Transcription factor motifs associated with anterior insula gene-expression in mood disorder Brains

Dhivya Arasappan[1]; Simon B. Eickhoff[2,3]; Charles B Nemeroff[4,5,6]; Hans A. Hofmann[7,8]; Mbemba Jabbi[4,5,7,9,*]

[1]Center for Biomedical Research Support, University of Texas at Austin

[2]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Germany

[3]Institute of Neuroscience and Medicine (INM-7), Research Centre Jülich, Germany

[4]Department of Psychiatry, Dell Medical School, University of Texas at Austin

[5]The Mulva Clinic for Neurosciences, Dell Medical School, University of Texas at Austin

[6]Institute of Early Life Adversity Research

[7]Institute of Neuroscience, University of Texas at Austin

[8]Department of Integrative Biology, the University of Texas at Austin

[9]Department of Psychology, University of Texas at Austin

*__Corresponding Author:__ mbemba.jabbi@austin.utexas.edu

Keywords: Brain, gene expression, transcription factors, behavior, mood disorders, RNA-sequencing

ABSTRACT

Background: Mood disorders represent a major cause of morbidity and mortality worldwide but their pathophysiology remains obscure. Gene expression profiles, especially with regards to the interactive influence of genes within the expression profiles underlying pathological brain phenotypes associated with mood disorders remains largely undefined.

Methods: Because the anterior insula is reduced in volume in patients with mood disorders, RNA was extracted from postmortem mood disorder samples and compared with unaffected control samples for RNA-sequencing identification of differentially expressed genes (DEGs) in *a*) bipolar disorder (BD; n=37), and *b*) major depressive disorder (MDD n=30) vs controls (n=33), and *c*) low vs high Axis-I comorbidity (a measure of psychiatric morbidity burden). Given the regulatory role of transcription factors (TFs) in gene expression via specific-DNA-binding domains (motifs), we used JASPAR TF binding database to identify TF-motifs.

Results: We found that DEGs in BD-vs-controls, MDD-vs-controls, and high-vs-low Axis-I comorbidity burden were associated with TF-motifs known to regulate expression of toll-like receptor signaling genes, cellular homeostatic-control genes, and genes involved in embryonic, cellular, organ and brain developmental processes.

Discussion*: Robust image-guided transcriptomics* was applied by targeting the gray matter volume reduction in the anterior insula in mood disorders, to guide independent postmortem identification of TF motifs regulating DEG. TF motifs were identified for immune, cellular, embryonic and neurodevelopmental processes.

Conclusion: Our results provided novel information about the hierarchical relationship between gene regulatory networks, the TFs that control them, and proximate underlying neuroanatomical phenotypes in mood disorders.

INTRODUCTION

Neurobiological blueprints for adaptive behavior are governed by genes, especially their coding regions, by mediating molecular processes critical for brain development and functions. These, in turn, bridge the functional gene expression effects on complex behavioral phenotypes. Transcription factors (TFs) are sequence-specific DNA-binding proteins that are also known as regulatory proteins. TFs regulate the expression of genes by recognizing and binding to specific DNA regulatory elements called DNA binding domains in the promoter region of genes (Changeaux 2017; Lambert et al. 2018). This attribute enables TFs to hierarchically regulate the expression of genes by controlling (i.e. promoting/activating or blocking/repressing) transcription of the adjacent structural genes into mRNA, and subsequently proteins (Changeaux 2017; Lambert et al. 2018). TFs are therefore in a functional elevated status in the hierarchy of gene expression repertoires because they cooperatively, or synergistically regulate genes encoding other TFs (Lambert et al. 2018). As such, TFs are thereby able to control transcription of their own and related structural genes (Kerszberg & Changeaux 1998; Changeaux 2017). Consequently, nonlinear relationships arise from this functional complexity among genes, and intricate gene-gene interactions often create functionally co-operative gene assemblies of varying hierarchical gene expression programs essential for biological processes (Kerszberg & Changeaux 1998; Tsigelny et al. 2013).

Based on the complex but crucial role of thousands of transcription factors in controlling gene regulation, dysregulation of TF-mediated gene expression programs have been hypothesized to contribute to a broad range of diseases (Lee & Young 2013), including neuropsychiatric disorders (Lee & Young 2013; Nord AS, Pattabiraman K, Visel A & Rubinstein JLR 2015; Changeaux 2017; Torre-Ubieta et al. 2018). The putative role of TF networks in

mood disorder pathogenesis is however not well defined. Because mood disorders are the most common neuropsychiatric syndromes (Yizhar 2012), and they collectively account for a high global burden of disease (Collins et al. 2010), the role of TFs in anterior insula gene expression profiles associated with affective dysfunction was examined.

Adult brain gray matter volumetric reductions in the anterior insula have been consistently identified in mood disorders (Goodkind et al. 2015; Wise et al. 2016). Such reduced gray matter volumes are predictive of cognitive impairment and affective dysfunction in both major depressive disorder and bipolar disorder (McTeague et al. 2017). Functional magnetic resonance imaging (fMRI) of the right anterior insula region has been reported to selectively predict the therapeutic response to psychotherapy and pharmacotherapy (McGrath et al. 2013). The current study tested the hypothesis that specific transcription factor gene expression profiles will be associated with gene expression profiles in the postmortem anterior insula cortex of mood disorder subjects relative to controls. The results indicate that relative to controls, DEGs in postmortem mood disorder tissue are associated with TF motifs known to regulate expression of toll-like receptor signaling genes, cellular homeostatic control genes, and genes involved in cellular and brain

developmental processes.

METHODS

Localization of brain gray matter loss in mood disorders

To identify the brain's most anatomically proximate regional involvement in mood pathology, with the goal of targeting the identified pathological sites for postmortem transcriptomics, a large scale meta-analysis of voxel based morphometry studies of gray matter loss in mood disorders

was first performed using the anatomical likelihood estimation approach. This approach models the spatial uncertainty associated with each reported location for significant between-group differences (Eickhoff et al. 2009), and further compute the convergence across all included experiments by the union of the ensuing probabilistic model relative to a null-distribution and thereby reflecting a random spatial association between the findings of different experiments (Eickhoff et al. 2012). The brain region exhibiting the most statistically significant gray matter loss associated with mood disorders using the anatomical likelihood estimation was the right anterior insula cortex gray matter (***Figure. 1A***) (Jabbi et al. unpublished observation). The identified anterior insula region exhibiting gray matter loss in mood disorder brain imaging cohorts was used for our *robust image-guided transcriptomics* by guiding dissection of the imaging identified reduced anterior insula volume in the independent postmortem mood disorder cohorts towards the goal of realizing a more anatomically precise RNA-seq study of the putative pathological tissue associated with mood disorder diagnoses.

