UNDERSTANDING FACTORS THAT AFFECT GENE CO-REGULATION PATTERNS

BY

TÁSSIA MANGETTI GONÇALVES

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Animal Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Professor Sandra Rodriguez-Zas, Chair
Professor Gustavo Caetano-Anolles
Associate Professor Maria Bonita Villamil
Dr. Dianelys Gonzalez-Pena, Zoetis

ABSTRACT

The study of traits involved with behavior disorders and immune system can be complex and challenging. Gene expression and network analysis can help to elucidate the main molecules and pathways that influence complex traits. The goal of these studies is to understand factors that affect gene co-regulation patterns related to attention-deficit/hyperactivity disorder (ADHD), drug addiction, depression-like behavior, and immune challenge. ADHD is a common neuropsychiatric disorder that affects people of different ages. People with ADHD also have a predisposition to high levels of physical activity, and some drugs like amphetamine, are used to treat these type of disorders. The molecular mechanisms that contribute to the effectiveness of amphetamine therapy to ameliorate the symptoms of attention-deficit hyperactivity disorder (ADHD) are partially understood. The first experiment aimed to understand the molecular profiles supporting the ameliorating effect of amphetamine in a mouse hyperactivity line that exhibits ADHD-like behavior using mice striatum transcriptome. The findings support the development of therapies to ameliorate ADHD-like behaviors that target gene sub-networks while minimizing the disruption in other pathways triggered by amphetamine-based treatments. The second experiment implicated in understand the association between inflammation mechanism of action and behavioral disorders such as depression-like behavior. The tryptophan (Trp) metabolic route has been linked with chronic inflammatory diseases and psychiatric disorders, being tryptophan-degrading enzyme Indoleamine 2,3-dioxygenase (IDO1) activated during the inflammatory process. The goal of this study was to understand pathways associated with depression-like symptoms after recovery from immune challenge and deficiency of a gene associated with depression-like symptoms. The results confirmed that the genes identified can be used as potential candidates to understand the mechanisms involved in the infection leading to inflammation and depression-like behavior. The third study is a continuation of the first experiment, where a systems biology approach was used to elucidate mechanisms that regulate functional processes and pathways that are shared between genes found in the differential expression analyses. Differential co-expression analysis was also used to compare the structure of two co-expression networks between the conditions. Both weighted gene co-expression network analysis (WGCNA) and differential co-expression analysis were powerful alternatives to find more answers about complex systems. Differences in gene expression correlation across treatments and lines helped to understand changes in connectivity associated with the factors studied.

I dedicate this work to my whole family, especially to those who will always be examples to me: my grandparents Ozana Gonçalves Reis, Erotildes Ribeiro Reis *(in memoriam)*, Julva de Araujo Mangetti, and Silvio Mangetti *(in memoriam).*

ACKNOWLEDGMENTS

I would like to first thank my advisor, Dr. Sandra Rodriguez-Zas, for the opportunity of being a PhD student in her Lab, for her support, assistance, and friendship throughout the completion of my thesis project. Thank you to Dr. Bruce Southey for his help during my PhD, and to the members of my thesis committee, Dr. Caetano-Anolles, Dr. Villamil, and Dr. Gonzalez-Pena, for contributing with their intellectual ideas during my preliminary exam. My thanks also go to my lab-mate Pan and past lab-mate Harris, for the friendship, collaboration, and support. I would like to express my gratitude to the awesome University of Illinois at Urbana-Champaign and Animal Sciences Department, especially to Carolyn Thomas for her kindness, patience, attention, and help to all the graduate students of this Department, and to Terry Kizer for his friendship, sympathy, and happiness. Thank you to the graduate students of this Department for letting me be part of the Graduate Student Association (GSA) as their social chair, and especially to the GSA team for this experience, for their friendship, and partnership during the events and meetings.

I would like to thank my Brazilian family in Urbana-Champaign that were extremely important to me during this time living abroad. I was very fortunate to meet friends that I can trust, and that I can count on for everything. A huge special thank you goes to my favorite "Gaucha" Lidia Sbaraini Arend and my favorite "Cearense" and roommate Rachel Braz Arcanjo that where everything to me, giving me emotional support, affect, patience, laughs, I cannot say how much crucial they are in my life. I also need to thank my Brazilian friends that I met in UIUC: Ricardo, Mainity, Marcos, Hans, Cibele, Nycole, Mari, Eduardo, Leticia, Juliana, Massae, Murilo, Fernanda, and Lucas.

I want to thank Gabriel Costa for being not just my friend, but my inspiration, my example of person and professional. Thank you to Ligia Medeiros Miguel for so many years of friendship and support. Thank you to Pamela Alexandre for being the "twin sister" that the scientific life gave to me, sharing not just a similar appearance but also the same challenges and happiness of life, always at the same time! I also need to thank Salete Bonato and Greg Rien for the friendship and emotional support. I would like to thank the friends that I met in the Unity Church and Spiritual Center for the blessed moments shared. I was also fortunate enough to find my darling Tom Dart, who makes my life even happier, sharing amazing moments of life, giving me love and support.

I am very thankful to my parents Mauricio Ribeiro Gonçalves and Silvia Mangetti Gonçalves for always encouraging me to follow my dreams, even if it makes me be so far away from home. I consider myself very blessed for having my parents as my examples of faith, honesty, strength, and kindness; they taught me how to be respectful, independent, and ethics in my life. Many thanks for all the amazing moments shared with me, for giving me support, confidence, wise advice, and for believing in my potential. I am very proud to be their daughter. I also need to thank my sisters Tatiana, Thais for the friendship and happy moments in family, also Daniel and Giulia, and my whole family. Thanks to those who have contributed in some way to this work.

Finally, I want to finish my acknowledgments with this Steve Jobs inspirational thoughts: “My favorite thing in life don’t cost any money. It’s really clear that the most precious resource we all have is time.”

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

Introduction

Many complex traits such as behavioral disorders, immune challenge, gene knockout, and use of drugs in immune response are influenced by the level of RNA transcribed from the genes. Studies of gene expression are offering insights into the pathways that influence these traits. The aim of the study in Chapter 2 was to advance the understanding of the molecular profiles supporting the ameliorating effect of amphetamine in a mouse hyperactivity line that exhibits ADHD-like behavior using mice striatum transcriptome. Gene network analysis of the molecular pathways disrupted by amphetamine treatment in hyperactivity and control lines relative to saline treatment uncovered interactions between amphetamine treatment and ADHD-like hyperactivity line. In Chapter 3, gene differential expression and networks of peripheral macrophage after sickness recovery from BCG challenge relative to saline control in IDO1-KO relative to WT mice were explored. After the gene expression analysis studied in Chapter 2, Chapter 4 reported a systems biology approach to understanding the molecular basis of the simultaneous effects of hyperactivity line and amphetamine treatment. A review of approaches to measure and analyze gene expression and systems biology, and a review of traits that are influenced by gene expression variation follows.

Gene expression measurements using RNA-seq platform

RNA is a small molecule similar to DNA, but with few chemical modifications that change its properties relative to DNA, as the presence of ribonucleic acid. There are many classes of RNA: mRNA, tRNA, rRNA, and others. They are involved in the evolution of life and also play functional and regulatory roles in extant organisms (Brosius and Raabe, 2016).

Transcription is the process of copying DNA and creating mRNA. In eukaryotes, the mRNA, an intermediary molecule which carries the genetic information from the cell nucleus to the cytoplasm is converted to a protein sequence by the rRNA in a process called translation. Many proteins can be translated from a single mRNA molecule (Brosius and Raabe, 2016).

Transcriptome analysis with microarrays was widely used a few years ago. This methodology is still used, but it presents some limitations. The mRNA is converted into a stable cDNA form to allow the transcriptome study as mRNA degrades easily. Then, cDNA is labeled

with fluorochrome dyes Cy3 (green) and Cy5 (red) (Govindarajan et al., 2012). Complementary sequences are fixed in the chip, and they react with the cDNA targets resulting in a signal from the hybridization. The presence of artifacts from hybridization, poor dye-based detection, low detection of RNA splice patterns, and no discovery of new genes are some disadvantages of using microarray (Govindarajan et al., 2012).

A new way to measure transcriptome composition and to discover new exons or genes is by next generation high-throughput sequencing of cDNA (Mortazavi et al., 2008). This technique is called RNA-sequencing (RNA-seq), and it results in sequence reads that will be individually mapped to the source genome and counted to obtain the number and density of reads corresponding to RNA from each new candidate gene, known exon, or splice event. If samples have enough reads, for example above 40 million, it is possible to detect and quantify RNAs (Mortazavi et al., 2008).

There is a variety of different pipelines available for the analysis of RNA-seq, the choice will depend on research goals and organisms to be studied (Conesa et al., 2016). Also, there are some aspects that should be evaluated before RNA-seq analysis, like the design of the experiment, library type, depth of sequencing, and number of biological replicates. The analysis starts with the selection of the mRNA using 3' polyadenylated (poly (A)) from the total RNA. Then the mRNA is fragmented and processed in double-stranded cDNA, adapters are fixed at one end (single-end) or both ends (paired-end), this step is called cDNA library preparation (Wang et al., 2009). After sequencing, data is acquired containing raw reads that will pass for a quality filtering due to possible biases that could be introduced during PCR amplification or other steps of the library preparation. FastQC, FASTX-Toolkit, and Trimmomatic are popular software used to quality control, the first one generate graphics to visualize the quality, and the others are used for filtering low-quality reads or bases and remove adaptor sequences (Conesa et al., 2016).

RNA-seq technique improves genome annotation, allow discoveries of biological functions, and detection of alternative splices and new splices, resulting in more information of complex regulatory mechanisms of RNA (Martin and Wang, 2011). The depth coverage of RNA-seq decreases the cost of the detection of gene expression, which facilitates proteomic studies and comparisons of peptides found by mass spectrometry with their respective gene sequence. RNA-Seq has also been used as an efficient method to identify SNPs in transcribed regions from different species (Chepelev et al., 2009). One disadvantage of this methodology is the bioinformatics

knowledge and resources that are necessary to the analyses and the interpretation of huge volume of data.

Bioinformatics analysis of gene expression studies

Once the data passed the quality control, next step is to align reads against a reference genome or assemble them when there is no reference genome available. A frequent issue for the alignment step is that reads may map in multiple sequences of the genome due to repetitive sequences or common domains of paralogous genes (Conesa et al., 2016). The use of paired-end, in which both ends of the fragment are sequenced, is a more precise technique that facilitates reducing the number of alignment errors (Wang et al., 2009).

Among various software available for mapping, TopHat2 is one of the most popular. This method identifies splice sites by first locating non-junction reads (those contained within exons) using Bowtie (http://bowtie-bio.sourceforge.net), and then splitting the unmapped reads and aligning them independently to identify exon junctions (Langmead et al., 2009). Bowtie is a mapping program that indexes the reference genome using the Burrows-Wheeler method which scan reads against a genome (Burrows and Wheeler, 1994). There are sequenced parts with incorrect base calls due to sequencing errors in low coverage regions. The reference genome is used by TopHat2 to call bases and captures missing sequences at the end of exons. TopHat2 also controls the gaps presented in exons using a parameter that merges different single exons when necessary (Trapnell et al., 2009). Pseudogenes are present in the reference genome and have high similarity with some functional genes that contain introns. Reads can be aligned in the pseudogene version of these genes by mistake, also may fail to align to structural variations of the genome. TopHat2 was implemented with a new algorithm that allows the right alignment of reads in indels (insertions and deletions), inversions and translocations (Kim et al., 2013).

After the alignment, it is necessary to generate an assembly for each condition if one is using a similar method to TopHat2. There is a genome-guided method that takes advantage of longer read lengths; it is called Cufflinks2 (Trapnell et al., 2012). Cufflinks2 identify splice variants of genes and reports only the sets of isoforms that are compatible with read data, maximizing the precision of finding the most accurate combination of paths for assembling (Garber et al., 2011). Cufflinks2 quantify the expression level based on a statistical model, for genes that present multiple alternative splice events, its algorithm provides all the full-length

transcript fragments that are necessary to explain the splicing events. The normalization of the count is based on the length of each transcript, computed as FPKMs (fragments per kilo base of exon per million fragments mapped). Cufflinks2 algorithm also excludes less abundant transcripts that can be artifacts, and quantify abundances using a reference annotation (Trapnell et al., 2012). All the assemblies are merged using Cuffmerge, which generates a file that is used as the basis for calculating gene expression of each condition. Many of the concepts and algorithms used for Cuffmerge are the same that was used in Cufflinks assembles of individual reads. It is also possible to integrate Cuffmerge merged assemble with a reference genome annotation (Trapnell et al., 2012).

The merged assembly file is used for differentially expression analysis. Cuffdiff2 algorithm accurately identifies gene and isoform expression and consider the variability present across biological replicates (Trapnell et al., 2012). The consequences of not controlling for variability are the over-prediction of differentially expressed genes or transcripts and high false-positive rates. Errors can also be introduced due uncertainty in the presence of alternative splicing or repetitive regions. Cuffdiff2 addresses these problems by modeling variability of fragments abundance. It implements a Poisson model where the estimated variability is the mean count across replicates. The P-value is calculated for any observed change in the count of fragments. It identifies differential expression at transcriptional or post-transcriptional level (Trapnell et al., 2012).

A different approach to analyze RNA-Seq data is using Kallisto (Bray et al., 2016). Aligning to a reference genome, and estimating the transcript abundances can be time-consuming in the analysis of RNA-seq. Kallisto is a two order methodology that is based on a pseudoaligment of reads to a reference genome, it combines information from k-mers within reads, and avoids to align individual bases, instead, it produces a list of transcripts that are compatible with each read. This program makes alignment using transcriptome de Bruijn graph constructed from k-mers (Compeau et al., 2011). For quantification, the method includes a model for bias and the expectation-maximization algorithm (Patro et al., 2014). The advantages of this software are the speed, the accuracy, and the simplicity.

For analysis of similar methods to Kallisto, it is possible to integrate transcript-level abundance estimates into count-based statistical inference using an R package called tximport (Soneson et al., 2015). In this software, evidence showed that the use of gene-level is more statistically robust than transcripts. This methodology analyzes gene-level abundance but takes

advantage of transcript-level abundance estimates to improve differentially expression results. The package tximport imports transcript length and abundance estimates from Kallisto, for example, and export count matrices (Soneson et al., 2015). There is also an option to average transcript length correction terms and apply them to common statistical R packages that use counts as input to analyze differential expressions like edgeR (Robinson et al., 2010) and DESeq2 (Love et al., 2014).

Both edgeR and DESeq2 are software packages available in R, the input data is a table of counts, where columns correspond to samples and rows to genes or transcripts (Love et al., 2014; Robinson et al., 2010). These packages determine if there are significant differences between counts across experimental conditions. To account for biological and technical variability edgeR uses an overdispersed Poisson model and empirical Bayes to improve the reliability of inference (Robinson et al., 2010). In DESeq2, each gene is described with a generalized linear model (GLM) of the negative binomial distribution, but it also uses empirical Bayes shrinkage for dispersion estimation (Love et al., 2014). The difference between DESeq2 and edgeR is that the first one estimates the width of the prior distribution from the data and controls the amount of shrinkage based on the observed properties of the data. The second one requires the prior degrees of freedom to weight the contribution of each gene and for the dispersion fit (Love et al., 2014). The advantage of using DESeq2 and edgeR is that they both provide the flexibility to analyze more complex designs, and present high sensitivity of the algorithms to controlled type-I error (Love et al., 2014). **Figure 1** shows a brief workflow of bioinformatics tools used to perform the analysis in the Chapters of this thesis.

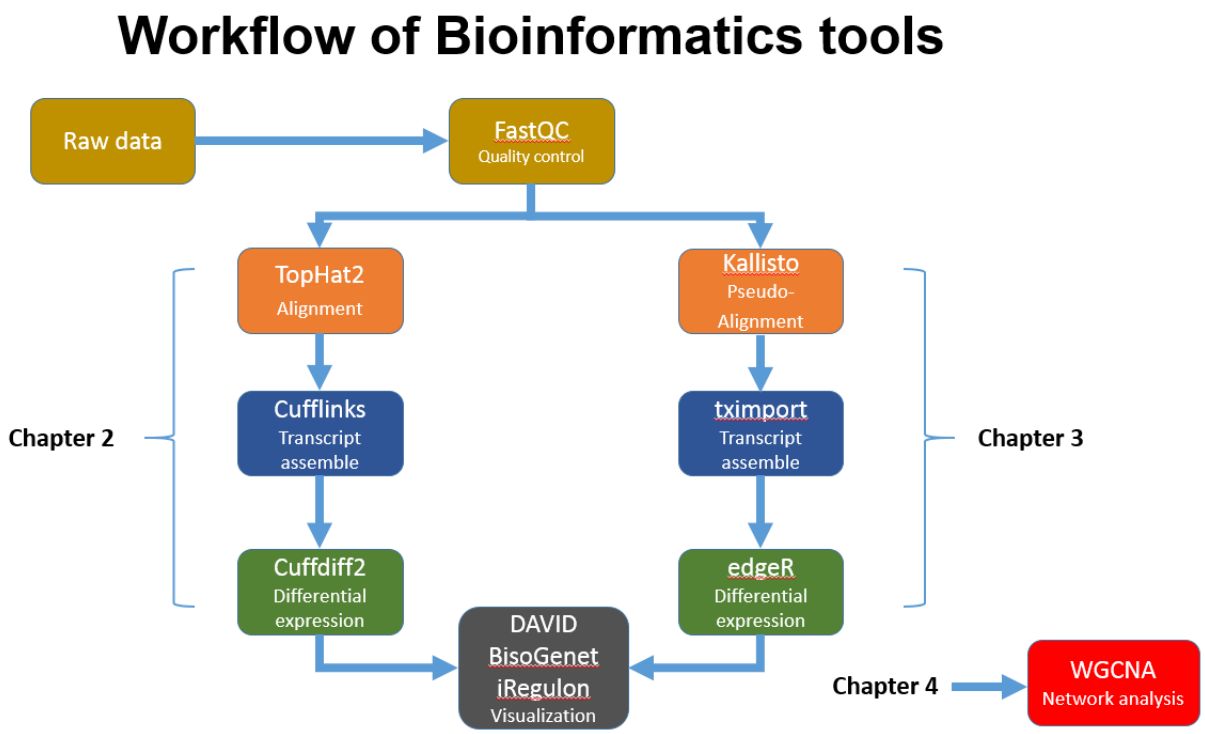

Figure 1. Workflow of bioinformatics tools used in the analysis of each Chapter.

The Kallisto methodology

Kallisto is considerate an ultra-fast alignment-free method, where alignments are not necessary to find the origins of reads (Bray et al., 2016). Kallisto does pseudo-alignments of reads by using hashing of k-mers (hash function is used for rapid data lookup, mapping data of any size to data with fixed size). Then, Kallisto combines the hashing k-mers with de Bruijn graph to perform the pseud-alignment, called transcriptome de Bruijn graph (T-DBG), which will be used to make the comparison of the reads to the transcripts (Bray et al., 2016). After the T-DBG is built, Kallisto maps each k-mer to the contig by using a hash table (a hash map that associate keys to values) and create the Kallisto index (Bray et al., 2016). The association of each k-mer in the T-DBG with transcripts will belong to a determined class called k-compatibility class. The intersection of all k-compatibility classes that reads are associated will give the transcript that the read is compatible (Bray et al., 2016). Kallisto takes advantage of the redundant information along the path of the graph and does not consider k-mers that have same equivalent class because the k-mers are redundant and provide no new information. Then, the intersection of the k-compatibility classes contains only non-redundant k-mers (Bray et al., 2016).

In comparison with Kallisto, Cufflinks has an alignment-dependent algorithms and is the most popular tool that makes novel transcript and isoforms discovery and quantification (Zhang et al., 2017). Although alignment-free methods like Kallisto are fastest for requiring less computational time, both methods showed strong concordance for highly expressed transcripts (Zhang et al., 2017). Cufflinks showed inferior accuracy performance than Kallisto, and also, FPKMs from Cufflinks produces largest variations in the estimation of quantification and underestimates the expression of absent transcript (Zhang et al., 2017). In general, Kallisto present more advantages related to speed and accuracy compared to Cufflinks to perform RNA-Seq analysis.

Interpretation of gene expression analyses through Annotation, Enrichment, and Networks

The use of technology like high-throughput next generation sequencing to 'omics' studies, usually result in large output files of genes. Understanding the biological meaning of large volume of data may be challenging. The software DAVID (the Database for Annotation, Visualization, and Integrated Discovery) allows analyzing many types of gene lists, coming from diverse platforms (Huang et al., 2009). It is a computational procedure that explores the biological meaning of the data. It provides annotation, enrichments, and have integrated data-mining environment. One can submit a list of genes for different purposes as genes that contain the same characteristics, genes for expression studies with fold changes >|2| or p-values <0.05, genes up or down regulated or lists of genes with an important statistical threshold to search for biological processes. Larger gene lists result in higher statistical power, higher sensitivity, consequently having more significant P-values and more enriched terms (Huang et al., 2009). One disadvantage of using DAVID is that it gets very limited if the gene list is small. The principle of the enrichment in DAVID analysis is that higher potential genes (enriched) are selected based on the background list and type of species that is provided. The enrichment is quantified by more than one method: Fisher's exact test, Binomial probability, and Hypergeometric distribution. Enriched Gene Ontology (GO) terms, for example, are the most statistically over-represented terms of the list, which increases the likelihood to identify important biological processes (Huang et al., 2009).

According to Huang et al. (2008), the enrichment score for groups is "the geometric mean of all the enrichment p-values (Expression Analysis Systematic Explorer - EASE scores) for each annotation term associated with the gene members in the group." EASE scores are calculated based

on Fisher exact test (Gonzalez-Pena et al., 2016a). To be a relative score, the P-values suffer minus log transformation. Higher group score results in more enriched terms for gene members of that group. To control for false discovery rate, DAVID provides three types of multiple testing corrections for the enriched P-values: Bonferroni, Benjamini, and FDR (Huang et al., 2008). For the Functional Annotation Chart, the significance of enriched clusters is evaluated with a modified Fisher's exact test (EASE score) (Huang et al., 2008).

The Gene Set Enrichment Analysis (GSEA) is another analytical method to interpret gene expression data. As DAVID, it also focuses on a set of genes that share the same biological functions or patterns. The difference between them is that instead of just use genes that show largest differences to determine the biological meanings, GSEA uses all genes that are expressed on the list. In GSEA, it is considerate a gene set *S* (originated from prior studies and biological knowledge) and a list of genes *L*. The aim is to analyze if genes of the sets tend to go toward to the top or to the bottom of the list, showing if it is correlated with phenotype and its distribution (Subramanian et al., 2005).

GSEA method has three principles: first, is the calculation of an enrichment score (ES), which means the level in where the sets *S* is overrepresented at the thorough list. The score increases when genes are part of the set, and it decreases when genes are not present in the set along the whole list. The ES is the maximum deviation from zero found on the list and is equivalent to a weighted Kolmogorov–Smirnov-like statistic (Hollander and Wolfe, 1999). Second, there is the estimation of the significance level of ES, where the nominal p-value of ES is calculated using an empirical phenotype-based permutation test method that maintains the correlation structure of data, generating a null distribution for the *ES* (Subramanian et al., 2005). Finally, the adjustment for multiple hypothesis testing is when the whole list of genes is evaluated, first it is normalized to account for the size of the set (normalized enrichment score - *NES*), then false positives are controlled by FDR (Subramanian et al., 2005). Also, significance levels are estimated for both positive and negative scoring gene sets.

Networks have been used for data visualization and to find biological-meaningful answers for the data (Martin et al., 2010). Networks are used as a graph representation of molecules interaction in a known relationship, where nodes and edges are used to represent functional relations and type of entity, respectively. Estimated values associated with P-values resulted from differential gene expression analysis might be integrated, then the size of node indicates the

differential expression P-value and colors represent over- or under-expression between treatments (Martin et al., 2010; Shannon et al., 2003). There are several tools for construction and visualization of networks as VisANT (Hu et al., 2005), Osprey (Breitkreutz et al., 2003), and BisoGenet (Martin et al., 2010). VisANT provides a visual interface, annotation, integrates different types of biological information based on association and connectivity data. Osprey is a similar application and uses a wide range of platforms; it can incorporate new interactions into existing networks, and also uses connectivity levels to filter networks. However, these methodologies present some limitations, for example, Osprey is not efficient for large-scale network analysis and VisANT's visual integration tools do not take any combination of operations in any order (Pavlopoulos et al., 2008). Bisogenet is a new tool for analyzing networks; it is different from the others because it can represent coding relations, for example, finding multiple isoforms of a gene as a result of alternative splicing (Martin et al., 2010). This program provides information of disease-related and tissue-specific, more regulatory information as transcription factors, and microRNA silencing relations, and consequently a more complete systems biology tool (Martin et al., 2010). BisoGenet is used as a Cytoscape plugin, and the information of annotation and pathways come from SysBiomics, a database that integrates many public data sources (Martin et al., 2010; Shannon et al., 2003). The statistical is based on cluster coefficient, and degree of the node depends on neighbor's relationship and distance of paths (Martin et al., 2010).

The use of gene expression analysis, with the study of different statistical methods and bioinformatics tools and finally network analysis are provided to understand how biological information underlies complex traits such the molecular architecture of behavioral disorders and drug addiction.

Weighted Gene Co-expression Network Analysis and differential co-expression analysis

Studies of systems biology can be applied to transcriptome data to elucidate more complex mechanisms of regulation (Langfelder and Horvath, 2007). The use of systems biology has been widely applied to diseases (Sengupta et al., 2009). Methods of correlation network analysis contribute to finding patterns among genes across the samples. Gene co-expression networks have been used to detect modules that are biologically meaningful and organized according to functional

processes and pathways. A methodology that uses gene co-expression is the Weighted Gene Co-expression Network Analysis (WGCNA) (Langfelder and Horvath, 2007).

The concept for fundamental networks varies in three types: 1) networks not related to modules, 2) network properties of a module (intramodular), and 3) relationship between modules (intermodular) (Zhang and Horvath, 2005). Gene co-expression networks only identify correlations, indicating what genes act simultaneously in the same processes. Networks can explain the relationship between *n* nodes, which can be any biomolecule like genes or proteins. Networks are specified by an *n*x*n* dimensional adjacency matrix A = (Aij), being Aij the connection strength from node i to j (Zhang and Horvath, 2005). The gene-gene similarity matrix describes the similarity between expression patterns across samples. Measures of correlation are applied to construct the networks (van Dam et al., 2017).

The resulted co-expression associations are used to construct modules. Modules are groups of co-expressed genes, identified using clustering methodologies as hierarchal clustering (van Dam et al., 2017). This clustering method divides each cluster into sub-clusters, resulting in a tree with branches where the branches are modules. An eigengene is a vector that best describes the patterns of all genes inside each module. As co-expressed genes are functionally related, functional enrichment analysis can be used to interpret the list of genes in each module. Final results will depend on the heterogeneity of samples, co-expression analysis decrease the statistical power to detect modules when limited to specific conditions to identify common co-expression patterns (van Dam et al., 2017).

Differential co-expression analysis detects modules that cannot be detected in the regular co-expression or differential expression methods. Co-expression networks identify correlations, and genes active in the same biological process, while the differential co-expression analysis identifies co-expressed genes under different conditions, such as treatments, identifying modules unique to the specific type of treatment (van Dam et al., 2017). Therefore, genes identified in this analysis are more likely to explain the differences between phenotypes. **Figure 2** shows the workflow of the WGCNA used in Chapter 4.

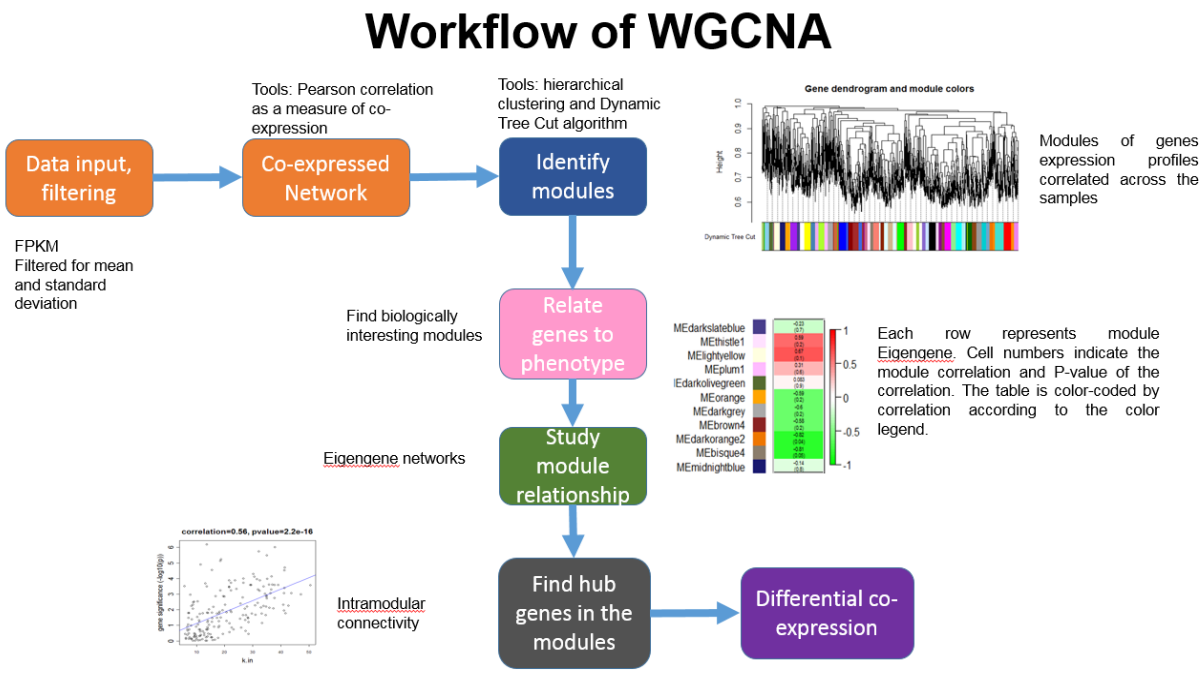

Figure 2. Workflow of WGCNA, including each step of the analysis.

Complex behavior and immune traits

Attention deficit/hyperactivity disorder and Drugs of Abuse

Attention deficit/hyperactivity disorder (ADHD) is a common mental disorder that occurs during the childhood and may continue into adulthood. It has a substantial impact on someone's life and despite it is characterized by high level of inattention, increased impulsivity and hyperactivity the diagnosis may not be so easy (Castellanos and Proal, 2012). The genetic predisposition for hyperactivity was already related to ADHD; the estimated heritability is approximately 76% in humans (Faraone et al., 2005). ADHD prevalence rate in the world is estimated at 5.9–7.1% for children and adolescents (Willcutt, 2012).

People with ADHD also have a predisposition to high levels of physical activity (hyperactivity). These levels of physical activity are influenced by complex genetic factors that are associated with behavioral diseases, which require further elucidations of neurobiological mechanisms. A mouse line selected for increased home cage activity (Hyperactivity line) is an effective model to understand genetic factors involved in neurobiological processes influencing behavioral disorders (Majdak et al., 2014). This line emerged from a selective breeding method to produce animal models with extreme values allowing the study of genetic disorders. This selection

was made in a population over many generations to understand how genes related to neurobiological circuits increase levels of physical activity and consequently the features of ADHD (Zombeck et al., 2011).

For the breeding selection, increased physical activity was measured by voluntary wheel running, open field behavior and forced treadmill running in mice. These traits are involved with neural circuits and act in different regions of the brain. The selection for increased home cage activity started from two populations: Hsd:ICR and Collaborative Cross G2:F1 (CC) mice (Chesler et al., 2008). Hsd:ICR was used because it was already explored for a selective breeding experiment for increased voluntary wheel running behavior (Swallow et al., 1998). The Collaborative Cross mice were used because it increases the genetic diversity due to the eight inbred strains that derived it: A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/H1LtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ (Beck, et al. 2000; Zombeck et al., 2011). Based on these populations emerged two lines, High-Active and Control (Zombeck et al., 2011).

After generations, the High-Active mice still demonstrated increased hyperactivity, being used as a model to study ADHD (Majdak et al., 2014). This study however also had another goal which was to demonstrate that the use of different doses of d-amphetamine can be therapeutic and ameliorate the behavioral symptoms of ADHD. Rhodes et al. (2004) (Rhodes et al., 2004) showed that exposure to methylphenidate made mice selectively bred to diminish running levels, while control lines that also received the same treatment increased wheel running. Majdak et al. (2014) (Majdak et al., 2014) had as a result that High-Active mouse line can be used to study neurobiological mechanisms related to hyperactivity suggesting that there is a response to amphetamine administration to it.

The administration of different drugs had already shown addiction effects using striatum tissue. Drug administration affects dopamine and serotonin, neurotransmitters that can be responsible for developing the addiction. Studies showed the increase in total serotonin concentration in brain tissues including striatum after drug (cocaine) administration in rats (Kirby et al., 2011; Krzascik et al., 2015). In a study of meta-analysis using in vivo imaging of amphetamine and methamphetamine to understand the dopamine system and dependence process in humans, the amphetamine decreased striatal dopamine transporter availability (Ashok et al., 2017). Drugs addiction effect is related to the striatum and its neurotransmitters.

Some genes already showed an important role in ADHD, a study of genotyping in Chinese children and adolescent suggested that *SLC6A1* gene variants may have a significant effect on the ADHD risk (Yuan et al., 2017). Another study with mice also showed that *SLC6A1* gene knock-out ($^{-/-}$) had hyperactivity and impaired sustained attention as phenotype (Chen et al., 2015). Growth factor receptor-bound protein isoform 10 (GRB10) is highly expressed in many brain regions rich in dopaminergic and serotonergic neurons, which are associated with ADHD (Franke et al., 2009). The gene cadherin 13 (CDH13) was associated with addiction vulnerability to substances like methamphetamine; it is expressed in regions of the brain that are related to ADHD and addiction behavior (Uhl et al., 2008).

Understanding immune response

The immune system is composed of dynamic and functional networks of specialized cell types and tissues that provide complex molecular and cellular events during an immune response. Gene expression analysis has been applied to all cell types of immune system, like cytokines and chemokines, to find the interaction between immune system components at the molecular level (Furman and Davis, 2015).

Gene expression and bioinformatics tools have also been used to addiction-related traits to help to elucidate mechanisms of the immune system and how they are influenced by multiple genetic and environmental factors. Differentially expressed genes identify genes in regions depending on the magnitude of the behavioral trait complexity. These analyses are important to find genetic markers that influence amphetamine-related responses and identify genes and their precise function in a given behavior, however, they are also challenging due to the complexity of drug-related traits. Using genomic approaches, it was possible to link the action of amphetamine in the dopaminergic system and find dopamine-related genes to drug-response (Phillips et al., 2008).

Another method is the use of gene knockout on behavioral responses to drugs and immune system. Addiction studies found answers related to signaling and cell support molecules involved in complex drug reward-related traits (Phillips et al., 2008). Gene expression analyses take advantage of this technology allowing to understand the difference between transgenic animals with a knockout and the wild-type version to study the immune system.

Advances in technology also allowed researchers to make detailed studies of the immunoglobulin (Ig) and T-cell receptor (TCR) repertoires responding to vaccines. The development of computational tools to integrate models to analyze human immune response accelerates the understanding of how the immune system works and consequently accelerates vaccine and other medical treatments (Furman and Davis, 2015).

The use of high-throughput sequencing approaches can be applied to the immune system to gives insights and identify immune biomarkers (Furman and Davis, 2015). The goal of these studies is to use information from gene expression analysis to understand the molecular basis of two complex traits: the simultaneous effect of amphetamine and ADHD, and immune response of macrophages.

Inflammation in depression and ADHD

Studies have been shown the relationship between inflammation and other conditions such as psychiatric disorders (Mitchell et al., 2017). Children and adolescents present high prevalence of conditions related to inflammation, like obesity, asthma, and allergy (Mitchell et al., 2017). Inflammation is a complex trait that can have a broad variety of reactions, including systemic reactions, migration and activation of leukocytes, and vascular responses (Mitchell et al., 2017). The key modulators of inflammation are the cytokines, which are also mediators between the immune system and the central nervous system (Mitchell et al., 2017). Studies show the link between inflammation, pro-inflammatory cytokines, and depression in adult humans (Dowlati et al., 2010). Interferon (IFN)-a therapy can induce depression in patients treated for hepatitis C (Lotrich et al., 2009). People with depression present higher concentration of interleukin 6 (Il-6) and TNF-a relative to no depressed people (Mitchell et al., 2017). Inflammation was also associated with bipolar disorder, schizophrenia, and ADHD (Mitchell et al., 2017). Low levels of Il-10 are associated with ADHD and schizophrenia (Mittleman et al., 1997). Increased Il-16 and Il-13 were related to ADHD symptoms of hyperactivity and inattention in humans, respectively (Oades et al., 2010).

ADHD and depression might be sharing common causative pathways. One example is the kynurenine metabolites that can modulate neurotransmission and have neuroprotective or neurotoxic properties (Myint et al., 2007). Pro-inflammatory cytokines induce the enzyme

indoleamine 2,3-dioxygenase (IDO) (Christmas et al., 2011). The increase of IDO1 activity may increase neurotoxic metabolites of the kynurenine pathways and cause tryptophan depletion, causing depression (Christmas et al., 2011). Metabolites of tryptophan/kynurenine pathway and cytokines activities were also related to children with ADHD symptoms in a medicated group (Oades et al., 2010).

ADHD in children might be related to maternal care during pregnancy, affecting the kynurenine levels. Maternal smoking during pregnancy is related to decreasing levels of kynurenine, and increased levels of 3-hydroxy-kynurenine (3HK) in children with ADHD (Oades, 2011). Medications and supplementations taken during pregnancy are also related to children that had ADHD compared to mothers of healthy children. Increasing levels of kynurenic acid in children with ADHD was associated with stressful events that occurred during pregnancy (Oades, 2011). Tryptophan levels tend to decrease in children with ADHD that consume seafood and spend time with parents (Oades, 2011). Also, increased child care in these cases also decreases tryptophan levels and cytokines Il-2, IFN-y, and S100 calcium Protein B (S100B). Diet with fish might be beneficial to a kid with ADHD and increase pro-inflammatory IL-2 and IFN-y (Oades, 2011). The tryptophan metabolism and cytokines may be influencing the occurrence of both ADHD and depression symptoms.

The overarching goal of this thesis is to understand factors that affect gene co-regulation patterns. Complementary methodology was applied to two RNA-seq data sets. One experiment aimed at understanding the pathways associated with high voluntary activity genotype and the central nervous system stimulant amphetamine. The other experiment aimed at understanding the pathways associated with depression-like symptoms after recovery from immune challenge and deficiency of a gene associated with depression-like symptoms. The first experiment profiled the striatum and the second experiment profiled macrophages in mice.

CHAPTER 2 - GENE NETWORK EVIDENCE OF INTERACTION BETWEEN AMPHETAMINE TREATMENT AND HYPERACTIVITY IN A MOUSE MODEL OF ATTENTION-DEFICIT HYPERACTIVITY DISORDER

Abstract

The psychostimulant amphetamine is frequently used in treatments to ameliorate the symptoms of attention-deficit hyperactivity disorder (ADHD). The molecular mechanisms that contribute to the effectiveness of this therapy are partially understood. To address this, the striatum of mice from a line (H) selected for hyperactivity that models ADHD and a contemporary control line (C) receiving either amphetamine treatment (A) or saline (S) treatments were studied using RNA-Seq. Among the 1,498 genes that exhibited significant (FDR adjusted P-value < 0.05) treatment-by-line interaction effect. These genes included a high number of neuropeptide coding genes and genes associated with neurogenesis, memory and neuroplasticity, cell adhesion, angiogenesis, and synaptic signaling. Lesser frequent genes are exhibiting interaction effects associated with energy and redox processes that exhibited opposite differential expression between A and S treatments across H and C lines genes. Network analysis uncovers a unique association between the synaptic signaling genes Nrgn and Rims1 that could explain the therapeutic effect of amphetamine on ADHD symptoms. Our study of gene co-regulatory network uncovered one mode of action by which amphetamine ameliorates the antagonistic disruption of Arc and Shank in the hyperactive line that exhibits ADHD-like behaviors. The network module including Shank1, Arc, Rims1, Dlgap1, Rapgef4, and Dlgap1, Rapgef4 suggests that treatment A decreases the differential expression of genes between the H and C lines by targeting whole subnetworks of connected genes, rather than isolated genes. Our findings support the development of therapies to ameliorate ADHD-like behaviors that target gene sub-networks while minimizing the disruption in other pathways triggered by amphetamine-based treatments.

Introduction

Amphetamine is a synthetic psychological and metabolic stimulant that enhances concentration and awareness. Individuals exposed to amphetamine exhibit over-activity levels disrupted sleep patterns, and reduced appetite (Weyandt et al., 2014). The effects of this stimulant have been harnessed in treatments to ameliorate the symptoms of behavioral disorders including attention-deficit hyperactivity disorder (ADHD) and in weight loss programs (Berman et al., 2009). Amphetamine usage has an increased risk for development of tolerance, abuse, dependence, and addiction (Weyandt et al., 2014). Both ADHD status and amphetamine effects involve dysregulation of neurological and molecular mechanisms in the striatum (Burns et al., 1993; Castellanos and Proal, 2012), a brain region that controls locomotor activities, voluntary behavior, spatial memory, and response and addiction to psychostimulants (Gruber and McDonald, 2012; Mizumori et al., 2009; Yager et al., 2015).

Given the therapeutic effect of amphetamine treatment on ADHD, mice "knocked out" for individual genes in the amphetamine addiction pathway have been used to understand the molecular mechanisms underlying ADHD. For example, heterozygous knockout mice for dopamine transporter hypofunctional (DAT+/-) exhibited hyperactivity, attentional and impulsivity deficits and molecular disruptions and these deficits were ameliorated with a low-dose chronic amphetamine treatment (Ashok et al., 2017; Mereu et al., 2017). Brain-derived neurotrophic factor (Bdnf) serves as a neurotransmitter and plays a role in the etiology of ADHD. Likewise, a conditional knockout of Bdnf ADHD exhibited high locomotor hyperactivity and ADHD behavioral characteristics (Sagvolden et al., 2005; Tsai, 2016). Another model encompasses mice null for adhesion G protein-coupled receptor L3 (Adgrl3), a gene associated with ADHD and susceptibility to addiction. Mice from this line took more time to become immobile and were immobile for shorter periods in a forced swimming test than wild-type mice and 11 genes associated with neuron structure and function were differentially expressed between these groups (Orsini et al., 2016).

Selective breeding strategies that address the polygenic nature of ADHD offer effective models of this disorder (Majdak et al., 2016; Yen et al., 2013). One ADHD model consists of a mouse line selected for hyperactivity not only run and traveled more in the cage but also underperformed in the open field behavior and Morris water maze relative to the control line (Majdak et al., 2016). A chronic low-dose amphetamine treatment reduced the hyperactivity and

motor impulsivity and the activation of neurons in the prefrontal cortex and cerebellum in the selected line (Majdak et al., 2016). This treatment had the opposite effect on the activation of neurons in the control line suggesting a potential interaction between amphetamine treatment and activity line. Another ADHD model involves a mouse line selected for low trait anxiety-related behavior that displays hyperactivity in home cages (Yen et al., 2015). In this model, acute low-dose amphetamine treatment was associated with a reduction of locomotor activity and inhibition of the activity of glycogen synthase kinase 3beta (Gsk3b) in the medial prefrontal cortex. A microarray study identified 75 genes differentially expressed in the spontaneously hypertensive rat line when exposed to amphetamine (dela Peña et al., 2015). These 75 genes were associated with angiogenesis, cell adhesion, apoptosis, and neural development.

The objective of this study is to advance the understanding of the molecular profiles supporting the ameliorating effect of amphetamine in a mouse hyperactivity line that exhibits ADHD-like behavior. A comprehensive study of the molecular patterns disrupted by amphetamine treatment in the striatum of a mouse model of ADHD provides a complete understanding of the mode of action of amphetamine therapy. This model provides insights into the dysregulation of molecular mechanisms associated with amphetamine (independent of hyperactivity line) and into dysregulation associated with hyperactivity line (independent of amphetamine treatment). Moreover, our comparative gene network analysis of the molecular pathways disrupted by amphetamine treatment in hyperactivity and control lines relative to saline treatment uncover interactions between amphetamine treatment and ADHD-like hyperactivity line.

