FINDING THE PATHOLOGY OF MAJOR DEPRESSION THROUGH EFFECTS ON GENE INTERACTION NETWORKS

by

Christopher Gaiteri

B.S. in Neuroscience, Washington and Lee University, 2005

Submitted to the Graduate Faculty of

The School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

School of Arts and Sciences

This dissertation was presented

by

Christopher Gaiteri

It was defended on

February 22, 2011

and approved by

Dr. Marc Sommer, Adjunct Associate Professor, Departmental of Neuroscience

Dr. Linda Rinaman, Professor, Department of Neuroscience

Dr. Jonathan Rubin, Professor, Department of Mathematics

Dr. George Tseng, Associate Professor, Department of Biostatistics

Dr. Michael Oldham, UCSF Fellow, Department of Neurology

Dissertation Advisor: Dr. Etienne Sibille, Associate Professor, Department of Psychiatry

FINDING THE PATHOLOGY OF MAJOR DEPRESSION THROUGH EFFECTS ON GENE INTERACTION NETWORKS

Christopher Gaiteri, PhD

University of Pittsburgh, 2011

Abstract:

The disease signature of major depressive disorder is distributed across multiple physical scales and investigative specialties, including genes, cells and brain regions. No single mechanism or pathway currently implicated in depression can reproduce its diverse clinical presentation, which compounds the difficulty in finding consistently disrupted molecular functions. We confront these key roadblocks to depression research - multi-scale and multi-factor pathology - by conducting parallel investigations at the levels of genes, neurons and brain regions, using transcriptome networks to identify collective patterns of dysfunction. Our findings highlight how the collusion of multi-system deficits can form a broad-based, yet variable pathology behind the depressed phenotype. For instance, in a variant of the classic lethality-centrality relationship, we show that in neuropsychiatric disorders including major depression, differentially expressed genes are pushed out to the periphery of gene networks. At the level of cellular function, we develop a molecular signature of depression based on cross-species analysis of human and mouse microarrays from depression-affected areas, and show that these genes form a tight module related to oligodendrocyte function and neuronal growth/structure. At the level of brain-region communication, we find a set of genes and hormones associated with the loss of feedback between the amygdala and anterior cingulate cortex, based on a novel assay of interregional expression synchronization termed “gene coordination”. These results indicate that in the absence of a single pathology, depression may be created by dysynergistic effects among genes, cell-types and brain regions, in what we term the “floodgate” model of

depression. Beyond our specific biological findings, these studies indicate that gene interaction networks are a coherent framework in which to understand the faint expression changes found in depression and complex neuropsychiatric disorders.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

I wish to thank the members of my thesis committee for their individual strengths and for how they have banded together to provide me with versatile guidance on an interdisciplinary thesis topic. The following trends in their counsel and questions stand out to me: The chair of my thesis committee, Dr. Marc Sommer had the challenging task of uniting commentary from committee members, which ran the gamut from mathematical details to mouse behavior. His varied expertise and calm guidance have been an excellent bridge between perspectives. When my daily involvement with graph theory algorithms threatened to become abstracted, Dr. Linda Rinaman has always brought me back to the core biological questions surrounding depression and ensured a respect for experimental complexity and uncertainty. During the years in which I've constructed models of the pre-Botzinger complex with Dr. Jonathan Rubin, I've developed great respect for his ability to see into dynamical systems and to formulate pristine arguments, even when dealing with complex systems. Those same traits are appreciated in the context of guiding my research in gene networks. My research has benefited from Dr. George Tseng's parallel investigation of differential expression analysis in many of the same depression microarray datasets. His ability to concisely state statistical tests of complex ideas and his sympathetic ear for the daily challenges of working with genome-wide datasets have invigorated me at trying moments in my analysis. Dr. Michael Oldham has been a source of inspiration for a number of years through his carefully framed interdisciplinary studies that do not follow the status quo, but push coexpression network analysis forward technically and conceptually.

Finally, I would like to thank Dr. Etienne Sibille for taking the risk of hiring a network specialist in a laboratory, and indeed an entire field of research, that is not historically network-based. He has always listened thoughtfully to my latest outside-the-box ideas and been a cheerful and encouraging advisor. I most admire how he can simultaneously live in the pragmatic world of procuring grants without losing sight of the overarching goals that must be pursued for patient care.

I would like to dedicate the work's best aspects to two groups of people who are landmarks in my life:

To my parents, who together have given me love, education, and a homozygous recessive mutation in the MC1R gene.

And, to my friends, whom I trust - you have laughed with me, inspired me, rescued me, forgiven me and (most frequently) stayed up all night with me: Mark Allen, Peter Djalaliev, David Gaiteri, Christopher Honey, Huan Jin, Sashi Marella and Weifei Zheng.

Abbreviations

5HT 5-hydroxytryptophan (serotonin precursor)
ACC/ACG : anterior cingulate cortex/gyrus
AD: antidepressant (drug)
AMY: amygdala
BDNF: brain-derived neurotrophic factor
BZ: benzodiazepine
CREB: cAMP response element binding protein
DLPFC: dorsolateral prefrontal cortex
DRN: dorsal raphe nucleus
DSM-IV: Diagnostic and Statistics Manual version 4
GABA(R): gamma-aminobutyric acid (receptor)
GR: glucocorticoid receptor
HAM-D: Hamilton rating of depression (scale): questionnaire of depression symptoms
HAT: histone acetyltransferase
HDACi: histone deacetylase inhibitor
HPA (axis): hypothalamus pituitary adrenal glands
IDO: indoleamine 2,3-dioxygenase
MAOi: monoamine oxidase inhibitor
MDD: major depressive disorder aka depression
miRNA: microRNA
MRS: proton magnetic resonance spectroscopy
NAc: nucleus accumbens
NE: norepinephrine
NMDA(R): n-methyl d-aspartate (receptor)
PET: positron emission tomography
PFC: prefrontal cortex
PPI: protein-protein interaction (network)
PKA: protein kinase A
ROC: receiver operating characteristic
shRNA: short hairpin RNA
SNRI: serotonin norepinephrine reuptake inhibitor
SSRI: selective serotonin reuptake inhibitor
STAR*D: Sequenced Treatment Alternatives to Relieve Depression
TCA: tri-cyclic amine (antidepressant)
TF: transcription factor
TNF-α: Tumor necrosis factor alpha
UCMS: unpredictable chronic mild stress (model of depression)
VTA: ventral tegmental area

1.0 INTRODUCTION

Fueled by high rates of major depressive disorder world-wide (Holden, 2000; Kessler *et al*, 2003), treatment response rates barely above placebo levels (Kirsch *et al*, 2008; Turner *et al*, 2008; Warden *et al*, 2007), and by the inability of any single molecule or pathway to account for the repertoire of depression symptoms (Lewis *et al*, 2010; Mehta *et al*, 2010), there is increasing pressure for depression researchers to find a coherent basis for depression pathology that unites the many hypothesized mechanisms (Covington *et al*, 2010; Holsboer, 2008; Krishnan and Nestler, 2010). Using gene coexpression networks to infer molecular relationships based on repeated microarrays, it is possible to create a literal framework for thousands of molecular interactions that recapitulates cellular dynamics in healthy and disease states. These networks have already been useful in understanding the transcriptome organization of several diseases (Horvath *et al*, 2006; Ray and Zhang, 2010; Torkamani *et al*, 2010; Wang *et al*, 2009). Here, we apply related network techniques to the highly debated and only faintly detectable transcriptome signature of major depressive disorder. We conduct these network-based investigations of major depression at the levels of genes, cell-types, and brain regions, because all of these physical scales show evidence of pathology. Our results implicate simultaneous multi-system deficits in depression, which may exert a cooperative effect in creating observed symptoms and pathology. Through these biological investigations we also demonstrate the power of network-based techniques for understanding the diffuse impact of neuropsychiatric disorders.

Because we take a relatively novel approach in the search for the biological basis of major depression, we first review evidence for and against several mainstream theories of depression

pathogenesis (Chapter 1.2), stressing combinatorial pathology. Then we review developments in coexpression network analysis which are relevant to understanding transcriptome-wide disease impact (Chapter 1.3). While coexpression network links are taken as a generic marker of related cellular function, because these links are fundamental to our analysis, we also include a review of specific biological processes that are represented by coexpression networks (Chapter 1.4).

Since coexpression networks (based on synchronized gene expression fluctuations) encapsulate thousands of biological processes (Ihmels *et al*, 2004; Nayak *et al*, 2009; Obayashi and Kinoshita, 2009; Pavlidis *et al*, 2004; Tsaparas *et al*, 2006), in Chapters 2-4 we use coexpression networks extracted from depression-affected regions of postmortem brains to infer strategically important biological processes and trends associated with depression. We then discuss the ramification of our specific findings for depression pathology and their concerted meaning for depression research at large (Chapter 5), advocating a new “floodgate” model of depression that is based on simultaneous dysregulation in several linked systems.

1.1 EPIDEMIOLOGY OF MAJOR DEPRESSIVE DISORDER

Major depressive disorder (MDD or depression) is distinct from transient mood swings and immediate reactions to stress. It involves a minimum two-week continuous period of at least five of the following symptoms: lowered mood for the majority of the day, diminished pleasure in daily activities, weight loss or gain, sleep disturbance, agitation or lethargy, fatigue, feelings of worthlessness or helplessness, impaired thought or memory, and recurring thoughts of self-harm or death (DSM-IV 2000). Depression is a common human psychiatric disorder and the leading cause of disability in North America, afflicting an estimated 18% of the population with an approximate lifetime incidence of 12% in men and 20% in women (Kessler *et al*, 2003). This distribution is subject to significant international variation – with the lowest rates generally in Asia and the highest

in Europe (Weissman *et al*, 1993), but in all instances, rates of depression are higher in females. The symptoms of depression are the greatest contributor to the "global burden of disease," (Holden, 2000) as calculated by total days lived with the disorder. It remains the fourth leading cause of worldwide disability, after accounting for higher mortality in other diseases. This ranking is expected to rise to second place by the year 2020, as current effective treatment for other diseases become more globally accessible.

Risk factors for depression Major risk factors for depression include the sex of an individual, previous history of the illness, genetic predisposition/family history, and chronic or acute stress (Fava and Kendler, 2000). Some combination of these can prompt a depressive episode, but the requisite combination varies by individual. The threshold for depression is sensitive to social support, religiosity, age, and life stressors (Caspi *et al*, 2003; Kendler *et al*, 2003a; Kendler *et al*, 2003b). These environmental factors interact with the genetics of depression – estimated at 33% heritance (Fava *et al*, 2000). This is a lower heritability than bipolar disorder, or schizophrenia, which adds to the difficulty in teasing apart contributory factors. Depression itself is a risk factor for the disorder, as untreated depression is likely to reoccur (Mueller *et al*, 1999). This is particularly problematic as a significant percentage of patients (varying from placebo levels of 30%, up to 40% depending on the study) never meet the criteria for complete remission and will commonly endure increasingly lengthy bouts of depression (Gaynes *et al*, 2008; Kirsch *et al*, 2008).

Personal and economic costs of depression Mortality estimates for depression vary from 5-15 percent, but 60 percent of suicide cases are associated with some combination of mood disorders (Mann 2003). In addition, depression is significantly comorbid with a variety of neurological and non-neurological diseases (Schultz *et al*, 2003). The cost of depression-related disability and death is over $40 billion annually in the U.S alone (Berto *et al*, 2000; Hu, 2006). While numerous trials and meta-analyses indicate the relatively low rate of response to antidepressant (AD) medication (Kirsch *et al*, 2008; Trivedi *et al*, 2007), the economic cost of the disorder is so great that every major therapeutic intervention other than Freudian psychoanalyses

is justified, even if the therapies themselves are rarely successful (Donohue and Pincus, 2007).

Limited pharmacological mechanisms in depression treatment: For the past fifty years, the vast majority of depression research has revolved around the various monoamine hypotheses of depression (see Chapter 1.2.1). However, recent meta-analysis shows that monoaminergic drug therapies are only slightly more effective than placebo and, in fact, fail FDA measures for clinical efficacy (Kirsch *et al*, 2008) when the severity of side-effects is included in the efficacy calculation. Numerous alternative hypotheses have been suggested for depression, but these are hampered by conflicting evidence and limited scope (see Chapter 1.2). Microarray studies have the potential to circumvent historical limitations on depression targets (Drigues *et al*, 2003; Takahashi *et al*, 2006). However, microarray results do not necessarily translate to depressed humans because they often lack naturalistic depression induction protocols or administer drugs to non-depressed populations. Furthermore, microarray analysis of human post-mortem data do not paint a consistent picture of the nature of the disease (Bezchlibnyk *et al*, 2001; Kang *et al*, 2007; Mehta *et al*, 2010; Sibille *et al*, 2004) potentially due to multifactorial multi-system nature of depression (Chapter 1.2).

The continued prominence of depression, the lack of effective pharmacological therapy (Gaynes *et al*, 2008) and the diversity of opinion on fundamental depression pathology present a substantive challenge to science. Depression symptoms have evaded a range of therapies through variable clinical presentation and diverse neural mechanisms (Nestler *et al*, 2002) in which the individual contribution of single elements is relatively small (Bosker *et al*; Uher *et al*). In a vicious cycle, the disease complexity makes it difficult to model, and thus models for disease are based on "serendipitous discovery of antidepressant treatment" (Vaidya and Duman, 2001). Because these models are mechanistically limited, the field of depression research is often left optimizing inadequate therapies to artificially framed questions (Nutt, 2006). Microarrays and graph theory of large biological networks may offer a way to place putative mechanisms of depression within a larger unbiased transcriptome-based perspective on the neuropathology of depression, based on

their success in bringing coherence to similarly complex disease networks (Guye *et al*, 2010; Micheloyannis *et al*, 2006; Smit *et al*, 2008; Srinivas *et al*, 2007; Stam *et al*, 2007; van Nas *et al*, 2009).

Complexity and heterogeneity obscure the neuropathology of depression Depression's continued toll on society is a function of multiple genetic and environmental susceptibilities that recruits a diverse cadre of further genetic factors to sustain the condition (Bauer *et al*, 2002). To date, most experiments have examined single aspects of the disease, but the complex causal factors in depression make it resistant to highly specific approaches. One immediate question is: Why not create sub-divisions of depression that have more homogeneous symptom groups that will be amenable to a pathology classification? However, clinical evidence does not strongly support this approach. In patients with repeated depressive episodes there is no correspondence of symptoms across episodes, preventing definitive clinical subdivisions that might have more consistent pathophysiology (Oquendo *et al*, 2004). There is some evidence to suggest that classes of antidepressants have distinct response rates in different DSM-IV classifications of depression (atypical, psychotic, bipolar etc) (Ayuso-Gutiérrezd, 2005). However, a meta-analysis of over 100 antidepressant drugs trials found no difference in response rates as an interaction of drug class and putative subtype (Cipriani *et al*, 2005).

Potential sub-populations in depression: There are clusters of depressed patients with putative biomarkers suggesting greater involvement in certain affected subsystems (Mössner *et al*, 2007). For instance, Shelton (2007) found distinctly lower PKA levels in a sub-set of depressed patients. However, these biological subdivisions do not map onto clinical subdivisions in a way that might allow a finer dissection of factors in depression. The STAR*D trial of 1500 MDD patients found that those with relapsing depression were more likely to have a family history of the disease (Hollon *et al*, 2006) and the depressive episodes of these patients are longer, more intense and incapacitating than non-familial depression cases (Fava *et al*, 2008). Thus, while there is some evidence for subtypes in depression, it is not coherent or significant enough to establish different

sub-groups for our analysis. Rather, depression continues to exhibit clinical and biochemical heterogeneity that contaminates analysis.

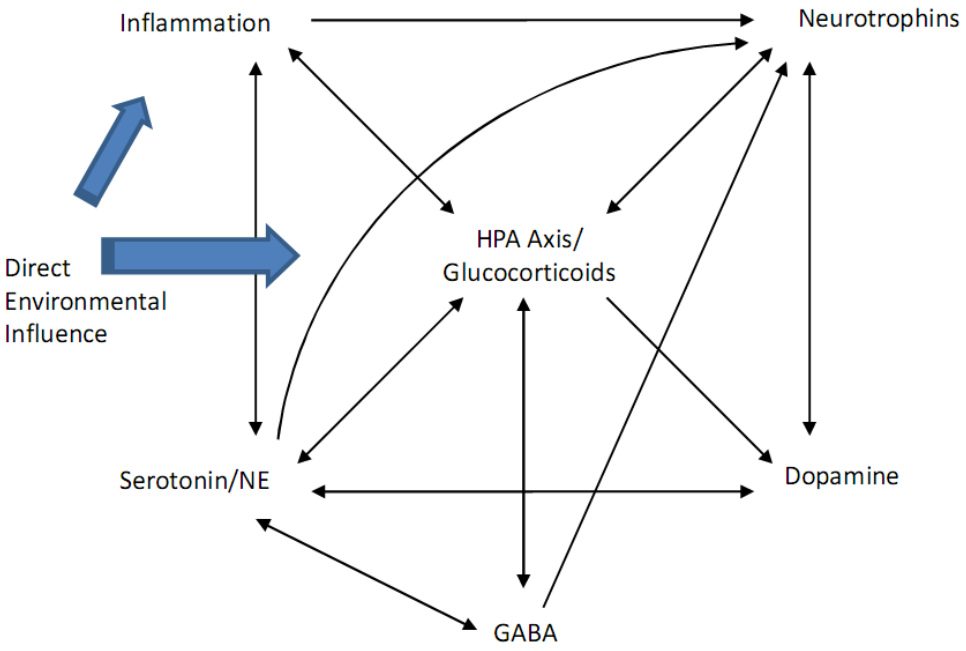
1.2 PUTATIVE DEPRESSION PATHOLOGY

A series of potential mechanisms of depression pathology (often identified as the "_____" hypothesis of depression) have been proposed to account for depression symptoms. A characteristic of depression research is that initial excitement is generated when some neural mechanism is found to be necessary for antidepressant efficacy, for instance the necessity of neurogenesis for SSRI efficacy (Santarelli *et al*, 2003; Surget *et al*, 2008a). However, several further tests are necessary to justify the validity of a hypothesized mechanism of depression, especially when evidence for that hypothesis comes from a drug intervention.

The following criteria are a useful checklist when evaluating the many competing hypotheses of depression. The fundamental criterion for a molecular mechanism of depression is that effects of the putative mechanism match some known aspect of depression pathology or symptoms. This may appear to be an obvious criterion; however, it is frequently difficult to confirm a hypothesized mechanism directly in depressed humans, so the evidence must be drawn from (biological) disease model systems. Reversal of depression symptoms by a particular drug or antagonist might be thought to be the gold-standard of relevance to pathology. However, from a genetic and morphological perspective, the actions of antidepressants are not the inverse of depression pathology (Surget *et al*, 2009). Similarly, inducing a depressive-like state by ablating a particular factor does not mean it is responsible for naturally occurring depression, due to developmental interactions and the typically global/extreme nature of these interventions. Finally, while various theories of depression compete for research funding, in reality they likely cooperate, so success of any one theory does not dismiss possible contributions of other mechanisms to

depression. Each current hypothesis of depression has variable performance with regard to these criteria for a successful theory of depression. Since different hypotheses interact in neural substrates, one potential approach to depression research would be to focus on an integrated theory of depression pathology (covered in Chapter 1.2.8, 5.4 based on interactions shown in Figure 1).

A



B

	HPA activation/ Glucocorticoids	Serotonin/NE Hypothesis	GABAergic Hypothesis	Dopamine Hypothesis	Neurotrophic Hypothesis	Inflammatory Hypothesis
HPA activation/ Glucocorticoids	**Blunted HPA feedback, higher amygdala reactivity** from decreased GR sensitivity or # in AMY, ACC, and PFC	**Dorsal raphe sensitivity changes** through CRF receptors	Stress decreases FC GABAergic cell #'s and PV+ cells in hippocampus	MCH , MSH and AgRP (neuropeptides) **activate mesolimbic dopa mine signaling**; chronic HPA activation **decreases PFC dopamine**	**Decreases neurogenesis**, through altered CREB levels leading to calcium and LTP changes, also decreased NG2 proliferation	Blunted **feedback permits continued inflammation**, despite normally mutually inhibitory regulations between these systems
Serotonin/NE Hypothesis	**May predispose to depression** through HPA reactivity via many possible/unclear mechanistic/devel links	**Depressed mood,** but only in conjunction with developmental effects, or secondary long-term depression effects, little evidence for direct effect	**Decrease in THP levels** by insufficiently activated & fewer $GABA_AR$'s; numerous 5HTR's directly modulate GABAergic neurons	Many serotonin receptors subtypes found throughout 3 dopamine pathways	Mainly coupled through **necessity for SSRI action**, not endogenous mechanisms	Can modify cytokine levels by cAMP, nfκB/AKT action, but demonstrated mainly through AD effects
GABAergic Hypothesis	**HPA hyperexcitability** through decreased inhibition to AMY and PVN, structural modifications and	Direct action of GABAerigic neurons on the dorsal raphe	γ2 het KO mouse indicates **depressed and anxious phenotype** due to developmental/ structural changes		**Decreased hippo neurogenesis by** reduced GABA-mediated activation of cell precursors - compounds BDNF deficit	
Dopamine Hypothesis	(indirect through GABA)	Both activation and inhibition of locus cere. and DRN, confirmed by dual reuptake inhibitors	**Disinhibition of amygdala** through action on multiple GABAergic subpopulations	**Sporadic depression** prior to Parkinson's onset, anhendonia from NAc deficits (hypothesized)		
Neurotrophic Hypothesis	Dendritic hypotrophy in limbic areas; altered hippo function may lead to **undamped amygdala activation**		BDNF postsynaptic release can modulate trkB receptors on presynaptic GABAergic cells	BDNF overexpression in NAc/VTA has prodepressive effect	Decreased hippo volumes and BNDF postmortem	
Inflammatory Hypothesis	**Sickness syndrome, excitotoxic glia death,** via reduced glia/glutamate scavenging and extrasynaptic NMDA R's activation	**Serotonin decrease quinolinic acid release** through IDO activation, leading to further inflammatory activation through glutamate			Various interleukins , TNF-α, and ROS **decrease hippocampal neurogenesis** through CREB mechanism	Inflammation accompanies depression and **increases BBB perm to immune cells**, but other depression effects require substrates

Figure 1 Overview of mechanistic interactions between various hypotheses of depression.

Because our findings emphasize the collective contribution of multiple deficits to the depressed phenotype, we include this schematic of direct interactions and the resulting depression effects. (A) Directed links between hypotheses are shown if there is a demonstrated endogenous pro-depressive interaction. (B) The chart should be read as: "The deficit in row *x* contributes to/interacts with the hypothesis in column *y* to create the following depression-related effect" (listed in bold). The main diagonal represents the contribution of a given hypothesis towards depression, in isolation from all other hypotheses. This is far from an exhaustive list of interactions, nor a replacement for Chapter 1.2 in describing the full effects and interactions of the hypotheses. We also attempt to limit these interactions to endogenous cases vs links driven by AD drugs – antidepressant drug actions do link many hypotheses, but would entail a separate different table of dozens of drugs and hundreds of molecular/downstream targets.

1.2.1 Monoamine hypotheses of depression

The vast majority of AD (antidepressant) medications operate on serotonin, norepinephrine or dopamine neurotransmitter systems. The initial evidence from the 1950's for neurotransmitter involvement in depression occurred in unrelated studies, when reserpine (which increases transport of several monoamines into the presynaptic cell) was found to increase depressive symptoms. The monoamine hypothesis became more entrenched in the 1960's as researchers continued to pursue a pharmacologically driven search for the roots of depression and found that the TCA drug (tri-cyclic amine) imipramine prevented NE (norepinephrine) reuptake (Iversen, 1965). In tracking down the exact mechanism of imipramine, researchers found it also inhibited serotonin reuptake, leading to the first SSRI's (selective serotonin reuptake inhibitors) (Ross and Renyi, 1969). SSRI's generally have fewer adverse side effects than the original non-selective TCA drugs (Freemantle and Mason, 2000) and nominally act to increase neurotransmission through increased serotonin levels in the synaptic cleft. In reality the exact mechanism of efficacy remains unclear because they operate through 15 receptor subtypes that mediate a combination of pre- and post-synaptic receptors and IP3 and cAMP mediated pathways, multiple ion channels, and serotonin synthesis, storage and transport (Belmaker and Agam, 2008). Related SNRI's prevent reuptake in both serotonin and norepinephrine and seem to have efficacy that is closer to TCA medication, alongside the improved tolerability profile of SSRIs, as shown in a meta-analysis of clinical trials (Papakostas *et al*, 2007).

Dopamine, the latest monoamine hypothesis: A more recent addition to the monoamine depression theories is that dopamine-mediated effects are a component of observed depression symptoms (Willner and Mitchell, 2002). The clinical basis for this hypothesis is observation of high levels of depression in pre-Parkinsonian patients and the pro-hedonic effect of some Parkinson's drugs (Gershon *et al*, 2007). It is also possible that symptoms of psychomotor retardation and lack of reward in social interactions could be related to decreased D2 receptor binding in the substantia

nigra and the ventral tegmental area (VTA) (Winter *et al*, 2007) and abnormal bursting patterns in VTA (Friedman *et al*, 2007). The basal ganglia's influence on the prefrontal cortex and the amygdala also make it an attractive substrate for altered function in depression (Floresco and Tse, 2007). Anhedonia's status as a core clinical feature of depression indicates that reward pathways may be affected in depression. Multiple assays of nucleus accumbens function indicate this dopaminergic reward-sensitive region may have depression-related function, as CREB levels, volumetric studies in depressed patients, animal social defeat models, and fMRI studies show accumbens activity is modulated by stress (for overview see Krishnan and Nestler (2010)). Collectively these findings implicate that nucleus accumbens and dopamine signaling may be involved in generating anhedonia observed depression.

Contradictions and circularity in the monoamine hypotheses: Despite its popularity, the monoamine hypothesis has relatively weak support from studies of actively depressed brains (as opposed to antidepressant mechanisms). One challenge to the monoamine hypotheses is that if depression simply consists of altered neurotransmitter levels, SSRI's should have effects within minutes (when synapses are saturated with serotonin) instead of the observed period of 2-4 weeks for clinical improvement (Duman *et al*, 1997), which implies synaptic connectivity changes are responsible for AD effects. Also, tryptophan depletion can induce depression in patients with a history of the disorder, but it does not do so in healthy subjects (Carpenter *et al*, 1998). A meta-analysis of depletion studies for serotonin, norepinephrine and dopamine indicates that serotonin depletion can create temporarily depressed mood in subjects with a family history of MDD or (more severely) in those with a personal history of MDD (Ruhé *et al*, 2007). Conversely, these studies indicate that monoaminergic depletion by itself is not a sufficient causal factor in depression as healthy controls with no history of MDD do not exhibit depressed mood in response to monoamine depletion. PET studies in major depression have attempted to answer this key issue of whether there are deficits in serotonin or serotonin receptors. The largest study to date (Parsey *et al*, 2006) with over 100 subjects found greater density of bound serotonin 1A receptors in depressed vs

control subjects, which is the opposite of what would be expected under the monoamine hypothesis. A series of other studies on serotonin 2A/B receptor binding have only moderate and conflicting results (Smith and Jakobsen, 2009). The lack of convincing evidence does not appear to stem from a lack of sensitivity in PET scans, as they readily detected differences in SSRI treated and untreated subjects (Voineskos *et al*, 2007).

Some of the theoretical appeal of the dopamine hypothesis is the system's ties to reward/anhedonia – relatively high-level perceptual and emotional states affected in depression. However, dorsal raphe neurons also respond in reward-related paradigms, self-stimulation of the dorsal raphe in primates alters reward valuation, and SSRi's in humans acutely alter reward responses (Kranz *et al*, 2010); thus non-dopaminergic systems also have potential to capture the anhedonia component of depression. Also, accumbens deficits are not specific to depression among neuropsychiatric disorders, and deep brain stimulation of accumbens is also useful for treatment of obsessive compulsive disorder (Denys *et al*, 2010). Nor is deep brain stimulation only effective on the accumbens, as stimulation of anterior cingulate and ventral striatum also have useful anti-depressant effects with no strict relationship of efficacy to accumbens proximity (Giacobbe *et al*, 2009). Attempts to see if dopamine is responsible for the effects of DBS (applied to DLPFC) using PET imaging found the dopamine binding did not increase post-DBS in any region. Furthermore, BDNF and CREB show opposite changes in NAc in response to social stress (in mice) as they do in the hippocampus. The justification for this is that CREB levels can mediate either adaptive or maladaptive neuronal plasticity. While it is entirely possible CREB leads to different symptom directions in this region than in the hippocampus, without a mechanistic explanation, this line of reasoning raises questions of falsifiability of NAc's relationship with depression. Thus, dopamine's relationship to depression is non-specific, and while it could contribute to depression symptoms, it is not implicated as a necessary component of the disorder.

1.2.2 Stress and the glucocorticoid hypothesis of depression

A classic definition of stress is “the non-specific response of the body to any demand for change” (Selye, 1936). In some cases the object of stress may be external, and in other cases the stress itself may be corporeal, such as an infection or tissue injury. Conversely “stress” as it is commonly used in the depression literature more properly refers to anything which contributes to an organism’s allostatic load, or allostatic load itself – the summed physiological changes stemming from a lifetime of stress (McEwen, 2003). The generality of the definition of stress is matched by the range of stressors that have relevance to depression – early life stress, chronic stress, acute stress, social stress and others.