Postmortem variable factor-analysis

To explore morbidity related gene expression profiles beyond classic case vs control comparisons, we conducted a factor analysis to identify higher-order composite variables included with the postmortem data. For each postmortem sample, the donor data includes specific mood disorder and comorbid lifetime-Axis-I comorbidity (i.e. number of lifetime-Axis-I diagnostic occurrences, e.g. (poly)-substance use disorders, psychosis, anxiety, eating disorders, etc. alongside the primary mood disorder diagnosis of BD or MDD); comorbid lifetime-Axis-III diagnoses (i.e. number of lifetime-Axis-III medical conditions such as diabetes, cancer, cardiovascular disease, etc.); cause of death (i.e. death by suicide, homicides, accidents or natural

death due to comorbid Axis-III medical conditions); and as specified by the medical examiner reports (e.g. blunt force trauma to the chest, gunshot, motor vehicle accident, drowning, hanging, etc.); demographics (race, age at death, sex, years of education, number of children/fecundity, and marital records); technical variables (brain-weight, postmortem interval 'i.e. the time that has elapsed since a person has died', pH, and RNA integrity number (RIN)); and toxicology (blood alcohol/blood narcotics levels) (see **Supplementary Table 1** & **Supplementary Table 2** for diagnostic sampling, demographics and postmortem qualitative data). Principal Axis Factoring (Oblimin Rotation with Kaizer Normalization) (Costello & Osborne 2005) was applied to identify higher-order factors explaining the differences in postmortem variables and included those variables with communalities of ≥ 0.5. The identified higher order factors included Axis-I comorbidity burden and suicide completion (here referred to as postmortem tissue quality measures, and age-related variables (i.e. age at death, marital status, number of children, and number of chronic medical diagnoses). Given our focus on identifying TF motifs for mood disorder comorbidity burden, we first conducted DEG for bipolar disorder vs controls, major depressive disorder vs controls, as well as high vs low Axis-I comorbidity by conducting a split-half comparison of the lower half scoring versus the higher scoring donors on this higher order variable representing Axis-I comorbidity.

Brain Dissection: The NIMH Human Brain Collection Core (HBCC) provided the *Postmortem* samples for which informed consents are acquired according to NIH IRB guidelines. Clinical characterization, neuropathology screening, and toxicology analyses followed previous protocols (Martin et al. 2006). We applied *robust image-guided transcriptomics* as follows: first, the region of interest targeted for dissection was defined as portion of right AIC encompassing the

identified reduced GMV in the completed meta-analysis by the authors. Electronic image slices of the imaging-defined GMV loss volumes were then shared with the HBCC neuropathologist who used these images to guide dissection of coronal tissue slabs of each postmortem donor brain at the NIH clinical center. The dissected regional volume corresponded to the anterior portion of the insula where the caudate and putamen are approximately equal in size and tissue was dissected from this section for each brain for use in RNA-sequencing (***Figure. 1A-D***). The region of interest results of gray matter volume reduction in right anterior insula previously observed (Jabbi et al. unpublished observations) to be associated with mood disorder diagnoses in the imaging meta-analysis (***Figure. 1A***, with rendered images showed on sagittal slide and 3-D reconstructed capture in ***Figure. 1B*)** was reconstructed manually in ITKSnap (http://www.itksnap.org/pmwiki/pmwiki.php) to guide independent postmortem dissection for RNA-sequencing.

RNA-Extraction: All dissected tissues were separately pulverized and 50 mg aliquoted from each sample for standardized total RNA processing. Specifically, RNeasy Lipid Tissue Mini Kit (50) was used for RNA purification using the 50 RNeasy Mini Spin Columns, Collection Tubes (1.5 ml and 2 ml), QIAzol Lysis Reagent, RNase-free Reagents and Buffers kit from Qiagen. DNase treatment was applied to the purified RNA using Qiagen RNase-Free DNase Set (50) kit consisting of 1500 Kunitz units RNase-free DNase I, RNase-free Buffer RDD, and RNase-free water for 50 RNA minipreps. After DNAse treatment, the purified RNA from the pulverized AIC tissue sections were used to determine RNA quality as measured in RNA integrity number (RIN) values using Agilent 6000 RNA Nano Kit consisting of the microfluidic chips, Agilent 6000 RNA Nano ladder and reagents on Agilent 2100 Bioanalyzer.

Samples with RIN < 6 were excluded and the 100 samples meeting inclusion were shipped directly from the NIMH HBCC core to the Genome Sequencing and Analysis Facility (GSAF: https://wikis.utexas.edu/display/GSAF/Home+Page) at the University of Texas, Austin, USA for RNA-sequencing.

Illumina-Sequencing, Read-Mapping and Gene-Quantification: Total RNA was extracted and only samples with RNA integrity numbers (RIN values) greater than 6 as confirmed using the Agilent Bioanayler were used for library preparation. First. Ribosomal RNA was depleted using RiboMinus Eukaryote kit from Life Technologies (Foster City, CA, USA) for RNA-Seq and confirmed using an Agilent Technologies' Bioanalyzer (Santa Clara, CA, USA). mRNA selection was completed using the Poly(A) purist kit from Thermofisher and paired-end libraries with average insert sizes of 200bp were obtained using NEBNext Ultra II Directional RNAs Library Prep kit from New England BioLabs. All 100 samples were processed and then sequenced on the Illumina HiSeq 4000, PE150, at the Genome Sequencing and Analysis Facility at UT Austin, USA.

30 million paired-end reads per sample (150 base pairs in length) were generated by sequencing runs of 4 samples per lane of the sequencer. Sequenced reads were assessed for quality with Fastqc (Andrews 2010) to specifically assess sequencing reads for median base quality, average base quality, sequence duplication, over-represented sequences and adapter contamination.

Differential Gene Expression Analysis: The reads were pseudo-aligned to the human reference transcriptome (GRCh38- gencode) using kallisto (Bray 2016), and gene-level abundances were

obtained. The abundances were normalized using DESeq2, and transformed with variance stabilizing transformation (defined here as a transformation or normalization technique that seeks to create more homoscedasticity, and thereby having a closer to constant variance in the dataset regardless of the mean expression value (Love et al. Genome Biology 2014). We performed differential expression analysis based on the negative binomial distribution for modeled gene counts using DESeq2. RIN-values were included in the DESeq2 design matrix as a covariate to control for the potentially confounding effects of RNA quality. The analysis controlled for group factors as well as possible individual outliers by removing genes with expression values of 0 in 80% or more of the samples. We performed differential expression analysis to identify DEGs in the following comparisons: BD vs controls; MDD vs controls, and in a pooled cohort of BD & MDD individuals with high vs low Axis-I comorbidity at absolute fold change >=2 and adjusted p-values <=0.1 were selected as significantly differentially expressed genes. The DESeq2 package by default calculates false discovery rate adjusted p-values according to Benjamini and Hochberg (Benjamin & Hochberg 1995). We used an adjusted p-value <=.1 as a cut-off to balance type-1 and type-2 error rates and allow more inclusive capture of regulatory elements/TFs associated with DEGs in for mood disorders across bipolar disorder and unipolar depression.

JASPAR 2018 TFs binding profiling: Given the role of transcription factors (TFs) in regulating gene expression via specific DNA-binding domains (motifs) in the gene promoters, we followed-up on our previous study (Jabbi et al. unpublished observations) that identified whole transcriptome gene expression profiles in mood disorders, with the aim to explore in more detail, mood disorder transcriptomics by identifying TF motifs. To this goal, we used JASPAR TF

binding database to identify motifs that were associated with DEGs in BD vs controls; MDD vs controls, and in pooled mood disorder individuals with high vs low Axis-I comorbidity.

