

Molecular Characterization of Depression Trait and State

One Sentence Summary: Integrating transcriptomic with various in-silico analyses identified cellular, molecular and putative biological causal pathways in trait and state depression

Authors: Rammohan Shukla^{1,2,*}, Dwight F. Newton^{2,6}, Akiko Sumitomo^{2,3,4}, Habil Zare⁵, Robert Mccullumsmith¹, David A. Lewis⁶, Toshifumi Tomoda^{2,3,4}, Etienne Sibille^{2,3,7,*}

Affiliations:

¹ Department of Neuroscience, University of Toledo, Toledo, Ohio, 43614, USA.

² Campbell Family Mental Health Research Institute of CAMH, Toronto, M5T 1R8 Canada.

³ Departments of Psychiatry, University of Toronto, Toronto, M5T 1R8 Canada.

⁴ Department of Research and Drug Discovery, Medical Innovation Center, Kyoto University, Graduate School of Medicine, Kyoto 606-8507, Japan.

⁵ Department of Computer Science, Texas state University, San Marcos, Texas, 78666, USA.

⁶ Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA.

⁷ Departments of Pharmacology and Toxicology, University of Toronto, M5T 1R8, Toronto, Canada.

*To whom correspondence should be addressed:

Etienne.Sibille@camh.ca

250 College Street, room 134, Toronto, ON M5T 1R8, Canada, Tel: 416-535-8501, ext 36571

Rammohan.Shukla@utoledo.edu

3000 Arlington Avenue, room 182, Toledo, Ohio 43614, USA, Tel: 419-383-4506

ABSTRACT

Major Depressive disorder (MDD) is a chronic and recurrent brain disorder characterized by episode and remission phases, and poor therapeutic responses. The molecular correlates of MDD have been investigated in case-control settings, but the biological changes associated with trait (regardless of episode/remission) or state (illness phases) remains largely unknown, hence preventing therapeutic opportunities. To address this gap, we generated transcriptome profiles in the subgenual anterior cingulate cortex of MDD subjects who died during a single or recurrent episode or when in remission. We show that biological changes associated with MDD trait (inflammation, immune activation, reduced bioenergetics) are distinct from those associated with MDD phases or state (neuronal structure and function, neurotransmission). On the cell-type level, gene variability in subsets of GABAergic interneurons positive for corticotropin-releasing hormone, somatostatin or vasoactive-intestinal peptide was associated with MDD phases. Applying a probabilistic Bayesian network approach, we next show that gene modules enriched for immune system activation, cytokine response and oxidative stress, may exert causal roles across MDD phases. Finally, using a database of drug-induced transcriptome perturbations, we show that MDD-induced changes in putative causal pathways are antagonized by families of drugs associated with clinical response, including dopaminergic and monoaminergic ligands, and uncover potential novel therapeutic targets. Collectively, these integrative transcriptome analyses provide novel insight into cellular and molecular pathologies associated with trait and state MDD, and a method of drug discovery focused on disease-causing pathways.

INTRODUCTION

Major depressive disorder (MDD) is the world leading cause of years lost due to disability, with an annual and lifetime prevalence of 6% and 18%, respectively (1). Episodic phases of MDD are characterized by heterogeneous symptoms, including low mood, anhedonia, cognitive impairments, and physiological symptoms (i.e. activity, weight) (2). For the majority of subjects, MDD follows a periodic trajectory of recurring depressive episodes of increasing severity, duration, and of progressive resistance to antidepressants, separated by gradually shortening and incomplete remission phases, leading to treatment-resistance and deteriorating functional fitness (fig.1A) (3). This clinical trajectory suggests the presence of an underlying trait-like and/or progressive neuropathology, which might differ from that of episode or remission “states” of MDD.

The subgenual anterior cingulate cortex (sgACC) lies at the intersection of bottom-up sensory input and top-down cortical control and is involved in integrated processing of emotions, including mood and reward. The activity of sgACC is increased in subjects with MDD, as well as in individuals with high neuroticism, fear of peer rejection and healthy humans during experimentally induced sadness (4–6). An effective antidepressant therapy reverses the hyperactivity of the sgACC, making it the target and suggested mediator for therapeutic effect of deep brain stimulation (7, 8). Magnetic resonance spectroscopy and transcranial magnetic stimulation studies suggest reduced gamma-aminobutyric acid (GABA) levels and cortical inhibition as a mechanism of sgACC dysregulation in MDD, which too appear to normalize after successful treatment (9–12). Large-scale transcriptomic studies in MDD postmortem sgACC samples demonstrate dysregulation in cytoskeleton, rearrangement of neuronal processes,

synaptic function and presynaptic neurotransmission (13, 14) associated with GABA and glutamate receptor signaling. At the cellular level, reduced glial and increased neuronal densities were reported, associated with reduction in axon and dendrites (15, 16). Reduced expression of markers for GABAergic interneurons targeting either the dendritic (e.g. somatostatin) or perisomatic (parvalbumin) compartments of pyramidal cells were reported, associated with reduced neurotrophic support (17–19).

To go beyond case-control studies and investigate molecular correlates of disease trait or state, we recently performed a large-scale proteomic study in a postmortem cohort of subjects in first or recurrent phase of episodes or remission, and in healthy controls (20). In addition to state-specific results, this study highlighted robust MDD trait-associated changes (i.e., regardless of episode or remission state), suggesting a continuous underlying pathology (20). However, limitations in detection and sensitivity of the proteomic approach precluded a full molecular characterization of the illness phases and putative progression. In contrast, RNAseq-based transcriptome analysis provides a broader snapshot of functional state and transcriptome-based systems biology analysis have proven appropriate for investigating complex diseases and putative causal pathways (21–23).

Here we hypothesize that RNAseq-based transcriptome profiling of human postmortem sgACC samples from patients in depressive or remission phases of MDD and in controls, combined with ontological, systems biology, Bayesian network and perturbagen-induced transcriptome analyses, would enable the molecular characterization of disease trait, state and chronicity, as well as putative causal biological changes and therapeutic targets.

RESULTS

MDD trait and state are characterized by distinct biological changes in sgACC

RNAseq-based expression profiles from sgACC samples were obtained from one control (C) and four cohorts of MDD subjects during a first episode (E_1), remission after first episode (R_1), recurrent episode (E_R) or remission after recurrent episodes (R_R) (fig.1A, [table S1](#)). Differential expression and biological pathway analyses were performed along the following group contrasts (fig.1A and Methods): 1) **MDD-all**: all MDD cohorts versus Controls, 2) **MDD-episode**: two MDD episode cohorts versus Controls, 3) **MDD-remission**: two remission cohorts versus controls, 4) **Episode/Remission**: two MDD episode cohorts versus two remission cohorts, 5) **MDD Progressive Episode**: monotonic increase or decrease across control and episode groups, 6) **MDD Progressive Remission**: monotonic increase or decrease across controls and remission groups; 7) **MDD-phasic**: capturing gene expression patterns coupled to the phasic changes across groups.

The differential expression analysis revealed distinct gene sets with expression changes matching the illness trait and state with the following features: Overall, upregulated genes were more common than downregulated genes across contrasts (fig.1B, [table S2](#)); The **MDD-all** contrast, representing trait, showed consistent up- or down-regulation patterns of gene expression regardless of episodes or remission phases compared to controls (fig.1C, top two panels); The **MDD-phasic** contrast revealed robust patterns of gene expression that oscillated with episodes and remission phases (see “waves” in bottom two panels in fig.1C). The MDD-phasic gene set overlapped substantially with the direct **Episode/Remission** contrasts (fig.1D), which captured the range of the “wave” pattern; Minimal overlap was observed between gene sets identified in

the **MDD-episode** and **MDD-remission** contrasts (fig.1D) representing MDD states, or between the **Progressive-episode** and **Progressive-remission** contrasts (less than 1%).

To summarize the vast data, gene expression changes across different contrasts were next analyzed at the level of biological pathways (p-value <0.01; FDR<0.25) and functional themes. Consistent with gene expression findings, upregulated pathways were more common than down-regulated pathways (fig.2, **Table S3**). For the **MDD-all contrast** defining the illness trait, multiple pathways associated with inflammation, immune system, angiogenesis and vascular growth factors were found to be upregulated. Other upregulated pathways included transcription associated with stress response and multiple aspects of protein modification processes (e.g., involving *JNK cascade*, *MAPK activity*, *auto-ubiquitination*; **table S3**). Few down-regulated pathways were observed and were associated with cellular bioenergetics (*Mitochondrial translation associated process*, *mitochondrial activity*, *ATP metabolic process*) and with *neuropeptide signalling*, including signalling and cell-type gene markers.

Results for the **MDD-episode contrast** were highly similar to **MDD-all** contrast (fig.2, Columns 1-2), including upregulated immune and inflammation-related pathways, upregulated translation related to stress response and aspect of protein modifications, as well as downregulated pathways related to mitochondrial bioenergetic functions. For neurotransmission, pathways associated with *Voltage gated potassium channels* and *transporter* activities were downregulated, in addition to the same prior neuropeptide signaling pathways. The **Progressive-episode** contrast was more selectively associated with upregulated pathways related to transcription, protein modifications, angiogenesis and immune system, suggesting increasing pathological changes related to these pathways with successive MDD episodes. Interestingly, several pathways related to cellular responses to hormones were selectively upregulated in this

contrast.

The **MDD-Remission contrast**, similar to the MDD-all and MDD-episode contrasts, was associated with upregulated inflammation- and immune-related pathways, although immune-related pathways were limited to innate versus both adaptive and innate immunity in the Episode contrast (fig.2, Columns 1-3). Notably, changes more specific to the **MDD-Remission** contrast included multiple upregulated pathways related to cytoskeleton, axonal and extracellular matrix, suggesting a structural cellular reorganization during remission phases of the illness. Changes associated with the **Progressive remission** contrast were associated with upregulated *bioenergetic and cellular respiration* pathways, upregulated cellular *organelle activity* (*Endoplasmic reticulum, Golgi bodies, Lysosomes Vacuoles and lumen, secretion*) (Fig.2), and with few pathways associated with inflammation and the immune system.

Finally, the direct **Episode/Remission contrast** was associated with downregulated bioenergetic-related pathways (*ATP, cellular respiration, and Mitochondrion-related pathways*), consistent with findings from the Episode- or Remission-contrasts. Surprisingly, we observed an additional massive downregulation of multiple pathways related to the structure and function of all neuronal compartments (*cytoskeleton, organelles, matrix, axon, dendrites synapse, channel, receptor/transmission dendrite, vesicle secretion, extracellular matrix*). These changes included pathways associated with multiple aspects of synaptic function and neurotransmission, such as downregulated neurotransmission mediated by glutamate, GABA, dopamine, catecholamines potassium and calcium voltage-gated channels, and multiple membrane transporter functions. This suggests the presence of multiple moderate state-dependent changes along different phases of episode and remission, affecting multiple aspects neuronal signaling, which are only detectable by direct comparison across opposite phases, and that are missed here when

comparing MDD subjects to control groups. Indeed, an analysis for changes matching the oscillatory nature of the disease trajectory (**MDD-phasic**) identified most functional themes of the Episode/Remission contrast (fig.2, Column 6), although to lower degrees than separate contrasts.

Together the results of the various group contrasts demonstrate the presence of two parallel pathological entities, one corresponding to a depressive trait, generally associated with inflammation, immune system activation and decreased bioenergetics, and a second state-dependent pathological entity matching the clinical phases of the illness, and associated with changes in the structure and function of cells, and with multiple aspects of neurotransmission, both generally negatively impacted during depressive states.