Epidemiology of stress and depression Stress is a major contributing factor towards entering a depressive episode, and stress may exert a cumulative effect on brain structures beginning early in life (Hammen, 2005; Mazure *et al*, 2002; Weaver, 2007) as it interacts with hundreds of relevant polymorphisms and other factors to determine individual predisposition towards depression (Caspi *et al*, 2003; Muglia *et al*, 2008). The relative contribution to depression of childhood stressors vs recent stressors is difficult to calculate for humans, due to the survey methods generally employed (Ensel and Lin, 2000). By regressing fMRI BOLD response in depressed patients to negative words against either recent or distal life stressors, Hsu (2010) found the severity of recent, but not distal life events, correlated with activation of several brain regions previously implicated in major depression. If males and females experience different amounts of stress, this could account for the sex difference in depression rates. However, stress estimates by sex vary as a function of subject age, and have frequently conflicting results (Hammen, 2005). Stressors clearly have different effects over the course of an organism’s life, as they interact with developing nervous systems, but unfortunately there is no clear threshold calculation that can integrate biomarkers for depression propensity with a weighted compilation of all stressful life events to predict depression onset. Therefore, the well-validated consensus in

depression research is that some variable amount of stress precipitates depression, but there is no precise calculation that predicts individual risk for depression, and thus stress is functionally considered to be a generic pro-depressive influence by the depression field, with many extended physiological effects.

Glucocorticoid signaling is a major transducer of physiological/emotional stress. Brain structures respond to stress through a feed-forward activation of the paraventricular nucleus of the hypothalamus, anterior pituitary and adrenal glands (HPA axis) mediated by corticotropin releasing hormone (CRH), adrenal cortical releasing hormone (ACTH) and cortisol, which all feed back on the brain through CRH receptors and glucocorticoid and mineralcorticoid receptors (GR/MR). Since glucocorticoids activate transcription factors that lead to transcription of hundred of genes, their downstream effects are vast (for reviews see (Bamberger *et al*, 1996; de Kloet *et al*, 2005; Dedovic *et al*, 2009; McEwen, 2007; Pariante and Lightman, 2008; Rodrigues *et al*, 2009)), activating dozens of major intracellular signaling pathways, including effects on ionotropic and metabotropic receptors, growth factors, enzymes, neurotransmitters, and cell morphology. Therefore, we present a brief overview of some of the primary effects of HPA activation as they relate to depression. Interactions with other hypotheses of depression are covered in their respective sections.

While absolute levels of cortisol do not have a consistent relationship with depression (Knorr *et al*, 2010), a classic finding is that depressed patients have a blunted response to cortisol release as shown by the DMT (dexamethasone suppression test), although this result is not specific to depression. (The DMT tracks cortisol feedback by using the synthetic glucocorticoid dexamethasone to induce suppression of natural cortisol production.) The glucocorticoid hypothesis of depression postulates that blunted feedback on cortisol and CRH production allows persistently high levels of glucocorticoids or abnormal diurnal regulation of cortisol. The inability to suppress cortisol release may be due to downregulation of glucocorticoid receptors in response to repeated stress, or receptor insensitivity due to polymorphisms in the receptor itself or anywhere in the associated nuclear transport (van Rossum *et al*, 2006). Chronically administering

corticosterone in mice, making the assumption that functionally increased cortisol levels mimic the induction of depression in humans, produces a behaviorally depressive-like state (David *et al*, 2009). Alternately, CRH receptor antagonists should decrease apparent feedback of HPA activation, and subsequent adaptive effects. CRH antagonists do reverse chronic mild stress-induced depression in mice to an equal transcriptome/behavioral extent as SSRI's (Surget *et al*, 2009), and may be effective in humans, though they currently have toxic side effects (Holsboer and Ising, 2008). The DMT response appears to be heritable and characteristic of subjects from depressed families, even if they are not themselves depressed (Modell *et al*, 1998), therefore glucocorticoid signaling fits as a heritable basis for MDD. Within depression-affected systems, glucocorticoid responses have predictive power – if DMT results do not normalize, recovered depression patients are at a high risk of relapse (Zobel *et al*, 2001). Therefore, even disregarding the downstream effects of stress on other systems, altered HPA activation correlates with depression induction and status, but the system itself cannot be identified as the direct substrate of depression.

Interactions and uncertainty in a pure glucocorticoid model of depression Because the HPA axis and glucocorticoid signal are the point of input for environmental and psychological stressors, it is difficult to isolate the direct pro-depressive effects of glucocorticoids from their complex secondary effects. It appears that the direct effects of glucocorticoids do not completely account for depression symptoms. For instance, elevated cortisol levels should accompany blunted feedback, and while this is a historic belief in neuropsychiatry (Hinkelmann *et al*, 2009), recent meta-analysis indicates that the easy-to-administer salivary cortisol tests the belief is based on, do not show depressed patients have elevated cortisol (Knorr *et al*, 2010). Similarly, polymorphisms in GR's that affect binding sensitivity do increase depression rates, but also appear to lessen the severity of cognitive symptoms in depression (Spijker and van Rossum, 2009), and furthermore might only be elevated in depressed populations with childhood abuse (Bet *et al*, 2009). Thus it is debatable if the immediate HPA activation and glucocorticoid levels themselves are the direct actuators of the

depressed brain state. While the glucocorticoid signaling hypothesis is highly relevant to depression pathogenesis, it may be more appropriate to see cortisol levels as an intermediary to important secondary mechanisms of depression (such as dendritic changes) driven by the HPA axis.

1.2.3 Neurotrophic hypothesis of depression

Synaptic plasticity and neurogenesis may play a role in depression pathology and antidepressant mechanisms of action. Neurogenesis in the subgranular zone of the hippocampus is often necessary for behavioral reversal of induced depression in mice (David *et al*, 2009; Santarelli *et al*, 2003). When it is not possible to directly test for neurogenesis, BDNF (brain-derived neurotrophic factor), which acts through TrkB receptors to inhibit MAPK- and AKT-mediated cell-death, is commonly used as a marker for neurotrophic effects. Electroconvulsive shock therapy, the most effective therapy for treatment-resistant depression, also increases neurogenesis along with levels of BDNF (Hellsten *et al*, 2005). Several monoamine antidepressants increase levels of BDNF (Nibuya *et al*, 1995), increasing TrkB receptor activity alone has an antidepressant effect, and antidepressants have lower efficacy in the BDNF inducible KO mouse (Monteggia *et al*, 2004). Thus, there is strong evidence for the role of neuro/synaptogenesis in antidepressant mechanisms.

The transcriptional control of BDNF reinforces its putative role in depression. BDNF has a CRE binding site, and CREB levels been shown to have specific pro- or anti-depressant effects depending on the brain region (Nair and Vaidya, 2006). This process may also be regulated through histone binding sites near BDNF's several promoters, and indeed those histones have been shown to be methylated in response to social stress (Tsankova *et al*, 2006). Reducing CREB levels actually has an antidepressant effect (Newton *et al*, 2002), but the result is difficult to interpret due to multiple transcriptional and area-specific effects. However, the conditional BNDF

knockout does show depressive behaviors, the magnitude of which are significantly higher in females, making it an attractive model of depression (Monteggia *et al*, 2007).

Ambiguous causality for AD effects in the neurotrophic hypothesis Neurogenesis appears to be more of a way-point on the road to antidepressant effect rather than the core mechanism of antidepressants, since halting neurogenesis does not create depressed behaviors in mice and they can still show depressed behavior while neurogenesis is ongoing (Airan *et al*, 2007; Surget *et al*, 2008a). However, neurogenesis may be necessary for some antidepressant drugs, and indeed increasing neurogenesis directly through injecting stem cells into the hippocampus does have antidepressant effects (Tfilin *et al*, 2010). Similarly, there is limited evidence that lower BNDF itself induces depression, as there are inconsistent postmortem brain findings (Chen *et al*, 2001b; Thompson Ray *et al*, 2010), though there are more consistent reports of low peripheral BDNF in depression that correlates with depression severity (Dell'Osso *et al*, 2010). A study by Angelucci (2005) suggested the valmet66 polymorphism in BDNF could unify the stress, neurogenesis, and hippocampal volume decreases in depression. However, a recent meta-analysis concluded that there is no effect of this particular polymorphism on rates of depression (Chen *et al*, 2008a). BDNF's involvement in a range of psychiatric disorders also makes it less likely that it is a specific mediator of depression, as opposed to a general mechanism for plasticity induction (Castrén, 2004) that is frequently associated with antidepressant efficacy (Sahay and Hen, 2007).

1.2.4 GABAergic hypothesis of depression

GABA-expressing cells are found in all depression-affected brain regions, including the hippocampus, frontal cortex and amygdala and modulate release of serotonin, norepinephrine and dopamine (Bowery, 1989; Takahashi *et al*, 2010). Both ionotropic $GABA_A$ and metabotropic $GABA_B$ receptors can have depressive-related behavioral effects; however, $GABA_BR$ knockouts display a simultaneous increase in depressive and anxiety behaviors, making them less likely to be future

antidepressant drug targets. The overarching idea in the GABAergic hypothesis is that independent GABA deficits in the frontal cortex, amygdala and hippocampus create a hyperexcitable PVN, which then leads to high cortisol levels and blunted HPA feedback (Luscher *et al*, 2010).

GABAergic deficits have been suspected in MDD for the last 30 years on the basis of blood and CSF (Gerner and Hare, 1981; Petty and Schlesser, 1981). This was long-standing support for GABA involvement because depression has extensive comorbidity with anxiety, which is highly responsive to BZ (benzodiazepines - $GABA_A$ receptor allosteric modulators) and BZ's are often adjunctive therapy for depression. Several microarray studies show GABA deficits in occipital, cingulate and pre-frontal cortex, in an approximate match for consensus dysregulated brain regions in MDD (see Luscher (2010) for review), but the results are highly region- and receptor-subtype specific. A more direct test of GABA activity showed differences in GABA-mediated cortical excitability in control vs euthymic vs treatment resistant depression patients (Levinson *et al*, 2010). Because different parameters of evoked responses to TMS are related to the timescale of $GABA_A$ or $GABA_B$ cortical inhibition, it appears that both euthymic and unmedicated actively depressed patients had abnormalities in $GABA_A$, while treatment-resistant patients also had $GABA_B$ deficits. A series of MRS (magnetic resonance spectroscopy) studies show evidence for lower GABA levels in brain regions including various segments of frontal cortex and anterior cingulate cortex (Bhagwagar *et al*, 2007; Sanacora *et al*, 1999) and that patients most resistant to depression had the lowest GABA levels in these areas (Price *et al*, 2009). Thus it appears that one or more types of GABA deficits are associated with depression pathology.

GABA-mediated depression animal models In addition to the observed GABA deficits in depressed patients, animal models of GABA deficits show several depression-related effects. The γ2 GABA channel heterozygous KO mouse, which has impaired neurogenesis and moderate postsynaptic $GABA_AR$ reductions, shows an anxious and depressive phenotype. There appears to be a developmental interaction, wherein the heterozygous KO only produces depressive behaviors

if activated early in development (Shen *et al*, 2010). In vivo, GABA deficits may be created through chronic corticosterone administration activation that leads to dendritic remodeling (Orchinik *et al*, 2001). Similarly, chronic restraint stress leads to loss of PV hippocampal interneurons, through a GR-driven nitric oxide mechanism (Hu *et al*, 2010). The decreased sensitivity to GABA in these regions may then exacerbate the HPA hyperactivation, which in turn leads to further down regulation of GABA.

Antidepressant action through GABA activity SSRI activity has reciprocal ties with GABAergic antidepressant mechanisms (for review see Croarkin (2010)). SSRI's interact with several 5HT receptors found on GABAergic cells and appear to control membrane receptor concentrations through trafficking proteins (Egeland *et al*, 2010). Their action also reverses the low GABA levels observed in occipital cortex of depressed patients (Sanacora *et al*, 2002). In addition to the direct action of AD's on GABAergic cells, they may also exert influence through neurosteroid or neurogenesis levels.

In the neurosteroid mechanism of action, SSRI's act on GABAergic cells to mildly increase production of allopregnenalone (THP) (Pinna *et al*, 2006), which acts through $GABA_AR$'s. Although this hypothesis has not yet been tested in humans, controlled increases in THP injections into the hippocampus lead to increased γ2 subunit expression and antidepressant effects (Nin *et al*, 2008; Uzunova *et al*, 2006). Also the time course of THP production matches the two-week onset latency of SSRI effects. Thus, actions of THP represent a viable avenue for antidepressant effects outside of the canonical serotonin pathways, through GABAR's.

The effect of $GABA_AR$'s activation in immature neurons in the hippocampus is excitatory due to the higher reversal potential for chloride in these neurons. The resultant calcium influx activates several kinases and ultimately CREB, which is found to be low in the hippocampi of depressed patients (Chen *et al*, 2001a) and upregulated in that region by antidepressants (Dowlatshahi *et al*, 1998). Thus GABA signaling links to the neurotrophic hypothesis of depression, as BDNF is transcribed in response to elevated CREB.

1.2.5 Inflammatory hypothesis of depression

Inflammatory activation and modulation are entwined with the etiology of major depression as risk factors, secondary effectors and bridges between other theories of depression. The initial proposal of the inflammatory hypothesis lacked causal evidence, but rather noted the high comorbidity of inflammation-driven processes such as coronary heart disease with depression, and also the similarity of depressive symptoms with behaviors of humans with immune response to viral infections (Smith, 1991). Now there is direct evidence of pro-depressive action of inflammation beyond these associations, as multiple SSRI's decrease pro-inflammatory cytokines (Sutcigil *et al*, 2007). Reciprocally, levels of TNF-α are associated with, and predictive of, SSRI response – patients with high levels of TNF-α are unlikely to show improvement (Eller *et al*, 2008; Kim *et al*, 2008).

Inflammatory modulators also feed back on serotonin levels by affecting the conversion of tryptophan to serotonin. Control over serotonin levels is likely the mechanism behind the association of interferon-α, and major depression. Studies of patients with hepatitis-C, which is treated with interferon-α have found that about half of all patients will develop depression during the course of their treatment. This is one of the few reliable pharmacological means to induce depression in subjects with no history of depression. The main accepted mechanism for immune-induced depression is the action of TNF-α on several enzymes that control conversion of tryptophan to serotonin (Dantzer *et al*, 2008). While this induced depression is responsive to SSRI's, the serotonin-inflammation link is not critically indicated in this form of depression because the somatic complaints of patients remain, even while the “psychological” factors remit. Other depressive mechanisms besides serotonin levels may be recruited by inflammation; IL-6 and TNF-α (Koo and Duman, 2008; Pucak and Kaplin, 2005) have been shown to act as breaks on neurogenensis in the hippocampus, probably due to their interaction with STAT and MAPK (Nakanishi *et al*, 2007; Whitney *et al*, 2009; Zhu *et al*, 2006). The depressive effects of interferon

could be triggered by buildup of other metabolites associated with tryptophan conversion that activate NMDA receptors (Müller and Schwarz, 2007) which further casts doubt on a pure serotonin mechanism behind inflammatory depression.

Inflammation seems to account for a significant portion of social withdrawal and lethargy associated with depression through a set of behavior patterns it triggers known as the “sickness syndrome”. These consist of behavioral and psychological symptoms associated with bacterial or viral infections – lack of interest in normal activities, lack of motivation, lack of appetite and irritability – all of which are components of the DSM-IV definition of depression. These symptoms are known to be mediated by TNF-α, IL1α/β, and IL1/6 (Kent *et al*, 1992; McDonald *et al*, 1987). Certain strains of mice injected with LPS (a bacterial component used to induce inflammation) show evidence of anhedonia in the sucrose-preference test, which is reduced with concomitant administration of SSRI's. The key brain areas associated with the long-term depression component of “sickness syndrome” are the hypothalamus, amygdala and hippocampus (Frenois *et al*, 2007). Because the sickness syndrome behaviors closely match those of depression, and many other hypotheses of depression have difficulty in accounting for the full range of DSM-IV depression symptoms, the inflammatory hypothesis could be an attractive explanation for depression if it integrates mechanistically with other theories.

The mechanism of where/how cytokines create the sickness behavior is still debated, as cytokine receptor densities are very low, but several studies implicate interaction with the HPA axis and particularly the paraventricular nucleus of the hypothalamus (Dantzer *et al*, 2008). Of course, depression and immune response-mediated sickness behavior are not identical, and while sickness behavior remits, depression is maintained for long periods and is not adaptive. This may be reflected in elevated cortisol and cytokines in many depression patients, which is abnormal because the HPA and immune systems are usually mutually inhibitory. Cytokines may break the normal HPA feedback loop by decreasing sensitivity to cortisol on the hypothalamus, through expression of less sensitive receptor subtypes, or certain individuals may possess mutations in

glucocorticoid receptors that naturally decrease feedback (Pariante and Miller, 2001; van Rossum *et al*, 2006; Zunszain *et al*, 2010). The blunted cortisol feedback may simultaneously be ineffective in reducing cytokine production, and thus both HPA and cytokine activity may be elevated.

One key piece of evidence for the inflammatory hypothesis, which is currently absent, would be large-scale clinical trials of anti-inflammatory drugs that specifically target inflammatory mediators in depression, either independently or in combination with other antidepressants. But there is some secondary evidence from a clinical trial of a TNF-α inhibitor (intended to treat psoriasis) that it also decreases depression rates (Tyring *et al*, 2006). Since there are plausible mechanistic links between the inflammatory hypothesis of depression, and monoaminergic neurotransmission, neurogenesis, HPA activity, and the behavioral symptoms of depression, it may be a common target of different pathological routes into depression or a bridge that links the activity of several systems into a pro-depressive force.

1.2.6 Epigenetic regulation associated with major depression

Epigenetic regulation refers to heritable non-sequence DNA modifications that can lead to altered expression levels (Waddington 1957). Two major epigenetic mechanisms, DNA methylation and histone acetylation, may both operate in major depression, but only histone acetylation appears likely to generate a coexpression signature because it tends to expose continuous regions of DNA wherein RNA polymerase may transcribe multiple genes consecutively. (Mechanisms of epigenetic regulation that may be detectable by coexpression analysis are discussed in Chapter 1.4.4.) Maladaptive epigenetic regulation is an attractive depression mechanism in that it links glucocorticoid signaling, effects of early life stress, and the neurotrophic hypothesis.

Early life stress and the level of maternal care in both rodents and primates are associated with epigenetic modifications that maintain a cellular memory of these early events and influence the propensity to develop depression later in life. Using the maternal separation paradigm as an

early life stressor, (Ladd *et al*, 2004) found increased density of glucocorticoid receptors in the PVN, but lower density in the hippocampus (Aisa *et al*, 2007). The increased sensitivity to glucocorticoids was associated with decreased mossy fiber density in the hippocampus, as would be expected from the action of glucocorticoids, and mildly decreased Morris water maze performance (Huot *et al*, 2002). Eye-blink conditioning appears to be a more telling assay of the long-term effects of maternal separation, as adult rats who underwent brief maternal separation had significantly impaired learning, which was accompanied by increased glucocorticoid receptors in the interpositus nucleus (a key nucleus mediating conditioned motor responses) (Wilber and Wellman, 2009). To more directly assess the culpability of epigenetic mechanisms in GR function, Weaver showed that the growth factor NGFI has a two-stage binding to the 17GR promoter, whose methylation status was responsive to levels of maternal care, probably due to 5HTR activation (Weaver *et al*, 2004; Weaver *et al*, 2007). In a striking cross-species study (McGowan *et al*, 2009) the glucocorticoid receptor promoter in the hippocampus was found to be methylated in both mice that underwent maternal neglect and in postmortem brain of patients with a history of childhood abuse. Tendency for pup abuse in rats may even be heritable through epigenetic mechanisms. When pups were raised by stressed and less attentive mothers they showed methylation near BNDF coding DNA in hippocampus and frontal cortex, and the abused pups also took less care of their own pups (Roth *et al*, 2009). This could be due to the pups mimicking parental behaviors, but even when pups were swiftly cross-fostered, they exhibited the neglectful behaviors in adulthood.

Therapeutic manipulation of epigenetic mechanisms The overall balance of HATs (histone acetyltransferases) and HDAC's (histone deactylases, which control regions of chromatin available for transcription) is a critical component of homeostasis, wherein excessive HAT activity is associated with cell-death and HDAC activity is associated with cancer (Barlev *et al*, 2001; Carrozza *et al*, 2003; Minucci and Pelicci, 2006). For instance BDNF levels show disease-related modulation through methylation of its promoter and actions of MeCP2. But BDNF is simultaneously under epigenetic control via HDAC5, which increases in models of social defeat, and is reversed

by imipramine (Tsankova *et al*, 2006). The major limitation in using HDAC inhibitors (HDACi's) to reverse the deactylation often associated with stressful events is that while the stress-related acetylation is fairly gene-specific, HDACi's act across the genome and have a high risk of increasing expression of at least some deleterious genes. Clinically, this non-specificity may be the source of valproic acid's (a HDACi used to treat mood disorders) many side effects and increased risk of various cancers.

The epigenetic modifications of BDNF and GABA receptors, and epigenetic modifications in postmortem depressed subjects, suggest that HDACi's could have antidepressant qualities. A straightforward test of this hypothesis would be to administer an HDACi to depressed patients. But sodium butyrate (NaB - an HDACi) which would be expected to have antidepressant properties, actually increases latency in the novelty induced hypophagia test (Gundersen and Blendy, 2009). While it did acutely increase acetylation in the hippocampus, it did not improve performance in the zero maze. A comprehensive collection of behavioral effects of the social defeat model of depression showed depression reversal with an HDACi, which had a similar expression profile to fluoxetine (Covington *et al*, 2009). From this preliminary evidence using a specific HDACi, it appears that the antidepressant function of monoaminergic drugs do not operate purely through HDACs, but that HDAC activity may replicate some anti-depressant associated activity.

1.2.7 Brain areas critically implicated in depression

Evidence from Conti (2007) suggests that depression and antidepressant activity may be brain region specific, so in Chapters 2-4 we analyze microarray samples from multiple putatively depression-affected brain regions. Likely due to the focus on developing antidepressant drugs, depression is traditionally examined within the framework of global neurotransmitter systems or circulating factors as opposed to specific brain regions. Depression impact on specific brain regions is covered incidentally in previous sections and commonly impacted regions are briefly

reviewed here as well.

Cortico-limbic network as crucial mediator of depression: In a direct attempt to quantify the effects of depression on the amygdala, a number of volumetric studies of post mortem brains have attempted to show a localized loss of cells in this area. These have met with limited success and often conflicting results as a volume reduction was described in some studies (Sheline *et al*, 1998), but does not appear in a number of other studies (Campbell and MacQueen, 2006). However, a decrease in glial density specifically was reported in the amygdala of depressed subjects (Bowley *et al*, 2002) and recently attributed to reduced oligodendrocyte numbers (Hamidi *et al*, 2004). Further validating the amygdala as a central mediator of limbic function, Stein (2007) used structural equation modeling of fMRI data to infer a network of limbic activation, with the amygdala as a central hub, transferring information to cingulate, orbitofrontal, insular, and dorsolateral prefrontal cortex and the parahippocampal gyrus. Examining the amygdala's activity through its divergent connectivity targets, two meta-analyses point towards a corticolimbic circuitry of depression (Mayberg *et al*, 1997; Seminowicz *et al*, 2004). This circuitry includes areas of the prefrontal cortex, the ACC, the hippocampus and anterior thalamic nuclei.

Numerous functional imaging studies have shown emotional reactivity is linked to the amygdala and that changes in this structure track recovery from depression (Rhodes *et al*, 2007; Robertson *et al*, 2007; Siegle *et al*, 2006). However, amygdala reactivity may also be a result of insufficient negative feedback from other brain regions. The strength of amygdala coupling (synchronization) with multiple frontal cortex areas correlates with the strength of emotional regulation in response to negative stimuli (Banks *et al*, 2007). Siegle (2007) found DLPFC was hypoactivated in response to negative words in depressed patients who had high undamped levels of amygdala activation in the task. The extended activation of amygdala in response to negative stimuli, and the hypoactivation of anterior cingulate fits with data from chronic restrain stress showing that the increased glucocorticoids levels lead to amygdala hypertropy (Vyas *et al*, 2002) and frontal cortex hypotrophy (Wellman, 2001).

Frontal cortex involvement in depression Evidence of frontal cortex emotional appraisal and feedback on amygdala may also be seen from individual variability in response to stress. Subjects who showed a significant stress response to a serial subtraction task occurring under social-stress had decreased activity in anterior cingulate and orbital frontal cortex (Pruessner *et al*, 2008). The deactivation of frontal regions and poor performance under stress are correlated with cortisol response levels (Kern *et al*, 2008). Depressed subjects also have a tendency towards rumination on negative events, which has only recently been directly quantified. In contrast to hypoactivation under stress response, depressed patients show higher orbitalfrontal, mPFC and anterior cingulate activation when ruminating (thinking about hopes, dreams and life trajectory) and activity in those areas is sustained when switching from rumination to a distractor task (Cooney *et al*, 2010; Johnson *et al*, 2009). Based on these studies, several areas of frontal cortex appear to be incorporated into stress evaluation, responsiveness, and the pathological activity patterns seen in depression.

The mechanism behind the PFC deficits may be glia-related as postmortem studies (Radley *et al*, 2004) are consistent with the effects of long-term elevated cortisol (Alonso 2000). This fits well with the decrease in apical dendrite arborization in the anterior cingulate observed after chronic restraint stress and the generally lower glucose metabolism in PFC of depressed patients, which is reversed by antidepressants (Baxter *et al*, 1989). The consensus mechanism behind glial destruction in turn is glutamate excitotoxicity from persistent immune activation (see Chapter 1.2.5).

Neurogenesis studies suggest hippocampal role in depression: Numerous genes linked to neurogenesis, dendritic arborization and spine formation, HPA regulation, and inflammatory processes/cytokines are preferentially altered in rodent hippocampus under a variety of stress paradigms (Alfonso *et al*, 2005). Chronic restraint stress reduced arborization in CA3 neurons and this appears to have functional consequences as these rats show reduced LTP in CA3 and dentate gyrus. These effects can be largely prevented by NMDA blockers. Mineur (2007) linked

performance in hippocampal and non-hippocampal dependent tasks following the UCMS protocol to the level of neurogenesis in the subgranual zone. While these deficits and lack of neurogenesis were reversed by AD treatment, there was a strong effect of sex and genetic strain on the results that does not permit these results to be generalized. In an attempt to increase the spatial-temporal resolution of such studies, Airan (2007) used voltage sensitive dye imaging to track activity propagation across mice after UCMS. They found lower levels of activation in the dentate gyrus of UCMS mice vs. controls that was reversed with SSRI and TCA AD treatment, but clinical improvement was not prevented by irradiation. The normally anxious BALB/c mice do not require hippocampal neurogenesis to recover from UCMS with SSRI's (Holick *et al*, 2008), and antidepressant effects achieved through exercise and environmental enrichment also do not require neurogenesis (Meshi *et al*, 2006). Thus it appears hippocampal neurogenesis is necessary for some antidepressant effects and closely linked with cognitive deficits in depression (Sahay *et al*, 2007).

1.2.8 Practical recommendations for more effective molecular hypotheses of depression

Why are there so many competing hypotheses for the biological basis of depression? We suggest that depression research is cast against a pair of major roadblocks in neuroscience research: one scientific and one human-based. The major scientific roadblock to greater understanding of depression (which would be evidenced by actual improvement in treatment response rates) is that there are no agreed upon landmark findings to guide depression research. Specifically, there is no consistently used animal model, behavioral test, biomarker, cell-type, brain region, gene, pathway, or neuronal function which is specifically associated with the depressive state in all patients. Without such a landmark finding, researchers are forced to look for mutually consistent sets of findings, which are then inferred to represent a concerted detriment to some

cognitive process underlying normal function. However, such a unified theory of depression correlates has not been produced, because there is little understanding of the interoperability of multiple neurotransmitters in creating mental states, and thus it is difficult to relate findings from different brain regions and experimental paradigms. The issue is compounded as depressed patients likely have more than one deficit, and perhaps even several minor deficits/abnormalities (Nikolaus *et al*, 2009). Therefore, researchers studying subjects who show deficits in a given system of interest, are in fact studying deficits in additional systems, but the degree to which multiple deficits are distinct or part of some concerted dysfunction is unclear. Thus, depression research has reached a point of such obfuscation that it is not even clear when two particular results are in conflict – frequently they simply contribute to an acausal morass of depression-related factoids, at best waiting for the post-hoc theoretical framework of a review paper, and at worst forming the literary equivalent of cosmic background radiation from an expanding universe of undirected depression studies.

Of course, mental function is based on a combination of many neural systems including micro and macro structural connectivity, receptor densities as determined through complex second messenger systems, hormonal regulation, and slow and fast time-scale neuronal excitation. Thus the imprecision and redundancy of depression research in reality is even greater than expected for classic neurotransmitter hypotheses, as there is no unified framework for how various depression-related deficits that are apparently scattered across the breadth of neuroscience may relate to each other or some common framework. To be more concrete, there is no agreed upon core mood circuitry, alterations in which are capable of generating depression in response to changes in levels of one or more neurotransmitters, through modulation of some specific neuronal function. Nor is there a single computational or theoretical model of what type of neural activity is associated with depressive brain-states. Indeed the closest approximation of what depression actually might be (as opposed to random manifestations that accompany it) is a non-mechanistic serotonin-based reinforcement learning model in a slug (Dayan and Huys, 2009).

However, lack of a unified theory of brain function does not prevent (treatment-evidenced) progress in other disorders, so why should it be necessary to understand integrated brain function before progress can be made on depression? Astonishingly, amid the thousands of studies showing specific fiber tracts, cell-types, neurotransmitters, hormone levels, there is almost no effort made to link different results into a conceptual framework which collapses these deficits down to simpler systems behind the clinical depression: symptoms of anhedonia, lack of motivation, and obsession with negative problems. While the definition of depression may indeed cover several different biologically distinct subtypes, it is not these biological divisions that allow results to accumulate and never integrate, but mutual indifference of researchers to theories other than their own. This is the second reason for the plethora of hypotheses. Thus the field continues to amass new observations of dysfunction, apparently in the hope that some single observation will revolutionize our understanding of depression, but more likely because grant funding is set up to support reductionist investigations, not conceptual synthesis. Despite the clear intellectual need for emergent integrated theories of depression, the economic, educational and technical challenges to creating a legitimate integrated model of depression (something more than an arrow diagram) has been sufficient to prevent progress in this key regard.

