissue. Over 65% of the transcripts identified as candidate genes in our human blood data were also regulated following GR agonist stimulation in mice (see section 3.2.6). Furthermore, a study by Jasinska and colleagues [99] determined gene expression profiles across eight brain regions (cerebellar vermis, pulvinar, head of caudate, hippocampus, occipital pole, orbitofrontal cortex, frontal pole, dorsolateral PFC) and peripheral blood in male monkeys. The primary focus of the study was to identify the correlation of brain and peripheral blood transcriptional profiles. Candidate transcripts were selected for mapping brain eQTLs in peripheral blood. The gene expression of 33% of the transcripts expressed in both brain and blood was found to be highly correlated. Furthermore, Sullivan et al. [212] report correlations of gene expression patterns available in whole blood and multiple brain tissues of around 50%.

Confounding factors of eQTL studies

Confounding factors affect the relationship between an independent and depend variable, thus evoking false correlations leading to incorrect results. Several eQTL studies have shown that confounding variables reduce the power to detect eQTLs [206]. Some confounders are known and others are unknown. Well known confounders include population differences, sample size and technical source of variation referred to as batch effects. Batch effects are systemic variation between groups of samples (batches) induced by experimental features that are not of biological interest [181], i.e. chip type, protocols (sample preparation, amplification, hybridization and labeling), platform, laboratory, staff (technician), storage (time and place), etc. [206, 130, 181]. They can be minimized or even avoided by a thoroughly planed experimental design, considering e.g. randomizing possibly present groups. This means, in cases-control studies, one should always avoid to process all cases on one day and all of the controls on another day. It will be impossible to distinguish the introduced batch effect from any real biological effects [118]. But many types of batches are unavoidable. For example, larger studies with huge sample sizes and have to be carried out over many month or even years [117]. A different issue are the unknown or hidden factors, which cannot directly be removed as they are not measured [206]. Normalization methods for gene expression microarrays have been widely used to adjust for such experimental artifacts between samples [130]. The methods remove systemic effects within or between microarrays (chips). However, they are not directly designed to remove batch effects. Therefore, batch effects may often remain after normalization. Consequently, multiple methods have been developed to remove batch effects, including PCA [181], empirical Bayes approach- called ComBat [102] and surrogate variable analysis [119].

Multiple-testing problem in eQTL studies

The multiple testing problem

Multiple-testing refers to any instance that tries to test a set of hypotheses at the same time. Take the case of $n = 100$ hypotheses to be tested simultaneously, using some level of significance α. For $\alpha = 0.05$, one expects five true hypotheses to be rejected. Further, if all tests are independent, the probability that at least one of the significant results is due to chance is given by $1 - (1 - \alpha)^n = 1 - 0.95^{100} = 0.994$ (family-wise error rate (FWER)). Thus, with 100 tests being considered, one has a 99,4% chance of observing at least one significant result, even if all of the test are actually not significant. In microarray studies the number if simultaneous tests is quite large, and if one does not take the multiplicity of tests into account, the probability that some of the true null hypotheses are rejected by chance alone is very high [184].

The Bonferroni correction

A very conservative method to corrected for multiple-testing is the Bonferroni correction, which reduces the number of FP and at the same time it also reduces the number of true discoveries. It sets the significance cut-off at α/n. For example, with $n = 100$ tests and $\alpha = 0.05$, the null hypothesis will only be rejected if the P value is less than 0.0005 [184].

	called significant	called not significant	total
null true	F	$m_0 - F$	m_0
alternative true	T	$m_1 - T$	m_1
total	S	$m - S$	m

Table 1.4.: Hypotheses definition. The table is based on [210].

Benjamini and Hochberg's false-discovery rate control

A less conservative method is the false discovery rate (FDR) correction introduced by Benjamini and Hochberg [9]. The FDR method has been widely used to detect differential gene expression in microarray experiments [210]. It considers the probability of one or more false positives (FP) discoveries among multiple tests and estimates the proportion of FP in the result [185]. Thus, FDR is a sensitive measure of the balance between the number of TP and FP [210].

Suppose we have m P values with various outcomes that can occur when a significance threshold is applied to them. Table 1.4 summarize these outcomes. F is the number of FP, T is the number of TP and S is the total number of features called significant. m_0 is the number of features that are true and $m_1 = m - m_0$ is the number of truly alternative features. The FDR is can be expressed as in Eq. 1.9, $E[^*]$ being the overall error measure in terms of an expected value [210].

$$FDR = E\left[\frac{F}{F+T}\right] = E\left[\frac{F}{S}\right] \tag{1.9}$$

,where FDR is defined to be 0 when $S = 0$ [210]. The FDR adjusted p-values are called q-values.

Resampling-based methods

Resampling-based multiple-testing corrections are the most commonly used methods to adjust the significance of differential gene expression between classes. Some common resampling-based methods include bootstrapping or permutation tests. These methods create a pseudo-data set, where the phenotype values are randomly sampled and reassigned to individuals with (e.g. bootstrap) or without (e.g. permutation) replacement. A statistical test is applied to the pseudo-data set [227]. Westfall and Young [233] suggested to compare the observed minimum P value for given pseudo-data set (pseudo-p-value) with the actual P value and record (R) the number of times the pseudo P value is equal or less than the actual one. This procedure is repeated several thousand times (R counters) and the proportion of resampled data sets with a minimum pseudo P value less or equal to an actual P value is the adjusted P value.

Multiple-testing correction for eQTLs

For eQTL studies a huge number of statistical tests will be examined. A typical study includes 500,000 genotyped or up to 8 million imputed SNPs and approximately 15,000 gene expression probes, resulting in $(500,000 - 8,000,000) \times 15,000 = 7.5 \times 10^9 - 1.2 \times 10^{11}$ tests. In this case a large number of false positive results is expected. To increase the amount of meaningful information obtained from eQTL studies, appropriate multiple-testing correction is crucial [243].

As eQTL data contains two kinds of data, genotyping as well as gene expression data, a two stage design is common. Multiple testing correction on both (i) SNP level and (ii) gene expression probe levels is performed. An according approach was used in this thesis (see section 2.5.2). Briefly, (i) one of the most commonly used method for multiple-testing correction across multiple genetic marker for each phenotype include the resampling-based procedure using permutation or bootstrap (here: permutation). The null distribution is simulated with permuted phenotype values. In detail, the phenotype values are randomly shuffled and reassigned to individuals without replacement. To find the maximal test statistic for each phenotype among SNPs a genome-wide scan is performed. The adjusted P value is the proportion of permutated phenotypes with a maximal test statistics greater than the actual test statistic of the original data [243]. The resampling-based test preserves the correlation structure of the SNPs (LD) and does not require any assumption of distribution for the test statistic [243]. In this thesis the permutation-based p-values for each phenotype were further adjusted by a Westfall-Young correction regarding the number of SNPs per probe. (ii) A second level of multiple-testing problem in eQTL studies are the multiple tests across gene expression traits. Therefore, the prior estimated adjusted empirical p-values (of the most significant association- independence) for each expression probe was used to determine a threshold for the adjusted P values across all gene expression probes by controlling for the FDR [243].

2. Material and Methods

2.1. Samples and study design

2.1.1. MPIP cohort

The subject pool for the eQTL analysis consisted of 164 male Caucasian individuals (90% of German origin) recruited within the MARS project [97]: 93 healthy probands (age = 40.2 ± 12.4 years; body mass index (BMI) = 24.9 ± 3.1 kg/m^2) and 71 in-patients with depressive disorders (age = 48.5±13.5 years; HAM-D = 25.3±8.0; BMI = 26.1±3.6 kg/m^2) treated at the hospital of the Max Planck Institute of Psychiatry in Munich, Germany (MPIP cohort). Only individuals not reporting a history of current psychiatric disorders as well as major neurological and general medical disorders were included in the control sample. Recruitment strategies and further characterization of the MPIP cohort have been described previously [148, 88]. Of these participants, 4 were excluded due genotyping problems. Baseline whole blood samples (for plasma and RNA) were obtained at 6pm after 2 hours of fasting and abstention from coffee and physical activity and immediately followed by oral administration of 1.5 mg dexamethasone. A second blood draw was performed three hours later at 9pm (see figure 3.3a). Cortisol and ACTH serum levels were determined using previously described radioimmunoassays [148]. Plasma dexamethasone concentrations were assessed in serum samples drawn at 9pm using Liquid chromatography-tandem mass spectrometry on API4000 (AB Sciex). The study was approved by the local ethics committee and all individuals gave written informed consent.

A subsample of 60 participants (29 patients with depressive disorders and 31 healthy probands; see table 2.1 for full description of diagnoses) out the 164 participants from the eQTL analysis were used to study differential gene expression between patients and controls. This subsample was further randomly subdivided into two samples, which were used as training and test set for classification.

2.1.2. MARS cohort

This sample included 1,005 unipolar depressed patients (561 female, 444 males; age = 48.15 ± 14.13 years; HAM-D = 25.68 ± 6.5), as well as 478 controls (298 females, 180 males; age = 47.83 ± 13.7 years), recruited for the MARS project at the MPIP in Munich, Germany. All included patients were of European descent. Recruitment strategies and further characterization including population stratification of the MARS cohort have been described previously [148, 88]. All individuals used within the eQTL study (MPIP cohort) were not part of this sample.

	Sample 1 (training set)		Sample 2 (test set)		
	Cases $n = 18$	Controls $n = 18$	Cases $n = 11$	Controls $n = 13$	P value
Mean Age	44.2 ± 14.3	43.2 ± 10.8	50.8 ± 15	37.7 ± 10.2	0.075
Mean BMI	24.5 ± 3.0	25.3 ± 3.5	26.1 ± 2.5	24.7 ± 4.5	0.612
Mean age of onset	33.3 ± 13.3		30.1 ± 13.2		0.548
Mean HAM-D	25.8 ± 7.9		23.9 ± 9.8		0.573
Bipolar disorder	5 (28%)		2 (18%)		0.649
Prev. episodes	2.7 (3.7)		4.6 (9.6)		0.486
Recurrent depression	13 (72%)		6 (55%)		0.331
Prev. suicide attempts	3 (17%)		4 (36%)		0.172
Family history	9 (50%)		7 (64%)		0.306
Response 5 weeks	9 (50%)		5 (45%)		0.306
Remission 5 weeks	7 (38%)		5 (45%)		0.125
Medication*					
TCA	4 (22%)		1 (9%)		0.418
SSRI	2 (11%)		4 (36%)		0.074
SNRI	7 (39%)		6 (55%)		0.283
NaSSA	6 (33%)		2 (18%)		0.454
SSRE	1 (6%)		0		0.448
Antipsychotics	6 (33%)		3 (27%)		0.856
Mood stabilizer	5 (28%)		2 (18%)		0.649
Lithium	2 (11%)		1 (9%)		0.927
Benzodiazepine	8 (44%)		8 (73%)		0.069

Table 2.1.: Sociodemographic and clinical characteristics of MPIP cohort subsample subdivided into training (sample 1) and test set (sample 2). ∗at RNA withdraw

2.1.3. DNS cohort[1]

The DNS samples were ascertained at the Duke University in Durham, NC, USA and all participants provided informed written consent prior to participation in accord with the guidelines of the Duke University Medical Center Institutional Review Board. All participants were in good general health and free of DNS exclusion criteria: (1) medical diagnosis of cancer, stroke, diabetes requiring insulin treatment, chronic kidney or liver disease or lifetime psychotic symptoms; (2) use of psychotropic, glucocorticoid or hypolipidemic medication, and (3) conditions affecting cerebral blood flow and metabolism (e.g., hypertension). Current DSM-IV Axis I and select Axis II disorders (Antisocial Personality Disorder and Borderline Personality Disorder) were assessed with the electronic Mini International Neuropsychiatric Interview [196] and Structured Clinical Interview for the DSM-IV Axis II (SCID-II) [61], respectively. These disorders are not exclusionary as the DNS seeks to

[1]Parts of the DNS methods have been published previously by our collaborators [175, 16]

establish broad variability in multiple behavioral phenotypes related to psychopathology.

On January 6th, 2014, 726 participants had overlapping fMRI and genetic data that was fully processed and used for these analyses. Of these participants, 79 were excluded due to scanner-related artifacts in fMRI data ($n = 6$), incidental structural brain abnormalities ($n = 2$), a large number of movement outliers in fMRI data ($n = 21$; see Artifact detection tool (ART) description below), inadequate signal in our amygdala regions of interest ($n = 14$; see coverage description below), poor behavioral performance ($n = 20$; accuracy lower than 75%), outlier status according to ancestrally-informative principal components ($n = 5$), scanner malfunctions ($n = 2$), incomplete fMRI data collection ($n = 1$), and failed genotyping at one genetic risk profile score (GRPS) polymorphisms (without a proxy of $r^2 > 0.90$; $n = 8$). Thus, all imaging genetics analyses were conducted in a final European-Americans (EUR-AM) subsample of 306 participants (age = 19.72 ± 1.23 years; 148 males; 63 with DSM-IV Axis I disorder) and a full sample of 647 participants (age = 19.65 ± 1.24 years; 285 males; 117 with DSM-IV Axis I disorder; 306 European Americans, 72 African Americans, 170 Asians, 37 Latino/as, and 62 of Other/Multiple racial origins according to self-reported ethnicity; for a full description of diagnoses present in the sample see table 2.2).

	EUR-AM $n = 306$	**Full sample** $n = 647$
Alcohol Abuse	22	41
Alcohol Dependence	19	31
Major Depressive Disorder	8	17
Marijuana Abuse	7	15
Gernalized Anxiety Disorder	7	11
Social Anxiety Disorder	3	8
Agoraphobia w/o Panic Disorder	6	8
Bipolar Disorder NOS	6	8
Marijuana Dependence	5	7
Bipolar II	3	6
OCD	4	6
Bulimia Nervosa	2	5
Panic Disorder	1	4
Dysthymia	0	1
PTSD	0	1
Anorexia Nervosa	0	1
Bipolar I	1	1
TOTAL	94	171

Table 2.2.: This table represents the number of diagnoses across DNS participants for the European-Americans (EUR-AM) and entire sample. Some individuals presented with comorbid status.

2.1.4. Mouse models

The animal experiments were carried out in the animal facilities of the MPIP in Munich, Germany. Male C57BL/6N mice at an age of 12 weeks (mean bodyweight 26.8 ± 0.1 g) were used for the dexamethasone-stimulation test (DEX-mouse). The experiment was performed twice with two separate batches of mice ($n = 22$ per batch). All mice were kept under 12h light/dark cycle and held under standard conditions. Food and tap water were available ad libitum. All efforts were made to minimize animal suffering during the experiment. The committee for the Care and Use of Laboratory animals of the Government of Upper Bavaria, Germany approved the protocols. Animals were injected i.p. with either vehicle ($n = 11$) or 10 mg/kg dexamethasone ($n = 11$) between 9am and 11am. Animals were sacrificed 4 hours post injection, blood was collected and the brains were carefully removed. The PFC (batch 1), HC (batch 1) and AM (batch 2) were dissected immediately according to standard protocols [202]. Briefly, Amygdala preparation: brains were cut into ca. 1 mm thick slices using a custom-mounting device. Amygdala (all subnuclei [167]) was manually dissected with a scalpel under visual control using a binocular microscope; HC and PFC preparation: brain regions were manually dissected from the whole brain by trained personnel. Dissected tissues were directly transferred into RNA lysis solution (Applied Biosystems, USA) and frozen at −80 °C. In addition, 300 μl of trunk blood (batch 1) was collected into microcentrifuge tubes containing PaxGene RNA stabilizer solution and frozen at −20 °C.

The chronic social stress mice sample includes 12 male CD1 mice (6 susceptible and 6 resilient mice). The chronic social stress procedure and mice sample have been described previously [110, 190, 189]. Brain slices of the hippocampal region were cut at 20 μm and thaw-mounted on membrane-coated slides (Carl Zeiss MicroImaging). Laser dissection of CA1 and dentate gyrus (DG) material was performed using a laser-capture microscope (for more details see Schmidt et al. [190]).

2.2. Gene expression data

2.2.1. MPIP cohort

Human whole blood of the MPIP cohort was collected using PAXgene Blood RNA Tubes (PreAnalytiX), processed as described previously [148] and hybridized to Illumina HumanHT-12 v3.0 Expression BeadChips. Samples had a mean RNA integrity number (RIN) of 7.97 ± 0.42 standard deviation (SD). The Illumina Bead Array Reader was used to scan the microarrays and summarized raw probe intensities were exported using Illumina's GenomeStudio v2011.1 Gene Expression module. Further processing was carried out using R version 2.14.0[2].

[2] http://www.r-project.org/

2.2.1.1. eQTL analysis

All 48,750 probes present on the microarray were first filtered by an Illumina detection P value of 0.01 in at least 10% of the samples, leaving 14,168 expressed probes for further analysis. Second, each transcript was transformed and normalized through variance stabilization and normalization (VSN) [125]. Third, technical batches were adjusted using ComBat [102] with fixed effects of amplification round. To further reduce batch effects baseline (6pm)and dexamethasone stimulated (9pm) RNA samples for each individual were processed within a single run. Finally for each probe, we constructed a linear model of the log fold change in gene expression between 6pm and 9pm standardized to 6pm controlling for age, disease status and BMI. Models were implemented in R using the "lm" function. The residuals (=GR-response residuals) from this regression were used as phenotype values in the following analyses. Including dexamethasone serum levels 3 hours following administration or the RIN factor, as additional covariates did not change the results. To control if significant eQTLs might be biased due to SNPs within the probes, the Illumina re-annotation pipeline [7] (ReMOAT version August 2009) was used to annotate SNPs (relying on UCSC dbSNP 126 table) that were located within the gene expression probe sequence. No bias of eQTL misclassifications due to such sequence polymorphisms in the probe region could be identified. For the GR-response eQTL analysis only transcripts that showed a difference in gene expression between the samplings at 6pm and 9pm with an absolute fold change ≥ 1.3 in at least 20% of all samples were categorized as robustly effected by dexamethasone stimulation ($n = 4,630$ transcripts) and further used in the analysis. The probe gene names were updated using the NCBI build 36 (hg18) Reference Sequence (RefSeq) [178] gene annotation table obtained from the UCSC table browser[3]. The positions of the probes were annotated using ReMOAT and only autosomal probes were used for the GR-response eQTL analysis ($n = 4,447$ autosomal probes).

2.2.1.2. Differential gene expression analysis

A subset ($n = 60$ individuals, i.e 120 RNA samples) from the raw microarray data of the entire MPIP cohort was extracted and first filtered by a Illumina detection P value of 0.01 in at least five individuals and secondly, transformed and normalized through VSN [125]. From a total of 48,750 transcripts 15,573 remained in the analysis having significant expression according to these criteria.

2.2.2. Mouse models

DEX-mouse RNA was extracted from whole blood using the PAXgene blood miRNA kit (PreAnalytiX) according to Krawiec et al. [111] and mouse brain regions using RNeasy Plus Universal Mini Kit (Qiagen) with standardized protocols. RNA was quality checked on the Agilent 2100 Bioanalyser, amplified using the Illumina Total Prep 96-Amplification kit (Life Technology) and then hybridized on Illumina MouseRef-8 v2.0 BeadChips. For each tissue

[3] http://hgdownload.soe.ucsc.edu/goldenPath/hg18/database/refGene.txt.gz

the samples were processed together (RNA amplification, hybridization and scanning). All samples had a mean RIN of 7.5 ± 0.2 SD for blood cells and a mean RIN of 9.2 ± 0.4 SD for brain tissues. All 25,697 probes present on the microarray were first filtered by an Illumina detection P value of 0.05 in at least 15% of the samples, leaving for blood 10,667, HC 16,838, PFC 16,576 and AM 14,890 expressed array probes for further analysis. Secondly, each transcript was transformed, normalized and batch corrected, in the same fashion as for the human gene expression data. An additional analysis was performed taking HC and PFC samples together. Therefore the quality control was repeated on the joint data set, leaving 16,536 expressed array probes. For differential gene expression analysis between the vehicle and dexamethasone treated animals linear regression models implemented in R were used on the normalized, transformed and batch corrected expression values for each tissue. For the combined analysis of PFC and HC repeated measures ANOVA was performed.

Full details of the RNA quality checks as well hybridization procedures from the chronic social stress sample have been previously described in [190], including details for the microarray gene expression analysis.

2.3. Genotype data

2.3.1. MPIP cohort

Human DNA of the MPIP cohort samples was isolated from EDTA blood samples using the Gentra Puregene Blood Kit (Qiagen) with standardized protocols. Genome-wide SNP genotyping was performed using Illumina Human610-Quad and Illumina Human660W-Quad Genotyping BeadChips according to the manufacturer's standard protocols. 582,539 genetic markers in 163 individuals could be successfully genotyped. Individuals with low genotyping rate ($< 98\%$) and SNPs showing significant deviation from HWE (P value $< 1 \times 10^{-5}$) were excluded. Similarly, low MAF ($< 10\%$) and SNPs with high rates of missing data ($> 2\%$) were excluded, resulting in 436,643 SNPs and 160 individuals for further analysis. In the 160 samples that passed the quality control imputation of additional variants was performed using IMPUTE v2 [94] on the basis of HapMap Utah Residents (CEPH) with Northern and Western European ancestry (CEU) Phase 3 [95] and 1,000 Genomes Project version June 2010 (hg18) CEU data for $\sim$ 8 million SNPs [54]. Imputed SNPs were excluded if their posterior probability averages were less than 90% for the most likely imputed genotype. SNPs were also excluded if their call rate was less than 98%, HWE P value was less than 1×10^{-5} and MAF $< 10\%$. This yielded a total of 2,011,895 SNPs. To annotate SNPs for the closest gene, we used Annovar version November 2011 [230] with the RefSeq gene annotation [178]. SNP coordinates are given according to hg18.

2.3.2. MARS cohort

Human DNA of the MARS cohort samples was extracted from EDTA blood samples using the Gentra Puregene Blood Kit (Qiagen) with standardized protocols. Whole-genome SNP genotyping for the MARS cohort was performed on Illumina Sentix Human-1, HumanHap300, Illumina Human610-Quad and HumanOmniExpress Genotyping BeadChips according to the manufacturer's standard protocols. Individuals as well as the genotype data have been subjected to the same quality control steps as the MPIP cohort (genotyping rate $< 98\%$,MAF $< 10\%$, HWE P value $< 1 \times 10^{-5}$, SNP missingness $< 98\%$). Missing genotype data were imputed via IMPUTE v2 [94] based the 1,000 Genomes Project version Nov. 2010 (hg19) ALL reference panel [54]. The GR/MDD eSNP profile was derived from loci associated with both GR-induced differences in gene expression and MDD, and included alleles from 20 of the 23 tagging eSNPs (3 SNPs diverged from HWE in the MARS sample, see table 3.4). Non-risk and risk alleles (according to association with depression in the PGC dataset) were coded 0 and 1, respectively, and summed in an additive fashion to create cumulative genetic risk profile scores (GRPS; 0,1,2). The MARS GRPSs ranged from 12-30. This genetic profile is highly correlated with a profile weighted based on the strength of association in the PGC dataset ($R = 0.99$); hence we report only the additive profile here for simplicity.

2.3.3. DNS cohort

Human DNA from participants of the DNS cohort was isolated from saliva derived from Oragene DNA self-collection kits (DNA Genotek) customized for 23andMe[4]. DNA extraction and genotyping were performed by the National Genetics Institute (NGI), a CLIA-certified clinical laboratory and subsidiary of Laboratory Corporation of America. The Illumina HumanOmniExpress BeadChips and a custom array containing an additional $\sim 300,000$ SNPs were used to provide genome-wide data. Due to differences in genotyping array content the DNS GRPSs included alleles from 19 of the 23 eSNPs (see table 3.4) and were coded in the same way as the MARS GRPSs. All SNPs used for the GRPSs had genotyping rates $< 97\%$, MAF $< 10\%$, HWE P value $< 1 \times 10^{-5}$ (see table 3.4). DNS GRPSs ranged from 10-28 and were normally distributed (see figure 3.10). To account for differences in ancestral background in the full sample, we used EIGENSTRAT (v, 5.0.1) [176] to generate principal components and included the first 5 components as covariates in the analysis. Five participants were outliers on these components (± 6 SD from the mean on one of the top ten components) and were hence excluded from analyses.

[4] www.23andme.com

2.4. DNS neuroimaging protocol[5]

2.4.1. BOLD fMRI paradigm

A widely used and reliable challenge paradigm was employed to elicit amygdala reactivity. The paradigm consists of 4 task block requiring face-matching interleaved with 5 control blocks requiring shape-matching (see figure 2.1). In each face-matching trial within a block, participants view a trio of faces derived from a standard set of facial affect pictures (expressing angry, fearful, surprised, or neutral emotions), and select which of 2 faces presented on the bottom row of the display matches the target stimulus presented on the top row. Each emotion-specific block (e.g., fearful facial expressions only) consists of 6 individual trials, balanced for gender of the face. Block order is pseudo-randomized across participants. Each of the 6 face trios is presented for 4 seconds with a variable inter-stimulus interval of 2-6 seconds; total block length is 48 seconds. In the shape-matching control blocks, participants view a trio of geometric shapes (i.e., circles, horizontal and vertical ellipses) and select which of 2 shapes displayed on the bottom matches the target shape presented on top. Each control block consists of 6 different shape trios presented for 4 seconds with a fixed inter-stimulus interval of 2 seconds, comprising a total block length of 36 seconds. The total paradigm is 390 seconds in duration. Reaction times and accuracy are recorded through an MR-compatible button box.

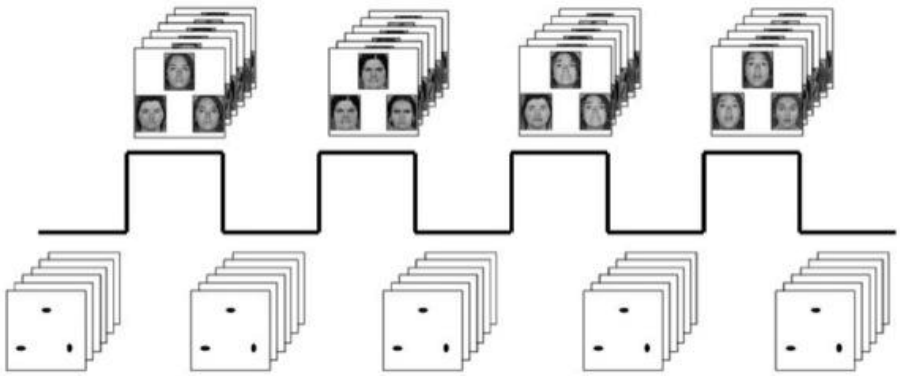


Figure 2.1.: DNS fMRI task. Participants completed four expression-specific (Neutral, Angry, Fear, Surprise) face-matching task blocks interleaved with five sensorimotor shape-matching control blocks. Order for task blocks was counterbalanced across participants.

[5]Parts of the DNS neuroimaging protocol have been published previously by our collaborators [175, 16]

2.4.2. BOLD fMRI acquisition

Participants were scanned using a research-dedicated GE MR750 3T scanner equipped with high-power high-duty-cycle 50-mT/m gradients at 200 T/m/s slew rate, and an eight-channel head coil for parallel imaging at high bandwidth up to 1MHz at the Duke-UNC Brain Imaging and Analysis Center. A semi-automated high-order shimming program was used to ensure global field homogeneity. A series of 34 interleaved axial functional slices aligned with the anterior commissure-posterior commissure (AC-PC) plane were acquired for full-brain coverage using an inverse-spiral pulse sequence to reduce susceptibility artifact (TR/TE/flip angle = 2000 ms / 30 ms / 60; FOV = 240 mm; 3.75 × 3.75 × 4 mm voxels (selected to provide whole brain coverage while maintaining adequate signal-to-noise and optimizing acquisition times); interslice skip = 0). Four initial RF excitations were performed (and discarded) to achieve steady-state equilibrium. To allow for spatial registration of each participant's data to a standard coordinate system, high-resolution three-dimensional structural images were acquired in 34 axial slices co-planar with the functional scans (TR/TE/flip angle = 7.7 s / 3.0 ms / 12; voxel size = 0.9 × 0.9 × 4 mm; FOV = 240 mm; interslice skip = 0).

2.4.3. BOLD fMRI data analysis

The general linear model of Statistical Parametric Mapping 8 (SPM8)[6] was used for whole-brain image analysis. Individual subject data were first realigned to the first volume in the time series to correct for head motion before being spatially normalized into the standard stereotactic space of the Montreal Neurological Institute (MNI) template using a 12-parameter affine model. Next, data were smoothed to minimize noise and residual differences in individual anatomy with a 6mm FWHM Gaussian filter. Voxel-wise signal intensities were ratio normalized to the whole-brain global mean. Then the ART[7] was used to generate regressors accounting for images due to large motion (i.e., > 0.6 mm relative to the previous time frame) or spikes (i.e., global mean intensity 2.5 standard deviations from the entire time series). Participants for whom more than 5% of acquisition volumes were flagged by ART ($n = 21$) were removed from analyses. An region of interest (ROI) mask (Automated Anatomical Labeling (AAL) atlas) from WFU pickatlas [134] was used to ensure adequate amygdala coverage for the face-matching and number-guessing tasks, respectively. Participants who had less than 90% coverage of the amygdala ($n = 14$) were excluded from analyses.

Following preprocessing steps outlined above, linear contrasts employing canonical hemodynamic response functions were used to estimate task-specific (i.e., "Angry & Fearful Faces > Neutral Faces", "Angry & Fearful > Shapes", "Neutral > Shapes") BOLD responses for each individual. The primary contrast of "Angry & Fearful > Neutral" was used to assay centromedial reactivity to cues that are conditioned social signals to threat in the envi-

[6] http://www.fil.ion.ucl.ac.uk/spm

[7] https://www.nitrc.org/docman/view.php/104/390/Artifact%20Detection%20Toolbo/%20Manual

ronment (i.e., angry and fearful expressions) relative to signals that do not convey threat information about the environment (i.e., neutral expressions). Post-hoc analyses using the "Angry & Fearful > Shapes" and "Neutral > Shapes" contrasts were used to discern if the association with GRPS reflected relatively decreased reactivity to angry and fearful expressions or increased reactivity to neutral expressions. Individual contrast images (i.e., weighted sum of the beta images) were used in second-level random effects models accounting for scan-to-scan and participant-to-participant variability to determine mean contrast-specific responses using one-sample t-tests. A voxel-level statistical threshold of P value < 0.05, family wise error corrected for multiple comparisons across the bilateral centromedial amygdala ROIs, and a cluster-level extent threshold of 10 contiguous voxels was applied to these analyses. The bilateral centromedial amygdala ROIs were defined using anatomical probability maps [4]. The centromedial ROI was chosen because it includes the central nucleus of the amygdala (CeA), which specifically functions to drive physiologic, attentive, and neuromodulatory responses to threat, as opposed to the basolateral complex of the amygdala (BLA), which primarily functions to relay information to the CeA; thus the expression of stress responsive behavior is more closely linked with the activity of the CeA and not the BLA [40, 116]. Human research using such distinctions has shown that ROIs encompassing the CeA or BLA differentially respond to stimuli and share different patterns of functional as well as structural connectivity [23, 59, 120].

BOLD parameter estimates from a cluster within the left centromedial amygdala ROI exhibiting a main effect for the "Angry & Fearful > Neutral" contrast were extracted using the VOI tool in SPM8 and exported for regression analyses in SPSS (v.18). No significant cluster emerged in the right centromedial amygdala. Extracting parameter estimates from clusters activated by our fMRI paradigm, rather than those specifically correlated with our independent variables of interest, precludes the possibility of any correlation coefficient inflation that may result when an explanatory covariate is used to select a region of interest. Our collaborators have successfully used this strategy in prior studies [16].

2.5. Statistical Analysis

2.5.1. Differential gene expression analysis

For differential expression analysis linear regression models implemented in R were used on the normalized log_2-transformed expression values from a subset of MPIP cohort ($n = 60$ individuals) at baseline, after GR-stimulation as well as their difference comparing cases and controls

Classifications were performed with the RF classification technique [124]. The sample was randomly subsetted into training (sample 1: 18 cases and 18 controls; see section 2.1 and table 2.1) and test set (sample 2: 11 cases and 13 controls; see section 2.1 and table 2.1). The RF provides the mean decrease Gini as an importance measure that calculates the quality of a split for each node of a tree by means of the Gini index. Each time a node is split on a variable, the Gini index for both descendent nodes is less than the ancestor

node. A higher mean decrease Gini represents a higher variable importance [20].

2.5.2. eQTL analysis

The eQTL analysis (MPIP cohort, $n = 164$) was restricted to those SNP-probe pairs that map within a region of 1Mb upstream or downstream of the gene expression probe, in order to detect *cis*-eQTLs. To measure the transcriptional response we used the log fold change in gene expression changes between 6pm (baseline) and 9pm (GR-stimulation) standardized to 6pm.

PLINK v1.07 [179] was used to test for *cis*-association between all imputed SNPs and transcriptional response. As eQTL data were composed of two kinds of data: genotyping and expression data, we used a two stage of multiple testing correction: (i) SNP level correction: for each *cis*-region (array probe) we preformed a permutation test. The sample identifiers in the gene expression data were shuffled in order to preserve the structure in the genotype data (LD). A total of 500,000 permutations were carried out per probe, i.e. maxT procedure of Westfall-Young [233]. (ii) Probe level correction: *Cis*-regions with an extensive LD structure will increase the number of false positive eQTLs [234]. Therefore we applied the Benjamini-Hochberg method to correct the maxT adjusted P value significance by using only the most significant and independent SNPs per probe (tag SNPs). The number of tag eSNPs per *cis*-region was identified by LD pruning and "clumping" the SNPs using the "clump" command in PLINK (using distance < 1 Mb and $r^2 \leq 0.2$ as setting). Each tag SNP forms a SNP bin, by aggregating all other SNPs into bins by tag SNP at $r^2 \leq 0.2$ and distance < 1Mb, such that all SNPs within a given bin were correlated to the tag SNP, but to any other tag SNP. We aim to limit the false-positive SNP-probe pairs to less than 5% and therefore we considered the FDR analogue of the P value (Q value) $< 5\%$ as statistically significant.

Validation of *cis*-eQTL results was carried out with a sample size-weighted Z-score meta-analysis [60] in an additional independent data set using peripheral blood samples (baseline and after GR-stimulation with 1.5 mg dexamethasone) of 58 individuals (21 male controls, 14 male cases and 23 female cases). We applied the same strategy as used in the discovery sample (MPIP cohort) to filter, normalize and batch correct the gene expression data. We adjusted the analysis for the same covariates plus gender; applied the same SNP quality control checks and performed the *cis*-eQTL mapping in PLINK. A *cis*-eQTL is validated if the meta-analysis P value is less than the actual maxT adjusted P value in the discovery sample.

The Genomic control inflation factor (λ_{gc}) [46] was calculated for every GR-response *cis*-eQTL gene expression probe ($n = 297$) based on (i) the genome-wide genotype data (λ_{gc}) and (ii) the genotype data within each of the *cis*-windows (λ_{cis}). Both inflation factors were computed in PLINK as median χ^2 statistic. The median λ_{gc} over all probes is 1 and the median λ_{cis} is 1.0099, which imply that no large inflation was present.

The combined set of the first two replicates of the RNA Polymerase II ChIA-PET data [121, 122] generated from K562 chronic myeloid leukemia cell lines ($n > 400,000$ interaction regions; other cell lines $n < 70,000$ interaction regions) were obtained from the UCSC genome browser[8]. Genomic coordinates of our GR-response eSNP bins were converted from hg18 to GRCh build 37 (hg19) using the UCSC genome browser liftOver tool[9] and the probe gene coordinates were updated with the hg19 RefSeq [178] gene table obtained from the UCSC table browser[10] (excluding 15 probe genes on hg19). To estimate the overlap of the direct chromatin interactions and GR-response eQTL bins (eSNP bin-probe gene combination) we tested if one ChIA-PET tag overlapped with the region of the eSNP bin ±10kb as well as the relevant array probe gene (10kb ± transcription start or end). To establish the null distribution, we permuted the distances between the GR-response eSNP bins and the transcription sites of the corresponding probe genes ($n = 270$ updated to hg19) and estimated the overlap with ChIA-PET interaction signals. We repeated the analysis 1,000 times and for each set we counted the number of genes with overlapping ChIA-PET data. Enrichment calculations with a permutation-based $FDR < 10\%$ were considered as statistically significant within the entire manuscript.

Transcription factor binding affinities for a set of SNPs were estimated using the web tool of the physical affinity-based method TRAP [183, 135]. We extracted a region of ±20bp around all *cis*-eSNPs from the human genome (hg18) as provided by UCSC (Bioconductor Annotation package[11]). For each binding matrix ($n = 904$) TRAP first calculates the affinity of the matrix for every SNP sequence (length of 41bp) and then transforms these affinities into P values. The P values for the SNPs are combined using Fisher's method to determine whether the affinities to specific TF are significantly higher than the ones from 2,000 random promoter sequences (matched for sequence length) after correction for multiple testing using Benjamini Hochberg method [215]. This analysis was applied to the reference and alternative allele, separately. To exclude possible effects due to different GC content between our eSNP and human promoter sequences, we generated random control data sets with RSA tools[12] matching (i) GC content = 42% (equal to eSNP sequences) and (ii) GC content = 50% (equal to random promoter sequences). We determined the TRAP affinity P values for both controls sets and observed similar P value distributions (Wilcoxon $P_{(i):(ii)} = 0.0002$) showing a pronounced difference to the GR-response P value distribution (Wilcoxon $P_{eQTL:(i)} = 5.57 \times 10^{-52}$ and Wilcoxon $P_{eQTL:(ii)} = 3.68 \times 10^{-50}$; no differences among the alleles). Therefore, we exclude confounding due to different GC contents between human promoter and GR-response sequences.

In order to calculate whether GR-response eSNPs would show allele-specific binding profiles (reference vs. alternate allele) we used the GR position-specific weight matrix present

[8] http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeGisChiaPet/
[9] http://genome.ucsc.edu/cgi-bin/hgLiftOver
[10] http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/refGene.txt.gz
[11] www.bioconductor.org/packages/2.12/data/annotation/BSgenome.Hsapiens.UCSC.hg18.html
[12] rsat.ulb.ac.be/rssa/

in Genomatix (Matrix Family Library version 8.4[13], vertebrates: 867 matrices in 182 families) and calculated the binding affinities separately for the two SNP alleles using the R package of TRAP[14]. The significance of the differences in the binding affinity between the opposite alleles was archived using 1,000 sets of random SNPs (sized matched) and determining the expected allele-bases differences.

To identify whether GR-response eSNPs were enriched for association with MDD we conducted a meta-analysis based on the Psychiatric Genomics Consortium (PGC) GWAS mega-analysis for MDD [133]. We used the "meta-analysis" function in PLINK assuming a fixed effect model in 17,846 individuals of European ancestry (8,864 cases with MDD and 8,982 controls) from 8 of the 9 studies included in the PGC data and excluded all samples from the initial PGC data (n = 18,759) that overlap with our MARS cohort (n = 376 cases and 537 controls), which we used as validation sample. The PGC used SNP data imputed to the 1,000 Genomes Project version June 2011. For comparability we converted all our SNP coordinates from hg18 to hg19 using the UCSC genome browser liftOver tool and created the overlap of all MDD SNPs and our imputed SNPs of high quality. To match the MAF distributions of the random SNP sets with our GR-response eQTL data we divided the SNPs into non-overlapping MAF bins, each of the width 0.05 as described previously [159]. We generated 1,000 sets (conditional on MAF and number of GR-response eSNPs overlapping with MDD associations; n = 3,492) of randomly drawn SNPs (without replacement) from the set of all imputed SNPs. For every set we counted the percentage of unique SNPs with a MDD meta-analysis P value ≤ 0.05. On this basis we constructed the null distribution and compared it to the observed percentage of eSNPs with a MDD meta-analysis P value ≤ 0.05 to measure the enrichment statistics. Additionally, we compared baseline eSNPs (see supplementary notes A.1.1) to GR-response eSNPs. Due to the different size of the overlap of GR-response and baseline eSNPs with the MDD meta-analysis data (GR-response eSNP overlap n = 3,492 vs. baseline eSNP overlap n = 28,861; 31,541 original number of eSNPs) we used the same strategy as described above for random SNPs. To further account for the genomic LD structure, we limited the analyses to tag SNPs (tag SNP = SNP showing the highest association per *cis*-eQTL bin) and generated 1,000 randomized SNP sets; conditional on MAF and each of the same size as the GR-response tag SNPs overlap with MDD associations (n = 285).

To identify whether GR-response eSNPs were enriched within risk loci associated with other psychiatric disorders we used the results of the PGC cross-disorder (CD) analysis (33,332 patients and 27,888 controls of European ancestry distributed among five disorders: SCZ, BPD, attention deficit-hyperactivity disorder (ADHD), autism spectrum disorder (ASD) and MDD) [37]. The PGC CD analysis applied a multinomial regression procedure and used SNP data imputed to the HapMap Phase 3 data (hg18). The data sets from the PGC CD associations and the sub-analysis results for BPD, SCZ, ASD and

[13] http://www.genomatix.de
[14] http://trap.molgen.mpg.de/cgi-bin/download.cgi

ADHD used within the CD associations were obtained from the PGC website[15]. We used the same enrichment analysis as described for MDD to assess whether SNPs with nominal CD associations ($P \leq 0.05$) and if SNPs with nominal associations among the four individual diseases (SCZ, BPD, ASD and ADHD) were enriched for GR-response eSNPs. Briefly, for the individual diseases and the CD associations, 1,000 randomized SNP sets were generated, each conditional on MAF and of the same size as the GR-response eSNPs overlap with CD data sets (n = 1,047). We summarized the enrichment results for these simulations in table 3.5.