JASPAR is an open access database of non-redundant, manually curated TF-binding profiles provided as a publicly available web framework (Khan A et al. 2018). The JASPAR 2018 version has an updated list of 579 non-redundant TF-binding profiles of the vertebrate taxonomy that are systematically stored as position frequency matrices (PFMs), which summarizes experimentally determined DNA sequences bound by an individual TFs by counting the number of occurrences of each nucleotide at each position within aligned TF-binding sites (Khan A et al. 2018). This JASPAR 2018 CORE vertebrate database of 579 PFMs was first used to predict TF-binding sites in the human genome, and then made available to the scientific community through the UCSC Genome Browser track data hub (http://jaspar.genereg.net/genome-tracks/) for use to identify specific TF-binding profiles.

EnrichR (Kuleshov et al. 2016) was used to identify TF motifs that were associated with our DEGs. We focused on the top 10 most significant TF motifs found in the database of 579 PFMs associated with the DEGs observed in our DESeq2 results dataset for each of the 3 analytical contrasts (i.e. BD vs controls; MDD vs controls; pooled mood disorder cohort with high Axis-I comorbidity vs pooled mood disorder cohort with low Axis-I comorbidity). TF motifs within the JASPAR 2018 that were found to be associated with DEGs at adjusted p-value cutoff of 0.1 (i.e. using the false discovery rate Benjamini and Hochberg multiple comparison method (Benjamini & Hochberg 1995) were identified as identified TFs associated with each of the 3 DEG contrasts (i.e. BD vs controls; MDD vs controls; low vs high Axis-I comorbidity) we selected the top 10 TF motifs (see **Figures 2&3**).

RESULTS

Postmortem Variability

Factor analysis identified two higher-order factors representing 1) number of Axis-1 diagnoses (e.g. substance use disorders or abuse/psychosis/anxiety, etc.), referred to here as *morbidity* and suicide-completion referred to as *mortality;* and 2) RIN values which were subsequently included in all differential gene expression analyses as covariates. Apart from race, no other demographic variables differed across groups [MDD & BD samples had more Caucasian donors whereas controls had more African American donors ($p<0.0001$, $F=12$)]. Covarying for race in subsequent ANOVAs, *Axis-I comorbidity burden* which clustered with mood disorder donor sample *suicide completion rate* in the explorative factor analysis was different across groups ($p<0.0001$, $F=30$); showing Bonferroni corrected pairwise-comparison differences between MDD > controls ($p<0.0001$); BD > controls ($p<0.0001$); and BD > MDD ($p = 0.004$).

DEGs and JASPAR TF motif identification of mood disorder group-related differential gene expression

We found 456 differentially expressed genes for the bipolar vs controls (**Supplementary Table 2**) and 2722 differentially expressed genes for major depressive disorder vs controls (**Supplementary Table 4**) in our analyses. Further analysis that includes DEG results from more than one brain region in the same mood disorder samples will help explain if our observed higher number of DEGs in major depressive vs control samples relative to bipolar disorder vs controls is pathophysiologically relevant in a region-specific manner. By first directly comparing TF motifs found in BD vs control DEGs, we found DEGs with associated motifs of *MIB2*, *Esrrb*, *TEAD2*, *ZBTB7A*, *MIR210*, *MZF1*, *KLF11*, *SMAD4*, *ZNF148*, and *TEAD4* TFs (***Figure. 2A***).

Direct comparisons between MDD vs controls DEGs identified associated motifs of *MIR210*, *TEAD4*, *SP4*, *TEAD2*, *TCFAP2A*, *SP1*, *SP3*, *PCBP1*, *HNF1B*, and *Esrrb* TFs (***Figure. 2B***).

Given our approach to examine the hierarchical relationship between TF motifs and DEGs profiles associated with the degree of psychiatric morbidity/Axis-I comorbidity in the two mood disorder cohorts, we first used principal axis factoring to identify *Axis I comorbidity (i.e. total number of psychiatric diagnoses),* and *manner of death by suicide or non-suicide* (together comprising the factor we refer to here as the Axis-I comorbidity burden) to compositely explain variability in (i) Axis-I comorbidity burden, and (ii) suicide completion. We included this factor in our analysis using a split half approach of high vs low Axis-I comorbidity burden to identify TF motifs for DEG profiles for Axis-I comorbidity burden. We found DEGs in the pooled BD & MDD donor samples (i.e. without control donors) and these DEGs were associated with TF motifs of *NFATC2*, *GABPA*, *HMAGA1*, *NR3C1*, *GTF2I*, *IRF2*, *POU1F1*, *SAMD9L*, *SNAI1*, and *CBEPB* (***Figure. 3***). Collectively, these TFs are known to regulate expression of toll-like receptor signaling genes, cellular homeostatic control genes, and genes involved in embryonic, cellular and neurodevelopmental processes.

Specifically, comparing BD vs controls revealed differentially expressed genes with associated motifs for TFs involved in regulating *a*) immune response and antigen processing (i.e. *MIB2*), *b*) master transcriptional repression and activation of a wide range of genes (i.e. *ZBTB7A*), *c*) Zinc Finger transcriptional control and cellular developmental processes and apoptosis (i.e. *ZBTB7A*, *MZF1*, *KLF11*, *ZNF148*), *d*) post-translational modification of gene expression and ion channel transporter expression (i.e. *SMAD4*, *MIR210*, *Esrrb*) and *e*) Hippo-signaling pathway, organ size control and regulation of cell proliferation and apoptosis (i.e. *TEAD2*, *TEAD4*) (***Figure 2A***). Comparing MDD vs controls revealed differentially expressed

genes with associated motifs for TFs involved in regulating *a*) post-translational modification of gene expression and ion channel transporter expression (i.e. *MIR210*, *Esrrb*) and *b*) Hippo-signaling pathway, organ size control and regulation of cell proliferation and apoptosis (i.e. *TEAD2*, *TEAD4*), *c*) Zinc Finger G Protein-Coupled receptor binding TF activity (i.e. *SP4, SP3, SP1*) implicated in mood disorders (Zhou et al. 2009; Shi et al. 20011; Shyn et al. 2011), *d*) TF super families involved in embryonic development, activation or repression of transcriptional activity of different sets of genes and RNA-binding (i.e. *TFAP2A, PCBP1*, *HNF1B*) (***Figure 2B***). Restricting our analysis to comparing mood disorder donors with high Axis-I comorbidity vs mood disorder donors scoring low on Axis-I comorbidity, we found differentially expressed genes with associated motifs for TFs involved in regulating *a*) inflammatory and immune response and toll-like receptor signaling, *b*) embryonic and cellular developmental processes, and *c*) cellular and peripheral homeostatic control (see ***Figure 3*** for specific regulatory functions for each of the identified TFs). When comparing DEGs in high vs low Axis-I comorbidity in mood disorders, we found 3 motifs for master TFs implicated in regulation an array of functions including general transcriptional regulation through interaction with basal transcriptional machinery (i.e. *G2FTI*), master transcriptional co-activation of several other TFs (i.e. *CREBBP*), and a master TF that regulates inflammatory, cellular developmental, and glucocorticoid gene expression processes (i.e. *NR3C1*) (see ***Figure 3*** for specific regulatory functions for each of the identified TFs).