Whatever the factors behind the failure of psychiatric research to find an effective description of depressive mechanisms, for the last 60 years depression research has never coalesced around a core dysfunction, in the way that schizophrenia research or Alzheimer's research have circled around GABA deficits or plaque accumulation. While those molecular markers carry their own ambiguity, at least they are a consistent finding and appear somehow tied to a core mechanism. While many supposed and real breakthroughs have occurred, when only 1/3 of depressed patients show complete remission, objectively, the scope of those “breakthroughs” has limited clinical impact. Since antidepressant activity is not tied to pathology or pathogenesis, and because depression may well consist of multi-system dysfunction, attempting to understand specific characteristics of the disorder on a specific level has limited relevance to the systemic

impact of depression on the brain. In a troublesome loop, lack of understanding about pathogenesis may have led the field to focus on needlessly complex secondary and adaptive features that further distract attempts to find convergent neural mechanisms of depression.

In this setting of competing hypotheses, only weakly associated with actual depression pathology, the results of Chapters 2-5, and in light of the exodus of pharmaceutical funding from neuropsychiatry (Miller, 2010) we offer a working philosophy of depression to guide the trajectory of future research. It may appear naive to propose a new theoretical paradigm in a field with so many competing hypotheses. However, given the enormous scientific challenge of depression and the undirected piles of evidence for depression deficits, adhering to this approach could provide great benefit with little additional funding. Futhermore, we make practical recommendations for how to construct an inclusive model of depression mechanisms via this philosophy (Chapter 5.4).

Working philosophy of depression: Consider all the major depression hypotheses (monoamine, neurotrophic, etc) to be correct. The goal of researchers operating under this working hypothesis will be to create a new style of depression research that draws connections between the different hypotheses, finding instances of mutual activation and convergent neural substrates of depression that could be activated by several pathways. The theoretical justification for this approach is the many cross-links between the various hypotheses of depression – activity which would be classified under one particular hypothesis of depression is likely feeding into “competing” mechanisms (Figure 1). Even disregarding the links between theories, it is clear from the common clinical endpoint of depressive symptoms that the various hypothesized mechanisms of depression are all pushing the brain state into a dynamic regime that produces depressive behaviors and mental states. The existence of several different (though probably related) paths into the depressive state indicates that depression could be a unitary concept (in terms of brain dynamics underlying recurrent negative thought patterns) but present clinically with several deceptively unrelated markers, that would seemingly indicate multiple pathologies.

By looking for interoperability of different hypotheses of depression, it appears likely that we would move closer to a unified understanding of the common neural effects of the various mechanisms, and therefore closer to understanding depression pathology. The intention for this working philosophy is that by understanding how depression mechanisms operate cooperatively, it will reduce the confusion caused by redundant research efforts, and hopefully develop insightful hybrids of existing theories. Based on the interactions and cumulative effects of various hypotheses of depression, and then incorporating the combinatorial effects observed in Chapters 2-4, we propose a testable theory of depression pathology under this philosophy, termed the floodgate model of depression (Chapter 5.4).

1.3 COEXPRESSION NETWORK STRUCTURE AND ANALYSIS

The purpose of this coexpression network overview is to highlight research which is representative of major trends in coexpression analysis. We only report on coexpression networks that are based on first-order Pearson correlations between genes/probe-sets across replicate samples. Many other networks, including various Bayesian networks, higher-order partial correlation networks and mutual information-based networks have been applied to microarray analysis and offer complementary information to coexpression networks, but are only mentioned as supporting information because human post-mortem sample sizes are usually insufficient for these other approaches and the research in Chapters 2-4 focuses on Pearson-based coexpression networks.

The past decade of coexpression research largely falls into two historical periods, each with a particular methodological and thematic focus. Initial studies (roughly years 2002-2004) demonstrated that gene-gene correlations represented the convergent influences of many biological control structures (covered separately in Chapter 1.4) and established the small-world

and scale-free connectivity properties of coexpression networks that are fundamental to future studies. The latest stage of coexpression research (2005-present) has been focused on understanding how topological properties of coexpression networks (which encapsulate many biological relationships) relate to cellular activity in healthy and disease states. Within disease coexpression research, key approaches are either hub-based, module-based, differential-coexpression based (examining altered network structure in disease), or use some combination of these approaches (see Chapters 1.3.1-1.3.4 for review of each approach).

Basic coexpression network structure The structure of coexpression networks – the transcriptome-wide configuration of correlation-based links between genes – reflects many underlying cellular processes (Chapter 1.4). However, the abstract network structure itself has properties that orchestrate the flow of information through molecular pathways. Discovery that metabolic networks (composed of substrate-ligand interactions) and protein-protein interactions networks (composed of proteins linked by physical interactions) were scale-free and small-world (Guelzim *et al*, 2002; Jeong *et al*, 2001) presaged coexpression network structure. There were early examples of large-scale clustering of gene expression profiles (synchronous fluctuations of multiple genes across replicate samples) (Eisen *et al*, 1998; Tavazoie *et al*, 1999), but later series demonstrated the formal small-world and scale-free structure of coexpression networks, based on Pearson correlations between all pairs of genes (Agrawal, 2002; Ihmels *et al*, 2004; Jordan *et al*, 2004; Pavlidis *et al*, 2004; van Noort *et al*, 2004). Scale-free and small-world topologies are common to man-made networks (Barabasi and Albert, 1999) and many natural systems including human neural and non-neural datasets (Horvath *et al*, 2006; Purmann *et al*, 2007). Their biological implementations, particularly coexpression networks, are characterized by highly clustered (mutually interconnected) “modular” communities of genes, and low average pathlength between nodes, courtesy of the wide-ranging connectivity of hub genes (see Chapter 1.3.1).

The small-world and scale-free organization of coexpression networks persists across species, with significant conservation of links between species (Lee *et al*, 2004; Prieto *et al*, 2008;

Tsaparas *et al*, 2006). Coexpression relationships and larger “modules” of mutually coexpressed links have been repeatedly shown to have common biological functions, usually through GO (gene ontology). Gene ontology is a curated hierarchical classification of individual gene properties using controlled vocabulary to specific gene function on multiple dimensions (Ashburner *et al*, 2000). It is a useful tool for functionally characterizing large numbers of genes, as it can return the most common biological classifications, and an estimate of what functions are over-represented in a group of genes compared to chance. Gene coexpression networks may reiterate or predict known biological organization as annotated by GO or other classification systems, and therefore can implicate new genes in disease through guilt-by-association algorithms or reveal disease insights through their connectivity structure (Nayak *et al*, 2009). Two structural properties of these networks in particular – hub nodes and modularity are central to subsequent understanding of transcription dynamics and potential disease mechanisms.

1.3.1 Leveraging hub connectivity as a functional marker of network activity

In the small-world framework, hub nodes are network elements with non-local connectivity that serve to bridge different communities in the network, thereby lowering the average pathlength without destroying the overall clustering structure in the network (Watts and Strogatz, 1998). Hubs in scale-free networks are rare nodes with connectivity to a significant portion of the network. Remaining “provincial” nodes in scale-free networks have relatively few connections while the overall connectivity distribution follows a power-law distribution. Information flow through scale-free networks is unlikely to be affected by random node deletion, but is especially vulnerable to targeted attacks on hubs (compared to random networks) (Albert *et al*, 2000). The particular vulnerability to targeted attack and the numerous instances of scale-free networks in natural systems has led to the concept of hub-targeting in small-world/scale-free networks as a potential disease mechanism, often known as the “lethality-centrality” relationship. This was supported for molecular networks by

the classic work of Jeong (2001) who showed that lethality of particular gene deletion in yeast was proportional to the connectivity of the node in the protein-protein network. Examples of how hub-targeting can lead to crucial functional impairment in humans also come from cellular and brain region networks with similar structural organization. For example, Stam (2007) showed that Alzheimer's disease is accompanied by selective loss of small-world hub connections in EEG-based networks and epileptic activity is accompanied by altered clustering coefficients in local microcircuits in the hippocampus (Kramer *et al*, 2008; Netoff *et al*, 2004).

Example hub-centric analysis The gene *ASPM* was previously implicated in control of cortical size when Horvarth et al. (2006), in a classic application of hub-gene based coexpresion analysis, predicted that it was central to formation of glioblastomas. In brief, the weighted gene coexpression network analysis (WGCNA) routines employed in this study used topological overlap (a module completion/detection algorithm) to identify modules in the gene-gene correlation matrix that have been weighted to suppress small magnitude correlations, which are likely to have been generated randomly. The modules are then assigned to one or more biological functions by hand or through gene ontology. In a network generated from glioblastoma data, this method identified a module enriched in oligodendrocyte genes. *ASPM* was the most connected gene in this module, meaning that fluctuations in its expression level closely matched those of other oligodentrocyte genes. This gene showed a 40-fold increase in expression level in glioblastomas vs control, and hence is far from the subtle marker of dysfunction typically seen in depression. However gene therapy silencing *ASPM* using shRNA did halt development of tumors in mice, perhaps effective in part because of the centrality of *ASPM* to neural precursor proliferation (Bikeye *et al*, 2010).

Unlike complex neuropsychiatric disorders, differential expression ratios found in cancer microarrays are quite large, but it is difficult to prioritize among possible candidates. Using the WGCNA procedure on testicular cancer samples, Wang (2009) found a module enriched in genes known to be associated with disease progression and selected several well-connected hubs (that were found to be related to cell-cycling) within this module as potential key disease mediators.

While their selection criteria were relatively ad-hoc, they avoided the danger of using WGNCA as a platform for molecular stargazing, by showing there is a strong relationship between a gene's connectivity and the severity of disease when that gene is differentially expressed. The authors simultaneously measured levels of micro-RNA (miRNA) and found that several differentially expressed miRNA's bound to the cancer hubs. While there was no statistical estimation of the likelihood of this occurrence, combined with evidence for common miRNA disregulation in multiple cancers, it points towards an expression control mechanism that utilizes coexpression hubs.

Causality of hub identity vs hub connectivity in centrality-lethality relationship The centrality-lethality/disease relationship has been validated in many systems, but it is unclear if highly connected nodes are intrinsically important, or if their biological importance stems from their numerous connections. This question is fundamentally a debate on the relative importance of pure network mechanisms vs specific biological explanations of the influence of hub nodes. Within gene networks, the debate takes the form of determining if hub node ablations are deleterious because the nodes are essential, or if it is merely because they act on numerous other nodes, some of which are likely to be essential (He and Zhang, 2006). Essential genes are defined as those genes which are found to be lethal in systematic knockout experiments (Winzeler *et al*, 1999). He (2006) (2006) found that the results from Jeong (2001) were consistent with a model of essentially protein-protein interactions, randomly distributed throughout the network. Since hub-node disruptions are more likely to intersect one of these crucial links, this would ostensibly account for their increase lethality over non-hubs. In a complementary test of the relevance of network structure to hub lethality, Zotenko (2008) found that deletion of essential genes was no more likely to affect flow of information in a network than would the deletion of an equal number of hubs. Furthermore deleting high betweenness nodes (which have far-flung network influence) in the protein-protein interaction network (PPI) was no more disruptive to information flow in protein-protein networks than was deleting an equal number of nodes based purely on degree. However, by incorporating many different topological measures, especially clustering information, it is possible to very accurately

create a filter that predicts which genes will be essential in yeast (del Rio *et al*, 2009). These results are from protein-protein interaction networks in yeast, and may not apply directly to coexpression or other networks with more heterogeneous control structure. But they do illustrate that the global network contribution to essential gene lethality is at least more nuanced than originally predicted.

1.3.2 Source of small-world and scale-free coexpression topology

Theories concerning the origin of scale-free and small-world structures naturally focus on the evolution of modularity and hub nodes – the hallmarks of these topologies. The cannonical process for producing scale-free networks is “preferential attachment”, a network generation process wherein new network nodes are gradually added to a seed network structure, and the probability of connecting to a particular node in the seed network is proportional to the number of connections it has in the seed network (Barabasi *et al*, 1999). In this model, nodes with many connections gain even more connections, and this rich-get-richer pattern produces a power-law connectivity distribution of connectivity, which is found in many naturally occurring networks. Preferential attachment does not necessarily produce the modularity observed in gene coexpression structure. Ravasz (2002) found evidence of a hierarchy of modules in many metabolic networks i.e. modules nested within modules, each of which was largely devoted to a particular cellular process. Coexpression networks may be created by systems that use hierarchical control (Chapter 1.4.2), but they themselves do not show rigid hierarchy of modules (Dorogovtsev *et al*, 2002; Jordan *et al*, 2004; Ravasz *et al*, 2002).

A theory that accounts for both the scale-free and modular organization of coexpression networks combines preferential attachment algorithms, with the gene duplication and divergence model (van Noort *et al*, 2004). The model consists of virtual genes that are linked if they share common transcription factors. When new genes are created by duplication of existing genes, then

the transcriptional control of that gene will likely resemble its predecessor (since the promoter regions of the two genes bind the same transcription factor) and thus gene duplication leads to clustered communities. Genes in the model also have a gradual rate of divergence from their paralogues due to random additions or deletions to transcription factors that control them. Any new regulatory interactions the gene gains are likely to be with the most connected elements of the network – the preferential attachment scenario. Together this theory reproduces the modularity (from similar connections via gene duplications) and scale-free connectivity (through preferential attachment of new regulatory relationships to hubs) found in coexpression networks.

1.3.3 Use of modules as functional markers in coexpression analysis

These two previous studies (Chapter 1.3.1) which rely on the idea of hubs as crucial mediators of disease also incorporate modularity to identify the most relevant hubs to particular disease-affected subsystems. A supporting finding behind this logic is that coexpression modules generated from heterogeneous tissues are distinctly enriched with markers of certain cellular populations (for example in brain for glia, oligodendrocytes and glutamatergic/GABAergic neurons). To rigorously support this claim Oldham et al. (2008) showed that modules are frequently enriched with markers associated with specific cell-type populations. The linkage between modules and morphological details is even more fine-grained, as illustrated by the finding that sub-modules within a mother GABA module are associated with specific cell sub-types, and sub-modules within a mitochondrial-focused module were localized to different cellular compartments in neurons (Winden *et al*, 2009). Thus, the modular components of coexpression networks appear to be recruited *en masse* to facilitate particular biological functions that are associated with, or even create, morphological diversity.

This concept was actively tested by knocking out *RGS4* and *DLX1/2*, which were well connected in their modules and linked to pre-natal death and schizophrenia, respectively (Winden

et al, 2009). Differential expression between control and KO samples was significant and largely confined within those modules wherein the knocked out genes were originally members. These interventions provide evidence that transcriptional control acts through the coexpression network, or the mechanisms that generate the network structure. Since coexpression networks are static, such comparisons of networks between two conditions can highlight genes that are more causally linked to observed effects. For instance, a comparative module-membership approach characterized a set of genes which are likely involved in neurogenesis. By contrasting membership in an astroglia-related module found in the subventricular zone of the caudate nucleus with the module membership of an astroglia-related module common to many brain areas, Oldham et al. (2008) produced a short list of candidate genes predicted to be associated with neurogenesis in the adult human brain. Using differences in connectivity (differences in module membership) to highlight genes related to a specific cellular process is a reversal of the standard logic that common functions reside in specific modules. As modules provide a specific biological context for the activity of hubs, so too does differential connectivity extend the relevance and power of module-based understanding of biological functions (see below).

1.3.4 Differential coexpression – using changing network structure to highlight disease effects

Many coexpression network-based studies use guilt-by-association algorithms to find new putative disease genes. These rely on identification of hub-genes or modules that have high connectivity to known disease genes. The methods are therefore not completely unbiased, as they rely on existing knowledge, and they are acausal, in that the putative disease genes may only be correlates of the essential disease process, because they have been selected from a static network structure. Differential coexpression algorithms skirt both of these issues by contrasting network connectivity in control and disease states to highlight genes associated with structural

alterations that occur between control and disease states. Differential coexpression algorithms are aimed to detect changes in the correlation values between the full correlation matrices generated by control data vs correlation matrices generated by disease data. Differential coexpression between two genes or groups of genes may or may not occur alongside differential expression, as the two measurements are mathematically independent. Studies have shown both an interesting collusion of differential connectivity with differential expression but also that differential connectivity can highlight causal factors that differential expression fails to detect (Reverter 2010). As detection and understanding of differential coexpression improves, it may be possible to track adaptive cellular processes as new modular functions are recruited/disbanded, or to determine the transcriptional control mechanisms behind these changes, which would be useful even in traditional microarray analysis. But these techniques to track the evolution of network structure in disease, and the potential for differential coexpression analysis, are tempered by several statistical challenges.

The increased detail of differential coexpression is accompanied by increased statistical challenges determining if two networks, each based on noisy correlation matrices, are significantly different. While multiple testing limitations are a constant concern in traditional microarrays, differential coexpression faces severe multiple testing challenges, because in a network of *n* genes there are (*n*^2/2-*n*) possible links (correlations) that may change between conditions. This huge number of possible individual changes has prompted most researchers to adopt techniques that aggregate differential coexpression on a per-gene, per-probeset or per-module basis, reducing the scale of multiple testing adjustments to those of traditional microarray experiments. Establishing high confidence coexpression links requires significantly more replicates than traditional expression-level experiments, but finding stable distinctions between correlation values requires even more replicates: the few existing differential coexpression analyses have used between 30-300 samples per group, which severely limits current applications. For these reasons, differential coexpression has been used for cancer datasets and knock-outs wherein the disease effects are

relatively large, compared to psychiatric datasets. However, there would be significant benefits in a more causally linked understanding of how disease impact percolates and reconfigures hundreds of interactions in neuropsychiatric disorders, if sufficiently large datasets become available.

Because differential coexpression analysis is recent and limited, almost every publication to date employs a different statistical test to find these changes in correlation. The methods may be broadly grouped into approaches that (1) identify significant correlations in control and disease networks separately, in which differential coexpression is defined as net gain or loss of correlation for particular genes (Choi *et al*, 2005; Fuller *et al*, 2007; Kostka and Spang, 2004; Lai *et al*, 2004; van Nas *et al*, 2009; Wong and Huk, 2008) or (2) approaches that stress the total amount of rewiring (differential correlations), regardless of changes in net connectivity (Hudson *et al*, 2009; Leonardson *et al*, 2010; Reverter *et al*, 2010; Tesson *et al*). While these two approaches may appear similar, the second approach skirts the troublesome issue of fixing individual cutoffs for control and disease networks, and facilitates simultaneous detection of both gain and loss of coexpression associated with single genes or pathways. Nevertheless, the first approach has also been used successfully, exemplified by Choi (2005) who used extensive permutation testing to conclusively demonstrate, for the first time, the existence of differential coexpression between control state networks and a collection of cancer datasets. To understand the contribution of network reconfiguration to disease, they identified modules of genes which were correlated only in control or cancer-state networks, which were generally related to control of cell-cycling. Thus, groups of genes which are differentially correlated, appear to correspond to disease processes.

What other biological pressures can produce differential coexpression? Cataloging how networks change across different organisms may improve our basic understanding of how network architecture supports normal function, which ultimately could lead to more informed analysis of disease states. Speciation, aging and sex appear to each be accompanied by specific network alterations (differential coexpression) that are related to the comparative function of organisms. In a study that simultaneously considered differential expression, differential coexpression, and

differential modules between humans and chimpanzees (Oldham *et al*, 2006) researchers found that genes over-expressed in particular brain regions tended to be coexpressed in modules. Furthermore, modular conservation (similarity of connections within modules of genes over-expressed in a particular brain region) reiterated evolutionary conservation between human and chimpanzee brain regions. This divergence in connectivity appears to be driven by divergence in genome sequence, as highly differentially coexpressed genes had significantly larger sequence changes between species than genes that maintained connectivity across species. Thus it appears that coexpression network organization corresponds to those biological processes that distinguish larger morphological and developmental features.

Identifying differential coexpression between different age-ranges in humans is difficult because thousands of genes have a robust existing correlation with age and because it is unclear where to segment the data into relevant epochs, from which to calculate correlation values. Since wavelets have the ability to contrast behavior across multiple scales, Gillis et al. (2009) used the Haar transform to compare coexpression patterns from several age groups ranging from pre-natal to geriatric. They found that even excluding genes which have expression-level correlations with age, there are significantly altered correlations between genes across multiple age-ranges. These appear to be very biologically coherent, as the GO over-representation for differentially coexpressed genes had significantly superior ROC characteristics, even to standard coexpression modules. These analysis of these GO processes is complex due to many possible pair-wise comparisons between age ranges, but they roughly reiterate known aging properties, for example with hormonal-mediated processes associated mainly with early life. Chromatin remodeling may be one process specifically linked to differential coexpression, as gene members of modules that became decorrelated with increasing age (in mice) were frequently colocalized on chromosomes (Southworth *et al*, 2009). Differential coexpression also occurs on an even shorter time-scale, as a study using rare human time-series microarrays over a 24-hour period found that thousands of genes in whole-blood samples experience significant rewiring in response to food and diurnal

rhythms (Leonardson *et al*, 2010). Thus it appears that differential coexpression is actively controlled within organisms and may be a component in implementing particular transcriptional programs.

Further indication that changes in microarray correlation structure relate to functional differences may be found in multi-tissue comparisons of male and female mice. van Nas (2009) noted that the majority of links were conserved in all tissues: the overlap in significant links varied from relatively low correlations of 0.5-0.7 (adipose and liver) to greater than 0.9 (brain and muscle). Within this broad similarity, simultaneous differential expression and coexpression analysis showed punctuate modules found only in males that were enriched in genes related to spermatogenesis. Aside from these specific differences, since the network structure remained intact, and there were a large number of differentially expressed genes in all tissues, this indicates that expression level is not necessarily dictated through network structure. Rather, based on evidence in the study, within species, the potential correlation structure remains in the background and may be activated (show increased expression) via transcription factors that act on all members of the module. While there is no apparent mathematical or biological principle that seems to consistently determine when differential expression and differential coexpression overlap, (or not), they do appear to regulate sets of functionally related genes in concert.

1.4 BIOLOGICAL COEXPRESSION MECHANISMS

Intro to transcriptional mechanisms behind coexpression Based on these numerous studies, coexpression links and higher-order coexpression network structures (modules) appear to reflect meaningful transcriptional programs that provide the cell with appropriate levels of mRNA and are devoted to some unitary/modular function. While this may be the purpose of correlated expression among functionally related genes, what are the physical mechanisms that produce

gene coexpression? Correlated gene expression reflects multiple influences from structural and dynamic processes in the cell. The most reproduced and influential mechanisms relevant to synchronous mRNA expression are common transcription factor binding sites, chromosome location, epigenetic modification and chromatin remodeling. These influences on transcription each have their own regulation, and these mechanisms interact and all contribute to the final output of correlated gene expression. Thus, coexpression links, or changes in coexpression structure are a proxy for many background regulatory processes. Therefore, to interpret results of gene coexpression studies, it is helpful to know what biological processes are encapsulated and measured by gene coexpression.

1.4.1 Genome organization is a foundation of coexpression

There is a long-standing observation in multiple species that genes with similar functions tend to be proximal to each other on the genome, specifically that adjacent genes are likely to share GO categories (Caron *et al*, 2001; Cohen *et al*, 2000). Since RNA polymerase apparently acts outside of the exact domains of individual genes (Boutanaev *et al*, 2002; Ebisuya *et al*, 2008), and since functionally related genes are often adjacent, these two mechanisms operating together could combine to regulate functionally related genes (Xu *et al*, 2009), which would ostensibly be advantageous for the organism, and which would then be detected as synchronous transcript fluctuations by microarray (Chen and Zhao, 2005).

The definition of what constitutes adjacent genes may have a significant influence on the odds of coexpression. RNA polymerase runs 3’ to 5’ on DNA, and thus genes are typically read off sequentially in one direction from a single strand of DNA. However there are many bidirectional promoters (Trinklein *et al*, 2004), that can lead to transcription of genes that are actually upstream of the promoter on the opposite strand of DNA. The operation and function of these bidirectional promoters are still debated: some studies find their gene products tend to be antiregulated (Chen

et al, 2010), though others have found a lack of coregulation (Ebisuya *et al*, 2008) and yet others have found positive correlations (Chen *et al*, 2010; Sémon and Duret, 2006; Trinklein *et al*, 2004). The function of genes transcribed through bidirectional promotors tend to be related to DNA repair and somatic cancers, (for instance BRCA2 has bidirectionally transcribed partners) though the reason for this association and physical mechanisms of bidirectional polymerase recruitment by a single transcription factor are unknown (Yang *et al*, 2007).

Analysis of how coexpression relates to inter-gene distance (on the same DNA strand) indicates that close proximity (under 50kb) leads to a relatively small, but highly significant increase in coexpression r=~.1 between genes (Baskerville and Bartel, 2005). If more than two genes are found in close proximity (known as a cluster of genes, and precise definitions of cluster varies between studies), then the odds of coexpression are much higher (r=~.8), at which level coexpression could be detected by microarrays (Ng *et al*, 2009), though this is likely organism- and tissue-specific. Despite the overall higher correlations between genes defined as coexpressed in microarrays (e.g. r>0.7) compared to the low average coexpression of adjacent genes, in humans, adjacent genes are twice as likely to have common GO categories as coexpressed genes (Purmann *et al*, 2007).

To quantify the relative influence of genome position on generating functionally related coexpressed genes vs other mechanisms of producing related pairs, Yanai (2009) compared transcription in two morphologically similar species of nematodes and found that coexpression of adjacent genes was largely species-specific, unless the genes were essential. The expression level of genes which changed genome position was strongly dictated by levels of its (new) neighbor. Similarly, Ebisuya (2008) found that when cells expressed early immediate genes (IEG's) in response to environmental stress, they also tended to coexpress “functionally unrelated” genes as well, which were close to the early immediate genes. Such “ripples” of transcription would produce another level of transcriptional control, if the position of genes on the genome represents a distance code for functional relationships. However, the ripples were also associated with histone

acetylation in the region of the IEGs, indicating cooperation between genome distances and another mechanism of coexpression.

Because adjacent genes interact with the histone and transcription factors which also have clustered activity on the genome, it seems possible that the functionality associated with adjacent genes could be epiphenomenal and completely accounted for by these other clustered functions. However, attempts to subtract out these interacting clustering mechanisms by tracking the location of transcription factor binding sites and using ChIP-chip measures to detect histone binding, could not account for the majority of coexpression in adjacent genes (Purmann *et al*, 2007). In summary the genome appears to be organized in such a way that it supports coexpression of functionally related genes. The intrinsic contribution of inter-gene distances to observed coexpression in microarrays may be small, but the adjacency of functionally related genes interacts with histone and transcription factor binding to facilitate expression of related genes.

However, the exact contribution of each mechanism with genome organization, and a consistent evolutionary explanation for how genome reordering leads to functional advantages (beyond the basic idea that coexpression of functionally related genes is beneficial) are currently lacking. Furthermore, conflicting and imprecise definitions of what constitutes a domain/family/cluster hamper direct comparisons of results. A coherent analysis of overlap between the intersecting clusters, defined by the several biological systems that interact with DNA, would be helpful in clarifying how these control system combine to produce appropriate expression levels. To prevent needless parallel development of similar ideas, or worse needlessly conflicting paradigms, methodological “shootouts” between various methods would be very informative. While some studies of cooperative coexpression mechanisms exist (Byrne *et al*, 2007; Ren *et al*, 2005; Zhan *et al*, 2006), a more typical pattern of research is to propose new methods without systematic comparison to older methods or application to multiple organisms.

1.4.2 Biological basis of coexpresssion – transcription factors

Levels of transcription factors (TF's), which bind to DNA and facilitate transcription, are a major determinant of mRNA levels (Jothi *et al*, 2009) and their nuclear receptors are the targets of 10% of all therapeutic drugs (Overington *et al*, 2006). Because each TF may have numerous DNA binding sites, they are hypothesized to be a major source of correlated gene expression (Allocco *et al*, 2004; Altman and Raychaudhuri, 2001; Brazma *et al*, 1998; Marco *et al*, 2009). By the transitive property, the study of coexpression networks will be in part the study of transcriptional regulatory networks. TF's themselves are subject to regulation (activation and inactivation) by other TF's in what are known as transcription factor networks, or more broadly, transcriptional regulatory networks (Babu *et al*, 2004) (Guelzim *et al*, 2002) (Yu *et al*, 2003). Therefore, properties and structures of transcriptional regulatory networks may be relevant to the endpoint mRNA expression and associated cellular states.

One important caveat to the details of TF activity is that most systematic studies are completed in *E.coli* or *S.Cerevisiae* since ChIP-chip assays are noisy (likely due to intersecting biological influences on TF binding) and therefore the highest quality TF maps are available for these organisms (Zhu *et al*, 2007). While it is hoped that principles from these organisms extend to humans, some facets, such as coexpression of genes for proteins found in complexes, (Zampieri *et al*, 2008) may not apply equally well to humans (Xulvi-Brunet and Li, 2009) and identification of TF binding sites in higher organisms is expected to be difficult and currently less accurate (Tompa *et al*, 2005).

The broad structure of TF networks is one of small in-degrees and large out-degrees, meaning that TF-TF regulation is relatively simple, but each TF controls many gene targets. This asymmetric nature of TF connectivity is also seen in coexpression networks, wherein sorting genes by their signed connectivity (number of positive and number of negative correlations) reveals gene hubs that code for transcription factors have large positive correlations with many genes and few

negative correlations (Gustin *et al*, 2008). The global connectivity of transcription factors (the set of TF-TF interactions and the gene targets of expressed TF's) and the distribution of network motifs are globally reconfigured between biological states in yeast (Luscombe *et al*, 2004; Ni *et al*, 2009). Most of the 150 TF's in yeast are expressed in response to more than one endogenous or environmental perturbation, indicating these networks use combinatorial logic to produce a larger number of cellular responses than expected under pure 1-to-1 TF-to-environment regulation.