Material and methods

Animal experiments

Generation 17 mice from a hyperactivity line selected for increased home cage activity (line H) and from a contemporary control line (line C) were studied (Majdak et al., 2014; Majdak et al., 2016; Zombeck et al., 2011). This selective breeding line constitutes a model for ADHD-like behaviors. Home cage activity is individually measured using video recording at approximately two months of age as the average distance traveled during two days (Zombeck et al., 2011). By generation 16, male H mice exhibited significantly higher activity (distance traveled, wheel running) and motor impulsivity (measured using false alarms of the Go/No-go test) than C

mice. Moreover, H mice exhibited significant lower motor coordination and learning skills measured using the accelerating rotarod task (Majdak et al., 2016).

Male mice were group-housed after weaning, and cages were placed in rooms monitored for temperature (21 ± 1 °C) and photo-period (12:12 Light:Dark; lights off at 8:00 AM). Food (Harlan Teklad, 7012) and water were provided *ad libitum*. At approximately five months of age, drug-naive mice were acclimated to their home cage for four days. The average activity traveled during days 2 and 3 of individual mice were used as a baseline measurement to confirm that the mice studied were representative of the H and C lines. On day 4, 1.5 hours into the Dark cycle (active phase for mice), one intraperitoneal injection of either 0.25 mg/kg of d-amphetamine (treatment A) or equal amount of saline (treatment S) was given to each mouse in the home cage over 3 days (Majdak et al., 2016). Two hours after injection, mice were decapitated followed by immediate removal of the brain. In total, 20 mice were analyzed across the 2 activity lines and the two amphetamine treatments. The Illinois Institutional Animal Care and Use Committee approved all animal procedures, and these procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

RNA-Seq analysis

Striatum was extracted and dissected on a chilled aluminum block, snap frozen on dry ice, and stored in a centrifuge tube at -80^{o}C following our published protocols (Caetano-Anollés et al., 2016; Saul et al., 2017). The RNase-free disposable pellet pestle (Fisher Scientific, Pittsburgh, PA, USA) was used for tissue homogenization, and RNeasy1 Lipid Tissue Mini Kit (Qiagen, Valencia, CA, USA) was used for RNA extraction. The isolated RNA was purified with DNase I (Qiagen, Valencia, CA, USA). Total RNA yield measured using Qubit1 2.0 (Life Technologies, Carlsbad, CA, USA) was > 14 ug per sample. The RNA Integrity Number (RIN) measured using Fragment Analyzer (Advanced Analytical Technologies Inc., Ankeny, IA, USA) was > 8 for all samples. Libraries from individual mouse striatum samples were prepared using Illumina's TruSeq Stranded RNAseq Sample Prep kit (Illumina Inc., San Diego, CA, USA) and 100nt-long paired-end reads from each animal separately were obtained. The libraries were sequenced using Illumina HiSeq 2500, and a TruSeq SBS sequencing kit version 4 (Illumina, San Diego, CA, USA) and the FASTQ files containing paired-end reads of length 100nt were generated and demultiplexed with the software Casava 1.8.2 (Illumina, San Diego, CA, USA).

The average Phred quality score of the reads was assessed using the software program FastQC (Andrews, 2010). The nucleotide quality score was > 30 across all read positions and was not trimmed (Nixon et al., 2015). The software program TopHat2 v2.1.1 was used to map the reads to the mouse genome assembly Genome Reference Consortium GRCm38 (Pruitt et al., 2006; Trapnell et al., 2009). Transcripts were assembled and abundance estimated using the software program Cufflinks (Trapnell et al., 2012) with default specifications: bias correction; weight reads mapping to multiple locations in the genome; upper quartile normalization of the number of fragments mapping to individual loci; use of reference genome annotation to provide additional information for assembly (Trapnell et al., 2012). Cuffdiff was used to test the changes in gene expression between groups (Trapnell et al., 2012).

Comparison among line-treatment combination groups offers insights into the interaction between line and treatment (Caetano-Anollés et al., 2015). Mouse groups are identified with a 2-letter acronym: the first letter denotes the line and the second letter denotes the treatment: H line and S treatment (HS group); H line and A treatment (HA group); C line and S treatment (CS group); and C line and A treatment (CA group). These four line-treatment groups enable the evaluation of six pairwise contrasts. The contrasts HA-HS and CA-CS offer information on gene dysregulation associated with amphetamine treatment within the H and C lines, respectively; HS-CS and HA-CA offer information on gene dysregulation associated with hyperactivity within the A and S treatments, respectively; and HS-CA and HA-CS offer information on gene dysregulation associated with simultaneous effects of hyperactivity line and amphetamine treatment. Among the six possible contrasts between hyperactivity line-amphetamine treatment groups, the three orthogonal contrasts that will be interpreted are CA-CS, HA-HS, HS-CS. Comparison between lines (irrespective of treatment) and between treatments (irrespective of the line) offer insights into the main effects of line and treatment. All genes with 5 or more reads per line-treatment group were analyzed to ensure adequate representation across groups. The Benjamini-Hochberg false discovery rate (FDR) was used to adjust the differential expression P-value for multiple testing (Benjamini and Hochberg, 1995).

Functional and network analysis

Functional analysis was used to aggregate information on the level and pattern of gene expression profiles between hyperactivity line-amphetamine treatment groups. Two

complementary analyses were undertaken. The enrichment of functional categories among the differentially expressed genes was evaluated using a hypergeometric test within the web-service Database for Annotation, Visualization and Integrated Discovery (DAVID; Version 6.8) (Huang et al., 2009). The enrichment of functional categories was also evaluated using a Kolmogorov test on the standardized expression profile of all genes available in the Gene Set Enrichment Analysis (GSEA) software (Subramanian et al., 2005). The functional categories assessed included Gene Ontology (GO) biological processes (BP), molecular functions (MF), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The Functional Annotation Tool (FAT) categories corresponding to the GO terms available in DAVID were used to minimize redundancies across the GO hierarchy (Caetano-Anollés et al., 2016). Categories were clustered in DAVID, and the geometric mean of the enrichment P-values (Enrichment Score, ES) was used as evidence supporting the category clusters. The mouse genome was used as background for testing and FDR was used to adjust the enrichment P-values for multiple testing (Pruitt et al., 2006).

Gene networks were reconstructed to uncover gene dysregulation on a system-level basis, using Bisogenet package within the Cytoscape environment (Delfino et al., 2011; Killcoyne et al., 2009; Martin et al., 2010; Serao et al., 2013). These networks depict gene relationships based on protein-protein interactions annotated in BIOGRID, HPRD, DIP, BIND, INTACT, and MINT databases (Keshava Prasad et al., 2008; Mathivanan et al., 2006). Networks encompassing 75 genes exhibiting significant differential expression in at least one of the 3 orthogonal contrasts between hyperactivity line-amphetamine treatment groups (CA-CS, HA-HS, HS-CS) were reconstructed to facilitate the comparison of disrupted networks across contrasts and interpretation (Caetano-Anollés et al., 2016; Gonzalez-Pena et al., 2016a; Gonzalez-Pena et al., 2016b). All genes in the final networks were associated with functional categories enriched among the genes exhibiting a significant interaction effect. The comparison and study of the expression profiles within networks reveal shared and distinct co-regulation patterns and complement the information from functional analysis.

Results and discussion

The sequencing of the striatum libraries produced a total of over 2.1 billion paired-end reads of length 100nt and on average 109 million reads per sample. Approximately 49 million pairs of reads were aligned to the mouse reference genome per sample on average amounting to

89.54% of concordant read pairs aligned. Approximately 19,450 genes were tested for gene expression analysis across the six contrasts between hyperactivity line-amphetamine treatment groups.

Amphetamine treatment-by-hyperactivity line interaction effects on the transcriptome

Our study of differential gene expression between hyperactivity line- amphetamine treatment groups, offered insights into transcriptome disruptions that are line- and amphetamine-dependent. Overall, 1,498 genes exhibited significant (FDR adjusted P-value < 0.05) treatment-by-line interaction effect. **Table 1** highlights the top differentially expressed genes exhibiting significant interaction effect (FDR-adjusted P-value < 1.xE-7). Within pairwise contrast, the numbers of significantly differentially expressed (FDR adjusted P-value < 0.05) genes were: 494 genes (HS-CS), 186 genes (HA-HS), 319 genes (HA-CA), 93 genes (CA-CS); 208 genes (HA-CS), and 1,084 genes (HS-CA).

Table 1. Most significant genes exhibiting amphetamine treatment-by-hyperactivity line interaction effect and log2(fold change) between combinations of treatment (A=amphetamine, S=saline) and line (H=hyperactivity, C=control) groups.

	Log2(Fold Change)[a]						Overall	
Gene Symbol	**CA-CS**	**HA-HS**	**HS-CA**	**HA-CS**	**HS-CS**	**HA-CA**	**Raw P-value**	**FDR P-value[b]**
LOC100861610	4.20	1.26	-1.73	3.73	2.47	-0.47	7.8E-13	7.0E-10
Neurod6	0.79	1.62	-1.99	0.41	-1.20	-0.38	7.8E-13	7.0E-10
3110035E14Rik	0.77	1.10	-1.45	0.41	-0.68	-0.36	7.8E-13	7.0E-10
C1ql3	0.65	1.01	-1.26	0.40	-0.61	-0.25	7.8E-13	7.0E-10
Prss12	0.65	1.09	-1.52	0.22	-0.88	-0.43	7.8E-13	7.0E-10
Nr4a2	0.57	1.26	-1.55	0.29	-0.98	-0.29	7.8E-13	7.0E-10
Ctgf	0.63	0.85	-1.42	0.06	-0.79	-0.57	7.8E-13	7.0E-10
Cck	0.42	1.22	-1.39	0.25	-0.97	-0.18	7.8E-13	7.0E-10
Nov	0.50	1.28	-1.64	0.14	-1.14	-0.36	7.8E-13	7.0E-10
Slc17a7	0.45	1.58	-1.87	0.16	-1.42	-0.29	7.8E-13	7.0E-10
Nptx1	0.40	0.80	-1.09	0.11	-0.69	-0.30	7.8E-13	7.0E-10
Hap1	-0.34	-0.42	0.79	0.02	0.45	0.36	7.8E-13	7.0E-10
Dlk1	-0.46	-0.71	1.24	0.07	0.78	0.53	7.8E-13	7.0E-10

Table 1 (cont.)

	Log2(Fold Change)[a]						Overall	
Gene Symbol	**CA-CS**	**HA-HS**	**HS-CA**	**HA-CS**	**HS-CS**	**HA-CA**	**Raw P-value**	**FDR P-value[b]**
Peg10	-0.52	-0.62	1.14	-0.01	0.62	0.51	7.8E-13	7.0E-10
Irs4	-0.87	-0.59	1.68	0.23	0.81	1.09	7.8E-13	7.0E-10
Gm9800	-1.54	1.46	0.25	0.17	-1.29	1.70	7.8E-13	7.0E-10
Avp	-1.70	-1.51	3.14	-0.07	1.44	1.63	7.8E-13	7.0E-10
Gm5506	-0.51	0.83	-1.86	-1.54	-2.37	-1.04	7.8E-13	7.0E-10
Oxt	-2.41	-1.65	3.66	-0.40	1.26	2.01	7.8E-13	7.0E-10
Cplx3	0.65	1.08	-1.46	0.26	-0.81	-0.39	1.6E-12	1.2E-09
Tbr1	0.45	1.25	-1.53	0.17	-1.08	-0.28	1.6E-12	1.2E-09
Sv2b	0.30	0.70	-0.95	0.05	-0.65	-0.26	1.6E-12	1.2E-09
Adcy1	0.31	0.76	-1.02	0.04	-0.72	-0.26	2.3E-12	1.5E-09
AW551984	-0.66	-0.48	1.33	0.19	0.67	0.84	2.3E-12	1.5E-09
Arhgap36	-0.80	-0.74	1.74	0.19	0.93	0.99	2.3E-12	1.5E-09
Gm11549	0.59	1.44	-1.78	0.25	-1.19	-0.34	3.1E-12	1.9E-09
Scube1	0.43	1.33	-1.53	0.23	-1.10	-0.20	6.8E-12	4.2E-09
Il1f9	-0.50	-1.05	1.13	-0.42	0.63	0.08	1.1E-11	6.2E-09
Lingo1	0.30	0.59	-0.80	0.08	-0.51	-0.21	1.1E-11	6.4E-09
Neurod2	0.43	1.54	-1.89	0.08	-1.46	-0.35	1.6E-11	8.6E-09
Satb2	0.47	1.43	-1.68	0.21	-1.22	-0.25	1.9E-11	9.6E-09
Galnt9	0.35	0.67	-0.86	0.16	-0.51	-0.19	1.9E-11	9.6E-09
Trh	-1.17	-0.57	2.20	0.47	1.03	1.64	1.9E-11	9.6E-09
Ttr	-1.63	-0.63	2.31	0.05	0.68	1.68	2.4E-11	1.2E-08
Nrn1	0.31	1.22	-1.26	0.26	-0.96	-0.05	3.5E-11	1.6E-08
Mical2	0.28	0.73	-0.81	0.19	-0.53	-0.09	3.6E-11	1.6E-08
Calb2	-0.70	-0.38	1.64	0.56	0.94	1.25	4.3E-11	1.9E-08
Pde1a	0.36	0.52	-0.66	0.23	-0.30	-0.14	4.9E-11	2.1E-08
Igsf1	-0.67	-0.65	1.30	-0.02	0.63	0.65	5.4E-11	2.3E-08
Mpped1	0.26	0.55	-0.68	0.12	-0.42	-0.13	5.9E-11	2.4E-08
Rtn4rl2	0.42	1.65	-2.03	0.04	-1.61	-0.38	9.9E-11	4.0E-08
Magel2	-0.50	-0.89	1.43	0.05	0.93	0.55	1.2E-10	4.6E-08
Resp18	-0.46	-0.30	0.91	0.14	0.45	0.60	1.3E-10	4.8E-08
Nts	-0.37	-0.67	1.09	0.05	0.72	0.42	1.5E-10	5.6E-08
Ecel1	-0.29	-0.49	0.77	-0.01	0.48	0.28	1.6E-10	5.7E-08
Baiap3	-0.68	-0.46	1.14	0.00	0.46	0.67	1.7E-10	6.0E-08
Mef2c	0.21	0.57	-0.73	0.05	-0.52	-0.16	1.7E-10	6.0E-08
Fam70a	-0.32	-0.42	0.85	0.11	0.52	0.43	1.7E-10	6.0E-08
Cdhr1	1.31	0.85	-0.93	1.23	0.38	-0.08	1.8E-10	6.0E-08

Table 1 (cont.)

	Log2(Fold Change)[a]						Overall	
Gene Symbol	**CA-CS**	**HA-HS**	**HS-CA**	**HA-CS**	**HS-CS**	**HA-CA**	**Raw P-value**	**FDR P-value[b]**
Serpini1	0.21	0.56	-0.66	0.12	-0.44	-0.09	1.8E-10	6.0E-08
Robo3	0.33	0.88	-1.10	0.11	-0.77	-0.23	1.9E-10	6.3E-08
3110047P20Rik	0.31	0.79	-1.12	-0.03	-0.82	-0.33	2.2E-10	6.9E-08
Prlr	-0.54	-0.57	1.30	0.19	0.76	0.73	2.3E-10	7.2E-08
Hba-a2	-1.42	-0.60	3.66	1.64	2.24	3.07	2.4E-10	7.2E-08

[a]Log_2(fold change) between different pairs of treatment-line groups: H=hyperactivity line; C=control line; A=amphetamine treatment; and S=saline treatment.

[b]False Discovery rate adjusted P-value for the overall line-by-treatment interaction.

The Venn diagram in **Figure 3** depicts the overlap of differentially expressed genes (FDR-adjusted P-value < 0.05) within each of four informative contrasts (CA-HA, CS-HS, CA-CS, and HA-HS). This diagram highlights the majority of the genes in the contrasts with the highest numbers of differentially expressed genes were unique to that contrast: 248 genes in CS-HS amounting to 33.4% of all differentially expressed genes, and 144 genes in CA-HA amounting to 19.4% of all differentially expressed genes. In addition, a large number of differentially expressed genes (107 genes amounting to 14.4% of all differentially expressed genes overlapped between both contrasts suggesting a prevalent gene dysregulation that is associated with differences between the H line model of ADHD-like behaviors and the C line.

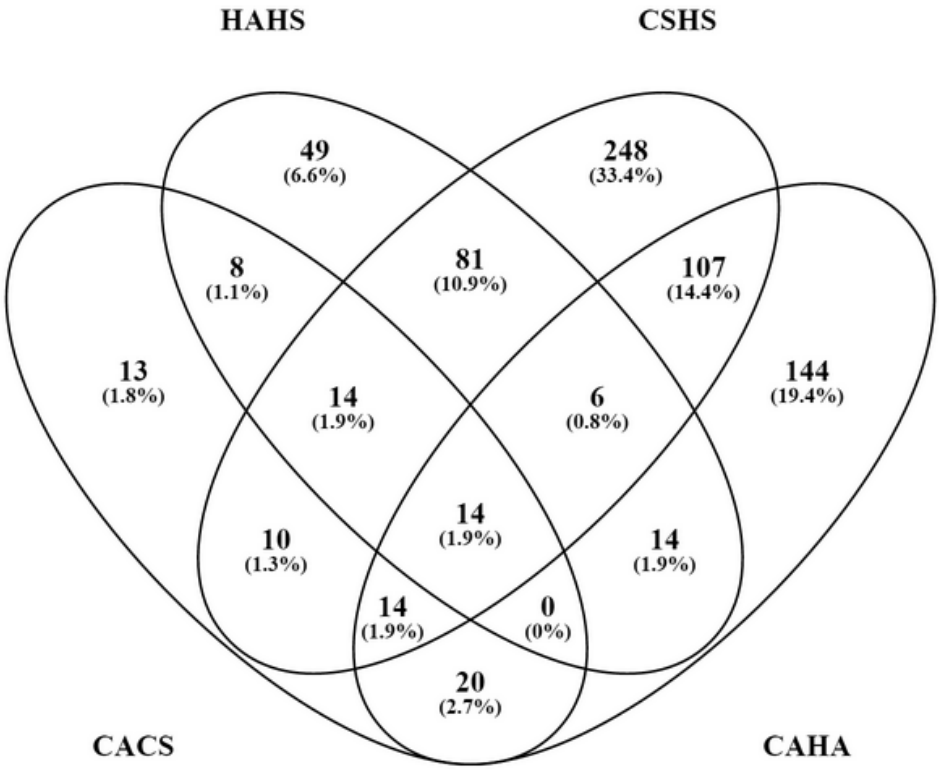

Figure 3. Venn diagram summarizing the number of differentially expressed genes (FDR-adjusted P-value < 0.05) within each of four informative contrasts between combinations of treatment (A=amphetamine, S=saline) and line (H=hyperactive, C=control) groups analyzed separately and the overlap. CAHA=CA-HA, CSHS=CS-HS, CACS=CA-CS, HAHS=HA-HS)

The lowest number of differentially expressed genes were detected in the CS-CA contrast, and 13 of these genes (1.8% of all differentially expressed genes) did not overlap with another contrast. This suggests that the A treatment used elicits modest disruption in the transcriptome compared to the S treatment in the C line. Most of the differentially expressed genes in the CA-CS contrast (52 genes amounting to 7% of the differentially expressed genes) overlapped with the CS-HS contrast. Both of these contrasts share the CS mouse group, and thus the response of these genes to CA and HS relative to CS support the effectiveness of the hyperactivity line model of ADHD to understand the responsive to amphetamine treatment.

As expected, the lower number of differentially expressed genes in HA-HS (186 genes) relative to CS-HS (494 genes) indicates that gene expression in HS is more similar to HA than to

CS mice. The Venn diagram highlights the large percentage of the differentially expressed genes in HA-HS that overlap with CS-HS (115 genes amounting to 15.5% of the differentially expressed genes). Overall, the overlap of differentially expressed genes between HA-HS and CS-HS indicates that the A treatment reverted some of the gene dysregulation associated with the hyperactivity ADHD-like line. This conclusion is also supported by the number of differentially expressed genes (319 genes) in HA-CA that is intermediate between HA-HS and HS-CS. The previous results are consistent with the therapeutic effects of low-dose amphetamine on ADHD-like symptoms and the interaction between the same amphetamine treatments and lines at earlier generation assessed at behavioral and protein levels. In early generations of the H line studied, amphetamine treatment lowered the locomotor activity in the line and increased the activity in the C line (Majdak et al., 2014). In later generations, a significant interaction between treatment and line on impulsivity was reported (Majdak et al., 2016). A significant interaction between anxiety behavior and amphetamine treatment on the activity of glycogen synthase kinase 3 beta (GSK3B) was also reported in another ADHD model that used a line selected for low trait anxiety behavior and amphetamine treatment (Yen et al., 2015).

The Venn diagram identifies 14 genes that were differentially expressed in the four contrasts (**Figure 3**). These genes are: Rho GTPase activating protein 36 (Arhgap36), arginine vasopressin (Avp), expressed sequence AW551984 (AW551984), BAI1 associated protein 3 (Baiap3), connective tissue growth factor (Ctgf), delta like non-canonical Notch ligand 1 (Dlk1), enolase 1B, retrotransposed (Gm5506), predicted gene 9800 (Gm9800), Huntingtin associated protein 1 (Hap1), immunoglobulin superfamily member 1 (Igsf1), insulin receptor substrate 4 (Irs4), neuronal pentraxin 1 (Nptx1), oxytocin/neurophysin (Oxt), and paternally expressed 10 (Peg10). From these 14 genes, 10 are among top 20 genes with lowest FDR P-value: Ctgf, Nptx1, Hap1, Dlk1, Peg10, Irs4, Gm9800, Avp, Gm5506, and Oxt. Also, most of the genes (Hap1, Dlk1, Peg10, Irs4, Avp, AW551984, Arhgap36, Igsf1, and Baiap3) present the same interaction profile "- - +". Ctgf and Nptx1 present the opposite profile "+ + -", while Gm9800 and Gm5506 present "- + -" profile. These interaction profiles, and especially the genes Avp, and Oxp are discussed in the next sections.

Interaction profiles

The sign of the log2(fold change) corresponding to the three orthogonal contrasts enabled the precise characterization of the interaction profiles. The "+" or "-" sign of each denotes over-expression or under-expression respectively in the first mice group relative to the second mouse group specified in the contrast. Profiles are expressed regarding the signs of the three contrasts in the following order: CA-CS, HA-HS, and HS-CS. For example, the profile "+ + -" corresponds to gene positive CA-CS (i.e., over-expression in CA relative to CS), positive HA-CS (i.e., over-expression in HA relative to HS), and negative HS-CS (i.e., under-expression in HS relative to CS mice). Alternatively, this profile can also be summarized as positive A-S treatment contrast across both lines and negative HS-CS. Likewise, the profile "- - +" corresponds to negative CA-CS, negative HA-HS, and positive HS-CS log2(fold change) respectively.

Profile "+ + -" was the most common profile (470 genes) among the 1,498 genes exhibiting significant treatment-by-line interaction effect. The "+ + -" profile includes positive CA-CS, positive HA-HS, and negative HS-CS. The frequencies of the remaining ordered profiles were 378 genes "- - +" (negative CA-CS, negative HA-HS, and positive HS-CS); 256 genes "+ - +" (positive CA-CS, negative HA-HS, and positive HS-CS); 167 genes "- + -" (negative CA-CS, positive HA-HS, and negative HS-CS); 97 genes "- - -" (negative CA-CS, negative HA-HS, and negative HS-CS); 80 genes "+ + +" (positive CA-CS, positive HA-HS, and positive HS-CS); 32 genes "- + +" (negative CA-CS, positive HA-HS, and positive HS-CS); and 17 genes "+ - -" (positive CA-CS, negative HA-HS, and negative HS-CS).

The predominance of the 2 opposite profiles "+ + -" and "- - +" across the 3 orthogonal contrasts CA-CS, HA-HS, and CS-HS highlights the impact of the differential gene expression between A and S treatments across both lines and the correspondence between A and S treatments with C and H lines receiving the S treatment. These profiles may trigger molecular mechanisms elucidating reports that low doses of amphetamine can decrease wheel running in mouse lines exhibiting standard activity levels (Williams and White, 1984) in addition to decreasing activity in hyperactive lines.

Amphetamine treatment-by-hyperactivity line interaction effects on individual gene expression profiles

The fundamental study of individual gene profiles thrust our understanding of the molecular mechanisms that exhibited amphetamine treatment and ADHD model line-specific profiles. An unexpected finding was the high number of genes coding for neuropeptides and hormone, and associated receptors exhibited either one of the two most common yet opposite profiles ("+ + -" and "- - +") across the three orthogonal contrasts CA-CS, HA-HS, and CS-HS (**Table 1**). Genes positive for the A-S treatment contrast in both lines and negative for HS-CS ("+ + -" profile) included: cholecystokinin (Cck), neurogranin (Nrgn), vasoactive intestinal polypeptide (Vip), somatostatin receptor 3 (Sstr3), and cholecystokinin B receptor (Cckbr). Genes negative for the A-S contrast in both lines and positive for HS-CS ("- - +" profile) included: Avp and Oxt (both common between the four contrasts in the Venn diagram presented in **Figure 3**), thyrotropin-releasing hormone (Trh), neurotensin (Nts), angiotensinogen (Agt), galanin (Gal), prolactin receptor (Prlr), and calcitonin receptor (Calcr). Our novel identification of a large number of hormone and neuropeptide genes exhibiting amphetamine treatment-by-hyperactivity line interaction is confirmed across numerous reports of dysregulation of neuropeptide genes, in particular, ADHD-related or amphetamine scenarios. Neuropeptides and hormones modulate the dopaminergic neurotransmitter systems and influence the effects of amphetamine. Nts blocks the hyperactivity induced by amphetamine treatment in rats (Boules et al., 2001). The observed over-expression of Nrgn in the amphetamine-treated mice relative to saline is consistent with increased levels of Nrgn in the frontal cortex of amphetamine-treated rats (Szabo et al., 2009). Cck diminishes the locomotor stimulatory effects of low amphetamine doses in rats (Weiss et al., 1988). Cck diminishes the locomotor stimulatory effects of low amphetamine doses in rats. Amphetamine exposure alters Oxt, and dopamine neurotransmission in prairie voles (Young et al., 2014) and Oxt lowers the locomotor activity in cocaine- and methamphetamine-exposed mice (Carson et al., 2013; Leong et al., 2016; Qi et al., 2008; Sarnyai and Kovács, 1994). Avp also reduces locomotor stimulation by amphetamine in mice (Chiu et al., 1998).

Two thyroid-related genes exhibited the "- - +" profile characterized by negative A-S treatment contrast in both lines: Trh and transthyretin (Ttr). Trh plays a role in stimulating the release of the thyroid-stimulating hormone, and Ttr plays a role in thyroid synthesis. Thyroid

hormones influence behavior and cognitive function and individuals that have lower levels of this hormone exhibit behaviors similar to ADHD (Builee and Hatherill, 2004). The under-expression of Trh in the A relative to the S group is consistent with lower levels of Trh in the striatum of rats 2 hours after amphetamine administration (Jaworska-Feil et al., 1995). Our findings demonstrate the potential impact of neuropeptide-driven therapies to ameliorate hyperactivity ADHD-like behaviors while minimizing the disruption in other pathways triggered by amphetamine-based treatments.

In addition to neuropeptide- and thyroid-related genes, other genes exhibiting interaction effect have been associated with amphetamine or other psychostimulants, and to ADHD or related comorbidities (**Table 1**). Genes participating in synaptic signaling that exhibited the "+ + -" profile characterized by positive CA-CS and HA-HS and negative HS-CS included: Bdnf, 5-hydroxytryptamine (serotonin) receptor (Htr5a), adenosine A1 receptor (Adora1), synapsin I (Syn1), syntaxin 1A (Stx1a), and synaptotagmin I, XIII and XVII (Syt1, Syt13, and Syt17). The profile of Bdnf is consistent with the well-established positive association between Bdnf levels and amphetamine administration (Meredith et al., 2002) and with higher striatal levels of Bdnf in wild-type mice relative to heterozygous Bdnf knockout mice after acute amphetamine injection (Saylor and McGinty, 2008). Consistent with the profile uncovered in this study, mutations and other disruptions in serotonin receptors including Htr5a have been linked to ADHD (Romanos et al., 2008). Also, Htr5a is over-expressed in the nucleus accumbens of mice exposed to the psychostimulant cocaine (Eipper-Mains et al., 2013). Likewise, in agreement with the pattern detected in this study, mutations in Adora1 have been associated with anxiety (Johansson et al., 2001) and the psychostimulant caffeine affects the level of Adora1 (Muñiz et al., 2016). Adora1 has also been associated with regulation of blood pressure and blood vessel diameter, hormone secretion, synaptic depression, and cognition (Yang et al., 2010). Ours is the first study to simultaneously link hyperactivity and amphetamine exposure to Hrt5a and Adora1.

Many genes associated with neuronal development presented the "+ + -" profile characterized by positive A-S treatment contrast across both lines and negative HS-CS (**Table 1**). These genes include transcription factor nuclear receptor subfamily 4, group A, member 2 (Nr4a2), and genes in the neuronal differentiation (Neurod) family (Neurod6, Neurod2, and Neurod1). Additional genes associated with neuronal development that exhibited positive CA-CS and negative HS-CS contrast included the transcription factors early growth response 2 (Egr2) and

members of the nuclear receptor (Nr) subfamily (Nr4a3, and Nr4a1). Egr2, Nr4a1, and Nr4a3 were also differentially expressed in a study of amphetamine-treated rats from the spontaneously hypertensive model of ADHD (dela Peña et al., 2015). The transcriptome study of the homozygous Lphn3 knockout model of ADHD uncovered disruption on genes that play a role in neuron structure and function (Orsini et al., 2016). Other transcription factors including SRY-box 1 (Sox1) and Forkhead box J1 (Foxj1) exhibited the reverse pattern, characterized by under-expression in A relative to S in both lines and were over-expressed in HS-CS (**Table 1**). Sox1 is associated with forebrain neuron development, and Foxj1 has been linked to brain development (Christie et al., 2013).

Genes participating in memory and learning processes that exhibited a significant interaction effect and a "+ + -" profile included Bdnf, neurogenic differentiation 2 (Neurod2), and Nrgn (**Table 1**). The positive A-S treatment contrast across lines exhibited by neurotrypsin (Prss12) is consistent with the higher focus observed in individuals exposed to amphetamine while Prss12 knockout mice exhibit impaired long-term memory formation and cognitive deficits and behavioral disorders in humans and mice (Schachtschneider et al., 2016). Our study confirms that ADHD and psychostimulant use are involved in disruption of neuronal development and cognition processes (Webb et al., 2009) and furthers the understanding that this effect expands across hyperactivity genotypes associated with ADHD.

In addition to genes associated with neurogenesis and cognitive processes, several ion transport genes associated with neuroplasticity presented the "+ + -" profile (**Table 1**). These genes include various ATPase Ca++ transporters (e.g., Atp2b1), potassium channel family members (e.g., Kcnk2), and solute carrier family members (e.g., Slc17a7). The solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 7 (Slc17a7, also known as vesicular glutamate transporter or VGluT1) plays a role in behavioral flexibility and adaptation (Granseth et al., 2015) that is typically associated with neuroplasticity in response to psychostimulants and ADHD-related disorders.

Several genes that participate in cell adhesion processes exhibited significant amphetamine treatment-by-hyperactivity line interaction effect including connective tissue growth factor (Ctgf), adhesion G protein-coupled receptor L3 (Adgrl3 or Lphn3), and various members of the collagen (Col) family: Col9a3, Col24a1, Col6a5, Col11a1, Col5a2, Col6a2, Col19a1, Col6a3, Col8a2, Col27a1, Col4a2, Col9a2, Col16a1, and Col5a1 (**Table 1**). The "+ + -" profile of Ctgf

characterized by positive A-S treatment contrast across both lines is consistent with the profile observed in amphetamine-treated rats from the spontaneously hypertensive model of ADHD (dela Peña et al., 2015). Adgrl3 or Lphn3 also exhibited the "+ + -" profile. This gene is also involved in cell-cell adhesion and has been associated with susceptibility to ADHD, addiction processes (Liu et al., 2006) and a homozygous Lphn3 knockout mouse line are used as ADHD model (Orsini et al., 2016). Nephroblastoma overexpressed gene (Nov), together with plexin domain containing 1 (Plxdc1) and R-spondin 3 (Rspo3) have the same profile as collagen genes Col4a2 and Col8a2, and all these genes are associated with angiogenesis processes. Consistent with our findings, genes involved in cell adhesion angiogenesis were differentially expressed in amphetamine-treated rats from the spontaneously hypertensive model of ADHD (dela Peña et al., 2015). Our findings ratify the proposition that physical activity stimulated by A treatment supports brain vasculature and blood flow and increases angiogenesis, and neurogenesis (Gapin et al., 2011).

Dopamine receptor 3 (Drd3) is among the genes exhibiting the second most frequent profile ("- - +" profile) that is characterized by negative A-S treatment contrast in both lines and positive HS-CS (**Table 1**). Drd3 is associated with regulation of blood pressure, behavioral response to the psychostimulant cocaine, and locomotor behavior similar to genes exhibiting the reverse trajectory. Our Drd3 finding is in agreement with the reported striatal over-expression of Drd3 in the spontaneously hypertensive rat model of ADHD relative to control rats (Li et al., 2007).

Among the third most common profile ("- + -" profile), ribosomal protein L29 (Rpl29) and Enolase 1B, retrotransposed (GM5506 or Eno1b) presented negative CA-CS, positive HS-CS, and negative HS-CS. The under-expression in H relative to C lines under S treatment is consistent with reports under-expression of Rpl29 in the Lphn3 null model of ADHD relative to wild-type mice (Orsini et al., 2016). Eno1b has been associated with neuronal diseases and is part of the glycolytic pathway that plays a role in ATP production in the brain (Sultana et al., 2006). The profile of Eno1b observed in the present study is consistent with the role of enolases in ATP production in the brain, where low energy availability in the brain may lead to cellular dysfunction (Sultana et al., 2006). LOC100042025 predicted as glyceraldehyde-3-phosphate dehydrogenase-like, transcript variant 2, also plays a role in glycolysis and follows a similar profile of positive HA-HS and negative HS-CS albeit no substantial expression differential in CA-CS.

Glyceraldehyde 3-phosphate dehydrogenase (Gapdh) exhibited less significant interaction effect including positive HS-CS. Gapdh participates in redox post-translational modifications (El

Kadmiri et al., 2014) and mice exposed to methamphetamine treatment exhibit higher GAPDH protein levels and had more oxidative stress in the brain (Mark et al., 2007; Nakajima et al., 2009). Amphetamine exposure is also associated with increased oxidative stress in rodent striatum (Frey et al., 2006; Wan et al., 2000). Also, in vitro studies detected that Gapdh increases cellular toxicity and enhances pathogenicity of intercellular transmission of Huntington disease (Mikhaylova et al., 2016). The energy required by neuronal processes to sustain high activity in H mice or in response to treatment A may trigger the dysregulation of the previously reviewed genes involved in energy pathways.

S-adenosylmethionine decarboxylase 2 (Amd2) and Glutathione S-transferase Pi 2 (Gstp2) were positive CA-CS, negative HA-HS, and positive HS-CS. The profile of these genes in related experiments is consistent with our findings. Amd2 is part of the polyamine system, and polyamine treatment has been associated with modulation of limbic dopamine function that is involved climbing and wheel running behavior in rodents (Hirsch et al., 1987). Gstp2 plays an important role in detoxification, protect cells from injury by toxic chemicals and from products of oxidative stress (Hayes et al., 2005). Forms of Gstp2 ameliorate the oxidative milieu especially prevalent in dopaminergic neurons (Beiswanger et al., 1994). Drd3 exhibited the same negative HA-HS and positive HS-CS profile as Gstp2 confirming the impact of H line and A treatment in the dopamine system.

Three related transcription factors, FBJ osteosarcoma oncogene (Fos), FBJ osteosarcoma oncogene B (Fosb), and fos-like antigen 2 (Fosl2) exhibited the most frequent profile characterized by positive CA-CS, negative HA-HS, and positive HS-CS. Our result agrees with reports of elevated Fosb expression in the striatum of rats after exposure to the psychostimulant Methylphenidate (Ritalin) that is commonly used to treat ADHD and with findings of Fos induction associated with amphetamine exposure in young mice (Chase et al., 2005; Labandeira-Garcia et al., 1994).

Our study of treatment-by-line interaction effects on gene expression enabled us to detect transcript profiles that are simultaneously amphetamine-treatment and hyperactivity line-dependent. This information will aid in the identification of molecular targets that can effectively ameliorate ADHD symptoms in consideration of psychostimulant treatments or alleviate the effects of psychostimulant use in consideration of genetic predispositions to behavioral disorders.

Amphetamine treatment-by-hyperactivity line interaction effects on molecular mechanisms

The investigation of molecular mechanisms encompassing multiple genes exhibiting significant interaction effects augmented our understanding of the pathways and processes that are affected by an amphetamine treatment and hyperactivity line-dependent manner. Highly enriched and informative clusters and descriptive GO and KEGG terms among the 1,498 genes exhibiting a significant amphetamine treatment-by-hyperactivity line interaction effect are listed in **Table 2**. Enriched categories include the GO BP terms forebrain development, synaptic transmission, neurogenesis and neuronal development, ion transport, peptide hormone secretion, memory and learning, response to stimuli including hormones, and regulation of sensory perception.

Table 2. Informative clusters (Enrichment Score ≥ 4) of descriptive DAVID FAT categories including Gene Ontology (GO) biological processes (BP) and molecular functions (MF) enriched among genes exhibiting significant (FDR adjusted P-value < 0.05) amphetamine treatment-by-hyperactivity line interaction effects.

Category	Term	Count[a]	Raw P-value	FDR P-value[b]
Cluster 1	**Enrichment Score: 14.81**			
BP	GO:0007399~nervous system development	268	8.45E-28	1.67E-24
BP	GO:0030900~forebrain development	63	6.70E-11	1.33E-07
Cluster 2	**Enrichment Score: 13.98**			
BP	GO:0007267~cell-cell signaling	176	3.52E-24	6.96E-21
BP	GO:0099536~synaptic signaling	101	6.05E-22	1.20E-18
Cluster 3	**Enrichment Score: 12.36**			
BP	GO:0007399~nervous system development	268	8.45E-28	1.67E-24
BP	GO:0022008~neurogenesis	203	6.27E-23	1.24E-19
BP	GO:0030182~neuron differentiation	172	1.29E-19	2.56E-16
Cluster 4	**Enrichment Score: 12.32**			
BP	GO:0043269~regulation of ion transport	94	1.04E-17	2.06E-14
MF	GO:0022843~voltage-gated cation channel activity	38	1.68E-14	2.78E-11
Cluster 6	**Enrichment Score: 7.73**			
BP	GO:0007610~behavior	105	1.59E-17	3.15E-14
BP	GO:0050890~cognition	51	1.19E-10	2.36E-07
BP	GO:0007611~learning or memory	45	3.59E-09	7.09E-06
Cluster 10	**Enrichment Score: 6.25**			
BP	GO:0009719~response to endogenous stimulus	165	4.85E-11	9.59E-08
BP	GO:0009725~response to hormone	94	8.87E-09	1.75E-05

Table 2 (cont.)

Category	Term	Count[a]	Raw P-value	FDR P-value[b]
Cluster 19	**Enrichment Score: 4.98**			
BP	GO:0051049~regulation of transport	205	4.34E-18	8.58E-15
BP	GO:0030072~peptide hormone secretion	44	1.70E-08	3.37E-05

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

Many of the functional categories exhibiting significant treatment-by-line interaction effects in this study have been previously associated with amphetamine exposure, hyperactivity, and ADHD (Gruber and McDonald, 2012). Neuronal development was enriched among the genes differentially expressed in the striatum of spontaneously hypertensive rats, an ADHD model, treated with amphetamine (dela Peña et al., 2015). The enrichment of genes associated with synaptic transmission was expected since ADHD treatment using amphetamine acts through disruption of this signaling system (Mazei-Robison and Blakely, 2006). The enrichment of the memory and learning term is consistent with the impaired working memory observed in congenic wiggling rats that are hyperactive, exhibit impulsive behaviors and are a model to study ADHD (Masuo et al., 2007). These functional analysis results advance the understanding of the etiology of ADHD that is characterized by poor performance in executive functions including attention, correct response time and working memory (Loos et al., 2010).

Several enriched categories among the genes exhibiting significant treatment-by-line interaction effects are consistent with categories indirectly associated with ADHD and amphetamine effects. These categories are consistent with previous functional studies of behaviors and phenotypes typically associated with amphetamine treatment or ADHD. For example, high levels of physical activity have been associated with anatomical and physiological alterations in the brain, like size, volume, and spatial memory (Phillips et al., 2014).

Many enriched clusters and categories shared genes that exhibited significant treatment-by-line interaction effects. This finding supports both the understanding that multiple processes are associated with the effects of ADHD and amphetamine and that these processes are highly interconnected by genes. For example, Oxt participates in the enrichment of many categories ranging from cellular homeostasis to behavior, such as social behavior, and can have a therapeutic effect on addiction (Carson, 2014). Likewise, myocyte enhancer factor 2C (Mef2c) is annotated to

various functional categories previously discussed including the Oxt signaling pathway, cellular response to parathyroid hormone stimulus, and transcription factor activity. Mef2c presented a profile of positive CA-CS, positive HA-HS, and negative HS-CS. In agreement with this profile, conditional deletion of Mef2c was associated with the disrupted balance between excitatory and inhibitory synapses in the developing brain, neurodevelopmental disorders, behavioral changes including hyperactivity and repetitive movements (Harrington et al., 2016). Over-expression of Mef2c in mice prefrontal cortex neurons improves cognition, behavior, and working memory (Mitchell et al., 2017).

In addition to the enrichment of categories among genes exhibiting significant treatment-by-line interaction effects, the enrichment of categories within the most frequent profiles across the three orthogonal contrasts was studied. Highly enriched and informative clusters of descriptive categories identified by DAVID for the profiles with an Enrichment Score ≥ 4 are presented in **Table 3**. The most enriched categories among all genes exhibiting a significant interaction effect are also enriched in the two most frequent profiles: positive CA-CS, positive HA-HS, and negative HS-CS ("+ + -" profile) and negative CA-CS, negative HA-HS, and positive HS-CS ("- - +" profile). The remaining profiles included less than 170 genes, and no category reached an Enrichment Score > 2.

Table 3. Highly enriched and informative clusters of DAVID FAT categories including Gene Ontology (GO) biological process (BP) and molecular function (MF) among genes exhibiting significant interaction effect and the 3 most common profiles across the 3 orthogonal contrast (("+ + -", "- - +", and "+ - +")).

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
	Positive CA-CS, positive HA-HS, and negative HS-CS ("+ + -" contrast)			
Cluster 1	**Enrichment Score: 11.13**			
BP	GO:0099536~synaptic signaling	49	2.0E-16	4.8E-13
BP	GO:0006836~neurotransmitter transport	19	3.5E-08	4.9E-06
Cluster 2	**Enrichment Score: 6.83**			
MF	GO:0005249~ voltage-gated potassium channel activity	16	6.4E-10	5.3E-07

Table 3 (cont.)

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
	Positive CA-CS, positive HA-HS, and negative HS-CS ("+ + -" contrast)			
BP	GO:0006811~ion transport	58	4.4E-07	3.8E-05
Cluster 3	**Enrichment Score: 6.49**			
BP	GO:0030900~forebrain development	30	1.8E-08	3.3E-06
Cluster 4	**Enrichment Score: 5.9**			
BP	GO:0022008~Neurogenesis	79	5.4E-12	3.9E-09
BP	GO:0048666~neuron development	61	8.5E-12	4.6E-09
Cluster 5	**Enrichment Score: 5.56**			
BP	GO:0007610~Behavior	43	1.1E-09	3.3E-07
BP	GO:0007611~learning or memory	22	3.4E-07	3.1E-05
Cluster 16	**Enrichment Score: 2.64**			
BP	GO:0015833~peptide transport	19	8.3E-05	3.8E-03
BP	GO:0046879~hormone secretion	20	1.7E-04	7.2E-03
Category	**Negative CA-CS, negative HA-HS, and positive HS-CS ("- - +" contrast)**	**Count**	**Raw P-value**	**FDR-adjusted P-value**
Cluster 1	**Enrichment Score: 4.94**			
BP	GO:0099536~synaptic signaling	26	6.0E-06	1.6E-03
Cluster 2	**Enrichment Score: 4.84**			
MF	GO:0005184~neuropeptide hormone activity	9	1.3E-08	1.0E-05
MF	GO:0005179~hormone activity	12	5.3E-06	1.0E-03
Cluster 3	**Enrichment Score: 3.78**			
BP	GO:0050801~ion homeostasis	30	5.9E-06	1.7E-03
BP	GO:0008015~blood circulation	20	1.2E-04	1.0E-02
BP	GO:0003013~circulatory system process	20	1.3E-04	1.1E-02
Cluster 4	**Enrichment Score: 3.39**			
BP	GO:0030900~forebrain development	19	1.3E-04	1.0E-02
BP	GO:0040011~Locomotion	43	4.3E-04	2.0E-02
Cluster 10	**Enrichment Score: 2.57**			
BP	GO:0046879~hormone secretion	16	5.4E-04	2.3E-02
BP	GO:0002790~peptide secretion	14	7.5E-04	2.9E-02
Cluster 21	**Enrichment Score: 2.19**			
BP	GO:1901698~response to nitrogen compound	33	1.8E-04	1.2E-02
BP	GO:0009725~response to hormone	27	8.1E-04	2.9E-02

Table 3 (cont.)

