Transcriptomics of the depressed and PTSD brain

Jing Zhang\textsuperscript{a}, Alfred P. Kaye\textsuperscript{b}, Jiawei Wang\textsuperscript{c,d}, Matthew J. Girgenti\textsuperscript{b,e,*}

\textsuperscript{a} Department of Computer Science, University of California- Irvine, Irvine, CA, USA
\textsuperscript{b} Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA
\textsuperscript{c} Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA
\textsuperscript{d} Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA
\textsuperscript{e} National Center for PTSD, U.S. Department of Veterans Affairs, USA

ARTICLE INFO

Keywords: PTSD
MDD
Stress transcriptomics
Postmortem brain
Animal behavior

ABSTRACT

Stress is the response of an organism to demands for change, yet excessive or chronic stress contributes to nearly all psychiatric disorders. The advent of high-throughput transcriptomic methods such as single cell RNA sequencing poses new opportunities to understand the neurobiology of stress, yet substantial barriers to understanding stress remain. Stress adaptation is an organismal survival mechanism conserved across all organisms, yet there is an infinity of potential stressful experiences. Unraveling shared and separate transcriptional programs for adapting to stressful experience remains a challenge, despite methodological and analytic advances. Here we review the state of the field focusing on the technologies used to study the transcriptome for the stress neurobiologist, and also attempt to identify central questions about the heterogeneity of stress for those applying transcriptomic approaches. We further explore how postmortem transcriptome studies aided by preclinical animal models are converging on common molecular pathways for adaptation to aversive experience. Finally, we discuss approaches to integrate large genomic datasets with human neuroimaging and other datasets.

1. Introduction

Neuropsychiatric disorders cause considerable disability and burden economically and to the health of people around the world. Significant efforts have been made to understand these disorders; however, their underlying molecular pathology remains elusive. Preclinical and clinical studies indicate that psychiatric disorders onset is multifactorial; part environment and part genetic and arise from structural and molecular changes in corticolimbic and mesolimbic circuits (Jovanovic et al., 2010; Jovanovic and Ressler, 2010; Shin and Liberzon, 2010). Specifically, aversive psychosocial stressors are known to cause major depressive disorder (MDD) and post-traumatic stress disorder (PTSD). There is now a substantial literature describing the genetic architecture and risk for developing MDD (Wray et al., 2018) and PTSD (Gelernter et al., 2019; Nievergelt et al., 2019). However, few of the identified risk variants have been functionally annotated.

The human genetic code contains 3 billion base pairs and an its annotation (introns, exons, intergenic regions, etc.) is still an area of extensive research (Kundaje et al., 2015). The transcriptome is the “read out” of the genetic code and studying it reduces the dimensions from 3 billion base pairs to approximately 25,000 coding transcripts.

Transcriptomic studies of neuropsychiatric disorders in both animal models and human postmortem tissue provide the best strategy for understanding how risk variants and stress exposure affect the molecular pathology of the central nervous system (CNS). Postmortem brain tissue provides us with a direct measurement of how risk variants may affect gene expression and perturbation of these genes in animal models provides critical functional validation. The transcriptome can be thought of a molecular phenotype of a particular trait or illness state in the same way as neurological disruptions such as plaques and tangles are phenotypic of Alzheimer’s disease. This analogy is particularly relevant as there are no macro-neurological features that distinguish patients with psychiatric disorders. For example, functional genomic studies of autism have consistently identified atypical gene co-expression networks across multiple ASD cohorts (Adhya et al., 2021; Wang et al., 2018).

There are currently major efforts to understand the molecular pathology of stress disorders. Preclinical work on stress disorders has been directed at understanding how much of the disease transcriptomic changes are recapitulated by each stress model (Scarpa et al., 2020). Increased collection of human postmortem tissue from donors with psychiatric illness has contributed to advances in understanding the
transcriptomic changes of MDD and PTSD and provided comprehensive genomic atlases which can be used to identify clinically relevant changes in animal model comparisons. Here we review the current state of technologies used in transcriptomic studies: from RNA-sequencing of bulk-tissue to single cell-types. We explore how animal models have contributed to our understanding of the molecular pathology of MDD and PTSD. And finally, we describe how postmortem genomics work has revolutionized our understanding of MDD and PTSD and how integrating these large, rich data sets with other human analyses such as neuroimaging are providing critical information necessary for better biomarker identification and therapeutic design.

2. Transcriptomic methods

Technical advances have made high-fidelity measurements of gene expression changes following stress at cellular resolution possible (Macosko et al., 2015), as well as enabling simultaneous large-scale studies of other cellular measurements (“multi-omics”) (Macaulay et al., 2017). This rapidly advancing methodological approach is yielding insights into fundamental building blocks of pathophysiology across biomedical fields. Here we provide a brief summary of approaches to transcriptomic measurements relevant to understanding stress in order to facilitate psychiatric neuroscience investigators who wish to adopt these approaches in the understanding of stress-related disorders.