What is the mechanism that allows for specific expression of the exact set of genes necessary for adaptation to external influence on a cell? A hierarchical organization in the transcription regulatory network would allow controlled expression of a coherent set of target genes (Ma *et al*, 2004; Yu *et al*, 2003). While hierarchical organization has not been conclusively dismissed (by examining the relationship of comparing clustering coefficient to degree (Barabasi and Oltvai, 2004), the presence of numerous feedback loops among transcription factors and their targets makes pure hierarchical feed-forward architecture unlikely. Despite the presence of feedback loops in regulatory networks, which would appear to prevent hierarchical specificity in gene targeting, it appears that transcription factor binding sites are organized on DNA in a manner suggestive of transcription that is highly targeted at specific selection of genes. Specifically, TF binding sights are most commonly found on specific chromosomes, and on those chromosomes for the specific TF there are select regions with high densities of binding sites, and within those regions binding sites are often adjacent to commonly coexpressed gene sets (Janga *et al*, 2008; Vogel *et al*, 2005).

Practical caveats in hierarchical control of expression by transcription factors The concept of a hierarchical control structure (transcription factor network) converting environmental and endogenous signals into appropriate and comprehensive sets of transcriptional activity is very elegant. While some studies indicate that transcription factor networks act in approximately this fashion, in practice there are limitations on the power of this mechanism to control expression. For instance, the transcription factors are assumed to regulate proximal genes, but may leave those

unaffected, and in fact regulate distant genes (Hartman *et al*, 2005). Alternately, deleting a transcription factor does not universally silence its targets (Gasch *et al*, 2000; Greenbaum *et al*, 2003). Contrary to assumptions behind algorithms that generate TF network structure, in some cases, the binding site is far away from the observed gene response (Carroll *et al*, 2005). Furthermore, transcriptional regulatory networks show an irregular pattern of conservation across species, wherein closely related species may have widely different regulatory interactions, while quite different species have similar networks (Venkataram and Fay, 2010). Thus many aspects of TF activity are not fully quantified, they form a likely, but not universal, basis for coexpression relationships.

Some of these points of uncertainty about TF function may be explained by additional mechanisms that redirect the TF targets under certain cellular regimes, or by variable transcription efficiencies under combinatorial TF control, but the point remains that TF networks are currently an incomplete mechanism for correlated gene expression. The broad solution to this uncertainty would be systematic investigation to characterize TFs by their network position, rates of activation and type of interaction. Any broad relationships between these variables would be useful in understanding how transcriptional programs are implemented, which would provide additional meaning behind observed coexpression relationships.

1.4.3 Do coexpression networks mirror protein-protein interaction networks?

What is the biological justification for coexpressing gene sets? While there are many explanations for how coexpression occurs mechanistically, and how it might have evolved, these do not necessarily provide a rationale for the prevalence of coexpression. Most commonly, researchers show that coexpressed genes have an overabundance of select GO terms, but that is as far as the characterization proceeds, with few exceptions (Oldham *et al*, 2008; Winden *et al*, 2009). One stoichiometric justification for the prevalence of genes regulated with equal proportions

(i.e. coexpressed genes) would be if those genes form protein complexes. If this is the case, then coexpression links should align with PPI (protein-protein interaction) links, since PPI links are determined by physical binding of proteins. Different organisms and methodologies presents conflicting evidence for the extent of link overlap between coexpression and PPI networks. The first examination of this relationship found a weak overlap between coexpression and PPI's (Ge *et al*, 2001), however, that was likely due to the essentially erroneous inclusion of dimerization (self-links) from the PPI, when no such phenomenon of self-regulation is possible to measure in microarrays (Xulvi-Brunet *et al*, 2009). Similarly, while Bhardwaj and Lu (2005) are often cited as supporting links between protein network structure and coexpression, careful examination of the actual evidence in that paper shows the associations are quite weak. But when using a selection of hub genes based on every MOE430 array in the GEO database, Winden (2009) found that coexpressed gene pairs were likely to be part of a protein complex, a finding supported by Oldham (2008). However, in an exhaustive comparison of yeast coexpression to PPI's, in this well annotated organism, there was little if any relationship between the global organization of the two networks (Xulvi-Brunet *et al*, 2009). These conflicting results indicate either variable conservation across species, small but highly significant overlap between coexpression and PPI's, crucial methodological differences in determining network overlap, or some combination of these factors. Systematically deciding when and if coexpressed genes take part in a protein complex could provide a more specific mechanistic link from coexpression to cellular function, rather than the often broad GO categories used most times to classify coexpressed gene function.

1.4.4 Biological mechanisms of coexpression – chromatin remodeling

Histones are a collection of proteins that bind to DNA and control transcription of between 2% and 10% of genes, by affecting transcription-factor access to DNA (Lee *et al*, 1993). Each nucleosome (multi-histone complex) can tightly bind a variable amount of DNA as a function of

small molecular groups attached to the molecular tail of the histone (Strahl and Allis, 2000). Since a histone or series of histones may expose or conceal several genes simultaneously, this may be a biological process that contributes to coexpression relationships. While there is little experimental literature describing a coexpression-histone relationship, there are no negative reports and two positive reports (Chen *et al*, 2005; Deng *et al*, 2010a). Since epigenetic mechanisms appear to be important in depression pathogenesis, we present an overview of potential coexpression mechanisms operating through histones.

Histones exert control over transcription of genes sets through a combination of pervasive transcription in tandem with HDACs and HATS (histone deacetylases and histone acetyl transferases) which operate like global gain controls on transcriptional efficiency. Since each nucleosome only binds 147 bp of DNA it appears unlikely that histones directly induce coexpression by binding and unbinding to several sequential genes, although no studies have directly addressed this. There are two broad trends in the literature for how histone-DNA interactions could lead to coexpression, which we term the passive and active mechanisms.

In the passive mechanism of histone-induced coexpression, histones appear to be clustered into groups on the chromosome which operate as road blocks to pervasive transcription (Chen *et al*, 2005). By consistently derailing polymerase at particular points during sequential gene transcription, histones may be responsible for defining coexpression membership of large groups of genes. Originally this role for histones was thought to explain coexpression of long swathes of unrelated genes (Boutanaev *et al*, 2002), but it appears that there are also common GO descriptions of genes demarcated by clusters of histones, so transcriptional proximity and histone boundaries may operate in tandem to promote transcription of genes sets with variable biological coherence (Batada *et al*, 2007). Generally it appears that regions of DNA tightly bound by histones (heterochromatin) are capped by insulators (also called boundary complexes/elements) that segregate DNA into sets of continuous domains, prone to coexpression (Chen *et al*, 2005; Deng *et al*, 2010b; Li *et al*, 2010). Also, such chromatin domains initially appeared to be rare and the genes

they contained varied completely between species (Valenzuela and Kamakaka, 2006). Now it appears that the majority of DNA is organized into domains and that these insulators may play an analogous role to promoters in selected sets of coexpressed genes (de Wit *et al*, 2008; Raab and Kamakaka, 2010).

The overall balance of HATs and HDACs is a critical component of homeostasis wherein excessive HAT activity is associated with cell death and HDAC activity is associated with cancer (Barlev *et al*, 2001; Carrozza *et al*, 2003; Minucci *et al*, 2006). Therefore, an active mechanism of histone-based coexpression would be through the balance of HDAC and HAT activity which could lead to synchronous expression of hundreds of genes. In general, sequential segments of chromatin tend to have similar acetylation status (Sproul *et al*, 2005). As opposed to looking at how histones may lead to coexpression by halting transcription, an inverse question would be: "are neighboring genes that are coexpressed bound by histones with identical acetylation status". If so, this would permit transcription of several neighboring (and ostensibly functionally related) genes. While adjacent histones are frequently acetylated or deacetylated together, the influence of identical histone acetylation status appears to only synchronize gene neighborhoods of approximately 4 genes in length (Deng *et al*, 2010a). Thus, while there are multiple ways that histone modifications could lead to coexpression, their contribution appears to be limited to small communities of genes.

Chromatin conformation and coexpression Insulators may also play a role in the most recently described mechanism behind coexpression: chromosome interactions. Through mechanisms that are currently unclear, during chromosome interactions, insulators or regulatory elements can interact with other regulatory elements or promoter sequences that may be physically distant on the same or different chromosome (Engel and Tanimoto, 2000). The mechanisms for bringing DNA strands together or linking them is unclear and debated, but these interactions occur frequently between specific sections of distant chromatin (Dekker, 2008). Chromatin interactions generate large DNA loops that encompass hundreds of genes. This higher-

order DNA structure is relevant to coexpression, because there is some evidence that these interaction sites occur at “transcriptional factories” wherein genes near the interaction site are highly transcribed (Sutherland and Bickmore, 2009). Because chromosomal interactions are transient, this could lead to consistent groups of genes with fluctuating expression levels: i.e. responsible for some portion of the coexpression signature. Since the discovery that these interactions are detectable and prevalent, the shift from using PCR (in the original 3C technology – “chromosome conformation capture”) to deep sequencing (in the current “Hi-C” technology) has improved the resolution at which chromosome interactions can be located (Lieberman-Aiden *et al*, 2009). Now it appears there may be a hierarchy of spatially adjacent regions on chromosomes in their natural 3D configuration. There is also evidence directly linking sites of chromatin interactions to coexpression of neighboring genes (Deng *et al*, 2010a). Either the transcriptional factories hypothesis or some other facilitation of expression upon chromosome interaction could be responsible for the observed high correlation of genes near chromosome interactions sites. The observation that 3D chromosome configuration may play a role in coexpression indicates that many previous studies, which focused on linear adjacency on chromosomes, likely missed a large portion of geometry-driven gene correlations.

1.4.5 Relationship of cellular coexpression mechanisms to measured expression correlations

Each microarray chip measures bulk mRNA levels derived from a variety of cell-types which are present in a given homogenized brain sample. Natural variability in the density of these various cell-type populations is another possible source of correlated genes expression. Under this paradigm, coexpression modules are generated by variability in numbers of specific cell-types across multiple brain samples. Because cell-types have similar levels of gene products, in microarrays across multiple subjects, these transcript sets will covary. In the same way, sub-

commmunities of cells under common physiological influence, with similar activity-driven transcription, could also generate correlated expression. To validate this bulk coexpression mechanism, it would be necessary to check if certain genes are coregulated within certain cell-types, and then also to check if these genes are correlated in homogenized microarray samples. This exact experiment has not been performed, but unique cell-type gene markers are known to exist, and some coexpression modules are enriched with markers for major cell-types (Oldham *et al*, 2008); therefore bulk coexpression (generated by cell-number variablity) appears to be a plausible source of detected gene correlations.

This source of coexpression is in some ways distinct from cellular mechanisms of coexpression, in that it represents fluctuations in the bulk quantity of mRNA from cell populations, rather than actively regulated processes within single cells. However, if this were the only source of correlated expression, there would be tight relationship between differential expression and coexpression. Namely, condition-specific modules would always be up- or down-regulated, when in fact they may not show differential expression at all (de la Fuente, 2010) - this indicates that internal cellular coregulation mechanisms make a contribution to coexpression. Furthermore, the mechanisms that maintain the transcriptional profiles of specific cell-types are many of the same mechanisms that generate cellular coexpression. Thus natural variability in the prevalence of certain cell-types could generate gene correlations, but the set of genes which covary by this mechanism may also stem from cellular mechanisms of coexpression.

1.4.6 Summary of transcription regulatory systems affecting coexpression

The regular structure of coexpression networks across multiple species is remarkable in light of the conglomeration of biological mechanisms that generate gene correlations. There is some seemingly incidental interaction between the correlation mechanisms (Lercher *et al*, 2003), but there is no overarching uniform control over coexpression mechanisms. While transcribing

genes in regulated fractions may be convenient from a cellular perspective, this is scarcely an explanation for how the assortment of regulatory mechanisms that form the physical basis for coexpression cooperate to generate appropriate levels of functionally related gene sets. The analysis of network structure generated by these mechanisms takes place in a parallel world of research that is oddly removed from the details of cellular mechanisms, (with rare exceptions, see Hudson (2009)) wherein consistent network geometry and carefully delineated network structure-functions are expected.

There have been no systematic studies that consider the contributions of TF's, histones, chromosome location, and epigenetic modifications to the final structure of coexpression networks. This is likely due to challenges of rallying molecular expertise to focus on a single organism and the tendency to work inside specialties that only cover a single mechanism. Since the various mechanisms responsible for coexpression do interact, and the global structure of coexpression networks recapitulates an orderly categorization of cellular functions, examining how multiple coexpression mechanisms are aligned to achieve a particular functional goal without counterproductive cross-talk could be helpful in knowing what types of dysfunction to look for in disease states.

1.5 OVERVIEW OF RELEVANT COEXPRESSION METHODS

This section provides definitions and rationale for statistical techniques used in building coexpression networks, with emphasis on the qualitative logic for quantative decisions. This is not a review of all coexpression techniques, only fundamental techniques directly applicable to Chapters 2-4. Complete detailed methods for individual studies may be found in Chapters 2-4, and this section is intended to be read in conjuction with them as a reference.

As detailed in Chapter 1.3, many biological processes lead to correlated expression levels among multiple genes. To build coexpression networks, these gene-gene correlations are commonly detected using Pearson correlations. The Pearson product-moment correlation (Pearson correlation, denoted ρ) maps the relationship of two variables (typically genes X and Y measured repeatedly across *n* subjects) onto [-1:1] as a function of the linearity of their relationship.

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

where the covariance of X and Y is defined as:

$$cov(X,Y) = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$$

and the standard deviation is defined as:

$$\sigma_X = \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

For a microarray dataset of *p*-genes by *n*-samples, the Pearson correlation between all *p* genes produces a symmetric *p* x *p* dimension correlation matrix *A,* whose entries a_{XY} equal the correlation of genes X and Y across all *n* samples. This matrix of correlations between all pairs of genes is referred to as the (raw) correlation matrix. After various transformations this becomes the adjacency matrix that describes the existence of links between pairs of genes. Note that we will refer to the entries of the correlation matrix *A* as a_{ij}, because while nomenclature for correlation is between variables X and Y, coexpression nomenclature refers to gene pairs *i* and *j*.

<u>Correlation properties relevant to coexpression networks</u>: A common worry in coexpression networks is that the Pearson correlation will be driven by spurious/random correlations between genes. While outlying values can drive correlations, the significance of Pearson correlation values increases with increasing sample size, as the likelihood of spurious correlations decreases (a process depicted in Figure 20A). The significance may be robustly calculated through a permutation procedure, or through the student's distribution, but the latter assumes the dataset is

large and normally distributed. Pearson correlation will not detect non-linear relationships between genes.

Using mutual information as a metric for gene-gene relationships could detect non-linear coupling, but even a recent algorithm that dramatically accelerates the calculation (Qiu *et al*, 2009), is still significantly slower than using Pearson correlation - a key shortcoming when huge adjacency matrices must be repeatedly calculated. Also, given the size of typical postmortem datasets, it is unlikely that non-linear relationships can be detected, and pragmatically, the gene sets with the highest mutual information and highest correlation scores are very similar (Steuer *et al*, 2002).

Link selection methods – transforming correlation matrices into adjacency matrices

Transforming microarray measurements of transcript expression level into networks of interactions is a critical task in gene network analysis, because techniques and parameter choices at this stage will influence the biological conclusions drawn from network structure. In the context of coexpression networks, the question of what constitutes a link is equivalent to the question of what correlation values represent true biological relationships. Since not all biological relationships are known, and indeed coexpression networks are often used to discover new relationships, calibrating correlation-based information to biological reality is challenging. It is possible to use known biological information, the network structure of inferred networks, or some combination of these to optimize the selection of which correlations are deemed coexpression links. Prior to selecting certain correlations as coexpression links, a larger question that should be addressed is why any threshold/weighting function should be applied to the basic correlation matrix. There are statistical, computational, and biological reasons to apply some filtration to the full correlation matrix.

Statistical rationale for thresholding: Due to random effects, there is some distribution of non-zero correlation coefficients for any dataset. This null distribution can be estimated by scrambling the rows of microarrays (in the standard microarray format rows are genes and columns are samples) and again calculating the correlation between all genes (examples of the

null distribution shown in Figure 20A). These null distributions can be used to estimate the likelihood that a correlation of given magnitude would be observed at random. As a practical guide, for postmortem samples sizes suitable for building gene networks (roughly n>=20), random correlations greater than 0.4 are almost never observed, although a specific function can be generated for any dataset that estimates the link false discovery rate. Thus there is strong evidence that the different levels of correlation may be filtered according to their corresponding levels of confidence.

Computational rationale for thresholding: Most inferred biological networks have a connectivity of ~1%. This means that of the (*n*^2-*n*)/2 possible unique connections between *n* nodes, only about 1% of them are "truly" utilized by biological systems. This limited connectivity is beneficial when computing statistical measures of connectivity, because the requisite computational time is a linear or non-linear function of the number of vertices and edges in the graph. Therefore, assuming 1% connectivity, it will be roughly 100 times faster to compute a statistic on a thresholded/binarized graph, vs full correlation matrix. If the time required is an exponential function of the number of edges, then it may be 10^6 or more times slower for typical algorithms to work on full matrices. When operating on networks with 20K-50K nodes, frequently encountered in microarrays, the dramatic increase in time required to compute classic graph statistics such as clustering coefficients and pathlength on full matrices is prohibitive.

Biological rationale for thresholding: Coexpression link strength varies with tissue/condition/platform, but based on large-scale analysis across tissues (Day *et al*, 2009; Prieto *et al*, 2008), there is some evidence for selective and robust correlation of certain genes, and thus a broad distinction between real and random/spurious correlation levels. But do smaller distinctions in correlation (for instance between 0.8 vs 0.6) matter to cellular function? While highly correlated genes are involved in common biological process, it is unclear how cells would utilize or create a full gradation of correlation values - the full correlation matrix. There is no evidence of active regulation of particular coexpression relationships down to a specific decimal place value. However

there is evidence for a continuum of biological reality tied to the continuum of correlation values, since using the scale-free criteria (see below) to weight the raw correlation matrix improved the clustering of biological functions (Zhang and Horvath, 2005).

Specific methods for thresholding/weighting gene correlations There are many possible functions that could be applied to correlation matrices to achieve the goal of filtering out irrelevant correlation values. Specific methods try to achieve this goal by relying on known biological information, the network structure of inferred networks, or some combination of these, to optimize the selection of which correlations are deemed coexpression links. These different approaches mean that there may not be a universal optimum method for filtering correlations. But for particular applications there may be an optimal method, which has useful features and irrelevant weaknesses. Several that have been used in associated literature are briefly reviewed here, and choice of methods in Chapters 2-4 will be discussed in light of these options.

Arbitrary correlation filters: The simplest way to filter out low correlations consists of setting all correlations with absolute value less than some threshold τ to zero and setting all correlations greater than τ equal to one – they become coexpression links:

$$A_{ij} = \begin{cases} 1 \text{ if } |a_{ij}| \geq \tau \\ 0 \text{ if } |a_{ij}| < \tau \end{cases}$$

For a dataset of a given size, an experienced researcher can use arbitrary cutoffs as a way to prototype network structure, but formally employing this method to generate networks is questionable. Using the highest 1% of all correlation values has been shown to select genes with related functions (Shi 2010, Lee 2004). Picking an arbitrary threshold in the range of what has been used by other studies may produce significant biological relationships between coexpressed genes, but that does not mean that the resulting network is the best representation of coexpression relationships. Even if a threshold is optimal in one study, there is no justification for using it as a universal threshold in other studies, as some coexpression networks may have different levels of connectivity (Reverter *et al*, 2006).

Network structure-based correlation filters: Biological networks show prevalence of non-random network structures (such as specific degree distributions and modularity). Therefore, assuming that one or more of these graph features are indicative of biological information, it is possible to select a threshold that maximizes these features in the network. Again, assuming the particular network feature is indicative of biological function, this threshold selection technique should maximize the biological information contained in a network. (This procedure is illustrated schematically in Figure 2.) The success of these methods therefore depends on the veracity with which the presence of some graph measure corresponds to biological interactions, and depends on also the technical implementation of the method. We provide an overview of three popular network structure-based correlations filters: *maximum clustering*, *maximum thresholding*, and *scale-free criterion* methods of threshold optimization.

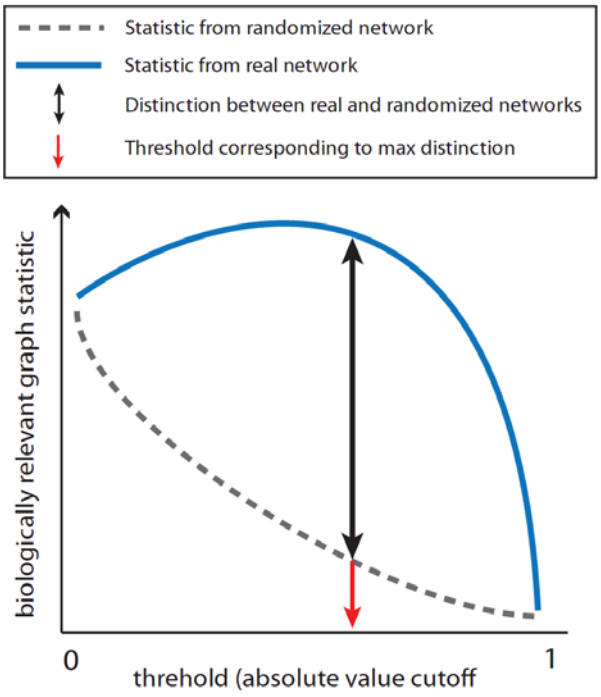


Figure 2 Typical behavior graph statistics calculated from real vs randomized networks over a range of possible threshold values.

Some techniques for correlation threshold optimization attempt to maximize a graph statistic (network feature) vs randomized networks, in order to predict the optimal threshold value.

Examples of relevant graph statistics could be clustering coefficient or synchronzability. See Figure 6A for an actual example of this strategy.

Maximum clustering thresholding: Biological networks are characterized by local communities - semi-isolated groups of nodes with dense interconnections, and these communities often correspond to functional units. The clustering coefficient (see formula below) is a classic method to detect evidence of local (functionally homogeneous) communities. Therefore the biological information in a network should be maximized by selecting the threshold with the maximum clustering coefficient (compared to degree-matched random networks).

Maximum modularity thresholding: The concept of modularity extends the concept of clustering to larger communities that may be separated by more than a single link. Since biological networks are characterized by modular structures of dedicated function, optimizing the threshold for maximum modularity will expose the community structure in a network. This point will occur somewhere between the extremes of a low threshold (in which all nodes are coupled) and a high threshold (wherein the network has too few nodes, or is overly fragmented) as depicted in Figure 20C. These goals are very general, and efficiently estimating what threshold will produce high modularity requires a graph measure that encapsulates the degree of modularity in the entire network. A measure that fits these criteria, “synchronizability” (see definition below) is a single graph statistic related to the number of nearly disconnected modules in a network and total network diameter. When it is calculated for a series of thresholds, it reaches a nadir at the point of maximum network modularity (Borate *et al*, 2009; Perkins and Langston, 2009).

Scale-free criterion: Since many biological networks are scale-free, this structure can be used to calibrate link selection. By applying cutoffs at a series of correlation values, and estimating how well the resulting networks fit the scale-free connectivity distribution, it is possible to find the threshold which creates the most truly scale-free network (where a scale-free network is defined by degree distribution $P(k) \approx ck^{\gamma}$, where k is the number of connections of a node (its degree) and $P(k)$ is the probability of a given number of connections). Two practical concerns with the standard

implementation of the scale-free criterion are that threshold is chosen based on the assumption that the scaling exponent (γ) of the network should be -1, when there are counter examples of scale-free molecular systems with scaling exponents other than -1. Also, the scale-free connectivity relationship is estimated with a log-log linear fit of the (assumed) scale-free connectivity distribution, when scale-free distributions with exponential cutoffs may be more appropriate. However scale-free networks with an exponential cutoff makes estimation of the scaling exponent more difficult, potentially lead to unrepresentative fits (Khanin and Wit, 2006; Zhang *et al*, 2005).

The scale-free criterion can be used to create weighted connectivity matrices that both decrease the contribution of low correlations and preserve the dynamic range of high correlations, by raising the raw correlation matrix to a power β>1, the such that the entries of the original correlation matrix *A* become $\left|A_{ij}\right|^{\beta}$. Just as a cutoff threshold may be selected by applying various cutoffs and evaluating the scale-free fit, so too can the exponent β be estimated by checking the scale-free fit for a sequence of β-values (Zhang *et al*, 2005). Preserving a range of correlation values can improve retrieval of modular communities compared to a binarized threshold.

Overview of other graph statistics employed:

The *clustering coefficient* is a measure of local community structure, that ranges from 0 (neighboring nodes unconnected) to 1 (all neighboring nodes connected to each other). For an unweighted graph (network) $G = (V, E)$ consisting of a set of vertices (nodes) V and edges (links) E, wherein a given vertex v_i has k_i neighbors with a total of e_{jk} edges between them, the clustering coefficient (CC) for node v_i is defined as:

$$CC_i = \frac{2\left|\{e_{jk}\}\right|}{k_i(k_i - 1)}$$

Betweeness centrality is a measure of the extended influence of a network node: it is proportional to how many shortest paths intersect a given node. If many shortest paths intersect a node, it is likely to be located in a bottleneck position, or else in the geometric center of the network. The formula for betweenness centrality (BC) for a given node v_i is

$$BC(v_i) = \sum_{j,k \neq i} \frac{\sigma_{jk}(v_i)}{\sigma_{jk}}$$

$\sigma_{jk}(v)$ is defined as the number of shortest paths from node v_j to node v_k that intersect node v_i, and σ_{jk} is the number of shortest paths from *j* to k. Thus, if a node lies on one of a small number of routes between other pairs of nodes, it will have a high betweenness centrality. Calculating betweenness centrality for each node in a network entails finding all shortest paths between all pairs of nodes, which can be done efficiently in unweighted graphs with Dijkstra's algorithm.

Synchronizability refers to the smallest positive eigenvalue of the Laplacian (Kirchoff) graph matrix. The associated eigenvector (the Fiedler vector) is often used for spectral clustering (though we do not use it here). The Laplacian graph matrix *L* is created from the binarized correlation matrix with the degree of each node listed on the main diagonal, and each connection between genes listed at -1 instead of +1:

$$L_{ij} = \begin{cases} \deg(v_i) & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and } A_{ij} = 1 \\ 0 & \text{otherwise} \end{cases}$$

Assortativity ranges from -1 to 1 and quantifies the likelihood that an edge will connect two nodes of similar degree (total connectivity level). If highly connected nodes are connected to other highly connected nodes, the network has positive assortativity, and if highly connected nodes are generally connected to provincial (low connectivity) nodes, the network has negative assortativity. There is a broad trend for technological networks to be structured such that they have positive assortativity, while biological networks generally have negative assortativity (Newman, 2003). The assortativity value *r* for a network with *M* edges, connecting nodes of degree (total connectivity) *j* and *k* is defined by Newman (2002) as:

$$r = \frac{M^{-1}\sum_{i=1}^{m} j_i k_i - \left[M^{-1}\sum_{i=1}^{m} \frac{1}{2}(j_i + k_i)\right]^2}{M^{-1}\sum_{i=1}^{m} \frac{1}{2}(j_i^2 + k_i^2) - \left[M^{-1}\sum_{i=1}^{m} \frac{1}{2}(j_i + k_i)\right]^2}$$

Examples of how to choose methods to select correlation thresholds: To illustrate the selection of appropriate coexpression methods, we briefly present some example rationale for choosing the particular coexpresion methods found Chapters 2 & 4. These decisions show how specific study goals dictate certain methods and how when possible we seek to show that results are robust, regardless of the specific method. Further details on study goals, methods, and efforts to ensure robustness are found in Chapters 2-4.

Consider a primary goal of Chapter 2: to understand if a set of genes chosen through extensive cross-species analysis form modular communities. Multiple methods and precautions were piled on top of each other to ensure that the inferred networks represented biologically driven interactions, and that the conclusions were robust against experimental noise. To establish that relationships between genes were similar in both mouse and human datasets, we were faced with the choice of using raw, weighted, or thresholded correlation matrices. Because the mouse sample size was small, fine gradations of correlation values were unlikely to be meaningful, and so we applied a threshold to generate mouse and human networks, whose structures we then showed to be similar. Working within the context of these small networks, we were less concerned about extracting information from all coexpression links, so we used the maximum clustering method to select the threshold. Finally, because we were searching for evidence of modules of glial- and neuronal-related genes based on the results of differential expression testing, to avoid any suggestion of bias, we show that these semi-distinct communities exist over a range of threshold values.

As a prerequisite to achieve the goals of Chapter 4 - locating differentially connected and differentially expressed genes in many coexpression networks - we sought a very general method that would produce biologically insightful networks for all datasets, but at the same time would not force them into a certain configuration. Because we did not want to make any *a priori* choices about the degree distribution of the resulting networks, and because the maximum clustering threshold method slows down considerably for large networks, and due to the results of a threshold

selection comparison (Borate *et al*, 2009), we used the maximum modularity method to select a threshold value for each dataset. We could have used weighted correlation matrices as the basis of all tests in the paper, but that would have critically slowed the computation of betweenness centrality and clustering coefficients, and since using a hard threshold still showed several strong effects, weighted adjacency matrices did not appear necessary in this case.

1.6 SUMMARY OF INTRODUCTION AND RESEARCH OVERVIEW

The biological basis of major depressive disorder is a shifting target for statistical analysis, as there are multiple hypotheses of depression (Chapter 1.2) which span several major fields of neuroscience research, including neurotransmitter systems, synaptic structure, endocrine function, and large-scale connectivity. Since it is likely that hundreds of genes contribute to the disorder, high-throughput analyses are an increasingly common attempt to find some coherent pathology. However, there is scant agreement between these studies, in part due to high false discovery rates associated with multiple sub-populations, a variable set of affected systems and a disease signal originating at different scales of investigation (Bosker *et al*, 2010; Mehta *et al*, 2010). We confront these roadblocks to understanding depression directly, by accepting that depression cannot be represented by a single model system, that it is generated by the action of many genes, and that it is accompanied by symptoms and effects on multiple physical scales. In the next three chapters, we use coexpression networks to conceptualize the activity of multiple genes in depression, and then apply network analysis in the context of modular cell processes (Chapter 2), brain region communication (Chapter 3) and global gene-based disease signals (Chapter 4). These applications of coexpression networks to neuropsychiatric datasets are a new avenue of understanding complex disorders, containing a literal framework in which to organize disease-related changes. The failure of other attempts to find a unified theory of depression indicates that

these coexpression-based analyses will not be a facile proof-of concept application, but will press the limits of statistical detection and biological interpretation. To show the concerted impact of our findings in expanding domains of influence, we discuss what the specific network-derived biological findings entail for the neurobiology of depression (Chapter 5), how multi-system interactions implicated in depression by our findings may trigger the “floodgate” model of depression, and new strategies for research design in complex disease research.