Validation of altered gene expression and pathway profiles

Results were validated using three approaches. First, a comparison to results obtained in a similar study performed by Pantazatos et al (24), using RNAseq in dorsolateral prefrontal cortex of MDD and control samples, revealed significant similarities in gene expression changes in that study with the MDD-episode (p-value=0.05), MDD-phasic (p-value=0.046) and Progressive episode (p-value=0.043) contrasts, and trend-level overlap of MDD-all (p-value=0.062). Second, we compared the current RNAseq results with previous mass spectrometry-based proteomics results in the same cohort (20). Gene and protein expression significantly correlated (n=3000 genes/proteins; $R=0.34$; $p=9.34 \times 10^{-94}$), and significant overlaps between protein and gene changes were observed for the MDD-episode (p-value=0.02), Episode/Remission (p-value=0.043), and trend-level overlap for the MDD-remission (p-value=0.062) contrasts. Third, technical validation was performed by qPCR for selected genes among affected pathways within

different contrasts (fig.3A). qPCR and RNAseq results were highly concordant (fig.3B).

RNA-seq data deconvolution identified a subpopulation of interneurons affected by phasic state changes in MDD

Our present data was derived from gray matter samples where relative cell proportions and associated biological signals are masked. To address this limitation, we used existing layer-specific single nucleus RNAseq data from ACC (from Allen brain atlas) and statistical methods to deconvolute the data. We identified 19 cell-type clusters (fig.4A), which we characterized in two stages: First, globally, based on known markers associated with pyramidal neurons (SLC17A4), interneurons (GAD1 or GAD2), astrocytes (GFAP), oligodendrocytes (OPALIN), microglia (CX3CR1) and neuroglia (PDGFRA) (fig.3B), and second, locally, based on the genes most enriched in each clusters (fig.4C). For instance, the GABAergic interneuron markers PVALB, SST, VIP and CRH were amongst the most enriched genes in cluster 5, 4, 7 and 8 respectively. Next, we identified sets of genes that clearly discriminated each of the 19 cell-type clusters (fig.4D, [table S4](#)). These discriminative gene sets were then used to estimate the relative proportion of each cell type in the MDD datasets obtained from each cohort, using a support vector machine (SVM) approach (25). Results show that cluster 8 displayed lower expression of its gene markers in episode compared to remission phases ($p < 0.042$, fig.4E, right panel), and that these changes followed a phasic trajectory between single and recurrent phases of the illness ($p < 0.016$, fig.4E, left panel). Cluster 8 corresponds to GABAergic interneurons expressing the SST, VIP and CRH gene markers (fig.4C). Notably, these markers correspond to the neuropeptide signaling GO group previously identified in the biological pathway analysis of the MDD-All, MDD-episode and MDD-Phasic contrasts (Fig.2). An analysis with layer-specific

single cell data further showed an enrichment of cluster 8 gene sets in in cortical layers 1 and 2/3 (fig.3F).

We next sought to validate these findings using gene expression datasets obtained in independent cohorts. We used a list of 566 genes obtained in a meta-analysis of altered gene expression in MDD across corticolimbic regions, including the sgACC (26) (table S5) and performed a hypergeometric overlap analysis in all 19 gene clusters (table S4). Within the interneuron clusters (Yellow in Fig 4B), the meta-analysis gene set overlapped significantly with the previously identified cluster 8 (q -value < 0.03), corresponding to interneurons expressing CRH, VIP and SST, hence validating the current results.

Next, we investigated the putative impact of altered Cluster 8 cell types (fig.4A-C) on overall biological changes. For this we focused on the most significant contrast (**Episode/Remission**) and compared gene expression differences with or without regressing out the variability associated with the putative cell-type proportions differences of cluster 8 (See Discussion). 721 (Up: 419, Down: 302 with respect to Remission) genes showed lower p -values after regression (table S6). These combined gene changes corresponded to the following downregulated biological pathways in episode: *voltage-gated potassium channel complex* ($q < 1.06 \times 10^{-2}$), *postsynaptic membrane* ($q < 2.7 \times 10^{-3}$), *asymmetric synapse* ($q < 3.88 \times 10^{-3}$), *axon-ensheathment* ($q < 4.8 \times 10^{-3}$) and *regulation of cell morphogenesis* ($q < 2.21 \times 10^{-2}$), and no upregulated pathways (table S6), consistent with the analysis of biological pathways associated with the Episode/Remission contrast (Fig. 2A, insert). Finally, to assess cluster 8 specific differential expression, we included the interaction between cluster 8 cell proportion and variables (i.e. Episode and remission state) of **Episode/Remission** contrast in a statistical linear model (27) we identified 75 upregulated and 175 downregulated (table S7, p -value < 0.01), corresponding to

upregulated *chemokine binding* (q-value < 0.04) and downregulated *nf-KappaB* (q-value < 0.01) pathways.

Taken together, inferring cell-type origin of gene expression revealed state-dependent phasic biological changes affecting upper layers GABAergic interneurons identified by the SST, VIP and CRH cell-type markers. Moreover, the data suggest that these cell type-dependent changes occur in synchrony with the broader reduced neuronal signaling phenotype observed in the MDD Episode/Remission contrast and may be directly affected by inflammation/immune-related events.

Investigating causality using Bayesian gene network analysis

Whereas the previous group contrasts and cell-based analyses provide biological and cellular insight into MDD state and phases, they do not inform on causal links. To identify putative causal pathways, we summarized the expression profiles across all five cohorts into modules of co-expressed genes using consensus weighted gene co-expression network analysis. This analysis reduces expression profiles across all five cohorts into common, more stable, and coherent functional modules of co-expressed genes that are comparable across all cohorts. The first principal component summarizing omnibus gene expression in each module (i.e. eigengene) was used to construct a Bayesian network. A Bayesian network models the probabilistic dependencies between gene modules and with the disease node, and is represented as a directed acyclic graph (DAG)(28–30), where edges represent the probabilistic dependencies between gene modules, organized from early to later module (referred as parent to child).

22 consensus modules were identified, containing from 31 to 4494 genes. 18 of these modules were significantly enriched for at least one gene ontology (GO) functional category at

q-value < 0.05 (fig.5A, table. S8) and were used to create the DAG (fig.5B). The direction of the edges and probabilistic dependencies between modules were first inspected for biological and theoretical consistency. For instance, module M05, the initial source of all modules, is enriched in pathways associated with *system development* (q-value<1.52x10⁻⁹) and connected to module M02 and M011, enriched in pathways associated with *nervous system development* (q-value<4.03x10⁻¹¹) and *anatomical structure morphogenesis* (q-value<6.5x10⁻³), respectively. Similarly, the known relationship between the immune system and cytokine (31) is captured by the high probability parent-child association between module M21, enriched with immune response (q-value<8.67x10⁻³), and module M10, enriched with pathways associated with cytokine response (q-value<4.79x10⁻²⁷).

The MDD node is linked to three DAG modules: module M13, enriched with *innate immune response* (q-value<1.60X10⁻¹⁷), module M10, enriched with *cytokine response* (q-value<1.31x10⁻²¹), and *response to stress* (q-value<1.72X10⁻¹⁷), and module M21, associated with *defense response* (q-value<1.04X10⁻²) and *stress response* (q-value<4.63X10⁻³) mostly associated with oxidative stress, as suggested by the child module M17, enriched with *response to oxidative stress* (q-value<4.6X10⁻³) and *cellular response to stress* (q-value<2.05x10⁻³, table. S8). Among the 3 MDD-associated modules, M10 had the highest probabilistic association to the disease module and is a child to disease-associated M13 and M21, making it the endpoint of the graph and the most disease-associated module. This is supported by the enrichment of M10 with upregulated pathways in all seven contrasts outlined above (fig.5A).

We next investigated the hypergeometric overlap of discriminant markers of the 19 cell-types clusters (fig.4D) with the MDD-associated modules. Module M10 was enriched in astrocyte (cluster 10, p < 0.02) and microglia (cluster 14, p < 5.3 x 10⁻⁹); module M13, in SST neurons

(cluster 04, $p < 0.03$) and microglia (cluster 14, $p < 5.55 \times 10^{-127}$) and module M21 was enriched in none of the cell markers. As an independent validation, we looked for hypergeometric overlap with cell specific gene set available from a previous study (32). Module M10 was enriched in microglia ($p < 0.01$), module M13 in microglia ($p < 2.23 \times 10^{-27}$) and module M21 in endothelial ($p < 0.03$). Overlap of module with lamina-specific gene sets (33) suggest the enrichment of all three modules in layer 1 (M10: $p < 1.2 \times 10^{-14}$; M13: $p < 2.29 \times 10^{-27}$; M21: $p < 0.03$) which also coincides with the significant enrichment of disease associated cluster 8 cells in superficial layer 1 (Figure 4F).

Together, the Bayesian network analysis organized the biological variability observed across the control and MDD cohorts (as measured by gene coexpression) into a coherent and directional graph. While these biological events occur in the context of the various clinical phases of MDD, the direct association of the disease module with three DAG modules suggests that immune and cellular stress functions associated with these modules are more proximal, and potentially causal, to the trajectory and clinical manifestation of the illness. Moreover, cell-enrichment analyses implicate superficial layer inhibitory neurons, glial and immune cells as potential mediators of putative causal biological events in MDD.

The MDD gene module identified in the Bayesian network is associated with known antidepressant drugs related pathways and novel drug targets

Based on the assumption that the MDD-associated DAG end-module M10 may have a causal role in the illness, we next investigated whether the M10 gene expression profile could be mimicked or antagonized by drugs. For this we probed the M10 associated gene-set against the connectivity map (cmap) database, a catalog of transcriptomic responses to multiple molecules in

diverse cell types. Despite many limitations of the interventions (simplified cell systems, single doses, acute drug exposure, few neuronal cell lines), the use of the cmap database has been instrumental in identifying novel therapeutic modalities in complex diseases (34), including in neurology and neuropsychiatry (35).

33 and 50 molecules were identified as antagonizing (fig.6) or mimicking (fig.7) the M10 gene expression profiles in neuronal cell lines, which, based on the signature reversion principal (34), may reflect therapeutic or MDD-inducing effects, respectively. Among the 33 M10-antagonizing molecules, three were either dopamine, its precursor (glutamyl-dopamine) or its structural variant (6-nitrodopamine), 15 (45.45%) have G coupled protein receptors (GPCRs) as their target class, of those 5 were dopaminergic, 4 were serotonergic receptors, and two targeted the serotonin and/or dopamine transporters. The other major class of the drug target includes enzyme (6/33), electrochemical transporter (4/33) and nuclear receptors (3/33). Notably, 7 of the identified drugs are either clinically used to treat neuropsychiatric disorders or identified as a potential drug in human postmortem studies (spermidine) (36).

Among the 50 MDD-mimicking molecules (fig. 7), 12 (24%) were of GPCRs target class, of those 6 were serotonin and 3 were dopamine receptors (mostly with opposing effects than compounds antagonizing the M10 expression profile), and one each for opioid, melatonin and adrenergic receptors. The other major class (> 10% of total) includes eraser (9/50, i.e. HDAC inhibitors) and protease (5/50).

In contrast, using the M16 gene module, which is part of the DAG, but not enriched in any of the MDD contrasts (fig.5A), only 7 and 9 molecules were identified with antagonizing and mimicking effects, respectively and the target class associated antagonizing molecules belonged to Eraser (3/7), nuclear receptor (2/7), kinase (1/7) and voltage gated channel (1/7).

