Investigation of genetic and molecular susceptibility factors for bipolar disorder

Cristiana Cruceanu

Department of Human Genetics
Faculty of Medicine
McGill University, Montreal, Canada
March 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Table of contents

Abstract/Resume

Abstract

Bipolar disorder (BD) is a psychiatric condition characterized by at least two episodes of clinically significantly disturbed mood, energy, and activity. Given its debilitating nature, lifetime prevalence and significant occurrence in the general population (1-2.5%), BD is a major public health concern. We know that both environmental and genetic factors contribute to BD susceptibility, with a relatively high contribution of heritable factors (estimates ranging from 60 to 85%). In spite of the strong support for the role of genetics in BD, molecular studies have had little success in replicating specific gene findings, likely as a result of phenotypic and genetic heterogeneity. Alterations in gene expression and regulatory mechanisms in the brain have been shown to play a role in BD. To date, most of the findings point to broad dysregulation across many pathways that are essential for brain function. Questions remain about dysregulation of the plethora of non-coding transcripts whose importance in brain biology has been recently demonstrated, but not characterized for BD. The approaches used in the studies that comprise this thesis were designed to shed light on some of the susceptibility factors for BD, as well as to follow-up on previously implicated pathways and regulatory systems. Thus, high-throughput genome-wide exploratory investigations including whole exome sequencing and transcriptome sequencing, as well as hypothesis-driven candidate gene studies and manipulations of *in vitro* systems were included in this body of work. One major finding is the role of G protein-coupled receptors (GPCRs), both at the genomic level through an enrichment of deleterious mutations carried by affected individuals in the genes encoding these receptors, as well as through a global dysregulation of RNA expression of these receptors in the BD brain. A second major finding is a role for synaptic genes, particularly Synapsin II (*SYN2*), in BD susceptibility and response to treatment with the classical mood stabilizer drug lithium. The results presented in this thesis represent significant contributions toward characterizing the BD susceptibility profile, and shed light on genetic, transcriptional, and epigenetic mechanisms for disease etiology, causality, and course of illness.

Résume

Les troubles bipolaires (TB) sont des troubles psychiatriques caractérisés par au moins deux épisodes où l'humeur, l'activité et l'énergie des patients sont cliniquement perturbés. Dû a sa nature débilitante et sa forte prévalence dans la population générale (1-2,5%), les TB sont des problèmes majeurs de santé publique. Des facteurs génétiques et environnementaux sont impliqués dans l'étiologie des TB, avec notamment une forte contribution de facteurs héréditaires (estimations allant de 60 à 85%). Malgré l'importance du déterminisme génétique dans la susceptibilité aux TB, les résultats des études gènes candidats ont été très peu répliqués, probablement à cause de l'hétérogénéité phénotypique et génétique, et de l'hérédité complexe des troubles. Des modifications de l'expression des gènes et des mécanismes de régulation dans le cerveau jouent également un rôle dans les TB. Aujourd'hui, la plupart des résultats pointent vers des dérèglements majeurs dans de nombreux mécanismes qui sont essentiels pour le fonctionnement du cerveau. Des questions demeurent sur la régulation d'un grand nombre de transcrits non codants (dont l'importance a été récemment démontré dans la biologie du cerveau) et sur la caractérisation de ces mécanismes dans les TB. Les approches utilisées dans mes travaux de thèse ont été conçues à la fois pour mettre en lumière des facteurs de susceptibilité aux TB, et pour caractériser plus finement des voies biologiques et des systèmes de régulations. Pour ce faire, des études haut débit à l'échelle du génome comprenant le séquençage d'exomes entiers et de transcriptomes, couplées à des études gènes candidats (basées sur des hypothèses) et des systèmes d'analyses in vitro ont été utilisé dans mes travaux. L'un des résultats majeurs de ces analyses est le rôle des récepteurs couplés aux protéines G (GPCR), à la fois au niveau génomique par un enrichissement de mutations délétères portées par des individus atteints de TB dans les gènes codant pour ces récepteurs, ainsi que par une dysrégulation globale de l'expression de l'ARN de ces récepteurs dans des cerveaux de patients atteints de TB. Un autre est le rôle des gènes synaptiques, dont notamment Synapsin II (*SYN2*), dans la susceptibilité aux TB et dans la réponse au traitement au lithium utilisé comme médicament régulateur de l'humeur. Les résultats présentés dans mes travaux de thèse contribuent de façon importante à la caractérisation du profil de sensibilité des TB, et apportent de nombreuses clés concernant les mécanismes génétiques, transcriptionnels, et épigénétiques impliqués dans l'étiologie, la causalité, et l'évolution des troubles.

Acknowledgements

This work was supported by funding grants from Génome Québec and the Canadian Institutes of Health Research. The thesis author was supported by generous scholarships from the McGill University Faculty of Medicine, the Network of Applied Medical Genetics (RMGA), the Fonds de Recherché en Santé du Québec (FRSQ), and the Canadian Institutes of Health Research (CIHR).

First and foremost I would like to thank my supervisors, Dr. Gustavo Turecki and Dr. Guy Rouleau, whose support and mentorship has allowed me to mature over the last few years as I learned from their excellent examples how to be a scientist. I hope to one day have made even a fraction of the impact you two have made on the field and the people you have mentored over the years. Thank you for giving me the opportunity to be a part of your incredible teams.

I must also acknowledge the guidance of Drs. Roberta Palmour and Martin Alda, members of my Supervisory Committee. Thank you for sharing your expertise, encouraging my progress, and challenging me to produce the best research I could.

The work I was involved with during my PhD would not have been possible without the support of people who have become like a family. The McGill Group for Suicide Studies consists of so many wonderful and talented individuals who have each taught me very important aspects of scientific research and team work. To Jacques Richard, thank you for being like a father to me, always ready to help me through any hurdle as well as force me to speak French. To Volodymyr Yerko and Jennie Yang, thank you for all of your technical and intellectual guidance and most importantly the positive atmosphere you create in the lab. To Laura Fiori, Erika Freemantle, Benoit Labonte, Corina Nagy, Gaby Torres-Platas, Juan Pablo Lopez, Jeff Gross, Kathryn Vaillancourt, and all the past and present members of the group, thank you for thoughtful discussion, excellent collaboration, and the most fun years of my life.

Special members of the Rouleau Lab bare mention as well. Patrick Dion, Helene Catoire, Pascale Hince, Daniel Rochefort, Dan Spiegelman, thank you for your technical support and guidance, as well as always making me feel welcome.

My time in Montreal would not have been the same, from the first week until the last, without my Montreal sisters. Noha Gerges, Gaby Torres-Platas, Olga Zurita and Bahar Kasaai, you have contributed enormously to my development and my happiness over these years through conversation, adventure and so much laughter! I am forever richer for having you in my life.

Finally and most importantly I must thank my family, whose support has been the most constant thing in my life, making me believe that any goal is achievable with them alongside. To my sister and best friend, Ana, for making me feel like I was never alone in the world. To my mother, Margareta, the most generous person I know, who has given me everything and if she could she would give me more in a heartbeat. To my father, Florentin, who taught me to question everything and ultimately instilled in me the passion for research. He would have been the proudest and most interested reader of this document. Finally, to Juan Pablo Lopez, who has been the best partner, collaborator, and friend I could ever imagine to find. I am a better scientist and human being because of you, and look forward to many years of growing together both professionally and personally.

In memory of my father, Florentin Ion Cruceanu.

List of Abbreviations

µM	microMolar
5-HTTLPR	serotonin neurotransporter
ABCG2	ATP-binding cassette-sub-family G-member2
ACC	Anterior Cingulate Cortex
ACTB	Beta Actin
ADCY2	Adenylate Cyclase 2 (Brain)
ADHD	Attention Deficit Hyperactive Disorder
ANCOVA	Analysis of covariance
ANK3	Ankyrin 3, Node Of Ranvier (Ankyrin G) gene
ANOVA	Analysis of Variance
AP-2α	adaptor-related protein complex 2, alpha 1 subunit
ARN	Acide Ribonucléique
AUC	Area under the curve
B3GALT2	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2
BA10	Brodmann Area 10
BA24	Brodmann Area 24
BA9	Brodmann Area 9
BD	Bipolar disorder
BD II	Bipolar disorder Type II
BDI	Bipolar disorder Type I
BDNF	Brain-derived neurotrophic factor
BRET	Bioluminescence Resonance Energy Transfer
BWA	Burrows-Wheeler Aligner
CACNA1B	calcium channel, voltage-dependent, N type, alpha 1B subunit
CACNA1C	calcium voltage-gated channel subunit alpha1 C
cAMP	Cyclic adenosine monophosphate
CD8	cluster of differentiation 8, a transmembrane glycoprotein
CDHA	Capital District Health Authority
cDNA	Complementary DNA
ChIP	chromatin immunoprecipitation
Chr	Chromosome
CHRM2	Cholinergic receptor, muscarinic 2
CIHR	Canadian Institutes of Health Research
ConLiGen	Consortium on Lithium Genetics
CpG	shorthand for 5'-C-phosphate-G-3'
CREB	cAMP response element-binding
CRF	Corticotropin Releasing Factor
CRH	corticotropin releasing hormone
CRHR2	corticotropin releasing hormone receptor 2
CRHR2-R384X	stop mutation affecting an Arginine amino acid at the C-terminal position 384 of the CRHR2 protein
CRHR2-WT	wild-type CRHR
CTRL	control
DBH	dopamine β-hydroxylase

DE	Differentially expressed
DGKH	Diacylglycerol kinase, beta
DIRAS2	DIRAS family, GTP-binding RAS-like 2
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DRD1	dopamine receptor D1
DRD2	dopamine receptor D2
DRD3	dopamine receptor D3
DRD5	D(1B) dopamine receptor)
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
EC_{50}	the concentration of a drug that gives half-maximal response
EGR1	early growth response 1
ELISA	enzyme-linked immunosorbent assay
EPPK1	epiplakin 1
ER	endoplasmic reticulum
EVS	Exome Variant Server
FBS	fetal bovine serum
FDR	False discovery rate
FGFR1	fibroblast growth factor receptor 1
FRSQ	Fonds de Recherché en Santé du Québec
FYN	FYN oncogene related to SRC, FGR, YES
GABA	gamma-Aminobutyric acid
GADL1	glutamate decarboxylase-like protein 1
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GATK	Genome Analysis Toolkit
GEO	Gene Expression Omnibus
GERP	Genomic Evolutionary Rate Profiling
GFAP	Glial fibrillary acidic protein
GFP	green fluorescent protein
GO	Gene Ontology
GPCR	G protein-coupled receptors (récepteurs couplés aux protéines G)
GR	glucocorticoid receptor
GRCh37	Genome Reference Consortium Human genome build 37
GRIA2	glutamate/alpha-amino-3-hydroxy-5-methyl-4-isoxazolpropionate (AMPA) receptor subunit 2
GRIN2B	NMDA receptor 2B subunit gene
GRM1	metabotropic glutamate receptor 1
GRM1-D508E	missense mutation changing an Aspartate to a Glutamate at amino acid 508 of the GRM1 protein
GRM1-WT	wild-type GRM1
GRM4	metabotropic glutamate receptor 4
GRM7	glutamate receptor 7
GSK3-β	glycogen synthase kinase 3βgene
GWAS	genome-wide association study (-ies)
H3K4	Histone 3 of the 4th Lysine residue (of a nucleosome)
H3K4me3	tri-methylation of the 4th lysine of histone 3

HBAset	bioinformatics too that assembles the Allen Human Brain Atlas expression data for each gene in an input set and computes an average expression level for each region.
HDAC	histone deacetylase
HEK293	human embryonic kidney cell line
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRH1	histamine receptor H1
HRP	horseradish peroxidase
HTR1B	5-hydroxytryptamine receptor 1B
HTR2A	serotonin receptor 2A
HTR3A	ionotropic 5-hydroxytryptamine receptor 3A
HTS	high-throughput sequencing
HTSeq	a Python package that provides infrastructure to process data from high-throughput sequencing assays
IMEM	Iscove's Modified Dulbecco's Medium
IMPA2	inositol monophosphatase gene
IP3	inositol(1,4,5)trisphosphate
LCL	lymphoblastoid cell lines
Li	Lithium
LiCl	Lithium Chloride
lincRNA	long intergenic non-coding RNA
LMO3 (*DAT1*)	LIM domain only 3 (rhombotin-like 2)
lncRNA	long non-coding RNA
LOD	logarithm of odds
LRT	Likelihood Ratio Test
MAP2	microtubule-associated protein 2
MC2R	melanocortin 2 receptor
MDD	Major depressive disorder
Mipp1	Multiple Inositol Polyphosphate Phosphatase
miRNA	microRNA
mM	miliMolar
mRNA	messenger RNA
MRPS25	mitochondrial ribosomal protein subunit 25
NaCl	Sodium Chloride
NCAN	Neurocan gene
NCBI	National Center for Biotechnology Information
ncRNA	non-coding RNA
NGS	normal goat serum
nM	nanomolar
NPC	Neural progenitor cells
NR1D1	nuclear receptor subfamily 1, group D, member 1
NRGN	neurogranin
ODZ4	teneurin transmembrane protein 4 (also known as TENM4)
OFC	olanzapine/fluoxetine combination
OR	Odds Ratio
PBS	Phosphate-buffered saline

PCDH9	protocadherin 9
PCR	polymerase chain reaction
PEI	polyethylenimine
PFC	Pre-frontal cortex
PMI	post-mortem interval
PO	Prolyl oligopeptidase
POLR2A	Polymerase (RNA) II (DNA directed) polypeptide A
poly(A)	polyadenylated
poly-dT	a short sequence of deoxy-thymine nucleotides
PolyPhen-2	Polymorphism Phenotyping v2
PPIF	gene that encodes cyclophilin D (CypD)
qPCR	Quantitative Real-time Polymerase Chain Reaction
qRT-PCR	Quantitative Real-time Polymerase Chain Reaction
rCBF	regional Cerebral Blood Flow
RDC	Research Diagnostic Criteria
RGS	Regulator of G protein Signaling
RIN	RNA integrity number
RMGA	Network of Applied Medical Genetics
RNA	Ribonucleic Acid
RNASeq	RNA sequencing
ROC (curve)	Receiver operating characteristic (curve)
RQ	Relative Quantification
RT	room temperature
RXFP1	Relaxin/insulin-like family peptide receptor 1
RXR	retinoid X receptor (interaction motifs)
SADS-L	Schedule for Affective Disorders and Schizophrenia – Lifetime
SEM	standard error of means
SESTD1	SEC14 and spectrin domains 1
SIFT	Sorting Intolerant from Tolerant
SK-N-AS	human neuroblastoma cell line
SLC6A4	serotonin transporter gene
SLC7A14	Solute carrier family 7, member 14
SNAP29	synaptosomal-associated protein, 29kDa
SNP	single nucleotide polymorphism
snRNA	small nuclear RNAs
SNV	single nucleotide variants
SPSS	software package used for statistical analysis
SSTR2	Somatostatin receptor 2
STEP-BD	Systematic Treatment Enhancement Program for Bipolar Disorder
SYN1	Synapsin I
SYN2	Synapsin II
SYN3	Synapsin III
SYNE1	Spectrin Repeat Containing, Nuclear Envelope 1
TB	Les troubles bipolaires
TH	tyrosine hydroxylase
TMM	weighted trimmed mean of log expression ratios

TRANK1	Tetratricopeptide Repeat And Ankyrin Repeat Containing 1
TrkB	Tropomyosin receptor kinase B
TSS	transcriptional start sites
TTN	titin
U-118 MG	human glioblastoma cell line
UTR	Untranslated Region
UTR3	3' Untranslated region
UTR5	5' Untranslated region
VWC2L	von Willebrand factor C domain containing protein 2-like
WES	whole exome sequencing
WGS	whole genome sequencing
XBP1	X-box binding protein 1

List of Figures

List of Tables

Preface

Contribution of Authors

This thesis is presented in the manuscript based format for a Doctoral Thesis, as described in the Thesis Preparation Guidelines by the Department of Graduate and Postdoctoral Studies. The studies described here were performed under the co-supervision of Drs. Gustavo Turecki and Guy Rouleau.

Chapter 1 contains a review of the literature covering topics such as the psychopathology and history, genetic and non-genetic susceptibility, and the most recent advances in the knowledge of bipolar disorder. The writing and literature review were performed by the thesis author under the supervision of Dr. Turecki. **Part 1.2** represents a Review manuscript published in 2011 in the journal *Current Opinion in Psychiatry*. Writing and editing for this was performed by the thesis author as well as Drs. Martin Alda, Guy A. Rouleau, Gustavo Turecki.

Chapter 2 is made up of two main parts. **Part 2.2** represents a manuscript currently under review in the journal *Nature Communications,* authored by Cristiana Cruceanu, Jean-Francois Schmouth, Susana Gabriela Torres-Platas, Juan Pablo Lopez, Amirthagowri Ambalavanan, Dan Spiegelman, Daniel Rochefort, Pascale Hince, Julie Gauthier, Ronald G. Lafrenière, Patrick A. Dion, Martin Alda, Gustavo Turecki, and Guy A. Rouleau. Experimental design, lab procedures, data analysis and interpretation, as well as manuscript writing were performed by the thesis author. J-FS, SGT-P, JPL, DR, and PC contributed to wet lab procedures. AA and DS performed bioinformatics analyses. JG, RGL, PAD, MA, GT, and GAR contributed to experimental design. All authors contributed to MS writing and editing. GT and GAR co-supervised the project.

Part 2.3 represents a manuscript published in November 2015 in the *American Journal of Psychiatry* by Cristiana Cruceanu, Powell Patrick Cheng Tan, Sanja Rogic, Juan Pablo Lopez, Susana Gabriela Torres-Platas, Carolina O. Gigek, Martin Alda, Guy A. Rouleau, Paul Pavlidis, and Gustavo Turecki. Experimental design, lab procedures, data analysis and interpretation, as well as manuscript writing were performed by the thesis author. PPCT and SR performed bioinformatics analyses under the supervision of PP. JPL, SGT-P and COG contributed to wet lab procedures. MA, GAR, PP and GT contributed to experimental design. All authors contributed to MS writing and editing. GT supervised.

Chapter 3 is made up of three main parts. **Part 3.2** represents a manuscript published on February 2012 in the journal *PLoS ONE* by Cristiana Cruceanu, Martin Alda, Paul Grof, Guy A.

Rouleau and Gustavo Turecki. Experimental design, all lab procedures, data analysis and interpretation, as well as manuscript writing were performed by the thesis author. MA, PG, GAR and GT contributed to experimental design. All authors contributed to MS writing and editing. GT supervised. **Part 3.3** represents a manuscript published in March 2013 in the *International Journal of Neurpsychopharmacology* by Cristiana Cruceanu, Martin Alda, Corina Nagy, Erika Freemantle, Guy A. Rouleau and Gustavo Turecki. Experimental design, lab procedures, data analysis and interpretation, as well as manuscript writing were performed by the thesis author. MA, CN, and EF contributed to data analysis. MA, GAR and GT contributed to experimental design. All authors contributed to MS writing and editing. GT supervised. **Part 3.4** represents a manuscript published in January 2013 in the journal *Neuropsychopharmacology* (Hot Topics section) by Cristiana Cruceanu, Erika Freemantle, Martin Alda, Guy A. Rouleau and Gustavo Turecki. Conceptualization, writing and editing for this was performed by all authors.

Original Contribution to Knowledge

The work presented in this thesis represents several significant contributions toward characterizing the global susceptibility profile in bipolar disorder. The approaches used here included both high-throughput genome-wide exploratory investigations and hypothesis-driven candidate gene studies.

One of the studies described herein (Chapter 2.2), and currently under review in the journal *Nature Communications*, is entitled "Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors". This investigation used whole exome sequencing to search for rare genetic susceptibility factors in 40 families with increased BD heritability as well as a group of singleton BD cases, and identified a significant enrichment of deleterious mutations in G protein-coupled receptor (GPCR) genes as well as significant regulatory roles for some of these receptors. Particularly, a major contribution was the characterization of functional effects of a rare protein-truncating mutation in the corticotropin releasing hormone receptor 2 (*CRHR2*).

The study published in the *American Journal of Psychiatry*, entitled "Transcriptome sequencing of the anterior cingulate in bipolar disorder: dysregulation of G protein-coupled receptors" (Chapter 2.3) was the first to report whole-transcriptome sequencing of both coding and non-coding RNAs in the BD brain. This investigation identified a global trend for down-regulation of gene expression in the BD brain, as well as an overrepresentation of genes involved in G protein-coupled receptor (GPCR) regulation. The latter corroborated the evidence implicating GPCRs in BD suggested by the mutation-discovery exome sequencing study.

Additional contributions were made toward describing the transcriptional and epigenetic dysregulation of the synaptic gene Synapsin II (*SYN2*) in the context of mood disorders and their

treatment. The manuscript published in *PloS ONE*, entitled "Synapsin II is involved in the molecular pathway of lithium treatment in bipolar disorder" (Chapter 3.2), examined the effect of lithium treatment on *SYN2* expression in cell line models, and showed for the first time that treatment with this drug selectively manipulated gene expression in cells from excellent-lithium-responders, a sub-population of patients with increased heritability.

Another study published in the *International Journal of Neuropsychopharmacology*, entitled "H3K4 tri-methylation in synapsin genes leads to different expression patterns in bipolar disorder and major depression" (Chapter 3.3), explored the transcriptional and epigenetic dysregulation of *SYN2* in the post-mortem brains of individuals with mood disorders. This work established a role for methylation enrichment of a histone modification (H3K4me3) in elevating levels of *SYN2* expression in the post-mortem BD brain.

Chapter 1: Introduction

Part 1.1: Preface

Bipolar disorder (BD) is a complex condition characterized by mood alterations and associated neurovegetative disturbances and changes in energy levels. It poses a significant burden on patients, their families, and society, and given its debilitating nature, lifetime prevalence (2.5% in Canada) and significant occurrence in the general population (1-2.5%), BD is a major public health concern.

The literature review presented in Chapter 1.2 focuses on early (prior to 2010) studies investigating genetic factors associated with BD susceptibility. These studies focused primarily on candidate genes, largely due to the available research tools, but in later years expanded to include a number of genome-wide linkage studies with the beginning of the genome-wide association study era. In addition, this chapter pays close attention to the work focused on a specific sub-phenotype of BD: excellent response to lithium treatment.

The literature review presented in Chapter 1.3 discusses more recent progress (after 2010) in BD research, with substantial focus on genome-wide approaches. The publication of the human genome and development of high-throughput sequencing (HTS) technology made it possible to query the totality of genomic mutation, transcriptomic variation, and epigenomic regulation in an individual. As these technologies represent a large portion of the work in this thesis, it is important to start by introducing all the recent advances in this field in BD.

Part 1.2: Response to treatment in bipolar disorder: recent molecular and genetic findings

Title: Response to treatment in bipolar disorder: recent molecular and genetic findings

Authors: Cristiana Cruceanu[1], Martin Alda[2], Guy Rouleau[3], Gustavo Turecki[1#]

[1] McGill Group for Suicide Studies, Douglas Hospital, McGill University, Montreal, Quebec, Canada, H4G 1X6.

[2] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

[3] Department of Medicine, Université de Montréal, Montreal, QC, Canada

[#] Corresponding author

Gustavo Turecki MD PhD

McGill Group for Suicide Studies

6875 LaSalle Blvd.

Verdun, Quebec, Canada

H4H 1R3

Tel: (514) 761-6131 x2369

Fax: (514) 762-3023

gustavo.turecki@mcgill.ca

Structured Abstract

Purpose of review: Bipolar disorder (BD) is a complex psychiatric condition that has been shown to carry a great degree of genetic loading. This review addresses current research in the genetics of treatment response in BD, with a focus on findings that have shaped our understanding of the changing direction of this field in light of recent technological advancements.

Recent findings: The recent publications in BD treatment response have helped consolidate or improve upon knowledge of susceptibility loci and genes in the field. There seems to be an increasing trend toward functionally assessing the role played by putative candidate genes and molecular factors modulating expression in BD, as well as a movement toward more global, pathway- and genome-wide-oriented research.

Summary: Genetic and molecular research to date in BD treatment response has not completely answered all the lingering questions in the field, but has contributed to the development of a more individual patient-based understanding of treatment. In order to apply these findings at a clinical level, more comprehensive treatment response studies are imperative, combining recent advances in high-throughput genomics with functional molecular research.

Keywords: bipolar disorder, lithium, genetics, GWAS

Introduction

Bipolar disorder (BD) is a psychiatric condition characterized by abnormal shifts in energy, activity levels, and mood. Given its debilitating nature, lifetime occurrence, and relatively high prevalence in the general population (1-2%), BD represents a major public health problem and an important topic in health research (1-3). BD has been shown to have a relatively high genetic risk component, with estimates ranging from 60 to 85% (1). However, BD is a complex genetic condition, fact made evident by the limited consensus findings from the linkage and association studies thus far (4, 5). Among the treatment options available, lithium (Li) salts are the most commonly prescribed, and are considered as the first-line mood stabilizer (6). Other highly prescribed medications for BD include, among others, valproate, carbamazepine, and lamotrigine (1, 6). Studies have shown that response to Li treatment runs in families, indicating a significant genetic component. Accordingly, phenotypic and genetic factors identified as predictors of treatment response should be used as guidelines when prescribing BD medication in order to increase the likelihood of treatment success.

The scope of this review is to outline the recent publications in the genetics of BD in reference to response to treatment. It addresses research ranging from candidate gene studies to more comprehensive genome-wide approaches, as well as the functional and expression-related research that goes hand in hand with the identification of susceptibility factors for the disorder.

GENETIC STUDIES

Family history is an important factor associated with treatment response in bipolar disorder as was shown primarily for lithium as well as other treatments. BD patients who respond well to lithium have shown higher genetic liability which led to a variety of studies focusing on the families of these

patients, and most of these studies have confirmed an increased frequency of BD among relatives (2, 3, 7-9). These reports have also revealed very low rates of other psychiatric disorders, like schizophrenia, among relatives of Li-responder patients (7, 10, 11), as well as familial clustering (8). Studies investigating family histories of patients who respond to other drugs such as lamotrigine and divaloprex (10, 12) suggest that these different treatments may be most effective in patients that are clinically and biologically distinct from Li-responders – possibly with distinct genetic profiles.

Previous work in BD genetics focused on linkage and association. Linkage work has identified over 40 chromosomal susceptibility regions (2), and several meta-analyses have been performed in an effort to compare these, but no significant genome-wide support was found for any loci (13, 14). Since sample heterogeneity is a major factor in the disagreement between studies, more homogenous phenotypes have been used in linkage studies. Our group has been focusing on Li response, finding interesting results for chromosomes 18, 15, and 7 (11, 15, 16). To complement linkage findings, association studies have been performed to test individual candidate genes in relation to response to Li (17-21). Some of the most promising findings from other research groups looking at treatment response involve a promoter polymorphism in the serotonin transporter gene (SLC6A4) (22), a promoter variant in the glycogen synthase kinase 3β (GSK3-β) gene (23), and the inositol monophosphatase gene (IMPA2) on chromosome 18 (24).

Candidate Gene Studies

The genetic studies in BD published in the past year have focused in large part on individual BD candidate genes, such as serotonin, dopamine or glutamate-related genes, as well as genes in the

GSK3-β pathway (1, 25). Others have pursued genes identified by the recent genome-wide association studies (GWAS) in BD (26-28). A few of the recent studies have taken a broader approach to look for associations with genome-wide effect.

Campos-de-Sousa *et al*. investigated the Rev-erb-α gene - nuclear receptor subfamily 1, group D, member 1 (*NR1D1*) - for single nucleotide polymorphism (SNP) associations with Li therapy in a sample of 170 BD patients followed for up to 27 years (29). The authors found no significant association with Li response, but the nonresponder group showed a significant increase in T allele frequency at rs2314339 (29).

Manchia *et al*. reported efforts to validate genes associated with BD in regards to Li response by looking for associations between polymorphisms in the dopamine receptors D1, D2 and D3 (*DRD1*, *DRD2*, *DRD3* respectively), LIM domain only 3 (rhombotin-like 2) (*LMO3* a.k.a. *DAT1*), the serotonin neurotransporter (5-*HTTLPR*), and the serotonin receptor 2A (*HTR2A*) and response to Li prophylaxis in a Sardinian sample of 155 unrelated BD probands (30). The same authors also queried the association and interaction effect of the *NR1D1* gene and the Diacylglycerol kinase, beta (*DGKH*) gene with response to Li prophylaxis in a sample of 199 Sardinian Li-responsive BD patients (31). Overall, the results from these two studies showed no significant associations with Li-response, which corroborates previous findings for these genes.

Szczepankiewicz *et al*. investigated the role of glutamatergic neurotransmission in Li response by investigating SNPs in the NMDA receptor 2B subunit gene (*GRIN2B*) in a sample of 105 BD patients treated with Li for at least 5 years and assessed for positive response (32). The gene was a promising candidate based on its chromosomal location (12p12) and evidence of altered protein expression in BD, however no significant associations were found (32). In a different study published

recently, the same authors found a putative association with two SNPs of the glutamatergic FYN oncogene related to SRC, FGR, YES (*FYN*) with BD in a cohort of 425 BD patients and 518 controls (33). This protein kinase is functionally related to NMDA receptors involved in signal transduction mediation in the BDNF/TrkB pathway commonly altered in BD (34). When analyzing specifically Li prophylaxis in a follow-up study of the same SNPs, the group found no association with the rs6916861 SNPs and only a marginal association with rs3730353 (35).

While the studies discussed above focused on response to Li, recent work has also investigated response to other drugs prescribed for BD, and some interesting significant associations have been described. Polymorphisms in the X-box binding protein 1 (*XBP1*), a gene involved in endoplasmic reticulum (ER) stress response, have previously been described as risk factors for BD (1). Kim *et al*. looked at the *XBP1*-116C/G SNP in relation to prophylactic treatment response to valproate in 51 BD patients (36). They showed that patients with the G-allele of *XBP1*-116C/G had better response to prophylactic valproate treatment compared to C-allele carriers, which is in accordance with *in vitro* data showing that the drug ameliorates the ER stress response compromised in G-allele carriers (36). In a randomized, double-blind study of 88 BDI patients treated with an olanzapine/fluoxetine combination (OFC) and 85 patients treated with lamotrigine, Perlis *et al*. genotyped 19 genes (37). SNPs within the dopamine D3 receptor (*DRD3*) and histamine H1 receptor (*HRH1*) genes were significantly associated with response to OFC. SNPs within the dopaminergic receptor *DRD2*, histamine receptor H1 (HRH1), dopamine β-hydroxylase (*DBH*), glucocorticoid receptor (GR), and melanocortin 2 receptor (*MC2R*) genes were significantly associated with response to lamotrigine (37).

Genome-Wide Association Studies

Over the last few years, researchers have moved toward a more comprehensive approach when investigating the role of genetic variation on BD susceptibility and drug response, using GWAS designs. Several GWAS have been published to date focusing on BD in general (26-28) but few loci have reached statistical significance, and overlap between findings has been minimal. One of the most interesting findings was the genome-wide association found by Baum *et al.* (27) between BD and the *DGKH* gene, which encodes a key protein in the Li-sensitive phosphatidyl inositol pathway. An attempt to replicate this result was reported by Squassina *et al.* (38) in a Sardinian sample of 197 BD patients of which 97 were characterized as Li-responders. However, neither the associations found by Baum *et al.* nor the expected association with Li response in BD could be validated (38).

One issue with large-scale GWAS is the high degree of phenotypic and genotypic heterogeneity among the BD patients, resulting from the large sample sizes required for sufficient statistical power. To address this problem, Perlis *et al.* used treatment response to classify patients in a more homogenous, though smaller, subgroup. They performed a GWAS in 1177 BD patients from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) cohort of whom 458 were Li-treated, as well as an additional replication cohort of 359 Li-responsive BD patients (39). Though no SNP passed the significance threshold for genome-wide association, the study pointed to several candidate genes. Of note was the gene for the glutamate/alpha-amino-3-hydroxy-5-methyl-4-isoxazolpropionate (AMPA) receptor subunit 2 (*GRIA2*) (39) which has been shown to be downregulated by chronic Li treatment in a human neuronal cell line (40).

Candidate gene study results showing association of particular genetic variants with treatment response suggest that eventually genetic markers may be used when selecting pharmacologic treatments for BD. However, more research is imperative in order to identify valid markers of

response and comprehend the complexity of the various response pathways. For this reason, the Consortium on Lithium Genetics (ConLiGen) was formed to bring together Li researchers from around the world to establish the largest sample to date for genome-wide studies of Li response. The group's collaborative effort boasts more than 1,200 patients characterized for response under a very stringent phenotype definition. In a publication released earlier this year, the consortium extended an invitation to all Li researchers to join in this effort (41).

FUNCTIONAL STUDIES

It has been well demonstrated through a wealth of genetic epidemiological studies that the susceptibility to develop BD is strongly influenced by genetic factors. However it is clear that BD is a complex disorder and the genetic studies completed to date have provided little insight into the underlying molecular pathology. In order to fully understand the nature of BD, it is essential to elucidate the pathways that are influenced by genetic variants, and the functional effects these have at the cell and organism levels leading to the clinical presentation. With this aim in mind, in a recently published study our group combined linkage with gene expression strategies (16). We initially performed a linkage study in 36 families (275 individuals, of which 132 were affected) ascertained through long-term Li-responsive BD probands. We found genome-wide linkage significance at 3 chromosomal regions (3p25, 3p14 and 14q11), and pursued these findings with a study of the brain expression of all the genes mapping to these regions in a separate cohort of post-mortem BD and control brains. Our findings point to an altered synaptic and mitochondrial functional profile in BD,

with some of the most interesting genes being synapsin II (*SYN2*) and mitochondrial ribosomal protein subunit 25 (*MRPS25*) (16).

Expression studies

In a recent effort to extend genetic susceptibility knowledge into a more global functional analysis was reported by Pedrosa *et al.* (42) who attempted to identify genes of interest in the GSK3-β pathway – a well-established Li target (1). To achieve this goal they used a chromatin immunoprecipitation (ChIP)-chip approach in fetal brains to capture all annotated human promoters bound by β-catenin, a transcription factor that is directly regulated by GSK3-β. They identified 640 genes, which included several genes of interest to BD: calcium channel, voltage-dependent, N type, alpha 1B subunit (*CACNA1B*), neurogranin (*NRGN*), synaptosomal-associated protein, 29kDa (*SNAP29*), fibroblast growth factor receptor 1 (*FGFR1*), and protocadherin 9 (*PCDH9*). Many of the other genes identified correlate with previous findings in schizophrenia and related psychiatric disorders (42). Thus, Pedrosa *et al*. showed that a significant number of BD candidate genes fit into a molecular pathway revolving around GSK3-β signalling.

In another pathway analysis approach, King *et al*. undertook a genetic screen for Li resistance in the social amoeba *Dictyostelium* in hopes of deciphering the molecular basis for Li's effectiveness as a mood stabilizer (43). Prolyl oligopeptidase (PO) - an enzyme altered in BD patients – is a modulator of Li sensitivity and a negative regulator of inositol(1,4,5)trisphosphate (IP3) synthesis, a Li-sensitive intracellular signal. The authors showed that in *Dictyostelium*, as well as in cultured human cells, PO acts via Multiple Inositol Polyphosphate Phosphatase (Mipp1) to modulate Li sensitivity through a gene regulatory network that converges on inositol metabolism (43). Kubota *et*

al. took a similar approach to the King *et al*. study by comparing gene expression data from brains of BD-like transgenic mice - the phenotype includes periodic activity change and altered circadian rhythm - with expression data obtained from post-mortem brains of BD patients to identify relevant biological pathways (44). They identified several genes differentially expressed in the brains of both species, however only one gene was consistently down-regulated in both humans and mice: PPIF. Since this gene encodes cyclophilin D (CypD), a component of the mitochondrial permeability transition pore, the authors continued by showing that a CypD inhibitor was effective in treating the bipolar-like behaviour in their mouse model, thus pointing to a potential treatment avenue involving CypD inhibition (44).

Post-transcriptional regulation

Other levels of regulation - such as microRNA (miRNAs) post-transcriptional interference - have been shown to be relevant in psychiatric disorders and though there have been limited studies thus far, it is important to incorporate these into treatment-responsive pathways. Zhou *et al*. have recently investigated miRNAs and their predicted effectors as targets for the long-term actions of mood stabilizers (45). They screened miRNA levels in Li- or valproate-treated rat hippocampi and showed altered levels for several miRNAs suspected to modulate the expression of brain-specific genes. Additionally, they identified miRNA target sequences amongst BD-risk genes such as dipeptidyl-peptidase 10, metabotropic glutamate receptor 7 (*GRM7*). Changes in expression of this gene were correlated with changes in miR-34a in primary cultures under Li or valproate treatment, confirming that miR-34a contributes to the effects of Li and valproate on *GRM7* (45).

In a similar approach in humans, Chen *et al*. queried the expression patterns of 13 miRNAs in 20 lymphoblastoid cell lines (from 10 BD patients and 10 corresponding discordant unaffected siblings) with or without Li treatment in culture. Seven miRNAs showed significant changes after treatment (46). Interestingly, miR-221 and miR-34a had also been identified by Zhou *et al*. in rat hippocampi, although expression was altered in the opposite direction. Another human study by Rong *et al*. took a candidate approach by focusing on one miRNA of interest: miRNA-134, a potential regulator of dendritic spine volume and synapse formation (47). In a sample of 21 BDI manic patients and matched controls they found that plasma miR-134 levels in drug free, 2-week medicated, and 4-week medicated BD patients were significantly decreased when compared with controls before treatment, and the level was increased following treatment (47). These results suggest that miR-134 may be a peripheral marker of mania and response to mood stabilizers in BD.

Conclusion

Response to treatment in bipolar disorder has a significant genetic component, as primarily shown for Li (1, 48). Factors such as clinical presentation, family history, genetic variants or biomarkers can predict response and should be used to make decisions on course of treatment in order to enhance long-term treatment success. The recent research presented in this review has contributed to the field by providing more information on potential molecular mechanisms involved in BD or on underlying neurobiological process associated with drug response, as well the mechanisms by which they influence gene expression and molecular pathways. These findings can be incorporated into a strategy for improving treatment, but do not completely answer the lingering questions regarding the aetiology of BD and treatment response.

More comprehensive treatment response studies need to be conducted combining high-throughput genomics in the form of treatment-specific GWAS and large scale re-sequencing, as well as assessments of the precise molecular functions of the genetic factors identified. The latter is essential since, as was seen from the wealth of genetic studies thus far, BD is a very complex disorder and the combined action of various relatively rare susceptibility factors likely results in this phenotype.

Acknowledgements

The authors declare no conflict of interest. The authors are generously supported by grants from the Canadian Institute of Health Research (CIHR).

Part 1.3: Genome-wide approaches in bipolar disorder in the era of high-throughput sequencing

History of common variation findings in bipolar disorder

While decades of genetic research in BD identified a multitude of candidate genes and loci, in the last 5-10 years the focus has shifted toward genome-wide approaches, in large part due to technological advances that made this possible. A large body of work in the genome-wide search for BD genetic susceptibility to date has consisted of linkage and genome-wide association studies (GWAS), which suppose that causality of this complex trait is derived from common variants in a small subset of genes. Linkage studies of BD have identified many susceptibility regions (49-59), and unfortunately, finding the right molecular approaches to narrow these chromosomal regions down to specific BD genes has been challenging.

The largest linkage study in BD, which included 972 pedigrees of mixed European ethnicity (60), included 2284 individuals with BD I, 498 individuals with BD II and 702 subjects with recurrent major depression. The strongest findings occurred at 6q21 and 9q21, both with modest non-parametric pairs LOD (logarithm of odds) scores of 3.4, which were not significant after correction for multiple testing. Although this study promised substantially increased statistical power, it was no more effective than previous work in elucidating the missing heritability in BD. Linkage is most effective in the case of limited locus heterogeneity, meaning that genetic variation is restricted to a small number of genomic loci. Given the linkage results in BD so far, this scenario appears less likely than originally anticipated. Alternatively, linkage analyses are effective when a single genetic variant is nearly sufficient to cause disease in all affected members of a large pedigree, as is the case for classical Mendelian transmission. Our knowledge of BD genetics so far render the latter unlikely. Not surprisingly given these caveats, the linkage findings to date have not been very strong and there has

been minimal replication across loci in different populations. This suggests that risk variants are not fully causal and that many regions in the genome are likely contributing to BD susceptibility (61).

The landmark completion of the Human Genome Project revealed that most genetic variation consists of single nucleotide polymorphisms (SNPs), which occur regularly throughout the genome (62, 63). Technological advances led to accurate and cost-effective methods to genotype SNPs throughout the genome in a highly automated assay that is widely known as a genome-wide association study (GWAS). Initial cohorts, with a few thousand cases and controls, were underpowered to detect genome-wide significant associations. As such, while genome-wide studies completed thus far have suggested an ever-growing number of genes and loci, relatively few have reached genome-wide significance levels, and most have not been replicated between studies (64, 65). Overall only a hand-full of genome-wide significant loci have been identified with the following genes most consistently replicated: *TRANK1*, *ANK3*, *ODZ4*, *CACNA1C*, and *NCAN* (66). These are largely the result of meta-analyses aimed at increasing statistical power by combining sample sets from multiple GWAS. The first BD meta-analysis was published by the Psychiatric Genomics Consortium, consisting of 11,974 cases and 51,792 controls (67), which found genome-wide significant evidence at two loci: the calcium voltage-gated channel subunit alpha1 C (*CACNA1C*) gene and the cell surface receptor protein teneurin transmembrane protein 4 (*TENM4* also known as *ODZ4*). Several subsequent meta-analyses have been published (68-71). As sample sizes increase, the hope is that a threshold will be crossed where further sample increases will lead to a regular, linear increase in genome-wide significant findings, as has happened in other complex disorders.

GWAS focus on sub-phenotypes: response to lithium treatment

The potential disadvantage of larger sample sizes is that increasing numbers will lead to increased genetic and phenotypic heterogeneity, which in turn will mask any significant effect of loci with lower penetrance. A strategy to overcome this has been to focus on sub-phenotypes of the disorder with a more homogeneous presentation in terms of symptoms, and presumably a genetic determination based in a smaller pool of loci. One of the most interesting such sub-phenotypes is treatment response, and to date, there have been four GWAS specifically investigating patients who respond favorably to treatment with lithium (and none with other BD drugs). Perlis *et al*. (39) described the first GWAS focusing on response to lithium in a longitudinal cohort of 458 BD subjects actively undergoing treatment. Unfortunately, through interesting suggestive findings emerged, not one locus was found to pass the significance threshold for genome-wide association.

A more recent study from a group in Taiwan reported an unusually strong association (Odds Ratio, OR=73.9) with a SNP in the glutamate decarboxylase-like protein 1 (*GADL1*) gene. The effect size was almost two orders of magnitude greater than was previously reported in BD case-control GWAS (72), and, perhaps not surprisingly, attempts at replication have repeatedly failed (73-77), making the relevance of this finding uncertain and in need for further confirmation. A third GWAS was reported in 2015 consisting of 2698 patients with self-reported lithium response, 1176 patients with clinically documented lithium response, and 8899 healthy controls recruited in Sweden and the United Kingdom (78). The two cohorts were analysed separately as well as meta-analysed, and when comparing lithium-responsive patients with controls, one imputed marker attained genome-wide significant association and was replicated. This was an intronic SNP on chromosome 2q31.2 in the gene SEC14 and spectrin domains 1 (*SESTD1*), which encodes a protein involved in regulation of phospholipids (78).

The largest and most recent study of lithium response included 2,563 patients collected by the Consortium for Lithium Genetics and phenotyped using a uniform retrospective lithium rating scale (79) called the "Alda Scale". Briefly, this scale classifies excellent lithium-responders according to primary diagnosis, episode recurrence risk, and long term stability while ongoing lithium treatment (CITE). This study found a genome-wide significant association at a locus on chromosome 21 that is flanked by two long non-coding RNAs (lncRNAs). Little is known about the function of this locus and further replication and additional work is needed to determine any potential causal relationship between the associated markers and the expression of lncRNAs (79). These initial GWAS show the more typical, modest effect sizes that require large samples sizes to be resolved, and point to the likelihood that lithium response, like BD, may also be a complex polygenic phenotype.

High-throughput sequencing (HTS) technology

GWAS have been successfully applied across a large range of complex traits, where they best characterize genetic variation that is common in the population (typically with minor allele frequency greater than 1–5%). However it is becoming clear that there are a substantial number of genetic variants that influence risk to BD – both common and rare – and most display only small to modest effect sizes, necessitating large-scale genetic studies to robustly identify novel risk factors. The minimal success in finding specific causal BD genes suggests that the past approaches have not been sufficient for psychiatric disorders (80), and also that the genetic architecture of BD is likely far more complex than previously thought (81). Linkage studies were limited in detecting variants since the markers used were too widely distributed throughout the genome. Conversely, GWAS provided tremendously increased accuracy, but since these studies required large sample sizes for statistical

power, they necessitated the use of heterogeneous clinical populations in spite of efforts to focus on sub-phenotypes.

Rare variation may be particularly relevant for the understanding of disease-related biology. By definition, rare variants are evolutionarily more recent and have had less time to be selected against by evolution (82) and thus may be more likely to be pathogenic compared with common variants. Therefore rare variation could provide a more direct and actionable insight into disease pathophysiology (83). Until recently it was very difficult to explore rare genetic variation on a meaningful scale, but fortunately high-throughput next-generation sequencing (HTS) technologies have become available and affordable in the last few years. These can provide a detailed snapshot of all genetic variations in an individual in a matter of days. Applications of HTS are broad, including both DNA and RNA sequencing, and allow the exploration of whole-genome genetic, epigenetic, and transcriptomic variability.

Genomic studies

Two major HTS assays are used to measure rare variants in clinical and research settings: whole genome sequencing (WGS) and whole exome sequencing (WES). The latter is a targeted approach focused on the approximately 1.5% of the genome that is transcribed into messenger RNA and translated into protein. WES has the advantage of lower cost, more manageable bioinformatics demands, and a focus on the more easily interpretable part of the genome where most variants with high penetrance are expected to be found. Although WES is currently the most widely used genetic assay of rare variants, it will ultimately be replaced by WGS as sequencing costs decrease and

bioinformatics capacity is expanded in research institutes around the world, since this technique allows for a far improved exploration of all genetic variation, whether coding or non-coding.

The role of rare variation in psychiatric disorders has been a topic of avid interest over the past couple of years, with ease of interpretation leading to particular emphasis on more heritable syndromes like intellectual disability, autism, schizophrenia, and BD (66). Several BD sequencing studies are ongoing, using both familial and case-control cohorts. The majority of these groups are part of the Bipolar Sequencing Consortium. Only a few family-based sequencing studies have been published so far (84-87), and the results have so far not converged on a specific gene. A number of research groups have been working on exome or whole genome sequencing of large case-control samples, with results expected in early 2016. Together, these case-control studies should have sufficient sample size (several thousand cases and controls) to identify some of the highest penetrance BD susceptibility factors (66). The next several years should reveal whether more penetrant mutations exist in BD, as has been found in other highly heritable psychiatric conditions like autism and intellectual disability (66)

Transcriptomic studies

Aside from the efforts to characterize the genetic susceptibility factors in BD, extensive work has focused on finding the genes and regulatory mechanisms altered in the BD brain. Several studies have shown gene expression dysregulation to play a major role in the aetiology of BD. Early work in this field consisted of candidate gene and genome-wide microarray expression analysis of post-mortem human brain (88), which identified some promissing genes (89, 90). Due to limited sample size and the confounding factors typically associated with post-mortem brain collections (91), very

few findings passed corrections for multiple testing and, consequently, limited replication has been achieved across studies (88, 90, 92). One major limitation of previous genome-wide transcriptome approaches relates to microarray technology, which falters in the level of sensitivity required for detecting low-abundance RNA or subtle transcript-level differences typical of brain tissue. Also, due to the pre-designed nature of this probe-based technology, it misses the majority of non-coding RNAs, unidentified coding transcripts, and differentially spliced genes and isoforms that might have relevance to disease risk and progression.

RNA sequencing (RNASeq), a technique based on HTS technologies, offers a solution in that it provides direct estimates of transcript abundance as well as nucleotide-level sequence. Differential expression can be measured both at the gene and transcript levels, thus providing unbiased and unparalleled evidence for novel RNAs detection (93) and regulatory mechanisms such as alternative splicing (94) which are not well represented on microarrays. Since this technology has become available, there has been strong interest in applying it to post-mortem brain research and psychiatric disorder cohorts. One of the first studies of this nature, and a useful resource for the field, was published by Webb *et al.* who performed transcriptome sequencing in ten post-mortem brain regions from ten psychiatrically healthy individuals (95). Interestingly, when preparing sequencing libraries they used both poly-dT and random hexamer primers in order to detect all RNA classes, including long non-coding (lncRNA), intronic and intergenic transcripts, and transcripts lacking polyadenylated (poly(A)) tails. They were able to detect nearly 40,000 coding and non-coding transcripts and compare and contrast transcriptional profiles between brain regions. Most importantly, they produced the first database of RNA expression in the human brain (95).

However interesting sequencing approaches are for characterizing the healthy brain, there is also great value in using this technology to elucidate the global brain transcriptional dysregulation

profile in psychiatric disorders like BD. The first study (96) to report the use of RNA sequencing in bipolar disorder was published in 2014 by Akula *et al* (96). They examined post-mortem dorsolateral prefrontal cortex (BA46) from 11 individuals diagnosed with bipolar disorder (BD) and from 11 age- and gender-matched controls. They obtained very good sequencing coverage which allowed them to quantify a large majority of known RNA molecules in the brain (~25,000). At a false discovery rate of 0.05%, they showed five genes to be differentially expressed. Some of these, like Prominin 1/CD133 and ATP-binding cassette-sub-family G-member2 (*ABCG2*), were previously-unidentified candidates for this disorder but which had previously been shown to play important roles in neuroplasticity. Though this group performed RNA selection based on presence of poly(A) tails prior to preparing sequencing, which biases against certain types of RNAs that do not receive this modification, they did manage to quantify some lncRNAs with a poly(A) tail, thus adding to our knowledge of the non-coding transcriptome in BD (96).

Two other studies (97, 98) followed in August 2015 showing transcriptome sequencing results in the anterior cingulate cortex (BA24) of BD post-mortem brains. The first (97), performed transcriptome sequencing on RNA from post-mortem brains from 13 individuals diagnosed with bipolar disorder (BD) and 13 matched controls at very high coverage, similar to the Akula *et al.* study. At a false discovery rate of 0.05%, 10 genes were found to be significantly differentially expressed. This was the first study to report unbiased transcriptome sequencing by performing RNA selection by ribosomal depletion, thus quantifying all coding as well as non-coding RNA transcripts regardless of poly(A) tail presence. This study represents Chapter 2.3 of this thesis and will be discussed in greater detail later. The second study (98), that was published around the same time by Zhao *et al.*, performed transcriptome sequencing using the same approach as the Akula et al. study in the post-mortem cingulate cortex from 35 schizophrenia patients, 35 bipolar disorder patients and 35 healthy controls.

This study had the benefit of increased sample size. However, with a 15-fold decreased sequencing coverage they were unable to detect any RNA transcripts that passed multiple testing corrections. Nonetheless, they reported 105 and 153 genes differentially expressed at a nominal p-value in schizophrenia and BD, respectively, and found that many of the genes differentially expressed in both disorders were concordant in their expression levels (98). This finding confirmed the hypothesis that there was a great degree of similarity at the molecular level between BD and other related psychiatric conditions.

Epigenomic studies

Two additional transcriptome studies have been published in recent years, both from the same group, showing RNA sequencing in postmortem brains (BA9 (99, 100) and BA24 (100)) from individuals who had suffered from schizophrenia (n=5), bipolar disorder (n=7), and controls (n=6). The contribution of these studies to our understanding of the transcriptional profile in BD is limited given the low sample size as well as extremely low sequencing coverage. However, it should be noted that the primary goal of this work was to characterize the DNA-methylome in these disorders, which is the global DNA methylation profile. This is interesting as DNA methylation, particularly at regions like promoters and enhancers, has been shown to be closely coupled with gene expression regulation. The researchers used their RNA sequencing data to explore both the overall pathways disrupted by the differential DNA methylation they identified (100) as well as to identify other regulatory mechanisms that may be affected by DNA methylation in BD, including microRNAs (99). Even though epigenetic regulation has been the topic of wide interest in the BD field, by-and-large the post-mortem brain studies that have been published thus far have focused on candidate genes or regions rather than exploratory genome-wide approaches. The two studies mentioned above are among very few exploring epigenetic mechanisms like DNA methylation genome-wide using sequencing technology.

The first study to explore genome-wide DNA methylation in the BD brain, by using CpG-island microarrays, aimed to identify DNA-methylation changes in the frontal cortex and germline associated with schizophrenia and BD (101). Their brain findings showed evidence for psychosis-associated (both BD and schizophrenia) DNA-methylation differences at loci involved in glutamatergic and GABAergic neurotransmission as well as brain development (101). Other more recent genome-wide DNA-methylation studies quantified these modifications in blood samples from BD individuals (102, 103). These identified thousands of differentially methylated regions preferentially located in promoters, 3'-UTRs and 5'-UTRs of genes (102), as well as showed that certain psychotropic drugs frequently used in BD were significantly associated with altered methylation signatures (103). Finally, work has also been done to characterize DNA methylation in transformed lymphoblastoid cell lines (LCL) from individuals with BD. Of note, a study focusing on well-characterized responders to lithium treatment, showed globally decreased DNA methylation in the cells of responders following treatment with lithium in culture, as well as in their relatives, demonstrating the interplay between genetic and epigenetic factors (104).

Part 1.4: Rationale and hypotheses

After several decades of research, our understanding of the role of genetic and environmental factors in conferring susceptibility to psychiatric disorders remains limited. Bipolar disorder (BD) is a typical example of a condition found to be highly heritable ($h^2 > 80\%$), but associated with only a few validated susceptibility loci (61, 105). Furthermore, the genetic predisposition is not fully penetrant (up to 70% based on twin studies (106)), implying that the remaining portion of the susceptibility is related to non-genetic factors. Quantitative genetic analyses point to the role of shared genetic factors and non-shared environment effects, but practically no effects of shared environment (105). These findings justify examination of the genetic predisposition in families, in which the genetic susceptibility is more homogeneous than in unrelated cases. This also justifies the exploration of the role that environment has played in combination with genetic predisposition, through the study of dysregulation in the brains of individuals. A combination of genetic and epigenetic approaches is warranted in order to piece together the multiple biological levels contributing to susceptibility and development of this complex condition.

Bipolar disorder is undoubtedly a very important public health concern given its significant prevalence in the population. It has devastating effects on individuals, who suffer from unusual shifts in mood, energy, activity levels, and reduced ability to carry out day-to-day tasks. Furthermore it can result in damaged relationships, poor job or school performance, and even suicide. Aside from the significant impact it has on individuals, there is also a considerable emotional and economic burden on society. Improving our understanding of the disorder ultimately implies improving our ability to manage and treat it. The work presented in this thesis seeks to contribute to the quest for elucidating the BD susceptibility profile through a combination of different approaches. We propose to test the

general **hypothesis** that susceptibility to bipolar disorder arises from an interaction between the genetic predisposition conferred by relatively rare loci of moderate-to-large effect, with environmental effects mediated via transcriptomic and epigenetic changes.

Part 1.5: Objectives

The **first primary objective** of this work was to investigate the genome-wide patterns of genetic as well as non-genetic susceptibility factors for bipolar disorder. Thus we used high-throughput next-generation sequencing approaches to investigate this on two different levels. **Firstly**, to explore the contribution of inherited DNA mutations we employed whole exome sequencing in families with increased loading of bipolar disorder. The cohort of families for this project has been collected by our group for several decades, and thoroughly characterized for a well-defined clinical subtype of BD – excellent lithium response. Sequencing these families allowed us to focus on a limited amount of genetic heterogeneity, and as a result, to tease out inherited variants with low frequency in the population but moderate to high penetrance in the families. The evidence from rare mutation in each family was meant to help paint a larger picture of the genes and pathways implicated in BD susceptibility and resolve some of the "missing heritability". **Secondly,** to investigate transcriptional dysregulation in the bipolar disorder brain at the level of the whole transcriptome, we used RNA sequencing in post-mortem brains obtained from the Douglas-Bell Canada Brain Bank. These brains were obtained from individuals who died suddenly by suicide or accidental means after having lived with bipolar disorder, and are thoroughly characterized and expertly diagnosed for both primary neuropathology as well as other relevant medical and environmental factors. The goal of this work was to identify genes differentially expressed between cases and controls that could contribute to

our understanding of the pathways and biological processes dysregulated in the brains of individuals who suffer from this devastating mental illness.

The **second primary objective** of this thesis was to follow a candidate-gene approach to resolve specific aspects of bipolar disorder susceptibility. Thus, we followed-up on previous research from our group that postulated a role for the synaptic gene Synapsin II (*SYN2*) in bipolar disorder susceptibility and potentially response to lithium treatment. This gene had been identified through a linkage study of a larger fraction of the lithium-response familial cohort used for our exome sequencing study. **Firstly**, we focused on the role of lithium treatment on modulating *SYN2* expression and explored this through *in vitro* treatment studies in model cell lines. **Secondly**, we investigated the dysregulation of *SYN2* as well as highly homologous sister genes *SYN1* and *SYN3* in the post-mortem brain again using brains from the Douglas-Bell Canada Brain Bank. Furthermore, we sought to elucidate part of the cellular mechanism mediating this dysregulation and explored one of the more common epigenetic modifications associated with gene expression: tri-methylation of the 4^{th} lysine of histone 3 (H3K4me3).

Chapter 2: High-throughput approaches to identify BD susceptibility

Part 2.1 Preface

Bipolar disorder is a complex trait, with heritability estimates from family, twin, and adoption studies ranging from 60-85%. These have suggested that there is a strong genetic component, as well as an unquestionable contribution of environmental factors to disease susceptibility. Unfortunately, as described in greater detail in the Introduction, the search for BD susceptibility factors has been a long and arduous process. It is now clear that multiple genetic susceptibility factors that act through a variety of dysregulated pathways are to blame for symptom development. With the fairly recent development and constant improvement of high-throughput sequencing (HTS) technology, it has now become possible to investigate multiple susceptibility factors for BD concurrently. The modern –omics fields (genomics, transcriptomics, epigenomics, etc.) are fast-paced, exciting, and offer an unprecedented opportunity for health research. Thus, the goal of the work presented in this chapter has been to take advantage of this technology and investigate BD susceptibility factors through high-throughput -omics approaches.

The work presented in Chapter 2.2 sought to address the question of genetic susceptibility in BD. There have been many studies aimed at identifying the causal genes for BD over the past four decades, and though some loci have been found, the larger part of the BD heritability is still 'missing'. The lessons from a great number of linkage and association studies whose focus has been primarily to find common variation has been that rare, private variation likely accounts for a large part of BD susceptibility. The development of high-throughput sequencing technology following the final sequencing mapping of the human genome on April 14, 2003 (62, 63) has made it possible in recent years to query the totality of genomic mutations in an individual, through either whole genome or whole exome sequencing. Because whole genome sequencing (WGS) is still prohibitively expensive (or was at the time we designed this study), we opted to query genetic variation present in the protein-coding portion of the genome through whole exome sequencing (WES). We hypothesized that BD is

caused by highly penetrant rare variants in many different genes across the population, and to avoid dilution of these likely small effects at the population level, we focused on well-characterized multiplex families. Thus, we performed WES in all affected individuals from 40 multi-generational families (3-8 individuals per family across 2-4 generations) and to identify BD susceptibility genes we prioritized rare variants segregating with affected status. The most interesting finding that emerged was an enrichment of putatively causal variants in genes belonging to the G protein-coupled receptor family, which are important drug targets and have previously been connected to psychiatric pathology. Furthermore, we followed up on the functional implications of some of the most deleterious mutations and showed targeted downstream GPCR dysregulation that could explain pathology for a nonsense mutation in the Corticotropin Releasing Hormone Receptor 2 gene (*CRHR2*).

The work presented in Chapter 2.3 sought to address the question of disease susceptibility through gene expression dysregulation in the BD brain. A large body of work has been undertaken by our group as well as others to characterize gene expression changes and alterations in regulatory mechanisms in psychiatric disorders including BD, mainly through candidate gene and a limited number of global microarray expression studies in postmortem brain. However, these previous studies had limitations in regards to the sensitivity of available technology, and questions remain about isoform-specific dysregulation of known genes as well as the plethora of non-coding transcripts whose importance has been demonstrated recently in the brain but not characterized for BD. In line with other high-throughput advances that have become available in recent years, transcriptome sequencing (also referred to as RNAseq) is a powerful technique that captures the complexity of gene expression, and greatly improves upon previous approaches in both accuracy and quantities of information. We performed RNAseq in fresh-frozen post-mortem brain tissue from the anterior cingulate gyrus from 13 BD cases and 13 matched psychiatrically-healthy sudden-death controls. One of the main findings of

this study was a global down-regulation of gene expression in the BD brain compared to controls, as well as an overrepresentation of genes involved in GPCR regulation identified from a gene ontology analysis of the entire set of differentially expressed genes. This finding was very exciting as it mirrored our global finding from the genomic investigation of BD. Though our most interesting gene from the WES study, *CRHR2*, did not show dysregulation that passed multiple testing significance corrections, other GPCRs equally interesting in terms of psychiatry emerged through this analysis, including *SSTR2* (somatostatin receptor 2), *CHRM2* (cholinergic receptor, muscarinic 2) and *RXFP1* (relaxin/insulin-like family peptide receptor 1). Furthermore, we followed-up the top genes by querying the effect of treatment with mood stabilizers commonly prescribed in BD through an *in vitro* study, and found evidence that these drugs affect the expression of several of these genes. Finally, we characterized the non-coding transcriptome in BD and identified the first long intergenic non-coding RNAs associated with BD.

This chapter presents a large body of information regarding genetic and gene expression dysregulation in BD, and globally points to an important role of GPCR genes and pathways. This has important implications in regards to fine-tuning our understanding of the dysregulated BD brain, as well as for identifying potential new drug target genes or pathways. Secondly, this large body of data serves as a resource for the scientific community as it can serve to answer many BD-related questions in the future, far beyond the characterization we have presented here.

Part 2.2: Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors

Title: Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors

Authors: Cristiana Cruceanu[1,2], Jean-Francois Schmouth[2], Susana Gabriela Torres-Platas[1], Juan Pablo Lopez[1], Amirthagowri Ambalavanan[2], Emmanuel Darcq[3], Florence Gross[3], Billy Breton[4], Dan Spiegelman[2], Daniel Rochefort[2], Pascale Hince[2], Julie Gauthier[2], Ronald G. Lafrenière[2], Patrick A. Dion[2], Brigitte L. Kieffer[3], Martin Alda[5], Gustavo Turecki[1#], Guy A. Rouleau[2#]

Affiliations:
1 McGill Group for Suicide Studies, Douglas Mental Health University Institute, Department of Psychiatry, McGill University, Montreal, QC, Canada, H4H 1R3.
2 Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada, H3A 2B4.
3 Department of Psychiatry, Faculty of Medicine, Douglas Hospital Research Center, McGill University, Montreal, QC, Canada, H4H 1R3.
4 Domain Therapeutics NA Inc. Montreal, QC, Canada, H4S 1Z9.
5 Department of Psychiatry, Dalhousie University, Halifax, NS, Canada, B3H 4R2.