2.0 PAPER #1: A MOLECULAR SIGNATURE OF DEPRESSION IN THE AMYGDALA

Etienne Sibille, Ph.D.

Yingjie Wang, M.S.

Jennifer Joeyen-Waldorf, B.S.

Chris Gaiteri, B.S.

Alexandre Surget, Ph.D.

Sunghee Oh, B.S.

Catherine Belzung, Ph.D.

George C. Tseng,Ph.D.

David A. Lewis, M.D.

From the Department of Psychiatry and Biostatistics, and Center For Neuroscience, University of Pittsburgh, PA; and CNRS, INSERM and Université François Rabelais, Tours, France.

Address correspondence and reprint request to Etienne Sibille, Department of Psychiatry, University of Pittsburgh, 3811 O'Hara Street, BST W1643, Pittsburgh, PA 15213; sibilleel@upmc.edu (E-mail).

Disclosure: Dr. Lewis has received research support from the Bristol-Myers Squibb Foundation, Curidium Ltd, Merck, and Pfizer and has served as a consultant to Astra-Zeneca,

Bristol-Myers Squibb, Merck, Pfizer, Roche, Sepracor, and Wyeth. Dr. Belzung has served as a consultant to Takeda. All other authors report no competing interests.

Supported by the following grants from the National Institute of Mental Health (NIMH): MH067721 (ES), MH077159 (ES) and MH45156 (DAL), and by a fellowship from the University of Pittsburgh Institute for Clinical Research Education and the Clinical and Translational Science (CG). The funding agency had no role in the study design, data collection and analysis, decision to publish and preparation of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health. We thank Jean-Philippe Guilloux, Ph.D., for his qPCR expertise.

Preface

Note on the contribution of CG to this paper: he is the sole author of all statistical analysis, methods, figures and text related to coexpression and network analysis/structure. He also read, commented on and edited the remainder of the document.

Abstract

Objective: Major depressive disorder (MDD) is a heterogeneous illness with a mostly uncharacterized pathology. Recent gene array attempts to identify the molecular underpinnings of the illness in human postmortem subjects have not yielded a consensus. Thus, we hypothesized that controlling several sources of clinical and technical variability, and supporting our analysis with array results from a parallel study in the unpredictable chronic mild stress (UCMS) rodent model of depression would facilitate identification of the molecular pathology of MDD.

Methods: Large-scale gene expression was monitored in anterior cingulate cortex (ACC) and amygdala (AMY) in paired male familial MDD and control subjects (n=14-16 pairs). Area dissections and analytical approaches were optimized. MDD results were compared to UCMS results, and confirmed by quantitative PCR and Western blot. Gene coexpression network analysis was performed on transcripts with conserved MDD-UCMS effects.

Results: Significant and bi-directional predictions of altered gene expression were identified in AMY between MDD and the UCMS model of depression. These effects were detected at the group level, and also identified a subgroup of depressed subjects with a more homogeneous molecular pathology. This phylogenetically-conserved “molecular signature” of MDD was reversed by antidepressants in mice, identified two distinct oligodendrocyte and neuronal phenotypes, and participated in highly cohesive and interactive gene coexpression networks.

Conclusion. These studies demonstrate that the biological liability to MDD is reflected in a persistent molecular pathology that affects the AMY, and supports the hypothesis of maladaptive changes in this brain region as a putative primary pathology in MDD.

Introduction

Although dysregulated serotonergic and stress pathways are contributing factors in MDD, and clinically-effective antidepressants were discovered over 50 years ago, the biological mechanisms of MDD remain mostly uncharacterized (Belmaker *et al*, 2008). Recent attempts at identifying the molecular pathology of the illness, based on large-scale gene arrays, have implicated several biological systems, but have not yielded a consensus, potentially due to differences in cohorts, brain areas investigated, and analytical approaches (Aston *et al*, 2005; Choudary *et al*, 2005; Kang *et al*, 2007; Sequeira *et al*, 2006; Sibille *et al*, 2004), and to a substantial clinical heterogeneity. In fact, MDD may correspond to a family of disorders, which may be identified based on more narrow clinical and biological definitions (Hasler *et al*, 2004). Consequently, we sought to address these challenges in a gene expression profiling study of MDD designed to *(i)* reduce the clinical heterogeneity of the human cohort, *(ii)* focus on a relevant neural network, *(iii)* control for the variability in gene expression intrinsic to each brain region, *(iv)* maximize true discovery in gene array approaches, and *(v)* utilize a parallel study in a more tractable animal model of depression and antidepressant reversal to support the analysis of the human results. UCMS is an informative model to study depression in animals, as it mimics in a naturalistic way the role of psychosocial environmental stressors in precipitating a depressive pathology and the timeframe of antidepressant response (Surget *et al*, 2008b; Willner, 2005). The random application of several environmental and social mild stressors for several weeks results in a syndrome that is reminiscent of symptoms of depression, including physiological changes, increased anxiety-like/fearfulness and altered agonistic behavior (Surget *et al*, 2008b).

Our studies focus on the ACC and AMY, as critical components of a corticolimbic circuit of mood regulation (Pezawas *et al*, 2005) that is affected in MDD (Seminowicz *et al*, 2004). Evidence supporting dysfunctions of these areas in MDD include: decreased ACC volume and altered activity (Botteron *et al*, 2002; Drevets *et al*, 2002; Drevets *et al*, 1997; Mayberg *et al*, 1999), decreased glial density and reduced (Cotter *et al*, 2001) or no change in neuronal size (Ongür *et*

al, 1998) in ACC, decreased glial density (Bowley *et al*, 2002) and fewer oligodendrocytes (Hamidi *et al*, 2004) in AMY, and abnormal processing of emotional stimuli and sustained AMY reactivity (Roberson-Nay *et al*, 2006; Sheline *et al*, 2001; Siegle *et al*, 2002). Accordingly, we tested the hypothesis that the biological liability to MDD would be reflected in a persistent molecular pathology affecting the AMY and/or ACC.

Methods

Subjects

Brain samples were obtained during autopsies conducted at the Allegheny County Medical Examiner's Office after consent from next-of-kin. Consensus DSM-IV diagnoses were made by an independent committee of experienced clinical research scientists, utilizing information from clinical records, toxicology exam and a standardized psychological autopsy. We analyzed 16 pairs (14 in AMY), consisting of white male subjects with familial MDD and normal comparison subject matched for age, sex and race (Table 1). The increased disease severity was supported by a longer average duration of illness in the familial MDD cohort compared to non-familial MDD subjects from the same brain donation program (9±2 years versus 3±1 years; Mean±sem; p=0.01). A symptom score was calculated based on the presence at time of death (1=unequivocal yes; 0.5=unsure or subthreshold; 0=unequivocal no) of nine MDD symptoms: depressed mood, anhedonia, appetite disturbance, sleep disturbance, psychomotor change, anergia, self-recrimination, diminished ability to concentrate or make decisions, and suicidality. All procedures were approved by the University of Pittsburgh's Institutional Review Board and Committee for Oversight of Research Involving the Dead. Detailed information is available in supplements.

Brain samples

Rostral AMY samples enriched in lateral, basolateral and basomedian nuclei were delineated as described (Hamidi *et al*, 2004) and dissected from frozen coronal blocks ~2-3cm

caudal to the temporal pole. ACC samples containing all six cortical layers were harvested from coronal sections in subgenual ACC. A pilot study revealed rostro-caudal variations in transcript levels in both areas (Figure 1A-B). Selected white matter (WM) samples in ACC (n=7) and AMY (n=4) were obtained for analysis of cellular origin of transcripts (Sibille *et al*, 2008).

Microarray samples

Total RNA were extracted from frozen samples stored in TRIZOL (Invitrogen, Carlsbad, CA) and processed for microarray analysis according to manufacturer's protocol (Affymetrix Inc., Santa Clara, CA). In brief, 2.5µg of total RNA was reverse-transcribed and converted into double-stranded cDNA. A biotinylated complementary RNA (cRNA) was transcribed *in vitro*, using an RNA polymerase T7 promoter site introduced during the reverse-transcription step. 20µg of fragmented labeled cRNA sample were hybridized onto Human Genome U133Plus-2.0 arrays, assessing 54,675 probesets or gene transcript levels. To reduce the influence of technical variability, paired samples were processed together, but different pairs were randomly distributed at each experimental step. For samples, hybridization and arrays quality control, probeset signals (i.e., transcript levels) were extracted with the Affymetrix GCOS software. For statistical analysis, Log_2-transformed probeset signal intensities were extracted and normalized with the Robust Multi-array Average (GC-RMA) algorithm. Probesets with GC-RMA data values below 12 displayed systematic co-regulation patterns corresponding to the array normalization procedure and were considered background signal, leaving ~25,859 probesets with detectable signal in ACC and/or AMY.

Microarray quality control

Individual scans were visually inspected for the presence of manufacturing defaults and hybridization artifacts. Quality control parameters were as follows: noise (RawQ, ACC, 1.47±0.34; AMY, 1.47±0.34), background (ACC, 44±8; AMY, 43±5), scale factor (ACC, 2.62±1.32; AMY, 4.41±0.97), 3'/5' Actin ratio (ACC, 2.97±0.98; AMY, 2.98±0.93) and 3'/5' GAPDH ratio (ACC,

1.23±0.217; AMY, 1.45±0.28). Consistent number of genes were detected across arrays (ACC, 46.8±2.2%; AMY, 48.2±3.3%), and consistent detection of BioB/C hybridization spiked controls. To further identify outlier samples, we used the “outlier detection” feature of the DNA-Chip Analyzer software, and analyzed correlation coefficients between array signals across all samples. Based on these results 16 pairs were retained for analysis in ACC and 14 pairs in AMY.

Rodent UCMS model

We have previously described a behavioral and microarray study in the UCMS model of depression in mice (Surget *et al*, 2008b). Behavioral and array results from that study were used here to support the analysis of data in the homologous brain areas in human subjects. In brief, BALB/c mice were subjected to various stressors according to a pseudo-random schedule for seven weeks. Drug [fluoxetine, 20mg/kg/day, or a corticotropin-releasing-factor 1 (Crf1r) antagonist (SSR125543), 20mg/kg/day] or vehicle treatments started on day 14 and continued until the end of UCMS, when microarray data (MOE430-2.0) were generated in cingulate cortex (CC) and AMY (lateral/basolateral nuclei). Stressors included: altered bedding (sawdust change, removal or damp; substitution with water, rat or cat feces); cage tilting or shaking; cage exchange (mice exposed to the empty cage of another male); induced defensive posture (repeated slight grips on the back) and altered light/dark cycle. Body weight and coat state were assessed weekly, as markers of the progression of the UCMS-evoked syndrome. The coat score combined results from different body parts (O=well-groomed, 1=unkempt). This index has been pharmacologically validated (Santarelli *et al*, 2003). Emotion-related and agonistic behaviors were measured at the end of UCMS using the novelty suppressed feeding (NSF) and the resident/intruder (RI) tests. The NSF test consists of providing food-deprived mice with a food pellet in a novel, aversive environment (a brightly lit enclosure). The latency to start feeding correlates with fearfulness and decreases after acute treatment with anxiolytic drugs or chronic antidepressant exposure, suggesting that mechanisms underlying changes in the latency to feed involve anxiety-like and

antidepressant-like processes. For the RI test, control and UCMS-treated mice were single-housed in new cages two days before testing. The opponent, a six-month-old BALB/c intruder, was placed into the home cage of the test animal (resident). Latencies and number of attacks were recorded for 10min. Detailed methods and results are described in (Surget *et al*, 2008b) and summarized in Table 2.

WM/GM analysis

Ratios of transcript levels between WM and gray matter (GM) samples were generated as described, and used as estimates of relative gene transcript enrichment in glia (WM/GM>1.5), neurons (WM/GM<-1.5) or both cellular population (-1.5<WM/GM<1.5). Details in supplements.

Array data statistical analysis

- *Selection of significant genes*. To maximize discovery, we opted for an analytical approach with initial low stringency, followed by a comparative analysis with UCMS array data. First we flagged any gene potentially affected in correlation with MDD, and then focused on cross-species identification of similar changes for orthologous probesets. The assumption was that MDD-related changes would manifest as weak effects, but that the conservation of such changes across species would provide independent lines of validation and thus facilitate their identification from background variability or unrelated changes. We have previously validated this approach in a study of gene expression correlates of aging between the human and mouse brain.

- *Statistical criteria*. Changes in gene expression can take different non-exclusive forms in large datasets and are difficult to assess with any single test. Thus, for our inclusive first step, genes were tested by parametric paired t-test, non-parametric paired Wilcoxon signed rank test and by analysis of variance (ANOVA), and taking into consideration several clinical covariate parameters. In ANOVA, the observed variance was partitioned into components explained by different explanatory variables (covariates). We applied the following ANOVA model for each gene,

$Y_{gi} = \alpha_g + \beta_g * S_i + \tau_g * R_i + \gamma_g * A_i + \delta_g * AD_i + \varepsilon_{gi}$, where $1 \leq g \leq G$ =26199 genes, $1 \leq i \leq I$ pairs, and with clinical covariate values of 1 if the condition is true and 0 otherwise (S_i, suicide; R_i, recurrence; A_i, co-morbid alcohol dependence; AD_i antidepressant). Since some of the covariates may not be statistically significant for each gene g, we applied a stepwise model selection by Akaike Information Criterion (Akaike, 1974) to find the final "best fit" model, and a corresponding adjusted p-value was calculated for each gene. Genes were selected if any of the three statistical criteria were less than 0.05 and if group differences in transcript levels (paired and unpaired) were greater than 20% using mean or median averaged values, corresponding to a difference at the low limit of qPCR confirmation. Genes were selected according to similar criteria in the UCMS dataset [ANOVA, $p<0.05$; changes >20%;]. The significance of the cross-species concordance was assessed by bootstrap resampling (using the same analytical procedures), where sample groups were repeatedly shuffled in equal proportion (control and experimental) thus removing the MDD or UCMS component from the analysis. The procedure was repeated 1,000 times to generate a null distribution so that p-values of the observed directional correlations could be assessed.

- *Directional correlations,* $r(D_1, D_2)$ were calculated in a pair of cross-species or cross-brain region datasets D1 and D2. Ortholog probesets were identified using the Netaffx webtool (Affymetrix Inc.) to link the human and mouse datasets. In contrast to traditional Pearson correlation, the directional correlation measures the fitness of significant genes in D1 to predict D2, and conversely of D2 to predict D1. For instance, the statistically significant genes in D1 were first selected and the Pearson correlation of the log-ratios of D1 and D2 in this restricted D1-significant gene set was calculated. Thus, directional correlations are not symmetric [$r(D1, D2) \neq r(D2, D1)$].

- *Group variability* in gene expression was assessed by two-group t-test using individual gene transcript variances as continuous variables.

Real-time quantitative PCR (qPCR)

qPCR was performed as previously described. Results were calculated as the geometric mean of relative intensities compared to three internal controls (actin, glyceraldehyde-3-phosphate dehydrogenase and cyclophilin). See details in supplements.

Western blot analysis

Proteins from paired samples were processed in quadruplicate using rabbit anti-actin (Sigma #A2066), and mouse anti-cyclic nucleotide 3' phosphodiesterase (CNP; SMI-91R, Covance) primary antibodies and IRDye® 800 anti-rabbit and 680 anti-mouse (LI-COR Biosciences) secondary antibodies. Signals were simultaneously detected using the LI-COR Odyssey® Infrared imaging system. CNP protein content was expressed relative to actin. See details in supplements.

Gene coexpression networks

Gene coexpression networks were built through Pearson correlation of expression patterns and optimized using clustering coefficient analysis and jackknife correlation. See details in Figure 4 and supplements.

Results

AMY-ACC altered gene expression in human MDD

Large-scale gene expression profiles were generated from sub-dissected AMY and ACC (Figure 1A-B) in postmortem brains of male familial MDD subjects and matched controls (Table 1). Overall correlations of gene transcript levels were significantly higher in matched pairs, compared to non-matched MDD-control pairs (Figure 1C), thus validating the pairing protocol at controlling non-disease related factors and reducing signal variability. 395 genes in ACC and 191 genes in AMY were identified by paired statistics or ANOVA models as differentially expressed in MDD (Figure S1 and 1D). A qPCR survey on adjacent tissue sections yielded highly concordant results

(Array-qPCR Pearson correlation r=0.88, $p<e^{-5}$, n=16 genes; Figure S1 and Table 3 for selected genes), confirming the sampling and technical reliability of the array data. Nevertheless, as results are expected to contain false positives, we hypothesized that identifying relevant findings within this large pool of genes would benefit from a comparison with equivalent data obtained in an animal model that recapitulates behavioral and pharmacological aspects of depression.

Profiles of altered gene transcripts are conserved between human MDD and the UCMS mouse model of depression, and are reversed by antidepressant treatments in mice

We previously reported that UCMS induces a depressive-like syndrome in mice, consisting of progressive deterioration in coat state, reduced weight gain, and increased agonistic and emotion-related behaviors, and have shown that both symptom dimensions were reversed by chronic administration of an effective (fluoxetine) or putative (Crf1r antagonist) antidepressant (Surget *et al*, 2008b). UCMS also induced region-specific patterns of altered gene expression in cingulate cortex (CC) and in the lateral/basolateral nuclei of the AMY, which were reversed by both drug treatments. These behavioral and molecular results are summarized in Table 2 and the array results from that study were used here to support the analysis of the human data. Specifically, we hypothesized that if cellular mechanisms underlying mood regulation were conserved across species, then altered transcriptome in human MDD would predict similar changes in mice after UCMS, and that the effects of antidepressants in mice would help separate the effects of MDD from those of drug exposure in humans.

We investigated the degree of conservation of altered transcript levels for ortholog genes between MDD and UCMS by analysis of directional correlations. Confirming our hypothesis, highly significant, reciprocal and consistent predictions of molecular changes were identified in AMY (Figures 1D-2A). Specifically, of 191 genes with altered transcript levels in MDD, expression levels in mice were available for 105 ortholog probesets. Changes for these 105 mouse probesets were significantly correlated with human MDD-related changes (r=0.29, p<0.005). Conversely, of 299

genes with altered transcripts in AMY of UCMS-mice, the 213 identified human orthologs revealed a reciprocal mouse-to-human significant correlation of similar amplitude ($r=0.29$, $p<0.00001$). Analysis of 1000 bootstrap resamplings demonstrated that the probability of obtaining the observed reciprocal concordance levels by chance was very low ($p<0.001$ for individual directional correlations; $p<0.001$ for concurrent positive findings in both directions). Markedly, the human-to-mouse correlations disappeared, after successful antidepressant treatments in UCMS-exposed mice (Figure 2A). Thus, the pharmacological reversal of the MDD-UCMS correlation by two different antidepressants (i.e., targeting serotonergic or neuroendocrine stress pathways) demonstrated that the molecular changes supporting the MDD-UCMS correlations in AMY were specific to the altered mood phenotype.

Toxicological screens identified the presence of antidepressants in 5 human subjects (4 subjects in AMY cohort), although these subjects were depressed at time of death, suggesting a lack of efficacy, suboptimal treatment, or treatment-resistance. Similar correlations were observed between that patient subgroup and UCMS ($r\sim0.35$), thus supporting the clinical evidence of a lack of antidepressant efficacy in these subjects, at least for genes underlying the UCMS-to-MDD correlation.

Conversely, UCMS-induced changes in mouse CC did not predict corresponding changes in human ACC ($r=0.10$), while human MDD-related changes were also unrelated to changes in mouse CC ($r=0.02$) (Figure 2A). These low and non-significant ACC-CC correlations could result from differential involvement of that brain area in MDD and UCMS, or reflect a low conservation of cingulate structure and function across species. To partly address this question, we took advantage of the robust differences in transcriptomes between ACC and AMY in human subjects (~20% of genes; >2-fold change, $p<0.01$), and between mouse CC and AMY (~10% of genes; >2-fold change, $p<0.01$), to estimate the degree of similarity in “molecular structure” between areas across species. We found highly significant and reciprocal correlations between human ACC/AMY and mouse CC/AMY differences (Mouse-to-human, $r=0.63$, $p<0.0001$; Human-to-mouse, $r=0.55$,

p<0.0001). These values did not differ between MDD subjects, UCMS-treated mice or control samples (r=~0.60, all comparisons, not shown). These results suggested that the lack of conserved depression-related findings in ACC was not due to overall differences in "molecular structures" of the AMY/ACC network across species, thus also highlighting the AMY specificity of the human-rodent correlation of the molecular impacts of MDD and UCMS.

AMY cross-species correlations of depression-related molecular changes identified a subgroup of human MDD subjects

Absent or weak mouse-human correlations in cingulate cortices could also arise from variable or opposite effects in subgroups of human subjects, resulting in a null group-effect. Indeed, despite our efforts to reduce the heterogeneity of the human cohort, MDD is by its clinical definition a heterogeneous disorder, and one may reasonably expect differences in molecular pathologies across subjects. Moreover, since the current analyses rely on large numbers of genes (178 in ACC and 213 in AMY), different gene sets may weigh differently across subjects; thus, correlation analyses in individual subjects may reveal features of cross-species predictions otherwise not available using combined group-values. Here, using subject-wise changes in transcript levels for the identified genes (Figure 1D, step 3), we confirmed the lack of conserved MDD/UCMS effect in ACC, as most human individual subjects displayed no cross-species correlation (Figures 1D and 2B, left panel). In AMY, however, directional correlations revealed a large heterogeneity in cross-species predictions, with half of the subjects displaying positive correlations, and the rest displaying either absent or negative correlations (Figure 2B, right panel). This difference from the ACC distribution was not explained by baseline changes, as the variability in gene expression of controls was comparable between ACC and AMY (p>0.2). Rather it was due to a selective increase in gene expression variability in MDD subjects in AMY (AMY: 50.3% higher gene transcript variance versus controls, p<0.01; ACC: 1% increase, p=0.97).

Notably, the subgroup of MDD subjects with positive UCMS correlation in AMY (denoted

here MDD^{UCMS}) did not differ in demographic parameters from controls or other MDD subjects (all p>0.05), and was not explained by differences in death by suicide, disease recurrence, antidepressant exposure or alcohol dependence relative to other MDD subjects (Figure 2B, bottom grid). The four MDD subjects with evidence of antidepressant exposure at time of death were all included in the MDD^{UCMS} subgroup, confirming the presence of a depression-related molecular profile in these subjects (i.e., positive correlation with UCMS), consistent with their clinical profile. Overall, MDD^{UCMS} subjects displayed a trend towards more depressive symptoms (7.4 versus 4.6 in the remaining MDD subjects, p=0.07). Interestingly, the two subjects with large negative correlations were among the only three MDD subjects who met requirements for remission or partial remission due to fewer depressive symptoms at time of death (Figure 2B, crossed circles). Together, these findings suggest that the degree of correlation between UCMS and MDD molecular changes in AMY may predict the severity of depression in human subjects. Indeed, a positive and significant correlation was observed between symptom numbers and UCMS/MDD correspondence (r=0.62; p=0.02; n=14 pairs), although this effect was partly driven by two remitted subjects (Figure S2). Finally, restricting the analysis to MDD^{UCMS} subjects (Figure 1D, step 4), we identified a larger number of genes with altered transcript levels in AMY (n~2100; 1139 orthologs), suggesting a greater homogeneity in molecular profiles within this subgroup. In the absence of demographic identifiers, we interpreted these findings as evidence for a subgroup of MDD subjects (MDD^{UCMS}) with a consistent AMY pathology, potentially reflecting a more severe form of the illness, and for which the UCMS rodent model provided significant predictability at the gene expression level.

Two distinct oligodendrocyte and neuronal depression-related phenotypes in AMY

To characterize putative biological events underlying the cross-species correlations of changes and to address the presence of false positives in single datasets, we focused on genes with confirmed changes across species. Selected genes had to be significantly affected by UCMS

and MDD, and reversed by antidepressant treatments in the mouse model (Figure 1D, step 4), thus tracking the altered mood phenotype and controlling for drug effects. Of 299 gene transcripts affected by UCMS in AMY, 61 were also significantly affected in MDD, mostly corresponding to changes in MDD^{UCMS} subjects. Several of these transcript changes corresponded to the same genes and were combined, reducing the selection to 44 genes. 38 of the 44 genes displayed changes in the same direction in both species. Finally, antidepressant treatments reversed changes for 32 of these genes in rodents (Table 3), together identifying a core set of genes, characterized by concordant MDD and UCMS effects and effective reversal by antidepressant treatments.

qPCR analysis on RNA extracted from adjacent tissue sections for 17 of these genes revealed a very high correlation with array results in the MDD^{UCMS} group ($r=0.95$, $p<0.00001$; Table 3), even if individual statistical significances for some genes were only at the trend level ($p=0.1$). To determine whether this represented a quantitative limitation of the qPCR assay or a lack of biological effect, we assessed changes in protein levels for *CNP*, one of the three genes with trend-level significance by qPCR. Quantitative western blot analysis revealed stable CNP protein levels over the postmortem interval covered in our study (PMI/protein, $r=0.01$; not shown), a high concordance with RNA levels ($r=0.76$ for all 14 pairs, $p=0.002$), and a significant downregulation in MDD^{UCMS} subjects (-21.5%, $p=0.01$; Figure 3). In concert with qPCR, these findings provided supporting evidence for the technical reliability and biological validity of the identified molecular profile described in Table 3.

Within the group of genes with suggested glial-enrichment of transcripts, genes were almost exclusively related to oligodendrocyte structure and function and were all downregulated (Table 3, bottom rows). This striking convergence of gene function and direction of biological effects strongly suggests the presence of a conserved phenotype negatively affecting oligodendrocytes in AMY under MDD and UCMS conditions. Conversely, genes with suggested neuronal enrichment of transcripts were mostly upregulated, and related to cellular maturation and

synaptic development, neurotransmission and signaling, and cell-cell and cell-matrix interactions (Table 3, upper rows), suggesting a putative increase in neuronal structure and function in AMY of MDD subjects.

Genes with conserved MDD- and UCMS-related changes participate in a highly cohesive and interactive gene coexpression network

We next investigated whether the identified genes represented various unrelated molecular findings or if they participate in shared cellular and biological functions (known as functional modules). It is possible to test these hypotheses by simultaneously inferring the interactions, or "links", between our identified genes. These links are based on synchronized fluctuations in gene expression across samples (i.e. "coexpression" link), which have been shown to correspond to shared biological functions (Lee *et al*, 2004). Indeed, gene networks built on coexpression links typically cluster in functional modules that correspond to specific cellular activities (Alexander *et al*, 2009; Lee *et al*, 2004; Zhang *et al*, 2005) and this organization persists across species (Bergmann *et al*, 2004). Hence, biological networks built on coexpression links are useful means to determine if genes share common functions, and represent here a bias-free and data-driven way to investigate putative unifying MDD-related cellular processes shared by our identified genes.

Accordingly, we used Pearson correlations to determine pair-wise coexpression links between the 32 identified genes, which were then used to build gene networks (See Supplements). To ensure that the coexpression links represented robust markers of biological gene interactions, we used clustering coefficient analysis and jackknife resampling methods to optimize our criteria for inclusion in the networks. Clustering coefficients estimate the density of local connections within functional modules and represent measures of network structure with wide applicability in brain networks (Sporns *et al*, 2004). Here, local modules were more connected than randomized networks (i.e. higher clustering coefficient; Figure 4A), indicating that the identified genes participate in shared biological functions. Pearson correlation values resulting in networks with the

largest differences in clustering coefficients compared to permutated networks, provide the most biological information and were retained here as optimized cutoff points to build gene networks (Dashed line in Figure 4A) (. Additionally, we used a jackknife resampling approach to remove spurious links and maximize the biological reliability of the network. The obtained bimodal distributions clearly segregated links as robust (i.e., survive jackknife resampling; Figure 4B, right columns) or spurious (i.e., do not survive jackknife; Left columns) in both species. Thus, gene networks were built using 100 links from the most robust groups in the jackknife histograms (Figure 4C-D), corresponding to clustering coefficients within the suggested range of optimized values (Figure 4A; >0.65 in human; >0.75 in mouse).

We report that the 32 identified genes formed a tightly clustered network (Figure 4C-D) with ~7 times more connections than random networks of similar sizes ($p<0.01$). Moreover, the overall clustering coefficients for each network were on average 77% higher than degree-matched randomly-selected reference networks ($p<0.001$). Results were highly similar for all conditions and in both species, thus strongly supporting the biological validity and reliability of the identified network. Although the organic representation of the networks showed some differences (Figure 4C-D), the internal topology was well conserved, with a ~40% concordance of individual links across species, or ~57% using correlation of “betweenness centrality”, a more general measure of network similarity (Girvan and Newman, 2002). Within this network, genes with suggested glial- or neuronal-enrichment of cellular origin of transcripts naturally segregated (Figure 4C-D), which was quantitatively reflected by higher intra- (glial-glial and neuronal-neuronal) than inter-connections (glial-neuronal) (Figure 4E). In summary, these results demonstrate that genes forming the identified molecular signature of depression belong to an existing and tightly connected gene network that is conserved across species and that reflects the interactive glial/neuronal cellular compartments of gray matter tissue, together suggesting an abnormal recruitment by the illness of existing cellular pathways.