To proof the significance of the GR/MDD tag SNPs for MDD we used a logistic regression model in R with the function "glm" to test the association of the GRPSs for disease status in the MARS cohort. Gender and age were used as covariates. To establish the null distribution we generated 1,000 random SNP profiles by swapping individual labels to provide new SNP profiles under the null hypothesis.

The interaction network (figure 3.11) was built by manual curation and literature mining, using the CIDeR database [115] and the yED software (yWorks GmbH, Tübingen, Germany). Gene products identified in the GR-response eQTL analysis were analyzed for physical interactions, regulatory interactions and association with psychiatric disorders.

We used the GR-response residuals from all array probes (n = 4447) to determine if the 25 GR/MDD array probes are more tightly co-regulated than 1,000 sets of randomly chosen transcripts. To realize this, we carried out a co-expression analysis in R using the function "dist" specifying the Euclidian distance as distance measure and calculated the mean distance of all pair-wise distances. We established the significance of co-expression network of the 25 GR/MDD array probes by testing the observed mean distance versus the null distributions created by calculating the mean distance of all pair-wise distances for 1,000 sets of 25 randomly chosen GR-response transcripts. Next, we determined the number of sets, having lower mean distances than the actual GR/MDD-relevant transcripts to measure the enrichment statistic.

Statistical analyses of the imaging data were completed using linear regression in SPSS to test the association of the DNS GRPSs to amygdala reactivity. To maintain variability but constrain the influence of extreme outliers, prior to analyses all imaging variables were winsorized (i.e., following data quality control procedures, outliers more than ± 3 SD were set at ±3 SD from the mean; for the "Angry and Fearful > Neutral faces" contrast, 13 outliers (2.01%) of the entire sample were moved to ±3 SD from the mean). Gender, psychiatric diagnosis (0,1) and age were entered as covariates for both EUR-AM and entire sample analyses. Five ancestrally-informative principal components that distinguish the sample were added as additional covariates in the analyses of the entire sample. We computed permutations (n = 1,000) in which we constructed randomly generated SNP

[15] http://pgc.unc.edu

profiles that were matched for MAF, amount of SNPs ($n = 19$) and constrained by the max LD observed within the sample.

Graphs were generated with Haploview [8], ggplot2 [235] and Circos[112].

2.6. Quantitative real-time PCR validation

Total RNA was reverse-transcribed to cDNA using random primers and the Superscript II reverse transcriptase (Invitrogen) for quantitative real-time PCR (qPCR) to validate microarray results. qPCR was carried out according to manufactures instructions using Roche-LightCycler 480 System (Roche Applied Science) and assays were designed using the Roche Universal Probe Library[16].

2.6.1. Differential gene expression analysis

Microarray results were validated separately in sample 1 (training set) and 2 (test set). We selected *FKBP5* and *BEST1* as genes showing both regulation with dexamethasone as well as differences between cases and controls (in the microarray data), and *TBP* as the endogenous control gene. For the both target genes both the regulation with dexamethasone as well as the differences between cases and controls could be validated (see figure 3.2). Sequences of primers used are summarized in table 2.3. All samples were run in duplicates and duplicates discordant in crossing points by more than 0.4 cycles, were excluded from the analysis.

Target gene	Primer set (5'-3')	UPL probe nr.
FKBP5	Forward: ccattgctttattggcctct	15
	Reverse: ggatatacgccaacatgttcaa	
BEST1	Forward: ttgattcaggctgttgtaggac	76
	Reverse: ctaggaagcggccaatgat	
TBP	Forward: ctttgcagtgacccagcat	67
	Reverse: ccagcaggacactgatcca	

Table 2.3.: List of primers and Universal probe library number used for the qPCR validation of differential expressed genes within a subset of the MPIP cohort ($n = 60$).

[16] http://qpcr.probefinder.com

2.6.2. eQTL analysis

Microarray results were validated for *ADORA3*- the probe with a significant GR-response eQTLs), *HIST2H2AA3/ HIST2H2AA4*- the probe with the most eSNPs overlapping with the meta-analysis results for MDD) and *TBP* as the endogenous control gene. The association between eSNPs and GR-stimulated gene expression of the two target genes could be validated using qPCR (see figure 3.4a,(b) and supplementary notes A.1.2). Sequences of primers used are summarized in table 2.4. All samples were run in duplicates and duplicates discordant in CT values by more than 0.2 cycles, were excluded from the analysis. Relative gene transcript levels were determined by Pfafflʼs equation [172] with:

$$ratio = \frac{(E_{gene})^{\triangle CT_{gene}(\text{baseline sample-GR-stimulated sample})}}{(E_{TBP})^{\triangle CT_{TBP}(\text{baseline sample-GR-stimulated sample})}} \tag{2.1}$$

These ratios were regressed against the SNPs of interest, while adjusting for age, BMI and disease status in R. qPCR ratios shown in figure 3.4b were calculated using the following equations:

$$pre = \frac{(E_{TBP})^{CT_{TBP}(\text{baseline sample})}}{(E_{gene})^{CT_{gene}(\text{baseline sample})}} \tag{2.2}$$

$$post = \frac{(E_{TBP})^{CT_{TBP}(\text{GR-stimulated sample})}}{(E_{gene})^{CT_{gene}(\text{GR-stimulated sample})}} \tag{2.3}$$

Target gene	Primer set (5'-3')	UPL probe nr.
ADORA3	Forward: tcatttgcagccaggtagc Reverse: tgcttgggtgtggtctatca	82
HIST2H2AA3/HIST2H2AA4 (short isoform)	Forward: cgacgaggaactgaacaagc Reverse: gcctggatgttaggcaagac	61
HIST2H2AA3/HIST2H2AA4 (long isoform)	Forward: aaggggcacctgtgaactc Reverse: gactgagagtggccagcatt	21
TBP	Forward: ctttgcagtgacccagcat Reverse: cgctggaactcgtctcacta	67

Table 2.4.: List of primers and Universal probe library number used for qPCR validation of GR-response eQTLs (MPIP cohort, $n = 160$).

3. Results

3.1. Genome-wide gene expression profiles following glucocorticoid stimulation in healthy volunteers and MDD patients[1]

To test whether reliable case-control differences can be identified following GR activation, we compared gene expression profiles between MDD cases and controls before and 3h after stimulation with dexamethasone in two independent samples (see section 2.1).

3.1.1. Dexamethasone effect on gene expression levels of MDD cases and healthy controls

In controls dexamethasone led to the overlapping regulation of 2,670 transcripts in both samples at P value ≤ 0.05 of which 42% ($n = 1,132$ array probes) were up-regulated and 58% ($n = 1,538$ array probes) down-regulated (see supplementary table 2 in Menke/Arloth et al. [148] for more details). In depressed patients only 1,151 transcripts were significantly regulated in both samples with 44% up-regulated transcripts ($n = 507$ array probes; see supplementary table 3 in Menke/Arloth et al. [148] for more details). In total 23% of the significant dexamethasone-induced gene expression changes observed in patients were also observed in controls.

3.1.2. Using gene expression profiles to classify MDD cases and controls

Classification was performed with the RF algorithm using sample 1 to train the model and sample 2 for testing (see subsection 2.1).

3.1.2.1. Baseline gene expression

We first performed a feature selection by including those transcripts that showed a significant difference in baseline gene expression between MDD cases and controls in the training set at P value ≤ 0.05 (uncorrected) and of absolute fold change ≥ 1.15. This resulted in 635 array probes that were used for classification with the RF algorithm. The RF was run using the best performing parameters, e.g. 1,000 trees (ntree) and 25 random features to

[1] A version of this chapter has been published in Menke/Arloth et al. [148].

build each tree (mtry)). This revealed an OOB error rate of 27.8% to classify MDD cases and controls with a sensitivity (see Eq. 1.5) of 72.2% and specificity (see Eq. 1.6) of 72.2%. However, using the test set (sample 2), the constructed prediction model only achieved an area under the curve (AUC) value of 0.56 and an accuracy of 41.7% (see Eq. 1.7; 10 out of 24 correctly classified).

3.1.2.2. GR-stimulated gene expression

To select features for classifying cases and controls using GR-stimulated gene expression measures we only kept those transcripts that showed a difference in gene expression change from baseline to GR-stimulation between MDD cases and controls at a nominal P value ≤ 0.05 and an absolute fold change following dexamethasone in controls ≥ 1.15. This resulted in 250 transcripts, which we used for classification with RF. Using the best performing parameters, e.g. $ntree = 1,500$ trees and $mtry = 240$, the RF algorithm revealed an OOB error rate of 16.7% to differentiate between MDD cases and controls with a sensitivity of 80% and specificity of 87.5%. Testing this model in sample 2 resulted in an AUC value of 0.73 and an accuracy of 79.2% (19 out of 24 correctly classified). Thus the predicator created by GR-stimulated gene expression levels performed much better than the predictor built from baseline gene expression (accuracy of 79.2% vs. 41.7%).

3.1.2.3. Reduction of the number of transcripts for classification

To identify the genes that contributed most to the classification in both samples, RF classification was performed without feature selection in all individuals together (not separated by samples). The analysis was repeated 10 times and the importance scores (Gini index) were averaged. Out of the 206 most important features, i.e. features having an average Gini index in the 10 repetitions ≥ 0.02, 19 transcripts also showed a significant differences (nominal P value ≤ 0.05) in dexamethasone regulation between MDD cases and controls in both samples (see table 3.1 and figure 3.1).

3.1.2.4. Validation of differentially regulated transcripts

We used qPCR to validate gene expression differences of two out of the 19 transcripts, e.g. *FKBP5* and*BEST1* in cases and controls in both samples (see figure 3.2). For *FKBP5* we could validate the significant association between the change in gene expression from baseline to GR-stimulation and disease status for both samples ($P_{\text{sample 1}} = 0.005$; $P_{\text{sample 2}} = 0.007$) as well as the main effect of GR-stimulation ($P_{\text{sample 1}} = 0.001$; $P_{\text{sample 2}} = 0.001$). For *BEST1* we could validate the regulation by dexamethasone ($P_{\text{sample 1 \& 2}} = 0.028$) and the effect of the disease status when analyzing both samples together ($P_{\text{sample 1 \& 2}} = 0.08$ two-sided and 0.041 one-sided), but not separately for each sample. However, the direction of the effect was the same in both samples and the observed lack of significance may reflect lack of power in the subsample analysis.

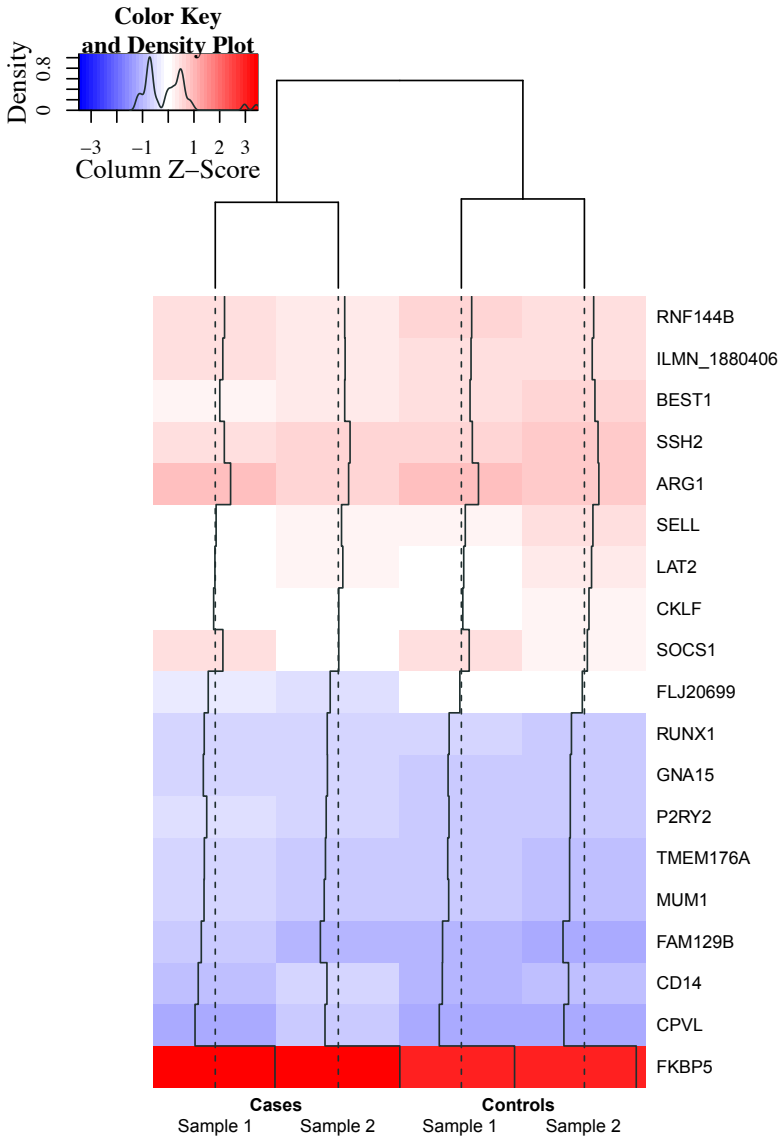


Figure 3.1.: This heatmap illustrates the gene expression changes (GR-stimulation/baseline mRNA levels) of the 19 significantly differentially regulated genes between MDD cases and controls. Red indicates an up-regulation (58% of the transcripts) following dexamethasone and blue a down-regulation (42% of the transcripts).

		Sample 1			Sample 2		
Probe gene	Probe id	FC_{Cases}	$FC_{Controls}$	P	FC_{Cases}	$FC_{Controls}$	P
ARG1	ILMN_1812281	1.70	2.09	0.03	1.56	2.04	0.003
BEST1	ILMN_1718982	1.33	1.62	0.003	1.43	1.80	0.01
CD14	ILMN_1740015	-1.25	-1.44	0.01	-1.06	-1.32	0.037
CKLF	ILMN_1712389	1.15	1.30	0.039	1.24	1.48	0.042
CPVL	ILMN_2400759	-1.34	-1.59	0.04	-1.11	-1.53	0.005
FAM129B	ILMN_1661755	-1.16	-1.43	0.00	-1.24	-1.58	0.005
FKBP5	ILMN_1778444	4.78	6.32	0.03	5.16	6.93	0.031
FLJ20699	ILMN_1692464	1.01	1.19	0.01	1.02	1.19	0.019
GNA15	ILMN_1773963	-1.11	-1.21	0.02	-1.05	-1.25	0.015
LAT2	ILMN_2326953	1.18	1.33	0.01	1.37	1.61	0.031
MUM1	ILMN_1764764	-1.09	-1.19	0.04	-1.13	-1.27	0.007
P2RY2	ILMN_2372915	-1.02	-1.18	0.02	-1.08	-1.25	0.002
RNF144B	ILMN_1752526	1.48	1.70	0.03	1.43	1.72	0.002
RUNX1	ILMN_1801504	-1.08	-1.17	0.03	-1.05	-1.20	0.023
SELL	ILMN_1724422	1.22	1.40	0.02	1.33	1.68	0.024
SOCS1	ILMN_1774733	1.43	1.58	0.04	1.24	1.40	0.017
SSH2	ILMN_1672834	1.47	1.74	0.01	1.62	1.99	0.04
TMEM176A	ILMN_1791511	-1.08	-1.22	0.04	-1.10	-1.27	0.019
no symbol	ILMN_1880406	1.42	1.67	0.05	1.44	1.65	0.047

Table 3.1.: List of the 19 transcripts contributing most to the classification algorithm.

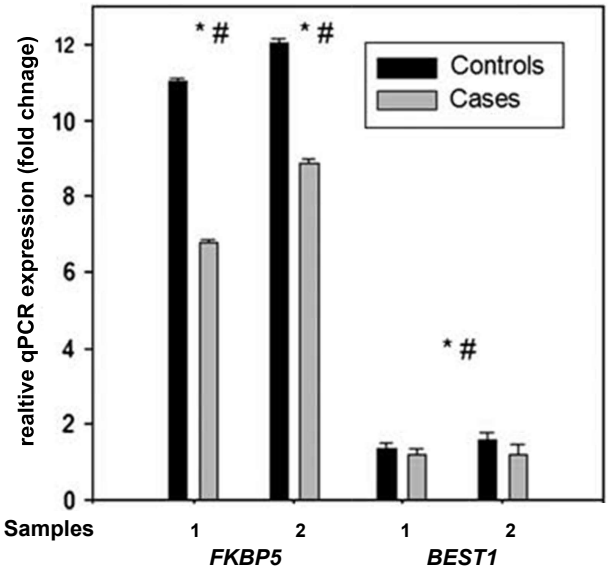
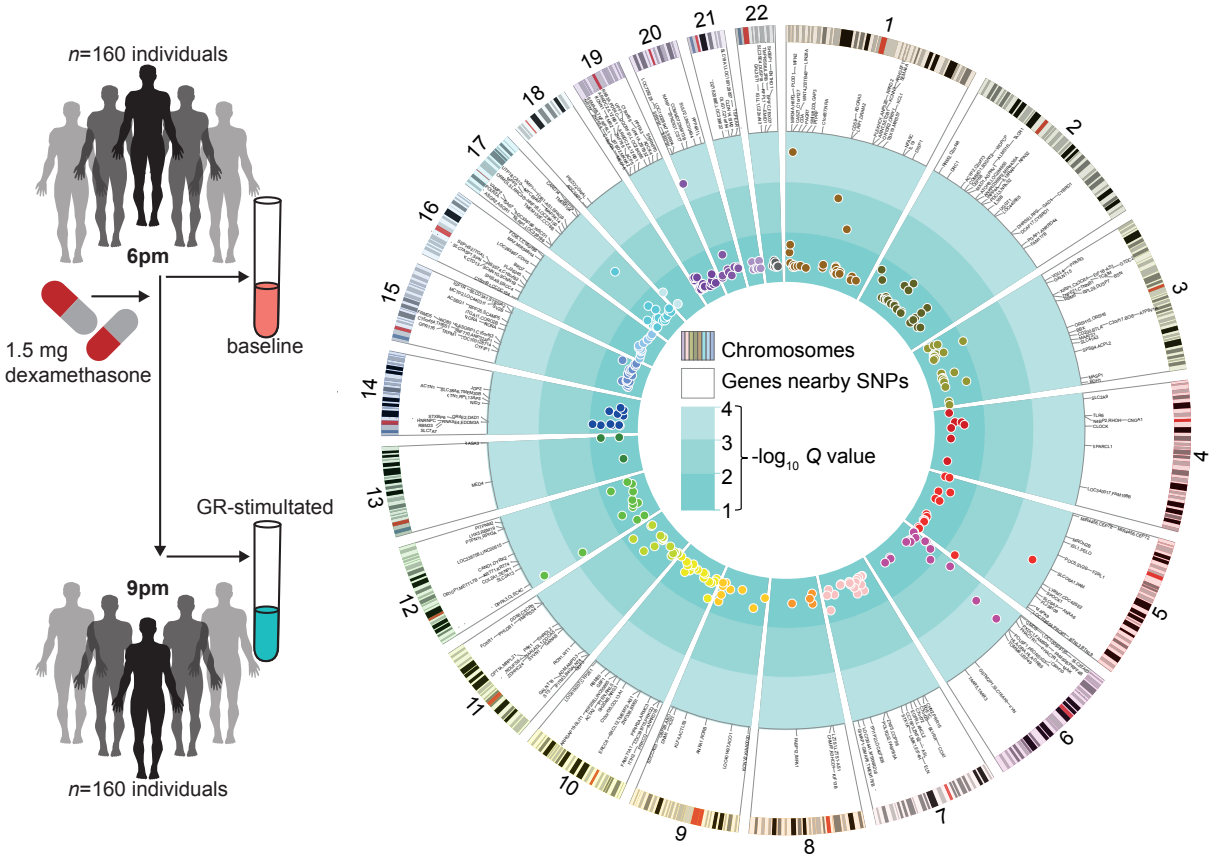


Figure 3.2.: Validation of the differential gene expression of *FKBP5* and *BEST1* and comparison of healthy controls and depressed cases for sample 1 (training set) and sample 2 (test set). A significant associations were observed for *FKBP5* (*) and significant main effects (#) of dexamethasone was observed for *FKBP5* when analyzing both samples separately. For *BEST1* we observed a significant main and association with disease status when analyzing both samples together.

3.2. Genetically determined differences in the immediate transcriptome response to stress predict risk-related brain function and psychiatric disorders

3.2.1. Genetic regulation of GR-stimulated gene expression

We first identified genetic variants that alter GR-stimulated gene expression changes, by adopting a stimulated eQTL approach (see figure 3.3a).

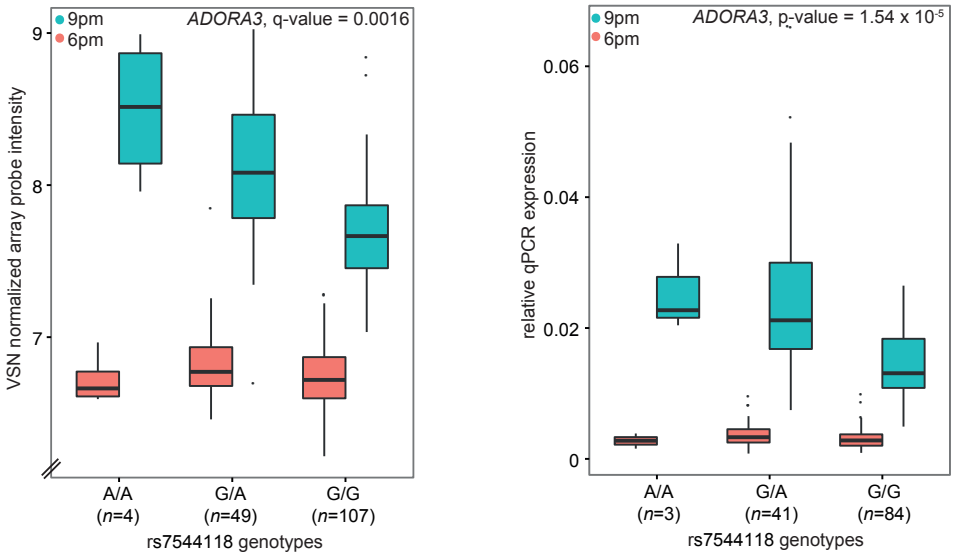


(a) Study design for GR-stimulated gene expression in whole blood of 160 male individuals from the MPIP cohort.

(b) Circularized Manhattan plot displaying *cis*-associations for GR-response set of eSNP bin probe combination (eQTL bin)s (n = 320) and their respective significance ($-log_{10}$ Q values). Displayed from the outer to the inner circle are the number of chromosomes, the ideograms for the human karyotype (hg18), genes nearby eSNPs and Manhattan plots for the eQTL bins that survived correction for multiple testing.

Figure 3.3.: Analysis of GR-response *cis*-eQTLs

Gene expression profiles from peripheral blood cells of 160 male individuals of the MPIP cohort (91 cases and 69 controls, see section 2.1) were obtained at baseline and three hours after stimulation with the selective GR agonist dexamethasone (see supplementary figure A.1) and combined with genome-wide SNP data. All individuals showed a strong endocrine response to dexamethasone (Cortisol: $F_{1,159} = 43.93$, $P = 5.02 \times 10^{-10}$ and ACTH: $F_{1,158} = 37.96$, $P = 5.76 \times 10^{-9}$, see supplementary figure A.2). After quality control, 4,447 gene expression probes that exhibited strong regulation following dexamethasone administration (absolute fold change in gene expression, i.e. baseline to three hours post-dexamethasone, $\geq$ 1.3 in at least 20% of all samples) were combined with genotype data of around 2 million imputed SNPs (see section 2.5.2). We used the log fold change in gene expression standardized to baseline values as phenotype and restricted the analysis to a 1Mb *cis*-region around each array probe. 3,820 GR-response-modulating *cis*-eQTLs (GR-response eQTLs) were identified, which remain significant after accounting for disease status, age, and BMI and correction for multiple testing (see section 2.2). These comprised 297 unique array probes and 3,662 unique SNPs. The 3,662 unique GR-response *cis*-expression SNP (eSNP)s can be summarized in terms of independent tag SNPs (tag SNP=SNP showing the highest association per bin) into 296 uncorrelated *cis*-eSNP bins, i.e. sets of SNPs in LD (see section 2.5.2). These 296 *cis*-eSNP bins correspond to 320 *cis*-eQTL bins, i.e. *cis*-eSNP bin-probe combinations (listed in supplementary table A.1 and see figure 3.3b, 3.4(a)(b) for illustration).



(a) Microarrays data.

(b) qPCR validation.

Figure 3.4.: Boxplots of human gene expression values for *ADORA3*, which is an example of a significant GR-response eQTL. Expression levels are stratified based on the eSNP genotypes for *ADORA3*. Baseline (6pm) measures are displayed in red and GR-stimulated measures (9pm) in blue. Microarray data could be validated by qPCR.

Including dexamethasone serum levels as covariate did not change the results, thus we could exclude any confounding effects of individual genetic and environmental variability of dexamethasone concentration.

To assess the robustness of these GR-response eQTLs, we validated these eQTLs in an independent sample of $n = 58$ (see section 2.5.2). Due to the small sample size of the validation set we applied a sample size-weighted Z-score meta-analysis. 72% of the GR-response eQTLs could be validated, i.e. showing a meta-analysis P value equal or less than the actual significance threshold (see section 2.5.2).

3.2.2. GR-response eSNPs are located in long-range enhancer regions

Next, we mapped the distances between eSNPs and array probes. In order to account for genomic LD structure, the following analyses are based on eSNP bins. We mapped the distance of the set of eSNPs in LD (eSNP bin)s from the 320 GR-response eQTL bins to the probe sequence of the respective regulated transcript and compared this to eSNP bin-probe distances for baseline *cis*-eSNP bins (see supplementary note A.1.1). GR-response eSNP bin-probe distance ($mean = 406$kb $\pm$ 303kb, $n = 320$ bins) was significantly longer (Wilcoxon $P = 1.03 \times 10^{-50}$) than baseline eSNP bin-probe distance ($mean = 149$kb $\pm$ 232kb, $n = 1,148$ bins; see figure 3.5, supplementary note A.1.1). The results indicate that GR stimulation is associated with significantly more long-range transcriptional regulation than baseline gene expression.

To evaluate whether the long-range regulation of GR-response expression quantitative trait loci (eQTLs) may also be associated with long-range physical chromatin interaction, we compared our data to data from a chromatin interaction analysis with paired-end tag sequencing (ChIA-PET [121]) generated by ENCODE [57] in leukemia cells line K562. For this, we examined whether regions containing the GR-response eSNP bin and the corresponding probe gene overlap with physically interacting ChIA-PET tags (see section 2.5.2). Twenty-five percent of the GR-response eSNP bin/probe gene combinations (i.e. 68 out of 270 probe genes) overlapped with chromatin interaction signals. Notably, we observed that our GR-response eSNP bin/probe gene pairs were more likely to colocalize with physically interacting ChIA-PET tag pairs than 1,000 equally sized sets of randomly distributed GR-response eSNP bin/probe gene pairs (enrichment ratio = 1.13, permutation-based FDR = 0.056; see section 2.5.2). Interestingly, restricting the analysis to more long-range eSNP bin/probe gene pairs (distance > 100kb) strengthened the enrichment of such colocalizations (enrichment ratio = 1.57, permutation-based FDR = 0.007).

Figure 3.6 illustrates such an example of long-range regulation accompanied by long-distance chromatin interaction: SNPs in the *CLOCK* locus regulate the GR-stimulated gene expression of the *PAICS* transcript, which is located about 900kb upstream of the *CLOCK* locus.

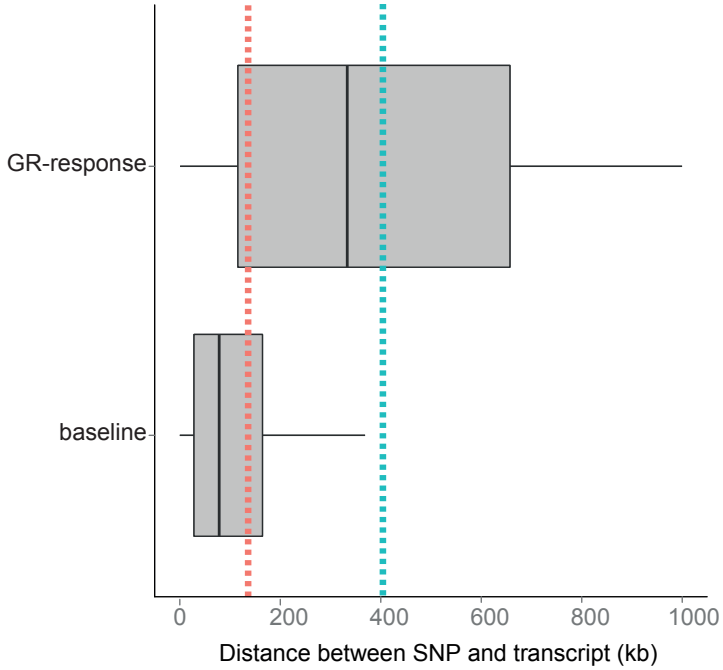


Figure 3.5.: Boxplots show the distance of eSNP to array probe for significant eQTL bins from the GR-response and the baseline analyses. The dotted red lines indicate the mean eSNP bin-probe distances

3.2.3. GR-response eSNPs influence predicted transcription factor binding affinity

To investigate whether specific TF binding motifs are enriched in the set of GR-response eSNPs, we compared these eSNP sequences to random human promoter sequences using TRAP [183] (see section 2.5.2). Consistent with evidence that dexamethasone selectively activates the GR (itself a TF), GR binding sites and TF binding sites that directly modulate GR signaling (such as AP1, CEBP, HNF3, HNF4, OCT1, STAT5A and STAT6 [103, 44]) were significantly over-represented in the GR-response eSNP sequence set (TRAP affinity P values $\leq 4.58 \times 10^{-17}$) as compared to the random sequences (see supplementary table A.2). Using only the unique tagging eSNPs ($n = 296$, representing the highest association per *cis*-eQTL bin) we also observed a significant enrichment of these TF binding sites.

Interestingly, we found not only that TF binding sites were enriched among the GR-response sequences, but also significant GR affinity differences between the opposite alleles ($t_{all} = -3.02$, $P_{all} = 0.0025$; $t_{taggingSNPs} = -2,23$, $P_{taggingSNPs} = 0.027$). It is thus likely that the observed eSNPs modulate GR-stimulated gene expression changes by altering the affinity of the DNA binding sites to the GR and its co-factors. Figure 3.7, provides one specific example of this possible mechanism by illustrating how the predicted changes in affinity to the GR for the sequence containing one intronic SNP (rs2460432) parallel the observed allele-based differences in GR-stimulated gene expression of the gene *ASL*.

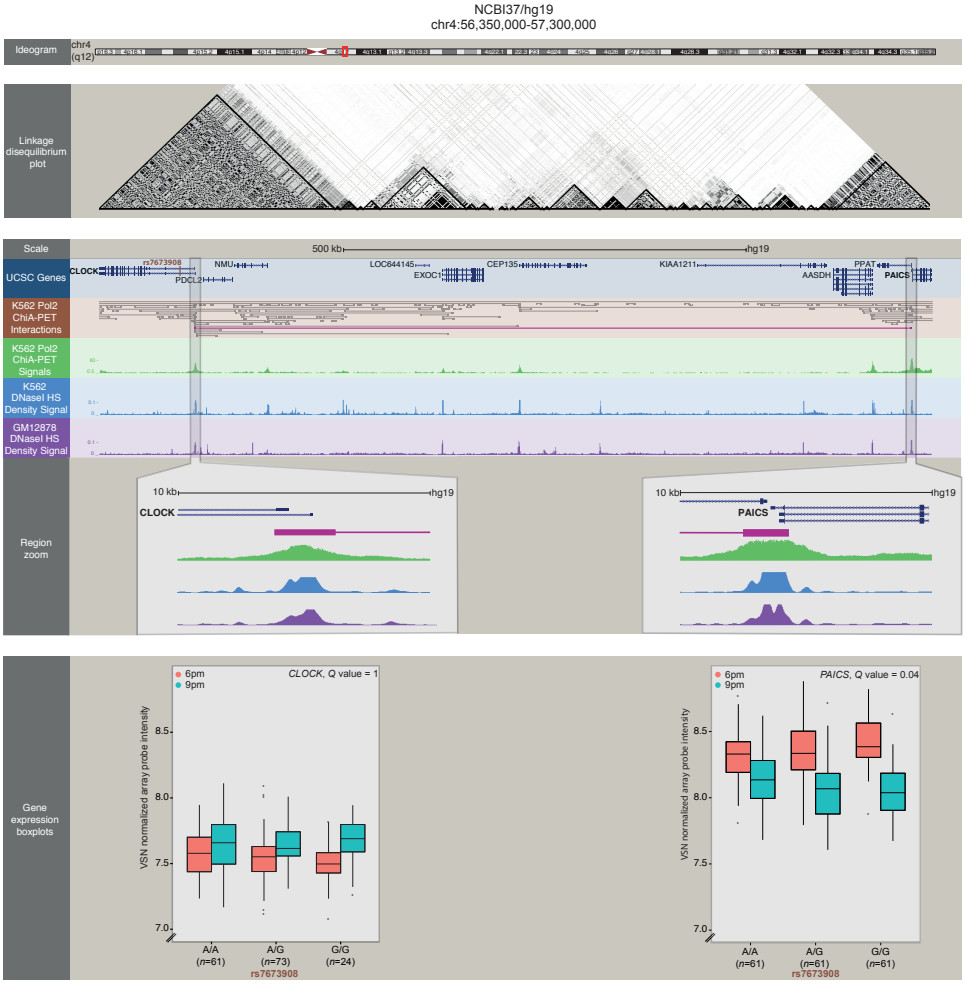
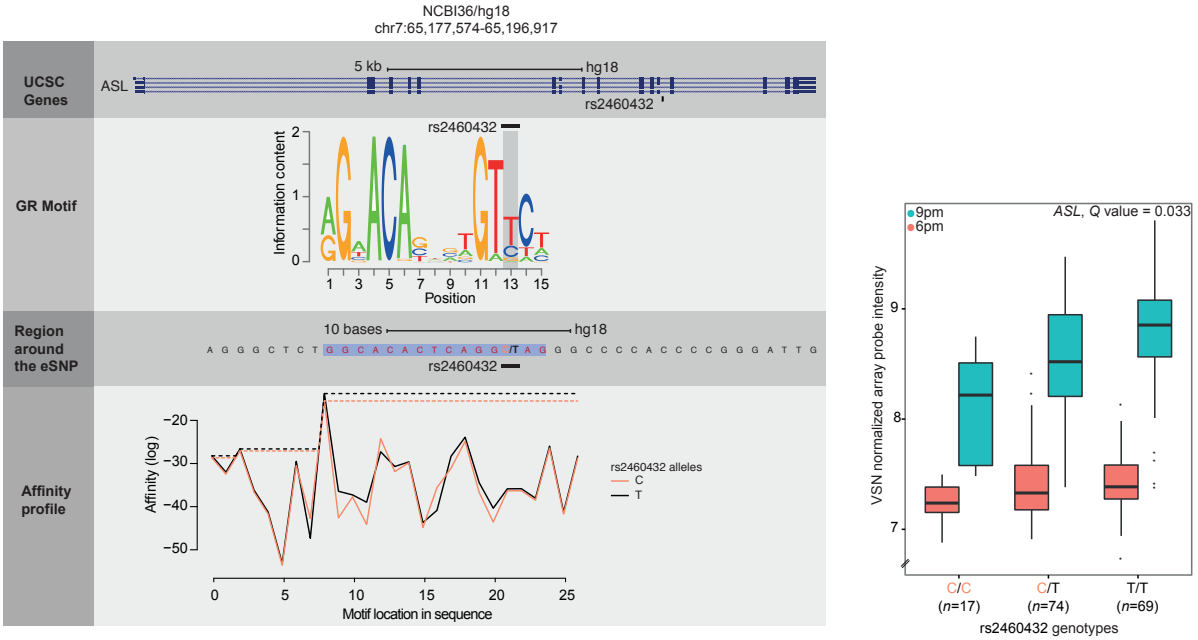


Figure 3.6.: Long-range chromatin interaction of GR-response eQTLs as exemplified by the region containing the *CLOCK* locus (chr4: 56,350,000-57,300,000; hg19).Top panel, ideogram for chromosome 4. A red box isolate the region shown (enlarged) in panel 2 and 3. Second panel, linkage disequilibrium plot based on r^2 for the SNPs surrounding the tagging SNP rs7673908 of the *CLOCK* locus ($r^2 = 0$ is shown in white, $0 < r^2 < 1$ is shown in shades of grey and $r^2 = 1$ is shown in black). Third panel, ChIA-PET tags from leukemia cells (brown and green tracks) validate a chromatin interaction between the *CLOCK* eSNP locus and the promoter of the regulated gene *PAICS* (relevant tags highlighted in purple). The paired ChIA-PET tags coincide with DNaseI hypersensitivity sites in the leukemia cell line (blue track) and blood cells (violet track). Magnified views of the region around the ChIA-PET tags are shown in the last track. Bottom panel, normalized expression intensity levels for *CLOCK* (left) and *PAICS* (right) at baseline (6pm; red) and after dexamethasone stimulation (9pm; blue), stratified by rs7673908 genotypes. Note that rs7673908 shows significant associations only with the GR-stimulated gene expression of *PAICS* but not with the expression of *CLOCK* (Q values were obtained by GR-response *cis*-eQTL analysis).



(a) TF binding analysis. **(b)** Microarray data.

Figure 3.7.: Transcription factor binding site analysis of the GR. (a) First panel, gene structure of the *ASL* gene and position of the eSNP rs2460432. Second panel, sequence motif of the GR affinity matrix and a box indicating the position of rs2460432 within the motif. Third panel, 20bp sequence around rs2460432 and a box indicating the GR binding site with the highest affinity. Fourth panel, C-allele of rs2460432 impairs GR binding affinity. Affinity values are estimated for both alleles (orange line for C-allele and black line for T-allele). Peaks (continuous lines) indicate possible binding sites of the GR with a reduction in the predicted binding affinity for the C-allele sequence (total affinity indicated by horizontal dashed lines: $A_C = 8.83^{-6}$, $A_T = 4.68^{-6}$). (b) SNP rs2460432 significantly influences GR-induced changes in *ASL* expression (q value obtained by GR-response *cis*-eQTL analysis). The C-allele, which has a lower predicated affinity (as shown in 3.7 is associated with a smaller GR-induced change in *ASL* transcription.

3.2.4. GR-response eSNPs are enriched in loci nominally associated with MDD

In order to test whether our functionally relevant eSNPs are over-represented among nominally significant GWAS results for MDD, we determined the overlap between the GR-response eSNPs and all MDD-associated markers reaching a meta-analysis P value ≤ 0.05 from an analysis including approximately 9,000 cases and the same number of controls (part of the PGC data [133]; see section 2.5.2). Permutation analysis (see section 2.5.2) predicted an expected mean overlap of 210 SNPs (6%, range 168 to 255, SD = 13.9) from 1,000 randomly selected SNP sets. For 1,000 randomly generated baseline *cis*-eSNP sets,

a mean overlap of 218 SNPs (6.2%, range 174 to 268, SD = 13.5) was calculated. Both estimates were significantly lower than the actual overlap of 282 (8%) SNPs from the GR-response eSNPs. No simulated random set nor baseline *cis*-eSNP set ever yielded the same or greater overlap with MDD-associated SNPs (enrichment $ratio_{random} = 1.34$, enrichment $ratio_{baseline} = 1.29$, permutation-based FDRs < 0.001; see figure 3.8).

The 282 GR-response eSNPs that overlap with MDD-associated SNPs correspond to 23 unique eSNP bins (reflecting 26 eQTL bins) that regulate 25 unique transcripts (see table 3.2). We call these 23 eSNP bins GR/MDD eSNP bins in the remaining manuscript to refer to GR-response eSNPs that also show a nominal association with MDD in this meta-analysis. If we restricted this analysis to tagging eSNPs only ($n = 285$; see section 2.5.2) we still observed an enrichment compared to 1,000 equally sized sets of random and random baseline tagging eSNP sets (enrichment $ratio_{random} = 1.42$, permutation-based $FDR_{random} = 0.059$ and enrichment $ratio_{baseline} = 1.31$, permutation-based $FDR_{baseline} = 0.082$), indicating that the enrichment is independent from LD structure.