Contact Information:
Corresponding Authors:
#Gustavo Turecki
Douglas Mental Health University Institute
6875, boulevard LaSalle, Room F-3125
Montreal, Quebec, H4H 1R3
Email: gustavo.turecki@mcgill.ca
Phone: 514-761-6131 x2369
Fax: 514-762-3023

#Guy Rouleau
Montreal Neurological Institute and Hospital, McGill University
3801 University Street, Office 636
Montreal, Quebec, H3A 2B4
Email: guy.rouleau@mcgill.ca
Phone: 514-398-2690
Fax: 514-398-8248

Abstract

Bipolar disorder (BD) is a prevalent mood disorder that tends to cluster in families. Despite high heritability estimates, few genetic susceptibility factors have been identified over decades of genetic research. One possible interpretation for the shortcomings of previous studies to detect causative genes is that BD is caused by highly penetrant rare variants in many genes. We explored this hypothesis by sequencing the exomes of affected individuals from each of 40 well-characterized multiplex families as well as a singleton replication cohort. We identified rare variants segregating with affected status in many interesting genes, and found an enrichment of deleterious variants in G protein-coupled receptor (GPCR) family genes, which are important drug targets. Furthermore, we showed targeted downstream GPCR dysregulation for some of the variants that may contribute to disease pathology. By focusing on rare variants in informative families we identified key biochemical pathways likely implicated in this complex disorder.

Introduction

Bipolar disorder (BD) is a psychiatric condition characterized by mood alterations that commonly associate with changes in energy, sleep, activity levels, and the ability to carry out day-to-day tasks. Given its debilitating nature, lifetime prevalence and significant occurrence in the general population (1-2%), BD is a major public health concern (107). It has been clearly shown that BD susceptibility is determined by both environmental and genetic factors, but in comparison to other common disorders BD has a relatively high heritability, with estimates ranging from 60 to 85% (1, 108).

Much of the BD genetics research to date has consisted of linkage and genome-wide association studies (GWAS), which hypothesize that rare or common variants, respectively, in a small subset of genes play a causal role in BD etiology. Linkage studies of BD have identified dozens of susceptibility regions across the genome, but these have not been independently replicated (81, 109). Finding the right molecular approaches to narrow these chromosomal regions down to specific BD genes has been challenging, given that these findings would commonly involve large candidate genomic regions that could not be effectively narrowed down by subsequent studies, possibly explained by the fact that linkage is most effective for traits with limited locus heterogeneity(61). GWAS, with the intrinsic requirement for very large sample sizes, have been best suited to identify common susceptibility variants, have had some level of success in recent years (reviewed in (61, 81, 110)). Single nucleotide polymorphisms (SNPs) in a few genes – i.e. *CACNA1C*, *ANK3*, *ODZ4*, *SYNE1*, and *TRANK1* – have been robustly associated with BD risk and replicated across studies(110). However, these findings only explain a small fraction of the BD heritability estimated through twin and adoption studies, suggesting that additional genetic variants – both common and rare – likely influence BD risk (81), and most effect sizes are small to modest, necessitating large-scale genetic

studies to robustly identify novel risk factors (61). Alternatively, these apparently small effects could result from averaging uncommon variants of larger effects across a number of heterogeneous individuals (111).

High-throughput sequencing (HTS) technologies have recently become available and affordable, providing a detailed snapshot of all genetic variations in an individual. Thus, recent studies have started to explore rare variants which could not be investigated before. Previous linkage and association studies unveiled only a small fraction of the estimated BD heritability; thus, the hypothesis that BD is caused by highly penetrant rare variants in a large number of different genes emerged. To pursue this, we sequenced the protein-coding portion of the genome, the exome, of individuals from a collection of 40 multiplex families with high incidence of BD that have been followed longitudinally for as long as 40 years. Family units consist of 3-7 affected individuals across 1-3 generations, with as many as 36 total individuals sampled per family. We focused on rare coding variants that segregated with affected status in families and found an enrichment of putatively damaging mutations in G protein-coupled receptors (GPCRs) among individuals affected with BD. Members from this family of integral membrane proteins have been associated with BD previously, and have been shown to be excellent drug targets. Two major downstream signaling pathways of GPCRs which are mediated by effectors such as cAMP and phosphatidylinositol may be involved in the pathophysiology of BD (112) as well as in the mechanism of some drugs commonly prescribed for this disorder (113, 114). Furthermore, we followed-up functionally on some highly penetrant variants and showed that a nonsense mutation in the GPCR gene corticotropin releasing hormone receptor 2 (*CRHR2*) had a number of downstream effects on cellular function and thus was likely to explain at least part of the disease causation in the affected family members.

Results

Sequencing statistics and quality control measures

High-throughput exome sequencing data from 186 individuals belonging to 40 multiplex families was annotated and analyzed for segregation. On average we sequenced 4.2 affected and 0.6 unaffected individuals per family. An example family is provided in **Figure 1** and all family information is provided in **Supplemental Table 1**. Furthermore, a replication cohort consisting of singleton BD cases for which family members could not be sampled (N=58) and a group of ethnically-matched non-psychiatric controls (N=69) were processed similarly. In terms of data quality, following Burrows-Wheeler alignment (115), on average 99% of reads were successfully aligned, resulting in an average 124X coverage. On average 94% of the exome or an average of 84,000 variants per individual were covered by 20 or more reads.

Variants identified within families

To investigate our hypothesis, i.e., that a burden effect, rather than common inherited variants, leads to an increase in BD susceptibility, we followed a variant filtration approach whereby at least three affected individuals within each family, and no controls, would carry a given rare variant. We also filtered variants by sequencing coverage, frequency in the population, and pathogenicity of mutation as described in **Supplemental Table 2**. Our approach identified more than 3000 individual genes across all 40 families. To test whether any particular pathways or biological processes were enriched through the genes identified, we performed a Gene Ontology analysis (**Supplemental Table 3**). This analysis identified a variety of different enriched processes, in line with the documented genetic complexity and heterogeneity of BD (116, 117).

Following quality-control data processing described in the Methods and following a family-by-family strategy as exemplified in **Supplemental Table 2**, we further prioritized on average 172 variants per family, of which on average 110 were missense single nucleotide variants (SNVs), 2 were stop gains or losses, 2 were frame-shifting insertions or deletions, 5.5 were exon splicing variants, and 38 mapped to either the 3' or the 5' untranslated regions (UTRs). Some variants matched more than one category depending on the isoform affected. To predict potential pathogenicity of missense mutations, we used three of the best-established algorithms, namely SIFT (Sorting Intolerant from Tolerant) (118), PolyPhen-2 (Polymorphism Phenotyping v2) (119) and Mutation Taster (120). These algorithms use different criteria to predict pathogenicity thus pointing to slightly different lists of variants as "probably damaging", the most inclusive and most highly cited being SIFT. We also used conservation prediction tools LRT (Likelihood Ratio Test) (121), PhyloP (122), and GERP (Genomic Evolutionary Rate Profiling) (123) under the assumption that conserved variants would be more likely to be of functional interest. The detailed counts are presented in **Table 1a**. Overall, 9.6 missense variants per family on average satisfied all six conditions and were thus considered to be of priority interest.

Mutations or genes identified across families

We also explored whether any mutations recurred across families (**Supplemental Table 4**), and counted 326 mutations present in two or more families according to the same filtering criteria presented above. Of these, 25 were present in three families and none in more than three families. Some genes, for example *TTN* (titin) and *EPPK1* (epiplakin 1), had an excess of recurring mutations, though this was not surprising given their very large size or presence of highly homologous repeats, respectively. There were 6 stop gain mutations, 5 frameshift insertions or deletions and 198 missense

mutations, as well as a number of mutations mapping to the 3' or 5' UTRs of genes. For the missense variants we also counted those predicted to be deleterious or conserved by the same six commonly used prediction tools, and the detailed counts are presented in **Table 1b**. Overall, 22 missense variants satisfied all six conditions. This analysis is of particular importance here as recurring variants may be technical artifacts resulting from the sequencing technology rather than truly rare variants occurring in these families. Thus, focusing on only those predicted to be pathogenic or conserved serves as a second line of quality filtering.

Mutation burden analysis – a role for G Protein-Coupled Receptors

When focusing our attention specifically on those variants with pathogenic potential, we found many interesting brain-related pathways represented. Among these, some examples are serotonin receptors, with mutations in *HTR3A* (ionotropic 5-hydroxytryptamine receptor 3A) and *HTR1B* (5-hydroxytryptamine receptor 1B), glutamate receptors, with mutations in *GRM1* (metabotropic glutamate receptor 1) and *GRM4* (metabotropic glutamate receptor 4), and dopaminergic genes, with mutations in *DRD5* (D(1B) dopamine receptor) and *TH* (tyrosine hydroxylase). We observed that many of the flagged mutations that were present in affected families mapped to G Protein-Coupled Receptor (GPCR) genes, for example *HTR1B*, *GRM1*, *GRM4*, and *DRD5*. These receptors are important both as regulators of brain functions, as well as potential drug targets. We saw that 38 of the families had at least one prioritized segregating variant that mapped to a GPCR gene, with almost five GPCR gene variants on average per family (**Supplemental Table 5**). We wanted to question whether mutations in GPCR genes occurred more frequently as compared to randomly chosen gene sets (n=100 sets) matched one-to-one by gene size and sequencing coverage. However, we noted that several mutations mapped to the highly polymorphic taste and olfactory receptor genes, which have been

known to lead to exaggerated false positive rates in mutation detection by exome sequencing. For this analysis we opted to exclude all taste and olfactory receptors when creating the matched gene sets. We provide a list of the GPCR genes considered in **Supplemental Table 6**. We found a statistically significant difference between the observed mutation counts in the GPCR genes and the expected mutation counts in the matched gene sets (**Table 2a**; fold change = 1.16; p-value = 0.000034). We considered "deleterious" variants to be nonsense, splice, and missense SNVs, as well as insertions or deletions. We found an increased number of deleterious variants in the GPCR genes as compared to the randomly chosen non-GPCR gene set (**Table 2a**; p-value = 0.000215), while variants that are less likely to be deleterious such as synonymous and 5' or 3'UTR variants did not have the same impact (**Table 2a**; p-value = 0.012565). This suggested that the overall effect was in part driven by the "deleterious" mutation classes. The strongest difference was in nonsense mutations, of which there were 3.2 times more in the GPCR genes than on average in the non-GPCR genes. Since exonic size does not necessarily account for the number of mutations that may be present within a gene, two other approaches were used to generate non-GPCR matched gene sets, both of which took into account predicted gene mutation tolerance. We opted to use both the Constraint Score (124) and the Residual Variation Intolerance Score (RVIS) (125) as they are based on different algorithms and each has its merits. We repeated the mutation burden analysis as above and obtained strikingly similar comparisons between mutation counts, thus validating our finding of an enrichment in deleterious mutations in GPCR genes and an even weaker effect of non-deleterious mutations compared to the size-matched comparisons (**Supplemental Table 7A and B**).

Mutation burden analysis in GPCR genes – replication cohort

In order to investigate the external validity of our findings we used a sample of singleton BD patients. This unrelated cohort was phenotyped similarly to our familial cohort, and consisted of 58 affected individuals who also had family history of BD and 69 ethnically-matched controls. The observed mutation counts in GPCR genes were significantly different in the BD cohort compared to the expected counts from the control group (**Table 2b**; X^2 p-value < 0.0001). Overall we saw an increase in deleterious GPCR mutations in the BD group (**Table 2b**; X^2 p-value < 0.0001), while no significant effect for the non-deleterious mutation classes. This finding in combination with the evidence from the family cohort suggested that the increase in deleterious mutations in GPCR genes was specific to BD individuals and thus these mutations might contribute to disease susceptibility, which led us to further explore the potential pathogenicity of some of the variants identified.

High penetrance variants in GPCR genes

Based on putatively deleterious status, different numbers of variants were prioritized for Sanger sequencing validation in each family. An example analysis is presented in **Supplemental Table 2**. In addition to the criteria described above, we prioritized variants that were shared by all affected members and absent from all unaffected, thus focusing our attention on the highest penetrance variants identifiable in each family. Of the variants we prioritized across all families, 13 (from eight families) were particularly interesting based on segregation pattern and gene function. These belonged to the class of genes that encode G protein-coupled receptors, and included important brain proteins *CRHR2* (corticotropin releasing hormone receptor 2), a member of the hypothalamic-pituitary-adrenal axis, *DRD5* (dopamine receptor D5), a dopaminergic receptor that stimulates adenylyl cyclase, and *GRM1* (Metabotropic glutamate receptor 1), a metabotropic glutamate receptor that functions by

activating phospholipase C, as well as a number of orphan GPCRs. We verified all variants by Sanger sequencing and found 5 that not only validated technically but also maintained segregation evidence when extending the analysis to the available family members not originally sequenced (**Figure 1, Supplemental Figures 1-4**). These five variants, summarized in **Table 3**, include two nonsense and three missense mutations. We decided to focus our follow-up strategy on better understanding the potential effect these variants may have on GPCR function in the cells of patients by expressing the mutations *in vitro*. Given the gene structure and size, the mutations in *CRHR2* and *GRM1* were the most feasible candidates for this strategy; thus we proceeded with these further.

CRHR2 nonsense mutation – effect on downstream GPCR partners

Cloning and confocal microscopy

The heterozygous variant identified in the gene *CRHR2* was found in family number 28 (pedigree in **Figure 1A**) through the filtering approach described and validated using Sanger sequencing in the extended family (**Figure 1B**) to maintain perfect segregation (**Supplemental Table 2A-B**). This nonsense mutation has been shown to modify an Arginine amino acid at the C-terminal position 384 of the CRHR2 protein and cause an early stop by removing the final 28 amino acids (**Figure 1C and 1D**). We used site-directed mutagenesis to replicate this mutation in a cDNA of the wild-type *CRHR2* (CRHR2-WT), from now on referred to as CRHR2-R384X. To ensure that the mutant receptor would be expressed as well as transported to the cellular membrane, we transfected CRHR2-WT and CRHR2-R384X mutant constructs into HEK293T cells and investigated expression with immunofluorescence followed by confocal microscopy. We showed that CRHR2-R384X was indeed expressed in this *in vitro* model and confirmed its presence within the plasma membrane (**Figure 2A-B, Supplemental Figure 5**) by co-localization with the membrane marker Wheat Germ

Agglutinin (WGA). However, we also showed reduced recruitment of the mutant at the plasma membrane and accumulation in the cytoplasm, which was contrary to the CRHR2-WT receptor that was more efficiently recruited to the membrane (**Figure 2A-B, Supplemental Figure 5**). Furthermore, since the mutation in the index family was heterozygous, we co-transfected CRHR2-WT and CRHR2-R384X constructs each bearing a flag or myc tag and vice versa, and showed that both receptors could co-exist in the same cellular membrane (**Supplemental Figure 6A-B**). Further investigations using confocal microscopy demonstrated the presence of both wild-type and mutant protein products at the cell membrane and also confirmed the abundance of the mutant protein in the cytoplasm (**Supplemental Figure 6C** green signal; Z-stack of deconvoluted images and Orthogonal plane view, **Supplemental Figure 6D**), suggesting a disruption in the mutant protein localization at the cell membrane.

Cell surface expression

The impact of the R384X nonsense mutation located at the C-terminus of the *CRHR2* receptor was further evaluated by quantifying plasma membrane localization. Using an ELISA assay revealing the N-terminal flag- tag of both CRHR2-WT and CRHR2-R384X, a significant difference in cell surface receptor expression was observed (**Figure 2C**; $F=425.7$ p-value< 0.0001), which correlated with the microscopy results obtained previously. These results were also used to establish transfection conditions allowing similar wild-type and mutated receptor expression levels in all ensuing biosensor experiments. Similar receptor expression levels were mandatory to allow appropriate biosensor data interpretation.

G-protein intracellular signaling

G proteins are divided into four main families: G_i/G_o, G_q, G_s, and G_{12} (126). The G_s family is known to have a role in activation of adenylate cyclase; thus, the G-protein heterotrimer biosensor for

Gs was first tested as CRHR2 is known to couple to cAMP production (127). For both receptor forms, a BRET (Bioluminescence Resonance Energy Transfer) signal decrease was measured upon stimulation with increasing concentrations of Corticotropin Releasing Factor (CRF) peptide, with similar potency (EC_{50} of 0.33 nM and 0.23 nM for CRHR2-WT and CRHR2-R384X, respectively) but with a better efficacy for the CRHR2-R384X mutant receptor (**Figure 3A; Supplemental Table 8**). There was a significant difference between the response curve EC50 for CRHR2-WT and CRHR2-R384X ($F=6.172$, p-value=0.0173) and an even stronger effect at the top of the curve ($F=173.4$, p-value< 0.0001). Remarkably, while no other G-protein alpha subunits tested seemed engaged by the wild-type, the CRHR2-R384X mutant also showed G_{i2} biosensor activation with EC_{50} of 57.1 nM (**Figure 3B; Supplemental Table 8**), suggesting emergence of a G_i-mediated inhibitory activity for the mutant receptor ($F= 5.135$, p-value=0.0292). No significant differences were found for G_q, G_{12}, or G_{13} (**Figure 3C-E; Supplemental Table 8**). Moreover, all $G_i/G_o/G_z$ family members presented the same activation, only in the presence of the CRHR2-R384X mutant (data not shown). In the absence of an agonist, constitutive activation was only observed for the G_s pathway, and solely for the mutant receptor.

Next, we tested whether the CRF-induced increase of G_i recruitment is present in human patients with the CRHR2-R384X mutation. To do so, we collected membranes from available transformed lymphoblastoid cell lines (LCLs) from individuals in the discovery family (Fam28) and completed a GTPγS assay with the CRF agonist (**Supplemental Figure 7**). Results from representative wild-type unaffected individuals (DNA IDs: 18070 and 18072) as well as affected heterozygous carriers (DNA IDs: 17004 and 19456) of the CRHR2-R834X mutation, show that even with one good copy of the receptor, the mutation significantly increases the GTPγS binding response in human cells. Because G_i is most abundant and has a faster GDP–GTP exchange rate compared to

other G proteins, the GTPγS assay detects primarily G_i-mediated signaling (128, 129). Thus, together with our observation of increased G_{i2} biosensor activation in HEK293 cells (**Figure 3B**), increased CRF-induced GTPγS binding in human cells (F=8.705, p-value=0.0037) strongly suggests that, in cells from CRHR2-R384X carrier patients, signal transduction properties of the receptor shift towards an inhibitory function (**Supplemental Figure 7**).

Since CRHR2 is known to couple to cAMP biosynthesis, we also tested G_s and G_i activation by CRHR2-R384X and CRHR2-WT via cAMP measurements in HEK293T cells co-transfected with our constructs and plasmids coding for cAMP biosensors and exposed to increasing amounts of CRF peptide. As expected, there was no difference in cAMP levels for the CRHR2-WT receptor (panel A) compared to Pertussis toxin (PTX) treatment, which is known to prevent the $G_{i/o}$ subunit G-proteins from interacting with their cognate G protein-coupled receptors (130). At high CRF concentrations however, we observed a small decrease in total cAMP levels in the presence of the CRHR2-R384X mutant (panel B, orange), where the addition of PTX could block the effect (panel B, green) (**Supplemental Figure 8**). Interestingly the EC50 of $G_{i/o}$ biosensor engagement by the CRHR2-R384X mutant corresponds with the development of this biphasic cAMP production phenomenon. At high CRF peptide concentrations, the CRHR2-R384X mutant seems to have the ability to physically engage the $G_{i/o}$ subfamily and as a result modulate cAMP levels (**Supplemental Figure 8**).

β-arrestin

In addition to the engagement of G-proteins, CRHR2 is known to recruit β-arrestin to the plasma membrane (131). The amino acid sequence deleted by the early stop codon in the CRHR2-R384X mutant contains multiple phosphorylation sites involved in β-arrestin interaction. Thus, the effect of the mutation was evaluated on β-arrestin-2 translocation. Results show that in the presence of agonist,

CRHR2-R384X is as potent but less efficient in engaging β-arrestin-2 translocation than the native receptor (EC_{50} of 178nM and 189nM for CRHR2-WT and CRHR2-R384X, respectively) (**Figure 3F; F= 0.3433, p-value= 0.5612**). While there is no significant difference in the EC50 of the WT and mutant response curves, there is a strongly significant difference at the top of the curve (F= 51.56, p-value< 0.0001). Thus, CRHR2-R384X shows less constitutive activity on the translocation phenomenon while CRHR2-WT does induce a constitutive basal translocation of β-arrestin (**Figure 3F**).

GRM1 missense mutation - effect on downstream GPCR partners

The *GRM1* gene encodes the metabotropic glutamate receptor 1 protein, a G protein-coupled receptor suggested to couple predominantly with the G_q/calcium pathway and to not recruit β-arrestin2. We identified a rare putatively deleterious missense mutation (**Table 3**) that segregated perfectly with affected status in a family (**Supplemental Figure 2**). We cloned this and investigated it with the same battery of tests used for the *CRHR2* mutation. The *GRM1* mutation (GRM1-D508E) showed no difference in either cell surface receptor expression levels (**Figure 4A-B, Supplemental Figure 9A-B**), G-protein activation or β-arrestin2 translocation (**Figure 4C-H**). Only the G_q biosensor responded to activation by glutamic acid, which is consistent with the literature (132), but there was no significant difference between the WT and the mutant. Further verification of these results via cAMP measurements in the absence or presence of forskolin, which is known to activate cAMP production without GPCR activation (133), showed no cAMP production by glutamate stimulation (no G_s activation) and no cAMP production inhibition (no G_i activity), respectively (**Supplemental Figure 10A-B**). Finally, since GRM1 primarily couples to G_q second messengers to interact with the IP3-Diacylglycerol(DAG) pathway (134, 135), which in turn activates the protein kinase C (PKC) pathway via intracellular rise in Calcium (Ca^{2+}) concentration, we measured phosphorylation activation based

on conformational change of PKC biosensors after phosphorylation. This has been suggested to be an indirect measurement of calcium oscillations as PLC and diacylglycerol fluctuate together with Ca^{2+} and phosphorylation (136, 137). Upon activation, GRM1 is expected to favor the cytoplasmic Ca^{2+} influx that will then activate endogenously expressed PKC (138) and phosphorylate the biosensor. However, like the results of G-protein activation, efficacy and the potency were similar between GRM1-WT and GRM1-D508E, confirming the initial finding (**Supplemental Figure 10C-D**). All statistical data presented in **Supplemental Table 8**.

Discussion

In this study we sequenced the exomes of multiple BD affected and unaffected individuals from 40 well-characterized families as well a group of singleton BD cases, for a total of 244 high quality whole exomes. Given the genotypic and phenotypic heterogeneity of BD demonstrated by several decades of research, we postulated that deleterious rare variants would be the most likely cause for disease transmission across generations (11, 48). We found several likely functional segregating variants in each family, like nonsense and some high impact missense mutations. Our unbiased search for rare variants identified essentially no homozygous variants, which was in part expected given that the most likely genetic transmission in these families is autosomal dominant (1, 109). There was no convincing overlap of specific variants, and very limited overlap of specific variant-carrying genes. This was not entirely surprising, given the genetic complexity of BD.

Thus, we further explored the rare mutation landscape in BD that could be identified through the exomes of the sequenced families. One of our most interesting findings was an enrichment of putatively damaging mutations in G protein-coupled receptors (GPCRs), with increased numbers of deleterious variants such as missense and nonsense mutations, compared to randomly-selected size-equivalent non-GPCR genes. Members from this family of integral membrane proteins have been previously associated with BD, and have been shown to be excellent drug targets. The most recent example is a study we completed using transcriptome sequencing where we have shown an enrichment of dysregulated GPCR genes in the post-mortem brains of individuals who lived with BD (97). Another example is the GWAS significant finding near the gene *ADCY2* (71), a member of the cAMP dependent GPCR pathway. Two major downstream signaling pathways of GPCRs which are mediated by effectors such as cAMP and phosphatidylinositol may be involved in the pathophysiology of BD (112) as well as in the mechanism of action of drugs commonly prescribed for this disorder (113, 139).

Though a large number of GPCRs encode highly polymorphic sensory receptors, the remaining 376 are of potential interest (140, 141). More than half have a known natural ligand while 150, known as orphan GPCRs, do not (140). Many of the known GPCRs are part of signaling pathways that render them relevant to downstream signaling dysregulation observed in BD (142). Among the signal transmission systems associated with GPCRs, monoaminergic and neuropeptidergic systems are believed to be dysregulated in BD.

One of the most interesting GPCR variants we identified was a premature stop in the C-terminus of the corticotropin releasing hormone receptor 2 (*CRHR2*), which removed the terminal 28 amino acids. This receptor is believed to be involved in stress response through the hypothalamic-pituitary-adrenal axis, a pathway that has been extensively investigated in BD and other mood disorders (143). Following treatment of cells with or without endogenous agonists for the CRHR2 receptor, and looking at proximal GPCR effectors such as G-proteins and β-arrestins, we aimed to evaluate the functional responses following activation of wild-type and mutant receptors *in vitro*. Based on our results, the rare premature stop in *CRHR2* creates a receptor with unique features compared to the wild-type. Firstly, we demonstrated the viability of the mutant receptor produced by this heterozygous variant, though with a decreased capacity to reach the plasma membrane. In the C-terminal tail of CRHR2 there is one 14-3-3 interaction motif (144, 145) removed by the mutation and two RXR interaction motifs (144, 145) of which one is compromised by the mutation. Since both of these motifs are responsible for retention of proteins in the endoplasmic reticulum, this may contribute to the lower expression of the CRHR2 truncated mutant observed at the cell surface. Secondly, we showed increased ligand ability to engage G_i G-protein alpha subunit members, higher constitutive activity for activation of the G_s biosensor, and a lower ligand-activated and constitutive activity for translocation via β-arrestin. At the plasma membrane, β-arrestin's involvement in G-protein signaling

shutdown, via interaction with the phosphorylated C-terminal tail of GPCRs, is known to be important for receptor desensitization (146). The premature R384X stop codon identified in this study removes 6 out of 8 putative phosphorylation sites at the C-terminal tail of CRHR2, likely resulting in a β-arrestin-driven decrease in membrane translocation. Additionally, loss of the 14-3-3 motif could impair complex formation with RGS (Regulator of G protein Signaling) family proteins (147) involved in modulating G-protein signaling. This could thereby favor the increased signal transduction via G_i activation observed in the CRHR2 mutant.

We also explored the downstream effect of a missense mutation in the *GRM1* (Metabotropic Glutamate Receptor 1) gene, which plays important roles in synaptic plasticity-related learning and memory (134) and has previously been associated with BD, schizophrenia, depression etc (148-150). Nonetheless, our tests demonstrated no difference in cell surface receptor expression levels, G-protein activation, PKC-dependent phosphorylation, or β-arrestin2 translocation. While our findings suggest that the *GRM1* mutation identified does not alter receptor function compared to the WT, an alternative interpretation is that this mutation may have a more subtle effect on the kinetics of GPCR activation that is undetectable by our assays, or that it impacts other aspects of GPCR function not tested such as plasma membrane localization or receptor half-life or desensitization. For example, we did not directly test calcium mobilization or calcium oscillations, processes which have previously been shown to be influenced by Group I mGluRs like GRM1 and GRM5 (151-153). Group I mGluRs are positively coupled to phospholipase C (PLC) and stimulate the production of diacylglycerol (DAG) and IP3. Because IP3 promotes the release of sequestered Ca^{2+} from intracellular stores, GRM1 is known for coupling to intracellular Ca^{2+} signaling (154). Our lack of evidence for a functional role of the GRM1 mutation is disappointing, especially in light of literature that links other mutations in this gene with susceptibility to bipolar disorder (149) and schizophrenia (148, 149). However it should be noted that

both studies found that mutations in the ligand-binding domain, where *GRM1-D508E* is located, tended to cluster more in control than cases and showed no significant association with disease. Our findings at this locus thus are consistent with the idea that certain domains in this very important GPCR may be more relevant for psychiatric disease development than others, and in no way weaken the possible link between metabotropic receptors, specifically GRM1, and psychiatric phenotypes. Furthermore, *GRM1-D508E* replaces a glutamic acid residue for aspartic acid, which would not influence the negative charge at that position and thus perhaps also not have a large impact on the general structure and function of the receptor, but rather may affect receptor maturation time or physical interaction with protein partners. Given that in our initial tests the mutation did not display a distinct pharmacological behavior compared to the WT we did not pursue this mutation further, but future research into additional cohorts would be very interesting in better positioning this rare mutation in its phenotypic context. A limitation of the present study is that the implications of the mutations in the *CRHR2* and *GRM1* genes were not explored in BD animal models. This approach would be interesting, as BD, like many psychiatric disorders, does not simply affect specialized cell types but rather the entire nervous system. However, given the challenges of modeling BD in animals (155, 156) this validation work was not attempted here, but future research of these GPCR mechanisms at the whole-organism level would be of interest. Secondly, from the outset our study was focused largely on discovery of rare protein-altering mutations and as a result was limited in the exploration of mutations that do not affect protein structure, such as 3'and 5' UTR mutations. This is in part due to our hypothesis, and in part because at present the algorithms available for predicting pathogenicity of such mutations are limited. Our current knowledge prevents us from accurately distinguishing between "putatively damaging" and non-damaging mutations in these regions. Nonetheless, rapid gains are being made toward our understanding of functional motifs in these

regions, and we encourage future research to revisit this and similar datasets to re-assess the impact of rare UTR mutations for disease susceptibility.

Using unprecedented depth in exploring rare variation, in this study we catalogued all putatively deleterious rare variants that segregated with bipolar disorder in 40 multiplex families, and found an enrichment of deleterious mutations in genes belonging to the GPCR family. Furthermore, we showed that a nonsense mutation in the GPCR gene *CRHR2* had a number of downstream effects on cellular function and thus was likely to explain at least part of the disease causation in the affected family members. There were many GPCR mutations discovered in our analyses that were not followed-through experimentally, but may have interesting functional implications. Through this work we have merely begun to explore the implications of the mutations identified in these families, but additional work by other research groups is warranted to fully elucidate the genetic transmission and dysregulated molecular pathways in BD.

Methods

Description of samples

Ethics approval for the use of human samples in this study was obtained from the Capital District Health Authority (CDHA) in Halifax, Nova Scotia. All subjects gave written informed consent to their participation in the study in regards to sample collection and the generation of lymphoblastoid cell lines (when applicable). No subjects had reduced capacity to consent.

The clinical assessments followed a strict procedure with blind interviews done by pairs of experienced clinicians using the Schedule for Affective Disorders and Schizophrenia – Lifetime version (SADS-L), followed by consensus diagnosis based on DSM-IV criteria. All interviewers underwent extensive training and established very good inter-rater reliability. Probands with a current diagnosis of BD (either type I or type II) were recruited from mood disorders clinics in Canada (Halifax, Ottawa and Hamilton), and all subjects are of Caucasian origin. Relatives were considered affected if they met criteria for BD type I or II, recurrent schizoaffective disorder of bipolar type, or recurrent unipolar depression. Families have been followed longitudinally for up to 40 years, allowing for unprecedented depth in clinical evaluation, including their clinical course and long term outcome. DNA samples collected from multigenerational family units consist of 3–7 affected individuals across 1–3 generations, with as many as 36 total individuals sampled per family. Our strategy in selecting individuals within each family for exome sequencing was to sequence all clearly affected individuals and in some circumstances some clear unaffected familial controls as well. Comorbid psychiatric phenotypes, age of the individual and potential carrier status were carefully considered when choosing family controls. Thus having excellent clinical information for each family member, followed over many years is of utmost importance. Availability of sufficient quantities of genomic DNA to conduct

the study was also important, especially the availability of blood DNA, which has a minimal risk of harboring *de novo* mutations that could be generated during cell line culturing.

Whole Exome Sequencing

Whole exome DNA was captured from total blood DNA using the SureSelect Human All Exon V4 in-solution capture kit (Agilent). Briefly, genomic DNA was sheared, size selected to roughly 150-250 base pairs, and the ends repaired and ligated to specific adapters and multiplexing indexes. Fragments were then incubated with SureSelect biotinylated RNA baits, and the RNA-DNA hybrids were purified using streptavidin-coated magnetic beads, at which point the targeted DNA fragments were briefly amplified by ≤15 PCR cycles. The libraries were then sequenced at the McGill University and Genome Québec Innovation Centre on the HiSeq2000 platform (Illumina) using 100bp pair-ended reads. Raw fastq files were aligned to NCBI human reference GRCh37 using Burrows-Wheeler Aligner (BWA) (115). Single nucleotide variants (SNVs) were called for each exome using primarily the Genome Analysis Toolkit (GATK) (157) and variants were annotated with the Annovar software (158). All computer code used for these analyses is available upon request. In order to prioritize variants with potentially important roles in the genetic susceptibility of BD, according to our hypothesis we firstly focused on rarity (≤1% allelic frequency in the population). Sequence information from three publicly-available repositories of sequencing datasets were used to assess the frequency of each variant in the general population: 1000 Genomes (159), Exome Variant Server (EVS) (160) and Complete Genomics (161). Furthermore, the Rouleau lab has sequenced over 1000 exomes to date - of which >800 are suitable controls for this project. These in-house controls allowed us to correct for any false positives from technical bias specific to the library preparation and sequencing platform.

Variant prioritization

Previous evidence from linkage analyses in this cohort suggested the most likely mode of inheritance, to be dominant (11). According to our hypothesis, we focused our analysis on rare variants. We filtered variants within a family by sequencing quality control metrics (Freeze Set Filter = PASS) and variant frequency in the population based on minor allele frequencies (MAF ≤2%) across all three of the following publicly available exome sequencing datasets: the 1000 Genomes (159), Exome Variant Server (EVS) (160) and Complete Genomics (161). Furthermore we focused on variants segregating with affected status within a family (shared by affected individuals in a family) and not shared with population controls. Other variables considered in the filtering criteria were coverage of variant (set to ≥4 reads), predicted variant function (nonsense, missense, splicing, etc.), as well as predicted damaging effect according to mutation prediction tools such as SIFT (Sorting Intolerant From Tolerant) (118), PolyPhen-2 (119), MutationTaster (120) and conservation predicted by tools such as LRT (Likelihood Ratio Test) (121), PhyloP (122) and GERP (Genomic Evolutionary Rate Profiling) (123). Variants were predicted to be deleterious if they had a SIFT mutation score of ≤0.05, "probably damaging" with a PolyPhen-2 mutation score >0.86, and predicted to be deleterious with a Mutation Taster score >0.9. The follow-up approach was to directly sequence prioritized variants using the classical Sanger method. This is necessary for two reasons: (1) to technically validate variants and confirm they are not false positive calls that emerged from sequencing or analysis irregularities; (2) to test that genuine variants only segregate across the exome-sequenced affected individuals of a family, and not across its unaffected members (from whom DNA was available). An example analysis from one family is provided in **Supplemental Table 2**.

Statistical approach to GPCR mutation burden analysis

To unbiasedly investigate the distribution of GPCR variant types we first obtained a list of all described GPCR genes from the GPCR Natural Variants Database (141), a total of 824 genes. From these we excluded all taste and olfactory receptor genes as these have a high level of heterogeneity that translates into an exaggerated false positive rate in mutation detection by exome sequencing – leaving a total of 376 genes. We provide a list of the GPCR genes considered in **Supplementary Table 6.** In order to have an unbiased comparison gene set, we generated 100 gene lists that were randomly selected except for the requirement that they match one-by-one by exonic size to each GPCR gene. We further generated matched gene sets by two published mutation burden algorithms, the Constraint Score (124) and the RVIS Score (125). These comparisons resulted in the use of 8750 unique genes for the size-matched comparison, 11,452 unique genes for the Constraint-matched comparison and 10,525 unique genes for the RVIS-matched comparison.

We performed variant filtering in all families according to the criteria above and then counted different variant types in each of the 101 gene sets (1 GPCR and 100 non-GPCR). These counts are summarized in **Table 2**. We deemed "deleterious" variants to be in/dels as well as stop, splice, and missense SNVs. We deemed "non-deleterious" variants to synonymous, 3'UTR, and 5'UTR SNVs. To determine whether there was any enrichment of mutations across the seven mutation classes analyzed, we compared the measured mutation counts (GPCR) to the expected mutation counts (non-GPCR) in each class using a Chi Square test. We repeated this analysis for just the 4 deleterious mutation classes or the 3 non-deleterious mutation classes to generate the p-values reported in Table 2a for the size-matched gene sets and **Supplemental Table 7a and 7b** for the gene sets matched by Constraint Score or RVIS Score respectively.

For the singleton replication cohort we filtered variants using the same approach as for the family analysis with the exception of the requirement that variants be shared across individuals. Average mutation counts within the GPCR genes were compared across the same classes described above between the BD (N=58) and the ethnically-matched CTRL (N=69) group after normalization by group size (i.e. BD normalized counts = counts/58*100; CTRL normalized counts = counts/69*100). The statistical analysis for **Table 2b** was completed as described for **Table 2a**.

Cloning

To clone the constructs of interest, human CD8 peptide leader sequence (1-21 amino acids, UniProt ID: P01732-1) followed by FLAG-tag or Myc-tag were used in place of endogenous GPCR signal peptide sequence. DNAs were cloned by Gibson assembly 63 at the BamHI/NotI site of pcDNA3.1 (Invitrogen) using gBlock DNA sequences (IDT DNA Technologies). For CRHR2, isoform alpha (UniProt ID: Q13324-1; GE Dharmacon clone ID: 7389734) was PCR-amplified from amino acid position 20 to 411 with Phusion DNA polymerase (New England Biolabs) and introduced at the NotI site of the pcDNA3.1-CD8-Flag and pcDNA3.1-CD8-Myc. For GRM1, isoform alpha (UniProt ID: Q13255-1, GE Dharmacon clone ID: 40080840) was PCR-amplified from amino acid position 19 to 1194 and cloned as above. Sequences modifications CRHR2-R384X (3222, 3223) and GRM1-D508E (3225, 3226) were introduced by site directed mutagenesis (Stratagene) and all clones were validated by Sanger sequencing.

Immunofluorescence and confocal microscopy

Human embryonic kidney cells (HEK293T) were obtained from the American Type Culture Collection (ATCC). All cells were authenticated by STR profiling and tested for mycoplasma contamination. For immunofluorescence, 12mm diameter cover slips were placed in 24-well plates and treated with POLY-L-LYSINE (final concentration of 0.001%) (Sigma) for 15 minutes at 37°C. Then 50,000 HEK293T cells were seeded and cultured in standard conditions until the next day when transfection of 0.1ug of pcDNA3.1 vectors containing the constructs of interest were transfected with jetPrime (Polyplus transfection) following the manufacturer's indications. Two days later the cells were washed with PBS and fixed with 4% paraformaldehyde for 15 minutes at room temperature (RT). Cells were then blocked in PBS with 10% normal goat serum (NGS) for 1 hour at RT. Co-localization of the CRHR2-WT and CRHR2-R384X constructs was performed using the following antibody combination: 1) primary mouse anti-flag (Sigma, 1:500) overnight at RT, followed by secondary Alexa Fluor 488-conjugated donkey anti-mouse (Sigma) 1:500 for 1 hour at RT; and 2) rabbit anti-myc (Sigma, 1:500) for 1 hour at RT, followed by secondary Alexa Fluor 555-conjugated donkey anti-rabbit (Sigma, 1:500) for 1 hour at RT. ToTo (Invitrogen) 1:300 was used for nuclei staining. Immunofluorescence experiments performed with each individually transfected constructs (i.e.: CRHR2-WT, CRHR2-R384X, GRM1-WT, and GRM1-D508E) were carried out using the following antibody combination: 1) primary mouse anti-myc (Sigma, 1:500), overnight at RT, followed by secondary Alexa Fluor 488-conjugated donkey anti-mouse (Sigma, 1:500), 1hour at RT; 2) primary rabbit anti-CALNEXIN (Abcam, 1:200), 1hour at RT, followed by secondary Alexa Fluor 555-conjugated donkey anti-rabbit (Sigma, 1:500), 1 hour at RT; and 3) WGA 633 (Invitrogen, 1:200), 10 minutes at RT. DAPI (Invitrogen, 1:50,000) was used for nuclei staining. Laser confocal microscopy was carried out with a FLIM LSM 710 confocal microscope. Higher resolution images were obtained at 63x optical magnification combined with a 5x confocal numerical zoom. Image series

were obtained by consecutive confocal scanning using the "scan mode" built in from ZEN (Zeiss). Confocal scanning was performed with 0.380-0.500 µm between each obtained image panels prior to deconvolution. Z-stack files were subsequently deconvoluted with the AutoQuant X3 deconvolution software using default settings (Media Cybernetics).

Bioluminescence Resonance Energy Transfer (BRET) assays by BioSens

Cell Culture

HEK293T cells were maintained in culture in DMEM (Wisent; without Sodium Pyruvate, with 4.5 g/L Glucose, with L-Glutamine) supplemented with Penicillin-Streptomycin (Wisent) and 10% fetal bovine serum (FBS) (Wisent). In each condition, cells were co-transfected with the cloned *CRHR2* or *GRM1* receptors and with one of the G_q, G_s, G_{12}, G_{13}, G_i, Protein G heterotrimer biosensor or with the β-arrestin2 biosensor. HEK293T cells were first transfected with polyethylenimine (PEI, PolyScience). Total DNA amount used for transfection was kept constant at 1 µg/mL of culture, thus salmon sperm DNA (Invitrogen) was used to supplement the coding plasmids (biosensor and receptor) – a 3:1 ratio of PEI:DNA was used. The DNA/PEI solution was incubated for 20 min at RT before adding to the cells pre-seeded in 96-well plates (White Opaque 96-well Microplates, PerkinElmer) pre-treated with poly-L-ornithine (Sigma) at a density of 35 000 cells per well.

Cell-surface ELISA experiments

ELISA experiments were performed 48 hours post-transfection with the cloning constructs. The DMEM medium was removed and cells were washed once with Tyrode-HEPES buffer (Sigma) and fixed with 3.7% paraformaldehyde (Electron Microscopy Sciences) for 15 minutes and then washed 2x with Tyrode-Hepes buffer. The cells were then blocked in Tyrode-HEPES + 1% BSA

(Sigma) for 1h at RT. Primary antibody anti-FLAG-HRP (Sigma) 1/20 000 was added to each well for a 1h incubation at RT, followed by 2 washes Tyrode-Hepes + 1% BSA + agitation for 5 min and another 2 washes Tyrode+25mM HEPES. SensoLyte Luminescence Peroxidase mix (AnaSpec) was added for a 5-10 min incubation and luminescence was measured with the Synergy Neo (BioTek Instruments, Inc., USA), without filter and at 0.4 sec/well.

BRET experiments

BRET (Bioluminescence Resonance Energy Transfer) experiments were designed and performed with the Biosens-All™ platform in collaboration with Domain Therapeutics NA Inc. (Montreal, Canada). BRET signals were recorded 48 hours post-transfection. The DMEM medium was removed and cells were washed once with Tyrode-HEPES buffer (Sigma), and then incubated in Tyrode-HEPES buffer and plated for 30 min at 37°C. Coelenterazine Prolume Purple (Methoxy e-CTZ) (Nanolight) was added to each well for a final concentration of 2.5µM. For increased accuracy in agonist testing, the test compound was added to each well using the HP D300 digital dispenser (Tecan) and 11 concentrations were used for each receptor-biosensor combination. Cells were then incubated with the test compound at RT for 5 min and BRET readings were collected with a 0.4 sec integration time on a Synergy NEO plate reader (filters: 400nm/70nm, 515nm/20nm). BRET signal was determined by calculating the ratio of the light emitted by GFP (515nm) over the light emitted by the luciferase (400nm).

Cyclic AMP (cAMP) assays

HEK293T cells were transfected with WT and mutant constructs according to previously-determined concentrations (20ng of CRHR2-WT, 500ng of CRHR2-R384X, 50ng of GRM1-WT or

50ng of GRM1-D508E) and co-transfected with plasmids coding for cAMP biosensors (modified EPAC biosensors (162)). Increasing amounts of CRF or Glutamate were added overnight to cells with and without 100ng/mL of Pertussis toxin (PTX) and the BRET assays were performed as described above. Experimental data were produced in singleton and curves were fitted using a dose-response with four parameters nonlinear fit.

Protein Kinase C (PKC) assays

HEK293T cells were transfected with WT and mutant constructs according to previously-determined concentrations (50ng of GRM1-WT or 50ng of GRM1-D508E) and co-transfected with plasmids coding for PKC biosensors. Increasing amounts of Glutamate were added overnight to cells and the BRET assays were performed as described. Experimental data were produced in singleton in two independent experiments, and curves were fitted using a dose-response curve with non-linear fit.

[35S]-GTPγS binding assay

Peripheral blood samples were obtained from family individuals following standard procedures, and Epstein-Barr virus-transformed β-lymphoblastoid cell lines (LCLs) were generated as described previously (163, 164). All cells were authenticated by STR profiling and tested for mycoplasma contamination. Cells were cultured and expanded in Iscove's Modified Dulbecco's Medium (IMEM) supplemented with 15% FBS, 1% Fungizone and 1% penicillin/streptomycin/glutamine (Invitrogen) in a 5% CO2 humidified incubator at 37°C, after which cell pellets were collected and frozen at -80°C. [S35]-GTPγS assays were performed on membrane preparations as previously described (165). Membranes were prepared by homogenizing cell pellets in ice-cold 0.25 M sucrose solution and then centrifuged at 2500g for 10 min. Supernatants were collected and diluted 10 times in buffer containing 50 mM TrisHCl (pH 7.4), 3 mM MgCl2, 100 mM NaCl, 0.2 mM EGTA, following which they were centrifuged at 23 000 g for 40 min. The pellets were homogenized in 400µL ice-cold sucrose solution (0.32 M) and kept at - 80°C. For each [35S]GTPγS binding assay, 5µg of protein per well was used (in triplicate). Samples were incubated with and without ligands, for 1 hour at 25°C in assay buffer containing 30 mM GDP and 0.1 nM [35S]GTPγS. Bound radioactivity was quantified using a liquid scintillation counter. Bmax and Kd values were calculated. Non-specific binding was defined as binding in the presence of 10 µM GTPγS, and binding in the absence of agonist was defined as the basal biding.

Acknowledgements

We thank the participants and their families for their contribution to this research. Furthermore, we would like to acknowledge Karine Lachapelle and Annie Levert from the laboratory of GAR for technical assistance.

Figures

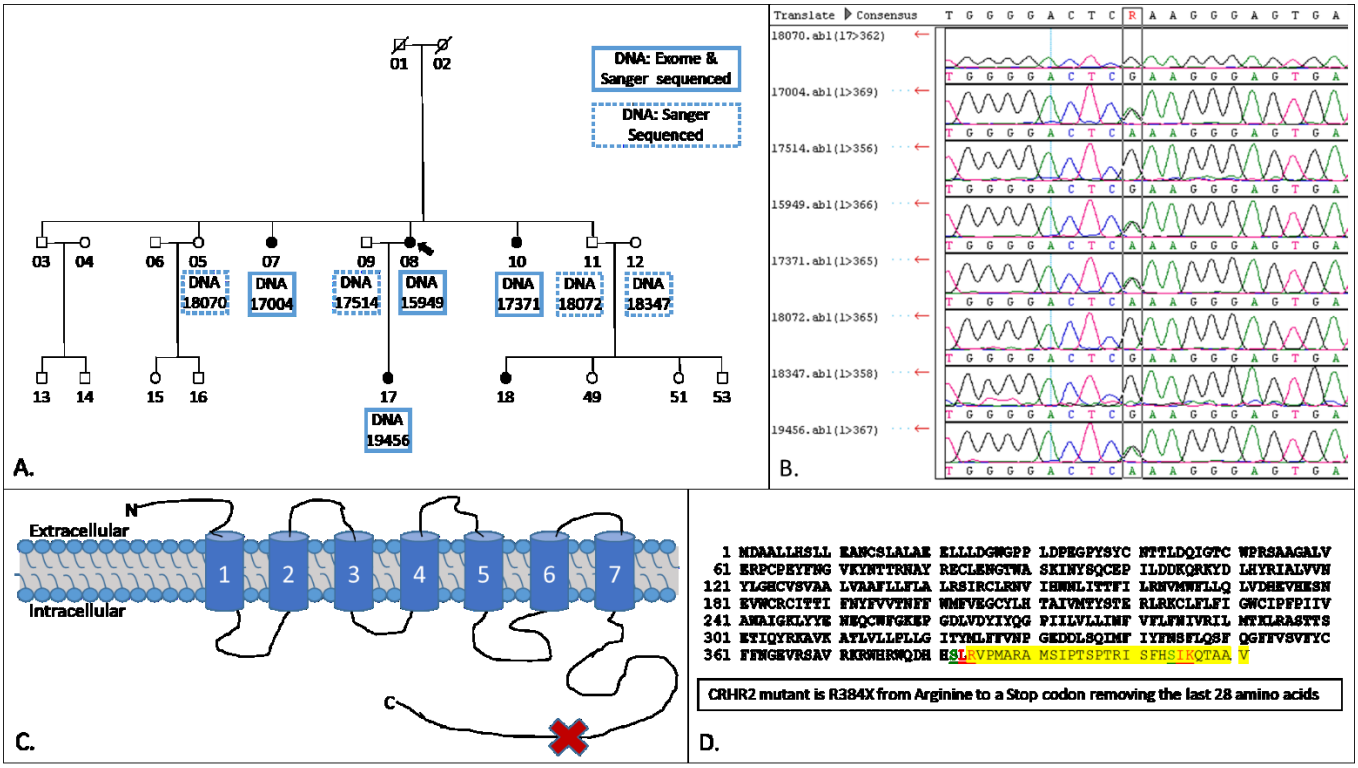


Figure 1. *CRHR2* nonsense mutation A. Pedigree indicating the samples that were exome sequenced or Sanger sequenced. **B.** Sanger sequencing traces for *CRHR2* mutation. **C.** Schematic of CRHR2 protein and *CRHR2-R384X* mutation. **D.** CRHR2 peptide sequence showing truncated portion (highlighted yellow) and lost motifs.

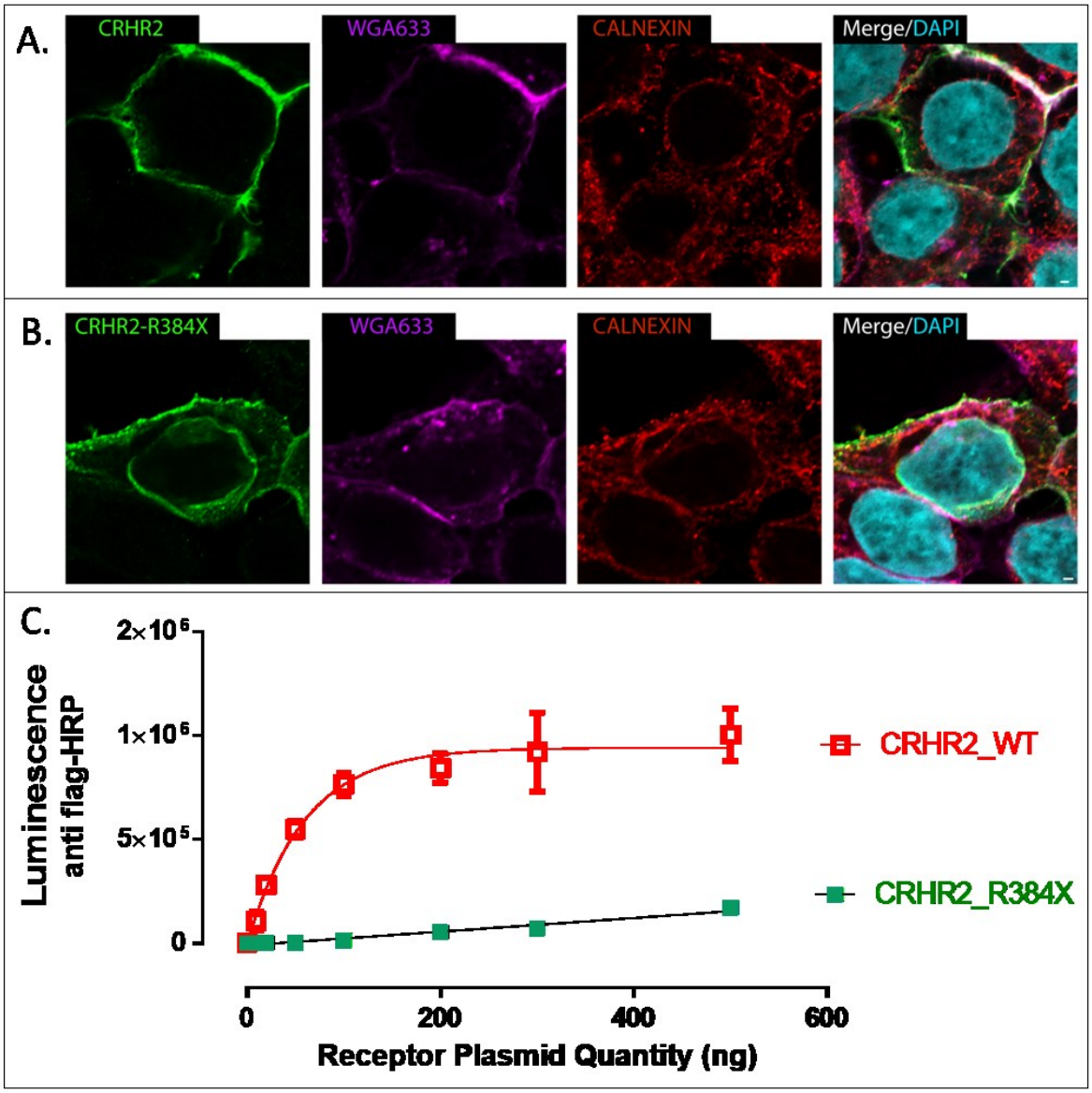


Figure 2. Differential localization of CRHR2-WT and CRHR2-R384X in HEK293T cells. **A.** High resolution, deconvoluted confocal microscopy images demonstrated the predominant presence of CRHR2-WT protein (green) at the membrane in HEK293T cells. Co-localization was observed at the membrane with WGA (purple), whereas little to no co-localization was observed with CALNEXIN (red), an endoplasmic reticulum marker (See also **Supplemental Figure 2A**). Scale bar, 1 µm. **B.** High resolution, deconvoluted confocal microscopy images demonstrated the predominant presence of CRHR2-R384X mutant protein (green) at both the membrane and in the cytoplasm of HEK293T cells. Immunofluorescence results suggest co-localization of the mutant protein with membrane marker WGA (purple) as well as endoplasmic reticulum marker CALNEXIN (red) (See also **Supplemental Figure 2B**). Scale bar, 1 µm. **C.** Transfection of different amounts of plasmid encoding *CRHR2-WT* or *CRHR2-R384X* in HEK293T cells followed by cell surface receptor expression quantification using an ELISA assay, showed a marked difference in membrane recruitment of the *CRHR2-R384X* mutant.

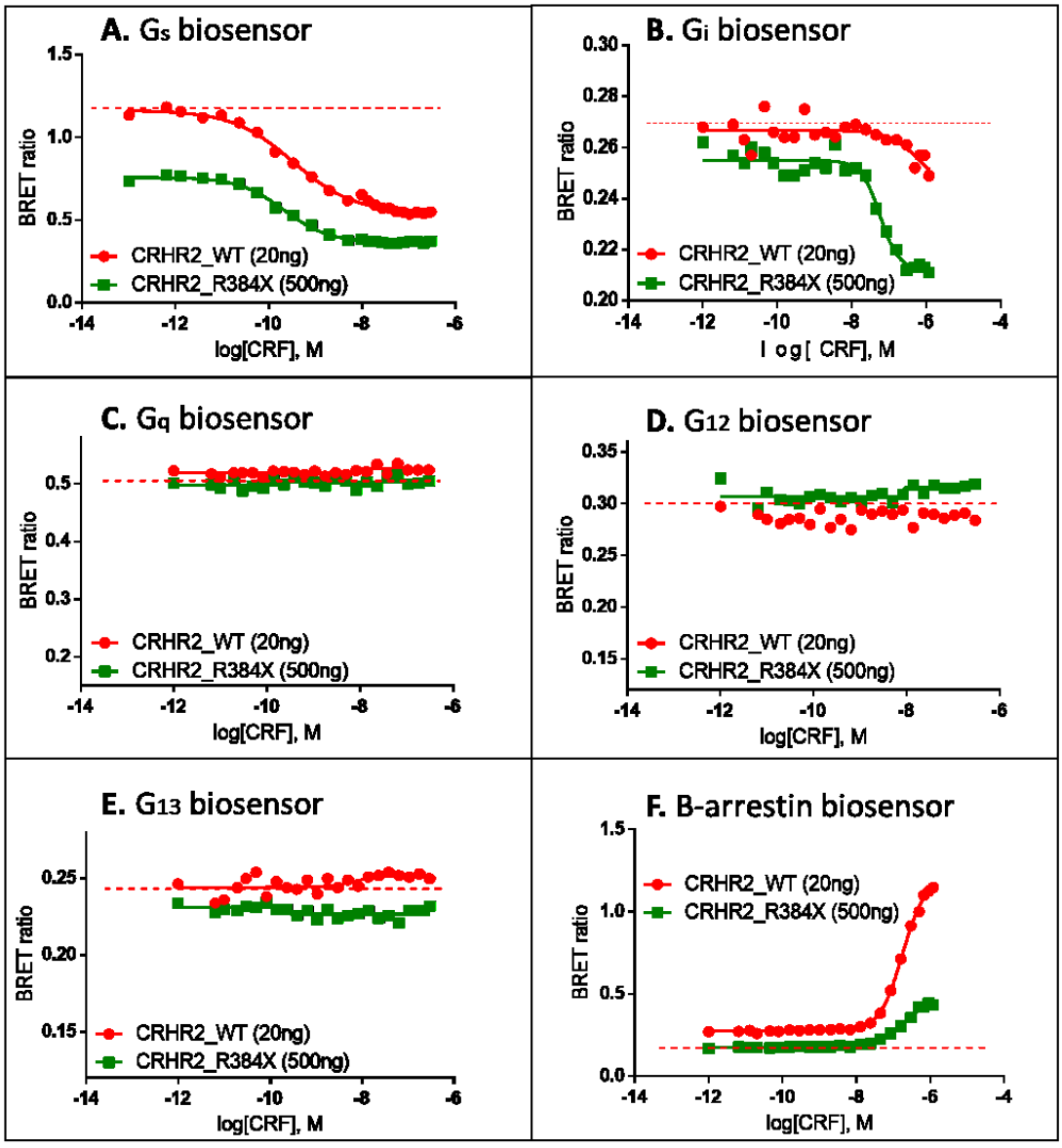


Figure 3. Biosensor modulation by CRHR2-WT and CRHR2-R384X. In HEK293T cells, 20ng of CRHR2-WT or 500ng of CRHR2-R384X were transfected with the plasmids coding for each of 5 different G Protein heterotrimer biosensors (**A-E**) or β-arrestin2 (**F**). Increasing amounts of CRF peptide was added and the BRET assay was performed as described. The red dashed line represents the BRET signal level for cells expressing the biosensor in absence of ligand (constitutive receptor activity is detected when the curve starts below this line). Experimental data were produced in singleton and curves were fitted using a dose-response with four parameters nonlinear fit. Graph is representative of three independent experiments (n=3).

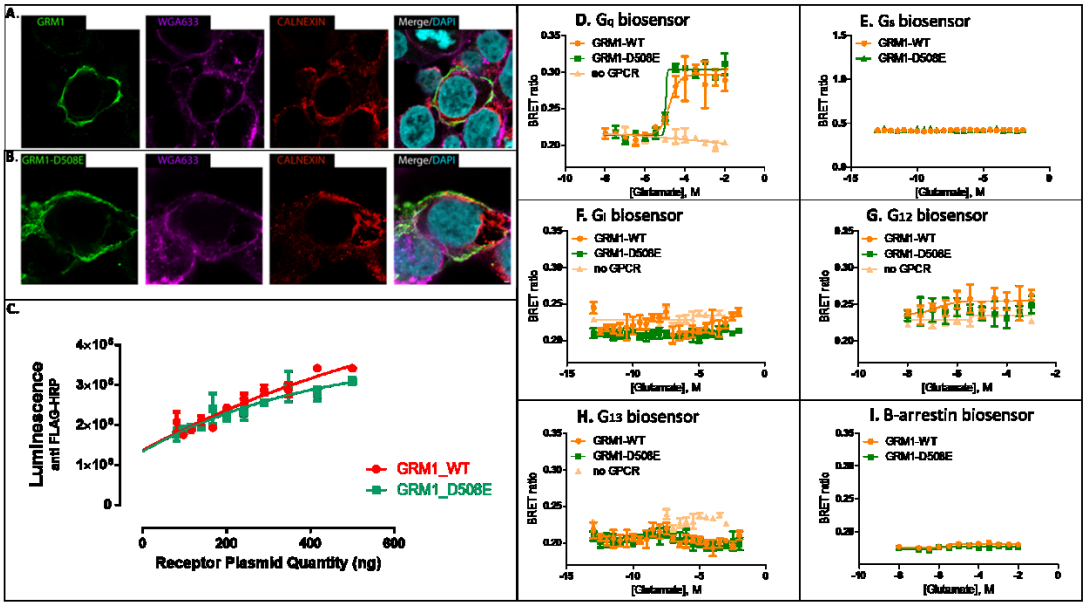


Figure 4. No effect of GRM1-D508E mutation on GPCR membrane localization or biosensor activation. A. High resolution, deconvoluted confocal microscopy images shows the predominant presence of GRM1-WT protein (green) at the membrane in HEK293T cells. Co-localization was observed at the membrane with WGA (purple), whereas reduced co-localization was observed with CALNEXIN (red), an endoplasmic reticulum marker (See also **Supplemental Figure 9A**). Scale bar, 1 µm. **B.** High resolution, deconvoluted confocal microscopy images demonstrated the predominant presence of GRM1-D508E mutant protein (green) at both the membrane and in the cytoplasm of HEK293T cells Co-localization was observed at the membrane with WGA (purple), whereas reduced co-localization was observed with CALNEXIN (red), an endoplasmic reticulum marker (See also **Supplemental Figure 9B**). Scale bar, 1 µm. **C.** Different amounts of plasmid encoding *CRHR2-WT* or *CRHR2-R384X* were transfected in HEK293T cells. Cell surface receptor expression was quantified using an ELISA assay, showing no difference between wild-type and mutant. Experimental data were produced in quadruplicate and curves were fitted using one-phase association nonlinear fit **D-I.** In HEK293T cells, 50ng of *GRM1-WT* or 50ng of *GRM1-D508E* were transfected with the plasmids coding for the G protein biosensors (**D-H**) or β-arrestin2 (**I**). Increasing amounts of glutamate were added and the BRET assay was performed as described. Experimental data were produced in singleton and curves were fitted using a dose-response with four parameters nonlinear fit. Graph is representative of three independent experiments (n=3).

Tables

Table 1. Counts of variants predicted to be pathogenic by commonly used algorithms.		
A. Segregating variants present in 3 or more affected individuals per family (Average per family).		
Algorithm	**Score Criteria**	**Count**
SIFT	≤0.05	55.5
PolyPhen V2	>0.86	30
MutationTaster	>0.9	27
LRT	>0.9995	38
PhyloP	>0.95	58
GERP	Positive	77
All six conditions		9,6
B. Mutations identified across two or more families (segregation in 3 or more affected individuals per family).		
Algorithm	**Score Criteria**	**Count**
SIFT	≤0.05	96
PolyPhen V2	>0.86	61
MutationTaster	>0.9	47
LRT	>0.9995	69
PhyloP	>0.95	98
GERP	Positive	135
All six conditions		22

Table 2A. GPCR vs. non-GPCR variant type distribution in familial BD cohort - genes matched by exonic size.						
		GPCR	**Non-GPCR (Avg of 100)**	**Fold Change (GPCR/Non)**	**Chi Square (p-value)**	
Missense	Deleterious	75	52.68	1.42	0.000034	0.000215
Splicing		0	3.86	0		
Nonsense		3	0.8	3.26		
In/dels		3	2.42	1.24		
UTR3	Non-deleterious	4	10.33	0.39		0.012565
UTR5		1	6.75	0.15		
Synonymous		44	35.43	1.24		
Total		**130**	**112**	**1.16**		

Table 2B. GPCR variant type distribution in the singleton BD cohort compared to matched controls (normalized counts by samples size).