Discussion

To identify the molecular pathology of MDD, a heterogeneous illness with multiple symptom dimensions, we focused on postmortem samples in a cohort with reduced clinical and demographic heterogeneity (familial MDD, male subjects; Table 1) and on two brain regions (AMY and ACC) within a neural network putatively involved in altered mood regulation, a core symptom of the illness. After controlling several sources of technical variability (Figure 1) and by relying on cross-species validation and antidepressant reversal of results, we now report the identification of reciprocal predictions of altered transcriptome between MDD and the UCMS rodent model of the disease (Figure 2A). These effects were detected at the group level, and also identified a subgroup of MDD subjects (MDD^{UCMS}; Figure 2B) with a more homogeneous molecular pathology and for which UCMS provided a means to identify individual genes with conserved changes. Specifically, changes in transcript levels of ~30 genes were similarly detected in human MDD and mouse UCMS, and were reversed by antidepressant treatments in mice, hence corresponding to a pool of genes affected in correlation with mood states. These genes belonged to an existing cohesive network (Figure 4) and suggested two distinct cellular phenotypes: decreased oligodendrocyte and upregulated neuronal structure and function (Table 3). Finally, the interconnections between the glial and neuronal components of the network suggested either a conserved cross-talk between the two phenotypes, or a common upstream mechanism. Together, the present studies confirm that the biological liability to MDD is reflected in a persistent molecular pathology affecting the AMY and support the hypothesis of maladaptive changes in AMY as a putative primary pathology in MDD.

A phylogenetically conserved molecular subtype of depression in AMY

To address past difficulties, we assumed that evidence for mechanisms of mood dysregulation would manifest as weak, but conserved signals between MDD and UCMS. Thus, critical analytical aspects were to first maintain a high discovery level and then rely on cross-

species validations of results, as previously described in a mouse-human comparative transcriptome study of brain aging (Sibille *et al*, 2007). The conserved changes observed here were confirmed by permutation analyses and of similar magnitude as age-related changes (~0.3-0.4 Pearson directional correlations). Thus, compared to the robust effects of aging, the molecular correlates of UCMS and MDD were surprisingly strong.

It is not known whether these conserved changes correspond to specific symptom dimensions (i.e., stress effects, or AMY-mediated anxiety component) or to a core pathology in MDD^{UCMS} patients, as available parameters did not correlate with the subgroup (Figure 2). Overall, MDD^{UCMS} subjects displayed more depressive symptoms at time of death. The fact that MDD subjects who met requirements for remission or partial remission displayed negative, or no correlation, also suggested that opposite changes may participate in both depressive and remission phenotypes. We propose that this molecular signature may represent a novel approach to categorize MDD, based on similarities of biological changes. This molecular view of disease heterogeneity is consistent with the notion that other MDD subjects may present altered functions in different brain regions (exerting control over AMY, for instance), yielding similar clinical phenotypes, but through pathogenic mechanisms remote from the AMY.

In contrast, no UCMS-MDD correlations in gene expression were identified in cingulate cortices. The complex evolutionary changes and potential differences in connectivity of this brain area may explain the lack of conserved effects, despite similarities in molecular structures. The ACC integrates input from cognitive and emotion-related sources, influences activities relating to decision making, and modulates neuroendocrine, motor and visceral responses (Paus, 2001). Phylogenetic specializations of the human ACC include increased size, more functional subdivisions, dense prefrontal cortex connections and cellular specializations (i.e. spindle cells) that allow distal connections with other brain regions (Allman *et al*, 2002; Allman *et al*, 2001), together reflecting the human capacity for higher integration of complex emotion and cognitive functions, compared to rodents (Paus, 2001). Alternatively, some ACC findings may relate to the

illness, but the UCMS model did not make their distinction possible from unrelated effects, as the model mimics only one of many putative pathogenic mechanisms in the illness. Hence, molecular data obtained in different models, based on cognitive- or reward-related symptoms for instance, combined with investigations of additional brain regions, could provide additional insights on molecular dysfunctions in MDD, as different symptoms dimensions likely correspond to dysfunctions in distinct neural networks (Belmaker *et al*, 2008; Nestler *et al*, 2002). Investigating selected brain areas in human subjects, in parallel to relevant animal models of symptom dimensions, may thus represent a fruitful approach to address the heterogeneity of the molecular pathology of MDD. Here, in the absence of quantitative differences in clinical features or demographic identifiers, the present findings identified a subgroup of MDD subjects (MDD^{UCMS}) with a homogeneous molecular pathology and for which UCMS provided a significant predictability at the gene expression level.

Oligodendrocyte and principal pyramidal neuronal changes in AMY in depression

The striking convergence of downregulated glial-related gene transcripts in MDD and UCMS clearly suggests a conserved phenotype selectively affecting oligodendrocytes (Table 3), consistent with reports of reduced oligodendrocyte number in AMY (Hamidi *et al*, 2004) and decreased oligodendrocyte-related gene expression in temporal cortex (Aston *et al*, 2005), thus establishing AMY oligodendrocyte alterations as a confirmed pathological finding in MDD. These changes appear more robust in AMY, since there were not observed in ACC/CC or frontal cortex (Choudary *et al*, 2005; Kang *et al*, 2007; Sequeira *et al*, 2006; Sibille *et al*, 2004), although see (Uranova *et al*, 2004).

The coexpression analyses revealed that the identified genes participate in a naturally-occurring tightly-linked functional network that includes glial and neuronal components (Figure 4), suggesting an abnormal recruitment by the illness of existing cellular pathways, although the identity and origin of the neuronal component is not known. What mechanisms might link these two

phenotypes, and be informative as to their cellular characterization? Results from the network analyses suggest that the two phenotypes may occur in concert, either through mutual interactions, or downstream from a common perturbation. We speculate that the converging effects of increased AMY recruitment (e.g. through excitatory glutamatergic neurotransmission originating from stress-related sensorimotor modalities), and elevated stress-induced glucocorticoid exposure, as occurring in UCMS and suggested in MDD, may provide a common synergistic mechanism.

Both effects would be consistent with the observed glial phenotype, in view of the known vulnerability of oligodendrocytes to glutamatergic excitotoxicity (Rosin *et al*, 2004) and inhibitory effect of glucocorticoids on oligodendrocyte proliferation (Banasr and Duman, 2007; Wennström *et al*, 2006). Accordingly, decreased oligodendrocyte gene transcripts may represent early evidence of combined excitotoxic insults and glucocorticoid inhibition, eventually leading over time to decreased oligodendrocyte numbers in MDD (Hamidi *et al*, 2004).

On the other hand, this putative synergistic mechanism would suggest AMY principal pyramidal cell as the likely source of the neuronal molecular pathology. Indeed, the convergence of increased ARHGAP6 (a RhoA inhibitor), CACNB2 (voltage-dependent Ca channel) and modulators of glutamatergic synaptic plasticity (CAMK2D, EGR1), coupled with increased components of cell-matrix remodeling (MATN2, CDH13 and CHSY1) suggest increased structural and functional dendritic/synaptic compartments. This interpretation is consistent with the reported increased dendritic branching in pyramidal neurons in the rat AMY after chronic mild stress (Vyas *et al*, 2002) and with the increased excitability of AMY basolateral pyramidal neurons after glucocorticoid exposure (Duvarci and Paré, 2007). Together with the absence of changes of interneuron markers, the known stimulatory effect of glucocorticoids on AMY principal cells (Duvarci *et al*, 2007)d and an expected activity-driven dendritic structural upregulation, the present results suggests an increased structure/function phenotype of AMY principal pyramidal cells in UCMS and MDD. In humans, these proposed neuronal changes may correspond to cellular and molecular correlates of increased AMY function in MDD (Drevets *et al*, 1997; Sheline *et al*, 2001; Siegle *et al*, 2002),

although the causes may be complex and combine (mal)adaptive mechanisms and intrinsic genetic vulnerability .(Hariri *et al*, 2002; Pezawas *et al*, 2005)

These results await replication in independent cohorts. The scarcity of well-characterized human postmortem brains of male familial MDD subjects prevented the direct replication in our brain bank, and it is not known whether different demographic and clinical parameters (i.e., female subjects, non-familial) will affect the nature and/or robustness of the molecular findings. Another limitation concerns the presence of antidepressant treatments. Here, antidepressant-treated subjects still met DSM-IV criteria for MDD, thus suggesting lack of efficacy, consistent with molecular profiles suggesting depressive-like states. Finally, numerous additional genes were identified than included in this report, but the putative association of these genes with MDD could not be confirmed in the absence of independent cohorts or of animal models for alternate pathogenic mechanisms in the illness.

TABLE AND FIGURE LEGENDS

	Major depression (MDD) group										Control group					
Pair	**Case**	**Age**	**PMI**	**pH**	**Storage**	**RIN**	**Suicide**	**Rec.**	**AD**	**Alc.**	**Case**	**Age**	**PMI**	**pH**	**Storage**	**RIN**
1	505	57	12.8	7.1	125	8.7	Y	Y	N	Y	685	57	14.0	6.6	98	8.0
2	513	25	13.1	6.9	123	7.0	Y	Y	N	N	789	23	20.0	7.0	80	7.4
3	598	69	5.9	7.3	111	8.8	Y	N	Y	N	795	69	12.0	6.6	102	8.2
4	600	64	9.9	6.7	110	7.1	Y	N	N	N	615	62	7.2	6.4	108	7.8
5*	613	59	15.6	7.0	108	8.7	Y	Y	N	Y	551	62	16.4	6.6	118	8.3
6*	619	55	18.8	6.9	107	7.9	Y	N	Y	N	634	53	16.0	7.0	105	8.1
7	698	59	13.0	6.8	96	9.0	Y	N	Y	N	713	58	37.5	7.0	94	8.4
8	783	63	11.5	6.5	82	8.8	N	Y	N	N	736	55	15.5	6.9	88	8.3
9	863	52	28.3	7.2	65	8.4	N	N	N	N	1086	51	24.2	6.8	27	8.1
10	868	47	10.5	6.8	64	8.1	N	Y	N	Y	857	48	16.6	6.7	66	8.9
11	926	57	19.0	7.0	20	7.3	N	N	N	Y	1122	56	15.4	6.7	49	7.9
12	1001	54	7.3	6.6	67	7.6	N	N	N	N	852	54	8.0	6.8	36	9.1
13	1049	48	5.4	6.6	30	8.4	N	N	Y	N	1067	49	6.0	6.6	77	8.2
14	809	50	20.0	6.9	77	8.5	N	Y	Y	N	1031	54	23.1	6.8	35	8.2
15	1060	30	11.1	6.6	97	8.3	Y	Y	N	Y	604	39	19.3	7.1	108	8.6
16	943	56	15.4	6.6	26	8.2	Y	Y	N	Y	1047	43	12.0	6.9	32	9.0
AMY average		52.4	13.1	6.8	77.9	8.2	50%	50%	29%	36%		51.6	16.5	6.7	73.0	8.3
Stdev		12.3	6.2	0.2	35.6	0.6						11.3	8.2	0.2	32.3	0.4
ACC average		53.0	13.6	6.8	81.6	8.2	56%	50%	31%	38%		52.0	16.5	6.8	76.3	8.2
Stdev		12.3	6.1	0.2	35.0	0.7						10.7	7.7	0.2	31.5	0.6

Table 1 (Table 1) MDD and control cohorts

Age (years); PMI, postmortem interval (Hours); Storage, freezer storage at -80°C (months); RIN, RNA integrity number; Rec., recurrent episode; AD, antidepressant at death; Alc., Alcohol dependence at time of death. (*) Samples from pairs 5 and 6 were not available in AMY, thus group parameters were compared separately for 16 pairs in ACC and 14 pairs in AMY. Average values for age, PMI, pH, storage time and RNA quality were not different between MDD and control groups (all $p>0.05$).

	Physical effects		Emotion-related behavior		Altered Gene Expression	
	Body weight	**Fur coat quality**	**Anxiety-like**	**Agonistic**	**CC**	**AMY**
Control animal	Increase *	n/c	n/a	n/a	n/a	n/a
UCMS	Smaller increase #	Decreased #	Increased #	Increased #	254 genes	299 genes
UCMS + Flx	Reversed *	Reversed *	Reversed *	Reversed *	75% reversal	72% reversal
UCMS + Crf1r	Reversed *	Reversed *	Reversed *	Reversed *	44% reversal	77% reversal

Table 2 (Table 2) Summary of physical and behavioral changes evoked in the UCMS mouse model of depression, and reversal by chronic exposure to two antidepressant treatments

Data is from (9). The physical effects, emotion-related changes and associated gene changes evoked by UCMS were reversed by chronic antidepressant exposure. Significant effects of UCMS and antidepressant reversal are respectively indicated by (#) and (*); n/c, no change; n/a, not applicable. Flx, fluoxetine. Crf1r, corticotropin-releasing factor receptor 1 antagonist, SSR125543. Behavior: n=18-19/group. Array: n=6 arrays/group; Statistical criteria, ANOVA $p<0.05$, group differences >20%.

	Gene Title	Gene Symbol	Entrez Gene	Human MDD		Mouse UCMS					Neuronal/ glial origin		qPCR (p)	qPCR (alr)	Synapse function & development	Neurotransmission & signalling	Cell-matrix interaction	Oligogendrocyte-enriched
				All MDD subjects (alr)	MDD^{UCMS} subjects (alr)	UCMS Effect (alr)	Flx reversal (%)	Residual UCMS effect (alr)	Crf1r reversal (%)	Residual UCMS effect (alr)	Log_2 (WM/GM)							
*	calcium channel, voltage-dependent, beta 2 subunit	CACNB2	783	0.02	0.49	0.37	91	0.03	84	0.06	-1.9	NEURONAL	0.004	0.27	X	X		
	Calcium/calmodulin-dependent protein kinase II delta	CAMK2D	817	0.08	1.07	0.54	51	0.27	50	0.27	-1.6		0.070	0.49		X		
	ankyrin repeat domain 43	ANKRD43	134548	0.06	0.28	0.44	107	-0.03	119	-0.08	-1.4							
	Rho GTPase activating protein 6	ARHGAP6	395	0.42	0.75	0.45	151	-0.23	87	0.06	-1.2		0.003	0.68	X			
	cadherin 13, H-cadherin (heart)	CDH13	1012	0.18	0.42	0.45	84	0.07	66	0.15	-1.1						X	
	diacylglycerol kinase, gamma 90kDa	DGKG	1608	0.34	0.44	0.46	151	-0.23	93	0.03	-0.9		0.031	0.17		X		
	early growth response 1	EGR1	1958	-0.29	-0.30	-0.29	48	-0.15	50	-0.15	-0.9		0.030	-0.58	X			
	neuronal pentraxin I	NPTX1	4884	0.1	0.35	0.35	99	0.00	38	0.22	-0.7		0.060	0.39		X		
*	V-jun sarcoma virus 17 oncogene homolog	JUN	3725	0.19	0.66	0.39	98	0.01	64	0.14	-0.6					X		
	protein phosphatase 1, regulatory (inhibitor) subunit 16	PPP1R16A	84988	-0.01	-0.23	-0.27	70	-0.08	83	-0.05	-0.5				X			
	transmembrane protein 17	TMEM17	200728	0.09	0.26	0.37	91	0.03	109	-0.03	-0.4							
	rabphilin 3A homolog	RPH3A	22895	-0.36	-0.47	-0.33	82	-0.06	53	-0.15	-0.3		0.090	-0.54	X			
	matrilin 2	MATN2	4147	0.33	0.64	0.80	111	-0.09	90	0.08	-0.3		0.005	0.49			X	
	zinc finger protein 703	ZNF703	80139	-0.09	-0.24	-0.33	66	-0.11	140	0.13	-0.3							
	chromosome 5 open reading frame 22	C5orf22	55322	0.02	0.60	0.34	88	0.04	98	0.01	-0.2							
*	potassium channel tetramerisation domain containing	KCTD12	115207	0.17	0.34	0.31	127	-0.09	49	0.16	-0.1		0.110	0.21		X		
*	P18SRP protein	P18SRP	285672	0.05	0.45	0.27	96	0.01	134	-0.09	-0.1							
	carbohydrate (chondroitin) synthase 1	CHSY1	22856	0.02	0.30	0.30	50	0.15	46	0.16	-0.1						X	
	nuclear factor I/B	NFIB	4781	0.08	0.41	0.50	153	-0.27	80	0.10	0.1					X		
*	integral membrane protein 2A	ITM2A	9452	-0.15	-0.41	-0.35	114	0.05	96	-0.01	0.1							
	chromosome 5 open reading frame 13	C5orf13	9315	-0.16	-0.45	-0.46	111	0.05	78	-0.10	0.2							
	zinc finger protein, multitype 1	ZFPM1	161882	0	-0.25	-0.58	62	-0.22	81	-0.11	0.4							
	copine family member IX	CPNE9	151835	-0.29	-0.65	-0.70	52	-0.34	53	-0.33	0.6		0.001	-1.00		X		
*	myelin basic protein	MBP	4155	-0.22	-0.41	-0.51	72	-0.14	73	-0.14	1.3		0.010	-0.95				X
	aspartylglucosaminidase	AGA	175	-0.26	-0.19	-0.38	112	0.05	69	-0.12	1.5		0.010	-0.60				X
*	2',3'-cyclic nucleotide 3' phosphodiesterase	CNP	1267	-0.12	-0.37	-0.56	58	-0.24	68	-0.18	2.1		0.085	-0.50				X
	breast carcinoma amplified sequence 1	BCAS1	8537	0.02	-0.41	-0.65	82	-0.12	89	-0.07	2.1		0.080	-0.64				
*	ectonucleotide pyrophosphatase/phosphodiesterase 2	ENPP2	5168	-0.12	-0.29	-0.50	107	0.04	130	0.15	2.1		0.118	-0.37				X
	plasma membrane proteolipid (plasmolipin)	PLLP	51090	-0.16	-0.40	-0.56	46	-0.30	62	-0.21	2.3							X
*	endothelial differentiation, lysophosphatidic acid GPCR	EDG2	1902	-0.10	-0.40	-0.85	61	-0.33	79	-0.18	2.7		0.080	-0.39				X
	G protein-coupled receptor 37	GPR37	2861	-0.31	-0.52	-0.50	69	-0.16	95	-0.03	2.9							X
	myelin-associated oligodendrocyte basic protein	MOBP	4336	-0.11	-0.51	-0.88	83	-0.15	88	-0.10	3.1	GLIAL	0.019	-0.76				X

Table 3 (Table 3) Core genes significantly affected in human MDD and mouse UCMS

(*) identify genes for which multiple probes were significantly and similarly affected in MDD and/or UCMS. Red and blue indicate significant up- and downregulation, respectively. % of reversal indicates the degree to which the drug treatments opposed the UCMS effect and brought transcript levels (alr, average log_2 ratios) back to control levels. “Residual effect” indicates changes in gene transcripts after antidepressant treatments in UCMS-exposed mice. “Log_2(WM/GM)” indicate relative glial (high LogR, bottom rows) to neuronal (low LogR, top rows) enrichment of gene transcript [See methods and (Sibille *et al*, 2008)], consistent with other large-scale categorization of cellular origin approaches (Oldham *et al*, 2008). qPCR p- and alr values correspond to MDD^{UCMS} samples.

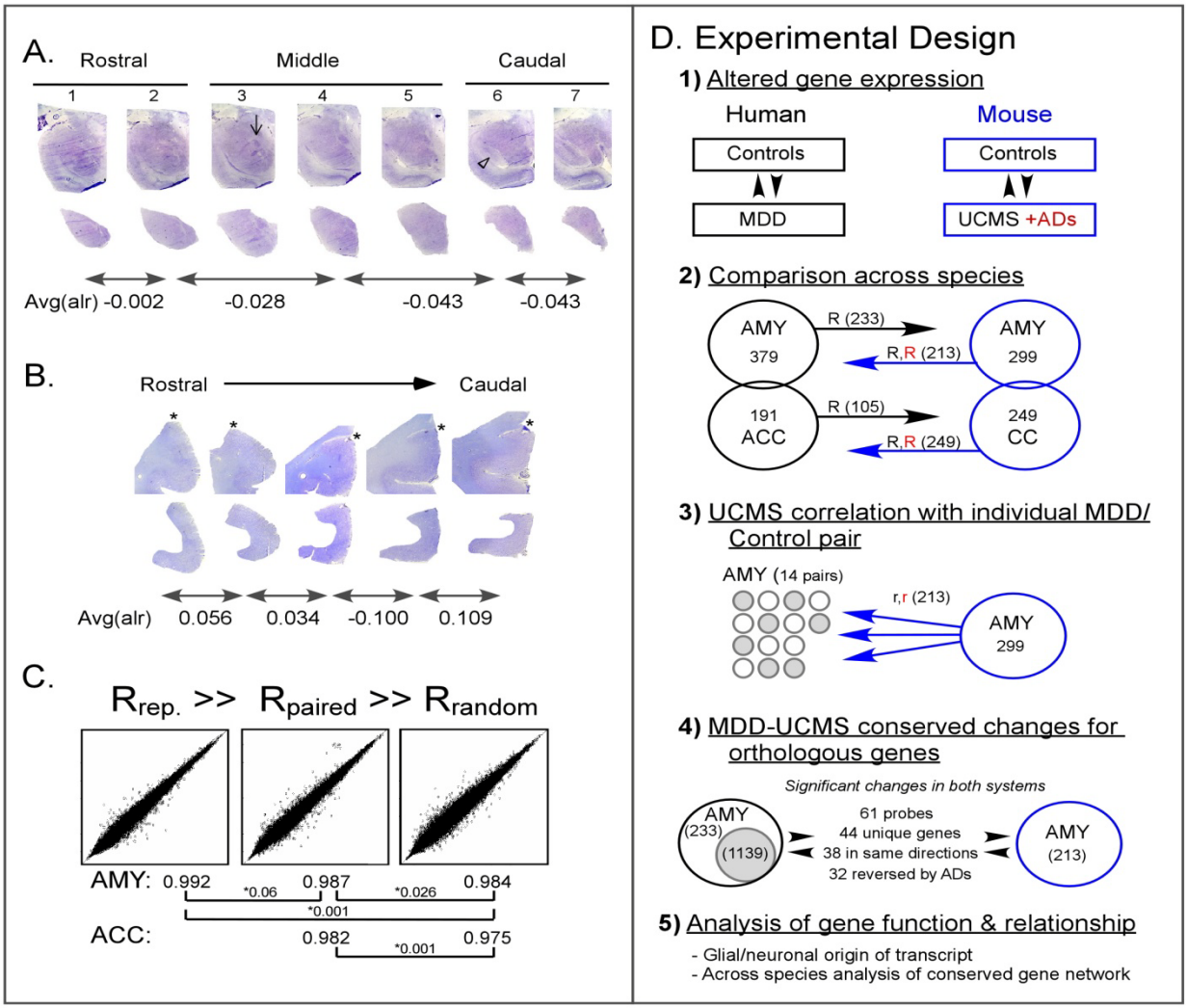


Figure 3 (Figure 1) AMY/ACC sampling pilot study, effect of pairing protocol on analytical sensitivity, and experimental outline

A) AMY micro-dissected samples (bottom) from serial 20µm sections 1mm apart were processed on microarrays along the rostral-caudal axis for sections 1, 2, 4, 6 and 7. Markers for the middle and caudal AMY were the dorso-lateral subnuclei of the basolateral nucleus (arrow) and the appearance of the hippocampus (arrowhead). The lowest variability was observed between levels 1 and 2, as indicated by the average changes for detected genes [Avg(alr)]. **B)** 20µm serial sections through the rostral subgenual ACC were initiated with the rostral tip of the subgenual ACC. Asterisks indicate the boundary between ACC and corpus callosum. Microarray samples were processed on the gray matter from 5 sub-dissected sections 1 mm apart (bottom). Although

absolute values were generally higher than in AMY, the lowest average variability in transcript levels was detected between levels 1 and 2 in rostral ACC. Accordingly, sampling was performed for all subjects at the levels of sections 1-2 in AMY and ACC. **C)** Representative Log_2-based correlation graphs for all detected transcripts illustrate the effect of the pairing protocol at reducing non-disease related effects. “Rep” indicates technical replicates. “Random” indicates non-paired MDD and control samples. *, $p<0.05$. *; p-values are from exact Wilcoxon signed-rank tests performed on ~26,000 gene transcripts with detectable levels.

D) Experimental design. Following analyses of altered gene expression in MDD and UCMS (1), the degree to which changes in one species predicted similar trends at the group level in the other species was assessed (2). In (3), UCMS predictions of gene changes in individual MDD subject were assessed. (4) Individual genes with conserved changes and AD reversal were selected between UCMS and all MDD subjects, or with the MDD^{UCMS} subgroup (Grey circle), and (5) assessed for known cellular origin, function and relationship. R, directional Pearson correlation coefficient, including after AD treatments (red). “Values” indicate genes with significant effects in (1). “(Values)” indicate the numbers of identified mouse-human orthologs among significant genes that were used for cross-species analyses.

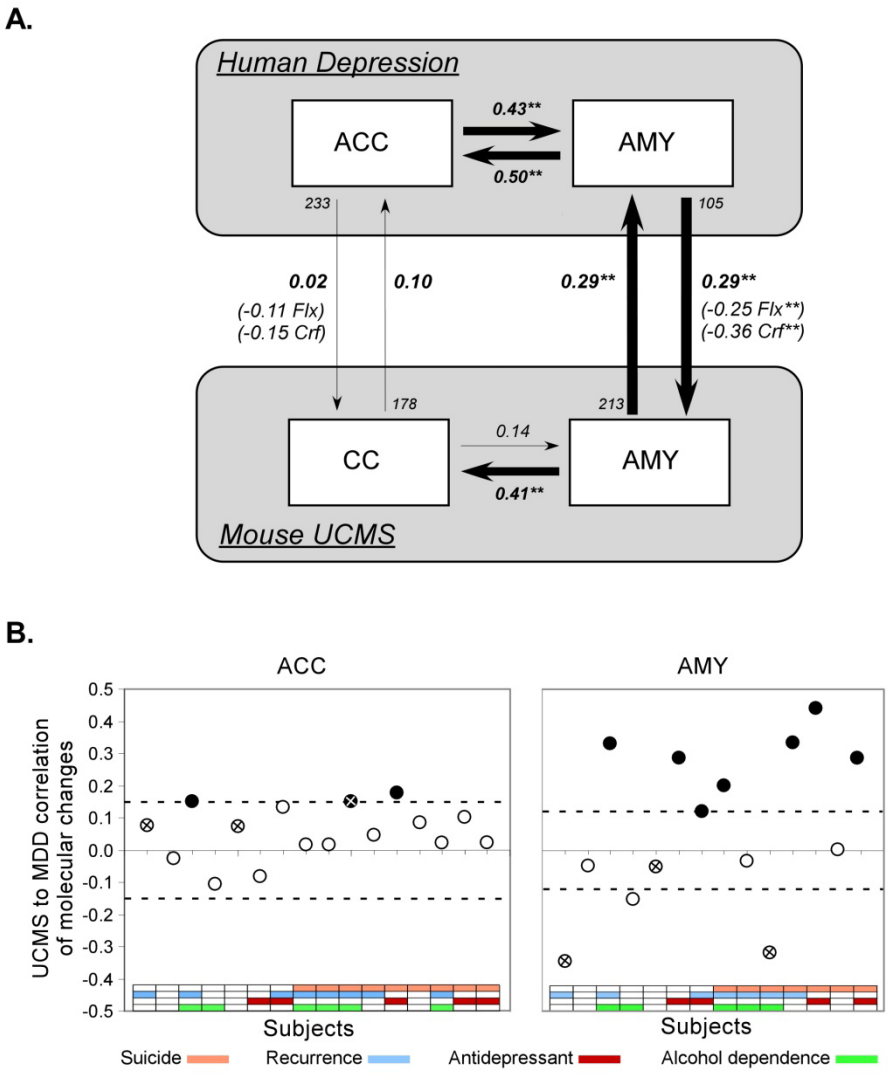


Figure 4 (Figure 2) Reciprocal prediction of altered AMY gene expression between human MDD and the mouse UCMS model of depression, and reversal by antidepressant drug treatments

A) Arrows indicate directional correlations between changes in transcript levels for genes identified in one area (origin of arrow) and changes for the same genes across areas within species, or within the same areas but across species (end of arrow). Numbers in italics at the origin of the vertical arrows indicate the numbers of genes significantly affected in that species and for which data was available for orthologous genes in the other species. Averaged group values per gene transcript were used here, when assessing MDD-UCMS correlation. Numbers in parentheses indicate levels of directional correlations between human MDD and UCMS-exposed mice

chronically treated with two antidepressants (Flx, fluoxetine; Crf, Crf1r antagonist). *, p<0.05; **, p<0.001. **B)** Correlation between UCMS and individual MDD subjects confirmed the absence of conserved effect in ACC, and identified a subgroup of human MDD subject with significant and positive cross-species correlations of altered transcript levels in the AMY. Subject-wise gene transcript changes were used here. Black dots indicate significant positive correlation. Crossed circles indicate subjects under partial or full remission at time of death. Dashed lines represent boundaries for significance of correlations.

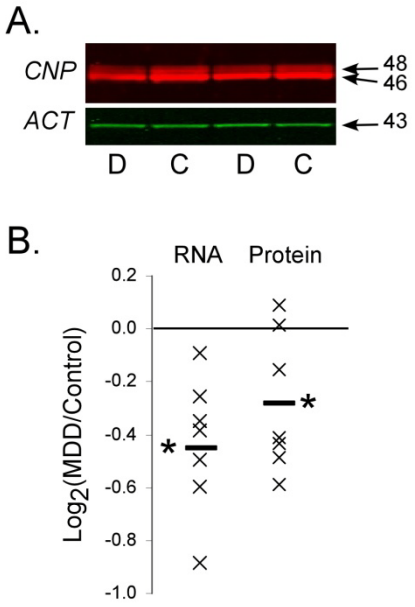


Figure 5 (Figure 3) Concordant *CNP* RNA and protein downregulation in MDD^{UCMS} subjects

A) Western blot analysis revealed the expected bands for the two CNP protein isoforms. D and C indicate replicate MDD and control samples from the same matched pair of subjects. **B)** Quantitative plot denoting significant *CNP* downregulation at the RNA (-29.2%; p<0.01; array results) and protein (-21.5%; p=0.01) levels in MDD^{UCMS} subjects. Both protein isoforms were similarly affected.