DISCUSSION

Understanding the molecular substrates for regional brain abnormalities underlying major mental illness such as mood disorders, which affects over 10% of the population across the lifetime, will

not only contribute to better mechanistic understanding of the neurobiological bases for behavioral pathologies (Nestler 2015; Sahin & Sur Science 2015; Torre-Ubieta et al. 2018; Jabbi & Nemeroff 2019) but will also be essential for novel drug design (Changeaux 2017).

In the present study, we applied *robust image-guided transcriptomics* by extracting RNA-samples from a precisely localized anterior insula region exhibiting the strongest degree of gray matter loss in mood disorders. This *robust image-guided transcriptomics* study therefore integrated RNA-seq characterization of gene expression abnormalities in an anatomically abnormal region in mood disorders and allowed testing of the hypothesis that transcription factor motifs will be associated with differential gene expression profiles in the postmortem anterior insula cortex of mood disorder subjects relative to controls. We found that comparing BD vs controls revealed DEGs with associated motifs for TFs involved in regulating immune response and antigen processing, master transcriptional repression/activation of a wide range of genes, Zinc Finger transcriptional control of cellular developmental processes, post-translational modification of gene expression and ion channel transporter expression and Hippo-signaling pathway involved in organ size control and regulation of cellular development. Comparing MDD vs controls revealed DEGs with associated motifs for TFs involved in regulating post-translational modification of gene expression and ion channel transporter expression and Hippo-signaling pathway, Zinc Finger G Protein-Coupled receptor binding TF activity that has been implicated in mood disorders (Zhou et al. 2009; Shi et al. 20011; Shyn et al. 2011), and TF super families involved in embryonic and cellular development. DEGs in high vs low Axis-I comorbidity revealed motifs for master TFs implicated in the regulation of an array of functions including general transcriptional regulation through interaction with basal transcriptional machinery, and in inflammatory, cellular developmental, and glucocorticoid gene expression

processes. Of interest, the two comparisons between BD vs controls and MDD vs controls identified TF motifs of similar pathways including four TFs (i.e. *MIR120*, *TEAD2*, *TEAD4* and *Esrrb*) suggesting possible specific as well as overlapping TF motifs for BD & MDD. If replicated, the identified TF motifs associated with DEGs in mood disorder diagnoses and overall disease morbidity will likely reveal a vulnerability in the general transcriptional pathway mechanism in mood disorder disease states.

The identified DEG-profile associated TF motifs are known to regulate expression of toll-like receptor signaling genes, cellular homeostatic control genes, and embryonic and cellular including neurodevelopmental gene networks found to be differentially expressed in BD vs controls, MDD vs controls, and in high Axis-I comorbid mood disorder individuals vs low Axis-I comorbidity mood disorder individuals. We found putative hierarchical TF regulatory involvement in the gene expression landscapes associated with BD and MDD diagnoses, such that enrichments of master transcription factors were predominantly associated with the gene expression landscape in elevated mood disorder morbidity, thereby echoing previous work (Tsigelny et al. 2013; Lee & Young 2013; Changeaux 2017; Lambert et al. 2018; Chen C et al. 2019).

Traditional analyses of differential gene expression often assumes that genes of similar average expression strengths will likely have similar dispersion without considering that different gene classes such as TFs that are hierarchically programmed to functionally regulate other gene classes (Love et al. 2014; Anders & Huber 2010). Our current identification of mood disorder DEG associated TF motifs known to regulate the expression of toll-like receptor signaling pathway genes, cellular homeostatic control genes, and embryonic as well as cellular and neurodevelopment pathway genes in mood disorder diagnostic phenotypes, suggests a

putative hierarchical involvement of the identified TFs in the neurobiological abnormalities underlying mood disorder phenotypes.

Our current *robust image-guided transcriptomics* approach illustrates associations rather than direct testing of the causal involvements of the identified TF motifs in *1*) the developmental or disease-driven emergence of the anatomical gray matter loss identified in the anterior insula of mood disorder individuals, and *2*) the observed postmortem differential gene expression landscape measured within the anterior insula in mood disorder morbidity. Anatomical abnormalities in the anterior insula are not specific to mood disorder diagnoses per se, but involved in several neuropsychiatric diagnoses including anxiety, psychosis, substance use and eating disorders (Goodkind et al. 2016; Janiri et al. 2019), suggesting the importance of including other more mood disorder-specific regional anatomical brain abnormalities in future transcriptomics studies. Furthermore, our approach to target the anterior insula gray region that exhibited a reduction in volume in living mood disorder individuals for tissue dissection and subsequent RNA-sequencing in an independent postmortem sample, does not address if the reduced insula volume was present in the currently studied postmortem mood disorder sample. Future studies need to explore integrating anatomical and gene expression abnormalities in the same postmortem brain samples across brain regions to directly link anatomical abnormalities with underlying molecular genetic regulatory abnormalities in the same postmortem brains.

In light of the limited number of samples for a study of complex biological variables, and the fact that the TF data are secondary derivations from the differential gene expression analyses, the reported results needs to be considered as exploratory findings that could be confounded by factors between cases and controls, despite our attempts to control for such factors. Our bulk tissue RNA-sequencing approach does not account for the role of specific cell-types in defining

TF regulatory involvements in mood disorder transcriptomics. Therefore, exploration of cell-type-specific regional brain analysis in mood disorders across the lifespan will also advance our understanding of the cell-type specific role of the identified TFs in the pathogenesis of mood disorders. Although our current study identified relevant TF involvement in regulatory gene expression signatures for mood disorders, inclusion of a comparison psychiatric cohort without mood symptoms such as schizophrenia in future studies will help address the specificity of the identified TF targets for mood disorder therapeutics. These limitations notwithstanding, our current focus on the anterior insula given its identification as the region exhibiting the strongest statistically significant gray matter loss associated with mood disorder diagnoses, provides a framework for integrative studies of anatomical and molecular abnormalities underlying prevalent brain diseases.

In conclusion, we applied *robust image-guided transcriptomics* to characterize the roles of specific TFs in phenotypic mood disorder morbidity. Our study provides a first delineation of gene pathways, and the TFs that may regulated them, underlying psychiatric diseases like mood disorders. Our results suggest that postmortem gene expression patterns are associated with (1) the magnitude of mood disorder morbidity, in a putative abnormal brain anatomically compromised region in mood disorder diagnoses; and (2) specific TF motif patterns known to regulate general transcriptional processes, immune responsive, cellular homeostatic control, embryonic and neuronal development functions. Together, these findings illustrate that studies of gene regulatory networks, in addition to studying the functional profiles of the genes these TF regulatory networks control/transcribe, has the potential to elucidate the hierarchical organizational principles of gene expression landscapes driving major psychiatric disorders and thereby accelerate novel pharmacological target discovery.

Author Contributions: MJ conceived and designed the studies, and acquired postmortem material from the NIMH HBCC. DA, SBE, and MJ performed the experiments and analyzed the data and results. DA and MJ drafted the manuscript and SBE, CBN, and HH contributed critically and substantially to the content/writing of the manuscript and interpretation of the findings.

Acknowledgements: The NIMH Human Brain Collection Core provided RNA-samples for donors and we thank the NIMH and Drs. Barbara Lipska, Stefano Marenco, Pavan Auluck and HBCC colleagues for the studied samples. We thank Wade Weber of Dell Medical School Psychiatry Department, UT Austin for assistance in preparing the manuscript, Dr. Mark Bond of Dell Medical School Psychiatry Department, UT Austin for statistical reviews, and Jessica Podnar and several GSAF colleagues for RNA-seq support.