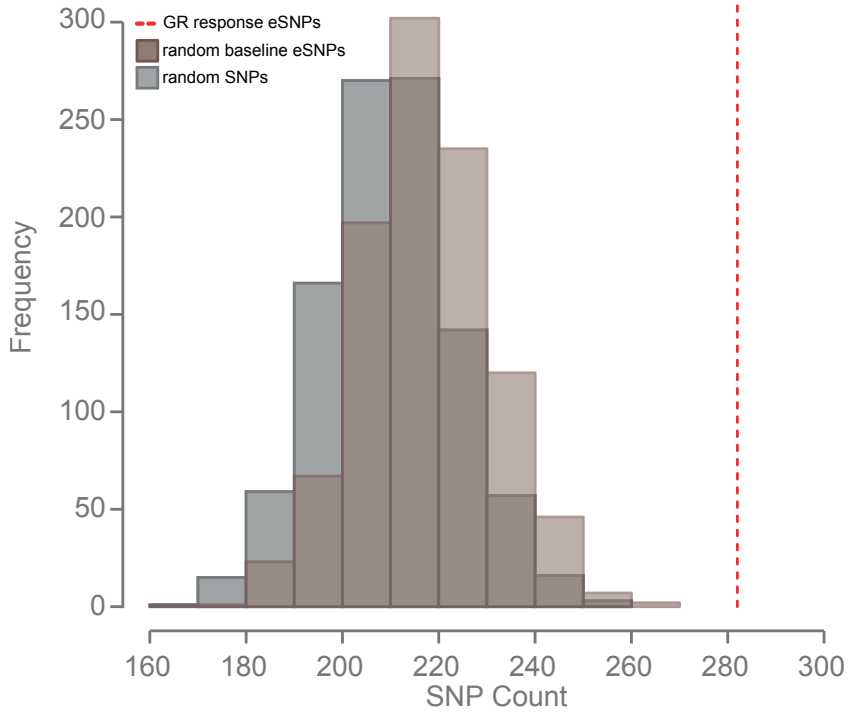


Figure 3.8.: GR-response eSNPs are enriched among variants associated with MDD (GR/MDD eSNPs). The dotted red line shows the number of GR-response eSNPs that overlap with SNPs in our meta-analysis for MDD. The distribution of the observed overlap for sets of 1,000 random SNPs (gray) and 1,000 random baseline eSNPs (brown) are represented as histograms.

	GR-response *cis*-eQTL data					PGC data						
	tag SNP	Genes nearby tag SNP	Probe id[a]	Probe gene	Q[b]	A1[c]	A2[d]	OR[e]	Risk[f]	P[g]	CD Association[h]	Net[i]
1	1-148440425	*PLEKHO1, ANP32E*	ILMN_1695435	*HIST2H2AA3/4*	0.006	T	G	1.09	T	0.013	CD, BPD, SCZ, ADHD	yes
2	19-40883657	*UPK1A, ZBTB32*	ILMN_1720542	*POLR2I*	0.044	C	G	0.91	G	0.001	BPD	yes
3	rs10002500	*CNGA1*	ILMN_1700306	*OCIAD2*	0.024	T	C	1.07	T	0.043		yes
4	rs10505733	*CLEC4C*	ILMN_1665457	*CLEC4C*	0.00021	A	C	0.94	C	0.021	SCZ	yes
	rs10505733	*CLEC4C*	ILMN_1682259	*CLEC4C*	0.00021	A	C	0.94	C	0.021	SCZ	yes
5	rs12432242	*SLC7A7*	ILMN_1810275	*SLC7A7*	0.041	T	C	0.94	C	0.008	CD, BPD	yes
6	rs12611262	*SEMA6B, TNFAIP8L1*	ILMN_1658486	*MRPL54*	0.046	T	C	1.06	T	0.022		yes
7	rs12620091	*ALMS1P*	ILMN_1662954	*CCT7*	0.047	T	C	0.95	C	0.022		yes
8	rs17239727	*BLVRA*	ILMN_2081335	*C7orf44*	0.024	A	G	0.94	G	0.022	CD	yes
9	rs1873625	*BSN*	ILMN_1705737	*IMPDH2*	0.048	A	C	0.94	C	0.018		yes
10	rs1981294	*LRIF1, DRAM2*	ILMN_1721989	*ATP5F1*	0.037	T	C	1.07	T	0.021	CD	yes
11	rs2072443	*TMEM176B*	ILMN_1791511	*TMEM176A*	0.036	T	C	1.05	T	0.034		yes
12	rs2269799	*SV2B*	ILMN_1663699	*SLCO3A1*	0.047	T	C	0.95	C	0.04		yes
13	rs2395891	*BTBD2, MKNK2*	ILMN_1721344	*MOB3A*	0.024	T	G	1.07	T	0.031	CD, BPD	yes
	rs2395891	*BTBD2, MKNK2*	ILMN_2347068	*MKNK2*	0.028	T	G	1.07	T	0.031	CD, BPD	yes
14	rs2422008	*WDPCP*	ILMN_1679268	*PELI1*	0.042	A	C	1.05	A	0.036	CD, ASD	yes
15	rs2956993	*GANAB*	ILMN_1746525	*FTH1*	0.044	T	G	0.95	G	0.032		yes
16	rs35288741	*NFASC*	ILMN_2094952	*NUAK2*	0.044	A	G	1.05	A	0.042		yes
17	rs6493387	*TRPM1*	ILMN_1778734	*MTMR15*	0.045	T	C	0.93	C	0.001	CD	no
18	rs6545924	*COMMD1, B3GNT2*	ILMN_1761242	*COMMD1*	0.045	T	G	1.06	T	0.018		yes
19	rs7194275	*C16orf91, CCDC154*	ILMN_1688749	*RPS2*	0.049	T	C	0.92	C	0.021	CD, BPD, SCZ	yes
20	rs7252014	*KCNN1*	ILMN_1766487	*LRRC25*	0.038	A	G	1.06	A	0.016		no
21	rs917585	*SLC6A7*	ILMN_1694686	*HMGXB3*	0.045	C	G	1.05	C	0.029	CD, SCZ	yes
22	rs9268671	*HLA-DRA, HLA-DRB5*	ILMN_1697499	*HLA-DRB5*	0.00021	A	G	0.95	G	0.031	CD, SCZ, ASD	yes
23	rs9268926	*HLA-DRA, HLA-DRB5*	ILMN_1697499	*HLA-DRB5*	0.012	A	G	0.92	G	0.041	CD, SCZ, ASD	yes
	rs9268926	*HLA-DRA, HLA-DRB5*	ILMN_2159694	*HLA-DRB4*	0.00073	A	G	0.92	G	0.041	CD, SCZ, ASD, ADHD	yes

Table 3.2.: List of 26 GR-response eSNP bins (23 tagging SNPs), representing the overlap of the 282 GR-response *cis*-eSNPs and SNPs from the meta-analysis for MDD (nominally associated with MDD; $n = 17,846$ samples- part of the PGC data).

[a] llumina probe identifier (Human HT-12 v3)
[b] Q values, which were obtained by GR-response *cis*-eQTL analysis
[c] code for allele 1 (reference allele, not necessary minor allele) in PGC data
[d] code for allele 2 in PGC data
[e] odds ratio in PGC data
[f] risk allele in PGC data
[g] meta analysis P value in PGC data
[h] probes that also had an eSNP associated with bipolar disorder (BPD), schizophrenia (SCZ), attention deficit-hyperactivity disorder (ADHD), autism spectrum disorder (ASD) or the cross disorders analysis (CD)
[i] probe genes that generated a tightly interconnected network in figure 3.11 (yes)

We next constructed a genetic risk profile score (GRPS) using the 23 GR/MDD tagging eSNPs for each individual, by adding up the individual's number of risk alleles of each SNP (see section2.5.2). These GRPSs are associated with MDD in an independent cohort ($Z = 3.76$, $P = 0.00017$; 1,005 MDD cases and 478 controls; see table 3.3 and section 2.5.2); specifically, individuals with higher GRPSs were overrepresented in the MDD group (see figure 3.9). The association was significantly enriched compared to 1,000 randomly generated SNP profiles (see section 2.5.2) with a permutation-based FDR = 0.008.

	GR-response *cis*-eQTL data			**MARS cohort**				
	tag SNP	Proxy for SNP	Genes nearby tag SNP	A1[a]	A2[b]	MAF	HWE[c]	Used
1	1-148440425	rs72694971 (renamed)	*PLEKHO1, ANP32E*	G	T	0.12	0.56	yes
2	19-40883657	rs73048504 (renamed)	*UPK1A, ZBTB32*	C	G	0.18	0.22	yes
3	rs10002500		*CNGA1*	T	C	0.13	0.58	yes
4	rs10505733		*CLEC4C*	C	A	0.29	0.42	yes
5	rs12432242		*SLC7A7*	C	T	0.39	0.87	yes
6	rs12611262		*SEMA6B, TNFAIP8L1*	T	C	0.39	0.59	yes
7	rs12620091	rs34874205 ($r^2 = 0.92$)	*ALMS1P*	C	T	0.37	<0.00001	no
8	rs17239727		*BLVRA*	T	C	0.21	0.48	yes
9	rs1873625		*BSN*	A	C	0.29	0.85	yes
10	rs1981294		*LRIF1, DRAM2*	C	T	0.17	0.47	yes
11	rs2072443		*TMEM176B*	T	C	0.41	0.75	yes
12	rs2269799		*SV2B*	C	T	0.32	0.23	yes
13	rs2395891		*BTBD2, MKNK2*	T	G	0.35	0.21	yes
14	rs2422008		*WDPCP*	A	C	0.43	1	yes
15	rs2956993		*GANAB*	G	T	0.38	0.30	yes
16	rs35288741		*NFASC*	G	A	0.35	0.25	yes
17	rs6493387		*TRPM1*	T	C	0.47	0.11	yes
18	rs6545924		*COMMD1, B3GNT2*	G	T	0.50	0.30	yes
19	rs7194275		*C16orf91, CCDC154*	C	T	0.12	0.0007	yes
20	rs7252014		*KCNN1*	A	G	0.48	0.054	yes
21	rs917585		*SLC6A7*	G	C	0.50	0.57	yes
22	rs9268671	rs116072659 (renamed)	*HLA-DRA, HLA-DRB5*	A	G	0.34	<0.00001	no
23	rs9268926	rs114766558 ($r^2 = 0.81$)	*HLA-DRA, HLA-DRB5*	G	A	0.31	<0.00001	no

Table 3.3.: GR/MDD tagging eSNPs and their proxy SNPs used to generate the cumulative risk allele profile in the MARS cohort. Three SNPs deviated from HWE (rs12620091, rs9268671 and rs9268926) and were excluded from the analysis. As result the remaining 20 SNPs were used to generate a profile.

[a] code for allele 1 (reference allele, not necessary minor allele) in MARS cohort
[b] code for allele 2 in MARS cohort
[c] Hardy-Weinberg test statistics (P values) in MARS cohort

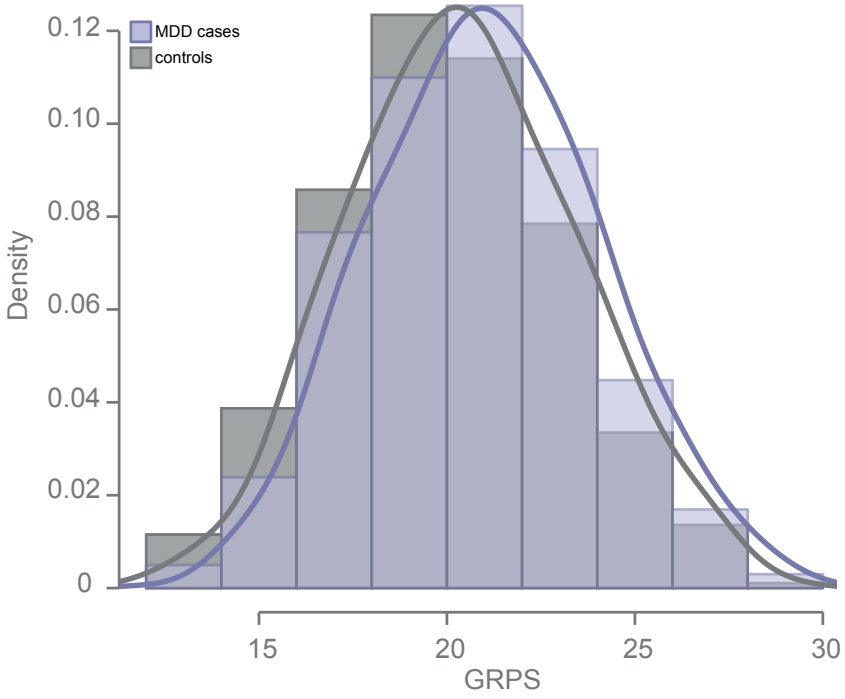


Figure 3.9.: The distribution of the GRPSs for an independent sample of MDD cases (violet) and controls (gray) are represented as histograms and kernel density curves. Individuals with depression display higher GRPSs.

3.2.5. Cumulative risk scores for the GR/MDD eSNPs correlate with dysfunctional amygdala reactivity[2]

To investigate the relationship between GR/MDD tagging eSNP GRPSs and variability in stress-related brain function, we applied an imaging genetics strategy to data from 647 participants (171 individuals with current or past DSM-IV Axis I disorders and 476 controls; 306 of participants were self-reported EUR-AM; see table 3.4) of the Duke Neurogenetics Study (DNS) (see section 2.1). Our analyses focused on centromedial amygdala reactivity to canonical threat-related angry and fearful facial expressions (see figure 3.10a), because this phenotype is clearly implicated in the etiology and pathophysiology of stress-related disorders, including depression [173]. Moreover, activity in the amygdala triggers a coordinated behavioral and physiologic response to threat. This includes activation of the stress hormone response via projections from the medial division of the central nucleus of the amygdala, (captured in our analysis by our centromedial amygdala region of interest) to the paraventricular nucleus of the hypothalamus [217].

A significant effect of GRPSs (see section 2.5.2) on centromedial amygdala responses to angry and fearful facial expression in comparison to neutral expressions was identified for the EUR-AM DNS subsample ($F_{1,301} = 7.06$, $P = 0.008$; see figure 3.10a and 3.10b) after correcting for age, sex, and the presence of an Axis I disorder. The effect was found in the entire sample as well, after accounting for population stratification ($F_{1,637} = 6.05$, $P = 0.014$; see supplementary figure A.3). In both the EUR-AM subsample and entire sample, individuals with higher GRPSs had blunted centromedial amygdala reactivity to angry and fearful facial expressions in comparison to neutral expressions relative to individuals with lower GRPSs (see figure 3.10b, supplementary figure A.3). Permutation analyses that formed random SNP profiles ($n = 1,000$; matched for MAF and not exceeding the maximum correlation among profile SNPs; see section 2.5.2) indicated that the actual GRPSs were more likely to be associated with these differences in amygdala reactivity than 1,000 sets of random SNP profiles (EUR-AM subsample: permutation-based $FDR = 0.003$; entire sample: permutation-based FDR = 0.012). Post-hoc analyses revealed that this differential effect was driven by a higher centromedial amygdala reactivity to neutral facial expressions in comparison to shapes in participants with higher GRPS (EUR-AM subsample: $F_{1,301} = 6.47$, $P = 0.011$; see figure 3.10c and entire sample: $F_{1,637} = 8.52$, $P = 0.004$; see supplementary figure A.3 and A.4b) while there were no effects of GRPS on amygdala response to angry and fearful facial expressions in comparison to shapes (EUR-AM subsample: $F_{1,301} = 0.2$, $P = 0.65$; see figure 3.10d and entire sample: $F_{1,637} = 0.09$, $P = 0.76$; see supplementary figure A.3 and and figure A.4a).

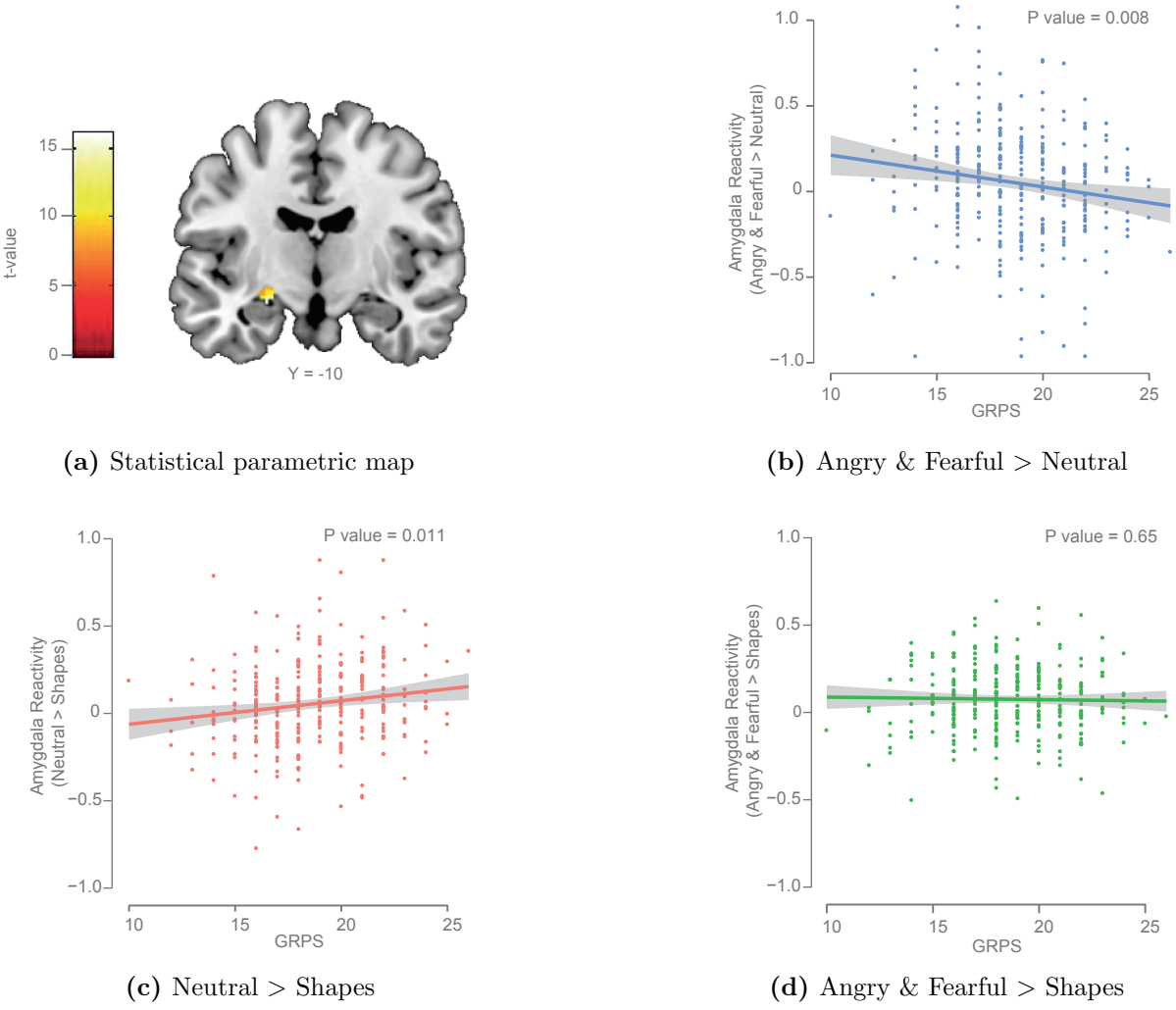
This pattern of altered amygdala reactivity in individuals with higher GRPSs is suggestive of impaired threat-related cue learning with inappropriately increased reactivity

[2]The imaging analysis in the DNS cohort was conducted in collaboration with Ryan Bogdan[+,*] and Ahmad R. Hariri[*].

[+] Department of Psychology, Washington University in St Louis, St Louis, MO, USA

[*] Department of Psychology and Neuroscience, Institute for Genome Sciences and Policy, Duke University, Durham, NC, USA

to neutral expressions, which do not convey threat [21, 160]. Thus, higher GRPS may be associated with non-specific or overgeneralized threat and stress responses, which are consistently observed in depression as well as other mood and anxiety disorders [21, 160].



(a) Statistical parametric map

(b) Angry & Fearful > Neutral

(c) Neutral > Shapes

(d) Angry & Fearful > Shapes

Figure 3.10.: GR-response eSNPs correlate with dysregulated threat-related amygdala reactivity. (a) Statistical parametric map illustrating left centromedial amygdala reactivity to the "Angry & Fearful > Neutral" contrast in the entire sample (15 contiguous voxels; max voxel MNI coordinate, $x = -24$, $y = -10$, $z = -14$, $t = 4.35$, $P = 7.76 \times 10^{-6}$). (b) Higher DNS GRPSs in the European-American subsample ($n = 306$) predicted amygdala reactivity to threat-related facial expressions in comparison to neutral facial expressions. Post-hoc analyses (c, d) revealed that GRPSs did not predict amygdala reactivity to threat-related expressions (d), but that higher GRPSs predicted elevated amygdala reactivity to neutral facial expressions (c), in comparison to non-face control stimuli. The 95% confidence interval is displayed as gray shaded band in (b-d).

	GR-response *cis*-eQTL data			DNS cohort								
						MAF		HWE[d]				
	tag SNP	Proxy for SNP[a]	Genes nearby tag SNP	A1[b]	A2[c]	EUR-AM	ALL	EUR-AM	AFR-AM	Latino/a	ASN1	ASN2	Used
1	1-148440425	rs11588837 ($r^2 = 0.96$)	*PLEKHO1, ANP32E*	A	G	0.15	0.34	0.48	0.95	0.34	0.99	0.72	yes
2	19-40883657	rs8106959 ($r^2 = 0.95$)	*KMT2B*	A	G	0.22	0.18	0.53	0.89	0.87	0.28	0.5	yes
3	rs10002500		*CNGA1*	T	C	0.1	0.19	0.28	0.74	0.65	0.48	0.5	yes
4	rs10505733	rs1894823 ($r^2 = 1$)	*CLEC4C*	T	C	0.31	0.28	0.34	0.4	0.16	0.14	0.35	yes
5	rs12432242	rs2281677 ($r^2 = 0.93$)	*SLC7A7*	A	G	0.38	0.39	0.96	0.29	0.04	0.16	0.31	yes
6	rs12611262		*SEMA6B, TNFAIP8L1*	T	C	0.37	0.44	0.49	0.84	0.57	0.26	0.55	yes
7	rs12620091	no Proxy											no
8	rs17239727	rs10229363 ($r^2 = 1$)	*BLVRA*	A	G	0.2	0.13	0.23	0.62	0.47	0.86	0.35	yes
9	rs1873625	rs9858280 ($r^2 = 1$)	*BSN*	T	C	0.37	0.28	0.39	0.6	0.71	0.52	0.24	yes
10	rs1981294	rs4838884 ($r^2 = 1$)	*LRIF1, DRAM2*	A	G	0.2	0.19	0.63	0.66	0.48	0.932	0.67	yes
11	rs2072443		*TMEM176B*	T	C	0.42	0.44	0.38	0.41	0.59	0.39	0.74	yes
12	rs2269799		*SV2B*	C	T	0.33	0.35	0.1	0.6	0.32	0.5	0.35	yes
13	rs2395891		*BTBD2, MKNK2*	T	G	0.34	0.38	0.49	0.18	0.26	0.3	0.03	yes
14	rs2422008		*WDPCP*	A	C	0.47	0.41	0.85	0.25	0.9	0.13	0.82	yes
15	rs2956993		*GANAB*	G	T	0.35	0.29	0.42	0.47	0.43	0.61	0.99	yes
16	rs35288741	rs7534993 ($r^2 = 0.95$)	*NFASC*	G	A	0.34	0.27	0.24	0.21	0.56	0.53	0.35	yes
17	rs6493387	rs12901022 ($r^2 = 1$)	*TRPM1*	C	T	0.48	0.46	0.79	0.44	0.41	0.94	0.82	yes
18	rs6545924	rs921320 ($r^2 = 1$)	*COMMD1, B3GNT2*	C	A	0.5	0.5	0.17	0.53	0.4	0.65	0.94	yes
19	rs7194275		*C16orf91, CCDC154*	C	T	0.19	0.19	0.5	0.92	0.73	0.051	1	yes
20	rs7252014		*KCNN1*	A	G	0.48	0.47	0.55	0.37	0.31	0.07	0.45	yes
21	rs917585	no Proxy											no
22	rs9268671	no Proxy											no
23	rs9268926	no Proxy											no

Table 3.4.: GR/MDD tagging eSNPs and their proxy SNPs used to generate the cumulative risk allele profile in the DNS cohort. Four SNPs did not have a proxy available (rs12620091, rs917585, rs9268671 and rs9268926). No SNPs deviated from HWE.

[a] r^2= LD for CEU population from 1KGP (> 0.90 for all subpopulations)
[b] code for allele 1 (reference allele, not necessary minor allele) in DNS cohort
[c] code for allele 2 in DNS cohort
[d] Hardy-Weinberg test statistics (P values) in DNS cohort (European-Americans (EUR-AM), East Asian (ASN), African American (AFR-AM))

3.2.6. Functional relevance of transcripts regulated by GR/MDD eSNPs

3.2.6.1. Network-based analysis of GR/MDD genes[3]

To validate our results, we investigated whether the probe genes ($n = 24$) regulated by the GR/MDD eSNPs are part of specific pathways that may be relevant for the pathophysiology of MDD. We were able to generate a tightly interconnected network containing 22 of the 24 gene products, based on manually curated relationships extracted from the scientific literature (see figure 3.11, supplementary table A.3). This network revealed that the 22 gene products show direct associations with mood disorders and response to antidepressant treatment in independent datasets. In addition, they are predominantly involved in pathways associated with ubiquitination and proteasome degradation and the inflammatory response - systems that have been implicated in the pathophysiology of MDD and in stress-related changes in synaptic plasticity [149, 214].

These gene products not only interact on the protein level but also appear to be co-regulated on a transcriptional level. Co-expression analysis (see section 2.5.2) from the GR-stimulated gene expression measures in blood cells from all individuals of the MPIP cohort ($n = 160$) identified that the 25 GR/MDD array probes are more tightly co-regulated than 1,000 sets of randomly chosen transcripts selected from all GR-stimulated transcripts (inverse enrichment $ratio = 1.04$, permutation-based $FDR = 0.078$). These data suggest that these 25 array probes (24 genes) not only functionally interact on the protein level but are also coordinated in their transcriptional response to GR activation or stress to perform an orchestrated function.

3.2.6.2. Convergent functional genomics: integrating human GR/MDD genes with relevant mouse models

To establish whether the transcripts regulated by acute GR activation in blood are also regulated in the brain in a similar timeframe, we investigated whether the orthologues of the 24 GR/MDD genes were differentially regulated in mouse blood and brain (PFC, HC, and AM) following dexamethasone administration (10 mg/kg dexamethasone i.p.)). In this experiment 15 of the 24 genes had a mouse orthologous gene, which were expressed above detection threshold (see supplementary table A.3). Ten (66.7%) of the 15 genes showed significant changes in transcriptional levels 4 hours after dexamethasone administration in one or more of the investigated brain regions, and all 15 genes were also regulated in mouse blood (see figure 3.12 right panel and supplementary table A.4).

In order to better link the 24 GR/MDD genes to actual risk for MDD and not only GR reactivity, we further investigated whether chronic social stress differentially regulates the same 24 GR/MDD genes in resilient vs. susceptible mice. In this animal model,

[3]The interaction network was conducted in collaboration with Goar Frishman* and Andreas Ruepp;*.
* Helmholtz Zentrum München - German Research Center for Environmental Health, Ingolstädter Landstrae 1, 85764 Neuherberg, Germany

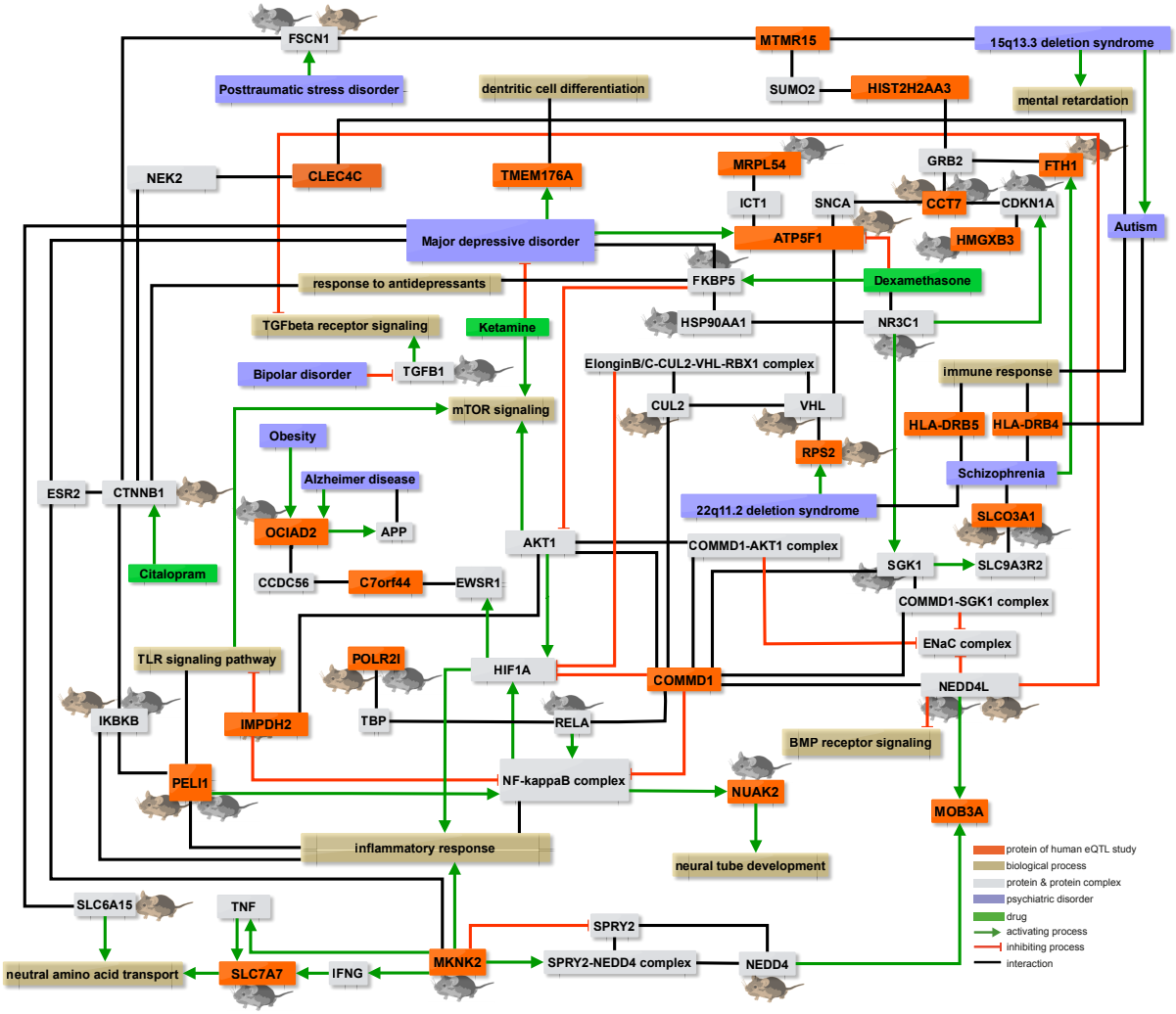


Figure 3.11.: For 22 of the 24 GR/MDD genes a tightly interconnected network was generated from manually curated experimental data derived from the scientific literature. Transcripts that were differently expressed in the GR-stimulation mouse model are marked by a gray mouse and those differential regulated in the chronic social stress mouse model are marked by a brown mouse.

mice were exposed to seven weeks of chronic social stress during adolescence. After five weeks of recovery, susceptible animals exhibit depression-like behavioral and endocrine phenotypes, while this is not the case for resilient animals [110, 190, 189]. In this experiment 11 orthologous genes were analyzed in HC samples (see supplementary table A.4). After 5 weeks recovery, 8 (72.7%) of 11 genes expressed in the CA1 area and/or the DG showed nominally significant (absolute $DiffScore > 13$) differential expression between resilient vs. susceptible animals (see supplementary table A.4). Interestingly, for 6 of these 8 genes, the difference in gene expression regulation between resilient and susceptible animals in the HC matched the direction of the difference in GR-induction in humans with the risk

allele vs. humans with the protective genotypes (see figure 3.12 left panel).

We also tested whether additional transcripts of proteins linked to the GR/MDD genes in the interaction network analysis were significantly regulated in either of the two mouse models. We identified a set of 15 additional transcripts from the network (see supplementary table A.5) that were regulated in brain by GR activation or were associated with risk or resilience to exposure to chronic social stress; these are marked in figure 3.11.

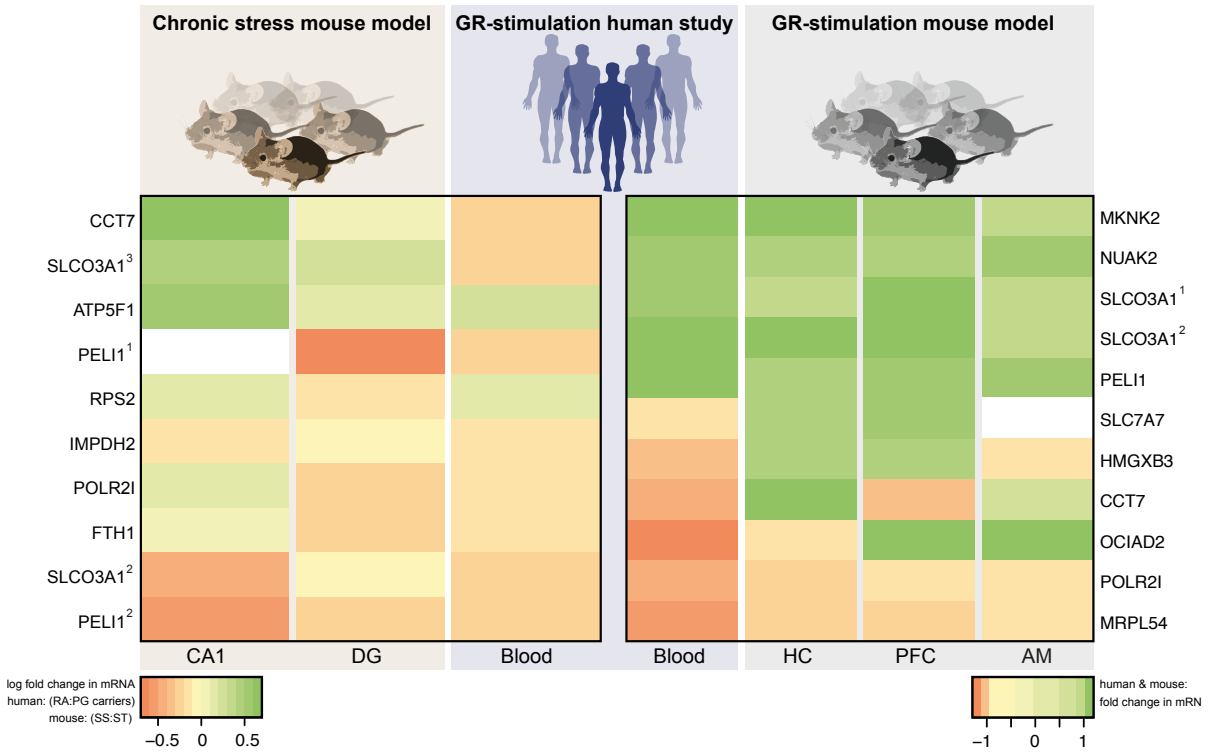
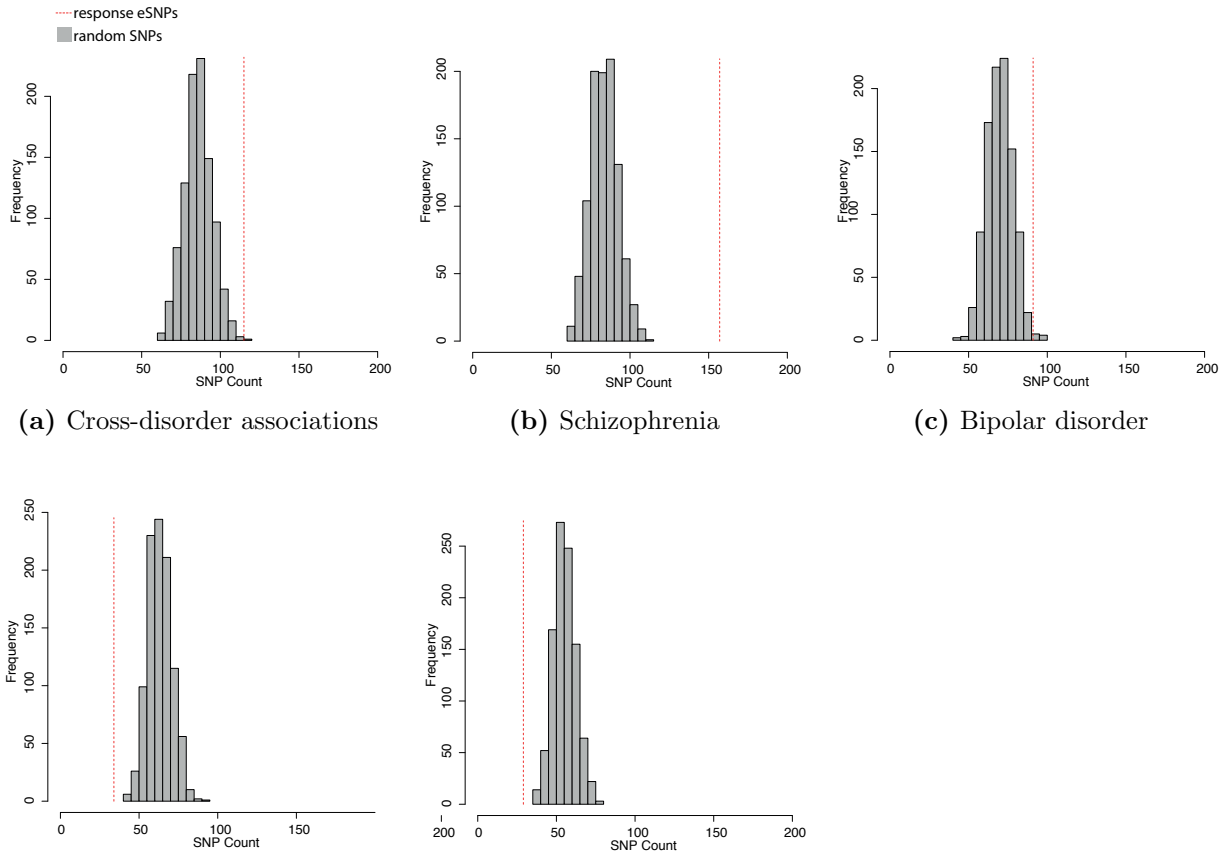


Figure 3.12.: Rows of the heatmaps correspond to genes, and columns correspond to the analyzed tissues. First heatmap: left panel, log fold change in gene expression between stress-susceptible (SS) vs. stress-resilient (SR) groups of mice (brown, $n = 12$ mice) in the CA1 and DG region of the hippocampus and difference in GR-induced expression of these genes in blood cells in humans between risk allele (RA) and protective genotype (PG) carriers (blue, $n = 160$ samples). Second heatmap: gene expression changes from baseline to GR-stimulation in mouse brain (gray, $n = 22$ mice) and human blood cells (blue). Investigated tissues are labeled within the bottom row of every heatmap (PFC, HC, and AM, CA1 area of HC and DG). Log fold change in mRNA levels are color coded as indicated in bottom of each heatmap.

3.2.7. GR-response eSNPs are enriched in loci associated with other psychiatric disorders

Recent studies suggest shared genetic risk loci for five major psychiatric disorders, which include MDD, BPD, SCZ, ADHD and ASD [37]. In fact, using GWAS data to identify SNP-based genetic correlation (rg_{SNP}), MDD has been shown to have relevant genetic correlation with BPD ($rg_{SNP} = 0.47$), SCZ ($rg_{SNP} = 0.43$) and ADHD ($rg_{SNP} = 0.32$) but not with ASD ($rg_{SNP} = 0.05$) [38].



(a) Cross-disorder associations **(b)** Schizophrenia **(c)** Bipolar disorder

(d) Autism spectrum disorder **(e)** Attention deficit-hyperactivity disorder

Figure 3.13.: The plots show the results of the GR-response eSNP enrichment analysis in GWAS for bipolar disorder (BPD), schizophrenia (SCZ), attention deficit-hyperactivity disorder (ADHD), autism spectrum disorder (ASD) and the cross-disorder associations (CDA). The dotted red lines show the actual overlap of GR-response eSNPs with the respective disease SNPs. The histograms represent the distribution of the overlap observed for sets of 1,000 random SNPs matched for MAF. GR-response eSNPs are enriched for CDA (permutation based $FDR = 0.001$), SCZ (permutation based $FDR < 0.001$) and BPD (permutation based $FDR = 0.009$) susceptibility loci.

Considering this shared genetic liability across at least some of these five psychiatric disorders and the fact that adverse life events confer risk for most of them [144], we tested for enrichment of our GR-response eSNPs in the association data set from the PGC cross-disorder analysis (33,332 cases and 27,888 controls; see section 2.5.2). We also tested for enrichment in each disorder independently (9,379 SCZ cases; 6,990 BPD cases; 840 ADHD cases, 1,947 ADHD trio cases; 161 ASD cases, 4,788 ASD trio cases; see figure 3.13 and table 3.5). There was evidence for significant GR-response eSNP enrichment in 1,000 randomly generated SNP sets in the cross-disorder associations (enrichment $ratio = 1.33$, permutation-based $FDR = 0.001$). Furthermore, there was evidence of GR-response eSNP enrichment in SCZ (enrichment $ratio = 1.87$, permutation-based $FDR < 0.001$) and BPD (enrichment $ratio = 1.29$, permutation-based $FDR = 0.009$). By contrast, a significant underrepresentation in ASD (inverse enrichment $ratio = 1.87$, permutation-based $FDR < 0.001$) and ADHD (inverse enrichment $ratio = 1.92$, permutation-based $FDR < 0.001$) was observed.Those findings are consistent with the evidence that SCZ and BPD have the highest co-heritability with MDD, while ASD and ADHD have the lowest co-heritability with MDD [38].