		BD	CTRL	Fold Change (BD/CTRL)	Chi Square (p-value)	
Missense	Deleterious	300	273.91	1.1	0.000197	0.000078
Splicing		18.97	7.25	2.62		
Nonsense		3.45	4.35	0.79		
In/dels		6.9	5.8	1.19		
UTR3	Non-deleterious	41.38	39.13	1.06		0.067464
UTR5		22.41	14.49	1.55		
Synonymous		205.17	228.99	0.9		
Total		**598.3**	**573.93**	**1.04**		

Table 3. Short-listed GPCR variants.

Family	Position	Reference Allele	Mutant Allele	Gene	Gene class	Variant Function	Exome Variant Server frequency
Fam28	Chr7:30693162	G	A	CRHR2	GPCR	Stop Gain	0.001461
Fam29	Chr1:168074103	G	A	GPR161	GPCR	Stop Gain	N/A
Fam38	Chr6:146678752	T	G	GRM1	GPCR	Missense	N/A
Fam33	Chr3:48678823	GGTT	G	CELSR3	GPCR	Non-frameshift	0.001838
Fam19	Chr8:37688966	G	A	GPR124	GPCR	Missense	0.008688

Part 2.3: Transcriptome sequencing of the anterior cingulate in bipolar disorder: dysregulation of G protein-coupled receptors

Title: Transcriptome sequencing of the anterior cingulate in bipolar disorder: dysregulation of G protein-coupled receptors

Authors: Cristiana Cruceanu[1,2,3,*], Powell Patrick Cheng Tan[4*], Sanja Rogic[4], Juan Pablo Lopez[1,2], Susana Gabriela Torres-Platas[2], Carolina O. Gigek[2], Martin Alda[5], Guy A. Rouleau[1,3], Paul Pavlidis[4], Gustavo Turecki[1,2,#]

Authors' affiliations:

1 Department of Human Genetics, McGill University, Montreal, QC, Canada

2 McGill Group for Suicide Studies & Douglas Research Institute, Montreal, QC, Canada

3 Montreal Neurological Institute, McGill University, Montreal, QC, Canada

4 Centre for High-Throughput Biology and Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

5 Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

*These authors contributed equally

#Corresponding Author
Douglas Mental Health Institute, McGill University
6875 LaSalle Blvd., Montreal, QC, H4H 1R3
Phone: (514) 761-6131 ext. 3366
Fax: (514) 762-3023

Word Count:
Abstract: 248
Introduction: 394
Body: 3453

Keywords: Bipolar disorder, RNA sequencing, transcriptome, non-coding RNA

Abstract

Gene expression dysregulation in the brain has been associated with bipolar disorder (BD) previously, through candidate gene and microarray expression studies, but questions remain about isoform-specific dysregulation, and the role of non-coding RNAs whose importance in the brain has been suggested recently but not yet characterized for BD. We used RNA sequencing (RNAseq), a powerful technique that captures the complexity of gene expression, in post-mortem tissue from the anterior cingulate cortex from 13 BD cases and 13 matched controls. We computed case-control differential expression and detected a global trend for downregulation in over 4000 differentially expressed transcripts, of which 10 were significant at a false discovery rate of ≤5%. Among the most significant results, we observed genes coding for Class A G protein-coupled receptors (GPCRs): *SSTR2* (somatostatin receptor 2), *CHRM2* (cholinergic receptor, muscarinic 2) and *RXFP1* (relaxin/insulin-like family peptide receptor 1). Interestingly, a gene ontology analysis of the entire set of differentially expressed genes pointed to an overrepresentation of genes involved in GPCR regulation. We followed-up the top genes by querying the effect of treatment with mood stabilizers commonly prescribed in BD, and found evidence that these drugs affect expression of our candidate genes. By using RNAseq in the post-mortem BD brain, we identified an interesting profile of GPCR dysregulation, pointed to several new BD genes, and characterized the non-coding transcriptome in BD. Our findings have important implications in regards to fine-tuning our understanding of the BD brain as well as for identifying potential new drug target pathways.

Introduction

Bipolar disorder (BD) is an episodic and debilitating mood disorder that affects approximately 1% of the general population (1, 109). Extensive work has been done to understand the role of genes and regulation in the BD brain. Candidate gene and genome-wide microarray expression analyses of post-mortem human brains have shown that transcriptional dysregulation plays a role in the aetiology of BD (for a review see (88)). Some notable genes have been identified through these studies (89, 90), however very few passed corrections for multiple testing and findings from these studies have largely not been replicated (88, 90, 92). RNA sequencing (RNASeq), a technique that takes advantage of the recent development of high-throughput sequencing technologies, offers a number of advantages in comparison to previous methodologies, such as microarray studies, in that it provides direct estimates of transcript abundance as well as nucleotide-level sequence. Differential expression can be measured both at the gene and transcript levels, thus providing unbiased and unparalleled evidence for novel RNAs (93) and regulatory mechanisms such as alternative splicing (94) which are not well represented on microarrays. Transcriptome analysis using RNASeq in BD is timely and important, not only given the power of the technology, but also given the need for greater understanding of the BD pathophysiology and development of effective treatment options.

In order to understand the role that coding and non-coding RNAs play in brain regulation and how their potential dysregulation could impact brain function and ultimately onset of bipolar disorder, we investigated gene expression changes in post-mortem brain tissue from bipolar disorder cases using RNAseq. While the precise neuroanatomical circuits of bipolar disorder are not exactly known and there are data supporting the involvement of diverse brain regions, there is strong support for the role of the anterior cingulate cortex in the regulation of ideo-affective and mood functions and thus in the neurobiology of bipolar disorder (166, 167). Consequently, we focused this post-mortem expression

study on this region, and found a global pattern of downregulation. Furthermore, we identified several differentially expressed genes, and followed-up these findings with an *in vitro* study that showed mood stabilizers lithium, carbamazepine, and valproate to modulate the expression of these transcripts. By using RNAseq we hope to have achieved a more comprehensive level of understanding of the BD brain and shed important light on the dysregulated mechanisms as well as the potential implications for treatment.

Materials and Methods

Post-mortem brain samples and high throughput transcriptome sequencing. Post-mortem brain tissue was obtained from the Douglas-Bell Canada Brain Bank (www.douglasbrainbank.ca) (**Supplementary Methods**). Cases in this study were individuals who had a diagnosis of BD type I or type II (N = 13). Controls (N = 13) had neither current nor past psychiatric diagnoses. Cases and controls were matched for refrigeration delay, age and brain pH (**Table S1**). RNA extraction and sequencing library preparation is described in detail in **Supplemental Methods**. All sequencing was completed on the Illumina HiSeq2000 platform using 100bp paired-end reads (**Table S2**). Reads were aligned to the human genome reference (hg19) using TopHat v2.0.8b (168). On average, 276M paired-end reads had a mapping quality of ≥50, and were used for gene- and isoform-level quantification. For gene-level quantification we employed HTSeq-count version 0.5.4p1(169) (**Figure S1)**. As validation we also ran Cufflinks v2.1.1 (93) for gene-level counts as well as for isoform-level counts. For differential expression analysis, fragment counts were normalized across libraries by using the weighted trimmed mean of log expression ratios (TMM) from the edgeR v3.0.8 R package (170). Furthermore, counts were corrected for heteroscedasticity by employing voom from the limma v3.14.4 R package (171). The linear model used to fit the data included diagnosis, post-mortem interval (PMI) and RNA integrity number (RIN) as covariates.

Gene Ontology analyses. Gene set enrichment analysis was performed using erminej v3.0.2 (http://erminej.chibi.ubc.ca/) (172) with a maximum gene set size of 300 and a minimum gene set size of 5, using the best scoring replicate. The precision-recall analysis was run for 10000 iterations on all the transcripts from the differential expression analysis of the HTSeq genes.

Brain region expression enrichment analysis. We used the HBAset tool (http://www.chibi.ubc.ca/~lfrench/HBAset/) (French and Pavlidis, in preparation). HBAset assembles the

Allen Human Brain Atlas (173) expression data for each gene in an input set and computes an average expression level for each region. It also computes a probability reflecting the degree of enrichment of expression compared to random background genes.

Comparison to prefrontal cortex external dataset. In order to compare our results with those of one previous transcriptome sequencing study in BD (96), raw count expression matrices deposited by Akula *et al.* were obtained from the Gene Expression Omnibus (GEO, GSE53239). Expression matrices from the two platforms described by *Akula et al*. (NISC1 and NISC2) were combined and batch-corrected by removing the first principal component. The first principal component contributed to the 20% of the variance and the scores were significantly different between the two platforms ($P < 0.001$, t-test). The list of differentially expressed transcripts was identified by applying the same procedure used for our data. We performed an over-representation analysis by compiling the list of downregulated transcripts ($p<0.01$) from one study and calculating the AUC against the entire list of downregulated transcript p-values from the other study. We repeated this analysis for upregulated transcripts. ROC curves were plotted with the pROC_1.7.2 R package (171) (**Supplemental Figure 5**).

Neural Progenitor Cell lines chronic drug treatment experiments. Human neural progenitor cells (NPCs), previously characterized (174), were maintained in standard conditions (**Supplemental Methods**). Chronic (1 week) treatments were performed with drugs commonly prescribed in BD: lithium (1mM), valproic acid (1mM), and carbamazepine (50μM), or no-drug control, after which cell pellets were collected and RNA was extracted. In order to validate the brain-like properties of NPCs, we also performed immunohistochemistry with neuron-specific and astrocyte-specific markers MAP2 and GFAP respectively (**Supplemental Methods**).

Quantitative Real-time Polymerase Chain Reaction (qRT-PCR). Brain RNA for RNAseq and qRT-PCR validation was used from the same original extraction. Complementary DNA (cDNA) synthesis and qRT-PCR was performed as previously described (175) (**Supplemental Methods**). We investigated the stability of common endogenous genes in each sample set and determined the most suitable to be *POLR2A* (Polymerase (RNA) II (DNA directed) polypeptide A) or *ACTB* (Beta Actin) using the NormFinder Algorithm (176) (**Table S3**). All graphical data are presented as the mean ± s.e.m. Statistical differences between groups were analyzed by Student's t-tests, Mann-Whitney tests, One-Way ANOVA with Dunnet's post-hoc corrections or Pearson's correlation coefficients. Statistical significance was calculated using GraphPad Prism5 and SPSS 20. A p-value of ≤ 0.05 was considered statistically significant, and ≤ 0.1 was considered suggestive of a trend for significance.

Results

Transcriptome sequencing in the anterior cingulate cortex (ACC) of bipolar disorder post-mortem brains

We used a directional library protocol that allows distinction of genes overlapping at the same chromosomal locus and a ribosomal-depletion transcriptome selection to allow for identification of non-poly(A)-tailed RNAs. On average, 318M 100bp paired-end reads were mapped per individual, of which on average 276M reads had a mapping quality of ≥50 based on TopHat (93) . For gene-level quantification we used HTSeq-count (169). We found on average 68.1M fragments mapping to 60,905 genes, while on average 64.6M fragments did not map to any genes from the reference annotation. This is expected as many reads map to introns undergoing splicing or as-yet uncharacterized transcribed regions (177). We removed RNA transcripts with zero or aberrantly high counts (e.g. RN7SL2, RN7SK) and were left with a total of 27,706 genes. About 61% of the fragments were attributed to protein-coding genes, while the remaining fragments were attributed to other RNA classes including lincRNAs, pseudogenes, antisense RNAs, etc. (**Figure S2**).

Gene-level differential expression – a global trend for downregulation

After quantifying gene-level expression using the HTSeq-count pipeline (169), we identified all differentially expressed (DE) transcripts (**Table S4**) and by-and-large we found a strikingly prominent global downregulation, with 70% of overall DE transcripts being down-regulated, (**Figure S3A**) and a comparable enrichment of downregulated genes among the top 100 genes ranked by p-value (72 of 100) (**Figure S3B**). Of these, 10 were significantly DE at a false discovery rate (FDR) ≤0.05 (**Table S4**). Interestingly, all 10 transcripts that passed the stringent FDR cutoff were downregulated and all

were protein-coding genes (**Table 1**). In order to replicate our findings obtained with HTSeq, we ran Cufflinks (**Table S5**), an alternate method for which there is currently no gold standard way of getting raw counts. However, its advantage is that in addition to gene-level, allows for isoform-level DE analysis (93). All 10 genes from the initial analysis were found to be differentially downregulated at FDR≤0.05, (**Table 1**). Following validation by qRT-PCR, we found that all 10 genes had fold changes in the expected direction, 8 of the 10 genes' expression values were correlated with the RNAseq data at a suggestive (p-value≤0.1) level and 8 were nominally significant (p-value≤0.05) with qRT-PCR (**Table 1, Figure S4**). These results support the accuracy of our expression quantification and DE analysis and the cutoffs we applied in representing the true transcriptomic landscape. Furthermore, while these genes are largely unstudied, it is worth noting that three (*RXFP1*, *SSTR2*, *CHRM2*) of the top genes belong to class A of the G protein-coupled receptor (GPCR) family of genes which potentially suggests similar functions.

Enrichment of G protein-coupled receptor pathways in the differentially expressed genes

In order to understand broader patterns of differentially expressed genes in the ACC, we performed a gene set enrichment analysis on the all genes from the HT-Seq DE analysis using the Precision-recall method in ErmineJ v3.0.2 (172). This method uses the ranks of gene scores rather than the gene scores for computing p-values for each gene set. Interestingly, the top two biological processes identified that also passed corrections for multiple testing were "G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger" (GO: 0007187, corrected p=5.24E-09) and "adenylate cyclase-modulating GPCR signaling pathway" (GO: 0007188, corrected p=2.62E-09) (**Table S6**). Finally, we investigated the top 10 candidates to see if they are enriched in expression in particular brain regions, considering them as a group. Using an "expression enrichment" tool for the

Allen Human Brain Atlas data, HBAset (French and Pavlidis, in preparation), we find that expression of the 10 candidate genes is significantly enriched in many cortical regions compared to random genes, including regions previously connected to BD (**Figure S5**), and in fact the ACC is one of the top significant regions (p-value=0.0013; **Table S7**). This suggests that concerted dysregulation of these genes might have effects on multiple cortical regions, including the cingulate cortex.

The effect of psychiatric drugs on identified genes

In order to explore the role of medication commonly used to treat BD on the dysregulated genes, we investigated the effect of lithium, valproate, and carbamazepine on the expression of the top differentially-expressed genes. We performed an *in vitro* chronic treatment in neural progenitor cell lines that express both neuron-specific marker *MAP2* (microtubule-associated protein 2) and astrocyte-specific marker *GFAP* (glial fibrillary acidic protein) (**Figure 1**). All three drugs significantly decreased expression of *CHRM2* and *VWC2L*, while increasing expression of *DIRAS2*. Interestingly, only valproate had an effect on expression of *SLC7A14* and *SSTR2*, where both genes were upregulated (**Figure 1**).

Isoform-level analysis

We used Cufflinks to determine isoform-level differential expression, and after removing transcripts with very high or low expression, 120,845 remained. No transcripts were differentially expressed at FDR≤0.05 (**Table S8**). However, many of the top ranked isoforms by p-value belonged to the same genes identified by gene-level analysis. For proof-of-principle, we wanted to see if we could detect any of these differences with qRT-PCR. Since the gene *B3GALT2* only has one isoform, we opted for

the second-ranked transcript, ENST00000445907, also known as Isoform 1 of the gene *CHRM2* (**Figure 2A**). We designed assays that would target all isoforms (**Figure 2B**), only ENST00000445907 (**Figure 2C**), and other isoforms besides ENST00000445907 (**Figure 2D**). We found that the total gene downregulation (p-value = 0.0004) is driven by the ENST00000445907 isoform (p-value =0.0002) and not maintained when this isoform is not targeted by the qRT-PCR assay (p-value =0.1065).

Analysis of non-coding RNAs

The approach to prepare sequencing libraries using ribosomal depletion as the transcriptome selection method allowed us to also retain non-poly(A)-tailed RNAs. These include potentially interesting RNA classes such as long intergenic non-coding RNAs (lincRNAs), antisense RNAs, small nuclear and small nucleolar RNAs, etc. The relative abundance of these non-coding RNA (ncRNA) classes is lower than coding RNAs (**Figure S2**), though much higher than previously suspected (177). A large number of reads was attributed to pseudogenes, a finding which has been previously reported (178). In the DE analysis no ncRNA passed (FDR≤0.05) corrections for multiple testing, however the top two DE ncRNAs, linc-KARS-3 (also known as TCONS_0024733) and linc-SFSWAP-3 (also known as TCONS_0021259) were ranked 14 and 16, respectively by p-value (**Table S4**). Interestingly, RP11-638F5.1 (also known as TCONS_0020164), ranked 23, mapped to the same genomic location on chromosome 12 as linc-SFSWAP-3 and appears to be a shorter isoform of the same locus, sharing two exons (**Figure S6B**). We found all three of these lincRNAs to be downregulated in the RNAseq data, and succeeded in showing the same effect by qRT-PCR analysis (linc-KARS-3, p-value = 0.0487) (**Figure 3**). For the Chr12 locus, an assay targeting the two shared exons resulted in a significant downregulation (p-value=0.0503) for linc-SFSWAP-3, while an assay querying just RP11-638F5.1

also resulted in a significant downregulation (p-value=0.0138) (**Figure 3**). For technical reasons, an assay specific to linc-SFSWAP-3 could not be designed (**Figure S6B**). All qRT-PCR results were significantly correlated with the RNAseq data. These results along with the fact that there was no significant correlation between the two qRT-PCR datasets, suggests that the RNASeq findings at this locus are independent of each other.

Comparison with Akula *et al.* prefrontal cortex study

Recently, another study was published using RNAseq in BD (96) and even though this work profiled a different brain region, namely prefrontal cortex, we wanted to investigate potential consistencies with our own findings as different cortical regions have sometimes been shown to have similar expression patterns(179, 180). Firstly, we performed a re-analysis of the combined and batch-corrected Akula *et al*. data from two platforms (Illumina GA-IIx and HISeq2000, N total=21) using the same pipeline used for the analysis of our data. When testing whether all DE transcripts from one dataset tended to be at the top in the other dataset (ranked by p-value) we found the significantly downregulated transcripts in our dataset (1,761 at $p<0.01$) to be enriched at the top of the downregulated transcripts in the Akula *et al*. study (AUC = 0.664) (**Figure S7A**). Likewise, the Akula *et al*. reanalyzed list of significantly downregulated transcripts (956 at $p<0.01$) were enriched in the top of the downregulated transcripts in our study (AUC = 0.611) (**Figure S7B**). On the other hand, upregulated transcripts showed no enrichment between the two datasets (AUC = 0.532 using our list of 324 upregulated transcripts and AUC = 0.533 using the list of 529 upregulated Akula *et al*. transcripts).

Discussion

In this study we investigated the transcriptome of individuals with bipolar disorder (BD) in post-mortem brain samples from the Anterior Cingulate Cortex (ACC, BA24) using RNASeq. Extensive evidence from microarray and candidate gene studies has demonstrated the role of transcriptional dysregulation in the aetiology of BD (88). Furthermore, accumulating evidence is starting to point to the key involvement of as-yet uncharacterized non-coding RNAs, in addition to the more commonly studied protein-coding RNAs in psychiatric disorders.

We showed excellent validation of our methods both at the bioinformatics level across two different pipelines, as well as at the molecular level with qRT-PCR. We identified a number of interesting dysregulated genes. By-and-large we found a strikingly prominent global downregulation, with all differentially expressed transcripts that passed multiple testing corrections (FDR <0.05) being downregulated as well as an enrichment of downregulated genes among the top 100 genes ranked by p-value and a much more consistent expression pattern across subjects in downregulated genes compared to upregulated ones. The top gene identified by both gene-level bioinformatics pipelines as well as the isoform-level analysis was *B3GALT2* (UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2) (181). This gene is a member of the beta-1,3-galactosyltransferase (beta3GalT) family which encodes type II membrane-bound glycoproteins. Though very little is known about *B3GALT2* and associations with BD have not yet been documented, members of this family have been shown to be primarily brain-expressed in the mouse (182). It is worth noting that Akula *et al.* (96), who recently published the only other RNAseq study of the BD brain, detected this gene to be dysregulated in their analyses of the PFC, and our re-analysis of their dataset also detected this gene to be similarly and significantly downregulated. Further work is

warranted to validate this top finding in other BD cohorts as well as to characterize its dysregulation in the BD brain.

One of the most interesting findings was the differential expression of G protein-coupled receptors (GPCRs), as suggested by the gene-set enrichment analysis that indicated an enrichment of G protein-coupled receptor pathways among the DE genes, and supported by the fact that three of the top DE genes, *SSTR2* (somatostatin receptor 2), *CHRM2* (cholinergic receptor, muscarinic 2) and *RXFP1* (relaxin/insulin-like family peptide receptor 1), are GPCRs belonging to Class A of this superfamily. GPCRs, noted drug targets in many disorders including those afflicting the brain (183), have previously been linked to mood disorders including BD (142). Of the top three GPCRs that we have identified in this study, only *CHRM2*, a muscarinic receptor defined by the binding of acetylcholine and involved in adenylate cyclase inhibition, phosphoinositide degeneration, and potassium channel mediation, has been previously linked to BD (184, 185). *SSTR2* mRNA levels have been shown to be decreased in the prefrontal cortices of schizophrenia patients (186) and to decrease in response to stress in animal models (187). Our study suggests that these three GPCRs in particular, as well as GPCRs as a whole, should be further investigated in BD.

In order to further explore the role of the identified dysregulated genes, we investigated the effect of mood stabilizers lithium, valproate, and carbamazepine on the expression of the most significant DE genes through an *in vitro* chronic treatment experiment in cultured neural progenitor cells (characterized previously (174)). These drugs were selected based on a long-standing history of documented efficacy in the clinical treatment of BD (1, 109). Since the majority of differentially expressed transcripts were significantly downregulated in the bipolar brain, we were interested in the possibility that the expression of these genes would be upregulated by mood stabilizers. The expression of *DIRAS2* was upregulated by all the three drugs tested, while the expression of *SSTR2*

and *SLC7A14* was significantly upregulated by valproate only. Very little is known about the implication of *DIRAS2* (DIRAS family, GTP-binding RAS-like 2) in the brain, as it has only been studied in adult Attention Deficit Hyperactive Disorder (ADHD) (188). Neither *SLC7A14* (solute carrier family 7, member 14) nor *SSTR2* (somatostatin receptor 2) have previously been investigated in BD or valproate treatment, but the latter is a GPCR belonging to class 4A that has been implicated in adaptive response to stress (187, 189). Furthermore, other somatostatins have been linked to BD genetics (190-192). While encouraging, these results provide but the first steps in the attempt to elucidate how commonly prescribed mood stabilizers may influence the expression of genes found dysregulated in the BD brain.

Finally, to our knowledge this is the first study in BD and more generally in psychiatry that uses ribosomal depletion in the preparation of RNA sequencing libraries, and thus can quantify all classes of RNA of 100b base pairs or longer, regardless of their possession of a poly(A) tail. Furthermore, since we used a very high coverage, we were able to detect some very lowly expressed transcripts including long intergenic non-coding RNAs (lincRNAs), antisense RNAs, small nuclear RNAs (snRNA), and other non-coding classes that are not as abundant as protein-coding transcripts. This analysis is of interest because more and more reports have emerged in the last few years documenting the importance of non-coding RNAs in normal brain development, maintenance, and aging (193-195), as well as a variety of conditions including neurodevelopmental disorders like autism (196, 197). Though they did not pass multiple testing corrections, the top three (ranked by p-value) most significant lincRNAs were significantly downregulated when validated using qRT-PCR, suggesting that these lincRNAs should be further investigated in BD. Unfortunately, we have no information from the literature to help us understand how dysregulation of these lincRNAs could be connected to the expression of the coding genes identified. Since the exploration of non-coding RNA species is still in

its infancy, undoubtedly computational tools will improve along with our understanding of these RNA classes, allowing us to extract even more valuable knowledge. Further work is warranted to fully exploit the abundance of information collected with total transcriptome sequencing analysis.

Acknowledgements

We are grateful for the contributions made by the families who consented to donate brain tissue to the Douglas-Bell Canada Brain Bank. We thank Dr. Carl Ernst at McGill University for his contribution of human NPCs. CC received a doctoral funding award from CIHR. GT is an FRQS chercheur national.

Conflict of Interest

The authors declare no conflict of interest.

Tables and Figures

Table 1: Top differentially expressed genes dysregulated in the BA24 of BD individuals. Gene-level expression quantification followed by differential expression analysis using the HTSeq-count identified 10 genes, all protein-coding, significant at a false discovery rate (FDR) ≤0.05. For bioinformatics validation of HTSeq findings we ran the Cufflinks pipeline, which identified the same 10 genes to be significant at FDR≤0.05, though in a slightly different rank order by p-value. For biological validation we performed qRT-PCR on these genes and showed expression of all in the expected direction with 8 being nominally significant (p-value≤0.05). Excellent correlation of expression values was achieved with both validation analyses. Legend: *** ≤0.001; ** ≤0.01; * ≤ 0.05; # ≤0.1; ns ≥ 0.1.

Gene info				RNAseq HT-Seq				RNASeq Cufflinks			qRT-PCR Validation		
Gene	Description	Position	Rank	FC	P-Value	Adjusted P-Value	Rank	FC	P-Value	Adjusted P-Value	FC	P-value Sig.	Correlation Sig.
B3GALT2	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2	Chr1:193148175-193155784	1	0,3824	5,49E-07	0,0083	1	0,3804	3,68E-07	0,0074	yes	***	#
CHRM2	cholinergic receptor, muscarinic 2	Chr7:136553416-136705002	2	0,3932	6,10E-07	0,0083	10	0,4296	1,64E-05	0,0424	yes	***	***
VWC2L	von Willebrand factor C domain containing protein 2-like	Chr2:215275789-215443683	3	0,3886	8,99E-07	0,0083	2	0,3795	5,71E-07	0,0074	yes	ns	**
RXFP1	relaxin/insulin-like family peptide receptor 1	Chr4:159236463-159574524	4	0,2855	2,29E-06	0,0159	6	0,2748	4,75E-06	0,0204	yes	*	***
SLC35F1	solute carrier family 35, member F1	Chr6:118228689-118638839	5	0,5843	3,14E-06	0,0174	3	0,5496	9,44E-07	0,0081	yes	ns	ns
RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated)	Chr15:38780304-38857776	6	0,3187	4,32E-06	0,0199	7	0,3536	8,50E-06	0,0313	yes	**	ns
SSTR2	somatostatin receptor 2	Chr17:71161151-71167185	7	0,2801	6,91E-06	0,0267	8	0,2664	1,42E-05	0,0417	yes	*	***
DIRAS2	DIRAS family, GTP-binding RAS-like 2	Chr9:93372114-93405386	8	0,2938	7,71E-06	0,0267	5	0,2835	4,23E-06	0,0204	yes	*	*
LRRC55	leucine rich repeat containing 55	Chr11:56949221-56959191	9	0,3786	1,35E-05	0,0393	9	0,4293	1,46E-05	0,0417	yes	**	***
SLC7A14	solute carrier family 7 (orphan transporter), member 14	Chr3:170182353-170303863	10	0,4573	1,42E-05	0,0393	4	0,4967	4,09E-06	0,0204	yes	*	***

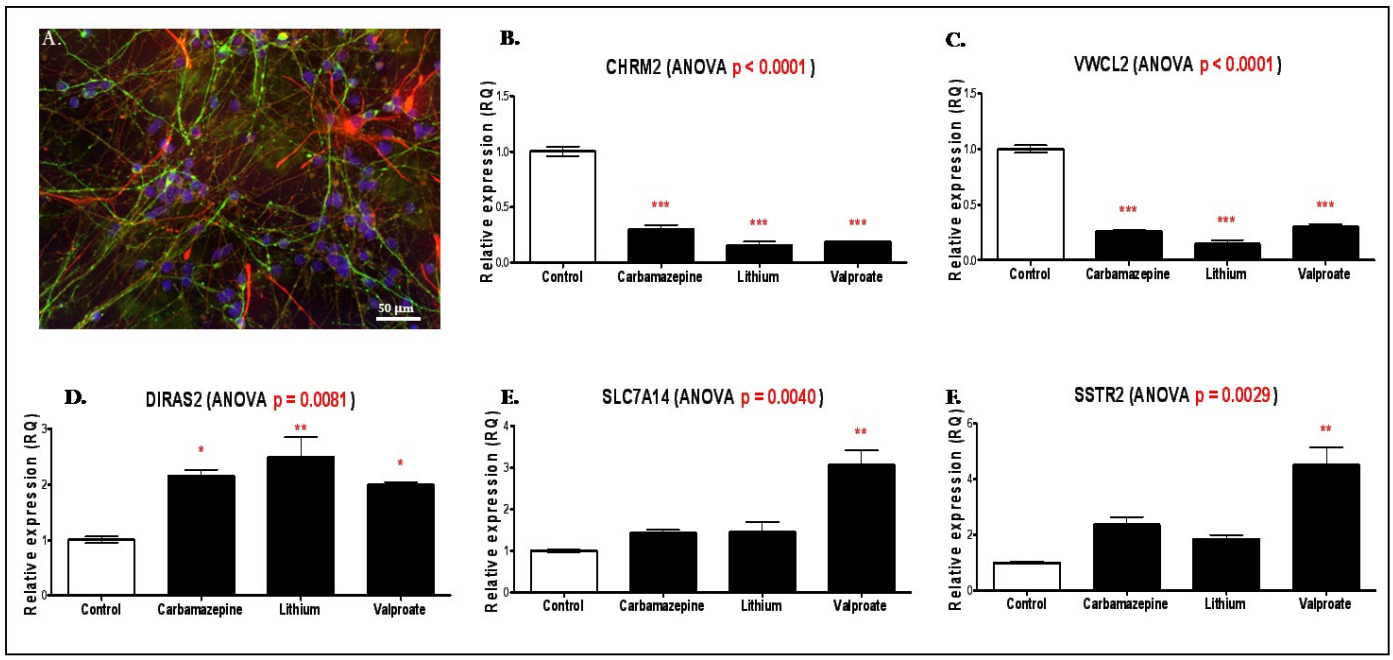


Figure 1: **Mood stabilizer treatment effects on top differentially expressed genes.** The effect of lithium (1mM), valproate (1mM), and carbamazepine (50µM) treatment on the expression of the top differentially expressed genes was quantified through an *in vitro* chronic treatment assay in neural progenitor cell lines. **A.** The cells represent a brain model in that they express either neuron-specific marker *MAP2* (microtubule-associated protein 2) or astrocyte-specific marker *GFAP* (glial fibrillary acidic protein) at the time of treatment start. **B-D.** All three drugs affected expression of *CHRM2*, *VWC2L*, and DIRAS2. **E-F**. Valproate had a specific upregulating effect on *SLC7A14* and *SSTR2*. Legend: *** ≤0.001; ** ≤0.01; * ≤ 0.05; # ≤0.1; ns ≥ 0.1.

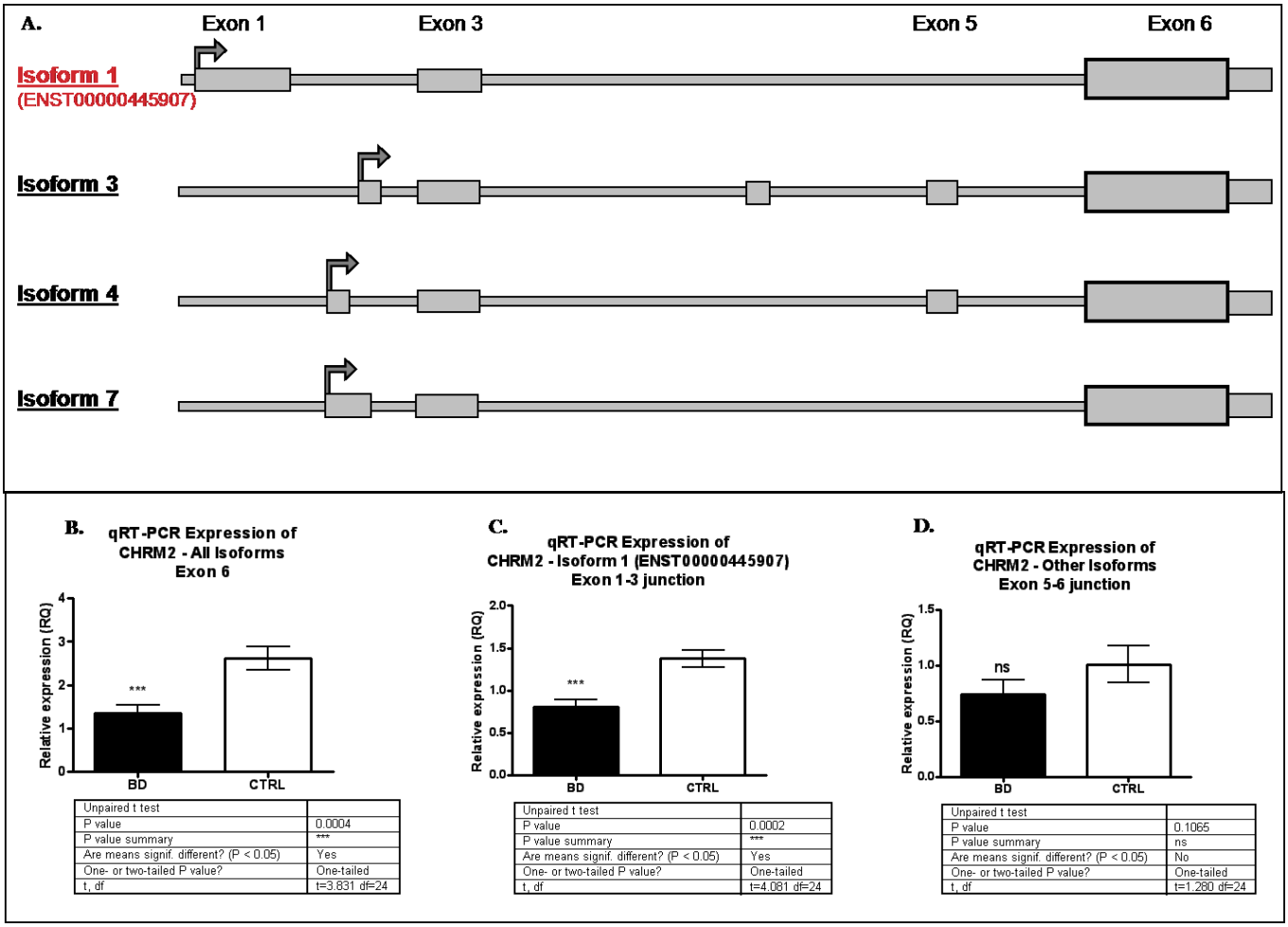
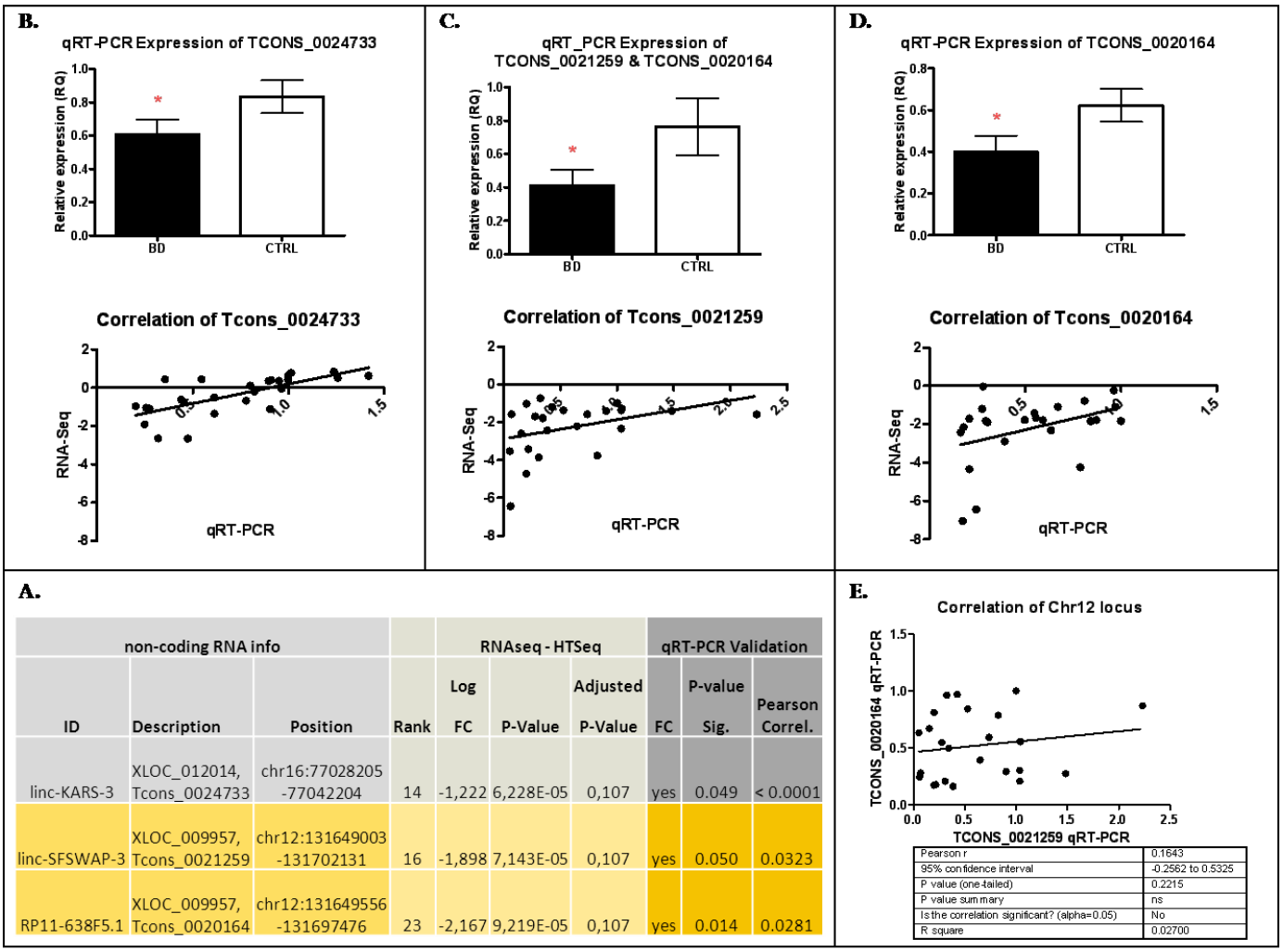


Figure 2: **Isoform-specific expression validation of CHRM2 by qRT-PCR**. **A.** Isoform structure of the gene CHRM2. **B.** qRT-PCR analysis of all isoforms of the gene shows a significant decrease in BD subjects (p-value = 0.0004). **C.** qRT-PCR validation results for only isoform 1, also known as ENST00000445907, show that the whole-gene effect is driven by this isoform (p-value =0.0002). **D.** A qRT-PCR assay excluding isoform 1 was not statistically significant (p-value =0.1065).



non-coding RNA info				RNAseq - HTSeq			qRT-PCR Validation		
ID	Description	Position	Rank	Log FC	P-Value	Adjusted P-Value	FC	P-value Sig.	Pearson Correl.
linc-KARS-3	XLOC_012014, Tcons_0024733	chr16:77028205 -77042204	14	-1,222	6,228E-05	0,107	yes	0.049	< 0.0001
linc-SFSWAP-3	XLOC_009957, Tcons_0021259	chr12:131649003 -131702131	16	-1,898	7,143E-05	0,107	yes	0.050	0.0323
RP11-638F5.1	XLOC_009957, Tcons_0020164	chr12:131649556 -131697476	23	-2,167	9,219E-05	0,107	yes	0.014	0.0281

Figure 3: **Non-coding RNA results and validation**. **A.** Table showing chromosomal locations and RNASeq results of top ranked lincRNAs, as well as subsequent qRT-PCR validation analysis results. **B.** Linc-KARS-3 (also known as TCONS_0024733) is significantly decreased in BD and the qRT-PCR data correlates significantly with the RNAseq data. **C.** An assay targeting both lincRNAs that map to the same Chr12 locus, linc-SFSWAP-3 (also known as TCONS_0021259) and RP11-638F5.1 (also known as TCONS_0020164), shows a significant decrease in BD and the values are significantly correlated with RNAseq data for linc-SFSWAP-3. **D.** An assay targeting just RP11-638F5.1 (also known as TCONS_0020164) shows a statistically significant decrease that is correlated significantly with the RNAseq data. **E.** Significant correlation of qRT-PCR results for both lincRNAs mapping to the Chr12 locus.

Chapter 3: Candidate gene studies

Part 3.1: Preface

Unlike Chapter 2, which takes a global approach at defining BD susceptibility, the work presented here has a candidate-gene focus motivated by previous evidence generated by our group (16). This work described a linkage study in 36 multiplex families ascertained through BD probands characterized for excellent response to treatment with lithium, which pointed to three chromosomal regions linked to BD: 3p25.1, 3p14.1, and 14q11.2. This was followed by gene expression studies in cortical regions from bipolar disorder post-mortem brains, in order to select specific genes for further investigation. The gene Synapsin II (*SYN2*), located at 3p25.1 was identified as one of the most interesting candidates, and at the mRNA expression level it was shown to be upregulated in the prefrontal cortices of patients (16).

The work in Chapter 3.2 follows up on the *SYN2* gene. Since in the original study the candidate genes were ascertained through a lithium-responsive cohort, we sought to determine if lithium treatment *in vitro* could modulate *SYN2* expression. To model the genetic background of Li-responder BD patients, long-term Li treatment assays were performed in B-lymphoblastoid cell lines from BD patients classified as excellent lithium responders, non-responders, or non-psychiatric controls. Additionally, to model brain expression patterns, treatment assays were performed in brain-specific cell lines and gene expression changes were assessed using quantitative real-time PCR. In both models, we found *SYN2* to be upregulated by the presence of lithium in cell culture (163) - which corroborated our previous findings in post-mortem BD brain samples (16).

The work in Chapter 3.3 seeks to further explore the finding of dysregulated synapsin expression in the bipolar disorder brain and explore potential regulatory mechanisms. The focus is on epigenetic regulatory mechanisms, specifically histone modifications in gene promoter regions. Alterations at this level of gene regulation have been implicated in previous investigations of a number

of psychiatric disorders, including BD and major depressive disorder (MDD). We started by analyzing expression of synapsin variants in the prefrontal cortex (Brodmann Area 10) of post-mortem brains from BD as well as MDD subjects compared to non-psychiatric controls. This added to our previous work by including all three synapsin genes (*SYN1*, *SYN2*, and *SYN3*) with alternative splicing resulting in several variants with high levels of homology. This was interesting because the other synapsin genes have also been postulated to play roles in the etiology of BD and other related disorders such as schizophrenia and autism.We then queried the potential associations with histone modifications. We showed distinct profiles for the genes' expression in the two related disorders, as well as a potential regulatory role for histone 3 lysine 4 tri-methylation (H3K4me3), a histone modification which is believed to open the chromatin structure in the promoter region close to the transcription start site and encourage active transcription.

Finally, Chapter 3.4 provides a summary of the current research relating to expression of synapsin genes and their regulation, particularly as applied to epigenetic modifications.

Part 3.2: Synapsin II is involved in the molecular pathway of lithium treatment in bipolar disorder

Title: Synapsin II is involved in the molecular pathway of lithium treatment in bipolar disorder

Authors:

Cristiana Cruceanu[1], Martin Alda[2], Paul Grof[3], Guy A. Rouleau[4], Gustavo Turecki[1#]

Affiliations:

[1] McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, QC, Canada.

[2] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada.

[3] Mood Disorders Centre of Ottawa, Ottawa, ON, Canada.

[4] Centre of Excellence in Neuromics, CHUM Research Center and the Department of Medicine, University of Montreal, Montreal, QC, Canada

Corresponding Author:
McGill Group for Suicide Studies
Douglas Mental Health University Institute
Frank B Common Pavilion
Room F-3125
6875 LaSalle Boulevard
Montreal (Quebec) H4H 1R3
Email: gustavo.turecki@mcgill.ca
Phone: 514-761-6131 ext: 2369
Fax: 514 762-3023

Abstract

Bipolar disorder (BD) is a debilitating psychiatric condition with a prevalence of 1-2% in the general population that is characterized by severe episodic shifts in mood ranging from depressive to manic episodes. One of the most common treatments is lithium (Li), with successful response in 30-60% of patients. Synapsin II (SYN2) is a neuronal phosphoprotein that we have previously identified as a possible candidate gene for the etiology of BD and/or response to Li treatment in a genome-wide linkage study focusing on BD patients characterized for excellent response to Li prophylaxis. In the present study we investigated the role of this gene in BD, particularly as it pertains to Li treatment. We investigated the effect of lithium treatment on the expression of SYN2 in lymphoblastoid cell lines from patients characterized as excellent Li-responders, non-responders, as well as non-psychiatric controls. Finally, we sought to determine if Li has a cell-type-specific effect on gene expression in neuronal-derived cell lines. In both *in vitro* models, we found *SYN2* to be modulated by the presence of Li. By focusing on Li-responsive BD we have identified a potential mechanism for Li response in some patients.

Key Words: Bipolar disorder, lithium, synapsin II, gene expression, treatment response

Introduction

Bipolar disorder (BD) is a major topic in health research given its debilitating nature, lifetime prevalence and significantly high occurrence in the general population (1-2%) (1).This psychiatric condition is characterized by abnormal shifts in energy, activity levels, mood, and one's ability to carry out routine tasks. In comparison to other psychiatric conditions, BD has been shown to have relatively high heritability, with estimates ranging from 60 to 85% (1, 108). One of the most common treatments of BD is lithium (Li), administered as metallic salts, due to its proven efficacy both as a short term intervention for manic episodes as well as a prophylactic against episode recurrence. The drug has been highly prescribed since the 1950s and 1960s when Mogens Schou showed its efficacy through a series of systematic trials with BD patients (198), and demonstrated a high success rate with approximately 30-60% of patients showing full or partial treatment response (199, 200).

Synapsin II (SYN2) is a gene that codes for a neuronal phosphoprotein involved in synaptic plasticity and transmission as well as synaptogenesis. It maps to chromosome 3p25 and has two known variants, IIa and IIb, which are highly expressed in nerve terminals in the majority of the adult brain (201) with demonstrated homology across numerous vertebrate and invertebrate organisms (202). The majority of brain regions co-express synapsin genes at similar levels, suggesting that they are functionally complementary (203), and though all synapsins have been primarily studied for their roles in the brain, the genes' expression is widespread in the peripheral nervous system. In non-neuronal cells, synapsins are mostly found in association with the cytoskeleton, where their involvement is likely at the level of vesicular trafficking (202). For example, Syn2 protein was isolated from rat as well as bovine chromaffin cells of the adrenal medulla (204, 205). Though limited work has been done on *SYN2* outside of neurons, expression of other synapsins has been shown in

undifferentiated astrocytes (206), osteoblasts (207), liver endosomes (208), epithelial cells (209), , as well as the cell lines HeLa and NIH/3T3 (210).

Given the multiple roles played by synapsins in neuronal cell function and maintenance, it may be hypothesized that disruption of these roles could result in the onset of pathological conditions. Indeed, knockout experiments have shown the absence of *SYN2* to induce epileptic-like seizures in mice (211, 212) and genetic mapping identified variants in the *SYN2* gene as significantly contributing to epilepsy predisposition (212, 213). Genetic association studies have also linked *SYN2* variants with schizophrenia, as shown in affected families of different genetic backgrounds (214-216). Data for BD are more limited, however. The only reported case-control analysis of SYN2 single nucleotide polymorphisms (SNPs) in individuals with BD comes from Wang *et al*. who studied the Han Chinese population but did not find any significant association (217). Additional work has been reported for *SYN2* at the protein or mRNA levels, where several studies showed significant dysregulation in alcoholism, Huntington's disease, and schizophrenia (218-220). In BD, Vawter *et al*. showed differential down-regulation of SYN2 protein levels in hippocampi of patients compared to non-psychiatric controls. We have recently published a linkage study in families ascertained through Li-responsive BD probands, where the *SYN2* gene was identified as one of the more interesting candidates (16). In the same study, at the mRNA expression level, *SYN2* was shown to be up-regulated in the prefrontal cortex of patients (16). In the present study, we hypothesize that the implication of *SYN2* in BD is more prominent in a subset of BD patients. Moreover, we predict that in such patients *SYN2* is more relevant to the response to lithium treatment.

To explore these hypotheses, we conducted a series of studies investigating the expression of *SYN2* in BD, particularly as it pertains to lithium treatment. Because this candidate gene was originally identified through a linkage study of lithium-responsive BD families, we investigated what effect

lithium treatment would have on the expression of *SYN2*. We performed *in vitro* long-term treatment studies in Epstein-Barr-virus transformed lymphoblastoid cell lines (LCLs) from BD patients characterized for excellent Li-response (as described previously) (7, 11, 164) in order to identify the effect of this drug in a model replicating the genetic background of response. In addition, we performed the same experiments with human neuroblastoma and glioblastoma cells to model the biological context.

Methods and Materials

I. Ethics statement. Ethics approval for the use of human samples in this study was obtained from the Capital District Health Authority (CDHA) in Halifax, Nova Scotia. All subjects gave written informed consent to their participation in the study in regards to sample collection and the generation of lymphoblastoid cell lines; no subjects had reduced capacity to consent. Sample collection and cell lines generation has been described previously (164, 221).

II. BD Li-response lymphoblastoid samples. Subjects were diagnosed with BD I and BD II according to both Research Diagnostic Criteria (RDC) and DSM-IV criteria, and followed prospectively at specialized clinics in Hamilton, Ottawa and Halifax (11). Their clinical course was characterized by a high number of manic and depressive episodes before Li treatment. The responders (n = 11) showed full stability on long-term Li monotherapy. The non-responders (n = 12) continued experiencing illness episodes in spite of good compliance documented by therapeutic blood levels. These are the same criteria as outlined previously (7, 11). Unaffected controls (n = 13) were matched for ethnic background and excluded if they had a history of BD, schizophrenia, or major depression. Peripheral blood samples were obtained from patients and controls following standard procedures and

Epstein-Barr virus-transformed β-lymphoblastoid cell lines were generated as described previously (164, 221).

III. Cell culture. To determine patient-specific effects of Li on target genes, *in vitro* assays were performed in LCLs from excellent Li-responders, non-responders, and healthy controls. Aliquots of frozen cell lines were stored in liquid nitrogen after Epstein-Barr virus transformation for each sample according to "LCL frozen storage" time until all samples were randomized, thawed for experiments, grown and processed in a sequential fashion as described below. This effectively ensures no difference in passage number between LCL samples and no batch effect. Cells were cultured in Iscove's Modified Dulbecco's Medium (IMEM) supplemented with 15% FBS, 1% Fungizone and 1% penicillin/streptomycin/glutamine (Invitrogen) in a 5% CO_2 humidified incubator at 37°C, in the continuous presence of 1.0 mM LiCl or vehicle (NaCl) for 7 days (164) after which cell pellets were collected and frozen at -80^{o}C. Experiments were performed in triplicate. Clinical and demographic characteristics of patient and control LCLs are listed in Table 1.

To determine cell-type-specific modulation of candidate genes in the brain, *in vitro* assays were performed in three cell lines: HEK293 (human embryonic kidney, ATCC CRL1573) as a non-brain control, SK-N-AS (human neuroblastoma, ATCC CRL2137), and U-118 MG (human glioblastoma; astrocytoma, ATCC HTB15). Cells were cultured at 37°C in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS, 100 U/ml penicillin and 100 μg/ml streptomycin (Invitrogen) in a 5% CO_2 humidified incubator at 37°C. For Li treatments, cells were grown in the continuous presence of 0.5 mM, 1.0 mM, or 2.0 mM LiCl or vehicle (NaCl) for 7 days after which cell pellets were collected and frozen at -80^{o}C. Experiments were performed in triplicate.

IV. Real-time PCR. Total RNA was extracted from frozen cell pellets using the RNeasy Mini Kit (QIAGEN). For synthesis of cDNA, M-MLV reverse transcriptase (Gibco, Burlington, Ontario) and oligo(dT)16 primers (Invitrogen) were used. Real-time PCR reactions were run in quadruplicate using an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) and the Power SYBR® Green PCR Master Mix (Applied Biosystems). Relative expression was calculated using the relative quantitation method (ΔΔCt) in the RQ Manager 1.2 software (Applied Biosystems) with GAPDH as an endogenous control.

V. Data analysis. Test coefficients and probability distributions were calculated using statistical software GraphPad Prism 5 and SPSS.

Results

Lithium affects gene expression in transformed lymphoblastoid cell lines (LCLs) distinctly in lithium responders compared to both non-responder BD patients and controls.

To determine patient-specific effects of Li on the target genes, *in vitro* assays were performed in Human Epstein-Barr virus–transformed LCLs from excellent Li-responders (R), non-responders (N) and controls without psychiatric history (C) (221). For long-term treatment, cells were cultured in the continuous presence of 1.0 mM treatment (LiCl) or vehicle (NaCl) for 7 days (164). Data in Figure 1 are presented as fold change between Li treatment and vehicle treatment values. We performed a ANCOVA analyses with "Age at Sampling" and "LCL frozen storage" as covariates, followed by Tukey's multiple comparison post-tests for group comparisons, but found no significant mean differences between the three groups: C vs. R, C vs. N, and R vs. N for either Synapsin II variant (SYN2a p=0.613, SYN2b p=0.691), as shown in Table 2.

Interestingly, there was a significant difference in the distribution of expression fold-change in the responder patient group as compared to the non-responders and the controls. LCLs from non-responder BD patients displayed the same distribution pattern as the controls whereas the Li-responder patient LCLs had a broader spectrum of expression than the other two groups. The same pattern was observed with the SYN2a variant shown in Figure 1.A (F-test P=0.001 for both C vs. R and N vs. R) as with the SYN2b variant shown in Figure 1.B (F-test P<0.001 for both C vs. R and N vs. R). Furthermore, the expression pattern was consistent across the two variants, with subjects showing consistently low or high expression in both the *SYN2a* and *SYN2b* variant. This was illustrated through the color-coding in Figure 1.

Environmental factors do not explain the variant effect of lithium in Responders.

Given the fact that in some patient LCLs both *SYN2a* and *SYN2b* were up-regulated by lithium treatment while in others the two variants were down-regulated, we attempted to elucidate the stratifying factors responsible for this behavior. Ethnic background did not differ across subjects as all were Caucasian of European descent, so this variable was not included in the analyses. We investigated a number of other factors including age of onset, initial Li prescription, and time on Li prior to DNA collection (Table 3). Furthermore, we investigated factors relating to psychiatric medication such as Li dosage and use of other medications, as well as family history of other psychiatric disorders. We determined normality of each dataset using a Shapiro-Wilk normality test and computed Pearson's correlations for normally distributed and Spearman's correlations for non-normally distributed datasets. None of the 15 potential environmental covariates showed significant correlations with either *SYN2a* or *SYN2b* expression values, demonstrating that the reported variance difference cannot be explained by these possible covariates (Table 3).

Synapsin II shows cell-type specific response to lithium treatment in neuroblastoma cells.

Since our previously reported brain expression results (16) were from homogenate tissue brain extracts, we set out to investigate a possible cell-type-specific effect of lithium treatment. As such, we used three cell lines representing neurons (SK-N-AS), glial cells (U-118 MG) and embryonic kidney cells as a non-central nervous system cell control (HEK293). In order to detect concentration-specific effects, three different concentrations of treatment (LiCl) or vehicle (NaCl) were used: 0.5 mM, 1.0 mM, and 2.0 mM – the values represent lower and higher ends of the therapeutic concentrations of lithium used clinically. SYN2a demonstrated a significant 33% increase in expression when treated with LiCl compared to vehicle at both of the two higher treatment concentrations: 1.0 mM and 2.0 mM ($P = 0.001$ and 0.035, respectively) in the neuronal cell line (Figure 2). A similar data set was collected for the *SYN2b* variant, but this had no significant change in expression in any of the conditions tested (Figure 3), suggesting that our findings are specific to *SYN2a*.

Discussion

Synapsin II is a candidate gene that was originally identified through a linkage study of Li-responsive BD families. This gene was also shown to be dysregulated in the post-mortem brains of patients with BD as compared to psychiatrically healthy controls in the same study (16). Thus, we were interested to investigate the effect of Li treatment on the expression of this gene. We did so in the genetic context of the disorder by treating with Li monotherapy Epstein-Barr virus-transformed lymphoblastoid cell lines from BD patients characterized as excellent Li-responders or non-responders, as well as healthy controls with no history of psychiatric disorders. We found that the

pattern of expression was significantly different in Li-responders compared to both non-responder BD patients as well as controls. However, the direction of change of expression was not uniform across subjects (Figure 1), resulting in no overall mean differences between groups. These data suggest that Li modulates *SYN2* expression in a way that is specific to Li-responders, possibly reflecting significant genetic heterogeneity.

The relevance of *SYN2* expression in peripheral cells compared to the central nervous system in BD patients is not clear from our findings, particularly since we saw no mean differences between Li-responders, non-responders, and controls. It is however clear from the literature that the gene is expressed, though at more basal levels, in lymphoblasts as well as many other cell types. Despite their peripheral origin, studying transformed LCLs offers the benefit of performing *in vitro* assays on cells from patients and studying putative factors in their endogenous expression context. However, results from these experiments should be considered with a level of scepticism, as the relevance of *SYN2* expression in this cell type is unclear.

Environmental factors could be involved in Li's regulatory role, which might account for the observed patient-specific effects in Li-responders. To investigate this possibility we computed correlations with a number of environmental factors relating to age of patients, Li therapy, and family history of other psychiatric disorders (for a complete list, refer to Table 3). However, none of the potential covariates correlated with *SYN2a* or *SYN2b* expression values, suggesting that the source of variation may be related to genetic or possibly epigenetic differences between patients. For example, variants in CREB genes (18) or GSK3-β (222), have been shown to associate with Li-treatment response. Similarly, it is possible that epigenetic factors may increase *SYN2* expression variance among patients. Though this is of interest, to our knowledge, no studies have investigated the role of Li treatment on epigenetic modifications in the human brain. However, valproate, another widely used

mood stabilizer, is well known for its inhibitory effect on histone deacetylases (HDACs) (223, 224) and therefore, it is possible that at least part of Li's action may be related to epigenetic regulation. Another epigenetic regulatory level where lithium's effect could be confounded is microRNA-mediated regulation. Studies in LCLs (46) and animal models (45) have shown the drug's global effect on this class of molecules. For a variety of biological reasons, each patient's LCLs could be enriched in a combination of regulatory factors which could then impact the response to Li treatment.

Since our LCL results do not automatically represent what is occurring in the brain, we sought to determine if Li would have a cell-type-specific effect on *SYN2* expression in model cell lines representative of the brain, and showed a significant change in the neuronal cell line SK-N-AS only (Figures 2 and 3). There was an effect at 1.0 and 2.0 mM Li, but not at 0.5 mM, suggesting that this concentration was not high enough to elicit a response. Interestingly, the effect was specific to the *SYN2a* variant (Figure 2), as the *SYN2b* variant remained unchanged between conditions (Figure 3). Originally, *SYN2* had been believed to display neuron-specific expression in the brain; however, further studies demonstrated the gene's expression in other cell types, though at considerably lower concentrations (225, 226). *SYN2* is expressed at basal levels in various cell types and thus lithium likely modulates its expression to a certain degree in these cells but perhaps not in a functionally-relevant manner. This is consistent with the fact that synapsins are evolutionarily conserved from humans to very primitive organisms and likely their expression has become more specialized in higher organisms through a loss of the ability to regulate other cellular functions but not necessarily through a complete loss of expression (202).

According to our results, in neurons, Li treatment significantly increases *SYN2* expression perhaps by also recruiting other neuron-specific transcription factors that bind to the gene's promoter such as EGR1 (early growth response 1), which has been suggested to regulate the gene (227), or AP-

2alpha, which has been shown to be regulated by lithium (228). Our results from LCLs are seemingly contradictory, as Li has an up-regulating effect on SYN2 in some patients, and a down-regulating effect in others. To interpret these results, one needs to consider that lithium acts as a mood stabilizer in patients who present both manic and depressive episodes. These clinical episodes are characterized by symptoms that are on opposite sides of the mood spectrum. Accordingly, manic patients present mood and neurovegetative activation, while depressed patients are characterized by a decreased mood levels and neurovegetative inhibition. Therefore, in order to be an effective mood stabilizer, Li needs to act by normalizing variance.

One interesting addition to this study would have been direct evidence for the effect of Li on *SYN2* expression in the central nervous system of BD patients. An ideal study would investigate the expression of *SYN2* variants in the post-mortem brains of BD patients who had been excellent responders to prophylactic Li for an extended period of time, so as to match the criteria used for our LCL samples. However, post-mortem brain donors with a history of BD are most often suicide completers. The literature provides extensive evidence for the anti-suicidal effects of Li prophylaxis through observational studies (229, 230), randomized controlled studies (231, 232) and meta-analyses (233, 234). Thus, such a study would be logistically quite challenging.

Another limitation of our study is the lack of protein-level evidence to support our mRNA-level findings. Such validation would be interesting in the pursuit of qualifying *SYN2* as a factor of potential pharmacological significance. However, the results presented here mainly point to *SYN2* as a new mediator of Li action. Perhaps by further investigating how *SYN2* is regulated we will also elucidate lithium's mode of action. There are likely several regulatory levels at play and clarifying them will be instrumental for our understanding of lithium response in BD, but as it stands the pharmacological application of this work is preliminary.

In conclusion, this is, to our knowledge, the first study attempting to determine the effect of Li treatment on mRNA-level expression of *SYN2*. We found a responder-specific effect of Li in LCLs from BD patients, suggesting that even though the gene is important for BD in general, there are genetic or epigenetic differences in Li responders that make them more susceptible to modulation of *SYN2*. Additionally, we showed that the effect of long-term treatment with Li is likely cell-type specific. As far as brain expression, our data suggest that the effect of lithium treatment is only significant in neuronal cells and not in astrocytic or glial cells. Support from additional cell types would be important to strengthen the validity of these conclusions. Our distinct findings for the two *SYN2* variants as well as the reported homology in sequence and function of the family of synapsin genes opens up the question of whether the other synapsins have a neuron-specific effect, as well as a patient-specific effect. Our study points to a very interesting player in response to Li prophylaxis, but more studies are required to decipher the full pathway of Li action that leads to its stabilizing effect in a large fraction of BD patients.

Acknowledgements

This work was supported by the Canadian Institute of Health Research (CIHR) grant 64410. We would like to thank Karine Lachapelle from the lab of GAR for technical assistance with the lymphoblastoid cell lines.

Tables

Table 1: **Lymphoblastoid cell line sample group demographics**. Data are presented as mean±SEM for non-psychiatric controls, bipolar disorder patients who are excellent lithium responders ("Responders") and bipolar disorder patients who do not respond to lithium treatment ("Non-Responders"). "Age at sampling" refers to the subject's age at the time blood was drawn. "LCL frozen storage" refers to the length of time of liquid nitrogen storage after Epstein-Barr-Virus transformation. "Age at onset" refers to the age at which patients were diagnosed with BD.

