Transcription factor motifs in anterior insula gene-expression in mood disorder

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

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

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

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

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

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

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

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

[8]Institute of Early Life Adversity Research

[9]Department of Integrative Biology, the 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: Organizational principles of brain structure and function are important for mediation of complex adaptive behavioral traits. When aspects of these hierarchical brain organizations go awry, maladaptive disease phenotypes with varying behavioral implications emerge. The hierarchical organizations of the neurobiological blueprints for adaptive behavior are governed by the role of genes, especially their expressive codes, in mediating molecular processes critical for the cellular and systems-level development and functions of the brain, which, in turn bridges the genetic effects on complex behavioral phenotypes. Although over ~10% of the United States adult population suffers from mood-related symptoms and disability, hierarchical gene expression modules underlying pathological brain phenotypes associated with diagnostic mood symptoms remains poorly undefined.

Methods: Here, we dissected anterior insula tissue identified to be reduced in volume in mood disorders and extracted RNA-samples from the dissected tissue in an independent postmortem mood disorder and unaffected control samples. We then performed RNA-sequencing with the RNA-samples to identify differentially expressed genes (DEGs) in lifetime diagnoses of **a**) bipolar disorder (BD; n=37), **b**) major depressive disorder (MDD; n=30) vs and **c**) non-psychiatric controls (n=33). 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. submitted) that identified whole transcriptome gene expression profiles in mood disorders, with the aim to explore in more detail, the regulatory logic of mood disorder transcriptomics by identifying TF motifs involved in regulating the previously identified mood disorder related gene expression landscape. To this goal, we used JASPAR TF binding database

to identify motifs for DEGs in BD vs controls; MDD vs controls, and in a pooled cohort of BD & MDD individuals with high Axis-I comorbidity vs low Axis-I comorbidity.

Results: We found that differential gene expressions (DEGs) in BD vs controls were enriched with motifs of *REL*, *MIR210*, *SMADD3*, *TEAD4*, *HINFP*, *AP1*, *NFE2*, *IRF9*, *RUNX2*, and *FOXA3* TFs. DEGs in MDD vs controls were found to be enriched with motifs of *SPI1*, *IRF9*, *LTF*, *FOS*, *FOXO3A*, *SND1*, *PPARA*, *MIR138*, and *CBEPA* TFs. Of interest, DEGs profiles for Axis-I comorbidity load which was factor analytically clustering with suicide mortality in the pooled BD & MDD postmortem donor samples (i.e. without control donors included in the analysis) were found to be enriched with motifs of *NFATC2*, *GABPA*, *HMAGA1*, *NR3C1*, *GTF2i*, *IRF2*, *POU1F1*, *SAMD9L*, *SNAI1*, and *CBEPB* TFs. 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.

Conclusion: Recent genomic advances are transforming our understanding of the molecular building blocks of complex behavioral phenotypes including mood disorders. Here, we applied *robust image-guided transcriptomics* by first targeting the most pronounced gray matter loss in the anterior insula underlying mood disorder diagnoses, to guide an independent postmortem RNA-seq study of TF motifs regulating gene-expression enrichments in mood disorders. We identified TF motifs for differentially expressed innate-immune, cellular homeostasis, embryonic and neurodevelopmental processes. Together, our results provide the foundation for better understanding of the hierarchical relationship between gene regulatory networks, the TFs that control them, and proximate neuroanatomical abnormalities underlying behavioral dysfunctions.

INTRODUCTION

Transcription factors (TFs) are sequence-specific DNA-binding factors also known as regulatory proteins that carries out their regulatory functions of genes by: a) recognizing, and b) binding to specific DNA regulatory elements called DNA binding domains present in the promoter region of genes (Changeaux 2017; Lambert et al. 2018). This DNA binding domain-specific regulatory attribute enables TFs to hierarchically regulate the functional expression of genes by controlling (i.e. promoting/activating or blocking/repressing) transcription of the adjacent structural genes into mRNA, and subsequently further structurally downstream into related proteins (Changeaux 2017; Lambert et al. 2018). TFs are therefore elevated in the hierarchy of gene expression repertoires because they cooperatively, or synergistically regulate genes encoding other TFs (Lambert et al. 2018), and, are thereby functionally 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 interactive relationships often create functionally co-operative gene assemblies of varying hierarchical trees of gene expression programs (Kerszberg & Changeaux 1998; Tsigelny et al. 2013).