DISCUSSION:

Unraveling the complex biological disturbances occurring in heterogeneous brain disorders, such as MDD, is a major challenge that is further complicated by the frequent trajectory of recurrent episode and remission phases of the illness. This symptomatic profile suggests that pathophysiological disturbances underlying MDD may wax and wane, however this contradicts the other clinical observation of progressive development of treatment-resistance and deteriorating functional fitness, suggesting a continuous and progressive pathophysiology. Here, using RNAseq-based transcriptome profiles from human postmortem sgACC samples obtained from MDD patients who died during depressive or remission phases, and from control subjects, we show the presence of two parallel pathological entities (fig. 1B). The first pathology corresponds to a stable MDD state, regardless of episode and remission and is associated with inflammation, immune activation, and reduced bioenergetics. The second is a state-dependent pathology, which consists of dynamic biological changes that match the clinical phases of the illness and that are mostly detected when directly comparing the transcriptomes of episode versus remission subjects. This state-related pathology is associated with neuronal structure and function, affecting both fast neurotransmitter (GABA, Glutamate) and neuromodulator (catecholamines, neuropeptides) systems, sets of changes that are downregulated during episodes and not observed or reversed during remission phases. Although these studies were performed using combined cortical gray matter and in cross-sectional cohorts, two additional *in silico* analyses provide perspective on cellular specificity and putative causal processes. First, cellular deconvolution approaches showed that variability in gene sets associated with GABAergic interneurons positive for corticotropin-releasing hormone, somatostatin or vasoactive-intestinal

peptide was associated with MDD phases. Second, a probabilistic Bayesian network approach showed that gene modules enriched for immune system activation, cytokine response and oxidative stress, may exert causal roles across MDD phases. Finally, using a database of drug-induced transcriptome perturbations, we show that MDD-induced changes in putative causal pathways are antagonized by families of drugs associated with clinical response, including dopaminergic and monoaminergic ligands, and suggest an innovative approach to uncover potential new therapeutic targets. Collectively, these integrative transcriptome analyses provide novel insight into cellular and molecular pathologies associated with trait and state MDD, and a method of drug discovery focused on disease-causing pathways.

Dissociating trait and state pathologies of MDD

The identification of two parallel pathological entities corresponding to trait and state MDD is graphically represented in the Figure 2 cluster graph, and supported by coherent results from the analysis of biological pathways associated with these patterns. Starting with the MDD trait pathology, the results show a robust link to inflammation and immune activation, consistent with prior reports (41). These events often recruit extensive signal transduction pathway protein modifications, such as phosphorylation of proteins in the MAPK pathway (ref), and increased demand on local blood supply, mediated by increased angiogenesis, which are both also associated here with MDD trait pathology. These findings are consistent with a large body of literature on MDD for inflammation (39), immune system activation (40), MAPK activation (ref), but less so for angiogenesis, as increasing angiogenesis has been proposed as a therapeutic avenue (ref).

Note that “trait” is defined based on statistical association with the MDD-All contrast, i.e.

subjects in episode or remission phases of the illness compared to controls, but some of the identified trait phenotypes displayed various degrees of significance in the separate contrasts. For instance, pathways related to innate immunity were consistently upregulated in all three contrasts, but changes in adaptive immunity and inflammatory response were limited to the MDD-All and MDD-episode, but not in the remission state, suggesting additional phasic state-dependent immune features. Reduced bioenergetics was identified as trait pathology, represented by decreased expression of genes implicated in mitochondrial structure and function, and in cellular energy production (i.e. ATP). This was observed in the MDD-All and MDD-Episode contrasts, but not the MDD-remission, suggesting a partial reversal of the phenotype during remission, as supported by significant results in the direct Episode/remission and phasic contrast analyses.

The MDD state pathology (as identified in group contrasts, but not the MDD-all contrast) was associated with classical neurotransmission changes often reported in depression (REFS), negatively affecting excitatory glutamatergic (NMDA and AMPA receptors), inhibitory GABAergic systems, multiple aspects of voltage-gated channel activities, and with reduced structural elements of neurotransmission (axons, dendrites, vesicle machinery, etc), consistent with anatomical studies showing reduced pyramidal neurons dendrites and sizes (Refs). The fact that these results are only observed when comparing opposite groups (i.e. Episode/Remission contrast) suggests moderate state-dependent changes affecting neuronal signaling that are more difficult to identify in traditional case-control studies and that would be more sensitive to cohort composition and sizes, potentially explaining past discrepancies across studies (refs). In contrast, neuronal structural components (cytoskeleton, matrix, axons) were restored during remission phases (i.e. significant in the MDD-Remission contrast), and a progressive upregulation of

intracellular vesicles associated with trafficking, and lysosomal and autophagy activities, was observed with successive episodes (i.e. Progressive-Remission contrast). These results show that the often-reported neurotransmission-related pathology of MDD is state-related and mostly resolved during remission of MDD, and that remission is further associated with restructuring of intracellular components, structural elements and extracellular cell to cell contacts.

CRH-, SST- and VIP-positive GABAergic cellular changes across MDD episodes and remission

The deconvolution of cell-type origin of gene expression obtained from gray matter samples (using independent sgACC single cell expression data) showed that a subpopulation of GABAergic inhibitory interneurons that express the CRH, SST and/or VIP neuropeptides are dynamically affected in synchrony with the episode and remission phases of MDD. Reduced CRH expression has been previously reported in subcortical brain regions, consistent with dysregulation of the hypothalamus pituitary adrenal stress axis in MDD, and we have previously reported reduced CRH expression in cortical layers as well (ref). Reduced SST expression has been reported by our group and others in several independent cohorts (refs). Reports on altered VIP expression in MDD are sparser (ref). Importantly, VIP, CRH and SST are markers of GABAergic interneuron subtypes that display overlapping patterns of expression, and that directly or indirectly inhibit pyramidal cell dendrites. These results are therefore consistent with a state-dependent MDD-related pathology that affects the regulation of excitatory input onto pyramidal cells, suggesting altered information processing in MDD (ref). At the sgACC regional level, reduced dendritic inhibition may alter the excitation inhibition balance and contribute to the elevated activity of this brain region during MDD episodes (refs).

The cellular deconvolution approach suggests variable proportions of a specific cell type between MDD episodes and remissions (Fig.3 D-E), but one could question how this can occur for interneurons, post-mitotic cells that seldom divide in the adult brain. Deconvolution analysis assumes that the gene expression of each cell type is linearly additive, making its contribution proportional to its fraction in bulk. Gene expression however, largely depends on other cell-level covariate, particularly cell size and function. For instance, we previously showed a reduced SST expression per cell in the sgACC of MDD subjects, rather than reduced SST-positive cell numbers (Refs). The association of change in cell size or of expression per cell for CRH, SST and VIP positive interneurons is also supported by facts that normalizing for putative cell-type differences overall enhanced the MDD-related significance of cell-to-cell signaling and neurotransmission, and that multiple genes implicated in cell structure (hence volume and function) are affected here in phase with MDD episode and remission states.

A putative causal role for immune function and inflammation in upper cortical layer cells in MDD

Bayesian network analysis is an approach that uses changes in association probabilities between gene modules across multiple network iterations to deduct potential causal flow. Applied to the dataset across control and MDD cohorts, this analysis suggests that activation of immune function, inflammation and oxidative stress originating from or affecting inhibitory neurons, glial, endothelial and immune cells in upper cortical layers may have causal role in MDD across episode and remission phases. The collective results may point to sustained immune activation, combined with other cellular stressors (oxidative stress, inflammation), whether of intrinsic or external origin, and implicating glial and endothelial cells, in turn affecting inhibitory neurons

located in upper cortical layers and regulating processing of cortical information. This collective pathology is maintained throughout the disease trajectory but appears to wax and wane with episodes and remission, as suggested by phasic changes in CRH-/VIP-/SST-neurons, reversal of neuronal structure during remission, shift from innate and adaptive immune activation during episodes receding to innate immunity only during remission, and paralleled by bioenergetics and cellular structure phase-dependent changes. In this way, the model suggests that the trait pathology associated with immune/inflammation increases the biological vulnerability of the sgACC, setting it up for relapse into episode state-like pathology, together demonstrating a novel type of plasticity associated with MDD.

It is intriguing that the gene expression profile of the Bayesian network gene module most associated with MDD can be antagonized or mimicked in cell lines by drugs and ligands that target the dopamine or monoamine systems, consistent with the clinic roles of these neurotransmitter systems in diseases and therapeutics. It is surprising that this information can be obtained from acute drug exposure in simple cell-based systems that lack the complexity of the brain and the timeframe of disease trajectory. This may reflect the fact that the complex interplays of multiple biological events across time and cell systems in MDD may recruit a unique combination of cellular processes, as identified by correlated gene coexpression patterns between complex modules obtained in combined gray matter tissue across disease states and those induced by specific compounds on simplified cell systems. This is also consistent with novel therapeutic or pro-disease targets suggested by this approach, which belong to complex biological regulators, such as epigenetic, nuclear receptors and protein modification regulators, rather than single target molecules. Importantly, these findings provide an independent confirmation of the validity of the causal gene modules and biological pathways identified in this

study. Moreover, the integrated approach applied in this study provides a framework and putative temporal sequence to assemble known hypotheses of MDD, namely suggesting immune- and cellular stress-related phenomenon as causal and upstream in series of event contributing to the state-dependent plasticity affecting excitatory and inhibitory neuron and transmission, as well as serotonin, glutamate and GABA dysfunction.

Limitations

These studies were performed in cross-sectional cohorts, and results suggesting causal links and sequence of events should be interpreted as hypothesis generating, to be tested in future studies. These studies were aimed at identifying broad biological events rather than specificities associated with demographic (sex, age) or clinical (antidepressant use, suicide). Finally, the unique feature of the current cohorts (remission and episodes) and availability of samples precluded the validation of the state-dependent aspects of the current results in independent cohort. Instead we have used available genomic datasets to provide complementary or additional supporting evidence for aspects of the current results.

MATERIAL AND METHODS:

Human postmortem brain samples. Postmortem brain samples were collected during routine autopsies performed at the Allegheny County Medical Examiner's Office with procedures approved by the University of Pittsburgh committee for oversight of study and clinical training involving the dead. Consensus DSM-IV diagnoses were made by an independent committee of experienced clinicians using information from structured interviews with family members, clinical records, toxicology results, and standardized psychological autopsies. A similar approach was employed to confirm the absence of a psychiatric diagnosis in comparison subjects. The MDD cohorts were carefully matched with controls to ensure that they did not differ in mean age, postmortem interval (PMI), brain pH or RNA integrity number (RIN). Ninety samples including 20 control subjects, 20 subjects in first MDD episode, 15 in remission after a single episode, 20 in recurrent episode and 15 in remission after recurring episodes (fig.1, table S1). Samples comprising all six cortical layers were collected from coronal sections, as previously described (42). Samples from the same cohort had been previously used in a large-scale proteomic study (20).

RNA extraction and sequencing library preparation. Total RNA was extracted from the sample homogenates using RNeasy Mini kit (Qiagen, Cat.No.74104) with in-column DNase treatment using RNAase-Free DNase (Qiagen, Cat.No.79254). Sequencing libraries were prepared using SMARTer Stranded Total RNA-seq kit (Clontech Laboratories, Cat. No. 634876). All steps involved were performed according to manufacturer's protocol.

Sequencing and data generation. Pooled libraries were sequenced in illumina HiSeq2500 to generate 2 x 100 paired end reads, which were aligned to human reference genome GRCh38

provided by Ensembl using HISAT2 aligner. Count data were then generated for the reads aligned to exons and transcripts using the GenomicFeature and GenomicAlignments packages in R and gene model (GTF file) provided by Ensembl. 9.7 million reads were obtained per sample, aligning to 21,684 genes.