	GR-response eSNPs	Random eSNPs		
	Count	Mean count[a]	Range	FDR[b]
CDA*	115 (11%)	86.5 ± 8.99 SD (8%)	61-119	0.001
BPD*	91 (9%)	70.36 ± 8.34 SD (7%)	44-100	0.009
SCZ*	157 (15%)	84.08 ± 8.79 SD (8%)	61-111	<0.001
ADHD*	29 (3%)	55.69 ± 7.14 SD (5%)	36-79	1
ASD*	34 (3%)	63.73 ± 7.62 SD (6%)	44-91	1
MDD#	282 (8%)	210 ± 13.9 SD (6%)	168-255	<0.001

Table 3.5.: Number (proportion) of GR-response eSNPs overlapping with SNPs from the PGC cross-disorder analyses (*) and meta-analysis for MDD (#).

[a] mean count (proportion) of the number of GR-response eSNPs observed for 1,000 random draws of 1,047* and 3,492# SNPs from bins matched for the MAF to the GWAS SNPs

[b] permutation-based FDR

4. Discussion

This doctoral thesis presented genome-wide results from both a differential gene expression as well as an eQTL study and shows that both analyses are important contributors in the understanding of the pathophysiology of stress-related psychiatric disorders, especially MDD.

The differential gene expression analysis showed that GR-stimulated gene expression changes in peripheral blood cells enable a much better discrimination between depressed patients and healthy controls than gene expression measures at baseline. Interpretation of the GR-stimulated gene expression profiles of the top differentially regulated genes (see figure 3.1 and table 3.1) led to a correct prediction for 79.2% of the tested samples, and thus outperformed baseline gene expression patterns with a rate of 41.7% correct predictions. This supports the previously described superiority of GR-stimulated over baseline gene expression differences as biomarkers for MDD [203]. Spijker and colleagues [203] investigated gene expression changes in peripheral blood monocytes following an *ex vivo* challenge with lipopolysaccharide, a strong immunogenic stimulus. This paradigm is less influenced by confounding variables than the here presented *in vivo* challenge test. However, the reported sensitivity and specificity values are comparable to the ones reported here in this thesis (sensitivity: 76.9% vs. 80% and specificity: 71.8% vs. 87.5%.). This suggests that *in vivo* stimulated gene expression patterns, which require less hands on time in the laboratory than *ex vivo* stimulation, could be better suited to further serve as potential biomarkers for depression-related GR function changes.

Endocrine measures could serve as an alternative to gene expression patterns to classify depressed cases and healthy controls. Such endocrine measures, e.g. cortisol or ACTH, showed a robust suppression in depressed patients as well as healthy controls after dexamethasone investigation. However, they failed to be a reliable discriminator between both groups (for more details please see [148]) in our data set.

Robust GR-induced reproducible gene expression profiles in both depressed patients and healthy controls were found. Interestingly, three of the most significantly GR-regulated genes in both groups (*FKBP5*, *DUSP1* and *ZBTB16*) have been previously reported to be involved in the development of stress-related psychiatric disorders or neuroprotection. Briefly, SNPs in *FKBP5* have been associated with a number of phenotypes related to mood disorders [12] and *FKBP5* gene expression differences have been reported for post-traumatic stress disorder [146, 193]. *DUSP1* has been shown to be more strongly expressed in postmortem hippocampus tissue from depressed patients as compared to healthy controls [55]. Additionally, it mediates stress-related depression-like behavior in rodents [55]. *ZBTB16* is a corticosteroid-responsive transcription factor. It is induced by stress exposure

in the brain and may have a neuroprotective function [170].

Of the 19 genes contributing most to our classifier (see figure 3.1 and table 3.1), *FKBP5* and *TMEM176A* have previously been associated with stress-related disorders. Alternated gene expression profiles in HC and PFC of *TMEM176A* were identified in a rat model of depression-like behavior [14]. Thus, these results support the importance of using GR-stimulated gene expression patterns as biomarkers for depression.

A limitation of this study was the discordant medication with antidepressants of the patients compared to un-medicated controls. This generally complicates downstream analyses of the observed associations since it is difficult to examine whether gene expression changes are related to treatment or disease status. But even if the observed differences were exclusively related to the effects of antidepressant drugs, this would not affect the observed superiority of GR-stimulated over baseline gene expression. To exclude possible medication effects, we post-hoc tested if the duration or number of antidepressant treatment/s at the time of RNA collection had an impact on gene expression regulation in all patients for the 19 genes that best differentiate between cases and controls (see table 3.1). We did not find any significant effect (see table 4.1) suggesting that antidepressant medication is not likely to have a major impact on the expression of those genes. To further exclude medication as a confounder of the reported case-control differences, a new experiment comparing medicated and un-medicated patients and controls has to be investigated as follow-up.

Probe id	Gene name	Type of medication	Duration of medication
ILMN_1718982	*BEST1*	0.69	0.82
ILMN_1724422	*SELL*	0.97	1
ILMN_1752526	*RNF144B*	0.39	0.21
ILMN_1773963	*GNA15*	0.08	0.09
ILMN_1661755	*FAM129B*	0.36	0.32
ILMN_2372915	*P2RY2*	0.08	0.34
ILMN_1672834	*SSH2*	0.85	0.27
ILMN_1712389	*CKLF*	0.79	0.34
ILMN_1791511	*TMEM176A*	0.99	0.62
ILMN_1692464	*FLJ20699*	0.65	0.49
ILMN_1764764	*MUM1*	0.35	0.4
ILMN_1778444	*FKBP5*	0.72	0.19
ILMN_1880406	no symbol	0.2	0.8
ILMN_1740015	*CD14*	0.9	0.47
ILMN_1774733	*SOCS1*	0.65	0.38
ILMN_1801504	*RUNX1*	0.78	0.68
ILMN_1812281	*ARG1*	0.86	0.29
ILMN_2326953	*LAT2*	0.64	0.94
ILMN_2400759	*CPVL*	0.22	0.07

Table 4.1.: List of the 19 transcripts contributing most to our classifier for MDD disease status. The duration or number of antidepressant treatment/s had no impact on gene expression regulation of the listed transcripts.

Antidepressant drugs have been reported to increase GR mRNA expression [28, 66, 87, 222], GR protein expression [28, 89, 113] and GR function as measured using the translocation of the receptor to the nucleus [28, 66, 87, 165] or dexamethasone- induced GR-mediated gene transcription [162, 163, 164, 165, 169]. This pattern of increased GR activation was not observed here. Moreover, the number of genes significantly regulated by dexamethasone in depressed patients was significantly lower than in controls (1.151 vs. 2.670 transcripts). In addition, the fold changes of the 19 genes used for classification (see table 3.1) were lower in patients than the ones in controls (mean fold change in patients: sample $1 = 1.19 \pm 1.51$, sample $2 = 1.24 \pm 1.72$ vs. controls: sample $1 = 1.23 \pm 1.51$,sample $2 = 1.27 \pm 1.78$). This observation could likely reflects baseline differences in GR sensitivity. This is further supported since *FKBP5* shows less regulation after dexamethasone in patients than in controls. *FKBP5* is a heatshock protein 90 associated co-chaperone of the GR and its expression is strongly induced by glucocorticoids via intronic steroid response elements as part of an intracellular short negative feedback loop for GR activity [223]. Its induction by GR activation has been proposed as a molecular indicator of GR sensitivity [98, 223].

On the other hand, this thesis also reported a genome-wide analysis of genetic variants that influence the GR-induced gene expression changes of *cis*-genes ($\pm$1Mb) in peripheral blood cells of 160 male individuals. Based on the results of previous reports [186] and our current sample size we chose not to engage to the analysis of *trans* variants.

This study showed that common variants in long-range enhancer elements alter the transcriptional responsiveness of a network of GR target genes to the GR, and that these variants cumulatively increase the risk for stress-related psychiatric disorders, including MDD. These findings suggest that the risk of developing MDD after adverse life events may be influenced by an individual's sensitivity to the downstream, transcriptional effects of cortisol released during the stressful adverse events. In addition, the findings suggest that the alterations in the very first transcriptional response to stress may influence how an individual processes stressful exposures. Indeed, the risk variants were also associated with altered centromedial amygdala reactivity to threat-related cues. Such abnormal neural processing of threat-related cues may mediate the increase in risk for MDD and other psychiatric disorders.

One of our notable genetic findings was that the distance between the GR-response eSNPs and the regulated probe is significantly longer than the distances previously reported for baseline eQTLs (mean distance of 406kb for GR-response eQTLs vs. 149kb baseline eQTLs in our dataset). Our data support and extend previous observations that indicated a long-range transcriptional regulation by the GR [200, 100, 79], i.e. GREs were generally distally distributed between upstream and downstream regions of the transcription start and end site. Particularly, 63% of the GREs were distal, i.e. further than 10kb from the transcription start and end site, whereas only 31% were proximal (within 5kb from the transcription start and end site) [100]. In fact, a combined analysis of our GR-response eQTLs and ChIA-PET data [121] from the ENCODE project [57] suggests that there could be a physical long-range interaction between the eSNP locus and the pro-

moter of the GR-regulated transcript for at least 25% of the GR-response eQTLs. This observation was more frequent than expect (permutation-based FDR = 0.056). Additional experiments that investigate the direct effects of the different alleles on the enhancer function are necessary to further validate this suggestion. The finding that the identified GR-response eSNPs do indeed tag GREs is supported by a strong enrichment of GREs and other GR-function-related TF binding sites [103] among GR-response eSNP sequences compared to random sequences. The potential long-range transcriptional regulations by GR activation further highlights the importance of using functional data for the annotation of GWAS signals, including those for MDD [142]. Proximity to or location within a gene does not necessarily indicate that associated SNPs regulate the expression of this gene or that they do not impact the regulation of more distant genes (see figure 3.6), even if there is convincing prior evidence for pathophysiological relevance (such as, in our case, for the *CLOCK* gene and MDD [166]).

Our results indicate that stimulated eQTL approaches that involve disease-risk relevant transcriptional stimuli (in our case GR activation and stress) are useful in identifying novel risk genes for common disorders. Previous studies have used eQTLs or DNA methylation QTLs (mQTLs) for the annotation of GWAS results [155, 238] and have indicated the importance of using eQTLs and mQTLs from disease-relevant tissues [68, 159]. While we do not observe a significant enrichment of baseline blood eQTLs, GR-response eQTLs from this tissue were enriched almost 1.34-fold among the variants associated with MDD. Our findings support the notion that not only the tissue but also the type of stimulation, e.g. mimicking aspects of stress in our experiments, can be relevant for using such QTL studies in annotating GWAS results.

While these common genetic variants were discovered in peripheral blood cells, we provided evidence for their importance in neural circuits that are critical for generating and regulating the stress axis response to adversity. First, using imaging genetics we demonstrated that the cumulative GR/MDD eSNP genetic risk profile predicts dysfunctional reactivity of the human amygdala. Second, the majority of the transcripts affected by these eSNPs in their GR-regulated gene expression in human blood were also regulated by short-term GR activation or following exposure to chronic stress in the mouse hippocampus, prefrontal cortex, or amygdala. In addition, 22 of the 24 genes formed a tightly interconnected network with numerous experimentally validated links to psychiatric and neurological disorders as well as antidepressant treatment (see figure 3.11 and supplementary table A.3). Next to inflammation, proteasome degradation was the pathway with the highest connectivity in our network. For example *PELI1*, *MKNK2*, *MOB3A* and *COMMD1*, all GR/MDD transcripts, are involved in ubiquitination. It has been shown that activation of GRs enhances ubiquitin/proteasome-mediated degradation of glutamate receptor subunits, and thereby mediates cognitive impairment induced by repeated stress exposure [239]. Genetic modulation of such effects may provide a mechanistic link between risk for psychiatric disorders and the genetic differences in GR-induced expression of ubiquitination-related genes observed in this study.

Most importantly, our GR-response eQTL analysis revealed an enrichment of these eSNPs among MDD-associated SNPs contrary to random SNP sets. This suggests that SNPs altering the first transcriptional response to stress also influence the risk for MDD. This association could be verified in an independent cohort and the increase in risk conferred by these functional variants may extend to SCZ and BPD. A recent study of *cis*-eQTLs in human cerebellumon for BPD reported a 1.32-fold enrichment of eSNPs over random sets [68], which is concordant with our GR-respone eSNP enrichment results for BPD (1.29-fold enrichment). A possible explanation for this agreement could be that our analysis takes advantages, since we only used a set of preselected, i.e., differentially regulated, probes. Alternatively, it is not clear to what extent results obtained on postmortem brain tissue are relevant for the disease in living subjects. Interestingly, when using the eQTL data to measure the enrichment of GR-response eSNPs in a disorder where blood is a more relevant tissue, like rheumatoid arthritis (RA), the eSNPs perform differently compared to MDD. In this case GR-response *cis*-eSNP show no enrichment (see figure 4.1). Thus, we further can conclude that GR-response eQTLs are more relevant for MDD and related disorders than for RA and future analyses may benefit from incorporating knowledge of *cis*-regulatory eSNPs from disease-risk relevant transcriptional stimuli.

The imaging genetics results provided one potential neural pathway by which GR/MDD eSNPs may increase risk for the development of stress-related psychopathology, including depression. Interestingly, GR/MDD eSNPs predict heightened amygdala reactivity to stimuli that do not inherently signal threat (i.e., neutral facial expressions); this suggests that GR/MDD eSNPs associated with the immediate transcriptome response to stress may impair the neural circuitry that supports the learning of threat-related cues and, possibly, thereby contribute to the overgeneralization of threat-related stress responses. Such overgeneralization may evoke stress responses in non-threatening situations and contribute to cognitive biases associated with the development of depression and other forms of psychopathology [29].

In summary, both analyses supported that studying GR-stimulated blood may help to give additional insights into disease etiology of stress-related psychiatric disorders, especially MDD. The results of the first part of this doctoral thesis underline the value of GR-stimulated gene expression profiles as a biomarkers for depression-related GR resistance. Studies in larger independent samples with different gender composition and different clinical settings will further explore the potential of the molecular dexamethasone-stimulation test as a biomarker helping to characterize subgroups within patient samples that fulfill current diagnostic criteria for a certain psychiatric category. The data presented in the last part of this thesis show that common genetic variants that change the GR-mediated immediate transcriptome response to stress are linked, in long-term, to both changes in neural processing of threat and increased risk for MDD and other psychiatric disorders. To our knowledge this is the first *in vivo* study of eQTLs that moderate the transcriptional response to glucocorticoids. Two previous studies reported GR-stimulated *cis*-eQTLs using

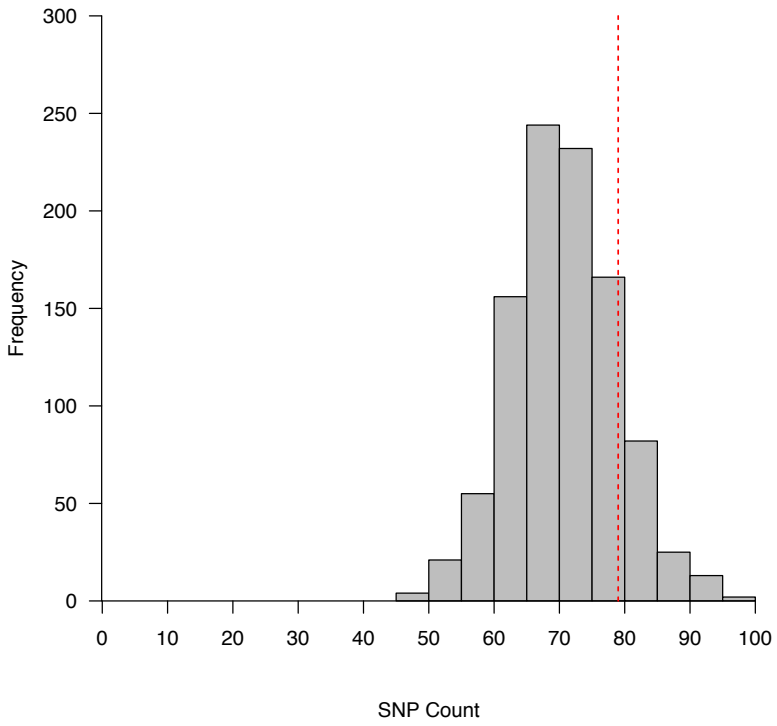


Figure 4.1.: An enrichment of GR-response eSNPs was not observed for a meta-analysis of rheumatoid arthritis (RA; n = 5,500 cases and 20,000 controls) [204], a disease in which glucocorticoids are one of the effective treatments [10]. Only 6.5% of the GR-response eSNPs (the dotted red line) were associated with RA at the significant level of 0.05. The distribution of the observed overlap for sets of 1,000 random SNPs is represented as histogram.

in vitro dexamethasone incubation in osteoblasts [75] and lymphoblastoid cell lines [137], respectively. Unfortunately, due to differences in study design, e.g. cell types, smaller *cis*-windows and in vitro conditions, which may have an effect on gene expression levels, our results are not directly comparable with these reports. Another important conclusion from the here reported data led further support to the notion of a possible shared genetic liability of some psychiatric disorders and specifically point to stress-responsive genes as common risk factors. Studies dissecting how these genetic variants alter the molecular, cellular, and neural response to glucocorticoids in the short- and long-term could inform the development of novel strategies for the prevention and treatment of stress-related psychiatric disorders.

A. Appendix

A.1. Supplementary Notes

A.1.1. Baseline cis-eQTLs

Using baseline gene expression of the 4,447 differently regulated autosomal array probes (absolute fold change ≥ 1.3 in at least 20% of all samples), 26,205 unique *cis*-SNPs and 764 gene expression probes corresponding to 31,541 *cis*-eQTLs were found to be significant after multiple testing correction with the same strategy as described for the GR-stimulated gene expression changes. The 26,205 unique eSNPs represent 1,010 uncorrelated eSNP bins (1,148 eSNP bin-probe combinations). 775 eQTL bins (68%) are located within 100 kb upstream or downstream from the array probe ends, 911 eQTL bins (79%) within 200 kb and only 237 eQTLs bins > 200 kb (21%; figure 3.5 main text).

A.1.2. qPCR validation results for GR-response eQTLs

HIST2H2AA3/4 was the array probe with the most eSNPs overlapping with the meta-analysis results for MDD. Two transcript variants encoding isoforms with a different 3'UTR length have been identified for *HIST2H2AA3, HIST2H2AA4*. The shorter gene product (isoform 1) is annotated by RefSeq while the alternatively spliced longer gene product (isoform 2) is annotated by Ensembl release 54 (HIST2H2AA3-001; ENST00000369161) and further predicted by AceView (HIST2H2AA3.aApr07-unspliced, HIST2H2AA4.aApr07-unspliced). This longer isoform is tagged by the significant Illumina probe (ILMN_1695435). Hence we designed two different assays- one covering the common part of both isoforms (assay 1) and the other tagging isoform 2 (assay 2). The expression levels measured with both assays were highly correlated (Spearman's test P value $< 1.5 \times 10^{-20}$, R = 74%). We could replicate a significant SNP effect in 137 samples with a P value of 0.012 using assay 1 with a genotypic model and P = 0.017 using a carrier model, with the same direction of change as in the expression array.

A.2. Supplementary Figures

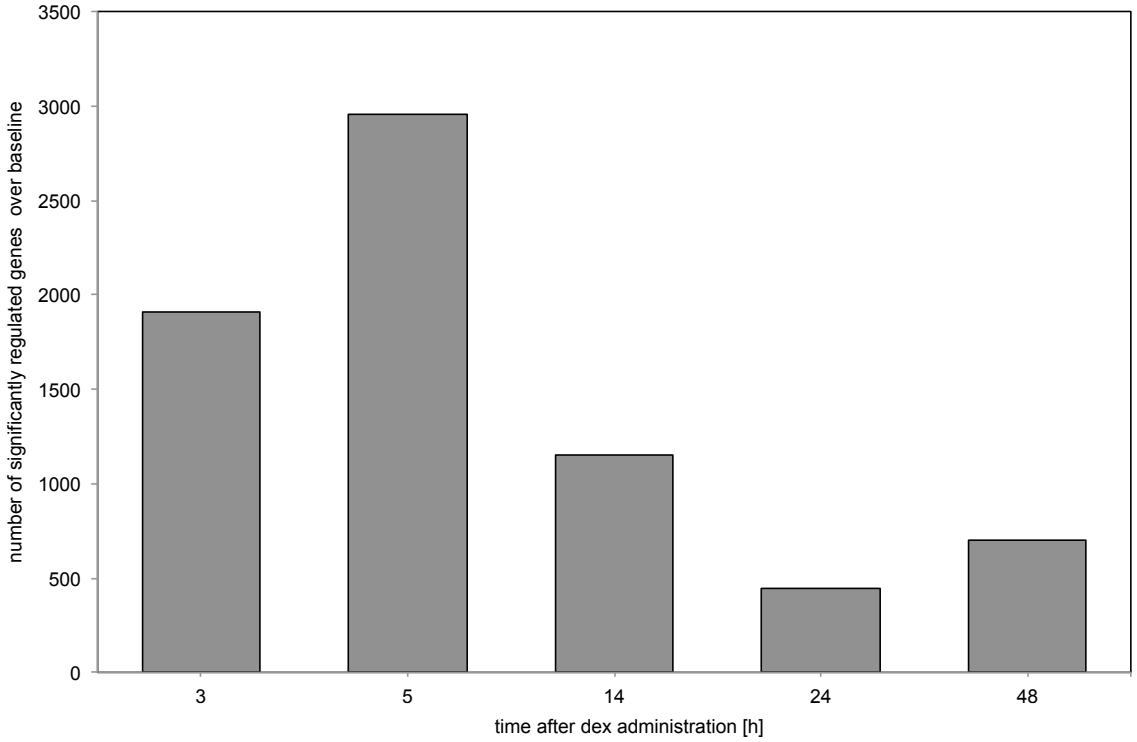
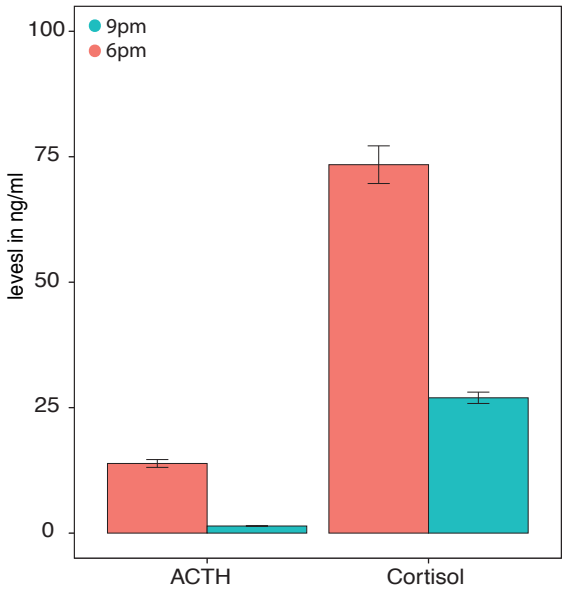
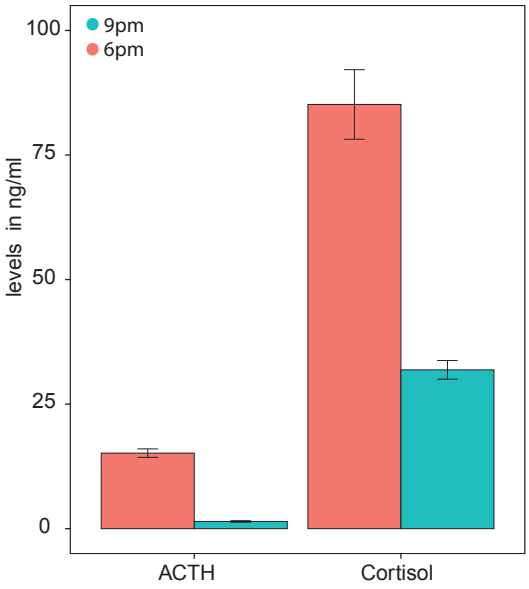


Figure A.1.: The number of genes that are differently expressed at several time points after administration of 1.5 mg dexamethasone relative to baseline in 4 healthy male individuals are shown. The height of the bars indicates the total number of transcripts with nominally significant changes from baseline gene expression. Baseline blood samples were obtained at 6pm. This evening time point was chosen so that the stimulation experiments took place during the quiescent period of the stress hormone system. Baseline blood draws were immediately followed by oral administration of dexamethasone. Additional blood samples were drawn at 9pm, 11pm the same day, 8am, 6pm the next day and 6pm at day 3. A comparison of baseline gene expression vs. gene expression after 3, 5, 14, 24 and 48 h shows an initial high number of gene expression changes, followed by a normalization within 24-48 hours. The highest number of differently expressed genes (highest bar in chart) was observed at 3 and 5 hours after dexamethasone ingestion. For practical reasons as well as to avoid secondary GR target effects, in the subsequent experiment we collected blood 3 hours after dexamethasone intake.



(a) Controls



(b) MDD patiens

Figure A.2.: Administration of dexamethasone resulted in a robust suppression of cortisol in all individuals. Cortisol levels were significantly suppressed in healthy controls (a; $F_{1,90} = 89.74, P = 3.57 \times 10^{-15}$) as well as in depressed patients (b; $F_{1,67} = 7.09, P = 0.0097$) 3h after dexamethasone challenge. Similar results were observed for ACTH, with a significant reduction in ACTH levels in healthy controls (a; $F_{1,91} = 43.96, P = 2.33 \times 10^{-9}$) and in depressed patients (b; $F_{1,65} = 9.75, P = 0.0027$) after 3h.

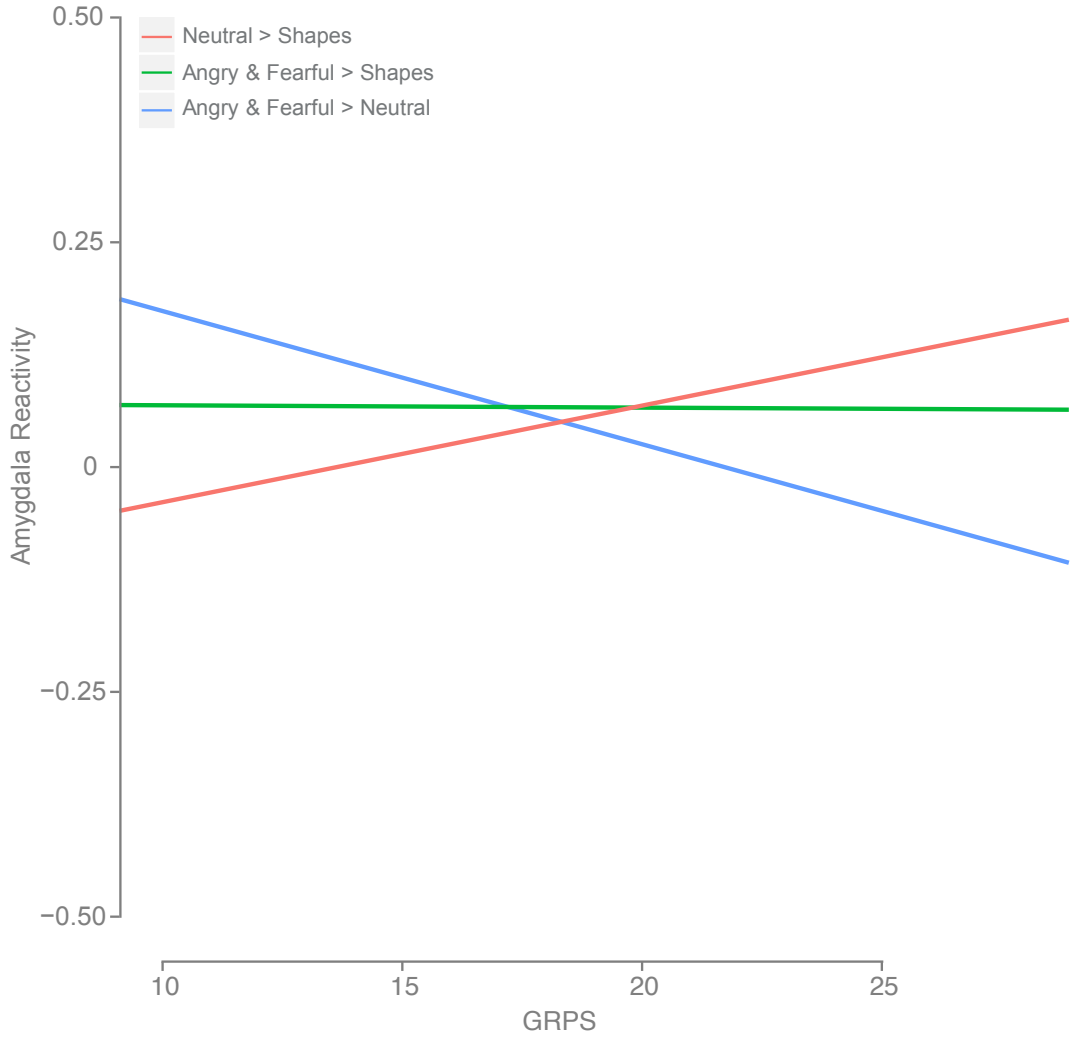
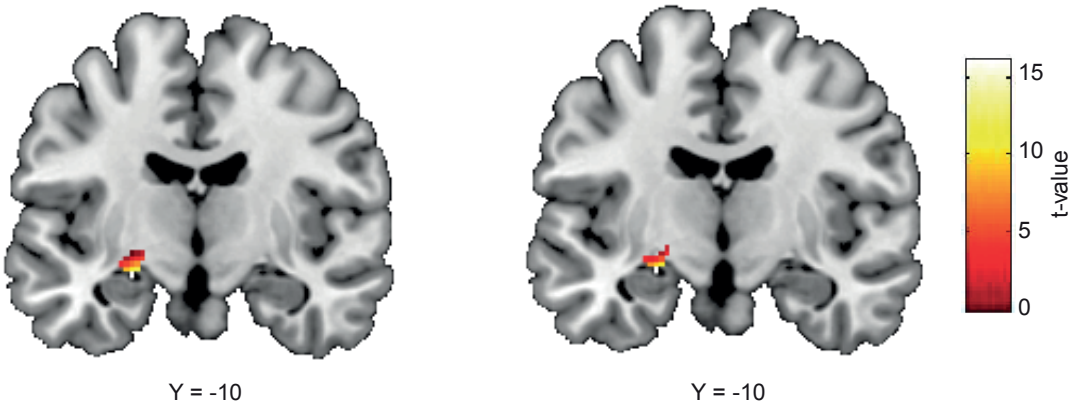


Figure A.3.: Elevated GRPSs correlate with dysfunctional amygdala reactivity in the entire DNS sample ($n = 647$). As was found in the European-American subsample, elevated GRPSs predicted blunted amygdala reactivity to threat-related expressions in comparison to neutral expressions in the entire sample when controlling for patterns of population stratification. Post-hoc analyses revealed that GRPS was not predictive of reactivity to threat-related expressions, but that higher GRPSs predicted elevated amygdala reactivity to neutral expressions, in comparison to non-face control stimuli.



(a) Angry & Fearful > Shape

(b) Neutral > Shape

Figure A.4.: Statistical parametric map illustrating amygdala reactivity. Main effects of post hoc contrasts for left centromedial amygdala reactivity used in imaging genetics analyses of GRPS in the entire sample. (a) "Angry & Fearful > Shape" (49 contiguous voxels; max voxel MNI coordinate, $x = -24$, $y = -10$, $z = -14$, $t = 22.59$, $P < 4.41 \times 10^{-16}$), and (b) "Neutral > Shape" (35 contiguous voxels; max voxel MNI coordinate, $x = -24$, $y = -10$, $z = -14$, $t = 10.73$, $P < 4.41 \times 10^{-16}$).

A.3. Supplementary Tables

Table A.1.: List of the 320 *cis*-eSNP bin-probe combinations (eQTL bins).

tag SNP	Ntag[1]	Genes nearby tag SNP	SNP Location	Chr.[2]	Probe id[3]	Probe gene	Q values[4]	FC[5]
rs9994839	0	*SPARCL1*	downstream	4	ILMN_1651354	*SPP1*	0.04493	1.14
rs3766746	7	*PLOD1*	intronic	1	ILMN_1651385	*MFN2*	0.04061	1.28
rs2295281	21	*MFN2*	intronic	1	ILMN_1651385	*MFN2*	0.00026	1.28
rs2460432	188	*ASL*	intronic	7	ILMN_1651950	*TPST1*	0.03325	2.3
rs2848122	0	*ANKRD36BP2,MIR4436A*	intergenic	2	ILMN_1652199	*IGKC*	0.04686	-1.16
rs12417156	1	*CHRDL2*	intronic	11	ILMN_1652753	*PAAF1*	0.01664	-1.17
rs17586305	8	*D21S2088E,LOC339622*	intergenic	21	ILMN_1653667	*TBX1*	0.03325	1.17
rs2828337	2	*D21S2088E,LOC339622*	intergenic	21	ILMN_1653667	*TBX1*	0.04686	1.17
rs231478	11	*MPP2*	intronic	17	ILMN_1653711	*FZD2*	0.01755	-1.22
rs4777959	2	*SLCO3A1,ST8SIA2*	intergenic	15	ILMN_1654735	*SLCO3A1*	0.02471	1.23
rs3865451	0	*ADCK4*	intronic	19	ILMN_1654875	*CLC*	0.04911	-1.36
rs3853657	0	*MIR4456,CEP72*	intergenic	5	ILMN_1655195	*SMA4*	0.03563	-1.17
rs7622109	27	*TCAIM*	intronic	3	ILMN_1655702	*ABHD5*	0.02774	1.51
rs7915524	0	*FAM171A1*	intronic	10	ILMN_1656378	*NMT2*	0.03563	-1.24
rs7256770	3	*ACP5*	upstream	19	ILMN_1656822	*DNM2*	0.04579	1.18
15-20476475	1	*CYFIP1*	intronic	15	ILMN_1657478	*MAGEL2*	0.04686	1.09
rs11227523	3	*ZDHHC24*	downstream	11	ILMN_1657701	*RBM4*	0.04735	-1.15
rs8082593	1	*UTP18,CA10*	intergenic	17	ILMN_1657884	*NME1*	0.04395	-1.1
rs3802984	0	*ODF3*	exonic	11	ILMN_1657932	*MUC6*	0.04861	1.11
rs6482235	23	*PIP4K2A,ARMC3*	intergenic	10	ILMN_1657977	*MSRB2*	0.03563	1.31
rs12611262	0	*SEMA6B,TNFAIP8L1*	intergenic	19	ILMN_1658486	*MRPL54*	0.04579	-1.24
rs524908	4	*FRMD5*	intronic	15	ILMN_1658743	*CCNDBP1*	0.04136	1.13
rs417557	0	*CLDN14,SIM2*	intergenic	21	ILMN_1661194	*CLDN14*	0.04686	1.14
rs12620091	0	*ALMS1P*	ncRNA_intronic	2	ILMN_1662954	*CCT7*	0.04686	-1.17
rs2282444	4	*TOMM6,USP49*	intergenic	6	ILMN_1663489	*UBR2*	0.04729	1.15
rs2269799	0	*SV2B*	intronic	15	ILMN_1663699	*SLCO3A1*	0.04686	1.11
rs698915	6	*RPRD2*	intronic	1	ILMN_1664706	*HIST2H3D*	0.03386	1.13
rs877836	0	*UQCRFS1,LOC284395*	intergenic	19	ILMN_1664920	*C19orf12*	0.0498	-1.12
rs12497322	1	*XIRP1,CX3CR1*	intergenic	3	ILMN_1665148	*RPSA*	0.04442	-1.36
rs10505733	6	*CLEC4C*	intronic	12	ILMN_1665457	*CLEC4C*	0.00021	1.19
rs4242902	5	*DPPA3,CLEC4C*	intergenic	12	ILMN_1665457	*CLEC4C*	0.0299	1.19
rs11055463	8	*DPPA3,CLEC4C*	intergenic	12	ILMN_1665457	*CLEC4C*	0.00164	1.19
rs11123840	8	*PDCL3,NPAS2*	intergenic	2	ILMN_1665877	*RNF149*	0.01182	1.42
rs4887017	0	*ACSBG1*	intronic	15	ILMN_1665887	*WDR61*	0.04395	-1.21
rs300035	0	*FOXL1,C16orf95*	intergenic	16	ILMN_1666594	*IRF8*	0.04493	-1.06
rs1893233	0	*PIEZO2,GNAL*	intergenic	18	ILMN_1667744	*MPPE1*	0.04196	1.19
rs7535902	0	*LIN28A*	intronic	1	ILMN_1668270	*ZDHHC18*	0.04729	1.34

[1]number of SNP within an eSNP bin
[2]eQTL chromosome
[3]llumina probe identifier (Human HT-12 v3)
[4]FDR analog of the P value
[5]fold change of GR-stimulated/baseline gene expression

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs10234768	7	*C7orf72*	intronic	7	ILMN_1669617	*GRB10*	0.04442	1.14
rs2178779	0	*OR4E2,DAD1*	intergenic	14	ILMN_1670272	*LRP10*	0.02929	1.31
rs6531673	2	*TLR6*	intronic	4	ILMN_1670931	*PDS5A*	0.03758	-1.1
rs12443981	0	*SEPHS2,ITGAL*	intergenic	16	ILMN_1671854	*ZNF48*	0.03758	-1.18
rs36074193	8	*TOB1-AS1,SPAG9*	intergenic	17	ILMN_1672004	*TOB1*	0.00198	1.46
rs885950	0	*POU5F1,PSORS1C3*	intergenic	6	ILMN_1673753	*ABCF1*	0.04488	-1.12
rs7955208	12	*PTPN11,RPH3A*	intergenic	12	ILMN_1674063	*OAS2*	0.01742	-1.25
rs2387976	17	*NANP*	intronic	20	ILMN_1674394	*C20orf3*	0.01664	1.37
3-43101876	17	*GTDC2*	intronic	3	ILMN_1674522	*HIGD1A*	0.04395	-1.07
rs12981801	1	*ZNF554*	downstream	19	ILMN_1674926	*C19orf35*	0.04686	1.37
rs60492890	1	*C19orf24*	upstream	19	ILMN_1674926	*C19orf35*	0.02774	1.37
rs4902681	0	*ACTN1*	intronic	14	ILMN_1675448	*ZFP36L1*	0.04729	-1.06
rs4767092	0	*LHX5,RBM19*	intergenic	12	ILMN_1675640	*OAS1*	0.04856	-1.16
rs10939637	9	*SLC2A9*	intronic	4	ILMN_1675844	*WDR1*	0.04727	1.18
rs9320357	0	*GSTM2P1,SLC16A10*	intergenic	6	ILMN_1676891	*CDC2L6*	0.04686	1.25
rs1562782	1	*ADM,AMPD3*	intergenic	11	ILMN_1678004	*TMEM41B*	0.04442	-1.06
rs2422008	3	*WDPCP*	intronic	2	ILMN_1679268	*PELI1*	0.04204	1.68
rs724208	3	*OLIG1,C21orf54*	intergenic	21	ILMN_1679476	*GART*	0.04493	-1.19
rs1532445	0	*LZTS1,LZTS1-AS1*	intergenic	8	ILMN_1679483	*INTS10*	0.04442	-1.11
rs4253082	4	*ERCC6*	intronic	10	ILMN_1679555	*TIMM23*	0.04579	-1.21
rs914314	62	*SYNDIG1,CST7*	intergenic	20	ILMN_1679826	*CST7*	0.03386	1.14
rs2296887	2	*GBF1*	UTR5	10	ILMN_1682165	*NT5C2*	0.04061	1.23
rs11055463	8	*DPPA3,CLEC4C*	intergenic	12	ILMN_1682259	*CLEC4C*	0.04395	1.12
rs10505733	6	*CLEC4C*	intronic	12	ILMN_1682259	*CLEC4C*	0.00021	1.12
rs11760934	0	*POLR2J2,FAM185A*	intergenic	7	ILMN_1682368	*LRWD1*	0.04729	1.26
rs59464952	1	*NLRP1,LOC339166*	intergenic	17	ILMN_1682761	*C17orf87*	0.04488	-1.18
rs10881678	1	*KIF20B,LINC00865*	intergenic	10	ILMN_1682799	*STAMBPL1*	0.04925	-1.21
rs3761821	13	*OBP2B,ABO*	intergenic	9	ILMN_1683498	*RPL7A*	0.03872	-1.16
rs17340646	2	*TPI1P2,LOC407835*	intergenic	7	ILMN_1683811	*TNPO3*	0.03386	1.18
rs1001073	0	*MASP1*	intronic	3	ILMN_1685722	*EIF4A2*	0.04442	-1.26
rs7099954	2	*ITIH2*	intronic	10	ILMN_1685774	*ATP5C1*	0.04061	-1.26
rs4924398	9	*GPR176*	intronic	15	ILMN_1686116	*THBS1*	0.04061	1.24
rs11638679	0	*C15orf54,THBS1*	intergenic	15	ILMN_1686116	*THBS1*	0.04715	1.24
rs5756520	5	*TMPRSS6,IL2RB*	intergenic	22	ILMN_1687306	*LGALS2*	0.04729	-1.18
rs4075428	7	*CIZ1*	intronic	9	ILMN_1687857	*ST6GALNAC4*	0.04686	1.07
rs2524379	4	*EMR2*	intronic	19	ILMN_1688152	*IL27RA*	0.04686	-1.2
rs7194275	0	*C16orf91,CCDC154*	intergenic	16	ILMN_1688749	*RPS2*	0.04856	-1.22
rs11645488	2	*FLJ26245*	intergenic	16	ILMN_1689327	*CR617556*	0.03563	-1.05
rs12335026	3	*FABP12,IMPA1*	intergenic	8	ILMN_1690586	*DQ579214*	0.03073	-1.28
rs2280516	0	*DRC1*	intronic	2	ILMN_1691090	*MPV17*	0.02446	-1.18
rs4147779	0	*NDUFS8*	downstream	11	ILMN_1692026	*SUV420H1*	0.04911	1.15
rs13285411	0	*DNM1*	intronic	9	ILMN_1692223	*LCN2*	0.04686	1.16
rs2359795	8	*JDP2*	intronic	14	ILMN_1694233	*ACYP1*	0.04061	-1.14
rs11672145	0	*ZNF799*	downstream	19	ILMN_1694325	*NFIX*	0.02607	1.06
rs917585	2	*SLC6A7*	intronic	5	ILMN_1694686	*HMGXB3*	0.04488	-1.14
rs4968392	2	*VMP1*	intronic	17	ILMN_1695157	*CA4*	0.04488	1.7