Based on these complex but crucial role of thousands of transcription factors in controlling gene regulatory programs that crucial for the maintenance of specific cell states in humans, misregulations of TF-mediated gene expression programs have been hypothesized to cause a broad range of diseases (Lee & Young 2013), including maladaptive behavioral processes rooted in brain developmental dysregulations in neuropsychiatric disorders (Lee & Young 2013; Nord AS, Pattabiraman K, Visel A & Rubinstein JLR 2015; Changeaux 2017; Torre-Ubieta et al. 2018). However, the role of TF networks in mood disorder morbidity is not well defined. Given that disorders of the human affective/mood system are the most common neuropsychiatric symptoms

(Yizhar 2012), and collectively account for one of the highest global burden of disease (Collins et al. 2010), we examined the role of TFs in anterior insula gene expression profiles associated with affective dysfunctions.

Rooted in its anatomical and physiological significance in affective pathology is the insula's transmodal cortical attributes and intricately complex connectivity with other brain regions through efferent fibers from subcortical pathways coupled with afferent fibers to the frontolimbic pathways (Mesulam 1998; Craig 2009; Jabbi et al. 2012). In line with this central anatomical status of the insula, in conjunction with findings of the importance of the anterior insula's structural and functional integrity in mediating the regulation of affective behavior and diagnoses of affective symptoms (Jabbi et al. 2012; Goodkind et al. 2015; Wise et al. 2016; McTeague et al. 2017; Jabbi et al. submitted), we test the hypothesis that given the anterior insula's transmodal cortical properties that places this brain region high up in the connectome and functional hierarchy of human frontolimbic cortical mediation of affective functions, transcription factor gene expression profiles will be enriched in the postmortem anterior insula cortex of mood disorder subjects relative to controls.

Although the developmental mechanism of brain modulation of affective states need further elucidation in humans, adult brain gray matter volumetric reductions centered in the anterior insula have been consistently identified in mood disorders (Goodkind et al. 2015; Wise et al. 2016; Jabbi et al. submitted). These reduced gray matter volumes are documented to predict cognitive control impairments and corresponding degree of affective dysfunction in both adult major depressive disorders and bipolar disorder (McTeague et al. 2017). At the functional level, hypometabolism of the right anterior insula region exhibiting volumetric reductions has been reported to predict remission to cognitive behavioral therapy (CBT) coupled with

pharmacotherapeutic response, while hypermetabolism of the right anterior insula predicted remission to pharmacotherapy and poor response to CBT in mood disorder patients (McGrath et al. 2013). To date however, specific gene-mediated neurobiological substrates for a) mood disorder morbidity, and 2) the putative role of TFs in gene expression profiles in the anterior insula involvement in mood symptomatology remains poorly defined, limiting our understanding of the molecular origin of the anterior insula's involvement in affective symptoms.

Here, we followed up on our previous *robust image-guided transcriptomics* approach that dissected postmortem tissue from the anterior insula region showing gray matter loss in mood disorder imaging and extracted RNA-samples from the postmortem dissected tissue. This was followed with a poly-A enrichment protocol followed by RNA-Sequencing library prep on a HiSeq 4000 PE150 for capturing 30 million paired end reads per sample RNA-sequencing to examine whole transcriptome defined differential gene expression in mood disorders across unique postmortem brain samples of non-psychiatric controls (n=33), those with a lifetime history of major depression (n=30) and bipolar disorder (n=37), i.e. 100 individual brain samples in total. With the aim to explore a more detail regulatory logic of mood disorder transcriptomics, we used the JASPAR 2018 TF binding database (Khan et al. 2018) to identify TF motifs involved in regulating the mood disorder related gene expression landscape. We found that relative to control donors, DEGs in postmortem mood disorder donors were enriched with TF motifs known to regulate expression of toll-like receptor signaling genes, cellular homeostatic control genes, and genes involved in cellular and neurodevelopmental 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, we first performed a large scale meta-analysis of voxel based morphometry studies of gray matter loss in mood disorders 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). By examining the brain region exhibiting the most statistically significant gray matter loss associated with mood disorder using the anatomical likelihood estimation, we identified the right anterior insula cortex gray matter to be most structurally reduced in volume (***Figure. 1A***) in mood disorder diagnoses relative to healthy controls in a living neuroimaging cohort (Jabbi et al. submitted). The identified anterior insula region exhibiting gray matter loss in living mood disorder 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