Experimental contrasts and differential expression analysis. After removing low expressed genes (mean row sum ≤ 5 counts across all 90 samples), 21,684 genes were included in differential expression analysis. The experimental design allowed the following contrasts to be examined using the DESeq2 R package: 1) **MDD-all**: all MDD cohorts versus Controls, 2) **MDD-episode**: two MDD episode cohorts versus Controls ($E_n|C$), 3) **MDD-remission**: two remission cohorts versus controls, 4) **Episode/Remission**: two MDD episode cohorts versus two remission cohorts ($E_n|R_n$), 5) **Progressive Episode**: monotonic increase or decrease across control and episode groups; $C \rightarrow E_1 \rightarrow R_R$, 6) **Progressive Remission**: monotonic increase or decrease across controls and remission groups; $C \rightarrow R_1 \rightarrow R_R$) and 7) **MDD-phasic**: capturing gene expression patterns coupled to the phasic changes across groups. The curve fitting for MDD-phasic was performed in two steps in R. For the first step. $pData\$MDD-Phasic = \sin(-\pi/2 + 2 * \pi/2 * pData\$group)$ adds the column named MDD-phasic to the sample phenotype data used to model the sinusoidal nature of the cohort. Here $-\pi/2$ and $\pi/2$ captures the down and up phase respectively. 2 denotes the step ensuring the model searches for alternate down and up pattern. “pData” denotes phenotype data table and group denotes the cohort where the order is preserved as $C \rightarrow E_1 \rightarrow R_1 \rightarrow E_R \rightarrow R_R$. The second step involved using the MDD-phasic column in the phenotype data as contrast (similar to other contrast) for finding genes associated with it. For all contrasts, Likelihood ratio test was employed with a full design of $\sim Age+Sex+PMI+Contrast$ and a reduced design of $\sim Age+Sex+PMI$ to remove the effect of age, sex and postmortem

interval (PMI). A p-value threshold of 0.05 was used.

Pathway enrichment analysis. Biological pathways affected in different contrasts were determined using gene set enrichment analysis (GSEA) (43). The 21,684 genes ranked by Wald statistic were tested against the three Gene Ontologies (GO): Biological Process (GOBP), Molecular Function (GOMF) and Cellular Component (GOCC). Updated gene set (pathway) lists were obtained from the Bader lab (http://download.baderlab.org/EM_Genesets/). To compare the effect of MDD pathology across different contrasts, the normalized enrichment score of significant pathways (P-value<0.05, q-value<0.25) was used to generate heatmaps.

A focused analysis of *a priori* functional themes was performed to better identify the character of the biological changes in the enrichment results. Specifically, we selected significantly enriched pathways which either 1) contained the name of the functional theme based on a text-based query or 2) were identified as child terms (nested pathways) of pre-selected parent terms which represent the theme.

Hypergeometric analysis. To look for significant overlap between gene sets we performed hypergeometric overlap using Geneoverlap package in R. Background of 21196 genes and significance cutoff of p-value < 0.05 was used for all the analysis.

Quantitative polymerase chain reaction (qPCR): Differentially expressed (DE) genes belonging to a leading-edge subset (core set of transcripts that accounts for the enrichment signal) in our enrichment analysis were used to validate the differential expression results using qPCR. Top 5 samples representing either control or diseased state were selected based on their expression profile. Total RNA (same as used for generating the sequencing libraries) were reverse transcribed to cDNA using PrimeScript RT Master Mix (TaKaRa). cDNA, primers, and TB Green Premix ExTaq (Tli RNaseH Plus) (TaKaRa) were mixed in 96-well PCR plate, and

qPCR was performed in triplicate using CFX96 Real-Time System (Bio-Rad). Results were normalized to GAPDH internal control. Primers used are described in Supplementary Table 2.

Cellular Deconvolution analysis: To estimate the cell type proportion and identify putative disease-associated cellular differences, we adapted the deconvolution analysis described by Baron et. al. (44) and implemented in bscqsc package in R. The analysis consists of five steps.

A) Identifying cell-type-specific marker-genes: Using the single nucleus data set for Anterior cingulate cortex available from Allen brain atlas, we first identified cluster of cell types (fig.3A) and markers specific to each cluster (table. S9) using SEURAT package in R for single cell analysis (45). In order to segregate cell clusters based on subtle differences in expression, the resolution parameter was set to 1.2

B) Building the reference basis matrix of marker-genes (fig.3D): This matrix contains the highly discriminatory marker-gene expression of each cell-type cluster averaged across all the cells of a given cell type cluster. For a given cluster, we considered a gene specific to a cluster only when it showed ≥ 3 -fold difference in expression when compared to its expression in other cell clusters.

C) Estimating proportions: The resulting reference matrix (fig.3D) was used to estimate cell proportion in different cohorts and contrast (fig.3E) using Support Vector Regression implemented by CIBERSORT package in R (46).

D) Adjusting the bulk tissue gene expression differences for proportion: We statistically regress out the effect of a cell types cluster (cluster 8) which showed the significant change in proportion in Episode/Remission contrast. This was implemented by expanding the design-model used for finding differential expression between episode and remission cohort by incorporating the estimated cell-type proportion (from step C) in the design matrix. Note that this step allows for differential expression analysis independent of variation introduced by cell proportion. Comparing the differentially expressed gene with and without the adjusted cell-type proportion

can be used to identify genes which show increased statistical significance (i.e. further reduced p-value) after removing the effect of cells showing highest change in proportion for a given contrast. These set of genes and the pathways associated with them can be considered as those influenced (thus consequential). **E) Finding cell-type specific differential expression**: As described previously (47) and implemented in the csSAM (a package in R) we use the two separate differential expression analyses, (i.e. with and without regressing the cell proportion) to fit a linear model that estimates interaction term between cell proportion and contrast to find cell type cluster specific differential expression

Bayesian Network analysis. To find probabilistic causal associations of disease states with gene co-expression modules representing different biological themes we used weighted gene coexpression network analysis (WGCNA) and Pigengene package in R. First, we used consensus WGCNA to generate co-expression modules common across samples and compared them between control and disease states. For each cohort, count data normalized based on size factor (using DESeq2 package in R) was used to create a matrix of pairwise Pearson correlations between genes, which was then transformed to a signed adjacency matrix using power $\beta=12$. To calculate the interconnectedness between genes, which defines “modules”, we derived the topological overlap, which gives biologically meaningful measurement of similarity between two genes, based on their co-expression relationship with all other genes. We then identified consensus modules between the control and the diseased cohorts using blockwiseConsensusModules function with consensusQuantile setting set to 0.50. We obtained 22-consensus modules with numeric label 1 to 22 and color labels for all. Label 0 and color grey was assigned to genes not assigned to any module and was removed from further analysis. The identified modules were functionally characterized using all three Gene Ontologies and top

pathways with Bonferroni corrected p-value < 0.05 were used to label the modules.

The first principal component (the eigengene) summarizes a given module by accounting for the maximum variability of all its constituent genes. It is used as a proxy for the overall gene expression of this module and to identify mechanisms associated with disease states.

Using the Pigengene R package, module eigengenes were used as random variables to train a Bayesian network that modeled the probabilistic dependencies between all modules. Additionally, experimental group was included as a categorical variable (disease) to simplify the inference. The disease node works as a source and has no parent terms above it. This is implemented by blacklisting all the incoming edges to the disease node. To search for an optimal Bayesian network fitting our data, we performed 1000 permutations (i.e. 1000 networks) with default parameters in the Pigengen package. Edges that appeared in greater than equal to 30% of the networks were used to construct the network.

Connectivity map (Cmap) analysis. Cmap contains expression profiles induced by ~19,000 different small molecules (perturbagens) in ~77 different cell lines (48). Genes in module 10 were submitted as query in the Cmap API (<https://clue.io/query>). The output file (cs_nlx476251.gct) containing the raw weighted connectivity scores for all molecules, or perturbagens, tested in different experimental conditions and cell types was used for further analysis. A connectivity scores summarizes the similarity between the query and signature profile of a perturbagen based on Kolmogorov-Smirnov enrichment statistics. To compare an observed connectivity score to all others in the database we calculate the percentile score, referred as “tau”. The weighted connectivity score ranges from -1 to +1 and accordingly the tau ranged from -100 to +100. Negative and positive scores indicate dissimilarity and similarity between query and perturbagen signatures, respectively. In this study we restricted our search space to all perturbagen tested in 9

hallmark cancer cell lines (A375, A549, HA1E, HCC515, HEPG2, HT29, MCF7, PC3, VACP) frequently used in drug repositioning studies and three neuronal cell lines (NEU, NEC.KCL, NPC) available in the Cmap database. We summarized the multiple experimental conditions under which a given perturbagen was tested in a given cell line separately for each cancer and neuronal cell lines based on the largest effect size magnitude between 33rd and 67th percentile. To select the top perturbagen candidate, we used the rank ordered row sum of tau score, separately for cancer and neuronal cell lines.

Drug target identification: Using the available two-dimensional (2D) structures (<http://lincsportal.ccs.miami.edu/SmallMolecules/catalog>) for top drugs associated with Module M10 and M16 molecules we searched for protein target and target class (protein family to which the target belongs) using SwissTargetPrediction, a web tool based working on structure function relationship of drugs and biomolecule. The frequency of different target class associated with the top 20 therapeutic drugs predicted based on neuronal cell lines was used to estimate the difference between drugs associated with M10 and M16 using chi-square test of independence. A drug can potentially dock to many protein targets of different target class and can hinder the calculation target class frequency. To avoid this issue we estimated the frequency based on the first predicted class target class for all the drugs.

FIGURE LEGENDS

Fig.1: RNAseq-based identification of state-dependent and phasic molecular changes in MDD. **A)** Clinical evidence frequently shows recurrent episodes of MDD (valleys) of increasing severity, reduced therapeutic response and shorter remission periods (crests), suggesting the presence of a progressive underlying pathology across phases (black arrow). **B)** Numbers of differentially expressed genes associated with different MDD phase contrasts ($p < 0.05$). The text (top) and graphics (bottom) describe the contrasts used in the text and figures. **MDD-all:** all MDD cohorts versus Controls. **MDD-episode:** two MDD episode cohorts versus Controls ($E_n|C$). **MDD-remission:** two remission cohorts versus controls. **Episode/Remission:** two MDD episode cohorts versus two remission cohorts ($E_n|R_n$). **MDD-phasic:** capturing the phasic up- and downregulated gene expression changes correlated with episode and remission phases. **Progressive Episode:** monotonic increase or decrease across control and episode groups; $C \rightarrow E_1 \rightarrow R_R$. **Progressive Remission:** monotonic increase or decrease across controls and remission groups; $C \rightarrow R_1 \rightarrow R_R$. **C)** Heatmap of gene expression changes corresponding to the MDD-all (top 2 panels) and MDD-phasic (bottom two panels) contrasts. The colored bars at the bottom of each panel are visual summaries of gene effects within the section above. The MDD-all contrast shows consistent up or downregulation of genes across the various phases of MDD, compared to controls. The MDD-phasic contrast shows “waves” of gene changes that coincide with the episode and remission phases of MDD. **D)** Venn Diagram showing intersection of up- and down-regulated genes associated with MDD-all, MDD-episode, MDD-remission, Episode|remission and Phasic contrast. Note the highest intersection between Episode|Remission and Phasic contrasts.

Fig.2: Profiles of biological pathway affected in the MDD group contrast and validation studies. Altered pathways associated with different biological themes (left) per contrast. The blue and yellow colors represent up and down regulated pathways scaled by enrichment scores, ranging from +1 (Blue) to -1 (Orange).

Fig.3: qPCR validation of differentially expressed genes and representative of affected pathways in each contrast. In each image (from **A1-A7**) left white box plots represent RNAseq data. Right grey box plots represent qPCR data (n=5/group) **A1) *RHOB***: belonging to GO group ‘angiogenesis’ and up regulated in all MDD subjects compared to control in MDD-all contrast. **A2) *CRYM***: belonging to GO group ‘mitochondria’ and ‘oxidation-reduction process’ and down-regulated in Episodic subjects compared to control in MDD-episode contrast. **A3) *ARID5B***: belonging to GO group ‘regulation of transcription, DNA-templated’ and progressively Up-regulated in episodic states. **A4) *CNTN2***: belonging to GO group ‘microtubule cytoskeleton organization’ and Up-regulated in Remission subjects compared to control in MDD-remission contrast. **A5) *DDX25***: belonging to GO group ‘regulation of translation’ and progressively Upregulated in remission states. **A6) *STX1A***: belonging to GO group ‘presynaptic membrane’ and Upregulated in Remission subjects compared to episodic subject in Episode|Remission contrast. **A7) *VEGFA***: belonging to GO group ‘angiogenesis’ and coupled with phasic changes involved in depression. P-values < 0.05 for all independent RNAseq and qPCR group comparisons. B) Correlation between RNAseq and qPCR was calculated for all the genes selected for validation (**A1-A7**). Data from all the gene were pooled and each point in the correlation plot represents one subject.