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs7755418	2	*FAM50B,PRPF4B*	intergenic	6	ILMN_1695311	*HLA-DMA*	0.04856	-1.37
1-148440425	199	*PLEKHO1,ANP32E*	intergenic	1	ILMN_1695435	*HIST2H2AA3/4*	0.00581	1.22
1-155229343	56	*ARHGEF11*	intronic	1	ILMN_1695576	*MRPL24*	0.04061	-1.25
rs12372446	0	*KRT71,KRT74*	intergenic	12	ILMN_1695812	*KRT72*	0.04493	-1.02
rs3771863	0	*TACR1*	intronic	2	ILMN_1696375	*TTC31*	0.04493	-1.16
rs6651024	1	*C6orf10*	intronic	6	ILMN_1697499	*HLA-DRB5*	0.04061	-1.07
rs1265758	13	*C6orf10*	intronic	6	ILMN_1697499	*HLA-DRB5*	0.04387	-1.07
rs9268926	196	*HLA-DRA,HLA-DRB5*	intergenic	6	ILMN_1697499	*HLA-DRB5*	0.01205	-1.07
rs9268671	392	*HLA-DRA,HLA-DRB5*	intergenic	6	ILMN_1697499	*HLA-DRB5*	0.00021	-1.07
rs11246074	0	*IFITM3,B4GALNT4*	intergenic	11	ILMN_1698519	*AL137655*	0.04625	1.25
rs10002500	8	*CNGA1*	intronic	4	ILMN_1700306	*OCIAD2*	0.02446	-1.32
rs17108932	3	*PTEN,RNLS*	intergenic	10	ILMN_1701134	*PTEN*	0.03563	1.56
rs2473263	4	*WNT4,ZBTB40*	intergenic	1	ILMN_1701603	*ALPL*	0.04398	1.42
rs10487531	0	*LOC730441,MTRNR2L6*	intergenic	7	ILMN_1701875	*ZYX*	0.04488	1.23
rs55776343	0	*SLC41A3*	intronic	3	ILMN_1702055	*ROPN1B*	0.04488	1.13
rs2305160	1	*NPAS2*	exonic	2	ILMN_1702806	*PDCL3*	0.04488	-1.15
rs3760352	0	*ASGR2,ASGR1*	intergenic	17	ILMN_1703433	*PLSCR3*	0.02607	-1.26
rs6070412	42	*PPP4R1L*	ncRNA_intronic	20	ILMN_1704079	*RBM38*	0.03563	1.11
rs35406858	8	*RPP25,SCAMP5*	intergenic	15	ILMN_1704477	*COX5A*	0.0384	-1.27
rs1873625	1	*BSN*	intronic	3	ILMN_1705737	*IMPDH2*	0.048	-1.24
rs4976450	12	*SPOCK1*	intronic	5	ILMN_1706539	*KDM3B*	0.04488	1.13
rs6439982	14	*SPSB4,ACPL2*	intergenic	3	ILMN_1706598	*ACPL2*	0.01664	1.25
rs9418982	2	*LOC619207,CYP2E1*	intergenic	10	ILMN_1708348	*ADAM8*	0.04686	1.24
rs717450	0	*XCL1*	downstream	1	ILMN_1709233	*F5*	0.04442	1.31
rs57988909	1	*OR5H15,OR5H6*	intergenic	3	ILMN_1710326	*CLDND1*	0.04442	-1.23
rs12497322	1	*XIRP1,CX3CR1*	intergenic	3	ILMN_1710885	*RPSA*	0.03872	-1.34
rs10489832	0	*OR10R2,OR6Y1*	intergenic	1	ILMN_1710937	*IFI16*	0.04686	1.29
rs8041381	0	*RORA*	intronic	15	ILMN_1711899	*ANXA2*	0.04811	-1.14
rs1647990	3	*RORA*	intronic	15	ILMN_1711899	*ANXA2*	0.03872	-1.14
rs624420	0	*CPT1A,MRPL21*	intergenic	11	ILMN_1711994	*TCIRG1*	0.04861	1.3
rs16858988	1	*GAD1*	intronic	2	ILMN_1712305	*CYBRD1*	0.03386	1.13
rs6749185	5	*CYBRD1*	intronic	2	ILMN_1712305	*CYBRD1*	0.03872	1.13
rs6433294	11	*DCAF17,CYBRD1*	intergenic	2	ILMN_1712305	*CYBRD1*	0.01664	1.13
rs166211	16	*LOC283867,CDH5*	intergenic	16	ILMN_1712389	*CKLF*	0.04294	1.3
rs7349097	0	*PTPRF*	intronic	1	ILMN_1714445	*SLC6A9*	0.04488	1.09
rs7826635	1	*CHMP7,R3HCC1*	intergenic	8	ILMN_1715969	*SLC25A37*	0.02774	1.24
rs809972	1	*MIR34A,H6PD*	intergenic	1	ILMN_1716465	*RBP7*	0.02504	1.33
rs2027237	0	*LOC728228*	downstream	20	ILMN_1717809	*RNF24*	0.04061	1.2
rs250145	0	*MAF,MIR548H4*	intergenic	16	ILMN_1719543	*MAF*	0.04493	-1.18
rs17031905	85	*INPP4A*	intronic	2	ILMN_1719756	*ZAP70*	0.04294	-1.2
19-40883657	5	*UPK1A,ZBTB32*	intergenic	19	ILMN_1720542	*POLR2I*	0.04442	-1.18
rs2395891	3	*BTBD2,MKNK2*	intergenic	19	ILMN_1721344	*MOB3A*	0.02446	1.15
rs11644259	15	*BRD7*	intronic	16	ILMN_1721349	*MAGT1*	0.04061	1.11
rs2363536	4	*CD53*	intronic	1	ILMN_1721989	*ATP5F1*	0.04196	-1.13
rs1981294	4	*LRIF1,DRAM2*	intergenic	1	ILMN_1721989	*ATP5F1*	0.03653	-1.13
rs3760352	0	*ASGR2,ASGR1*	intergenic	17	ILMN_1722900	*EIF4A1*	0.04061	-1.2

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs2730355	3	*GALNT15*	intronic	3	ILMN_1723414	*HACL1*	0.0299	-1.17
rs12423255	0	*PITPNM2*	upstream	12	ILMN_1725187	*PITPNM2*	0.04061	1.17
rs538645	0	*DDX6,CXCR5*	intergenic	11	ILMN_1726306	*HMBS*	0.04395	1.03
rs9525211	46	*RASA3*	intronic	13	ILMN_1727389	*CDC16*	0.04061	-1.14
rs2161343	0	*FLJ38109*	ncRNA_intronic	5	ILMN_1728742	*C5orf4*	0.04398	1.13
rs3826440	11	*POLR2A*	intronic	17	ILMN_1731546	*RPL26*	0.04395	-1.25
rs1859441	89	*COL2A1,SENP1*	intergenic	12	ILMN_1731666	*ZNF641*	0.03325	1.05
rs4075428	7	*CIZ1*	intronic	9	ILMN_1732049	*DPM2*	0.02607	1.12
rs1559155	0	*PPFIA3*	intronic	19	ILMN_1732053	*SNRNP70*	0.04856	-1.12
rs7544118	1	*ADORA3*	intronic	1	ILMN_1733259	*ADORA3*	0.00164	2.08
rs10906402	0	*LOC399715,PRKCQ*	intergenic	10	ILMN_1733421	*PRKCQ*	0.04488	-1.2
rs35406858	8	*RPP25,SCAMP5*	intergenic	15	ILMN_1733696	*IMP3*	0.04186	-1.27
1-20817476	5	*CDA*	intronic	1	ILMN_1734231	*DDOST*	0.04856	-1.19
rs11083620	0	*C19orf69*	upstream	19	ILMN_1734878	*CD79A*	0.04061	-1.3
rs10039049	1	*ANXA6*	intronic	5	ILMN_1736567	*CD74*	0.04686	-1.25
rs11760186	18	*PHACTR1*	intronic	6	ILMN_1736982	*PHACTR1*	0.005	1.19
rs1228529	5	*PHACTR1*	intronic	6	ILMN_1736982	*PHACTR1*	0.03872	1.19
rs2420147	0	*LYRM7,CDC42SE2*	intergenic	5	ILMN_1737343	*FNIP1*	0.04847	1.25
rs10835861	3	*RCN1,WT1*	intergenic	11	ILMN_1737806	*BX648962*	0.04686	1.05
rs12216600	0	*PDE1C*	intronic	7	ILMN_1737947	*LSM5*	0.04686	-1.2
rs12206258	2	*GMDS*	intronic	6	ILMN_1738401	*FOXC1*	0.02366	1.42
rs12128782	3	*TBX19,MIR557*	intergenic	1	ILMN_1739103	*MPZL1*	0.04061	1.24
rs6709463	4	*FAM117B*	intronic	2	ILMN_1739942	*FAM117B*	0.03728	-1.07
rs2812500	0	*C10orf35,COL13A1*	intergenic	10	ILMN_1740633	*PRF1*	0.04442	1.09
rs7556661	91	*ARNT,SETDB1*	intergenic	1	ILMN_1741200	*RFX5*	0.03563	-1.3
rs2209313	15	*SIRPB1*	intronic	20	ILMN_1742442	*SIRPB1*	0.00043	1.15
rs4845143	9	*IL19*	intronic	1	ILMN_1742601	*CR1*	0.04847	1.3
rs12438495	0	*IGF1R*	intronic	15	ILMN_1744023	*IGF1R*	0.04686	1.17
rs2672027	1	*MIR4456,CEP72*	intergenic	5	ILMN_1744210	*SDHA*	0.04387	-1.16
rs11078835	0	*GAS7*	intronic	17	ILMN_1745994	*GAS7*	0.04488	1.14
rs10790231	23	*TMPRSS4*	intronic	11	ILMN_1746516	*RPS25*	0.01596	-1.25
rs4442562	1	*FOXR1*	intronic	11	ILMN_1746516	*RPS25*	0.04686	-1.25
rs2956993	0	*GANAB*	intronic	11	ILMN_1746525	*FTH1*	0.04395	1.11
17-8005504	0	*VAMP2*	intronic	17	ILMN_1746883	*SSAT2*	0.04395	-1.17
rs6607302	0	*HNF1B,LOC284100*	intergenic	17	ILMN_1748651	*PSMB3*	0.03563	1.12
rs2385067	0	*TMEM104*	intronic	17	ILMN_1748797	*GRB2*	0.04061	1.21
rs17304079	5	*RBM6*	intronic	3	ILMN_1749662	*GPX1*	0.04442	1.09
rs4889991	0	*CARD14*	intronic	17	ILMN_1749722	*RNF213*	0.03728	1.13
rs1317577	2	*N4BP2,RHOH*	intergenic	4	ILMN_1750507	*RPL9*	0.02774	-1.41
rs6581076	1	*OR10P1,METTL7B*	intergenic	12	ILMN_1750636	*RPS26*	0.03563	-1.21
rs639459	2	*C7orf25*	upstream	7	ILMN_1751051	*C7orf25*	0.04686	1.16
rs11672145	0	*ZNF799*	downstream	19	ILMN_1751571	*RAD23A*	0.03981	-1.05
rs9526443	15	*MED4*	intronic	13	ILMN_1751708	*ITM2B*	0.04442	1.27
rs4990638	0	*TMEM132E,CCT6B*	intergenic	17	ILMN_1752520	*SLFN11*	0.03364	-1.2
rs2568032	1	*ST5*	intronic	11	ILMN_1752988	*C11orf17*	0.04488	-1.18
rs7089504	0	*PRKCQ*	intronic	10	ILMN_1754178	*GDI2*	0.04686	-1.08

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs11264449	1	*SEMA4A*	intronic	1	ILMN_1755123	*GBA*	0.04488	1.19
rs843631	0	*ACYP2,C2orf73*	intergenic	2	ILMN_1755883	*RPS27A*	0.04861	-1.36
rs492799	0	*NAALADL1,CDCA5*	intergenic	11	ILMN_1756204	*RPS6KA4*	0.02774	-1.17
rs2567342	3	*BDH1*	intronic	3	ILMN_1756360	*RPL35A*	0.04442	-1.22
rs200891	0	*LOC100289473,SIRPA*	intergenic	20	ILMN_1758146	*SIRPA*	0.04442	1.4
rs8047140	0	*KCTD13*	intronic	16	ILMN_1759008	*ZNF689*	0.04196	-1.24
rs734570	0	*SNRNP70*	intronic	19	ILMN_1759436	*NOSIP*	0.03121	-1.24
rs11238359	7	*EGFR,LANCL2*	intergenic	7	ILMN_1760338	*SUMO2*	0.04061	-1.1
rs213637	0	*PAQR7*	UTR3	1	ILMN_1760556	*BC041843*	0.048	1.12
rs6545924	0	*COMMD1,B3GNT2*	intergenic	2	ILMN_1761242	*COMMD1*	0.04488	-1.22
rs3793243	0	*STX1A*	intronic	7	ILMN_1761387		0.04061	1.31
rs444297	0	*LIMK1,EIF4H*	intergenic	7	ILMN_1761387		0.04395	1.31
rs1352312	1	*MAP3K14*	intronic	17	ILMN_1762678	*NMT1*	0.048	-1.13
rs10781518	6	*SDCCAG3*	intronic	9	ILMN_1764239	*PMPCA*	0.04488	-1.12
rs59562633	4	*LOC100128714*	ncRNA_intronic	15	ILMN_1764549	*UBE3A*	0.04395	-1.12
rs17178720	9	*UGGT1*	exonic	2	ILMN_1765122	*MAP3K2*	0.04442	1.17
rs1379868	2	*NRTN*	intronic	19	ILMN_1766125	*LONP1*	0.02446	-1.19
rs7252014	0	*KCNN1*	intronic	19	ILMN_1766487	*LRRC25*	0.0384	1.1
rs4845143	9	*IL19*	intronic	1	ILMN_1767193	*CR1*	0.04061	1.22
rs3015983	0	*PAK1*	intronic	11	ILMN_1767365	*PAK1*	0.04686	1.22
rs7870685	5	*LOC401497,ACO1*	intergenic	9	ILMN_1767980	*LOC401497*	0.02179	1.05
rs17034661	0	*VGLL4*	intronic	3	ILMN_1768480	*VGLL4*	0.04811	-1.16
rs12441390	4	*RASGRP1,C15orf53*	intergenic	15	ILMN_1768958	*RASGRP1*	0.04493	-1.17
rs698915	6	*RPRD2*	intronic	1	ILMN_1769027	*CDC42SE1*	0.04395	1.22
rs11707455	8	*BBX*	intronic	3	ILMN_1771333	*CD47*	0.02504	-1.17
rs7673908	50	*CLOCK*	intronic	4	ILMN_1773760	*PAICS*	0.04493	-1.19
rs6749185	5	*CYBRD1*	intronic	2	ILMN_1773847	*DYNC1I2*	0.04686	-1.16
rs1007122	5	*COMMD7,DNMT3B*	intergenic	20	ILMN_1774250	*PLUNC*	0.04686	1.06
rs4309551	1	*RRM2,C2orf48*	intergenic	2	ILMN_1775011	*NOL10*	0.01664	1.1
rs60157471	0	*IL36B*	intronic	2	ILMN_1775501	*IL1B*	0.04727	-1.47
rs13090	0	*MED16*	exonic	19	ILMN_1777190	*CFD*	0.04395	1.11
rs12548608	6	*KIF13B*	intronic	8	ILMN_1778226	*EXTL3*	0.03563	1.44
rs6493387	3	*TRPM1*	intronic	15	ILMN_1778734	*MTMR15*	0.04488	-1.16
rs12886153	0	*KTN1,RPL13AP3*	intergenic	14	ILMN_1780132	*PELI2*	0.03653	1.17
rs12891572	4	*HNRNPC*	upstream	14	ILMN_1780533	*RNASE6*	0.04204	1.4
rs12433896	5	*RNASE4,EDDM3A*	intergenic	14	ILMN_1780533	*RNASE6*	0.01755	1.4
rs6543137	1	*IL18RAP*	intronic	2	ILMN_1781700	*IL18R1*	0.01742	1.99
rs4845391	0	*KCNN3*	intronic	1	ILMN_1782057	*ATP8B2*	0.04686	-1.21
rs9525211	46	*RASA3*	intronic	13	ILMN_1782292	*LAMP1*	0.01755	1.12
rs6904470	0	*TAAR5,TAAR3*	intergenic	6	ILMN_1782621	*RPS12*	0.04442	-1.09
rs91710	0	*ARRDC2,IL12RB1*	intergenic	19	ILMN_1782977	*UBA52*	0.04671	1.08
rs17654580	0	*ARRDC2,IL12RB1*	intergenic	19	ILMN_1782977	*UBA52*	0.04398	1.08
rs4452682	0	*SLC22A23*	intronic	6	ILMN_1783158	*LY6G6F*	0.04856	1.12
rs11249756	4	*BTNL3,BTNL9*	intergenic	5	ILMN_1783795	*BTNL3*	0.00581	1.16
rs2329844	9	*TSPEAR*	intronic	21	ILMN_1785179	*UBE2G2*	0.03325	-1.25
rs57057834	53	*SLC19A1,LOC100129027*	intergenic	21	ILMN_1785179	*UBE2G2*	0.04729	-1.25
rs17280306	40	*ZNF621,CTNNB1*	intergenic	3	ILMN_1786242	*RPL14*	0.04442	-1.32

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs10885031	1	*RBM20*	intronic	10	ILMN_1787378	*ADD3*	0.04488	-1.05
rs10180924	5	*ATOH8,LOC284950*	intergenic	2	ILMN_1790692	*GNLY*	0.01664	1.02
rs2856728	0	*ELN*	intronic	7	ILMN_1791375	*STAG3L2*	0.04686	-1.25
rs17173596	28	*GIMAP1-GIMAP5,TMEM176B*	intergenic	7	ILMN_1791511	*TMEM176A*	0.01749	-1.15
rs2072443	0	*TMEM176B*	exonic	7	ILMN_1791511	*TMEM176A*	0.03563	-1.15
rs2388881	0	*MCTP2,LOC440311*	intergenic	15	ILMN_1792682	*MCTP2*	0.04442	1.29
rs2938387	7	*PPARG*	intronic	3	ILMN_1793724	*C3orf31*	0.03563	-1.18
rs760657	0	*BPIFC,FBXO7*	intergenic	22	ILMN_1793934	*PISD*	0.03872	1.35
rs1280984	5	*CASZ1,C1orf127*	intergenic	1	ILMN_1794165	*PGD*	0.01755	1.29
rs3825073	0	*SYVN1*	intronic	11	ILMN_1794364	*CTSW*	0.04395	-1.05
rs11176799	0	*CAND1,DYRK2*	intergenic	12	ILMN_1794588	*DYRK2*	0.03758	-1.32
rs3758587	0	*ARHGAP19-SLIT1*	ncRNA_intronic	10	ILMN_1794914	*UBTD1*	0.02929	1.15
rs10746914	4	*ANXA1,RORB*	intergenic	9	ILMN_1795228	*ZFAND5*	0.04387	1.07
rs7071536	2	*ANKRD16*	downstream	10	ILMN_1795467	*LOC399715*	0.01755	1.19
rs4433629	2	*LOC338758,LINC00615*	intergenic	12	ILMN_1795835	*LOC338758*	0.04488	-1.16
rs1423738	1	*HS3ST4,C16orf82*	intergenic	16	ILMN_1798204	*IL21R*	0.04488	-1.16
10-44400544	0	*CXCL12,TMEM72-AS1*	intergenic	10	ILMN_1798533	*ZNF22*	0.04488	-1.16
rs10931765	2	*PGAP1,ANKRD44*	intergenic	2	ILMN_1798543	*STK17B*	0.04121	1.42
rs158391	2	*ZNF33B,BMS1*	intergenic	10	ILMN_1799208	*CSGALNACT2*	0.03758	1.23
rs2234768	5	*ACTA2*	intronic	10	ILMN_1799848	*ANKRD22*	0.04579	1.43
rs1610037	0	*ADCYAP1*	UTR3	18	ILMN_1803676	*ENOSF1*	0.01807	-1.14
rs42931	1	*GAL3ST1*	intronic	22	ILMN_1803925	*MTMR3*	0.04061	1.26
rs12705071	0	*ZNF3,COPS6*	intergenic	7	ILMN_1804530	*ARPC1B*	0.04856	1.08
rs4795402	0	*ORMDL3,LRRC3C*	intergenic	17	ILMN_1805636	*PERLD1*	0.04488	-1.14
19-18810229	8	*UPF1*	intronic	19	ILMN_1805693	*GMIP*	0.04442	1.21
5-50827997	7	*ISL1,PELO*	intergenic	5	ILMN_1806651	*PARP8*	0.04686	1.25
rs9931197	31	*SCNN1G,SCNN1B*	intergenic	16	ILMN_1806908	*PRKCB*	0.04398	1.14
rs55678304	3	*RAB3A,PDE4C*	intergenic	19	ILMN_1807277	*IFI30*	0.04493	-1.18
rs2749883	3	*NID2*	intronic	14	ILMN_1807925	*GNG2*	0.048	1.04
rs5752890	2	*EMID1*	intronic	22	ILMN_1809433	*XBP1*	0.03563	-1.17
rs4985671	0	*LOC339166,WSCD1*	intergenic	17	ILMN_1810045	*NLRP1*	0.03563	1.16
rs12432242	5	*SLC7A7*	intronic	14	ILMN_1810275	*SLC7A7*	0.04061	-1.01
rs9873175	0	*RPL29,DUSP7*	intergenic	3	ILMN_1811063	*RPL29*	0.04811	-1.27
rs7198922	2	*SHISA9,ERCC4*	intergenic	16	ILMN_1814808	*BFAR*	0.04686	-1.15
rs7027886	4	*KCNV2,KIAA0020*	intergenic	9	ILMN_1818149		0.03728	-1.16
rs34771359	3	*CHN2,PRR15*	intergenic	7	ILMN_1821876	*AK024143*	0.04686	1.07
rs5994328	0	*SLC35E4,DUSP18*	intergenic	22	ILMN_1832879	*CD237904*	0.04727	1.36
rs2253693	0	*SIRPB1*	UTR3	20	ILMN_1841622	*AI655567*	0.04395	1.35
rs10784359	88	*SLC2A13*	intronic	12	ILMN_1859584	*AK026751*	0.0181	1.22
rs17108932	3	*PTEN,RNLS*	intergenic	10	ILMN_1880406	*PTEN*	0.03563	1.62
rs6001675	2	*ENTHD1*	intronic	22	ILMN_1883491	*AI970822*	0.04686	1.08
rs34764163	0	*DISP1*	intronic	1	ILMN_1901666	*AI445566*	0.04488	1.07
rs9329125	0	*LOC728554,PROP1*	intergenic	5	ILMN_1910550	*DR980253*	0.04729	-1.14
rs1529505	12	*F2RL1*	UTR5	5	ILMN_2041190	*F2RL1*	0.03758	1.19
rs17239727	174	*BLVRA*	intronic	7	ILMN_2081335	*C7orf44*	0.02446	1.18
rs2240516	0	*COA1*	intronic	7	ILMN_2081335	*C7orf44*	0.03325	1.18
rs6433294	11	*DCAF17,CYBRD1*	intergenic	2	ILMN_2087692	*CYBRD1*	0.02388	1.14

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs2387976	17	*NANP*	intronic	20	ILMN_2091792	*ENTPD6*	0.03872	-1.18
rs1562782	1	*ADM,AMPD3*	intergenic	11	ILMN_2093500	*ZBED5*	0.03563	-1.12
rs35288741	1	*NFASC*	intronic	1	ILMN_2094952	*NUAK2*	0.04387	1.15
rs4924543	1	*ZNF770,ANP32AP1*	intergenic	15	ILMN_2103547	*GOLGA8A*	0.03328	-1.26
rs7796045	83	*CCT6P3,ZNF92*	intergenic	7	ILMN_2118663	*ERV3*	0.02446	1.21
rs13332660	0	*SEZ6L2*	intronic	16	ILMN_2125747	*CORO1A*	0.04061	1.12
rs1408069	2	*KLF4,ACTL7B*	intergenic	9	ILMN_2137789	*KLF4*	0.03782	-1.11
rs425181	0	*C1orf87,NFIA*	intergenic	1	ILMN_2143148	*TM2D1*	0.04811	-1.06
rs8007588	4	*STXBP6*	intronic	14	ILMN_2148944	*ADCY4*	0.02723	1.26
rs13022989	0	*LOC440905*	ncRNA_intronic	2	ILMN_2156982	*IMP4*	0.04729	-1.19
rs9268926	196	*HLA-DRA,HLA-DRB5*	intergenic	6	ILMN_2159694	*HLA-DRB4*	0.00073	-1.08
rs749326	0	*SH3BP1*	intronic	22	ILMN_2162367	*DMC1*	0.04727	1.07
rs633683	0	*PHLDB1*	intronic	11	ILMN_2181241	*RPL23AP64*	0.04442	-1.22
rs9503750	0	*PXDC1,FAM50B*	intergenic	6	ILMN_2186806	*HLA-B*	0.02504	1.16
rs4688030	2	*MAATS1*	intronic	3	ILMN_2187718	*COX17*	0.04442	-1.16
rs7678870	11	*LOC340017,FAM198B*	intergenic	4	ILMN_2190851	*PPID*	0.04686	1.1
rs531815	30	*MAK*	intronic	6	ILMN_2209115	*MAK*	0.01664	1.13
rs2049490	0	*POC5,SV2C*	intergenic	5	ILMN_2221507	*F2R*	0.04686	1.09
rs7173954	0	*INO80*	intronic	15	ILMN_2234758	*SRP14*	0.04061	-1.11
rs2237250	5	*FYN*	intronic	6	ILMN_2249920	*FYN*	0.04398	-1.19
rs4653108	0	*SMIM12,DLGAP3*	intergenic	1	ILMN_2260500	*KIAA0319L*	0.04625	1.34
rs2277628	1	*MYCBPAP*	intronic	17	ILMN_2263718	*SPAG9*	0.02504	1.19
rs62262832	3	*C3orf17,BOC*	intergenic	3	ILMN_2286514	*GTPBP8*	0.04488	-1.18
rs2712353	0	*ATP6V1A*	intronic	3	ILMN_2286514	*GTPBP8*	0.04686	-1.18
rs6439982	14	*SPSB4,ACPL2*	intergenic	3	ILMN_2306955	*ACPL2*	0.04488	1.26
rs171803	405	*SLCO6A1,PAM*	intergenic	5	ILMN_2313901	*PAM*	0.00026	1.13
rs7256770	3	*ACP5*	upstream	19	ILMN_2339377	*DNM2*	0.02774	1.11
rs11859842	1	*SLC7A5P1,SPN*	intergenic	16	ILMN_2344373	*MVP*	0.03728	1.1
rs1041898	5	*SULF2,LINC00494*	intergenic	20	ILMN_2345142	*SULF2*	0.03386	1.17
rs2395891	3	*BTBD2,MKNK2*	intergenic	19	ILMN_2347068	*MKNK2*	0.02774	1.18
rs325828	13	*MROH2B*	intronic	5	ILMN_2357577	*PRKAA1*	0.02774	-1
rs9503168	0	*LOC100508120*	ncRNA_intronic	6	ILMN_2376205	*LTB*	0.04811	-1.36
rs10039049	1	*ANXA6*	intronic	5	ILMN_2379644	*CD74*	0.04488	-1.28
rs12766521	16	*SH2D4B,NRG3*	intergenic	10	ILMN_2380494	*ANXA11*	0.04442	1.22
rs4075678	1	*GALNT18*	intronic	11	ILMN_2380946	*EIF4G2*	0.04061	-1.01
rs736020	2	*DHRS9,LRP2*	intergenic	2	ILMN_2384181	*DHRS9*	0.04061	1.47
rs8033385	5	*ITGA11,CORO2B*	intergenic	15	ILMN_2386530	*RPLP1*	0.04061	-1.3
rs4845143	9	*IL19*	intronic	1	ILMN_2388112	*CR1*	0.04488	1.24
rs11686934	1	*MXD1,ASPRV1*	intergenic	2	ILMN_2388466	*TIA1*	0.04488	-1.1
rs17849707	2	*CEP68*	exonic	2	ILMN_2388605	*ACTR2*	0.04811	1.06
rs745749	0	*MAPK9*	intronic	5	ILMN_2390227	*TBC1D9B*	0.04442	-1.12
rs661552	2	*Sep-09*	intronic	17	ILMN_2391912	*SEC14L1*	0.04488	1.33
rs5763241	0	*RFPL1*	downstream	22	ILMN_2393169	*THOC5*	0.04515	1.15
rs131430	0	*IGLL1,C22orf43*	intergenic	22	ILMN_2393765	*IGLL1*	0.04488	-1.13
rs8015121	4	*RBM23*	intronic	14	ILMN_2403889	*PRMT5*	0.01205	1.1
rs6070412	42	*PPP4R1L*	ncRNA_intronic	20	ILMN_2404049	*RBM38*	0.03758	1.1
rs2371129	94	*EIF1B-AS1*	ncRNA_intronic	3	ILMN_2404850	*RPL14*	0.00997	-1.14

tag SNP	Ntag	Genes nearby tag SNP	SNP Location	Chr.	Probe id	Probe gene	Q values	FC
rs17834472	13	*SLC38A6,TMEM30B*	intergenic	14	ILMN_2410516	*PPM1A*	0.02471	1.14
rs4075428	7	*CIZ1*	intronic	9	ILMN_2413064	*ST6GALNAC4*	0.01888	1.08
rs166211	16	*LOC283867,CDH5*	intergenic	16	ILMN_2414027	*CKLF*	0.04392	1.27
3-113567114	3	*CD200,BTLA*	intergenic	3	ILMN_2415786	*CD96*	0.04488	-1.29

Table A.2.: List of transcription factor binding sites enriched within the sequences of GR-response *cis*-eSNPs. Transcription factor affinities were calculated using TRAP (http://trap.molgen.mpg.de/). For all GR-response *cis*-eSNPs a region of 20bp was used for the analysis. The list is ranked according to the most enriched factors.

ID[6]	Name[7]	Sequences with the reference alleles			Sequences with the alternative alleles		
		χ^2[8]	Combined P[9]	Corrected P[10]	χ^2[8]	Combined P[9]	Corrected P[10]
M01230	*ZNF333*	29329	<1e-321	<1e-321	27243	<1e-321	<1e-321
M00980	*TBP*	22500	<1e-321	<1e-321	20684	<1e-321	<1e-321
M00486	*PAX2*	16213	<1e-321	<1e-321	14555	<1e-321	<1e-321
M01107	*RUSH1A*	15509	<1e-321	<1e-321	15033	<1e-321	<1e-321
M01181	*NKX32*	14829	<1e-321	<1e-321	12921	<1e-321	<1e-321
M00489	*NKX62*	14763	<1e-321	<1e-321	14267	<1e-321	<1e-321
M01162	*OG2*	14540	<1e-321	<1e-321	13820	<1e-321	<1e-321
M00630	*FOXM1*	14433	<1e-321	<1e-321	13248	<1e-321	<1e-321
M01281	*NFAT1*	13992	<1e-321	<1e-321	12693	<1e-321	<1e-321
M01275	*IPF1*	13952	<1e-321	<1e-321	12937	<1e-321	<1e-321
M00493	*STAT5A*	13531	<1e-321	<1e-321	12277	<1e-321	<1e-321
M01032	*HNF4*	13184	<1e-321	<1e-321	11774	<1e-321	<1e-321
M01653	*HMGIY*	13135	<1e-321	<1e-321	12740	<1e-321	<1e-321
M00100	*CDXA*	12004	<1e-321	<1e-321	11957	<1e-321	<1e-321
M00137	*OCT1*	11468	<1e-321	<1e-321	11462	<1e-321	<1e-321
M01131	*SOX10*	11426	<1e-321	<1e-321	10355	4.8e-236	2.1e-234
M00395	*HOXA3*	11132	1e-309	<1e-321	11511	<1e-321	<1e-321
M00624	*DBP*	10963	4.4e-293	2e-291	10516	9e-251	4.1e-249
M00921	*GR*	10896	1.2e-286	5.1e-285	10194	1.2e-221	4.7e-220
M01292	*HOXA13*	10803	1.1e-277	4.6e-276	10560	7.6e-255	3.6e-253
M00500	*STAT6*	10726	2e-270	8e-269	9863	5.1e-193	1.9e-191
M00744	*POU1F1*	10234	3.3e-225	1.3e-223	10212	3.2e-223	1.3e-221
M00912	*CEBP*	10173	8.2e-220	3e-218	9664	1.7e-176	5.9e-175
M01665	*IRF8*	10138	9.3e-217	3.2e-215	9147	4.1e-136	1.2e-134
M01117	*OTX*	9897	5.9e-196	2e-194	9895	9.9e-196	3.7e-194
M01227	*MAFB*	9578	1.5e-169	4.6e-168	8420	2.6e-86	6.1e-85
M00962	*AR*	9461	3.3e-160	9.8e-159	8949	1.2e-121	3.3e-120
M00109	*CEBPB*	9381	6.4e-154	1.9e-152	9330	5.3e-150	1.7e-148
M00268	*XFD2*	9249	8.4e-144	2.3e-142	9065	4.3e-130	1.3e-128
M00482	*PITX2*	8901	2.9e-118	7.5e-117	8880	8e-117	2.2e-115
M00690	*AP3*	8854	5.8e-115	1.4e-113	8335	4.3e-81	9.3e-80
M00498	*STAT4*	8691	8e-104	2e-102	8137	2e-69	3.8e-68
M00285	*TCF11*	8651	3.5e-101	8.3e-100	8428	9e-87	2.1e-85
M00396	*EN1*	8569	8.4e-96	1.9e-94	8596	1.4e-97	3.8e-96
M01294	*PROP1*	8544	3.4e-94	7.6e-93	8590	3.6e-97	9.1e-96
M00672	*TEF*	8505	1.1e-91	2.3e-90	8398	6.1e-85	1.4e-83
M01067	*GFI1*	8459	9.6e-89	2e-87	8256	2.4e-76	4.9e-75
M00987	*FOXP1*	8386	3.2e-84	6.7e-83	8362	1.1e-82	2.4e-81
M00252	*TATA*	8244	1.2e-75	2.5e-74	7693	4.7e-46	7e-45
M01483	*DBX1*	8197	7.7e-73	1.5e-71	8241	1.9e-75	3.7e-74
M00960	*PR*	8159	1.2e-70	2.3e-69	7906	9.7e-57	1.6e-55
M00451	*NKX3A*	8136	2.5e-69	4.7e-68	8294	1.4e-78	2.9e-77
M01232	*SATB1*	8097	3.7e-67	6.7e-66	8157	1.5e-70	3e-69
M00463	*POU3F2*	8096	4.5e-67	8e-66	8082	2.7e-66	5e-65
M00116	*CEBPA*	8076	5.9e-66	1e-64	8161	9.2e-71	1.8e-69

[6]transcription factor accession number according to the TRANSFAC database
[7]name of transcription factor
[8]χ^2-distribution for each eSNP sequence combined using Fisher's method
[9]P value for the χ^2-distribution
[10]multiple testing corrected combined P values using Benjamini-Hochberg method

M00267	*XFD1*	7916	3e-57	5e-56	7939	1.6e-58	2.9e-57
M01654	*DRI1*	7912	4.3e-57	7.2e-56	7959	1.4e-59	2.5e-58
M00293	*FREAC7*	7876	3.6e-55	6e-54	7793	5.6e-51	9e-50
M01432	*HOXD8*	7875	3.7e-55	6e-54	7916	2.7e-57	4.7e-56
M00465	*POU6F1*	7842	2e-53	3.2e-52	7862	1.8e-54	3e-53
M00742	*HFH4*	7797	3.9e-51	6e-50	7790	8.2e-51	1.3e-49
M00510	*LHX3*	7758	3e-49	4.5e-48	7792	6.2e-51	9.8e-50
M01137	*FOXO3A*	7711	6.3e-47	9.4e-46	7766	1.2e-49	1.9e-48
M01464	*HOXA10*	7682	1.6e-45	2.4e-44	7731	7e-48	1.1e-46
M01316	*TST1*	7628	5.2e-43	7.4e-42	7682	1.6e-45	2.3e-44
M00096	*PBX1*	7616	2e-42	2.7e-41	7613	2.8e-42	3.8e-41
M01473	*BRN4*	7555	1.2e-39	1.6e-38	7616	2e-42	2.7e-41
M00351	*GATA3*	7549	2.2e-39	2.9e-38	7512	9.9e-38	1.3e-36
M01308	*SOX4*	7543	4.2e-39	5.5e-38	7178	2.2e-24	2e-23
M00809	*FOX*	7533	1.1e-38	1.5e-37	7524	2.8e-38	3.7e-37
M01420	*NCX*	7531	1.3e-38	1.7e-37	7556	1e-39	1.4e-38
M00269	*XFD3*	7512	9.8e-38	1.2e-36	7512	9.6e-38	1.3e-36
M00130	*FOXD3*	7485	1.4e-36	1.8e-35	7492	7.3e-37	9.2e-36
M00925	*AP1*	7475	3.7e-36	4.6e-35	6993	3.6e-18	2.6e-17
M00496	*STAT1*	7436	1.8e-34	2.2e-33	7164	6.6e-24	5.9e-23
M00423	*FOXJ2*	7434	2.2e-34	2.6e-33	7439	1.3e-34	1.6e-33
M01399	*HB24*	7415	1.4e-33	1.6e-32	7509	1.3e-37	1.6e-36
M01148	*DMRT3*	7397	7.6e-33	8.8e-32	7393	1.1e-32	1.3e-31
M00148	*SRY*	7396	8.3e-33	9.5e-32	7349	6.7e-31	7.1e-30
M00795	*OCT*	7371	8.7e-32	9.9e-31	7402	4.9e-33	5.7e-32
M00616	*AFP1*	7364	1.7e-31	1.9e-30	7335	2.6e-30	2.7e-29
M01599	*FOXP3*	7330	3.7e-30	4.1e-29	7319	1.1e-29	1.1e-28
M01469	*NKX61*	7324	6.5e-30	7.1e-29	7426	4.7e-34	5.7e-33
M01147	*DMRT2*	7315	1.5e-29	1.6e-28	7384	2.6e-32	3e-31
M01377	*IRXB3*	7311	2.3e-29	2.4e-28	7360	2.4e-31	2.6e-30
M01659	*CDX2*	7308	2.7e-29	2.8e-28	7304	4e-29	4e-28
M00082	*EVI1*	7295	8.9e-29	9.2e-28	7153	1.7e-23	1.5e-22
M00129	*HFH1*	7293	1.1e-28	1.1e-27	7258	2.5e-27	2.4e-26
M00241	*NKX25*	7290	1.4e-28	1.4e-27	7333	2.9e-30	3e-29
M00394	*MSX1*	7269	9.3e-28	9e-27	7405	3.6e-33	4.3e-32
M00802	*PIT1*	7264	1.4e-27	1.3e-26	7374	6.7e-32	7.6e-31
M01472	*IRX5*	7227	3.5e-26	3.3e-25	7244	8.4e-27	8e-26
M01324	*OCTAMER*	7227	3.6e-26	3.4e-25	7220	6.6e-26	6.2e-25
M00416	*CART1*	7216	9e-26	8.2e-25	7204	2.5e-25	2.3e-24
M01375	*HOXD10*	7199	3.7e-25	3.3e-24	7060	2.6e-20	2.1e-19
M00405	*MMEF2*	7196	4.9e-25	4.4e-24	7263	1.5e-27	1.5e-26
M01391	*PAX6*	7190	8e-25	7e-24	7250	4.9e-27	4.7e-26
M00406	*HMEF2*	7140	5e-23	4.4e-22	7168	4.9e-24	4.4e-23
M01439	*DLX1*	7108	5.9e-22	5.1e-21	7183	1.4e-24	1.3e-23
M00724	*HNF3ALPHA*	7108	6.4e-22	5.4e-21	7163	7.3e-24	6.4e-23
M01321	*HOXC8*	7101	1.1e-21	9.1e-21	7177	2.4e-24	2.2e-23
M01405	*IRX2*	7083	4.2e-21	3.5e-20	7053	4.2e-20	3.3e-19
M00006	*MEF2*	7079	6e-21	4.9e-20	7116	3.3e-22	2.8e-21
M00639	*HNF6*	7067	1.5e-20	1.2e-19	7098	1.3e-21	1.1e-20
M00318	*LPOLYA*	7054	4e-20	3.2e-19	7101	1e-21	8.8e-21
M00132	*HNF1*	7053	4.2e-20	3.4e-19	7023	3.9e-19	3e-18
M01363	*LMX1B*	7051	4.9e-20	3.9e-19	7087	3.2e-21	2.6e-20
M01410	*IRX4*	7049	5.6e-20	4.4e-19	7081	5.1e-21	4.1e-20
M01146	*DMRT1*	7037	1.4e-19	1.1e-18	6994	3.4e-18	2.5e-17
M01149	*DMRT4*	7020	5e-19	3.7e-18	6965	2.7e-17	1.8e-16
M00102	*CDP*	7012	9e-19	6.6e-18	7057	3.1e-20	2.5e-19
M00334	*DTYPEPA*	6990	4.4e-18	3.2e-17	6898	2.5e-15	1.6e-14
M01353	*LHX5*	6987	5.6e-18	4e-17	7038	1.3e-19	1e-18
M01268	*FXR*	6977	1.1e-17	7.7e-17	6439	4.5e-05	0.00019
M00991	*CDX*	6975	1.3e-17	8.9e-17	6953	5.9e-17	4e-16
M00311	*ATATA*	6967	2.2e-17	1.5e-16	6990	4.3e-18	3.1e-17
M00310	*APOLYA*	6967	2.3e-17	1.6e-16	6998	2.6e-18	1.9e-17
M00045	*E4BP4*	6962	3.2e-17	2.2e-16	6893	3.4e-15	2.2e-14
M00478	*CDC5*	6950	7.4e-17	5e-16	6954	5.7e-17	3.9e-16
M01409	*LMX1*	6947	9.2e-17	6.1e-16	6963	3.1e-17	2.1e-16