For each postmortem sample, the donor data includes specific mood disorder and comorbid lifetime-Axis-I diagnoses (i.e. # of lifetime-Axis-I diagnostic occurrences such as (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. # of lifetime-Axis-III medical conditions such as diabetes, cancer, cardiovascular disease, etc.); manner of death (i.e. death by suicide, homicides, accidents or natural death due to comorbid Axis-III medical conditions); and cause of death 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-index, pH, and RIN); and toxicology (blood alcohol/blood narcotics levels). We applied Principal Axis Factoring (Oblimin Rotation with Kaizer Normalization) (Costello & Osborne 2005) to identify higher-order factors explaining the differences in postmortem variables and included those variables with communalities of ≥ 0.5.

Brain regional tissue dissection

For *robust image-guided transcriptomics*, spatial maps of the localized anterior insula shown to be structurally reduced in the imaging meta-analysis of mood disorder diagnoses were *1*) electronically provided in 3-Dimentional images to the neuropathologist for *2*) dissection of independent postmortem tissue samples from the localized region. 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 following previous protocols (Lipska et al. 2006). The region of interest results of gray matter volume reduction in right anterior insula reported earlier (Jabbi et al. submitted) 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. Coronal tissue slabs of each postmortem donor brain, this corresponded to the anterior portion of the insula where the caudate and putamen are approximately equal in size and extending further in the anterior tip were dissected by the HBCC neuropathologist (***Figure. 1C-D***).

RNA extraction and processing

All dissected tissues were pulverized and 50mg were aliquoted and used for standardized total RNA extraction using Agilent Technologies RNA 6000 Nano Kit and only samples with RNA integrity numbers (RIN) which is a measure of RNA postmortem RNA quality were included in the study (see Table 1B for average RIN values per group). 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 as previously described in Jabbi et al. submitted. This purified RNA from the pulverized anterior insula tissue sections was used to determine RNA quality as measured in RNA integrity numbers (RIN) using Agilent 6000 RNA Nano Kit consisting of the microfluidic chips, Agilent 6000 RNA Nano ladder and reagents on Agilent 2100 Bioanalyzer. Samples with RNA integrity number (RIN) < 6 were excluded.

RNA-Sequencing and analysis

Illumina-Sequencing, Read-Mapping and Gene-Quantification: Total RNA was extracted and only samples with RNA integrity numbers > 5 were used. 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). The 100 samples were processed using TruSeq RNA Library Prep Kit v2 and sequenced on the Illumina HiSeq 4000 at the Genome Sequencing and Analysis Facility (GSAF) at the University of Texas, Austin, USA. Paired-end libraries with average insert sizes of 200bp were obtained using NEBNext Ultra II Directions Library Prep kit from New England BioLabs and mRNA selection was done used the Poly(A) purist kit from Thermofisher. 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).

Differential Gene Expression Analysis: The reads were pseudo-aligned to the human reference transcriptome (GRCh38- gencode) using kallisto (Kallisto, 2019), 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).

JASPAR 2018 TFs binding profiling: 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.

Here, we used the RESTful API programmatic access to JASPAR 2018 version using Python to access 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 diagnostic comorbidity vs pooled mood disorder cohort with low Axis-I diagnostic comorbidity).