Funding: This work was supported by the Dell Medical School, UT Austin Mulva Neuroscience Clinics Startup funds for MJabbi; DA and CBN are supported by the National Institutes of Health (NIH) and HHofmann is supported by NSF-DEB 1638861 and NSF-IOS 1326187. The funding bodies nor any other entities were involved in the design of the study and collection, analysis, and interpretation of data presented here.

Competing Interests:
Mbemba Jabbi, none.

Dhivya Arasappan, none.

Simon Eickhoff none.

Hans Hofmann, none.

Charles B Nemeroff:
Research/Grants: National Institutes of Health (NIH)

Consulting (*last three years*): Xhale, Takeda, Taisho Pharmaceutical Inc., Signant Health,, Sunovion Pharmaceuticals Inc., Janssen Research & Development LLC, Magstim, Inc., Navitor Pharmaceuticals, Inc., TC MSO, Inc., Intra-Cellular Therapies, Inc., EMA Wellness, Gerson Lehrman Group (GLG), Acadia Pharmaceuticals

Stockholder: Xhale, Celgene, Seattle Genetics, Abbvie, OPKO Health, Inc., Antares, BI Gen Holdings, Inc., Corcept Therapeutics Pharmaceuticals Company, TC MSO, Inc., Trends in Pharma Development, LLC, EMA Wellness

Scientific Advisory Boards: American Foundation for Suicide Prevention (AFSP), Brain and Behavior Research Foundation (BBRF), Xhale, Anxiety Disorders Association of America (ADAA), Skyland Trail, Signant Health, Laureate Institute for Brain Research (LIBR), Inc.

Board of Directors: Gratitude America, ADAA, Xhale Smart, Inc.
Income sources or equity of $10,000 or more: American Psychiatric Publishing, Xhale, Signant Health, CME Outfitters, Intra-Cellular Therapies, Inc., Magstim, EMA Wellness

Patents: Method and devices for transdermal delivery of lithium (*US 6,375,990B1*)
Method of assessing antidepressant drug therapy via transport inhibition of monoamine neurotransmitters by ex vivo assay (*US 7,148,027B2*)
Compounds, Compositions, Methods of Synthesis, and Methods of Treatment (CRF Receptor Binding Ligand) (*US 8,551, 996 B2*)

Speakers Bureau: None

REFERENCES

Anders S, Huber W 2010. "Differential expression analysis for sequence count data." Genome Biology. 11, R106; 10.1186/gb-2010-11-10-r106, http://genomebiology.com/2010/11/10/R106/.

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data; http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bray, N., Pimentel, H., Melsted, P. et al. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34:525–527

Benjamini, Yoav; Hochberg, Yosef. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B. 1995; 57(1):289–300.

Changeux JP. Climbing Brain Levels of Organisation from Genes to Consciousness. Trends Cogn Sci. 2017;21:168-181.

Chen C et a. The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. Sci Transl Med. 2018;10(472).

Craig AD. How do you feel--now? The anterior insula and human awareness. Nat Rev Neurosci. 2009;10:59-70.

Collins PY et a. Grand challenges in global mental health. Nature. 2011;475(7354):27-30.

Costello AB and Osborne JW. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Practical Assessment, Research & Evaluation. 2005;10:1-9.

de la Torre-Ubieta L, Stein JL, Won H, Opland CK, Liang D, Lu D et al. The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis. Cell. 2018;172(1-2):289-304.e18.

Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum Brain Mapp. 2009;30:2907-26.

Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT. Activation likelihood estimation meta-analysis revisited. Neuroimage. 2012;59:2349-61.

Goodkind M et al. Identification of a common neurobiological substrate for mental Illness. JAMA Psychiat. 2015;72:305–315.

Harris RM, Hofmann HA. Neurogenomics of behavioral plasticity. Adv Exp Med Biol. 2014; 781:149-68.

Jabbi M, Kippenhan JS, Kohn P, Marenco S, Mervis CB, Morris CA et al. The Williams syndrome chromosome 7q11.23 hemideletion confers hypersocial, anxious personality coupled with altered insula structure and function. Proc Natl Acad Sci U S A. 2012;109:E860-6.

Jabbi M, Chen Q, Turner N, Kohn P, White M, Kippenhan JS et al. Variation in the Williams syndrome GTF2I gene and anxiety proneness interactively affect prefrontal cortical response to aversive stimuli. Transl Psychiatry. 2015;5:e622.

Jabbi M, Arasappan D, Eickhoff SB, Strakowski SM, Nemeroff CB, Hofmann HA (in revision). See preprint: Neuro-transcriptomic signatures for mood disorder morbidity and suicide mortality. https://www.biorxiv.org/content/10.1101/762492v1.

Jabbi M, Nemeroff CB. Convergent neurobiological predictors of mood and anxiety symptoms and treatment response. Expert Rev Neurother. 2019;19:587-597.

Janiri D, Moser DA, Doucet GE, Luber MJ, Rasgon A, Lee WH, Murrough JW, Sani G, Eickhoff SB, Frangou S. Shared Neural Phenotypes for Mood and Anxiety Disorders: A Meta-analysis of 226 Task-Related Functional Imaging Studies. JAMA Psychiatry. 2019:1-8.

Kessler RC et al. Prevalence, Severity, and Comorbidity of Twelve-month DSM-IV Disorders in the National Comorbidity Survey Replication (NCS- R). Archives of General Psychiatry. 2005;62:617–627.

Kerszberg M, Changeux JP. A simple molecular model of neurulation. Bioessays. 1998;20:758-70.

Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 2018; 46:D1284.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016; 44:W90-7.

Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M. The Human Transcription Factors. Cell. 2018;175(2):598-599.

Lipska BK, Deep-Soboslay A, Weickert CS, Hyde TM, Martin CE, Herman MM et al. Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. Biol Psychiatry. 2006;60(6):650-8.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE, Franco AR. Toward a neuroimaging treatment selection biomarker for major depressive disorder. JAMA Psychiatry. 2013;70(8):821-9.

McTeague LM, Huemer J, Carreon DM, Jiang Y, Eickhoff SB, Etkin A. Identification of Common Neural Circuit Disruptions in Cognitive Control Across Psychiatric Disorders. Am J Psychiatry. 2017;174:676-685

Murray CJL et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis of Global Burden of Disease Study 2010. The Lancet. 2012;380:2197–2223.

Nestler EJ. Role of the Brain's Reward Circuitry in Depression: Transcriptional Mechanisms. Int Rev Neurobiol. 2015;124:151-70.

Nord AS, Pattabiraman K, Visel A, Rubenstein JLR Genomic perspectives of transcriptional regulation in forebrain development. Neuron. 2015;85(1):27-47.
O'Connell LA, Hofmann HA. Evolution of a vertebrate social decision-making network. Science. 2012;336:1154-7.

Rose EJ, Donohoe G . Brain vs behavior: an effect size comparison of neuroimaging and cognitive studies of genetic risk for schizophrenia. Schizophr Bull 2013;39:518–526.

Sahin M, Sur M. Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. Science. 2015;350(6263).

Shi J et al. Genome-wide association study of recurrent early-onset major depressive disorder. Mol Psychiatry. 2011; 16:193-201.