M01406	*HOXC6*	6934	2.3e-16	1.5e-15	7023	4e-19	3e-18
M01150	*DMRT5*	6919	6.2e-16	4.1e-15	6779	4.3e-12	2.6e-11
M01318	*IRX3*	6896	2.8e-15	1.8e-14	6929	3.2e-16	2.1e-15
M01408	*BRN3C*	6877	9.8e-15	6.4e-14	6941	1.4e-16	9e-16
M01451	*HOXB8*	6876	1.1e-14	6.9e-14	7006	1.4e-18	1.1e-17
M00260	*HLF*	6871	1.5e-14	9.4e-14	6850	5.6e-14	3.5e-13
M01392	*HOXA6*	6867	2e-14	1.2e-13	6944	1.1e-16	7.5e-16
M00294	*HFH8*	6865	2.2e-14	1.4e-13	6902	2e-15	1.3e-14
M00042	*SOX5*	6802	1.1e-12	6.7e-12	6741	3.8e-11	2.2e-10
M00407	*RSRFC4*	6800	1.2e-12	7.4e-12	6818	4.1e-13	2.6e-12
M00789	*GATA*	6779	4.2e-12	2.5e-11	6885	5.9e-15	3.7e-14
M00791	*HNF3*	6754	1.8e-11	1.1e-10	6757	1.5e-11	8.9e-11
M01426	*HOXB9*	6746	2.9e-11	1.7e-10	6730	7e-11	4e-10
M00131	*HNF3B*	6743	3.4e-11	1.9e-10	6694	5e-10	2.7e-09
M01125	*OCT4*	6742	3.6e-11	2.1e-10	6755	1.7e-11	9.8e-11
M00725	*HP1SITEFACTOR*	6735	5.2e-11	3e-10	6757	1.5e-11	8.9e-11
M01418	*LIM1*	6708	2.3e-10	1.3e-09	6779	4.1e-12	2.5e-11
M01400	*DLX3*	6701	3.4e-10	1.9e-09	6765	9.7e-12	5.8e-11
M00292	*FREAC4*	6687	7.2e-10	3.9e-09	6593	7.8e-08	3.9e-07
M01356	*PMX2B*	6672	1.6e-09	8.5e-09	6714	1.7e-10	9.4e-10
M01431	*BARX2*	6671	1.7e-09	8.9e-09	6720	1.2e-10	6.9e-10
M01446	*BARHL2*	6665	2.3e-09	1.2e-08	6654	3.8e-09	2.1e-08
M00734	*CIZ*	6648	5.3e-09	2.8e-08	6640	7.9e-09	4.2e-08
M00228	*VBP*	6648	5.4e-09	2.9e-08	6574	1.8e-07	9e-07
M00485	*NKX22*	6644	6.5e-09	3.4e-08	6623	1.8e-08	9.5e-08
M01211	*PARP*	6633	1.1e-08	5.9e-08	6523	1.7e-06	7.7e-06
M01486	*DLX7*	6632	1.2e-08	6.2e-08	6696	4.5e-10	2.5e-09
M01328	*ISL2*	6627	1.5e-08	7.9e-08	6723	1.1e-10	5.9e-10
M00291	*FREAC3*	6597	6.2e-08	3.1e-07	6523	1.7e-06	7.7e-06
M00145	*BRN2*	6593	7.7e-08	3.8e-07	6717	1.4e-10	7.7e-10
M01348	*K2B*	6593	7.7e-08	3.8e-07	6557	3.9e-07	1.9e-06
M01396	*HOXB7*	6573	1.9e-07	9.2e-07	6631	1.2e-08	6.5e-08
M00640	*HOXA4*	6567	2.5e-07	1.2e-06	6557	3.9e-07	1.9e-06
M00103	*CLOX*	6556	4e-07	2e-06	6581	1.3e-07	6.7e-07
M00622	*CEBPGAMMA*	6555	4.2e-07	2.1e-06	6573	1.9e-07	9.3e-07
M00157	*RORA2*	6548	5.8e-07	2.8e-06	6439	4.5e-05	0.00019
M01487	*HOXA1*	6546	6.2e-07	3e-06	6580	1.4e-07	7e-07
M01454	*HOXC5*	6536	9.5e-07	4.5e-06	6573	1.9e-07	9.2e-07
M01460	*HOXB6*	6531	1.2e-06	5.6e-06	6594	7.3e-08	3.7e-07
M01319	*HOXB5*	6516	2.3e-06	1e-05	6543	7.1e-07	3.4e-06
M01413	*HMX3*	6512	2.7e-06	1.2e-05	6504	3.6e-06	1.6e-05
M01458	*UNCX4.1*	6512	2.7e-06	1.2e-05	6563	3e-07	1.4e-06
M01429	*HOMEZ*	6507	3.3e-06	1.5e-05	6438	4.6e-05	0.00019
M00410	*SOX9*	6507	3.3e-06	1.5e-05	6263	0.0088	0.031
M01394	*HOXA7*	6500	4.3e-06	1.9e-05	6540	8.1e-07	3.8e-06
M00403	*AMEF2*	6499	4.4e-06	2e-05	6517	2.2e-06	9.8e-06
M01369	*HOXC4*	6490	6.5e-06	2.9e-05	6538	9e-07	4.2e-06
M01416	*HOXC9*	6481	9.2e-06	4.1e-05	6578	1.5e-07	7.4e-07
M01016	*SOX17*	6480	9.6e-06	4.2e-05	6460	2e-05	8.8e-05
M01378	*HOXA11*	6469	1.4e-05	6.2e-05	6528	1.4e-06	6.4e-06
M01381	*OBOX5*	6464	1.8e-05	7.6e-05	6557	3.9e-07	1.9e-06
M00289	*HFH3*	6442	4e-05	0.00017	6472	1.3e-05	5.7e-05
M01661	*HBP1*	6439	4.4e-05	0.00019	6429	6.4e-05	0.00027
M01359	*DOBOX4*	6438	4.6e-05	0.00019	6359	0.00063	0.0024
M01435	*PSX1*	6434	5.3e-05	0.00022	6435	5.2e-05	0.00022
M01244	*HSF2*	6422	8e-05	0.00033	6374	0.00041	0.0016
M01478	*CPHX*	6418	9.2e-05	0.00038	6521	1.8e-06	8.3e-06
M01351	*HOXA9*	6418	9.4e-05	0.00039	6423	7.9e-05	0.00033
M01423	*ARX*	6412	0.00011	0.00046	6416	1e-04	0.00042
M01360	*DBX2*	6408	0.00013	0.00052	6441	4.1e-05	0.00018
M01335	*VSX1*	6406	0.00014	0.00057	6428	6.5e-05	0.00027
M01183	*BCL6*	6402	0.00016	0.00065	6376	0.00037	0.0014
M01323	*OTP*	6393	0.00022	0.00087	6409	0.00013	0.00052
M00437	*CHX10*	6372	0.00043	0.0017	6402	0.00016	0.00066
M01345	*SIX6*	6371	0.00045	0.0017	6377	0.00036	0.0014

M01315	*NKX52*	6367	5e-04	0.0019	6360	0.00061	0.0023
M01251	*E2F1*	6360	0.00062	0.0024	6284	0.0053	0.019
M01327	*VAX2*	6356	7e-04	0.0027	6389	0.00025	0.001
M01424	*HOXB4*	6355	0.00072	0.0028	6387	0.00026	0.001
M01151	*DMRT7*	6341	0.0011	0.0042	6290	0.0045	0.016
M00462	*GATA6*	6333	0.0014	0.0052	6402	0.00016	0.00065
M01329	*HOXC11*	6333	0.0014	0.0052	6380	0.00033	0.0013
M01341	*MSX3*	6329	0.0016	0.0059	6358	0.00066	0.0025
M00805	*LEF1*	6304	0.0031	0.011	6268	0.0078	0.028
M01368	*OCT2*	6301	0.0034	0.013	6328	0.0016	0.0061
M01355	*ALX3*	6286	0.005	0.018	6296	0.0039	0.014
M01402	*HOXA2*	6282	0.0055	0.02	6307	0.0029	0.011
M01286	*SOX*	6275	0.0067	0.024	6264	0.0086	0.031
M01476	*POU2F3*	6257	0.01	0.036	6288	0.0048	0.017
M01326	*GSH2*	6246	0.013	0.047	6313	0.0024	0.009

Table A.3.: Annotation of all interactions in figure 3.11 with the respective literature references supporting the interaction network.

Subject	Relation type	Object	PubMed ID	CIDeR ID
15q13.3 microdeletion syndrome	affects activity of	*Autism*	19289393	46184
15q13.3 microdeletion syndrome	affects activity of	*mental retardation*	19289393	46185
22q11.2 deletion syndrome	affects activity of	*Schizophrenia*	17008057	46202
22q11.2 deletion syndrome	increases expression of	*RPS2*	17008057	46203
AKT1	increases activity of	*mTOR signaling*	18924132	32977
AKT1	increases quantity of	*HIF1A*	12764143	32458
AKT1	is part of	*COMMD1-AKT1 complex*	20237237	32241
Alzheimer disease	affects expression of	*APP (amyloid beta peptide)*	24270855	46214
Alzheimer disease	increases expression of	*OCIAD2*	24270855	46204
ATP5F1	Interacts (localizes) with	*ICT1*	20186120	46187
ATP5F1	interacts (colocalizes) with	*SNCA*	18614564	31888
ATP5F1	interacts (colocalizes) with	*VHL*	17353931	31896
C7orf44	Interacts (localizes) with	*CCDC56*	23260140	46189
CCDC56	Interacts (localizes) with	*OCIAD2*	23260140	46190
CCT7	interacts (colocalizes) with	*CDKN1A*	21900206	32886
CDKN1A	interacts (colocalizes) with	*HMGXB3*	21900206	32897
Citalopram	increases quantity of	*CTNNB1*	22634067	32972
CLEC4C	affects activity of	*immune response*	23606632	46851
CLEC4C	interacts (colocalizes) with	*NEK2*	12386167	46849
COMMD1	affects activity of	*NEDD4L*	20237237	32238
COMMD1	decreases activity of	*NF-kappaB complex*	19220812	32316
COMMD1	decreases expression of	*HIF1A*	20458141	32449
COMMD1	interacts (colocalizes) with	*CUL2*	17183367	31950
COMMD1	interacts (colocalizes) with	*SGK1*	20237237	31547
COMMD1	is part of	*COMMD1-SGK1 complex*	20237237	32931
COMMD1	interacts (colocalizes) with	*AKT1*	20237237	32235
COMMD1	is part of	*COMMD1-AKT1 complex*	20237237	32902
COMMD1-AKT1 complex	decreases activity of	*EnaC complex*	20237237	32237
COMMD1-SGK1 complex	decreases activity of	*EnaC complex*	20237237	32236
CTNNB1	interacts (colocalizes) with	*NEK2*	18086858	46845
CUL2	is part of	*ElonginB/C-CUL2-VHL-RBX1 complex*	21942715	31944
Dexamethasone	decreases quantity of	*ATP5F1*	21060993	31540
Dexamethasone	increases expression of	*FKBP5*	12519866	32803
Dexamethasone	interacts (colocalizes) with	*NR3C1*	16189295	31548
CUL2-VHL-RBX1 complex	decreases activity of	*HIF1A*	21942715	31945
ESR2	Interacts (localizes) with	*CTNNB1*	21182203	46192
ESR2 (polymorphism)	cooccurs with	*Major depressive disorder*	22901010	46208
EWSR1	affects expression of	*C7orf44*	20442286	41802
FAN1 (MTMR15)	co-occurs with	*15q13.3 microdeletion syndrome*	19289393	46186
FAN1 (MTMR15)	Interacts (localizes) with	*FSCN1*	20603015	46194
FAN1 (MTMR15)	Interacts (localizes) with	*SUMO2*	21693764	46191
FKBP5	decreases activity of	*AKT1*	22590527	32808
FKBP5	interacts (colocalizes) with	*HSP90AA1*	19560279	32795
FKBP5 (polymorphism)	cooccurs with	*Major depressive disorder*	15565110	19744
FKBP5 (polymorphism)	cooccurs with	*response to antidepressants*	15565110	19606
FSCN1	Interacts (localizes) with	*CTNNB1*	10026156	46195
FTH1	Interacts (localizes) with	*GRB2*	21988832	46197
GRB2	Interacts (localizes) with	*CCT7*	19380743	46198
GRB2	Interacts (localizes) with	*HIST2H2AA3*	12577067	46196
HIF1A	increases activity of	*inflammatory response*	12628185	32471
HIF1A	increases expression of	*EWSR1*	20442286	41801
HLA-DRB4 (alelle variant)	co-occurs with	*Autism*	12039413	41905
HLA-DRB4 (polymorphism)	cooccurs with	*Schizophrenia*	9713902	41901
HLA-DRB5 (haplotype)	affects activity of	*Schizophrenia*	17001352	46209
HLA-DRB5 (polymorphism)	affects activity of	*immune response*	24075919	46210
IFNG	increases expression of	*SLC7A7*	15280038	32261
IKBKB	affects activity of	*inflammatory response*	18626576	46211
IKBKB	Interacts (localizes) with	*CTNNB1*	11527961	46200
immune response	affects expression of	*HLA-DRB4*	9952022	41903
IMPDH2	decreases activity of	*NF-kappaB complex*	21460227	32324
IMPDH2	decreases activity of	*TLR signaling pathway*	21460227	32546
IMPDH2	interacts (colocalizes) with	*AKT1*	10930578	32903
Ketamine	decreases activity of	*Major depressive disorder*	22205190	32978
Ketamine	increases activity of	*mTOR signaling*	20724638	32809
Major depression disorder	increases expression of	*TMEM176A*	20830301	46212
Major depressive disorder	increases expression of	*ATP5F1*	22832852	31537
MKNK2	decreases activity of	*SPRY2*	19864419	32242
MKNK2	increases activity of	*inflammatory response*	10559880	32362
MKNK2	increases quantity of	*IFNG*	10559880	32361
MKNK2	increases quantity of	*SPRY2-NEDD4 complex*	19864419	32231
MKNK2	increases quantity of	*TNF*	10559880	32357
MKNK2	Interacts (localizes) with	*ESR2*	11013076	46199
MRPL54	Interacts (localizes) with	*ICT1*	20186120	46188
NEDD4	increases ubiquitination of	*MOB3A*	19953087	32436
NEDD4	is part of	*SPRY2-NEDD4 complex*	19864419	32233
NEDD4L	decreases activity of	*BMP receptor signaling*	15496141	32523
NEDD4L	decreases activity of	*EnaC complex*	20237237	32239
NEDD4L	increases ubiquitination of	*MOB3A*	19953087	32436
NF-kappaB complex	affects activity of	*inflammatory response*	22726116	32319
NF-kappaB complex	increases expression of	*HIF1A*	18432192	32463
NF-kappaB complex	increases expression of	*NUAK2*	15345718	32395

NR3C1	increases expression of	*CDKN1A*	9442036	32989
NR3C1	increases expression of	*SGK1*	16189295	31549
NR3C1	interacts (colocalizes) with	*HSP90AA1*	19560279	32793
NUAK2	increases activity of	*neural tube development*	22689267	32406
Obesity	increases expression of	*OCIAD2*	20020228	46206
OCIAD2	increases quaintity of	*APP (amyloid beta peptide)*	24270855	46205
PELI1	affects activity of	*inflammatory response*	19734906	32252
PELI1	affects activity of	*TLR signaling pathway*	19734906	32562
PELI1	increases activity of	*NF-kappaB complex*	19734906	32251
PELI1	Interacts (localizes) with	*IKBKB*	21204785	46201
POLR2I	interacts (colocalizes) with	*TBP*	20482850	32611
Post-traumatic stress disorder	increases methylation of	*FSCN1*	23630272	46215
RELA	increases activity of	*NF-kappaB complex*	12213593	32952
RELA	interacts (colocalizes) with	*COMMD1*	20048074	31555
RELA	interacts (colocalizes) with	*TBP*	7706261	31648
response to antidepressants	affects quantity of	*CTNNB1*	22472057	32974
RPS2	Interacts (localizes) with	*VHL*	17353931	46217
Schizophrenia	increases expression of	*FTH1*	18191109	46218
SGK1	increases transport of	*SLC9A3R2*	11751930	31677
SGK1	increases transport of	*SLC9A3R2*	11751930	31677
SGK1	is part of	*COMMD1-SGK1 complex*	20237237	32240
SLC6A15	increases activity of	*neutral amino acid transport*	21521612	32887
SLC6A15 (polymorphism)	cooccurs with	*Major depressive disorder*	21521612	22768
SLC7A7	increases activity of	*neutral amino acid transport*	9878049	32965
SLCO3A1	interacts (colocalizes) with	*SLC9A3R2*	15553237	31675
SLCO3A1 (polymorphism)	cooccurs with	*Schizophrenia*	18521091	32933
SNCA	interacts (colocalizes) with	*CCT7*	18614564	31889
SPRY2	interacts (colocalizes) with	*NEDD4*	19864419	32232
SPRY2	is part of	*SPRY2-NEDD4 complex*	19864419	32244
SUMO2	Interacts (localizes) with	*HIST2H2AA3*	21693764	46193
TLR signaling pathway	increases activity of	*mTOR signaling*	18924132	32871
TMEM176A	affects activity of	*dendritic cell differentiation*	20501748	46216
TNF	increases expression of	*SLC7A7*	11742806	32266
VHL	interacts (colocalizes) with	*CUL2*	9122164	31934
VHL	is part of	*CUL2-VHL-RBX1 complex*	21942715	31941

Table A.4.: List of the gene expression data from two mouse experiments compared to human eQTL study. The orthologous genes in mice were compared to the human probe genes of the GR-response *cis*-eQTLs that showed overlap with the meta-analysis for MDD and tested for differential expression between baseline and GR-stimulated expression in HC, PFC and AM, as well as for differential expression between stress resilient and susceptible groups of mice in CA1 and DG. Transcripts regulated in brain are marked with #.

			Humans				
	Network[$]	Probe gene[a]	Probe id[b]	FC[c]	FC_{Risk}[d]	FC_{NoRisk}[e]	change[f]
1	yes	ATP5F1	ILMN_1721989	-1.13	-0.12	-0.33	0.21
2	yes	C7orf44	ILMN_2081335	1.18	0.18	0.36	0.17
3	yes	CCT7	ILMN_1662954	-1.17	-0.4	-0.2	-0.2
4	yes	CLEC4C	ILMN_1665457	1.19	0.69	0.21	-0.48
		CLEC4C	ILMN_1682259	1.12	0.46	0.14	-0.32
5	yes	COMMD1	ILMN_1761242	-1.22	-0.18	-0.34	0.16
6	yes	FTH1	ILMN_1746525	1.11	0.29	0.11	-0.18
7	yes	HIST2H2AA3.HIST2H2AA4	ILMN_1695435	1.22	0.34	-0.02	-0.36
8	yes	HLA-DRB4	ILMN_2159694	-1.08	-0.03	-0.23	-0.2
9	yes	HLA-DRB5	ILMN_1697499	-1.07	0.04	-0.32	-0.36
10	yes	HMGXB3	ILMN_1694686	-1.14	-0.3	-0.13	0.16
11	yes	IMPDH2	ILMN_1705737	-1.24	-0.36	-0.2	-0.16
12	no	LRRC25	ILMN_1766487	1.1	-0.06	0.21	0.15
13	yes	MKNK2	ILMN_2347068	1.18	0.44	0.22	-0.21
14	yes	MOB3A	ILMN_1721344	1.15	0.42	0.2	-0.22
15	yes	MRPL54	ILMN_1658486	-1.24	-0.45	-0.27	-0.18
16	yes	MTMR15	ILMN_1778734	-1.16	-0.27	-0.16	0.11
17	yes	NUAK2	ILMN_2094952	1.15	0.1	0.3	0.19
18	yes	OCIAD2	ILMN_1700306	-1.32	0.24	-0.36	-0.61
19	yes	PELI1	ILMN_1679268	1.68	0.98	0.69	-0.29
20	yes	POLR2I	ILMN_1720542	-1.18	-0.29	-0.11	-0.19
21	yes	RPS2	ILMN_1688749	-1.22	0.14	-0.3	0.17
22	yes	SLC7A7	ILMN_1810275	-1.01	-0.23	0.03	-0.2
23	yes	SLCO3A1	ILMN_1663699	1.11	0.34	0.12	-0.22
24	yes	TMEM176A	ILMN_1791511	-1.15	-0.02	-0.24	-0.22

[*] expressed below background-signal
[$] Probe genes that generated a tightly interconnected network in 3.11 (indicated by: yes)
[a] Human array probe gene
[b] Illumina probe identifier for HumanHT-12 v3
[c, d, e] fold change of GR-stimulated/baseline gene expression (=deltaDex) for all samples, risk genotype carriers and nonrisk genotype carriers in human blood
[f] deltaDex change from risk to nonrisk genotype carriers in human blood
[g] orthologous genes in mice
[h,u] array quality checks
[i] Illumina probe identifier for MouseRef-8 v2
[j, l, n ,p ,r] fold change of GR-stimulated/baseline gene expression in each mice tissue (blood, HC, PFC, HC and PFC together, AM)
[k, m, o, q, l, s] nominal P value for differentially regulated mRNA Expression of deltaDex in each mice tissue (blood, HC, PFC, HC and PFC together, AM)
[U] Illumina probe identifier for MouseRef-8 v1
[v, w] score for differential expression between resilient and stress susceptible groups of mice (DS=DiffScore> ±13 equal to P value=0.05)
[x, y] fold change of general expression between resilient and stress susceptible groups of mice

	GR-stimulation mouse model						
	Orthologe Genes[g]	Array QC[h]	Probe id[i]	FC_{Blood}[j]	P_{Blood}[k]	FC_{HC}[l]	P_{HC}[m]
1	Atp5f1		ILMN_2790486	-1.11	0.00669	1.04	0.412
2	no orthologous gene	NA	NA	NA	NA	NA	NA
3	# Cct7		ILMN_1233793	1.14	0.00353	1.23	0.0428
4	Clec4c		ILMN_2959372	-1.14	0.00418	*	*
5	Commd1		ILMN_1215639	1.05	0.24	-1.01	0.93
	Commd1		ILMN_1231658	-1.19	9.66e-05	-1.02	0.503
6	Fth1		ILMN_2876066	1.16	0.109	-1	0.978
	Fth1		ILMN_2876071	1.23	0.0276	-1.04	0.732
7	Hist1h2al	no Ilum. Probe	NA	NA	NA	NA	NA
8	no orthologous gene	NA	NA	NA	NA	NA	NA
9	no orthologous gene	NA	NA	NA	NA	NA	NA
10	# Hmgxb3		ILMN_1246992	-1.17	0.00727	1.1	0.0632
11	Impdh2		ILMN_2588398	-1.43	2.36e-06	1.02	0.697
12	Lrrc25		ILMN_2715800	*	*	*	*
13	# Mknk2		ILMN_2733887	-1.19	0.0381	1.16	0.00162
14	Mob3a	no Ilum. Probe	NA	NA	NA	NA	NA
15	# Mrpl54		ILMN_2755519	-1.36	3.89e-05	-1.06	0.141
16	Mtmr15	no Ilum. Probe	NA	NA	NA	NA	NA
17	# Nuak2		ILMN_2680038	-1.24	0.00153	1.06	0.14
18	# Ociad2		ILMN_2943722	1.21	0.0388	-1.03	0.739
19	# Peli1		ILMN_1239770	-1.22	0.000412	1.07	0.279
	Peli1		ILMN_2627441	-1.26	0.00012	1.07	0.406
20	# Polr2i		ILMN_2666438	-1.02	0.597	-1.06	0.0287
21	Rps2		ILMN_2946616	-1.22	0.0435	-1.09	0.125
22	# Slc7a7		ILMN_1240318	*	*	1.06	0.288
	Slc7a7		ILMN_2723826	*	*	1	0.953
23	Slco3a1		ILMN_1235635	*	*	1.16	0.219
	# Slco3a1		ILMN_1235735	1.07	0.00475	1.04	0.402
	# Slco3a1		ILMN_2663230	1.05	0.105	1.23	0.000565
24	Tmem176a	no Ilum. Probe	NA	NA	NA	NA	NA

	GR-stimulation mouse model					
	FC_{PFC}[n]	P_{PFC}[o]	$FC_{PFC\$HC}$[p]	$P_{PFC\&HC}$[q]	FC_{AM}[r]	P_{AM}[s]
1	1.08	0.204	1.05	0.2463	1.05	0.294
2	NA	NA	NA	NA	NA	NA
3	-1.12	0.142	1.03	0.6176	1	0.968
4	*	*	*	*	*	*
5	-1.08	0.261	-1.06	0.284	1.02	0.723
	1.01	0.813	-1.01	0.7038	1	0.918
6	1.01	0.899	1.01	0.9968	-1.04	0.306
	-1.06	0.419	-1.05	0.455	-1.02	0.581
7	NA	NA	NA	NA	NA	NA
8	NA	NA	NA	NA	NA	NA
9	NA	NA	NA	NA	NA	NA
10	1.07	0.218	1.08	0.02834	-1.03	0.373
11	1.04	0.325	1.02	0.4948	-1.04	0.358
12	*	*	*	*	*	*
13	1.14	0.00155	1.14	3.847e-05	1.05	0.226
14	NA	NA	NA	NA	NA	NA
15	-1.08	0.0877	-1.08	0.02002	-1.01	0.787
16	NA	NA	NA	NA	NA	NA
17	1.1	0.0274	1.08	0.004751	1.13	0.00145
18	1.39	0.000124	1.15	0.01646	1.34	2.45e-06
19	1.12	0.0489	1.09	0.05299	1.11	0.0154
	1.02	0.643	1.04	0.3519	1.09	0.0667
20	-1.02	0.424	-1.04	0.01715	-1.02	0.376
21	1.08	0.231	1.01	0.8909	1.01	0.783
22	1.15	0.00107	1.09	0.00193	*	*
	1.06	0.12	1.03	0.2904	1.02	0.562
23	-1.02	0.805	1.06	0.3859	1.03	0.62
	1.19	0.000628	1.11	0.001138	1.03	0.509
	1.37	3.29e-05	1.28	6.379e-08	1.05	0.365
24	NA	NA	NA	NA	NA	NA

			Chronic social stress mouse model				
	Orthologe Genes	Array QC[t]	Probe id[u]	$DS_{CA1}{}^{v}$	$change_{CA1}{}^{w}$	$DS_{DG}{}^{x}$	$change_{DG}{}^{y}$
1	# Atp5f1		ILMN_2790486	-87.35	0.51	-5.67	0.11
2	no orthologous gene	NA	NA	NA	NA	NA	NA
3	# Cct7		ILMN_1233793	-259.81	0.71	-5.01	0.06
4	Clec4c	no Ilum. Probe	NA	NA	NA	NA	NA
5	Commd1		ILMN_1215639	*	*	-0.41	0.04
	Commd1		ILMN_1231658	-3.41	0.3	-4.38	0.09
6	# Fth1		ILMN_2744442	-0.92	0.02	30.91	-0.21
7	Hist1h2al	no Ilum. Probe	NA	NA	NA	NA	NA
8	no orthologous gene	NA	NA	NA	NA	NA	NA
9	no orthologous gene	NA	NA	NA	NA	NA	NA
10	Hmgxb3	no Ilum. Probe	NA	NA	NA	NA	NA
11	# Impdh2		ILMN_2588399	15.23	-0.14	4.16	-0.06
12	Lrrc25		ILMN_2715800	*	*	*	*
13	Mknk2		ILMN_2733887	2.44	-0.29	3.73	-0.27
	Mknk2		ILMN_2776441	*	*	*	*
14	Mob3a	no Ilum. Probe	NA	NA	NA	NA	NA
15	Mrpl54		ILMN_2755519	-2.67	0.22	-4.65	0.11
16	Mtmr15	no Ilum. Probe	NA	NA	NA	NA	NA
17	Nuak2	no Ilum. Probe	NA	NA	NA	NA	NA
18	Ociad2	no Ilum. Probe	NA	NA	NA	NA	NA
19	# Peli1		ILMN_1239770	*	*	18.18	-1.09
	# Peli1		ILMN_2627441	2.82	-0.55	17.92	-0.26
20	# Polr2i		ILMN_2666438	-3.05	0.16	16.44	-0.2
21	# Rps2		ILMN_2717549	-14.56	0.12	24.45	-0.13
22	Slc7a7		ILMN_1240318	*	*	*	*
	Slc7a7		ILMN_2690187	*	*	*	*
23	# Slco3a1		ILMN_1235635	-16.46	0.45	-35.34	0.29
	Slco3a1		ILMN_1235735	*	*	*	*
	Slco3a1		ILMN_2663230	13.38	-0.47	0.68	-0.03
24	Tmem176a	no Ilum. Probe	NA	NA	NA	NA	NA

Table A.5.: This list represents the set of genes of the interaction network in figure 3.11 including genes supported by the eQTL study ("eQTL") as well as the additional genes ("add") from the network analysis. The probes for orthologous mouse genes on gene expression arrays were tested for differential expression in the CA1 and DG region of the HC between stress-resilient and stress-susceptible groups of mice. The orthologous genes were also tested for significant regulation in the HC, PFC and the AM after stimulation with dexamethasone. The human genes *APP*, *C7ORF44*, *HLA-DRB4* and *HLA-DRB5* had no orthologs in mouse.

			Chronic social stress mouse model			GR-stimulation mouse model				
	Network	P gene[a]	Probe id[b]	DS_{CA1}[c]	DS_{DG}[d]	Probe id[e]	P_{HC}[f]	P_{PFC}[g]	$P_{PFC\&HC}$[h]	P_{AM}[i]
1	add	Akt1	ILMN1213935	*	*	ILMN1213935	*	*	*	*
2	eQTL	Atp5f1	ILMN_2790486	-87.35	-5.67	ILMN_2790486	0.412	0.204	0.2463	0.294
3	add	Ccdc56	ILMN_2932518	*	*	ILMN_2932518	0.458	0.767	0.4927	0.0676
4	eQTL	Cct7	ILMN_1233793	-259.81	-5.01	ILMN_1233793	0.0428	0.142	0.6176	0.968
	add	Cdkn1a	ILMN_2634083	-0.26	-2.86	ILMN_2634083	4.29e-14	1.59e-15	3.877e-28	2.17e-18
	add	Cdkn1a	ILMN_1214871	0.06	-0.85	no Ilum. Probe	NA	NA	NA	NA
	add	Cdkn1a	no Ilum. Probe	NA	NA	ILMN_2846775	1.46e-14	1.33e-13	5.184e-27	5.98e-17
5	add	Cdkn1a	no Ilum. Probe	NA	NA	ILMN_2846776	5.5e-16	3.06e-15	2.511e-30	4.28e-16
6	eQTL	Clec4c	no Ilum. Probe	NA	NA	ILMN_2959372	*	*	*	*
	eQTL	Commd1	ILMN_1215639	*	-0.41	ILMN_1215639	0.93	0.261	0.284	0.723
7	eQTL	Commd1	ILMN_1231658	-3.41	-4.38	ILMN_1231658	0.503	0.813	0.7038	0.918
	add	Ctnnb1	no Ilum. Probe	NA	NA	ILMN_2616556	0.671	0.519	0.7962	0.312
	add	Ctnnb1	ILMN_2696575	371.3	205.5	ILMN_2696575	0.524	0.0561	0.2291	0.568
8	add	Ctnnb1	no Ilum. Probe	NA	NA	ILMN_2994460	0.0794	0.246	0.6201	0.951
	add	Cul2	ILMN_1238615	2.51	-36.65	ILMN_1238615	0.438	0.53	0.8926	0.476
9	add	Cul2	no Ilum. Probe	NA	NA	ILMN_2987369	0.236	0.459	0.1312	0.629
	add	Esr2	ILMN_1236030	0.2	-5.55	ILMN_1236030	*	*	*	0.948
	add	Esr2	no Ilum. Probe	NA	NA	ILMN_3041053	*	*	*	*
10	add	Esr2	no Ilum. Probe	NA	NA	ILMN_3115826	*	*	*	*
	add	Ewsr1	ILMN_1219609	*	*	ILMN_1219609	*	*	*	*
11	add	Ewsr1	ILMN_2769308	*	*	ILMN_2769308	*	*	*	*
12	add	Fkbp5	ILMN_2718266	0.81	0.03	ILMN_2718266	5.05e-13	1.41e-13	1.223e-24	4.95e-20
13	add	Fscn1	ILMN_2630605	2.19	-53.54	ILMN_2630605	0.00419	0.5	0.003726	0.596
	eQTL	Fth1	ILMN_2744442	-0.92	30.91	ILMN_2876066	0.978	0.899	0.9968	0.306
14	eQTL	Fth1				ILMN_2876071	0.732	0.419	0.455	0.581
15	add	Grb2	ILMN_1222450	-10.32	-4.16	ILMN_2995537	0.579	0.689	0.6062	0.409
16	add	HIF1A	no Ilum. Probe	NA	NA	ILMN_2852034	0.505	0.11	0.06288	0.837
17	eQTL	Hist1h2al	no Ilum. Probe	NA	NA	no Ilum. Probe	NA	NA	NA	NA
18	eQTL	Hmgxb3	no Ilum. Probe	NA	NA	ILMN_1246992	0.0632	0.218	0.02834	0.373
	add	Hsp90aa1	no Ilum. Probe	NA	NA	ILMN_2662557	*	*	*	*
19	add	Hsp90aa1	no Ilum. Probe	NA	NA	ILMN_2752883	0.285	0.0204	0.3996	0.714
20	add	Ict1	ILMN_2642063	-10.32	-4.16	ILMN_2642063	0.271	0.735	0.3128	0.082
21	add	Ifng	ILMN_2685712	*	*	ILMN_2685712	*	*	*	*
22	add	Ikbkb	ILMN_2589557	4.41	25.13	ILMN_2589556	0.0291	0.938	0.08341	0.525
23	eQTL	Impdh2	ILMN_2588399	15.23	4.16	ILMN_2588398	0.697	0.325	0.4948	0.358
	eQTL	Mknk2	ILMN_2733887	2.44	3.73	ILMN_2733887	0.00162	0.00155	3.847e-05	0.226
24	eQTL	Mknk2	ILMN_2776441	*	*					
25	eQTL	Mob3a	no Ilum. Probe	NA	NA	no Ilum. Probe	NA	NA	NA	NA
26	eQTL	Mrpl54	ILMN_2755519	-2.67	-4.65	ILMN_2755519	0.141	0.0877	0.02002	0.787
27	eQTL	Mtmr15	no Ilum. Probe	NA	NA	no Ilum. Probe	NA	NA	NA	NA
28	add	Nedd4	ILMN_2594344	-9.54	13.46	ILMN_2594344	0.075	0.513	0.07812	0.196
29	add	Nedd4l	ILMN_2604457	34.36	-2.76	ILMN_2878501	8.15e-06	2.91e-06	1.274e-10	0.000183
30	add	Nek2	no Ilum. Probe	NA	NA	no Ilum. Probe	NA	NA	NA	NA
31	add	Nr3c1	ILMN_2740568	*	*	ILMN_2740568	0.0389	7.46e-07	3.807e-07	NA
32	eQTL	Nuak2	no Ilum. Probe	NA	NA	ILMN_2680038	0.14	0.0274	0.004751	0.00145
33	eQTL	Ociad2	no Ilum. Probe	NA	NA	ILMN_2943722	0.739	0.000124	0.01646	2.45e-06
	eQTL	Peli1	ILMN_1239770	*	18.18	ILMN_1239770	0.279	0.0489	0.05299	0.0154
34	eQTL	Peli1	ILMN_2627441	2.82	17.92	ILMN_2627441	0.406	0.643	0.3519	0.0667
35	eQTL	Polr2i	ILMN_2666438	-3.05	16.44	ILMN_2666438	0.0287	0.424	0.01715	0.376
36	add	Rela	ILMN_2740859	-7.16	-11.26	ILMN_2740859	0.61	0.909	0.6022	0.0311
37	eQTL	Rps2	ILMN_2717549	-14.56	24.45	ILMN_2946616	0.125	0.231	0.8909	0.783
38	add	Sgk1	ILMN_1213954	*	*	ILMN_1213954	1.23e-07	2.47e-09	4.711e-15	3.21e-12
	add	Slc6a15	ILMN_1245258	-102.16	-0.25	ILMN_1245258	0.75	0.715	0.8139	0.594
	add	Slc6a15	ILMN_1258914	0.72	4.5	ILMN_1258914	0.913	0.426	0.4772	0.753
39	add	Slc6a15	no Ilum. Probe	NA	NA	ILMN_2689230	0.825	0.624	0.7441	0.527
	eQTL	Slc7a7	ILMN_1240318	*	*	ILMN_1240318	0.288	0.00107	0.00193	*
40	eQTL	Slc7a7	ILMN_2690187	*	*	ILMN_2723826	0.953	0.12	0.2904	0.562
	add	Slc9a3r2	ILMN_1231582	*	*	no Ilum. Probe	NA	NA	NA	NA
	add	Slc9a3r2	ILMN_1218241	0.51	*	ILMN_1218241	*	0.755	*	*
41	add	Slc9a3r2	ILMN_2710274	*	*	ILMN_2710274	0.174	0.319	0.1818	0.328
	eQTL	Slco3a1	ILMN_1235635	-16.46	-35.34	ILMN_1235635	0.219	0.805	0.3859	0.62
	eQTL	Slco3a1	ILMN_1235735	*	*	ILMN_1235735	0.402	0.000628	0.001138	0.509
42	eQTL	Slco3a1	ILMN_2663230	13.38	0.68	ILMN_2663230	0.000565	3.29e-05	6.379e-08	0.365
	add	Snca	ILMN_3059393	*	*	ILMN_3059393	0.937	0.873	0.7163	0.393
	add	Snca	ILMN_3136638	*	*	ILMN_3136638	0.295	0.9	0.5863	0.709
43	add	Snca	ILMN_3161601	*	*	ILMN_3161601	0.71	0.695	0.7955	0.767

44	add	Spry2	ILMN_2749464	-0.27	2.29	ILMN_2749464	0.67	0.308	0.6795	NA
	add	Sumo2	ILMN_2715025	-8.62	-1.77	ILMN_2980331	0.938	0.123	0.221	0.703
45	add	Sumo2	ILMN_1221126	-0.44	7.72	no Ilum. Probe	NA	NA	NA	NA
	add	Tbp	ILMN_2626740	*	*	ILMN_2869461	0.991	0.112	0.4598	0.943
46	add	Tbp	ILMN_2613569	*	*	no Ilum. Probe	NA	NA	NA	NA
47	add	Tgfb1	ILMN_2711461	*	*	ILMN_2711461	0.048	0.0763	0.02671	0.00015
48	eQTL	Tmem176a	no Ilum. Probe	NA	NA	no Ilum. Probe	NA	NA	NA	NA
49	add	Tnf	ILMN_2467245	*	*	ILMN_2899863	*	*	*	*
50	add	Vhlh	ILMN_2518546	-119.09	-10.06	ILMN_2518546	0.244	0.392	0.7012	0.787

* expressed below background-signal
a Mouse array probe gene
b Illumina probe identifier for MouseRef-8 v1
c,d score for differential expression between resilient and stress susceptible groups of mice (for a P value of 0.05, DS=DiffScore $> \pm 13$)
e Illumina probe identifier for MouseRef-8 v2
f,g,h,i nominal P value for differentially regulated mRNA Expression of the ratio GR-stimulated/baseline

List of Abbreviations

1KGP 1,000 Genomes Project
5-HTR2A 5-hydroxytryptamine receptor 2A
5-HTTLPR 5-hydroxytriptamine-transporter-linked polymorphic region
ACC accuracy
ACTH adrenocorticotrophic hormone
ADHD attention deficit-hyperactivity disorder
AM amygdala
AP1 activating protein-1
ART Artifact detection tool
ASD autism spectrum disorder
ASN East Asian
ATP adenosine triphosphate
AUC area under the curve
BDI Beck depression invertury
BDNF brain-derived neurotrophic factor
BLA basolateral complex of the amygdala
BMI body mass index
BOLD blood oxygenation level-dependent
bp base pairs
BPD bipolar disorder
cAMP cyclic adenosine monophosphate
CD cross-disorder
CDA cross-disorder associations
cDNA complementary DNA
CeA central nucleus of the amygdala
CES-D center for epidemiological studies depression scale
CEU Utah Residents (CEPH) with Northern and Western European ancestry
Cg25 cingulate cortex
ChIA-PET chromatin interaction analysis by paired-end tag sequencing
ChIP chromatin immunoprecipitation
CI confidence interval
CMS chronic mild stress
COMT catechol-O-methyltransferase
CREB1 cAMP responsive element binding protein 1
CRF corticotropin releasing factor
CRH corticotrophin releasing hormone

CRHR1 corticotropin releasing hormone receptor 1
DAOA D-amino acid oxidase activator
DG dentate gyrus
DHSs deoxyribonuclease I hypersensitive sites
DISC-1 disrupted in schizophrenia-1
DNA deoxyribonucleic acid
DNaseI deoxyribonuclease I
DNS Duke Neurogenetics Study
DSM-IV-TR diagnostic and statistical manual of mental disorders forth edition, text revision
DST dexamethasone suppression test
EEG electroencephalography
ENCODE encyclopedia of DNA elements
eQTL bin set of eSNP bin probe combination
eQTL expression quantitative trait locus
eQTLs expression quantitative trait loci
eSNP bin set of eSNPs in LD
eSNP expression SNP
EUR-AM European-Americans
FDR false discovery rate
FKBP5 FK506 binding protein 5
fMRI functional magnetic resonance imaging
FN fase negatives
FP false positives
FST forced swim test
FWER family-wise error rate
GABA gamma-aminobutyric acid
GAD1 glutamate decarboxylase 1
GR glucocorticoid receptor
GREs glucocorticoid response elements
GRIA3 glutamate receptor, ionotropic, AMPA subunit 3
GRPS genetic risk profile score
GTEx Genotype-Tissue Expression
GWAS genome-wide association study
HAM-D Hamilton rating scale for depression
HC hippocampus
HLA human leukocyte antigen
HMM hidden Markov model
HPA hypothalamic-pituitary-adrenal
HWE Hardy-Weinberg equilibrium
ICD-10 international classification of diseases
kb kilo base pairs
LCLs lymphoblastoid cell lines

LD linkage disequilibrium
LH learned helplessness
LM linear models
MAF minor allele frequency
MAOA monoamine oxidase A
MARS Munich antidepressant response signature
Mb mega base pairs
MDD major depressive disorder
MDS multidimensional scaling
MEG magnetoencephalography
MHC major histocompatibility complex
miRNA microRNA
MPIP Max-Planck Institute of Psychiatry
mQTLs DNA methylation QTLs
MR mineralocorticoid receptor
MRI magnetic resonance imaging
mRNA messenger RNA
MuTHER Multiple Tissue Human Expression Resource
NaSSA noradrenergic and specific serotonergic antidepressant
NFκB nuclear factor kappa B
NR3C1 nuclear receptor subfamily 3, group C, member 1
NTRK2 neurotrophic tyrosine kinase, receptor, type 2
OOB out-of-bag
P2RX7 purinergic receptor P2X, ligand-gated ion channel, 7
PC principal components
PCA principal component analysis
PET positron emission tomography
PFC prefrontal cortex
PGC Psychiatric Genomics Consortium
PHQ-9 patient health questionnaire
qPCR quantitative real-time PCR
QTL quantitative trait locus
RA rheumatoid arthritis
RefSeq Reference Sequence
RF random forest
RIN RNA integrity number
RNA-seq RNA sequencing
RNA ribonucleic acid
ROI region of interest
SAGE serial analysis of gene expression
SCZ schizophrenia
SD standard deviation
SLC6A4 solute carrier family 6 (neurotransmitter transporter), member 4

SNP single nucleotide polymorphism
SNRI serotonin and noradrenaline re-uptake inhibitor
SSRE selective serotonin re-uptake enhancer
SSRI selective serotonin re-uptake inhibitor
TCA tricyclic antidepressants
TF transcription factors
TN true negatives
TNR true negative rate
TP true prosives
TPH2 tryptophan hydroxylase 2
TPR true positive rate
TST tail suspension test
VSN variance stabilization and normalization
WHO world health organisation

Bibliography

[1] G. R. Abecasis, E. Noguchi, A. Heinzmann, J. A. Traherne, S. Bhattacharyya, N. I. Leaves, G. G. Anderson, Y. Zhang, N. J. Lench, A. Carey, L. R. Cardon, M. F. Moffatt, and W. O. Cookson. Extent and distribution of linkage disequilibrium in three genomic regions. *American journal of human genetics*, 68(1):191–197, Jan. 2001.