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
Category	**Positive CA-CS, negative HA-HS, and positive HS-CS ("+ - +" contrast)**	**Count**	**Raw P-value**	**FDR-adjusted P-value**
Cluster 1	**Enrichment Score: 2.74**			
BP	GO:0003341~cilium movement	9	9.7E-08	3.2E-04
BP	GO:0007018~microtubule-based movement	12	1.1E-05	1.8E-02

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

The functional categories shared among the genes exhibiting the "+ + -" and "- - +" profiles included synaptic signaling, hormone secretion and peptide transport, and neurogenesis. The overlap in enrichment categories between the two opposite profiles and the interaction confirm the overlapping processes and overlapping genes corresponding to these processes disrupted by amphetamine and hyperactivity. The only highly enriched cluster of categories among the genes exhibiting the "+ + -" profile that was not detected at a significant level among the genes exhibiting the opposite profile "- - +" was GO BP memory/learning and cognition. Genes in this cluster include Nrgn, Vip, Bdnf, Neurod2, Adora1, amphiphysin (Amph), two members of the family of p21 proteins (Pak1 and Pak7), two members of the solute carrier family (Slc17a7 and Slc8a2), cannabinoid receptor 1 (Cnr1). Our finding is consistent with the effect of low amphetamine doses such as that used in the present study in enhancing focus and concentration and being used as a treatment to ameliorate behavioral disruptions in individuals with ADHD and in rodent models of ADHD.

In addition to the functional analysis of the most common profiles across the three orthogonal contrasts, functional analysis of each orthogonal contrast separately enabled to uncover specific disruption of processes and functions. Most of the functional categories detected by the DAVID analysis of individual orthogonal contrasts confirm the categories enriched in the most frequent profiles across the three contrasts. These categories included: behavior, learning or memory, regulation of neurotransmitter levels, response to chemical stimulus, and ion homeostasis. The substantial consistency of enriched molecular mechanisms among the three orthogonal contrasts that characterize the interaction indicates that amphetamine treatment and hyperactivity line do not impact the processes being disrupted but rather the interaction impacts

the individual gene expression profiles within each disrupted process. The two categories enriched in a single contrast (neuropeptide hormone system in CA-CS and ion homeostasis in HS-CS) appear to be an artifact of the number of differentially expressed genes analyzed within contrast because these categories were also detected in the functional analysis of the profile across the three contrast. Several enriched categories detected among the genes differentially express within individual orthogonal contrasts that were not directly detected in the functional analysis of profiles involving three orthogonal contrast were nevertheless indirectly correlated with the categories detected in the analysis of profiles. For example, the category GO MF GO:0008201 heparin binding is related to regulation of blood pressure and GO BP GO:0007631 feeding behavior is related to hormone and neuropeptide activity. Additional insights from the GSEA analysis of the standardized log2(fold change) of all genes within contrast include the enrichment of ribosomal pathway terms. The GO and KEGG ribosomal categories were enriched among the genes negative for CA-CS and positive for HA-HS and HS-CS. The ribosome term enrichment pattern detected is in agreement with the potentiation of amphetamine-induced locomotion by ribosome inactivating protein (Saporin) treatment (Jeltsch et al., 2004).

Two unexpected discoveries were drawn from the comparison of enriched categories between the two most common and opposite profiles ("+ + -" and "- - +") among the genes exhibiting amphetamine treatment–by-hyperactivity line interaction effects. Our first discovery is that genes are associated with the same or similar processes. **Table 3** highlights the overlap in enriched categories (e.g., GO BPs synaptic signaling, forebrain development, hormone secretion) between "+ + -" and "- - +". A similar overlap of enriched processes was observed between the profiles "-+-" and "+-+" at lower significance levels. These opposite or reverse profiles of genes annotated to the same mechanisms reflect possible feedback regulation (**Table 3**). Our second discovery is that processes associated with genes in a particular pair of opposite trajectories (e.g. "+ + -" or "- - +") were not identified among the genes in a distinct trajectory (e.g. "- + -" and "+ - +"). For example, **Table 3** includes the enriched GO BP cilium movement that was only enriched in the **"+ - +"** profile but was not enriched in profiles "- + -" and "+ - +". These observations confirm the strong and clear impact of hyperactivity line and amphetamine treatment on gene profiles and associated processes.

Our study of treatment-by-line interaction effects on pathways and processes helped us to identify mechanisms that are predominantly and concurrently affected by amphetamine and line

effects. These findings support the discovery of pathways that can be targeted to address ADHD behaviors in a psychostimulant treatment-dependent fashion or to understand psychostimulant effects in consideration of ADHD indicators.

Amphetamine treatment -by-hyperactivity line interaction effects on gene networks

The results previously discussed offered insights into amphetamine treatment-by-hyperactivity line interaction effects on individual genes and groups of genes belonging to the same functional terms. Network analysis complemented the previous analyses and helped us to elucidate the impact of the interaction on the relationship between genes. Gene networks were used to depict the relationship among the 75 genes exhibiting significant interaction effect (FDR-adjusted P-value < 0.05) and log2(fold change) >|2| and P-value < 0.001 in at least one of the three orthogonal contrasts. Individual gene co-regulation networks were developed for the three orthogonal contrasts: CA-CS (**Figure 4**), HA-HS (**Figure 5**) and HS-CS (**Figure 6**). To facilitate interpretation, the contrast HA-CA (**Figure 7**) is also presented. Among the 75 genes considered, the inferred networks focus on the largest group of directly interconnected genes or core genes. Genes directly connected to the core genes that did not reach significant differential expression in at least one contrast or were not reliably quantified are also included in the network to facilitate the identification of indirect gene co-regulation.

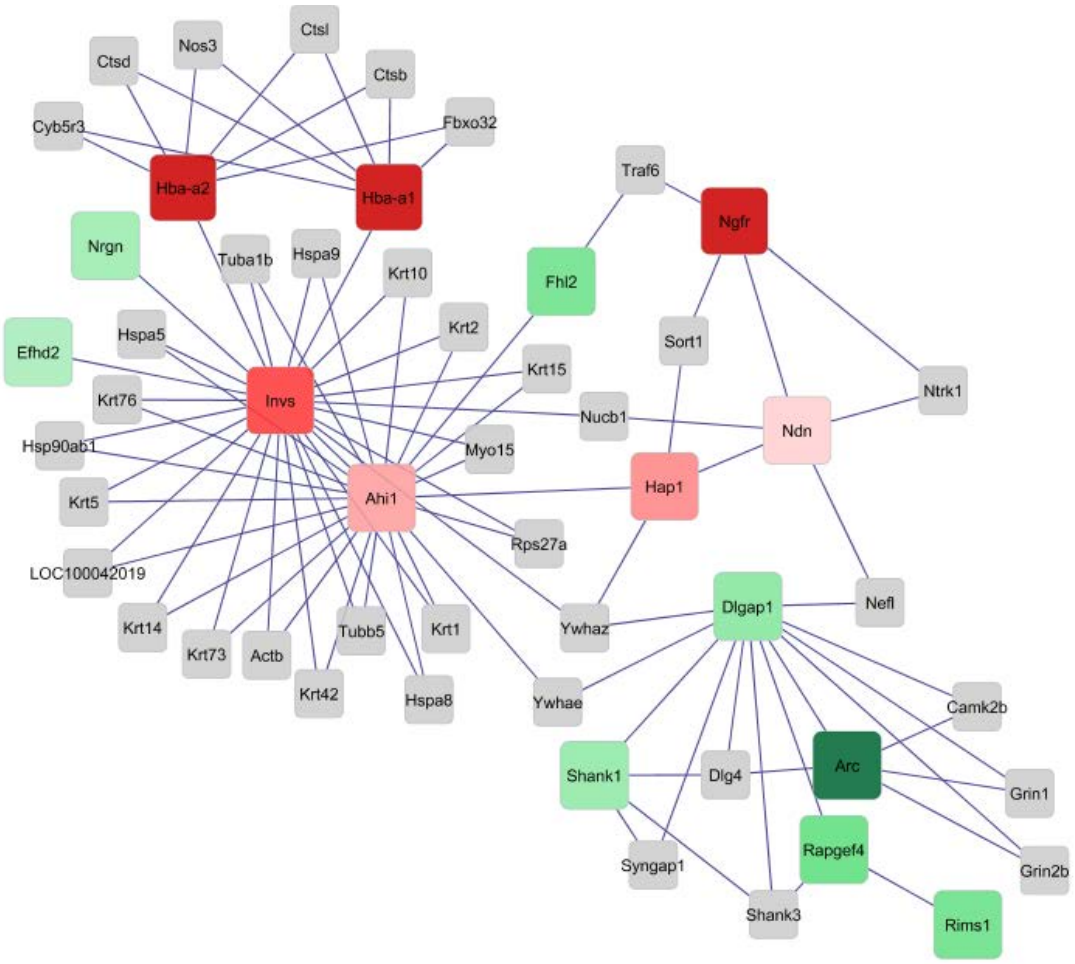

Figure 4. Over- (green) and under- (red) expressed genes in the contrast between control amphetamine and control saline (CA-CS) groups for the network of genes exhibiting significant - amphetamine treatment-by-hyperactivity line by interaction effect (FDR-adjusted P-value < 0.05) and at least one significant pairwise contrast between line-treatment groups (log2 (fold change) >|2| and P-value < 0.001). Color intensity is positively correlated with fold change. Gray connecting genes did not exhibit significant interaction effects. Genes: nerve growth factor receptor (Ngfr), neurogranin (Nrgn), inversin (Invs), Abelson helper integration site 1 (Ahi1), Fhl2, Huntingtin associated protein 1 (Hap1), necdin (Ndn), DLG associated protein 1 (Dlgap1), SH3 and multiple ankyrin repeat domains 1 (Shank1), activity regulated cytoskeleton-associated protein (Arc), Rapgef4, hemoglobin a (Hba-a1 and Hba-a2), and regulating synaptic membrane exocytosis 1 (Rims1).

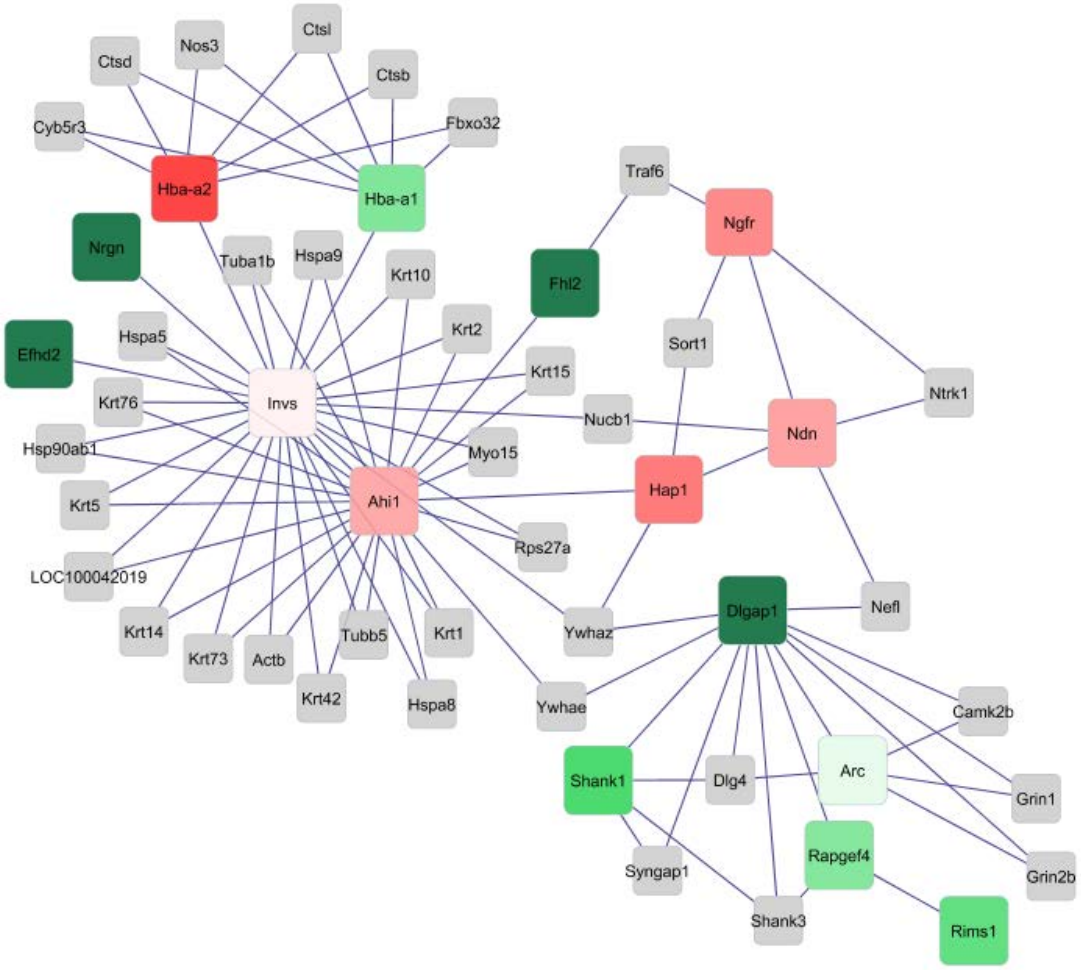

Figure 5. Under- (red) and over- (green) expressed genes in the contrast between hyperactive amphetamine and hyperactive saline (HA-HS) groups for the network of genes exhibiting significant amphetamine treatment-by-hyperactivity line interaction effect (FDR-adjusted P-value < 0.05) and at least on significant pairwise contrast between line-treatment groups (log2 (fold change) > |2| and P-value < 0.001). Color intensity is positively correlated with fold change. Gray connecting genes did not exhibit significant interaction effects. Genes: nerve growth factor receptor (Ngfr), neurogranin (Nrgn), inversin (Invs), Abelson helper integration site 1 (Ahi1), Fhl2, Huntingtin-associated protein 1 (Hap1), necdin (Ndn), DLG associated protein 1 (Dlgap1), SH3 and multiple ankyrin repeat domains 1 (Shank1), activity regulated cytoskeleton-associated protein (Arc), Rapgef4, hemoglobin a (Hba-a1 and Hba-a2), and regulating synaptic membrane exocytosis 1 (Rims1).

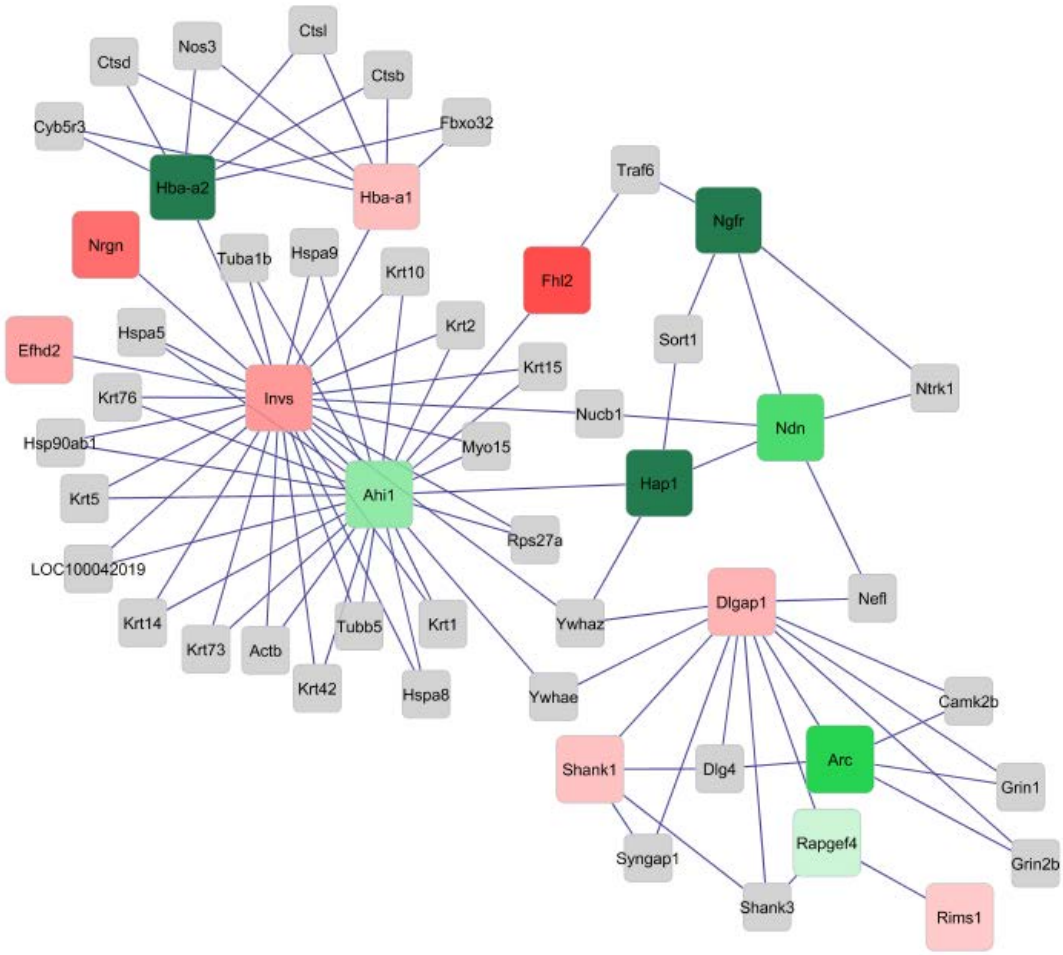

Figure 6. Over- (green) and under- (red) expressed genes in the contrast between hyperactive saline and control saline (HS-CS) groups for the network of genes exhibiting significant amphetamine treatment-by-hyperactivity line interaction effect (FDR-adjusted P-value < 0.05) and at least one significant pairwise contrast between line-treatment groups (log2 (fold change) >|2| and P-value < 0.001). Color intensity is positively correlated with fold change. Gray connecting genes did not exhibit significant interaction effects. Genes: nerve growth factor receptor (Ngfr), neurogranin (Nrgn), inversin (Invs), Abelson helper integration site 1 (Ahi1), Fhl2, Huntingtin-associated protein 1 (Hap1), necdin (Ndn), DLG associated protein 1 (Dlgap1), SH3 and multiple ankyrin repeat domains 1 (Shank1), activity regulated cytoskeleton-associated protein (Arc), Rapgef4, hemoglobin a (Hba-a1 and Hba-a2), and regulating synaptic membrane exocytosis 1 (Rims1).

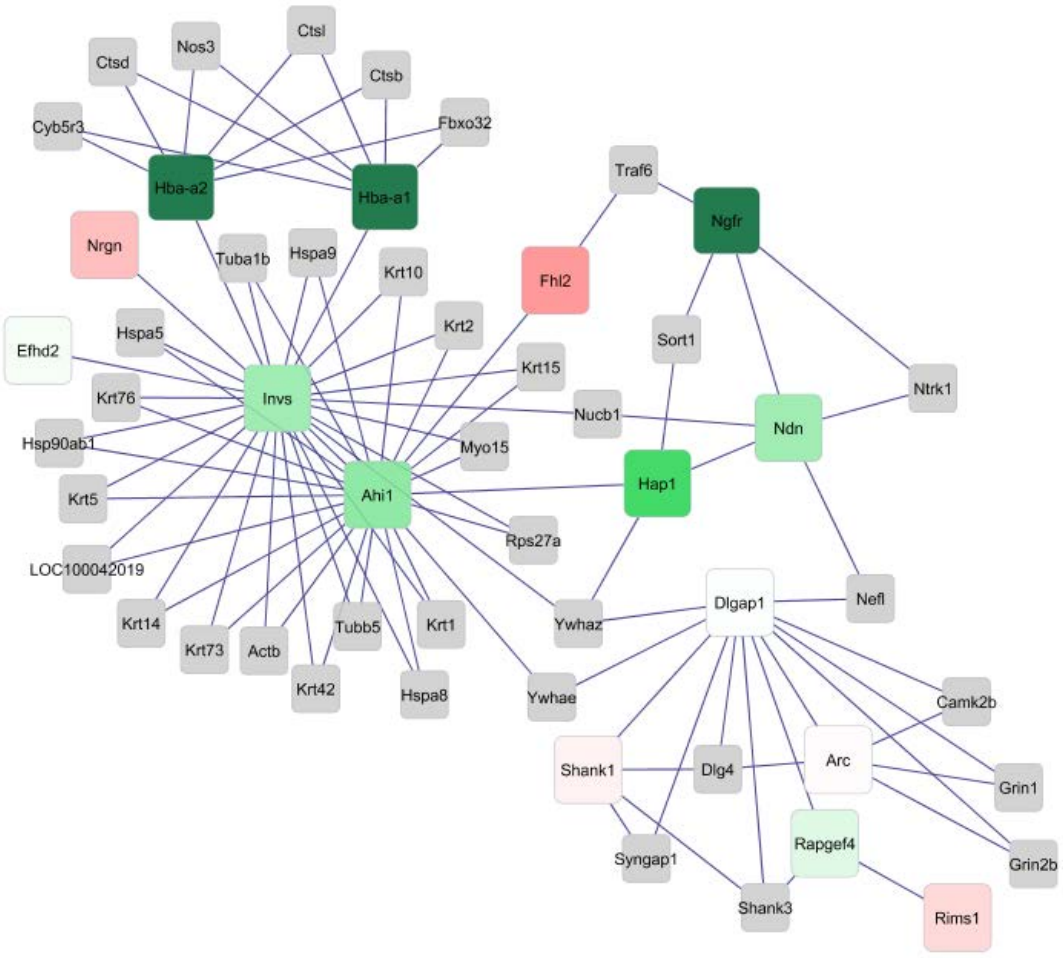

Figure 7. Over- (green) and under- (red) expressed genes in the contrast between hyperactive amphetamine and control amphetamine (HA-CA) groups for the network of genes exhibiting significant amphetamine treatment-by-hyperactivity line interaction effect (FDR-adjusted P-value < 0.05) and at least one significant pairwise contrast between line-treatment groups (log2 (fold change) >|2| and P-value < 0.001). Color intensity is positively correlated with fold change. Gray connecting genes did not exhibit significant interaction effects. Genes: nerve growth factor receptor (Ngfr), neurogranin (Nrgn), inversin (Invs), Abelson helper integration site 1 (Ahi1), Fhl2, Huntingtin-associated protein 1 (Hap1), necdin (Ndn), DLG associated protein 1 (Dlgap1), SH3 and multiple ankyrin repeat domains 1 (Shank1), activity regulated cytoskeleton-associated protein (Arc), Rapgef4, hemoglobin a (Hba-a1 and Hba-a2), and regulating synaptic membrane exocytosis 1 (Rims1).

The edges in the gene networks represent known connections between genes based on curated databases of molecular interactions used by Bisogenet and the rectangular nodes represent the genes that exhibited significant interaction effects. The intensity of colors indicates the strength of the positive (green) and negative (red) contrast of gene expression between mouse groups. Within contrast, positively connected genes are identified by edges connecting nodes of the same color (green-to-green and red-to-red) whereas negatively connected genes are identified by edges connecting nodes of different color (green-to-red).

The genes differentially expressed between line-treatment groups in the network are: nerve growth factor receptor (Ngfr), Nrgn, Inversin (Invs), Abelson helper integration site 1 (Ahi1), Fhl2, Hap1 (common gene between the four contrasts in the Venn diagram presented in **Figure 3**), necdin (Ndn), DLG associated protein 1 (Dlgap1), SH3 and multiple ankyrin repeat domains 1 (Shank1), activity regulated cytoskeleton-associated protein (Arc), Rapgef4, hemoglobin a (Hba-a1 and Hba-a2), and regulating synaptic membrane exocytosis 1 (Rims1). Ngfr is a neurotrophin receptor that plays a role in axonal growth. The genes depicted in the networks have been associated with angiogenesis, cell adhesion, apoptosis, and neuronal development processes. Consistent with our findings, these processes were also enriched in a study of the effects of amphetamine on the striatum of spontaneously hypertensive rat, an ADHD model (dela Peña et al., 2015).

Consideration of the GO BP categories associated with the genes in each network aids in understanding both their connection and the nature of this connection. The integration of gene interaction and differential expression depicted in the network enabled the visualization of co-regulation profiles. Nrgn and Rims1 positively correlated expression because both genes have the same pattern of differential expression (nodes with equal color) across all four networks. Moreover, the direction of the differential expression within contrast is consistent with the known role of these genes. Nrgn and Rims1 are over-expressed in the A relative to the S treatment groups across lines and under-expressed in the H relative to the C line across treatment. Both genes Nrgn and Rims1 play a role in calcium availability and transport. Nrgn is a neuropeptide positively associated with neural plasticity and negatively associated with cognitive deficit that binds to calmodulin decreasing the availability of calcium (Hoffma et al., 2014). Rims1 plays a role in synaptic potentiation and inhibition of postsynaptic potential through controls the neuronal voltage-gated calcium channels activity (Weiss et al., 2011). Calcium channel is essential to

neuronal signaling and was associated with substance abuse and emotional behaviors (Martínez-Rivera et al., 2017). The trend of these synaptic signaling genes observed in the gene networks could serve to explain the therapeutic effect of amphetamine in ADHD symptoms.

The inferred network depicts the relationship between 4 genes associated with neurogenesis: Ndn, Ahi1, Ngfr, and Hap1. In addition, Ndn has also been associated with sensory perception of pain and Ahi1 plays a role in the regulation of behavior. Ndn and Hap1 are directly associated in our network, and our findings support the positively correlated role of these genes because both genes have the same pattern of differential expression (nodes with equal color) across all four networks. The direction of the differential expression of Ndn, Ahi1, Ngfr, and Hap1 within contrast is consistent with the known role of these genes. Ndn, Ahi1, Ngfr, and Hap1 are over-expressed in the A relative to the S treatment groups across lines and under-expressed in the H relative to the C line across treatment. Our network comparison suggests that the use of amphetamine to treat ADHD may be linked to the opposite trend between H and A on these neurogenesis genes.

Further understanding of the capability of amphetamine to alleviate ADHD symptoms was gained from the juxtaposition of networks across contrasts across genes participating in multiple processes. Disruptions in Arc and Shank1 have been linked to psychostimulants and are involved with drug response in the mouse striatum (Buonaguro et al., 2017). Arc has been annotated to the amphetamine addiction KEGG pathway and is also associated with regulation of synaptic potentiation (Biever et al., 2016; Tan et al., 2000). Shank1 has been correlated with social and vocalization behavior, memory and synapse maturation. Our networks offer insights into the treatment- and line- dependent relationship between these genes across lines. Within line, Arc and Shank1 have a consistent green pattern indicating over-expression A relative to S groups, consistent with known gene function. However, Arc and Shank1 have opposite patterns in the contrasts of lines within S treatment and are not highly differentially expressed in the contrasts of lines within the A treatment. These results suggest that differences in expression between lines in these genes observed in the S treatment are diminished by the A treatment. Our network study uncovered one mode of action of amphetamine that ameliorates the antagonistic disruption of Arc and Shank in the hyperactive line.

The comparison of the networks across contrasts complements the understanding of the differential response of the striatum transcriptome to amphetamine treatment between lines. All

except one gene (Hba-a1) in the CA-CS and HA-HS networks exhibit the same differential expression indicating that amphetamine had similar effects on both hyperactivity lines (**Figs 2** and **3**). On the other hand, more genes exhibit opposite differential expression between HS-CS and HA-CA (**Figs 4** and **5**, respectively). Our network comparison depicts the interaction between amphetamine treatment and hyperactivity line in the disruption of gene profiles that could disrupt gene connections.

The majority of the direct connections between genes affected by the treatment-by-line interaction across all four networks are among genes with the same pattern (e.g., edges connecting red nodes or edges connecting green nodes). The notable exception is Invs and the connection of this gene with Nrgn, Efhd2, and Ahi1. In all contrasts except HS-CS, Invs and Ahi1 share the same expression pattern and Invs and Nrgn have opposite expression patterns (**Figure 6**). This is the only contrast that does not involve the A treatment and network comparison enabled us to uncover a critical and unique effect of amphetamine.

The comparative network analysis identified sub-networks of genes that were disrupted by the amphetamine treatment. The sub-network including Shank1, Arc, Rims1, Dlgap1, Rapgef4 and Dlgap1, Rapgef4 (the latter two genes are associated with the GO BP regulation of synaptic signaling) does not exhibit substantial differential expression in HA-CA (**Figure 7**), in comparison to all other contrasts. This finding suggests that treatment A decreases the differential expression of genes between the H and C lines by targeting whole subnetworks of connected genes, rather than isolated genes. The HA-CA network (**Figure 7**) had the highest number of unchanged genes expression across all four networks. This depiction suggests that for the genes in the network, the low amphetamine dose used in this study was able to revert profiles dysregulated in the H line without major dysregulation in the C line.

Network analysis highlighted the synergistic effect of the A treatment and H line in the co-regulation of 3 genes associated with vascular and blood systems: Invs, Hba-a1, and Hba-a2. Disruption of genes associated with angiogenesis processes was also reported in a study of amphetamine and an ADHD model using the spontaneously hypertensive rat (dela Peña et al., 2015). Hemoglobin is expressed in neurons and is essential to neuronal respiration and vascular system (Stankiewicz et al., 2015). All three genes, Invs, Hba-a1, and Hba-a2 were over-expressed in CA relative to CS (**Figure 4**) and under-expressed in CA relative to HA (**Figure 7**). Thus, the level of expression of these genes was HA > CA > CS, and this order suggests synergism between

the H line and A treatment. These genes support the circulation of blood carrying oxygen which in turn support the higher activity of mice in the H line and higher alertness, focus, intensity and duration of training with lower fatigue of mice receiving the A treatment.

Five genes tend to be hub genes: Dlgap1, Hba-a1, Hba-a2, Invs, and Ahi1. In **Figures 4** and **5** that compares treatments between lines, the over-expressed (green) genes tended to be in the periphery (outside) the networks whereas under-expressed (red) genes tend to be hub genes in the middle of networks. **Figure 4** has one green hub (Dlgap1) and four red hubs (Hba-a1, Hba-a2, Invs, and Ahi1), while **Figure 5** has two green hubs (Dlgap1 and Hba-a1) and three red hubs (Hba-a2, Invs, and Ahi1). The same is not applied to **Figures 6** and **7**, where the color of these hubs change completely. **Figure 6** has two green hubs (Hba-a2 and Ahi1) and three red hubs (Dlgap1, Hba-a1, and Invs), while **Figure 7** has only green hubs (Hba-a1, Hba-a2, Invs, and Ahi1) and Dlgap1 tend to be neutralized. The hubs Invs and Ahi1 share similar connected genes, and most of them are from two families: keratins (Krt1, Krt2, Krt5, Krt10, Krt14, Krt15, Krt42, Krt73, and Krt76) and heat shock proteins (Hspa5, Hspa8, Hspa9, and Hspab1). Keratins are intermediate filament proteins that can be found in neurons and glial cells and may play a role in the neural differentiation (Podgorniak et al., 2015). Intermediate filament proteins and heat shock proteins were associated with mouse pituitary gland under heat stress situations like elevated heart rate and sweating (Memon et al., 2016). The hubs Hba-a1 and Hba-a2 share genes from peptidase family (Ctsd, Ctsl, and Ctsb). The association of hemoglobin with peptidase might be related to mechanisms and controls of cerebral protein catabolism (Banay-Schwartz et al., 1985). The changes that occur in the hub genes might be linked to the activity of gray genes that did not exhibit significant interaction effects that are in some way also modified by the A treatment.

Our study of treatment-by-line interaction effects on gene co-regulation enabled the detection of co-expression profiles that are simultaneously amphetamine- and line-dependent. This information will aid in the identification of co-regulated targets that can be used to develop personalized ADHD or psychostimulant exposure therapies. The ensuing step is to study transcriptome profiles that associated with changes in either amphetamine treatment or hyperactivity line.

Effect of hyperactivity line on the striatum transcriptome

Our study of genes exhibiting significant H -C line contrast uncovered transcriptome disruptions in the ADHD model that is independent of the amphetamine treatment. Among the 860 significantly differentially expressed genes in the striatum, 239 genes did not exhibit significant (FDR-adjusted P-value < 0.05) interaction effects. **Table 4** lists the most differentially expressed genes (FDR-adjusted P-value < 0.05 and log2 (fold change) > |1.2|) that exhibit main line effect without significant treatment-by-line interaction.

Table 4. Top genes differentially expressed (FDR-adjusted P-value < 0.05) and log2(fold change) > |1.2|) between mice in control (C) and hyperactivity line (H) genotypes that do not exhibit significant interaction effect.

Gene Symbol	Description	Log2(C/H)	Raw P-value	FDR adjusted P-value[a]
Ubash3a	ubiquitin associated and SH3 domain containing, A	2.38	5.0E-05	2.0E-03
Olfr658	olfactory receptor 658	-2.70	5.0E-05	2.0E-03
Syce2	synaptonemal complex central element protein 2	1.59	2.0E-04	6.5E-03
Ninj2	ninjurin 2	-1.32	2.0E-04	6.5E-03
A530076I17Rik	RIKEN cDNA A530076I17 gene	-1.84	2.5E-04	7.6E-03
Gm2808	predicted gene, 2808	5.32	3.0E-04	8.8E-03
Nup62cl	nucleoporin 62 C-terminal like	-1.41	3.0E-04	8.8E-03
Ppp1r2-ps2	protein phosphatase 1, regulatory (inhibitor) subunit 2, pseudogene 2	5.03	4.0E-04	1.1E-02
Nlrp6	NLR family, pyrin domain containing 6	1.30	5.5E-04	1.4E-02
Gbp1	guanylate binding protein 1	-2.09	7.0E-04	1.7E-02
5031414D18Rik	RIKEN cDNA 5031414D18 gene	-3.81	7.0E-04	1.7E-02
P2rx1	purinergic receptor P2X, ligand-gated ion channel, 1	1.31	9.0E-04	2.1E-02
4930519G04Rik	RIKEN cDNA 4930519G04 gene	2.58	1.5E-03	3.0E-02
Il1a	interleukin 1 alpha	1.40	1.6E-03	3.1E-02
Gm17660	predicted gene, 17660	2.17	2.6E-03	4.6E-02
1700030C10Rik	RIKEN cDNA 1700030C10 gene	-1.87	2.7E-03	4.8E-02
E230025N22Rik	RIKEN cDNA E230025N22 gene	1.45	2.8E-03	4.9E-02

[a]False Discovery Rate adjusted P-value.

Among the top differentially expressed genes between the H and C groups listed in **Table 4**, several genes are associated with immune response. Guanylate binding protein 1 (Gbp1, also known as guanylate binding protein 2b or Gbp2) is associated with cellular response to interferon beta and gamma and immune response. Gbp1 is over-expressed in the H-C line contrast, and this profile is consistent with reports that the cytokines interferon gamma and alpha (highly homologous to beta) decrease the activity in mouse (Crnic and Segall, 1992). Another gene associated with immune response, Il1a was under-expressed in H-C. The reverse association is in agreement with the identification of a mutation in this gene associated with children exhibiting Tourette syndrome that is characterized by motor tics and high risk for ADHD (He et al., 2015). NLR family, pyrin domain containing 6 (Nlrp6) was under-expressed in H-C and consistent with our finding this gene negatively regulates innate immunity and facilitate the recovery after peripheral nerve injury (Anand et al., 2012).

Some genes differentially expressed in H-C has been associated in some cases with locomotor, behavioral, and ADHD-related disorders through methylation epigenetic mechanisms. Ninjurin 2 (Ninj2) exhibits positive H-C line contrast, and this gene is involved in cell adhesion, a process that is enriched among the genes over-expressed in HS-CS. Our findings are consistent with the identification of differential methylation in Ninj2 in children exhibiting ADHD symptoms (Wilmot et al., 2016). Likewise, 5031414D18Rik codes for a RUN and cysteine-rich domain containing beclin 1 interacting protein like (Rubcnl-like) and is over-expressed in H-C. Rubcnl is a known methylated promoter (Leal and Gulley 2017). Nucleoporin 62 C-terminal like (Nup62cl) encompasses a CG island that exhibits rapid or moderation methylation dynamics (Gendrel et al., 2012). These findings support the hypothesis of epigenetic mechanisms playing important role in hyperactivity and potentially ADHD symptoms.

Protein phosphatase 1, regulatory (inhibitor) subunit 2 (Ppp1r2) is enriched in GABAergic neurons (Belforte and Nakazawa, 2011) that are inhibited in ADHD models (Brennan and Arnsten 2008). This association is consistent with the under-expression of Ppp1r2 in H-C found in our study. Also, Ppp1r2 treatment has been associated with memory suppressor in rats (Yang and Wilson, 2015) and this is in agreement with the enrichment of memory and learning processes among genes differentially expressed in HS-CS. Another gene associated with memory and learning under-expressed in H-C was purinergic receptor P2X, ligand-gated ion channel, 1 (P2rx1).

Our result is in agreement with reports of over-expression of P2rx1 in mice with improved memory and learning (Meng et al., 2009).

Unexpectedly, the functional analysis of genes exhibiting hyperactivity line effect alone helped us identified processes that were also identified in the treatment-by-line interaction study. DAVID analysis for functional enrichment uncovered 3 clusters reached an Enrichment Score > 1.5 (**Table 5**). The lower level of significant relative to the interaction enrichment may be associated with the limited number of genes exhibiting hyperactivity only effect in the absence of significant interaction. The enriched categories were consistent with those uncovered among genes exhibiting interaction effect and included neurogenesis and neuron development and actin cytoskeleton organization. These categories have also been associated with ADHD symptoms and related phenotypes. Actin cytoskeleton was enriched in the GSEA analysis of genes under-expressed in the HS-CS contrast. The link between dysregulation of actin cytoskeleton processes and ADHD has been recognized (Lesch et al., 2008; Shiow et al., 2009). Moreover, actin cytoskeleton processes are associated with neuronal development and this together with neurogenesis and axogenesis were enriched among the genes differentially expressed in the HS-CS contrast (**Table 2**). These novel results demonstrate that the dysregulation of some genes in the neurogenesis and neuronal development processes in the hyperactivity line are not alleviated by amphetamine since the dysregulation is consistent across treatments whereas the dysregulation of other genes in the same processes is affected by the amphetamine treatment received. Our study of the main effect of hyperactivity line on individual gene and biological processes support the identification of molecular targets that can be used to enhance treatments of ADHD and related behavioral disorders.

Table 5. Informative clusters (Enrichment Score ≥ 1.5) of descriptive DAVID FAT categories including Gene Ontology (GO) biological processes (BP) enriched among genes exhibiting significant (FDR adjusted P-value < 0.05) hyperactivity line effect excluding genes exhibiting interaction effect.

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
Cluster 1	**Enrichment Score: 2.01**			
BP	GO:0044089~positive regulation of cellular component biogenesis	12	6.4E-3	5.4E-1
Cluster 2	**Enrichment Score: 1.59**			
BP	GO:0007010~cytoskeleton organization	19	4.2E-2	8.0E-1
BP	GO:0030036~actin cytoskeleton organization	11	6.4E-2	8.4E-1
Cluster 3	**Enrichment Score: 1.5**			
BP	GO:0022008~neurogenesis	32	6.9E-4	5.1E-1
BP	GO:0048666~neuron development	24	9.0E-4	5.1E-1
BP	GO:0010977~negative regulation of neuron projection development	7	5.0E-3	5.0E-1

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

Effect of amphetamine treatment on the striatum transcriptome

The identification of genes that exhibited significant A-S treatment contrast permitted the identification of disruptions in the striatum associated with amphetamine exposure that is independent of the genetic background for hyperactivity in mouse as ADHD model. Among the 183 genes exhibiting significant A-S treatment, two genes did not exhibit significant treatment-by-line interaction (FDR-adjusted P-value < 0.05): potassium inwardly-rectifying channel, subfamily J, member 6 (Kcnj6) and RAR-related orphan receptor alpha (Rora).

Kcnj6 exhibited positive A-S treatment contrast (log2 (fold change A/S) =0.37, FDR-adjusted P-value < 0.038). Kcnj6 plays a role in the Oxy signaling pathway, morphine addiction pathway, potassium ion transport, neurotransmitter, and in the dopaminergic and GABAergic synapse. Mutations on one of the G-protein-gated inwardly-rectifying potassium (Girk) channel subunits, Girk2, are related to analgesic requirements the therapeutic effects of Girk channel inhibitors in the treatment of methamphetamine dependence and alcoholism have been proposed (Sugaya et al., 2013).

Rora is involved in amphetamine metabolism (Yarosh et al., 2015), plays a role in angiogenesis, learning, and positive regulation of circadian rhythm (Sun et al., 2015). Mutations in Rora are associated with impaired in motor functions and motor coordination (Lalonde and Strazielle, 2015). Rora was positive for the A-S contrast (log2 (fold change A/S) = 0.32, FDR-adjusted P-value < 0.034) observed in this study confirms the effects that A elicits on Rora expression.

Our study uncovered a large number of gene expression profiles exhibiting amphetamine effects on a hyperactive line-dependent manner (interaction effects) whereas substantially fewer genes exhibited amphetamine exposure effect on a line independent manner. These novel findings highlight the need to develop therapies for psychostimulant use that are dependent on the genetic background associated with ADHD and related disorders.

CHAPTER 3 - DIFFERENTIAL TRANSCRIPTOME ANALYSIS AND NETWORKS BETWEEN IDO1-KNOCKOUT AND WILD-TYPE MICE IN MACROPHAGES UNDER IMMUNE CHALLENGE

Abstract

Inflammation has been implicated in behavioral disorders such as depression-like behavior. The tryptophan-degrading enzyme Indoleamine 2,3-dioxygenase (IDO1) is part of the tryptophan metabolic route, that is linked to infectious diseases and psychiatric disorders. The effect of IDO1 depletion on the peripheral macrophages transcriptome after Bacille Calmette Guérin (BCG)-infection has not been studied. Transcriptome and transcription factors (TF) network analyses were performed in the macrophages from IDO1-knockout (IDO1-KO) mice relative to wild-type (WT) mice under both BCG-infected and saline (Sal) treatment. Among the 546 differentially expressed genes exhibiting interaction effect (P-value < 0.05), sphingosine kinase 1 (Sphk1) plays a role in immune processes by activating of tumor necrosis factors (TNF)-alpha signaling and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB) pathways, and serum amyloid a1 (Saa1) that is related to depressive-like behavior. Both genes were under-expressed in BCG-infected compared to Sal treated IDO1-KO mice. The comparison between IDO1-KO relative to WT genotype irrespective of challenge suggested that the macrophages may find a way to compensate the absence of IDO1 restoring the immune response. Also, 8,798 genes were significantly differentially expressed (FDR-adjusted P-value < 0.05) between BCG-infected relative to Sal treated mice. Among them, ATP binding cassette subfamily b member (Tap1) was under-expressed in the BCG-infected relative to Sal treated mice, while histamine receptor h1 (Hrh1) was over-expressed. The functional analysis confirmed the participation of genes in significant GO terms (FDR-adjusted P-value < 0.05) associated with microbial response and host defense (i.e., GO BP response to virus; interaction with host), and macrophage metabolism (i.e., GO BP regulation of cysteine-type endopeptidase activity). Regulatory network analysis uncovered the enrichment of the transcription factors: signal transducer and activator of transcription 1 (Stat1), Rela proto-oncogene, NF-kB subunit (Rela), forkhead box g1 (Foxg1) and forkhead box l1 (Foxl1). Our findings can be used as potential candidates to understand the mechanisms involved in the BCG-infection leading to inflammation and depression-like behavior.

Introduction

Many diseases including heart diseases, obesity, and AIDS, are characterized by chronic and systemic inflammations. The tryptophan (Trp) metabolic route has been linked with infectious, autoimmune, chronic inflammatory diseases, and psychiatric disorders (Baumgartner et al., 2017; Dantzer et al., 2008). The enzymes present in this pathway are regulated by cytokines, and their interaction may have a potential role in cardiovascular diseases (Baumgartner et al., 2017). The tryptophan-degrading enzyme Indoleamine 2,3-dioxygenase (IDO1) is activated during the inflammation process and impairs T cell proliferation (O'Connor et al., 2009). Some studies showed that IDO1 could suppress immune responses to infections and also inhibit local inflammation and autoimmunity (Put et al., 2016).