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 [race was significantly different in that the 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 load* 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 vs controls ($p<0.0001$); BD vs controls ($p<0.0001$); and BD vs MDD ($p = 0.004$).

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

By first directly comparing TF motifs found in BD vs control DEGs, we found DEGs were enriched with motifs of *REL*, *MIR210*, *SMADD3*, *TEAD4*, *HINFP*, *AP1*, *NFE2*, *IRF9*, *RUNX2*, and *FOXA3* TFs (***Figure. 2A***). Direct comparisons between MDD vs controls identified enriched motifs of *SPI1*, *IRF9*, *LTF*, *FOS*, *FOXO3A*, *SND1*, *PPARA*, *MIR138*, and *CBEPA* TFs (***Figure. 2B***). Of interest, the two comparisons between BD vs controls and MDD vs controls identified TF motifs of similar pathways but the list of top ten TF genes did not overlap across the two comparisons suggesting possible diagnostic specific involvements of these motifs.

Given our approach to examine the hierarchical relationship between TF motifs and DEGs profiles associated with the degree of psychiatric morbidity/Axis-I diagnostic comorbidity in the two mood disorder cohorts, we first used principal axis factoring to identify *Axis I diagnostic-load (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 load) to compositely explain variability in (i) <u>Axis-I comorbidity load</u>, and (ii) <u>suicide completion</u>. We included this factor in our analysis using a split half approach of high vs low Axis-I comorbidity load to identify TF motifs for DEG profiles for Axis-I comorbidity load and found DEGs profiles in the pooled BD & MDD donor samples (i.e. without control donors) and found enriched 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 a differential gene expression landscape associated with enriched motifs for TFs involved in regulating *a*) inflammatory and immune response and toll-like receptor signaling (i.e. *REL* and *IRF9*), *b*) cellular developmental processes and apoptosis (i.e. *REL*, *SMAD3*, *TEAD4*, *RUNX2*, *NFE2*, and *FOXA3*), *c*) cellular and peripheral homeostatic control (i.e. *SMAD3*, *JUN*, *RUNX2*, and *FOXA3*) and *d*) post transcriptional regulation and DNA-methylation (i.e. *HINFP* and *MIR210*) (***Figure 2A***). Comparing MM vs controls revealed a differential gene expression landscape associated with enriched motifs for TFs involved in regulating *a*) inflammatory and immune response and toll-like receptor signaling (i.e. *IRF9I*, *LTF*, and *SND1*), *b*) cellular developmental processes and apoptosis (i.e. *SPI1*, *LTF*, *FOXA3*, *FOS*, *PPARA*, and *SND1*), *c*) cellular and peripheral homeostatic control (i.e. *CEBPA*, *NR5A2*, *PPARA*, and *SPI1)* and *d*) post transcriptional regulation and DNA-methylation (i.e. *SND1* and *MIR138-1*) (***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 a differential gene expression landscape associated with enriched 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

Recent multidisciplinary research is contributing to advancing our understanding of the structural and functional bases for human brain substrates for behavioral functions but in-depth understanding of the hierarchical links between the molecular building blocks of brain structure and functions supporting complex adaptive behavioral phenotypes remains poorly defined. Understanding the molecular substrates for specific regional brain abnormalities underlying behavioral dysfunctions such as prevalent 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 EJ International Review of Neurobiology 2015; Sahin & Sur Science 2015; Torre-Ubieta et al. 2018; Jabbi & Nemeroff 2019) but will also be essential for advances in novel drug design (Changeaux 2017).