Recent advances in our understanding of the genomic structure of the mammalian CNS have been directed by the development of high throughput sequencing approaches such as RNA-seq. RNA-seq provides three to four magnitudes more information than previous gene expression systems. Deep sequencing has been applied to many different genomic levels including genomic DNA sequencing (DNA-seq) (Yu et al., 2017), non-coding RNA profiling (smRNA-seq) (Choi et al., 2005; Xu et al., 2011; Zhou et al., 2021), DNA-protein interactions (ChIP-seq) (Barski et al., 2007) and DNA structure (e.g., open chromatin) (ATAC-seq) (Bryois et al., 2018). The actual process of RNA-seq is detailed elsewhere but we will briefly review here. RNA is extracted and ribosome depleted before cDNA synthesis. Ribosome depletion has become the gold standard for RNA sample preparation as it removes ribosome depleted before cDNA synthesis. Ribosome depletion has obscured other relevant molecular changes. Each cDNA fragment (usually detailed elsewhere but we will briefly review here. RNA is extracted and ribosome depleted before cDNA synthesis. Ribosome depletion has become the gold standard for RNA sample preparation as it removes ribosome depleted before cDNA synthesis. Ribosome depletion has obscured other relevant molecular changes. Each cDNA fragment (usually...
create a contrast between two conditions (stressed and unstressed). Importantly the effects of acute and chronic stress are frequently different in terms of transcriptomic changes (Simmons et al., 2020). One large gene expression atlas of different kinds of stress in rodent brain meta-analyzed gene expression data from 18 projects studying the effect of stress on the transcriptome of the brain (Flati et al., 2020). They observed a consistent pattern between acute and chronic stressors and brain gene expression. GO analysis of genes modulated a short time after stress (acute) enriched for genes involved in enzymatic activities and transcription factors. At longer time points after stress (chronic) functional annotation of DEGs revealed enrichment for structural changes such as axonogenesis and synapse organization. These findings are consistent with known structural and plasticity responses to chronic and acute stress.

Acute stress offers the ability to identify direct gene expression mediators of persistent changes after stress (Floriou-Servou et al., 2021). A recent study of the ventral hippocampus after exposure to three different types of acute stressors found that gene expression patterns related to different stressors were quite divergent, but that there were some common changes across different stress types (Floriou-Servou et al., 2018). Since acute stress responses evolve over hours to days, future studies and modeling efforts may facilitate understanding the temporal evolution and key elements of stress responsivity.

While chronic stress manipulations vary, the most common have either been repetitive unpredictable sensory stimuli (Girgenti et al., 2019), social defeat stress (Bagot et al., 2016), or early life stress (Peña et al., 2017). Social defeat stress, in which a mouse is exposed repeatedly to another aggressive mouse over 10 days, leads to differential expression of large numbers of genes throughout limbic circuits (Bagot et al., 2016). One advantage of this model is the ethological nature of the stress manipulation, as well the ability to distinguish resilient vs susceptible animals by their level of future social interaction. In transcriptomic studies, this has been used to identify gene networks and key gene expression programs. In one remarkable study, early life stress induced widespread and persistent changes in gene expression in the ventral tegmental area (Peña et al., 2017). Interestingly, this response was orchestrated by a brief change in expression of the transcription factor OTX2 which was not sustained into adult gene expression changes. Importantly, these studies highlight the critical need of tying gene expression changes to neuronal activity as both studies provide causal evidence for specific genes in the control of the stress response. Chronic stress has an impact on the individual cell’s regulation of gene expression. One group performed single cell RNA-seq on each component of the HPA axis and identified a previously unknown subcluster of cells that were Agrpα+ and which play an important role in the plasticity and adaptation processes of chronic stress in the adrenal cortex and highlights the need to move this work to the single cell-type level (Lopez et al., 2021).

However, neurobiological research on stress has reached a major obstacle - the very “non-specific” nature of the phenomenon has blocked further study (Simmons et al., 2020). The wide variety of ways to induce stress lead to diverse and often contradictory responses in neural circuits. Stress is typically induced in neurobiological studies by the standardized application of different aversive experiences – say, 4 h of physical restraint (Conrad et al., 1999) or ten days of a mouse being placed for 10 min per day with an aggressive mouse (Golden et al., 2011). These approaches have yielded tremendous insight into the neurobiology of stress but are difficult to compare across conditions. In the next section, we address one potential approach to this problem.

4. Exploring stress space with transcriptomics

Understanding the nature of stressful experiences requires environmental manipulations that vary across multiple dimensions: sensory variability, higher-order unpredictability, controllability, aversion, chronicity (McGonagle and Kessler, 1990; Peters et al., 1998; Seligman, 1972). The space of potentially stressful events is infinite, with contemporary human stress induced by diverse socioeconomic factors (Lantz et al., 2005) such as social isolation, threat of danger, grief, economic resource instability, diminished appetitive stimuli, loss of control and increased threat of violence. In human psychological studies, the observation that stressors of specific types can induce diverse behavioral manifestations has been an important facet of study going back to the beginning of the field. Yet in animal neurobiology research, the ability to compare across causal manipulations of distinct stressors has been limited by the difficulty of exploring an infinite space of stressors (although see recent efforts at comparison across acute stressors (Floriou-Servou et al., 2018)).

Typically, neurobiologists choose one or a few distinct types of stress, which can be thought of as points in the space of potential stressors. Yet, comparing neurobiological manifestations between two classes of stressors is equivalent to choosing two points, and does not permit inference into the nature of the structure of stressors within that space. Given that stress typically constitutes a series of aversive experiences over time, each individual incident within the stressful experience can be considered as a small directed perturbation in stress space. One quantitative way to conceptualize the transcriptomics of stress is as a mapping from the space of potential stressful (uncertainty- or aversion-inducing) (Peters et al., 2017) experiences to the space of cellular states.

In some ways, our understanding of transcriptomics of stress is thus limited by our ability to control and elicit stress in model systems. As discussed in the previous section, stress in animal models is usually either a single acute or chronic stressor with contrasts between stressed and unstressed conditions. However, this level of quantitative and individualized description has been limited by the lack of theoretical model for stress in which to localize transcriptomics data. In order to truly understand the commonalities across behavioral stress and its cell-type specific transcriptomic consequences, it may be necessary to develop improved behavioral programs and modeling approaches as well as make comparisons with human postmortem tissue of stress disorders. Work is currently underway to identify the molecular intersections between animal models of stress and postmortem tissue genomics work. One study compared a large postmortem, MDD cohort to three separate stress paradigms to identify whether one model better captured the molecular pathology of MDD than another (Scarpa et al., 2020). Interestingly, the authors found that each model captured different aspects of the depression molecular signature. The authors suggest that this is not an impediment but an opportunity to expand the types of mouse models of MDD. Because all three behaviors in this study capture parts of the MDD molecular profile, future work should be expanded into other stress disorders such as PTSD where a complete molecular profile of the animal models is still lacking. Further, future work in the depression model area should be to expand this into cell type-specific work. Despite the tremendous progress in transcriptomic methods, linking new bioinformatic methods for understanding cellular state trajectories after stress to an improved quantitative understanding of stress itself may be essential to understanding how stress alters neural circuit function.