Macrophages are immune cells that are found in all tissues in a variety of processes like development, homeostasis, and mainly immune responses to pathogens (Wynn et al., 2013). During inflammation, macrophages are involved in the phagocytosis of dead cells and the protection of the host through innate immunity (Martinez and Gordon, 2014). The acquisition of enhanced antimicrobial resistance is due to the interaction of macrophages with T and B lymphocytes (Martinez and Gordon, 2014). Disruptions in the blood-brain barrier may enable the migration of immune cells from the peripheral blood and elicit inflammation in the central nervous system. Macrophages exist in between two extremes with distinct functions: M1 macrophages defend the body in the presence of pathogens and tumor cells whereas M2 macrophages are related to inflammatory responses and adaptive immunity (Wang et al., 2014).

The higher activity of IDO1 in response to the immune challenge has been associated with behavioral disorders. IDO1 plays a role in macrophage differentiation and because of the involvement in mechanisms of tumor immune tolerance, IDO1 is considered a target for therapeutic studies (Wang et al., 2014). Transcriptome studies showed that macrophages regulate T cell responses via up-regulation of IDO1 and this process is associated with behavioral disorders (Gonzalez-Pena et al., 2016b).

Bacilli Calmette Guerin (BCG) is an intracellular pathogen and an attenuated strain of *Mycobacterium bovis* that is used in tuberculosis vaccines (O'Connor et al., 2009). Challenge with BCG triggers in mice immune response, inflammation in the brain, activation of IDO1 in the brain, and is associated with depression-like behavioral disorders (Dantzer et al., 2008; Rodriguez-Zas et al., 2015). Despite recovery from sickness phenotypes within one week, depression-like

symptoms may remain up to one month after challenge (Gonzalez-Pena et al., 2016a). No study has simultaneously profiled the expression of genes in macrophages from wild-type and IDO1-KO mice after sickness recovery from BCG challenge.

The first objective of this study was to identify gene differential expression of peripheral macrophage after sickness recovery from BCG challenge relative to saline control in IDO1-KO relative to WT mice. Second, to identify regulatory network and enrichments that uncover transcription factors. Finally, to find potential candidates to understand the relationship between inflammation caused by BCG-infection and depression-like behavior.

Material and methods

Animal experiments

Male mice, with approximately 22 weeks old, from the C57BL/6J genotype (wild-type or WT) and from an IDO1-knockout genotype that had a C57Bl/6J background (IDO1-KO) were studied. Mice were housed in individual cages in a controlled environment, under a normal 12:12 h light/dark cycle, 23°C of temperature, 45% of humidity. Water and food (Teklad 8640 chow, Harlan Laboratories, Indianapolis, IN, USA) were served *ad libitum*. Briefly, during three weeks prior to the challenge, mice were acclimated to the light cycle and facility before the injection with BCG or Sal. Mice were individually handled every day for few minutes before the challenge.

Half of the mice from each genotype were challenged with BCG, and half received Saline (Sal) treatment following protocols (Dantzer et al., 2008; Gonzalez-Pena et al., 2016a; Gonzalez-Pena et al., 2016b; Nixon et al., 2015). Before inoculation, each vial's reconstitution was made according to provider protocols using preservative-free saline. Live attenuated mycobacteria TICE BCG (50 mg wet weight of lyophilized culture containing 1x108 colony forming units or CFU/vial; Organon Teknika Corp. LLC, USA Inc., USA) was used (Gonzalez-Pena et al., 2016a; Gonzalez-Pena et al., 2016b). Mice were injected once with either 10 mg/mouse (BCG-challenged group, n = 12) or sterile saline solution (Control group, n = 12) administered via intraperitoneal injection (Rodriguez-Zas et al., 2015). No peritoneal mycobacteria infection was detected in mice by day 20 after intraperitoneal infection with BCG. However, there was a dissemination of infection to other organs, including the lung, spleen, and bone marrow (Ufimtseva, 2015).

Mice were euthanized seven days after BCG challenge. The endpoint was chosen based on prior studies that showed the recovery from sickness yet persistence of depressive-like symptoms seven days after challenge (Rodriguez-Zas et al., 2015). Mice were euthanized by trained personnel with CO_2 asphyxiation; everything was made to minimize suffering (Gonzalez-Pena et al., 2016b). Animal experimental procedures were performed under the approval of the Illinois Institutional Animal Care and Use Committee and were in accordance with the NIH Guide for the Care and Use of Laboratory Animals.

Macrophages were extracted from the peritoneum (abdominal cavity) using the previous protocols (Gonçalves and Mosser, 2015; Nixon et al., 2015). Mice had their abdomens disinfected, the skin retracted, and the peritoneal cavity was flushed with Hank's Balanced Salt Solution (cold harvest medium). The peritoneal fluid was centrifuged, and the resulting cell pellet was resuspended and plated. After 2 h incubation, the medium was aspired to remove non-adherent cells. The surviving adherent cells constituted the non-thioglycollate elicited peritoneal macrophages and cells from individual mice were frozen in Trizol at -80°C for RNA extraction (Gelsthorpe et al., 1997).

Flow cytometry validation of macrophages isolation encompassed cell staining with primary fluorescent antibodies for two primary markers for macrophages: CD11b (Integrin Subunit Alpha M) and CD45 (protein tyrosine phosphatase, receptor type, C antibody) (Gonzalez-Pena et al., 2016a; Gonzalez-Pena et al., 2016b). Fc receptors were blocked by incubation with anti-CD16/CD32 antibody before incubation with eBioscience anti-CD11b and anti-CD45 antibodies (eBioscience Inc., San Diego, CA). Surface receptor expression was identified using a Biosciences LSR II Flow Cytometry Analyzer with BD FACSDiva software (BD Biosciences, San Jose, CA). Antibody gating was determined using isotype-stained controls.

RNA extraction from macrophages followed the Trypsin procedure using a total RNA Kit (Omega Bio-Tek, Norcross, GA) and a DNase step to remove DNA contamination (Gelsthorpe et al., 1997). The RNA Integrity Numbers were accessed by using the Agilent 2100 Bioanalyzer with RNA Pico chip (Agilent Technologies, Palo Alto, CA).

Differential expression analysis

RNA libraries from individual mouse were sequenced using an Illumina HiSeq 2500 (Illumina, San Diego, CA) and 100nt long paired-end reads were obtained (Gonzalez-Pena et al.,

2016a). FastQC (Babraham Institute, 2013) was used for reading quality control (Andrews, 2010). Quality control analysis indicated Phred nucleotide quality score above 30 and sequences were not trimmed or filtered (Gonzalez-Pena et al., 2016a).

Reads were mapped to the Genome Reference Consortium GRCm38 mouse assembly using software Kallisto (Bray et al., 2016) with the Illumina iGenomes package (mm10; http://support.illumina.com/sequencing/sequencing_software/igenome.html) (Pruitt et al., 2006; Trapnell et al., 2009). After the quantification of transcript isoforms' abundance into counts, an R package tximport was used to combine counts with mapping information to quantify gene expression (Soneson et al., 2015). The edgeR package was used to describe the gene counts using a linear model including the main factors of treatment (BCG vs. Sal) and mice genotype (IDO1-KO vs. WT) and the interaction between treatment and genotype (Robinson et al., 2010). The Benjamini-Hochberg false discovery rate (FDR) was used to adjust the differential expression P-value for multiple testing (Reiner et al., 2003).

Functional and Transcription Factor analyses

Functional analysis was explored to find enriched categories among genes exhibiting interaction effect and main effects of treatment and genotype. The functional analysis used the Database for Annotation, Visualization and Integrated Discovery (DAVID; Version 6.8). Functional categories used in the analysis were Gene Ontology (GO) biological processes (BP), molecular functions (MF), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Functional Annotation Tool (FAT) classes were used in DAVID to facilitate interpretation (Caetano-Anollés et al., 2016). Expression Analysis Systematic Explorer (EASE) scores were used to measure the category enrichment, computed using a one-tailed jackknifed Fisher hypergeometric exact test. The statistical significance for each cluster of functional categories was evaluated using an Enrichment Score (ES) computed as the geometric mean (-log10 geometric mean of the cluster members EASE scores) (Gonzalez-Pena et al., 2016a). The *Mus musculus* genome was used as background for testing, and FDR was used to adjust the enrichment P-values for multiple testing (Pruitt et al., 2006; Trapnell et al., 2009).

Transcription factors (TF) enriched among genes exhibiting significant differential expression (FDR adjusted P-value < 0.05) were identified using iRegulon (Shannon et al., 2003; Verfaillie et al., 2014). This routine is a plugin available in the Cytoscape environment that detects

gene expression regulators (regulons) and gene targets using a motif prediction parameters for enrichment score threshold (Shannon et al., 2003; Verfaillie et al., 2014). The information from many databases including GeneSigDB, Ganesh clusters, and MSigDB is integrated to identify transcription factors and target genes (Culhane et al., 2011; Liberzon et al., 2015). The Area Under the cumulative Recovery Curve obtained from the database search is used to compute the normalized enrichment score (NES) for each transcription factor (Shannon et al., 2003; Verfaillie et al., 2014).

Results and Discussion

The level of the RNA integrity indicator (RIN) was highly comparable across the four groups of samples. The average RIN for the samples corresponding to the groups WT-Sal, WT-BCG, KO-Sal, and KO-BCG were: 9.7, 9.5, 9.4, and 9.6, respectively. Similarly, the average number of reads was similar across the four groups of samples. The average number of individual reads per sample corresponding to the groups Sal-WT, BCG-WT, Sal-KO, and BCG-KO were: 55,243410.3, 59,182973.4, 62,929589.2, and 62,315095.8, respectively. The number of genes per sample detected from the mapping of reads to the genome was similar across groups. The average number of detected genes per sample corresponding to the groups Sal-WT, BCG-WT, Sal-KO, and BCG-KO were: 22508, 22463, 22762, and 22701. The relative difference between the groups with highest and lowest number of genes was 1%.

Treatment-by-genotype interaction effects on the macrophage transcriptome

Immune treatment-by-IDO1 genotype interaction effects were identified on 65 genes (P-value < 0.005). **Table 6** highlights the genes with most significant interaction effect (P-value < 5x10-4). Among these are genes related to inflammatory response and immune response including: interleukin 10 (Il10), sphingosine kinase 1 (Sphk1), serum amyloid A1 (Saa1), C-C motif chemokine ligand 7 (Ccl7), C-X-C motif chemokine ligand 1 (Cxcl1), trophoblast glycoprotein (Tpbg), heparin binding epidermal growth factor (Egf) like growth factor (Hbegf), and cystathionine gamma-lyase (Cth).

Table 6. Most significant (P-value < 5x10-4) genes exhibiting treatment-by-genotype interaction effect between combinations of treatment (BCG=infected, Sal=saline) and genotype (KO=IDO1-KO, WT=wild type) groups.

	Overall		Log2(Fold Change)[a]					
Gene Symbol	**P-value**	**FDR P-value**	**BCG(WT)-Sal(WT)**	**BCG(KO)-Sal(KO)**	**Sal(KO)-Sal(WT)**	**BCG(KO)-BCG(WT)**	**BCG(KO)-Sal(WT)**	**Sal(KO)-BCG(WT)**
Il10	2.61E-05	3.70E-01	-1.55	-5.2	1.44	-2.23	-3.78	2.99
Creb3l1	1.27E-04	6.10E-01	-1.08	-2.41	0.63	-0.7	-1.8	1.72
Hbegf	1.41E-04	6.10E-01	-1.2	-3.54	1.46	-0.89	-2.12	2.66
Cth	2.37E-04	6.10E-01	-1.58	-3.46	0.38	-1.51	-3.1	1.96
Cxcl1	2.54E-04	6.10E-01	-1.39	-3.82	0.98	-1.47	-2.87	2.38
Sphk1	2.93E-04	6.10E-01	-1.27	-3.32	1.32	-0.73	-2.02	2.6
Ccl7	3.24E-04	6.10E-01	-0.4	-2.61	1.69	-0.53	-0.95	2.09
Tpbg	3.68E-04	6.10E-01	-1.1	-3.24	0.86	-1.29	-2.39	1.97
Saa1	3.82E-04	6.10E-01	-0.45	-3.46	1.36	NA[c]	-2.12	1.81

[a]Log_2(fold change) between different pairs of treatment-genotype groups

[b]False Discovery rate adjusted P-value for the overall treatment-by-genotype interaction.

[c]Not available

Many studies have associated inflammatory response to depression (Wiener et al., 2017). Peripheral levels of inflammatory cytokines might be associated with depression symptoms (Capuron and Miller, 2011). People with bipolar disorder had peripheral levels of Il10 positively correlated with functional impairment of the disorder (Wiener et al., 2017). Il10 is an immunosuppressive cytokine produced by almost all cells of the innate and adaptive arms of the immune system and can be anti- and pro-inflammatory. Il10 is engaged in the activation of the JAK-STAT signaling pathway and is involved with NF-kappa-B (NF-kB) activity (Ma et al., 2015). Il10 was under-expressed in WT relative to IDO1-KO mice in macrophages in a previous study with the same mice genotype (Gonzalez-Pena et al., 2016b).

Sphk1 plays an important role in inflammatory, anti-apoptotic, and immune processes by activating of tumor necrosis factors (TNF)-alpha signaling and the NF-kB pathways (Alvarez et al., 2010). Sphk1 is an isoenzyme that generates sphingosine-1-phosphate (S1P) in the central nervous system (CNS), S1P is a key regulator of cell death and survival, and the decrease of Sphk1

in hippocampus might be associated with neurological problems (Alvarez et al., 2010; Zhang et al., 2013). Reduced levels of SP1 and abnormal sphingolipid metabolism were reported in the brain with Alzheimer's disease (Zhang et al., 2013). Consistent with this, the present results showed over-expression of Sphk1 in the Sal-WT and Sal-KO, compared with BCG-WT and BCG-KO, which might be associated with a less protective activity of Sphk1 occurring during inflammation.

Serum amyloid A (Saa) is an apolipoprotein used as a biomarker of infection and inflammation (Ma et al., 2015). A study characterized Saa as a novel mechanism for regulation of inflammation and immunity by an acute-phase protein by activating IL-23/IL-17 pathway (Ma et al., 2015). Saa1 is highly expressed in the inflammation response pathway and tissue injury (Jang et al., 2017). The presence of Saa1 was associated with depressive-like behavior (Jang et al., 2017). Consistent with this result, another study found a gene from the same family (Saa3) in mice brain under BCG-immune challenge (Gonzalez-Pena et al., 2016a).

Chemokines are chemotactic cytokines coordinators of the immune system (Leighton et al., 2017). Ccl7 and Cxcl1 are from C-C subfamily of chemokines, and CXC subfamily of chemokines, respectively. Chemokines of these both subfamilies have a relationship with cognitive and neuromodulatory functions (Leighton et al., 2017). Ccl7 recruits pro-inflammatory cells and drive the response to an inflammatory activated state that contributes to depression through many mechanisms including enhancement of neurotoxicity, disruption of neurotransmitter systems, and aberrant tryptophan metabolism (Dantzer et al., 2008; Stuart et al., 2015). The up-regulation of Cxcl1 in CNS or peripherally occurs in response to stressor factors as peripheral lipopolysaccharide (LPS) injection (Girotti et al., 2011; Stuart et al., 2015). Ccl7 and Cxcl1 were both over-expressed in Sal-(KO) relative to Sal-(WT).

A tropic factor that is induced by inflammatory stimuli is known as Hbegf, which is highly induced in astrocytes during multiple sclerosis lesion formation (Schenk et al., 2013). In the brain, neurotrophins are essential for neurons survival, this gene is considerate important for neural development, and promoting survival of dopaminergic neurons (Oyagi et al., 2009). Hbegf knockout in mice's brain exhibited molecular biological characteristics of abnormal behavior similar to psychiatric disorders as depression and schizophrenia (Oyagi et al., 2009).

Cth produces the enzyme cystathionine gamma-lyase (Cse) that transforms cystathionine into cysteine. Cse is involved in the hydrogen sulfide pathway (H2S) biosynthesis, and H2S could be an inflammatory mediator in human's macrophages via the NF-kB/ERK pathway (Badiei et al.,

2015; Liu et al., 2017a). This gene was up-regulated in macrophages of mice under LPS challenge, Cse/H_2S biosynthesis ameliorated the inflammatory response (Liu et al., 2017a). The Cse/H_2S system had a protective function against the LPS-induced inflammation in another study in mice's vascular cells under immune challenge (Bourque et al., 2017). This study showed over-expression of Cth in the Sal-WT and Sal-KO, compared with BCG-WT and BCG-KO, more studies should be done to understand why the activity of Cth is low during inflammation.

Creb3l1 and Tpbg exhibited a significant treatment-by-genotype interaction. Creb3l1 may have a role blocking the proliferation of virus-infected cells and inhibiting virus activity (Denard et al., 2011). Tpbg is a member of glycoprotein containing leucine-rich repeats that are used as a target antibody-mediated immune responses (Harrop et al., 2006). Creb3l1 and Tpbg were under-expressed in BCG relative to Sal treated mice in both genotypes. In this study, some genes that have a protective function have decreased expression during inflammation.

Functional categories enriched among genes exhibiting treatment-by-genotype interaction effects

Enriched functional categories using 546 genes exhibiting an interaction effect (P-value <0.05) were investigated by DAVID. **Table 7** summarizes informative clusters that had Enrichment Score > 3. Enriched functional categories confirm the treatment-by-genotype interaction effect on genes involved in the regulation of the GO terms immune system, inflammatory response, generation of neurons, cardiovascular system development, regulation of cellular component movement, and response to lipopolysaccharide identified among genes with a significant interaction effect.

In addition to Cxc1, ccl7, and Il10 discussed before; six more genes in the chemokine subfamily were part of the enriched GO term positive regulation of immune system process category. Other important genes present in this enriched GO term are: Poliovirus Receptor (Pvr), Toll-Like Receptor 1 (Tlr1), Cd14 Molecule (Cd14), and Leucine-Rich Repeat Containing G Protein-Coupled Receptor 4 (Lgr4). This enrichment is in accordance with another study of BCG-challenged and control mice in microglia and macrophages (Gonzalez-Pena et al., 2016a).

Some genes that were enriched in many GO terms are essential to understanding regulation of immune system process. Pvr plays an important role in microbial infection acting as a cellular

receptor for poliovirus replication mediates natural killer (NK) cell adhesion and leads to immunological synapse between the target cell and NK cell (Fuchs et al., 2004). Tlr1 is from the Toll-like receptor (TLR) family, important to develop effective immunity and acts with Cd14 in response to triacylated lipopeptides (Oliveira-Nascimento et al., 2012). The Cd14 mediates inflammatory response, innate immune response to bacterial lipopolysaccharide and is expressed in monocytes/macrophages. Proinflammatory cytokine secretion leads to NF-kB activation (Arroyo-Espliguero et al., 2004). Lgr4 act as negative regulator of innate immunity and is important for autoimmune diseases (Du et al., 2013).

From the enrichment of the category generation of neurons, some genes presented in this GO term were discussed in association with behavior disorders, like SH3 and multiple ankyrin repeat domains 3 (Shank3), Wnt family member 2 (Wnt2), and solute carrier family 6 member 4 (Slc6a4). The Shank3 is an intracellular scaffolding protein associated with neurotransmitter receptors, synapse formation, schizophrenia, and autism (Südhof, 2008). Wnt2 is a secreted growth factor important for developmental processes and multiple biological functions like cellular differentiation, and regulation of neuronal migration was associated with autism (Won et al., 2013). Slc6a4 is a member of neurotransmitter transports' family associated with transport of serotonin, is a target of psychomotor stimulants, and has been related to depression (Margoob and Mushtaq, 2011).

The GO term cardiovascular system development included two genes identified for interaction effect, Hbegf, and Cth, among others important for this system like ryanodine receptor 2 (Ryr2) and adam metallopeptidase domain 19 (Adam19). Ryr2 is found in the cardiac muscle sarcoplasmic reticulum and is a component of calcium channel, important for the cardiac muscle contraction (Lanner et al., 2010). Adam19 has main roles in heart development, neurogenesis, angiogenesis, and release of proteins like EGF receptor ligands (Zhou et al., 2004).

From the genes annotated to the GO term response to lipopolysaccharide, Cxcl1, Il10, and Admt are also involved in other pathways like GO terms regulation of cell proliferation and positive regulation of the metabolic process. Other genes that are participating in the GO term response to lipopolysaccharide is platelet factor 4 (Pf4). The Pf4 is a regulator of vascular and immune biology has a direct role in B-cell differentiation in bone marrow, has antimicrobial activity and inhibits hematopoiesis, angiogenesis and T-cell function (Field et al., 2017).

GO terms that are involved in categories such as regulation of leukocyte chemotaxis, response to peptide, leukocyte activation, regulation of transmembrane receptor protein serine/threonine kinase signaling pathway, and peptidyl-tyrosine phosphorylation were enriched in this study in response to immune challenge. The treatment-by-genotype interaction effects on pathways and processes helped to identify mechanisms and specific genes that can be targeted to address depression-like behaviors after inflammation or to understand effects of IDO1 knockout in the whole process.

Table 7. Informative clusters (Enrichment Score > 3) of descriptive DAVID FAT categories including Gene Ontology (GO) biological processes (BP) and KEGG pathway enriched among 546 genes exhibiting significant treatment-by-genotype interaction effects.

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
Cluster 1	**Enrichment Score: 16.44**			
BP FAT	GO:0051272~positive regulation of cellular component movement	46	2.21E-15	4.30E-12
Cluster 3	**Enrichment Score: 8.87**			
BP FAT	GO:0002684~positive regulation of immune system process	48	6.00E-09	1.16E-05
Cluster 4	**Enrichment Score: 8.64**			
BP FAT	GO:0006954~inflammatory response	43	9.10E-10	1.76E-06
Cluster 12	**Enrichment Score: 5.25**			
BP FAT	GO:0042981~regulation of apoptotic process	69	1.26E-07	2.43E-04
Cluster 14	**Enrichment Score: 4.75**			
BP FAT	GO:0048699~generation of neurons	66	3.28E-06	6.34E-03
Cluster 15	**Enrichment Score: 4.64**			
BP FAT	GO:0072358~cardiovascular system development	56	5.03E-08	9.73E-05
Cluster 16	**Enrichment Score: 4.38**			
BP FAT	GO:0032496~response to lipopolysaccharide	26	1.27E-05	2.47E-02
Cluster 20	**Enrichment Score: 4.08**			
BP FAT	GO:0018108~peptidyl-tyrosine phosphorylation	21	9.83E-06	1.90E-02
Cluster 25	**Enrichment Score: 3.89**			
BP FAT	GO:0090092~regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	18	1.26E-05	2.43E-02
Cluster 26	**Enrichment Score: 3.82**			
BP FAT	GO:0002688~regulation of leukocyte chemotaxis	13	1.15E-05	2.22E-02

Table 7 (cont.)

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
Cluster 36	**Enrichment Score: 3.19**			
KEGG PATHWAY	mmu04390:Hippo signaling pathway	16	2.40E-05	3.06E-02

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

Transcription factor network analysis of genes exhibiting treatment-by-genotype interaction effects

The study of TFs is important to detect regulators of the genes associated with treatment-by-genotype interaction effects. Among the 65 genes identified (P-value < 0.005) by interaction effects, 20 genes with P-value < 0.002 were present in the network analysis. **Figure 8** shows the regulatory networks for TF exhibiting a normalized enrichment score (NES) > 3.0, the top enriched TFs with higher NES were selected. **Table 8** lists the enriched TFs exhibiting a normalized enrichment score > 2.0 among the 65 genes identified (P-value < 0.005) by interaction effects.

Table 8. Transcription factors (TF) enriched (normalized enrichment score NES > 3.0) among differentially expressed (number N of target genes = 65) genes exhibiting significant treatment-by-genotype interaction effects at P-value <0.005.

Transcription factor	NES[a]	Target Gene N[b]
Pparg	5.186	29
Ugp2	4.726	9
E2f1	4.479	27
Jazf1	4.346	11
Klf13	4.217	15
Spib	4.189	10
Hnrnph3	4.170	23
Pura	4.151	28
Mafa	3.877	14
Pou2f1	3.589	8
Sod1	3.509	5
Hand2	3.482	30
Foxc2	3.440	17

Table 8 (cont.)

Transcription factor	NES[a]	Target Gene N[b]
Hey1	3.422	23
Ing4	3.41	7
Kdm5a	3.366	17
Mtf1	3.215	12
Gata2	3.207	10
Jdp2	3.169	17
Socs4	3.139	13
Zcchc14	3.114	20
Hand2	3.111	11
Purg	3.033	13
Tfap4	3.024	35
Nr4a2	2.996	11
Mylk	2.978	41
Smad4	2.956	15
Tead3	2.910	25
Mettl14	2.880	9
Nxph3	2.870	9
Ahr	2.836	12

[a]Normalized enrichment score

[b]Number N of target genes

Among genes identified by iRegulon network analysis (**Figure 8**), the Creb3l1, Hbegf, Sphk1, and Tpbg were already discussed in the treatment-by-genotype interaction effects on the macrophage transcriptome section. The most enriched TF peroxisome proliferator-activated receptor gamma (Pparg) is involves in antidepressant-like activity and is a regulator of adipogenesis, being an important parameter in the obesity, which is associated with inflammation (Donma and Donma, 2016). Pparg is regulating both Hbegf and Sphk1. Both TFs UDP-Glucose Pyrophosphorylase 2 (Ugp2) and JAZF Zinc Finger 1 (Jazf1) are regulating the gene Creb3l1. The gene E2F transcription factor 1 (E2fr) promote expression of pro-apoptotic factors and is upregulated during neuroinflammation (Wu et al., 2014). The TF E2fr is regulating Tpbg.

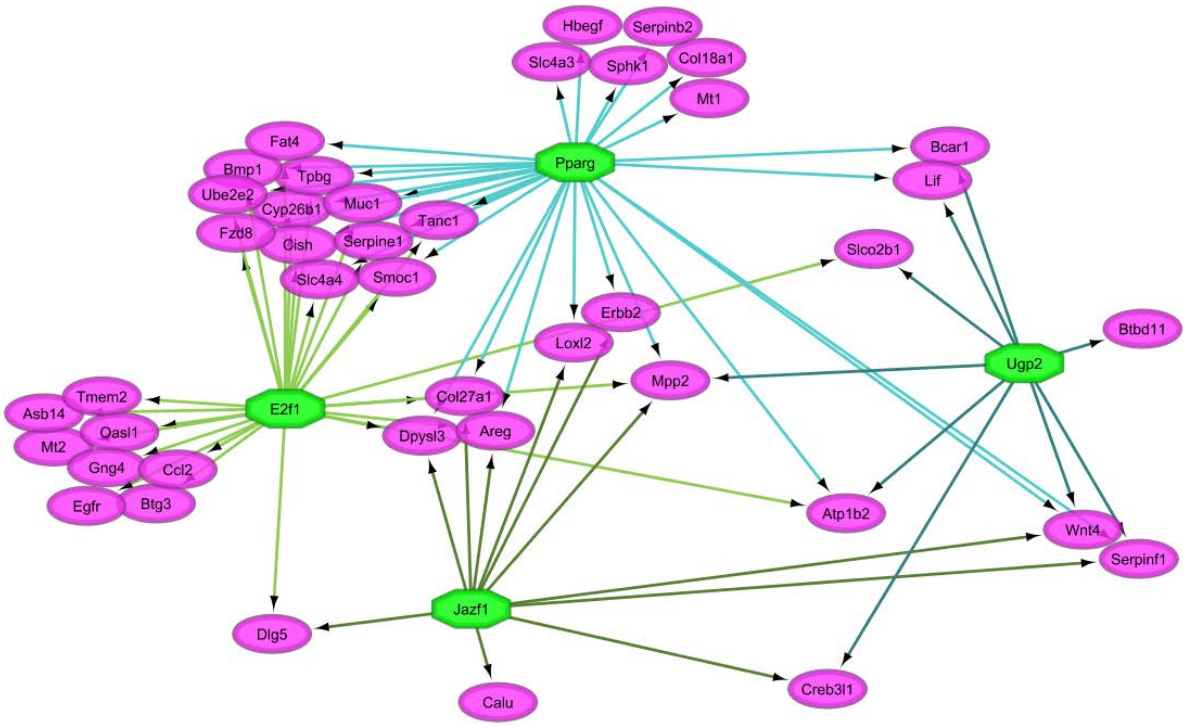

Figure 8. Regulatory network of transcription factors (normalized enrichment score > 3.0) and the target genes exhibiting significant treatment-by-genotype interaction effects at P-value <0.005. Green octagons denote TFs. The pink ovals are targets.

The identification of TFs with regulatory networks confirmed the findings discussed in the treatment-by-genotype interaction effects on the macrophage transcriptome section and both regulons and regulators could be potential candidates to understand the mechanisms involved in the interaction effects.

Effect of IDO1 genotype on the macrophage transcriptome

The processes and pathways of immune systems activating macrophages are very complex. Having a gene knockout in macrophages may be important to understand molecular mechanisms affecting the regulation and dysregulation of specific genes. This study aims to understand the regulation of gene expression in the macrophages in the presence and absence of IDO1, a gene that is highly regulated by cytokines (Baumgartner et al., 2017). **Table 9** lists four significant (FDR-adjusted P-value < 0.05) differentially expressed genes between mice with IDO1-KO and WT line.

The genes differentially expressed might be related to other mechanisms of control occurring in the macrophage to compensate for the absence of IDO1 in the cell. Macrophages are not the only immune cells that produce cytokines, other cells, like NK cells, have the same function (Martinez and Gordon, 2014). IDO1 deficiency in dendritic cells leads to decrease of regulatory T cells (Tregs) and Il10 in serum. However, controversial results for IDO1 deletion in the inflammatory response were also existed (Baumgartner et al., 2017).

In this study, a gene from transmembrane protein's (Tmem) family, Tmem181, was under-expressed in the IDO1-KO relative to the WT genotype group. Tmem181 is a receptor for cytolethal distending toxins (CDTs), which are potent toxins secreted by bacteria or virus. Mutant Tmem181 may have a role in the rate-limiting intoxication, where the CDT can bind the protein of Tmem181's cell surface (Carette et al., 2009). This gene might be essential to understanding pathogenesis and mechanisms of pathogen intoxication in the macrophages. The Tmem181 activity may be affected by the lack of IDO1 in the inflammation response in the IDO1-KO genotype group.

Another differentially expressed gene that might be involved in the suppression of IDO1 is the dynein light chain tctex-type 1 (Dynlt1). Dynein chains interact with viral proteins and are involved in the cellular cytoplasmic transportation mechanism that transport organelles along the microtubules, permitting the virus to travel and reach their replication sites (Merino-Gracia et al., 2011). Dynein motor is essential for the successful viral infection. The over-expression of dynein light chains could increase the infection of papillomavirus, while the Dynlt1 depletion can inhibit the infection of this virus (Schneider et al., 2011). Dynlt1was under-expressed in the IDO1-KO relative to WT genotype group, giving a higher anti-viral response in the macrophages of IDO1-KO group.

The differentially expressed gene Relt like 1 (Rell1) is homolog/paralog of Relt, from the TNF-receptor superfamily. TNF is important for regulating the immune response, inflammation, and for participating in many processes in the cell (Cusick et al., 2010). Rell1 has been reported as a stimulator of T-cell proliferation and responsible for activating an apoptotic pathway in epithelial cells (Cusick et al., 2010). IDO1 over-expression can also induce cellular death, by producing kynurenine (Kyn), which has pro-apoptotic properties (Asghar et al., 2017). On the other hand, supplemented Kyns can suffer cytotoxic and anti-proliferative effects caused by human

T cells (Asghar et al., 2017). Although the mechanisms involving T cells and IDO1 are still ambiguous, Rell1 and IDO1 might play a common role in apoptosis.

Midline1 (Mid1) was associated with the airway hyper-responsiveness (AHR), a chronic airway inflammation (Foster et al., 2017). The inhibition of the activity of Mid1 can decrease the expression and production of some interleukins, and also decrease AHR (Foster et al., 2017). This gene can indirectly promote type 2 helper t cells (Th2) responses, which are essential for the immune system (Foster et al., 2017). Mid1 was over-expressed in the IDO-KO relative to the WT group, which might be related to a mechanism that suppresses the absence of IDO1.

The number of genes differentially expressed prevented the identification of enriched categories using DAVID. The results of differentially expressed genes showed that few genes are involved in the differences between IDO1 phenotype and WT, suggesting that the macrophages may find a way to compensate the absence of IDO1 restoring the cell function. The mechanisms involved with this possible suppression can be the activation of inflammatory cells and induction of immune tolerance by divergent pathways.

Table 9. Differentially expressed genes (FDR-adjusted P-value < 0.05) between mice in wild-type (WT) and IDO1 knocked out (KO) genotypes.

Gene Symbol	Description	Log2(IDO1-KO/WT)	Raw P-value	FDR adjusted P-value[a]
Tmem181b-ps	Transmembrane protein 181B, pseudogene	-0.91	1.60E-42	2.30E-38
Dynlt1b	Dynein light chain Tctex-type 1	-1	2.00E-21	1.40E-17
Rell1	RELT-like protein 1 precursor	0.4	1.50E-07	7.30E-04
Mid1	PREDICTED: E3 ubiquitin-protein ligase Midline-1	0.78	4.20E-06	1.50E-02
G530011O06Rik	RIKEN cDNA G530011O06 gene	1.03	7.4E-05	2.1E-01
Rps4l	Ribosomal protein S4-like	0.36	9.4E-05	2.2E-01
Srpx	Sushi-repeat-containing protein SRPX precursor	1.20	3.0E-04	5.3E-01
Pde8a	High affinity cAMP-specific and IBMX-insensitive 3\',5\'-cyclic phosphodiesterase 8A	-0.22	3.2E-04	5.3E-01
Gm33272	Predicted gene, 33272	-0.72	3.3E-04	5.3E-01

[a]False Discovery Rate adjusted P-value.

Effect of BCG infection on the macrophage transcriptome

The transcriptome of macrophages BCG-infected compared to the Sal unchallenged mice, allowed to find many possible pathways involved in the recovery process post-infection. Among the 8,798 significant differentially expressed genes (FDR-adjusted P-value < 0.05), 4,530 genes were under-expressed in the BCG relative to Sal treatment and 4,268 genes were over-expressed in the BCG relative to Sal treatment. Most significant genes that were under-expressed and over-expressed in BCG relative to Sal treatment are listed in **Tables 10**, and **11**, respectively.

Table 10. Top 20 differentially expressed genes (FDR-adjusted P-value < 0.05) under-expressed in the BCG relative to Sal treatment.

Gene Symbol	Description	Log2(BCG/Sal)	Raw P-value	FDR adjusted P-value[a]
Ly6i	PREDICTED: lymphocyte antigen 6I	-9.20	1.0E-66	1.5E-62
Ly6c2	PREDICTED: lymphocyte antigen 6C2	-8.09	6.9E-66	5.0E-62
Zbp1	Z-DNA-binding protein 1	-4.43	2.2E-60	1.1E-56
Fcgr1	high affinity immunoglobulin gamma Fc receptor I precursor	-5.62	1.2E-59	3.5E-56
Pira6	paired-Ig-like receptor A6	-4.68	1.2E-59	3.5E-56
AW112010	expressed sequence AW112010	-4.79	9.5E-59	2.3E-55
Spon1	spondin-1 precursor	-6.99	3.2E-57	5.8E-54
Smpdl3b	acid sphingomyelinase-like phosphodiesterase 3b precursor	-4.30	3.3E-57	5.8E-54
Lair1	leukocyte-associated immunoglobulin-like receptor 1	-4.23	3.7E-57	5.8E-54
Gpr141	PREDICTED: probable G-protein coupled receptor 141	-6.33	1.8E-56	2.6E-53
Gbp2	guanylate-binding protein 1	-4.38	1.7E-55	2.2E-52
Tap1	antigen peptide transporter 1	-2.87	1.0E-54	1.2E-51
Ms4a6b	membrane-spanning 4-domains subfamily A member 6B	-3.68	1.4E-54	1.5E-51
Fcgr4	low-affinity immunoglobulin gamma Fc region receptor III-A precursor	-4.08	1.8E-54	1.8E-51
I830127L07Rik	PREDICTED: lymphocyte antigen 6A-2/6E-1	-8.85	6.9E-54	6.6E-51
Aif1	PREDICTED: allograft inflammatory factor 1	-4.45	7.7E-54	6.9E-51
Hrh2	PREDICTED: histamine H2 receptor	-4.16	1.8E-53	1.5E-50
Il18bp	PREDICTED: interleukin-18-binding protein	-3.48	2.4E-53	1.9E-50
Bst1	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2 precursor	-2.34	4.6E-53	3.5E-50
Lst1	leukocyte-specific transcript 1 protein	-4.69	7.6E-53	5.5E-50

[a]False Discovery Rate adjusted P-value.

Many under-expressed genes in the BCG relative to Sal treatment (**Table 10**), are associated with microbial response and host defense, macrophage structure, function, or metabolism. The genes Z-DNA binding protein 1 (Zbp1), leukocyte-associated immunoglobulin-like receptor 1 (Lair1), high affinity immunoglobulin gamma Fc receptor I (Fcgr1), low affinity immunoglobulin gamma Fc region receptor III-a precursor (Fcgr4), allograft inflammatory factor 1 (Aif1), and guanylate binding protein 2 (Gbp2) have the host defense function in common. The genes spondin 1 (Spon1), sphingomyelin phosphodiesterase acid like 3b (Smpdl3b), transporter 1, ATP binding cassette subfamily b member (Tap1), share the macrophage component function.

The Zbp1 is a Z-DNA binding protein that interacts with RIP homotypic interaction motif (Rhim), a domain that is present in the protein encoded by Zbp1 and may lead to the activation of NF-kB, caspase 8-dependent apoptosis, and interferon production (Upton and Kaiser, 2017). Zbp1 plays a role in the innate immunity and limit some virus infection (Upton and Kaiser, 2017). The absence of this gene can cause damage in parasite degradation by the system, augmenting the parasite number, and parasite survival (Pittman et al., 2016). The under-expression of this gene in the BCG relative to Sal treatment is different from studies that found Zbp1 more expressed during the infection, nevertheless, this decrease of expression might be related to a lack of activation that is induced by IFN-y, favoring the parasite growth (Pittman et al., 2016).

Lair1 belongs to an immunoglobulin superfamily and produces a signal that downregulates NK cells, T cells, and B cells (Poggi and Zocchi, 2014). Lair1 is present in leukocytes, and the decrease or absence of this gene can cause autoimmune or neoplastic diseases, like B cell chronic leukemia (Poggi and Zocchi, 2014). Lair1 was recently classified as an antibody, which are proteins of the immune system that respond to the presence of bacteria, virus, and parasites (Hsieh and Higgins, 2017).

Fcgr1 is a receptor that has a high affinity for antibodies, binding to immunoglobulin-G (IgG) in neutrophils and is related to the immune system (Minett et al., 2016). As a leukocyte surface antigen, Fcgr1a plays a role in the innate and adaptive responses (Farias et al., 2014). The expression of this gene in microglia was associated with neurodegenerative pathologies like dementia, and bad cognition (Minett et al., 2016). The expression of Fcgr1a can be induced by inflammatory cytokines in the presence of bacteria and is over-expressed during infection on the surface of neutrophils (Farias et al., 2014). This gene may have a potential therapeutic role and may be used as a biomarker for infection or sepsis, and as measurer of systemic inflammatory

responses (Farias et al., 2014). Fcgr1a was under-expressed in the BCG relative to Sal treatment. This profile is opposite to that previously reported suggesting that the expression of this gene can also be negatively affected by the presence of high level of inflammation. Another gene from the CD molecules family was under-expressed in the BCG relative to the Sal treatment Fcgr4is part of NK cells and have a cytotoxic function (Suzuki et al., 2017). A study showed that low levels of $CD56^{+}CD16^{-}$ are associated with depression, sleep disturbance, and vulnerability to viral infection (Suzuki et al., 2017).

The gene Aif1 induces macrophages activation and might be related to bipolar disorder and sleep disorder (Cai et al., 2017; Kittel-Schneider et al., 2017; Wadhwa et al., 2017). In this findings, Aif1 is under-expressed in the in the BCG relative to Sal treatment. In a study of depression-related symptoms, Aif1, that is also a glial cell marker in the brain, had elevated expression in a group under LPS challenge. In the same study, IDO was up-regulated in the brain of animals that suffered LPS treatment, suggesting that the peripheral immune challenge can cause neuroinflammation and kynurenine production (Parrott et al., 2016).

Gbp2 is essential for antibacterial defenses, inhibiting bacterial replication, and providing host protection against different types of pathogens (Li et al., 2017b). Mutation in this gene was associated with a migraine, and the absence of Gbp2 expression was associated with no migraine (Jiang et al., 2015). In macrophages, genes from Gbp family can also promote caspase-11 activation when infected with bacteria, regulating inflammation and cell death, mice deficient of Gbp have higher mortality in LPS infection than compared with the WT mice (Finethy et al., 2017). In these results, Gbp2 is under-expressed in the BCG relative to the Sal treatment.

Some differentially expressed genes were also related to macrophage component function. Spon1 encodes the protein F-spondin that is induced by neuronal injury and damage the binding of cells to the extracellular matrix protein (ECM) (Jahanshad et al., 2013). Spon1 interacts with receptors' family of alipoliprotein E (Apoe), which might be a strong genetic risk factor for Alzheimer's disease (Jahanshad et al., 2013). The over-expression of Spon1 is related to cognitive improvements (Jahanshad et al., 2013). The under-expression of Spon1 in the BCG relative to the Sal treatment might be associated with a difficulty to recover the normal functions caused by the infection.

The Smpdl3b is involved in the innate immunity signaling, and TLR-induced signaling processes have a role in the lipid metabolism, and macrophages that are deficient in this gene

present a hyper-responsiveness to TLR (Heinz et al., 2015). Smpdl3b is also a high phosphodiesterase active in macrophages surface, and a reduction in the expression of this gene causes reduction of the macrophage order (Heinz et al., 2015). Cells with Smpdl3b depletion present hyper-inflammatory response in the presence of lipid supplementation (Heinz et al., 2015). The under-expression of this gene in the BCG-infected group might be related to the damage caused by inflammation, compared to the Sal treated group that presents normal levels of expression of Smpdl3b.

Tap1 is from the ATP-binding cassette (ABC) transporters family, and has a role in the transportation of peptides across the ER membrane, the depletion of this gene is related to the impairment of assembly and transport of MHC class I (Nelson et al., 2013). Loss of MHC class I shows a deficit of memory, being important in synaptic plasticity and cognition, and might be related to depression (Nelson et al., 2013). These results are in agreement with findings of this study where Tap1 was under-expressed in the BCG relative to Sal treatment, suggesting a presence of depression-like behavior.

The results of under-expressed genes showed that many genes are involved in the differences between mice infected with BCG and with Sal. Some of the top genes presented in **Table 10** were discussed and showed direct mechanisms of action in the host defense that might be damaged, by destroying the action against the microbial infection or preventing protective mechanisms to improve the impaired functions caused by the infection. Some genes were related to neurodegenerative disorders or depression, confirming the hypothesis that the BCG infection might have a role in the cause of depression-like behaviors. Also, another suggestion for this results is that the mechanisms involved in the inflammation process can be influencing other organs in a similar way.

Table 11 lists the top 20 genes over-expressed in the BCG compared with Sal treatment. Genes over-expressed in the BCG relative to the Sal treatment, similarly to those under-expressed have roles related to the microbial response and host defense, and macrophage structure. The genes lysozyme 1 (Lyz1) and aquaporin 9 (Aqp9) have the microbial response function. The genes mitochondrial calcium uniporter (Mcu) and histamine receptor h1 (Hrh1) are associated with macrophage structure.

Lyz1 encodes lysozymes, that are antimicrobial agents, and related to immune system, playing a role in the phagosome, and activation of NLR family pyrin domain containing 3 (Nlrp3),

which regulates inflammation, apoptosis, and cytokine production (Vance, 2010). Digestive enzymes activate immune sensors inside the cell, and the resistance of bacteria to lysozyme avoid the recognition of their presence by Nlrp3 inflammasome (Shimada et al., 2010; Vance, 2010). Lysozyme resistance facilitates the bacteria to evade the inflammasome (Vance, 2010). Also, lysozyme is important for innate response and upregulated against infection by pathogens (Gao et al., 2017). Lyz1 is over-expressed in the BCG relative to the Sal treatment. Aqp9 is from aquaporin's (Aqp) family, and the high expression of Aqp9 was associated with infectious disease and systemic inflammatory response syndrome compared with healthy controls (Holm et al., 2016). In agreement with this study, Aqp9 is over-expressed in the BCG relative to the Sal treatment. The over-expression of both Lyz1 and Aqp9 in the BCG relative to the Sal treatment might be confirming the presence of inflammation and immune system reaction.