Here, we applied *robust image-guided transcriptomics* by extracting RNA-samples from brain tissue dissected from a precisely localized anterior insula region exhibiting the strongest degree of gray matter loss in association with mood disorder diagnoses. By extracting tissue samples from this brain region in an independent postmortem sample, this *robust image-guided transcriptomics* study therefore integrated RNA-seq characterization of gene expression abnormalities in a region known to harbor anatomical abnormalities in mood disorders. Such an integrative *robust image-guided transcriptomics* approach has the potential to examine hierarchical gene expression profiles in terms of TF motifs, in a putative anatomically aberrant

brain region in mood disorders. To this end, we used the JASPAR 2018 database to specifically search for statistically significant TF motifs in the identified gene expression profiles for specific mood disorder diagnostic phenotypes and identified enriched TF motifs known to regulate expression of toll-like receptor signaling genes, cellular homeostatic control genes, and embryonic and cellular including neurodevelopmental gene networks in BD relative to controls, MDD relative to controls, and in high Axis-I comorbid mood disorder individuals relative to low Axis-I diagnostic comorbid 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 only associated with the gene expression landscape in elevated mood disorder morbidity and 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 *a*) TFs that are hierarchically programmed to functionally regulate *b*) other gene classes that do not control other genes (Love et al. 2014; Anders & Huber 2010). Departing from this traditional approach, we determined which TF motifs stored as PFMs in the JASPAR 2018 database as TF-binding profiles of the vertebrate taxonomy are associated with the transcriptomic repertoire of differentially expressed genetic profiles in our anterior insula RNA-seq study of mood disorder morbidity. Our identification of enriched 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 this anterior insula anatomical locale 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 (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 exhibiting 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 is present in the 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.

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 assays in mood disorders across the lifespan will also advance understanding of the role of embryonic and neurodevelopment TFs in the onset of mood disorder diagnostic association with regional gray matter loss. While our current study identified relevant TF involvement in regulatory gene expression signatures for mood disorders, including a comparison psychiatric cohort with no mood symptoms like 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 neurotranscriptomics. Our findings successfully demonstrated that the gene expression landscape for degree of mood disorder morbidity, in an anatomical abnormal region in mood disorders, are associated with enriched TF motifs known to regulate immune responsive, cellular homeostatic control, and embryonic as well as cellular and neurodevelopment genetic functions. Together, these findings illustrates that studies of gene regulatory networks, in addition to studying the functional profiles of actual genes these TF regulatory networks control/transcribe, has the potential to elucidate the hierarchical organizational principles of gene expression landscapes in complex behavioral brain diseases and thereby accelerate target discovery. If replicated, our observed results of specific TF-network involvements in the gene expression repertoire for diagnostic mood disorder phenotypes – in brain anatomical locus showing reduced gray matter abnormalities in mood disorders – in line with the clinical relevance of transcriptional regulation and dysregulations in the emergence of complex diseases, may advance mechanistic insights and guide novel neurotherapeutics in mood and other dysfunctions of the nervous system.

FIGURE 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.