5. Bioinformatic approaches to brain transcriptomic studies

5.1. Differential gene and isoform identification

The original, most direct and informative analyses of transcriptomic data involve identifying lists of differentially expressed genes between sample groups such as discrete brain regions, diseased tissue, or even between the two sexes. RPKM or FPKMs counts for individual transcripts are normalized based on the total number of effective counts for a particular gene in each sample. Programs such as EdgeR(Robinson et al., 2010), DESeq (Love et al., 2014) and Limma/voom (Law et al., 2014) perform well for most data sets and are effective in ranking differentially expressed genes. EdgeR and DESeq2 both assume that no transcripts are
differentially expressed. DESeq uses a geometric normalization strategy where a scaling factor for a given sample is computed as the median of the ratio for each gene and its read count over its geometric mean across all samples. EdgeR calculates a weighted mean of log ratios between sample cohorts after exclusion of the most expressed genes and genes with the largest log ratios. Linma voom was historically used to analyze microarray data but has been updated to work with sequencing data. Linma/voom matches distributions of gene counts across samples to normalize expression (Shahjaman et al., 2020).

Alternative splicing (AS) is a critical component of gene expression regulation. Through the inclusion or exclusion of exons and intronic sequences, the transcriptome adds to the diversity of the proteome (Pan et al., 2008). AS can have a profound effect on gene function and different isoforms of the same gene can be involved in different or even opposing functions. AS can be measured using traditional RNA-seq methods. Numerous bioinformatic tools have been developed for identifying alternative splicing patterns from RNA-seq including LeafCutter (Li et al., 2018), MISO (Ratz et al., 2010), rMATs (Shen et al., 2014), MAJIQ (Vaquero-Garcia et al., 2016), RSEM (Li and Dewey, 2011), Kallisto (Bray et al., 2016), and Salmon (Patro et al., 2017). While RNA-seq data is generally analyzed on the gene level these packages can be used to analyze this data on the exon level to detect and quantify novel splicing events. AS is usually reported as percent spliced, a metric of the percentage of how efficiently sequences of interest are spliced into a full transcript (Li et al., 2018). A recent study (Gandal et al., 2018) by the Geschwind group identified that more than 25% of the human frontal cortex exhibits differential alternative splicing and that isoform-level changes captured the largest disease effects for donors with autism spectrum disorder, schizophrenia and bipolar disorder. While there is a high degree of common polygenicity in these disorders these findings suggest that alternative splicing of the transcriptome most differentiates these disorders from one another.

Traditionally, RNA-seq is performed on small fragments of RNA (75-150bp). Current detection methods for identifying AS events relies on concatenation of these small sequence fragments. The ability to sequence longer pieces of cDNA are allowing for more accurate identification of splicing events. These so-called long read sequencing assays are allowing for more accurate identification of splicing events. These so-called long read sequencing assays are allowing for more accurate identification of splicing events. These so-called long read sequencing assays include LeafCutter (Li et al., 2018), MISO (Ratz et al., 2010), rMATs (Shen et al., 2014), MAJIQ (Vaquero-Garcia et al., 2016), RSEM (Li and Dewey, 2011), Kallisto (Bray et al., 2016), and Salmon (Patro et al., 2017). While RNA-seq data is generally analyzed on the gene level these packages can be used to analyze this data on the exon level to detect and quantify novel splicing events. AS is usually reported as percent spliced, a metric of the percentage of how efficiently sequences of interest are spliced into a full transcript (Li et al., 2018). A recent study (Gandal et al., 2018) by the Geschwind group identified that more than 25% of the human frontal cortex exhibits differential alternative splicing and that isoform-level changes captured the largest disease effects for donors with autism spectrum disorder, schizophrenia and bipolar disorder. While there is a high degree of common polygenicity in these disorders these findings suggest that alternative splicing of the transcriptome most differentiates these disorders from one another.

5.3. Gene Co-Expression analysis and deep learning

Transcriptional regulation plays a key role in many homeostatic functions and in disease states. Traditionally, bioinformatic approaches to transcriptomic datasets have focused on the identification of genomewide significant differentially expressed genes and is the logical extension from the traditional candidate gene studies that predominated early gene profiling studies. However, additional tools have been developed to provide far more information about transcriptional regulation and organization in tissue. One approach that has been consistently popular is gene co-expression network analysis. The underlying biological principle of this analysis is that genes do not operate alone and likely require co-regulated transcriptional control. Thus, gene levels with correlated levels of expression across cohorts are grouped into network modules and differences in module connectivity and membership provide insight into dysregulated gene assemblies. Further, these modules can also be enriched for cell-type specific markers, allowing for identification of the cell types likely responsible for the aberrant transcriptional differences.