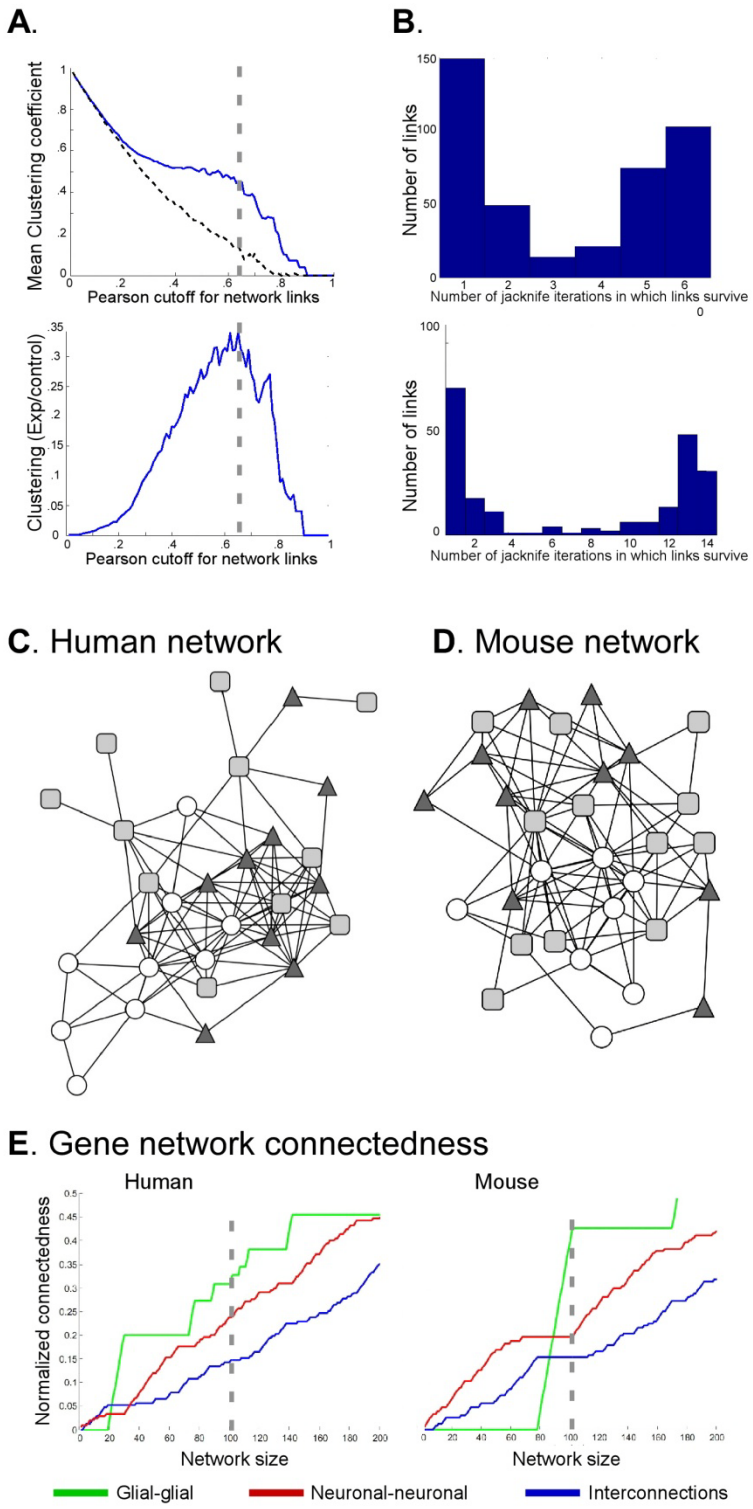


Figure 6 (Figure 4) A conserved and tightly clustered gene coexpression network with distinct glial and neuronal components underlies the identified molecular signature of depression.

A-B) Optimized parameters for building gene coexpression networks. **A)** (Top) The mean clustering coefficient (CC) for permutated degree-matched control networks decreases monotonically (dashed line). However, clustering in experimental networks showed a relative increase compared to permuted values at cutoff values in the region of 0.5-0.8 for both human (shown) and mouse (not shown) networks, indicating biologically relevant network structure(Elo *et al*, 2007). (Bottom) Optimum Pearson correlation cutoff for mean clustering coefficients (Dashed lines) were estimated by subtracting values of permutated degree-matched networks from actual values. Representative plots are from human control subjects. **B)** The histograms of link robustness under jackknife correlation demonstrated the common bimodal distributions between robust (right columns) or spurious (left columns) links in human (bottom) and mouse (top) networks. **C-D)** Organic representation of the coexpression networks formed by genes supporting the molecular signature of depression in humans (**C**) and mice (**D**), based on parameters optimized in (A-B). White circles indicate glial-enriched gene transcripts; Light to dark grey shapes indicate genes with increasing enrichment in neuronal origin of transcripts. Links between nodes represent coexpression links. **E)** Plots of glial-glial, neuronal-neuronal and neuronal-glial connectedness (i.e., number of connection within groups of genes divided by the total number of possible connections between those genes) as a function of network sizes (i.e. number of links) The visual segregation of glial-enriched and neuronal-enriched genes observed in (**C-D**) was reflected by increased numbers of connections within glial or neuronal groups compared to glial-neuronal connections (i.e. vertical distances for given network sizes). The sudden emergence of glial networks reflects the modularity of coexpression networks formed by fewer genes (and fewer samples in the mouse study) and is a strong indicator of unified function. The vertical hashed bars represent the chosen optimized cut-offs for network design using robust and inclusion of representative glial-glial, neuronal-neuronal and neuronal-glial connectedness.

3.0 PAPER #2: ALTERED GENE SYNCHRONY SUGGESTS A COMBINED HORMONE-MEDIATED DYSREGULATED STATE IN MAJOR DEPRESSION

Chris Gaiteri[1,2], Jean-Philippe Guilloux[1,3], David A. Lewis[1,2] & Etienne Sibille[1,2,*]

[1] Department of Psychiatry, [2] Center for Neuroscience, University of Pittsburgh, Pittsburgh, Pennsylvania, USA, [3] Université Paris-Sud EA 3544, Faculté de Pharmacie, Châtenay-Malabry cedex F-92296, France.

* Corresponding author. Department of Psychiatry, University of Pittsburgh, 3811 O'Hara street, BST W1643, Pittsburgh, PA 15213, USA. Tel.; 1-412-624-0804; E-mail: sibilleel@upmc.edu.

Abstract

Coordinated gene transcript levels across tissues (denoted “gene synchrony”) reflect converging influences of genetic, biochemical and environmental factors; hence they are informative of the biological state of an individual. So could brain gene synchrony also integrate the multiple factors engaged in neuropsychiatric disorders and reveal underlying pathologies? Using bootstrapped Pearson correlation for transcript levels for the same genes across distinct brain areas, we report robust gene transcript synchrony between the amygdala and cingulate cortex in the human postmortem brain of normal control subjects (n=14; Control/Permutated data, $p<0.000001$). Coordinated expression was confirmed across distinct prefrontal cortex areas in a separate cohort (n=19 subjects) and affected different gene sets, potentially reflecting regional network- and function-dependent transcriptional programs. Genewise regional transcript coordination was independent of age-related changes and array technical parameters. Robust shifts in amygdala-cingulate gene synchrony were observed in subjects with major depressive disorder (MDD, denoted here “depression”) (n=14; MDD/Permutated data, $p<0.000001$), significantly affecting between 100 and 250 individual genes (10-30% false discovery rate). Biological networks and signal transduction pathways corresponding to the identified gene set suggested putative dysregulated functions for several hormone-type factors previously implicated in depression (insulin, interleukin-1, thyroid hormone, estradiol and glucocorticoids; $p<0.01$ for association with depression-related networks). In summary, we showed that coordinated gene expression across brain areas may represent a novel molecular probe for brain structure/function that is sensitive to disease condition, suggesting the presence of a distinct and integrated hormone-mediated corticolimbic homeostatic, although maladaptive and pathological, state in major depression.

Introduction

Major depression affects more individuals than all other psychiatric illnesses combined, and a significant number of patients do not remit after pharmacological or behavioral treatment (Belmaker *et al*, 2008), hence inflicting a continuous toll on affected individuals and on society (Warden *et al*, 2007). Changes in the coordinated function of a neural network comprising cortical and subcortical brain areas are thought to underlie the mood regulation deficit in depression (Seminowicz *et al*, 2004). The functional connectivity between two critical components of this corticolimbic circuitry, the amygdala and anterior cingulate cortex, potentially mediates the relay of emotion-related information for cortical processing, and feedback regulation on amygdala activity (Pezawas *et al*, 2005). In control non-depressed subjects, the volume, function and connectivity of these two areas are affected by serotonin-related gene variants (Hariri *et al*, 2002; Pezawas *et al*, 2005), together suggesting that this pathway may be recruited in diseases of altered mood. Indeed, recent findings suggest an increased task-related recruitment of rostral cingulate and decreased coactivation of amygdala and cingulate in depressed patients(Matthews *et al*, 2008). Moreover, studies in depression suggest functional, cellular and molecular pathologies in both areas (Cotter *et al*, 2001; Drevets *et al*, 1997; Hamidi *et al*, 2004). So, alterations in the intrinsic circuitry of the amygdala and anterior cingulate cortex may result in altered connectivity, deficient cingulate feedback regulation on amygdala function and abnormal processing of emotion-related stimuli in depression (Sheline *et al*, 2001; Siegle *et al*, 2002).

How can brain region activities be investigated at the molecular and gene levels? We hypothesized that pathological mechanisms leading to depression may affect the coordination of gene expression patterns in the brain, and tested this hypothesis within a set of related brain areas, the amygdala and the subgenual anterior cingulate cortex. Our assay makes use of the fact that correlations in gene transcript levels across samples and datasets (“coexpression” or “coregulation”) represent intrinsic attributes of cellular and biological systems (Pavlidis *et al*,

2004; Prieto *et al*, 2008), including in the human brain (Bullmore and Sporns, 2009; Oldham *et al*, 2008). As expected for inter-related biological systems, many genes show variability in expression that do not reflect measurement error and that are consistently identified in a range of tissues and organisms(Lee *et al*, 2004; Mijalski *et al*, 2005; Prieto *et al*, 2008). These correlated relationships are *(i)* driven by various molecular mechanisms, genetic make-up and function-dependent synchronization, *(ii)* central to cellular function, *(iii)* link genes of common biological functions (Lee *et al*, 2004; Mijalski *et al*, 2005), and thus can be used to create gene interactions networks (Dobrin *et al*, 2009). Hence, based on indications that correlated expression profiles might serve as markers of cellular or tissue relationships, we investigated synchronized expression across two regions implicated in the altered mood component of major depression. Confirming our hypotheses, we show that gene-wise coordinated transcript levels is a robust component of expression across regions within subjects, and that major depression is associated with significant gene-specific alterations in amygdala-cingulate gene coordination.

Results

Large-scale gene transcript synchrony across brain areas

Transcripts for a particular gene are synchronized between two regions if they display significantly higher correlation across brain regions compared to permutated data. Here, using gene array data in the human postmortem brain of control subjects (Sibille *et al*, 2009) (n=14), we demonstrate that a large number of genes displayed positive correlations of transcript levels between amygdala and cingulate (Figure 1A-D), resulting in a unimodal distribution (Median r=0.32) (Figure 1E). In contrast, the distribution of the permutated data, in which the subject linkage across regions was scrambled centered on the null correlation (Dashed line in Figure 1E; r=0.014; Control/Permutated, $p<0.000001$). As no similar dataset are currently available (See Discussion), we investigated the presence of regional gene synchrony across a different set of brain areas. Large-scale gene synchrony was confirmed between two pre-frontal cortical

regions in an independent cohort (Erraji-Benchekroun *et al*, 2005) (n=19 subjects), and affected different genes (Figure 1F-G).

Gene coordination is not a result of age-related changes in gene transcript levels

Age-related genes could produce correlations between gene expression levels across samples because numerous genes have strong correlations with age (Erraji-Benchekroun *et al*, 2005) and our subjects are acquired from a range of ages. It is therefore possible that the permutation destroys the global shift towards higher correlation (Figure 1E) simply because it scrambles the age-induced correlations. To investigate this putative effect, we subtracted age as a contributing factor and recomputed regional gene transcript correlations (see Methods and Figure 2B-D for details on age-detrending). As figure 2a shows, the entire shift towards positive correlation values across regions via gene coordination was achieved without any age-correlations in the expression profiles, demonstrating that the gene expression correlations with age have only an exceedingly small influence on regional gene coordination.

A genetic confound for these effects was also extremely unlikely as allelic frequencies for very large gene numbers would need to be similar within cohort, while being significantly different across cohorts, in order to generate the observed differences. Moreover, transcripts with reported impact of genetic variant on array hybridization and signal level (Sliwerska *et al*, 2007) displayed low synchrony in both sets of brain regions (i.e. COMT, $r^2 < 0.2$, $p > 0.05$), thus ruling out technical confounders. Finally, regional gene coordination was observed across a heterogeneous cohort and could not distinguish the contribution of individual demographic or clinical factors (e.g. male/female difference, drug exposure; Supplementary Table 1).

Gene transcript coordination is not a result of microarray gene expression protocols

RMA-based methods (as used here) have superior low-level detection capability

compared to GCOS-MAS5 (Galfalvy *et al*, 2003); however, to ensure that normalization method was not responsible for the observed correlations, we recomputed the cross-area correlations using MAS5-normalized data. We observed similar and highly significant shifts ($p<.000001$) to positive levels of coordination for both amygdala-cingulate and prefrontal cortex regions using MAS5-normalized data, as we did with GCRMA-normalized data (not shown). Moreover, normalization-based differences in the estimated level of coordination were consistent low for highly coordinated genes, which are the focus of this study.

Technical reliability of array data by quantitative real-time PCR (qPCR)

qPCR confirmation typically relies on measuring group differences 20% or greater in magnitude and averaged over multiple samples. Accordingly, the technical reliability of the array data used here was previously validated by independent qPCR measurements (array/qPCR correlation ≥0.75; n=~20 genes) (Sibille *et al*, 2009). While independent verification of differential expression changes between groups is a preliminary condition to establish confidence level in the quality of the gene array dataset before pursuing gene coregulation studies, measures of coordinated gene expression rely on changes of small magnitude across individual samples (Main text; Figure 1A-C) that are typically within the margin of technical variability in qPCR. For instance, we performed qPCR on 18 additional genes with coordinated regional expression (ATP5G1, CRHBP, MAOB, NFE2L2, PHKB, POLR2E, PRKAG2, RXRA, SAT1, AACS, CAP1, CDC42, CRYZ, GRLF1, IRF2BP2, NEFL, PAPOLA and SCN2A2). Samples were run in quadruplicates based on three internal controls (ACT, GAPDH and CYC). Transcript changes associated with variable coregulation levels resulted into lower coefficients of variation by array quantification ($CV_{array}=0.19$ in amygdala and cingulate; n=18 genes) compared to qPCR values ($CV_{qPCR}=0.37$; $p<0.005$ in amygdala; $p<0.0001$ in cingulate, two-group *t*-tests). This higher CV_{qPCR} potentially reflects the exponential amplification of PCR

reactions compared to the linear hybridization detection by arrays, and effectively limits the potential of confirming coregulation by qPCR in this and related studies. These results are consistent with other studies on microarray-based gene network, where qPCR is typically used to validate the mean absolute expression level (Choi *et al*, 2005; Day *et al*, 2009). Instead, the validity and biological relevance of coordinated gene expression typically relies on extensive alternative forms of conformation through combining datasets and through permutation testing procedures (both performed here) (Mijalski *et al*, 2005; Pavlidis *et al*, 2004), while the larger accuracy of gene networks has been confirmed through functional convergence across groups of affected genes (Lee *et al*, 2004; Oldham *et al*, 2008; Prieto *et al*, 2008).

Altered amygdala-cingulate corticolimbic regional gene synchrony in depression

We next tested the hypothesis that pathological mechanisms leading to depression may affect the coordination of gene expression patterns between the amygdala and cingulate cortex. Array data from subjects with major depression (Sibille *et al*, 2009) (n=14; Supplementary Table 1) displayed a similar amygdala-cingulate gene transcript right-shifted correlation distribution (Figure 3A; median r=0.32; MDD/Permutated data, $p<0.000001$). Transcript synchrony was similar in control and depressed subjects (Gray dots in Figure 3B*;* $p<0.000001$); hence independently confirming amygdala-cingulate regional gene synchrony in human subjects. Relying on permutation testing procedures to ensure statistical significance of coregulation measures and after controlling the false discovery rate (FDR) (10-30%), as many as 94 gene transcripts displayed robust significant loss of regional synchrony in depression (Blue in Figure 3b-c; from greater to 0.7 to less than 0.2 values), while over 180 displayed significant gain of synchrony in depression (Red in Figure 3B-C; from less than -0.7 to greater than -0.2 values) (Supplementary Table 2).

Genes with altered transcript synchrony participate in biological networks and signal transduction pathways modulated by hormonal factors previously implicated in depression

To gain insight into biological functions affected by these changes, we overlaid the top 80 genes (~20% FDR) under both conditions (gain or loss of synchrony) onto the global molecular network of the Ingenuity knowledge database. The top gene networks in both categories displayed similarities, as they included numerous signal transduction and transcription components of the mitogen-activated protein kinase pathway (Table 1) and other genes previously implicated in depression (CREB1 (Nestler *et al*, 2002), SAT1(Sequeira *et al*, 2006)). The unbiased inclusion of additional nodes significantly linked to depression-affected genes identified insulin – a recurrently-suggested contributor to neuropsychiatric disorders - as a putative modulator for both networks (Figure 3d; Table 1). The pro-inflammatory cytokine, interleukin 1 (IL1) was identified as a second putative modulator ($p<0.0001$) for the top network formed by genes with elevated synchrony in depression (Figure 3d; Table 1). Additional biological modulators identified in the top three networks in each category included thyroid hormone, a clinically-useful antidepressant-augmenting agent (Bauer *et al*, 2002), and beta-estradiol, the major brain estrogen. All associations of the identified modulators with the top gene networks were significant ($p<0.05$), as assessed by bootstrap resampling. These associations were also selective ($p<0.01$), as assessed by repeated testing of the Ingenuity database with random gene lists of equivalent or variable sizes of selected genes (60 to 200 genes). Pathways, biological functions and diseases associated with altered gene synchrony are summarized in Table 1. Notably, glucocorticoid receptor signaling was the top canonical pathway associated with three of the top networks, linking stress hormone-related events - a well-characterized causative factor (Holsboer, 2000) - to the deregulated molecular state in depression.

Discussion

Our findings demonstrate that regional gene synchrony, as measured by gene-wise correlated transcript levels across brain regions within individuals, is a major component of gene expression patterns in the human brain (Figure 1). These patterns were not explained by genetic, age or microarray effects, and appeared driven by correlation within subjects, as scrambling the data across subjects abolished them (Figure 1e-f, Figure 2). Thus we speculate that regional gene synchrony may partly reflect an integrated molecular output of function-, and dysfunction-dependent, regulation of brain areas.

While data from the amygdala/cingulate cortex and from the two prefrontal cortex areas supports the contention that gene coordination may reflect an overall, or network-specific, concerted brain region function, our results are independent of these larger hypotheses, as we only considered here the functional significance of those genes which show significant depression-related alteration in gene synchrony between two regions known to be functionally affected by depression. Accordingly, by bootstrapping correlations and controlling the FDR, we identified a robust and conservative collection of genes that displayed significant gains or losses of amygdala-cingulate gene transcript coordination in subjects with depression (Figure 3) (which is distinct from mean absolute expression level changes; See Comments section). These gene sets implicated shifts in intracellular signaling, metabolism and cell growth/structure, and suggested the implication of several biological modulators previously associated with depression (Table 1).

Notably, changes in amygdala-cingulate gene synchrony suggest a combined dysregulated function for several hormone-type modulators (Figure 3d; Table 1), which together summarize several key hypotheses for pathophysiological mechanisms in depression. As postmortem studies preclude investigating short-term events, we propose that the present

findings may correspond to a stable, chronic and adaptive *de novo*, although pathological, state. The integrative nature of this deregulated state departs from reductionist approaches and has critical implications for our understanding and modeling of pathological mechanisms of depression.

What may underlie regional gene synchrony?

As coregulated gene expression reflect the influences of genetic, biochemical and environmental factors (Elo *et al*, 2007; Lee *et al*, 2004), we speculate that the observed gene synchrony across regions may reflect a molecular balance of local brain systems that is achieved over time (days-months) through the coordinated function and continuous feedback of interacting brain regions. For instance, starting at the cellular level, the rate of neuronal firing is determined by the molecular composition of local neuronal circuits. The cumulative electric signals of single neurons with neuronal ensembles oscillate on various timeframes, supporting regional brain function and underlying correlated functions across regions. In turn, the translation and integration of neuronal activity by intracellular signaling cascades is influenced by coordinated activities across functionally-related brain regions. Additionally, broad and long-acting modulators (hormones-type factors) modulate transcriptional programs (through nuclear receptors, for instance) and interact/modify this conversion of neuronal activity into cellular changes over time and across areas. According to this model, the disturbances in biological rhythms observed in depression (circadian, hormonal cycles) and the known role of environmental exposure (stress, disease) in precipitating disease episodes, will influence the degree of cellular exposure to hormone-type modulators, and may potentially result in an altered, yet stable, molecular balance. This suggested mechanism resembles a “decanalization” process that has been proposed for complex disorders, where chronic shifts in various regulating factors converge to induce and maintain a departure from the biologically-optimized

healthy organism into a distinct, stable and maladaptive pathological state (Gibson, 2009).

A combined hormone-mediated disease pathology in major depression

Consistent with the above-proposed model, the unbiased analysis of biological modulators associated with networks formed by genes with altered amygdala-cingulate coregulation identified factors previously implicated in depression or in its treatment (insulin, beta-estradiol, thyroid hormone, IL1 and glucocorticoids; Table 1), although notably, none of them would be sufficient to reasonably explain the presence of the illness in heterogeneous clinical cohorts (Belmaker *et al*, 2008; Nestler *et al*, 2002). ***(1)*** Insulin shared potential control over genes forming the most robust networks under conditions of gain and loss of synchrony (Figure 3), suggesting that deregulation of this homeostatic modulator may participate in mediating pathological changes in depression. Insulin has been suggested as a potential mediator of metabolic changes in neuropsychiatric disorders (Altar *et al*, 2008), whereas insulin-resistance is more frequent in subjects with familial depression (Lewis *et al*, 1983). ***(2)*** Thyroid hormone influences brain physiology through regulating basal metabolism and neuronal maturation. Low thyroid function is associated with increased incidence of depression, while thyroid adjuvant therapy augments antidepressant therapy, potentially through deactivation of limbic regions (Bauer and Whybrow, 2003). ***(3)*** Although associated with mood changes in female subjects, the relevance of altered estradiol function to this male group is underscored by local aromatase–mediated conversion of testosterone to estradiol, including in the amygdala, where it modulates anxiety and depressive-like behaviors (Walf and Frye, 2006). ***(4)*** IL1 is a potent pro-inflammatory cytokine, which mediates aspects of the “sickness behavior”, a syndrome sharing similarities with major depression (Dantzer *et al*, 2008). Interestingly, IL1 was associated here with increased regional gene synchrony in depression, suggesting a gain-of-function mechanism consistent with IL1 recruitment and role. ***(5)*** Finally, as potential core

inducing-factors of the illness (Holsboer, 2000), glucocorticoids (and stress) are known to modulate the functions of all other identified factors, resulting among others in altered blood brain barrier function, decreased glucose uptake, immune activation, and disrupted sex hormone cycling or release (Bauer *et al*, 2003; Dantzer *et al*, 2008; Goshen and Yirmiya, 2009). Circulating interleukin and other cytokines also affect insulin function (Dantzer *et al*, 2008), together suggesting that a complex interplay of disrupted hormone-mediated regulations of organs and cell ensembles may occur in depression.

Hence, as contributing roles in depression are separately consistent for all identified factors, it is conceivable to envision a model where sustained environmental and lifestyle changes induce chronic adaptive changes in several systems (insulin, sex-hormones and thyroid-related functions), which now interacts with individual genetic make-up or additional environmental disturbances (stress or infection). Thus this model connects and potentially synergizes distinct and previously-proposed pathophysiological mechanisms for depression. Accordingly, pathways to a depressive state are not likely to be explained by any single factor, but may reflect the disruption of several hormone-type factors acting on different timeframes (cyclic, constant, phasic or induced). Here our results suggest that in depression these factors may converge on intracellular pathways (e.g., MAPK pathway), mitochondria and energy metabolism, and on other neurochemical pathways, such as SAT1 and polyamines (Sequeira *et al*, 2006), resulting in altered function and cell structure (growth, adhesion) within the amygdala-cingulate network (Table 1). In conclusion, we propose that the identified departures in corticolimbic regional gene synchrony represent an integrated gene/molecular signature of a *de novo* maladaptive and pathological state in subjects with major depression.

Limitations and comments

The present findings demonstrate that regional gene coordination represents a biological

feature of the human brain across related areas, and that alterations in this phenomenon are useful for measuring integrated multi-scale effects in complex disorders such as depression. Our results provide the basis for further mapping of gene coordination structure onto specific functional brain networks. Indeed, the extent to which gene coordination follow the boundaries of known anatomical/functional networks is still to be determined. For instance, the amygdala is an anatomical and functional hub (Stein *et al*, 2007) and we may expect positive large-scale gene synchrony with other, but not all areas.

Similarly, the question of a control brain region in depression is often discussed but not clarified. Indeed, it is not know whether the primary pathology of the illness is region-specific or widespread and data has been provided for both cases (chi Hsiung *et al*, 2003; Sibille *et al*, 2009). So in short, there is no consensus for a "control" brain region in depression. An additional practical limitation is that no similar datasets are currently available to define the limits of the effect of MDD on regional gene coordination. Finally, critical to our findings, our results do not depend on other regions being affected or not, but instead provide information on gene coordination in areas of a network that is affected in depression. Whether other brain regions are affected is an important scientific question, but for which the answer is complementary rather than necessary for the current study.

With regard to validating gene coexpression, it is becoming increasingly clear that there are different types of independent validation of array data for: 1) differential expression level, 2) coregulation and 3) functional implications, which in turn require different analytical approaches. Coregulation relies on changes of small magnitude across samples that are difficult to replicate by qPCR. Our results provide a technical reason for the usual absence of such confirmatory approaches in coregulation studies, which is that the variability of qPCR measurement is higher than the one observed in array data. Such information had to our knowledge not yet been provided in the rapidly growing field of coregulation studies. Instead, coregulation methods rely

on assessing the probability of confirming the observed effects in other array datasets and on the probability of results belonging to common biological pathways. Here, we provide very robust statistical findings for these two types of validation, using bootstrap and other resampling statistical approaches. An implication of these observations is that it may not be wise to rely on single key genes as modulators of the observed effects, as the statistical reliability of any individual genes is moderate, compared to the robust statistical significance of coregulated or functionally-related gene groups. Here we relied on a process of convergent confirmation of mediators of depression, across groups of genes through the well-validated Ingenuity's literature-based database.

The mostly positive coordinated patterns may be surprising, based on known biological interactions between areas and cell types. Studies in regions with well-characterized neurotransmitter structures, such as raphe/cortex or substantia nigra/nucleus accumbens may help resolve this question. Notably, the proposed assay (regional gene synchrony) does not identify a single area of origin of disease-related changes (an intrinsic limitation of the approach), but rather suggests changes in factors supporting synchronization of gene function across brain areas in depression.

Other factors are likely engaged in the illness, for which the size and composition of the cohorts did not allow us to identify. For instance, early developmental events and indirect modulation (e.g. through monoamine regulation) may be more challenging to identify and are not necessarily well characterized in currently available functional gene networks. Finally, although we ruled out the contribution of several factors (genetic variants, tissue-specific programs, age) we only speculate that function-dependent regulation may be at play in supporting the depression-related correlation shifts. Hence, it will be of critical interest to assess whether correlated patterns return to control states in remitted subjects and if such patterns are measurable in rodent models of the illness. Finally, it is notable that robust coordinated patterns

were observed in a relatively small and heterogeneous cohort of subjects. Indeed, it is likely that demographic and clinical parameters, such as sex, race, lifetime stress exposure and antidepressant exposure for instance, will influence regional gene synchrony, although much larger cohorts and multi-region gene arrays datasets will be necessary to investigate the full extent of these effects. Despite all of these potentially confounding influences, gene coordination remains a strong influence on patterns of gene expression across areas, and for which alterations in depressed subjects correspond to known and suspected abnormalities in the illness.

Materials and Methods

Cohort description and array parameters

Human cohort 1 (amygdala and cingulate) includes samples from 28 white male subjects: 14 control subjects and 14 subjects with familial major depression. Subject description, array sampling and parameters were previously described (Sibille *et al*, 2009). In brief, brain samples were obtained at the Allegheny County Medical Examiner's Office (Pittsburgh) after written consent from next-of-kin. Consensus DSM-IV diagnoses were made by an independent committee of experienced clinical research scientists, utilizing information from clinical records, toxicology exam and a standardized psychological autopsy. Depressed and normal comparison subjects were matched for age, sex and race (Table S1). Amygdala samples were dissected from frozen coronal blocks ~2-3cm caudal to the temporal pole and were enriched in lateral, basolateral and basomedian nuclei. Cingulate samples were harvested from coronal sections in subgenual cingulate and contained all six cortical layers. All procedures were approved by the University of Pittsburgh's Institutional Review Board and by the Committee for Oversight of Research Involving the Dead. The second human cohort (BA9 and 47) includes 19 control subjects. Subjects description, array parameters and data are available in (Sibille *et al*, 2004).

Details about arrays processing and parameters are summarized in the supplemental files.