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

FIGURES

Figure 1

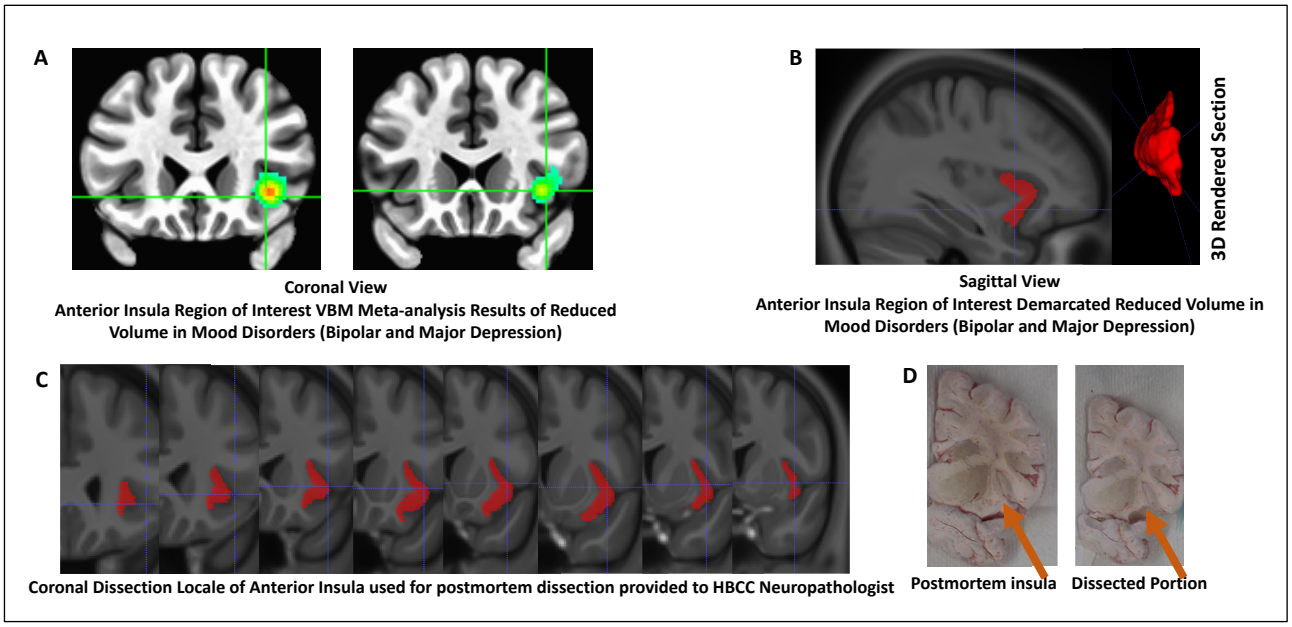