	Controls (C)	**Responders (R)**	**Non-Responders (N)**	**Group differences (p≤0.05)**
Subjects (M/F)	13 (3/10)	11 (5/6)	12 (3/9)	Not Significant
Age at DNA sampling (yr)	31±4.7	53.5±4.3	47.9±3.8	C vs. R and C vs. N
LCL frozen storage (yr)	3.7±0.3	7.8±1.0	6.8±0.5	C vs. R and C vs. N
Age at onset (yr)	n/a	32.6±3.5	29.8±3.5	Not Significant

Table 2: **Lithium response in lymphoblastoid cell line samples**. ANCOVA analysis was performed to compare the three groups (Controls, Responders, and Non-responders to lithium treatment) separately for *SYN2a* and *SYN2b* expression. The variables "Age at sampling" and "LCL frozen storage" were used as covariates.

	SYN2a			*SYN2b*		
	Control / Non-Responder	**Control / Responder**	**Non-Responder / Responder**	**Control / Non-Responder**	**Control / Responder**	**Non-Responder / Responder**
ANCOVA p-value	0.867			0.916		
Tukey's Test	0.231	1.315	1.112	0.108	0.993	1.123
F-test	0.897	0.001**	0.001**	0.839	0.0009***	0.0008***

Table 3: **Correlations of covariates with RQ expression values in excellent lithium responders**. To try and explain the distribution abnormal of Syn2 expression in Li-responders we computed correlations between RQ values and 15 potential covariates relating to age at sampling, onset, treatment start, etc., lithium treatment, as well as family history of other psychiatric disorders. (No samples had any family history of schizophrenia.) Normality of distribution was determined using the Shapiro-Wilk normality test and correlations were determined using Pearson's or Spearman's tests accordingly. No significant correlations were found with any of these variables.

	Shapiro- Wilk Normality		*SYN2a RQ*			*SYN2b RQ*		
	p-value	**Normal distrib.**	**Pearson coefficient**	**Spearman coefficient**	**p-value**	**Pearson coefficient**	**Spearman coefficient**	**p-value**
LCL frozen storage (yr)	0.2207	Yes	0.211		0.533	0.237		0.482
Age at DNA sampling (yr)	0.8805	Yes	-0.070		0.838	-0.065		0.849
Age at Onset (yr)	0.6388	Yes	0.217		0.521	0.158		0.644
Age at first treatment Li (yr)	0.3542	Yes	-0.404		0.320	-0.449		0.264
Time b/w onset and DNA collection	0.2446	Yes	-0.291		0.385	-0.227		0.502
Li Treatment response Score	0.2172	Yes	-0.129		0.705	-0.185		0.585
Episodes before Li	0.0114	No		-0.527	0.145		-0.527	0.145
Time on Li treatment (yr)	0.0456	No		0.477	0.194		0.477	0.194
Li dose at DNA sampling	0.5553	Yes	-0.437		0.239	-0.432		0.246
Number of other psych drugs	0.169	Yes	0.277		0.470	0.345		0.364
Family History Depression	0.0012	No		0.015	0.965		0.015	0.965
Family History Bipolar Disorder	0.0085	No		-0.193	0.569		-0.193	0.569
Family History Schizophrenia								
Family History Anxiety	< 0.0001	No		0.100	0.770		0.100	0.770
Family History Alcoholism	0.0004	No		0.438	0.178		0.438	0.178

Figures

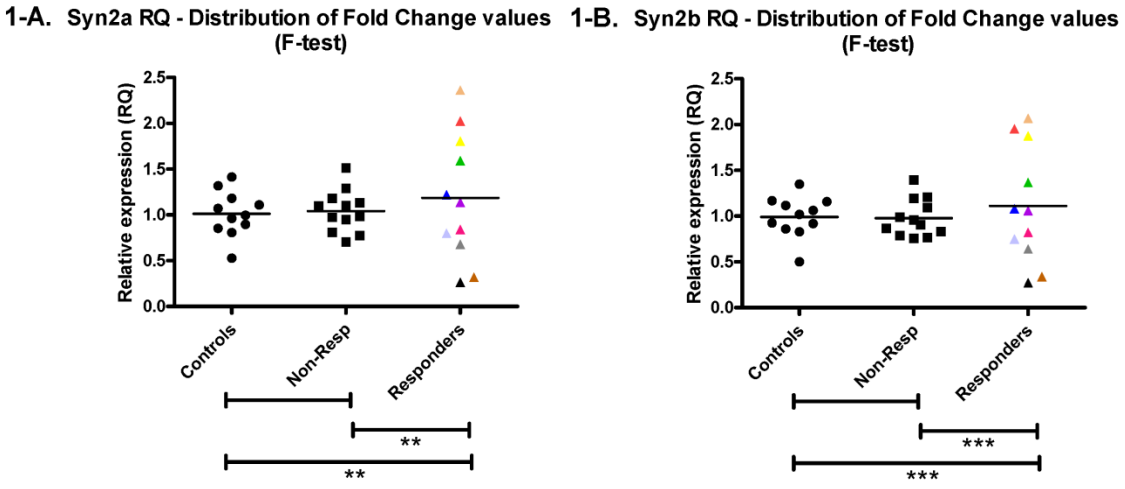


Figure 1: **Lymphoblastoid cell line expression**. Relative Quantification (RQ) values from qRT-PCR relative to *GAPDH* as an endogenous control. The groups compared are non-psychiatric controls, bipolar disorder patients without positive response to lithium (Non-Resp) and bipolar disorder patients with excellent response to lithium. The expression analyses were performed with separate primer sets for *SYN2a* (left) and *SYN2b* (right). The asterisks refer to F-test p-values depicting the differences in distribution between the individual expression changes in each group (** p-value≤0.001; *** value≤ 0.0001). There were no significant mean group differences, as indicated in Table 2.

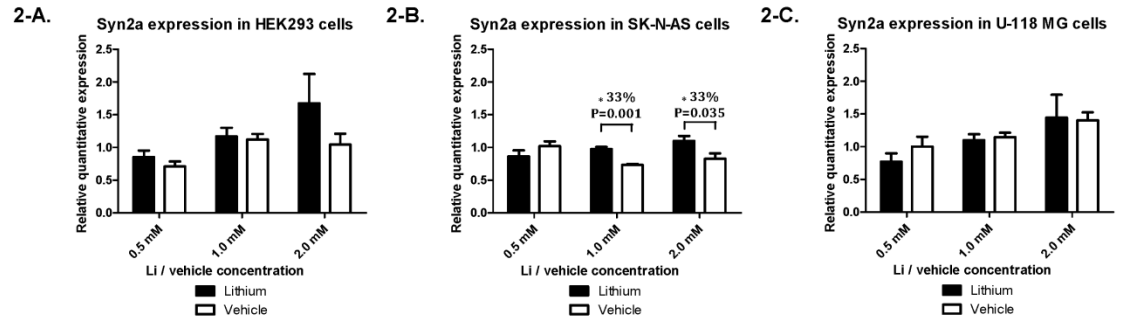


Figure 2: **Cell lines expression for SYN2a**. Expression in (A) HEK293 embryonic kidney cells, (B) SK-N-AS neuroblastoma cells, and (C) and U-118 MG glioblastoma/astrocytoma cells for the Synapsin IIa variant compared to *GAPDH*. P-values depicting the mean differences between 3 independent experiments for each cell line at each of the 3 treatment concentration of either lithium or vehicle (0.5mM, 1.0mM, and 2.0mM).

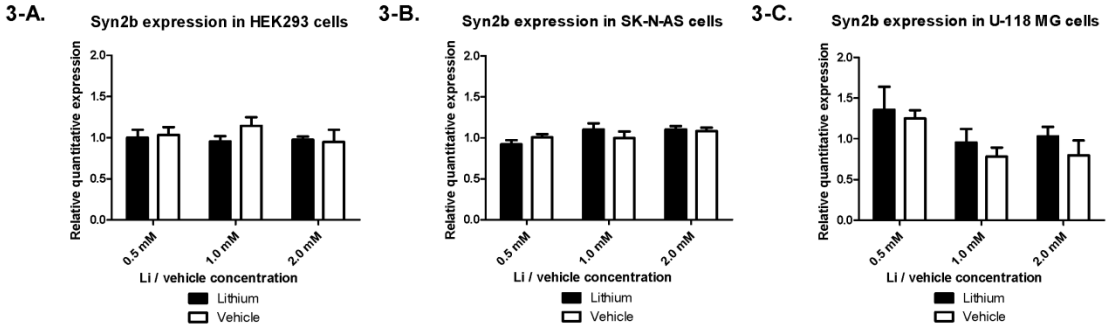


Figure 3: **Cell lines expression for SYN2b**. Expression in (A) HEK293 embryonic kidney cells, (B) SK-N-AS neuroblastoma cells, and (C) and U-118 MG glioblastoma/astrocytoma cells for the Synapsin IIb variant compared to *GAPDH*. P-values depicting the mean differences between 3 independent experiments for each cell line at each of the 3 treatment concentration of either lithium or vehicle (0.5mM, 1.0mM, and 2.0mM).

Part 3.3: H3K4 tri-methylation in synapsin genes leads to different expression patterns in bipolar disorder and major depression

Title: H3K4 tri-methylation in synapsin genes leads to different expression patterns in bipolar disorder and major depression

Authors: Cristiana Cruceanu[1], Martin Alda[2], Corina Nagy[1], Guy A. Rouleau[3], Gustavo Turecki[1#]

[1] McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, QC, Canada.

[2] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada.

[3] Centre of Excellence in Neuromics, CHUM Research Center and the Department of Medicine, University of Montreal, Montreal, QC, Canada

Corresponding Author:
McGill Group for Suicide Studies
Douglas Mental Health University Institute
Frank B Common Pavilion
Room F-3125
6875 LaSalle Boulevard
Montreal (Quebec) H4H 1R3
Email: gustavo.turecki@mcgill.ca
Phone: 514-761-6131 ext: 2369
Fax: 514 762-3023

Short Title: Epigenetic regulation of synapsins in BD and MDD

Category of manuscript submission: Regular Research Article

Statistical Summary
Word Count (abstract): 184
(body): 3355
Number of references: 66
Number of Figures: 4
Number of Tables: 5

Abstract

The synapsin family of neuronal phosphoproteins is composed of three genes (*SYN1*, *SYN2*, and *SYN3*) with alternative splicing resulting in a number of variants with various levels of homology. These genes have been postulated to play significant roles in several neuropsychiatric disorders, including bipolar disorder (BD), schizophrenia and epilepsy. Epigenetic regulatory mechanisms, such as histone modifications in gene regulatory regions, have also been proposed to play a role in a number of psychiatric disorders, including bipolar disorder and major depressive disorder. One of the best characterized histone modifications is histone 3 lysine 4 tri-methylation (H3K4me3), an epigenetic mark shown to be highly enriched at transcriptional start sites and associated with active transcription. In the present study we have quantified the expression of transcript variants of the three synapsin genes and investigated their relationship to H3K4me3 promoter enrichment in post-mortem brain samples. We found that histone modification marks were significantly increased in bipolar disorder and major depression, and this effect was correlated with significant increases in gene expression. Our findings suggest that synapsin dysregulation in mood disorders is mediated in part by epigenetic regulatory mechanisms.

Key Words: Bipolar disorder, synapsin, gene expression, epigenetics, H3K4me3

Introduction

The synapsin family of neuronal phosphoproteins is composed of three genes (synapsins I, II, and III) with alternative splicing giving rise to 10 reported variants expressed at various developmental time points and in various cell types (235, 236). The genes are involved in synaptogenesis, synaptic transmission, and synaptic plasticity (237). Of the three synapsin genes, synapsin I (*SYN1*) and synapsin II (*SYN2*) are predominantly expressed by mature neurons, where they have been shown to associate with the cytoplasmic surface of synaptic vesicles and to represent over 6% of their protein content (202, 238-241). *SYN1* maps to chromosome Xp11.23 and has two known variants, Ia and Ib (201, 202), and *SYN2* maps to chromosome 3p25 and has two known variants, IIa and IIb (202). Both *SYN1* and *SYN2* are differentially expressed in nerve terminals in the majority of the adult brain with demonstrated homology across numerous vertebrate and invertebrate organisms (201, 202, 238). Synapsin III (*SYN3*) maps to chromosome 22q12.3 and has been shown to produce up to 6 variants, though not all are expressed in the adult brain (235, 238). Its expression is much lower than that of synapsins I or II (242). The full-length synapsin III protein (isoform IIIa) exhibits protein homology with the other two synapsins and consequently possible functional homology as well, while the other variants have been shown to have developmentally specific-expression and the majority to be limited to foetal neuron expression (235, 238). The only other *SYN3* variant that shows adult expression in the human brain is *SYN3g*. The function of the *SYN3* variants is not as well understood as that of *SYN1* or *SYN2*, but it has been suggested to be mainly localized to regions outside of the synapse in the adult brain and function in neurogenesis and synaptic plasticity (237). The majority of brain regions jointly express synapsin variants at similar levels, suggesting that they are functionally complementary (203), however deleting each of the three synapsin genes produces different phenotypes, indicating that the various gene products must differ in their function to some degree (211, 212, 243, 244).

Synapsin genes have been proposed to play roles in several psychiatric disorders such as schizophrenia, bipolar disorder (BD) and epilepsy (202, 245) in both genetic (214-216) and functional studies (3, 218-220, 246-248) Given the evidence suggesting differential expression of synapsin genes in association with psychiatric phenotypes, it is interesting to study potential regulatory mechanisms that may underlie these changes. In this study, we set out to investigate epigenetic mechanisms, specifically the role of histone modifications, in explaining differential synapsin expression in bipolar disorder.

Epigenetic modifications have been investigated in various psychiatric phenotypes, including schizophrenia (249, 250), autism (238), major depression (248) and suicide (250). Interestingly, valproate, one of the most commonly used mood stabilizers in bipolar disorder, is an inhibitor of histone deacetylases (251-254), and thus it is possible that its stabilizing role in the disorder is mediated through inhibition of histone deacetylases. One of the best understood epigenetic mechanisms is histone methylation, particularly the tri-methylation of the 4^{th} lysine tail on histone 3 (H3K4me3) (255). This modification has been shown to be most abundant at transcriptional start sites (TSS) of genes and has been associated with increased transcription (256-258). H3K4me3 functions by opening up the chromatin and allowing transcriptional machinery to bind to the promoter region of genes, thus leading to the initiation of transcription. Enrichment of this mark typically leads to an increase in expression levels (258-260).

In this study, we analyzed expression of *SYN1a*, *SYN1b*, *SYN2a*, *SYN2b*, *SYN3a*, and *SYN3g* in post-mortem brains from BD patients, focusing on Brodmann Area 10 (BA10) of the prefrontal cortex (PFC). Our choice to focus on the PFC was based on studies showing its importance in mood regulation as well as documented deficits in PFC-mediated working memory and executive function in BD patients (261-263). In addition, imaging studies have shown abnormalities in PFC biochemistry

and function in BD patients during manic and depressive episodes, as well as during euthymia, suggesting the possibility of persistent neuropsychological deficits in BD (264, 265). Furthermore, the mediofrontal cortex has been previously linked to mood regulation in bipolar disorder. A study comparing BD patients with their at-risk but healthy siblings showed rCBF decreases in this region (BA9/10) in patients but an increase in their siblings, suggesting that this brain region may be involved in BD (266).

Since BD is characterized by alternating episodes of depression and mania, and the BD subjects investigated in this study died by suicide during a depressive episode, we included a comparison group of subjects with major depressive disorder (MDD) in order to control for possible effects that may be associated with depressive symptomatology. We compared both groups with a group of matched psychiatrically healthy controls.

Methods and Materials

I. Subjects

Post-mortem prefrontal cortex brain tissue from Brodmann Area 10 (BA10) used in this study was obtained from the Quebec Suicide Brain Bank (QSBB) (QSBB; www.douglasrecherche.qc.ca/suicide) as described elsewhere (3, 267). Clinical information, toxicology and history of psychoactive prescription drugs were collected for both cases and controls. These data were found to have no influence on our results; a detailed discussion is presented in Supplementary materials. All procedures in this study were approved by the ethics review board of our institution. Cases in this study were individuals who had a diagnosis of BD type I or type II (N = 13) or MDD (N=18) and died by suicide. Controls were individuals who died suddenly, and could not have undergone any resuscitation

procedures or other type of medical intervention (N = 14). Controls had neither current nor past psychiatric diagnoses. There were no significant group differences in gender, age, post-mortem delay, pH, and RNA integrity numbers (Table 1). We chose to focus on BA10 as a representative prefrontal cortex region and extracted total RNA from post-mortem brains.

II. Gene expression.

Total messenger RNA (mRNA) was extracted from frozen brain tissue using the RNeasy Lipid Tissue Mini Kit (Qiagen). For synthesis of complementary DNA (cDNA), M-MLV reverse transcriptase (Gibco, Burlington, Ontario) and oligo(dT)16 primers (Invitrogen) were used.

III. Chromatin Immunoprecipitation (ChIP).

DNA for chromatin immunoprecipitation was prepared from BA10 of post-mortem brain tissues (regions adjacent to those selected for mRNA experiments) as described by Matevossian and Akbarian (268). Briefly, 80 mg of tissue was cleaved between adjacent nucleosomes with micrococcal nuclease (Sigma Aldrich). A portion of selected intact nucleosomes was treated with anti-H3K4me3 antibody (Millipore) and purified with protein G agarose beads (Millipore). The remainder was used as input control. Both input and bound fractions were digested with proteinase K before purifying DNA by phenol/chloroform extraction (268, 269).

IV. Quantitative Real-Time PCR (qRT-PCR).

Samples were run on the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) in quadruplicate using standard qRT-PCR conditions and the TaqMan Fast Master Mix or the Power SYBR® Green PCR Master Mix (Applied Biosystems) as applicable. Relative expression for both mRNA and ChIP was calculated using the relative quantitation method ($\Delta\Delta$Ct) with GAPDH as an endogenous control in the RQ Manager 1.2 software. TaqMan assays were used for gene expression (Applied Biosystems). Expression values are presented as RQ (relative quantification) values

throughout the manuscript and they represent $2^{-\Delta\Delta Ct}$ metrics in reference to a pooled calibrator sample. For ChIP quantification, ratios of bound/input fractions were calculated for each sample by using custom SYBR Green primers designed (IDT) in the promoter region ~ 500bp upstream of the transcription start site. Primer sequences are available upon request.

V. Data analysis.

Test coefficients and probability distributions were calculated using statistical software GraphPad Prism 5 and SPSS. Before any other statistical computation or graphical representation of results, outlier analyses were performed for each dataset. For this reason, select subjects may be missing from analyses on a case-by-case basis. For qPCR experiments, relative quantitation was performed with *GAPDH* as an endogenous control in the RQ Manager 1.2 software (Applied Biosystems).

Results

Synapsin I and Synapsin II have different expression profiles in BD and MDD

Demographic and postmortem characteristics of the subjects included in the post-mortem expression study are reported in Table 1. As there were no significant differences between groups in these variables, we performed one-way ANOVA analyses followed by Tukey's post-hoc tests to assess the differences in expression between subjects with BD and controls as well as MDD and controls for the six synapsin variants (*SYN1a*, *SYN1b*, *SYN2a*, *SYN2b*, *SYN3a*, and *SYN3g*) that are expressed in the adult human brain and that are structurally and functionally similar (202). As shown in Figure 1.a and 1.b, the *SYN1a* variant was differentially upreglated in both BD and MDD (ANOVA P-value = 0.0045), while the *SYN1b* variant was only significantly upregulated in MDD (ANOVA P-value = 0.0172). These results suggest distinct patterns between the two *SYN1* variants.

The result for *SYN2* showed opposing expression patterns for the two variants. As shown in Figure 1.c and 1.d, *SYN2a* was significantly upregulated in BD with no effect in MDD (ANOVA P-value = 0.0001), while the converse was true for *SYN2b* (ANOVA P-value = 0.0005). Considering that gene expression changes in the brain are usually subtle, we note that these significant results were accompanied by fairly high fold changes of 2.47 and 1.80 respectively. Furthermore, these differences between BD and MDD are highly significant when comparing the groups to one another. For *SYN2a*, the BD group has an average RQ expression value 2.81 times higher than the MDD group, while for *SYN2b* the MDD groups has an average RQ expression value 2.23 times higher than the BD group (refer to Table 2 for significance coefficients).

For *SYN3* we only detected *SYN3a* and *SYN3g* at quantifiable levels in our brain samples. The two variants have perfect homology in regards to their coding exons, though at the mRNA level *SYN3g* expresses an additional exon at the 5'end. However, we did not detect differential expression in either the *SYN3a* variant (Fig 2.e) or the *SYN3g* variant (Fig 2.f) (ANOVA P-value = 0.2121 and 0.1551 respectively).

Synapsin II expression is modulated by H3K4me3 enrichment at the promoter region distinctly for BD and MDD

Given that expression of synapsin variants was increased in BD cases, we chose to investigate whether these changes were epigenetically regulated. We investigated levels of tri-methylation of the 4th lysine tail of histone 3 (H3K4me3) using chromatin immunoprecipitation and designed primers for each independent promoter in the first 500bp upstream of the TSS, since H3K4me3 has been shown to be enriched in this region. The *SYN1a* and *SYN1b* variants share a promoter (Supplemental Figure 1). As shown in Figure 2.a, this promoter was highly enriched in the MDD group with no change in the

BD group (ANOVA P-value = 0.005). There was also a significant difference when comparing the BD and MDD groups to one another, with a fold change of 3.22 (see Table 3 for significance coefficients). However, when following up this analysis with a Pearson's correlation between expression and H3K4me3 enrichment RQ values (Figure 3.a and 3.b), we found no significant effect. For simplicity all three diagnostic groups were included in this analysis since the expression patterns were very similar for the *SYN1a* and *SYN1b* variants, however we found that separate analyses by diagnostic status (as for *SYN2* below) yield the same non-significant correlation results (data not shown).

The *SYN2a* and the *SYN2b* variants also share a promoter (Supplemental Figure 1), which was significantly highly enriched in the H3K4me3 modification (ANOVA P-value = 0.0187) as shown in Figure 2.b. Only the BD group, though, was significantly different from controls in the Tukey's post-hoc test. Given the divergent expression of variants in the two disorders, Pearson's correlations were computed on the groups that had significantly different gene expression effects – BD-CTRL for *SYN2a* and MDD-CTRL for *SYN2b* – and these correlations were highly significant (Figures 3.c and 3.d).

Discussion

In this study we investigated expression patterns of synapsin variants and possible epigenetic regulatory mechanisms in the prefrontal cortex (BA10) of post-mortem brains from patients with BD, as well as MDD and controls with no psychiatric history. We focused on the PFC because of its involvement in mood regulation, working memory and executive function (261-263). Overall, we found that synapsins Ia and IIa were up-regulated in the BD brain samples. The most striking gene expression finding was for *SYN2*, where the gene was over-expressed in BD compared to controls, and

this effect was accounted for by the longer variant, *SYN2a*. The converse was found in post-mortem brains from patients with MDD, where we saw a significant up-regulation of the *SYN2b* variant, but no change for *SYN2a*. This expression difference between the two disorders may not be etiologically relevant, considering evidence that synapsin variants have overlapping function in the brain (202). However, when looking to identify functional individualities in various synapsin isoforms, Gitler *et al.* found a unique role for *SYN2a* during synaptic activity at glutamatergic synapses (270). This is of potential relevance, as alterations in glutamatergic transmission and plasticity have been indicated in BD (271-274). Furthermore, in a separate investigation of the effect of lithium treatment on synapsin expression in neuronal cell lines we found that this mood stabilizer classically used in BD treatment affected *SYN2a* but not *SYN2b* expression (163). Based on this evidence, our findings could reflect a subtle but distinct mechanism of regulation of the *SYN2* gene in the brains of patients with different mood disorders.

The second part of our study was to determine whether the observed up-regulation in gene expression was mediated through epigenetic modifications. To our knowledge, no previous studies have tried to identify histone modifications in the synapsin genes in relation to mood disorders, so we quantified H3K4me3 levels in the promoter regions of synapsin variants. Overall, we showed an increase in H3K4me3 levels at synapsin promoters in mood disorders, with patterns that are disease-specific. For the *SYN1* variants there was no significant correlation between mRNA expression and H3K4me3 enrichment. Though both the gene expression and the epigenetic findings for *SYN1a* and *SYN1b* are interesting, the two appear to be independent phenomena or part of a much more complex mechanism.

The most interesting epigenetic finding was the enrichment of H3K4me3 at the *SYN2* promoter. Unlike the *SYN1* data, the H3K4me3 enrichment in the *SYN2* promoter correlated with the expression

up-regulation shown for the individual variants on disease-specific lines. This finding suggests that gene expression of *SYN2a* in BD, and *SYN2b* in MDD, are regulated, at least in part, by changes in H3K4me3 levels at the *SYN2* promoter. H3K4me3 is a marker for open chromatin and subsequent enhanced expression, so once the chromatin has been opened, transcription levels are dependent on transcription factors binding. The promoter region where we detected H3K4me3 enrichment is between 176bp and 395bp upstream of the transcription start site. Our attempt to design primers in regions closer to the TSS did not yield quantifiable H3K4me3 levels. Interestingly, within this region there are two binding sites for the transcription factor AP-2α (adaptor-related protein complex 2, alpha 1 subunit). These sites were first identified by Petersohn *et al.* through DNA-protein binding assays *in vitro* (227) and the direct role of AP-2α in regulating SYN2 expression was validated through knock-down experiments in primary midbrain embryonic mouse neurons by Skoblenick *et al.* (275). The latter showed an increase in neuronal *SYN2* expression mediated through AP-2α following dopamine D1 receptor stimulation or dopamine D2 receptor inhibition (275). As dopamine dysfunction has been well characterized in both BD and MDD (276), AP-2α is a likely candidate for mediating the role of *SYN2* in these disorders. Furthermore, AP-2α has been shown to be regulated by lithium and carbamazepine (249, 277), two common mood stabilizer treatments used for BD, as well as by antidepressants like citalopram and imipramine (278).

Although the H3K4me3 findings are of interest, considering that the two *SYN2* variants share a promoter, the disease-specific expression cannot alone be explained by this epigenetic mechanism. Since the *SYN2* variants are only dissimilar at the 3' end, other regulatory mechanisms could explain the differential expression of these two *SYN2* transcripts in BD and MDD. One such mechanism could be microRNA regulation, a class of regulatory molecules that frequently act at 3' sites and have been shown to be dysregulated in bipolar disorder post-mortem brains (279-281).

As with all post-mortem brain studies, there are technical limitations to take into account, such as the relatively small sample size and the possible confounders associated with using frozen tissue for expression studies. To account for this we ensured that the three diagnostic groups had no significant differences in brain pH, post-mortem delay, as well as RNA integrity for expression studies (Table 1). Furthermore, as explained in greater detail in the Supplemental methods, we performed thorough post-mortem investigations on all subjects in an attempt to gather all the relevant medical history information as well as toxicology analyses at time of death. No significant effect of these potential covariates was identified in this study in regards to gene expression or epigenetic modifications (Supplemental methods).

Another limitation of this study is that we only investigated one epigenetic modification to try to explain our gene expression findings. It has been noted in the literature that epigenetic mechanisms seem to work in concert (259) and accordingly it is entirely possible that H3K4me3 enrichment is only one piece of the puzzle, particularly concerning the results for *SYN1* variants. The present study serves to demonstrate the involvement of epigenetic mechanisms in synapsin gene regulation in mood disorders, but it would be interesting to follow up our findings with a more in-depth look at various levels of epigenetic regulation not just in terms of histone modifications but also DNA methylation.

The main findings of this study are two-fold. Firstly, we showed distinct synapsin profiles for BD and MDD post-mortem brain mRNA expression. These findings are interesting because they potentially indicate a molecular marker for distinguishing the two clinically similar disorders. Secondly, we showed that for *SYN2* the changes in expression are correlated with enrichment of H3K4me3, an epigenetic mark associated with transcriptional activation. To our knowledge, this is the first study to identify an epigenetic mechanism to be involved in the regulation of this gene. As with any molecular studies of disease, independent replication in additional post-mortem sample sets is

extremely important to validate that the findings are truly relevant for the disorder and do not merely characterize the studied population. Future studies are warranted to understand the extent of epigenetic regulation of the *SYN2* gene in bipolar disorder, as well as the processes by which the *SYN2a* and *SYN2b* variants are distinctly expressed in the prefrontal cortex.

Acknowledgements

This work was supported by the Canadian Institute of Health Research (CIHR) grant 64410. The authors report no biomedical financial interests or potential conflicts of interest.

Figures

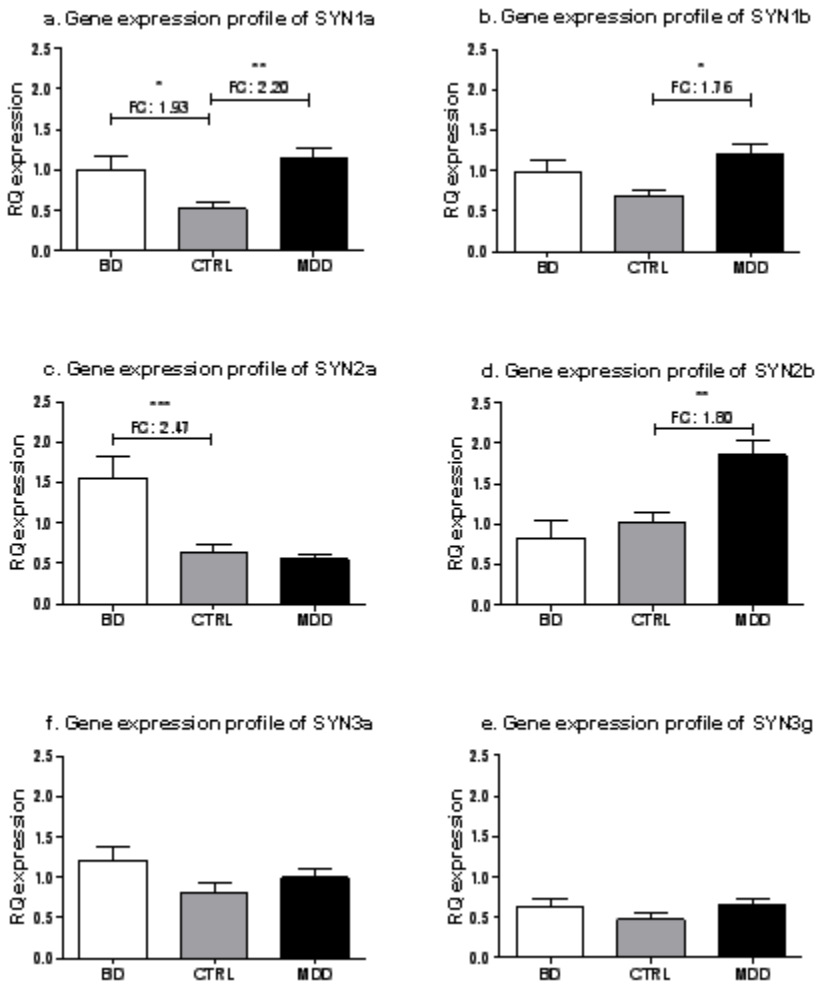


Figure 1: Brain expression RQ values from qRT-PCR relative to *GAPDH* as an endogenous control. Data are presented as RQ expression values, which represent $2^{-\Delta\Delta Ct}$ metrics. The groups compared are bipolar disorder (BD), major depressive disorder (MDD) and non-psychiatric controls (CTRL) using ANOVA analyses followed by Tukey's post-tests. A, Relative quantitative expression for variant *SYN1a* (ANOVA P-value = 0.0045; After outlier analysis BD n=13 , MDD n=15 , CTRL n=11); B, Relative quantitative expression for variant *SYN1b* (ANOVA P-value = 0.0172; After outlier analysis BD n=13, MDD n=15, CTRL n=11); C, Relative quantitative expression for variant *SYN2a* (ANOVA P-value = 0.0001; After outlier analysis BD n=13, MDD n=12, CTRL n=13); D, Relative quantitative expression for variant *SYN2b* (ANOVA P-value = 0.0005; After outlier analysis BD n=13, MDD n=15, CTRL n=11); E, Relative quantitative expression for variant *SYN3a* (ANOVA P-value = 0.2121; After outlier analysis BD n=12, MDD n=12, CTRL n=12) and F, Relative quantitative expression for variant *SYN3g* (ANOVA P-value = 0.1551; After outlier analysis BD n=12, MDD n=14, CTRL n=12).

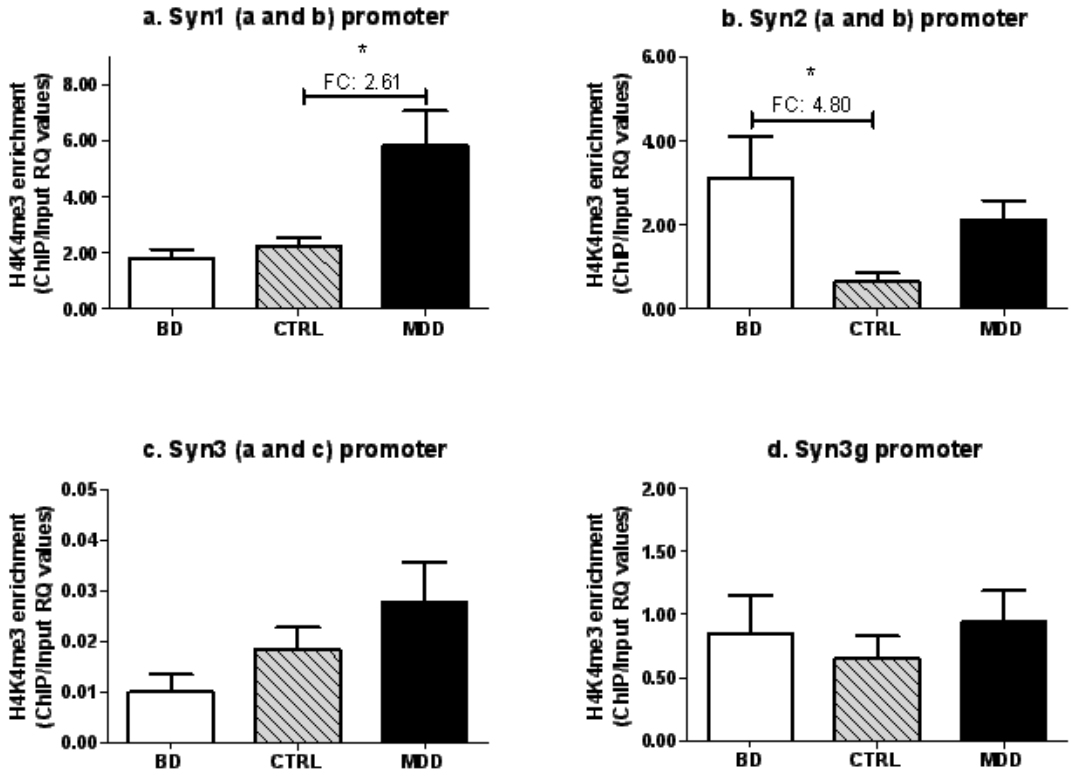


Figure 2: Histone 3 lysine 4 tri-methylation (H3K4me3) enrichment results for four different promoter regions representing the specific synapsin variants. Data are presented as RQ expression values, which represent $2^{-\Delta\Delta Ct}$ metrics. The groups compared are bipolar disorder (BD), major depressive disorder (MDD) and non-psychiatric controls (CTRL) using ANOVA analyses followed by Tukey's post-tests. A, H3K4me3 enrichment for the shared promoter of variants *SYN1a* and *SYN1b* (ANOVA P-value = 0.005; After outlier analysis BD n=12, MDD n=15, CTRL n=12); B, H3K4me3 enrichment for the shared promoter of variants *SYN2a* and *SYN2b* (ANOVA P-value = 0.0187; After outlier analysis BD n=9, MDD n=8, CTRL n=10).

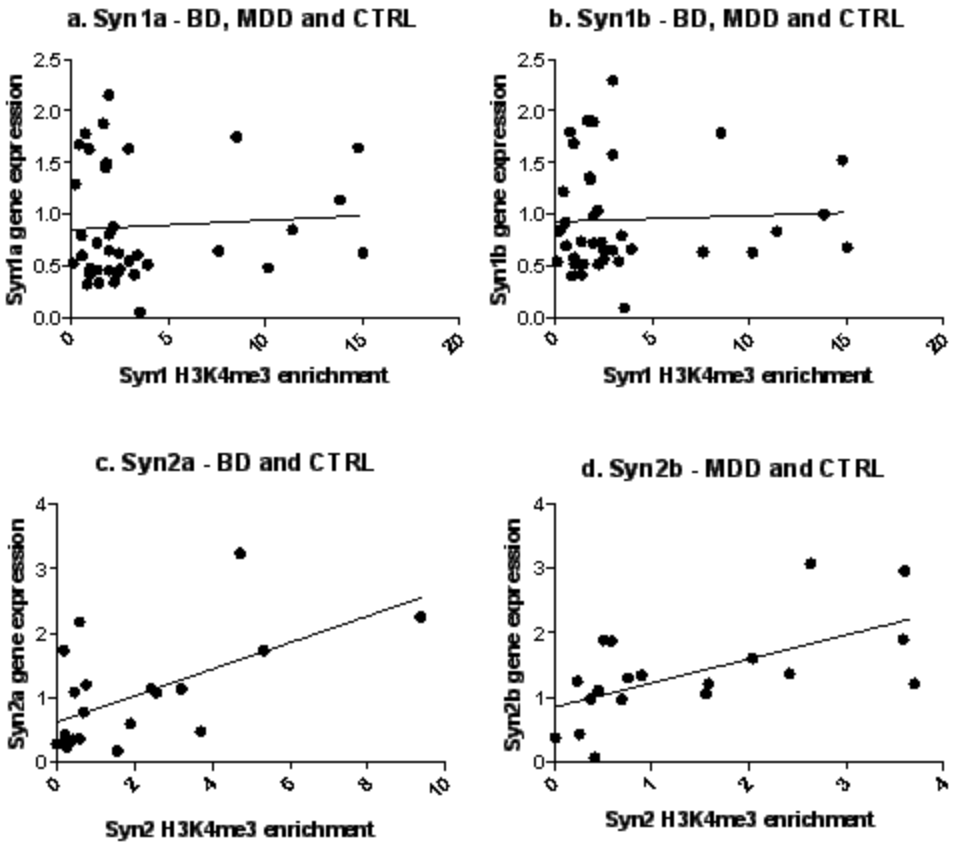


Figure 3: Pearson's correlations of gene expression RQ values versus H3K4me3 enrichment RQ values at the promoter region of the various synapsin variants. a) For the correlation between the *SYN1a* variant expression and the *SYN1* promoter H3K4me3 enrichment, the two-tailed P-value is 0.6833 (not significant); b) For the SYN1b variant the same correlation is also not significant, with a P-value of 0.7825. For the *SYN2* variants, since gene expression was so discrepant across diagnostic groups, with each variant showing an effect in a different disorder, correlations were computed accordingly. c) For the *SYN2a* variant, the correlation for the BD and CTRL groups had a two-tailed P-value of 0.0052 and d) for the *SYN2b* variant the correlation for the MDD and CTRL groups had a two-tailed P-value of 0.0054.

Tables

Table 1: Brain sample group demographics (presented as mean±SEM) for BD (bipolar disorder), MDD (major depressive disorder) and controls (CTRL). Group differences were computed using One-way ANOVA.

Status	Gender	Age	Brain pH	Post-Mortem Delay	RNA Integrity No.
BD	9 M / 4 F	44.00 ± 4.05	6.63 ± 0.07	30.38 ± 6.31	6,63 ± 0,30
MDD	11 M / 7 F	52.00 ± 3.81	6.72 ± 0.06	20.28 ± 4.32	6,34 ± 0,21
CTRL	12 M / 3 F	41.73 ± 6.04	6.56 ± 0.05	24.03 ± 4.62	6,48 ± 0,18
Group differences	ns	0,076	0,161	0,458	0,563

Table 2: Gene expression results. For each of the six synapsin variants, gene expression was quantified using qRT-PCR. Data was presented as RQ expression values, which represent $2^{-\Delta\Delta Ct}$ metrics One-way analyses of variance (ANOVA) and Tukey's post-hoc tests were computed for the three diagnostic groups: bipolar disorder (BD), control (CTRL) and major depressive disorder (MDD). P-values are presented along with significance levels (* p-val≤0.05, ** p-val≤0.001, ** p-val≤0.0001).

		BD vs CTRL	MDD vs CTRL	BD vs MDD
SYN1a	ANOVA p-value	0,0045**		
	Tukey's Multiple Comparison Test (q)	3,575	4,844	1,089
	Significance	*	**	ns
SYN1b	ANOVA p-value	0,0172*		
	Tukey's Multiple Comparison Test (q)	2,356	4,244	1,736
	Significance	ns	*	ns
SYN2a	ANOVA p-value	0,0001***		
	Tukey's Multiple Comparison Test (q)	5,661	0,487	6,037
	Significance	***	ns	***
SYN2b	ANOVA p-value	0,0005***		
	Tukey's Multiple Comparison Test (q)	1,049	4,624	5,626
	Significance	ns	**	***
SYN3a	ANOVA p-value	0,2121		
	Tukey's Multiple Comparison Test (q)	2,098	2,278	0,054
	Significance	ns	ns	ns
SYN3g	ANOVA p-value	0,1551		
	Tukey's Multiple Comparison Test (q)	2,798	1,329	1,567
	Significance	ns	ns	ns

Table 3: Histone 3 lysine 4 tri-methylation (H3K4me3) enrichment results. For each of the four different synapsin promoter regions, ChIP or Input enrichment was quantified using qRT-PCR. Here we report ChIP/Input ratios of RQ expression values, which represent $2^{-\Delta\Delta Ct}$ metrics. One-way analyses of variance (ANOVA) and Tukey's post-hoc tests were computed for the three diagnostic groups: bipolar disorder (BD), control (CTRL) and major depressive disorder (MDD). P-values are presented along with significance levels (* p-val≤0.05, ** p-val≤0.001, ** p-val≤0.0001).

		BD vs CTRL	MDD vs CTRL	BD vs MDD
SyYN1 a+b	ANOVA p-value	0,005**		
	Tukey's Multiple Comparison Test (q)	0,4217	4,046	4,211
	Significance	ns	*	*
SYN2 a+b	ANOVA p-value	0,0187*		
	Tukey's Multiple Comparison Test (q)	4,267	2,366	1,501
	Significance	*	ns	ns

Part 3.4: Epigenetic regulation of synapsin genes in mood disorders

Title: Epigenetic regulation of synapsin genes in mood disorders

Authors: Cristiana Cruceanu[1], Erika Freemantle[1], Martin Alda[2], Guy A. Rouleau[3], Gustavo Turecki[1#]

[1] McGill Group for Suicide Studies, Douglas Mental Health University Institute, McGill University, Montreal, QC, Canada.

[2] Department of Psychiatry, Dalhousie University, Halifax, NS, Canada.

[3] Centre of Excellence in Neuromics, CHUM Research Center and the Department of Medicine, University of Montreal, Montreal, QC, Canada

Corresponding Author:
McGill Group for Suicide Studies
Douglas Mental Health University Institute
Frank B Common Pavilion
Room F-3125
6875 LaSalle Boulevard
Montreal (Quebec) H4H 1R3
Email: gustavo.turecki@mcgill.ca
Phone: 514-761-6131 ext: 2369
Fax: 514 762-3023

Short Title: Synapsin epigenetics

Category of manuscript submission: Hot Topic

Statistical Summary

Word Count (Body): 636

(Figure legend): 189

Number of references: 9

Number of Figures: 1

The synapsins are a family of neuronal phosphoproteins consisting of *SYN1* at chrXp11.3, *SYN2* at chr3p25, and *SYN3* at chr22q12.3 with alternative splicing leading to as many as ten isoforms. They are involved in synaptic transmission and plasticity, as well as various stages of neurodevelopment including axon outgrowth and synapse formation (202). All synapsins are highly concentrated at presynaptic nerve terminals of central neurons and associated with the cytoplasmic surface of synaptic vesicles, but *SYN3* has markedly distinct developmental expression and subcellular distribution, suggesting divergent function (282). Therefore not surprisingly, a role for synapsins in neuropsychiatry has been suggested, and indeed, several studies have indicated that genetic variants at these genes can be associated with epilepsy, autism, schizophrenia, and bipolar disorder (BD) (202). Furthermore, mRNA- and protein-level post-mortem brain studies have suggested dysregulation of these genes in both BD and major depression (MDD) (175, 202). Thus, the study of mechanisms responsible for this dysregulation in mood disorders becomes pertinent.

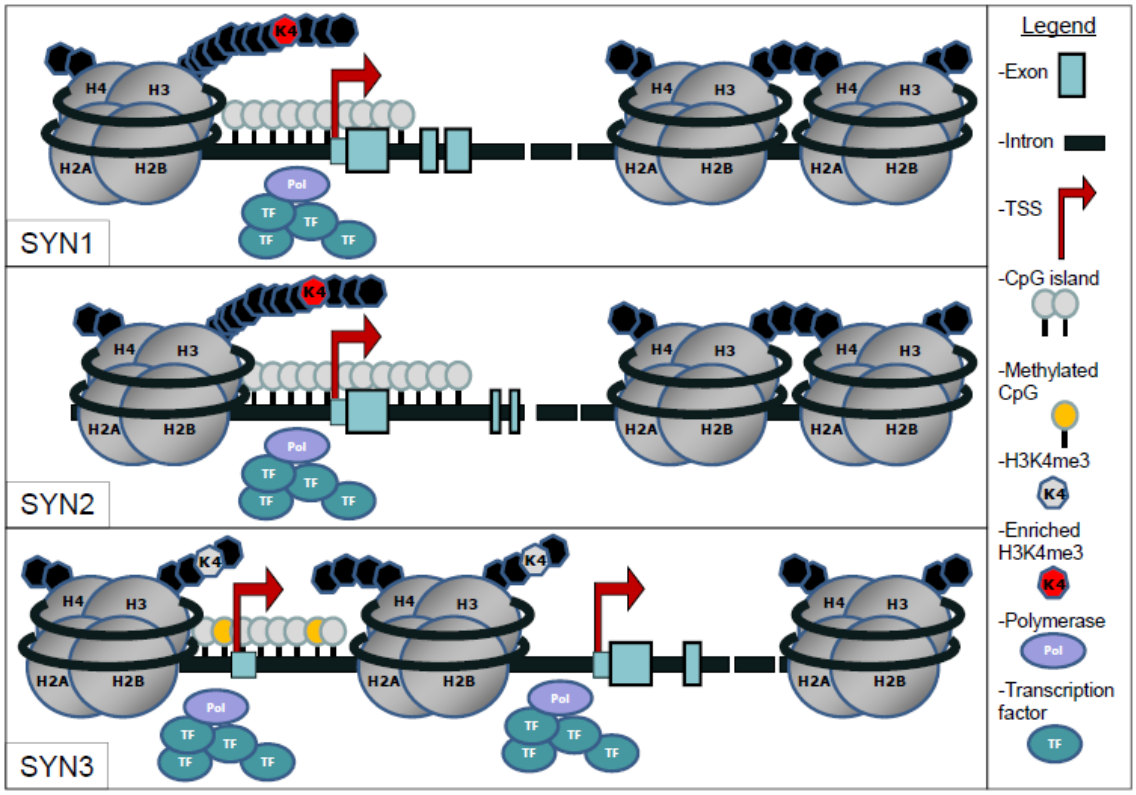
In the last few years, evidence has emerged suggesting that epigenetics play a role in neuropsychiatric disorders (283), thus it is plausible that the dysregulation observed in synapsin expression could be attributed in part to epigenetic mechanisms. We found evidence that enrichment of H3K4me3 – an epigenetic mark associated with increased transcription – at the promoters of *SYN1* and *SYN2* but not *SYN3* is correlated with increased expression of these genes in the prefrontal cortex of patients with BD and MDD compared to controls (175) – see Figure. These findings are encouraging, but future research should better characterize these mechanisms by exploring the role of other chromatin epigenetic marks and brain-region specificity. In addition, the role of DNA methylation, an equally important epigenetic mechanism, should be investigated. *In silico* analyses have detected rich CpG islands at the proximal promoters of *SYN1* (845bp) and *SYN2* (975bp), as well as at a distal promoter of *SYN3* (613bp) (284) – see Figure. To date there is no evidence in the

literature of altered DNA methylation at synapsins in mood disorders, though one study of a single schizophrenia patient suggests potentially variably methylated sites in the distal CpG island of *SYN3* (285). Interestingly, the CpG islands at *SYN1* and *SYN2* are immediately preceded by regions of enriched H3K4me3 in mood disorders (175) – see Figure. The same is not true for *SYN3*, and considering this gene's distinct expression profile and potential implication throughout neurogenesis (282), perhaps different mechanisms regulate *SYN3*. The Figure illustrates our current knowledge of the synapsin genes' structure as well as the epigenetic mechanisms that have been identified in psychiatric disorders to date.

In conclusion, brain expression differences seen in synapsin genes in mood disorders may be explained in part by differences in H3K4me3. These results need additional and independent confirmation. Moreover, considering that promoter DNA methylation can modulate gene expression and lead to neuropsychiatric phenotypes, a study of DNA methylation patterns at the synapsin promoters is warranted. Based on the growing evidence suggesting that epigenetic mechanisms may be involved in altered regulation of synapsins in mood disorders, it would be of interest to study these genes as potential therapeutic targets or biomarkers of treatment response. Evidence is starting to emerge pointing to epigenetic marks as potential biomarkers of treatment response. For instance, for brain-derived neurotrophic factor (*BDNF*), Lopez *et al.* showed that promoter H3K27me3 levels could serve as a biomarker of response to citalopram in MDD (286), and D'Addario *et al.* found distinct DNA methylation patterns at the *BDNF* promoter in BD patients depending on mood-stabilizer and antidepressant therapy (287). Though no such evidence has yet emerged for synapsins, a recent study showed that lithium, one of the most commonly prescribed drugs for BD, can modulate *SYN2* expression in neuronal cell type (163). Thus, an investigation of synapsin epigenetics in the brain

compared to the periphery would be an interesting next step in elucidating their potential to serve as biomarkers for mood disorders or their treatment.

Figure



Potential epigenetic mechanisms at the promoter regions of synapsin genes. **Upper panel**: The *SYN1* gene (chrX:47,431,300-47,479,256). *In silico* analysis predicts a CpG island 845bp in size at the 5'end of the gene (chrX:47,478,671-47,479,515) that spans from -259bp upstream of the transcription start site TSS) to +586bp downstream. Evidence from chromatin immunoprecipitation assays for H3K4-trimethylation suggests that this epigenetic mark is enriched in mood disorders around roughly -200bp to -350bp upstream of the TSS. **Middle panel**: The *SYN2* gene (chr3:12,045,862-12,233,532). The first 3 coding exons are represented here. *In silico* analysis predicts a CpG island of 975bp at the 5'end of the gene (chr3:12,045,653-12,046,627) that spans from -208bp upstream of the TSS to +767bp downstream. Evidence from chromatin immunoprecipitation assays for H3K4-trimethylation suggests that this epigenetic mark is enriched in mood disorders around roughly -175bp to -400bp upstream of the TSS. **Bottom panel**: The *SYN3* gene (chr22:32,908,540-33,402,809). There is no predicted CpG island at the proximal promoter, but at a distal promoter upstream an alternative non-coding first exon there is a CpG island 613bp in size. Certain CpGs within this island have been suggested to be variably methylated in schizophrenia.

Chapter 4: Discussion

Part 4.1: Discussion of results and implications

The focus of this work is on bipolar disorder (BD), a severe psychiatric illness with recurrent, often chronic clinical presentation and a strong genetic basis. It typically develops in genetically predisposed individuals in their late teens or early 20s and follows a recurrent or chronic lifelong course. It is associated with high morbidity and mortality due to suicide and other comorbidities, and poses a major burden to individuals and society. In fact it is one of top causes of morbidity and disability worldwide. Patients are often diagnosed after considerable delay and their treatment is commonly selected on a trial-and-error basis, as is the case for many other psychiatric conditions. It is known, however, that proper treatment can lead in many instances to full recovery and can minimize the impact of the illness on brain structure and function.

Given the impact that BD has on individuals and society, it represents an important area of research, and susceptibility factors for this disorder need to be resolved in order for improvement in detection and treatment outcome to follow. Thus, the work presented in this thesis, and completed as part of my PhD degree research component, aims at elucidating the susceptibility factors for BD using a combination of different approaches. The general hypothesis tested was that BD susceptibility arises from an interaction between the genetic predisposition conferred by relatively rare loci of moderate-to-large effect, in interaction with environmental effects mediated via transcriptomic and epigenetic changes. The combination of studies that make up this thesis represent a major contribution in the field as they shed important light on aspects of genetic and non-genetic susceptibility factors for BD as well as the neurobiological changes that happen in the brain as a result of BD.

The experimental design and findings are divided into two important and different approaches. Chapter 2 explores the first overarching aim, which was to investigate the genome-wide patterns of genetic and non-genetic susceptibility factors for BD. A major theme for this work was the finding of

a role for G Protein-Coupled Receptors (GPCRs) in BD. Chapter 3 follows a candidate-gene approach towards resolving aspects of BD susceptibility, and narrows in on the role of synaptic neurobiology in BD and treatment response. Importantly, this research is not presented according to the same timeline it was published, and thus the candidate gene studies in Chapter 3 do not directly follow the genome-wide findings in Chapter 2. Nonetheless, parallels can be drawn across the chapters and both contribute to the literature in important ways.

The first objective of this work, which is presented in Chapter 2, was to investigate genome-wide genetic and non-genetic BD susceptibility factors. Thus we used high-throughput next-generation sequencing approaches to investigate this on two different levels: the DNA through exome sequencing and the RNA through transcriptome sequencing. These investigations would not have been possible only a decade ago, before the revolutionary development of high-throughput next-generation sequencing technologies, which began to emerge after the publication of the complete human genome in April 2004 (62, 63). Since then, the rapid development of these technologies which allow nucleotide-level reading of nucleic acids, and the ensuing applications across many biological levels and fields, has led to impressive advances.

The high heritability of BD indicated a strong genetic component, and to elucidate this, a large part of the research in BD genetics to date has consisted of classical linkage or genome-wide association studies, thereby focusing on common susceptibility variants. Although common variants are informative, they explain only a small fraction of the predicted BD heritability, suggesting a considerable contribution comes from genetic variation of rare and private frequency in the population. In Chapter 2.2 we explored the contribution of inherited DNA mutations in families with increased loading of BD. The selection of families for this part was paramount as they belonged to a long-standing cohort that was thoroughly characterized for a well-defined clinical subtype of BD –

excellent response to lithium monotherapy. The importance of selecting families from this cohort, which our group has been following for several decades, was to allow us to limit the amount of genetic heterogeneity since a lot of susceptibility would be shared, and thus tease out inherited variants of low frequency in the population but moderate to high penetrance in the families.

We hoped that with the evidence from rare variation in each family we could paint a larger picture of the genes and pathways implicated in BD susceptibility and resolve some of the "missing heritability". Individuals were selected from 40 family units consisting of 3-7 affected individuals across 1-3 generations (N=244), and exome sequencing was performed to identify all variation within the protein-coding portion of the genome. We focused on rare coding variants that segregated with affected status in families, and what we found was indeed very interesting, namely an enrichment of putatively damaging mutations in G protein-coupled receptors (GPCRs) among individuals affected with BD. Our finding is relevant because members from this family of integral membrane proteins have been previously associated with BD (112, 142, 288, 289), and have been shown to be excellent targets for psychiatrically-relevant drugs (290). For example, two of the largest and best-characterized downstream signaling pathways of GPCRs, mediated by cAMP and phosphatidylinositol, have been implicated both in the pathophysiology of BD (112), as well as in the mechanism of action of common BD drugs (113, 139). Unfortunately we could not show an effect that was specific to lithium-response in these families, suggesting that the significant findings regarding GPCRs in our cohort were disease-related rather than treatment-response-related. However the involvement of GPCRs does not stop at drug mediation, and in fact many of the known GPCRs are essential components of signaling pathways which have been implicated in BD (142), including the monoaminergic and neuropeptidergic signal transmission systems (291, 292).

Additionally, in this study we further investigated the functionality of a nonsense mutation in the GPCR gene corticotropin releasing hormone receptor 2 (*CRHR2*). This mutation had a number of downstream effects on cellular function, and thus was likely to play a causal role in the disease in the affected family members. This is a receptor for the corticotropin releasing hormone (CRH) and Urocortins 1, 2,and 3 (293), and is involved in stress response through the hypothalamic-pituitary-adrenal axis, a pathway that has been extensively investigated in BD and other mood disorders (143). We showed that the truncated receptor, which lost 28 amino acids from the C-terminus, was translated and recruited to the plasma membrane, but with significantly reduced efficiency. Secondly, we showed increased ability to engage G_i G-protein alpha subunit members following ligand binding, higher constitutive activity for activation of G_s subunits, and a lower ligand-activated and constitutive activity for translocation via β-arrestin. The latter likely explained the decreased membrane translocation of the mutant, since 6 out of 8 putative phosphorylation sites at the C-terminal tail of *CRHR2* were lost in the mutant, and β-arrestin is known to interact with the phosphorylated C-terminal tail of GPCRs which is known to be important for receptor desensitization (146). The potential role for the *CRHR2* receptor in BD has been previously proposed by De Luca *et al*. who showed that haplotype variation at the *CRHR2* locus is associated with suicidal behavior in BD (294). Our study confidently positions this receptor as a major risk factor for BD and shows some of the downstream effects of a penetrant mutation at this locus.

In Chapter 2.3 we investigated the whole transcriptional dysregulation in BD using RNA sequencing in post-mortem brains obtained from the Douglas-Bell Canada Brain Bank. This is a wonderful resource, consisting of nearly 3000 human brains preserved under optimal conditions for research and thoroughly characterized for medical history including some of the most comprehensive post-mortem psychological autopsies in the world (295). This aspect is essential for molecular

psychiatry research, as accurate primary diagnoses and comorbidities are the heart of successful investigations.

Previous investigations characterizing gene expression changes and alterations in regulatory mechanisms in post-mortem brains included mainly candidate gene and microarray expression studies, which had several limitations relating to accuracy and sensitivity of available technology, leading to reduced replication across studies (88, 90, 92). Thus, questions remained about the many subtle gene expression changes in the brain as well as isoform-specific dysregulation of known genes and non-coding transcripts whose importance had been demonstrated recently in the brain but not characterized for BD. The goal of our study was to identify genes differentially expressed between cases and controls that could contribute to our understanding of the pathways and biological processes dysregulated in the brains of affected individuals. Furthermore, we aimed to characterize not just the coding transcriptome, but shed light on the non-coding transcriptome in hopes of reaching a greatly improved scope compared to previous studies. To this end, we performed ribosomal-depletion RNA sequencing, which removes the highly abundant ribosomal RNAs prior to sequencing library preparation, but otherwise unbiasedly leaves behind all other RNA species expressed in the cells sampled (296). Our resulting publication represents the first unbiased and comprehensive analysis of the entire transcriptome in BD, including all non-coding RNA species of approximately 200 base pairs or longer in length, such as long non-coding intergenic RNAs (lincRNAs).

In addition to reproducing findings for several previously-implicated genes, we identified many new candidate genes for BD, and were the first to identify non-coding RNAs as being involved in the pathology of BD. One of the major findings was a global down-regulation of gene expression in the anterior cingulate cortex, an important brain region for BD (166, 167). A pathway analysis also demonstrated an overrepresentation of genes involved in G protein-coupled receptor (GPCR)

regulation, which mirrored the findings from the mutation-discovery WES study in Chapter 2.2. Though our most interesting gene from the WES study, *CRHR2*, did not show dysregulation that passed thresholds for transcriptome-wide significance, other GPCRs equally interesting in the psychiatry field emerged through this analysis, including *SSTR2* (somatostatin receptor 2) (187, 189), *CHRM2* (cholinergic receptor, muscarinic 2) (184, 185), and *RXFP1* (relaxin/insulin-like family peptide receptor 1). Furthermore, our results support the involvement of these genes in the mechanism of action of the commonly prescribed BD drugs lithium, valproic acid, and carbamazepine.

A final and very important contribution of this study was a high-coverage catalogue of both gene-and isoform-level expression of transcripts in the BD brain. The latter allowed us to contribute to the general knowledge of the splicing landscape in the BD brain. We also were able to unbiasedly catalogue the non-coding transcriptome in the BD brain and identified some of the first lincRNAs in BD. Overall this work opens new doors for the study of BD both from the clinical as well as the basic research point of view.

The second primary objective of this thesis was to follow a candidate-gene approach to resolving aspects of BD susceptibility. This was contrary but also complementary to the exploratory focus of the first half of the thesis which was to investigate the genome-wide patterns of genetic as well as non-genetic susceptibility factors for BD. Nonetheless, candidate and hypothesis-driven approaches are equally important in that they allow more fine-tuned investigations of particular processes relevant to disease. In this case, the primary focus was synapse-related neurobiology, an important aspect of the BD susceptibility landscape. Specifically, we narrowed in on the role of the synapsin family of neuronal phosphoproteins composed of three genes (synapsins I, II, and III) and involved in synaptogenesis, synaptic transmission, and synaptic plasticity (237). We were not the first to study these genes in psychiatry however. In fact they had been proposed to play roles in several

psychiatric disorders including schizophrenia, BD and epilepsy (202, 245) through both genetic (214-216) and functional studies (3, 218-220, 246-248).

The work in Chapter 3.2 represents a follow-up of previous research from our group that postulated a role for Synapsin II (*SYN2*) in BD susceptibility and potentially response to lithium treatment. This gene had been identified through a linkage study of a larger fraction of the lithium-response familial cohort used for our WES study presented in Chapter 2.2. Since *SYN2* has been implicated in synaptic plasticity and transmission, synaptogenesis, and other major aspects of brain function and maintenance (237), it may be hypothesized that disruption of these roles could result in the onset of pathological conditions that may be mediated by drug treatment. Thus the first research focus was to understand how lithium treatment may modulate *SYN2* gene expression, which we explored through *in vitro* treatment studies in model cell lines. Firstly, long-term treatments were performed in Epstein-Barr-virus transformed lymphoblastoid cell lines (LCLs) from BD patients characterized for excellent lithium-response (7, 11, 164) in order to identify the effect of this drug in a model replicating the genetic background of response. In addition, the same experiments were repeated in human neuroblastoma and glioblastoma cell lines to model the biological context of brain cell types.

One of the major findings was a responder-specific effect of lithium in LCLs from BD patients, suggesting that while this gene is important for BD in general, there are genetic or epigenetic differences in the highly genetically homogeneous group of Li responders that makes them even more susceptible to gene expression modulation, at least at the *SYN2* locus. Additionally, our results suggested that the effect of long-term treatment with Li may be cell-type specific, as this was only significant in neuronal cells, but not in astrocytic or glial cells. Even more interestingly, lithium affected *SYN2a* but not *SYN2b* expression. This study points to *SYN2* isoforms as very interesting

players in the pathways of lithium-response, but further research is required to decipher the full pathway of Li action that leads to its clinically-relevant mood stabilizing capabilities in BD patients.

To follow-up on the role of this system in BD, in Chapter 3.3 we investigated gene expression dysregulation of *SYN2* and its highly homologous sister genes *SYN1* and *SYN3* in the prefrontal cortex using post-mortem brains from the Douglas-Bell Canada Brain Bank. The most important results were obtained with *SYN2*, which was overexpressed in BD compared to controls. This up-regulation was accounted for by the longer variant, *SYN2a*, while the converse was found in post-mortem brains from patients with MDD, who displayed up-regulated expression for the *SYN2b* variant, with no change for *SYN2a*. This expression difference between the two disorders may or may not be etiologically relevant, considering evidence that synapsin variants have overlapping function in the brain (202), but it was interesting in light of the isoform-specific effect of lithium-treatment demonstrated in Chapter 3.2. The distinct synapsin mRNA expression profiles in the BD and MDD post-mortem brain are interesting because they may give clues to a molecular marker for distinguishing the two clinically similar disorders.