A popular tool for this type of analysis is Weighted Gene Co-Expression Network Analysis (WGCNA) (Langfelder and Horvath, 2008). This tool has been used to study co-expression patterns in normal brain tissue (Oldham et al., 2008), major depressive disorder (Labonté et al., 2017) and PTSD (Girgenti et al., 2021). Importantly, co-expression analysis can be used in prioritizing illness-state DEGs and linking these to genetic risk. For example, the largest postmortem study of PTSD brain tissue identified a PTSD associated co-expression network with numerous down regulated molecular key drivers important in interneuron function (Girgenti et al., 2021). Using a global, step-wise approach the authors combined transcriptome-wide imputation (transcriptome-wide association (TWAS)) using eQTL data from the GTEx portal and integrated it with the largest PTSD GWAS generated by the Million Veteran Program. One of the molecular key drivers identified in their WGCNA modules, ELF1, achieved transcriptome-wide significance for PTSD by TWAS. This was an important first step in identifying convergences between cortical co-transcriptomics and PTSD genetic signals, a pipeline that can be exploited to identify high confidence molecular targets for potential therapeutics for PTSD. TWAS is quickly becoming one of the most popular transcriptomic analysis tools in stress disorders (Stein et al., 2021,Hucks et al., 2020) and its application and development has been reviewed extensively elsewhere (Cano-Ganez and Trynka, 2020; Chatzinakos et al., 2021). Deep learning technology is also being applied to large transcriptomic screens to identify gene-gene relationships (Yuan and Bar-Joseph, 2019). This study designed a novel pipeline which encoded single cell type gene expression data and followed that up with deep neural network analysis. This framework allows for detection of a diverse number of relevant biological phenomenon including transcription factor target prediction and identification of disease-related causal, co-expressed genes.
5.4. Single cell transcriptomics in brain

With an immense number of cell types and cellular connections, the brain is arguably the most complex organ in the human body. The recent single cell sequencing revolution opens a new avenue for dissecting the complexity of the human brain at the single cell level. As we have touched on previously, in no realm is the development of bioinformatics tools more crucial or expanding more rapidly than those for single cell type analysis. In this section, we will describe broad techniques and methods for conducting single cell analysis in brain tissue.

5.4.1. Preprocessing and quality control (QC)

Computational modeling of scRNA-seq data is challenging due to its ultra-high dimensionality, low capture efficiency, and high level of technical noise. Therefore, a series of QC steps are vital for accurate downstream analyses. For instance, bulk RNA-seq QC tools, such as FastQC, can be employed to check the sequencing quality of scRNA-seq data. Then cells with very few genes detected, mitochondrial genes or extremely high portion of reads mapped to the spike-ins will be removed. Further, several tools have been developed to further remove doublets to avoid artificial libraries generated from more than one cell (Bais and Kostka, 2019; Bernstein et al., 2020; McGinnis et al., 2019; Wolock et al., 2019). In addition, multi-sample scRNA-seq analysis usually suffers from severe batch effects when integrating data generated by distinct operators at different times or from multiple laboratories using disparate protocols and sequencing platforms. Recent computational methods such as MNM and KBET can be used to correct batch effects for improved performance (Böttner et al., 2019; HaghiVerdi et al., 2018). Single cell type transcriptomics datasets are prone to noise and circumstantially there has been advances in tools for processing these datasets particularly in the areas of cluster and drop out analysis (Bouland et al., 2021).

5.4.2. Cell level analysis

Normalization is usually carried out to correct unwanted biases (e.g., dropout, sequencing depth, and capture efficiency). Recently, a generalized linear model has been widely used to omit the need for heuristic parameterizations in normalization (e.g., pseudocount addition and log-transformation), benefiting downstream analytical tasks such as variable gene selection, dimensional reduction, and differential expression (Hafemeister and Satija, 2019). Afterward, highly variable genes are selected as informative features to perform various dimension reducing techniques. To identify distinct cell populations, we can perform either supervised clustering methods based on known marker genes or identify de novo cell types using unsupervised algorithms based on k-means, hierarchical clustering, density-based clustering and graph-based clustering. In addition, to accommodate a continuous spectrum of cellular status, cell trajectories and pseudotime can be reconstructed based on scRNA-seq using various tools, such as Monocle (Qiu et al., 2017), Waterfall (Shin et al., 2015), Wishbone (Setty et al., 2016), and Cell-Router (Rocha et al., 2018).

5.4.3. Gene level analysis

Distinct from bulk level analysis, scRNA-seq data can identify differentially expressed genes (DEGs) between disease and controls within each subpopulation or group of cells. Although traditional methods designed for bulk RNA-seq DE analysis can be used, recently several methods have been specifically developed for scRNA-seq data with improved performance, such as MAST (Finak et al., 2015), SCDE (Kharchenko et al., 2014), and DEsingle (Miao et al., 2018).

5.4.4. Network level analysis

Various network inferences, such as gene co-expression network and gene regulatory network (GRN) have been widely conducted using bulk RNA-seq data. A number of existing tools such as WGCNA and SCENIC can be directly applied to scRNA-seq data although caution is warranted due to the higher level of noise in single cell experiments. Several recent tools, such as PIDD (Chan et al., 2017) and CNNC (Yuan and Bar-Joseph, 2019), have been developed specifically for GRN construction using scRNA-seq data. Additionally, inter-cellular communication is particularly critical for the development and functionality of the human brain and one obvious advantage of scRNA-seq data is the ability to systematically dissect the cellular connectome. Recently several tools have been developed to construct cell-to-cell communication networks, such as cellPhoneDB (Efremova et al., 2020), CellChat (Jin et al., 2021), and cellTalkDB (Shao et al., 2020), from single cell data.

Recently, several high profile papers have begun mapping the single cell transcriptomic contribution to the human and mouse brain (Damaris et al., 2018; Hodge et al., 2019; Lake et al., 2018). These studies have challenged our traditional views of the number of potential neuronal subtypes found in the brain. For example, the Hodge et al. paper identified 75 distinct cell type clusters, including 24 types of excitatory neurons and 45 types of inhibitory neurons which precise laminar placement. Most of these clusters represent new, previously unknown cell subtypes. While Hodge is considered the standard by which other brain single cell type studies are being compared, one caveat of the study was the general lack of identification of novel non-neuronal cell subtypes. For example, it is likely that microglia would cluster into at least two transcriptional cell types (activated and quiescent) and there is good evidence of at least 2 separate oligodendrocyte populations in human cortex defined by myelination status (Kleijn et al., 2019). The classification of all cell types in the brain is one of the primary goals of neuroscience and will require the creation of additional bioinformatic tools to parse out the different cell types and subtypes. Further, this work could also help to identify molecules converging between different species (e.g., mice and non-human primates) as these are used extensively as models of human disease.