Mcu is a mediator of ER calcium uptake by mitochondria, regulates metabolism, cell viability, and ROS production (Gu et al., 2017). Mcu silencing can correct calcium defect associated with an increase of excessive mitochondrial calcium (Wang et al., 2017). This defect is caused by the clearance of apoptotic cells by phagocytes (Wang et al., 2017). Mice that were Mcu deficient in macrophages were protected from pulmonary fibrosis (Gu et al., 2017). Mcu is over-expressed in the BCG relative to the Sal treatment suggesting damage in the immune response.

Hrh1 encodes a histamine G-coupled receptor (H1R) and is part of the histaminergic system (Wright et al., 2017). Some genes of this system play a role in the neuroinflammation, cognition, sleep, and attention (Wright et al., 2017). Hrh1 was associated with NF-kb signaling, and cytokine production (Virakul et al., 2016). The Hrh1 was over-expressed in in the BCG relative to the Sal treatment. Genes (FDR-adjusted P-value < 0.05) over-expressed in the BCG relative to the Sal treatment showed the potential relationship between inflammation and behavioral problems. These results can be used to understand the role of treatment in depression-like behaviors.

Table 11. Top 20 differentially expressed genes (FDR-adjusted P-value < 0.05) over-expressed in the BCG relative to Sal treatment.

Gene Symbol	Description	Log2(BCG/Sal)	Raw P-value	FDR adjusted P-value[a]
Lyz1	lysozyme C-1 precursor	1.29	8.8E-27	3.7E-25
Mcu	PREDICTED: calcium uniporter protein, mitochondrial	0.81	7.1E-26	2.8E-24
Neurl1a	E3 ubiquitin-protein ligase NEURL1	1.76	8.4E-26	3.2E-24
Fam198b	PREDICTED: protein FAM198B	2.07	1.1E-24	3.7E-23
Hrh1	histamine H1 receptor	1.65	2.3E-24	7.5E-23
Aqp9	aquaporin-9	1.41	2.3E-24	7.7E-23
Dcstamp	PREDICTED: dendritic cell-specific transmembrane protein	2.27	7.0E-24	2.2E-22
Ear2	eosinophil cationic protein 2 precursor	2.30	8.0E-24	2.5E-22
LOC108168644	PREDICTED: proline-rich protein 36-like	2.24	9.2E-24	2.9E-22
Cdkn1c	PREDICTED: cyclin-dependent kinase inhibitor 1C	1.79	2.5E-23	7.6E-22
Cd200r3	PREDICTED: cell surface glycoprotein CD200 receptor 3	2.10	2.8E-23	8.4E-22
Tspyl2	PREDICTED: testis-specific Y-encoded-like protein 2	0.88	2.9E-23	8.6E-22
Epas1	endothelial PAS domain-containing protein 1	2.03	1.9E-22	5.1E-21
Cpeb1	PREDICTED: cytoplasmic polyadenylation element-binding protein 1	2.16	2.8E-22	7.5E-21
Atp6v0a1	V-type proton ATPase 116 kDa subunit a	1.10	4.6E-22	1.2E-20
Hpn	serine protease hepsin	2.03	8.7E-22	2.2E-20
Gbe1	1,4-alpha-glucan-branching enzyme	1.28	1.2E-21	2.8E-20
Lrrc32	PREDICTED: leucine-rich repeat-containing protein 32	2.39	2.3E-21	5.5E-20
F13a1	coagulation factor XIII A chain precursor	1.81	6.5E-21	1.5E-19
Nav2	PREDICTED: neuron navigator 2	1.70	7.5E-21	1.7E-19

[a]False Discovery Rate adjusted P-value.

Functional analysis of genes exhibiting BCG treatment effect

The functional analysis of genes exhibiting significant BCG infection effects can complement our findings in the differential expression analysis. However, different from the pattern assumed for the differentially expressed analysis of BCG-infected effects, the direction of regulation in over-expression or under-expression was not separated to find informative categories related to BCG-challenged. Functional analysis of 2,255 differentially expressed genes (FDR-adjusted P-value < 0.05) with log2 (fold change) >|1| was implemented using DAVID. Most informative clusters (Enrichment Score > 2) of descriptive DAVID FAT categories including Gene

Ontology (GO) biological processes (BP), molecular functions (MF), and KEGG pathways enriched among genes exhibiting significant effects are described in **Table 12**.

Enriched functional categories confirm the patterns previously discussed including GO BP terms related to microbial response and host defense, and macrophage structure or metabolism. For microbial response and host defense the following categories were enriched: response to virus, negative regulation of inflammatory response, response to lipid, positive regulation of cytokine production involved in immune response, antigen processing and presentation of peptide antigen via MHC class I, positive regulation of T cell mediated cytotoxicity, positive regulation of cell death, myeloid cell activation involved in immune response, negative regulation of viral life cycle, interaction with host, positive regulation of response to wounding, antigen processing, and presentation of endogenous peptide antigen (**Table 12**).

Myocyte enhancer factor 2c (Mef2c), toll-like receptor 3 (Tlr3), and S100 calcium binding protein a8 (S100a8) are part of many functional categories. Mef2c was associated with cognition, behavior, and memory (Mitchell et al., 2017). The expression of genes of TLR family is associated with the major depressive disorder in the presence of inflammation (Hung et al., 2017). The presence of gene S100a8 in inflammation is in accordance with a study that had S100a8 over-expressed in response to BCG challenge in microglia cells (Gonzalez-Pena et al., 2016a). Another gene from the same family of S100a8, S100a9, was also associated with inflammation and depression (Ha et al., 2010).

For functional categories related to macrophage structure, cellular function, and metabolism, the following GO BP and MF terms were enriched: regulation of GTPase activity, cellular response to peptide hormone stimulus, nucleoside-triphosphatase regulator activity, regulation of peptidyl-tyrosine phosphorylation, cellular calcium ion homeostasis, phosphatidylinositol-mediated signaling, regulation of cysteine-type endopeptidase activity, and regulation of peptidase activity (**Table 12**).

In agreement with this study, genes related to metabolism of regulatory peptides, inflammatory response, and immune system were associated with major depressive disorder (Song et al., 2013). Alanyl aminopeptidase, membrane (Anpep) and matrix metallopeptidase 8 (Mmp8) are associated with both inflammatory response and depression (Song et al., 2013). In this study, these genes are enriching the significant GO BP term anatomical structure formation involved in

morphogenesis. Pathways related to inflammatory responses, energy metabolism, including hormonal signals were also associated with depression in humans (Carboni et al., 2016).

Seven KEGG pathways were enriched in this analysis: graft-versus-host disease; antigen processing, and presentation; type I diabetes mellitus; axon guidance; chemokine signaling pathway; ECM-receptor interaction; focal adhesion. These results are highlighting the host defense mechanism that is acting in the presence of infection. The KEGG pathways together suggest that the animals might have behavior-like symptoms. First, because genes related to neuronal function, axon guidance, and synaptic vesicle cycle were associated with symptoms of depression (Brites and Fernandes, 2015; Carboni et al., 2016). Second, because focal adhesion and chemokine signaling pathway were also described as significantly enriched for major depressive disorder (Carboni et al., 2016). Finally, Type I diabetes mellitus was significantly enriched in the KEGG pathway, which is consistent with the findings that glucose metabolism, gluconeogenesis, and glycolysis are associated with the disorder. Also, the risk for type 2 diabetes mellitus may be increased in patients with depression (Carboni et al., 2016).

The enrichment of categories among genes exhibiting significant BCG effects relative to Sal treatment using GSEA confirmed and complemented the results of GO terms related to immune system. These findings allowed a better understanding of the processes that might be happening in the macrophage during the inflammation response affecting the whole immune system and consequently developing depression-like behaviors.

Table 12. Most informative clusters (Enrichment Score > 2) of descriptive DAVID FAT categories including Gene Ontology (GO) biological processes (BP), molecular functions (MF), and KEGG pathways enriched among genes exhibiting significant (FDR-adjusted P-value < 0.05) and with log2(fold change) >|1| between mice infected with BCG relative to Sal treatment.

Category	Term	Count[a]	Raw P-value	FDR P-value[b]
Cluster 11	**Enrichment Score: 15.49**			
BP_FAT	GO:0009615~response to virus	86	1.04E-12	2.11E-09
Cluster 14	**Enrichment Score: 11.10**			
BP_FAT	GO:0043087~regulation of GTPase activity	55	1.3E-05	2.64E-02
Cluster 15	**Enrichment Score: 10.80**			
BP_FAT	GO:0050728~negative regulation of inflammatory response	33	2.38E-07	4.82E-04

Table 12 (cont.)

Category	Term	Count[a]	Raw P-value	FDR P-value[b]
Cluster 17	**Enrichment Score: 8.62**			
BP_FAT	GO:0001916~positive regulation of T cell mediated cytotoxicity	11	1.98E-05	4.01E-02
Cluster 20	**Enrichment Score: 8.18**			
BP_FAT	GO:0010942~positive regulation of cell death	115	1.9E-09	3.85E-06
Cluster 22	**Enrichment Score: 7.42**			
BP_FAT	GO:0071375~cellular response to peptide hormone stimulus	50	1.12E-05	2.26E-02
Cluster 27	**Enrichment Score: 6.70**			
MF_FAT	GO:0060589~nucleoside-triphosphatase regulator activity	58	8.46E-07	1.45E-03
Cluster 29	**Enrichment Score: 6.52**			
BP_FAT	GO:0050730~regulation of peptidyl-tyrosine phosphorylation	52	1.14E-07	2.30E-04
Cluster 31	**Enrichment Score: 6.37**			
BP_FAT	GO:0033993~response to lipid	162	1.1E-08	2.22E-05
Cluster 34	**Enrichment Score: 6.25**			
BP_FAT	GO:0002720~positive regulation of cytokine production involved in immune response	15	3.67E-06	7.43E-03
Cluster 36	**Enrichment Score: 6.10**			
BP_FAT	GO:0006874~cellular calcium ion homeostasis	73	1.4E-05	2.83E-02
Cluster 37	**Enrichment Score: 5.94**			
BP_FAT	GO:0002474~antigen processing and presentation of peptide antigen via MHC class I	26	3.63E-08	7.36E-05
KEGG_PATHWAY	mmu05332:Graft-versus-host disease	21	7.43E-07	9.82E-04
KEGG_PATHWAY	mmu04612:Antigen processing and presentation	25	1.62E-05	2.15E-02
KEGG_PATHWAY	mmu04940:Type I diabetes mellitus	21	1.7E-05	2.24E-02
Cluster 40	**Enrichment Score: 5.58**			
KEGG_PATHWAY	mmu04360:Axon guidance	43	3.17E-10	4.19E-07
Cluster 41	**Enrichment Score: 5.562**			
BP_FAT	GO:0002275~myeloid cell activation involved in immune response	24	1.57E-06	3.17E-03
Cluster 43	**Enrichment Score: 5.43**			
BP_FAT	GO:1903901~negative regulation of viral life cycle	24	4.22E-06	8.55E-03
Cluster 45	**Enrichment Score: 5.31**			
BP_FAT	GO:0048015~phosphatidylinositol-mediated signaling	31	2.39E-05	4.84E-02
Cluster 47	**Enrichment Score: 5.14**			
KEGG_PATHWAY	mmu04062:Chemokine signaling pathway	49	4.9E-07	6.48E-04
Cluster 53	**Enrichment Score: 4.75**			
KEGG_PATHWAY	mmu04512:ECM-receptor interaction	28	1.78E-06	2.35E-03
KEGG_PATHWAY	mmu04510:Focal adhesion	49	2.65E-06	3.50E-03

Table 12 (cont.)

Category	Term	Count[a]	Raw P-value	FDR P-value[b]
Cluster 62	**Enrichment Score: 4.24**			
BP_FAT	GO:2000116~regulation of cysteine-type endopeptidase activity	45	2.25E-05	4.55E-02
Cluster 79	**Enrichment Score: 3.24**			
BP_FAT	GO:0052547~regulation of peptidase activity	67	3.02E-06	6.12E-03
Cluster 84	**Enrichment Score: 3.09**			
BP_FAT	GO:0051701~interaction with host	24	1.61E-05	3.26E-02
Cluster 96	**Enrichment Score: 2.87**			
BP_FAT	GO:1903036~positive regulation of response to wounding	20	1.26E-05	2.55E-02
Cluster 122	**Enrichment Score: 2.47**			
BP_FAT	GO:0002483~antigen processing and presentation of endogenous peptide antigen	11	7.4E-07	1.50E-03

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

Transcription factor network analysis of genes exhibiting BCG treatment effect

The understanding of TFs that affect the macrophage transcriptome is important to find sets co-dysregulated genes associated with BCG infection effects. The top significant genes (4,455) that were differentially expressed (FDR adjusted P-value < 2.0E-6) for BCG infection effects were used for transcription factor discoveries, without separating by the direction of expression (**Table 13**). The top enriched TF (NES > 3.0) in this analysis was the interferon regulatory factor 4 (Irf4), important in to mediate inflammation (Al Mamun et al., 2018). The second enriched TF, Irf9, is also from the IRF family of transcription factors, and both **Tables 13** and **14** showed this gene between the top enriched TFs.

All the 4,530 genes under-expressed (FDR-adjusted P-value < 0.05) in the BCG relative to the Sal treatment, and all the 4,268 genes over-expressed (FDR-adjusted P-value < 0.05) in the BCG relative to the Sal treatment were used for co-regulation analysis. The results of the TFs identified for both under- and over-expressed are in **Table 14**. Among these enriched TFs (NES > 3.0), the ETS transcription factor (Elk1) is the top enriched TF for the under-expressed genes. The Elk1 is a nuclear effector of the ERK cascade, which plays a role in synaptic plasticity and memory (Thiels et al., 2002). Network analysis in another study that leads inflammation to major depressive disorder in twins revealed that both Pparg (top enriched TF for interaction effects section) and the

top enriched TF for over-expressed genes, MYC proto-oncogenes, BHLH transcription factor (Myc), are interacting through the Jun proto-oncogene, AP-1 transcription factor subunit (Jun), important in inflammatory responses and depression-like behavior (Malki et al., 2016). In the same study of network analysis, Elk1 also interacts with Jun (Malki et al., 2016).

Table 13. Transcription factors (TF) enriched (normalized enrichment score NES > 2.0) among differentially expressed genes (FDR adjusted P-value < 2.0E-6; number N of target genes = 4,455) genes in BCG-infected relative to Saline treated mice.

Transcription factor	**NES**[a]	**Target Gene N**[b]
Irf4	5.232	2186
Irf9	5.230	1058
Sirt6	3.656	664
Ikzf1	3.383	965
Snapc4	3.333	461
Zfp503	3.317	184
Nf1	3.238	457
Gata2	3.238	1719
Pura	2.932	1093
Rrn3	2.888	1578
Ywhaz	2.882	260
Srf	2.824	237
Klf4	2.698	2584
Gm4204	2.677	950
Bcl3	2.646	338
Stat6	2.602	1131
Elk1	2.560	806
Cebpe	2.559	300
Ets1	2.557	605
Hes5	2.543	1334
Ubp1	2.514	1624
Bach2	2.504	1107
Stat1	2.503	599
Ar	2.494	245
Foxo1	2.444	48
Atf5	2.443	553
Runx1	2.437	1582

[a]Normalized enrichment score

[b]Number N of target genes

Table 14. Transcription factors (TF) enriched (normalized enrichment score NES > 3.0) among under-expressed (number N of target genes = 4,530) and over-expressed (number N of target genes = 4,268) genes in BCG-infected relative to Saline treated mice.

Under-expressed genes			**Over-expressed genes**		
TF[a]	NES[b]	Target Gene N[c]	TF	NES	Target Gene N
Elk1	5.530	3031	Myc	3.502	1048
Zfp143	4.308	1513	Smad1	3.416	94
Yy1	4.118	1371	Tef	3.314	1147
Sp3	3.766	2394			
Irf9	3.657	1152			
Rfx2	3.631	1789			
Mlx	3.019	238			

[a]Transcription Factor

[b]Normalized enrichment score

[c]Number N of target genes

Network analysis were performed on more stringent number of differentially expressed genes to identify the transcription factors that are influencing the BCG infection effects. Among the 4,530 genes under-expressed (FDR-adjusted P-value < 0.05) in the BCG relative to the Sal treatment, 150 genes with log2 (fold change) < -3 were selected for co-regulation analysis. Among the 4,268 genes over-expressed in the BCG relative to the Sal treatment, 105 genes with log2 (fold change) > 2.5 were selected for co-regulation analysis. **Figures 9** and **10** show the regulatory networks for TF exhibiting a normalized enrichment score > 3.0, the top enriched TFs (> 30 targets) with higher number of targets were selected, however, to facilitate visualization, they were filtered to 30 targets for each TF in the network. **Table 15** lists the enriched TFs exhibiting a normalized enrichment score > 3.0 among genes under-expressed and over-expressed (FDR-adjusted P-value < 0.05) in the BCG relative to the Sal treatment.

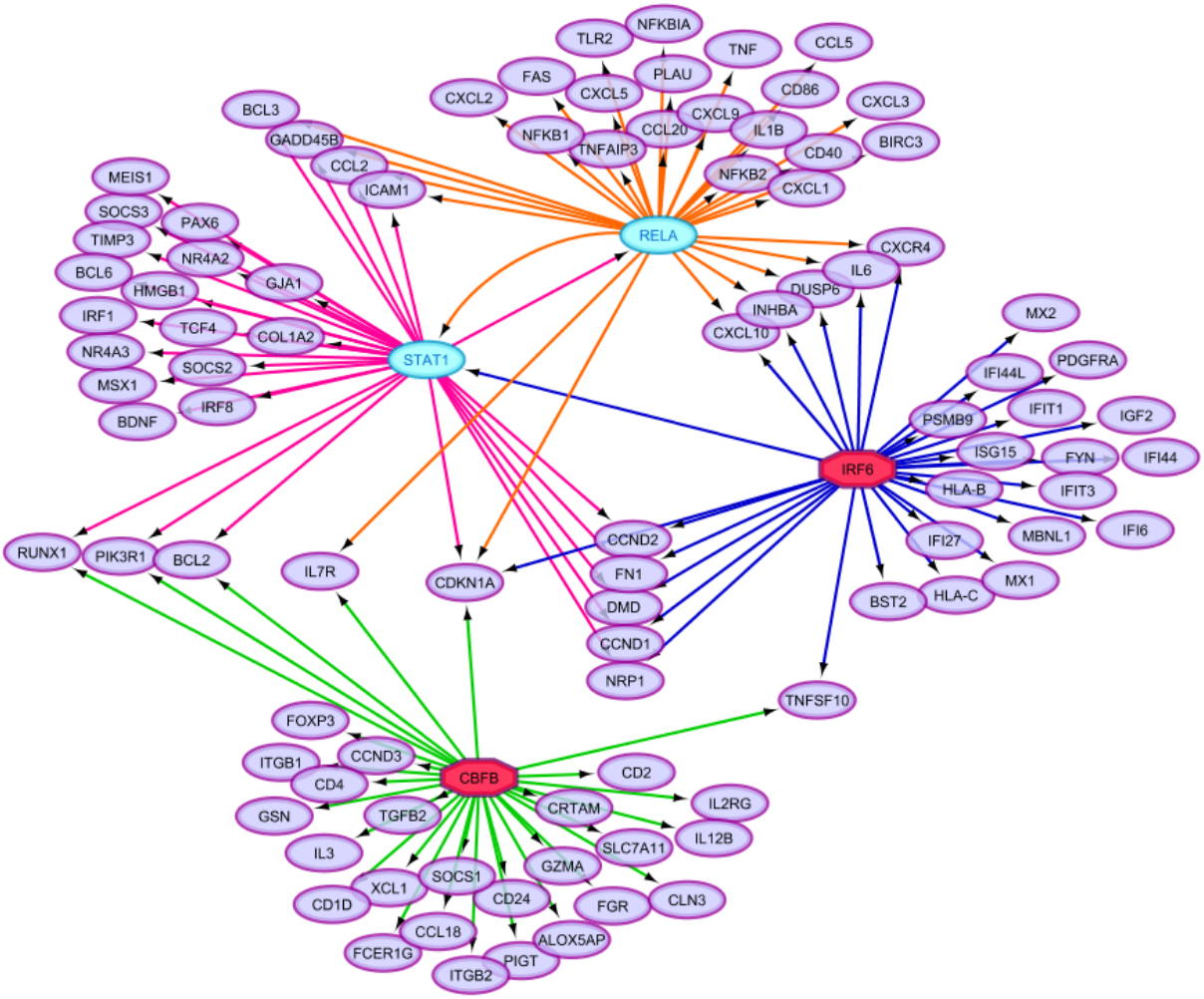

Figure 9. Regulatory network of transcription factors (normalized enrichment score > 3.0 and > 30 targets) and the 30 target genes under-expressed in BCG relative to Saline-treated mice. Red octagons denote TFs. The blue ovals are TFs that are targeting each other. The purple ovals are targets.

The co-regulatory network results are in agreement and complement the results discussed in the other sections. The regulatory network for under-expressed target genes (**Figure 9**) includes signal transducer and activator of transcription 1 (Stat1; normalized enrichment score = 5.9), Rela proto-oncogene, NF-kB Subunit (Rela; normalized enrichment score = 4.9), core-binding factor beta subunit (Cbfb; normalized enrichment score =3.7), and interferon regulatory factor 6 (Irf6; normalized enrichment score = 3.6) as the predicted regulons (**Table 15**).

Stat1 is from STAT protein family and can be activated by interferons (IFNs) and cytokines. Specifically, IFN-a, which is associated with depression symptoms, can lead to the phosphorylation of Stat1 and Stat2 (Borsini et al., 2017). Both IFN-alpha and Stat1-mediated production of cytokine is involved in cell apoptosis process in the brain and neurogenesis (Borsini et al., 2017). The presence of high levels of Stat1 was associated with severe major depressive disorder and dysregulation of neurotransmitter systems (Hoyo-Becerra et al., 2015). Stat1 was a hub gene connected with Janus kinase 2 (Jak2), signal transducer, an activator of transcription 3 (Stat3), and nuclear factor erythroid derived2 (Nfe2) in network analysis, being over-expressed in the BCG-infected compared to Sal treatment group within macrophages and microglia (Gonzalez-Pena et al., 2016a). Stat1 is regulating a gene that serves as a neurotransmitter, brain-derived neurotrophic factor (Bdnf), which is considerate a marker for depression treatment (Martocchia et al., 2014). Stat1 is being regulated by the predicted regulons Irf6 and is also both regulating and being regulated by Rela.

The gene Rela is a subunit of the canonical pathway NF-kB, which is essential for B cell maturation and activation (Milanovic et al., 2017). Rela is an abundant form in NF-kB, important to play a role in the innate and adaptive immune system, inflammation and cancer (Martincuks et al., 2017). Rela and Stat1 regulate five genes in common: B-cell CLL/lymphoma 3 (Bcl3), growth arrest and DNA damage-inducible beta (Gadd45b), C-C motif chemokine ligand 2 (Ccl2), intercellular adhesion molecule 1 (Icam1), and cyclin-dependent kinase inhibitor 1a (Cdkn1a). Cdkn1a is regulated by the four predicted regulons or TFs.

As the Irf4 and Irf 9 found as enriched TFs in **Tables 13** and **14**, the Irf6 is also from the interferon transcription factor (IRF) family, which primarily regulate immune system maturation, playing a role in the expression of pro- and anti-inflammatory cytokines (Li et al., 2017a). Irf6 plays a role in the macrophage activation, and is related to metabolic disorders and indirectly mediate TLR signaling (Li et al., 2017a). In the network, Irf6 is regulating many genes related to immune system like Il6 (**Figure 9**).

The TF Cbfb regulates genes involved in hematopoiesis and osteogenesis and is constituted of genes from RUNX family (Morita et al., 2017). The inhibition of RUNX has anti-tumor potential, implicating in cell death pathway (Morita et al., 2017). Cbfb regulates the target suppressor of cytokine signaling 1 (Socs1), that is a negative regulator of Tlr4 signaling, and a negative regulator of Trl4 has been associated with depression (Hung et al., 2017). The four TFs

enriched among under-expressed genes between mice infected with BCG relative to Sal treatments implicated in the finding of candidate regulons that can be potentially used as biomarkers to understand the molecular mechanisms involving inflammation and depression-like behavior.

Four enriched TFs were selected and had > 30 target genes among the genes over-expressed in the BCG-infected relative to the Sal treatment group (**Figure 10**): forkhead box g1 (Foxg1; normalized enrichment score = 3.8), forkhead box l1 (Foxl1; normalized enrichment score = 3.7), upstream transcription factor 1 (Usf1; normalized enrichment score = 3.6), and TEA domain transcription factor 1 (Tead1; normalized enrichment score = 3.1) (**Table 15**).

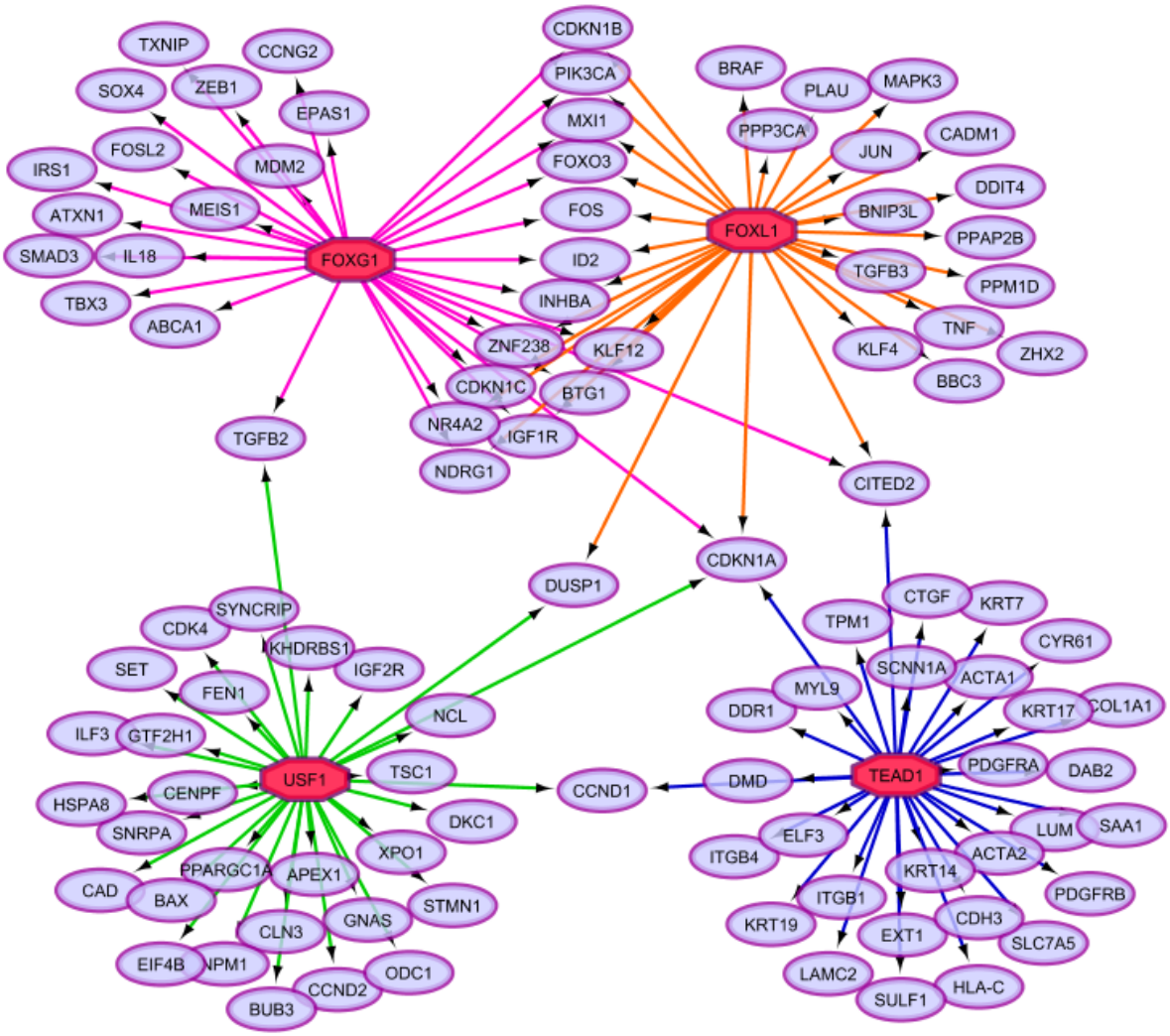

Figure 10. Regulatory network of transcription factors (normalized enrichment score > 3.0 and > 30 targets) and the 30 target genes over-expressed in BCG-infected relative to Saline treated mice. Red octagons denote TFs. The purple ovals are targets.

Foxg1 plays a role in the brain development and may lead to human neurological disorders (Kumamoto and Hanashima, 2017). This TF is a candidate gene for two neurological disorders, West and Rett syndromes, and in epilepsy, suggesting the participation of Foxg1 in the interneuron development (Yang et al., 2017). Also, Foxg1 regulates SMAD/transforming growth factor beta (TGF-B) pathway that leads to antidepressants mechanisms of action (Kinsler et al., 2010). Deficiency of Foxg1 may reflect in differences in behavior, such as hyperactivity, and can lead to impairment of antidepressant response (Kinsler et al., 2010). In agreement with these findings, Foxg1 is connected to the SMAD family member 3 (Smad3) suggesting a possible neurological effect during inflammation in the BCG-infected relative to Sal treatment group (**Figure 10**).

Although Foxl1 also belongs to the FOX family, this gene plays a different role in the immune system (Zhong et al., 2017). An over-expression of Foxl1 is related to inhibition of cell proliferation and migration, downregulation of Wnt/B-catenin signaling pathway and can also promote TNF-related apoptosis-inducing ligand (TRAIL) (Li et al., 2016; Zhong et al., 2017). In **Figure 10**, both Foxg1 and Foxl1 are regulating fos proto-oncogene, ap-1 transcription factor subunit (Fos or c-fos), which is related to antidepressant treatment, high neuronal activity and behavioral disorders (Fan et al., 2017; Zhu et al., 2017). Foxg1 and Foxl1 regulate fos and more fifteen genes in common, and from these, one is regulated by all four regulons (Cdkn1a), and one is regulated by three regulons (Cbp/P300 interacting transactivator with Glu/Asp-rich carboxy-terminal domain 2 (Cited2)). Cdkn1a is a cell cycle regulator, related to clock genes (Sato et al., 2017). Cited2 is a repressor of macrophage pro-inflammatory activation (Kim et al., 2017).

Usf1 also is a TF that binds the gene Bdnf promoter IV in the brain of rats (Cortes-Mendoza et al., 2013; Ren et al., 2016; Spohrer et al., 2017). Usf1 is regulating the gene Cad, mutations in this gene were involved with epileptic encephalopathy (Koch et al., 2017). Tead1 belongs to TEAD transcription factor family and plays an important role in development and Hippo signaling pathway, which regulates cell growth, proliferation, and homeostasis (Lin et al., 2017). In **Figure 10**, Tead1 is regulating Saa1, a gene that was discussed in the interaction section and is related to depressive-like behavior in brain (Jang et al., 2017). The study of regulons with regulatory networks confirmed the findings discussed in other sections and implicated in the identification of transcription factors that can be potential candidates to understand the mechanisms involved in the BCG-infection leading to inflammation and depression-like behavior.

Table 15. Transcription factors (TF) enriched (normalized enrichment score NES > 3.0) among under-expressed (number N of target genes = 150) and over-expressed (number N of target genes = 105) genes in BCG-infected relative to Saline treated mice.

Under-expressed genes			**Over-expressed genes**		
TF[a]	NES[b]	Target Gene N[c]	TF	NES	Target Gene N
Stat1	5.982	94	Zbtb18	4.111	29
Rela	4.936	46	Foxg1	3.795	30
Cbfb	4.344	39	Foxl1	3.728	42
Ywhae	4.146	25	Foxn4	3.613	40
Ikzf1	3.999	18	Pou3f4	3.596	25
Tppp	3.817	15	Usf1	3.562	35
Bcl6	3.676	14	Sox10	3.426	17
Irf6	3.573	56	Ing4	3.348	14
Gata5	3.443	16	Kdm4d	3.331	17
Jun	3.341	31	Tead1	3.135	49
Mettl3	3.307	15	Stat1	3.123	30
Ets1	3.033	12	Jun	3.064	19
Trp53	3.004	8	Smad4	3.026	17

[a]Transcription Factor

[b]Normalized enrichment score

[c]Number N of target genes

According to **Table 15**, the gene Jun proto-oncogene, AP-1 transcription factor subunit (Jun) was also among the top enriched TF for under-expressed genes in BCG-infected relative to Saline-treated mice and had > 30 targets. Jun, as discussed at the beginning of this section, was associated with major depressive disorder, neuroplasticity and neurogenesis in a rat model of depression in the prefrontal cortex (Malki et al., 2015). Among the top TF enriched over-expressed genes, the forkhead box N4 (Foxn4) and Stat1 also had > 30 targets. Foxn4 positively regulates the orphan nuclear receptor subfamily 4 group A member 2 (Nr4a2), which is essential for the ventral midbrain dopaminergic neurons and is related to Parkinson's disease (Jiang and Xiang, 2009). Despite zinc finger and BTB domain containing 18 (Zbtb18) was not in the top enriched TFs, this gene had 29 targets and had the highest score for NES among the TFs for over-expressed genes in BCG-infected relative to Saline-treated mice. Zbtb18 is associated with microdeletion syndrome, characterized by intellectual disability (De Munnik et al., 2014). Both Jun and Stat1

might be important TFs in this study for being enriched for under-expressed and over-expressed genes in BCG-infected relative to Saline-treated mice.

Comparison between results of two different pipelines

The analysis of the contrast BCG(WT)-Sal(WT) was performed using both Tuxedo pipeline (TopHat, Cufflinks, Cuffdiff) and Kallisto, tximport and edgeR to better identify how different these two methodologies are. The number of significant differentially expressed genes (FDR adjusted P-value <0.05) identified with Cuffdiff was 4,137, while edgeR identified 6,843. The edgeR method had 2,681 genes that were not present in Cuffdiff analysis, and Cuffdiff had 995 genes not present in edgeR analysis. The values of P-value, FDR-adjusted P-values, and fold changes from Cuffdiff were compared with edgeR using 11,029 genes that were common between both analyses. Both initial files were filtered by non-annotated genes, low or no expressed genes. Then, data from both analyses were combined by gene Id. P-value, FDR-adjusted P-values, and fold changes were logged 10 transformed to normalize data. However, the data were still not normally distributed, and Spearman method was used to calculate correlation in the SAS software (Saxton and Institute 2004).

The Spearman method results showed 87% of correlation (P-value < 0.0001) between adjusted P-values, 98% of correlation (P-value < 0.0001) between log10 (fold changes), and also 87% of correlation (P-value < 0.0001) between P-values for both analyses. The results showed that both pipelines, having Cuffdiff and edgeR for differential expression analysis, share a high number of significant genes, with a similar level of P-value, FDR-adjusted P-values, and fold changes. Despite a lower number of differentially expressed genes in Cuffdiff, both methods showed a strong correlation in the results.

In summary, our study of individual gene expression patterns, functional analysis, and TF confirmed that genes associated with immune pathways remain dysregulated seven days after BCG treatment. The association between macrophage transcriptome and depression-like symptoms may support reports that immune cells from the peripheral blood can surpass the blood-brain barrier and elicit central nervous system inflammation. The detection of genes that exhibited interaction effect suggests that the impact of BCG challenge after recovery in the macrophages is different

between IDO1-KO and WT mice. Several genes and pathways associated with depression and neurological disorders that exhibit depression symptoms further aid in understanding the depression-like symptoms that mice present even after sickness recovery from BCG challenge.

CHAPTER 4 - GENE CO-EXPRESSION NETWORK ANALYSIS TO FIND MODULAR ORGANIZATION OF COMPLEX TRAITS

Abstract

Systems biology approaches are used to elucidate mechanisms that orchestrate cellular functions and regulate complex traits. Gene co-expression networks are useful for interpretation of large datasets. Networks modules can reveal functional processes and pathways that are shared between genes. Differential co-expression analysis compares the structure of two co-expression networks between conditions, in this case, providing insight into alterations in the regulatory systems between genotype lines and between drug treatments. Weighted gene co-expression network analysis identified modules related to the selected line (H) for hyperactivity and drug treatment of amphetamine (A). Co-expressed genes in module grey were significantly correlated with genotype, and functional analysis enriched the GO BP terms regulation of cell death and Wnt signaling pathway. Among the members of module grey, the genes frizzled class receptor 9 (Fzd9), insulin-like growth factor (Igf1), and tachykinin receptor 1 (Tacr1) are related to behavior, learning, and memory, and genes peptidylprolyl isomerase F (cyclophilin F) (Ppif) and cytochrome c oxidase subunit IV isoform 2 (Cox4i2) is associated with neurodegenerative disorders. Genes in modules tan, cyan, and salmon were significantly correlated with treatment. The tan module had GO BP terms related to nervous system development, while cyan module had regulation of nitrogen compound metabolic process and salmon module had regulation of multicellular organismal process as significant. Among genes in the salmon module, the gene Jun proto-oncogene (Jun) is specifically related with an amphetamine addiction, and the genes neurogenic differentiation 1 (Neurod1), purine-nucleoside phosphorylase (Pnp), and xanthine dehydrogenase (Xdh) are related with response to the drug, drug binding, and drug metabolism, respectively. In the cyan module, the gene cAMP responsive element binding protein 3-like 1 (Creb3l1) is related to the dopaminergic synapse, Huntington's disease, and amphetamine addiction. From the tan module, the gene oxytocin (Oxt) plays a role in response to amphetamine and behavior, while the gene kallikrein related-peptidase 8 (Klk8) is important in the nervous system development, behavior, learning or memory. Four co-expressed genes found in the significant modules, Creb3l1, Neurod1, Ppif and Oxt also exhibited significant (FDR adjusted P-value < 0.05) treatment-by-line interaction effect in Chapter 2. Differentially co-expressed results identified EPH receptor B3

(Ephb3), important for the nervous system, as negatively differentially connected to mice receiving the A relative to the S treatment. The gene Ppif was positively differentially connected to mice from C relative to the H line and is possibly one of the most important genes of this analysis for being differentially expressed, significantly co-expressed in the grey module, and differentially co-expressed.

Introduction

Gene expression analysis helped to understand the molecular basis of the simultaneous effects of genotype, using mice from the line (H) selected for hyperactivity, a model of attention deficit hyperactivity disorder (ADHD), and of amphetamine treatment (A). Systems biology is used in this study to understand better the mechanisms underlying the hyperactivity line-amphetamine treatment.

The systems biology methods of correlation network analysis and identification of hub genes that contribute to gene patterns across the samples have been widely applied to diseases like heart disease, and diabetes (Sengupta et al., 2009; Tang et al., 2018). The networks can be constructed based on the data from gene expression analysis. Gene co-expression networks have been used to detect modules that are biologically meaningful and to find the relationship between modules. These modules are composed of highly interconnected genes organized according to functional processes and pathways. Weighted Gene Co-expression Network Analysis (WGCNA) is a software package in the R environment that uses gene co-expression information to uncover networks (Langfelder and Horvath, 2007). WGCNA routines were used to identify gene modules related correlated to the hyperactivity line or amphetamine treatment.

Another approach known as differential co-expression analysis compares the structure of two co-expression networks and connectivity between conditions, providing insight into alterations in the regulatory systems (de la Fuente, 2010). This method has been successfully used to identify differential co-expression in cancer and to identify transcription factors and microRNAs related with phenotypes (Choi et al., 2005).

The objective of this study was to identify modules and genes associated with the selected line (H) for hyperactivity and drug treatment of amphetamine (A) using WGCNA and differential co-expression analysis. Another goal was to compare the results of these analyses with the results

of differential gene expression found in Chapter 2 and complement our insights, as these are powerful alternatives to find more answers about complex systems.

Materials and Methods

Data collection and sequencing

Mice from a hyperactivity line (H) selected for increased home cage activity and from a contemporary control line (C) were studied (Majdak et al., 2014; Majdak et al., 2016; Zombeck et al., 2011). Home cage activity was individually measured, and male H mice exhibited significantly higher activity and motor impulsivity compared to C mice. Room conditions of temperature, light, diet, and management followed previous protocols (Majdak et al., 2016). At approximately five months of age, one intraperitoneal injection of either 0.25 mg/kg of d-amphetamine (treatment A) or equal amount of saline (treatment S) was given to each mouse in the home cage over three days (Majdak et al., 2016). Two hours after injection, mice were decapitated, and brains were removed immediately. In total, 20 mice were analyzed across the 2 activity lines and the two amphetamine treatments. All animal procedures were approved by the Illinois Institutional Animal Care and Use Committee and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Striatum was extracted and dissected and stored in a centrifuge tube at -80^{o}C following published protocols (Caetano-Anollés et al., 2016; Saul et al., 2017). RNA was extracted, isolated, and purified. Total RNA yield measured using Qubit1 2.0 (Life Technologies, Carlsbad, CA, USA) was > 14 ug per sample. The RNA Integrity Number (RIN) measured using Fragment Analyzer (Advanced Analytical Technologies Inc., Ankeny, IA, USA) was > 8 for all samples. The libraries were prepared and sequenced using Illumina HiSeq 2500. FASTQ files containing paired-end reads of length 100nt were generated and demultiplexed with the software Casava 1.8.2 (Illumina, San Diego, CA, USA).

The software program FastQC was used to assess Phred quality score of the reads (Andrews, 2010). The nucleotide quality score was > 30 across all read positions and was not trimmed (Nixon et al., 2015). The software program TopHat2 v2.1.1 was used to map the reads to the mouse genome assembly Genome Reference Consortium GRCm38 (Pruitt et al., 2006; Trapnell et al., 2009). Transcripts were assembled and abundance estimated using the software

program Cufflinks2 (Trapnell et al., 2012) with default specifications, generating fragment per kilobase per million mapped reads (FPKM) for each sample.

Weighted Gene Co-expression Network Analysis

The WGCNA package in R was used for the analysis of co-expression networks (Zhang and Horvath, 2005). The input data was the table of FPKM values generated by Cufflinks2 per annotated gene and sample. The expression on 44,796 genes and "CUFFs" (short transcripts assembled by Cufflinks or transfrags in GTF format, normally not annotated) across 20 samples genes were available. This information was filtered to remove genes with low expression level or variation across samples as per recommendations in the software manual. Genes that had mean expression < 0.1 and standard deviation < 0.3 across samples were excluded as well all "CUFFs". This filtering reduced the number of genes available for network analysis to 2,193 genes. The purpose was to eliminate low-expressed or non-varying genes that are considerate noise in WGCNA analysis (Zhang and Horvath, 2005).

After filtering, the FPKM values were log transformed (log2 (FPKM +1)). Although normalization does not make a lot of difference for WGCNA analysis, the log-transformed FPKM values were also normalized by mean and standard deviation across all samples (van Dam et al., 2017). The genes that remained on the list were evaluated according to their connectivity. The Pearson correlation coefficient was used as a measure for co-expression, and a threshold was given to this correlation to find networks that influence the trait. Genes that had significant pairwise expression profiles associated with samples/phenotype were considered connected (Zhang and Horvath, 2005).

In a network, nodes are considerate biomolecular species as metabolites, proteins, and genes, while the edges represent the interaction between nodes (de la Fuente, 2010). The probability that a node is connected with another "k" node decays following the power law, indicating that large networks are organized into scale-free networks (Zhang and Horvath, 2005). The power-law distribution is also called scale-free topology and can be explained using a network growth model. In other words, the free-scale topology depends on the growth of the network; new nodes are preferentially connected to already existing nodes (Albert et al., 2000). Highly connected

nodes are called hubs, and scale-free networks are dominated by heterogeneous topology detecting few hubs that are more essential in the system (Zhang and Horvath, 2005).

Next step was choosing a set of soft-thresholding power by calling the network topology analysis function (Zhang and Horvath, 2005). A soft threshold determines that the network will be weighted by assigning a connection weight to each gene pairs (Zhang and Horvath, 2005). Constructing a weighted gene network entails the choice of the soft thresholding power β to which co-expression similarity is raised to calculate adjacency. A measure of similarity between the gene expression profiles was defined (Zhang and Horvath, 2005). In this study, the absolute value of the Pearson correlation sij = |cor(i, j)| was used. Each co-expression network corresponds to an adjacency matrix that encodes the connection strength between each pair of nodes (Zhang and Horvath, 2005). The resulting adjacency matrix was used to define a measure of node dissimilarity (distance). The node dissimilarity measure was used as input of a clustering method to define network modules (clusters of nodes) (Zhang and Horvath, 2005). An important aim of co-expression network analysis is to detect subsets of nodes (modules) that are tightly connected to each other. The use of 'soft' adjacency functions is to avoid the disadvantages of hard thresholding (Zhang and Horvath, 2005).