Fig.4: Cell-type deconvolution of gray matter RNAseq reveals gene expression changes in synchrony with MDD phases for CRH-, SST- and VIP-positive GABAergic interneurons.

A) Clustering of human anterior cingulate cortex single nuclei RNAseq data from Allen Brain Atlas. B) Clusters were characterized globally using known markers of Pyramidal neurons, interneurons, astrocytes, Oligodendrocytes, microglia and polydendrocytes (see results section for markers). C) Clusters were characterizing locally using the top enriched expressed genes (upregulated) by cluster. C₁) Cluster 5 enriched with PVALB. C₂) Cluster 4 enriched with SST marker. C₃) Cluster 7 enriched with VIP. C₄) Cluster 8 enriched with CRH. Note that some clusters have overlapping expression of different markers. For instance, cluster 8 (CRH) is also enriched in SST and VIP. D) Heatmap of highly discriminative gene, i.e. markers used as reference for estimating relative cell proportions in (E). Rows represent genes and columns represent the cluster identified in (A). Only representative clusters are shown here. For a heatmap of all the clusters see supplementary fig.X. E) Proportion differences of cluster 8 cell types in Episode/remission contrast (top) and phasic contrast (bottom). The y-axis shows the relative proportion of cluster 8 cell types while the x-axis represents different cohort used in the study F) Layer specific distribution of cluster 8 neurons (yellow). Note the high enrichment of cluster 8 neurons in supragranular layer 1 and 2/3.

Fig.5: Prioritizing putative causal gene modules in MDD using Bayesian network: A)

Consensus gene modules were identified using the combined RNAseq data across diseases phases and characterized using GO enrichment. GO terms highly enriched for each module are shown on the left, and $-\log_{10}$ p-values ($\text{fdr} < 0.05$) of enrichment tests for the GO term are shown on the right. The middle panels show the enrichment of modules in the MDD contrasts (From

Fig.1B). The blue and yellow circle shown enrichment in up- and down- regulated genes for each module. The size of circles is proportional to the $-\log_{10}$ of p-value associated the enrichment. Note that module ME03, ME08, ME19 and ME22 were not associated with any GO terms and were removed from further analysis. **B)** Directed acyclic graph (DAG) obtained after fitting Bayesian network to the eigengenes of characterized consensus modules, shown as nodes. The thickness of the edges connecting the nodes is proportional to the number of times, in percent, the edge was detected in 1000 permutation used to fit the Bayesian network. The connection of “Disease” node (methods) to its child nodes are shown in red.

Fig.6: Connectivity mapping of antagonizing molecules: Dissimilarity (antagonizing effect), represented by tau (see methods), of drug-induced transcriptomic profiles and module 10 gene sets is shown. Drug-induced profiles are derived neuronal cell lines. Significant Tau values of <90 are shown in increasing shades of blue. Molecules in red are associated with either monoamine or catecholamine transmitter while those in green with other known neurotransmitter. The target class column represents predicted protein targets using SwissTargetPrediction. Mode of action and notes taken from pubchem

Fig.7: Connectivity mapping of disease mimicking molecules: Similarity (**disease mimicking effect**), represented by tau (see methods), of drug-induced transcriptomic profiles and module 10 gene sets is shown. Drug-induced profiles are derived neuronal cell lines. Significant Tau values of >90 are shown in increasing shades of orange. The target class column represents predicted protein targets using SwissTargetPrediction. Mode of action and notes taken from pubchem.

REFERENCES:

1. G. S. Malhi, J. J. Mann, Depression., *Lancet (London, England)* **392**, 2299–2312 (2018).
2. C. Otte, S. M. Gold, B. W. Penninx, C. M. Pariante, A. Etkin, M. Fava, D. C. Mohr, A. F. Schatzberg, Major depressive disorder, *Nat. Rev. Dis. Prim.* **2**, 16065 (2016).
3. E. Sibille, B. French, Biological substrates underpinning diagnosis of major depression, *Int. J. Neuropsychopharmacol.* (2013), doi:10.1017/S1461145713000436.
4. B. W. Haas, K. Omura, R. T. Constable, T. Canli, Emotional Conflict and Neuroticism: Personality-Dependent Activation in the Amygdala and Subgenual Anterior Cingulate, *Behav. Neurosci.* (2007), doi:10.1037/0735-7044.121.2.249.
5. C. L. Masten, N. I. Eisenberger, L. A. Borofsky, K. Mcnealy, J. H. Pfeifer, M. Dapretto, Subgenual anterior cingulate responses to peer rejection: A marker of adolescents' risk for depression, *Dev. Psychopathol.* (2011), doi:10.1017/S0954579410000799.
6. H. S. Mayberg, M. Liotti, S. K. Brannan, S. McGinnis, R. K. Mahurin, P. A. Jerabek, J. A. Silva, J. L. Tekell, C. C. Martin, J. L. Lancaster, P. T. Fox, Reciprocal limbic-cortical function and negative mood: Converging PET findings in depression and normal sadness, *Am. J. Psychiatry* (1999).
7. H. S. Mayberg, A. M. Lozano, V. Voon, H. E. McNeely, D. Seminowicz, C. Hamani, J. M. Schwalb, S. H. Kennedy, Deep brain stimulation for treatment-resistant depression, *Neuron* **45**, 651–660 (2005).
8. A. Merkl, W. J. Neumann, J. Huebl, S. Aust, A. Horn, J. K. Krauss, I. Dziobek, J. Kuhn, G. H. Schneider, M. Bajbouj, A. A. Kühn, Modulation of Beta-Band Activity in the Subgenual Anterior Cingulate Cortex during Emotional Empathy in Treatment-Resistant Depression, *Cereb. Cortex* (2016), doi:10.1093/cercor/bhv100.

9. G. Hasler, J. W. van der Veen, T. Tumonis, N. Meyers, J. Shen, W. C. Drevets, Reduced Prefrontal Glutamate/Glutamine and γ -Aminobutyric Acid Levels in Major Depression Determined Using Proton Magnetic Resonance Spectroscopy, *Arch. Gen. Psychiatry* **64**, 193 (2007).
10. G. Sanacora, G. F. Mason, D. L. Rothman, K. L. Behar, F. Hyder, O. A. C. Petroff, R. M. Berman, D. S. Charney, J. H. Krystal, Reduced cortical γ -aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy, *Arch. Gen. Psychiatry* (1999), doi:10.1001/archpsyc.56.11.1043.
11. G. Sanacora, G. F. Mason, D. L. Rothman, J. H. Krystal, Increased occipital cortex GABA concentrations in depressed patients after therapy with selective serotonin reuptake inhibitors, *Am. J. Psychiatry* (2002), doi:10.1176/appi.ajp.159.4.663.
12. G. Sanacora, D. L. Rothman, G. Mason, J. H. Krystal, in *Annals of the New York Academy of Sciences*, (2003).
13. A. Sequeira, F. Mamdani, C. Ernst, M. P. Vawter, W. E. Bunney, V. Lebel, S. Rehal, T. Klempan, A. Gratton, C. Benkelfat, G. A. Rouleau, N. Mechawar, G. Turecki, Global brain gene expression analysis links Glutamatergic and GABAergic alterations to suicide and major depression, *PLoS One* (2009), doi:10.1371/journal.pone.0006585.
14. V. Duric, M. Banasr, C. A. Stockmeier, A. A. Simen, S. S. Newton, J. C. Overholser, G. J. Jurjus, L. Dieter, R. S. Duman, Altered expression of synapse and glutamate related genes in post-mortem hippocampus of depressed subjects, *Int. J. Neuropsychopharmacol.* (2013), doi:10.1017/S1461145712000016.
15. D. Ongur, W. C. Drevets, J. L. Price, Glial reduction in the subgenual prefrontal cortex in mood disorders, *Proc. Natl. Acad. Sci.* (2002), doi:10.1073/pnas.95.22.13290.

16. A. Medina, S. J. Watson, W. Bunney, R. M. Myers, A. Schatzberg, J. Barchas, H. Akil, R. C. Thompson, Evidence for alterations of the glial syncytial function in major depressive disorder, *J. Psychiatr. Res.* (2016), doi:10.1016/j.jpsychires.2015.10.010.
17. E. Sibille, H. M. Morris, R. S. Kota, D. A. Lewis, GABA-related transcripts in the dorsolateral prefrontal cortex in mood disorders., *Int. J. Neuropsychopharmacol.* **14**, 721–34 (2011).
18. A. Tripp, R. S. Kota, D. A. Lewis, E. Sibille, Reduced somatostatin in subgenual anterior cingulate cortex in major depression., *Neurobiol. Dis.* **42**, 116–124 (2011).
19. A. Tripp, H. Oh, J.-P. P. Guilloux, K. Martinowich, D. A. Lewis, E. Sibille, Brain-derived neurotrophic factor signaling and subgenual anterior cingulate cortex dysfunction in major depressive disorder, *Am. J. Psychiatry* **169**, 1194–202 (2012).
20. E. Scifo, M. Pabba, F. Kapadia, T. Ma, D. A. Lewis, G. C. Tseng, E. Sibille, Sustained Molecular Pathology Across Episodes and Remission in Major Depressive Disorder, *Biol. Psychiatry* (2018), doi:10.1016/j.biopsych.2017.08.008.
21. Y. Chen, J. Zhu, P. Y. Lum, X. Yang, S. Pinto, D. J. MacNeil, C. Zhang, J. Lamb, S. Edwards, S. K. Sieberts, A. Leonardson, L. W. Castellini, S. Wang, M. F. Champy, B. Zhang, V. Emilsson, S. Doss, A. Ghazalpour, S. Horvath, T. A. Drake, A. J. Lusis, E. E. Schadt, Variations in DNA elucidate molecular networks that cause disease, *Nature* (2008), doi:10.1038/nature06757.
22. B. Zhang, C. Gaiteri, L.-G. G. Bodea, Z. Wang, J. McElwee, A. A. A. Podtelezhnikov, C. Zhang, T. Xie, L. Tran, R. Dobrin, E. Fluder, B. Clurman, S. Melquist, M. Narayanan, C. Suver, H. Shah, M. Mahajan, T. Gillis, J. Mysore, M. E. E. MacDonald, J. R. R. Lamb, D. A. A. Bennett, C. Molony, D. J. J. Stone, V. Gudnason, A. J. J. Myers, E. E. E. Schadt, H.