[2] H. Abusamra. A Comparative Study of Feature Selection and Classification Methods for Gene Expression Data of Glioma. *Procedia Computer Science*, 2013.

[3] J. Alonso, M. C. Angermeyer, S. Bernert, R. Bruffaerts, T. S. Brugha, H. Bryson, G. de Girolamo, R. Graaf, K. Demyttenaere, I. Gasquet, J. M. Haro, S. J. Katz, R. C. Kessler, V. Kovess, J. P. Lépine, J. Ormel, G. Polidori, L. J. Russo, G. Vilagut, J. Almansa, S. Arbabzadeh-Bouchez, J. Autonell, M. Bernal, M. A. Buist-Bouwman, M. Codony, A. Domingo-Salvany, M. Ferrer, S. S. Joo, M. Martínez-Alonso, H. Matschinger, F. Mazzi, Z. Morgan, P. Morosini, C. Palacín, B. Romera, N. Taub, W. A. M. Vollebergh, and ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta psychiatrica Scandinavica*, 109(Suppl. 420):21–27, 2004.

[4] K. Amunts, O. Kedo, M. Kindler, P. Pieperhoff, H. Mohlberg, N. J. Shah, U. Habel, F. Schneider, and K. Zilles. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anatomy and embryology*, 210(5-6):343–352, Dec. 2005.

[5] L. Andrade, J. J. Caraveo-Anduaga, P. Berglund, R. V. Bijl, R. De Graaf, W. Vollebergh, E. Dragomirecka, R. Kohn, M. Keller, R. C. Kessler, N. Kawakami, C. Kiliç, D. Offord, T. B. Ustun, and H.-U. Wittchen. The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *International journal of methods in psychiatric research*, 12(1):3–21, 2003.

[6] A. P. Association. Diagnostic And Statistical Manual Of Mental Disorders DSM-IV-TR Fourth Edition (Text Revision) Author: American Psychiatr. *Amer Psychiatric Pub*, 2000.

[7] N. L. Barbosa-Morais, M. J. Dunning, A. S. Samarajiwa, J. F. J. Darot, M. E. Ritchie, A. G. Lynch, and S. Tavaré. A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. *Nucleic Acids Research*, 38, Jan. 2010.

[8] J. C. Barrett, B. Fry, J. Maller, and M. J. Daly. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2):263–265, Jan. 2005.

[9] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1):289–300, 1995.

[10] J. W. J. Bijlsma. Disease control with glucocorticoid therapy in rheumatoid arthritis. *Rheumatology (Oxford, England)*, 51 Suppl 4:iv9–iv13, June 2012.

[11] E. Binder, D. Salyakina, P. Lichtner, and G. Wochnik. Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. *Nature genetics*, 2004.

[12] E. B. Binder. The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology*, 34 Suppl 1:S186–95, Dec. 2009.

[13] K. Björkqvist. Social defeat as a stressor in humans. *Physiology & behavior*, 73(3):435–442, June 2001.

[14] E. Blaveri, F. Kelly, A. Mallei, K. Harris, A. Taylor, J. Reid, M. Razzoli, L. Carboni, C. Piubelli, L. Musazzi, G. Racagni, A. Mathé, M. Popoli, E. Domenici, and S. Bates. Expression profiling of a genetic animal model of depression reveals novel molecular pathways underlying depressive-like behaviours. *PloS one*, 5(9):e12596, 2010.

[15] A. Blomhoff, M. Olsson, S. Johansson, H. E. Akselsen, F. Pociot, J. Nerup, I. Kockum, A. Cambon-Thomsen, E. Thorsby, D. E. Undlien, and B. A. Lie. Linkage disequilibrium and haplotype blocks in the MHC vary in an HLA haplotype specific manner assessed mainly by DRB1*03 and DRB1*04 haplotypes. *Genes and immunity*, 7(2):130–140, Mar. 2006.

[16] R. Bogdan, D. E. Williamson, and A. R. Hariri. Mineralocorticoid receptor Iso/Val (rs5522) genotype moderates the association between previous childhood emotional neglect and amygdala reactivity. *American Journal of Psychiatry*, 169(5):515–522, May 2012.

[17] J. R. Bostwick and C. H. Le. Pharmacogenetics and Depression: Realized Dream or Great Expectation? *US Pharmacist*, 2011.

[18] M. P. Bowley, W. C. Drevets, D. Ongür, and J. L. Price. Low glial numbers in the amygdala in major depressive disorder. *Biological Psychiatry*, 52(5):404–412, Sept. 2002.

[19] R. G. Bradley, E. B. Binder, M. P. Epstein, Y. Tang, H. P. Nair, W. Liu, C. F. Gillespie, T. Berg, M. Evces, D. J. Newport, Z. N. Stowe, C. M. Heim, C. B. Nemeroff, A. Schwartz, J. F. Cubells, and K. J. Ressler. Influence of Child Abuse on Adult Depression: Moderation by the Corticotropin-Releasing Hormone Receptor Gene. *Archives of General Psychiatry*, 65(2):190–200, Feb. 2008.

[20] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct. 2001.

[21] J. C. Britton, S. Lissek, C. Grillon, M. A. Norcross, and D. S. Pine. Development of anxiety: the role of threat appraisal and fear learning. *Depression and anxiety*, 28(1):5–17, Jan. 2011.

[22] J. Brookfield. Q&A: promise and pitfalls of genome-wide association studies. *BMC biology*, 2010.

[23] V. M. Brown, K. S. Labar, C. C. Haswell, A. L. Gold, G. McCarthy, R. A. Morey, and M.-A. M. Workgrp. Altered Resting-State Functional Connectivity of Basolateral and Centromedial Amygdala Complexes in Posttraumatic Stress Disorder. *Neuropsychopharmacology*, 39(2):351–359, Jan. 2014.

[24] S. Browning. Multilocus association mapping using variable-length markov chains. *The American Journal of Human Genetics*, 2006.

[25] J. H. Byrne. *Neuroscience Online: An Electronic Textbook for the Neurosciences*. Department of Neurobiology and Anatomy; University of Texas Medical School at Houston, 1997.

[26] C. Cai, P. Langfelder, T. F. Fuller, M. C. Oldham, R. Luo, L. H. van den Berg, R. A. Ophoff, and S. Horvath. Is human blood a good surrogate for brain tissue in transcriptional studies? *BMC genomics*, 11:589, 2010.

[27] B. Carroll. Dexamethasone suppression test for depression. *Advances in biochemical psychopharmacology*, 1984.

[28] L. A. Carvalho and C. M. Pariante. In vitro modulation of the glucocorticoid receptor by antidepressants. *Stress (Amsterdam, Netherlands)*, 11(6):411–424, Nov. 2008.

[29] A. Caspi, A. R. Hariri, A. Holmes, R. Uher, and T. E. Moffitt. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *American Journal of Psychiatry*, 167(5):509–527, May 2010.

[30] A. Caspi, K. Sugden, T. E. Moffitt, A. Taylor, I. W. Craig, H. Harrington, J. McClay, J. Mill, J. Martin, A. Braithwaite, and R. Poulton. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science (New York, NY)*, 301(5631):386–389, July 2003.

[31] V. Cheung, L. Conlin, T. Weber, and M. Arcaro. Natural variation in human gene expression assessed in lymphoblastoid cells. *Nature genetics*, 2003.

[32] V. G. Cheung and R. S. Spielman. Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nature Publishing Group*, 10(9):595–604, Sept. 2009.

[33] H. Chun and S. Keles. Expression Quantitative Trait Loci Mapping With Multivariate Sparse Partial Least Squares Regression. *Genetics*, 2009.

[34] D. C. Clark, S. vonAmmon Cavanaugh, and R. D. Gibbons. The core symptoms of depression in medical and psychiatric patients. *The Journal of nervous and mental disease*, 171(12):705–713, Dec. 1983.

[35] S. D. Cohen, L. Norris, K. Acquaviva, R. A. Peterson, and P. L. Kimmel. Screening, diagnosis, and treatment of depression in patients with end-stage renal disease. *Clinical journal of the American Society of Nephrology : CJASN*, 2(6):1332–1342, Nov. 2007.

[36] W. Cookson, L. Liang, G. Abecasis, M. Moffatt, and M. Lathrop. Mapping complex disease traits with global gene expression. *Nature Publishing Group*, 10(3):184–194, Mar. 2009.

[37] Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet*, 381(9875):1371–1379, Apr. 2013.

[38] Cross-Disorder Group of the Psychiatric Genomics Consortium, S. H. Lee, S. Ripke, B. M. Neale, S. V. Faraone, S. M. Purcell, R. H. Perlis, B. J. Mowry, A. Thapar, M. E. Goddard, J. S. Witte, D. Absher, I. Agartz, H. Akil, F. Amin, O. A. Andreassen, A. Anjorin, R. Anney, V. Anttila, D. E. Arking, P. Asherson, M. H. Azevedo, L. Backlund, J. A. Badner, A. J. Bailey, T. Banaschewski, J. D. Barchas, M. R. Barnes, T. B. Barrett, N. Bass, A. Battaglia, M. Bauer, M. Bayés, F. Bellivier, S. E. Bergen, W. Berrettini, C. Betancur, T. Bettecken, J. Biederman, E. B. Binder, D. W. Black, D. H. R. Blackwood, C. S. Bloss, M. Boehnke, D. I. Boomsma, G. Breen, R. Breuer, R. Bruggeman, P. Cormican, N. G. Buccola, J. K. Buitelaar, W. E. Bunney, J. D. Buxbaum, W. F. Byerley, E. M. Byrne, S. Caesar, W. Cahn, R. M. Cantor, M. Casas, A. Chakravarti, K. Chambert, K. Choudhury, S. Cichon, C. R. Cloninger, D. A. Collier, E. H. Cook, H. Coon, B. Cormand, A. Corvin, W. H. Coryell, D. W. Craig, I. W. Craig, J. Crosbie, M. L. Cuccaro, D. Curtis, D. Czamara, S. Datta, G. Dawson, R. Day, E. J. De Geus, F. Degenhardt, S. Djurovic,

G. J. Donohoe, A. E. Doyle, J. Duan, F. Dudbridge, E. Duketis, R. P. Ebstein, H. J. Edenberg, J. Elia, S. Ennis, B. Etain, A. Fanous, A. E. Farmer, I. N. Ferrier, M. Flickinger, E. Fombonne, T. Foroud, J. Frank, B. Franke, C. Fraser, R. Freedman, N. B. Freimer, C. M. Freitag, M. Friedl, L. Frisén, L. Gallagher, P. V. Gejman, L. Georgieva, E. S. Gershon, D. H. Geschwind, I. Giegling, M. Gill, S. D. Gordon, K. Gordon-Smith, E. K. Green, T. A. Greenwood, D. E. Grice, M. Gross, D. Grozeva, W. Guan, H. Gurling, L. De Haan, J. L. Haines, H. Hakonarson, J. Hallmayer, S. P. Hamilton, M. L. Hamshere, T. F. Hansen, A. M. Hartmann, M. Hautzinger, A. C. Heath, A. K. Henders, S. Herms, I. B. Hickie, M. Hipolito, S. Hoefels, P. A. Holmans, F. Holsboer, W. J. Hoogendijk, J.-J. Hottenga, C. M. Hultman, V. Hus, A. Ingason, M. Ising, S. Jamain, E. G. Jones, I. Jones, L. Jones, J.-Y. Tzeng, A. K. Kähler, R. S. Kahn, R. Kandaswamy, M. C. Keller, J. L. Kennedy, E. Kenny, L. Kent, Y. Kim, G. K. Kirov, S. M. Klauck, L. Klei, J. A. Knowles, M. A. Kohli, D. L. Koller, B. Konte, A. Korszun, L. Krabbendam, R. Krasucki, J. Kuntsi, P. Kwan, M. Landén, N. Langstrom, M. Lathrop, J. Lawrence, W. B. Lawson, M. Leboyer, D. H. Ledbetter, P. H. Lee, T. Lencz, K.-P. Lesch, D. F. Levinson, C. M. Lewis, J. Li, P. Lichtenstein, J. A. Lieberman, D.-Y. Lin, D. H. Linszen, C. Liu, F. W. Lohoff, S. K. Loo, C. Lord, J. K. Lowe, S. Lucae, D. J. MacIntyre, P. A. F. Madden, E. Maestrini, P. K. E. Magnusson, P. B. Mahon, W. Maier, A. K. Malhotra, S. M. Mane, C. L. Martin, N. G. Martin, M. Mattheisen, K. Matthews, M. Mattingsdal, S. A. McCarroll, K. A. McGhee, J. J. McGough, P. J. McGrath, P. McGuffin, M. G. McInnis, A. McIntosh, R. McKinney, A. W. McLean, F. J. McMahon, W. M. McMahon, A. McQuillin, H. Medeiros, S. E. Medland, S. Meier, I. Melle, F. Meng, J. Meyer, C. M. Middeldorp, L. Middleton, V. Milanova, and A. Miranda. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics*, 45(9):984–994, Sept. 2013.

[39] B. Crosson, A. Ford, K. M. McGregor, M. Meinzer, S. Cheshkov, X. Li, D. Walker-Batson, and R. W. Briggs. Functional imaging and related techniques: an introduction for rehabilitation researchers. *Journal of rehabilitation research and development*, 47(2):vii–xxxiv, 2010.

[40] M. Davis and P. J. Whalen. The amygdala: vigilance and emotion. *Molecular Psychiatry*, 6(1):13–34, Jan. 2001.

[41] S. de Jong. *Expression QTL analysis of top loci from GWAS meta-analysis highlights additional schizophrenia candidate genes.* PhD thesis, University Medical Center Utrecht, Utrecht, The Netherlands, 2011.

[42] E. R. de Kloet, M. Joëls, and F. Holsboer. Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience*, 6(6):463–475, 2005.

[43] E. R. de Kloet, H. Karst, and M. Joëls. Corticosteroid hormones in the central stress response: quick-and-slow. *Frontiers in neuroendocrinology*, 29(2):268–272, May 2008.

[44] G. Deblois and V. Giguère. Nuclear receptor location analyses in mammalian genomes: from gene regulation to regulatory networks. *Molecular endocrinology (Baltimore, Md.)*, 22(9):1999–2011, Sept. 2008.

[45] J. M. Deussing. Animal models of depression. *Drug discovery today: disease models*, 3(4):375–383, Dec. 2006.

[46] B. Devlin and K. Roeder. Genomic control for association studies. *Biometrics*, 1999.

[47] A. Dimas, S. Deutsch, and B. Stranger. Common Regulatory Variation Impacts Gene Expression in a Cell Type–Dependent Manner. *Science (New York, NY)*, 2009.

[48] A. Dixon, L. Liang, M. Moffatt, W. Chen, and S. Heath. A genome-wide association study of global gene expression. *Nature genetics*, 2007.

[49] C. Dong, M.-L. Wong, and J. Licinio. Sequence variations of ABCB1, SLC6A2, SLC6A3, SLC6A4, CREB1, CRHR1 and NTRK2: association with major depression and antidepressant response in Mexican-Americans. *Molecular Psychiatry*, 14(12):1105–1118, Dec. 2009.

[50] L. Du, D. Bakish, A. Ravindran, and P. D. Hrdina. MAO-A gene polymorphisms are associated with major depression and sleep disturbance in males. *Neuroreport*, 15(13):2097–2101, Sept. 2004.

[51] S. Dudoit, J. Fridlyand, and T. P. Speed. Comparison of Discrimination Methods for the Classification of Tumors Using Gene Expression Data. *Journal of the American Statistical Association*, 97(457):77–87, Mar. 2002.

[52] C. H. Duman. Models of depression. *Vitamins and hormones*, 82:1–21, 2010.

[53] R. S. Duman and G. K. Aghajanian. Synaptic dysfunction in depression: potential therapeutic targets. *Science (New York, NY)*, 338(6103):68–72, Oct. 2012.

[54] R. M. Durbin, D. L. Altshuler, R. M. Durbin, G. R. Abecasis, D. R. Bentley, A. Chakravarti, A. G. Clark, F. S. Collins, F. M. De La Vega, P. Donnelly, M. Egholm, P. Flicek, S. B. Gabriel, R. A. Gibbs, B. M. Knoppers, E. S. Lander, H. Lehrach, E. R. Mardis, G. A. McVean, D. A. Nickerson, L. Peltonen, A. J. Schafer, S. T. Sherry, J. Wang, R. K. Wilson, R. A. Gibbs, D. Deiros, M. Metzker, D. Muzny, J. Reid, D. Wheeler, J. Wang, J. Li, M. Jian, G. Li, R. Li, H. Liang, G. Tian, B. Wang, J. Wang, W. Wang, H. Yang, X. Zhang, H. Zheng, E. S. Lander, D. L. Altshuler, L. Ambrogio, T. Bloom, K. Cibulskis, T. J. Fennell, S. B. Gabriel, D. B. Jaffe, E. Shefler, C. L. Sougnez, D. R. Bentley, N. Gormley, S. Humphray, Z. Kingsbury, P. Koko-Gonzales, J. Stone, K. J. McKernan, G. L. Costa, J. K. Ichikawa, C. C. Lee, R. Sudbrak, H. Lehrach, T. A. Borodina, A. Dahl, A. N. Davydov, P. Marquardt, F. Mertes, W. Nietfeld, P. Rosenstiel, S. Schreiber, A. V. Soldatov, B. Timmermann, M. Tolzmann, M. Egholm, J. Affourtit, D. Ashworth, S. Attiya,

M. Bachorski, E. Buglione, A. Burke, A. Caprio, C. Celone, S. Clark, D. Conners, B. Desany, L. Gu, L. Guccione, K. Kao, A. Kebbel, J. Knowlton, M. Labrecque, L. McDade, C. Mealmaker, M. Minderman, A. Nawrocki, F. Niazi, K. Pareja, R. Ramenani, D. Riches, W. Song, C. Turcotte, S. Wang, E. R. Mardis, R. K. Wilson, D. Dooling, L. Fulton, R. Fulton, G. Weinstock, R. M. Durbin, J. Burton, D. M. Carter, C. Churcher, A. Coffey, A. Cox, A. Palotie, M. Quail, T. Skelly, J. Stalker, H. P. Swerdlow, D. Turner, A. De Witte, S. Giles, R. A. Gibbs, D. Wheeler, M. Bainbridge, D. Challis, A. Sabo, F. Yu, J. Yu, J. Wang, X. Fang, X. Guo, R. Li, Y. Li, R. Luo, S. Tai, H. Wu, H. Zheng, X. Zheng, Y. Zhou, G. Li, J. Wang, H. Yang, G. T. Marth, E. P. Garrison, W. Huang, A. Indap, D. Kural, W.-P. Lee, W. Fung Leong, A. R. Quinlan, C. Stewart, M. P. Stromberg, A. N. Ward, J. Wu, C. Lee, R. E. Mills, X. Shi, M. J. Daly, M. A. DePristo, D. L. Altshuler, A. D. Ball, E. Banks, T. Bloom, B. L. Browning, K. Cibulskis, T. J. Fennell, K. V. Garimella, S. R. Grossman, R. E. Handsaker, M. Hanna, C. Hartl, D. B. Jaffe, A. M. Kernytsky, J. M. Korn, H. Li, J. R. Maguire, S. A. McCarroll, A. McKenna, J. C. Nemesh, A. A. Philippakis, R. E. Poplin, A. Price, M. A. Rivas, P. C. Sabeti, S. F. Schaffner, E. Shefler, I. A. Shlyakhter, D. N. Cooper, E. V. Ball, M. Mort, A. D. Phillips, P. D. Stenson, J. Sebat, V. Makarov, K. Ye, S. C. Yoon, C. D. Bustamante, A. G. Clark, A. Boyko, J. Degenhardt, S. Gravel, R. N. Gutenkunst, M. Kaganovich, A. Keinan, P. Lacroute, X. Ma, A. Reynolds, L. Clarke, P. Flicek, F. Cunningham, J. Herrero, S. Keenen, E. Kulesha, R. Leinonen, W. M. McLaren, R. Radhakrishnan, R. E. Smith, V. Zalunin, X. Zheng-Bradley, J. O. Korbel, A. M. Stütz, S. Humphray, M. Bauer, R. Keira Cheetham, T. Cox, M. Eberle, T. James, S. Kahn, L. Murray, A. Chakravarti, K. Ye, F. M. De La Vega, Y. Fu, F. C. L. Hyland, J. M. Manning, S. F. McLaughlin, and H. Peckham. A map of human genome variation from population-scale sequencing. *Nature*, 467(7319):1061–1073, Oct. 2010.

[55] V. Duric, M. Banasr, P. Licznerski, H. D. Schmidt, C. A. Stockmeier, A. A. Simen, S. S. Newton, and R. S. Duman. A negative regulator of MAP kinase causes depressive behavior. *Nature medicine*, 16(11):1328–1332, Nov. 2010.

[56] V. Emilsson, G. Thorleifsson, B. Zhang, and A. Leonardson. Genetics of gene expression and its effect on disease. *Nature*, 2008.

[57] ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS biology*, 9(4):e1001046, Apr. 2011.

[58] ENCODE Project Consortium, B. E. Bernstein, E. Birney, I. Dunham, E. D. Green, C. Gunter, and M. Snyder. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414):57–74, Sept. 2012.

[59] A. Etkin, K. C. Klemenhagen, J. T. Dudman, M. T. Rogan, R. Hen, E. R. Kandel, and J. Hirsch. Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron*, 44(6):1043–1055, 2004.

[60] E. Evangelou and J. P. A. Ioannidis. Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews Genetics*, 14(6):379–389, June 2013.

[61] M. B. First, M. Gibbon, R. L. Spitzer, J. W. B. Williams, and L. S. Benjamin. Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II). *Washington, D.C.: American Psychiatric Press, Inc*, 1997.

[62] J. Flint and K. S. Kendler. The genetics of major depression. *Neuron*, 81(3):484–503, Feb. 2014.

[63] T. Flutre, X. Wen, J. Pritchard, and M. Stephens. A statistical framework for joint eQTL analysis in multiple tissues. *PLoS genetics*, 9(5):e1003486, May 2013.

[64] K. A. Frazer, S. S. Murray, N. J. Schork, and E. J. Topol. Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, 10(4):241–251, Apr. 2009.

[65] T. Frodl, H.-J. Möller, and E. Meisenzahl. Neuroimaging genetics: new perspectives in research on major depression? *Acta psychiatrica Scandinavica*, 118(5):363–372, Nov. 2008.

[66] H. Funato, A. Kobayashi, and Y. Watanabe. Differential effects of antidepressants on dexamethasone-induced nuclear translocation and expression of glucocorticoid receptor. *Brain research*, 1117(1):125–134, Oct. 2006.

[67] T. S. Furey. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nature Reviews Genetics*, 13(12):840–852, Dec. 2012.

[68] E. R. Gamazon, J. A. Badner, L. Cheng, C. Zhang, D. Zhang, N. J. Cox, E. S. Gershon, J. R. Kelsoe, T. A. Greenwood, C. M. Nievergelt, C. Chen, R. McKinney, P. D. Shilling, N. J. Schork, E. N. Smith, C. S. Bloss, J. I. Nurnberger, H. J. Edenberg, T. Foroud, D. L. Koller, W. A. Scheftner, W. Coryell, J. Rice, W. B. Lawson, E. A. Nwulia, M. Hipolito, W. Byerley, F. J. McMahon, T. G. Schulze, W. H. Berrettini, J. B. Potash, P. P. Zandi, P. B. Mahon, M. G. McInnis, S. Zöllner, P. Zhang, D. W. Craig, S. Szelinger, T. B. Barrett, and C. Liu. Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Molecular Psychiatry*, Jan. 2012.

[69] X. Gao, T. Haritunians, P. Marjoram, R. McKean-Cowdin, M. Torres, K. D. Taylor, J. I. Rotter, W. J. Gauderman, and R. Varma. Genotype Imputation for Latinos Using the HapMap and 1000 Genomes Project Reference Panels. *Frontiers in genetics*, 3:117, 2012.

[70] P. G. Giresi, J. Kim, R. M. McDaniell, V. R. Iyer, and J. D. Lieb. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Research*, 17(6):877–885, June 2007.

[71] A. Gladkevich, H. F. Kauffman, and J. Korf. Lymphocytes as a neural probe: potential for studying psychiatric disorders. *Progress in neuro-psychopharmacology & biological psychiatry*, 28(3):559–576, May 2004.

[72] T. Goltser-Dubner, E. Galili-Weisstub, and R. H. Segman. Genetics of unipolar major depressive disorder. *The Israel journal of psychiatry and related sciences*, 47(1):72–82, 2010.

[73] H. Göring, J. Curran, M. Johnson, and T. Dyer. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nature genetics*, 2007.

[74] L. Grapes, M. Firat, J. Dekkers, and M. Rothschild. Optimal haplotype structure for linkage disequilibrium-based fine mapping of quantitative trait loci using identity by descent. *Genetics*, 2006.

[75] E. Grundberg, V. Adoue, T. Kwan, B. Ge, Q. L. Duan, K. C. L. Lam, V. Koka, A. Kindmark, S. T. Weiss, K. Tantisira, H. Mallmin, B. A. Raby, O. Nilsson, and T. Pastinen. Global analysis of the impact of environmental perturbation on cis-regulation of gene expression. *PLoS genetics*, 7(1):e1001279, 2011.

[76] E. Grundberg, K. S. Small, Å. K. Hedman, A. C. Nica, A. Buil, S. Keildson, J. T. Bell, T.-P. Yang, E. Meduri, A. Barrett, J. Nisbett, M. Sekowska, A. Wilk, S.-Y. Shin, D. Glass, M. Travers, J. L. Min, S. Ring, K. Ho, G. Thorleifsson, A. Kong, U. Thorsteindottir, C. Ainali, A. S. Dimas, N. Hassanali, C. Ingle, D. Knowles, M. Krestyaninova, C. E. Lowe, P. di Meglio, S. B. Montgomery, L. Parts, S. Potter, G. Surdulescu, L. Tsaprouni, S. Tsoka, V. Bataille, R. Durbin, F. O. Nestle, S. O'Rahilly, N. Soranzo, C. M. Lindgren, K. T. Zondervan, K. R. Ahmadi, E. E. Schadt, K. Stefansson, G. D. Smith, M. I. McCarthy, P. Deloukas, E. T. Dermitzakis, T. D. Spector, and Multiple Tissue Human Expression Resource (MuTHER) Consortium. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nature genetics*, 44(10):1084–1089, Oct. 2012.

[77] GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nature genetics*, 45(6):580–585, June 2013.

[78] R. E. Gur and R. C. Gur. Functional magnetic resonance imaging in schizophrenia. *Dialogues in clinical neuroscience*, 12(3):333–343, 2010.

[79] O. Hakim, M.-H. Sung, T. C. Voss, E. Splinter, S. John, P. J. Sabo, R. E. Thurman, J. A. Stamatoyannopoulos, W. de Laat, and G. L. Hager. Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. *Genome Research*, 21(5):697–706, May 2011.

[80] E. Halperin, G. Kimmel, and R. Shamir. Tag SNP selection in genotype data for maximizing SNP prediction accuracy. *Bioinformatics*, 21 Suppl 1:i195–203, June 2005.

[81] R. Hashimoto, T. Numakawa, T. Ohnishi, E. Kumamaru, Y. Yagasaki, T. Ishimoto, T. Mori, K. Nemoto, N. Adachi, A. Izumi, S. Chiba, H. Noguchi, T. Suzuki, N. Iwata, N. Ozaki, T. Taguchi, A. Kamiya, A. Kosuga, M. Tatsumi, K. Kamijima, D. R. Weinberger, A. Sawa, and H. Kunugi. Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. *Human molecular genetics*, 15(20):3024–3033, 2006.

[82] G. Hasler, W. C. Drevets, H. K. Manji, and D. S. Charney. Discovering endophenotypes for major depression. *Neuropsychopharmacology*, 29(10):1765–1781, Oct. 2004.

[83] M. E. Hawley and K. K. Kidd. HAPLO: a program using the EM algorithm to estimate the frequencies of multi-site haplotypes. *J Hered*, 86(5):409–411, 1995.

[84] A. Heils, A. Teufel, S. Petri, G. Stöber, P. Riederer, D. Bengel, and K. P. Lesch. Allelic variation of human serotonin transporter gene expression. *Journal of neurochemistry*, 66(6):2621–2624, June 1996.

[85] C. Heim and C. B. Nemeroff. The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. *Biological Psychiatry*, 49(12):1023–1039, June 2001.

[86] A. Heinz, D. W. Jones, C. Mazzanti, D. Goldman, P. Ragan, D. Hommer, M. Linnoila, and D. R. Weinberger. A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. *Biological Psychiatry*, 47(7):643–649, Apr. 2000.

[87] A. Heiske, J. Jesberg, J.-C. Krieg, and H. Vedder. Differential effects of antidepressants on glucocorticoid receptors in human primary blood cells and human monocytic U-937 cells. *Neuropsychopharmacology*, 28(4):807–817, Apr. 2003.

[88] J. M. Hennings, T. Owashi, E. B. Binder, S. Horstmann, A. Menke, S. Kloiber, T. Dose, B. Wollweber, D. Spieler, T. Messer, R. Lutz, H. Künzel, T. Bierner, T. Pollmächer, H. Pfister, T. Nickel, A. Sonntag, M. Uhr, M. Ising, F. Holsboer, and S. Lucae. Clinical characteristics and treatment outcome in a representative sample of depressed inpatients - findings from the Munich Antidepressant Response Signature (MARS) project. *J Psychiatr Res*, 43(3):215–229, Jan. 2009.

[89] M. Héry, A. Sémont, M. P. Fache, M. Faudon, and F. Héry. The effects of serotonin on glucocorticoid receptor binding in rat raphe nuclei and hippocampal cells in culture. *Journal of neurochemistry*, 74(1):406–413, Jan. 2000.

[90] J. M. Hettema. Genetics of Depression. *Focus*, 8(3):316–322, July 2010.

[91] J. M. Hettema, S. S. An, M. C. Neale, J. Bukszar, E. J. C. G. van den Oord, K. S. Kendler, and X. Chen. Association between glutamic acid decarboxylase genes and anxiety disorders, major depression, and neuroticism. *Molecular Psychiatry*, 11(8):752–762, Aug. 2006.

[92] R. M. Hirschfeld. History and evolution of the monoamine hypothesis of depression. *The Journal of clinical psychiatry*, 61 Suppl 6:4–6, 2000.

[93] R. Horton, L. Wilming, V. Rand, R. Lovering, E. Bruford, V. Khodiyar, M. Lush, S. Povey, C. Talbot, M. WrighO, H. Wain, J. Trowsdale, A. Ziegler, and S. Beck. Gene map of the extended human MHC : Abstract : Nature Reviews Genetics. *Nature Publishing Group*, 5(12):889–899, 2004.

[94] B. N. Howie, P. Donnelly, and J. Marchini. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics*, 5(6):e1000529, June 2009.

[95] International HapMap Consortium. The International HapMap Project. *Nature*, 426(6968):789–796, Dec. 2003.

[96] International HapMap Consortium. A haplotype map of the human genome. *Nature*, 437(7063):1299–1320, Oct. 2005.

[97] M. Ising, S. Lucae, E. B. Binder, T. Bettecken, M. Uhr, S. Ripke, M. A. Kohli, J. M. Hennings, S. Horstmann, S. Kloiber, A. Menke, B. Bondy, R. Rupprecht, K. Domschke, B. T. Baune, V. Arolt, A. J. Rush, F. Holsboer, and B. Müller-Myhsok. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. *Archives of General Psychiatry*, 66(9):966–975, Sept. 2009.

[98] T. Jääskeläinen, H. Makkonen, and J. J. Palvimo. Steroid up-regulation of FKBP51 and its role in hormone signaling. *Current Opinion in Pharmacology*, 11(4):326–331, Aug. 2011.

[99] A. J. Jasinska, S. Service, O.-W. Choi, J. DeYoung, O. Grujic, S.-Y. Kong, M. J. Jorgensen, J. Bailey, S. Breidenthal, L. A. Fairbanks, R. P. Woods, J. D. Jentsch, and N. B. Freimer. Identification of brain transcriptional variation reproduced in peripheral blood: an approach for mapping brain expression traits. *Human molecular genetics*, 18(22):4415–4427, 2009.

[100] S. John, P. J. Sabo, R. E. Thurman, M.-H. Sung, S. C. Biddie, T. A. Johnson, G. L. Hager, and J. A. Stamatoyannopoulos. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nature genetics*, 43(3):264–268, Jan. 2011.

[101] G. C. L. Johnson, L. Esposito, B. J. Barratt, A. N. Smith, J. Heward, G. D. Genova, H. Ueda, H. J. Cordell, I. A. Eaves, F. Dudbridge, R. C. J. Twells, F. Payne, W. Hughes, S. Nutland, H. Stevens, P. Carr, E. Tuomilehto-Wolf, J. Tuomilehto, S. C. L. Gough, D. G. Clayton, and J. A. Todd. Haplotype tagging for the identification of common disease genes - Nature Genetics. *Nature genetics*, 29(2):233–237, Oct. 2001.

[102] W. E. Johnson and L. Cheng. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, 8:118–127, Jan. 2007.

[103] O. Kassel and P. Herrlich. Crosstalk between the glucocorticoid receptor and other transcription factors: molecular aspects. *Molecular and cellular endocrinology*, 275(1-2):13–29, Sept. 2007.

[104] K. S. Kendler, M. Gatz, C. O. Gardner, and N. L. Pedersen. A Swedish national twin study of lifetime major depression. *The American journal of psychiatry*, 163(1):109–114, Jan. 2006.

[105] K. S. Kendler and C. A. Prescott. A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry*, 56(1):39–44, Jan. 1999.

[106] R. Kessler. Sex and depression in the National Comorbidity Survey I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective disorders*, 29(2-3):85–96, Nov. 1993.

[107] R. C. Kessler, P. Berglund, O. Demler, R. Jin, D. Koretz, K. R. Merikangas, A. J. Rush, E. E. Walters, P. S. Wang, and National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA : the journal of the American Medical Association*, 289(23):3095–3105, June 2003.

[108] R. C. Kessler, P. Berglund, O. Demler, R. Jin, K. R. Merikangas, and E. E. Walters. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6):593–602, June 2005.

[109] G. D. Kitsios and E. Zintzaras. Genome-wide association studies: hypothesis-"free" or "engaged"? *Translational Research*, 2009.

[110] M. A. Kohli, S. Lucae, P. G. Saemann, M. V. Schmidt, A. Demirkan, K. Hek, D. Czamara, M. Alexander, D. Salyakina, S. Ripke, D. Hoehn, M. Specht, A. Menke, J. Hennings, A. Heck, C. Wolf, M. Ising, S. Schreiber, M. Czisch, M. B. Müller, M. Uhr, T. Bettecken, A. Becker, J. Schramm, M. Rietschel, W. Maier, B. Bradley, K. J. Ressler, M. M. Nöthen, S. Cichon, I. W. Craig, G. Breen, C. M. Lewis, A. Hofman, H. Tiemeier, C. M. van Duijn, F. Holsboer, B. Müller-Myhsok, and E. B. Binder. The neuronal transporter gene SLC6A15 confers risk to major depression. *Neuron*, 70(2):252–265, Apr. 2011.

[111] J. A. Krawiec, H. Chen, S. Alom-Ruiz, and M. Jaye. Modified PAXgene (TM) method allows for isolation of high-integrity total RNA from microlitre volumes of mouse whole blood. *Laboratory Animals*, 43(4):394–398, Oct. 2009.

[112] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra. Circos: an information aesthetic for comparative genomics. *Genome Research*, 19(9):1639–1645, Sept. 2009.