Additionally, in this study we sought to elucidate part of the regulatory mechanism mediating the gene expression dysregulation identified at the synapsin loci, and explored one of the more common epigenetic modifications associated with gene expression: tri-methylation of the 4th lysine of histone 3 (H3K4me3). Epigenetic modifications have been investigated in various psychiatric phenotypes, including schizophrenia (249, 250), autism (238), major depression (248) and suicide (250) and are considered to be important mediators of disease. The particular epigenetic mark we focused on has been shown to be most abundant at transcriptional start sites of genes (256-258). It functions by opening up the chromatin and allowing transcriptional machinery to bind to the promoter region of genes, leading to the initiation of transcription and thus increased gene expression (256-258).

We showed that for both *SYN2* isoforms the changes in expression were correlated with enrichment of H3K4me3, in what was at the time the first study to identify an epigenetic mechanism to be involved in the regulation of this gene. Only recently has there been another report of an epigenetic mediator of *SYN2* in psychiatry, specifically showing a role for DNA methylation in the pathophysiology of suicide at this locus (297). From this work we concluded that brain expression differences seen in synapsin genes in mood disorders may be explained in part by differences in H3K4me3. However, this also opened the question of what other epigenetic mechanisms might be involved in the regulation of synapsins and other important synapse-related factors in mood disorders. To address this issue, in Chapter 3.4 we discussed how future research should better characterize these mechanisms by exploring the role of other chromatin epigenetic marks and DNA methylation, as well as characterizing their brain-region specificity.

The findings from Chapter 3, relating to synapsins and synaptic neurobiology do not follow specifically from the role of GPCRs characterized in Chapter 2. However, both make important contributions to the BD susceptibility field, and there is a strong connection between these findings in terms of neurobiology. In fact many important regulators of synaptic signaling and plasticity are GPCRs. These receptors are necessary for functional neurotransmission throughout the central nervous system, as they control neurophysiological processes ranging from movement to mood (298). At neuronal synapses, GPCRs and G proteins work together to regulate key aspects of neurotransmitter release, synaptic transmission, and synaptic plasticity, processes which are necessary for central nervous system physiology and behavior (298, 299).

The findings presented in this thesis and shared with the international research community through peer-reviewed journals represent some important contributions to our understanding of susceptibility factors for bipolar disorder. However, given the complexity of mental illness in general

and of BD particularly, there is still work to be done toward the identification of causal factors and characterization of symptom mediating factors. The future of BD is bright; however the cure is not yet within reach. Many large-cohort high-throughput studies using whole exome or whole genome sequencing are ongoing and expected to report results in the next years. These studies – including work from the Bipolar Sequencing Consortium which combines nearly every research group in the world currently collecting genomic sequencing data on BD individuals, including our own – are expected to identify most of the rare susceptibility factors and thus key in on the most important genes and pathways for BD genetic susceptibility. However this will not elucidate the complete picture for BD as the mediation of both genetic and environmental factors through transcriptomic and epigenomic mechanisms is clearly important. Given the difficulty of collecting post-mortem brain samples, it is not likely that many cohorts will emerge beyond those already discussed in this report. Instead, several groups are currently working on characterizing transcriptional differences across brain regions. In fact, we have also obtained post-mortem brain RNA sequencing data from the hippocampus, a brain region strongly implicated in BD and other mood disturbances. Many groups are working toward characterizing the epigenome in BD, both in the brain and peripherally. We have an on-going project aiming to follow up on the transcriptome findings from the anterior cingulate cortex by characterizing the genome-wide DNA methylation profile in the same brain samples in the hopes of desciphering a more intricate susceptibility landscape. Given the complex nature of BD, it is becoming increasingly clear that the future of research in this field consists of a combination of complementary approaches and systems biology interpretations of large data sets.

Part 4.2: Conclusions and Future Directions

Bipolar disorder (BD) is a complex mental illness, with causality being determined by contributions from both genetic and environmental susceptibility factors. To date, most of the findings point to broad dysregulation across many neurobiological pathways that are essential for brain function. To better characterize the genetic and epigenetic susceptibility landscape in BD, the work presented in this thesis employed a combination of high-throughput and candidate approaches to query the various levels of dysregulation contributing to BD susceptibility, causality, and course of illness. Major findings include the characterization of rare genomic variation in well-characterized families with increased loading of BD, characterization of the global coding as well as non-coding transcriptomic landscape in the post-mortem brains of individuals with BD, and a role for synaptic genes in BD susceptibility and response to treatment with the classical mood stabilizer drug lithium. These contributions not only serve to fill a major gap in our knowledge of the BD susceptibility profile, but are some of the first reports in psychiatry harnessing the power of high-throughput sequencing technology.

Appendices

Appendix 1: Significant contributions by the thesis author to other projects

First-author publications not included in the thesis:

1. No evidence for GADL1 variation as a bipolar disorder susceptibility factor in a Caucasian lithium-responsive cohort. **Cruceanu C**, Alda M, Dion PA, Turecki G, Rouleau GA. *Am J Psychiatry*. 2015 Jan 172(1):94-5.
2. Family-based exome-sequencing approach identifies rare susceptibility variants for lithium-responsive bipolar disorder. **Cruceanu C**, Ambalavanan A, Spiegelman D, Gauthier J, Lafrenière RG, Dion PA, Alda M, Turecki G, Rouleau GA. *Genome*. 2013 Oct 56(10):634-40.
3. Lithium: a key to the genetics of bipolar disorder. **Cruceanu C**, Alda M, Turecki G. *Genome Med*. 2009 Aug 1(8):79
4. Pharmacogenetics of lithium response. **C Cruceanu**, CL Lara, M Alda, G Turecki. *Psychiatric Annals* 38. 2008

Contributions to other projects not directly relevant for the thesis topic:

5. Biomarker discovery: quantification of microRNAs and other small non-coding RNAs using next generation sequencing. Lopez JP, Diallo A, **Cruceanu C**, Fiori LM, Laboissiere S, Guillet I, Fontaine J, Ragoussis J, Benes V, Turecki G, Ernst C. BMC Med Genomics. 2015 Jul 8:35.
6. Effects of postmortem interval on biomolecule integrity in the brain. Nagy C, Maheu M, Lopez JP, Vaillancourt K, **Cruceanu C**, Gross JA, Arnovitz M, Mechawar N, Turecki G. J Neuropathol Exp Neurol. 2015 May 74(5):459-69
7. Novel integrative genomic tool for interrogating lithium response in bipolar disorder. Hunsberger JG, Chibane FL, Elkahloun AG, Henderson R, Singh R, Lawson J, **Cruceanu C**, Nagarajan V, Turecki G, Squassina A, Medeiros CD, Del Zompo M, Rouleau GA, Alda M, Chuang DM. Transl Psychiatry. 2015 Feb 5:e504.
8. Evidence for increased microglial priming and macrophage recruitment in the dorsal anterior cingulate white matter of depressed suicides. Torres-Platas SG, **Cruceanu C**, Chen GG, Turecki G, Mechawar N. *Brain Behav Immun*. 2014 Nov 42:50-9.
9. miR-1202 is a primate-specific and brain-enriched microRNA involved in major depression and antidepressant treatment. Lopez JP, Lim R, **Cruceanu C**, Crapper L, Fasano C, Labonte B, Maussion G, Yang JP, Yerko V, Vigneault E, El Mestikawy S, Mechawar N, Pavlidis P, Turecki G. *Nat Med*. 2014 Jul 20(7):764-8.
10. Decreased global methylation in patients with bipolar disorder who respond to lithium. Huzayyin AA, Andreazza AC, Turecki G, **Cruceanu C**, Rouleau GA, Alda M, Young LT. *Int J Neuropsychopharmacol*. 2014 Apr 17(4):561-9.

11. Morphometric characterization of microglial phenotypes in human cerebral cortex. Torres-Platas SG, Comeau S, Rachalski A, Bo GD, **Cruceanu C**, Turecki G, Giros B, Mechawar N. *J Neuroinflammation*. 2014 Jan 21;11:12.
12. Analysis of oxysterols and cholesterol in prefrontal cortex of suicides. Freemantle E, Chen GG, **Cruceanu C**, Mechawar N, Turecki G. *Int J Neuropsychopharmacol*. 2013 Jul 16(6):1241-9.
13. Implication of synapse-related genes in bipolar disorder by linkage and gene expression analyses. Lopez de Lara C, Jaitovich-Groisman I, **Cruceanu C**, Mamdani F, Lebel V, Yerko V, Beck A, Young LT, Rouleau G, Grof P, Alda M, Turecki G. *Int J Neuropsychopharmacol*. 2010 Nov 13(10);1297-410.

Articles published by the Consortium for Lithium Genetics, of which the thesis author is an author:

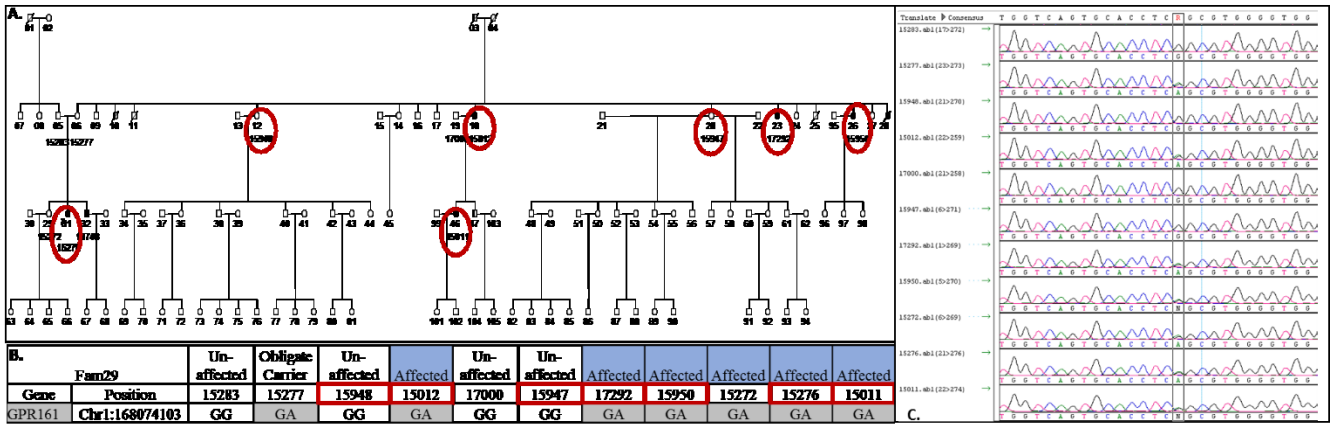
14. Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. Hou L, Heilbronner U, Degenhardt F, Adli M, Akiyama K, Akula N, Ardau R, Arias B, Backlund L, Banzato CE, Benabarre A, Bengesser S, Bhattacharjee AK, Biernacka JM, Birner A, Brichant-Petitjean C, Bui ET, Cervantes P, Chen GB, Chen HC, Chillotti C, Cichon S, Clark SR, Colom F, Cousins DA, **Cruceanu C**, Czerski PM, Dantas CR, Dayer A, Étain B, Falkai P, Forstner AJ, Frisén L, Fullerton JM, Gard S, Garnham JS, Goes FS, Grof P, Gruber O, Hashimoto R, Hauser J, Herms S, Hoffmann P, Hofmann A, Jamain S, Jiménez E, Kahn JP, Kassem L, Kittel-Schneider S, Kliwicki S, König B, Kusumi I, Lackner N, Laje G, Landén M, Lavebratt C, Leboyer M, Leckband SG, Jaramillo CA, MacQueen G, Manchia M, Martinsson L, Mattheisen M, McCarthy MJ, McElroy SL, Mitjans M, Mondimore FM, Monteleone P, Nievergelt CM, Nöthen MM, Ösby U, Ozaki N, Perlis RH, Pfennig A, Reich-Erkelenz D, Rouleau GA, Schofield PR, Schubert KO, Schweizer BW, Seemüller F, Severino G, Shekhtman T, Shilling PD, Shimoda K, Simhandl C, Slaney CM, Smoller JW, Squassina A, Stamm T, Stopkova P, Tighe SK, Tortorella A, Turecki G, Volkert J, Witt S, Wright A, Young LT, Zandi PP, Potash JB, DePaulo JR, Bauer M, Reininghaus EZ, Novák T, Aubry JM, Maj M, Baune BT, Mitchell PB, Vieta E, Frye MA, Rybakowski JK, Kuo PH, Kato T, Grigoroiu-Serbanescu M, Reif A, Del Zompo M, Bellivier F, Schalling M, Wray NR, Kelsoe JR, Alda M, Rietschel M, McMahon FJ, Schulze TG. *Lancet*. 2016 Jan 21. [Epub ahead of print]
15. Genome-wide analysis implicates microRNAs and their target genes in the development of bipolar disorder. Forstner AJ, Hofmann A, Maaser A, Sumer S, Khudayberdiev S, Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, Mattheisen M, Schumacher J, Breuer R, Meier S, Herms S, Hoffmann P, Lacour A, Witt SH, Reif A, Müller-Myhsok B, Lucae S, Maier W, Schwarz M, Vedder H, Kammerer-Ciernioch J, Pfennig A, Bauer M, Hautzinger M, Moebus S, Priebe L, Sivalingam S, Verhaert A, Schulz H, Czerski PM, Hauser J, Lissowska J, Szeszenia-Dabrowska N, Brennan P, McKay JD, Wright A, Mitchell PB, Fullerton JM, Schofield PR, Montgomery GW, Medland SE, Gordon SD, Martin NG, Krasnov V, Chuchalin A, Babadjanova G, Pantelejeva G, Abramova LI, Tiganov AS, Polonikov A, Khusnutdinova E, Alda M, **Cruceanu C**, Rouleau GA, Turecki G, Laprise C, Rivas F, Mayoral F, Kogevinas M,

Grigoroiu-Serbanescu M, Propping P, Becker T, Rietschel M, Cichon S, Schratt G, Nöthen MM. *Transl Psychiatry*. 2015 Nov 10;5:e678.

16. Variant GADL1 and response to lithium in bipolar I disorder. **Consortium on Lithium Genetics**, Hou L, Heilbronner U, Rietschel M, Kato T, Kuo PH, McMahon FJ, Schulze TG. *N Engl J Med*. 2014 May 370(19):1857-9.

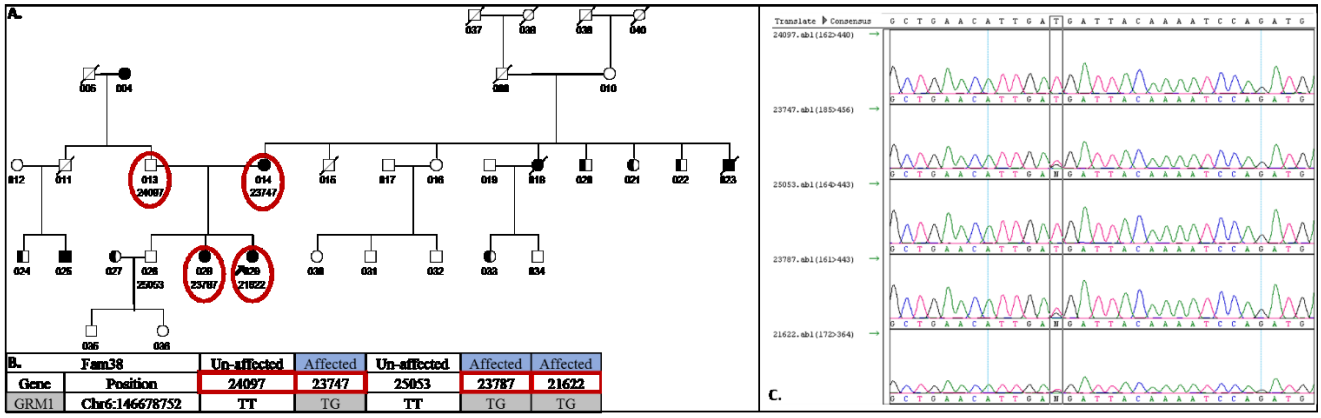
Appendix 2: Supplemental material: "Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors"

Supplemental Figures



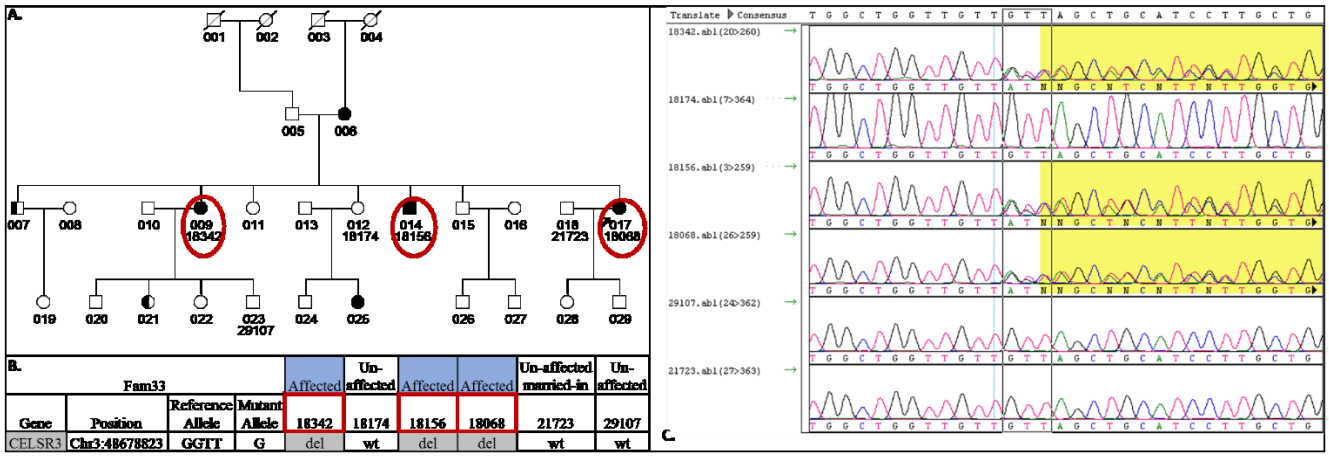
B. Fam29		Un-affected	Obligate Carrier	Un-affected	Affected	Un-affected	Un-affected	Affected	Affected	Affected	Affected	Affected
Gene	Position	15283	15277	15948	15012	17000	15947	17292	15950	15272	15276	15011
GPR161	Chr1:168074103	GG	GA	GG	GA	GG	GG	GA	GA	GA	GA	GA

Supplemental Figure 1: A. Pedigree of family number 29. **B.** Descriptive results of Sanger sequencing validation of the missense mutation n the GPCR gene *GPR161*. **C.** Sanger sequencing tracks for the *GPR161* mutation.

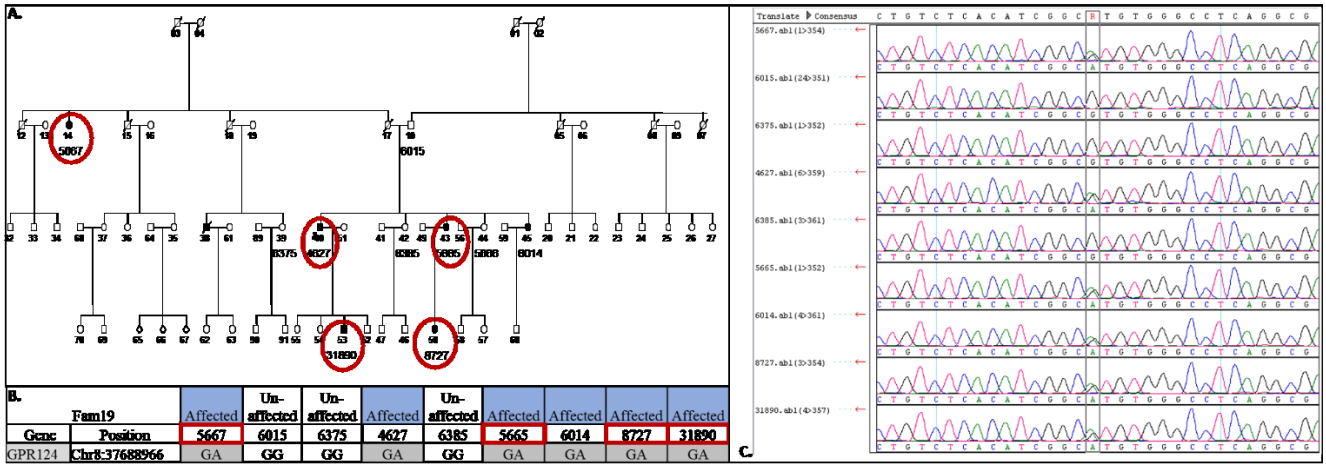


B. Fam38		Un-affected	Affected	Un-affected	Affected	Affected
Gene	Position	24097	23747	25053	23787	21622
GRM1	Chr6:146678752	TT	TG	TT	TG	TG

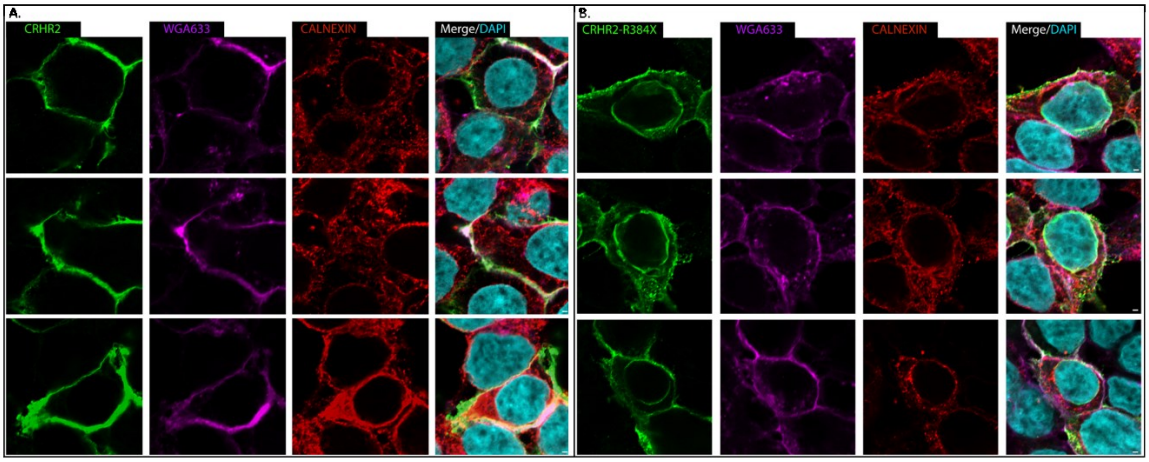
Supplemental Figure 2: A. Pedigree of family number 38. **B.** Descriptive results of Sanger sequencing validation of the missense mutation n the GPCR gene *GRM1*. **C.** Sanger sequencing tracks for the *GRM1* mutation.



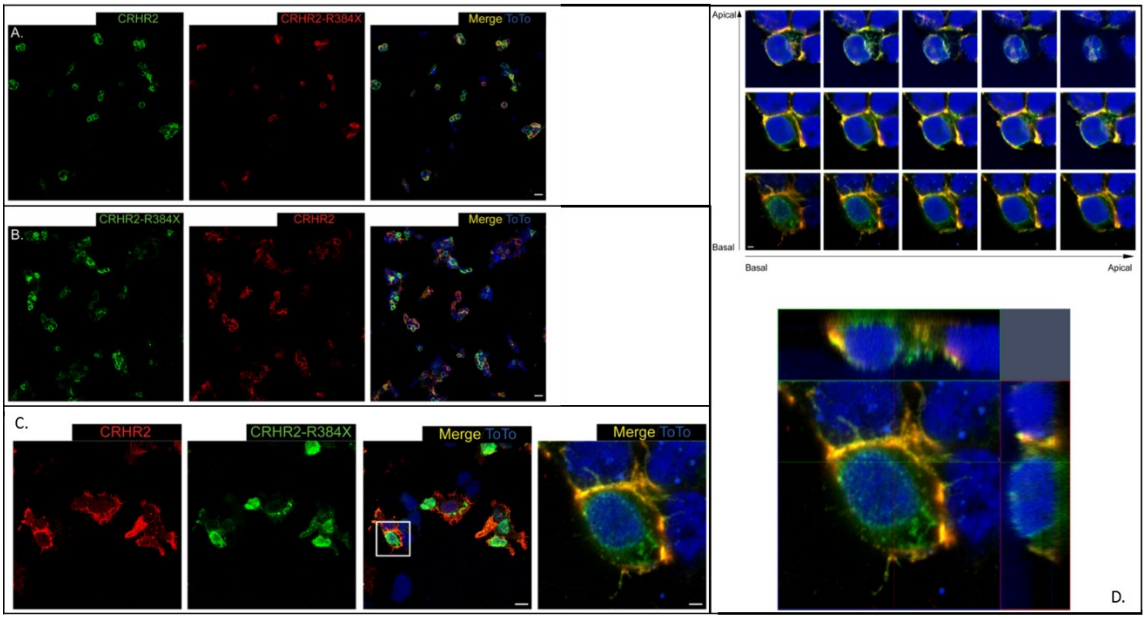
Supplemental Figure 3: A. Pedigree of family number 33. **B.** Descriptive results of Sanger sequencing validation of the missense mutation n the GPCR gene *CELSR3*. **C.** Sanger sequencing tracks for the *CELSR3* mutation.



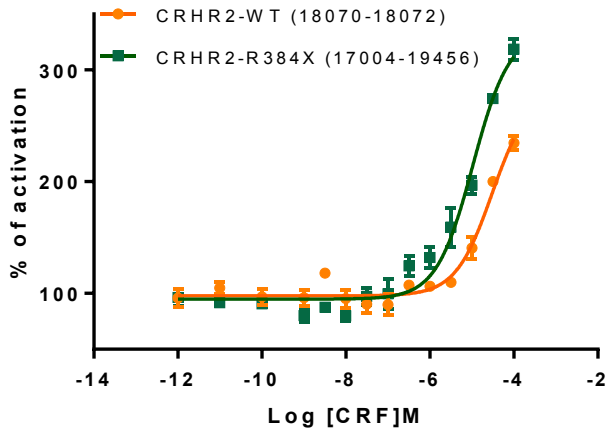
Supplemental Figure 4: A. Pedigree of family number 19. **B.** Descriptive results of Sanger sequencing validation of the missense mutation n the GPCR gene *GPR124*. **C.** Sanger sequencing tracks for the *GPR124* mutation.



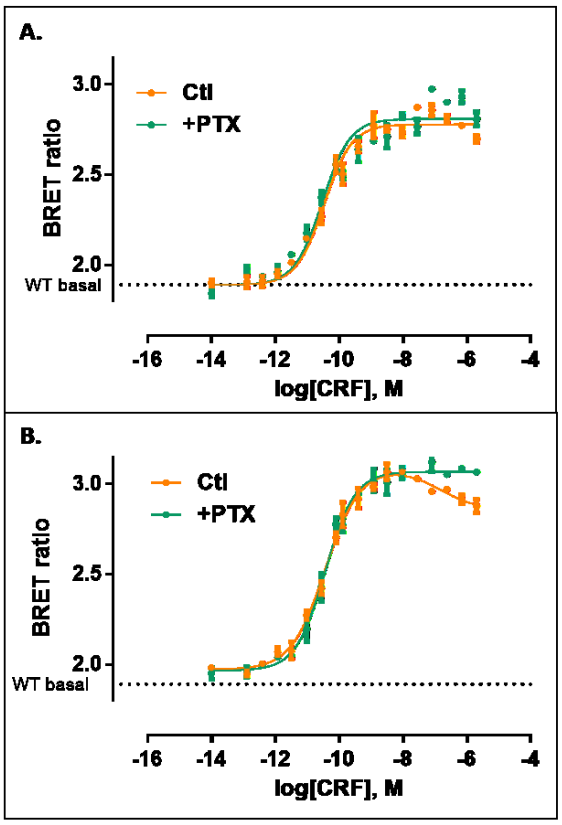
Supplemental Figure 5: **CRHR2-WT and CRHR2-R384X protein localization at the membrane in HEK293T cells**. **A.** A panel of three different HEK293T cells expressing transfected *CRHR2-WT* (green) showing the protein's presence at the membrane, exemplified by co-localization with WGA633 (purple) in three separately transfected HEK293T cells. Limited co-localization with CALNEXIN (red) an endoplasmic reticulum marker was observed. Far right panel: white represents area of merged staining, cyan represents nuclear staining using DAPI. Scale bar, 1 μm. **B.** A panel of three different HEK293T cells expressing transfected *CRHR2-R384X* is shown. CRHR2-R384X (green) was shown to be present at the membrane, as exemplified by co-localization with WGA633 (purple), but also throughout the cytoplasm, as exemplified by co-localization with CALNEXIN (red), an endoplasmic reticulum marker. Far right panel: white represents area of merged staining, cyan represents nuclear staining using DAPI. Scale bar, 1 μm.



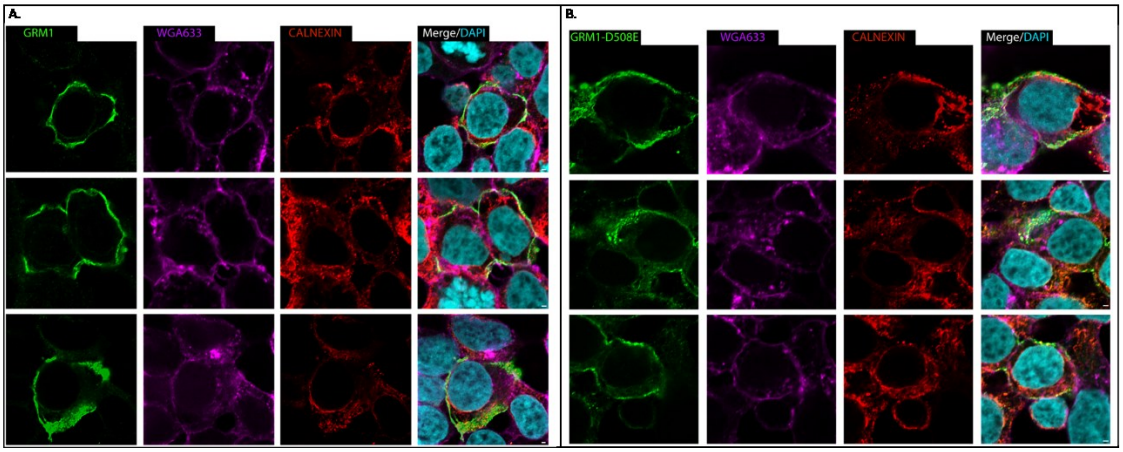
Supplemental Figure 6. CRHR2 and CRHR2-R384X proteins expression localized at the membrane in HEK293T cells. CRHR2 expression was evaluated using myc- and flag-tagged constructs for both CRHR2-WT and CRHR2-R384X proteins. **A.** Confocal microscopy imaging showing that the CRHR2-WT (green) and CRHR2-R384X (red) protein localized at the cell membrane in HEK293T cells. **B.** Confocal microscopy imaging obtained in a reciprocal experiment suggested that the CRHR2-WT (red) and CRHR2-R384X (green) protein localized at the cell membrane in HEK293T. Scale bars, 20 µm. C. Confocal microscopy imaging demonstrated the presence of CRHR2-R384X mutant proteins (green) both at the membrane and in the cytoplasm of HEK293T cells, whereas the CRHR2-WT protein (red) was predominantly found at the membrane. Scale bar, 10 µm. **D.** Deconvoluted confocal images revealed cytoplasmic expression of CRHR2-R384X mutant protein. (Top) Panel of serial images obtained through different Z-plane revealed a discrepancy between the expressions of CRHR2-R384X in comparison to CRHR2-WT protein. CRHR2-R384X is found at the membrane and in the cytoplasm of HEK293T cells (green), whereas CRHR2-WT (red) is found exclusively at the membrane. The image series depict the cell most basal Z-plane in the lower left and the most apical Z-plane in the upper right. Scale bar, 2 µm. (Bottom) Orthogonal plane of the image series depicted in (**C**); Green represents CRHR2-R384X staining; Red represents CRHR2-WT staining; Yellow represents area of merge staining; Blue represents nuclear staining using ToTo. The green and red lines represent X and Y positions respectively, the blue line represents the Z position.



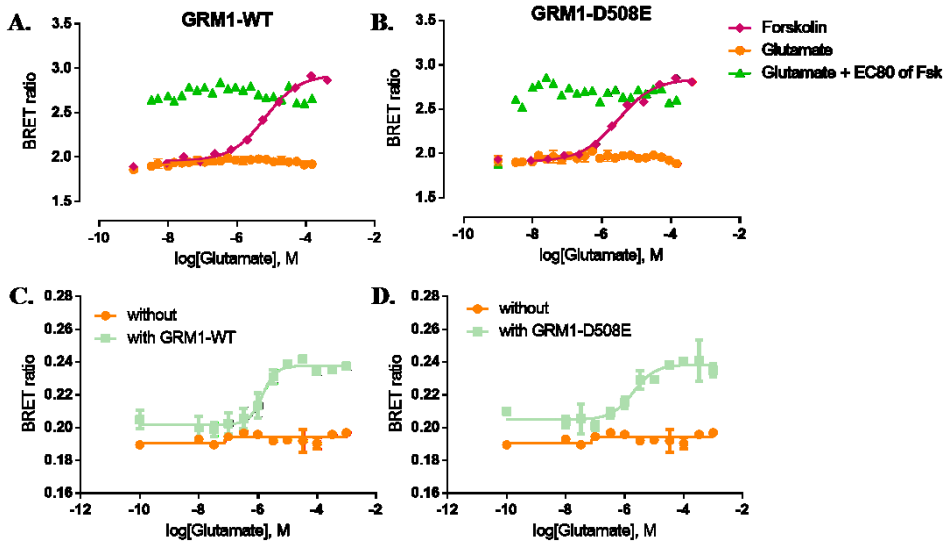
Supplemental Figure 7: Altered agonist-induced CRHR2 signaling in human cells from patients with the CRHR2-R384X mutation and control family members wild-type at the CRHR2 locus. [^{35}S]-GTPγS binding assay shows significantly stronger CRHR2 activation by the agonist Corticotropin Releasing Factor (CRF) in membranes prepared from Epstein-Barr virus-transformed lymphoblastoid cell lines from individuals in Fam28 who were carriers for the *CRHR2-R384X* mutation (individuals 17004, 19456; n=2) as compared to wild-type family controls (individuals: 18070,18072; n=2). Family pedigree presented in **Figure 1A**. Graphs represent the accumulation of two independent experiments (n=2) and curves were fitted using a dose-response nonlinear fit. Data are presented as mean ± SEM. EC_{50} and Emax values were as follows: control group (24.240 μM, 268.1%), mutant group (10.580 μM, 334.5%).



Supplemental Figure 8: Cyclic AMP (cAMP) modulation by CRHR2-WT and CRHR2-R384X. In HEK293T cells, 20ng of CRHR2-WT **(A)** or 500ng of CRHR2-R384X **(B)** were transfected with plasmids coding for cAMP biosensors. Increasing amounts of CRF peptide were added overnight to cells with (green curve) and without (orange cure) 100ng/mL of Pertussis toxin (PTX) and the BRET assay was performed as described. The black dashed line represents the BRET signal level for cells expressing the biosensor in absence of ligand (constitutive receptor activity is detected when the curve starts above this line). Graphs represent the accumulation of two independent experiments (n=2) and curves were fitted using a dose-response nonlinear fit.



Supplemental Figure 9: **GRM1-WT and GRM1-D508E protein localization at the membrane in HEK293T cells**. **A.** A panel of three different HEK293T cells expressing transfected GRM1-WT (green), showing the protein's presence at the membrane, exemplified by co-localization with WGA633 (purple). Little to no co-localization with CALNEXIN (red) an endoplasmic reticulum marker was observed. Far right panel: white represents area of merged staining, cyan represents nuclear staining using DAPI. Scale bar, 1 µm. **B.** A panel of three different HEK293T cells expressing transfected GRM1-D508E (green) showing the protein's presence at the membrane, exemplified by co-localization with WGA633 (purple). Little to no co-localization with CALNEXIN (red) was also observed. Far right panel: white represents area of merged staining, and cyan represents nuclear staining using DAPI. Scale bar, 1 µm.



Supplemental Figure 10: Second messenger activation by GRM1. In HEK293T cells, 50ng of GRM1-WT or 50ng of GRM1-D508E were transfected with the plasmids coding for the cAMP **(A-B)** and PKC biosensor **(C-D)**. Increasing amounts of Glutamate was added and the BRET assay was performed as described. Graphs represent the combination of two independent experiments (n=2) and curves were fitted using a dose-response nonlinear fit.

Supplemental Tables

Table S1. Family structure

Family	Affected sequenced	Unaffected sequenced	Total
FAM1	3	0	3
FAM2	3	0	3
FAM3	3	1	4
FAM4	7	0	7
FAM5	3	2	5
FAM6	5	1	6
FAM7	4	1	5
FAM8	8	0	8
FAM9	3	0	3
FAM10	5	0	5
FAM11	4	0	4
FAM12	5	1	6
FAM13	4	1	5
FAM14	5	0	5
FAM15	3	0	3
FAM16	4	0	4
FAM17	3	1	4
FAM18	4	1	5
FAM19	4	0	4
FAM20	3	1	4
FAM21	2	2	4
FAM22	3	2	5
FAM23	6	1	7
FAM24	3	0	3
FAM25	7	0	7
FAM26	3	1	4
FAM27	4	1	5
FAM28	4	0	4
FAM29	5	2	7
FAM30	4	0	4
FAM31	3	1	4
FAM32	3	0	3
FAM33	3	0	3
FAM34	5	1	6
FAM35	4	0	4
FAM36	4	1	5
FAM37	7	0	7
FAM38	3	1	4
FAM39	4	0	4
FAM40	3	0	3
Total	163	23	186
Average	4,175	0,625	4,65

Table S2

Table S2:

(A) Sanger sequecing validation results of 7 variants prioritized in (B).
The stop mutation in the gene CRHR2 was the only variant to show perfect segregation across additional family members included I the Sanger sequencing follow-up. This data is presented in Figure 1B.

(B) Filtering criteria and identified variants

Filtering criteria
Freeze Set Filter: Pass
Frequency in 1000 Genomes, Complete Genomics, Exome Variant Server: ≤0.02
Family member total coverage: ≥4 per individual reads
Family Affected with variant: 4 (of 4 sequenced)
Family Unaffected with variant: 0 (of 0 sequenced)
Non family controls w/ variant: 0
Excluded variant classes: Intronic, Intergenic, downstream, synonymous
Total variants: 38

Prioritization criteria
Any stop or splicing variant
Any nonsynonymous variant with SIFT, PolyPhen, and LTR Scores within predictive tresholds.
Variants/genes selected for Sanger Sequencing validation: 7

Exome Seq.	yes	yes	yes	yes	no	no	no	no
Diagnosis	Affected	Affected	Affected	Affected	Unaffected	Unaffected	Unaffected	Unaffected
Individ ID	17004	15949	17371	19456	18070	18072	17514	18347
Genotype	AC	AC	AC	AC	AC	AC	AA	AA
	TC	TC	TC	TC	TC	TT	TT	TT
	AG	AG	AG	AG	AG	AA	AA	AA
	CT	CT	CT	CT	CC	CT	CC	CC
	TA	TA	TA	TA	TA	TT	TT	TT
→→	GA	GA	GA	GA	GG	GG	GG	GG
	GA	GA	GA	GA	GA	GG	GG	GG

Family ID	Chr	Position	Reference	Mutant Al	Variant Class	Variant Function Type	Gene	Variant Type	Detailed A	DBSNP (Cl	DBSNP (N	Thousand	CG Freque	EVS Frequ	SIFT score	PolyPhen	LRT score	Mutation T	PhyloP sc	GERP++ sc
Fam28	11	4660921	C	T	SNP	upstream	OR51D1					0.0023								
Fam28	11	5021117	G	C	SNP	exonic	OR51L1	nonsynonymous_SNV	OR51L1:NM_001004755:exon1:(			0.0027		0.002696	0	0.808	0.99969	0.015711	0.921119	1.23
Fam28	11	8111646	A	C	SNP	exonic	TUB	nonsynonymous_SNV	TUB:NM_177972:exon3:c.A121C:p.K41Q|TUB:NM_003320:exon						0	0.868	0.999992	0.99988	0.997585	4.68
Fam28	11	8190553	C	T	SNP	UTR5	RIC3													
Fam28	11	44939539	T	C	SNP	exonic	TSPAN18	nonsynonymous_SNV	TSPAN18:NM_130783:exon5:c.T275C:p.L92P						0	0.997	1	0.999716	0.997076	4.89
Fam28	11	113848522	G	A	SNP	exonic	HTR3A	nonsynonymous_SNV	HTR3A:NM_000869:e	rs1177930		0.0005	0.014	0.003625	0	0.003	0.997753	0.615497	0.936178	2.48
Fam28	13	36748957	C	T	SNP	exonic	CCDC169	nonsynonymous_SNV	SOHLH2:NM_017826:exon7:c.G			0.0018		0.001208	0.36	0.024	0.537172	0.029526	0.059562	-1.87
Fam28	15	65490682	C	T	SNP	exonic	CILP	nonsynonymous_SNV	CILP:NM_003613:exon9:c.G1942A:p.E648K					0.000279	0.03	0.73	0.999997	0.892118	0.998764	5.15
Fam28	15	89173656	T	G	SNP	UTR3	AEN													
Fam28	16	90025620	C	T	SNP	exonic	DEF8	nonsynonymous_SNV	DEF8:NM_017702:exon6:c.C571			0.0014		0.000324	0					
Fam28	17	73917631	T	G	SNP	exonic	FBF1	nonsynonymous_SNV	FBF1:NM_001080542:exon15:c.A1492C:p.T498P						0.18					
Fam28	17	74075101	G	A	SNP	upstream	ZACN													
Fam28	17	74562128	C	T	SNP	UTR3	ST6GALNAC2					0.01								
Fam28	19	2431843	C	A	SNP	exonic	LMNB2	nonsynonymous_SNV	LMNB2:NM_032737:exon10:c.G1588T:p.G530C						0.04	0.024	0.97356	0.020596	0.790603	1.93
Fam28	2	25611152	A	G	SNP	exonic	DTNB	nonsynonymous_SNV	DTNB:NM_183360:exon17:c.T1654C:p.S552P|DTNB:NM_001256						0.04					
Fam28	2	42275294	G	A	SNP	UTR5	PKDCC													
Fam28	2	162661063	A	G	SNP	exonic	SLC4A10	nonsynonymous_SNV	SLC4A10:NM_001178016:exon4:c.A268G:p.R90G|SLC4A10:NM_(						0.19					
Fam28	2	179257199	A	G	SNP	exonic	OSBPL6	nonsynonymous_SNV	OSBPL6:NM_032523:exon23:c.A2507G:p.N836S|OSBPL6:NM_00						0	0.997	1	0.999553	0.999016	5.52
Fam28	22	31320999	G	T	SNP	ncRNA_exonic	MORC2-AS1					0.0009								
Fam28	22	36003307	C	T	SNP	UTR3	MB							0.000093						
Fam28	22	42221817	C	A	SNP	exonic	CCDC134	nonsynonymous_SNV	CCDC134:NM_024821:exon7:c.C680A:p.S227Y						0	0.669	1	0.990285	0.998825	4.13
Fam28	3	15298590	C	T	SNP	exonic	SH3BP5	nonsynonymous_SNV	SH3BP5:NM_001018009:exon8:c.G449A:p.C150Y|SH3					0.001115	0.01	0.994	0.999462	0.791352	0.998936	5.3
Fam28	3	81699051	C	T	SNP	exonic	GBE1	nonsynonymous_SNV	GBE1:NM_000158:exon4:c.G451A:p.G151R						0					
Fam28	3	194118743	G	A	SNP	exonic	GP5	nonsynonymous_SNV	GP5:NM_004488:exon2:c.C269T:p.A90V						0.18	0.042	0.92293	0.00619	0.903402	-0.676
Fam28	4	76955947	T	A	SNP	exonic	CXCL11	stoploss_SNV	CXCL11:NM_005409:(	rs6175719		0.0009		0.001584	0.96	0.392133	0.830085	0	0.896254	1.92
Fam28	5	168690620	T	C	SNP	ncRNA_exonic	MIR585							0.000646						
Fam28	5	176005407	C	T	SNP	exonic	CDHR2	nonsynonymous_SNV	CDHR2:NM_001171976:exon16:c.C1616T:p.T539M|CDHR2:NM_(						0.13	0.996	1	0.867074	0.974239	3.73
Fam28	6	33554429	C	G	SNP	ncRNA_exonic	LINC00336													
Fam28	7	30693162	G	A	SNP	exonic	CRHR2	stopgain_SNV	CRHR2:NM_001883:e	rs8192492				0.001394	1	0.732426	1	1	0.980947	2.81
Fam28	8	8750666	C	G	SNP	UTR5	MFHAS1													
Fam28	8	146220807	T	C	SNP	ncRNA_exonic	TMED10P1													
Fam28	9	13125340	T	C	SNP	exonic	MPDZ	nonsynonymous_SNV	MPDZ:NM_003829:exon34:c.A4682G:p.H1561R						0.43					
Fam28	9	75567927	G	T	SNP	UTR5	ALDH1A1								0.06					
Fam28	9	107288675	G	C	SNP	exonic	OR13C4	nonsynonymous_SNV	OR13C4:NM_001001919:exon1:c.C816G:p.N272K						1	0.004	0.240199	0.002756	0.221804	-0.19
Fam28	9	139916016	G	A	SNP	exonic	ABCA2	nonsynonymous_SNV	ABCA2:NM_001606:exon8:c.C725T:p.P242L|ABCA2:N					0.000509	0.31	0.025	1.30E-05	0.002822	0.930981	-0.132
Fam28	X	50350384	G	A	SNP	exonic|splicing	SHROOM4	nonsynonymous_SNV	SHROOM4:NM_020717:exon6:c.C3758T:p.A1253V						0.01	0.319717	0.02877	5.87E-04	0.991702	2.82
Fam28	X	73070208	A	G	SNP	ncRNA_exonic	XIST													
Fam28	X	135496326	C	A	SNP	exonic	GPR112	nonsynonymous_SNV	GPR112:NM_153834:exon25:c.(			0.0018		0.003767	0	0.14	0.773235	0.001918	0.76667	1.77

Table S3 (rows 1-50 of 302): Gene Ontology analysis of all genes carrying mutations shared by three or more affected family individuals												
Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferroni	Benjamini	FDR
GOTERM_MF_FAT	GO:0032559~adenyl ribonucleotide binding	360	11.01591187	6.61E-12	NM_014003, NM_052988, …	2270	1497	12983	1.375402788	1.02E-08	1.02E-08	1.10E-08
GOTERM_MF_FAT	GO:0030554~adenyl nucleotide binding	376	11.50550796	6.92E-12	NM_014003, NM_052988, …	2270	1577	12983	1.363657645	1.07E-08	5.34E-09	1.15E-08
GOTERM_MF_FAT	GO:0001882~nucleoside binding	382	11.68910649	1.08E-11	NM_052988, NM_006488, …	2270	1612	12983	1.35533772	1.66E-08	5.54E-09	1.80E-08
GOTERM_MF_FAT	GO:0001883~purine nucleoside binding	379	11.59730722	1.53E-11	NM_014003, NM_052988, NM_003640, NM_006488, …	2270	1601	12983	1.353932702	2.36E-08	5.89E-09	2.55E-08
GOTERM_MF_FAT	GO:0005524~ATP binding	354	10.83231334	1.66E-11	NM_014003, NM_052988, NM_003640, NM_006488, …	2270	1477	12983	1.37079328	2.57E-08	5.13E-09	2.77E-08
GOTERM_CC_FAT	GO:0005856~cytoskeleton	322	9.853121175	3.03E-09	NM_002487, NM_012144, NM_024991, NM_003450, …	2227	1381	12782	1.3382609	2.16E-06	2.16E-06	4.57E-06
GOTERM_CC_FAT	GO:0005578~proteinaceous extracellular matrix	97	2.968176255	1.52E-08	NM_002291, NM_002293, NM_002900, NM_031889, …	2227	320	12782	1.739804109	1.09E-05	5.43E-06	2.30E-05
GOTERM_MF_FAT	GO:0017076~purine nucleotide binding	423	12.94369645	1.94E-08	NM_052988, NM_006488, NM_006484, NM_016457, …	2270	1918	12983	1.261365547	3.00E-05	4.99E-06	3.24E-05
GOTERM_CC_FAT	GO:0031012~extracellular matrix	102	3.121175031	2.59E-08	NM_002291, NM_002293, NM_002900, NM_031889, …	2227	345	12782	1.696913375	1.85E-05	6.16E-06	3.92E-05
GOTERM_MF_FAT	GO:0032555~purine ribonucleotide binding	406	12.42350061	2.99E-08	NM_052988, NM_006488, NM_006484, NM_016457, …	2270	1836	12983	1.264743793	4.61E-05	6.59E-06	4.99E-05
GOTERM_MF_FAT	GO:0032553~ribonucleotide binding	406	12.42350061	2.99E-08	NM_052988, NM_006488, NM_006484, NM_016457, …	2270	1836	12983	1.264743793	4.61E-05	6.59E-06	4.99E-05
GOTERM_MF_FAT	GO:0005509~calcium ion binding	223	6.82374541	6.45E-08	NM_003737, NM_138769, NM_031882, NM_032457, …	2270	919	12983	1.387837287	9.95E-05	1.24E-05	1.08E-04
GOTERM_BP_FAT	GO:0022610~biological adhesion	174	5.324357405	2.79E-07	NM_002291, NM_003737, NM_002293, NM_031882, …	2346	701	13528	1.431320255	0.001206248	0.001206248	5.23E-04
GOTERM_BP_FAT	GO:0007155~cell adhesion	173	5.29375765	4.14E-07	NM_002291, NM_003737, NM_002293, NM_031882, …	2346	700	13528	1.425127268	0.001787989	8.94E-04	7.76E-04
GOTERM_CC_FAT	GO:0044430~cytoskeletal part	224	6.854345165	5.71E-07	NM_002487, NM_012144, NM_032524, NM_206862, …	2227	952	12782	1.350484693	4.07E-04	1.02E-04	8.63E-04
GOTERM_MF_FAT	GO:0043169~cation binding	830	25.39779682	7.33E-07	NM_182931, NM_021777, NM_032522, NM_007055, …	2270	4179	12983	1.135938767	0.00113	1.26E-04	0.001222348
GOTERM_MF_FAT	GO:0043167~ion binding	841	25.73439412	7.54E-07	NM_182931, NM_021777, NM_032522, NM_007055, …	2270	4241	12983	1.134166782	0.00116128	1.16E-04	0.001256204
GOTERM_MF_FAT	GO:0005085~guanyl-nucleotide exchange factor activity	52	1.591187271	1.12E-06	NM_020820, NM_001034853, NM_007200, …	2270	152	12983	1.956631115	0.001717969	1.56E-04	0.001858909
GOTERM_MF_FAT	GO:0000166~nucleotide binding	471	14.4124847	1.69E-06	NM_052988, NM_006488, NM_005094, NM_006484, …	2270	2245	12983	1.19992406	0.002605594	2.17E-04	0.002820594
GOTERM_MF_FAT	GO:0003774~motor activity	49	1.499388005	1.80E-06	NM_017539, NM_012144, NM_001127180, NM_017596, …	2270	142	12983	1.973589998	0.002768371	2.13E-04	0.002997045
GOTERM_BP_FAT	GO:0007156~homophilic cell adhesion	46	1.407588739	1.83E-06	NM_001007540, NM_152750, NM_003737, NM_001447, …	2346	131	13528	2.02484658	0.007864674	0.002628461	0.003423256
GOTERM_MF_FAT	GO:0046872~metal ion binding	819	25.06119951	1.97E-06	NM_182931, NM_021777, NM_032522, NM_007055, …	2270	4140	12983	1.13144321	0.003026931	2.17E-04	0.003277382
GOTERM_MF_FAT	GO:0003777~microtubule motor activity	32	0.979192166	2.01E-06	NM_017539, NM_017596, NM_015656, NM_001369, …	2270	77	12983	2.37688655	0.003098485	2.07E-04	0.003354976

GOTERM_BP_FAT	GO:0007018~microtubule-based movement	40	1.223990208	7.66E-06	NM_004181, NM_017539, NM_138769, NM_017596, …	2346	113	13528	2.041207403	0.032554214	0.008239838	0.014348139
GOTERM_CC_FAT	GO:0005604~basement membrane	31	0.948592411	8.15E-06	NM_002291, NM_021229, NM_015831, NM_002293,	2227	78	12782	2.281107158	0.005792681	0.00116123	0.012317191
GOTERM_CC_FAT	GO:0015630~microtubule cytoskeleton	136	4.161566707	8.95E-06	NM_002487, NM_012144, NM_018451, NM_023019, …	2227	549	12782	1.421821772	0.006364342	0.001063547	0.013536541
GOTERM_MF_FAT	GO:0004674~protein serine/threonine kinase activity	110	3.365973072	1.92E-05	NM_207189, NM_052988, NM_014720, NM_006484, …	2270	430	12983	1.463098043	0.029205051	0.001850786	0.032039251
GOTERM_CC_FAT	GO:0005874~microtubule	75	2.29498164	4.17E-05	NM_012144, NM_004434, NM_015656, NM_006640, …	2227	274	12782	1.571047431	0.029297509	0.004238884	0.06302787
GOTERM_MF_FAT	GO:0030695~GTPase regulator activity	103	3.151774786	4.19E-05	NM_001034853, NM_003835, NM_001134382, …	2270	404	12983	1.458159593	0.062625563	0.003797036	0.069894213
GOTERM_MF_FAT	GO:0005089~Rho guanyl-nucleotide exchange factor activity	28	0.856793146	7.20E-05	NM_020820, NM_007200, NM_173728, NM_014786, …	2270	74	12983	2.164090963	0.105049209	0.006146949	0.119917738
GOTERM_MF_FAT	GO:0060589~nucleoside-triphosphatase regulator activity	103	3.151774786	1.06E-04	NM_001034853, NM_003835, NM_001134382, …	2270	413	12983	1.426383719	0.150351034	0.008538705	0.175993877
GOTERM_BP_FAT	GO:0001539~ciliary or flagellar motility	10	0.305997552	1.20E-04	NM_017539, NM_173628, NM_015512, NM_178019, …	2346	14	13528	4.118864937	0.405672865	0.098833349	0.225339784
GOTERM_CC_FAT	GO:0044420~extracellular matrix part	38	1.162790698	1.22E-04	NM_002291, NM_152753, NM_021229, NM_015831, …	2227	117	12782	1.864130581	0.082998629	0.010772347	0.183547771
GOTERM_MF_FAT	GO:0004672~protein kinase activity	142	4.345165239	1.24E-04	NM_207189, NM_052988, NM_003640, NM_014720, …	2270	606	12983	1.340185516	0.174217148	0.009525514	0.206737572
GOTERM_MF_FAT	GO:0005088~Ras guanyl-nucleotide exchange factor activity	31	0.948592411	1.25E-04	NM_020820, NM_007200, NM_173728, NM_014786, …	2270	88	12983	2.014782739	0.17476783	0.009105459	0.207457275
GOTERM_BP_FAT	GO:0051056~regulation of small GTPase mediated signal transduction	68	2.080783354	1.37E-04	NM_020820, NM_014914, NM_007200, NM_015556, …	2346	252	13528	1.556015643	0.446617257	0.093910839	0.256213034
GOTERM_CC_FAT	GO:0005605~basal lamina	11	0.336597307	1.48E-04	NM_005559, NM_198129, NM_002291, NM_015831, …	2227	17	12782	3.713832906	0.100467223	0.011695491	0.224245414
GOTERM_MF_FAT	GO:0005083~small GTPase regulator activity	73	2.23378213	1.50E-04	NM_020820, NM_014914, NM_007200, NM_001134382, …	2270	274	12983	1.523777292	0.206987996	0.010486309	0.250415799
GOTERM_BP_FAT	GO:0043467~regulation of generation of precursor metabolites and energy	15	0.458996328	1.92E-04	NM_000162, NM_001079817, NM_003749, NM_006208, …	2346	30	13528	2.883205456	0.563221488	0.111599586	0.358489072
GOTERM_BP_FAT	GO:0006468~protein amino acid phosphorylation	152	4.651162791	1.95E-04	NM_052988, NM_003640, NM_178313, NM_014720, …	2346	667	13528	1.314084646	0.569830152	0.100077737	0.365075311
GOTERM_MF_FAT	GO:0016887~ATPase activity	85	2.600979192	2.22E-04	NM_014003, NM_173694, NM_080282, NM_080284, …	2270	334	12983	1.455531668	0.289924269	0.014775993	0.369473342
GOTERM_CC_FAT	GO:0030286~dynein complex	16	0.489596083	2.60E-04	NM_017539, NM_012144, NM_001369, NM_003777, …	2227	34	12782	2.700969386	0.169206412	0.01836663	0.392277406
GOTERM_BP_FAT	GO:0007017~microtubule-based process	67	2.050183599	2.75E-04	NM_138769, NM_015656, NM_012291, NM_003777, …	2346	253	13528	1.527073246	0.695579202	0.123789894	0.514328983
GOTERM_BP_FAT	GO:0010675~regulation of cellular carbohydrate metabolic process	17	0.520195838	2.98E-04	NM_000162, NM_001079817, NM_003749, …	2346	38	13528	2.579710145	0.723759029	0.120716732	0.556218766
GOTERM_BP_FAT	GO:0010906~regulation of glucose metabolic process	16	0.489596083	3.63E-04	NM_000162, NM_001079817, NM_003749, NM_002625, …	2346	35	13528	2.63607356	0.79199583	0.13302505	0.678468006

GOTERM_BP_FAT	GO:0006109~regulation of carbohydrate metabolic process	17	0.520195838	4.24E-04	NM_000162, NM_001079817, NM_003749, NM_002625, …	2346	39	13528	2.513563731	0.840164835	0.141699509	0.791835534
GOTERM_MF_FAT	GO:0003779~actin binding	81	2.478580171	6.91E-04	NM_020441, NM_178313, NM_012307, NM_023923, …	2270	326	12983	1.421073755	0.655791785	0.043464958	1.146401142
GOTERM_BP_FAT	GO:0016337~cell-cell adhesion	70	2.141982864	7.39E-04	NM_001007540, NM_003737, NM_001447, NM_031882, …	2346	276	13528	1.462495521	0.958953697	0.217780257	1.374853028

Table S4 (Rows 1-50 of 667). Mutations that recurred across families (3 or more affected segregation)

Family ID	Chr	Position	Reference Allele	Mutant Allele	Variant class	Gene symbol	Detailed annotation of the variant	1000 Genomes	Complete Genomics	EVS	SIFT	Poly Phen	LRT	Mutation Taster	PhyloP	GERP	Family affected wildtype	Family affected w variant	Family controls wildtype	Family controls w variant
FAM29	1	12726169	G	A	SNP	AADACL4	exonic:nonsynonymous_SNV:AADACL4:NM_001013630:exon4:c.G647A:p.R216Q	0,01		0,004767	0,07	0,981	0,998448	0,004302	0,922559	1,7	2	3	2	0
FAM35	1	12726169	G	A	SNP	AADACL4	exonic:nonsynonymous_SNV:AADACL4:NM_001013630:exon4:c.G647A:p.R216Q	0,01		0,004767	0,07	0,981	0,998448	0,004302	0,922559	1,7	1	3	0	0
FAM14	17	79093270	C	T	SNP	AATK	exonic:nonsynonymous_SNV:AATK:NM_004920:exon12:c.G3685A:p.A1229T|exonic:nonsynonymous_SNV:AATK:NM_001080395:exon13:c.G3994A:p.A1332T	0,01	0,007		0						1	4	0	0
FAM18	17	79093270	C	T	SNP	AATK	exonic:nonsynonymous_SNV:AATK:NM_004920:exon12:c.G3685A:p.A1229T|exonic:nonsynonymous_SNV:AATK:NM_001080395:exon13:c.G3994A:p.A1332T	0,01	0,007		0						1	3	1	0
FAM7	17	67081278	A	G	SNP	ABCA6	exonic:nonsynonymous_SNV:ABCA6:NM_080284:exon32:c.T4075C:p.C1359R	0,01		0,014224		1	0,999975	0,999982	0,998185	4,41	1	3	1	0
FAM8	17	67081278	A	G	SNP	ABCA6	exonic:nonsynonymous_SNV:ABCA6:NM_080284:exon32:c.T4075C:p.C1359R	0,01		0,014224		1	0,999975	0,999982	0,998185	4,41	5	3	0	0
FAM13	1	55085648	C	T	SNP	ACOT11| FAM151A	exonic:nonsynonymous_SNV:FAM151A:NM_176782:exon2:c.G151A:p.D51N|intronic:ACOT11:NM_015547	0,0041		0,006305	0,31	0,995	0,999998	0,422797	0,998155	4,25	0	4	1	0

FAM20	1	55085648	C	T	SNP	ACOT11 |FAM151A	exonic:nonsynonymous_SNV:FAM151A:NM_176782:exon2:c.G151A:p.D51N|intronic:ACOT11:NM_015547	0,0041		0,006305	0,31	0,995	0,999998	0,422797	0,998155	4,25	0	3	1	0
FAM13	2	111806832	T	A	SNP	ACOXL	exonic:stopgain_SNV:ACOXL:NM_001142807:exon16:c.T1407A:p.C469X	9,00E-04		0,000538	1	0,597218	0,999996	1	0,158844	-0,717	1	3	1	0
FAM37	2	111806832	T	A	SNP	ACOXL	exonic:stopgain_SNV:ACOXL:NM_001142807:exon16:c.T1407A:p.C469X	9,00E-04		0,000538	1	0,597218	0,999996	1	0,158844	-0,717	4	3	0	0
FAM6	16	20787240	T	C	SNP	ACSM3	exonic:nonsynonymous_SNV:ACSM3:NM_202000:exon3:c.T299C:p.L100P|exonic:nonsynonymous_SNV:ACSM3:NM_005622:exon3:c.T299C:p.L100P	0,01		0,008306	0	0,998	0,999999	0,998114	0,998433	5,57	2	3	1	0
FAM24	16	20787240	T	C	SNP	ACSM3	exonic:nonsynonymous_SNV:ACSM3:NM_202000:exon3:c.T299C:p.L100P|exonic:nonsynonymous_SNV:ACSM3:NM_005622:exon3:c.T299C:p.L100P	0,01		0,008306	0	0,998	0,999999	0,998114	0,998433	5,57	0	3	0	0
FAM13	2	114699940	A	AT	insertion	ACTR3	intronic_splicing:ACTR3:NM_005721(NM_005721:exon8:c.858+4->T)	0,01		0,013042							1	3	1	0
FAM37	2	114699940	A	AT	insertion	ACTR3	intronic_splicing:ACTR3:NM_005721(NM_005721:exon8:c.858+4->T)	0,01		0,013042							4	3	0	0
FAM34	5	178563002	C	T	SNP	ADAMTS2	exonic:nonsynonymous_SNV:ADAMTS2:NM_014244:exon13:c.G1993A:p.G665R	0,01		0,010303	0,09	0,004	0,99964	0,191149	0,977355	3,83	0	5	1	0
FAM39	5	178563002	C	T	SNP	ADAMTS2	exonic:nonsynonymous_SNV:ADAMTS2:NM_014244:exon13:c.G1993A:p.G665R	0,01		0,010303	0,09	0,004	0,99964	0,191149	0,977355	3,83	1	3	0	0