- Neumann, J. Zhu, V. Emilsson, Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease, *Cell* **153**, 707–720 (2013).
23. V. Emilsson, G. Thorleifsson, B. Zhang, A. S. Leonardson, F. Zink, J. Zhu, S. Carlson, A. Helgason, G. B. Walters, S. Gunnarsdottir, M. Mouy, V. Steinthorsdottir, G. H. Eiriksdottir, G. Bjornsdottir, I. Reynisdottir, D. Gudbjartsson, A. Helgadottir, A. Jonasdottir, A. Jonasdottir, U. Styrkarsdottir, S. Gretarsdottir, K. P. Magnusson, H. Stefansson, R. Fossdal, K. Kristjansson, H. G. Gislason, T. Stefansson, B. G. Leifsson, U. Thorsteinsdottir, J. R. Lamb, J. R. Gulcher, M. L. Reitman, A. Kong, E. E. Schadt, K. Stefansson, Genetics of gene expression and its effect on disease, *Nature* (2008), doi:10.1038/nature06758.
24. S. P. Pantazatos, Y.-Y. Huang, G. B. Rosoklija, A. J. Dwork, V. Arango, J. J. Mann, Whole-transcriptome brain expression and exon-usage profiling in major depression and suicide: evidence for altered glial, endothelial and ATPase activity., *Mol. Psychiatry* **22**, 760–773 (2017).
25. A. M. Newman, C. L. Liu, M. R. Green, A. J. Gentles, W. Feng, Y. Xu, C. D. Hoang, M. Diehn, A. A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, *Nat. Methods* **12**, 453–457 (2015).
26. Y. Ding, L.-C. Chang, X. Wang, J.-P. Guilloux, J. Parrish, H. Oh, B. J. French, D. A. Lewis, G. C. Tseng, E. Sibille, Molecular and Genetic Characterization of Depression: Overlap with other Psychiatric Disorders and Aging., *Mol. neuropsychiatry* **1**, 1–12 (2015).
27. S. S. Shen-Orr, R. Tibshirani, P. Khatri, D. L. Bodian, F. Staedtler, N. M. Perry, T. Hastie, M. M. Sarwal, M. M. Davis, A. J. Butte, Cell type-specific gene expression differences in complex tissues, *Nat. Methods* (2010), doi:10.1038/nmeth.1439.
28. P. Langfelder, S. Horvath, Eigengene networks for studying the relationships between co-

- expression modules, *BMC Syst. Biol.* **1**, 54 (2007).
29. S. Mostafavi, C. Gaiteri, S. E. Sullivan, C. C. White, S. Tasaki, J. Xu, M. Taga, H. U. Klein, E. Patrick, V. Komashko, C. McCabe, R. Smith, E. M. Bradshaw, D. E. Root, A. Regev, L. Yu, L. B. Chibnik, J. A. Schneider, T. L. Young-Pearse, D. A. Bennett, P. L. De Jager, A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease, *Nat. Neurosci.* (2018), doi:10.1038/s41593-018-0154-9.
 30. R. Agrahari, A. Foroushani, T. R. Docking, L. Chang, G. Duns, M. Hudoba, A. Karsan, H. Zare, Applications of Bayesian network models in predicting types of hematological malignancies, *Sci. Rep.* (2018), doi:10.1038/s41598-018-24758-5.
 31. H. Anisman, Z. Merali, Cytokines, stress and depressive illness: Brain-immune interactions *Ann. Med.* (2003), doi:10.1080/07853890310004075.
 32. S. Darmanis, S. A. Sloan, Y. Zhang, M. Enge, C. Caneda, L. M. Shuer, M. G. Hayden Gephart, B. A. Barres, S. R. Quake, A survey of human brain transcriptome diversity at the single cell level., *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7285–90 (2015).
 33. Z. He, D. Han, O. Efimova, P. Guijarro, Q. Yu, A. Oleksiak, S. Jiang, K. Anokhin, B. Velichkovsky, S. Grünewald, P. Khaitovich, Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques, *Nat. Neurosci.* (2017), doi:10.1038/nn.4548.
 34. J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J.-P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, T. R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.,

- Science* **313**, 1929–35 (2006).
35. C. R. Sullivan, R. H. Koene, K. Hasselfeld, S. M. O'Donovan, A. Ramsey, R. E. McCullumsmith, Neuron-specific deficits of bioenergetic processes in the dorsolateral prefrontal cortex in schizophrenia, *Mol. Psychiatry* (2019), doi:10.1038/s41380-018-0035-3.
 36. J. Gross, G. Turecki, Suicide and the Polyamine System, *CNS Neurol. Disord. - Drug Targets* (2013), doi:10.2174/18715273113129990095.
 37. S. H. Mellon, O. M. Wolkowitz, M. D. Schonemann, E. S. Epel, R. Rosser, H. B. Burke, L. Mahan, V. I. Reus, D. Stamatiou, C.-C. Liew, S. W. Cole, Alterations in leukocyte transcriptional control pathway activity associated with major depressive disorder and antidepressant treatment., *Transl. Psychiatry* **6**, e821 (2016).
 38. T. Xie, M. G. Stathopoulou, F. De Andrés, G. Siest, H. Murray, M. Martin, J. Cobaleda, A. Delgado, J. Lamont, E. Penás-Liedó, A. Llerena, S. Visvikis-Siest, VEGF-related polymorphisms identified by GWAS and risk for major depression, *Transl. Psychiatry* (2017), doi:10.1038/tp.2017.36.
 39. A. H. Miller, C. L. Raison, The role of inflammation in depression: From evolutionary imperative to modern treatment target *Nat. Rev. Immunol.* (2016), doi:10.1038/nri.2015.5.
 40. B. E. Leonard, The Concept of Depression as a Dysfunction of the Immune System, *Curr. Immunol. Rev.* (2010), doi:10.2174/157339510791823835.
 41. C. H. Liu, G. Z. Zhang, B. Li, M. Li, M. Woelfer, M. Walter, L. Wang, Role of inflammation in depression relapse *J. Neuroinflammation* (2019), doi:10.1186/s12974-019-1475-7.
 42. E. Sibille, Y. Wang, J. Joeyen-Waldorf, C. Gaiteri, A. Surget, S. Oh, C. Belzung, G. C. Tseng, D. A. Lewis, A Molecular Signature of Depression in the Amygdala, *Am. J. Psychiatry* **166**, 1011–1024 (2009).

43. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–50 (2005).
44. M. Baron, A. Veres, S. L. Wolock, A. L. Faust, R. Gaujoux, A. Vetere, J. H. Ryu, B. K. Wagner, S. S. Shen-Orr, A. M. Klein, D. A. Melton, I. Yanai, A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure, *Cell Syst.* (2016), doi:10.1016/j.cels.2016.08.011.
45. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive Integration of Single-Cell Data, *Cell* (2019), doi:10.1016/j.cell.2019.05.031.
46. A. M. Newman, C. L. Liu, M. R. Green, A. J. Gentles, W. Feng, Y. Xu, C. D. Hoang, M. Diehn, A. A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, *Nat. Methods* **12**, 453–457 (2015).
47. S. S. Shen-Orr, R. Tibshirani, P. Khatri, D. L. Bodian, F. Staedtler, N. M. Perry, T. Hastie, M. M. Sarwal, M. M. Davis, A. J. Butte, Cell type-specific gene expression differences in complex tissues, *Nat. Methods* **7**, 287–289 (2010).
48. A. Subramanian, R. Narayan, S. M. Corsello, D. D. Peck, T. E. Natoli, X. Lu, J. Gould, J. F. Davis, A. A. Tubelli, J. K. Asiedu, D. L. Lahr, J. E. Hirschman, Z. Liu, M. Donahue, B. Julian, M. Khan, D. Wadden, I. C. Smith, D. Lam, A. Liberzon, C. Toder, M. Bagul, M. Orzechowski, O. M. Enache, F. Piccioni, S. A. Johnson, N. J. Lyons, A. H. Berger, A. F. Shamji, A. N. Brooks, A. Vrcic, C. Flynn, J. Rosains, D. Y. Takeda, R. Hu, D. Davison, J. Lamb, K. Ardlie, L. Hogstrom, P. Greenside, N. S. Gray, P. A. Clemons, S. Silver, X. X.

Wu, W. N. Zhao, W. Read-Button, X. X. Wu, S. J. Haggarty, L. V. Ronco, J. S. Boehm, S. L. Schreiber, J. G. Doench, J. A. Bittker, D. E. Root, B. Wong, T. R. Golub, A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles, *Cell* (2017), doi:10.1016/j.cell.2017.10.049.