6. Transcriptomic insights into the molecular pathology of stress implicates interneuron and microglial dysfunction

Postmortem studies of the brains of psychiatric patients has proved essential in understanding the molecular effects of these disorders. This work has been aided by the creation of many large brain banks collecting tissue across numerous neurological and neuropsychiatric disorders. Further, this work has expanded rapidly by the establishment of several large consortia such as Common Mind (Fromer et al., 2016), BrainSeq (Consortium et al., 2018), and PsychENCODE (Wang et al., 2018) which are compiling large disparate genomics datasets and uniformly analyzing them to increase the number of brains analyzed.

Recent postmortem studies of stress disorders including major depressive disorder (MDD) (Duric et al., 2010; Kang et al., 2012; Labonte et al., 2017; Ota et al., 2014; Seney et al., 2018) and post-traumatic stress disorder (PTSD) (Girgenti et al., 2021; Licznierski et al., 2015; Logue et al., 2021; Young et al., 2015) have identified several hundred differentially expressed (DE) genes. Most postmortem studies of the MDD brain have focused on the prefrontal cortex, hippocampus and nucleus accumbens. Evidence has consistently pointed to GABAergic interneuron dysfunction, specifically the subtype that express the neuromodulatory somatostatin (SST), as having a significant role in MDD pathology (Levinson et al., 2010; Luscher et al., 2011). Transcript levels of SST have been reported to be reduced in the dorsolateral prefrontal cortex (DLPFC) and the anterior cingulate cortex (ACC) across cortical layers and the amygdala (Seney et al., 2015; Tripp et al., 2011). Studies in knock out mice of SST demonstrate a causal role in several MDD behavioral models, increased corticosterone, and reduction in brain-derived neurotrophic factor (BDNF) and Gad67 transcripts (Sounier and Sibille, 2014). Reductions in interneuron marker genes including parvalbumin (PV), VIP, and SST have also been reported in schizophrenia and autism spectrum disorder (Chung et al., 2018; Gandal et al., 2018a). However, there is inconsistent evidence of reduced PV and VIP expression in MDD (Fee et al., 2017) indicating significant
differences in the cell types driving the molecular pathology of these disorders.

Nearly 50% of patients diagnosed with PTSD are comorbid for MDD (Flory and Yehuda, 2015; Kessler et al., 1995; Rytwinski et al., 2013). Recently, the largest postmortem study of PTSD frontal cortex was performed using a matched MDD cohort to disentangle the differences between the two disorders (Girgenti et al., 2021). Similar to MDD, this study also discovered GABAergic signaling deficits in PTSD and identified the interneuron-specific transporter ELFN1 as significantly down regulated and associated with PTSD by transcriptomic imputation using the largest PTSD GWAS from the Million Veteran Program (Fig. 1). While this decrease in SST expression would seem to be a function of comorbidity between MDD and PTSD, formally testing the differentially expressed (DE) gene overlap revealed few DEGs in common: 111 DEGs in the DLPFC, 67 in the mOFC, 14 in the dACC, and 0 in the sgPFC with the most significant being ceruloplasmin (Cp), c-c motif chemokine ligand 2 (Ccl2), and ADAM Metallopeptidase with Thrombospondin Type 1 Motif 2 (Adamts2) (Girgenti et al., 2021). Further, the transcriptomic correlation between PTSD, MDD, and an aggregated, uniformely analyzed MDD profile from the PsychENCODE consortium (Gandal et al., 2018a) revealed no significant correlation between the transcriptomic patterns of the two disorders suggesting different molecular pathologies. However, it should be noted that there was also a lack of correlation between the MDD cohorts of both studies as well. While there are many possible reasons for this, we believe it stems from technical variance between meta-analyzed microarray data with RNA-seq data and differences in MDD diagnosis criteria. The meta-analysis by Gandal was from multiple labs and different brain banks making it likely that MDD cases varied especially for depression subtype, ancestry, trauma load, etc. Twelve DEGs were shared between cohorts and were consistent in fold change direction. Perhaps the most interesting common DEG was corticotropin releasing hormone 1 (CRH). CRH has been a fundamental neurobiological correlate of stress disorders and MDD in particular for twenty years (Holsboer, 2000; Pariante and Miller, 2001). It has been extensively studied for its role in regulating HPA axis function in the stressed brain (Menke, 2019) and the robustness of these findings are highlighted by its appearance in the two largely disparate datasets. On the other hand, a separate PTSD transcriptomic study found in this issue (Logue et al., 2021) compared expression changes between their PTSD cohort and the Girgenti PTSD cohort. 50% of this cohort overlaps with donors in the Girgenti study and they found that 17% of their nominally significant DEGs overlapped with the Girgenti cohort- significantly greater than would be expected by chance. Additionally, the Girgenti cohort is made up of tissue from two brain banks (UPMC and the VA National PTSD Brain Bank) and despite no overlap between donors in the UPMC group, the Logue cohort had significant overlap in DEGs with this group alone. Further, the Logue cohort found significant global correlation (r = 0.75, p < 2.2 × 10^{-16}) overall between their transcriptomic signature and the Girgenti cohort. The strong correlation between two PTSD transcriptomic data

![Cell type-specific transcriptomic changes in PTSD frontal cortex.](image-url)

**Fig. 1.** Cell type-specific transcriptomic changes in PTSD frontal cortex. A. GABA-related key drivers and transcripts exhibiting regional (DLPFC) down regulation in PTSD postmortem brain. B. Microglial marker genes are down regulated in PTSD frontal cortex.
sets versus weak correlation between two MDD transcriptomic datasets implies greater differences not only between MDD and PTSD as disorders but also in the way MDD is diagnosed or presents and suggests a need to better evaluate how we diagnose depression in future brain bank collections.