FAM14	7	31104520	A	G	SNP	ADCYAP1R1	exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199637:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199636:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199635:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001118:exon3:c.A125G:p.N42S	0,0046		0,003307	0,08	0	0,998592	0,212813	0,998673	4	2	3	0	0
FAM19	7	31104520	A	G	SNP	ADCYAP1R1	exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199637:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199636:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199635:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001118:exon3:c.A125G:p.N42S	0,0046		0,003307	0,08	0	0,998592	0,212813	0,998673	4	1	3	0	0
FAM14	8	67344765	C	T	SNP	ADHFE1	exonic:nonsynonymous_SNV:ADHFE1:NM_144650:exon1:c.C14T:p.A5V	0,01	0,014	0,003537	0,02	0,521542	0,885528	0,025509	0,885983	0,978	2	3	0	0
FAM40	8	67344765	C	T	SNP	ADHFE1	exonic:nonsynonymous_SNV:ADHFE1:NM_144650:exon1:c.C14T:p.A5V	0,01	0,014	0,003537	0,02	0,521542	0,885528	0,025509	0,885983	0,978	0	3	0	0
FAM8	20	60884154	G	A	SNP	ADRM1|LAMA5	UTR3:LAMA5:NM_005560|downstream:ADRM1:NM_175573|downstream:ADRM1:NM_007002	NA	0,007								2	4	0	0

FAM33	20	60884154	G	A	SNP	ADRM1|LAMA5	UTR3:LAMA5:NM_005560|downstream:ADRM1:NM_175573|downstream:ADRM1:NM_007002	NA	0,007								0	3	0	0
FAM4	11	47711820	A	G	SNP	AGBL2	exonic:nonsynonymous_SNV:AGBL2:NM_024783:exon10:c.T1439C:p.L480P	0,01	0,007	0,012387	0	0,967	1	0,979202	0,999115	5,61	2	4	0	0
FAM38	11	47711820	A	G	SNP	AGBL2	exonic:nonsynonymous_SNV:AGBL2:NM_024783:exon10:c.T1439C:p.L480P	0,01	0,007	0,012387	0	0,967	1	0,979202	0,999115	5,61	0	3	1	0
FAM14	7	134730252	C	T	SNP	AGBL3	exonic:nonsynonymous_SNV:AGBL3:NM_178563:exon10:c.C1655T:p.T552M	0,0037		0,009417	0,01						2	3	0	0
FAM25	7	134730252	C	T	SNP	AGBL3	exonic:nonsynonymous_SNV:AGBL3:NM_178563:exon10:c.C1655T:p.T552M	0,0037		0,009417	0,01						4	3	0	0
FAM24	6	135748451	T	C	SNP	AHI1	intronic_splicing:AHI1:NM_017651(NM_017651:exon20:c.2624-6A>G)|intronic_splicing:AHI1:NM_001134832(NM_001134832:exon20:c.2624-6A>G)|intronic_splicing:AHI1:NM_001134831(NM_001134831:exon21:c.2624-6A>G)|intronic_splicing:AHI1:NM_001134830(NM_001134830:exon19:c.2624-6A>G)	0,01		0,015981							0	3	0	0
FAM31	6	135748451	T	C	SNP	AHI1	intronic_splicing:AHI1:NM_017651(NM_017651:exon20:c.2624-6A>G)|intronic_splicing:AHI1:NM_001134832(NM_001134832:exon20	0,01		0,015981							0	3	1	0

							:c.2624-6A>G)|intronic_splicing:AHI1:NM_001134831(NM_001134831:exon21:c.2624-6A>G)|intronic_splicing:AHI1:NM_001134830(NM_001134830:exon19:c.2624-6A>G)												
FAM8	9	133996655	C	T	SNP	AIF1L	UTR3:AIF1L:NM_031426|UTR3:AIF1L:NM_001185096|UTR3:AIF1L:NM_001185095	0,01	0,007								2	6	0	0
FAM20	9	133996655	C	T	SNP	AIF1L	UTR3:AIF1L:NM_031426|UTR3:AIF1L:NM_001185096|UTR3:AIF1L:NM_001185095	0,01	0,007								0	3	1	0
FAM10	1	95448400	A	C	SNP	ALG14	UTR3:ALG14:NM_144988	0,0027	0,007								2	3	0	0
FAM40	1	95448400	A	C	SNP	ALG14	UTR3:ALG14:NM_144988	0,0027	0,007								0	3	0	0
FAM10	11	26463582	C	T	SNP	ANO3	exonic:nonsynonymous_SNV:ANO3:NM_031418:exon2:c.C164T:p.S55F	0,0018		0,006074	0	0,975	0,999999	0,983811	0,999241	5,29	2	3	0	0
FAM14	11	26463582	C	T	SNP	ANO3	exonic:nonsynonymous_SNV:ANO3:NM_031418:exon2:c.C164T:p.S55F	0,0018		0,006074	0	0,975	0,999999	0,983811	0,999241	5,29	2	3	0	0
FAM28	11	26682183	C	T	SNP	ANO3	UTR3:ANO3:NM_031418	0,01	0,014								0	4	0	0
FAM33	11	26682183	C	T	SNP	ANO3	UTR3:ANO3:NM_031418	0,01	0,014								0	3	0	0
FAM8	2	242135499	C	A	SNP	ANO7	intronic:ANO7:NM_001001891|UTR3:ANO7:NM_001001666	0,01	0,007								3	5	0	0
FAM35	2	242135499	C	A	SNP	ANO7	intronic:ANO7:NM_001001891|UTR3:ANO7:NM_001001666	0,01	0,007								1	3	0	0
FAM37	9	72064582	A	C	SNP	APBA1	exonic:nonsynonymous_SNV:APBA1:NM_001163:exon10:c.T2099G:p.L700R	NA			0	0,995	1	0,999996	0,999045	5,54	1	6	0	0
FAM39	9	72064582	A	C	SNP	APBA1	exonic:nonsynonymous_SNV:APBA1:NM_001163:exo	NA			0	0,995	1	0,999996	0,999045	5,54	0	4	0	0

							n10:c.T2099G:p.L700R												
FAM9	1	161018896	C	T	SNP	ARHGAP30	exonic:nonsynonymous_SNV:ARHGAP30:NM_181720:exon12:c.G1915A:p.G639R|exonic:nonsynonymous_SNV:ARHGAP30:NM_001025598:exon12:c.G1915A:p.G639R	0,01	0,007	0,014686	0,23	0	0,414415	0,274025	0,130608	0,95	0	3	0	0
FAM40	1	161018896	C	T	SNP	ARHGAP30	exonic:nonsynonymous_SNV:ARHGAP30:NM_181720:exon12:c.G1915A:p.G639R|exonic:nonsynonymous_SNV:ARHGAP30:NM_001025598:exon12:c.G1915A:p.G639R	0,01	0,007	0,014686	0,23	0	0,414415	0,274025	0,130608	0,95	0	3	0	0
FAM6	1	231114915	A	T	SNP	ARV1|TTC13	exonic:nonsynonymous_SNV:ARV1:NM_022786:exon1:c.A64T:p.T22S|upstream:TTC13:NM_024525|upstream:TTC13:NM_001122835	5,00E-04		0,001538	0,45	0	0,456656	0,000387	0,146545	0,487	0	5	1	0
FAM25	1	231114915	A	T	SNP	ARV1|TTC13	exonic:nonsynonymous_SNV:ARV1:NM_022786:exon1:c.A64T:p.T22S|upstream:TTC13:NM_024525|upstream:TTC13:NM_001122835	5,00E-04		0,001538	0,45	0	0,456656	0,000387	0,146545	0,487	3	4	0	0
FAM30	14	67807215	G	A	SNP	ATP6V1D	exonic:nonsynonymous_SNV:ATP6V1D:NM_015994:exon8:c.C544T:p.R182C	0,0014		0,002384	0,01	0,008	1	0,999986	0,977251	3,73	1	3	0	0
FAM35	14	67807215	G	A	SNP	ATP6V1D	exonic:nonsynonymous_SNV:ATP6V1D:NM_015994:exon8:c.C544T:p.R182C	0,0014		0,002384	0,01	0,008	1	0,999986	0,977251	3,73	1	3	0	0
FAM1	7	97921941	C	T	SNP	BAIAP2L1|BRI3	intronic:BRI3:NM_001159491|UTR3:BAIAP2L1:NM_018842	5,00E-04		0,006132							0	3	0	0

FAM4	**7**	**97921941**	**C**	**T**	**SNP**	**BAIAP2L1|BRI3**	**intronic:BRI3:NM_001159491|UTR3:BAIAP2L1:NM_018842**	**0,0005**		**0,006132**							**3**	**3**	**0**	**0**
FAM23	**12**	**56995856**	**G**	**C**	**SNP**	**BAZ2A**	**exonic:nonsynonymous_SNV:BAZ2A:NM_013449:exon20:c.C3551G:p.S1184C**	**0,0037**		**0,009415**	**0,04**						**0**	**4**	**1**	**0**

Table S5 (Rows 1-50 of 201). GPCR variants present among all filtered variants present in 3 or more affected individuals per family.																			
Family ID	Chr	Position	Ref Allele	Mutant Allele	Variant class	Gene symbol	Detailed annotation of the variant	1000 Genomes	Complete Genomics	EVS	SIFT	Poly Phen	LRT	Mutation Taster	PhyloP	GERP	Family affected wildtype	Family affected w variant	Family controls wildtype	Family controls w variant
FAM14	7	31104520	A	G	SNP	ADCYAP1R1	exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199637:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199636:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199635:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001118:exon3:c.A125G:p.N42S	0,0046		0,003307	0,08	0	0,998592	0,212813	0,998673	4	2	3	0	0
FAM19	7	31104520	A	G	SNP	ADCYAP1R1	exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199637:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199636:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001199635:exon3:c.A125G:p.N42S|exonic:nonsynonymous_SNV:ADCYAP1R1:NM_001118:exon3:c.A125G:p.N42S	0,0046		0,003307	0,08	0	0,998592	0,212813	0,998673	4	1	3	0	0
FAM35	4	3770178	G	T	SNP	ADRA2C	UTR3:ADRA2C:NM_000683	9,00E-04									1	3	0	0
FAM33	3	148459552	G	T	SNP	AGTR1	exonic:nonsynonymous_SNV:AGTR1:NM_032049:exon3:c.G817T:p.A273S|exonic:nonsynonymous_SNV:AGTR1:NM_031850:exon4:c.G835T:p.A279S|exonic:nonsynonymous_SNV:AGTR1:NM_009585:exon2:c.G730T:p.A244S|exonic:nonsynonymous_SNV:AGTR1:NM_004835:exon3:c.G835T:p.A279S|exonic:nonsynonymous_SNV:AGTR1:NM_000685:exon3:c.G730T:p.A244S	9,00E-04		0,003537	0	0,988	1	0,999154	0,999497	5,02	0	3	0	0
FAM38	3	148459130	T	C	SNP	AGTR1	exonic:nonsynonymous_SNV:AGTR1:NM_032049:exon3:c.T395C:p.I132T|exonic:nonsynonymous_SNV:AGTR1:NM_031850:exon4:c.T413C:p.I138T|exonic:nonsynonymous_SNV:AGTR1:NM_009585:exon2:c.T308C:p.I103T|exonic:nonsynonymous_SNV:AGTR1:NM_004835:exon3:c.T413C:p.I138T|exonic:nonsynonymous_SNV:AGTR1:NM_000685:exon3:c.T308C:p.I103T	0,0005		0,00123	0,11	0,001	0,994176	0,132654	0,997629	5,13	0	3	1	0

FAM13	1	206224999	T	C	SNP	AVPR1B	exonic:nonsynonymous_SNV:AVPR1B:NM_000707:exon1:c.T559C:p.W187R	0,0018		0,002849	0,03	0,998	0,999999	0,637963	0,975398	4,68	1	3	1	0
FAM32	6	70071382	C	T	SNP	BAI3	exonic:nonsynonymous_SNV:BAI3:NM_001704:exon29:c.C4217T:p.T1406M	NA		0,000154	0	0,05	0,999993	0,849736	0,98352	4,41	0	3	0	0
FAM18	11	6292790	G	A	SNP	CCKBR	UTR3:CCKBR:NM_176875	NA									1	3	1	0
FAM23	3	46399633	G	A	SNP	CCR2	exonic:nonsynonymous_SNV:CCR2:NM_001123396:exon2:c.G615A:p.M205I|exonic:nonsynonymous_SNV:CCR2:NM_001123041:exon2:c.G615A:p.M205I	5,00E-04			0	0,923	0,992864	0,00724	0,999408	4,9	2	4	1	0
FAM18	3	39374956	G	A	SNP	CCR8	UTR3:CCR8:NM_005201	0,0014									1	3	1	0
FAM4	3	45942554	A	G	SNP	CCR9	exonic:nonsynonymous_SNV:CCR9:NM_031200:exon3:c.A274G:p.I92V|exonic:nonsynonymous_SNV:CCR9:NM_006641:exon2:c.A238G:p.I80V|exonic:nonsynonymous_SNV:CCR9:NM_001256369:exon4:c.A238G:p.I80V|intronic:LZTFL1:NM_001276379|intronic:LZTFL1:NM_001276378	0,01	0,014	0,006766	0,48	0,037	0,999033	0,094623	0,880884	2,56	1	5	0	0
FAM4	22	46932580	C	A	SNP	CELSR1	exonic:nonsynonymous_SNV:CELSR1:NM_014246:exon1:c.G488T:p.R163M	NA			0,06	0,530191	0,988809	3,03E-04	0,939331	0,707	3	3	0	0
FAM15	3	48677191	G	A	SNP	CELSR3	exonic:nonsynonymous_SNV:CELSR3:NM_001407:exon34:c.C9827T:p.P3276L	0,01		0,006229	0	0,699879	0,95572	0,649392	0,988832	3,49	0	3	0	0
FAM27	3	48677858	C	T	SNP	CELSR3	exonic:nonsynonymous_SNV:CELSR3:NM_001407:exon34:c.G9160A:p.G3054R	NA		0,000231	0,01	0,578966	0,858536	0,265904	0,857777	0,709	1	3	1	0
FAM33	3	48678823	GGT	G	deletion	CELSR3	exonic:nonframeshift_deletion:CELSR3:NM_001407:exon33:c.8956_8958del:p.2986_2986del	NA		0,001997							0	3	0	0
FAM34	3	48698929	G	A	SNP	CELSR3	exonic:nonsynonymous_SNV:CELSR3:NM_001407:exon1:c.C1139T:p.P380L	NA		7,80E-05	0	0,748169	1	0,938763	0,999614	5,23	2	3	1	0
FAM25	7	136700373	G	T	SNP	CHRM2	exonic:nonsynonymous_SNV:CHRM2:NM_001006632:exon3:c.G761T:p.G254V|exonic:nonsynonymous_SNV:CHRM2:NM_001006631:exon4:c.G761T:p.G254V|exonic:nonsynonymous_SNV:CHRM2:NM_001006630:exon4:c.G761T:p.G254V|exonic:nonsynonymous_SNV:CHRM2:NM_001006629:exon2:c.G761T:p.G254V|exonic:nonsynonymous_SNV:CHRM2:NM_001006628:exon3:c.G761T:p.G254V|exonic:nonsynonymous_SNV:CHRM2:NM_001006627:exon3:c.G761T:p.G254V|exonic:nonsynonymous_SN	NA		7,70E-05	0,23	0	0,996614	0,972873	0,987726	3,97	4	3	0	0

							V:CHRM2:NM_001006626:exon5:c.G761T:p.G254V|exonic:nonsynonymous_SNV:CHRM2:NM_000739:exon4:c.G761T:p.G254V|ncRNA_intronic:LOC349160:NR_046103												
FAM29	15	34355792	G	A	SNP	CHRM5	exonic:nonsynonymous_SNV:CHRM5:NM_012125:exon3:c.G874A:p.A292T	NA		0,000154	0,83	0	0,703999	0,000557	0,032732	-3,43	2	3	2	0
FAM29	15	34355796	A	G	SNP	CHRM5	exonic:nonsynonymous_SNV:CHRM5:NM_012125:exon3:c.A878G:p.N293S	NA			1	0	0,988711	0,000281	0,999092	5,07	2	3	2	0
FAM15	6	88853594	G	A	SNP	CNR1	exonic:nonsynonymous_SNV:CNR1:NM_033181:exon2:c.C1301T:p.T434M|exonic:nonsynonymous_SNV:CNR1:NM_016083:exon2:c.C1400T:p.T467M|exonic:nonsynonymous_SNV:CNR1:NM_001160259:exon2:c.C1400T:p.T467M|exonic:nonsynonymous_SNV:CNR1:NM_001160258:exon4:c.C1400T:p.T467M|exonic:nonsynonymous_SNV:CNR1:NM_001160226:exon3:c.C1400T:p.T467M	NA		7,70E-05	0	0,907	1	0,999993	0,999761	5,59	0	3	0	0
FAM34	17	43906973	G	A	SNP	CRHR1	exonic:nonsynonymous_SNV:CRHR1:NM_001145146:exon6:c.G481A:p.V161M|intronic:CRHR1:NM_004382|intronic:CRHR1:NM_001256299|intronic:CRHR1:NM_001145148|intronic:CRHR1:NM_001145147	0,01		0,008979	0,05						2	3	1	0
FAM28	7	30693162	G	A	SNP	CRHR2	exonic:stopgain_SNV:CRHR2:NM_001883:exon12:c.C1150T:p.R384X|exonic:stopgain_SNV:CRHR2:NM_001202482:exon12:c.C1147T:p.R383X|exonic:stopgain_SNV:CRHR2:NM_001202481:exon14:c.C1108T:p.R370X|exonic:stopgain_SNV:CRHR2:NM_001202475:exon13:c.C1231T:p.R411X|UTR3:CRHR2:NM_001202483	NA		0,001461	1	0,732426	1	1	0,980947	2,81	0	4	0	0
FAM28	2	219029932	C	T	SNP	CXCR1	exonic:nonsynonymous_SNV:CXCR1:NM_000634:exon2:c.G3A:p.M1I	0,0014		0,005997	0,1	0,662	0,006133	0,787448	0,997049	-5,04	1	3	0	0
FAM8	11	637536	GCCGCCGACCTCCT	G	deletion	DRD4	exonic:frameshift_deletion:DRD4:NM_000797:exon1:c.233_245del:p.78_82del	NA		0,008361							5	3	0	0
FAM33	4	9784833	A	T	SNP	DRD5	exonic:nonsynonymous_SNV:DRD5:NM_000798:exon1:c.A1180T:p.I394F	NA			0,22	0,007	0,973119	0,10693	0,024153	-3,25	0	3	0	0
FAM33	4	9784834	T	C	SNP	DRD5	exonic:nonsynonymous_SNV:DRD5:NM_000798:exon1:c.T1181C:p.I394T	NA			0,04	0,003	0,973119	0,028814	0,996888	4,3	0	3	0	0

FAM16	13	78470568	G	A	SNP	EDNRB	UTR3:EDNRB:NM_003991|UTR3:EDNRB:NM_001201397|UTR3:EDNRB:NM_001122659|UTR3:EDNRB:NM_000115	0,01	0,007	0,009066							1	3	0	0
FAM16	19	6928269	A	T	SNP	EMR1	exonic:nonsynonymous_SNV:EMR1:NM_001256252:exon16:c.A2180T:p.E727V|intronic:EMR1:NM_001974|intronic:EMR1:NM_001256255|intronic:EMR1:NM_001256254|intronic:EMR1:NM_001256253	0,01		0,008458	0						1	3	0	0
FAM38	19	6928269	A	T	SNP	EMR1	exonic:nonsynonymous_SNV:EMR1:NM_001256252:exon16:c.A2180T:p.E727V|intronic:EMR1:NM_001974|intronic:EMR1:NM_001256255|intronic:EMR1:NM_001256254|intronic:EMR1:NM_001256253	0,01		0,008458	0						0	3	1	0
FAM7	19	14752344	G	C	SNP	EMR3	exonic:nonsynonymous_SNV:EMR3:NM_032571:exon10:c.C1135G:p.L379V	0,01	0,007	0,017151	0	0,518305	0,846531	0,004035	0,959916	1,92	1	3	1	0
FAM10	5	76114881	C	G	SNP	F2RL1	UTR5:F2RL1:NM_005242	NA									2	3	0	0
FAM11	17	74073386	C	G	SNP	GALR2	exonic:nonsynonymous_SNV:GALR2:NM_003857:exon2:c.C1038G:p.S346R	0,0023	0,007	0,002495	0,03	0,27	0,966931	0,693029	0,949885	2,66	1	3	0	0
FAM12	17	79768764	C	T	SNP	GCGR	exonic:nonsynonymous_SNV:GCGR:NM_000160:exon4:c.C227T:p.T76M	5,00E-04		0,000876	0,12						2	3	1	0
FAM8	7	31018855	G	A	SNP	GHRHR	exonic:nonsynonymous_SNV:GHRHR:NM_000823:exon13:c.G1268A:p.C423Y	NA		0,000231	0	0,999	0,919521	0,270907	0,9996	4,88	5	3	0	0
FAM23	9	132816257	A	G	SNP	GPR107	exonic:nonsynonymous_SNV:GPR107:NM_020960:exon1:c.A46G:p.R16G|exonic:nonsynonymous_SNV:GPR107:NM_001136558:exon1:c.A46G:p.R16G|exonic:nonsynonymous_SNV:GPR107:NM_001136557:exon1:c.A46G:p.R16G	NA			0,11	0,013	0,349815	0,000134	0,876982	0,886	3	3	1	0
FAM18	6	47647808	A	G	SNP	GPR111	exonic_splicing:nonsynonymous_SNV:GPR111:NM_153839:exon6:c.A269G:p.E90G	NA			0,12	0	0,060653	0,040745	0,016748	-7,23	1	3	1	0
FAM18	6	47648033	A	G	SNP	GPR111	exonic:nonsynonymous_SNV:GPR111:NM_153839:exon6:c.A494G:p.K165R	0,01	0,014	0,012385	0,28	0,183	0,998732	0,006593	0,998988	3,99	1	3	1	0
FAM33	6	47645624	A	G	SNP	GPR111	intronic_splicing:GPR111:NM_153839(NM_153839:exon4:c.115+3A>G)	NA									0	3	0	0
FAM8	X	135427828	A	G	SNP	GPR112	exonic:nonsynonymous_SNV:GPR112:NM_153834:exon6:c.A1963G:p.I655V	NA		0,000568	0,01	0,24	0,838319	8,00E-05	0,986727	2,73	2	6	0	0
FAM28	X	135496326	C	A	SNP	GPR112	exonic:nonsynonymous_SNV:GPR112:NM_153834:exon25:c.C9045A:p.S3015R	0,0018		0,003597	0	0,14	0,773235	0,001918	0,76667	1,77	0	4	0	0
FAM33	X	135429562	T	C	SNP	GPR112	exonic:nonsynonymous_SNV:GPR112:NM_153834:exon6	NA		0,002651	0,12	0,002	0,788608	0,001065	0,849002	1,39	0	3	0	0

							:c.T3697C:p.S1233P												
FAM7	2	26532034	T	G	SNP	GPR113	intronic:GPR113:NM_153835|UTR3:GPR113:NM_001145169	0,01	0,007	0,00657							1	3	1	0
FAM25	6	46852013	C	T	SNP	GPR116	intronic_splicing:GPR116:NM_015234(NM_015234:exon6:c.329-5G>A)|intronic_splicing:GPR116:NM_001098518(NM_001098518:exon6:c.329-5G>A)	0,0023		0,003921							4	3	0	0
FAM25	10	134942698	C	A	SNP	GPR123	exonic:nonsynonymous_SNV:GPR123:NM_001083909:exon7:c.C1366A:p.P456T	0,01	0,007	0,00719	0	0,877	0,998135	0,59848	0,153805	-5,43	4	3	0	0
FAM19	8	37688966	G	A	SNP	GPR124	exonic:nonsynonymous_SNV:GPR124:NM_032777:exon8:c.G958A:p.V320M	0,0032	0,007	0,008688	0,15	0,474	0,999596	0,340237	0,862921	0,543	0	4	0	0
FAM32	8	37688966	G	A	SNP	GPR124	exonic:nonsynonymous_SNV:GPR124:NM_032777:exon8:c.G958A:p.V320M	0,0032	0,007	0,008688	0,15	0,474	0,999596	0,340237	0,862921	0,543	0	3	0	0
FAM18	17	72368627	G	A	SNP	GPR142	exonic:nonsynonymous_SNV:GPR142:NM_181790:exon4:c.G1277A:p.R426Q	NA		0,000846	0,02	0,986	0,970281	0,743801	0,95023	2,21	1	3	1	0
FAM20	9	127215935	G	A	SNP	GPR144	exonic:stopgain_SNV:GPR144:NM_001161808:exon4:c.G959A:p.W320X	NA			1						0	3	1	0
FAM37	5	145895394	G	A	SNP	GPR151	exonic:stopgain_SNV:GPR151:NM_194251:exon1:c.C283T:p.R95X	0,01		0,005382	1	0,734626	0,999995	1	0,994325	4,53	1	6	0	0
FAM14	1	6310562	C	G	SNP	GPR153	exonic:nonsynonymous_SNV:GPR153:NM_207370:exon5:c.G1102C:p.G368R	0,0041		0,00831	0,65	0	0,995534	0,327189	0,972586	2,3	1	4	0	0

Table S6: List of all GPCR genes used for mutation burden comparisons.	
ADCYAP1R1	Pituitary adenylate cyclase-activating polypeptide type I receptor
ADORA1	Adenosine receptor A1
ADORA2A	Adenosine receptor A2a
ADORA2B	Adenosine receptor A2b
ADORA3	Adenosine A3 receptor
ADRA1A	Alpha-1A adrenergic receptor
ADRA1B	Alpha-1B adrenergic receptor
ADRA1D	Alpha-1D adrenergic receptor
ADRA2A	Alpha-2A adrenergic receptor
ADRA2B	Alpha-2B adrenergic receptor
ADRA2C	Alpha-2C adrenergic receptor
ADRB1	Beta-1 adrenergic receptor
ADRB2	Beta-2 adrenergic receptor
ADRB3	Beta-3 adrenergic receptor
AGTR1	Type-1 angiotensin II receptor
AGTR2	Type-2 angiotensin II receptor
APLNR	Apelin receptor
AVPR1A	Vasopressin V1a receptor
AVPR1B	Vasopressin V1b receptor
AVPR2	Vasopressin V2 receptor
BAI1	Brain-specific angiogenesis inhibitor 1
BAI2	Brain-specific angiogenesis inhibitor 2
BAI3	Brain-specific angiogenesis inhibitor 3
BDKRB1	B1 bradykinin receptor
BDKRB2	B2 bradykinin receptor
BRS3	Bombesin receptor subtype-3
C3AR1	C3a anaphylatoxin chemotactic receptor
C5AR1	C5a anaphylatoxin chemotactic receptor
CALCR	Calcitonin receptor
CALCRL	Calcitonin gene-related peptide type 1 receptor
CASR	Extracellular calcium-sensing receptor
CCBP2	Chemokine-binding protein 2
CCKAR	Cholecystokinin receptor type A
CCKBR	Gastrin/cholecystokinin type B receptor
CCR1	C-C chemokine receptor type 1
CCR10	C-C chemokine receptor type 10
CCR2	C-C chemokine receptor type 2
CCR3	C-C chemokine receptor type 3
CCR4	C-C chemokine receptor type 4
CCR5	C-C chemokine receptor type 5
CCR6	C-C chemokine receptor type 6
CCR7	C-C chemokine receptor type 7
CCR8	C-C chemokine receptor type 8
CCR9	C-C chemokine receptor type 9
CCRL1	C-C chemokine receptor type 11
CCRL2	C-C chemokine receptor-like 2
CD97	CD97 antigen
CELSR1	Cadherin EGF LAG seven-pass G-type receptor 1
CELSR2	Cadherin EGF LAG seven-pass G-type receptor 2
CELSR3	Cadherin EGF LAG seven-pass G-type receptor 3
CHRM1	Muscarinic acetylcholine receptor M1
CHRM2	Muscarinic acetylcholine receptor M2
CHRM3	Muscarinic acetylcholine receptor M3
CHRM4	Muscarinic acetylcholine receptor M4
CHRM5	Muscarinic acetylcholine receptor M5
CMKLR1	Chemokine receptor-like 1
CNR1	Cannabinoid receptor 1
CNR2	Cannabinoid receptor 2

CRHR1	Corticotropin-releasing factor receptor 1
CRHR2	Corticotropin-releasing factor receptor 2
CX3CR1	CX3C chemokine receptor 1
CXCR3	C-X-C chemokine receptor type 3
CXCR4	C-X-C chemokine receptor type 4
CXCR5	C-X-C chemokine receptor type 5
CXCR6	C-X-C chemokine receptor type 6
CXCR7	C-X-C chemokine receptor type 7
CYSLTR1	Cysteinyl leukotriene receptor 1
CYSLTR2	Cysteinyl leukotriene receptor 2
DARC	Duffy antigen/chemokine receptor
DRD1	D(1A) dopamine receptor
DRD2	D(2) dopamine receptor
DRD3	D(3) dopamine receptor
DRD4	D(4) dopamine receptor
DRD5	D(1B) dopamine receptor
EDNRA	Endothelin-1 receptor
EDNRB	Endothelin B receptor
ELTD1	EGF, latrophilin and seven transmembrane domain-containing protein 1
EMR1	EGF-like module-containing mucin-like hormone receptor-like 1
EMR2	EGF-like module-containing mucin-like hormone receptor-like 2
EMR3	EGF-like module-containing mucin-like hormone receptor-like 3
EMR4P	Putative EGF-like module-containing mucin-like hormone receptor-like 4
F2R	Proteinase-activated receptor 1
F2RL1	Proteinase-activated receptor 2
F2RL2	Proteinase-activated receptor 3
F2RL3	Proteinase-activated receptor 4
FFAR1	Free fatty acid receptor 1
FFAR2	Free fatty acid receptor 2
FFAR3	Free fatty acid receptor 3
FPR1	fMet-Leu-Phe receptor
FPR2	N-formyl peptide receptor 2
FPR3	N-formyl peptide receptor 3
FSHR	Follicle-stimulating hormone receptor
FZD1	Frizzled-1
FZD10	Frizzled-10
FZD2	Frizzled-2
FZD3	Frizzled-3
FZD4	Frizzled-4
FZD5	Frizzled-5
FZD6	Frizzled-6
FZD7	Frizzled-7
FZD8	Frizzled-8
FZD9	Frizzled-9
GABBR1	Gamma-aminobutyric acid type B receptor subunit 1
GABBR2	Gamma-aminobutyric acid type B receptor subunit 2
GALR1	Galanin receptor type 1
GALR2	Galanin receptor type 2
GALR3	Galanin receptor type 3
GCGR	Glucagon receptor
GHRHR	Growth hormone-releasing hormone receptor
GHSR	Growth hormone secretagogue receptor type 1
GIPR	Gastric inhibitory polypeptide receptor
GLP1R	Glucagon-like peptide 1 receptor
GLP2R	Glucagon-like peptide 2 receptor
GNRHR	Gonadotropin-releasing hormone receptor
GNRHR2	Putative gonadotropin-releasing hormone II receptor
GPBAR1	G protein-coupled bile acid receptor 1
GPER	G protein-coupled estrogen receptor 1
GPR1	Probable G protein-coupled receptor 1
GPR101	Probable G protein-coupled receptor 101

GPR110	Probable G protein-coupled receptor 110
GPR111	Probable G protein-coupled receptor 111
GPR112	Probable G protein-coupled receptor 112
GPR113	Probable G protein-coupled receptor 113
GPR114	Probable G protein-coupled receptor 114
GPR115	Probable G protein-coupled receptor 115
GPR116	Probable G protein-coupled receptor 116
GPR119	Glucose-dependent insulinotropic receptor
GPR12	G protein-coupled receptor 12
GPR120	G protein-coupled receptor 120
GPR123	Probable G protein-coupled receptor 123
GPR124	Probable G protein-coupled receptor 124
GPR125	Probable G protein-coupled receptor 125
GPR126	Probable G protein-coupled receptor 126
GPR128	Probable G protein-coupled receptor 128
GPR132	Probable G protein-coupled receptor 132
GPR133	Probable G protein-coupled receptor 133
GPR135	Probable G protein-coupled receptor 135
GPR139	Probable G protein-coupled receptor 139
GPR141	Probable G protein-coupled receptor 141
GPR142	Probable G protein-coupled receptor 142
GPR143	G protein-coupled receptor 143
GPR144	Probable G protein-coupled receptor 144
GPR146	Probable G protein-coupled receptor 146
GPR148	Probable G protein-coupled receptor 148
GPR149	Probable G protein-coupled receptor 149
GPR15	G protein-coupled receptor 15
GPR150	Probable G protein-coupled receptor 150
GPR151	Probable G protein-coupled receptor 151
GPR152	Probable G protein-coupled receptor 152
GPR153	Probable G protein-coupled receptor 153
GPR156	Probable G protein-coupled receptor 156
GPR157	Probable G protein-coupled receptor 157
GPR158	Probable G protein-coupled receptor 158
GPR160	Probable G protein-coupled receptor 160
GPR161	G protein-coupled receptor 161
GPR162	Probable G protein-coupled receptor 162
GPR17	Uracil nucleotide/cysteinyl leukotriene receptor
GPR171	Probable G protein-coupled receptor 171
GPR173	Probable G protein-coupled receptor 173
GPR174	Probable G protein-coupled receptor 174
GPR176	Probable G protein-coupled receptor 176
GPR179	Probable G protein-coupled receptor 179
GPR18	N-arachidonyl glycine receptor
GPR182	G protein-coupled receptor 182
GPR183	G protein-coupled receptor 183
GPR19	Probable G protein-coupled receptor 19
GPR20	Probable G protein-coupled receptor 20
GPR21	Probable G protein-coupled receptor 21
GPR22	Probable G protein-coupled receptor 22
GPR25	Probable G protein-coupled receptor 25
GPR26	Probable G protein-coupled receptor 26
GPR27	Probable G protein-coupled receptor 27
GPR3	G protein-coupled receptor 3
GPR31	Probable G protein-coupled receptor 31
GPR32	Probable G protein-coupled receptor 32
GPR33	Probable G protein-coupled receptor 33
GPR34	Probable G protein-coupled receptor 34
GPR35	G protein-coupled receptor 35
GPR37	Probable G protein-coupled receptor 37
GPR37L1	Endothelin B receptor-like protein 2

GPR39	G protein-coupled receptor 39
GPR4	G protein-coupled receptor 4
GPR42P	Putative G protein-coupled receptor 42
GPR44	Putative G protein-coupled receptor 44
GPR45	Probable G protein-coupled receptor 45
GPR50	Melatonin-related receptor
GPR52	Probable G protein-coupled receptor 52
GPR55	Probable G protein-coupled receptor 55
GPR56	G protein-coupled receptor 56
GPR6	G protein-coupled receptor 6
GPR61	Probable G protein-coupled receptor 61
GPR62	Probable G protein-coupled receptor 62
GPR63	Probable G protein-coupled receptor 63
GPR64	G protein-coupled receptor 64
GPR65	Psychosine receptor
GPR68	Ovarian cancer G protein-coupled receptor 1
GPR75	Probable G protein-coupled receptor 75
GPR77	C5a anaphylatoxin chemotactic receptor C5L2
GPR78	Probable G protein-coupled receptor 78
GPR81	G protein-coupled receptor 81
GPR82	Probable G protein-coupled receptor 82
GPR83	Probable G protein-coupled receptor 83
GPR84	G protein-coupled receptor 84
GPR85	Probable G protein-coupled receptor 85
GPR87	G protein-coupled receptor 87
GPR88	Probable G protein-coupled receptor 88
GPR97	Probable G protein-coupled receptor 97
GPR98	G protein-coupled receptor 98
GPRC5A	Retinoic acid-induced protein 3
GPRC5B	G protein-coupled receptor family C group 5 member B
GPRC5C	G protein-coupled receptor family C group 5 member C
GPRC5D	G protein-coupled receptor family C group 5 member D
GPRC6A	G protein-coupled receptor family C group 6 member A
GRM1	Metabotropic glutamate receptor 1
GRM2	Metabotropic glutamate receptor 2
GRM3	Metabotropic glutamate receptor 3
GRM4	Metabotropic glutamate receptor 4
GRM5	Metabotropic glutamate receptor 5
GRM6	Metabotropic glutamate receptor 6
GRM7	Metabotropic glutamate receptor 7
GRM8	Metabotropic glutamate receptor 8
GRPR	Gastrin-releasing peptide receptor
HCRTR1	Orexin receptor type 1
HCRTR2	Orexin receptor type 2
HRH1	Histamine H1 receptor
HRH2	Histamine H2 receptor
HRH3	Histamine H3 receptor
HRH4	Histamine H4 receptor
HTR1A	5-hydroxytryptamine receptor 1A
HTR1B	5-hydroxytryptamine receptor 1B
HTR1D	5-hydroxytryptamine receptor 1D
HTR1E	5-hydroxytryptamine receptor 1E
HTR1F	5-hydroxytryptamine receptor 1F
HTR2A	5-hydroxytryptamine receptor 2A
HTR2B	5-hydroxytryptamine receptor 2B
HTR2C	5-hydroxytryptamine receptor 2C
HTR4	5-hydroxytryptamine receptor 4
HTR5A	5-hydroxytryptamine receptor 5A
HTR6	5-hydroxytryptamine receptor 6
HTR7	5-hydroxytryptamine receptor 7
IL8RA	High affinity interleukin-8 receptor A

IL8RB	High affinity interleukin-8 receptor B
KISS1R	KiSS-1 receptor
LGR4	Leucine-rich repeat-containing G protein-coupled receptor 4
LGR5	Leucine-rich repeat-containing G protein-coupled receptor 5
LGR6	Leucine-rich repeat-containing G protein-coupled receptor 6
LHCGR	Lutropin-choriogonadotropic hormone receptor
LPAR1	Lysophosphatidic acid receptor 1
LPAR2	Lysophosphatidic acid receptor 2
LPAR3	Lysophosphatidic acid receptor 3
LPAR4	Lysophosphatidic acid receptor 4
LPAR5	Lysophosphatidic acid receptor 5
LPHN1	Latrophilin-1
LPHN2	Latrophilin-2
LPHN3	Latrophilin-3
LTB4R	Leukotriene B4 receptor 1
LTB4R2	Leukotriene B4 receptor 2
MAS1	MAS proto-oncogene
MAS1L	Mas-related G protein-coupled receptor MRG
MC1R	Melanocyte-stimulating hormone receptor
MC2R	Adrenocorticotropic hormone receptor
MC3R	Melanocortin receptor 3
MC4R	Melanocortin receptor 4
MC5R	Melanocortin receptor 5
MCHR1	Melanin-concentrating hormone receptor 1
MCHR2	Melanin-concentrating hormone receptor 2
MLNR	Motilin receptor
MRGPRD	Mas-related G protein-coupled receptor member D
MRGPRE	Mas-related G protein-coupled receptor member E
MRGPRF	Mas-related G protein-coupled receptor member F
MRGPRG	Mas-related G protein-coupled receptor member G
MRGPRX1	Mas-related G protein-coupled receptor member X1
MRGPRX2	Mas-related G protein-coupled receptor member X2
MRGPRX3	Mas-related G protein-coupled receptor member X3
MRGPRX4	Mas-related G protein-coupled receptor member X4
MTNR1A	Melatonin receptor type 1A
MTNR1B	Melatonin receptor type 1B
NIACR1	Niacin receptor 1
NIACR2	G protein-coupled receptor 109B
NMBR	Neuromedin-B receptor
NMUR1	Neuromedin-U receptor 1
NMUR2	Neuromedin-U receptor 2
NPBWR1	Neuropeptides B/W receptor type 1
NPBWR2	Neuropeptides B/W receptor type 2
NPFFR1	Neuropeptide FF receptor 1
NPFFR2	Neuropeptide FF receptor 2
NPSR1	Neuropeptide S receptor
NPY1R	Neuropeptide Y receptor type 1
NPY2R	Neuropeptide Y receptor type 2
NPY5R	Neuropeptide Y receptor type 5
NPY6R	Putative neuropeptide Y receptor type 6
NTSR1	Neurotensin receptor type 1
NTSR2	Neurotensin receptor type 2
OPN1LW	Red-sensitive opsin
OPN1MW	Green-sensitive opsin
OPN1SW	Blue-sensitive opsin
OPN3	Opsin-3
OPN4	Melanopsin
OPN5	Opsin-5
OPRD1	Delta-type opioid receptor
OPRK1	Kappa-type opioid receptor
OPRL1	Nociceptin receptor

OPRM1	Mu-type opioid receptor
OXER1	Oxoeicosanoid receptor 1
OXGR1	2-oxoglutarate receptor 1
OXTR	Oxytocin receptor
P2RY1	P2Y purinoceptor 1
P2RY10	Putative P2Y purinoceptor 10
P2RY11	P2Y purinoceptor 11
P2RY12	P2Y purinoceptor 12
P2RY13	P2Y purinoceptor 13
P2RY14	P2Y purinoceptor 14
P2RY2	P2Y purinoceptor 2
P2RY4	P2Y purinoceptor 4
P2RY5	Oleoyl-L-alpha-lysophosphatidic acid receptor
P2RY6	P2Y purinoceptor 6
P2RY8	P2Y purinoceptor 8
PPYR1	Neuropeptide Y receptor type 4
PRLHR	Prolactin-releasing peptide receptor
PROKR1	Prokineticin receptor 1
PROKR2	Prokineticin receptor 2
PTAFR	Platelet-activating factor receptor
PTGDR	Prostaglandin D2 receptor
PTGER1	Prostaglandin E2 receptor EP1 subtype
PTGER2	Prostaglandin E2 receptor EP2 subtype
PTGER3	Prostaglandin E2 receptor EP3 subtype
PTGER4	Prostaglandin E2 receptor EP4 subtype
PTGFR	Prostaglandin F2-alpha receptor
PTGIR	Prostacyclin receptor
PTH1R	Parathyroid hormone/parathyroid hormone-related peptide receptor
PTH2R	Parathyroid hormone 2 receptor
QRFPR	Pyroglutamylated RFamide peptide receptor
RGR	RPE-retinal G protein-coupled receptor
RHO	Rhodopsin
RRH	Visual pigment-like receptor peropsin
RXFP1	Relaxin receptor 1
RXFP2	Relaxin receptor 2
RXFP3	Relaxin-3 receptor 1
RXFP4	Relaxin-3 receptor 2
S1PR1	Sphingosine 1-phosphate receptor 1
S1PR2	Sphingosine 1-phosphate receptor 2
S1PR3	Sphingosine 1-phosphate receptor 3
S1PR4	Sphingosine 1-phosphate receptor 4
S1PR5	Sphingosine 1-phosphate receptor 5
SCTR	Secretin receptor
SMO	Smoothened homolog
SSTR1	Somatostatin receptor type 1
SSTR2	Somatostatin receptor type 2
SSTR3	Somatostatin receptor type 3
SSTR4	Somatostatin receptor type 4
SSTR5	Somatostatin receptor type 5
SUCNR1	Succinate receptor 1
TAAR1	Trace amine-associated receptor 1
TAAR2	Trace amine-associated receptor 2
TAAR3	Putative trace amine-associated receptor 3
TAAR5	Trace amine-associated receptor 5
TAAR6	Trace amine-associated receptor 6
TAAR8	Trace amine-associated receptor 8
TAAR9	Trace amine-associated receptor 9
TACR1	Substance-P receptor
TACR2	Substance-K receptor
TACR3	Neuromedin-K receptor
TBXA2R	Thromboxane A2 receptor

TRHR	Thyrotropin-releasing hormone receptor
TSHR	Thyrotropin receptor
UTS2R	Urotensin II receptor
VIPR1	Vasoactive intestinal polypeptide receptor 1
VIPR2	Vasoactive intestinal polypeptide receptor 2
VN1R1	Vomeronasal type-1 receptor 1
VN1R2	Vomeronasal type-1 receptor 2
VN1R3	Vomeronasal type-1 receptor 3
VN1R4	Vomeronasal type-1 receptor 4
VN1R5	Vomeronasal type-1 receptor 5
VN2R1P	Putative calcium-sensing receptor-like 1
VNRL4	Putative vomeronasal receptor-like protein 4
XCR1	Chemokine XC receptor 1

Table S7A. GPCR vs. non-GPCR variant type distribution in familial BD cohort - genes matched by Constraint Score.

		GPCR	Non-GPCR (Avg of 100)	Fold Change (GPCR/Non)	Chi Square (p-value)	
Missense	Deleterious	75	58.50	1.28		2.79E-03
Splicing		0	4.64	0.00		
Nonsense		3	0.92	3.26		
In/dels		3	2.52	1.19	**1.33E-03**	
UTR3	Non-deleterious	4	8.60	0.47		0.05
UTR5		1	6.03	0.17		
Synonymous		44	37.77	1.16		
Total		**130**	**119**	**1.09**		

Table S7B. GPCR vs. non-GPCR variant type distribution in familial BD cohort - genes matched by RVIS Score.

		GPCR	Non-GPCR (Avg of 100)	Fold Change (GPCR/Non)	Chi Square (p-value)	
Missense	Deleterious	75	58.76	1.28		6.03E-03
Splicing		0	3.92	0.00		
Nonsense		3	1.00	3.26		
In/dels		3	2.72	1.10	**1.56E-03**	
UTR3	Non-deleterious	4	9.09	0.44		0.03
UTR5		1	6.08	0.16		
Synonymous		44	35.85	1.23		
Total		**130**	**117**	**1.11**		

Supplemental Table 8: Statistical results comparing Non-linear curves for all figures. Data presented as distribution F-tests between WT and Mutant expressing cells where primarily EC50 values were compared. In relevant cases comparisons of distribution in the top of the curve are presented. Where no activation of the second messenger was detected statistical results were marked as n/a.

			Non-linear Curve EC50 comparison		Non-linear Curve Top comparison	
Experiment	**Gene**	**Figure**	**F (DFn, DFd)**	**p-value**	**F (DFn, DFd)**	**p-value**
ELISA	CRHR2-WT vs. CRHR2-R384X	Figure 2C	425.7 (1,57)	< 0.0001		
Gs second messenger BRET	CRHR2-WT vs. CRHR2-R384X	Figure 3A	6.172 (1,40)	0.0173	173.4 (1,40)	< 0.0001
Gi second messenger BRET	CRHR2-WT vs. CRHR2-R384X	Figure 3B	5.135 (1,38)	0.0292		
Gq second messenger BRET	CRHR2-WT vs. CRHR2-R384X	Figure 3C	no activation	n/a		
G12 second messenger BRET	CRHR2-WT vs. CRHR2-R384X	Figure 3D	no activation	n/a		
G13 second messenger BRET	CRHR2-WT vs. CRHR2-R384X	Figure 3E	no activation	n/a		
B-Arrestin second messenger BRET	CRHR2-WT vs. CRHR2-R384X	Figure 3F	0.3433 (1,40)	0.5612	51.56 (1,40)	< 0.0001
ELISA	GRM1-WT vs. GRM1-D508E	Figure 4C	no activation	n/a		
Gq second messenger BRET	GRM1-WT vs. GRM1-D508E	Figure 4D	no activation	n/a		
Gs second messenger BRET	GRM1-WT vs. GRM1-D508E	Figure 4E	no activation	n/a		
Gi second messenger BRET	GRM1-WT vs. GRM1-D508E	Figure 4F	no activation	n/a		
G12 second messenger BRET	GRM1-WT vs. GRM1-D508E	Figure 4G	no activation	n/a		
G13 second messenger BRET	GRM1-WT vs. GRM1-D508E	Figure 4H	no activation	n/a		
B-Arrestin second messenger BRET	GRM1-WT vs. GRM1-D508E	Figure 4I	no activation	n/a		
GTPyS BRET	CRHR2-WT vs. CRHR2-R384X	Supp. Figure 7	8.705 (1,137)	0.0037		
cAMP BRET	CRHR2-WT	Supp. Figure 8A	0.1974 (1,68)	0.6582		
cAMP BRET	CRHR2-R384X	Supp. Figure 8B	13.36 (1,64)	0.0005		
cAMP BRET	GRM1-WT	Supp. Figure 10A	no activation	n/a		
cAMP BRET	GRM1-D508E	Supp. Figure 10B	no activation	n/a		
PKC BRET	GRM1-WT vs. GRM1-D508E	Supp. Figure 10C-D	0.6176 (1,40)	0.4366		

Appendix 3: Supplemental material: "Transcriptome Sequencing of the Anterior Cingulate in Bipolar Disorder: Dysregulation of G Protein-Coupled Receptors"

Supplemental Methods

Post-mortem brain samples

Postmortem brain tissue was obtained from the Douglas-Bell Canada Brain Bank (www.douglasbrainbank.ca). This facility collects brains from subjects who died by suicide as well as from psychiatrically healthy control subjects. Once a family accepts to make a donation, a series of interviews known as psychological autopsies (295) are carried out, whereby information is obtained by means of structured interviews on psychiatric history (Axis I and Axis II), psychological traits, development, life events and history of trauma/abuse. These lengthy interviews are then complemented by information from medical charts, police and coroner records. In addition, extensive demographic and medical information is collected which includes history of medical treatment (300, 301). Psychological autopsies were performed post-mortem on both cases and controls by a panel of psychiatrists and diagnoses were assigned based on DSM-IV criteria. The control group was composed of individuals who died suddenly from accidental causes or myocardial infarction, and could not have undergone any resuscitation procedures or other type of medical intervention. Controls had no history of psychopathology, including suicidal behavior or major mood or psychotic disorders (**Supplementary Table 1**). Brains were rapidly preserved upon arriving at the Brain Bank, and the left hemisphere was cut into consecutive 1 cm-thick coronal sections that were snap-frozen and stored at -80˚C. Dissections from thick frozen sections were performed on dry ice, following well-established anatomic landmarks. Specifically, grey matter was dissected from the dACC,

adjacent to the dorsal part of the genu of the corpus callosum (BA24) (302, 303). The anterior region immediately dorsal to the genu of the corpus callosum was located as shown by Hersher *et al.* in Figure 1 (304) and 1-cm^3 tissue blocks were removed while maintaining the tissue on dry ice until RNA extraction was performed. Cases in this study were individuals who had a diagnosis of BD type I or type II (N = 13). Controls (N = 13) had neither current nor past psychiatric diagnoses. Cases and controls were matched for refrigeration delay, age, brain pH, and RNA integrity number, and there were no group differences in these variables. Refrigeration delay refers to the difference between the estimated time of death (determined by the pathologist through external body examination details) and the time at which the brain was refrigerated.

High throughput transcriptome sequencing

Total RNA was extracted from brain tissue sections using the RNeasy system (Qiagen). RNA quality and concentrations were measured on a Nanodrop 2000 Spectrophotometer and an Agilent 2100 Bioanalyzer. In order to maintain the long non-coding RNA fraction that does not contain a poly(A) tail, we selected RNA for sequencing using ribosomal depletion. A starting amount of 4ug total RNA was used according to the RiboZero (Epicentre) protocol for ribosomal depletion. Briefly, the total RNA was incubated with ribosomal (rRNA) sequence-specific 5'-biotin labeled oligonucleotide probes. Following probe hybridization, the rRNA/probe complex was removed from the sample with streptavidin-coated magnetic beads, leaving behind only 10-20% of the total RNA fraction. This fraction was used to create RNAseq libraries following the TruSeq dUTP degradation-based directional protocol (Illumina). All sequencing for this project was

carried out at the Genome Quebec Innovation Center using the Illumina HiSeq 2000 platform. In order to achieve adequate coverage of the transcriptome including lowly expressed transcript variants, one library was sequenced per lane (**Supplementary Table 2**). Throughout the library preparation a randomization process was used to ensure that no batch effects were generated. Briefly, we identified four batching stages each with different samples per batch: ribosomal depletion (n=6), library preparation (n=8) and sequencing flow cell (n=8). We randomized the samples in each batch and then tested correlations with possible confounders: Diagnosis, pH Value, Post-mortem Delay, Gender, Age, Cause of Death and RIN. This ensured that there would be no batch effects going into the experiment.

Bioinformatics analyses

Alignment. Following high-throughput sequencing, 100bp paired-end reads were aligned to the human genome reference (hg19) using TopHat v2.0.8b (http://tophat.cbcb.umd.edu/) (168) with a mate insert distance of 75 bp (-r) and library type fr-firststrand. Those reads that passed mapping quality of at least 50 were used for gene and transcript quantification.

Quantification. Gene annotations were assembled by combining the annotations from the Illumina iGenomes UCSC (hg19) which corresponds to Ensembl annotations downloaded on March 6, 2013 (Ensembl release 70) (https://support.illumina.com/sequencing/ sequencing_software/igenome.ilmn). SmallRNA annotation files were downloaded from miRBase release 19. Additional

lincRNA annotations were obtained from the lincRNA catalog stringent set downloaded on Sep 20 2013 (http://www.broadinstitute.org/ genome_bio/human_lincrnas/sites/default/files/lincRNA_catalog/lincRNAs_transcripts_stringentSet.gtf) (305).

For gene-level quantification we used HTSeq-count version 0.5.4p1 (http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html) to count fragments that overlap genes identified through the annotations described (169). HTSeq-count was ran with the intersection-nonempty mode and reverse strand parameters for each sample and the results were combined to form a count matrix of 60,905 transcribed RNAs across 26 samples (**Supplementary Figure 1)**. As validation, we also ran Cufflinks v2.1.1 (http://cufflinks.cbcb.umd.edu/) (93) to count fragments at the gene as well as transcript level using the same gene annotation files as for HTSeq-count with parameters --multi-read-correct and --library-type fr-firststrand for each sample and the results were combined to form a FPKM matrix of 60,327 transcribed RNAs across 26 samples (**Supplementary Figure 1)**. Due to differences between the tools' counting algorithms, the FPKM matrix was approximated to a count matrix where gene lengths were obtained by summing the exon length for each gene using the hg19 ensGene table in the GenomicFeatures_1.12.2 R package (http://www.bioconductor.org/packages/release/bioc/html/ GenomicFeatures.html). The library size for each sample was estimated using the number of mapped reads in the BAM file using 'samtools view -c' (http://samtools.sourceforge.net/) command.

Differential expression analysis. All whole-transcript and isoform matrices were analyzed separately. For each transcript, we summed the mapped fragments across all samples. We

removed those transcripts with no mapped fragments. In addition, those transcripts whose total is greater than 34.7 million mapped fragments (or 1% of the total for all the transcripts) were also removed. Fragments were normalized across libraries by using the weighted trimmed mean of log expression ratios (TMM) from the edgeR v3.0.8 R package (170). Furthermore, genes and isoforms with low counts were removed by keeping only those which have counts of at least 0.2 CPM (counts per million) in at least 8 samples per group. Counts were corrected for heteroscedasticity by employing voom from the limma v3.14.4 R package (171). The linear model used to fit the data included diagnosis, postmortem interval (PMI) and RNA integrity number (RIN) as covariates. Gene annotations were incorporated using the biomaRt v2.14.0 R package (http://www.bioconductor.org/packages/release/bioc/html/biomaRt.html).

External validation cohort analysis. We obtained RNAseq data (98) from the Stanley Neuropathology Consortium Integrative Database (SNCID) Array Collection consisting of 61 thoroughly characterized samples (BD=26, CTRL=35) from the anterior cingulate cortex described previously (306). Fragments were aligned to the human (hg19) reference genome using STAR_2.4.0h (307) fragments were mapped to genes using featureCounts from the subread-1.4.6 package (308, 309) with a minimum quality of 50. Ensembl gene annotations were obtained from the Illumina iGenomes UCSC (hg19) (Ensembl release 70, March/6/2013) (https://support.illumina.com/sequencing/sequencing_software/ igenome.ilmn). Gene expression levels were normalized using the TMM method followed by employing voom from the limma v3.14.4 R package (171, 307).

Comparison to dorsolateral prefrontal cortex (DLPFC) external dataset. In order to compare our results with those of one previous transcriptome sequencing study in BD

(96), raw count expression matrices deposited by Akula *et al.* were obtained from the Gene Expression Omnibus (GEO, GSE53239). Expression matrices from the two platforms described by *Akula et al*. (NISC1 and NISC2) were combined and batch-corrected by removing the first principal component. The first principal component contributed to the 20% of the variance and the scores were significantly different between the two platforms ($P < 0.001$, t-test). The list of differentially expressed transcripts was identified by applying the same procedure used for our data. We performed an over-representation analysis by compiling the list of downregulated transcripts ($p<0.01$) from one study and calculating the AUC against the entire list of downregulated transcript p-values from the other study. We repeated this analysis for upregulated transcripts. ROC curves were plotted with the pROC_1.7.2 R package (310).

Neural Progenitor Cell lines chronic drug treatment experiments

Cell culture and treatments. Human neural progenitor cells (NPCs) derived from induced pluripotent stem cell (iPSC) line GM08330 obtained from a healthy male and previously characterized (174), were generously provided by Dr. Stephen Haggarty. NPCs were maintained on culture plates coated with 200µg/ml Poly-L-ornithine hydrobromide (Sigma) and 5mg/ml laminin (Sigma) and maintained in media with 70% DMEM (Invitrogen), 30% Ham's F12 (Mediatech), 1x penicillin/streptomycin (Invitrogen) and supplemented with B-27 (Invitrogen). During expansion cells were grown in media containing 20ng/ml of human EGF (Sigma), FGF (R&D Systems) and 5µg/ml heparin (Sigma). To induce neural differentiation, cells were allowed to reach 90% confluence before growth factors were removed. At this point chronic (1 week) treatments were performed with drugs commonly prescribed in bipolar disorder: lithium, valproic acid,

and carbamazepine. To find adequate drug concentrations for treatment, cells were screened for cytotoxic effects by measuring the activity of mitochondrial dehydrogenase using the 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide (MTT) assay (Sigma-Aldrich Co) with three different concentrations tested for each drug in accordance with the literature and estimates of the correspondence to clinical treatment levels in patients. For lithium, concentrations of 0.5M, 1.0M, and 2.0M were tested. For valproic acid, concentrations of 0.5M, 1.0M, and 2.0M were tested. For carbamazepine, concentrations of 25mM, 50mM, and 100mM were tested. No significant toxicity was detected at any concentration, thus, cells were treated with 1.0 mM lithium chloride (Li), 1.0 mM valproic acid (VPA), 50uM carbamazepine (CBZ), or no-drug control for one week, after which cell pellets were collected and RNA was extracted. All experiments were performed in triplicate.

Immunohistochemistry. In order to validate the neuronal and astroglial properties of neural progenitor cell lines, we performed immunohistochemistry with neuron-specific and astrocyte-specific markers MAP2 and GFAP respectively. Cover slips were washed 3 times for 5 minutes in TBS + 0.05% tween and incubated for 20 minutes in a solution of 1% bovine serum albumin (BSA) and 0.2% Triton in PBS. This was followed by an hour pre-incubation in a solution of 1% BSA in PBS containing 5% Normal goat serum (NGS) before being transferred one hour in the same solution containing anti- Map2 (1:1, 0000, Abcam, Cambridge, MA, USA) and GFAP (1:1000, Dako, Burlington, ON, CA)) antibodies for two hours. Cover slips were then incubated 2h with secondary goat anti-rabbit antibody coupled to Texas red (1:1000, Abcam, Cambridge, MA, USA) and a donkey anti-mouse antibody coupled with the florophore FITC (1:1000, Vector

Laboratories Inc., Burlington, ON, Canada). Sections were mounted on glass slides, and coverslipped with Prolong gold with Dapi (Fisher Scientific Inc., Pittsburgh, PA, USA).

Quantitative Real-time Polymerase Chain Reaction (qRT-PCR)

Total RNA was extracted from frozen brain tissue using the RNeasy Lipid Tissue Mini Kit (Qiagen) and from frozen cell pellets using the RNeasy Mini Kit (Qiagen). Brain RNA for validation of RNASeq results was from the same original extraction. Synthesis of cDNA was performed in triplicated, using M-MLV reverse transcriptase (Gibco, Burlington, Ontario) along with oligo(dT)16 primers (Invitrogen) and random hexamers (IDT DNA) in a 1:1 ratio. Real-time PCR reactions were run in quadruplicate using an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) and the iTaq Universal SYBR Green Supermix (BioRad). Extensive characterization of all SYBR Green assays was undertaken to ensure single-product specificity and efficiency compatibility with endogenous controls (Data not shown). All primer sequences and reaction parameters are available upon request. Relative expression was calculated using the relative quantitation method (ΔΔCt) in the RQ Manager 1.2 software (Applied Biosystems). We investigated the stability of various endogenous genes prior to performing qRT-PCR experiments in each sample set and determined the most suitable endogenous gene using the NormFinder Algorithm (176) (**Supplemental Table 3**). All qRT-PCR experiments were reported with POLR2A (Polymerase (RNA) II (DNA directed) polypeptide A) or ACTB (Beta Actin) as endogenous control.

Supplemental Figures

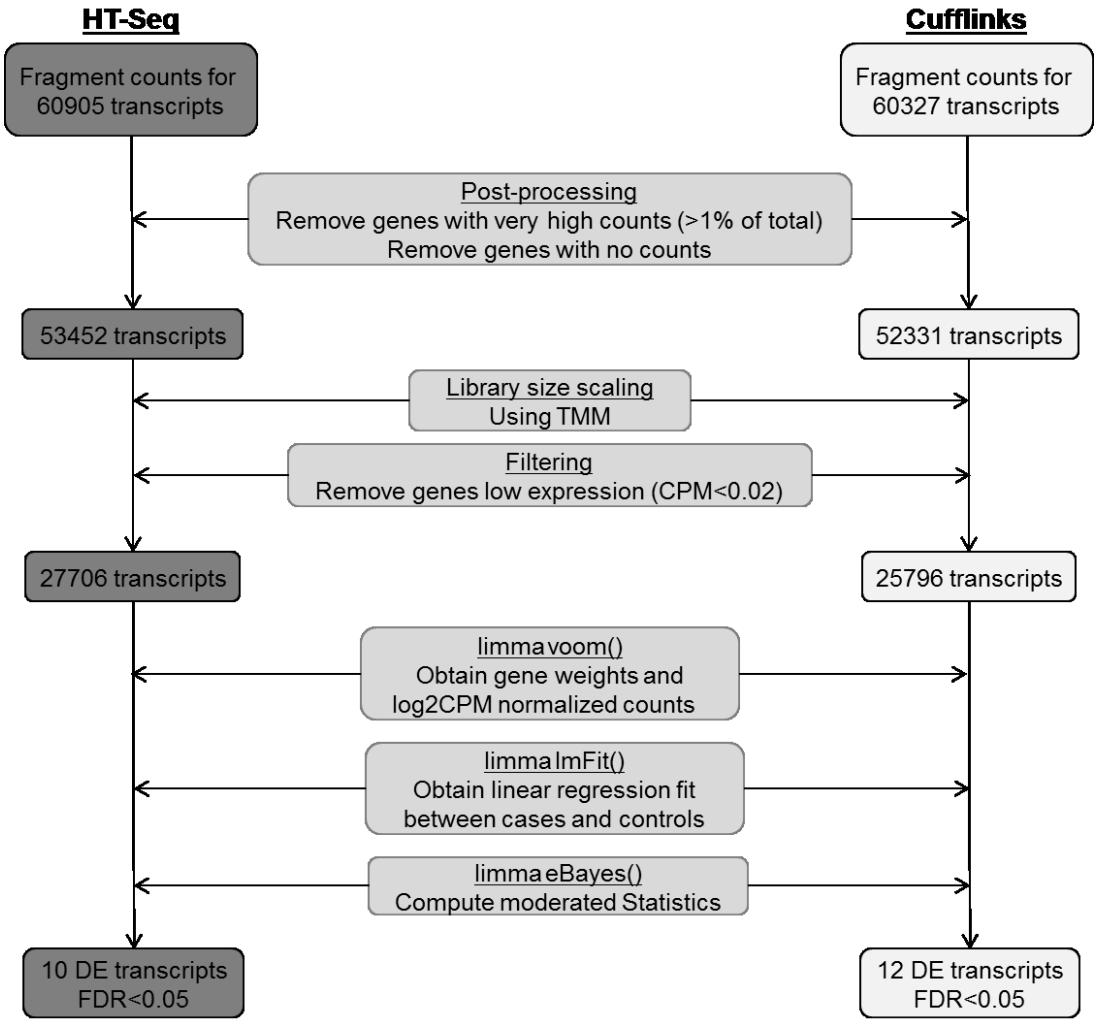


Figure S1: Gene-level differential expression bioinformatics analyses – consistency between two methods.