Connectivity, clustering coefficient, and topological overlap were calculated for the networks (Zhang and Horvath, 2005). The connectivity or row sum is the sum of the correlations between a gene and all the other genes of the network. The adjacency matrix encodes the connection strengths between pairs of nodes or genes, and this matrix is generated by the Pearson correlation between all genes elevated to the beta parameter, which is chosen based on topology scale-free index (R^2) (Zhang and Horvath, 2005).

The groups of genes that have the same degree of overlap is determined by counting the number of neighbors (first-order interactions) that a pair of nodes shares (Zhang and Horvath, 2005). The topological overlap matrix (TOM) is calculated from this, and a value between 0 and 1 (connected =1, all neighbors of a node are also neighbors of the other node; unconnected=0, two nodes do not share any neighbors) is attributed to each pair of nodes (Zhang and Horvath, 2005). The TOM matrix was used as input for an average linkage hierarchical clustering step based on Euclidean distances method that results in a dendrogram or tree. A height cutoff value was chosen in the dendrogram to decide where to cut the tree branches; the cutoff choice is guided by the TOM plot that is generated and results in the number of modules. Large height values result in big

modules, and small values result in small tight modules. The resulting branches correspond to gene modules (Zhang and Horvath, 2005).

In WGCNA, co-expression modules are represented by eigengene networks that contain genes that are more representative of the gene expression profile instead of using all genes. A module eigengene is equivalent to a first principal component of a given module, and usually explains more than 50% of the variance of the module expressions (Langfelder and Horvath, 2007).

The association of individual genes with the traits (genotype and treatment) was made by Gene Significance (GS), which is the absolute value of correlation between the gene and the trait. The correlation of the module eigengene and the gene expression profile was also quantified, this measure is called module membership (MM) (Zhang and Horvath, 2005).

In this study two traits were considered for final interpretation: the line H versus line C, and treatment A versus treatment S. Statistically significant modules selected and the genes present in these modules were studied for functional enrichment using a hypergeometric test, the web-service Database for Annotation, Visualization and Integrated Discovery (DAVID; Version 6.8) (Huang et al., 2009). The functional categories Gene Ontology (GO) biological processes (BP), molecular functions (MF), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were assessed. The geometric mean of the enrichment P-values (Enrichment Score, ES) was used as evidence supporting the category clusters. The mouse genome was used as background for testing and FDR was used to adjust the enrichment P-values for multiple testing (Pruitt et al., 2006).

Differential co-expression analysis

In Chapter 2, the differential expression between two traits (line and treatment) and the interaction between them were discussed. In differential expression analysis, genes were significantly different for both line (H and C) and treatment (A and S) based on the observation of the mean expression level of genes. A new approach called differential co-expression was applied in this study to find genes significantly different for both line and treatment based on the observation of the correlation between the expression levels of two or more genes (de la Fuente, 2010).

The difference of connectivity between two or more genes in two different conditions, (between H and C and between A and S) can play an important role in the phenotypes (line and treatment). The difference from co-expression analysis is that the networks will be constructed from two different subgroups of samples (two conditions) from a prior WGCNA. In total, 18 samples and 2,193 genes were used in WGCNA analysis, and for differential co-expression analysis the number of samples and genes were the same (5 Control-Amphetamine (CA), 4 Control-Saline (CS), 5 Hyperactive-Amphetamine (HA), 4 Hyperactive-Saline (HS)). The Pearson correlation was used to quantify the association between two gene expression levels. The calculation of correlation of the expression between a pair of genes was made separately for the fours subgroups: H, C A and S, and the connectivity were compared between lines and between treatments separately.

The same step used in WGCNA was repeated for each subgroup. An adjacency matrix that measures the connection strengths was constructed, and the power for each subgroup was selected based on the free topology criterion (de la Fuente, 2010). Then, the connectivity was calculated as a function of the soft-thresholding power.

After finding the connectivity of each subgroup, the connectivity estimates were normalized to facilitate the comparison between subgroups (range of values for normalized connectivity was 0-1). Normalization encompasses the dividing each gene connectivity by the maximum connectivity of that subgroup (Fuller et al., 2007). The measure to calculate the differential connectivity between subgroups is $DiffK(i) = K_1(i) - K_2(i)$. The results are the differences in the connectivity between two subgroups, and the differences are considerate significant if *DiffK* is $< |0.4|$ (Fuller et al., 2007). Negative values of this difference imply that line C or treatment S connectivity (K_2) is greater than line H or treatment A (K_1) connectivity. The differentially co-expressed genes were analyzed for functional enrichment using DAVID.

Results and Discussion

Weighted Gene Co-expression Network Analysis

Filtering of genes. A CVS file including columns corresponding to the 20 samples from the four genotype-treatment groups: CA, CS, HA, HS. This file contained a total of 44,796 rows corresponding to genes and entries denoted CUFFS that had non-zero FPKM in at least one group.

All entries denoted CUFFS were deleted because of unknown annotation leaving 4,530 genes. Another filter was applied per gene row. Genes that had no expression value in at least ten samples were removed leaving 3,472 genes for analysis. Further filtering for computational reasons encompassed removal of genes that had mean expression < 0.1 and standard deviation < 0.3 were removed, leaving 2,193 genes for WGCNA analysis. From these 2,193 genes, 333 exhibited significant interaction effect, 54 were differentially expressed in the main effect of A treatment, and 194 genes were differentially expressed in the main effect of H genotype at FDR adjusted P-value < 0.05.

Filtering and clustering of samples. **Figure 11** shows the clustering of 20 samples based on the log-transformed FPKM of 2,193 genes after pre-processing with the objective of outlier detection. Two samples (CS411 and HS499) were creating a separated branch from the main one, and for this reason, these two samples were eliminated of the analysis as they might be confounding factors in the attempt to create one big simple cluster. **Figure 12** presents the cluster of the remaining 18 samples based on 2,193 genes used for WGCNA analysis.

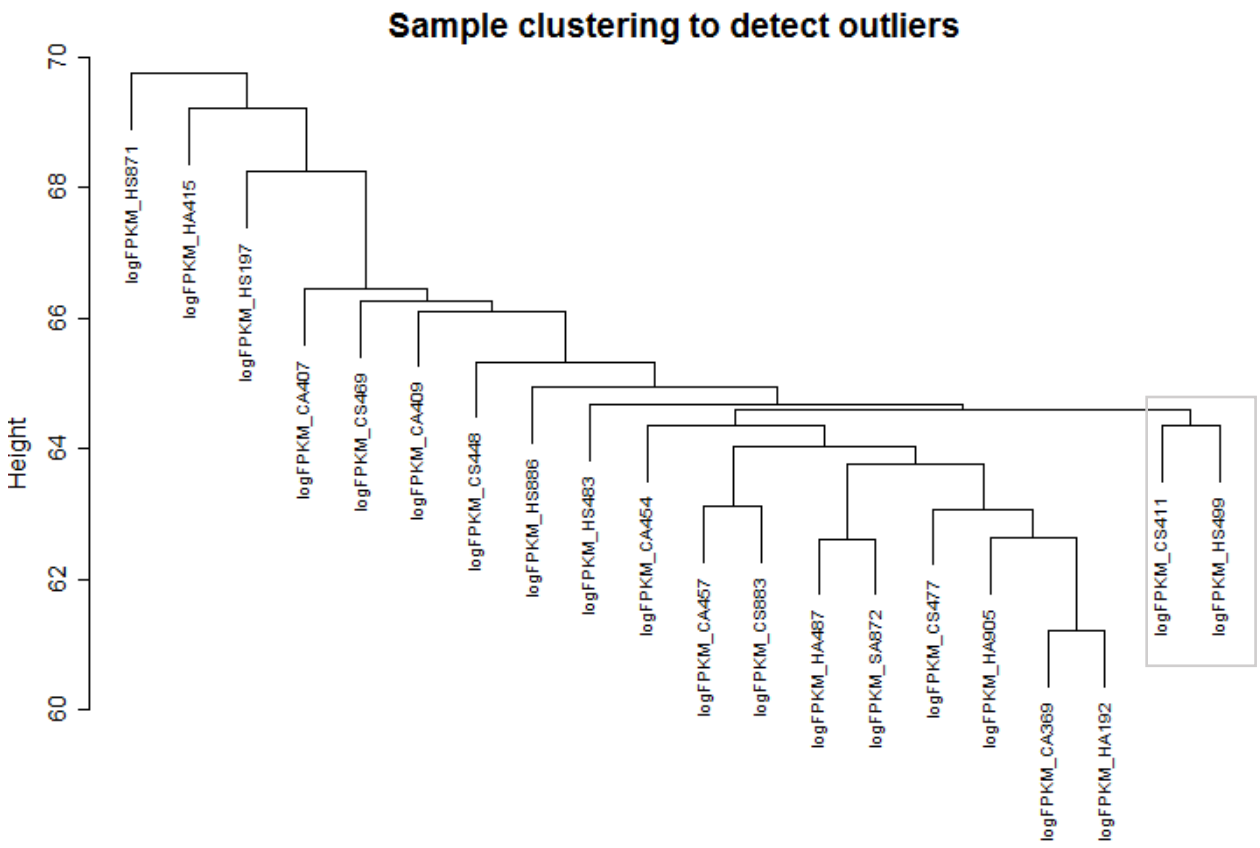

Figure 11: Clustering dendrogram of 20 genotypes by treatment samples (5 Control-Amphetamine (CA), 5 Control-Saline (CS), 5 Hyperactive-Amphetamine (HA), and 5 Hyperactive-Saline (HS)) based on their Euclidean distance using 2,193 genes. Height corresponds to the cut of the tree. The leaves of the tree correspond to the samples. LogFPKM means log transformed FPKMs.

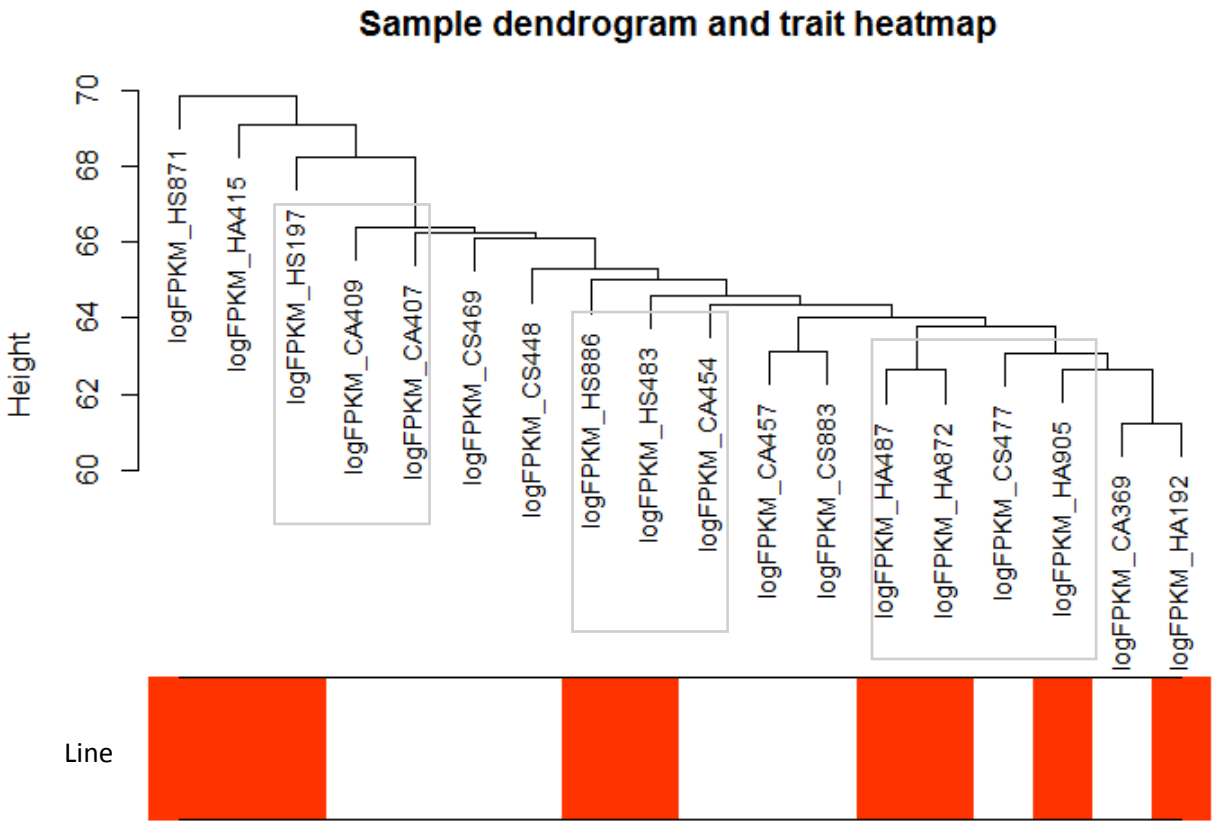

Figure 12: Clustering dendrogram of 18 genotypes by treatment samples (5 Control-Amphetamine (CA), 4 Control-Saline (CS), 5 Hyperactive-Amphetamine (HA), and 4 Hyperactive-Saline (HS)) based on their Euclidean distance using 2,193 genes. The first color band represents line (red indicates line H and white indicates line C). Height corresponds to the cut of the tree. The leaves of the tree correspond to the samples. LogFPKM means log transformed FPKMs.

The clustering in **Figure 12** includes three grey boxes showing that HS samples tend to cluster close to the CA samples (samples HS197 with CA409 and CA407; samples HS886, HS483 and CA454). Also, the HA samples tend to cluster close to CS samples (samples HA487, HA872, HA905, and CS477). These trends could suggest that the amphetamine treatment used in this study affects the transcripts of control mice similar to the hyperactivity genotype (CA ~ HS). Also, the amphetamine treatment used in this study had opposite associations with the transcript level between both mice genotypes. The amphetamine treatment on the hyperactive line was associated with transcript levels similar to saline treatment on control mice (HA ~ CS). These results are consistent with the findings in Chapter 2.

Gene network analysis in WGCNA focuses on the clustering of genes based on their gene expression, rather than the clustering of samples. An approach similar was used to compare the gene expression pattern using a cluster of samples to prove that samples that clustered together had transcriptional patterns (Svahn et al., 2016).

The information in **Figure 13** was used to select a set of soft-thresholding power by calling the network topology analysis function. The selected power was 6, the lowest power for which the scale-free topology fit index curve flattens out upon reaching a high value (in this case, roughly 0.85).

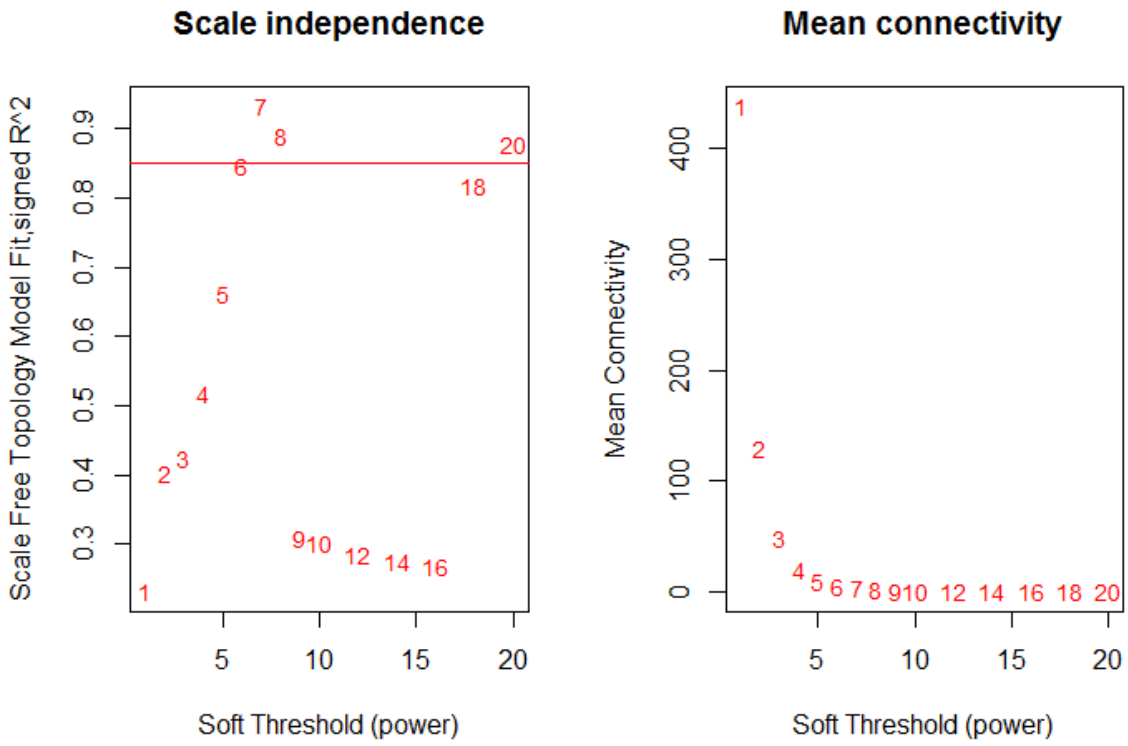

Figure 13: Analysis of network topology for various soft-thresholding powers. The left panel shows the scale-free fit index (y-axis) as a function of the soft-thresholding power (x-axis). The right panel displays the mean connectivity (degree, y-axis) as a function of the soft-thresholding power (x-axis).

The selection of the soft-thresholding power enabled the identification of gene modules. These modules encompass genes that exhibit highly correlated expression profiles across the samples or high topological overlap. Both **Figures 14** and 15 show a hierarchical clustering dendrogram of the eigengenes based on the dissimilarity. **Figure 14** shows the dendrogram of the

network application and provides a simple visual comparison of module assignments (branch cuttings) based on the dynamic hybrid branch cutting method. **Figure 15** is more used in the case that one wants to merge such modules since their genes are highly co-expressed.

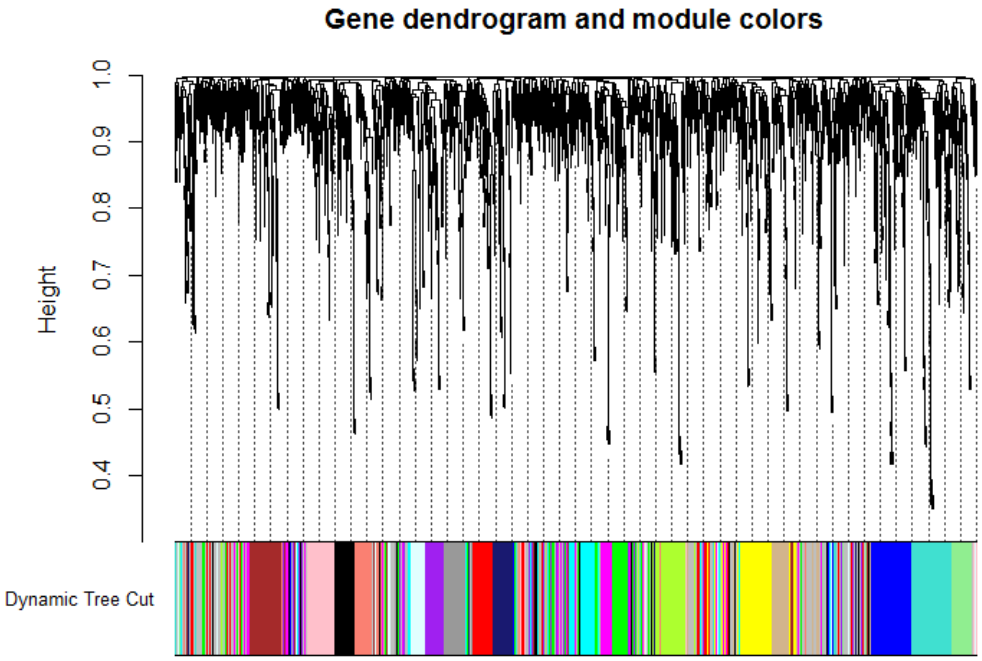

Figure 14: Clustering dendrogram of genes and gene modules identified using different colors. Height corresponds to the cut of the tree. Dynamic tree cut results from hierarchical linkage clustering. The color-band below the dendrogram denotes the modules, which are defined as branches in the dendrogram.

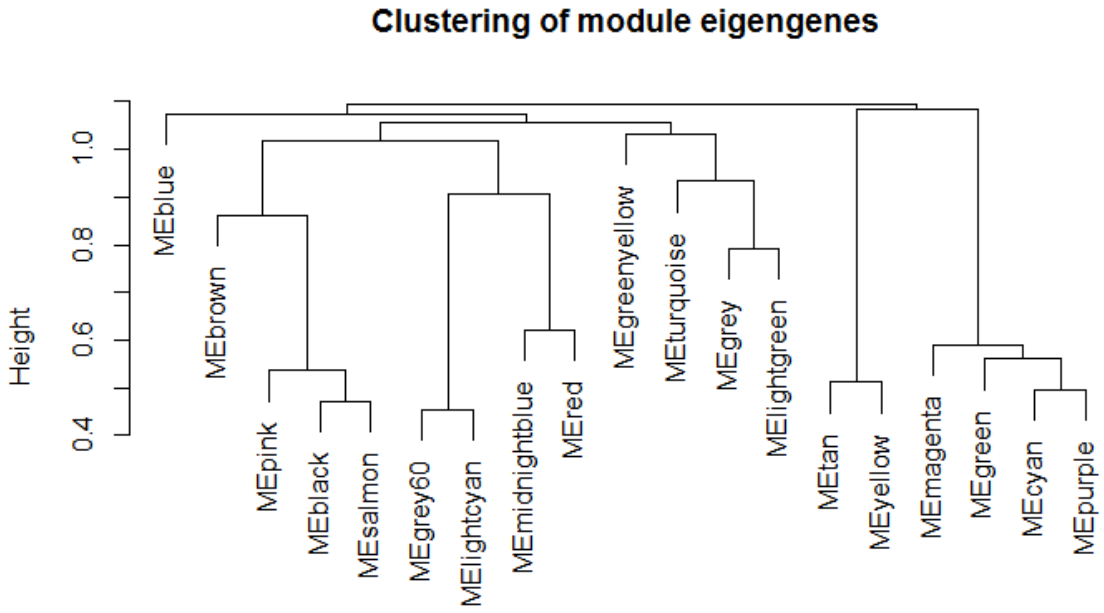

Figure 15. Clustering dendrogram of genes, with dissimilarity based on the topological overlap, with the original module colors. Height corresponds to the cut of the tree. ME before the name of colors means module eigengenes.

Table 16 lists the number of genes in each of 19 modules depicted in **Figure 15.** The list of the correlation estimates and P-values between the 19 module eigenvalues and the genotype and treatment indicators is presented in **Figure 16**. The C and H genotypes were coded 0 and 1, respectively and the S and A treatments were coded 0 and 1, respectively, to compute correlation. Thus, a -0.5 correlation between transcript eigenvalue for the cyan module and treatment indicates that higher levels of gene expression in this module are associated with the S relative to the A treatment. One correlation between module eigenvalues and genotypes (grey module) had a P-value < 0.1 and the estimate were -0.62 indicating that higher gene expression levels were observed in the C relative to the H genotype. Three modules (tan, cyan, and salmon) exhibited a significant (P-value $\leq$ 0.1) correlation between the eigenvalue transcripts. The tan and cyan modules had negative correlations whereas salmon had a positive correlation. The grey, tan, cyan and salmon modules included 247, 97, 90, and 95 genes respectively (**Table 16**).

Table 16. Number of genes in each module of **Figure 15**:

Dynamic Colors						
Black	Blue	Brown	Cyan	Green	Green Yellow	Grey
112	142	134	90	119	99	247
Grey60	Light Cyan	Light Green	Magenta	Midnight Blue	Pink	Purple
85	88	80	106	89	107	104
Red	Salmon	Tan	Turquoise	Yellow		
118	95	99	152	124		

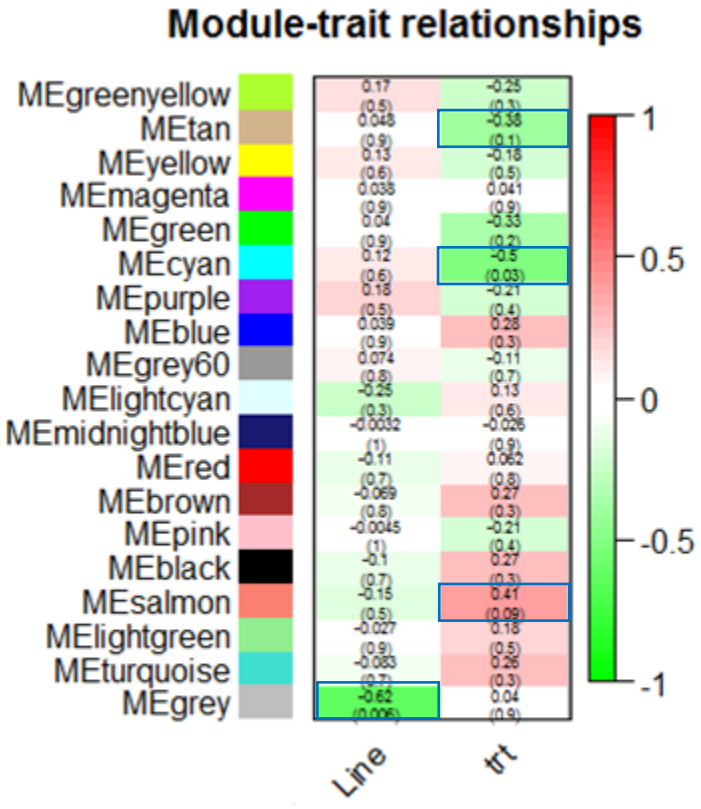

Figure 16. The matrix of module-trait correlations (first cell entry) and P-values (second cell entry in parenthesis) between traits (columns) and module eigenvalues (rows). The color scale on the right indicates the sign and strength of the correlation estimate from –1 (green) to 1 (red). In the figure, the significant modules are highlighted (P-value $\leq$ 0.1). "Line" means genotype and "trt" means treatment. ME before the name of colors means module eigengenes.

The plots in **Figure 17** depict the gene significance (GS) versus module membership (MM) for the modules that exhibited a significant (P-value $\leq$ 0.1) association between the eigengene and genotype (grey, **Figure 17A**) or treatment (tan, cyan and salmon, **Figure 17B**, **C**, and **D**, respectively). The GS and MM measures allow identifying genes that correlate with the traits as well as high module membership in interesting modules. Here, the grey module was the only significant module and presented moderate correlation with genotype. Despite significant, the modules salmon, cyan, and tan presented weak correlation with treatment.

A) B)

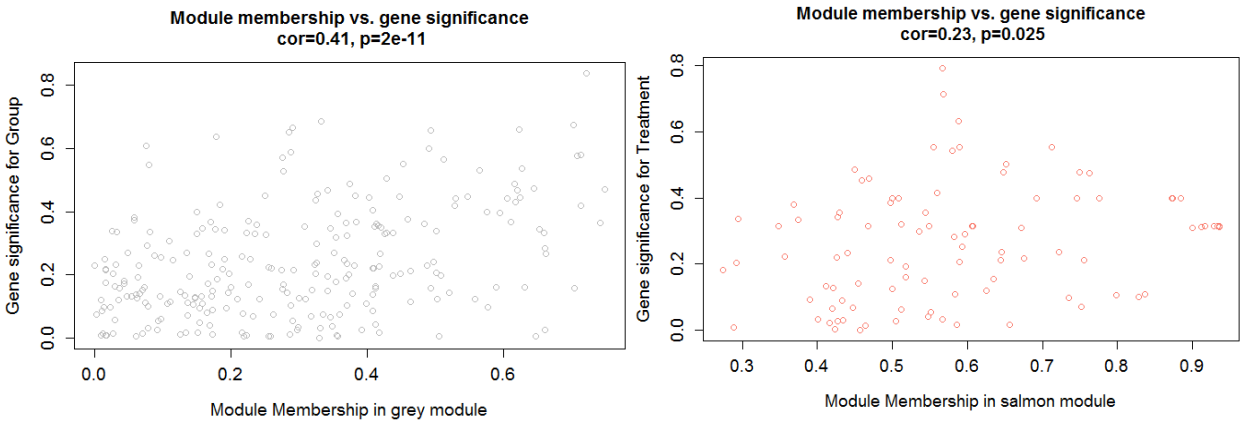

Figure 17. Gene significance (GS) versus module membership (MM) for four modules exhibiting significant (P-value $\leq$ 0.1) association between the eigenvalue expression and genotype or treatment. A) Grey module presented moderate correlation with genotype. B) Salmon module salmon presented weak correlation with treatment. C) Cyan module presented weak correlation with treatment. D) Tan module presented weak correlation with treatment. “Cor” means correlation and “p” means P-value.

Figure 17 (cont.)

C) D)

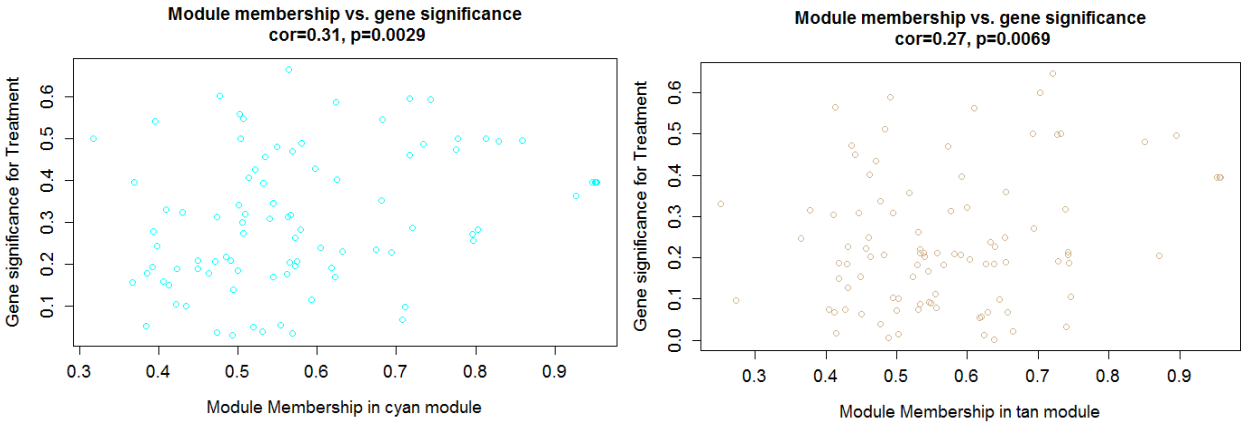

Figure 17 (continued). Gene significance (GS) versus module membership (MM) for four modules exhibiting significant (P-value $\leq$ 0.1) association between the eigenvalue expression and genotype or treatment. A) Grey module presented moderate correlation with genotype. B) Salmon module salmon presented weak correlation with treatment. C) Cyan module presented weak correlation with treatment. D) Tan module presented weak correlation with treatment. "Cor" means correlation and "p" means P-value.

Genes and enriched function with module associated with genotype or treatment

Table 17 summarizes the most significant clusters of enriched GO BP terms within the grey module that had eigenvalue expression significantly correlated with genotype, like regulation of cell death and Wnt signaling pathway. **Table 18** summarizes the most significant clusters of enriched GO BP terms within the tan, cyan and salmon modules that had eigenvalue expression significantly correlated with treatment. The tan module had GO BP terms related to nervous system development, while cyan module had regulation of nitrogen compound metabolic process and salmon module had regulation of multicellular organismal process as significant.

Table 17. Enriched clusters (FDR adjusted P-value <0.05) of Gene Ontology (GO) biological processes (BP) among the 247 genes in the grey module that had eigenvalue expression significantly associated with genotype.

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
Cluster 1	**Enrichment Score: 4.13**			
BP	GO:0010941~regulation of cell death	34	2.64E-05	0.049
Cluster 3	**Enrichment Score: 3.18**			
BP	GO:0030278~regulation of ossification	12	8.21E-06	0.015
Cluster 6	**Enrichment Score: 2.49**			
BP	GO:0016055~Wnt signaling pathway	15	2.50E-05	0.046
BP	GO:0198738~cell-cell signaling by wnt	15	2.57E-05	0.048

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

Table 18. Clusters of Gene Ontology (GO) biological processes (BP) among the 95 genes in the salmon module, among the 90 genes in the cyan module, and among the 99 genes in the tan module that had eigenvalue expression significantly associated with treatment.

Category	Term	Count[a]	Raw P-value	FDR-adjusted P-value[b]
Salmon				
Cluster 1	**Enrichment Score: 2.13**			
BP	GO:2000026~regulation of multicellular organismal development	16	0.0013	2.29
BP	GO:0051241~negative regulation of multicellular organismal process	12	0.0016	2.79
Cyan				
Cluster 1	**Enrichment Score: 1.137**			
BP	GO:0051171~regulation of nitrogen compound metabolic process	24	0.025	34.99
Tan				
Cluster 1	**Enrichment Score: 1.92**			
BP	GO:0051960~regulation of nervous system development	16	4.43E-06	0.01
BP	GO:0008285~negative regulation of cell proliferation	12	6.81E-05	0.12
BP	GO:0007399~nervous system development	23	7.76E-05	0.14

[a]Number of genes in the enriched category.

[b]False Discovery rate adjusted P-value.

The top genes of each module were selected based on their function that may be important for the trait, and are presented in **Table 19**. Four of these genes also exhibited significant (FDR adjusted P-value < 0.05) treatment-by-line interaction effect. One of these genes (Oxt) was already discussed in Chapter 2.

Table 19. Top important genes found in the four significant modules of the analysis and P-values and FDR adjusted P-values according to the comparison with treatment-by-line interaction effect and the main effect of genotype (grey module) and treatment (salmon, cyan, and tan modules) in the differential expression analysis of Chapter 2.

Gene Symbol	Gene Name	P-value	FDR adjusted P-value[a]
Grey Module			
Fzd9	Frizzled class receptor 9	0.333	0.868
Igf1	Insulin-like growth factor	0.902	0.999
Tacr1	Tachykinin receptor 1	0.556	0.989
Atf4	Activating transcription factor 4	0.771	0.999
Bche	Butyrylcholinesterase	0.366	0.896
Cat	Catalase	0.668	0.999
Ppif	Peptidylprolyl isomerase F (cyclophilin F)	2.7E-06	**1.4E-04**
Cox4i2	Cytochrome c oxidase subunit IV isoform 2	0.917	0.999
Salmon Module			
Jun	Jun proto-oncogene	0.546	0.999
Neurod1	Neurogenic differentiation 1	6.6E-09	1.0E-06
Pnp	Purine-nucleoside phosphorylase	0.432	0.999
Xdh	Xanthine dehydrogenase	0.542	0.999
Cyan Module			
Creb3l1	cAMP responsive element binding protein 3-like 1	2.2E-03	2.9E-02
Emx2	Empty spiracles homeobox 2	0.775	0.999
Ppia	Peptidylprolyl Isomerase A	0.440	0.999
Gm3244	Predicted pseudogene 3244	0.546	0.999
Tgm2	Transglutaminase 2, C polypeptide	0.353	0.999
Tan Module			
Oxt	Oxytocin	7.8E-13	**7.0E-10**
Klk8	Kallikrein related-peptidase 8	0.559	0.999
Cxcr4	Chemokine (C-X-C motif) receptor 4	0.336	0.999

[a]False Discovery rate adjusted P-value.

The identification of gene modules by WGCNA gives more information about the gene expression dataset that may be ignored just by doing differentially expression analysis. These results prove this concept by revealing important genes for the main effects of genotype and treatment, mostly not discussed before. Within the grey module, that was significantly correlated with C line relative to H line, the genes frizzled class receptor 9 (Fzd9), insulin-like growth factor

(Igf1), and tachykinin receptor 1 (Tacr1) were selected for being related to behavior, learning, and memory according to DAVID Functional Annotation Table. In agreement with this study, Fzd9 (-/-) mutant mice in hippocampus presented severe deficits on tests of visuospatial learning/memory, and increased apoptosis in developing dentate gyrus (Zhao et al., 2005). Igf1 is involved in synaptic plasticity and acts on Igf1-receptor, which deletion causes impaired memory formation and learning (Liu et al., 2017b). Tacr1 is also known as a neurokinin-1 receptor (Nk1r) in the brain. Nk1r activation in the hippocampus plays a role in neuropeptides that have excitatory effects, and in the striatum, Tacr1 mediates learning and memory (Liu et al., 2017b). The gene activating transcription factor 4 (Atf4) is related to dopaminergic synapse and amphetamine addiction, and the gene butyrylcholinesterase (Bche) and catalase (Cat), with response to drug cocaine, and alcohol, respectively (e Silva et al., 2018; Mattes et al., 1997; Pavlovsky et al., 2013; Valera et al., 2013). Atf4 is induced by amphetamine and stress in rat striatum, and the overexpression of this genes is involved with addiction-like and emotional behavior (Green et al., 2008; Pavlovsky et al., 2013). Peptidylprolyl isomerase F (cyclophilin F) (Ppif) encodes cyclophilin D, which deficiency in improves learning, memory and synaptic function in Alzheimer's disease (Du et al., 2008). Cytochrome c oxidase subunit IV isoform 2 (Cox4i2) is up-regulated in astrocytes in hypoxic and toxic conditions and is involved with increased necrosis rates in a mice model of Huntington's disease (Misiak et al., 2010). The fact that all these genes are correlated with C line relative to H line might suggest an alteration in the activity of these genes in the H line.

Genes found in these modules can be acting similarly or together with differentially expressed genes. For example, both neuropeptide vasopressin (Avp) and Tacr1 are important in the social interaction (Nelson et al., 2017). Avp, already discussed in Chapter 2, exhibiting significant treatment-by-line interaction effect, can also influence social, affective and addictive behaviors (Nelson et al., 2017). The increase in the expression of Avp and Tacr1 are associated with a decrease in social interaction (Nelson et al., 2017). This study identifies genes inside modules that are important for the complex system, which makes analysis of co-expression essential to complement the results already found.

Among genes in the salmon module that was significantly correlated with mice receiving A relative to S treatment, the gene Jun proto-oncogene, AP-1 transcription factor subunit (Jun) is related to an amphetamine addiction, exposition to amphetamine increases brain expression of Jun, which is in agreement with the correlation found in this module (Persico et al., 1993). The purine-

nucleoside phosphorylase (Pnp) deficiency causes apoptosis of activated T-cells, which consequently turns Pnp in the chemotherapeutic target for T-cell proliferative disorders (Gebre et al., 2017). Xanthine dehydrogenase (Xdh) is related to oxidative metabolism of purines and the increase in the expression of this gene is involved with oxidative stress in the brain of mice model for diabetes (Aliciguzel et al., 2003). The activity of neurogenic differentiation 1 (Neurod1) decreases after drug (morphine) administration, impairing contextual memory, regulating the opioid tolerance's development (Li et al., 2014). These genes are correlated with A relative to S treatment which might suggest a drug addiction behavior and an alteration in the purine metabolism and due to the A treatment. Just Neurod1 exhibited significant treatment-by-line interaction effect; the other genes were not differentially expressed in a previous analysis of the main effect of in mice receiving the A relative to the S treatment.

The cyan module that was significantly correlated with mice receiving S relative to A treatment also presented important genes that complement results found in Chapter 2, like the gene cAMP responsive element binding protein 3-like 1 (Creb3l1), that is a putative transcription factor of gene Avp (involved with addiction behavior before) in rat hypothalamus, and also called old astrocyte specifically induced substance (Oasis) when expressed in astrocytes, and may be important for protection against neuronal damage (Chihara et al., 2009; Greenwood et al., 2015). Mutations in the gene empty spiracles homeobox 2 (Emx2) may cause central nervous system malformation called schizencephaly and can include a motor or cognitive impairment and epilepsy (Dies et al., 2013). Peptidylprolyl isomerase A (Ppia) is also called cyclophilin A, and the presence extracellular Ppia mediates neuroinflammation in both patients and mouse model of sclerosis (Pasetto et al., 2017). The up-regulation of transglutaminase 2, C polypeptide (Tgm2) in different brain regions, is implicated with several neurodegenerative disorders, like Huntington's, Alzheimer's and Parkinson's diseases (Lai et al., 2008). The presence of these genes correlated with mice receiving S relative to A treatment suggests protection or damage of neuronal functions.

Finally, from the tan module that was also significantly correlated with mice receiving S relative to A treatment, oxytocin (Oxt) was selected for playing a role in response to amphetamine and behavior as discussed before. Kallikrein related-peptidase 8 (Klk8) that is related to nervous system development, behavior, learning or memory according to DAVID Functional Annotation Table. Klk8-knockout mice in hippocampus present impaired spatial working memory and anxiety-related behavior (Hirata et al., 2001; Tamura et al., 2017). And gene chemokine (C-X-C

motif) receptor 4 (Cxcr4) that modulates brain inflammation and neurotransmitter release, is down-regulated in the presence of drug anti-depressant administration in the hippocampus and the frontal cortex of adult rats in an animal model of depression (Trojan et al., 2017).

From the genes selected in the significant modules, four of them, one present in the grey module (Ppif), one in the salmon module (Neurod1), one in the cyan module (Creb3l1), and one in the tan module (Oxt), were differentially expressed in Chapter 2 and co-expressed in this analysis. Although Ppif, Neurod1, Creb3l1, and Oxt exhibited significant treatment-by-line interaction effect, just Oxt was discussed due to a large number of significant genes.

Differential Co-expression Analysis

Both WGCNA and differential co-expression analysis is useful for finding genes related to the traits of this study. The WGCNA showed genes in modules that are significantly and correlated with genotype and treatment. In WGCNA, all samples, irrespective of the trait (genotype or treatment) are used for the module construction (Fuller et al., 2007). In differential co-expression analysis what changes is that now we have subgroups to compare. Our aim in this analysis was to identify differentially connected genes between genotypes and between treatments. The genes that have differential connectivity may have a rewiring in response to environmental changes (Fuller et al., 2007).

Differential co-expression analysis revealed that 1,968 genes showed differences in the connectivity between the A and S treatments. **Table 20** is showing the top 10 of 87 genes that were meaningful (DiffK > |0.4|) and positively differentially connected, meaning that 87 genes had greater connectivity in A treatment than in S treatment. **Table 21** is showing the top 10 of 188 genes that were meaningful (DiffK > |0.4|) and negatively differentially connected, meaning that 188 genes had greater connectivity in S treatment than in A treatment.

Table 20. Top 10 genes that were meaningful (DiffK > |0.4|) and positively differentially connected in mice receiving the A relative to the S treatment.

Gene Symbol	Connectivity in A treatment	Connectivity in S treatment	Norm.[a] in A treatment	Norm. in S treatment	DiffK[b]
Slc35c1	55.236	4.902	0.989	0.072	0.917
Zfpm2	55.842	5.737	1	0.085	0.915
Ccdc3	55.842	7.656	1	0.113	0.887
Atoh8	55.842	8.244	1	0.122	0.878
Efna1	55.842	9.530	1	0.141	0.859
Nrtn	51.593	4.489	0.924	0.066	0.857
9330133O14Rik	55.842	12.126	1	0.179	0.821
Foxk1	55.842	12.339	1	0.182	0.818
Zfp316	55.842	12.755	1	0.188	0.812
Copz2	55.842	12.806	1	0.189	0.811

[a] Normalization dividing connectivities by the maximum value

[b] Differential co-expression

Table 21. Top 10 genes that were meaningful (DiffK > |0.4|) and negatively differentially connected in mice receiving the A relative to the S treatment.

Gene Symbol	Connectivity in A treatment	Connectivity in S treatment	Norm.[a] in A treatment	Norm. in S treatment	DiffK[b]
Phax	6.655	67.694	0.119	1	-0.881
Anxa2	6.244	67.167	0.112	0.992	-0.880
Rai2	6.694	67.694	0.119	1	-0.880
Zfp39	6.895	67.167	0.123	0.992	-0.869
Bmp2	7.608	67.694	0.136	1	-0.864
Gna15	7.910	67.167	0.142	0.992	-0.850
Ephb3	8.596	67.694	0.154	1	-0.846
C1qtnf6	8.637	67.694	0.155	1	-0.845
Lysmd1	8.744	67.694	0.157	1	-0.843
Gpn3	8.879	67.694	0.159	1	-0.841

[a] Normalization dividing connectivities by the maximum value

[b] Differential co-expression

This analysis also revealed that 2,128 genes showed differences in the connectivity between the H and C lines. **Table 22** is showing the top 10 of that 109 genes that were meaningful (DiffK > |0.4|) and positively differentially connected, meaning that 109 genes had greater connectivity in C line than in H line. **Table 23** is showing the top 10 of 50 genes that were

meaningful (DiffK > |0.4|) and negatively differentially connected, meaning that 50 genes had greater connectivity in H line than in C line.