Shyn SI et al. Novel loci for major depression identified by genome-wide association study of SequencedTreatment Alternatives to Relieve Depression and meta-analysis of three studies. Mol Psychiatry. 2011;16:202-15.

Tsigelny IF, Kouznetsova VL, Baitaluk M, Changeux JP. A hierarchical coherent-gene-group model for brain development. Genes Brain Behav. 2013;12:147-65.

Wise T et al. Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: evidence from voxel-based meta-analysis. Molecular Psychiatry. 2016;22: 1455–1463.

Yizhar O. Optogenetic insights into social behavior function. Biol Psychiatry. 2012;71:1075-80.

Zhou X, Tang W, Greenwood TA, Guo S, He L, Geyer MA, Kelsoe JR. Transcription factor SP4 is a susceptibility gene for bipolar disorder. PLoS One. 2009;4:e5196.

FIGURE & TABLE LEGENDS

Figure 1 A-B. Gray matter loss in anterior insula cortex of BD and MDD vs controls as identified with largescale voxel-based morphometry imaging meta-analysis. A) illustrates localized anterior insula gray matter reduced area associated with mood disorder diagnoses on a coronal section; and B) illustrates the sagittal reconstruction of the reduced volume in A as well as the 3-dimentional view of the reduced volume. The anatomical information in A and B illustrated slide by slide on coronal sections in C) and these images are used to guide postmortem dissection of anterior insula tissue from an independent sample in D.

Figure 2. Anterior Insula postmortem TF motifs for gene expression profiles in BD and MDD relative to controls. A) Gene expression profiles in BD vs Controls are illustrated to be associated with the top ten TF motifs for genes that are implicated in regulating inflammatory and immune responses, early embryonic and cellular development, and in post transcriptional gene expression. B) Gene expression profiles in MDD vs Controls shown to be associated with the top ten TF motifs known to regulate inflammatory and toll-like receptor signaling, cellular development and peripheral homeostatic/hormonal signaling and post transcriptional regulation/DNA methylation. C) Venn diagram for TF motifs specific for BD vs controls, MDD vs controls and the overlapping TFs motifs for which were found for both BD vs controls and MDD vs controls.

Figure 3. List of TF genes showing JASPAR 2018 identified associated motifs for gene expression profiles for high compared to low Axis-I comorbidity burden in BD and MDD.

The identified TFs include putative master transcriptional regulators (NR3C1, GTF2I and CBEPB) as well as regulators of cytokine gene expression (NFATC2 and IRF2), viral mediated functional immune-related gene expression (GABPA and HMGA1), and in regulating early embryonic (SNAI1), cellular (SAMD9L) and hormonal development (POU1F1).

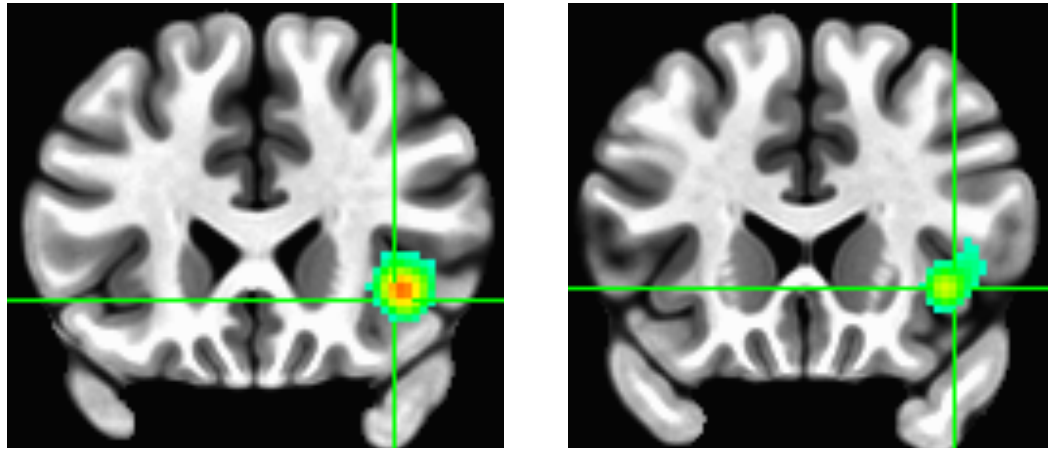
Supplementary Table 2 Legend

Abbreviations: MDD, major depressive disorder; GMV, gray matter volume; AIC, anterior insula cortex; PMI, postmortem index; Ph, measure of acidity; RIN, RNA integrity number which is a measure of RNA quality.

Supplementary Table 3-5 Legend

Abbreviations: baseMean, normalized read counts of all samples; log2Foldchange, effect size estimate; lfcSE, standard error of the log2foldchange; stat, Wald statistical test values; pvalue, uncorrected p-value; p-adj, corrected p-value. Fold change is calculated as the ratio of mean expression in high psychiatric morbidity & suicide mortality score to the mean expression in low psychiatric morbidity & suicide mortality score.

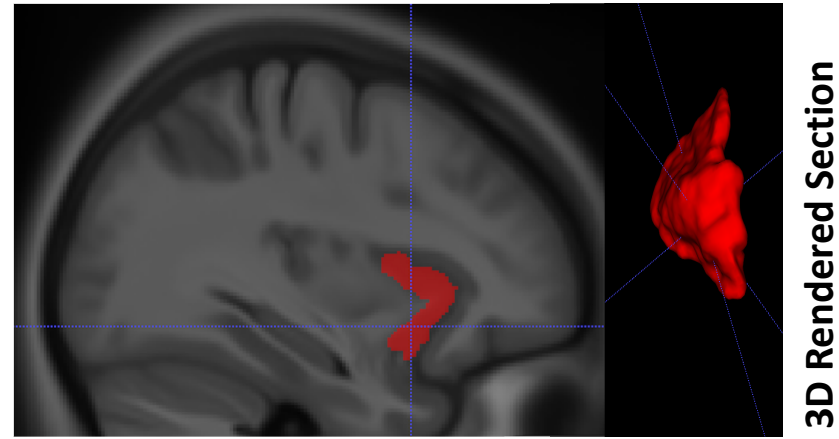
A



Coronal View

Anterior Insula Region of Interest VBM Meta-analysis Results of Reduced Volume in Mood Disorders (Bipolar and Major Depression)