There is considerable evidence for the role of microglia in many neural processes including the stress response (Wohleb et al., 2016). Chronic stress induced overactivation of microglia has been implicated in reduced neurogenesis in the hippocampus and in synaptic protein reduction in prefrontal cortex of subjects with MDD (Wohleb, 2016). A role for microglial dysfunction in PTSD has been suspected based on blood transcriptomic work identifying inflammatory and immune related gene dysregulation (Passos et al., 2015). Recent work has identified dysregulation of microglia in the CNS PTSD transcriptome. The immune gene UBA-7 was identified by TWAS as associating with PTSD and is a significant transcriptomic key driver in females with PTSD (Cirgenti et al., 2021)(Fig. 1). Another study using this same dataset stratified their cohorts (PTSD and neurotypical controls) by normal and high BMI for each sex (Stone et al., 2020). They identified numerous DEGs across comparisons and identified the cytokine IL-1B as a putative upstream regulator of transcription in PTSD males with high BMI. Previous work in the periphery of PTSD subjects has identified regulation of IL-1B (Passos et al., 2015). However, high BMI alone has not been shown to regulate IL-1B levels in human prefrontal cortex (Lauridsen et al., 2017), suggesting a possible molecular intersection between PTSD and BMI in brain and may imply further functional implications as genetic variation in IL-1B has been linked to risk of PTSD in males (Hovhannisyan et al., 2017). Taken together, these transcriptomic findings point to vulnerabilities in neuroimmune function as promoting behavioral and neurobiological consequences of stress disorders.

7. Cell-type specificity of stress induced transcriptomic changes

One of the goals of gene regulation biology is to understand the contribution individual cell types have in regulating the transcriptome. The human brain is made of many diverse cell types each contributing its own unique molecular signature to a “bulk-tissue” RNA-seq experiment (Newman et al., 2019). One possible confound is that shifts in cell type proportions may accompany a particular neuropsychiatric disorder and that may affect our ability to translate a particular gene expression difference. There are currently several methods for cell type deconvolution of gene expression from bulk-tissue RNA-seq data including CIBERSORTx (Newman et al., 2019), Braininblender (Hagenauer et al., 2018), and Bisque (Jew et al., 2020). These programs quantify relative population proportions from bulk tissue transcriptomics by using cell type-specific expression profiles derived from single -cell/nuclei sequencing data from specific regions as a background dataset. CIBERSORTx was used in a study examining the transcriptome of the PTSD frontal cortex (Girgenti et al., 2021). The study identified regional changes in cell type proportions, most notably a significant increase in excitatory neurons and a decrease in microglia. Bioinformatic inference of cell type proportions and gene expression from bulk-tissue RNA-seq is limited, however. There are subtypes of interneurons and layer specific excitatory neurons that are generally missed and undoubtedly reflect important information on how gene regulation is altered in a given disorder. To overcome this hurdle several laser cell-capture microscopy techniques (LCM) have been developed. A recent LCM study of gene expression isolated granule cells from the dentate gyrus in a large cohort of schizophrenia, bipolar disorder, and major depressive postmortem tissue (Jaffe et al., 2020). Few MDD DEGs (7) were identified. The authors then compared a subset of their MDD cohort treated with an SSRI to those that were not. Interestingly, none of the DEGs identified were involved in adult neurogenesis, a process where there is considerable evidence of SSRIs function in treating depression (Eisch and Petrik, 2012). This finding and the relatively few genes identified across all diagnostic groups, highlights the limits of this technology. In addition, LCM suffers from antibody specificity issues and throughput - while it is relatively easy to isolate granule cells in the hippocampus it is much more difficult to adopt the described LCM methods to a more cell-diffuse region such as the prefrontal cortex.

To more robustly identify cell type-specific changes, isolation of individual cells or their nuclei is necessary for transcriptomic interrogation. Single-nuclei transcriptomics (snRNA-seq) is an emerging technology that has been used to study postmortem brains from several neuropsychiatric and neurological disorders however, there are currently few studies that have used this on postmortem tissue of stress disorders. The first study using postmortem diPFC of donors with MDD used snRNA-seq on 80,000 nuclei from 17 MDD donors and 17 normal controls (Nagy et al., 2020). The study used 10X Genomics Chromium version 2 and isolated mRNA from approximately 3000 nuclei per sample. They identified 26 clusters (transcriptional cell types) with most (60%) having significant numbers of DEGs. They identified 96 DEGs almost 50% of which occurred in the excitatory neuron and oligodendrocyte precursor cells (OPC) clusters. Three DEGs had been identified in previous MDD postmortem studies (FADS2, CKB, and KAZN) and 26 have been previously linked to mental illness including the synaptic genes GRIN2A and Synapsin 1. The authors point out that cell-type proportions are difficult to estimate from snRNA-seq datasets in large part because of the inconsistency in dissections. However, deconvolution of a bulk-tissue RNA-seq dataset from a separate, large MDD cohort did find reductions in the oligodendrocyte and OPC populations of the diPFC(Girgenti et al., 2021). This study (Nagy et al., 2020) was the first to use single cell type gene expression profiling and sets the groundwork necessary for future studies focused on disentangling the role of each cell type in the diseased and neurotypical brain.

8. Integration of neuroimaging and transcriptomics of stress disorders

Human neuroimaging using functional magnetic resonance imaging (MRI) and positron-emission tomography (PET) scanning have proved indispensable to understanding how brain regions are connected and formed and the role that particular molecules play in disease state. This interest has led to the creation of large consortia, such as ENIGMA whose goal is correlating genetic variation with image-derived phenotypes (Thompson et al., 2020). The ENIGMA consortium has made substantial progress in linking neuroimaging derived structural changes with genetic research through over 200 studies across a wide-array of neuropsychiatric disorders.