Figure 2

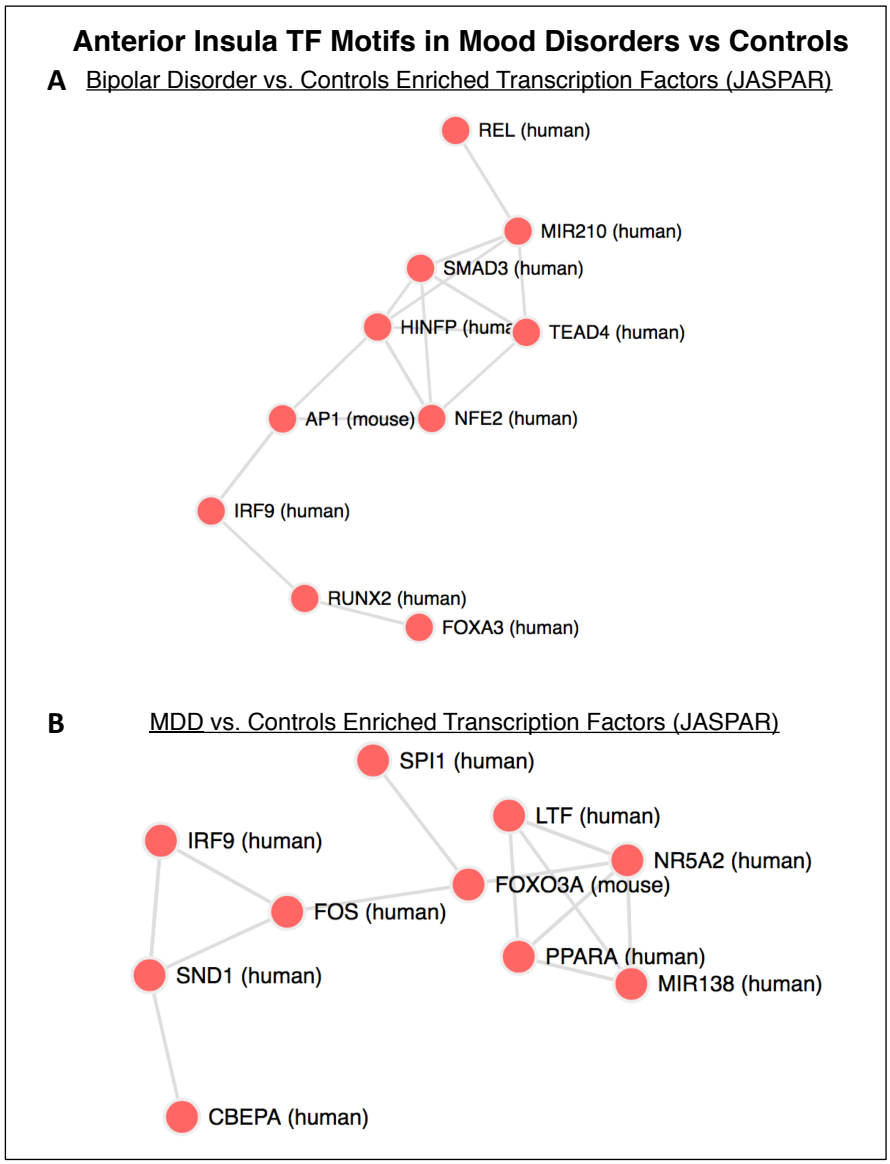
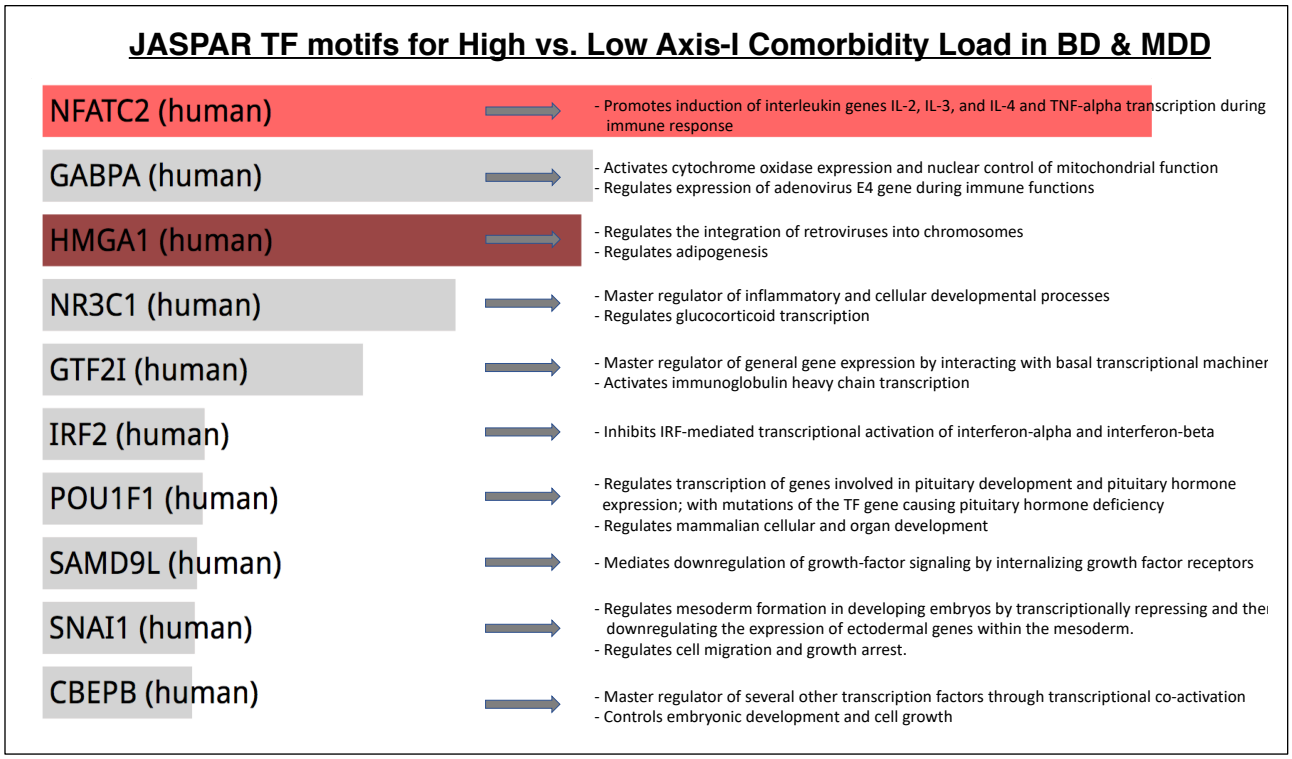