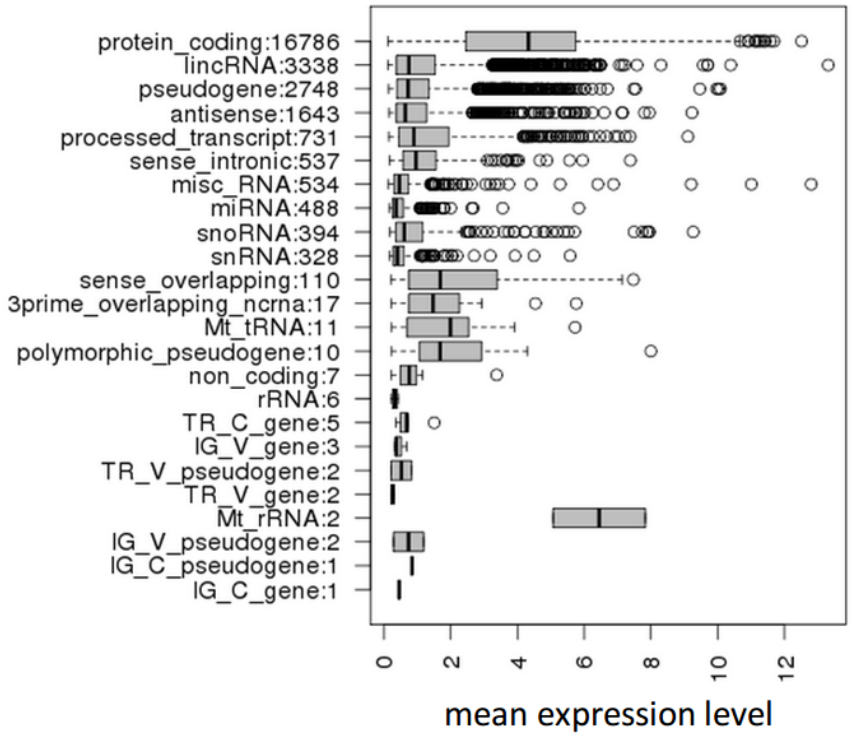


Figure S2: Mean expression statistics for the different RNA classes after filtering. The numbers after each RNA class indicate the number of genes that belong to that class.

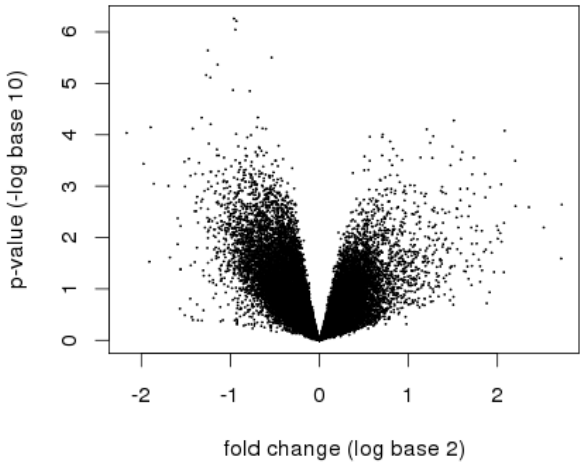


Figure S3A: Volcano plot that shows the overall transcript differential expression for all genes from the HT-Seq analysis. There is a stronger effect for downregulated (negative fold change) as opposed to upregulated genes (positive fold change).

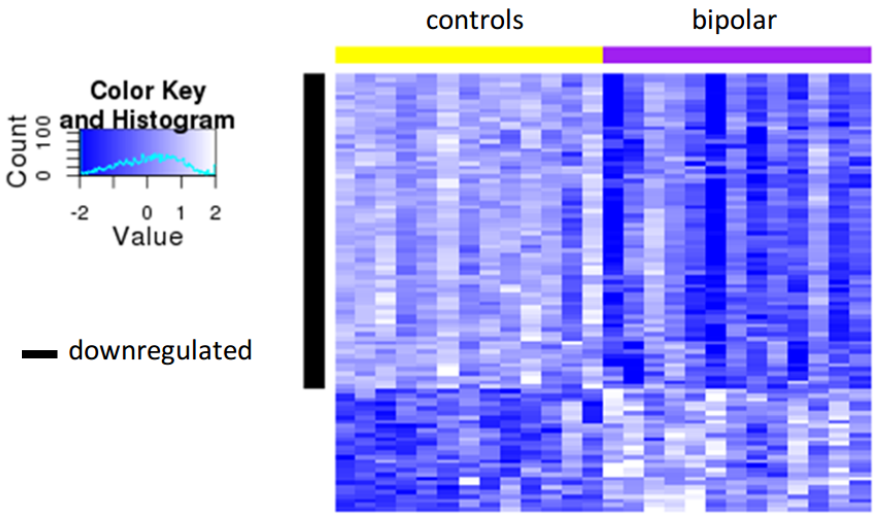


Figure S3B: Hierarchical clustering of the top 100 transcripts (ranked by increasing p-value) across controls and bipolar samples, 72 of which are downregulated. Expression levels have been mean centered and normalized.

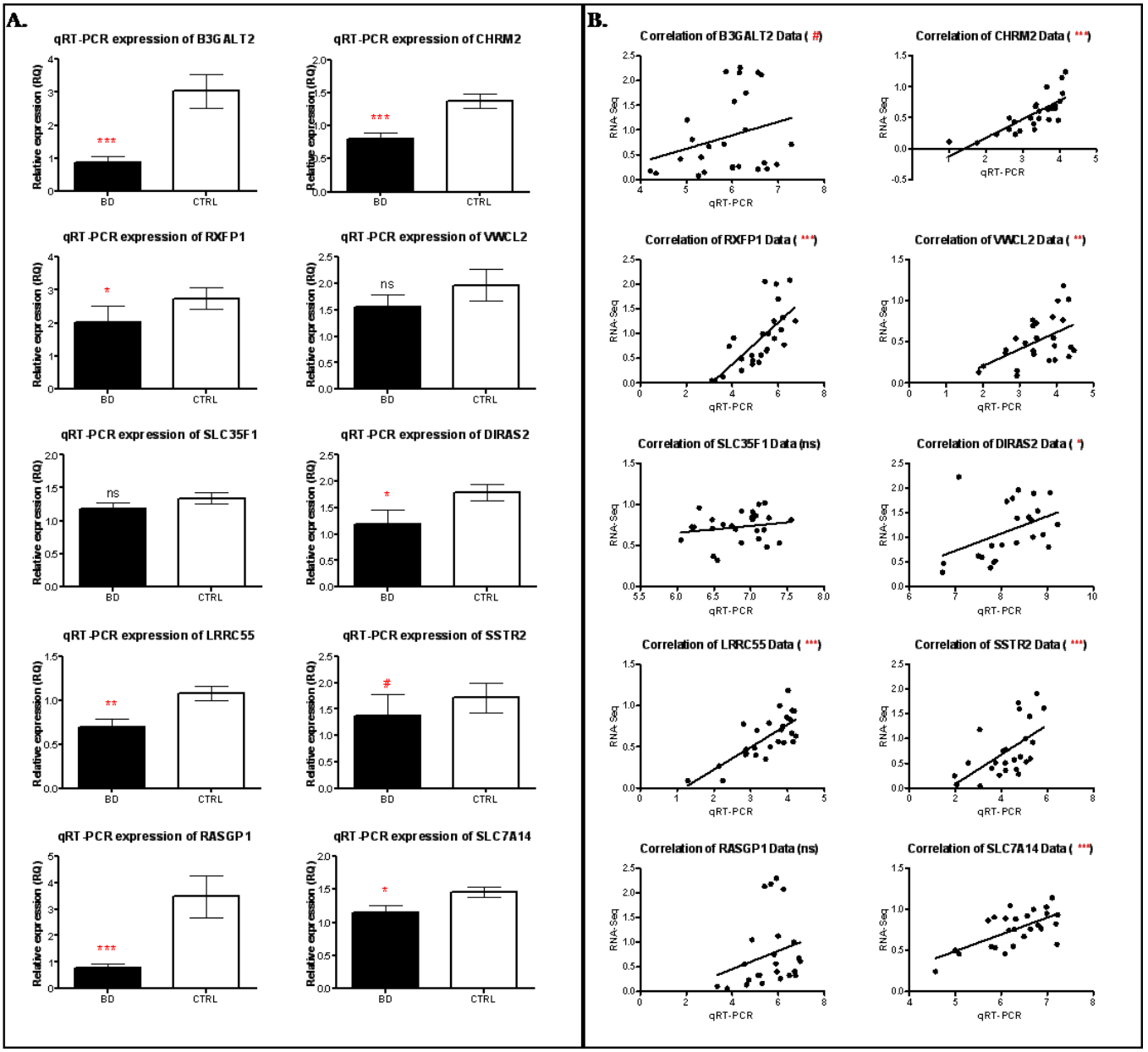


Figure S4: A: qRT-PCR results in BA24. **B**: Correlations between RNASeq and qRT-PCR expression values.

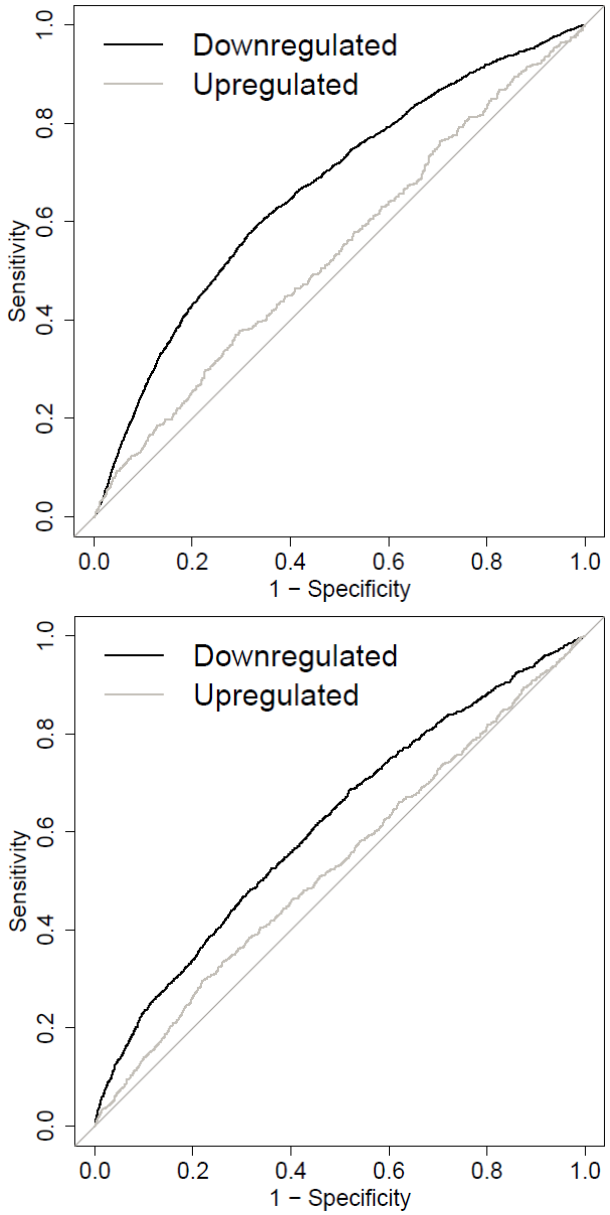


Figure S5: A. Left, Downregulated transcripts tended to be downregulated in a separate study. Top downregulated transcripts (p-value < 0.01) are overrepresented among the top downregulated transcripts in Akula *et al.* **B. Right**, Top downregulated transcripts from Akula et al. (p-value < 0.01) are also overrepresented among the top downregulated transcripts.

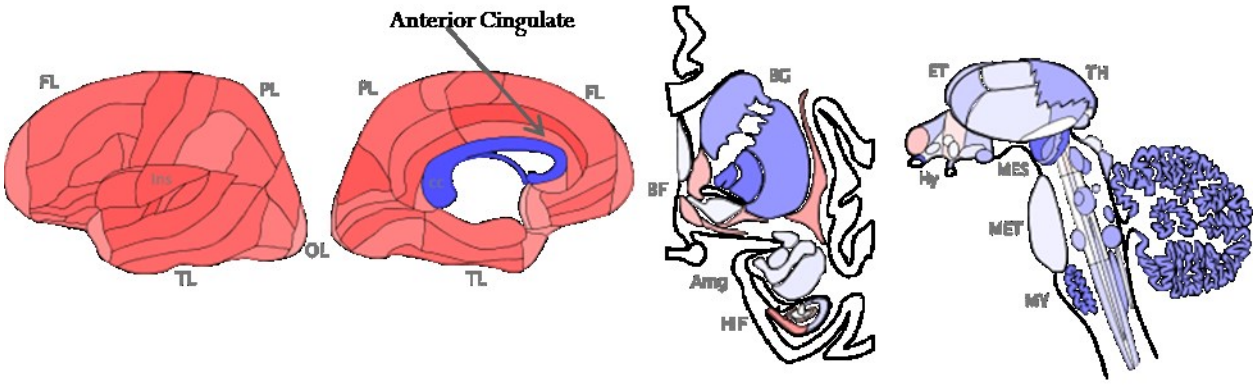


Figure S6: HBAset gene set enrichment of the ten candidate genes considered as a group. The brain regions considered are shown schematically (right to left) for the lateral and medial surfaces of the cortex; basal ganglia and deep temporal lobe; and midbrain, hindbrain and cerebellum at left. Red indicates enrichment, blue indicates "de-enrichment". Image generated by HBASet.

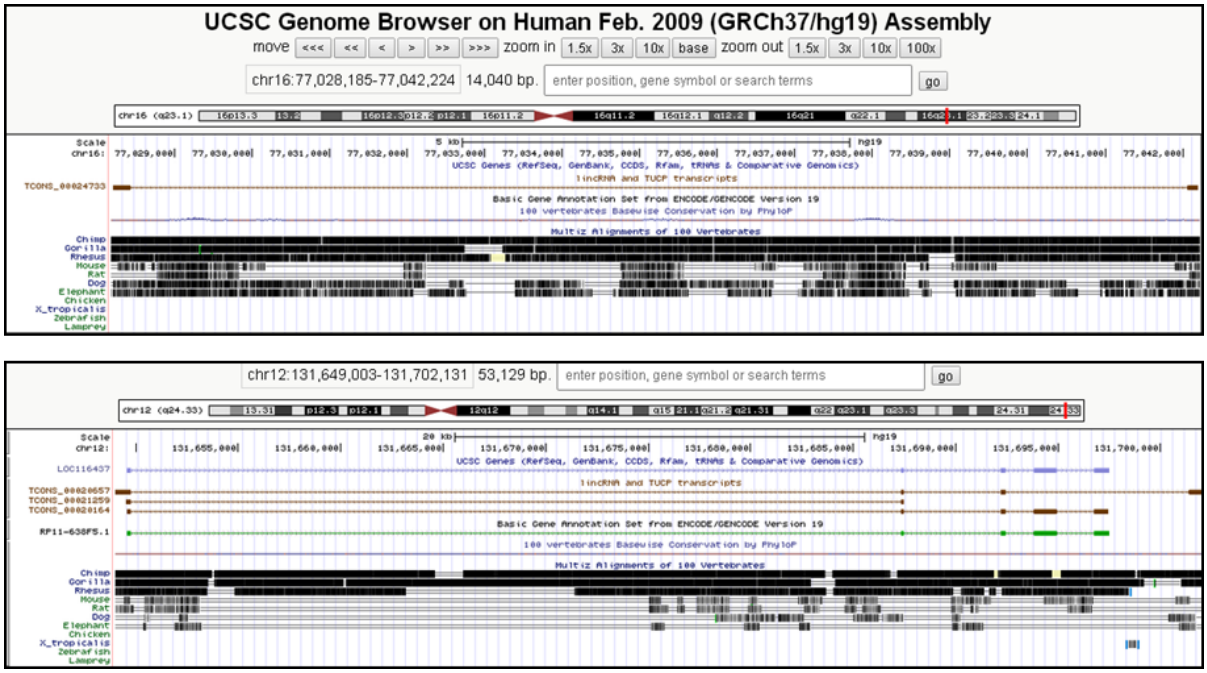


Figure S7: A. Top, Genomic location of linc-KARS-3 (also known as TCONS_0024733). **B. Bottom**, Genomic location of the Chr12 ncRNA locus. linc-SFSWAP-3 (also known as TCONS_0021259) and RP11-638F5.1 (also known as TCONS_0020164) share two exons. Due to the exon-intron distribution at this locus, a qRT-PCR assay (**Figure 3**) specific to linc-SFSWAP-3 could not be designed.

Supplemental Tables

Table S1 and Table S2

Table S1: Demographics for RNAseq Study. Brain sample gender ratios and group demographics means (presented as mean±SEM) for BD (bipolar disorder) and controls (CTRL). Fisher's exact tests and two-tailed t-tests showed that there are no significant differences between groups for any of these variables.

Status	Gender	Age	Post-Mortem Delay	Brain pH	RIN	Race	Method of death	Smoking	Alcohol toxicology
BD	9 M / 4 F	44.00 ± 4.05	30.38 ± 6.31	6.63 ± 0.07	5.7 ± 0.25	12 White Caucasian/ 1 Asian	12 Suicide/ 1 Accidental	6 Non-s/ 6 Smoker/ 1	3 positive/ 10 NA
CTRL	11 M / 2 F	40.15 ± 6.43	23.58 ± 5.31	6.60 ± 0.06	5.9 ± 0.13	13 White Caucasian	7 Accidental / 6 Natural	8 Non-s/ 2 Smoker/ 3	2 positive/ 11 NA
Group dif	ns	ns	ns	ns	ns	ns	ns	ns	ns

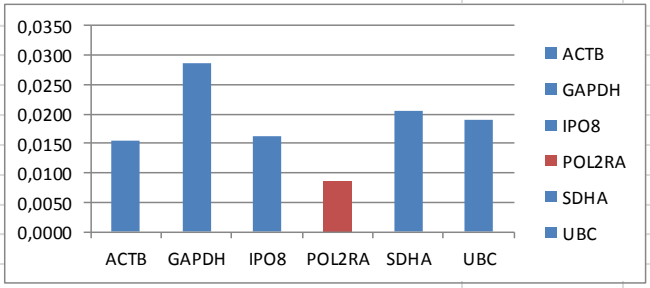
Table S2: RNA-seq sequencing quality statistics

	Number of Bases	Average Quality	Total aligned reads	Reads aligned (mapq ≥50)	Fragments mapped to a gene	Number of fragments with no feature
Average	37 240 575 315	32	391 389 863	276 055 408	69 791 000	67 473 000
Min	26 528 481 800	29	243 938 727	168 996 569	45 527 000	45 125 000
Max	43 572 794 600	34	502 743 638	349 224 334	97 733 000	102 670 000

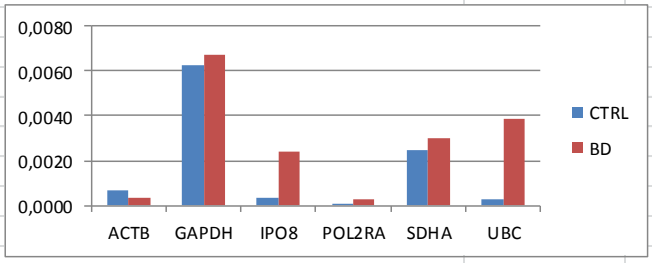
Table S3: Selection of endogenous gene for qRT-PCR analysis using the NormFinder algorithm.

Best gene	POL2Ra
Stability value	0,0086
Best combination of two genes	POL2RA and UBC
Stability value for best combination of two genes	0,0085

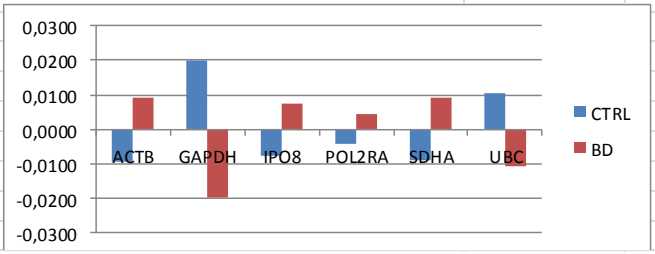
Gene name	**Stability value**
ACTB	0,0154
GAPDH	0,0287
IPO8	0,0163
POL2RA	0,0086
SDHA	0,0206
UBC	0,0191



Intragroup variation	CTRL	BD
Group identifier	2	1
ACTB	0,0007	0,0003
GAPDH	0,0062	0,0067
IPO8	0,0003	0,0024
POL2RA	0,0001	0,0003
SDHA	0,0025	0,0030
UBC	0,0003	0,0038



Intergroup variation	CTRL	BD
Group identifier	2	1
ACTB	-0,0093	0,0093
GAPDH	0,0199	-0,0199
IPO8	-0,0076	0,0076
POL2RA	-0,0044	0,0044
SDHA	-0,0091	0,0091
UBC	0,0105	-0,0105



TableS4 (Rows 1-50 of 4000+): HTSeq-BA24-Genes

ID	logFC	FC	AveExpr	P.Value	adj.P.Val	gene_name	chr	start	end	strand	exon_len	gene_biotype
ENSG00000162630	-0,96117	0,5136402	5,895223	5,49E-07	0,008299	B3GALT2	chr1	193148175	193155784	-	3274	protein_coding
ENSG00000181072	-0,93347	0,5235967	3,274726	6,10E-07	0,008299	CHRM2	chr7	136553416	136705002	+	8322	protein_coding
ENSG00000174453	-0,94513	0,5193833	3,494275	8,99E-07	0,008299	VWC2L	chr2	215275789	215443683	+	5230	protein_coding
ENSG00000171509	-1,25355	0,419416	5,16441	2,29E-06	0,01586	RXFP1	chr4	159236463	159574524	+	5654	protein_coding
ENSG00000196376	-0,53743	0,6889994	6,842219	3,14E-06	0,017416	SLC35F1	chr6	118228689	118638839	+	4852	protein_coding
ENSG00000172575	-1,1435	0,4526602	5,675151	4,32E-06	0,019942	RASGRP1	chr15	38780304	38857776	-	6090	protein_coding
ENSG00000180616	-1,27247	0,4139496	4,287914	6,91E-06	0,026701	SSTR2	chr17	71161151	71167185	+	2265	protein_coding
ENSG00000165023	-1,22479	0,4278606	7,93277	7,71E-06	0,026701	DIRAS2	chr9	93372114	93405386	-	4385	protein_coding
ENSG00000183908	-0,97132	0,5100394	3,419871	1,35E-05	0,039257	LRRC55	chr11	56949221	56959191	+	5410	protein_coding
ENSG00000013293	-0,78236	0,581416	6,308857	1,42E-05	0,039257	SLC7A14	chr3	170182353	170303863	-	5478	protein_coding
ENSG00000151079	-0,69283	0,6186412	5,24168	4,61E-05	0,10741	KCNA6	chr12	4918342	4960277	+	5977	protein_coding
ENSG00000158258	-1,32379	0,3994832	6,863653	4,67E-05	0,10741	CLSTN2	chr3	139654027	140286919	+	5276	protein_coding
ENSG00000184779	1,509091	2,8463063	-1,33647	5,29E-05	0,10741	RPS17	chr15	82821158	82824972	-	3815	protein_coding
XLOC_012014	-1,2217	0,4287782	-0,3699	6,23E-05	0,10741	linc-KARS-3	chr16	77028205	77042204	-	373	NA
ENSG00000145545	-0,70226	0,6146105	4,417428	7,10E-05	0,10741	SRD5A1	chr5	6633456	6669675	+	2891	protein_coding
XLOC_009957	-1,89756	0,2683976	-2,44613	7,14E-05	0,10741	linc-SFSWAP-3	chr12	131649003	131702131	+	3766	NA
ENSG00000164619	-0,63896	0,6421746	4,79133	7,53E-05	0,10741	BMPER	chr7	33944523	34195484	+	5889	protein_coding
ENSG00000261179	-1,42397	0,3726853	6,795147	7,60E-05	0,10741	RP11-13L2.4	chr3	140290548	140296239	+	5692	sense_overlapping
ENSG00000113361	-0,60168	0,658984	5,429288	7,69E-05	0,10741	CDH6	chr5	31193857	31329253	+	12379	protein_coding
ENSG00000069011	1,205601	2,3063338	-2,07157	7,83E-05	0,10741	PITX1	chr5	134362615	134370503	-	4321	protein_coding
ENSG00000239731	2,079604	4,2269109	-2,52932	8,35E-05	0,10741	Metazoa_SRP	chr10	32281018	32281309	-	292	misc_RNA
ENSG00000204603	-2,16733	0,2226222	-2,47476	9,22E-05	0,10741	RP11-638F5.1	chr12	131649556	131697476	+	2562	lincRNA
ENSG00000177519	-0,93586	0,5227296	2,404516	9,65E-05	0,10741	RPRM	chr2	154333852	154335322	-	1471	protein_coding
ENSG00000164106	0,711408	1,6374013	5,799231	9,87E-05	0,10741	SCRG1	chr4	174309299	174327531	-	1949	protein_coding
ENSG00000095596	-0,76997	0,5864305	1,056976	0,000102	0,10741	CYP26A1	chr10	94833232	94837647	+	2485	protein_coding
ENSG00000108691	1,279232	2,4270971	0,567342	0,000107	0,10741	CCL2	chr17	32582237	32584222	+	1986	protein_coding
ENSG00000089159	0,566408	1,4808322	4,077544	0,000109	0,10741	PXN	chr12	120648250	120703574	-	7452	protein_coding
ENSG00000133874	0,704903	1,6300355	0,332056	0,000111	0,10741	RNF122	chr8	33405273	33424643	-	1868	protein_coding
ENSG00000151025	-0,81002	0,5703725	6,51317	0,000112	0,10741	GPR158	chr10	25463991	25891155	+	7610	protein_coding
ENSG00000245532	0,793831	1,7336726	9,563369	0,000135	0,124871	NEAT1	chr11	65190245	65213011	+	22767	lincRNA
ENSG00000239899	1,243862	2,368316	3,116584	0,000144	0,128426	Metazoa_SRP	chr2	11724899	11725176	+	278	misc_RNA
ENSG00000236841	-1,24325	0,4224191	-0,99567	0,000149	0,128704	AC007750.5	chr2	163018280	163029426	+	3694	antisense
ENSG00000144057	-0,45937	0,7273034	6,229791	0,000164	0,129027	ST6GAL2	chr2	107418056	107503564	-	7708	protein_coding
ENSG00000153820	-0,91175	0,5315409	6,43476	0,000168	0,129027	SPHKAP	chr2	228844666	229046361	-	7009	protein_coding
ENSG00000257058	1,493456	2,8156265	-3,08577	0,00017	0,129027	RP11-864I4.4	chr11	62313471	62315171	+	391	antisense
ENSG00000074590	-0,67968	0,6243016	7,429975	0,000174	0,129027	NUAK1	chr12	106457118	106533811	-	7008	protein_coding
ENSG00000166250	-0,67015	0,6284427	1,800775	0,000182	0,129027	CLMP	chr11	122943035	123065989	-	2635	protein_coding
ENSG00000162636	-0,83206	0,5617248	6,392874	0,000185	0,129027	FAM102B	chr1	109102711	109187522	+	9355	protein_coding
ENSG00000232150	0,826539	1,7734258	-1,70037	0,000188	0,129027	ST13P4	chr13	50746225	50747317	+	1093	pseudogene
ENSG00000260248	-1,15182	0,4500574	2,443684	0,000193	0,129027	RP11-143K11.1	chr17	71171622	71172772	+	1151	lincRNA
ENSG00000250305	-0,65141	0,6366595	6,120821	0,000206	0,129027	KIAA1456	chr8	12803151	12889012	+	14668	protein_coding
ENSG00000253151	-0,71749	0,6081544	3,161234	0,000217	0,129027	RP11-628E19.3	chr8	56438745	56446511	+	865	lincRNA
ENSG00000243562	1,603826	3,0394824	-2,08061	0,000218	0,129027	Metazoa_SRP	chr11	440406	440693	-	288	misc_RNA
ENSG00000165966	-0,59548	0,661823	4,256131	0,000223	0,129027	PDZRN4	chr12	41582250	41968392	+	4586	protein_coding
ENSG00000182752	-0,69683	0,6169248	0,7028	0,000245	0,129027	PAPPA	chr9	118916083	119164601	+	11573	protein_coding
ENSG00000185518	-1,12969	0,457013	8,61567	0,000252	0,129027	SV2B	chr15	91643180	91844539	+	12454	protein_coding
ENSG00000154133	0,703606	1,6285707	2,79628	0,000252	0,129027	ROBO4	chr11	124753587	124768396	-	8121	protein_coding
ENSG00000251621	-1,15078	0,4503816	0,79855	0,000253	0,129027	AC009487.5	chr2	162280526	162285285	+	560	processed_transcript
ENSG00000082482	-0,62661	0,6476959	4,055485	0,000253	0,129027	KCNK2	chr1	215179118	215410436	+	3910	protein_coding

TableS5 (Rows 1-50 of 4000+): Cufflinks-BA24-Genes

ID	logFC	FC	AveExpr	P.Value	adj.P.Val	gene_nam	chr	start	end	strand	exon_leng	gene_biotype
ENSG00000162630	-0,96658	0,511716	4,576052	3,68E-07	0,007361	B3GALT2	chr1	193148175	193155784	-	3274	protein_coding
ENSG00000174453	-0,96883	0,510921	2,38326	5,71E-07	0,007361	VWC2L	chr2	215275789	215443683	+	5230	protein_coding
ENSG00000196376	-0,59865	0,660373	5,41456	9,44E-07	0,008118	SLC35F1	chr6	118228689	118638839	+	4852	protein_coding
ENSG00000013293	-0,69966	0,615716	5,121528	4,08E-06	0,020412	SLC7A14	chr3	170182353	170303863	-	5478	protein_coding
ENSG00000165023	-1,26044	0,417416	6,518672	4,23E-06	0,020412	DIRAS2	chr9	93372114	93405386	-	4385	protein_coding
ENSG00000171509	-1,29182	0,408435	4,581476	4,75E-06	0,020412	RXFP1	chr4	159236463	159574524	+	5654	protein_coding
ENSG00000172575	-1,03968	0,486435	4,874315	8,50E-06	0,031341	RASGRP1	chr15	38780304	38857776	-	6090	protein_coding
ENSG00000180616	-1,32262	0,399809	3,159663	1,42E-05	0,041743	SSTR2	chr17	71161151	71167185	+	2265	protein_coding
ENSG00000183908	-0,84555	0,556497	2,099353	1,46E-05	0,041743	LRRC55	chr11	56949221	56959191	+	5410	protein_coding
ENSG00000181072	-0,84494	0,556732	2,905614	1,64E-05	0,042385	CHRM2	chr7	136553416	136705002	+	8322	protein_coding
ENSG00000158258	-1,38028	0,384143	5,596402	2,15E-05	0,047787	CLSTN2	chr3	139654027	140286919	+	5276	protein_coding
ENSG00000151079	-0,73083	0,602556	3,983128	2,22E-05	0,047787	KCNA6	chr12	4918342	4960277	+	5977	protein_coding
ENSG00000145545	-0,7846	0,580512	3,262328	3,71E-05	0,073715	SRD5A1	chr5	6633456	6669675	+	2891	protein_coding
ENSG00000236841	-1,1127	0,462427	-2,52514	4,20E-05	0,077427	AC007750.	chr2	163018280	163029426	+	3694	antisense
ENSG00000177519	-0,97743	0,507884	1,188633	5,84E-05	0,100417	RPRM	chr2	154333852	154335322	-	1471	protein_coding
ENSG00000261179	-1,41904	0,373962	5,305104	7,11E-05	0,10709	RP11-13L2	chr3	140290548	140296239	+	5692	sense_overlapping
ENSG00000095596	-0,81839	0,567074	-0,31265	7,38E-05	0,10709	CYP26A1	chr10	94833232	94837647	+	2485	protein_coding
ENSG00000260248	-1,19819	0,435821	1,091551	7,67E-05	0,10709	RP11-143k	chr17	71171622	71172772	+	1151	lincRNA
ENSG00000204603	-1,80486	0,28621	-0,71377	8,13E-05	0,10709	RP11-638F	chr12	131649556	131697476	+	2562	lincRNA
ENSG00000153820	-0,9449	0,519464	4,987785	9,27E-05	0,10709	SPHKAP	chr2	228844666	229046361	-	7009	protein_coding
ENSG00000177600	1,215552	2,322296	7,6247	9,52E-05	0,10709	RPLP2	chr11	809647	812880	+	1955	protein_coding
XLOC_011183	-0,97759	0,507826	8,943447	9,93E-05	0,10709	linc-GABR	chr15	25247918	25281705	+	18011	NA
ENSG00000089159	0,624803	1,542001	3,820395	0,0001	0,10709	PXN	chr12	120648250	120703574	-	7452	protein_coding
ENSG00000151025	-0,84561	0,556474	5,312076	0,000108	0,10709	GPR158	chr10	25463991	25891155	+	7610	protein_coding
ENSG00000135750	-0,96845	0,511054	5,229359	0,000108	0,10709	KCNK1	chr1	233749750	233808258	+	3542	protein_coding
ENSG00000135324	-0,88483	0,54155	2,542705	0,00011	0,10709	MRAP2	chr6	84743475	84800600	+	2153	protein_coding
ENSG00000239731	2,067146	4,190568	-2,44493	0,000112	0,10709	Metazoa_	chr10	32281018	32281309	-	292	misc_RNA
ENSG00000184779	1,43599	2,705678	0,469222	0,000116	0,107167	RPS17	chr15	82821158	82824972	-	3815	protein_coding
ENSG00000206384	-0,73239	0,601906	0,941591	0,000129	0,107237	COL6A6	chr3	130279178	130396999	+	9733	protein_coding
ENSG00000118946	-0,5175	0,698579	6,392425	0,000131	0,107237	PCDH17	chr13	58205944	58303445	+	8242	protein_coding
ENSG00000239899	1,308106	2,476163	3,297013	0,000133	0,107237	Metazoa_	chr2	11724899	11725176	+	278	misc_RNA
ENSG00000101290	-0,40759	0,753881	7,558833	0,000133	0,107237	CDS2	chr20	5107432	5178533	+	11893	protein_coding
ENSG00000250305	-0,66761	0,629547	6,527246	0,000153	0,108637	KIAA1456	chr8	12803151	12889012	+	14668	protein_coding
ENSG00000163873	-0,70474	0,613553	5,321023	0,000157	0,108637	GRIK3	chr1	37261128	37499730	-	10111	protein_coding
ENSG00000153234	-1,00512	0,498228	2,117121	0,000164	0,108637	NR4A2	chr2	157180944	157198860	-	3967	protein_coding
ENSG00000162636	-0,8881	0,540324	5,675245	0,000165	0,108637	FAM102B	chr1	109102711	109187522	+	9355	protein_coding
ENSG00000258384	1,536425	2,900749	-0,53594	0,000166	0,108637	AC068831.	chr15	91495469	91498455	-	1858	antisense
XLOC_010514	-0,87586	0,544929	-0,00992	0,000178	0,108637	linc-CDC1(	chr13	114567162	114569790	+	1985	NA
ENSG00000185518	-1,09659	0,467621	8,15982	0,000179	0,108637	SV2B	chr15	91643180	91844539	+	12454	protein_coding
ENSG00000254531	-0,62256	0,649517	1,225614	0,00018	0,108637	AP001816.	chr4	102268937	102270040	+	790	protein_coding
ENSG00000171724	-0,95776	0,514855	4,77451	0,000195	0,108637	VAT1L	chr16	77822427	78014004	+	4137	protein_coding
ENSG00000185477	-0,59066	0,664038	3,270929	0,0002	0,108637	GPRIN3	chr4	90165429	90229161	-	6352	protein_coding
ENSG00000113361	-0,6532	0,635868	5,086784	0,000201	0,108637	CDH6	chr5	31193857	31329253	+	12379	protein_coding
ENSG00000263911	1,510615	2,849315	-1,94895	0,00021	0,108637	Metazoa_	chr20	43509479	43509775	-	297	misc_RNA
ENSG00000120833	-0,58365	0,667275	3,658073	0,00021	0,108637	SOCS2	chr12	93963590	93977263	+	6776	protein_coding
ENSG00000175175	-0,83519	0,560509	3,242135	0,00021	0,108637	PPM1E	chr17	56833230	57058983	+	2988	protein_coding
ENSG00000175906	-0,96581	0,511989	1,241434	0,000222	0,108637	ARL4D	chr17	41476327	41478492	+	1599	protein_coding
ENSG00000253719	-0,46062	0,726674	5,810747	0,000226	0,108637	ATXN7L3B	chr12	74931551	74935223	+	3673	protein_coding
ENSG00000133083	-0,80562	0,572117	8,432597	0,000231	0,108637	DCLK1	chr13	36345478	36705443	-	11609	protein_coding

Table S6: Validation in an external cohort obtained from the Stanley Neuropathology Consortium Integrative Database (SNCID) Collection consisting of 61 samples (BD=26, CTRL=35) from the anterior cingulate cortex.

Gene info		RNAseq HT-Seq				External Validation		
Gene	Description	Rank	FC	P-Value	Adjusted P-Value	FC	P-Value	P-value Sig.
B3GALT2	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2	1	0,51364	5,49E-07	0,0083	0,8666	0,062372	#
CHRM2	cholinergic receptor, muscarinic 2	2	0,523597	6,10E-07	0,0083	0,7782	0,047624	*
VWC2L	von Willebrand factor C domain containing protein 2-like	3	0,519383	8,99E-07	0,0083	0,8118	0,03395	*
RXFP1	relaxin/insulin-like family peptide receptor 1	4	0,419416	2,29E-06	0,0159	0,7911	0,006974	**
SLC35F1	solute carrier family 35, member F1	5	0,688999	3,14E-06	0,0174	0,8959	0,01303	*
RASGRP1	RAS guanyl releasing protein 1 (calcium and DAG-regulated)	6	0,45266	4,32E-06	0,0199	0,8022	0,007153	**
SSTR2	somatostatin receptor 2	7	0,41395	6,91E-06	0,0267	0,7587	0,012301	*
DIRAS2	DIRAS family, GTP-binding RAS-like 2	8	0,427861	7,71E-06	0,0267	0,7990	0,000181	**
LRRC55	leucine rich repeat containing 55	9	0,510039	1,35E-05	0,0393	0,6714	0,00225	**
SLC7A14	solute carrier family 7 (orphan transporter), member 14	10	0,581416	1,42E-05	0,0393	0,7431	0,014632	*

TableS7 (Rows 1-50 or 5000+): ermineJ-GO-HTSeqBA24

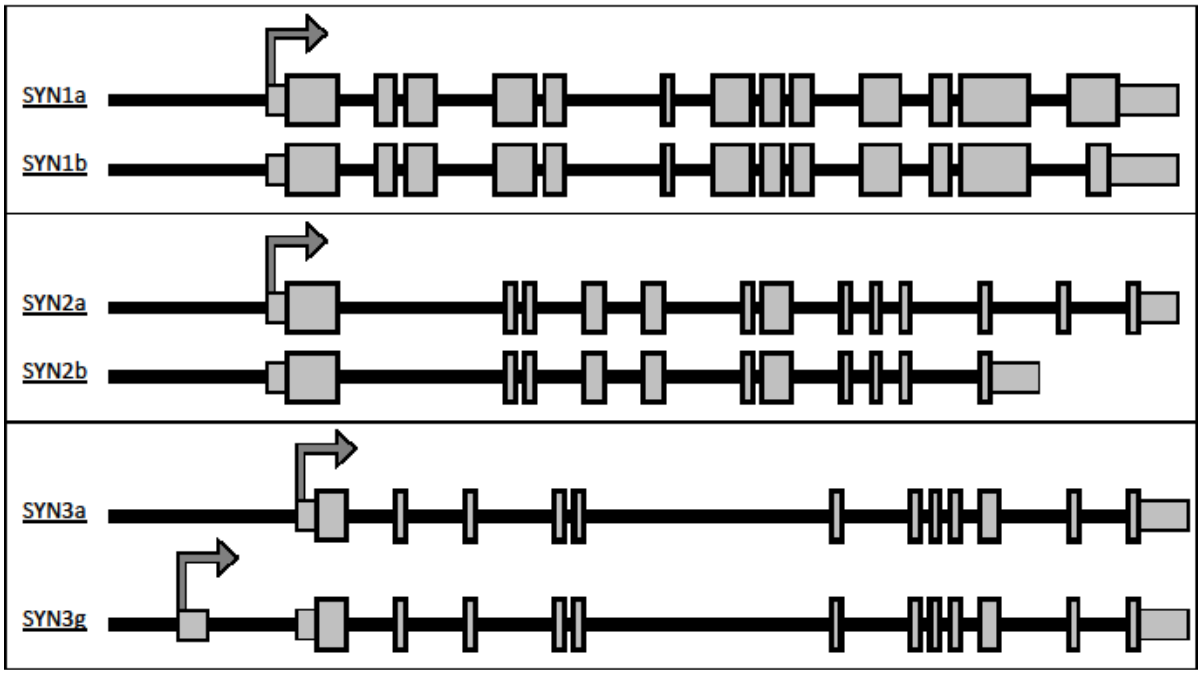
Name	ID	NumProbes	NumGenes	RawScore	Pval	CorrectedPvalue	MFPvalue	CorrectedMFPvalue	Multifunctionality	Same as	GeneMembers
G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messeng	GO:0007187	123	122	0,02816829	1,00E-12	5,24E-09	1,00E-04	0,5236	0,953		ADCY1|ADCY2|ADCY3|ADCY4|ADCY5|ADCY6|ADCY7|ADCY8|ADCY9|ADCYAP1|ADM2|ADOR
adenylate cyclase-modulating G-protein coupled receptor signaling pathway	GO:0007188	91	90	0,02841462	1,00E-12	2,62E-09	1,00E-04	0,2618	0,958		ADCY1|ADCY2|ADCY3|ADCY4|ADCY5|ADCY6|ADCY7|ADCY8|ADCY9|ADCYAP1|ADM2|ADOR
carbohydrate biosynthetic process	GO:0016051	220	220	0,02377319	4,00E-04	0,69813333	5,00E-04	0,87266667	0,943		ABCC5|ACADM|ACAN|AGL|AGRN|AKT1|AKT2|ALDOA|ALDOB|ALDOC|ALG1|ALG10|ALG10
G-protein coupled acetylcholine receptor signaling pathway	GO:0007213	11	11	0,04762129	9,00E-04	1	1,00E-03	1	0,599		ADRBK1|AGRN|CDKSR1|CHRM1|CHRM2|CHRM4|CHRM5|GNA15|GNAI2|GNB1|PLCB1|
monovalent inorganic cation transport	GO:0015672	192	192	0,02127609	1,90E-03	1	1,90E-03	1	0,913		ABCC8|ABCC9|ANK1|AQP1|ASIC1|ASIC2|ASIC3|ASIC4|ATP1A2|ATP1A4|ATP1B4|ATP4A|AT
regulation of neurological system process	GO:0031644	201	201	0,02093689	1,90E-03	1	9,30E-03	1	0,992		ABHD6|ACHE|ADCYAP1|ADIPOQ|ADORA1|ADORA2A|ADRA1A|ADRA2C|AGT|ANAPC2|APO
regulation of transmission of nerve impulse	GO:0051969	189	189	0,02022386	1,90E-03	1	9,30E-03	1	0,992		ABHD6|ACHE|ADCYAP1|ADIPOQ|ADORA1|ADORA2A|ADRA1A|AGT|ANAPC2|APOE|ARC|A
phospholipase C-activating G-protein coupled receptor signaling pathway	GO:0007200	53	53	0,01725157	2,50E-03	1	3,60E-03	1	0,86		ADRA1A|ADRA2A|AGT|AGTR1|CCKAR|CCKBR|CHRM1|CHRM2|CRHR1|CXCR2|DRD1|DRD2|
peristalsis	GO:0030432	8	8	0,02012637	2,60E-03	1	3,10E-03	1	0,678		AGT|DLG1|DRD1|DRD2|GDNF|P2RX2|SSTR2|TSHZ3|
negative regulation of BMP signaling pathway	GO:0030514	34	34	0,01857237	3,10E-03	1	4,70E-03	1	0,762		BMPER|CAV1|CER1|CHRD|DAND5|DKK1|FBN1|FSTL3|FZD1|GREM1|HIPK2|HTRA1|HTRA3|L
parturition	GO:0007567	15	15	0,0239492	3,20E-03	1	1,60E-03	1	0,677		CCL2|CD55|CRH|CRHR1|CYP1A1|EDN1|EDNRA|HPGD|MAFF|OXTR|PLA2G4B|PLA2G4C|PTG
oligosaccharide biosynthetic process	GO:0009312	44	44	0,02579917	3,30E-03	1	2,30E-03	1	0,659		ALG1|ALG10|ALG10B|ALG11|ALG12|ALG13|ALG14|ALG2|ALG3|ALG5|ALG6|ALG8|ALG9|B3G
mast cell activation involved in immune response	GO:0002279	11	11	0,01625389	3,50E-03	1	3,70E-03	1	0,605	GO:0043303|mast cell degranulation,	ADORA3|CPLX2|KIT|LAT|LAT2|LYN|PIK3CD|PIK3CG|RASGRP1|S100A13|YWHAZ|
mast cell degranulation	GO:0043303	11	11	0,01625389	3,50E-03	1	3,70E-03	1	0,605	GO:0002279|mast cell activation involved in immune response,	ADORA3|CPLX2|KIT|LAT|LAT2|LYN|PIK3CD|PIK3CG|RASGRP1|S100A13|YWHAZ|
regulation of BMP signaling pathway	GO:0030510	58	58	0,01457543	3,70E-03	1	8,90E-03	1	0,87		ACVR2A|ACVRL1|BMP4|BMPER|BMPR2|CAV1|CER1|CHRD|CYR61|DAND5|DKK1|ENG|FBN1
cellular response to estradiol stimulus	GO:0071392	9	9	0,01695279	3,70E-03	1	3,10E-03	1	0,604		AQP4|CRHBP|IL10|MSX2|SFRP1|SSTR1|SSTR2|SSTR3|TNFRSF1A|
cellular cation homeostasis	GO:0030003	299	298	0,02599925	3,90E-03	1	0,0273	1	0,968		ABCB6|ABCB7|ABCG2|ACO1|ADCYAP1|ADM|ADRA1A|AGT|AGTR1|ALAS2|ANK2|ANXA7|A
oligosaccharide metabolic process	GO:0009311	66	66	0,02105716	4,30E-03	1	4,20E-03	1	0,761		ALG1|ALG10|ALG10B|ALG11|ALG12|ALG13|ALG14|ALG2|ALG3|ALG5|ALG6|ALG8|ALG9|B3G
cognition	GO:0050890	161	161	0,01909151	4,70E-03	1	0,0104	1	0,983		AAAS|ABI2|ADCY1|ADCY8|ADORA1|ADRA1B|ADRB1|AFF2|AMPH|APBB1|APP|ARC|ASIC1|
mast cell mediated immunity	GO:0002448	12	12	0,01513717	5,30E-03	1	4,40E-03	1	0,589		ADORA3|CPLX2|KIT|LAT|LAT2|LYN|PIK3CD|PIK3CG|RASGRP1|S100A13|SERPINB9|YWHAZ|
phasic smooth muscle contraction	GO:0014821	14	14	0,0124832	5,50E-03	1	6,00E-03	1	0,778		AGT|DLG1|DRD1|DRD2|EDN1|EDN3|EDNRB|GDNF|HTR1D|HTR2B|P2RX2|SSTR2|TACR2|TSH
vesicle transport along microtubule	GO:0047496	16	16	0,01433987	6,10E-03	1	5,90E-03	1	0,632		CLN3|DYNC1I1|FYCO1|HAP1|HTT|KIF13A|KIF3A|KIF3B|KIF5B|KIFAP3|MAP2K1|NDE1|NDEL
inflammatory response to antigenic stimulus	GO:0002437	15	15	0,01204632	6,30E-03	1	6,80E-03	1	0,695		AHCY|AK7|GATA3|HLA-DRB1|HLA-DRB5|HMGB1|IL1RN|IL20RB|IL5RA|NOTCH1|NOTCH2|NR
blood vessel endothelial cell proliferation involved in sprouting angiogenesis	GO:0002043	5	5	0,01364997	6,30E-03	1	5,10E-03	1	0,332		ACVRL1|BMP4|BMPER|ITGB1BP1|NRARP|
endothelial cell activation	GO:0042118	5	5	0,01369578	6,30E-03	1	5,10E-03	1	0,24		APOLD1|BMPER|P2RX4|PRMT5|SMAD4|
regulation of synaptic transmission	GO:0050804	169	169	0,01927424	6,30E-03	1	0,0181	1	0,992		ABHD6|ACHE|ADCYAP1|ADIPOQ|ADORA1|ADORA2A|ADRA1A|AGT|ANAPC2|APOE|ARC|A
cellular response to estrogen stimulus	GO:0071391	14	14	0,01150663	7,00E-03	1	7,30E-03	1	0,694		AQP4|CRHBP|ESR1|IL10|MDM2|MSX2|RARA|SERPINB9|SFRP1|SSTR1|SSTR2|SSTR3|TNFRSF
mast cell activation	GO:0045576	16	16	0,01213354	7,20E-03	1	8,80E-03	1	0,574		ADORA3|CD48|CPLX2|FCER1G|KIT|LAT|LAT2|LCP2|LYN|NDRG1|PIK3CD|PIK3CG|RASGRP1|
homophilic cell adhesion	GO:0007156	133	133	0,01628813	7,60E-03	1	7,50E-03	1	7,99E-05		AMIGO1|AMIGO2|CADM1|CADM3|CD84|CDH1|CDH10|CDH11|CDH12|CDH13|CDH15|CDH1
locomotory behavior	GO:0007626	135	135	0,01601077	7,60E-03	1	0,0114	1	0,971		ABAT|ADAM22|ADCY5|ADORA2A|ADRA1B|AGTPBP1|ALS2|ANKH|APBA1|APBA2|APLP2|AI
astrocyte cell migration	GO:0043615	5	5	0,01077758	7,70E-03	1	8,00E-03	1	0,329		APCDD1|CCL2|CCL3|HEXB|MMP14|
learning or memory	GO:0007611	148	148	0,01756549	8,20E-03	1	0,0137	1	0,986		AAAS|ABI2|ADCY1|ADCY8|ADRA1B|ADRB1|AFF2|AMPH|APBB1|APP|ARC|ASIC1|ATAD1|A
cellular metal ion homeostasis	GO:0006875	257	256	0,02226245	8,50E-03	1	0,034	1	0,97		ABCB6|ABCB7|ABCG2|ACO1|ADCYAP1|ADM|ADRA1A|AGT|AGTR1|ALAS2|ANK2|ANXA7|A
glycoprotein biosynthetic process	GO:0009101	259	259	0,02232127	8,50E-03	1	0,034	1	0,912		A4GALT|ABO|ALG1|ALG10|ALG10B|ALG11|ALG12|ALG13|ALG14|ALG2|ALG3|ALG5|ALG6|AI
forebrain development	GO:0030900	274	273	0,02276116	8,50E-03	1	0,034	1	0,99		ADCYAP1|AGTPBP1|ALDH1A2|ALDH1A3|ANKS1B|APAF1|APLP1|APLP2|APP|AQP1|ARHGAF
regulation of membrane potential	GO:0042391	258	258	0,02240617	8,50E-03	1	0,034	1	0,98		ABCB5|ACSBG1|ACTN2|ADAM22|ADCYAP1|ADIPOQ|ADORA1|ADORA2A|ADRA1A|ADRB1|A
nerve growth factor receptor signaling pathway	GO:0048011	265	264	0,02272008	8,50E-03	1	0,034	1	0,948		AATF|ABR|ADAM17|ADCY1|ADCY2|ADCY3|ADCY4|ADCY5|ADCY6|ADCY7|ADCY8|ADCY9|AE
organelle transport along microtubule	GO:0072384	31	31	0,01046764	8,70E-03	1	8,10E-03	1	0,765		ARHGAP21|BICD1|BICD2|CDC42|CLN3|COPG1|DYNC1I1|FYCO1|HAP1|HTT|KIF13A|KIF1B|KI
regulation of heart contraction	GO:0008016	118	116	0,01477603	9,10E-03	1	0,0196	1	0,998		ADA|ADM|ADORA1|ADORA3|ADRA1A|ADRA1B|ADRA1D|ADRB1|ADRBK1|ANK2|APLN|ASI
positive regulation of neuron differentiation	GO:0045666	56	56	0,01119033	9,20E-03	1	0,0115	1	0,855		ACTR3|ADRA2B|ADRA2C|ASCL1|BDNF|BMP2|BMP4|BMP6|BMP7|CDON|DAB1|DMD|DUOX
cellular response to glucocorticoid stimulus	GO:0071385	16	16	0,01002494	9,50E-03	1	0,0106	1	0,703		ACVR1|ADCYAP1|ANXA1|AQP1|ARG1|CASP9|CRH|REST|SSTR2|SSTR3|SSTR4|STAR|STC1|T
ribosomal small subunit assembly	GO:0000028	9	9	0,00941952	9,90E-03	1	7,70E-03	1	0,296		ERAL1|RPL38|RPS14|RPS15|RPS17|RPS19|RPS25|RPS6|RPSA|
neurotransmitter transport	GO:0006836	98	97	0,01302967	9,90E-03	1	0,0108	1	0,818		ABAT|ALDH5A1|ATP1A2|BAIAP3|BLOC1S6|BRSK1|CADPS|CADPS2|CDK5|CLN8|CPLX1|CPLX
negative regulation of transmembrane receptor protein serine/threonine kinase signaling p	GO:0090101	93	93	0,01363981	9,90E-03	1	0,0108	1	0,956		ACVR1|ADAMTSL2|ASPN|BAMBI|BCL9L|BMPER|CAV1|CAV2|CER1|CHRD|CHST11|CIDEA|D
potassium ion transport	GO:0006813	88	88	0,0131038	1,00E-02	1	8,80E-03	1	0,643		ABCC8|ABCC9|AQP1|ATP1A2|ATP1A4|CDK2|CDKN1B|CHP1|HCN1|HPN|KCNA1|KCNA2|KC
adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway	GO:0007193	42	41	0,01000946	0,0101	1	0,0142	1	0,748		ADCY1|ADCY2|ADCY3|ADCY4|ADCY5|ADCY6|ADCY7|ADCY8|ADCY9|ADORA1|ADRA2A|APIT
negative regulation of leukocyte chemotaxis	GO:0002689	6	6	0,00886829	0,0105	1	9,90E-03	1	0,649		C5|C5AR2|CCL2|GREM1|NBL1|SLIT2|
regulation of glial cell apoptotic process	GO:0034350	6	6	0,00887661	0,0105	1	9,90E-03	1	0,653	GO:0034351|negative regulation of glial cell apoptotic process,	CCL2|GAS6|PRKCA|PRKCD|PRKCH|PRKCI|
negative regulation of glial cell apoptotic process	GO:0034351	6	6	0,00887661	0,0105	1	9,90E-03	1	0,653	GO:0034350|regulation of glial cell apoptotic process,	CCL2|GAS6|PRKCA|PRKCD|PRKCH|PRKCI|
maternal process involved in parturition	GO:0060137	6	6	0,00841456	0,0105	1	0,0145	1	0,619		CCL2|CD55|CYP1A1|EDN1|EDNRA|OXTR|

TableS8 (Rows 1-50 of 600+): HBAset

name	donors	AUC	pValue	EnclosingRegion	childrenC
Choroid Plexus of the lateral ventricle	3	0,118587522	0,001277182	NaN	0
superior occipital gyrus, Right, superior bank of gyı	2	0,853653957	0,001302273	Occipital Lobe	0
postcentral gyrus, Right, superior lateral aspect of	2	0,853755055	0,001302273	Parietal Lobe	0
Cingulate gyrus, frontal part, Left, superior bank of	6	0,854207587	0,001302273	Limbic Lobe	0
Long Insular Gyri, Left	6	0,855911804	0,001302273	Insula	0
planum polare, Right	2	0,857370499	0,001302273	Temporal Lobe	0
supraparietal lobule, Right, superior bank of gyrus	2	0,859666859	0,001302273	Parietal Lobe	0
Heschl's gyrus, Left	6	0,865236857	0,001302273	Temporal Lobe	0
superior frontal gyrus, Right, medial bank of gyrus	2	0,842513961	0,001353901	Frontal Lobe	0
precentral gyrus, Right, bank of the central sulcus	2	0,843029078	0,001353901	Frontal Lobe	0
planum temporale, Right	2	0,843212016	0,001353901	Temporal Lobe	0
postcentral gyrus, Right, inferior lateral aspect of g	2	0,843645292	0,001353901	Parietal Lobe	0
inferior temporal gyrus, Right, lateral bank of gyrus	2	0,845758714	0,001353901	Temporal Lobe	0
superior temporal gyrus, Right, lateral bank of gyru	1	0,84676969	0,001353901	Temporal Lobe	0
precentral gyrus, Right, superior lateral aspect of g	2	0,847679569	0,001353901	Frontal Lobe	0
Heschl's gyrus, Right	2	0,849744849	0,001353901	Temporal Lobe	0
cuneus, Left, peristriate	6	0,839081456	0,001580016	Occipital Lobe	0
precentral gyrus, Right, bank of the precentral sulc	2	0,837122087	0,001683664	Frontal Lobe	0
Short Insular Gyri, Left	6	0,83547564	0,00176331	Insula	0
corpus callosum	6	0,165790487	0,001808197	NaN	0
Inferior frontal gyrus, triangular part, Right	2	0,832052763	0,001869845	Frontal Lobe	0
paracentral lobule, anterior part, Left, superior bar	5	0,832524552	0,001869845	Frontal Lobe	0
Cingulate gyrus, parietal part, Left, inferior bank of	5	0,824797805	0,001959049	Limbic Lobe	0
precuneus, Right, inferior lateral bank of gyrus	2	0,825298479	0,001959049	Parietal Lobe	0
inferior frontal gyrus, orbital part, Right	2	0,825394762	0,001959049	Frontal Lobe	0
middle temporal gyrus, Right, superior bank of gyr	2	0,825673984	0,001959049	Temporal Lobe	0
superior temporal gyrus, Right, inferior bank of gyı	2	0,825688427	0,001959049	Temporal Lobe	0
postcentral gyrus, Right, bank of the central sulcus	2	0,82595802	0,001959049	Parietal Lobe	0
transverse gyri, Left	5	0,827647795	0,001959049	Temporal Lobe	0
precentral gyrus, Left, inferior lateral aspect of gyr	6	0,828480647	0,001959049	Frontal Lobe	0
precentral gyrus, Right, inferior lateral aspect of gy	2	0,828629886	0,001959049	Frontal Lobe	0
middle frontal gyrus, Right, superior bank of gyrus	2	0,828817639	0,001959049	Frontal Lobe	0
superior frontal gyrus, Left, medial bank of gyrus	6	0,820532448	0,001991146	Frontal Lobe	0
planum polare, Left	6	0,820729829	0,001991146	Temporal Lobe	0
superior temporal gyrus, Left, lateral bank of gyrus	6	0,820821298	0,001991146	Temporal Lobe	0
precentral gyrus, Left, bank of the central sulcus	5	0,821572309	0,001991146	Frontal Lobe	0
Long Insular Gyri, Right	2	0,821808203	0,001991146	Insula	0
postcentral gyrus, Left, bank of the central sulcus	6	0,822867321	0,001991146	Parietal Lobe	0
inferior temporal gyrus, Right, bank of mts	2	0,823127287	0,001991146	Temporal Lobe	0
paracentral lobule, anterior part, Left, inferior ban	6	0,823714616	0,001991146	Frontal Lobe	0
lateral orbital gyrus, Right	2	0,819608126	0,001995753	Frontal Lobe	0
Cingulate gyrus, frontal part, Right, superior bank o	2	0,819844021	0,001995753	Limbic Lobe	0
angular gyrus, Right, inferior bank of gyrus	2	0,817013287	0,002246797	Parietal Lobe	0
lingual gyrus, Left, peristriate	6	0,815564221	0,002322024	Occipital Lobe	0
fusiform gyrus, Right, bank of the its	2	0,815573849	0,002322024	Temporal Lobe	0
inferior rostral gyrus, Right	2	0,814230695	0,0024401	Frontal Lobe	0
middle frontal gyrus, Right, inferior bank of gyrus	2	0,811149624	0,00244742	Frontal Lobe	0
frontal operculum, Left	6	0,811245908	0,00244742	Frontal Lobe	0
precentral gyrus, Left, bank of the precentral sulcu	6	0,811674369	0,00244742	Frontal Lobe	0
paracentral lobule, anterior part, Right, superior ba	2	0,811736954	0,00244742	Frontal Lobe	0

Appendix 4: Supplemental material: "H3K4 tri-methylation in synapsin genes leads to different expression patterns in bipolar disorder and major depression"

Supplemental Materials



Supplemental Figure 1: Gene structures of adult brain-expressed synapsins. The *SYN1a* and *SYN1b* variants are identical in all but the 3'end where the last exon of *SYN1a* is longer and the 3'UTR is shorter. The *SYN2a* and *SYN2b* variants are identical except in all but the 3'end where *SYN2a* has two extra coding exons and a completely different 3'UTR. The *SYN3a* and *SYN3g* variants are identical in the coding regions but *SYN3g* has an additional non-coding exon in the 5'end and consequently the two variants have distinct promoters.

Supplemental Table 1: Correlations between possible confounding factors and RQ expression values relative to *GAPDH* as an endogenous control. Linear regression was used for gender and Pearson's tests were used for age, pH, post-mortem delay, and RNA Integrity numbers. Only age is significantly correlated with some of the expression values, specifically for *SYN1b*, *SYN3a*, and *SYN3g* (see *). When age was included as a covariate in an ANCOVA analysis of differences between diagnostic groups, there was no change in significance levels from the reported ANOVA results in Table 2.

		Gender	Age	Brain pH	Post-Mortem Delay	RNA Integrity No
SYN1a RQ	r2	0,033	0,27	0,176	0,187	-0,006
	p-value	0,250	0,084	0,265	0,236	0,970
SYN1b RQ	r2	0,046	0,327	0,177	0,112	0,044
	p-value	0,168	*0.033	0,256	0,475	0,791
SYN2a RQ	r2	0,068	-0,129	-0,169	-0,107	-0,040
	p-value	0,103	0,421	0,291	0,505	0,810
SYN2b RQ	r2	0,000	0,033	0,072	-0,154	0,146
	p-value	0,996	0,83	0,64	0,319	0,375
SYN3a RQ	r2	0,023	0,324	0,09	0,059	-0,063
	p-value	0,339	*0.039	0,578	0,716	0,716
SYN3g RQ	r2	0,027	0,401	0,154	0,080	0,043
	p-value	0,307	*0.009	0,337	0,618	0,797

Supplemental Table 2: Correlations between possible confounding factors and ChIP/Input values for the different promoter regions. Linear regression was used for gender and Pearson's tests were used for age, pH, and post-mortem delay. Only gender is significantly correlated with the H3K4me3 enrichment at the *SYN1a+b* promoter (see *). When gender was included as a covariate in an ANCOVA analysis of differences between diagnostic groups, there was no change in significance levels from the reported ANOVA results in Table 3.