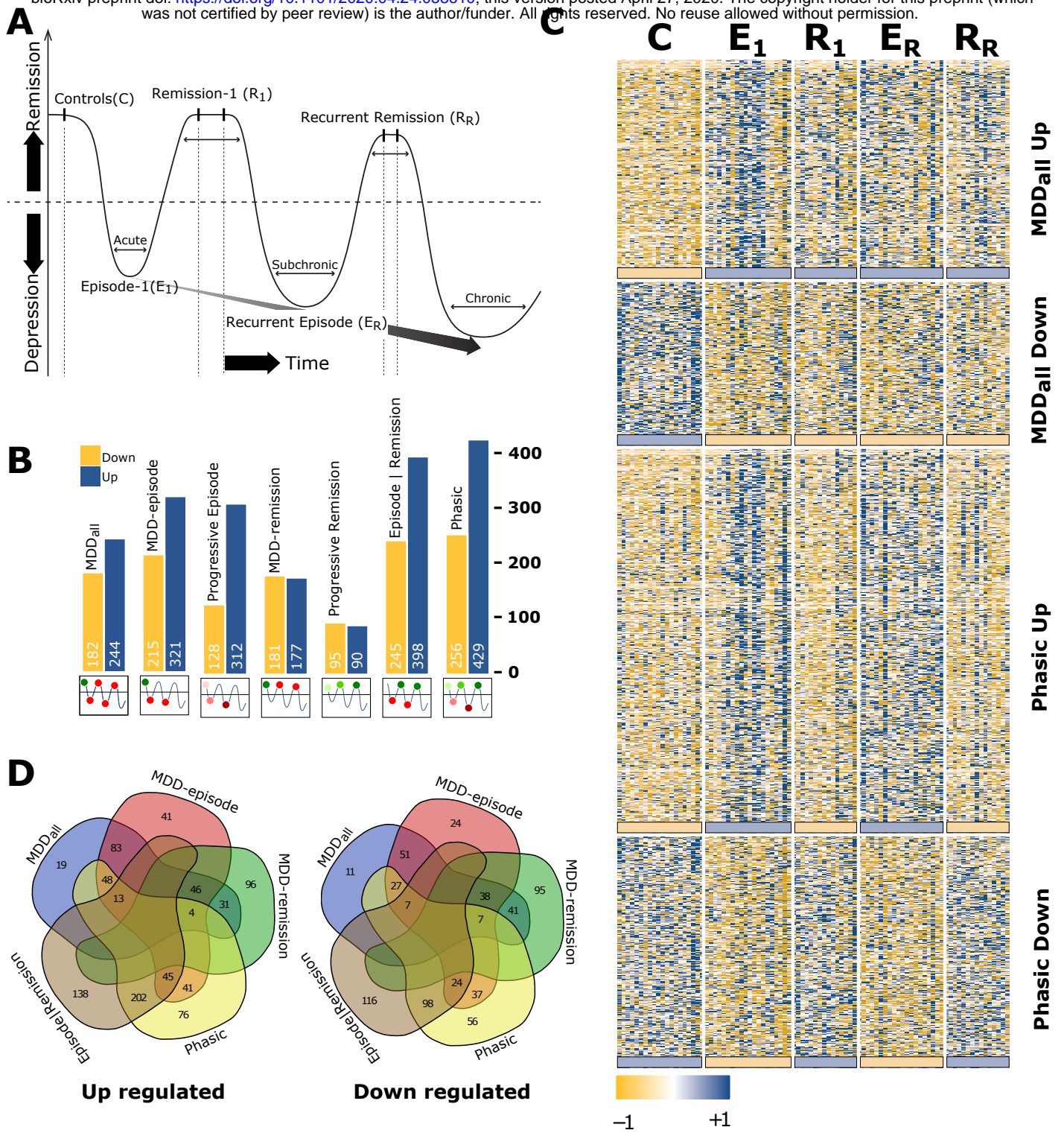


Figure 1

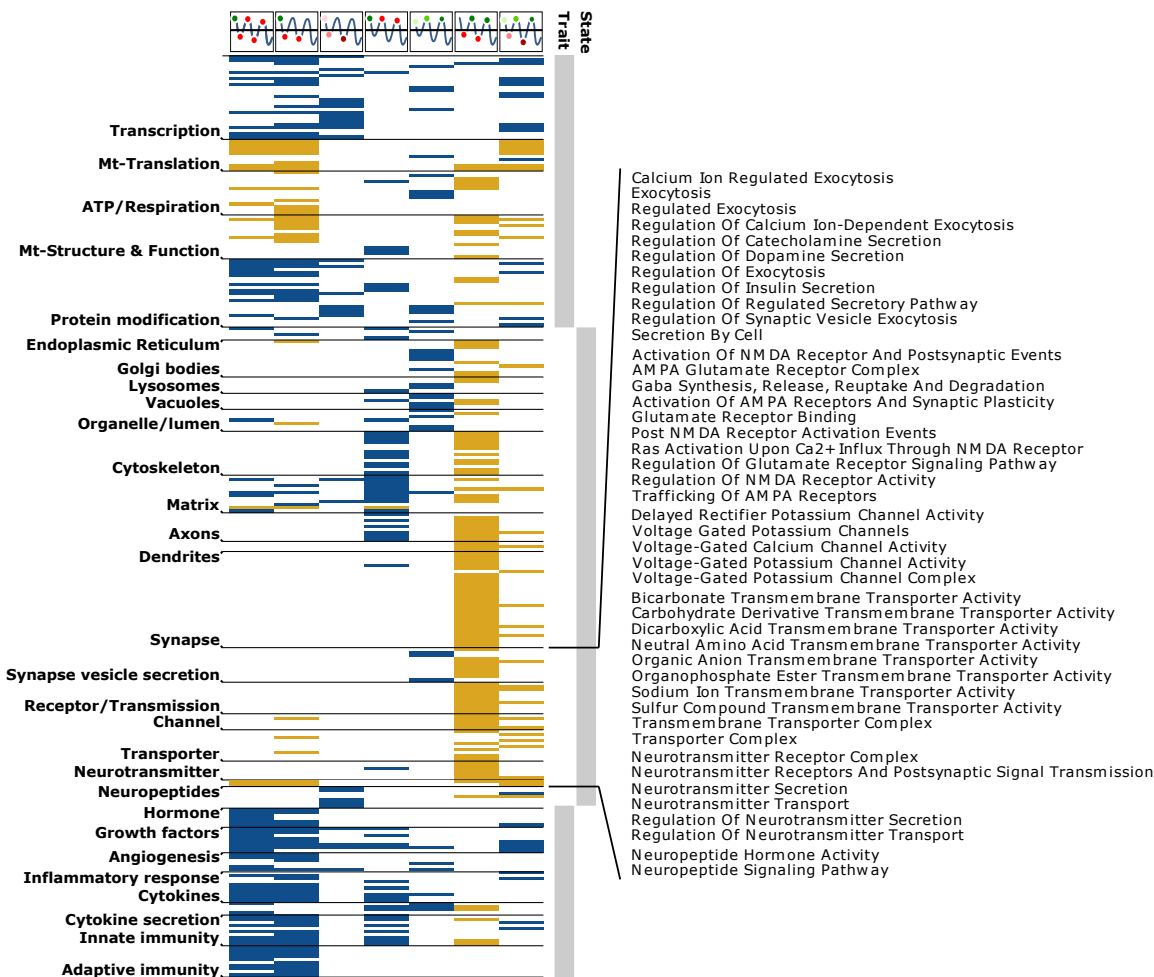


Figure 2

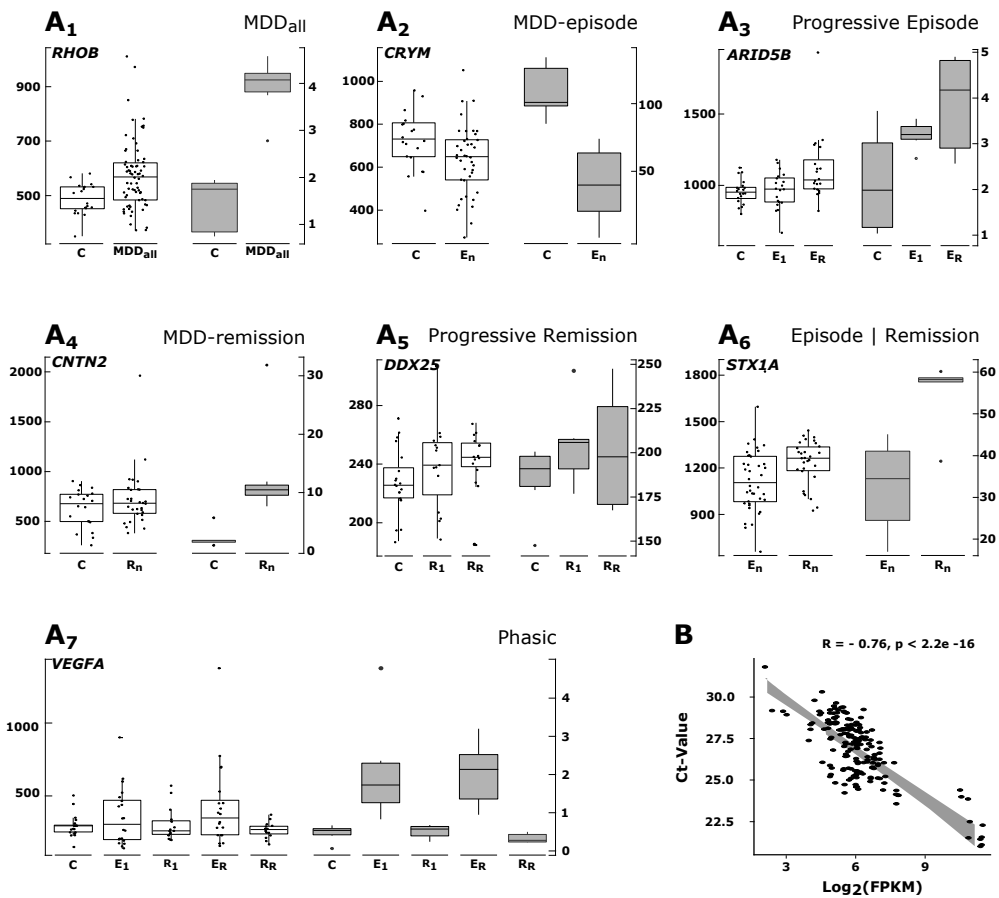


Figure 3

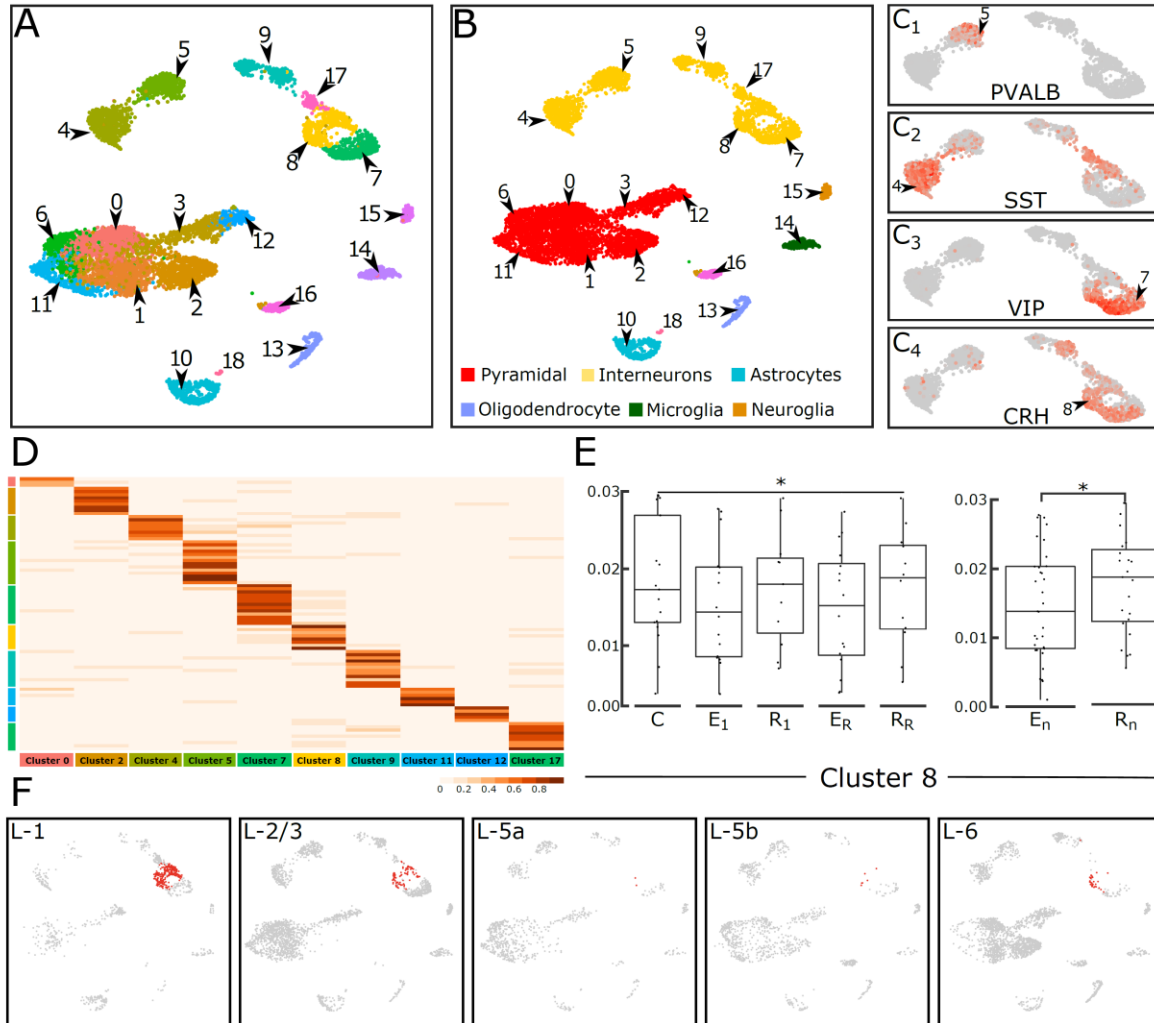


Figure 4

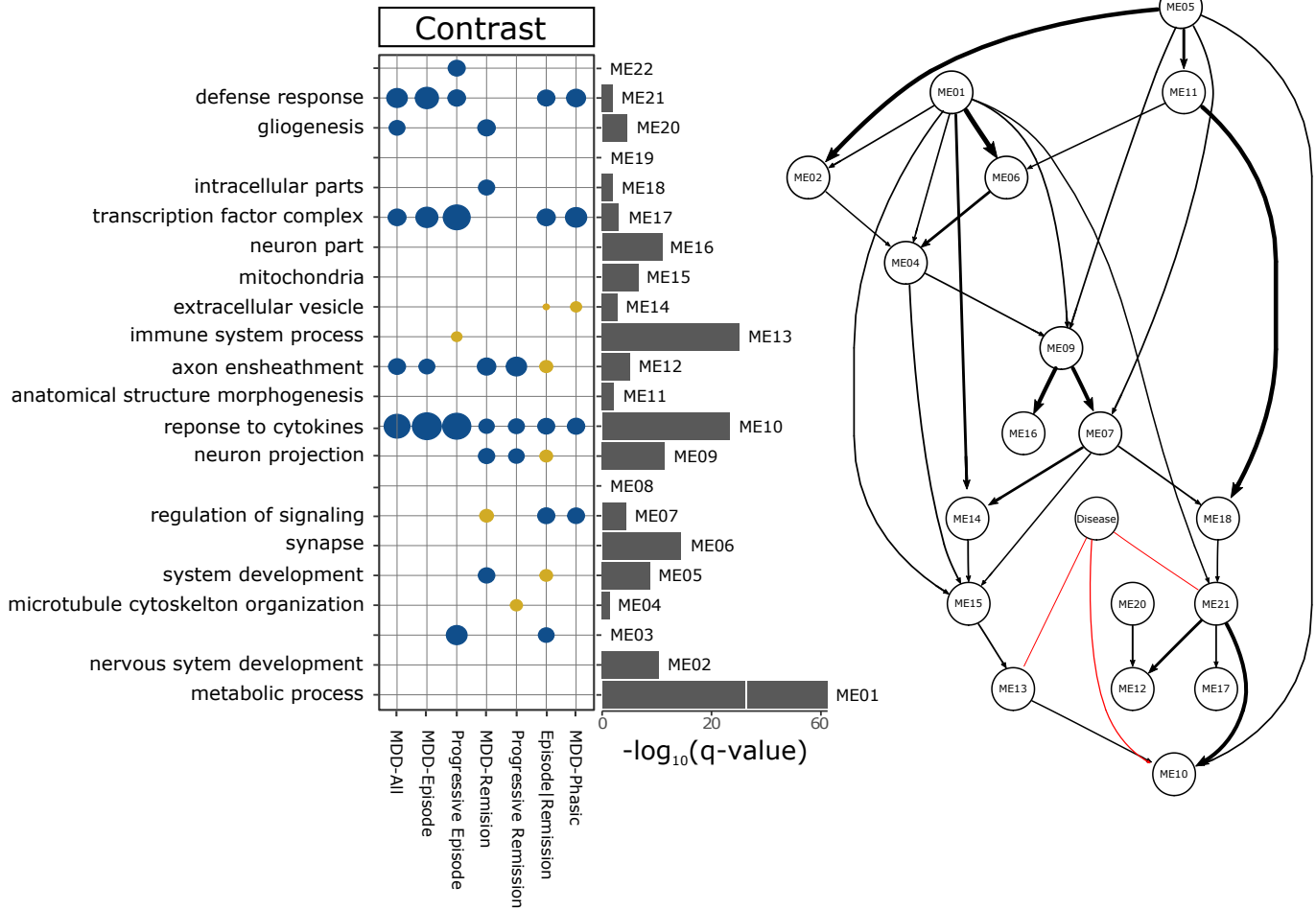


Figure 5

Molecules	NEU	KCI	NPC	Target	Target Class	Mode of action	Notes
Chlorpromazine				DRD2/DRD4/HTR2A	Family A GPCR	Antagonist	Antipsychotic
YM-09151-2				DRD2/DRD4/HTR2A/ADRA1A/DRD3	Family A GPCR	Antagonist	Atypical antipsychotic; Same as Nemenopride
Cycloheximide				FKBP1A	Isomerase	Blocks translation	—
Clonidine				ADRA2A/ADRA2C/ADRA2B	Family A GPCR	Agonist (presynaptic)	Used for blood pressure; ADHD
Valproxan				CA2/CA1 (GABA/Glu/Nach)	Lyase	Affects GABA; Blocking voltage-gated Na-Ch; inhibits histone deacetylases	Valproate; Used in epilepsy and BPD
Nemonapride				DRD2/DRD3/DRD4/HTR2A/ADRA1A	Family A GPCR	Antagonist	Atypical antipsychotic
EHPG-piperazine				ADRA2A/ADRA2B/ADRA2C	Family A GPCR	Affects adrenergic signalling	—
Zonisamide				CA2 (Ca ch./GABA/Glu)	Lyase	Sulfonamide; Antagonist	Used in epilepsy, bipolar, Parkinson's
BRD-A09053961				FYN/EGFR/LCK	Kinase	—	—
Allantoxanamide				ADORA1/ADORA2A	Family A GPCR	Uricase inhibitor	Increase uric acid
BRD-K99554241				MTNR1A/MTNR1B	Family A GPCR	Melatonin Rec antagonist	—
BRD-K97591839				SLC6A3 (DAT)	Electrochemical transporter	DAT inhibitor	—
Enevalproate				FNTA/FNTB/SRD5A2	Enzyme/Enzyme/Oxidoreductase	—	Valproate metabolite; Used in epilepsy and BPD
BRD-K72015216				SLC6A3 (DAT)/SLC6A4 (SERT)	Electrochemical transporter	Blocks DAT and SERT	Increase 5HT/DA (5HT2AR)
BRD-K27237442				DRD1	Family A GPCR	5-HT2A/2C antagonists, D1 antagonists and 5-HT1A agonists	D1 PET tracer
Lithocholic-acid				VDR/NR1H4/GPBAR1/AKR1B10	Nuclear receptor/Family A GPCR/Enzyme	Fat solubilization for absorption	Bile acid; Acts as detergent
Troglitazone				PPARG/HTR2B/TBXAS1/MAOB	Nuclear receptor/Family A GPCR/Cytochrome P450/Oxidoreductase	PPARG agonist	Used in Type 2 diabetes
Tetrabenazine				SLC18A2 (VMAT2)/ADRA1A/ADRA2A	Electrochemical transporter/Family A GPCR	Monoamine vesicle storage inhibitor	—
Clofibrac-acid				PPARA	Nuclear receptor	PPAR alpha agonist	—
FK-866				NAMPT	Enzyme	Inhibitor of nicotinamide phosphoribosyltransferase	—
Dopamine				KDM4E/DRD5/MTNR1A/SLC6A2	Eraser/Family A GPCR/Electrochemical transporter	Multiple actions	DA/NE agonist
BRD-K84252391				RCOR1/KDM1A/SIGMAR1/DRD2	Eraser/Membrane receptor/Family A GPCR	Multiple actions	Binds Sigma, DA and 5HT receptors
Ellagic-acid				ERBB2/AKR1B1/CCND1	Kinase/Enzyme/Kinase	Free-radical inhibiting antioxidant	—
Pyridine-2,4-dicarboxylic-acid				KDM4E/EGLN2/FTO/EGLN1	Eraser/Enzyme/Oxidoreductase	Hypoxia-inducible transcription factors (HIFs) inhibitor	—
Glutamyl-dopamine				SELL	Adhesion	Dopamine precursor	—
BRD-K23628492				BCHE	Hydrolase	Non-specific choline esterase	Similar to AChE
Beta-alanine				KDM4E	Eraser	Epigenetic regulators, increase GABA levels	—
Mevalonic-acid				PGD/MVD	Enzyme	—	Related for statin effects
Spermidine				CA2	Lyase	Regulate membrane potential, NOS inhibition	Identified in postmortem studies
BRD-K98336812				TRPV1	Voltage-gated ion channel	—	Chlorpromazine metabolite
BRD-A27924917				GABBR2/GABBR1	Family C GPCR	GABA-B Receptor Antagonists	—
Rosiglitazone				TBXAS1/MAOB/CA2/AGTR1	Cytochrome P450/Oxidoreductase/Lyase/Family A GPCR	Antihyperglycemic and anti-inflammatory activities	Antidiabetic agents
6-nitrodopamine				SNCA	Unclassified	Nitrosylated DA	—