Anterior Insula TF Motifs in Mood Disorders vs Controls
A
Bipolar Disorder vs. Controls Enriched Transcription Factors (JASPAR)
REL (human)
MIR210 (human)
SMAD3 (human)
HINFP (huma
TEAD4 (human)
AP1 (mouse)
NFE2 (human)
IRF9 (human)
RUNX2 (human)
FOXA3 (human)
B
MDD vs. Controls Enriched Transcription Factors (JASPAR)
SPI1 (human)
LTF (human)
IRF9 (human)
NR5A2 (human)
FOXO3A (mouse)
FOS (human)
PPARA (human)
SND1 (human)
MIR138 (human)
CBEPA (human)

Figure 3


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 machiner
- 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 the 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

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.

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

Chen C, Meng Q, Xia Y, Ding C, Wang L, Dai R, Cheng L, Gunaratne P, Gibbs RA, Min S, Coarfa C, Reid JG, Zhang C, Jiao C, Jiang Y, Giase G, Thomas A, Fitzgerald D, Brunetti T, Shieh A, Xia C, Wang Y, Wang Y, Badner JA, Gershon ES, White KP, Liu C. The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. *Sci Transl Med*. 2018; 10(472) [available online].

Craig AD. How do you feel--now? The anterior insula and human awareness. *Nat Rev Neurosci*. 2009 Jan;10(1):59-70.

Collins PY, Patel V, Joestl SS, March D, Insel TR, Daar AS; Scientific Advisory Board and the Executive Committee of the Grand Challenges on Global Mental Health, Anderson W, Dhansay MA, Phillips A, Shurin S, Walport M, Ewart W, Savill SJ, Bordin IA, Costello EJ, Durkin M, Fairburn C, Glass RI, Hall W, Huang Y, Hyman SE, Jamison K, Kaaya S, Kapur S, Kleinman A, Ogunniyi A, Otero-Ojeda A, Poo MM, Ravindranath V, Sahakian BJ, Saxena S, Singer PA, Stein DJ. Grand challenges in global mental health. *Nature*. 2011 Jul 6;475(7354):27-30.

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

de la Torre-Ubieta L, Stein JL, Won H, Opland CK, Liang D, Lu D, Geschwind DH.
The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis. *Cell*. 2018 Jan 11;172(1-2):289-304.e18.

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

Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT 2012. Activation likelihood estimation meta-analysis revisited. *Neuroimage*. 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, Meyer-Lindenberg A, Berman KF. 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(14):E860-6.

Jabbi M, Chen Q, Turner N, Kohn P, White M, Kippenhan JS, Dickinson D, Kolachana B, Mattay V, Weinberger DR, Berman KF. 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. (*In Revision*) 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 Jun;19(6):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.

Kallisto: https://www.ncbi.nlm.nih.gov/pubmed/27043002.

Kessler RC et al. (2005). Prevalence, Severity, and Comorbidity of Twelve-month DSM-IV Disorders in the National Comorbidity Survey Replication (NCS- R). *Archives of General Psychiatry*, *62*(6): 617–627.

Kerszberg M, Changeux JP. A simple molecular model of neurulation. *Bioessays*. 1998 Sep;20(9):758-70.

Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Chèneby J, Kulkarni SR, Tan G, Baranasic D, Arenillas DJ, Sandelin A, Vandepoele K, Lenhard B, Ballester B, Wasserman WW, Parcy F, Mathelier A. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* 2018; 46(D1):D1284.

Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT. The Human Transcription Factors. *Cell*. 2018 Oct 4;175(2):598-599.

Lipska BK, Deep-Soboslay A, Weickert CS, Hyde TM, Martin CE, Herman MM, Kleinman JE. Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biol Psychiatry*. 2006 Sep 15;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, Craddock RC, Mayberg HS. 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(7):676-685

Mesulam MM. From sensation to cognition. *Brain*. 1998 Jun;121 (Pt 6):1013-52.

Murray CJL et al. (2012). 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*, *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 Jan 7;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).

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 Jun 15;71(12):1075-80.