Table 22. Top 10 genes that were meaningful (DiffK > |0.4|) and positively differentially connected in mice from C relative to the H line.

Gene Symbol	Connectivity in C line	Connectivity in H line	Norm.[a] in C line	Norm. in H line	DiffK[b]
Mterf	75.542	37.040	0.834	0.333	0.501
Tmco1	87.790	51.887	0.969	0.467	0.503
Zfp748	73.043	33.689	0.807	0.303	0.504
Ccdc80	75.542	36.678	0.834	0.329	0.505
Slc19a2	75.542	36.619	0.834	0.329	0.505
Ppif	66.891	25.925	0.739	0.233	0.506
Trim62	73.043	33.478	0.807	0.301	0.506
Ddc	73.043	33.196	0.807	0.298	0.508
Haus4	73.043	32.944	0.807	0.296	0.511
2310008H04Rik	87.790	50.954	0.970	0.459	0.512

[a] Normalization dividing connectivities by the maximum value

[b] Differential co-expression

Table 23. Top 10 genes that were meaningful (DiffK > |0.4|) and negatively differentially connected in mice from C relative to the H line.

Gene Symbol	Connectivity in C line	Connectivity in H line	Norm.[a] in C line	Norm. in H line	DiffK[b]
Hspe1	23.311	111.164	0.257	1	-0.742
A330050F15Rik	26.913	111.164	0.297	1	-0.703
Chchd8	27.003	111.164	0.298	1	-0.702
Sbk1	27.076	111.164	0.299	1	-0.701
Pcdh20	27.456	111.164	0.303	1	-0.697
Dolk	27.476	111.164	0.303	1	-0.696
Mki67ip	27.917	111.164	0.308	1	-0.692
Snora65	27.930	111.164	0.308	1	-0.691
Hhip	28.757	111.164	0.318	1	-0.682
Cotl1	29.037	111.164	0.321	1	-0.679

[a] Normalization dividing connectivities by the maximum value

[b] Differential co-expression

From the most meaningful (DiffK > |0.4|) and positively differentially connected to mice receiving the A relative to the S treatment, the gene atonal bhlh transcription factor 8 (Atoh8) may be important for the implication in specification and differentiation of neuronal cell lineages in the brain (Yao et al., 2010). For the negatively differentially connected to mice receiving the A relative to the S treatment, EPH receptor B3 (Ephb3) is an ephrin receptor and mediate numerous developmental processes, particularly in the nervous system, playing a deleterious role in synaptic stability and plasticity after brain injury (Perez et al., 2016).

From the most meaningful (DiffK > |0.4|) and positively differentially connected to H-C line, the gene Ppif is a member of the peptidyl-prolyl cis-trans isomerase (PPIase) family. Among the related pathways of Ppif are calcium signaling pathway and respiratory electron transport according to DAVID Functional Annotation Table. This gene is probably one of the most important for this analysis because Ppif was differentially expressed in Chapter 2, significantly co-expressed in the module grey, and differentially co-expressed. For the most meaningful (DiffK > |0.4|) and negatively differentially connected to mice from C relative to the H line, SH3 domain binding kinase 1 (Sbk1) is related to signal-transduction pathways and control of brain development according to DAVID Functional Annotation Table. The number of genes positively and negatively differentially connected to mice receiving the A relative to the S treatment and mice from C relative to the H line prevented the identification of enriched categories using DAVID. However, DAVID Functional Annotation Table was used to identify what function is related to each gene.

Comparisons between analyses

Weighted Gene Co-expression Network Analysis vs. Differential Expression Analysis

Among the 2,193 genes analyzed in WGCNA, 54 were differentially expressed using Cuffdiff software (Chapter 2) for the main effect of A treatment, and 194 genes were differentially expressed in the main effect of H genotype, and 333 genes exhibited significant treatment-by-line interaction effect at FDR adjusted P-value < 0.05. Among the 194 genes differentially expressed across genotypes, the significant grey module included five genes: sarcospan (Sspn), lymphocyte cytosolic protein 1 (Lcp1), angiotensin II receptor-associated protein (Agtrap), insulin-like growth factor binding protein 6 (Igfbp6), and cerebellin 1 precursor (Cbln1). Among these, Cbln1 must

be highlighted as being required for synapse integrity and synaptic plasticity (Hirai et al., 2005). Among the genes in module grey, 41 presented significant interaction effect, as the Ppif, already discussed, neuronal differentiation 6 (Neurod6), that might be involved in the development and differentiation of the nervous system, and synuclein gamma (Sncg) that is a paralog of synuclein alpha (Sncg), involved in the pathogenesis of neurodegenerative diseases like Parkinson's disease (Khan et al., 2017; Siddiqui et al., 2016).

Among the modules including genes correlated with treatment, the salmon module did not encompass differentially expressed genes. The tan module presented 12 genes significant for interaction effects, including insulin receptor substrate 4 (Irs4), and Oxt. The cyan module presented 13 genes that were significant for interaction effects, including Creb3l1 and family with sequence similarity 19 member A1, C-C motif chemokine-like (Fam19a1)). The gene Fam19a1 encodes TAFA proteins, which are neurokines that act as regulators of immune and nervous cells (Tang et al., 2004). In general, the WGCNA analysis gives complementary information about the gene expression dataset and helps to focus on genes that might be important.

Weighted Gene Co-expression Network Analysis vs. Differential Co-expression Analysis

Differential co-expression analysis identified 87 genes that were meaningful and positively differentially connected, while 188 genes were meaningful and negatively differentially connected between mice receiving the A relative to the S treatment. From the positively differentially connected, seven were present in the tan module, and three were present in the cyan module. From the negatively differentially connected, ten genes were also present in the salmon module, one in the tan module, and two in the cyan module.

This analysis also revealed that 109 genes were meaningful and positively differentially connected, while 50 genes that were meaningful and negatively differentially connected between mice from C relative to the H line. From the positively differentially connected, just three genes were also present in the grey module, including the gene Ppif. No gene was both co-expressed in the grey module and negatively differentially connected between mice from C relative to the H line. The differential co-expression analysis also added information to the WGCNA analysis.

The study of weighted gene co-expression network and differential co-expression analysis offered insights that complemented the findings from the differential expression, functional and network analyses in Chapter 2. The study of correlations between genes and treatments or lines, and correlations between genes identified genes that exhibited similar patterns of expression irrespectively of the individual level of differential expression. Differences in gene expression correlation across treatments and lines help to understand changes in connectivity associated with the factors studied.

REFERENCES

Al Mamun, Abdullah, Anjali Chauhan, Haifu Yu, Yan Xu, Romana Sharmeen, and Fudong Liu. 2018. 'Interferon regulatory factor 4/5 signaling impacts on microglial activation after ischemic stroke in mice', *European Journal of Neuroscience*, 47: 140-49.

Albert, Réka, Hawoong Jeong, and Albert-László Barabási. 2000. 'Error and attack tolerance of complex networks', *arXiv preprint cond-mat/0008064*.

Aliciguzel, Yakup, Ikbal Ozen, Mutay Aslan, and Umit Karayalcin. 2003. 'Activities of xanthine oxidoreductase and antioxidant enzymes in different tissues of diabetic rats', *The Journal of laboratory and clinical medicine*, 142: 172-77.

Alvarez, S. E., K. B. Harikumar, N. C. Hait, J. Allegood, G. M. Strub, E. Y. Kim, M. Maceyka, H. Jiang, C. Luo, T. Kordula, S. Milstien, and S. Spiegel. 2010. 'Sphingosine-1-phosphate is a missing cofactor for the E3 ubiquitin ligase TRAF2', *Nature*, 465: 1084-8.

Anand, Paras K, RK Subbarao Malireddi, John R Lukens, Peter Vogel, John Bertin, Mohamed Lamkanfi, and Thirumala-Devi Kanneganti. 2012. 'NLRP6 negatively regulates innate immunity and host defence against bacterial pathogens', *Nature*, 488: 389-93.

Andrews, Simon. 2010. "FastQC: a quality control tool for high throughput sequence data." In.

Arroyo-Espliguero, R, P Avanzas, S Jeffery, and JC Kaski. 2004. 'CD14 and toll-like receptor 4: a link between infection and acute coronary events?', *Heart*, 90: 983-88.

Asghar, K., J. Brain, J. M. Palmer, S. Douglass, F. M. A. Naemi, G. O'Boyle, J. Kirby, and S. Ali. 2017. 'Potential role of indoleamine 2,3-dioxygenase in primary biliary cirrhosis', *Oncol Lett*, 14: 5497-504.

Ashok, A. H., Y. Mizuno, N. D. Volkow, and O. D. Howes. 2017. 'Association of Stimulant Use With Dopaminergic Alterations in Users of Cocaine, Amphetamine, or Methamphetamine: A Systematic Review and Meta-analysis', *JAMA Psychiatry*.

Badiei, Alireza, Steven Gieseg, Sian Davies, Mohd Izani Othman, and Madhav Bhatia. 2015. 'LPS Up-Regulates Cystathionine γ-Lyase Gene Expression in Primary Human Macrophages via NF-κB/ERK Pathway', *Inflammation & Allergy-Drug Targets (Formerly Current Drug Targets-Inflammation & Allergy)*, 14: 99-104.

Banay-Schwartz, M, F Bracco, D Dahl, T DeGuzman, V Turk, and A Lajtha. 1985. 'The pH dependence of breakdown of various purified brain proteins by cathepsin D preparations', *Neurochemistry international*, 7: 607-14.

Baumgartner, R., M. J. Forteza, and D. F. J. Ketelhuth. 2017. 'The interplay between cytokines and the Kynurenine pathway in inflammation and atherosclerosis', *Cytokine*.

Beck, Jon A, Sarah Lloyd, Majid Hafezparast, Moyha Lennon-Pierce, Janan T Eppig, Michael FW Festing, and Elizabeth MC Fisher. 2000. 'Genealogies of mouse inbred strains', *Nature genetics*, 24: 23.

Beiswanger, CM, MH Diegmann, RF Novak, MA Philbert, TL Graessle, KR Reuhl, and HE Lowndes. 1994. 'Developmental changes in the cellular distribution of glutathione and glutathione S-transferases in the murine nervous system', *Neurotoxicology*, 16: 425-40.

Belforte, Juan E, and Kazu Nakazawa. 2011. 'Genetically Engineered Mice for Schizophrenia Research', *Animal Models of Schizophrenia and Related Disorders*: 231-42.

Benjamini, Yoav, and Yosef Hochberg. 1995. 'Controlling the false discovery rate: a practical and powerful approach to multiple testing', *Journal of the royal statistical society. Series B (Methodological)*: 289-300.

Berman, Steven M, Ronald Kuczenski, James T McCracken, and Edythe D London. 2009. "Potential adverse effects of amphetamine treatment on brain and behavior: a review." In.: Nature Publishing Group.

Biever, Anne, Jihane Boubaker-Vitre, Laura Cutando, Irene Gracia-Rubio, Mauro Costa-Mattioli, Emma Puighermanal, and Emmanuel Valjent. 2016. 'Repeated exposure to d-amphetamine decreases global protein synthesis and regulates the translation of a subset of mRNAs in the striatum', *Frontiers in molecular neuroscience*, 9.

Borsini, A., A. Cattaneo, C. Malpighi, S. Thuret, N. A. Harrison, M. R. C. ImmunoPsychiatry Consortium, P. A. Zunszain, and C. M. Pariante. 2017. 'Interferon-Alpha Reduces Human Hippocampal Neurogenesis and Increases Apoptosis via Activation of Distinct STAT1-Dependent Mechanisms', *Int J Neuropsychopharmacol*.

Boules, Mona, Lewis Warrington, Abdul Fauq, Daniel McCormick, and Elliott Richelson. 2001. 'A novel neurotensin analog blocks cocaine-and D-amphetamine-induced hyperactivity', *European journal of pharmacology*, 426: 73-76.

Bourque, C., Y. Zhang, M. Fu, M. Racine, A. Greasley, Y. Pei, L. Wu, R. Wang, and G. Yang. 2017. 'H2S protects lipopolysaccharide-induced inflammation by blocking NFkappaB transactivation in endothelial cells', *Toxicol Appl Pharmacol*, 338: 20-29.

Bray, NL, H Pimentel, P Melsted, and L Pachter. 2016. "Near-optimal probabilistic RNA-seq quantification. Nat. 2016;(5): 525-7| Biotechnol 34." In.: PubMed Abstract Publisher Full Text.

Breitkreutz, Bobby-Joe, Chris Stark, and Mike Tyers. 2003. 'Osprey: a network visualization system', *Genome biology*, 4: R22.

Brennan, Avis R, and Amy FT Arnsten. 2008. 'Neuronal mechanisms underlying attention deficit hyperactivity disorder', *Annals of the New York Academy of Sciences*, 1129: 236-45.

Brites, D., and A. Fernandes. 2015. 'Neuroinflammation and Depression: Microglia Activation, Extracellular Microvesicles and microRNA Dysregulation', *Front Cell Neurosci*, 9: 476.

Brosius, Jürgen, and Carsten A Raabe. 2016. 'What is an RNA? A top layer for RNA classification', *RNA biology*, 13: 140-44.

Builee, TL, and JR Hatherill. 2004. 'The role of polyhalogenated aromatic hydrocarbons on thyroid hormone disruption and cognitive function: a review', *Drug and chemical toxicology*, 27: 405-24.

Buonaguro, Elisabetta Filomena, Felice Iasevoli, Federica Marmo, Anna Eramo, Gianmarco Latte, Camilla Avagliano, Carmine Tomasetti, and Andrea de Bartolomeis. 2017. 'Re-arrangements of gene transcripts at glutamatergic synapses after prolonged treatments with antipsychotics: A putative link with synaptic remodeling', *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 76: 29-41.

Burns, Lindsay H, Trevor W Robbins, and Barry J Everitt. 1993. 'Differential effects of excitotoxic lesions of the basolateral amygdala, ventral subiculum and medial prefrontal cortex on responding with conditioned reinforcement and locomotor activity potentiated by intra-accumbens infusions ofd-amphetamine', *Behavioural brain research*, 55: 167-83.

Burrows, Michael, and David J Wheeler. 1994. 'A block-sorting lossless data compression algorithm'.

Caetano-Anollés, Kelsey, Sanjibita Mishra, and Sandra L Rodriguez-Zas. 2015. 'Synergistic and antagonistic interplay between Myostatin gene expression and physical activity levels on gene expression patterns in Triceps Brachii muscles of C57/BL6 mice', *PLoS One*, 10: e0116828.

Caetano-Anollés, Kelsey, Justin S Rhodes, Theodore Garland Jr, Sam D Perez, Alvaro G Hernandez, Bruce R Southey, and Sandra L Rodriguez-Zas. 2016. 'Cerebellum transcriptome of mice bred for high voluntary activity offers insights into locomotor control and reward-dependent behaviors', *PLoS One*, 11: e0167095.

Cai, Hao, Xiao-Dong Zhu, Jian-Yang Ao, Bo-Gen Ye, Yuan-Yuan Zhang, Zong-Tao Chai, Cheng-Hao Wang, Wen-Kai Shi, Man-qing Cao, and Xiao-Long Li. 2017. 'Colony stimulating factor 1-induced AIF1 expression in tumor-associated macrophages enhances the progression of hepatocellular carcinoma', *OncoImmunology*: 00-00.

Capuron, Lucile, and Andrew H Miller. 2011. 'Immune system to brain signaling: neuropsychopharmacological implications', *Pharmacology & therapeutics*, 130: 226-38.

Carboni, L., T. P. Nguyen, and L. Caberlotto. 2016. 'Systems biology integration of proteomic data in rodent models of depression reveals involvement of the immune response and glutamatergic signaling', *Proteomics Clin Appl*, 10: 1254-63.
Carette, Jan E, Carla P Guimaraes, Malini Varadarajan, Annie S Park, Irene Wuethrich, Alzbeta Godarova, Maciej Kotecki, Brent H Cochran, Eric Spooner, and Hidde L Ploegh. 2009. 'Haploid genetic screens in human cells identify host factors used by pathogens', *Science*, 326: 1231-35.
Carson, Dean S. 2014. 'Oxytocin Treatment for Amphetamine-Induced Social Impairments', *Journal of Neuroscience*, 34: 14503-05.
Carson, Dean S, Adam J Guastella, Emily R Taylor, and Iain S McGregor. 2013. 'A brief history of oxytocin and its role in modulating psychostimulant effects', *Journal of Psychopharmacology*: 0269881112473788.
Castellanos, F Xavier, and Erika Proal. 2012. 'Large-scale brain systems in ADHD: beyond the prefrontal–striatal model', *Trends in cognitive sciences*, 16: 17-26.
Chase, TD, N Carrey, RE Brown, and M Wilkinson. 2005. 'Methylphenidate regulates c-fos and fosB expression in multiple regions of the immature rat brain', *Developmental brain research*, 156: 1-12.
Chen, Long, Xiaobo Yang, Xiaoyong Zhou, Cuicui Wang, Xue Gong, Biqin Chen, and Yinghui Chen. 2015. 'Hyperactivity and impaired attention in Gamma aminobutyric acid transporter subtype 1 gene knockout mice', *Acta neuropsychiatrica*, 27: 368-74.
Chepelev, Iouri, Gang Wei, Qingsong Tang, and Keji Zhao. 2009. 'Detection of single nucleotide variations in expressed exons of the human genome using RNA-Seq', *Nucleic acids research*, 37: e106-e06.
Chesler, Elissa J, Darla R Miller, Lisa R Branstetter, Leslie D Galloway, Barbara L Jackson, Vivek M Philip, Brynn H Voy, Cymbeline T Culiat, David W Threadgill, and Robert W Williams. 2008. 'The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics', *Mammalian Genome*, 19: 382-89.
Chihara, Kazuyasu, Atsushi Saito, Tomohiko Murakami, Shin-ichiro Hino, Yuri Aoki, Hiroshi Sekiya, Yuji Aikawa, Akio Wanaka, and Kazunori Imaizumi. 2009. 'Increased vulnerability of hippocampal pyramidal neurons to the toxicity of kainic acid in OASIS-deficient mice', *Journal of neurochemistry*, 110: 956-65.
Chiu, Joanne, Harold Kalant, and Dzung Anh Lê. 1998. 'Vasopressin opposes locomotor stimulation by ethanol, cocaine and amphetamine in mice', *European journal of pharmacology*, 355: 11-17.
Choi, Jung Kyoon, Ungsik Yu, Ook Joon Yoo, and Sangsoo Kim. 2005. 'Differential coexpression analysis using microarray data and its application to human cancer', *Bioinformatics*, 21: 4348-55.
Christie, Kimberly J, Ben Emery, Mark Denham, Helena Bujalka, Holly S Cate, and Ann M Turnley. 2013. 'Transcriptional regulation and specification of neural stem cells.' in, *Transcriptional and Translational Regulation of Stem Cells* (Springer).
Christmas, David M, JP Potokar, and Simon JC Davies. 2011. 'A biological pathway linking inflammation and depression: activation of indoleamine 2, 3-dioxygenase', *Neuropsychiatric disease and treatment*, 7: 431.
Compeau, Phillip EC, Pavel A Pevzner, and Glenn Tesler. 2011. 'How to apply de Bruijn graphs to genome assembly', *Nature biotechnology*, 29: 987-91.
Conesa, Ana, Pedro Madrigal, Sonia Tarazona, David Gomez-Cabrero, Alejandra Cervera, Andrew McPherson, Michał Wojciech Szcześniak, Daniel J Gaffney, Laura L Elo, and Xuegong Zhang. 2016. 'A survey of best practices for RNA-seq data analysis', *Genome biology*, 17: 13.
Cortes-Mendoza, J., S. Diaz de Leon-Guerrero, G. Pedraza-Alva, and L. Perez-Martinez. 2013. 'Shaping synaptic plasticity: the role of activity-mediated epigenetic regulation on gene transcription', *Int J Dev Neurosci*, 31: 359-69.
Crnic, Linda S, and Mark A Segall. 1992. 'Behavioral effects of mouse interferons-α and-γ and human interferon-α in mice', *Brain research*, 590: 277-84.

Culhane, Aedín C, Markus S Schröder, Razvan Sultana, Shaita C Picard, Enzo N Martinelli, Caroline Kelly, Benjamin Haibe-Kains, Misha Kapushesky, Anne-Alyssa St Pierre, and William Flahive. 2011. 'GeneSigDB: a manually curated database and resource for analysis of gene expression signatures', *Nucleic acids research*, 40: D1060-D66.

Cusick, J. K., A. Mustian, K. Goldberg, and M. E. Reyland. 2010. 'RELT induces cellular death in HEK 293 epithelial cells', *Cell Immunol*, 261: 1-8.

Dantzer, R., J. C. O'Connor, G. G. Freund, R. W. Johnson, and K. W. Kelley. 2008. 'From inflammation to sickness and depression: when the immune system subjugates the brain', *Nat Rev Neurosci*, 9: 46-56.

de la Fuente, Alberto. 2010. 'From ‘differential expression’to ‘differential networking’–identification of dysfunctional regulatory networks in diseases', *Trends in genetics*, 26: 326-33.

De Munnik, Sonja A, Sixto García-Miñaúr, Alexander Hoischen, Bregje W Van Bon, Kym M Boycott, Jeroen Schoots, Lies H Hoefsloot, Nine VAM Knoers, Ernie MHF Bongers, and Han G Brunner. 2014. 'A de novo non-sense mutation in ZBTB18 in a patient with features of the 1q43q44 microdeletion syndrome', *European Journal of Human Genetics*, 22: 844.

dela Peña, Ike, June Bryan de la Peña, Bung-Nyun Kim, Doug Hyun Han, Minsoo Noh, and Jae Hoon Cheong. 2015. 'Gene expression profiling in the striatum of amphetamine-treated spontaneously hypertensive rats which showed amphetamine conditioned place preference and self-administration', *Archives of pharmacal research*, 38: 865-75.

Delfino, KR, NVL Serão, BR Southey, and SL Rodriguez-Zas. 2011. 'Therapy-, gender-and race-specific microRNA markers, target genes and networks related to glioblastoma recurrence and survival', *Cancer Genomics-Proteomics*, 8: 173-83.

Denard, B., J. Seemann, Q. Chen, A. Gay, H. Huang, Y. Chen, and J. Ye. 2011. 'The membrane-bound transcription factor CREB3L1 is activated in response to virus infection to inhibit proliferation of virus-infected cells', *Cell Host Microbe*, 10: 65-74.

Dies, Kira A, Adria Bodell, Fuki M Hisama, Chao-Yu Guo, Brenda Barry, Bernard S Chang, A James Barkovich, and Christopher A Walsh. 2013. 'Schizencephaly: association with young maternal age, alcohol use, and lack of prenatal care', *Journal of child neurology*, 28: 198-203.

Donma, MM, and O Donma. 2016. 'Promising link between selenium and peroxisome proliferator activated receptor gamma in the treatment protocols of obesity as well as depression', *Medical hypotheses*, 89: 79-83.

Dowlati, Yekta, Nathan Herrmann, Walter Swardfager, Helena Liu, Lauren Sham, Elyse K Reim, and Krista L Lanctôt. 2010. 'A meta-analysis of cytokines in major depression', *Biological psychiatry*, 67: 446-57.

Du, Bing, Weijia Luo, Ruimei Li, Binghe Tan, Honghui Han, Xiaoling Lu, Dali Li, Min Qian, Dekai Zhang, and Yongxiang Zhao. 2013. 'Lgr4/Gpr48 negatively regulates TLR2/4-associated pattern recognition and innate immunity by targeting CD14 expression', *Journal of Biological Chemistry*, 288: 15131-41.

Du, Heng, Lan Guo, Fang Fang, Doris Chen, Alexander A Sosunov, Guy M McKhann, Yilin Yan, Chunyu Wang, Hong Zhang, and Jeffery D Molkentin. 2008. 'Cyclophilin D deficiency attenuates mitochondrial and neuronal perturbation and ameliorates learning and memory in Alzheimer's disease', *Nature medicine*, 14: 1097.

e Silva, Liana Dantas da Costa, Patrícia Pereira, Gabriela Gregory Regner, Fernanda Brião Menezes Boaretto, Cleonice Hoffmann, Pricila Pflüger, Lucas Lima da Silva, Luiza Reinhardt Steffens, Ana Moira Morás, and Dinara Jaqueline Moura. 2018. 'DNA damage and oxidative stress induced by seizures are decreased by anticonvulsant and neuroprotective effects of lobeline, a candidate to treat alcoholism', *Metabolic brain disease*, 33: 53-61.

Eipper-Mains, Jodi E, Drew D Kiraly, Michael O Duff, Michael J Horowitz, C Joel McManus, Betty A Eipper, Brenton R Graveley, and Richard E Mains. 2013. 'Effects of cocaine and withdrawal on the mouse nucleus accumbens transcriptome', *Genes, Brain and Behavior*, 12: 21-33.

El Kadmiri, N, I Slassi, B El Moutawakil, S Nadifi, A Tadevosyan, A Hachem, and A Soukri. 2014. 'Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Alzheimer's disease', *Pathologie Biologie*, 62: 333-36.

Fan, Jie, Bing Jin Li, Xue Feng Wang, Li Li Zhong, and Ran Ji Cui. 2017. 'Ghrelin produces antidepressant-like effect in the estrogen deficient mice', *Oncotarget*, 8: 58964.

Faraone, Stephen V, Roy H Perlis, Alysa E Doyle, Jordan W Smoller, Jennifer J Goralnick, Meredith A Holmgren, and Pamela Sklar. 2005. 'Molecular genetics of attention-deficit/hyperactivity disorder', *Biological psychiatry*, 57: 1313-23.

Farias, M. G., N. P. de Lucena, S. Dal Bo, and S. M. de Castro. 2014. 'Neutrophil CD64 expression as an important diagnostic marker of infection and sepsis in hospital patients', *J Immunol Methods*, 414: 65-8.

Field, David J, Angela A Aggrey-Amable, Sara K Blick, Sara K Ture, Andrew Johanson, Scott J Cameron, Sukanya Roy, and Craig N Morrell. 2017. 'Platelet factor 4 increases bone marrow B cell development and differentiation', *Immunologic research*, 65: 1089-94.

Finethy, Ryan, Sarah Luoma, Nichole Orench-Rivera, Eric M Feeley, Arun K Haldar, Masahiro Yamamoto, Thirumala-Devi Kanneganti, Meta J Kuehn, and Jörn Coers. 2017. 'Inflammasome Activation by Bacterial Outer Membrane Vesicles Requires Guanylate Binding Proteins', *MBio*, 8: e01188-17.

Foster, P. S., S. Maltby, H. F. Rosenberg, H. L. Tay, S. P. Hogan, A. M. Collison, M. Yang, G. E. Kaiko, P. M. Hansbro, R. K. Kumar, and J. Mattes. 2017. 'Modeling TH 2 responses and airway inflammation to understand fundamental mechanisms regulating the pathogenesis of asthma', *Immunol Rev*, 278: 20-40.

Franke, Barbara, Benjamin M Neale, and Stephen V Faraone. 2009. 'Genome-wide association studies in ADHD', *Human genetics*, 126: 13-50.

Frey, Benício Noronha, Márcio Rodrigo Martins, Fabrícia Cardoso Petronilho, Felipe Dal-Pizzol, João Quevedo, and Flávio Kapczinski. 2006. 'Increased oxidative stress after repeated amphetamine exposure: possible relevance as a model of mania', *Bipolar disorders*, 8: 275-80.

Fuchs, Anja, Marina Cella, Emanuele Giurisato, Andrey S Shaw, and Marco Colonna. 2004. 'Cutting edge: CD96 (tactile) promotes NK cell-target cell adhesion by interacting with the poliovirus receptor (CD155)', *The Journal of Immunology*, 172: 3994-98.

Fuller, Tova F, Anatole Ghazalpour, Jason E Aten, Thomas A Drake, Aldons J Lusis, and Steve Horvath. 2007. 'Weighted gene coexpression network analysis strategies applied to mouse weight', *Mammalian Genome*, 18: 463-72.

Furman, David, and Mark M Davis. 2015. 'New approaches to understanding the immune response to vaccination and infection', *Vaccine*, 33: 5271-81.

Gao, Xuejiao, Mengyao Guo, Zecai Zhang, Peng Shen, Zhengtao Yang, and Naisheng Zhang. 2017. 'Baicalin promotes the bacteriostatic activity of lysozyme on S. aureus in mammary glands and neutrophilic granulocytes in mice', *Oncotarget*, 8: 19894.

Gapin, Jennifer I, Jeffrey D Labban, and Jennifer L Etnier. 2011. 'The effects of physical activity on attention deficit hyperactivity disorder symptoms: The evidence', *Preventive Medicine*, 52: S70-S74.

Garber, Manuel, Manfred G Grabherr, Mitchell Guttman, and Cole Trapnell. 2011. 'Computational methods for transcriptome annotation and quantification using RNA-seq', *Nature methods*, 8: 469-77.

Gebre, Sara T, Scott A Cameron, Lei Li, YS Babu, and Vern L Schramm. 2017. 'Intracellular rebinding of transition-state analogues provides extended in vivo inhibition lifetimes on human purine nucleoside phosphorylase', *Journal of Biological Chemistry*, 292: 15907-15.

Gelsthorpe, Adrian R, Keith Gelsthorpe, and Robert J Sokol. 1997. 'Rapid isolation of total RNA from small samples of hypocelluar dense connective tissues', *Biotechniques*, 22: 11082-86.

Gendrel, Anne-Valerie, Anwyn Apedaile, Heather Coker, Ausma Termanis, Ilona Zvetkova, Jonathan Godwin, Y Amy Tang, Derek Huntley, Giovanni Montana, and Steven Taylor. 2012. 'Smchd1-dependent and-independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome', *Developmental cell*, 23: 265-79.

Girotti, Milena, Jennifer J Donegan, and David A Morilak. 2011. 'Chronic intermittent cold stress sensitizes neuro-immune reactivity in the rat brain', *Psychoneuroendocrinology*, 36: 1164-74.

Gonçalves, Ricardo, and David M Mosser. 2015. 'The isolation and characterization of murine macrophages', *Current protocols in immunology*: 14.1. 1-14.1. 16.

Gonzalez-Pena, Dianelys, Scott E Nixon, Jason C O'Connor, Bruce R Southey, Marcus A Lawson, Robert H McCusker, Tania Borras, Debbie Machuca, Alvaro G Hernandez, and Robert Dantzer. 2016. 'Microglia transcriptome changes in a model of depressive behavior after immune challenge', *PLoS One*, 11: e0150858.

Gonzalez-Pena, Dianelys, Scott E Nixon, Bruce R Southey, Marcus A Lawson, Robert H McCusker, Alvaro G Hernandez, Robert Dantzer, Keith W Kelley, and Sandra L Rodriguez-Zas. 2016. 'Differential transcriptome networks between IDO1-knockout and wild-type mice in brain microglia and macrophages', *PLoS One*, 11: e0157727.

Govindarajan, Rajeshwar, Jeyapradha Duraiyan, Karunakaran Kaliyappan, and Murugesan Palanisamy. 2012. 'Microarray and its applications', *Journal of pharmacy & bioallied sciences*, 4: S310.

Granseth, Björn, Fredrik K Andersson, and Sarah H Lindström. 2015. 'The initial stage of reversal learning is impaired in mice hemizygous for the vesicular glutamate transporter (VGluT1)', *Genes, Brain and Behavior*, 14: 477-85.

Green, Thomas A, Imran N Alibhai, Stephen Unterberg, Rachael L Neve, Subroto Ghose, Carol A Tamminga, and Eric J Nestler. 2008. 'Induction of activating transcription factors (ATFs) ATF2, ATF3, and ATF4 in the nucleus accumbens and their regulation of emotional behavior', *Journal of Neuroscience*, 28: 2025-32.

Greenwood, Mingkwan, Michael P Greenwood, Andre S Mecawi, Su Yi Loh, José Antunes Rodrigues, Julian FR Paton, and David Murphy. 2015. 'Transcription factor CREB3L1 mediates cAMP and glucocorticoid regulation of arginine vasopressin gene transcription in the rat hypothalamus', *Molecular brain*, 8: 68.

Gruber, A. J., and R. J. McDonald. 2012. 'Context, emotion, and the strategic pursuit of goals: interactions among multiple brain systems controlling motivated behavior', *Front Behav Neurosci*, 6: 50.

Gu, Linlin, Jennifer L Larson-Casey, and A Brent Carter. 2017. 'Macrophages utilize the mitochondrial calcium uniporter for profibrotic polarization', *The FASEB Journal*: fj. 201601371R.

Ha, Tae-Young, Keun-A Chang, Jeong a Kim, Hye-Sun Kim, Seonghan Kim, Young Hae Chong, and Yoo-Hun Suh. 2010. 'S100a9 knockdown decreases the memory impairment and the neuropathology in Tg2576 mice, AD animal model', *PLoS One*, 5: e8840.

Harrington, Adam J, Aram Raissi, Kacey Rajkovich, Stefano Berto, Jaswinder Kumar, Gemma Molinaro, Jonathan Raduazzo, Yuhong Guo, Kris Loerwald, and Genevieve Konopka. 2016. 'MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to neurodevelopmental disorders', *eLife*, 5: e20059.

Harrop, R., M. G. Ryan, K. A. Myers, I. Redchenko, S. M. Kingsman, and M. W. Carroll. 2006. 'Active treatment of murine tumors with a highly attenuated vaccinia virus expressing the tumor associated antigen 5T4 (TroVax) is CD4+ T cell dependent and antibody mediated', *Cancer Immunol Immunother*, 55: 1081-90.

Hayes, John D, Jack U Flanagan, and Ian R Jowsey. 2005. 'Glutathione transferases', *Annu. Rev. Pharmacol. Toxicol.*, 45: 51-88.

He, Fan, Xiaohui Shao, Mingji Yi, Yu Wang, Chuan-Yue Wang, and Shiguo Liu. 2015. 'Association of IL-1α rs17561 and IL-1 RN rs315952 polymorphisms with Tourette syndrome: a family-based study', *International journal of clinical and experimental pathology*, 8: 4182.

Heinz, L. X., C. L. Baumann, M. S. Koberlin, B. Snijder, R. Gawish, G. Shui, O. Sharif, I. M. Aspalter, A. C. Muller, R. K. Kandasamy, F. P. Breitwieser, A. Pichlmair, M. Bruckner, M. Rebsamen, S. Bluml, T. Karonitsch, A. Fauster, J. Colinge, K. L. Bennett, S. Knapp, M. R. Wenk, and G. Superti-Furga. 2015. 'The Lipid-Modifying Enzyme SMPDL3B Negatively Regulates Innate Immunity', *Cell Rep*, 11: 1919-28.

Hirai, Hirokazu, Zhen Pang, Dashi Bao, Taisuke Miyazaki, Leyi Li, Eriko Miura, Jennifer Parris, Yongqi Rong, Masahiko Watanabe, and Michisuke Yuzaki. 2005. 'Cbln1 is essential for synaptic integrity and plasticity in the cerebellum', *Nature neuroscience*, 8: 1534.

Hirata, Akio, Shigetaka Yoshida, Naoko Inoue, Kazumasa Matsumoto-Miyai, Ayako Ninomiya, Manabu Taniguchi, Tomohiro Matsuyama, Keiko Kato, Hisashi Iizasa, and Yuki Kataoka. 2001. 'Abnormalities of synapses and neurons in the hippocampus of neuropsin-deficient mice', *Molecular and Cellular Neuroscience*, 17: 600-10.

Hirsch, SR, R Richardson-Andrews, B Costall, ME Kelly, J De Belleroche, and RJ Naylor. 1987. 'The effects of some polyamines on putative behavioural indices of mesolimbic versus striatal dopaminergic function', *Psychopharmacology*, 93: 101-04.

Hoffman, Laurel, Anuja Chandrasekar, Xu Wang, John A Putkey, and M Neal Waxham. 2014. 'Neurogranin alters the structure and calcium binding properties of calmodulin', *Journal of Biological Chemistry*, 289: 14644-55.

Hollander, Myles, and Douglas A Wolfe. 1999. "Nonparametric statistical methods/by Myles Hollander, Douglas A. Wolfe." In.

Holm, A., K. E. Magnusson, and E. Vikstrom. 2016. 'Pseudomonas aeruginosa N-3-oxo-dodecanoyl-homoserine Lactone Elicits Changes in Cell Volume, Morphology, and AQP9 Characteristics in Macrophages', *Front Cell Infect Microbiol*, 6: 32.

Hoyo-Becerra, Carolina, Zijian Liu, Jinghong Yao, Britta Kaltwasser, Guido Gerken, Dirk M Hermann, and Joerg F Schlaak. 2015. 'Rapid regulation of depression-associated genes in a new mouse model mimicking interferon-α-related depression in hepatitis C virus infection', *Molecular neurobiology*, 52: 318-29.

Hsieh, F. L., and M. K. Higgins. 2017. 'The structure of a LAIR1-containing human antibody reveals a novel mechanism of antigen recognition', *eLife*, 6.

Hu, Zhenjun, Joe Mellor, Jie Wu, Takuji Yamada, Dustin Holloway, and Charles DeLisi. 2005. 'VisANT: data-integrating visual framework for biological networks and modules', *Nucleic acids research*, 33: W352-W57.

Huang, Da Wei, Brad T Sherman, and Richard A Lempicki. 2009. 'Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources', *Nature protocols*, 4: 44-57.

Hung, Y. Y., C. C. Lin, H. Y. Kang, and T. L. Huang. 2017. 'TNFAIP3, a negative regulator of the TLR signaling pathway, is a potential predictive biomarker of response to antidepressant treatment in major depressive disorder', *Brain Behav Immun*, 59: 265-72.

Jahanshad, Neda, Priya Rajagopalan, Xue Hua, Derrek P Hibar, Talia M Nir, Arthur W Toga, Clifford R Jack, Andrew J Saykin, Robert C Green, and Michael W Weiner. 2013. 'Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity', *Proceedings of the National Academy of Sciences*, 110: 4768-73.

Jang, W. Y., B. R. Lee, J. Jeong, Y. Sung, M. Choi, P. Song, H. Kim, S. Jang, H. Kim, K. I. Joo, J. W. Lee, Y. S. Choo, E. Kim, and Z. Y. Ryoo. 2017. 'Overexpression of serum amyloid a 1 induces depressive-like behavior in mice', *Brain Res*, 1654: 55-65.

Jaworska-Feil, L, B Budziszewska, and W Lason. 1995. 'The effects of repeated amphetamine administration on the thyrotropin-releasing hormone level. Its release and receptors in the rat brain', *Neuropeptides*, 29: 171-76.

Jeltsch, Hélène, Christine Lazarus, Brigitte Cosquer, Rodrigue Galani, and Jean-Christophe Cassel. 2004. 'No facilitation of amphetamine-or cocaine-induced hyperactivity in adult rats after various 192 IgG-saporin lesions in the basal forebrain', *Brain research*, 1029: 259-71.

Jiang, Haisong, and Mengqing Xiang. 2009. 'Subtype specification of GABAergic amacrine cells by the orphan nuclear receptor Nr4a2/Nurr1', *Journal of Neuroscience*, 29: 10449-59.
Jiang, Yue, Rong Wu, Chen Chen, Zhi-Fei You, Xingguang Luo, and Xiao-Ping Wang. 2015. 'Six novel rare non-synonymous mutations for migraine without aura identified by exome sequencing', *Journal of neurogenetics*, 29: 188-94.
Johansson, Björn, Linda Halldner, Thomas V Dunwiddie, Susan A Masino, Wolfgang Poelchen, Lydia Giménez-Llort, Rosa M Escorihuela, Alberto Fernández-Teruel, Zsuzsanna Wiesenfeld-Hallin, and Xiao-Jun Xu. 2001. 'Hyperalgesia, anxiety, and decreased hypoxic neuroprotection in mice lacking the adenosine A1 receptor', *Proceedings of the National Academy of Sciences*, 98: 9407-12.
Keshava Prasad, TS, Renu Goel, Kumaran Kandasamy, Shivakumar Keerthikumar, Sameer Kumar, Suresh Mathivanan, Deepthi Telikicherla, Rajesh Raju, Beema Shafreen, and Abhilash Venugopal. 2008. 'Human protein reference database—2009 update', *Nucleic acids research*, 37: D767-D72.
Khan, Shabana, Simon RW Stott, Audrey Chabrat, Anna M Truckenbrodt, Bradley Spencer-Dene, Klaus-Armin Nave, François Guillemot, Martin Levesque, and Siew-Lan Ang. 2017. 'Survival of a novel subset of midbrain dopaminergic neurons projecting to the lateral septum is dependent on NeuroD proteins', *Journal of Neuroscience*, 37: 2305-16.
Killcoyne, Sarah, Gregory W Carter, Jennifer Smith, and John Boyle. 2009. 'Cytoscape: a community-based framework for network modeling', *Protein Networks and Pathway Analysis*: 219-39.
Kim, Daehwan, Geo Pertea, Cole Trapnell, Harold Pimentel, Ryan Kelley, and Steven L Salzberg. 2013. 'TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions', *Genome biology*, 14: R36.
Kim, G. D., R. Das, X. Rao, J. Zhong, J. A. Deiuliis, D. L. Ramirez-Bergeron, S. Rajagopalan, and G. H. Mahabeleshwar. 2017. 'CITED2 restrains pro-inflammatory macrophage activation and response', *Mol Cell Biol*.
Kinsler, R., M. M. Taylor, N. M. Flores, J. J. Leffert, and R. D. Beech. 2010. 'Altered response to antidepressant treatment in FoxG1 heterozygous knockout mice', *Synapse*, 64: 169-71.
Kirby, L. G., F. D. Zeeb, and C. A. Winstanley. 2011. 'Contributions of serotonin in addiction vulnerability', *Neuropharmacology*, 61: 421-32.
Kittel-Schneider, Sarah, Max Hilscher, Claus-Jürgen Scholz, Heike Weber, Lena Grünewald, Ricarda Schwarz, Andreas G Chiocchetti, and Andreas Reif. 2017. 'Lithium-induced gene expression alterations in two peripheral cell models of bipolar disorder', *The World Journal of Biological Psychiatry*: 1-14.
Koch, J., J. A. Mayr, B. Alhaddad, C. Rauscher, J. Bierau, R. Kovacs-Nagy, K. L. Coene, I. Bader, M. Holzhacker, H. Prokisch, H. Venselaar, R. A. Wevers, F. Distelmaier, T. Polster, S. Leiz, C. Betzler, T. M. Strom, W. Sperl, T. Meitinger, S. B. Wortmann, and T. B. Haack. 2017. 'CAD mutations and uridine-responsive epileptic encephalopathy', *Brain*, 140: 279-86.
Krzascik, P., M. E. Zajda, and M. D. Majewska. 2015. 'The neurosteroid dehydroepiandrosterone sulfate, but not androsterone, enhances the antidepressant effect of cocaine examined in the forced swim test--Possible role of serotonergic neurotransmission', *Horm Behav*, 70: 64-72.
Kumamoto, T., and C. Hanashima. 2017. 'Evolutionary conservation and conversion of Foxg1 function in brain development', *Dev Growth Differ*, 59: 258-69.
Labandeira-Garcia, JL, JP Tobio, and MJ Guerra. 1994. 'Comparison between normal developing striatum and developing striatal grafts using drug-induced Fos expression and neuron-specific enolase immunohistochemistry', *Neuroscience*, 60: 399-415.
Lai, Thung-S, Yusha Liu, Tim Tucker, Kurt R Daniel, David C Sane, Eric Toone, James R Burke, Warren J Strittmatter, and Charles S Greenberg. 2008. 'Identification of chemical inhibitors to human tissue transglutaminase by screening existing drug libraries', *Chemistry & biology*, 15: 969-78.
Lalonde, Robert, and Catherine Strazielle. 2015. 'Behavioral effects of neonatal lesions on the cerebellar system', *International Journal of Developmental Neuroscience*, 43: 58-65.