Figure 6

Molecules	NEU	NEU,KCI	NPC	Target	Target class	Mode of Action	Notes
Tacedinaline				HDAC3/ HDAC2/ HDAC1	Eraser	HDAC inhibitor	Epigenetic regulator
BG-1010				HTR2B/ HTR2A	Family A GPCR	-	-
BML-210				HDAC3/ HDAC2/ HDAC1	Eraser	HDAC inhibitor	Epigenetic regulator
BRD-A70541707				SLC6A3	Electrochemical transporter	-	-
BG-1024				ABL1/ HDAC/ Mapk14	Kinase/ Nuclear receptor/ kinase	-	-
Oxyphenbutazone				PTGS1	Oxidoreductase	Blocks prostaglandin synthesis	-
BRD-K99636700				-	-	-	-
NSC-3852				METAP2	Protease	HDAC inhibitor	Epigenetic regulator
BRD-K31344914				OPRM1	Family A GPCR	-	-
BG-1032				ATR	Kinase	-	-
Geldanamycin				TRAP1/ HSP90AA1/ HSP90AB1	Other cytosolic protein	Inhibits cytosolic chaperone (HSP90)	-
NO-ASA				PDE3A	Phosphodiesterase	Platelet aggregation Inhibitors	-
CCT036477				VDR	Nuclear receptor	Targets Wnt pathway	-
BRD-K39096267				MMP2	Protease	-	-
BG-1001				KCNMA1	Voltage-gated ion channel	-	-
BRD-K75971499				KCNH2	Voltage-gated ion channel	-	-
HDAC3-selective				-	-	-	-
Amiodarone				CHRM4/ ERBB2/ HTR2B/ FYN/ THRA	GPCR/ Nuclear receptor/ kinase/ electrochemical transporter	HERG blocker	-
Wortmannin				PRKDC/ PIK3CA/ PIK3R1/ MYLK	Kinase/ Enzyme	Immunosuppressants act by inhibiting DNA synthesis	-
Fenretinide				RBP4	Secreted protein	Activates retinoic acid receptors inducing apoptosis	-
Trichostatin-a				SF3B3/ HDAC3/ NCOR2	Eraser	HDAC inhibitor	Epigenetic regulator
KCI				-	-	Maintains intracellular tonicity	-
LY-294002				PDE5A/PIM1	Phosphodiesterase/Kinase	Prevent the normal catalytic reaction	-
MG-132				CAPN2/CTSK/PSMB2/NFKBIA	Protease./Other cytosolic protein	Inhibits cysteine endopeptidases	-
BRD-A11009626				DRD2/DRD1	Family A GPCR	-	-
BRD-K88575585				MTNR1A	Family A GPCR	-	-
BL003-049-HDAC3				HDAC1	Eraser	Appetite suppressants	-
BRD-K59915259				HTR7	Family A GPCR	HDAC inhibitor	Epigenetic regulator
Etoposide				MMP2	Protease	Inhibits DNA synthesis	-
Plumbagin				EP300	Writer	Augments immune response	-
Vorinostat				HDAC3	Eraser	HDAC inhibitor	Epigenetic regulator
BG-1009				HDAC1	Eraser	HDAC inhibitor	Epigenetic regulator
p-Azido-PE-TFMPP				HTR1A	Family A GPCR	HDAC inhibitor	Epigenetic regulator
Pyrvinium				AR	Nuclear receptor	Anthelmintic agent which acts to kill pinworms	5HT1A receptor labeling
SKF-83959				DRD1/DRD2	Family A GPCR	-	-
lonomycin				CES2	Enzyme	Acts as an agonist at the D1-D2 dopamine receptor	-
Niguldipine				ADRA2C	Family A GPCR	Increases calcium ion permeability of cell membrane	-
BG-1016				HDAC1/HDAC2/HDAC3	Eraser	Inhibits calcium influx through cellular membranes	-
BRD-A78236793				DRD2	Family A GPCR	HDAC inhibitor	Epigenetic regulator
Siperone				HTR2B	Family A GPCR	-	-
Mevastatin				HMGCR	Oxidoreductase	Blocks dopamine action	-
JS-K				FNTA/FNTB	Enzyme	Inhibit the growth of bacteria	-
Ac-Leu-Leu-Nle-CHO				CAPN1/CAPNS1/CTSB	Protease	Nitric oxide donor, induces autophagy	-
BRD-K61894884				-	-	Inhibits cysteine endopeptidases	-
BRD-K47535255				CNR1/CNR2	Family A GPCR	-	-
Terreic-acid(-)				CYP19A1	Cytochrome P450	Cannabinoid receptor	-
BRD-K32307229				HDAC2/HDAC6/HDAC8	Eraser	Inactivates MurA by covalently attaching to Cys115	-
BRD-K06209536				IDO1	Enzyme	HDAC inhibitor	Epigenetic regulator
Rotenone				MT-ND4/HTR6/CYP2C19	Oxidoreductase/Family A GPCR/Cytochrome P450	Aminophenol urease inhibitors	-
Suramin				HDAC3	Eraser	Mitochondrial NADH:ubiquinone reductase inhibitor	-
						Blocks binding of growth factors	-

Figure 7