There is currently much interest in integrating neuroimaging with gene expression states rather than allelic variation (Arnatkevičiute et al., 2019; Fornito et al., 2018). Multiple factors outside of genetic architecture can affect a genes activity and as a result, the mechanisms through which variants influence how a phenotype manifests are unclear. This work correlates gene expression levels with variation in one or more imaging-derived phenotypes. Modern advances in large gene expression assays have allowed these studies to move from single gene correlates to genome-wide level. Gene expression assays provide a direct measure of genes activity and combined with neuroimaging can help resolve how spatial variation on the molecular level manifests on the structural level. Large multi-region transcriptomic atlases such as the Allen Human Brain Atlas which is comprised of measures for more than 20,000 genes from over 3702 spatially distinct brain tissue samples offers the greatest range of coverage for healthy brain gene expression. There is now unprecedented capacity to link gene function to structural brain organization particularly as relates to canonical resting-state networks; fiber connections between discrete regions; and perhaps most importantly for this review pathological changes in brain disorders.

Recent studies linking stress disorders by integrating human imaging with postmortem brain transcriptomics has been quite fruitful in advancing our understanding of the neurobiology of major depression.
and PTSD. Indeed, work in this area has shown that there is significantly higher cortical mGluR5 availability in PTSD in vivo using a radioisotop that binds to mGluR5 protein (Holmes et al., 2017). Dysfunction of the glutamate system has been implicated in trauma and stress psychopathology. Up regulation of the transcript SHANK1, a mGluR5 cell membrane anchor was identified in a concurrent postmortem transcriptomic study. Another recent study using PET imaging of the microglial marker TSPO found prefrontal-limbic availability was lower in patients with PTSD (Bhatt et al., 2020). The TSPO and microglial-associated genes TSPOAPI (an upstream regulator of TSPO) and TNFRSF14 were found to be down regulated in postmortem PTSD cortex (Fig. 1). These findings suggest that PTSD is associated with suppression of the immune response (through microglia) and not through neuroimmune activation as previously thought. Taken together, these studies highlight how combining neuroimaging with gene expression studies can link structural findings to a molecular mechanisms.

9. Conclusion

Transcriptomic studies have provided critical information on the pathophysiology of stress disorders. While genetic studies have illuminated the inherited genetic risks for these disorders, transcriptomics has provided a window into understanding the functional output of these risks and of the illness itself. Our current understanding of the molecular pathology of MDD suggests disruptions in GABAergic signaling, specifically in the SST interneuron subtype (Bajbouj et al., 2006; Levinson et al., 2010; Sanacora et al., 1999). SST expression is reduced in the frontal cortex and other regions and its reduction appears to correlate with symptom severity (Seney et al., 2015; Tripp et al., 2011). SST expression is also reduced in the frontal cortex of PTSD subjects and suggests a molecular intersection between these two highly comorbid disorders. Further, neuroimmune dysfunction has been reported in both MDD and PTSD though neuroimaging of live subjects and suggests that this occurs in opposing directions, with MDD being characterized with increases in inflammatory signaling and PTSD with immune suppression and reduced microglial signaling (Bhatt et al., 2020). It should be noted that there are currently no strongly identified genetic risk loci associated with GABAergic signaling or microglia, highlighting the need to move genetic work to functional genomic studies.

Future work in the transcriptomics of stress disorders should be focused in three directions: single cell RNA-sequencing, multi -omics integration and comparative overlap of animal models with postmortem findings. The brain is comprised of a myriad of cell types and it is critical that we understand how these individual cells contribute to the mechanistic function of the cell types involved, as some gene pathways may have effects in only single cell types or opposite effects in different cell types. Bulk-tissue transcriptomics would miss the roles of such genes and pathways. Multi-omic approaches should be explored as well. Currently, most genetic variants fall within non-coding regions and their impact on the transcriptome and by extension cell biology is obscured. snATAC-seq and HiC assays will be crucial for understanding how the genomic and nuclear structure of the cell regulates the transcriptome. Spatial transcriptomics is emerging as one of the most promising genomic technologies. The spatial organization of the brain defines many of its functions with different regions exhibiting differing patterns of cell morphology and physiology. Outside approaches such 3-dimensional organoids and exosomes will also be critical as will identifying the animal models with the closest molecular changes to those measured in human (postmortem) brain. There is much more to be gained through transcriptomic work in the field of stress-disorders and basic research in this area is needed to identify promising therapeutic targets.

CRediT authorship contribution statement

Jing Zhang: Writing – review & editing. Alfred P. Kaye: Writing – original draft. Jiawei Wang: Writing – original draft. Matthew J. Girgenti: Conceptualization, Funding acquisition, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors report no financial or biomedical COIs.

Acknowledgements

This work was supported by the Department of Veterans Affairs, Veteran Health Administration, VISN1 Career Development Award (M. J.G). The Glenn H. Greenberg Fund for Research on Stress and Resilience, 10x Genomics Pilot Grant (A.P.K.). The National Center for PTSD (A.P.K. and M.J.G), and Brain and Behavior Research Foundation Young Investigator Awards (A.P.K. and M.J.G.), and NIMH (MH123896 to J. Z.). The views expressed here are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs (VA) or the U.S. government. The data used for the analyses described in this manuscript were obtained from Girgenti et al. (Girgenti et al., 2021). The authors declare no biomedical financial interests or potential conflicts of interest.

References

Adhya, D., Swarup, V., Nagy, R., Dutan, L., Shum, C., Valencia-Alarcon, E.P., Jozwik, K. M., Mender, M.A., Horder, J., Loth, E., Nowoslaw, P., Lee, I., Skuse, D., Flinter, F.A., Murphy, D., McAlonan, G., Geschwind, D.H., Price, J., Carroll, J., Srivastava, D.P., Baron-Cohen, S., 2021. Atypical neurogenesis in induced pluripotent stem cells from autistic individuals. Biol. Psychiatry. 89, 486–496. https://doi.org/10.1016/j.biopsycho.2020.06.014.