Langfelder, Peter, and Steve Horvath. 2007. 'Eigengene networks for studying the relationships between co-expression modules', *BMC systems biology*, 1: 54.
Langmead, Ben, Cole Trapnell, Mihai Pop, and Steven L Salzberg. 2009. 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome', *Genome biology*, 10: R25.
Lanner, Johanna T, Dimitra K Georgiou, Aditya D Joshi, and Susan L Hamilton. 2010. 'Ryanodine receptors: structure, expression, molecular details, and function in calcium release', *Cold Spring Harbor perspectives in biology*, 2: a003996.
Leal, Sixto M, and Margaret L Gulley. 2017. 'Current and Emerging Molecular Tests for Human Papillomavirus–Related Neoplasia in the Genomic Era', *The Journal of Molecular Diagnostics*.
Leighton, S. P., L. Nerurkar, R. Krishnadas, C. Johnman, G. J. Graham, and J. Cavanagh. 2017. 'Chemokines in depression in health and in inflammatory illness: a systematic review and meta-analysis', *Mol Psychiatry*.
Leong, Kah-Chung, Luyi Zhou, Shannon M Ghee, Ronald E See, and Carmela M Reichel. 2016. 'Oxytocin decreases cocaine taking, cocaine seeking, and locomotor activity in female rats', *Experimental and clinical psychopharmacology*, 24: 55.
Lesch, Klaus-Peter, Nina Timmesfeld, Tobias J Renner, Rebecca Halperin, Christoph Röser, T Trang Nguyen, David W Craig, Jasmin Romanos, Monika Heine, and Jobst Meyer. 2008. 'Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies', *Journal of neural transmission*, 115: 1573-85.
Li, C., W. Ying, Z. Huang, T. Brehm, A. Morin, A. T. Vella, and B. Zhou. 2017. 'IRF6 Regulates Alternative Activation by Suppressing PPARgamma in Male Murine Macrophages', *Endocrinology*, 158: 2837-47.
Li, Peng, Wei Jiang, Qin Yu, Wang Liu, Ping Zhou, Jun Li, Junjie Xu, Bo Xu, Fengchao Wang, and Feng Shao. 2017. 'Ubiquitination and degradation of GBPs by a Shigella effector to suppress host defence', *Nature*, 551: nature24467.
Li, Qi, Gang Lu, GE Antonio, YT Mak, John A Rudd, Ming Fan, and David T Yew. 2007. 'The usefulness of the spontaneously hypertensive rat to model attention-deficit/hyperactivity disorder (ADHD) may be explained by the differential expression of dopamine-related genes in the brain', *Neurochemistry international*, 50: 848-57.
Li, W., S. He, Y. Zhou, Y. Li, J. Hao, X. Zhou, F. Wang, Y. Zhang, Z. Huang, Z. Li, H. H. Loh, P. Y. Law, and H. Zheng. 2014. 'Neurod1 modulates opioid antinociceptive tolerance via two distinct mechanisms', *Biol Psychiatry*, 76: 775-84.
Li, Y., D. Yu, W. Sheng, H. Song, Y. Li, and M. Dong. 2016. 'Co-expression of FOXL1 and PP2A inhibits proliferation inducing apoptosis in pancreatic cancer cells via promoting TRAIL and reducing phosphorylated MYC', *Oncol Rep*, 35: 2198-206.
Liberzon, A, C Birger, H Thorvaldsdottir, M Ghandi, JP Mesirov, and P Tamayo. "The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015; 1 (6): 417–25." In.
Lin, K. C., H. W. Park, and K. L. Guan. 2017. 'Regulation of the Hippo Pathway Transcription Factor TEAD', *Trends Biochem Sci*, 42: 862-72.
Liu, Qing-Rong, Tomas Drgon, Catherine Johnson, Donna Walther, Judith Hess, and George R Uhl. 2006. 'Addiction molecular genetics: 639,401 SNP whole genome association identifies many “cell adhesion” genes', *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 141: 918-25.
Liu, S., X. Wang, L. Pan, W. Wu, D. Yang, M. Qin, W. Jia, C. Xiao, F. Long, J. Ge, X. Liu, and Y. Zhu. 2017. 'Endogenous hydrogen sulfide regulates histone demethylase JMJD3-mediated inflammatory response in LPS-stimulated macrophages and in a mouse model of LPS-induced septic shock', *Biochem Pharmacol*.
Liu, Zhihui, Zijun Chen, Congping Shang, Fei Yan, Yingchao Shi, Jiajing Zhang, Baole Qu, Hailin Han, Yanying Wang, and Dapeng Li. 2017. 'IGF1-Dependent Synaptic Plasticity of Mitral Cells in Olfactory Memory during Social Learning', *Neuron*, 95: 106-22. e5.

Loos, Maarten, Jorn Staal, Anton NM Schoffelmeer, August B Smit, Sabine Spijker, and Tommy Pattij. 2010. 'Inhibitory control and response latency differences between C57BL/6J and DBA/2J mice in a Go/No-Go and 5-choice serial reaction time task and strain-specific responsivity to amphetamine', *Behavioural brain research*, 214: 216-24.

Lotrich, Francis E, Robert E Ferrell, Mordechai Rabinovitz, and Bruce G Pollock. 2009. 'Risk for depression during interferon-alpha treatment is affected by the serotonin transporter polymorphism', *Biological psychiatry*, 65: 344-48.

Love, M. I., W. Huber, and S. Anders. 2014. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, 15: 550.

Ma, X., W. Yan, H. Zheng, Q. Du, L. Zhang, Y. Ban, N. Li, and F. Wei. 2015. 'Regulation of IL-10 and IL-12 production and function in macrophages and dendritic cells', *F1000Res*, 4.

Majdak, P., P. J. Bucko, A. L. Holloway, T. K. Bhattacharya, E. K. DeYoung, C. N. Kilby, J. A. Zombeck, and J. S. Rhodes. 2014. 'Behavioral and pharmacological evaluation of a selectively bred mouse model of home cage hyperactivity', *Behav Genet*, 44: 516-34.

Majdak, Petra, John R Ossyra, Jessica M Ossyra, Adam J Cobert, Gabrielle C Hofmann, Stephen Tse, Brent Panozzo, Elizabeth L Grogan, Anastassia Sorokina, and Justin S Rhodes. 2016. 'A new mouse model of ADHD for medication development', *Scientific reports*, 6: 39472.

Malki, Karim, O Pain, MG Tosto, E Du Rietz, L Carboni, and LC Schalkwyk. 2015. 'Identification of genes and gene pathways associated with major depressive disorder by integrative brain analysis of rat and human prefrontal cortex transcriptomes', *Translational psychiatry*, 5: e519.

Malki, Karim, E Koritskaya, F Harris, K Bryson, M Herbster, and MG Tosto. 2016. 'Epigenetic differences in monozygotic twins discordant for major depressive disorder', *Translational psychiatry*, 6: e839.

Margoob, Mushtaq A, and Dhuha Mushtaq. 2011. 'Serotonin transporter gene polymorphism and psychiatric disorders: Is there a link?', *Indian journal of psychiatry*, 53: 289.

Mark, Karla A, Maria S Quinton, Shelley J Russek, and Bryan K Yamamoto. 2007. 'Dynamic changes in vesicular glutamate transporter 1 function and expression related to methamphetamine-induced glutamate release', *Journal of Neuroscience*, 27: 6823-31.

Martin, Alexander, Maria E Ochagavia, Laya C Rabasa, Jamilet Miranda, Jorge Fernandez-de-Cossio, and Ricardo Bringas. 2010. 'BisoGenet: a new tool for gene network building, visualization and analysis', *BMC bioinformatics*, 11: 91.

Martin, Jeffrey A, and Zhong Wang. 2011. 'Next-generation transcriptome assembly', *Nature reviews genetics*, 12: 671-82.

Martincuks, A., K. Andryka, A. Kuster, H. Schmitz-Van de Leur, M. Komorowski, and G. Muller-Newen. 2017. 'Nuclear translocation of STAT3 and NF-kappaB are independent of each other but NF-kappaB supports expression and activation of STAT3', *Cell Signal*, 32: 36-47.

Martínez-Rivera, A, J Hao, TF Tropea, TP Giordano, M Kosovsky, RC Rice, A Lee, RL Huganir, J Striessnig, and NA Addy. 2017. 'Enhancing VTA Cav1. 3 L-type Ca2+ channel activity promotes cocaine and mood-related behaviors via overlapping AMPA receptor mechanisms in the nucleus accumbens', *Molecular psychiatry*.

Martinez, Fernando O, and Siamon Gordon. 2014. 'The M1 and M2 paradigm of macrophage activation: time for reassessment', *F1000prime reports*, 6.

Martocchia, A., M. Curto, S. Scaccianoce, F. Comite, D. Xenos, C. Nasca, G. M. Falaschi, S. Ferracuti, P. Girardi, F. Nicoletti, and P. Falaschi. 2014. 'Effects of escitalopram on serum BDNF levels in elderly patients with depression: a preliminary report', *Aging Clin Exp Res*, 26: 461-4.

Masuo, Yoshinori, Masami Ishido, Masatoshi Morita, Hirofumi Sawa, Kazuo Nagashima, and Etsuo Niki. 2007. 'Behavioural characteristics and gene expression in the hyperactive wiggling (Wig) rat', *European Journal of Neuroscience*, 25: 3659-66.

Mathivanan, Suresh, Balamurugan Periaswamy, TKB Gandhi, Kumaran Kandasamy, Shubha Suresh, Riaz Mohmood, YL Ramachandra, and Akhilesh Pandey. 2006. 'An evaluation of human protein-protein interaction data in the public domain', *BMC bioinformatics*, 7: S19.

Mattes, CE, TJ Lynch, A Singh, RM Bradley, PA Kellaris, RO Brady, and KL Dretchen. 1997. 'Therapeutic use of butyrylcholinesterase for cocaine intoxication', *Toxicology and applied pharmacology*, 145: 372-80.
Mazei-Robison, MS, and RD Blakely. 2006. 'ADHD and the dopamine transporter: are there reasons to pay attention?', *Neurotransmitter transporters*: 373-415.
Memon, Shahar Bano, Li Lian, Javaid Ali Gadahi, and Wang Genlin. 2016. 'Proteomic response of mouse pituitary gland under heat stress revealed active regulation of stress responsive proteins', *Journal of thermal biology*, 61: 82-90.
Meng, Bo, Shujia Zhu, Shijia Li, Qingwen Zeng, and Bing Mei. 2009. 'Global view of the mechanisms of improved learning and memory capability in mice with music-exposure by microarray', *Brain research bulletin*, 80: 36-44.
Meredith, Gloria E, Shannon Callen, and Deborah A Scheuer. 2002. 'Brain-derived neurotrophic factor expression is increased in the rat amygdala, piriform cortex and hypothalamus following repeated amphetamine administration', *Brain research*, 949: 218-27.
Mereu, M, G Contarini, EF Buonaguro, G Latte, F Manago, F Iasevoli, A de Bartolomeis, and F Papaleo. 2017. 'Dopamine transporter (DAT) genetic hypofunction in mice produces dimensional alterations consistent with ADHD but not schizophrenia or bipolar disorder', *Neuropharmacology*.
Merino-Gracia, J., M. F. Garcia-Mayoral, and I. Rodriguez-Crespo. 2011. 'The association of viral proteins with host cell dynein components during virus infection', *FEBS J*, 278: 2997-3011.
Mikhaylova, Elena R, Vladimir F Lazarev, Alina D Nikotina, Boris A Margulis, and Irina V Guzhova. 2016. 'Glyceraldehyde 3-phosphate dehydrogenase augments the intercellular transmission and toxicity of polyglutamine aggregates in a cell model of Huntington disease', *Journal of neurochemistry*, 136: 1052-63.
Milanovic, M., N. Heise, N. S. De Silva, M. M. Anderson, K. Silva, A. Carette, F. Orelli, G. Bhagat, and U. Klein. 2017. 'Differential requirements for the canonical NF-kappaB transcription factors c-REL and RELA during the generation and activation of mature B cells', *Immunol Cell Biol*, 95: 261-71.
Minett, T., J. Classey, F. E. Matthews, M. Fahrenhold, M. Taga, C. Brayne, P. G. Ince, J. A. Nicoll, D. Boche, and Cfas Mrc. 2016. 'Microglial immunophenotype in dementia with Alzheimer's pathology', *J Neuroinflammation*, 13: 135.
Misiak, Magdalena, Shilpee Singh, Sascha Drewlo, Cordian Beyer, and Susanne Arnold. 2010. 'Brain region-specific vulnerability of astrocytes in response to 3-nitropropionic acid is mediated by cytochrome c oxidase isoform expression', *Cell and tissue research*, 341: 83-93.
Mitchell, AC, B Javidfar, V Pothula, D Ibi, EY Shen, CJ Peter, LK Bicks, T Fehr, Y Jiang, and KJ Brennand. 2017. 'MEF2C transcription factor is associated with the genetic and epigenetic risk architecture of schizophrenia and improves cognition in mice', *Molecular psychiatry*.
Mittleman, Barbara B, F Xavier Castellanos, Leslie K Jacobsen, Judith L Rapoport, Susan E Swedo, and Gene M Shearer. 1997. 'Cerebrospinal fluid cytokines in pediatric neuropsychiatric disease', *The Journal of Immunology*, 159: 2994-99.
Mizumori, S. J., C. B. Puryear, and A. K. Martig. 2009. 'Basal ganglia contributions to adaptive navigation', *Behav Brain Res*, 199: 32-42.
Morita, K., M. Noura, C. Tokushige, S. Maeda, H. Kiyose, G. Kashiwazaki, J. Taniguchi, T. Bando, K. Yoshida, T. Ozaki, H. Matsuo, S. Ogawa, P. P. Liu, T. Nakahata, H. Sugiyama, S. Adachi, and Y. Kamikubo. 2017. 'Autonomous feedback loop of RUNX1-p53-CBFB in acute myeloid leukemia cells', *Sci Rep*, 7: 16604.
Mortazavi, Ali, Brian A Williams, Kenneth McCue, Lorian Schaeffer, and Barbara Wold. 2008. 'Mapping and quantifying mammalian transcriptomes by RNA-Seq', *Nature methods*, 5: 621-28.
Muñiz, Javier, Gimena Gomez, Betina González, María Rivero-Echeto, Jean Cadet, Edgar García-Rill, Francisco Urbano, and Veronica Bisagno. 2016. 'Combined Effects of Simultaneous Exposure to Caffeine and Cocaine in the Mouse Striatum', *Neurotoxicity research*, 29.

Myint, Aye-Mu, Yong Ku Kim, Robert Verkerk, Simon Scharpé, Harry Steinbusch, and Brian Leonard. 2007. 'Kynurenine pathway in major depression: evidence of impaired neuroprotection', *Journal of affective disorders*, 98: 143-51.

Nakajima, Hidemitsu, Wataru Amano, Takeya Kubo, Ayano Fukuhara, Hideshi Ihara, Yasu-Taka Azuma, Hisao Tajima, Takashi Inui, Akira Sawa, and Tadayoshi Takeuchi. 2009. 'Glyceraldehyde-3-phosphate dehydrogenase aggregate formation participates in oxidative stress-induced cell death', *Journal of Biological Chemistry*, 284: 34331-41.

Nelson, Britta S, Michelle K Sequeira, and Jesse R Schank. 2017. 'Bidirectional relationship between alcohol intake and sensitivity to social defeat: association with Tacr1 and Avp expression', *Addiction biology*.

Nelson, P Austin, Jennifer R Sage, Suzanne C Wood, Christopher M Davenport, Stephan G Anagnostaras, and Lisa M Boulanger. 2013. 'MHC class I immune proteins are critical for hippocampus-dependent memory and gate NMDAR-dependent hippocampal long-term depression', *Learning & Memory*, 20: 505-17.

Nixon, Scott E, Dianelys González-Peña, Marcus A Lawson, Robert H McCusker, Alvaro G Hernandez, Jason C O'Connor, Robert Dantzer, Keith W Kelley, and Sandra L Rodriguez-Zas. 2015. 'Analytical workflow profiling gene expression in murine macrophages', *Journal of bioinformatics and computational biology*, 13: 1550010.

O'Connor, Jason C, Marcus A Lawson, Caroline André, Eileen M Briley, Sandra S Szegedi, Jacques Lestage, Nathalie Castanon, Miles Herkenham, Robert Dantzer, and Keith W Kelley. 2009. 'Induction of IDO by bacille Calmette-Guerin is responsible for development of murine depressive-like behavior', *The Journal of Immunology*, 182: 3202-12.

Oades, Robert D. 2011. 'An exploration of the associations of pregnancy and perinatal features with cytokines and tryptophan/kynurenine metabolism in children with attention-deficit hyperactivity disorder (ADHD)', *ADHD Attention Deficit and Hyperactivity Disorders*, 3: 301-18.

Oades, Robert D, Maria R Dauvermann, Benno G Schimmelmann, Markus J Schwarz, and Aye-Mu Myint. 2010. 'Attention-deficit hyperactivity disorder (ADHD) and glial integrity: S100B, cytokines and kynurenine metabolism-effects of medication', *Behavioral and brain functions*, 6: 29.

Oliveira-Nascimento, Laura, Paola Massari, and Lee M Wetzler. 2012. 'The role of TLR2 in infection and immunity', *Frontiers in immunology*, 3.

Orsini, Caitlin A, Barry Setlow, Michael DeJesus, Stacy Galaviz, Kimberly Loesch, Thomas Ioerger, and Deeann Wallis. 2016. 'Behavioral and transcriptomic profiling of mice null for Lphn3, a gene implicated in ADHD and addiction', *Molecular genetics & genomic medicine*, 4: 322-43.

Oyagi, A., Y. Oida, K. Kakefuda, M. Shimazawa, N. Shioda, S. Moriguchi, K. Kitaichi, D. Nanba, K. Yamaguchi, Y. Furuta, K. Fukunaga, S. Higashiyama, and H. Hara. 2009. 'Generation and characterization of conditional heparin-binding EGF-like growth factor knockout mice', *PLoS One*, 4: e7461.

Parrott, J. M., L. Redus, and J. C. O'Connor. 2016. 'Kynurenine metabolic balance is disrupted in the hippocampus following peripheral lipopolysaccharide challenge', *J Neuroinflammation*, 13: 124.

Pasetto, Laura, Silvia Pozzi, Mariachiara Castelnovo, Manuela Basso, Alvaro G Estevez, Stefano Fumagalli, Maria Grazia De Simoni, Valeria Castellaneta, Paolo Bigini, and Elena Restelli. 2017. 'Targeting extracellular cyclophilin a reduces neuroinflammation and extends survival in a mouse model of amyotrophic lateral sclerosis', *Journal of Neuroscience*, 37: 1413-27.

Patro, Rob, Stephen M Mount, and Carl Kingsford. 2014. 'Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms', *Nature biotechnology*, 32: 462-64.

Pavlopoulos, Georgios A, Anna-Lynn Wegener, and Reinhard Schneider. 2008. 'A survey of visualization tools for biological network analysis', *Biodata mining*, 1: 12.

Pavlovsky, Ashly A, Darren Boehning, Dingge Li, Yafang Zhang, Xiuzhen Fan, and Thomas A Green. 2013. 'Psychological stress, cocaine and natural reward each induce endoplasmic reticulum stress genes in rat brain', *Neuroscience*, 246: 160-69.

Perez, Enmanuel J, Maria L Cepero, Sebastian U Perez, Joseph T Coyle, Thomas J Sick, and Daniel J Liebl. 2016. 'EphB3 signaling propagates synaptic dysfunction in the traumatic injured brain', *Neurobiology of disease*, 94: 73-84.

Persico, Antonio M, Charles W Schindler, Bruce F O'Hara, Michael T Brannock, and George R Uhl. 1993. 'Brain transcription factor expression: effects of acute and chronic amphetamine and injection stress', *Molecular brain research*, 20: 91-100.

Phillips, C., M. A. Baktir, M. Srivatsan, and A. Salehi. 2014. 'Neuroprotective effects of physical activity on the brain: a closer look at trophic factor signaling', *Front Cell Neurosci*, 8: 170.

Phillips, Tamara J, Helen M Kamens, and Jeanna M Wheeler. 2008. 'Behavioral genetic contributions to the study of addiction-related amphetamine effects', *Neuroscience & Biobehavioral Reviews*, 32: 707-59.

Pittman, K. J., P. W. Cervantes, and L. J. Knoll. 2016. 'Z-DNA Binding Protein Mediates Host Control of Toxoplasma gondii Infection', *Infect Immun*, 84: 3063-70.

Podgorniak, Tomasz, Massimo Milan, Jose Marti Pujolar, Gregory E Maes, Luca Bargelloni, Eric De Oliveira, Fabien Pierron, and Francoise Daverat. 2015. 'Differences in brain gene transcription profiles advocate for an important role of cognitive function in upstream migration and water obstacles crossing in European eel', *BMC Genomics*, 16: 378.

Poggi, A., and M. R. Zocchi. 2014. 'NK cell autoreactivity and autoimmune diseases', *Front Immunol*, 5: 27.

Pruitt, Kim D, Tatiana Tatusova, and Donna R Maglott. 2006. 'NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins', *Nucleic acids research*, 35: D61-D65.

Put, Karen, Ellen Brisse, Anneleen Avau, Maya Imbrechts, Tania Mitera, Rik Janssens, Paul Proost, Francesca Fallarino, Carine H Wouters, and Patrick Matthys. 2016. 'IDO1 Deficiency does not affect disease in mouse models of systemic juvenile idiopathic arthritis and secondary hemophagocytic lymphohistiocytosis', *PLoS One*, 11: e0150075.

Qi, Jia, Jing-Yu Yang, Ming Song, Yan Li, Fang Wang, and Chun-Fu Wu. 2008. 'Inhibition by oxytocin of methamphetamine-induced hyperactivity related to dopamine turnover in the mesolimbic region in mice', *Naunyn-Schmiedeberg's archives of pharmacology*, 376: 441-48.

Reiner, Anat, Daniel Yekutieli, and Yoav Benjamini. 2003. 'Identifying differentially expressed genes using false discovery rate controlling procedures', *Bioinformatics*, 19: 368-75.

Ren, YQ, QH Li, and LB Liu. 2016. 'USF1 prompt melanoma through upregulating TGF-β signaling pathway', *European review for medical and pharmacological sciences*, 20: 3592-98.

Rhodes, Sinead M, David R Coghill, and Keith Matthews. 2004. 'Methylphenidate restores visual memory, but not working memory function in attention deficit-hyperkinetic disorder', *Psychopharmacology*, 175: 319-30.

Robinson, Mark D, Davis J McCarthy, and Gordon K Smyth. 2010. 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics*, 26: 139-40.

Rodriguez-Zas, Sandra L, Scott E Nixon, Marcus A Lawson, Robert H Mccusker, Bruce R Southey, Jason C O'Connor, Robert Dantzer, and Keith W Kelley. 2015. 'Advancing the understanding of behaviors associated with Bacille Calmette Guérin infection using multivariate analysis', *Brain, behavior, and immunity*, 44: 176-86.

Romanos, M, C Freitag, C Jacob, DW Craig, A Dempfle, TT Nguyen, R Halperin, S Walitza, TJ Renner, and C Seitz. 2008. 'Genome-wide linkage analysis of ADHD using high-density SNP arrays: novel loci at 5q13. 1 and 14q12', *Molecular psychiatry*, 13: 522-30.

Sagvolden, Terje, Vivienne A Russell, Heidi Aase, Espen Borgå Johansen, and Mehdi Farshbaf. 2005. 'Rodent models of attention-deficit/hyperactivity disorder', *Biological psychiatry*, 57: 1239-47.

Sarnyai, Zoltán, and Gábor L Kovács. 1994. 'Role of oxytocin in the neuroadaptation to drugs of abuse', *Psychoneuroendocrinology*, 19: 85-117.
Sato, F., A. Kohsaka, K. Takahashi, S. Otao, Y. Kitada, Y. Iwasaki, and Y. Muragaki. 2017. 'Smad3 and Bmal1 regulate p21 and S100A4 expression in myocardial stromal fibroblasts via TNF-alpha', *Histochem Cell Biol*, 148: 617-24.
Saul, MC, Petra Majdak, Samuel Perez, Matthew Reilly, T Garland, and Justin S Rhodes. 2017. 'High motivation for exercise is associated with altered chromatin regulators of monoamine receptor gene expression in the striatum of selectively bred mice', *Genes, Brain and Behavior*, 16: 328-41.
Saxton, Arnold Myron, and SAS Institute. 2004. *Genetic analysis of complex traits using SAS* (SAS Institute Cary, NC).
Saylor, Alicia J, and Jacqueline F McGinty. 2008. 'Amphetamine-induced locomotion and gene expression are altered in BDNF heterozygous mice', *Genes, Brain and Behavior*, 7: 906-14.
Schachtschneider, Kyle M, Yingkai Liu, Laurie A Rund, Ole Madsen, Rodney W Johnson, Martien AM Groenen, and Lawrence B Schook. 2016. 'Impact of neonatal iron deficiency on hippocampal DNA methylation and gene transcription in a porcine biomedical model of cognitive development', *BMC Genomics*, 17: 856.
Schenk, G. J., S. Dijkstra, A. J. van het Hof, S. M. van der Pol, J. A. Drexhage, P. van der Valk, A. Reijerkerk, J. van Horssen, and H. E. de Vries. 2013. 'Roles for HB-EGF and CD9 in multiple sclerosis', *Glia*, 61: 1890-905.
Schneider, M. A., G. A. Spoden, L. Florin, and C. Lambert. 2011. 'Identification of the dynein light chains required for human papillomavirus infection', *Cell Microbiol*, 13: 32-46.
Sengupta, Urmi, Sanchaita Ukil, Nevenka Dimitrova, and Shipra Agrawal. 2009. 'Expression-based network biology identifies alteration in key regulatory pathways of type 2 diabetes and associated risk/complications', *PLoS One*, 4: e8100.
Serao, Nick VL, Dianelys Gonzalez-Pena, Jonathan E Beever, German A Bollero, Bruce R Southey, Daniel B Faulkner, and Sandra L Rodriguez-Zas. 2013. 'Bivariate genome-wide association analysis of the growth and intake components of feed efficiency', *PLoS One*, 8: e78530.
Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. 'Cytoscape: a software environment for integrated models of biomolecular interaction networks', *Genome research*, 13: 2498-504.
Shimada, Takahiro, Bong Goo Park, Andrea J Wolf, Constantinos Brikos, Helen S Goodridge, Courtney A Becker, Christopher N Reyes, Edward A Miao, Alan Aderem, and Friedrich Götz. 2010. 'Staphylococcus aureus evades lysozyme-based peptidoglycan digestion that links phagocytosis, inflammasome activation, and IL-1β secretion', *Cell host & microbe*, 7: 38-49.
Shiow, Lawrence R, Kenneth Paris, Matthew C Akana, Jason G Cyster, Ricardo U Sorensen, and Jennifer M Puck. 2009. 'Severe combined immunodeficiency (SCID) and attention deficit hyperactivity disorder (ADHD) associated with a Coronin-1A mutation and a chromosome 16p11. 2 deletion', *Clinical immunology*, 131: 24-30.
Siddiqui, Irum Javaid, Nashaiman Pervaiz, and Amir Ali Abbasi. 2016. 'The Parkinson Disease gene SNCA: Evolutionary and structural insights with pathological implication', *Scientific reports*, 6: 24475.
Soneson, Charlotte, Michael I Love, and Mark D Robinson. 2015. 'Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences', *F1000Research*, 4.
Song, G. G., J. H. Kim, and Y. H. Lee. 2013. 'Genome-wide pathway analysis in major depressive disorder', *J Mol Neurosci*, 51: 428-36.
Spohrer, S., R. Gross, L. Nalbach, L. Schwind, H. Stumpf, M. D. Menger, E. Ampofo, M. Montenarh, and C. Gotz. 2017. 'Functional interplay between the transcription factors USF1 and PDX-1 and protein kinase CK2 in pancreatic beta-cells', *Sci Rep*, 7: 16367.
Stankiewicz, Adrian M, Joanna Goscik, Alicja Majewska, Artur H Swiergiel, and Grzegorz R Juszczak. 2015. 'The effect of acute and chronic social stress on the hippocampal transcriptome in mice', *PLoS One*, 10: e0142195.

Stuart, M. J., G. Singhal, and B. T. Baune. 2015. 'Systematic Review of the Neurobiological Relevance of Chemokines to Psychiatric Disorders', *Front Cell Neurosci*, 9: 357.
Subramanian, Aravind, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, and Eric S Lander. 2005. 'Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles', *Proceedings of the National Academy of Sciences*, 102: 15545-50.
Südhof, Thomas C. 2008. 'Neuroligins and neurexins link synaptic function to cognitive disease', *Nature*, 455: 903-11.
Sugaya, Nagisa, Toru Kobayashi, and Kazutaka Ikeda. 2013. 'Role of GIRK Channels in Addictive Substance Effects', *Journal of Drug and Alcohol Research*: 122-32.
Sultana, Rukhsana, Debra Boyd-Kimball, H Fai Poon, Jian Cai, William M Pierce, Jon B Klein, Michael Merchant, William R Markesbery, and D Allan Butterfield. 2006. 'Redox proteomics identification of oxidized proteins in Alzheimer's disease hippocampus and cerebellum: an approach to understand pathological and biochemical alterations in AD', *Neurobiology of aging*, 27: 1564-76.
Sun, Ye, Chi-Hsiu Liu, John Paul SanGiovanni, Lucy P Evans, Katherine T Tian, Bing Zhang, Andreas Stahl, William T Pu, Theodore M Kamenecka, and Laura A Solt. 2015. 'Nuclear receptor RORα regulates pathologic retinal angiogenesis by modulating SOCS3-dependent inflammation', *Proceedings of the National Academy of Sciences*, 112: 10401-06.
Suzuki, H., J. Savitz, T. Kent Teague, S. K. Gandhapudi, C. Tan, M. Misaki, B. A. McKinney, M. R. Irwin, W. C. Drevets, and J. Bodurka. 2017. 'Altered populations of natural killer cells, cytotoxic T lymphocytes, and regulatory T cells in major depressive disorder: Association with sleep disturbance', *Brain Behav Immun*, 66: 193-200.
Svahn, Sara L, Leif Väremo, Britt G Gabrielsson, Eduard Peris, Intawat Nookaew, Louise Grahnemo, Ann-Sofie Sandberg, Ingrid Wernstedt Asterholm, John-Olov Jansson, and Jens Nielsen. 2016. 'Six tissue transcriptomics reveals specific immune suppression in spleen by dietary polyunsaturated fatty acids', *PLoS One*, 11: e0155099.
Swallow, John G, Patrick A Carter, and Theodore Garland. 1998. 'Artificial selection for increased wheel-running behavior in house mice', *Behavior genetics*, 28: 227-37.
Szabo, Steven T, Rodrigo Machado-Vieira, Peixiong Yuan, Yun Wang, Yanling Wei, Cynthia Falke, Chiara Cirelli, Giulio Tononi, Husseini K Manji, and Jing Du. 2009. 'Glutamate receptors as targets of protein kinase C in the pathophysiology and treatment of animal models of mania', *Neuropharmacology*, 56: 47-55.
Tamura, Hideki, Sadao Shiosaka, and Shota Morikawa. 2017. 'Trophic modulation of gamma oscillations: The key role of processing protease for Neuregulin-1 and BDNF precursors', *Neurochemistry international*.
Tan, Andrew, Rosario Moratalla, Gregory L Lyford, Paul Worley, and Ann M Graybiel. 2000. 'The Activity-Regulated Cytoskeletal-Associated Protein Arc Is Expressed in Different Striosome-Matrix Patterns Following Exposure to Amphetamine and Cocaine', *Journal of neurochemistry*, 74: 2074-78.
Tang, Y Tom, Peter Emtage, Walter D Funk, Tianhua Hu, Matthew Arterburn, Emily EJ Park, and Fabio Rupp. 2004. 'TAFA: a novel secreted family with conserved cysteine residues and restricted expression in the brain', *Genomics*, 83: 727-34.
Tang, Yu, Zun-Ping Ke, Yi-Gen Peng, and Ping-Tai Cai. 'Coexpression Analysis Reveals Key Gene Modules and Pathway of Human Coronary Heart Disease', *Journal of Cellular Biochemistry*.
Thiels, Edda, Beatriz I Kanterewicz, Eric D Norman, James M Trzaskos, and Eric Klann. 2002. 'Long-term depression in the adult hippocampus in vivo involves activation of extracellular signal-regulated kinase and phosphorylation of Elk-1', *Journal of Neuroscience*, 22: 2054-62.
Trapnell, Cole, Lior Pachter, and Steven L Salzberg. 2009. 'TopHat: discovering splice junctions with RNA-Seq', *Bioinformatics*, 25: 1105-11.

Trapnell, Cole, Adam Roberts, Loyal Goff, Geo Pertea, Daehwan Kim, David R Kelley, Harold Pimentel, Steven L Salzberg, John L Rinn, and Lior Pachter. 2012. 'Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks', *Nature protocols*, 7: 562-78.

Trojan, Ewa, Joanna Slusarczyk, Katarzyna Chamera, Katarzyna Kotarska, Katarzyna Głombik, Marta Kubera, Wladyslaw Lason, and Agnieszka Basta-Kaim. 2017. 'The Modulatory Properties of Chronic Antidepressant Drugs Treatment on the Brain Chemokine–Chemokine Receptor Network: A Molecular Study in an Animal Model of Depression', *Frontiers in pharmacology*, 8: 779.

Tsai, Shih-Jen. 2016. 'Role of neurotrophic factors in attention deficit hyperactivity disorder', *Cytokine & Growth Factor Reviews*.

Ufimtseva, Elena. 2015. 'Mycobacterium-host cell relationships in granulomatous lesions in a mouse model of latent tuberculous infection', *BioMed research international*, 2015.

Uhl, George R, Tomas Drgon, Catherine Johnson, Oluwatosin O Fatusin, Qing-Rong Liu, Carlo Contoreggi, Chuan-Yun Li, Kari Buck, and John Crabbe. 2008. '"Higher order" addiction molecular genetics: convergent data from genome-wide association in humans and mice', *Biochemical pharmacology*, 75: 98-111.

Upton, J. W., and W. J. Kaiser. 2017. 'DAI Another Way: Necroptotic Control of Viral Infection', *Cell Host Microbe*, 21: 290-93.

Valera, Elvira, Richard Dargusch, Pamela A Maher, and David Schubert. 2013. 'Modulation of 5-lipoxygenase in proteotoxicity and Alzheimer's disease', *Journal of Neuroscience*, 33: 10512-25.

van Dam, S., U. Vosa, A. van der Graaf, L. Franke, and J. P. de Magalhaes. 2017. 'Gene co-expression analysis for functional classification and gene-disease predictions', *Brief Bioinform.*

Vance, R. E. 2010. 'Inflammasome activation: how macrophages watch what they eat', *Cell Host Microbe*, 7: 3-5.

Verfaillie, Annelien, Hana Imrichová, Bram Van de Sande, Laura Standaert, Valerie Christiaens, Gert Hulselmans, Koen Herten, Marina Naval Sanchez, Delphine Potier, and Dmitry Svetlichnyy. 2014. 'iRegulon: from a gene list to a gene regulatory network using large motif and track collections', *PLoS computational biology*, 10: e1003731.

Virakul, S., T. Phetsuksiri, C. van Holten-Neelen, B. Schrijver, L. van Steensel, V. A. Dalm, D. Paridaens, W. A. van den Bosch, P. M. van Hagen, and W. A. Dik. 2016. 'Histamine induces NF-kappaB controlled cytokine secretion by orbital fibroblasts via histamine receptor type-1', *Exp Eye Res*, 147: 85-93.

Wadhwa, M., P. Kumari, G. Chauhan, K. Roy, S. Alam, K. Kishore, K. Ray, and U. Panjwani. 2017. 'Sleep deprivation induces spatial memory impairment by altered hippocampus neuroinflammatory responses and glial cells activation in rats', *J Neuroimmunol*, 312: 38-48.

Wan, Fang-Jung, Hui-Ching Lin, Kun-Lun Huang, Ching-Jiunn Tseng, and Chih-Shung Wong. 2000. 'Systemic administration of d-amphetamine induces long-lasting oxidative stress in the rat striatum', *Life sciences*, 66: 205-12.

Wang, Xian-Feng, Hong-Sheng Wang, Hao Wang, Fan Zhang, Ke-Fang Wang, Qiang Guo, Ge Zhang, Shao-Hui Cai, and Jun Du. 2014. 'The role of indoleamine 2, 3-dioxygenase (IDO) in immune tolerance: focus on macrophage polarization of THP-1 cells', *Cellular immunology*, 289: 42-48.

Wang, Y., M. Subramanian, A. Yurdagul, Jr., V. C. Barbosa-Lorenzi, B. Cai, J. de Juan-Sanz, T. A. Ryan, M. Nomura, F. R. Maxfield, and I. Tabas. 2017. 'Mitochondrial Fission Promotes the Continued Clearance of Apoptotic Cells by Macrophages', *Cell*, 171: 331-45 e22.

Wang, Zhong, Mark Gerstein, and Michael Snyder. 2009. 'RNA-Seq: a revolutionary tool for transcriptomics', *Nature reviews genetics*, 10: 57-63.

Webb, Katharine J, William HJ Norton, Dietrich Trümbach, Annemarie H Meijer, Jovica Ninkovic, Stefanie Topp, Daniel Heck, Carsten Marr, Wolfgang Wurst, and Fabian J Theis. 2009. 'Zebrafish reward mutants reveal novel transcripts mediating the behavioral effects of amphetamine', *Genome biology*, 10: R81.

Weiss, Friedbert, David J Tanzer, and Aaron Ettenberg. 1988. 'Opposite actions of CCK-8 on amphetamine-induced hyperlocomotion and stereotypy following intracerebroventricular and intra-accumbens injections in rats', *Pharmacology Biochemistry and Behavior*, 30: 309-17.
Weiss, Norbert, Alejandro Sandoval, Shigeki Kyonaka, Ricardo Felix, Yasuo Mori, and Michel De Waard. 2011. 'Rim1 modulates direct G-protein regulation of CaV2. 2 channels', *Pflügers Archiv-European Journal of Physiology*, 461: 447-59.
Weyandt, Lisa L, Danielle R Oster, Marisa E Marraccini, Bergljot Gyda Gudmundsdottir, Bailey A Munro, Brynheld Martinez Zavras, and Ben Kuhar. 2014. 'Pharmacological interventions for adolescents and adults with ADHD: stimulant and nonstimulant medications and misuse of prescription stimulants', *Psychology research and behavior management*, 7: 223-49.
Wiener, C. D., F. P. Moreira, T. A. Cardoso, T. C. Mondin, P. V. da Silva Magalhaes, F. Kapczinski, L. D. de Mattos Souza, R. A. da Silva, J. P. Oses, and K. Jansen. 2017. 'Inflammatory cytokines and functional impairment in drug-free subjects with mood disorder', *J Neuroimmunol*, 307: 33-36.
Willcutt, Erik G. 2012. 'The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review', *Neurotherapeutics*, 9: 490-99.
Williams, Julia L, and Jason M White. 1984. 'The effects of amphetamine and scopolamine on adjunctive drinking and wheel-running in rats', *Psychopharmacology*, 82: 360-67.
Wilmot, Beth, Rebecca Fry, Lisa Smeester, Erica D Musser, Jonathan Mill, and Joel T Nigg. 2016. 'Methylomic analysis of salivary DNA in childhood ADHD identifies altered DNA methylation in VIPR2', *Journal of Child Psychology and Psychiatry*, 57: 152-60.
Won, Hyejung, Won Mah, and Eunjoon Kim. 2013. 'Autism spectrum disorder causes, mechanisms, and treatments: focus on neuronal synapses', *Frontiers in molecular neuroscience*, 6.
Wright, C., J. H. Shin, A. Rajpurohit, A. Deep-Soboslay, L. Collado-Torres, N. J. Brandon, T. M. Hyde, J. E. Kleinman, A. E. Jaffe, A. J. Cross, and D. R. Weinberger. 2017. 'Altered expression of histamine signaling genes in autism spectrum disorder', *Transl Psychiatry*, 7: e1126.
Wu, Junfang, Zaorui Zhao, Boris Sabirzhanov, Bogdan A Stoica, Alok Kumar, Tao Luo, Jacob Skovira, and Alan I Faden. 2014. 'Spinal cord injury causes brain inflammation associated with cognitive and affective changes: role of cell cycle pathways', *Journal of Neuroscience*, 34: 10989-1006.
Wynn, Thomas A, Ajay Chawla, and Jeffrey W Pollard. 2013. 'Origins and Hallmarks of macrophages: development, homeostasis, and disease', *Nature*, 496: 445.
Yager, Lindsay M, Aaron F Garcia, Amanda M Wunsch, and Susan M Ferguson. 2015. 'The ins and outs of the striatum: role in drug addiction', *Neuroscience*, 301: 529-41.
Yang, Amy, Emma Childs, Abraham A Palmer, and Harriet De Wit. 2010. 'More on ADORA', *Psychopharmacology*, 212: 699-700.
Yang, Y., W. Shen, Y. Ni, Y. Su, Z. Yang, and C. Zhao. 2017. 'Impaired Interneuron Development after Foxg1 Disruption', *Cereb Cortex*, 27: 793-808.
Yang, Yisheng, and Megan J Wilson. 2015. 'Lhx9 gene expression during early limb development in mice requires the FGF signalling pathway', *Gene Expression Patterns*, 19: 45-51.
Yao, Jihua, Jingyao Zhou, Qiaoling Liu, Daru Lu, Lu Wang, Xiaojing Qiao, and William Jia. 2010. 'Atoh8, a bHLH transcription factor, is required for the development of retina and skeletal muscle in zebrafish', *PLoS One*, 5: e10945.
Yarosh, Haley L, Shashwath A Meda, Harriet De Wit, Amy B Hart, and Godfrey D Pearlson. 2015. 'Multivariate analysis of subjective responses to d-amphetamine in healthy volunteers finds novel genetic pathway associations', *Psychopharmacology*, 232: 2781.
Yen, Yi-Chun, Elmira Anderzhanova, Mirjam Bunck, Julia Schuller, Rainer Landgraf, and Carsten T Wotjak. 2013. 'Co-segregation of hyperactivity, active coping styles, and cognitive dysfunction in mice selectively bred for low levels of anxiety', *Frontiers in behavioral neuroscience*, 7.
Yen, Yi-Chun, Nils C Gassen, Andreas Zellner, Theo Rein, Rainer Landgraf, Carsten T Wotjak, and Elmira Anderzhanova. 2015. 'Glycogen synthase kinase-3β inhibition in the medial prefrontal cortex mediates paradoxical amphetamine action in a mouse model of ADHD', *Frontiers in behavioral neuroscience*, 9.

Young, Kimberly A, Yan Liu, Kyle L Gobrogge, Hui Wang, and Zuoxin Wang. 2014. 'Oxytocin reverses amphetamine-induced deficits in social bonding: evidence for an interaction with nucleus accumbens dopamine', *Journal of Neuroscience*, 34: 8499-506.

Yuan, Fang-fen, Xue Gu, Xin Huang, Yan Zhong, and Jing Wu. 2017. 'SLC6A1 gene involvement in susceptibility to attention-deficit/hyperactivity disorder: A case-control study and gene-environment interaction', *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 77: 202-08.

Zhang, Bin, and Steve Horvath. 2005. 'A general framework for weighted gene co-expression network analysis', *Statistical applications in genetics and molecular biology*, 4.

Zhang, Chi, Baohong Zhang, Lih-Ling Lin, and Shanrong Zhao. 2017. 'Evaluation and comparison of computational tools for RNA-seq isoform quantification', *BMC Genomics*, 18: 583.

Zhang, Yuan, Qian Yu, Tian-bao Lai, Yang Yang, Gang Li, and Sheng-gang Sun. 2013. 'Effects of small interfering RNA targeting sphingosine kinase-1 gene on the animal model of Alzheimer's disease', *Journal of Huazhong University of Science and Technology [Medical Sciences]*, 33: 427-32.

Zhao, Chunjie, Carmen Avilés, Regina A Abel, C Robert Almli, Patrick McQuillen, and Samuel J Pleasure. 2005. 'Hippocampal and visuospatial learning defects in mice with a deletion of frizzled 9, a gene in the Williams syndrome deletion interval', *Development*, 132: 2917-27.

Zhong, J., H. Wang, J. Yu, J. Zhang, and H. Wang. 2017. 'Overexpression of Forkhead Box L1 (FOXL1) Inhibits the Proliferation and Invasion of Breast Cancer Cells', *Oncol Res*, 25: 959-65.

Zhou, Hong-Ming, Gisela Weskamp, Valérie Chesneau, Umut Sahin, Andrea Vortkamp, Keisuke Horiuchi, Riccardo Chiusaroli, Rebecca Hahn, David Wilkes, and Peter Fisher. 2004. 'Essential role for ADAM19 in cardiovascular morphogenesis', *Molecular and cellular biology*, 24: 96-104.

Zhu, S., Z. A. Cordner, J. Xiong, C. T. Chiu, A. Artola, Y. Zuo, A. D. Nelson, T. Y. Kim, N. Zaika, B. M. Woolums, E. J. Hess, X. Wang, D. M. Chuang, M. M. Pletnikov, P. M. Jenkins, K. L. Tamashiro, and C. A. Ross. 2017. 'Genetic disruption of ankyrin-G in adult mouse forebrain causes cortical synapse alteration and behavior reminiscent of bipolar disorder', *Proc Natl Acad Sci U S A*, 114: 10479-84.

Zombeck, Jonathan A, Erin K DeYoung, Weronika J Brzezinska, and Justin S Rhodes. 2011. 'Selective breeding for increased home cage physical activity in collaborative cross and Hsd: ICR mice', *Behavior genetics*, 41: 571-82.