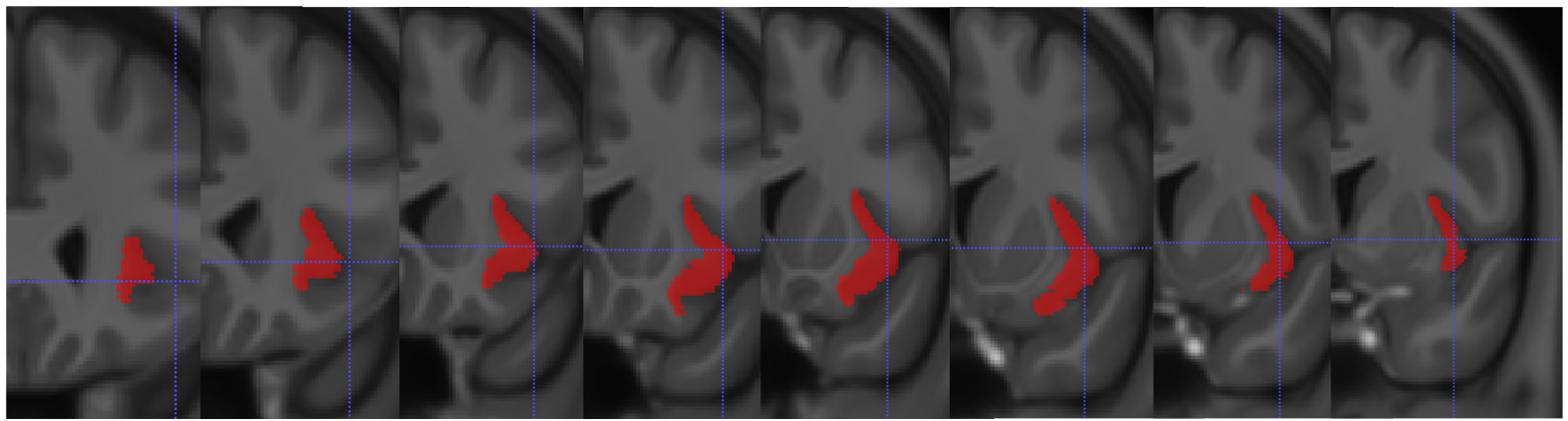
B



Sagittal View

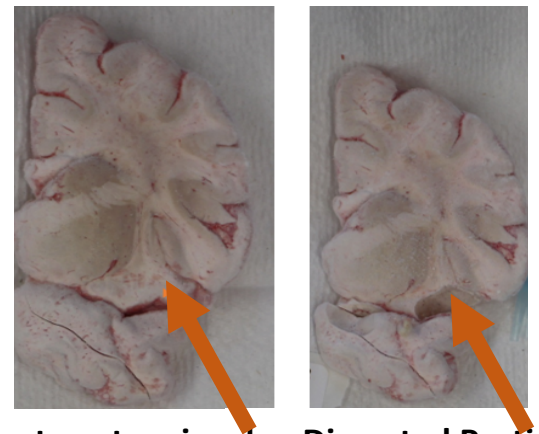
Anterior Insula Region of Interest Demarcated Reduced Volume in Mood Disorders (Bipolar and Major Depression)

C



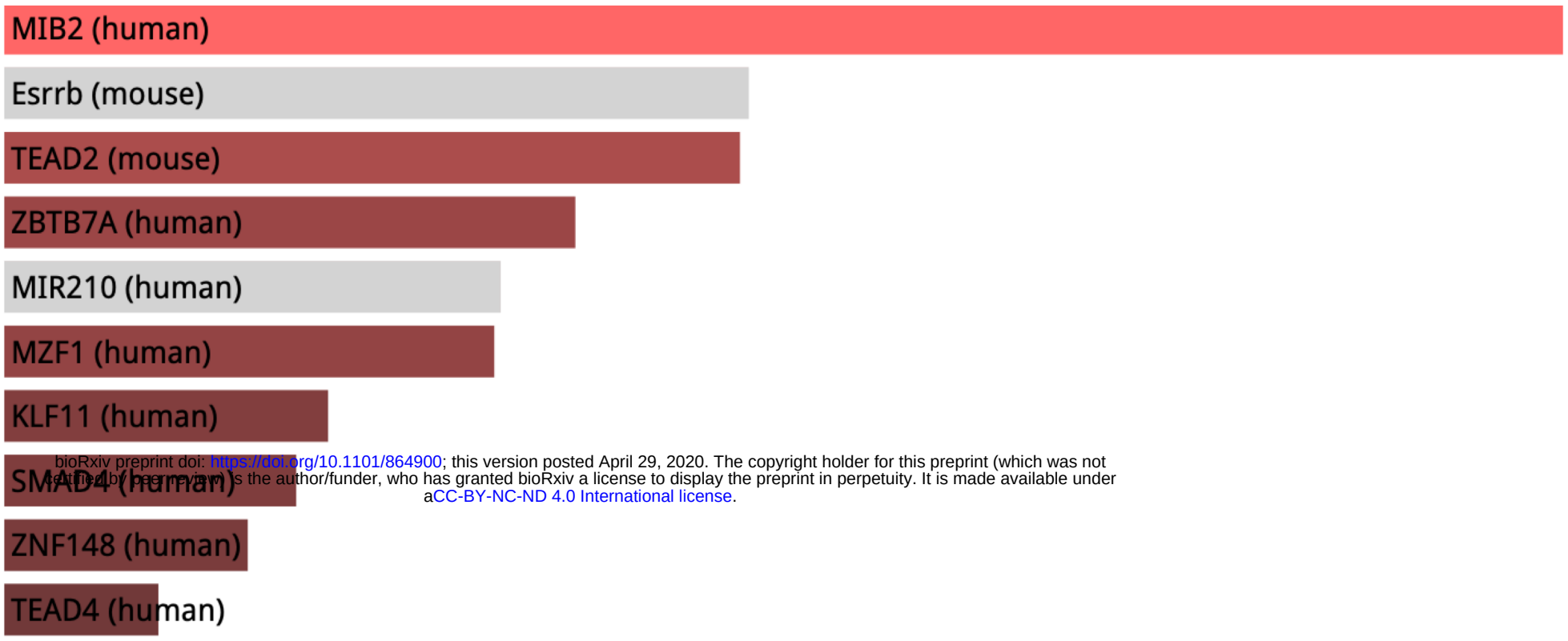
Coronal Dissection Locale of Anterior Insula used for postmortem dissection provided to HBCC Neuropathologist