Arambanicveiciene, A., Fulcher, B.D., Forntino, A., 2019. A practical guide to linking brain-wide gene expression and neuroimaging data. Neuroimage 189, 353–367. https://doi.org/10.1016/j.neuroimage.2019.01.011.

Bagot, R.C., Cates, H.M., Purushothaman, I., Lorsch, Z.S., Walker, D.M., Wang, J., Huang, X., Schuler, O.M., Mao, I., Pella, C.J., Heller, E.A., Isler, O., Wang, M., Song, W., Stein, J., Liao, Y., Doyle, M.A., Scobie, K.N., Sun, H.S., Neve, R.L., Geschwind, D., Dong, Y., Shen, L., Zhang, B., Nestler, E.J., 2016. Circuit-wide transcriptional profiling reveals brain region-specific genes networks regulating depression susceptibility. Neuron 90, 969–983. https://doi.org/10.1016/j.neuron.2016.04.015.

Bais, A.S., Kotaoka, D., 2019. Sceq: computational annotation of doublets in single-cell RNA sequencing data. Bioinformatics 36, 1150–1158. https://doi.org/10.1093/bioinformatics/btz698.

Bajbouj, M., Lisanby, S.H., Lang, U.E., Danker-Hopfe, H., Heuser, I., Neu, P., 2006. Evidence for impaired cortical inhibition in patients with unipolar major depression. Biol. Psychiatry. 59, 395–400. https://doi.org/10.1016/j.biopsych.2005.07.036.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., Zhao, K., 2007. High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837. https://doi.org/10.1016/j.cell.2007.05.009.

Bernstein, N.J., Fong, N.L., Lam, I., Roy, M.A., Hendrickson, D.G., Kelley, D.R., 2020. Solo: doublet identification in single-cell RNA-seq via semi-supervised deep learning. Cell Syst 11, 95–101. https://doi.org/10.1016/j.cels.2020.05.010 e5.

Bhatt, S., Hillmer, A.T., Girgenti, M.J., Rusowicz, A., Kapinos, M., Nabulsi, N., Huang, Y., Matsuky, D., Angarita, G.A., Esterlis, I., Davis, M.T., Southwick, S.M., Friedman, M. J., Girgenti, M.J., Friedman, M.J., Duman, R.S., Krystal, J.H., Duman, R.S., Caron, R.E., Krystal, J.H., Pietrzak, R.H., Cogrove, K.P., 2020. PTSD is associated with neuroimmune suppression: evidence from PET imaging and postmortem transcriptomic studies. Nat. Commun. 11, 2360. https://doi.org/10.1038/s41467-020-15229-5.

Bouland, G.A., Mahtouz, A., Reinders, M.J.T., 2021. Differential dropout analysis captures biological variation in single-cell RNA sequencing data. Bioxiv. https://doi.org/10.1101/2021.01.11.429187, 2021, 02.01.429187.

Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34, 525–527. https://doi.org/10.1038/nbt.3519.

Bryant, J., Garret, M.E., Song, L., Safi, A., Giusti-Rodriguez, P., Johnson, G.D., Shieh, A. W., Bull, A., Fullard, J.F., Rouxson, P., Sklar, P., Akbarian, S., Haroutunian, V., Stockmeier, C.A., Wray, G.A., White, K.P., Liu, C., Reddy, T.E., Ashley-Koch, A., Sullivan, P.F., Crawford, G.E., 2018. Evaluation of chromatin accessibility in prefrontal cortex of individuals with schizophrenia. Nat. Neurosci. 1158. https://doi.org/10.1038/s41467-018-05379-y.

Böttner, M., Miao, Z., Wolf, F.A., Treichmann, S.A., Theis, F.J., 2019. A test metric for assessing single-cell RNA-seq batch correction. Nat. Methods 16, 43–49. https://doi.org/10.1038/s41592-018-01924-1.

Cano-Gomez, E., Trynka, G., 2020. From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. Front. Genet. 11, 424. https://doi.org/10.3389/fgene.2020.00424.
Kleijn, K.M.A.D., Zuure, W.A., Peijnenborg, J., Heuvelmans, J.M., Martens, G.J.M., 2019. Jovanovic, T., Ressler, K.J., 2010. How the neurocircuitry and genetics of fear inhibition may inform our understanding of PTSD. Am. J. Psychiatr. 167, 648–662. https://doi.org/10.1038/nmeth.1528.

Lake, B.W., Zhao, J., Pajukanta, P., Halperin, E., 2020. Accurate estimation of cell composition in bulk RNA sequencing data using artificial nearest neighbors. Cell Syst 8, 315. https://doi.org/10.1016/j.cels.2020.04.015.

Meyer, E.R., Shimizu, T., Ye, Z., Strocchi, F., Schafer, A., Lin, C., Kier, J., Brouwer, R.G., Ferguson, S., et al., 2015. A highly sensitive multisystem discovery pipeline for the integration of single-cell RNA-seq data. Nature protocols 10, 326–344.

McGinnis, C.S., Murrow, L.M., Gartner, Z.J., 2019. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst 8, 315. https://doi.org/10.1016/j.cels.2018.10.019.

Mcloughlin, D., 2011. ITH: detecting intratumoral heterogeneity in RNA sequencing data. Genome biology 12, R106.

Meyers, L.A., Dekker, O., 2009. Analysis of somatic mutations in a genome-wide context. Nature 462, 898–903.

Meyerson, M., Wedge, D.C., Heath, I.K., Eshof, J., Narod, S.A., Speed, T.P., et al., 2008. Sensitive identification of somatic mutations in cancer genomes. Nature 452, 634–638.

Meyerson, M., Sanger, F., Corbett, S.E., et al., 2007. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.

Meyerson, M., Lee, C.S., Sanger, F., et al., 2006. A single-exon map of human genetic variation. Nature 449, 570–574.