		Gender	Age	Brain pH	Post-Mortem Delay
SYN1a+b promoter	r2	0,161	-0,085	-0,064	-0,154
	p-value	*0.010	0,602	0,694	0,343
SYN2a+b promoter	r2	0,039	0,005	-0,122	-0,06
	p-value	0,313	0,979	0,536	0,761

Supplemental Methods. Sample characterization:

Brains were collected in collaboration with the Quebec Coroner's Office after consent was obtained from next-of-kin and samples from brain tissue, peripheral blood and urine were collected for toxicological analysis. Two to 4 months later families were contacted and the person best acquainted with the deceased was recruited to undergo a series of structured interviews known as psychological autopsies (295). The interviews were supplemented with information from archival material obtained from hospitals, the Coroner's office and other relevant sources. Following the interviews, clinical vignettes were produced and assessed by a panel of clinicians to generate DSM-IV diagnoses.

The controls were specifically selected to be psychiatrically healthy according to psychiatric autopsies and thus they had no history of psychiatric medication prescriptions. The effect of psychoactive drugs on synapsin gene expression and promoter H3K4me3 enrichment values was investigated both in terms of medical prescription history and toxicology at the time of death. Antidepressants were reported to be prescribed in 46% of the BD group and 40% of the MDD group in the last 3 months before death, and toxicology reports detected these drugs in 15 and 13% respectively. There was no significant correlation (Spearman's test) with expression or H3K4me3 enrichment values. Lithium was reported to be prescribed to 0% of the BD group and 6.7% of the MDD group in the last 3 months before death though in the toxicology report showed lithium in 2 of the 13 BD patients (15%) and none of the MDD or CTRL. Lifetime medication reports indicate some history of lithium in 38% of the BD group. This is unlikely affect gene expression levels at time of death, but it can explain why the patients had access to the drug. Toxicology levels of Li, but not current (last 3 months) or lifetime prescription history showed a significant correlation with expression of *SYN1a* (p-value = 0.025), *SYN1b* (p-value = 0.037), and *SYN3a* (p-value = 0.038), though with only 2 of 41 total subjects represented, no conclusion can be drawn as to its biological effect on synapsin gene expression.

Toxicology reports were also analyzed for tobacco, non-prescription drugs (cocaine, methamphetamine, opiates and cannabinoids detected), and alcohol use. There are no records of tobacco in toxicology reports. Non-prescription drugs were detected in 15% of the BD group, 33% of the MDD group and 15% of the CTRL group; however there was no significant correlation with synapsin gene expression or H3K4me3 promoter enrichment values. Alcohol was detected in 39% of the BD group, 40% of the MDD group and 0% of the controls. Spearman's tests revealed a significant correlation with expression of *SYN1a* (p-value = 0.025), *SYN1b* (p-value = 0.020), and *SYN3a* (p-value = 0.018). This is not surprising as presence of alcohol was restricted to the BD and MDD groups as a result of the ascertainment bias of selecting psychiatrically clean controls; they were also screened for alcohol abuse and dependence problems. If the correlation is computed without the control group, significance is lost: *SYN1a* (p-value = 0.345), *SYN1b* (p-value = 0.181), and *SYN3a* (p-value = 0.102).

Bibliography

1. Alda M, Hajek T, Calkin C, O'Donovan C. Treatment of bipolar disorder: new perspectives. Ann Med. 2009;41(3):186-96.
2. Cruceanu C, Alda M, Turecki G. Lithium: a key to the genetics of bipolar disorder. Genome Med. 2009;1(8):79.
3. Cruceanu C, Lopez de Lara C, Alda M, Turecki G. Pharmacogenetics of Lithium Response. Psychiatric Annals Online. 2008;38(6):1-6.
4. Craddock N, Sklar P. Genetics of bipolar disorder: successful start to a long journey. Trends Genet. 2009;25(2):99-105.
5. Schulze TG. Genetic research into bipolar disorder: the need for a research framework that integrates sophisticated molecular biology and clinically informed phenotype characterization. Psychiatr Clin North Am. 2010;33(1):67-82.
6. Grof P. Sixty years of lithium responders. Neuropsychobiology. 2010;62(1):8-16.
7. Grof P, Alda M, Grof E, Zvolsky P, Walsh M. Lithium response and genetics of affective disorders. J Affect Disord. 1994;32(2):85-95.
8. Grof P, Duffy A, Cavazzoni P, Grof E, Garnham J, MacDougall M, et al. Is response to prophylactic lithium a familial trait? J Clin Psychiatry. 2002;63(10):942-7.
9. Rybakowski JK. Lithium: sixty years thereafter. Neuropsychobiology. 2010;62(1):5-7.
10. Passmore MJ, Garnham J, Duffy A, MacDougall M, Munro A, Slaney C, et al. Phenotypic spectra of bipolar disorder in responders to lithium versus lamotrigine. Bipolar disorders. 2003;5(2):110-4.
11. Turecki G, Grof P, Grof E, D'Souza V, Lebuis L, Marineau C, et al. Mapping susceptibility genes for bipolar disorder: a pharmacogenetic approach based on excellent response to lithium. Mol Psychiatry. 2001;6(5):570-8.
12. Swann AC, Bowden CL, Morris D, Calabrese JR, Petty F, Small J, et al. Depression during mania. Treatment response to lithium or divalproex. Arch Gen Psychiatry. 1997;54(1):37-42.
13. Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Mol Psychiatry. 2002;7(4):405-11.
14. Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, Gill M, Nurnberger JI, Jr., et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. Am J Hum Genet. 2003;73(1):49-62.
15. Turecki G, Grof P, Cavazzoni P, Duffy A, Grof E, Martin R, et al. Lithium responsive bipolar disorder, unilineality, and chromosome 18: A linkage study. Am J Med Genet. 1999;88(4):411-5.
16. Lopez de Lara C, Jaitovich-Groisman I, Cruceanu C, Mamdani F, Lebel V, Yerko V, et al. Implication of synapse-related genes in bipolar disorder by linkage and gene expression analyses. Int J Neuropsychopharmacol. 2010;13(10):1397-410.
17. Duffy A, Turecki G, Grof P, Cavazzoni P, Grof E, Joober R, et al. Association and linkage studies of candidate genes involved in GABAergic neurotransmission in lithium-responsive bipolar disorder. J Psychiatry Neurosci. 2000;25(4):353-8.
18. Mamdani F, Alda M, Grof P, Young LT, Rouleau G, Turecki G. Lithium response and genetic variation in the CREB family of genes. Am J Med Genet B Neuropsychiatr Genet. 2008;147B(4):500-4.
19. Mamdani F, Sequeira A, Alda M, Grof P, Rouleau G, Turecki G. No association between the PREP gene and lithium responsive bipolar disorder. BMC Psychiatry. 2007;7:9.

20. Cavazzoni P, Alda M, Turecki G, Rouleau G, Grof E, Martin R, et al. Lithium-responsive affective disorders: no association with the tyrosine hydroxylase gene. Psychiatry research. 1996;64(2):91-6.
21. Turecki G, Alda M, Grof P, Martin R, Cavazzoni PA, Duffy A, et al. No association between chromosome-18 markers and lithium-responsive affective disorders. Psychiatry research. 1996;63(1):17-23.
22. Serretti A, Lilli R, Lorenzi C, Lattuada E, Cusin C, Smeraldi E. Serotonin transporter gene (5-HTTLPR) and major psychoses. Mol Psychiatry. 2002;7(1):95-9.
23. Adli M, Hollinde DL, Stamm T, Wiethoff K, Tsahuridu M, Kirchheiner J, et al. Response to lithium augmentation in depression is associated with the glycogen synthase kinase 3-beta -50T/C single nucleotide polymorphism. Biol Psychiatry. 2007;62(11):1295-302.
24. Sjoholt G, Ebstein RP, Lie RT, Berle JO, Mallet J, Deleuze JF, et al. Examination of IMPA1 and IMPA2 genes in manic-depressive patients: association between IMPA2 promoter polymorphisms and bipolar disorder. Mol Psychiatry. 2004;9(6):621-9.
25. Serretti A, Mandelli L. The genetics of bipolar disorder: genome 'hot regions,' genes, new potential candidates and future directions. Mol Psychiatry. 2008;13(8):742-71.
26. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661-78.
27. Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry. 2008;13(2):197-207.
28. Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, et al. Whole-genome association study of bipolar disorder. Mol Psychiatry. 2008;13(6):558-69.
29. Campos-de-Sousa S, Guindalini C, Tondo L, Munro J, Osborne S, Floris G, et al. Nuclear receptor rev-erb-{alpha} circadian gene variants and lithium carbonate prophylaxis in bipolar affective disorder. J Biol Rhythms. 2010;25(2):132-7.
30. Manchia M, Congiu D, Squassina A, Lampus S, Ardau R, Chillotti C, et al. No association between lithium full responders and the DRD1, DRD2, DRD3, DAT1, 5-HTTLPR and HTR2A genes in a Sardinian sample. Psychiatry research. 2009;169(2):164-6.
31. Manchia M, Squassina A, Congiu D, Chillotti C, Ardau R, Severino G, et al. Interacting genes in lithium prophylaxis: preliminary results of an exploratory analysis on the role of DGKH and NR1D1 gene polymorphisms in 199 Sardinian bipolar patients. Neurosci Lett. 2009;467(2):67-71.
32. Szczepankiewicz A, Skibinska M, Suwalska A, Hauser J, Rybakowski JK. No association of three GRIN2B polymorphisms with lithium response in bipolar patients. Pharmacol Rep. 2009;61(3):448-52.
33. Szczepankiewicz A, Rybakowski JK, Skibinska M, Dmitrzak-Weglarz M, Leszczynska-Rodziewicz A, Wilkosc M, et al. FYN kinase gene: another glutamatergic gene associated with bipolar disorder? Neuropsychobiology. 2009;59(3):178-83.
34. Rybakowski JK, Suwalska A, Skibinska M, Szczepankiewicz A, Leszczynska-Rodziewicz A, Permoda A, et al. Prophylactic lithium response and polymorphism of the brain-derived neurotrophic factor gene. Pharmacopsychiatry. 2005;38(4):166-70.
35. Szczepankiewicz A, Skibinska M, Suwalska A, Hauser J, Rybakowski JK. The association study of three FYN polymorphisms with prophylactic lithium response in bipolar patients. Hum Psychopharmacol. 2009;24(4):287-91.
36. Kim B, Kim CY, Lee MJ, Joo YH. Preliminary evidence on the association between XBP1-116C/G polymorphism and response to prophylactic treatment with valproate in bipolar disorders. Psychiatry research. 2009;168(3):209-12.

37. Perlis RH, Adams DH, Fijal B, Sutton VK, Farmen M, Breier A, et al. Genetic association study of treatment response with olanzapine/fluoxetine combination or lamotrigine in bipolar I depression. J Clin Psychiatry. 2010;71(5):599-605.
38. Squassina A, Manchia M, Congiu D, Severino G, Chillotti C, Ardau R, et al. The diacylglycerol kinase eta gene and bipolar disorder: a replication study in a Sardinian sample. Mol Psychiatry. 2009;14(4):350-1.
39. Perlis RH, Smoller JW, Ferreira MA, McQuillin A, Bass N, Lawrence J, et al. A genomewide association study of response to lithium for prevention of recurrence in bipolar disorder. The American journal of psychiatry. 2009;166(6):718-25.
40. Seelan RS, Khalyfa A, Lakshmanan J, Casanova MF, Parthasarathy RN. Deciphering the lithium transcriptome: microarray profiling of lithium-modulated gene expression in human neuronal cells. Neuroscience. 2008;151(4):1184-97.
41. Schulze TG, Alda M, Adli M, Akula N, Ardau R, Bui ET, et al. The International Consortium on Lithium Genetics (ConLiGen): an initiative by the NIMH and IGSLI to study the genetic basis of response to lithium treatment. Neuropsychobiology. 2010;62(1):72-8.
42. Pedrosa E, Shah A, Tenore C, Capogna M, Villa C, Guo X, et al. beta-catenin promoter ChIP-chip reveals potential schizophrenia and bipolar disorder gene network. J Neurogenet. 2010;24(4):182-93.
43. King J, Keim M, Teo R, Weening KE, Kapur M, McQuillan K, et al. Genetic control of lithium sensitivity and regulation of inositol biosynthetic genes. PloS one. 2010;5(6):e11151.
44. Kubota M, Kasahara T, Iwamoto K, Komori A, Ishiwata M, Miyauchi T, et al. Therapeutic implications of down-regulation of cyclophilin D in bipolar disorder. Int J Neuropsychopharmacol. 2010;13(10):1355-68.
45. Zhou R, Yuan P, Wang Y, Hunsberger JG, Elkahloun A, Wei Y, et al. Evidence for selective microRNAs and their effectors as common long-term targets for the actions of mood stabilizers. Neuropsychopharmacology. 2009;34(6):1395-405.
46. Chen H, Wang N, Burmeister M, McInnis MG. MicroRNA expression changes in lymphoblastoid cell lines in response to lithium treatment. Int J Neuropsychopharmacol. 2009:1-7.
47. Rong H, Liu TB, Yang KJ, Yang HC, Wu DH, Liao CP, et al. MicroRNA-134 plasma levels before and after treatment for bipolar mania. J Psychiatr Res. 2011;45(1):92-5.
48. Grof P, Duffy A, Alda M, Hajek T. Lithium response across generations. Acta Psychiatr Scand. 2009;120(5):378-85.
49. Aita VM, Liu J, Knowles JA, Terwilliger JD, Baltazar R, Grunn A, et al. A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. Am J Hum Genet. 1999;64(1):210-7.
50. Blackwood DH, He L, Morris SW, McLean A, Whitton C, Thomson M, et al. A locus for bipolar affective disorder on chromosome 4p. Nat Genet. 1996;12(4):427-30.
51. Ewald H, Flint T, Kruse TA, Mors O. A genome-wide scan shows significant linkage between bipolar disorder and chromosome 12q24.3 and suggestive linkage to chromosomes 1p22-21, 4p16, 6q14-22, 10q26 and 16p13.3. Mol Psychiatry. 2002;7(7):734-44.
52. Morissette J, Villeneuve A, Bordeleau L, Rochette D, Laberge C, Gagne B, et al. Genome-wide search for linkage of bipolar affective disorders in a very large pedigree derived from a homogeneous population in quebec points to a locus of major effect on chromosome 12q23-q24. Am J Med Genet. 1999;88(5):567-87.
53. Degn B, Lundorf MD, Wang A, Vang M, Mors O, Kruse TA, et al. Further evidence for a bipolar risk gene on chromosome 12q24 suggested by investigation of haplotype sharing and allelic association in patients from the Faroe Islands. Mol Psychiatry. 2001;6(4):450-5.

54. Detera-Wadleigh SD, Badner JA, Berrettini WH, Yoshikawa T, Goldin LR, Turner G, et al. A high-density genome scan detects evidence for a bipolar-disorder susceptibility locus on 13q32 and other potential loci on 1q32 and 18p11.2. Proc Natl Acad Sci U S A. 1999;96(10):5604-9.
55. Badenhop RF, Moses MJ, Scimone A, Mitchell PB, Ewen KR, Rosso A, et al. A genome screen of a large bipolar affective disorder pedigree supports evidence for a susceptibility locus on chromosome 13q. Mol Psychiatry. 2001;6(4):396-403.
56. Berrettini WH, Ferraro TN, Goldin LR, Weeks DE, Detera-Wadleigh S, Nurnberger JI, Jr., et al. Chromosome 18 DNA markers and manic-depressive illness: evidence for a susceptibility gene. Proc Natl Acad Sci U S A. 1994;91(13):5918-21.
57. Freimer NB, Reus VI, Escamilla MA, McInnes LA, Spesny M, Leon P, et al. Genetic mapping using haplotype, association and linkage methods suggests a locus for severe bipolar disorder (BPI) at 18q22-q23. Nat Genet. 1996;12(4):436-41.
58. Kelsoe JR, Spence MA, Loetscher E, Foguet M, Sadovnick AD, Remick RA, et al. A genome survey indicates a possible susceptibility locus for bipolar disorder on chromosome 22. Proc Natl Acad Sci U S A. 2001;98(2):585-90.
59. Newton JR. Linked gene ontology categories are novel and differ from associated gene ontology categories for the bipolar disorders. Psychiatr Genet. 2007;17(1):29-34.
60. Badner JA, Koller D, Foroud T, Edenberg H, Nurnberger JI, Jr., Zandi PP, et al. Genome-wide linkage analysis of 972 bipolar pedigrees using single-nucleotide polymorphisms. Mol Psychiatry. 2012;17(8):818-26.
61. Neale BM, Sklar P. Genetic analysis of schizophrenia and bipolar disorder reveals polygenicity but also suggests new directions for molecular interrogation. Current opinion in neurobiology. 2015;30C:131-8.
62. Finishing the euchromatic sequence of the human genome. Nature. 2004;431(7011):931-45.
63. Schmutz J, Wheeler J, Grimwood J, Dickson M, Yang J, Caoile C, et al. Quality assessment of the human genome sequence. Nature. 2004;429(6990):365-8.
64. Askland K, Read C, Moore J. Pathways-based analyses of whole-genome association study data in bipolar disorder reveal genes mediating ion channel activity and synaptic neurotransmission. Hum Genet. 2009;125(1):63-79.
65. Marmol F. Lithium: bipolar disorder and neurodegenerative diseases Possible cellular mechanisms of the therapeutic effects of lithium. Prog Neuropsychopharmacol Biol Psychiatry. 2008;32(8):1761-71.
66. Goes FS. Genetics of Bipolar Disorder: Recent Update and Future Directions. Psychiatr Clin North Am. 2016;39(1):139-55.
67. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet. 2011;43(10):977-83.
68. Cichon S, Muhleisen TW, Degenhardt FA, Mattheisen M, Miro X, Strohmaier J, et al. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. Am J Hum Genet. 2011;88(3):372-81.
69. Chen DT, Jiang X, Akula N, Shugart YY, Wendland JR, Steele CJ, et al. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. Mol Psychiatry. 2013;18(2):195-205.
70. Green EK, Hamshere M, Forty L, Gordon-Smith K, Fraser C, Russell E, et al. Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. Mol Psychiatry. 2013;18(12):1302-7.

71. Muhleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. Nature communications. 2014;5:3339.
72. Chen CH, Lee CS, Lee MT, Ouyang WC, Chen CC, Chong MY, et al. Variant GADL1 and response to lithium therapy in bipolar I disorder. The New England journal of medicine. 2014;370(2):119-28.
73. Ikeda M, Kondo K, Iwata N. Variant GADL1 and response to lithium in bipolar I disorder. The New England journal of medicine. 2014;370(19):1856-7.
74. Vlachadis N, Vrachnis N, Economou E. Variant GADL1 and response to lithium in bipolar I disorder. The New England journal of medicine. 2014;370(19):1856.
75. Birnbaum R, Shin JH, Weinberger D. Variant GADL1 and response to lithium in bipolar I disorder. The New England journal of medicine. 2014;370(19):1855-6.
76. Lee CS, Cheng AT. Variant GADL1 and response to lithium in bipolar I disorder. The New England journal of medicine. 2014;370(19):1859-60.
77. Cruceanu C, Alda M, Dion PA, Turecki G, Rouleau GA. No evidence for GADL1 variation as a bipolar disorder susceptibility factor in a Caucasian lithium-responsive cohort. The American journal of psychiatry. 2015;172(1):94-5.
78. Song J, Bergen SE, Di Florio A, Karlsson R, Charney A, Ruderfer DM, et al. Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. Mol Psychiatry. 2015.
79. Hou L, Heilbronner U, Degenhardt F, Adli M, Akiyama K, Akula N, et al. Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. Lancet. 2016.
80. Alaerts M, Del-Favero J. Searching genetic risk factors for schizophrenia and bipolar disorder: learn from the past and back to the future. Hum Mutat. 2009;30(8):1139-52.
81. Craddock N, Sklar P. Genetics of bipolar disorder. Lancet. 2013;381(9878):1654-62.
82. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science. 2012;337(6090):64-9.
83. Lupski JR, Belmont JW, Boerwinkle E, Gibbs RA. Clan genomics and the complex architecture of human disease. Cell. 2011;147(1):32-43.
84. Collins AL, Kim Y, Szatkiewicz JP, Bloom RJ, Hilliard CE, Quackenbush CR, et al. Identifying bipolar disorder susceptibility loci in a densely affected pedigree. Mol Psychiatry. 2013;18(12):1245-6.
85. Georgi B, Craig D, Kember RL, Liu W, Lindquist I, Nasser S, et al. Genomic view of bipolar disorder revealed by whole genome sequencing in a genetic isolate. PLoS genetics. 2014;10(3):e1004229.
86. Cruceanu C, Ambalavanan A, Spiegelman D, Gauthier J, Lafreniere RG, Dion PA, et al. Family-based exome-sequencing approach identifies rare susceptibility variants for lithium-responsive bipolar disorder. Genome / National Research Council Canada = Genome / Conseil national de recherches Canada. 2013;56(10):634-40.
87. Ament SA, Szelinger S, Glusman G, Ashworth J, Hou L, Akula N, et al. Rare variants in neuronal excitability genes influence risk for bipolar disorder. Proc Natl Acad Sci U S A. 2015;112(11):3576-81.
88. Seifuddin F, Pirooznia M, Judy JT, Goes FS, Potash JB, Zandi PP. Systematic review of genome-wide gene expression studies of bipolar disorder. BMC Psychiatry. 2013;13:213.

89. Konradi C, Sillivan SE, Clay HB. Mitochondria, oligodendrocytes and inflammation in bipolar disorder: evidence from transcriptome studies points to intriguing parallels with multiple sclerosis. Neurobiology of disease. 2012;45(1):37-47.
90. Elashoff M, Higgs BW, Yolken RH, Knable MB, Weis S, Webster MJ, et al. Meta-analysis of 12 genomic studies in bipolar disorder. Journal of molecular neuroscience : MN. 2007;31(3):221-43.
91. Leonard S, Logel J, Luthman D, Casanova M, Kirch D, Freedman R. Biological stability of mRNA isolated from human postmortem brain collections. Biol Psychiatry. 1993;33(6):456-66.
92. Munkholm K, Vinberg M, Berk M, Kessing LV. State-related alterations of gene expression in bipolar disorder: a systematic review. Bipolar disorders. 2012;14(7):684-96.
93. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature biotechnology. 2010;28(5):511-5.
94. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews Genetics. 2009;10(1):57-63.
95. Webb A, Papp AC, Curtis A, Newman LC, Pietrzak M, Seweryn M, et al. RNA sequencing of transcriptomes in human brain regions: protein-coding and non-coding RNAs, isoforms and alleles. BMC genomics. 2015;16:990.
96. Akula N, Barb J, Jiang X, Wendland JR, Choi KH, Sen SK, et al. RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. Mol Psychiatry. 2014.
97. Cruceanu C, Tan PP, Rogic S, Lopez JP, Torres-Platas SG, Gigek CO, et al. Transcriptome Sequencing of the Anterior Cingulate in Bipolar Disorder: Dysregulation of G Protein-Coupled Receptors. The American journal of psychiatry. 2015:appiajp201514101279.
98. Zhao Z, Xu J, Chen J, Kim S, Reimers M, Bacanu SA, et al. Transcriptome sequencing and genome-wide association analyses reveal lysosomal function and actin cytoskeleton remodeling in schizophrenia and bipolar disorder. Mol Psychiatry. 2015;20(5):563-72.
99. Zhao H, Xu J, Pang L, Zhang Y, Fan H, Liu L, et al. Genome-wide DNA methylome reveals the dysfunction of intronic microRNAs in major psychosis. BMC medical genomics. 2015;8:62.
100. Xiao Y, Camarillo C, Ping Y, Arana TB, Zhao H, Thompson PM, et al. The DNA methylome and transcriptome of different brain regions in schizophrenia and bipolar disorder. PloS one. 2014;9(4):e95875.
101. Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S, Bouchard L, et al. Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. Am J Hum Genet. 2008;82(3):696-711.
102. Li Y, Camarillo C, Xu J, Arana TB, Xiao Y, Zhao Z, et al. Genome-wide methylome analyses reveal novel epigenetic regulation patterns in schizophrenia and bipolar disorder. BioMed research international. 2015;2015:201587.
103. Houtepen LC, van Bergen AH, Vinkers CH, Boks MP. DNA methylation signatures of mood stabilizers and antipsychotics in bipolar disorder. Epigenomics. 2016;8(2):197-208.
104. Huzayyin AA, Andreazza AC, Turecki G, Cruceanu C, Rouleau GA, Alda M, et al. Decreased global methylation in patients with bipolar disorder who respond to lithium. Int J Neuropsychopharmacol. 2014;17(4):561-9.
105. McGuffin P, Katz R. The genetics of depression and manic-depressive disorder. Br J Psychiatry. 1989;155:294-304.
106. Craddock N, Jones I. Genetics of bipolar disorder. Journal of medical genetics. 1999;36(8):585-94.

107. Mamdani F, Groisman IJ, Alda M, Turecki G. Pharmacogenetics and bipolar disorder. Pharmacogenomics J. 2004;4(3):161-70.
108. Smoller JW, Finn CT. Family, twin, and adoption studies of bipolar disorder. Am J Med Genet C Semin Med Genet. 2003;123C(1):48-58.
109. Cruceanu C, Alda M, Rouleau G, Turecki G. Response to treatment in bipolar disorder. Current opinion in psychiatry. 2011;24(1):24-8.
110. Shinozaki G, Potash JB. New developments in the genetics of bipolar disorder. Current psychiatry reports. 2014;16(11):493.
111. Manchia M, Cullis J, Turecki G, Rouleau GA, Uher R, Alda M. The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. PloS one. 2013;8(10):e76295.
112. Gould TD, Manji HK. Signaling networks in the pathophysiology and treatment of mood disorders. Journal of psychosomatic research. 2002;53(2):687-97.
113. Xu X, Muller-Taubenberger A, Adley KE, Pawolleck N, Lee VW, Wiedemann C, et al. Attenuation of phospholipid signaling provides a novel mechanism for the action of valproic acid. Eukaryotic cell. 2007;6(6):899-906.
114. Alda M. Lithium in the treatment of bipolar disorder: pharmacology and pharmacogenetics. Mol Psychiatry. 2015;20(6):661-70.
115. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.
116. Ament SA, Szelinger S, Glusman G, Ashworth J, Hou L, Akula N, et al. Rare variants in neuronal excitability genes influence risk for bipolar disorder. Proc Natl Acad Sci U S A. 2015;112(11):3576-81.
117. Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. Nature. 2014;506(7487):185-90.
118. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols. 2009;4(7):1073-81.
119. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nature methods. 2010;7(4):248-9.
120. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nature methods. 2010;7(8):575-6.
121. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome research. 2009;19(9):1553-61.
122. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome research. 2010;20(1):110-21.
123. Cooper GM, Stone EA, Asimenos G, Program NCS, Green ED, Batzoglou S, et al. Distribution and intensity of constraint in mammalian genomic sequence. Genome research. 2005;15(7):901-13.
124. Samocha KE, Robinson EB, Sanders SJ, Stevens C, Sabo A, McGrath LM, et al. A framework for the interpretation of de novo mutation in human disease. Nat Genet. 2014;46(9):944-50.
125. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. PLoS genetics. 2013;9(8):e1003709.
126. Nature Reviews Drug Discovery GQP. The state of GPCR research in 2004. Nat Rev Drug Discov. 2004;3(7):575, 7-626.
127. Lovenberg TW, Liaw CW, Grigoriadis DE, Clevenger W, Chalmers DT, De Souza EB, et al. Cloning and characterization of a functionally distinct corticotropin-releasing factor receptor subtype from rat brain. Proc Natl Acad Sci U S A. 1995;92(3):836-40.

128. Milligan G. Principles: extending the utility of [35S]GTP gamma S binding assays. Trends Pharmacol Sci. 2003;24(2):87-90.
129. Zhang R, Xie X. Tools for GPCR drug discovery. Acta Pharmacol Sin. 2012;33(3):372-84.
130. Mangmool S, Kurose H. G(i/o) protein-dependent and -independent actions of Pertussis Toxin (PTX). Toxins (Basel). 2011;3(7):884-99.
131. Markovic D, Punn A, Lehnert H, Grammatopoulos DK. Intracellular mechanisms regulating corticotropin-releasing hormone receptor-2beta endocytosis and interaction with extracellularly regulated kinase 1/2 and p38 mitogen-activated protein kinase signaling cascades. Molecular endocrinology. 2008;22(3):689-706.
132. Tateyama M, Kubo Y. Coupling profile of the metabotropic glutamate receptor 1alpha is regulated by the C-terminal domain. Molecular and cellular neurosciences. 2007;34(3):445-52.
133. Murphy MG, Horackova M. Short-term exposure of intact ventricular myocytes to isoproterenol results in desensitization of adenylate cyclase. Eur J Pharmacol. 1994;268(2):121-8.
134. Bhattacharyya S. Inside story of Group I Metabotropic Glutamate Receptors (mGluRs). Int J Biochem Cell Biol. 2016.
135. Abdul-Ghani MA, Valiante TA, Carlen PL, Pennefather PS. Metabotropic glutamate receptors coupled to IP3 production mediate inhibition of IAHP in rat dentate granule neurons. J Neurophysiol. 1996;76(4):2691-700.
136. Cullen PJ. Calcium signalling: the ups and downs of protein kinase C. Curr Biol. 2003;13(18):R699-701.
137. Violin JD, Zhang J, Tsien RY, Newton AC. A genetically encoded fluorescent reporter reveals oscillatory phosphorylation by protein kinase C. J Cell Biol. 2003;161(5):899-909.
138. Nanou E, El Manira A. Mechanisms of modulation of AMPA-induced Na+-activated K+ current by mGluR1. J Neurophysiol. 2010;103(1):441-5.
139. Racagni G, Popoli M. Cellular and molecular mechanisms in the long-term action of antidepressants. Dialogues in clinical neuroscience. 2008;10(4):385-400.
140. Wise A, Jupe SC, Rees S. The identification of ligands at orphan G-protein coupled receptors. Annual review of pharmacology and toxicology. 2004;44:43-66.
141. Kazius J, Wurdinger K, van Iterson M, Kok J, Back T, Ijzerman AP. GPCR NaVa database: natural variants in human G protein-coupled receptors. Hum Mutat. 2008;29(1):39-44.
142. Tomita H, Ziegler ME, Kim HB, Evans SJ, Choudary PV, Li JZ, et al. G protein-linked signaling pathways in bipolar and major depressive disorders. Frontiers in genetics. 2013;4:297.
143. Fries GR, Vasconcelos-Moreno MP, Gubert C, dos Santos BT, Sartori J, Eisele B, et al. Hypothalamic-pituitary-adrenal axis dysfunction and illness progression in bipolar disorder. Int J Neuropsychopharmacol. 2014;18(1).
144. Okamoto Y, Shikano S. Phosphorylation-dependent C-terminal binding of 14-3-3 proteins promotes cell surface expression of HIV co-receptor GPR15. J Biol Chem. 2011;286(9):7171-81.
145. Maurice P, Guillaume JL, Benleulmi-Chaachoua A, Daulat AM, Kamal M, Jockers R. GPCR-interacting proteins, major players of GPCR function. Adv Pharmacol. 2011;62:349-80.
146. Benovic JL, Kuhn H, Weyand I, Codina J, Caron MG, Lefkowitz RJ. Functional desensitization of the isolated beta-adrenergic receptor by the beta-adrenergic receptor kinase: potential role of an analog of the retinal protein arrestin (48-kDa protein). Proc Natl Acad Sci U S A. 1987;84(24):8879-82.
147. Benzing T, Yaffe MB, Arnould T, Sellin L, Schermer B, Schilling B, et al. 14-3-3 interacts with regulator of G protein signaling proteins and modulates their activity. J Biol Chem. 2000;275(36):28167-72.

148. Ayoub MA, Angelicheva D, Vile D, Chandler D, Morar B, Cavanaugh JA, et al. Deleterious GRM1 mutations in schizophrenia. PloS one. 2012;7(3):e32849.
149. Frank RA, McRae AF, Pocklington AJ, van de Lagemaat LN, Navarro P, Croning MD, et al. Clustered coding variants in the glutamate receptor complexes of individuals with schizophrenia and bipolar disorder. PloS one. 2011;6(4):e19011.
150. Menke A, Samann P, Kloiber S, Czamara D, Lucae S, Hennings J, et al. Polymorphisms within the metabotropic glutamate receptor 1 gene are associated with depression phenotypes. Psychoneuroendocrinology. 2012;37(4):565-75.
151. Tamura A, Yamada N, Yaguchi Y, Machida Y, Mori I, Osanai M. Both neurons and astrocytes exhibited tetrodotoxin-resistant metabotropic glutamate receptor-dependent spontaneous slow Ca2+ oscillations in striatum. PloS one. 2014;9(1):e85351.
152. Liu F, Zhou R, Yan H, Yin H, Wu X, Tan Y, et al. Metabotropic glutamate receptor 5 modulates calcium oscillation and innate immune response induced by lipopolysaccharide in microglial cell. Neuroscience. 2014;281C:24-34.
153. Cho HP, Garcia-Barrantes PM, Brogan JT, Hopkins CR, Niswender CM, Rodriguez AL, et al. Chemical modulation of mutant mGlu1 receptors derived from deleterious GRM1 mutations found in schizophrenics. ACS Chem Biol. 2014;9(10):2334-46.
154. Coutinho V, Knopfel T. Metabotropic glutamate receptors: electrical and chemical signaling properties. Neuroscientist. 2002;8(6):551-61.
155. Sharma AN, Fries GR, Galvez JF, Valvassori SS, Soares JC, Carvalho AF, et al. Modeling mania in preclinical settings: A comprehensive review. Prog Neuropsychopharmacol Biol Psychiatry. 2016;66:22-34.
156. Cosgrove VE, Kelsoe JR, Suppes T. Toward a Valid Animal Model of Bipolar Disorder: How the Research Domain Criteria Help Bridge the Clinical-Basic Science Divide. Biol Psychiatry. 2016;79(1):62-70.
157. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010;20(9):1297-303.
158. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 2010;38(16):e164.
159. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.
160. ExomeVariantServer. NHLBI GO Exome Sequencing Project (ESP) Seattle, WA [updated 2014]. Available from: http://evs.gs.washington.edu/EVS/.
161. Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science. 2010;327(5961):78-81.
162. Leduc M, Breton B, Gales C, Le Gouill C, Bouvier M, Chemtob S, et al. Functional selectivity of natural and synthetic prostaglandin EP4 receptor ligands. J Pharmacol Exp Ther. 2009;331(1):297-307.
163. Cruceanu C, Alda M, Grof P, Rouleau GA, Turecki G. Synapsin II is involved in the molecular pathway of lithium treatment in bipolar disorder. PloS one. 2012;7(2):e32680.
164. Sun X, Young LT, Wang JF, Grof P, Turecki G, Rouleau GA, et al. Identification of lithium-regulated genes in cultured lymphoblasts of lithium responsive subjects with bipolar disorder. Neuropsychopharmacology. 2004;29(4):799-804.
165. Pradhan AA, Becker JA, Scherrer G, Tryoen-Toth P, Filliol D, Matifas A, et al. In vivo delta opioid receptor internalization controls behavioral effects of agonists. PloS one. 2009;4(5):e5425.

166. Cerullo MA, Adler CM, Delbello MP, Strakowski SM. The functional neuroanatomy of bipolar disorder. International review of psychiatry. 2009;21(4):314-22.
167. Lim CS, Baldessarini RJ, Vieta E, Yucel M, Bora E, Sim K. Longitudinal neuroimaging and neuropsychological changes in bipolar disorder patients: review of the evidence. Neuroscience and biobehavioral reviews. 2013;37(3):418-35.
168. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology. 2013;14(4):R36.
169. Anders SP, P.T.; Huber, W. HTSeq – A Python framework to work with high-throughput sequencing data. bioRxiv. 2014.
170. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology. 2010;11(3):R25.
171. Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome biology. 2014;15(2):R29.
172. Gillis J, Mistry M, Pavlidis P. Gene function analysis in complex data sets using ErmineJ. Nature protocols. 2010;5(6):1148-59.
173. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012;489(7416):391-9.
174. Sheridan SD, Theriault KM, Reis SA, Zhou F, Madison JM, Daheron L, et al. Epigenetic characterization of the FMR1 gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile X syndrome. PloS one. 2011;6(10):e26203.
175. Cruceanu C, Alda M, Nagy C, Freemantle E, Rouleau GA, Turecki G. H3K4 tri-methylation in synapsin genes leads to different expression patterns in bipolar disorder and major depression. Int J Neuropsychopharmacol. 2013;16(2):289-99.
176. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer research. 2004;64(15):5245-50.
177. Kapranov P, St Laurent G, Raz T, Ozsolak F, Reynolds CP, Sorensen PH, et al. The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' un-annotated RNA. BMC biology. 2010;8:149.
178. Kalyana-Sundaram S, Kumar-Sinha C, Shankar S, Robinson DR, Wu YM, Cao X, et al. Expressed pseudogenes in the transcriptional landscape of human cancers. Cell. 2012;149(7):1622-34.
179. Sequeira A, Mamdani F, Ernst C, Vawter MP, Bunney WE, Lebel V, et al. Global brain gene expression analysis links glutamatergic and GABAergic alterations to suicide and major depression. PloS one. 2009;4(8):e6585.
180. Ernst C, Dumoulin P, Cabot S, Erickson J, Turecki G. SNAT1 and a family with high rates of suicidal behavior. Neuroscience. 2009;162(2):415-22.
181. Amado M, Almeida R, Carneiro F, Levery SB, Holmes EH, Nomoto M, et al. A family of human beta3-galactosyltransferases. Characterization of four members of a UDP-galactose:beta-N-acetyl-glucosamine/beta-nacetyl-galactosamine beta-1,3-galactosyltransferase family. J Biol Chem. 1998;273(21):12770-8.
182. Zhu D, Shen A, Wang Y, Gu X, Gu J. Developmental regulation of beta-1,3-galactosyltransferase-1 gene expression in mouse brain. FEBS letters. 2003;538(1-3):163-7.
183. Conn PJ, Christopoulos A, Lindsley CW. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. Nat Rev Drug Discov. 2009;8(1):41-54.

184. Cannon DM, Klaver JK, Gandhi SK, Solorio G, Peck SA, Erickson K, et al. Genetic variation in cholinergic muscarinic-2 receptor gene modulates M2 receptor binding in vivo and accounts for reduced binding in bipolar disorder. Mol Psychiatry. 2011;16(4):407-18.
185. Gibbons AS, Scarr E, McLean C, Sundram S, Dean B. Decreased muscarinic receptor binding in the frontal cortex of bipolar disorder and major depressive disorder subjects. J Affect Disord. 2009;116(3):184-91.
186. Beneyto M, Morris HM, Rovensky KC, Lewis DA. Lamina- and cell-specific alterations in cortical somatostatin receptor 2 mRNA expression in schizophrenia. Neuropharmacology. 2012;62(3):1598-605.
187. Nanda SA, Qi C, Roseboom PH, Kalin NH. Predator stress induces behavioral inhibition and amygdala somatostatin receptor 2 gene expression. Genes, brain, and behavior. 2008;7(6):639-48.
188. Reif A, Nguyen TT, Weissflog L, Jacob CP, Romanos M, Renner TJ, et al. DIRAS2 is associated with adult ADHD, related traits, and co-morbid disorders. Neuropsychopharmacology. 2011;36(11):2318-27.
189. Stengel A, Rivier J, Tache Y. Modulation of the adaptive response to stress by brain activation of selective somatostatin receptor subtypes. Peptides. 2013;42:70-7.
190. Nyegaard M, Borglum AD, Bruun TG, Collier DA, Russ C, Mors O, et al. Novel polymorphisms in the somatostatin receptor 5 (SSTR5) gene associated with bipolar affective disorder. Mol Psychiatry. 2002;7(7):745-54.
191. Kato T. Molecular genetics of bipolar disorder and depression. Psychiatry and clinical neurosciences. 2007;61(1):3-19.
192. Ryan MM, Lockstone HE, Huffaker SJ, Wayland MT, Webster MJ, Bahn S. Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes. Mol Psychiatry. 2006;11(10):965-78.
193. He Z, Bammann H, Han D, Xie G, Khaitovich P. Conserved expression of lincRNA during human and macaque prefrontal cortex development and maturation. Rna. 2014.
194. Ng SY, Bogu GK, Soh BS, Stanton LW. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. Molecular cell. 2013;51(3):349-59.
195. Earls LR, Westmoreland JJ, Zakharenko SS. Non-coding RNA regulation of synaptic plasticity and memory: Implications for aging. Ageing research reviews. 2014.
196. van D, II, Gordebeke PM, Khoshab N, Tiesinga PH, Buitelaar JK, Kozicz T, et al. Long non-coding RNAs in neurodevelopmental disorders. Frontiers in molecular neuroscience. 2013;6:53.
197. Wilkinson B, Campbell DB. Contribution of long noncoding RNAs to autism spectrum disorder risk. International review of neurobiology. 2013;113:35-59.
198. Schou M. [Preventive lithium treatment in manic-depressive disease: experiences and progress in recent years]. Nervenarzt. 1983;54(7):331-9.
199. Baldessarini RJ, Tondo L. Does lithium treatment still work? Evidence of stable responses over three decades. Arch Gen Psychiatry. 2000;57(2):187-90.
200. Garnham J, Munro A, Slaney C, Macdougall M, Passmore M, Duffy A, et al. Prophylactic treatment response in bipolar disorder: results of a naturalistic observation study. J Affect Disord. 2007;104(1-3):185-90.
201. Kile BM, Guillot TS, Venton BJ, Wetsel WC, Augustine GJ, Wightman RM. Synapsins differentially control dopamine and serotonin release. J Neurosci. 2010;30(29):9762-70.
202. Cesca F, Baldelli P, Valtorta F, Benfenati F. The synapsins: key actors of synapse function and plasticity. Prog Neurobiol. 2010;91(4):313-48.
203. Ullrich B, Sudhof TC. Differential distributions of novel synaptotagmins: comparison to synapsins. Neuropharmacology. 1995;34(11):1371-7.

204. Haycock JW, Greengard P, Browning MD. Cholinergic regulation of protein III phosphorylation in bovine adrenal chromaffin cells. J Neurosci. 1988;8(9):3233-9.
205. Browning MD, Huang CK, Greengard P. Similarities between protein IIIa and protein IIIb, two prominent synaptic vesicle-associated phosphoproteins. J Neurosci. 1987;7(3):847-53.
206. Maienschein V, Marxen M, Volknandt W, Zimmermann H. A plethora of presynaptic proteins associated with ATP-storing organelles in cultured astrocytes. Glia. 1999;26(3):233-44.
207. Bhangu PS, Genever PG, Spencer GJ, Grewal TS, Skerry TM. Evidence for targeted vesicular glutamate exocytosis in osteoblasts. Bone. 2001;29(1):16-23.
208. Matovcik LM, Karapetian O, Czernik AJ, Marino CR, Kinder BK, Gorelick FS. Antibodies to an epitope on synapsin I detect a protein associated with the endocytic compartment in non-neuronal cells. Eur J Cell Biol. 1994;65(2):327-40.
209. Bustos R, Kolen ER, Braiterman L, Baines AJ, Gorelick FS, Hubbard AL. Synapsin I is expressed in epithelial cells: localization to a unique trans-Golgi compartment. J Cell Sci. 2001;114(Pt 20):3695-704.
210. Hurley SL, Brown DL, Cheetham JJ. Cytoskeletal interactions of synapsin I in non-neuronal cells. Biochem Biophys Res Commun. 2004;317(1):16-23.
211. Li L, Chin LS, Shupliakov O, Brodin L, Sihra TS, Hvalby O, et al. Impairment of synaptic vesicle clustering and of synaptic transmission, and increased seizure propensity, in synapsin I-deficient mice. Proc Natl Acad Sci U S A. 1995;92(20):9235-9.
212. Rosahl TW, Spillane D, Missler M, Herz J, Selig DK, Wolff JR, et al. Essential functions of synapsins I and II in synaptic vesicle regulation. Nature. 1995;375(6531):488-93.
213. Lakhan R, Kalita J, Misra UK, Kumari R, Mittal B. Association of intronic polymorphism rs3773364 A>G in synapsin-2 gene with idiopathic epilepsy. Synapse. 2010;64(5):403-8.
214. Chen Q, He G, Qin W, Chen QY, Zhao XZ, Duan SW, et al. Family-based association study of synapsin II and schizophrenia. Am J Hum Genet. 2004;75(5):873-7.
215. Lee HJ, Song JY, Kim JW, Jin SY, Hong MS, Park JK, et al. Association study of polymorphisms in synaptic vesicle-associated genes, SYN2 and CPLX2, with schizophrenia. Behav Brain Funct. 2005;1:15.
216. Saviouk V, Moreau MP, Tereshchenko IV, Brzustowicz LM. Association of synapsin 2 with schizophrenia in families of Northern European ancestry. Schizophr Res. 2007;96(1-3):100-11.
217. Wang Y, Yu L, Zhao T, Xu J, Liu Z, Liu Y, et al. No association between bipolar disorder and syngr1 or synapsin II polymorphisms in the Han Chinese population. Psychiatry research. 2009;169(2):167-8.
218. Grebb JA, Greengard P. An analysis of synapsin II, a neuronal phosphoprotein, in postmortem brain tissue from alcoholic and neuropsychiatrically ill adults and medically ill children and young adults. Arch Gen Psychiatry. 1990;47(12):1149-56.
219. Vawter MP, Thatcher L, Usen N, Hyde TM, Kleinman JE, Freed WJ. Reduction of synapsin in the hippocampus of patients with bipolar disorder and schizophrenia. Mol Psychiatry. 2002;7(6):571-8.
220. Mirnics K, Middleton FA, Marquez A, Lewis DA, Levitt P. Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. Neuron. 2000;28(1):53-67.
221. Alda M, Keller D, Grof E, Turecki G, Cavazzoni P, Duffy A, et al. Is lithium response related to G(s)alpha levels in transformed lymphoblasts from subjects with bipolar disorder? J Affect Disord. 2001;65(2):117-22.
222. Benedetti F, Serretti A, Pontiggia A, Bernasconi A, Lorenzi C, Colombo C, et al. Long-term response to lithium salts in bipolar illness is influenced by the glycogen synthase kinase 3-beta -50 T/C SNP. Neurosci Lett. 2005;376(1):51-5.

223. Phiel CJ, Zhang F, Huang EY, Guenther MG, Lazar MA, Klein PS. Histone deacetylase is a direct target of valproic acid, a potent anticonvulsant, mood stabilizer, and teratogen. J Biol Chem. 2001;276(39):36734-41.
224. Gottlicher M, Minucci S, Zhu P, Kramer OH, Schimpf A, Giavara S, et al. Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. EMBO J. 2001;20(24):6969-78.
225. Valtorta F, Benfenati F, Greengard P. Structure and function of the synapsins. J Biol Chem. 1992;267(11):7195-8.
226. De Camilli P, Benfenati F, Valtorta F, Greengard P. The synapsins. Annu Rev Cell Biol. 1990;6:433-60.
227. Petersohn D, Schoch S, Brinkmann DR, Thiel G. The human synapsin II gene promoter. Possible role for the transcription factor zif268/egr-1, polyoma enhancer activator 3, and AP2. J Biol Chem. 1995;270(41):24361-9.
228. Rao JS, Rapoport SI, Bosetti F. Decrease in the AP-2 DNA-binding activity and in the protein expression of AP-2 alpha and AP-2 beta in frontal cortex of rats treated with lithium for 6 weeks. Neuropsychopharmacology. 2005;30(11):2006-13.
229. Goodwin FK, Fireman B, Simon GE, Hunkeler EM, Lee J, Revicki D. Suicide risk in bipolar disorder during treatment with lithium and divalproex. JAMA : the journal of the American Medical Association. 2003;290(11):1467-73.
230. Baldessarini RJ, Tondo L, Faedda GL, Suppes TR, Floris G, Rudas N. Effects of the rate of discontinuing lithium maintenance treatment in bipolar disorders. J Clin Psychiatry. 1996;57(10):441-8.
231. Thies-Flechtner K, Muller-Oerlinghausen B, Seibert W, Walther A, Greil W. Effect of prophylactic treatment on suicide risk in patients with major affective disorders. Data from a randomized prospective trial. Pharmacopsychiatry. 1996;29(3):103-7.
232. Muller-Oerlinghausen B, Muser-Causemann B, Volk J. Suicides and parasuicides in a high-risk patient group on and off lithium long-term medication. J Affect Disord. 1992;25(4):261-9.
233. Cipriani A, Pretty H, Hawton K, Geddes JR. Lithium in the prevention of suicidal behavior and all-cause mortality in patients with mood disorders: a systematic review of randomized trials. The American journal of psychiatry. 2005;162(10):1805-19.
234. Baldessarini RJ, Tondo L, Davis P, Pompili M, Goodwin FK, Hennen J. Decreased risk of suicides and attempts during long-term lithium treatment: a meta-analytic review. Bipolar disorders. 2006;8(5 Pt 2):625-39.
235. Porton B, Kao HT, Greengard P. Characterization of transcripts from the synapsin III gene locus. J Neurochem. 1999;73(6):2266-71.
236. Sudhof TC, Czernik AJ, Kao HT, Takei K, Johnston PA, Horiuchi A, et al. Synapsins: mosaics of shared and individual domains in a family of synaptic vesicle phosphoproteins. Science. 1989;245(4925):1474-80.
237. Corradi A, Zanardi A, Giacomini C, Onofri F, Valtorta F, Zoli M, et al. Synapsin-I- and synapsin-II-null mice display an increased age-dependent cognitive impairment. J Cell Sci. 2008;121(Pt 18):3042-51.
238. Grafodatskaya D, Chung B, Szatmari P, Weksberg R. Autism spectrum disorders and epigenetics. J Am Acad Child Adolesc Psychiatry. 2010;49(8):794-809.
239. Ferreira A, Kao HT, Feng J, Rapoport M, Greengard P. Synapsin III: developmental expression, subcellular localization, and role in axon formation. J Neurosci. 2000;20(10):3736-44.

240. Huttner WB, Schiebler W, Greengard P, De Camilli P. Synapsin I (protein I), a nerve terminal-specific phosphoprotein. III. Its association with synaptic vesicles studied in a highly purified synaptic vesicle preparation. J Cell Biol. 1983;96(5):1374-88.
241. Takamori S, Holt M, Stenius K, Lemke EA, Gronborg M, Riedel D, et al. Molecular anatomy of a trafficking organelle. Cell. 2006;127(4):831-46.
242. Kao HT, Porton B, Czernik AJ, Feng J, Yiu G, Haring M, et al. A third member of the synapsin gene family. Proc Natl Acad Sci U S A. 1998;95(8):4667-72.
243. Feng J, Chi P, Blanpied TA, Xu Y, Magarinos AM, Ferreira A, et al. Regulation of neurotransmitter release by synapsin III. J Neurosci. 2002;22(11):4372-80.
244. Gitler D, Takagishi Y, Feng J, Ren Y, Rodriguiz RM, Wetsel WC, et al. Different presynaptic roles of synapsins at excitatory and inhibitory synapses. J Neurosci. 2004;24(50):11368-80.
245. Fassio A, Raimondi A, Lignani G, Benfenati F, Baldelli P. Synapsins: from synapse to network hyperexcitability and epilepsy. Semin Cell Dev Biol. 2011;22(4):408-15.
246. Browning MD, Dudek EM, Rapier JL, Leonard S, Freedman R. Significant reductions in synapsin but not synaptophysin specific activity in the brains of some schizophrenics. Biol Psychiatry. 1993;34(8):529-35.
247. Nowakowski C, Kaufmann WA, Adlassnig C, Maier H, Salimi K, Jellinger KA, et al. Reduction of chromogranin B-like immunoreactivity in distinct subregions of the hippocampus from individuals with schizophrenia. Schizophr Res. 2002;58(1):43-53.
248. Schroeder M, Krebs MO, Bleich S, Frieling H. Epigenetics and depression: current challenges and new therapeutic options. Current opinion in psychiatry. 2010;23(6):588-92.
249. Akbarian S. Epigenetics of schizophrenia. Curr Top Behav Neurosci. 2010;4:611-28.
250. Akbarian S. The molecular pathology of schizophrenia--focus on histone and DNA modifications. Brain Res Bull. 2010;83(3-4):103-7.
251. Machado-Vieira R, Ibrahim L, Zarate CA, Jr. Histone Deacetylases and Mood Disorders: Epigenetic Programming in Gene-Environment Interactions. CNS Neurosci Ther. 2010.
252. Arent CO, Valvassori SS, Fries GR, Stertz L, Ferreira CL, Lopes-Borges J, et al. Neuroanatomical profile of antimaniac effects of histone deacetylases inhibitors. Mol Neurobiol. 2011;43(3):207-14.
253. Tsankova N, Renthal W, Kumar A, Nestler EJ. Epigenetic regulation in psychiatric disorders. Nat Rev Neurosci. 2007;8(5):355-67.
254. Kielland A, Erisir A, Walaas SI, Heggelund P. Synapsin utilization differs among functional classes of synapses on thalamocortical cells. J Neurosci. 2006;26(21):5786-93.
255. Berger SL. The complex language of chromatin regulation during transcription. Nature. 2007;447(7143):407-12.
256. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol. 2004;6(1):73-7.
257. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell. 2005;120(2):169-81.
258. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, et al. Active genes are tri-methylated at K4 of histone H3. Nature. 2002;419(6905):407-11.
259. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693-705.
260. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21(3):381-95.

261. Blumberg HP, Leung HC, Skudlarski P, Lacadie CM, Fredericks CA, Harris BC, et al. A functional magnetic resonance imaging study of bipolar disorder: state- and trait-related dysfunction in ventral prefrontal cortices. Arch Gen Psychiatry. 2003;60(6):601-9.
262. Quraishi S, Frangou S. Neuropsychology of bipolar disorder: a review. J Affect Disord. 2002;72(3):209-26.
263. Robinson JL, Monkul ES, Tordesillas-Gutierrez D, Franklin C, Bearden CE, Fox PT, et al. Fronto-limbic circuitry in euthymic bipolar disorder: evidence for prefrontal hyperactivation. Psychiatry research. 2008;164(2):106-13.
264. Malhi GS, Ivanovski B, Hadzi-Pavlovic D, Mitchell PB, Vieta E, Sachdev P. Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. Bipolar disorders. 2007;9(1-2):114-25.
265. Thompson JM, Gallagher P, Hughes JH, Watson S, Gray JM, Ferrier IN, et al. Neurocognitive impairment in euthymic patients with bipolar affective disorder. Br J Psychiatry. 2005;186:32-40.
266. Kruger S, Alda M, Young LT, Goldapple K, Parikh S, Mayberg HS. Risk and resilience markers in bipolar disorder: brain responses to emotional challenge in bipolar patients and their healthy siblings. The American journal of psychiatry. 2006;163(2):257-64.
267. Klempan TA, Rujescu D, Merette C, Himmelman C, Sequeira A, Canetti L, et al. Profiling brain expression of the spermidine/spermine N1-acetyltransferase 1 (SAT1) gene in suicide. Am J Med Genet B Neuropsychiatr Genet. 2009;150B(7):934-43.
268. Matevossian A, Akbarian S. A chromatin assay for human brain tissue. J Vis Exp. 2008(13).
269. Fiori LM, Turecki G. Epigenetic regulation of spermidine/spermine N1-acetyltransferase (SAT1) in suicide. J Psychiatr Res. 2011;45(9):1229-35.
270. Gitler D, Cheng Q, Greengard P, Augustine GJ. Synapsin IIa controls the reserve pool of glutamatergic synaptic vesicles. J Neurosci. 2008;28(43):10835-43.
271. Eastwood SL, Harrison PJ. Markers of glutamate synaptic transmission and plasticity are increased in the anterior cingulate cortex in bipolar disorder. Biol Psychiatry. 2010;67(11):1010-6.
272. Sanacora G, Zarate CA, Krystal JH, Manji HK. Targeting the glutamatergic system to develop novel, improved therapeutics for mood disorders. Nat Rev Drug Discov. 2008;7(5):426-37.
273. Chen G, Henter ID, Manji HK. Presynaptic glutamatergic dysfunction in bipolar disorder. Biol Psychiatry. 2010;67(11):1007-9.
274. Ongur D, Jensen JE, Prescot AP, Stork C, Lundy M, Cohen BM, et al. Abnormal glutamatergic neurotransmission and neuronal-glial interactions in acute mania. Biol Psychiatry. 2008;64(8):718-26.
275. Skoblenick KJ, Argintaru N, Xu Y, Dyck BA, Basu D, Tan ML, et al. Role of AP-2alpha transcription factor in the regulation of synapsin II gene expression by dopamine D1 and D2 receptors. Journal of molecular neuroscience : MN. 2010;41(2):267-77.
276. Hashimoto K, Sawa A, Iyo M. Increased levels of glutamate in brains from patients with mood disorders. Biol Psychiatry. 2007;62(11):1310-6.
277. Rao JS, Bazinet RP, Rapoport SI, Lee HJ. Chronic administration of carbamazepine down-regulates AP-2 DNA-binding activity and AP-2alpha protein expression in rat frontal cortex. Biol Psychiatry. 2007;61(2):154-61.
278. Damberg M, Ekblom J, Oreland L. Chronic pharmacological treatment with certain antidepressants alters the expression and DNA-binding activity of transcription factor AP-2. Life Sci. 2000;68(6):669-78.

279. Moreau MP, Bruse SE, David-Rus R, Buyske S, Brzustowicz LM. Altered microRNA expression profiles in postmortem brain samples from individuals with schizophrenia and bipolar disorder. Biol Psychiatry. 2011;69(2):188-93.
280. Kim AH, Reimers M, Maher B, Williamson V, McMichael O, McClay JL, et al. MicroRNA expression profiling in the prefrontal cortex of individuals affected with schizophrenia and bipolar disorders. Schizophr Res. 2010;124(1-3):183-91.
281. Miller BH, Wahlestedt C. MicroRNA dysregulation in psychiatric disease. Brain Res. 2010;1338:89-99.
282. Pieribone VA, Porton B, Rendon B, Feng J, Greengard P, Kao HT. Expression of synapsin III in nerve terminals and neurogenic regions of the adult brain. The Journal of comparative neurology. 2002;454(2):105-14.
283. Jiang Y, Langley B, Lubin FD, Renthal W, Wood MA, Yasui DH, et al. Epigenetics in the nervous system. J Neurosci. 2008;28(46):11753-9.
284. Miklem G, Hillier L. UCSC CpG islands track (unpublished) based on Gardiner-Garden M. and Frommer M. CpG islands in vertebrate genomes. J Mol Biol. 2012;196(2):261-82.
285. Murphy BC, O'Reilly RL, Singh SM. DNA methylation and mRNA expression of SYN III, a candidate gene for schizophrenia. BMC Med Genet. 2008;9:115.
286. Lopez JP, Mamdani F, Labonte B, Beaulieu MM, Yang JP, Berlim MT, et al. Epigenetic regulation of BDNF expression according to antidepressant response. Mol Psychiatry. 2012.
287. D'Addario C, Dell'Osso B, Palazzo MC, Benatti B, Lietti L, Cattaneo E, et al. Selective DNA methylation of BDNF promoter in bipolar disorder: differences among patients with BDI and BDII. Neuropsychopharmacology. 2012;37(7):1647-55.
288. Jope RS, Song L, Li PP, Young LT, Kish SJ, Pacheco MA, et al. The phosphoinositide signal transduction system is impaired in bipolar affective disorder brain. J Neurochem. 1996;66(6):2402-9.
289. Chang A, Li PP, Warsh JJ. cAMP-Dependent protein kinase (PKA) subunit mRNA levels in postmortem brain from patients with bipolar affective disorder (BD). Brain research Molecular brain research. 2003;116(1-2):27-37.
290. Donati RJ, Rasenick MM. G protein signaling and the molecular basis of antidepressant action. Life Sci. 2003;73(1):1-17.
291. Elhwuegi AS. Central monoamines and their role in major depression. Prog Neuropsychopharmacol Biol Psychiatry. 2004;28(3):435-51.
292. Carvajal C, Dumont Y, Quirion R. Neuropeptide y: role in emotion and alcohol dependence. CNS & neurological disorders drug targets. 2006;5(2):181-95.
293. Richard D, Lin Q, Timofeeva E. The corticotropin-releasing factor family of peptides and CRF receptors: their roles in the regulation of energy balance. European journal of pharmacology. 2002;440(2-3):189-97.
294. De Luca V, Tharmalingam S, Kennedy JL. Association study between the corticotropin-releasing hormone receptor 2 gene and suicidality in bipolar disorder. European psychiatry : the journal of the Association of European Psychiatrists. 2007;22(5):282-7.
295. Dumais A, Lesage AD, Alda M, Rouleau G, Dumont M, Chawky N, et al. Risk factors for suicide completion in major depression: a case-control study of impulsive and aggressive behaviors in men. The American journal of psychiatry. 2005;162(11):2116-24.
296. Cui P, Lin Q, Ding F, Xin C, Gong W, Zhang L, et al. A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. Genomics. 2010;96(5):259-65.
297. Schneider E, El Hajj N, Muller F, Navarro B, Haaf T. Epigenetic Dysregulation in the Prefrontal Cortex of Suicide Completers. Cytogenetic and genome research. 2015;146(1):19-27.

298. Betke KM, Wells CA, Hamm HE. GPCR mediated regulation of synaptic transmission. Prog Neurobiol. 2012;96(3):304-21.
299. Hamm HE. The many faces of G protein signaling. J Biol Chem. 1998;273(2):669-72.
300. McGirr A, Tousignant M, Routhier D, Pouliot L, Chawky N, Margolese HC, et al. Risk factors for completed suicide in schizophrenia and other chronic psychotic disorders: a case-control study. Schizophr Res. 2006;84(1):132-43.
301. McGirr A, Paris J, Lesage A, Renaud J, Turecki G. Risk factors for suicide completion in borderline personality disorder: a case-control study of cluster B comorbidity and impulsive aggression. J Clin Psychiatry. 2007;68(5):721-9.
302. Gittins R, Harrison PJ. A quantitative morphometric study of the human anterior cingulate cortex. Brain Res. 2004;1013(2):212-22.
303. Vogt BA, Nimchinsky EA, Vogt LJ, Hof PR. Human cingulate cortex: surface features, flat maps, and cytoarchitecture. The Journal of comparative neurology. 1995;359(3):490-506.
304. Hercher C, Parent M, Flores C, Canetti L, Turecki G, Mechawar N. Alcohol dependence-related increase of glial cell density in the anterior cingulate cortex of suicide completers. J Psychiatry Neurosci. 2009;34(4):281-8.
305. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes & development. 2011;25(18):1915-27.
306. Kim S, Webster MJ. The stanley neuropathology consortium integrative database: a novel, web-based tool for exploring neuropathological markers in psychiatric disorders and the biological processes associated with abnormalities of those markers. Neuropsychopharmacology. 2010;35(2):473-82.
307. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
308. Consortium SM-I. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. Nature biotechnology. 2014;32(9):903-14.
309. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.
310. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics. 2011;12:77.