D

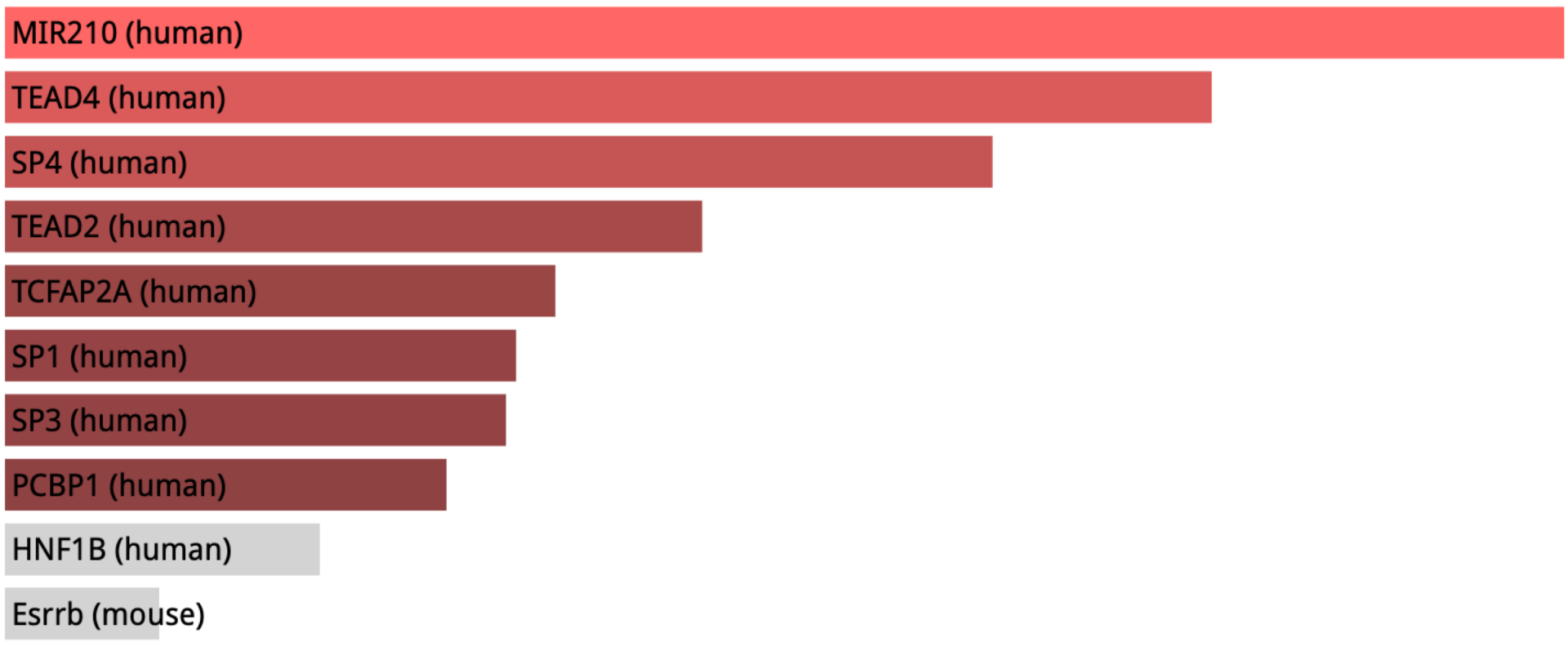


Anterior Insula TF Motifs in Mood Disorders vs Controls

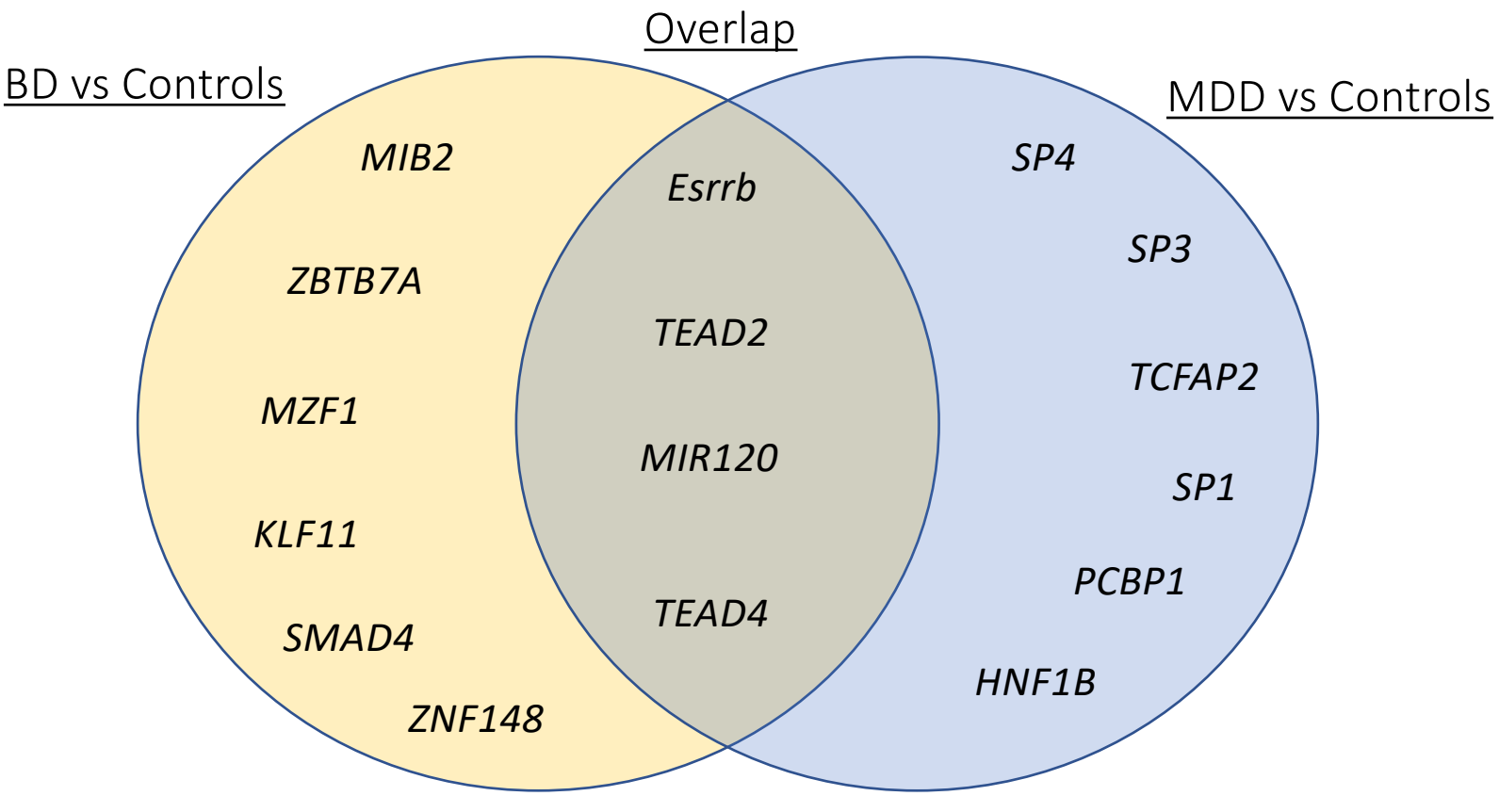
A Bipolar Disorder vs. Controls Enriched Transcription Factors (JASPAR)



B MDD vs. Controls Enriched Transcription Factors (JASPAR)



C



JASPAR TF motifs for High vs. Low Axis-I Comorbidity Load in BD & MDD

NFATC2 (human)	- Promotes induction of interleukin genes IL-2, IL-3, and IL-4 and TNF-alpha transcription during immune response
GABPA (human)	- Activates cytochrome oxidase expression and nuclear control of mitochondrial function - Regulates expression of adenovirus E4 gene during immune functions
HMGA1 (human)	- Regulates the integration of retroviruses into chromosomes - Regulates adipogenesis
NR3C1 (human)	- Master regulator of inflammatory and cellular developmental processes - Regulates glucocorticoid transcription
GTF2I (human)	- Master regulator of general gene expression by interacting with basal transcriptional machinery - Activates immunoglobulin heavy chain transcription
IRF2 (human)	- Inhibits IRF-mediated transcriptional activation of interferon-alpha and interferon-beta
POU1F1 (human)	- Regulates transcription of genes involved in pituitary development and pituitary hormone expression; with mutations of the TF gene causing pituitary hormone deficiency - Regulates mammalian cellular and organ development
SAMD9L (human)	- Mediates downregulation of growth-factor signaling by internalizing growth factor receptors
SNAI1 (human)	- Regulates mesoderm formation in developing embryos by transcriptionally repressing and thereby downregulating the expression of ectodermal genes within the mesoderm. - Regulates cell migration and growth arrest.
CBEPB (human)	- Master regulator of several other transcription factors through transcriptional co-activation - Controls embryonic development and cell growth