[113] M. Lai, J. A. McCormick, K. E. Chapman, P. A. T. Kelly, J. R. Seckl, and J. L. W. Yau. Differential regulation of corticosteroid receptors by monoamine neurotransmitters and antidepressant drugs in primary hippocampal culture. *Neuroscience*, 118(4):975–984, 2003.

[114] G. Laje, S. Paddock, H. Manji, A. J. Rush, A. F. Wilson, D. Charney, and F. J. McMahon. Genetic markers of suicidal ideation emerging during citalopram treatment of major depression. *The American journal of psychiatry*, 164(10):1530–1538, Oct. 2007.

[115] M. Lechner, V. Höhn, B. Brauner, I. Dunger, G. Fobo, G. Frishman, C. Montrone, G. Kastenmüller, B. Waegele, and A. Ruepp. CIDeR: multifactorial interaction networks in human diseases. *Genome biology*, 13(7):R62, July 2012.

[116] J. LeDoux. The amygdala. *Current Biology*, 17(20):R868–R874, 2007.

[117] S. Lee, J. Jeong, Y. Kwak, and S. K. Park. Depression research: where are we now? *Molecular brain*, 2010.

[118] J. T. Leek, R. B. Scharpf, H. C. Bravo, D. Simcha, B. Langmead, W. E. Johnson, D. Geman, K. Baggerly, and R. A. Irizarry. Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Publishing Group*, 11(10):733–739, Oct. 2010.

[119] J. T. Leek and J. D. Storey. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS genetics*, 3(9):1724–1735, Sept. 2007.

[120] Y. Lerner, N. Singer, T. Gonen, Y. Weintraub, O. Cohen, N. Rubin, L. G. Ungerleider, and T. Hendler. Feeling without Seeing? Engagement of Ventral, but Not Dorsal, Amygdala during Unaware Exposure to Emotional Faces. *Journal of Cognitive Neuroscience*, 24(3):531–542, Mar. 2012.

[121] G. Li, M. J. Fullwood, H. Xu, F. H. Mulawadi, S. Velkov, V. Vega, P. N. Ariyaratne, Y. B. Mohamed, H.-S. Ooi, C. Tennakoon, C.-L. Wei, Y. Ruan, and W.-K. Sung. ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. *Genome biology*, 11(2):R22, 2010.

[122] G. Li, X. Ruan, R. K. Auerbach, K. S. Sandhu, M. Zheng, P. Wang, H. M. Poh, Y. Goh, J. Lim, J. Zhang, H. S. Sim, S. Q. Peh, F. H. Mulawadi, C. T. Ong, Y. L. Orlov, S. Hong, Z. Zhang, S. Landt, D. Raha, G. Euskirchen, C.-L. Wei, W. Ge, H. Wang, C. Davis, K. I. Fisher-Aylor, A. Mortazavi, M. Gerstein, T. Gingeras, B. Wold, Y. Sun, M. J. Fullwood, E. Cheung, E. Liu, W.-K. Sung, M. Snyder, and

Y. Ruan. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell*, 148(1-2):84–98, Jan. 2012.

[123] Y. Li, C. J. Willer, J. Ding, P. Scheet, and G. R. Abecasis. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic epidemiology*, 34(8):816–834, Dec. 2010.

[124] A. Liaw and M. Wiener. Classification and Regression by randomForest. *R news*, 2002.

[125] S. M. Lin, P. Du, W. Huber, and W. A. Kibbe. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Research*, 36(2):e11, Feb. 2008.

[126] Z. Liu, F. Zhu, G. Wang, Z. Xiao, H. Wang, J. Tang, X. Wang, D. Qiu, W. Liu, Z. Cao, and W. Li. Association of corticotropin-releasing hormone receptor1 gene SNP and haplotype with major depression. *Neuroscience letters*, 404(3):358–362, Sept. 2006.

[127] M. T. Lowy, A. T. Reder, J. P. Antel, and H. Y. Meltzer. Glucocorticoid resistance in depression: the dexamethasone suppression test and lymphocyte sensitivity to dexamethasone. *The American journal of psychiatry*, 141(11):1365–1370, Nov. 1984.

[128] S. Lucae, D. Salyakina, N. Barden, M. Harvey, B. Gagne, M. Labbe, E. Binder, M. Uhr, M. Paez-Pereda, I. Sillaber, M. Ising, T. Brueckl, R. Lieb, F. Holsboer, and B. Mueller-Myhsok. P2RX7, a gene coding for a purinergic ligand-gated ion channel, is associated with major depressive disorder. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*, 141B(7):740–740, 2006.

[129] I. Lucki. The forced swimming test as a model for core and component behavioral effects of antidepressant drugs. *Behavioural pharmacology*, 8(6-7):523–532, Nov. 1997.

[130] J. Luo, M. Schumacher, A. Scherer, D. Sanoudou, D. Megherbi, T. Davison, T. Shi, W. Tong, L. Shi, H. Hong, C. Zhao, F. Elloumi, W. Shi, R. Thomas, S. Lin, G. Tillinghast, G. Liu, Y. Zhou, D. Herman, Y. Li, Y. Deng, H. Fang, P. Bushel, M. Woods, and J. Zhang. A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. *Pharmacogenomics Journal*, 10(4):S48–S61, Oct. 2010.

[131] G. M. MacQueen, S. Campbell, B. S. McEwen, K. Macdonald, S. Amano, R. T. Joffe, C. Nahmias, and L. T. Young. Course of illness, hippocampal function, and hippocampal volume in major depression. *Proceedings of the National Academy of Sciences of the United States of America*, 100(3):1387–1392, Feb. 2003.

[132] P. Madrigal and P. Krajewski. Current bioinformatic approaches to identify DNase I hypersensitive sites and genomic footprints from DNase-seq data. *Frontiers in genetics*, 3:230, 2012.

[133] Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium. A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, Apr. 2012.

[134] J. A. Maldjian, P. J. Laurienti, R. A. Kraft, and J. H. Burdette. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3):1233–1239, July 2003.

[135] T. Manke, H. G. Roider, and M. Vingron. Statistical modeling of transcription factor binding affinities predicts regulatory interactions. *Plos Computational Biology*, 4(3):– , 2008.

[136] T. A. Manolio, F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, J. H. Cho, A. E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C. N. Rotimi, M. Slatkin, D. Valle, A. S. Whittemore, M. Boehnke, A. G. Clark, E. E. Eichler, G. Gibson, J. L. Haines, T. F. C. Mackay, S. A. McCarroll, and P. M. Visscher. Finding the missing heritability of complex diseases. *Nature*, 461(7265):747–753, Oct. 2009.

[137] J. C. Maranville, F. Luca, A. L. Richards, X. Wen, D. B. Witonsky, S. Baxter, M. Stephens, and A. Di Rienzo. Interactions between glucocorticoid treatment and cis-regulatory polymorphisms contribute to cellular response phenotypes. *PLoS genetics*, 7(7):e1002162–, July 2011.

[138] J. Marchini and B. Howie. Genotype imputation for genome-wide association studies. *Nature Publishing Group*, 11(7):499–511, Feb. 2010.

[139] J. Marchini, B. Howie, S. Myers, G. McVean, and P. Donnelly. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature genetics*, 39(7):906–913, July 2007.

[140] W. N. Marsden. Synaptic plasticity in depression: molecular, cellular and functional correlates. *Progress in neuro-psychopharmacology & biological psychiatry*, 43:168–184, June 2013.

[141] I. Massat, D. Souery, J. Del-Favero, M. Nothen, D. Blackwood, W. Muir, R. Kaneva, A. Serretti, C. Lorenzi, M. Rietschel, V. Milanova, G. N. Papadimitriou, D. Dikeos, C. Van Broekhoven, and J. Mendlewicz. Association between COMT (Val158Met) functional polymorphism and early onset in patients with major depressive disorder in a European multicenter genetic association study. *Molecular Psychiatry*, 10(6):598–605, June 2005.

[142] M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P. Reynolds, R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutyavin, S. Stehling-Sun, A. K. Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen, S. Neph, P. J. Sabo, S. Heimfeld, A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, and J. A. Stamatoyannopoulos. Systematic localization of common disease-associated variation in regulatory DNA. *Science (New York, NY)*, 337(6099):1190–1195, Sept. 2012.

[143] M. I. McCarthy, G. R. Abecasis, L. R. Cardon, D. B. Goldstein, J. Little, J. P. A. Ioannidis, and J. N. Hirschhorn. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9(5):356–369, May 2008.

[144] J. McCauley, D. E. Kern, K. Kolodner, L. Dill, A. F. Schroeder, H. K. DeChant, J. Ryden, L. R. Derogatis, and E. B. Bass. Clinical characteristics of women with a history of childhood abuse: unhealed wounds. *JAMA : the journal of the American Medical Association*, 277(17):1362–1368, May 1997.

[145] L. I. McKay and J. A. Cidlowski. Molecular control of immune/inflammatory responses: interactions between nuclear factor-kappa B and steroid receptor-signaling pathways. *Endocrine reviews*, 20(4):435–459, Aug. 1999.

[146] D. Mehta and E. Binder. Gene x Environment vulnerability factors for PTSD: The HPA axis. *Neuropharmacology*, 2011.

[147] D. Mehta, A. Menke, and E. B. Binder. Gene expression studies in major depression. *Current psychiatry reports*, 12(2):135–144, Apr. 2010.

[148] A. Menke, J. Arloth, B. Pütz, P. Weber, T. Klengel, D. Mehta, M. Gonik, M. Rex-Haffner, J. Rubel, M. Uhr, S. Lucae, J. M. Deussing, B. Müller-Myhsok, F. Holsboer, and E. B. Binder. Dexamethasone Stimulated Gene Expression in Peripheral Blood is a Sensitive Marker for Glucocorticoid Receptor Resistance in Depressed Patients. *Neuropsychopharmacology*, 37:1455–1464, Jan. 2012.

[149] A. H. Miller, V. Maletic, and C. L. Raison. Inflammation and its discontents: the role of cytokines in the pathophysiology of major depression. *Biological Psychiatry*, 65(9):732–741, May 2009.

[150] J. L. Min, J. M. Taylor, J. B. Richards, T. Watts, F. H. Pettersson, J. Broxholme, K. R. Ahmadi, G. L. Surdulescu, E. Lowy, C. Gieger, C. Newton-Cheh, M. Perola, N. Soranzo, I. Surakka, C. M. Lindgren, J. Ragoussis, A. P. Morris, L. R. Cardon, T. D. Spector, and K. T. Zondervan. The Use of Genome-Wide eQTL Associations in Lymphoblastoid Cell Lines to Identify Novel Genetic Pathways Involved in Complex Traits. *PloS one*, 6(7):e22070, July 2011.

[151] M. Mistry. *Meta-analyses of expression profiling data in the postmortem human brain.* PhD thesis, University of British Columbia, Vancouver, Canada, 2012.

[152] M. Mistry, J. Gillis, and P. Pavlidis. Meta-analysis of gene coexpression networks in the post-mortem prefrontal cortex of patients with schizophrenia and unaffected controls. *BMC neuroscience*, 14:105, 2013.

[153] N. D. Mitchell and G. B. Baker. An update on the role of glutamate in the pathophysiology of depression. *Acta psychiatrica Scandinavica*, 122(3):192–210, Sept. 2010.

[154] S. Modell, A. Yassouridis, J. Huber, and F. Holsboer. Corticosteroid receptor function is decreased in depressed patients. *Neuroendocrinology*, 65(3):216–222, Mar. 1997.

[155] S. B. Montgomery and E. T. Dermitzakis. From expression QTLs to personalized transcriptomics. *Nature Reviews Genetics*, 12(4):277–282, 2011.

[156] C. J. L. Murray, T. Vos, R. Lozano, M. Naghavi, A. D. Flaxman, C. Michaud, M. Ezzati, K. Shibuya, J. A. Salomon, S. Abdalla, V. Aboyans, J. Abraham, I. Ackerman, R. Aggarwal, S. Y. Ahn, M. K. Ali, M. Alvarado, H. R. Anderson, L. M. Anderson, K. G. Andrews, C. Atkinson, L. M. Baddour, A. N. Bahalim, S. Barker-Collo, L. H. Barrero, D. H. Bartels, M.-G. Basáñez, A. Baxter, M. L. Bell, E. J. Benjamin, D. Bennett, E. Bernabé, K. Bhalla, B. Bhandari, B. Bikbov, A. Bin Abdulhak, G. Birbeck, J. A. Black, H. Blencowe, J. D. Blore, F. Blyth, I. Bolliger, A. Bonaventure, S. Boufous, R. Bourne, M. Boussinesq, T. Braithwaite, C. Brayne, L. Bridgett, S. Brooker, P. Brooks, T. S. Brugha, C. Bryan-Hancock, C. Bucello, R. Buchbinder, G. Buckle, C. M. Budke, M. Burch, P. Burney, R. Burstein, B. Calabria, B. Campbell, C. E. Canter, H. Carabin, J. Carapetis, L. Carmona, C. Cella, F. Charlson, H. Chen, A. T.-A. Cheng, D. Chou, S. S. Chugh, L. E. Coffeng, S. D. Colan, S. Colquhoun, K. E. Colson, J. Condon, M. D. Connor, L. T. Cooper, M. Corriere, M. Cortinovis, K. C. de Vaccaro, W. Couser, B. C. Cowie, M. H. Criqui, M. Cross, K. C. Dabhadkar, M. Dahiya, N. Dahodwala, J. Damsere-Derry, G. Danaei, A. Davis, D. De Leo, L. Degenhardt, R. Dellavalle, A. Delossantos, J. Denenberg, S. Derrett, D. C. Des Jarlais, S. D. Dharmaratne, M. Dherani, C. Diaz-Torne, H. Dolk, E. R. Dorsey, T. Driscoll, H. Duber, B. Ebel, K. Edmond, A. Elbaz, S. E. Ali, H. Erskine, P. J. Erwin, P. Espindola, S. E. Ewoigbokhan, F. Farzadfar, V. Feigin, D. T. Felson, A. Ferrari, C. P. Ferri, E. M. Fèvre, M. M. Finucane, S. Flaxman, L. Flood, K. Foreman, M. H. Forouzanfar, F. G. R. Fowkes, M. Fransen, M. K. Freeman, B. J. Gabbe, S. E. Gabriel, E. Gakidou, H. A. Ganatra, B. Garcia, F. Gaspari, R. F. Gillum, G. Gmel, D. Gonzalez-Medina, R. Gosselin, R. Grainger, B. Grant, J. Groeger, F. Guillemin, D. Gunnell, R. Gupta, J. Haagsma, H. Hagan, Y. A. Halasa, W. Hall, D. Haring, J. M. Haro, J. E. Harrison, R. Havmoeller, R. J. Hay, H. Higashi, C. Hill, B. Hoen, H. Hoffman, P. J. Hotez, D. Hoy, J. J. Huang, S. E. Ibeanusi, K. H. Jacobsen, S. L. James, D. Jarvis, R. Jasrasaria, S. Jayaraman, N. Johns, J. B. Jonas, G. Karthikeyan, N. Kassebaum,

N. Kawakami, A. Keren, J.-P. Khoo, C. H. King, L. M. Knowlton, O. Kobusingye, A. Koranteng, R. Krishnamurthi, F. Laden, R. Lalloo, L. L. Laslett, T. Lathlean, J. L. Leasher, Y. Y. Lee, J. Leigh, D. Levinson, S. S. Lim, E. Limb, J. K. Lin, M. Lipnick, S. E. Lipshultz, W. Liu, M. Loane, S. L. Ohno, R. Lyons, J. Mabweijano, M. F. MacIntyre, R. Malekzadeh, L. Mallinger, S. Manivannan, W. Marcenes, L. March, D. J. Margolis, G. B. Marks, R. Marks, A. Matsumori, R. Matzopoulos, B. M. Mayosi, J. H. McAnulty, M. M. McDermott, N. McGill, J. McGrath, M. E. Medina-Mora, M. Meltzer, G. A. Mensah, T. R. Merriman, A.-C. Meyer, V. Miglioli, M. Miller, T. R. Miller, P. B. Mitchell, C. Mock, A. O. Mocumbi, T. E. Moffitt, A. A. Mokdad, L. Monasta, M. Montico, and Moradi-... Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380(9859):2197–2223, Dec. 2012.

[157] A. J. Myers, J. R. Gibbs, J. A. Webster, K. Rohrer, A. Zhao, L. Marlowe, M. Kaleem, D. Leung, L. Bryden, P. Nath, V. L. Zismann, K. Joshipura, M. J. Huentelman, D. Hu-Lince, K. D. Coon, D. W. Craig, J. V. Pearson, P. Holmans, C. B. Heward, E. M. Reiman, D. Stephan, and J. Hardy. A survey of genetic human cortical gene expression. *Nature genetics*, 2007.

[158] C. B. Nemeroff. The corticotropin-releasing factor (CRF) hypothesis of. *Molecular Psychiatry*, 1996.

[159] D. L. Nicolae, E. Gamazon, W. Zhang, S. Duan, M. E. Dolan, and N. J. Cox. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS genetics*, 6(4):e1000888, Apr. 2010.

[160] L. Oliveira, C. D. Ladouceur, M. L. Phillips, M. Brammer, and J. Mourao-Miranda. What does brain response to neutral faces tell us about major depression? evidence from machine learning and fMRI. *PloS one*, 8(4):e60121, 2013.

[161] A. Özçift. Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis. *Computers in Biology and Medicine*, 41(5):265–271, May 2011.

[162] C. M. Pariante, R. B. Kim, A. Makoff, and R. W. Kerwin. Antidepressant fluoxetine enhances glucocorticoid receptor function in vitro by modulating membrane steroid transporters. *British journal of pharmacology*, 139(6):1111–1118, July 2003.

[163] C. M. Pariante, A. Makoff, S. Lovestone, S. Feroli, A. Heyden, A. H. Miller, and R. W. Kerwin. Antidepressants enhance glucocorticoid receptor function in vitro by modulating the membrane steroid transporters. *British journal of pharmacology*, 134(6):1335–1343, Nov. 2001.

[164] C. M. Pariante and A. H. Miller. Glucocorticoid receptors in major depression: relevance to pathophysiology and treatment. *Biological Psychiatry*, 49(5):391–404, Mar. 2001.

[165] C. M. Pariante, B. D. Pearce, T. L. Pisell, M. J. Owens, and A. H. Miller. Steroid-independent translocation of the glucocorticoid receptor by the antidepressant desipramine. *Molecular pharmacology*, 52(4):571–581, Oct. 1997.

[166] T. Partonen. Clock gene variants in mood and anxiety disorders. *Journal of Neural Transmission*, 119(10):1133–1145, Oct. 2012.

[167] G. Paxinos and K. B. J. Franklin. The mouse brain in stereotaxic coordinates. *Gulf Professional Publishing*, 2004.

[168] T. A. Pearson and T. A. Manolio. How to Interpret a Genome-wide Association Study. *JAMA : the journal of the American Medical Association*, 299(11):1335–1344, Mar. 2008.

[169] M. C. Pepin, M. V. Govindan, and N. Barden. Increased glucocorticoid receptor gene promoter activity after antidepressant treatment. *Molecular pharmacology*, 41(6):1016–1022, June 1992.

[170] M. Peppi, S. G. Kujawa, and W. F. Sewell. A corticosteroid-responsive transcription factor, promyelocytic leukemia zinc finger protein, mediates protection of the cochlea from acoustic trauma. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(2):735–741, Jan. 2011.

[171] L. Pezawas, A. Meyer-Lindenberg, E. M. Drabant, B. A. Verchinski, K. E. Munoz, B. S. Kolachana, M. F. Egan, V. S. Mattay, A. R. Hariri, and D. R. Weinberger. 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nature Neuroscience*, 8(6):828–834, June 2005.

[172] M. W. Pfaffl. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29(9):e45, May 2001.

[173] M. L. Phillips, W. C. Drevets, S. L. Rauch, and R. Lane. Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biological Psychiatry*, 54(5):515–528, Sept. 2003.

[174] C. Pittenger and R. S. Duman. Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology*, 33(1):88–109, Jan. 2008.

[175] A. A. Prather, R. Bogdan, and A. R. Hariri. Impact of sleep quality on amygdala reactivity, negative affect, and perceived stress. *Psychosomatic medicine*, 75(4):350–358, May 2013.

[176] A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, 38(8):904–909, Aug. 2006.

[177] J. K. Pritchard, M. Stephens, N. A. Rosenberg, and P. Donnelly. Association mapping in structured populations. *American journal of human genetics*, 67(1):170–181, July 2000.

[178] K. D. Pruitt, T. Tatusova, G. R. Brown, and D. R. Maglott. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Research*, 40(Database issue):D130–5, Jan. 2012.

[179] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*, 81(3):559–575, Sept. 2007.

[180] G. Rajkowska, J. J. Miguel-Hidalgo, J. Wei, G. Dilley, S. D. Pittman, H. Y. Meltzer, J. C. Overholser, B. L. Roth, and C. A. Stockmeier. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biological Psychiatry*, 45(9):1085–1098, May 1999.

[181] S. E. Reese, K. J. Archer, T. M. Therneau, E. J. Atkinson, C. M. Vachon, M. de Andrade, J.-P. A. Kocher, and J. E. Eckel-Passow. A new statistic for identifying batch effects in high-throughput genomic data that uses guided principal component analysis. *Bioinformatics*, 29(22):2877–2883, 2013.

[182] K. J. Ressler and H. S. Mayberg. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nature Neuroscience*, 10(9):1116–1124, Sept. 2007.

[183] H. G. Roider, A. Kanhere, T. Manke, and M. Vingron. Predicting transcription factor affinities to DNA from a biophysical model. *Bioinformatics*, 23(2):134–141, 2007.

[184] J. P. Romano, A. M. Shaikh, and M. Wolf. Multiple testing. *New Palgrave*, 2010.

[185] D. Salyakina. *Candidate Gene Association Testing in the Dissection of Genetic Causes for Depressive Disorders and the Response to Antidepressant Treatment*. PhD thesis, Technische Universität, München, Germany, 2007.

[186] E. Schadt, C. Molony, E. Chudin, K. Hao, and X. Yang. Mapping the genetic architecture of gene expression in human liver. *PLoS biology*, 2008.

[187] P. Scheet and M. Stephens. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *The American Journal of Human Genetics*, 2006.

[188] J. J. Schildkraut. The catecholamine hypothesis of affective disorders: a review of supporting evidence. *The American journal of psychiatry*, 122(5):509–522, Nov. 1965.

[189] M. V. Schmidt, V. Sterlemann, K. Ganea, C. Liebl, S. Alam, D. Harbich, M. Greetfeld, M. Uhr, F. Holsboer, and M. B. Müller. Persistent neuroendocrine and behavioral effects of a novel, etiologically relevant mouse paradigm for chronic social stress during adolescence. *Psychoneuroendocrinology*, 32(5):417–429, June 2007.

[190] M. V. Schmidt, D. Trümbach, P. Weber, K. Wagner, S. H. Scharf, C. Liebl, N. Datson, C. Namendorf, T. Gerlach, C. Kühne, M. Uhr, J. M. Deussing, W. Wurst, E. B. Binder, F. Holsboer, and M. B. Müller. Individual stress vulnerability is predicted by short-term memory and AMPA receptor subunit ratio in the hippocampus. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(50):16949–16958, Dec. 2010.

[191] D. Schoepf. *Psychiatric Disorders - New Frontiers in Affective Disorders*. InTech, Mar. 2013.

[192] J. Schumacher, R. A. Jamra, T. Becker, S. Ohlraun, N. Klopp, E. B. Binder, T. G. Schulze, M. Deschner, C. Schmäl, S. Höfels, A. Zobel, T. Illig, P. Propping, F. Holsboer, M. Rietschel, M. M. Nöthen, and S. Cichon. Evidence for a relationship between genetic variants at the brain-derived neurotrophic factor (BDNF) locus and major depression. *Biological Psychiatry*, 58(4):307–314, Aug. 2005.

[193] R. H. Segman, N. Shefi, T. Goltser-Dubner, N. Friedman, N. Kaminski, and A. Y. Shalev. Peripheral blood mononuclear cell gene expression profiles identify emergent post-traumatic stress disorder among trauma survivors. *Molecular Psychiatry*, 10(5):500–13– 425, May 2005.

[194] M. E. Seligman, R. A. Rosellini, and M. J. Kozak. Learned helplessness in the rat: time course, immunization, and reversibility. *Journal of comparative and physiological psychology*, 88(2):542–547, Feb. 1975.

[195] A. Serretti, R. Calati, L. Mandelli, and D. De Ronchi. Serotonin Transporter Gene Variants and Behavior: A Comprehensive Review. *Current Drug Targets*, 7(12):1659–1669, Dec. 2006.

[196] D. V. Sheehan, Y. Lecrubier, K. H. Sheehan, P. Amorim, J. Janavs, E. Weiller, T. Hergueta, R. Baker, and G. C. Dunbar. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of clinical psychiatry*, 59 Suppl 20:22–33–quiz 34–57, 1998.

[197] Y. I. Sheline. 3D MRI studies of neuroanatomic changes in unipolar major depression: The role of stress and medical comorbidity. *Biological Psychiatry*, 48(8):791–800, 2000.

[198] M. Skipper, U. Weiss, and N. Gray. Plasticity. *Nature*, 465(7299):703–703, 2010.

[199] M. Slatkin. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, 9(6):477–485, June 2008.

[200] A. So, C. Chaivorapol, E. Bolton, and H. Li. Determinants of cell-and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS genetics*, 2007.

[201] L. Song and G. E. Crawford. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor protocols*, 2010(2):pdb.prot5384, Feb. 2010.

[202] S. Spijker. *Dissection of Rodent Brain Regions*. Humana Press, Totowa, NJ, Apr. 2011.

[203] S. Spijker, J. S. Van Zanten, S. de Jong, B. W. J. H. Penninx, R. van Dyck, F. G. Zitman, J. H. Smit, B. Ylstra, A. B. Smit, and W. J. G. Hoogendijk. Stimulated gene expression profiles as a blood marker of major depressive disorder. *Biological Psychiatry*, 68(2):179–186, July 2010.

[204] E. A. Stahl, S. Raychaudhuri, E. F. Remmers, G. Xie, S. Eyre, B. P. Thomson, Y. Li, F. A. S. Kurreeman, A. Zhernakova, A. Hinks, C. Guiducci, R. Chen, L. Alfredsson, C. I. Amos, K. G. Ardlie, BIRAC Consortium, A. Barton, J. Bowes, E. Brouwer, N. P. Burtt, J. J. Catanese, J. Coblyn, M. J. H. Coenen, K. H. Costenbader, L. A. Criswell, J. B. A. Crusius, J. Cui, P. I. W. de Bakker, P. L. De Jager, B. Ding, P. Emery, E. Flynn, P. Harrison, L. J. Hocking, T. W. J. Huizinga, D. L. Kastner, X. Ke, A. T. Lee, X. Liu, P. Martin, A. W. Morgan, L. Padyukov, M. D. Posthumus, T. R. D. J. Radstake, D. M. Reid, M. Seielstad, M. F. Seldin, N. A. Shadick, S. Steer, P. P. Tak, W. Thomson, A. H. M. van der Helm-van Mil, I. E. van der Horst-Bruinsma, C. E. van der Schoot, P. L. C. M. van Riel, M. E. Weinblatt, A. G. Wilson, G. J. Wolbink, B. P. Wordsworth, YEAR Consortium, C. Wijmenga, E. W. Karlson, R. E. M. Toes, N. de Vries, A. B. Begovich, J. Worthington, K. A. Siminovitch, P. K. Gregersen, L. Klareskog, and R. M. Plenge. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nature genetics*, 42(6):508–514, June 2010.

[205] T. Steckler, N. H. Kalin, and J. Reul. Handbook of Stress and the Brain Part 1: The Neurobiology of Stress: The Neurobiology of Stress. *Elsevier*, 15, Part 1, 2005.

[206] O. Stegle, L. Parts, R. Durbin, and J. Winn. A Bayesian Framework to Account for Complex Non-Genetic Factors in Gene Expression Levels Greatly. *stegle.info*, 2010.

[207] J. C. Stephens, J. A. Schneider, D. A. Tanguay, J. Choi, T. Acharya, S. E. Stanley, R. Jiang, C. J. Messer, A. Chew, J. H. Han, J. Duan, J. L. Carr, M. S. Lee, B. Koshy, A. M. Kumar, G. Zhang, W. R. Newell, A. Windemuth, C. Xu, T. S. Kalbfleisch, S. L. Shaner, K. Arnold, V. Schulz, C. M. Drysdale, K. Nandabalan, R. S. Judson,

G. Ruano, and G. F. Vovis. Haplotype variation and linkage disequilibrium in 313 human genes. *Science (New York, NY)*, 293(5529):489–493, July 2001.

[208] L. Steru, R. Chermat, B. Thierry, and P. Simon. The tail suspension test: a new method for screening antidepressants in mice. *Psychopharmacology*, 85(3):367–370, 1985.

[209] C. A. Stockmeier, G. J. Mahajan, L. C. Konick, J. C. Overholser, G. J. Jurjus, H. Y. Meltzer, H. B. M. Uylings, L. Friedman, and G. Rajkowska. Cellular changes in the postmortem hippocampus in major depression. *Biological Psychiatry*, 56(9):640–650, Nov. 2004.

[210] J. D. Storey and R. Tibshirani. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16):9440–9445, 2003.

[211] B. E. Stranger, A. C. Nica, M. S. Forrest, A. Dimas, C. P. Bird, C. Beazley, C. E. Ingle, M. Dunning, P. Flicek, D. Koller, S. Montgomery, S. Tavaré, P. Deloukas, and E. T. Dermitzakis. Population genomics of human gene expression. *Nature genetics*, 39(10):1217–1224, Sept. 2007.

[212] P. Sullivan, C. Fan, and C. Perou. Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet B (Neuropsychiatric Genetics)*, 141:B:261–268, 2006.

[213] P. F. Sullivan, M. C. Neale, and K. S. Kendler. Genetic Epidemiology of Major Depression: Review and Meta-Analysis. *The American journal of psychiatry*, 157(10):1552–1562, Oct. 2000.

[214] H.-C. Tai and E. M. Schuman. Ubiquitin, the proteasome and protein degradation in neuronal function and dysfunction. *Nature Reviews Neuroscience*, 9(11):826–838, Nov. 2008.

[215] M. Thomas-Chollier, A. Hufton, M. Heinig, S. OKeeffe, N. E. Masri, H. G. Roider, T. Manke, and M. Vingron. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nature protocols*, 6(12):1860–1869, Dec. 2011.

[216] R. M. Twyman and S. B. Primrose. Techniques patents for SNP genotyping. *Pharmacogenomics*, 4(1):67–79, 2003.

[217] Y. M. Ulrich-Lai and J. P. Herman. Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience*, 10(6):397–409, June 2009.

[218] T. B. Ustün, J. L. Ayuso-Mateos, S. Chatterji, C. Mathers, and C. J. L. Murray. Global burden of depressive disorders in the year 2000. *The British journal of psychiatry : the journal of mental science*, 184:386–392, May 2004.

[219] S. J. Utge. *A Study of Candidate Genes in Depression and Disturbed Sleep*. FI, 2012.

[220] M.-J. van Tol, N. J. A. van der Wee, O. A. van den Heuvel, M. M. A. Nielen, L. R. Demenescu, A. Aleman, R. Renken, M. A. van Buchem, F. G. Zitman, and D. J. Veltman. Regional brain volume in depression and anxiety disorders. *Archives of General Psychiatry*, 67(10):1002–1011, Oct. 2010.

[221] D. van West, F. Van Den Eede, J. Del-Favero, D. Souery, K.-F. Norrback, C. Van Duijn, S. Sluijs, R. Adolfsson, J. Mendlewicz, D. Deboutte, C. Van Broeckhoven, and S. Claes. Glucocorticoid receptor gene-based SNP analysis in patients with recurrent major depression. *Neuropsychopharmacology*, 31(3):620–627, Mar. 2006.

[222] H. Vedder, U. Bening-Abu-Shach, S. Lanquillon, and J. C. Krieg. Regulation of glucocorticoid receptor-mRNA in human blood cells by amitriptyline and dexamethasone. *J Psychiatr Res*, 33(4):303–308, July 1999.

[223] H. Vermeer, B. I. Hendriks-Stegeman, B. van der Burg, S. C. van Buul-Offers, and M. Jansen. Glucocorticoid-induced increase in lymphocytic FKBP51 messenger ribonucleic acid expression: a potential marker for glucocorticoid sensitivity, potency, and bioavailability. *The Journal of clinical endocrinology and metabolism*, 88(1):277–284, Jan. 2003.

[224] J.-B. Veyrieras, S. Kudaravalli, S. Y. Kim, E. T. Dermitzakis, Y. Gilad, M. Stephens, and J. K. Pritchard. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS genetics*, 4(10):e1000214, Oct. 2008.

[225] M. Via, C. Gignoux, and E. G. Burchard. The 1000 Genomes Project: new opportunities for research and social challenges. *Genome medicine*, 2(1):3, 2010.

[226] P. M. Visscher. Sizing up human height variation. *Nature genetics*, 40(5):489–490, May 2008.

[227] B. D. Wagner, G. O. Zerbe, S. Mexal, and S. S. Leonard. Permutation-based adjustments for the significance of partial regression coefficients in microarray data analysis. *Genetic epidemiology*, 32(1):1–8, Jan. 2008.

[228] J. D. Wall and J. K. Pritchard. Haplotype blocks and linkage disequilibrium in the human genome. *Nature Reviews Genetics*, 4(8):587–597, Aug. 2003.

[229] C. Wang, X. Zhan, J. Bragg-Gresham, H. M. Kang, D. Stambolian, E. Y. Chew, K. E. Branham, J. Heckenlively, R. Fulton, R. K. Wilson, E. R. Mardis, X. Lin, A. Swaroop, S. Zoellner, G. R. Abecasis, and F. Study. Ancestry estimation and control of population stratification for sequence-based association studies. *Nature genetics*, 46(4):409–+, Apr. 2014.

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

[231] Y.-M. Wang, P. Zhou, L.-Y. Wang, Z.-H. Li, Y.-N. Zhang, and Y.-X. Zhang. Correlation between DNase I hypersensitive site distribution and gene expression in HeLa S3 cells. *PloS one*, 7(8):e42414, 2012.

[232] D. Warden, A. J. Rush, M. H. Trivedi, M. Fava, and S. R. Wisniewski. The STAR*D Project results: a comprehensive review of findings. *Current psychiatry reports*, 9(6):449–459, Dec. 2007.

[233] P. H. WESTFALL and S. S. YOUNG. On Adjusting P-Values for Multiplicity. *Biometrics*, 49(3):941–944, Sept. 1993.

[234] H.-J. Westra and L. Franke. From genome to function by studying eQTLs. *Biochimica et biophysica acta*, May 2014.

[235] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer Publishing Company, Incorporated, Aug. 2009.

[236] N. R. Wray, J. Yang, B. J. Hayes, A. L. Price, M. E. Goddard, and P. M. Visscher. Pitfalls of predicting complex traits from SNPs. *Nature Reviews Genetics*, 14(7):507–515, July 2013.

[237] F. A. Wright, P. F. Sullivan, A. I. Brooks, F. Zou, W. Sun, K. Xia, V. Madar, R. Jansen, W. Chung, Y.-H. Zhou, A. Abdellaoui, S. Batista, C. Butler, G. Chen, T.-H. Chen, D. D'Ambrosio, P. Gallins, M. J. Ha, J.-J. Hottenga, S. Huang, M. Kattenberg, J. Kochar, C. M. Middeldorp, A. Qu, A. Shabalin, J. Tischfield, L. Todd, J.-Y. Tzeng, G. Van Grootheest, J. M. Vink, Q. Wang, W. Wang, W. Wang, G. Willemsen, J. H. Smit, E. J. De Geus, Z. Yin, B. W. J. H. Penninx, and D. I. Boomsma. Heritability and genomics of gene expression in peripheral blood. *Nature genetics*, 46(5):430–437, May 2014.

[238] H. Yaguchi, K. Togawa, M. Moritani, and M. Itakura. Identification of candidate genes in the type 2 diabetes modifier locus using expression QTL. *Genomics*, 85(5):591–599, 2005.

[239] E. Y. Yuen, J. Wei, W. Liu, P. Zhong, X. Li, and Z. Yan. Repeated stress causes cognitive impairment by suppressing glutamate receptor expression and function in prefrontal cortex. *Neuron*, 73(5):962–977, Mar. 2012.

[240] T. Zeller, P. Wild, S. Szymczak, M. Rotival, A. Schillert, R. Castagne, S. Maouche, M. Germain, K. Lackner, H. Rossmann, M. Eleftheriadis, C. R. Sinning, R. B. Schnabel, E. Lubos, D. Mennerich, W. Rust, C. Perret, C. Proust, V. Nicaud, J. Loscalzo, N. H. bner, D. Tregouet, T. M. nzel, A. Ziegler, L. Tiret, S. Blankenberg, and F. o.

Cambien. Genetics and Beyond – The Transcriptome of Human Monocytes and Disease Susceptibility. *PloS one*, 5(5):1–15, May 2010.

[241] B. Zhang and S. Horvath. A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, 4:Article17, 2005.

[242] J. Zhang, H. M. Poh, S. Q. Peh, Y. Y. Sia, G. Li, F. H. Mulawadi, Y. Goh, M. J. Fullwood, W.-K. Sung, X. Ruan, and Y. Ruan. ChIA-PET analysis of transcriptional chromatin interactions. *Methods (San Diego, Calif)*, 58(3):289–299, Nov. 2012.

[243] X. Zhang, S. Huang, W. Sun, and W. Wang. Rapid and robust resampling-based multiple-testing correction with application in a genome-wide expression quantitative trait loci study. *Genetics*, 190(4):1511–1520, Apr. 2012.

[244] H. Zhong, J. Beaulaurier, P. Y. Lum, C. Molony, X. Yang, D. J. Macneil, D. T. Weingarth, B. Zhang, D. Greenawalt, R. Dobrin, K. Hao, S. Woo, C. Fabre-Suver, S. Qian, M. R. Tota, M. P. Keller, C. M. Kendziorski, B. S. Yandell, V. Castro, A. D. Attie, L. M. Kaplan, and E. E. Schadt. Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS genetics*, 6(5):e1000932, May 2010.

[245] Y. Zhou and N. C. Danbolt. Glutamate as a neurotransmitter in the healthy brain. *Journal of Neural Transmission*, Mar. 2014.

[246] P. Zill, T. C. Baghai, P. Zwanzger, C. Schule, D. Eser, R. Rupprecht, H.-J. Möller, B. Bondy, and M. Ackenheil. SNP and haplotype analysis of a novel tryptophan hydroxylase isoform (TPH2) gene provide evidence for association with major depression. *Molecular Psychiatry*, 9(11):1030–1036, Nov. 2004.

[247] K. T. Zondervan and L. R. Cardon. The complex interplay among factors that influence allelic association. *Nature Reviews Genetics*, 5(2):89–100, Feb. 2004.

Acknowledgements

This work was carried out in the Department of Translational Research in Psychiatry at the Max-Planck Institute of Psychiatry in Munich. First of all, I would like to thank Professor Dr. Dr. Dr. Florian Holsboer and Dr. Elisabeth B. Binder for the opportunity to work on this project. I would like to thank my advisor Dr. Elisabeth Binder for her guidance and support of this work. It was an honor to work under her and gain from her huge knowledge and her excellent supervision. Thank you for giving me the opportunity to participate in this great project. I would also like to express my gratitude to my thesis advisor Professor Dr. John Parsch and PD Dr. Mathias V. Schmidt as well as the rest of my thesis committee, who kindly agreed to referee this thesis.

This work would not have been possible without the great number of even greater collaborators. I am much obliged to Ryan Bogdan (Washington University and Duke University), Ahmad R. Hariri (Duke University), Goar Frishman (Helmholtz Zentrum München) and Andreas Ruepp (Helmholtz Zentrum München), Dietrich Trümbach (Helmholtz Zentrum München) and Wolfgang Wurst (Helmholtz Zentrum München). I am grateful to the members of the Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (PGC) for sharing their data.

Thanks to all former and present members of our group for making the department a great place to work. Especially, Dr. Peter Weber, Nazanin Karbalai, Simone Röh, Dr. Divya Mehta, Professor Dr. Bertram Müller-Myhsok, Dr. Darina Czamara, Dr. Andre Altmann for useful discussions; Dr. Peter Weber, Nazanin Karbalai, Simone Röh for proof-reading parts of this thesis; Klaus V. Wagner, Dr. Mathias V. Schmidt for performing the animal experiments; Monika Rex-Haffner, Anne Löschner, Dr. Manfred Uhr, Micheal Ködel, Dr. Peter Weber, Dr. Divya Mehta, Dr. Thomas Bettecken and Stefan Darchinger for excellent technical support and generating microarray data; Dr. Andreas Menke, Dr. Torsten Klengel, Laura Preis and Dr. Andreas Eichelkraut for recruitment and sample ascertainment.

My great appreciation goes to all my friends for their support and encouragement throughout the whole time. Finally and most importantly, I am deeply grateful to Captain Knauer for always believing in me and never giving up.

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt wurde.

Des Weiteren erkläre ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

Die vorliegende Dissertation liegt weder ganz, noch in wesentlichen Teilen einer anderen Prüfungskommission vor.

München, den 05. August 2014