Statistical methods

Gene-wise bootstrapped Pearson correlation r-values were used to ensure an accurate estimation of the "real" underlying distribution and to avoid spurious findings due to outlying data. Because gene coordination relies on multiple samples to form a single measure, we used the percentile bootstrap method to ensure that the shifts were robust and significant ($p<0.05$) and then applied the Benjamini-Hochberg FDR (Benjamini *et al*, 2001). To increase the power of the analysis, we considered alterations of coordination in genes with high correlations in at least one condition, which are indicative of inter-regional communication (Dobrin *et al*, 2009).

While methods exist to optimize cutoff values in coexpression networks (Elo *et al*, 2007), there are no analogous mathematical methods for coordinated expression for the same genes. Therefore we used a 0.7 r-value cutoff (resulting in=3244 probesets with cross-area links) that was indicated as an optimal balance of false positives and false negatives for within-area amygdala and cingulate networks (Elo *et al*, 2007), and results did not significantly vary for alternative cutoffs (+/-0.1) (not shown). To generate p-values that quantify the depression-related shift in gene coordination, we used the percentile bootstrap method. These p-values for shifts in correlation were estimated using 20,000 bootstrap resamples of the raw data, at which point p-values were stable.

Methods for eliminating age correlation in microarray data

While baseline comparison of age (Figure 2B) did not show any influence of age-correlated genes on the amygdala-cingulate expression correlations (r=-0.01), to avoid ambiguity in the source of gene coordination we detrended any linear relationship with age in both amygdala and cingulate data (eliminating the possible influence of any large magnitude y-

values shown in figure 2B on gene coordination). Because the final correlations used to assess gene coordination were bootstrapped, to detrend the data we removed any linear relationship with age in each of the 20,000 bootstrap instances and used these detrended resamples to generate a histogram of amygdala-cingulate expression profiles as before (see Figure 2A).

Biological pathway, gene network and modulator analyses

Selected genes were overlaid on the global molecular network developed from information contained in the Ingenuity Pathway knowledge base (www.ingenuity.com). This network is composed of ~2 million literature-based biological links between genes and bioactive molecules, and sub-networks are built on genes of interest based on their connectivity within this global network. Gene networks were limited to 35 nodes. The score for a network takes into account the relative numbers of network eligible molecules, of molecules analyzed and the total number of molecules in Ingenuity's knowledge base. These scores are based on the hypergeometric distribution and represent the negative log of the right-tailed Fisher's Exact Test p-value. Disease links are based on literature-based association with illness. The major functions of gene clusters were determined by DAVID functional clustering (http://david.abcc.ncifcrf.gov).

To assess whether the association of the identified biological modulators with the top networks was specific ($p<0.01$), we resampled the Ingenuity database 100 times of with random gene lists of equivalent sizes. The process was repeated with variable sizes of selected genes (60 to 200 genes). Finally, the probability of finding the identified modulators in the top gene networks was assessed bootstrap resampling.

Funding

Supported by the following grants from the National Institute of Mental Health (NIMH): MH077159 (ES), MH084060 (ES) and MH084053 (DAL), and by a fellowship from the

University of Pittsburgh Institute for Clinical Research Education and the Clinical and Translational Science (CG). The funding agency had no role in the study design, data collection and analysis, decision to publish and preparation of the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

Acknowledgments. We thank Dr. George C. Tseng for additional statistical support.

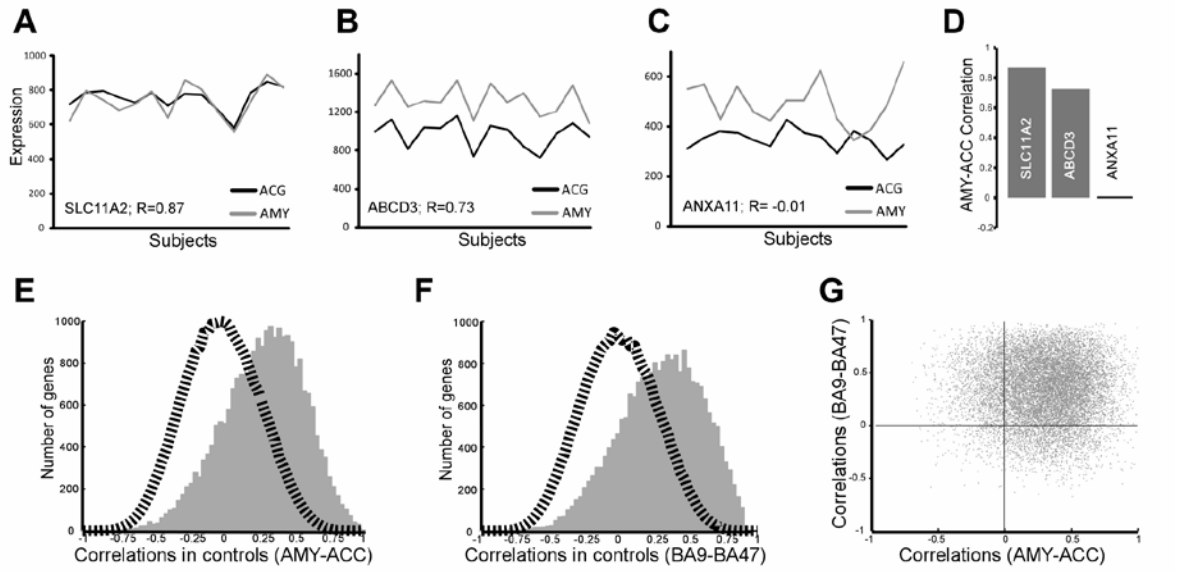


Figure 7 (Figure 1) Correlated genewise transcript levels across brain regions.

A-D Examples of within-subject positive, negative and absent amygdala-cingulate gene synchrony. **E** The right-shifted histogram of genewise transcript correlation suggests that the majority of genes are similarly regulated in both areas. The permutated data (dashed line) is centered on zero, indicating that gene coordination is subject-specific. **F** A similar pattern of gene synchrony was observed between two areas of the prefrontal cortex in an independent cohort ("BA", Brodman area). **G** The lack of correlations in the extent of gene synchrony

between the amygdala-cingulate and prefrontal cortical areas demonstrate that different sets of genes present coordinated transcript levels in the two different sets of brain areas (**G**, R=0.002). AMY, amygdala, ACC, anterior cingulate cortex.

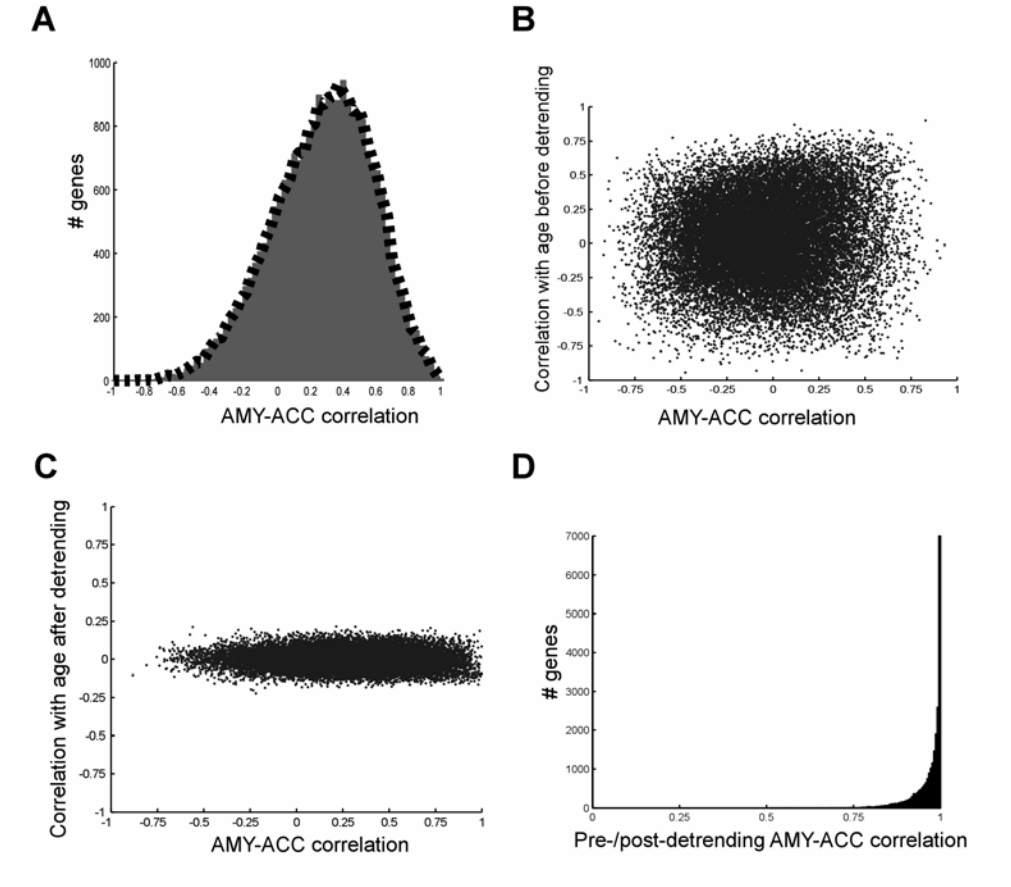


Figure 8 (Figure 2) Age-related genes do not significantly influence gene coordination.

A Age-detrended bootstrapped estimates of gene coordination were not significantly different from the null permutated model (distribution outlined by black dashed line) and did not decrease the overall levels of gene coordination. **B-C** Relationships between amygdala-cingulate coordination and age correlation, before (**B**) and after (**C**) removal of any age-correlation indicated that age-detrending did not affect amygdala-cingulate regional coordination. As shown in **C,** the distribution of age correlations is centered on zero and highly compressed compared to

figure **B**. The residual spread in the y-axis is due to inevitable, small, randomly occurring correlations while resampling the detrended data. **D** Histogram of correlations between original and detrended amygdala-cingulate coordination levels, showing that the vast majority of genes retain highly similar regional correlation, thus demonstrating an overall very small contribution of age to regional gene synchrony.

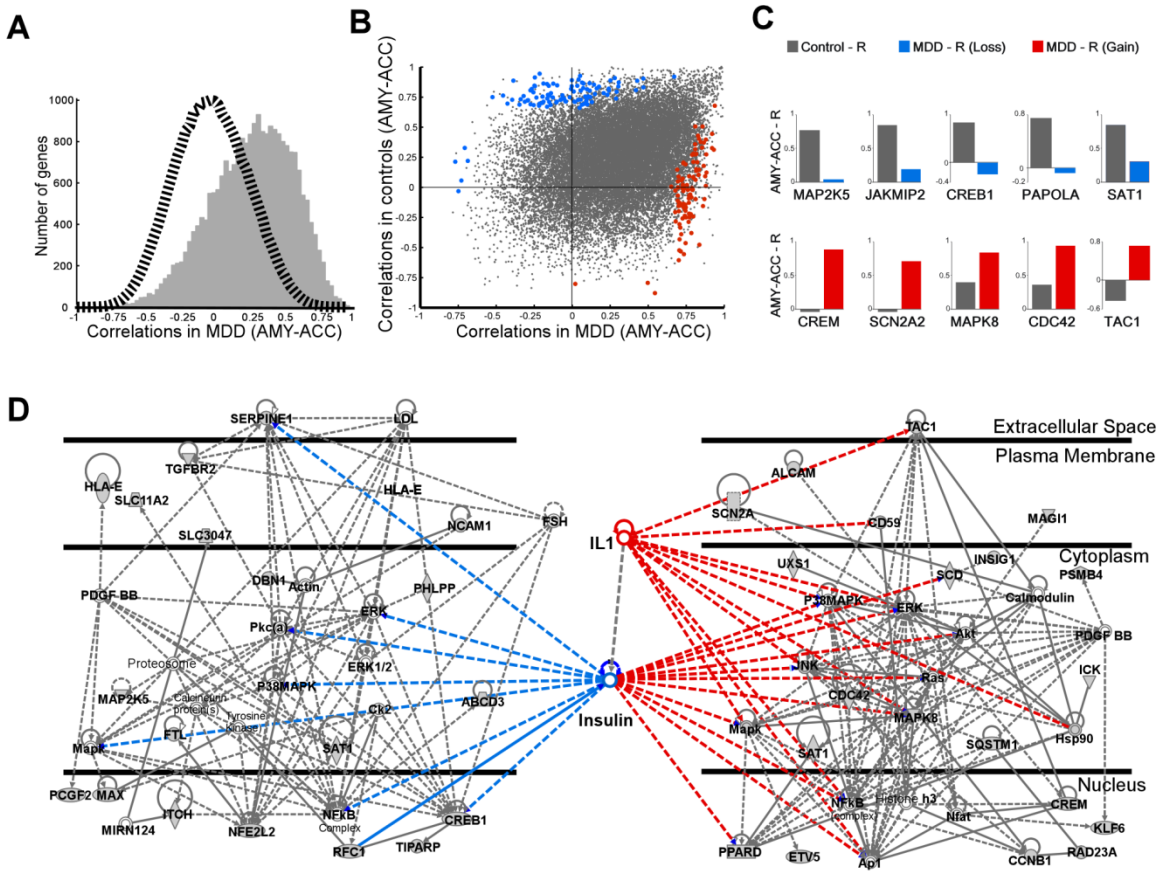


Figure 9 (Figure 3) Altered amygdala-cingulate regional gene synchrony in subject with major depression.

A Global right-shifted histogram of gene synchrony in subjects affected with depression. **B** Comparing gene synchrony between control and depressed subjects confirmed that genes are similarly regulated in both groups (Gray dots) and identified numerous genes with robust decrease (Blue dots: High R in controls, low R in depression) or increase (Red dots: low R in

controls, high R in depression) in amygdala-cingulate synchrony. **C** Single gene examples. **D** The top genes and molecule interaction networks built on genes selected with increased (right) or decreased (left) amygdala-cingulate synchrony in depression share similarities in signal transduction components and were linked through insulin, a homeostatic modulator with significant links to both networks (**Table 1**). IL1 was significantly connected to the network build in genes with increased regional synchrony in depression. Grey, depression-affected genes; White, genes or bioactive molecules significantly connected to the network.

Table 1. Top 3 biological networks formed by genes with gain of loss of amygdala-cingulate gene synchrony.

Net-work	Genes & Molecules in Network	Score	Focus Genes	Canonical pathways	Major function	Disease
HIGH AMY-ACC gene transcript synchrony in MDD (GAIN OF FUNCTION)						
1	Akt, **ALCA**M, Ap1, Calmodulin, **CCNB1, CD59, CDC42, CREM**, ERK, **ETV5**, Histone h3, Hsp90, **ICK, IL1, INSIG1, INSULIN**, Jnk, **KLF6, MAGI1**, Mapk, **MAPK8**, Nfat, NFkB (complex), P38 MAPK, PDGF BB, **PPARD, PSMB4, RAD23A**, Ras, **SAT1, SCD, SCN2A, SQSTM1, TAC1, UXS1**	40	20	**Glucocorticoid receptor signaling**; B-cell receptor signaling; PPAR signaling; Xenobiotic metabolism; interleukin signaling	Protein kinase cascade; intracellular signaling	Genetic disorder
2	**AACS, ABCB10, AMACR**, C14ORF106, CABC1, CBLL1, **CCDC106**, CDKL3, CREBL2, **CRYZ**, FAM62A, **FBXO9**, HNF4A, **IER5**, LRRC8C, **MRPL18**, MRPS15, MRPS27, **NDFIP1**, PLA2G12B, **RGNEF**, RNASE4, **RPS18, SCG2**, SH3BGRL2, SREBF1, STK11, TMEM87B, TP53, TPRKB, **TRIM4**, TUBE1, **WDR77**, YWHAZ, ZNF175	27	14		Mitochondrion, intracellular organelle, cellular metabolism, ribosome	Neurolo-gical disorder
3	ALDH1B1, **ATF7IP2**, **BETA-ESTRADIOL**, CNN2, CORO1C, CTNNA2, CTNNB1, DSC3, F2, **GLRB**, GPR137B, GRB2, IFNB1, IFRD2, IKBKE, LAGE3, LOC284230, LRRFIP2, **MAOB**, MYC, **PDHB, PFDN5, PKP4, PUM1, RPL35**, RPL21, SLC11A1, SLC16A5, SP1, **SPTLC1, TAGLN3, TGOLN2, UBAP2L**, VCL, **ZXDB**	24	13		cell-cell junction, cell adhesion, actin-binding	-
LOW AMY-ACC gene transcript synchrony in MDD (LOSS OF NORMAL CONTROL FUNCTION)						
1	**ABCD3**, Actin, Calcineurin protein(s), Ck2, **CREB1, DBN1**, ERK, ERK1/2, FSH, **FTL, HLA-E, INSULIN, ITCH**, LDL, **MAP2K5**, Mapk, **MAX**, MIRN124, **NCAM1, NFE2L2**, NFkB (complex), P38 MAPK, **PCGF2**, PDGF BB, **PHLPP**, Pkc(s), Proteasome, **RFC1, SAT1, SERPINE1, SLC11A2, SLC30A7, TGFBR2, TIPARP**, tyrosine kinase	41	19	**Glucocorticoid receptor signaling**; Xenobiotic metabolism, NRF2-mediated oxidative stress response; B-Cell receptor signaling; PPAR signaling	Protein kinase; intra-cellular signaling; Phosphoprotein	-
2	**ABHD5**, Akt, **BETA-ESTRADIOL**, **CTNNA1**, CTNNB1, **DDX21, GPNMB, HNRPDL, HR**, IL4, $IL13_{GC}$, **JAKMIP2**, MAP3K1, MIRNLET7G, MYC, **PDK1, PNN**, RELA, **SAR1B, SLC26A2**, SMAD3, **SMCR7L, SOCS4**, SPRR2B, SPRR2G, **STAG2, TFCP2**, TGFB1, **THYROID HORMONE, TNS1, TUBB6**, YWHAG, ZFP161, **ZMYM2, ZNF337**	40	19	**Glucocorticoid receptor signaling**	Regulation of cellular process: transcription, metabolism, cell growth	-
3	ADSS, APEH, **ARMC6**, C11ORF58, CRIPT, DLG4, **EML1, GGCX**, GRID2, GTPBP3, GYS2, HLA-B, HNF4A, HSPC152, **LILRB3, LRRC1**, LRRC40, MPP1, **NLGN4X**, OGDH, **PAPOLA, PPP1R10**, PPP1R11, PPP1R3D, PPP2R5B, PTPRG, **RIOK2**, SEC11A, **SEPT2, SGCE**, SLC25A1, SRP68, **STYXL1**, TPP2, **WDR42A**	25	13		Protein phosphatase activity; Cytoskeleton	-

In double-underline are biological modulators significantly connected to the network over a range of FDR's and unlikely to be selected at random ($p<0.01$). Depression-affected genes are in bold. Other included genes/molecules displayed significant interactions with Depression-affected genes in network. "Canonical pathways" contain genes linked to ≥25% of nodes in networks. AMY, amygdala; ACC, anterior cingulate cortex; MDD, major depressive disorder.

doi:10.1371/journal.pone.0009970.t001

Table 4 (Table 1) Top 3 biological networks formed by genes with gain of loss of amygdala-cingulate gene synchrony

. In green are biological modulators significantly connected to the network over a range of FDR's and unlikely to be selected at random ($p<0.01$). Depression-affected genes are in bold. Other included genes/molecules displayed significant interactions with Depression-affected genes in network. "Canonical pathways" contain genes linked to ≥25 % of nodes in networks. AMY, amygdala; ACC, anterior cingulate cortex; MDD, major depressive disorder.

4.0 PAPER #3: DIFFERENTIALLY EXPRESSED GENES IN DEPRESSION AND OTHER NEUROPSYCHIATRIC DISORDERS ARE DISTRIBUTED ON THE PERIPHERY OF RESILIENT GENE COEXPRESSION NETWORKS

Chris Gaiteri and Etienne Sibille*

Department of Psychiatry, Center for Neuroscience

3811 O'Hara Street, University of Pittsburgh, Pittsburgh, PA 15213

*To whom correspondence should be addressed. E-mail:sibilleel@upmc.edu.

Tel (412) 624-0804 Fax (412) 624-9910

Abstract

Gene networks built on coexpression links provide a novel perspective on complex molecular systems, which are only beginning to be investigated in neuropsychiatric disorders. Accordingly we tested whether the multi-system changes observed in depression may be reflected in strategic changes to gene network structure. In a related hypothesis, we searched for characteristic position and connectivity among differentially expressed genes that would be informative of disease processes. Using several depression-related human postmortem transcriptome datasets, we show (1) that genes assemble into small-world and scale-free networks in control subjects, (2) that this efficient network topology is largely resilient to changes in depressed subjects, and (3) that differentially expressed genes are positioned on the perimeter of coexpression networks. Similar results were observed in a mouse model of depression, and also in bipolar- and schizophrenia-related coexpression networks. Importantly, we show that baseline expression variability contributes to the propensity of genes to be network hubs and/or to be differentially expressed in disease. In summary, the small-world and scale-free properties of gene networks appear to constrain the extent to which a gene may be differentially expressed in depression and other neuropsychiatric disorders. Previous studies in large-scale networks often reported a centrality-lethality relationship, in which pathology is mediated by hub nodes, so the diffuse peripheral localization of disease-related genes observed here may be specific to neuropsychiatric disorders, as it was conserved across several neuropsychiatric disorders.

Introduction

The pathophysiology of major depression is hypothesized to involve dysregulation at the levels of genes, neurons and brain regions (Belmaker *et al*, 2008; Sibille *et al*, 2009), but the size, complexity, and interaction of the affected systems make it difficult to determine the disease impact of single modulators. Indeed, it appears that the disorder is the cumulative result of multi-system disturbances (Gaiteri *et al*, 2010), which moves the brain function into a pathological dynamical state. Gene microarrays have the potential to overcome this complexity by simultaneously measuring levels of many different gene transcripts. However, this larger window into cellular activity has not always led to more consistent results, as different laboratories, brain regions, and model systems implicate a divergent set of pathological mechanisms in depression (Mehta *et al*, 2010).

Here, we propose using coexpression-based gene networks, which encompass many types of molecular interactions, as a contextual biological framework that may highlight common features of suspected disease genes in depression and other neuropsychiatric disorders. In coexpression networks, the links between genes (nodes) are determined by the extent of their correlated pattern of expression across multiple samples (measured by Pearson correlation; see methods) and are thought to result from a variety of biological relationships between genes, including common transcription factors or adjacent genome position (Allocco *et al*, 2004; Marco *et al*, 2009; Purmann *et al*, 2007). Gene networks derived from different tissues and species consistently show stereotypical “small-world” and “scale-free” network architecture (Carlson *et al*, 2006; Oldham *et al*, 2008). In small-world networks, nodes (genes) are typically strongly clustered into local communities that support biological sub-processes (Lee *et al*, 2004). The connectivity distribution in scale-free networks is highly heterogeneous: most genes are “provincial”, with only a few connections, while rare “hub” genes provide efficient global connectivity by linking together many distant genes.

Networks with these structural characteristics are robust to the deletion of random

nodes, but critically sensitive to disrupted function or “attacks” targeted at the rare hub nodes (Albert *et al*, 2000). Accordingly, at the level of molecular interactions, there are numerous demonstrations of close relationships between network connectivity and disease activity (Barrenas *et al*, 2009; Feldman *et al*, 2008; Jeong *et al*, 2001). Specifically, pathology-related targets may occupy strategic positions within these networks, which are poised to interrupt normal cellular function (Yanashima *et al*, 2009; Zotenko *et al*, 2008). Hub-nodes may be intrinsically disease-targeted, or merely frequently associated with disease, due to their far-flung connections (Goh *et al*, 2007; He *et al*, 2006). In either case, the collusion of network structure and pathology are important to understanding the relevance and priority of disease-related changes.

Therefore, to understand how networks of molecular interactions may broadly direct transcription changes in depression, we investigated gene coexpression network structure in control and depressed subjects using postmortem transcriptome datasets. Based on the prevalence of the lethality-centrality relationship in disease-related networks, we hypothesized that network changes between control and disease-state networks would be centered around coexpression hubs, Second we hypothesized that differentially expressed genes would have a characteristic position and connectivity level in those networks. Mathematically, the first hypothesis tests the relationship of differential *co*expression to network connectivity, while the second hypothesis tests the relationship of differential expression to network connectivity. These “hybrid” expression-and-network hypotheses were tested using postmorterm microarray datasets from depressed and healthy subjects. Because we seek to establish general principles of differential expression, we show the findings are consistent in an animal model of depression and applicable to a broader class of neuropsychiatric disorders, by including schizophrenia and bipolar disorder array datasets.

While the basic structure of these networks is small-world and scale-free, connectivity changes in disease are not targeted at network hubs – the network topology is surprisingly

resilient effects of depression. Relatedly, tests of our second hypothesis show that differentially-expressed genes are positioned on the perimeter of the coexpression network – relatively far from the critical network core. To understand what may mediate these effects, we explore the relationship of variability in gene transcript expression levels to network connectivity. Based on these results, we propose a general model of altered transcription in neuropsychiatric disorders and speculate on how the diffuse and peripheral localization of disease-related genes may relate to the difficulty in finding consistently differentially expressed genes across studies.

RESULTS

Small-world gene network structure in human and mouse brain transcriptome datasets

In unweighted coexpression networks, each node is a single gene/probe-set and each link represents a correlation between gene expression profiles above some threshold. These networks were generated with established techniques to minimize false-positive links (under 1%), that optimize the threshold for maximal biological information (see Figure S2 and supplemental information). As expected based on many previous studies (Bergmann *et al*, 2004; Jordan *et al*, 2004; Tsaparas *et al*, 2006; van Noort *et al*, 2004), we find that all datasets examined here have approximately scale-free and small-world characteristics (Table S1). These characteristics of the "consensus" network generated from both control and disease-state samples are illustrated in Figure 1 using gene expression from the amygdala of human postmortem subjects. First, the distribution of gene links (degree) fits a power-law on log-log scale (R=0.82), indicating an approximately scale-free connectivity distribution (Figure 1A). Second, the distribution of path-lengths (i.e. number of links) between any two genes is similar to that of a randomized network (Figure 1B). Third, genes were clustered into local communities with a high number of mutual interconnections (Figure 1C), compared to a randomized network with identical degree distribution and number of links. The combination of high clustering (Figure 1C) and low average pathlength (Figure 1B) is the hallmark of small-world networks. These

networks also showed uniformly positive *assortativity,* meaning that hubs were likely connected to other hubs, and provincial nodes to other provincial nodes (Table S1). A schematic of network structure that incorporates all of these characteristics is presented in Figure 1D.

Resilient small-world gene network structure in major depression

Since the basic combined structure of both control and disease coexpression networks is scale-free and small-world, it is possible that disease genes propagate pathological activity by altering the connections of influential hub genes - a concept that has been validated in other biological systems (Bullmore *et al*, 2009). Our specific test for this is to compare gain or loss of connections (gain or loss of correlation-based links going from control to disease-state networks) for hub genes compared to provincial genes (Figure 2A). Technically, this means contrasting the amount of differential coexpression for hubs vs provincial nodes. This comparison (Figure 2B,C) shows that hub connections are not preferentially disrupted in the example of the amygdala network. In fact, hubs experience less average rewiring between control and depressed states than do provincial nodes. Also, the total number of differential connectivity between control and depressed networks lies within the disease permutated bounds (Figure 2B), indicating that apparent connectivity changes in disease are indistinguishable from variability in the Pearson correlations (Figure 2C).

We find no evidence of hub targeting in any dataset (using $p<.05$ as the criterion for hub-targeting in all cases) including gene networks derived from an animal model of depression and bipolar and schizophrenia datasets. These results demonstrate that the pathology of depression is not created through differential connectivity targeted at hub nodes (at least in the current datasets using this methodology). While there are hundreds of ways in which control and disease-state networks may differ, in this critical measure (hub link targeting) coexpession networks are resilient to changes in human depression.

A hybrid approach linking differentially-expressed genes with coexpression networks

To address our second hypothesis relating to connectivity characteristics of differentially expressed genes within coexpression networks, we use a constant network structure derived from the combined control and disease samples, supported by observation of similar connectivity in both conditions. We observed a robust trend between differential expression and connectivity, wherein genes with the lowest p-values for differential expression display very low connectivity, while genes with non-significant disease effect are progressively more connected. Evidence for this relationship (Figure 3A) was generated by (1) sorting p-values for differential expression from least to greatest numeric value, (2) binning the ordered p-values into 100 groups/percentiles, and (3) computing the mean connectivity of each p-value group/percentile. We use this binned percentile method because it facilitates comparisons across datasets, which each has a unique distribution of p-values. Significance of connectivity was tested by repeatedly selecting an equivalent number of genes at random from that particular dataset to generate a distribution of expected connectivity. Performing all analyses using only control samples to generate the network structure does not significantly alter the trends (not shown), but since it does lead to higher false-discovery rates on networks (illustrated in supplementary Figure S2) we used the consensus network generated by combining control and disease samples.

The consensus relationship of connectivity to differential expression across all datasets (defined here as meta-connectivity; estimated by combining the p-values for under- or over-connectivity for each percentile across all datasets, see methods) shown in Figure 3F indicates a strong and stereotypical transcription response to disease that is closely related to gene network structure. The collective analysis (Figure 3F) also suggests a stereotypical connectivity for the entire continuum of differentially expressed genes, with the most consistent finding across datasets being low connectivity for low p-values DE genes. This was very consistent for the top 10% of differentially expressed genes across datasets even when controlling for FDR (10%). Greater than expected connectivity is also observed for some moderately differentially expressed genes, but in a less consistent manner (see fewer number of significant points at

middle percentiles in Figure 3F), due to lack of this broad trend in specific datasets, for instance Figure 3A. Similarly, there is an inconsistent drop in connectivity for ultra-high p-values, but this is not supported across datasets, unlike the major finding of robust low connectivity of low p-value genes trend.

The meta-analysis results in Figures 3F shows that the connectivity of disease genes is highly non-random, but we also investigated the broader question of whether these trends were specific to disease. To answer this question, we evaluate the same p-value to connectivity trend for pseudo-groups, each containing 50/50 combinations of control and disease samples (Figure 4A). The range of expected results from this permutation testing (Figure 4B) shows that the non-random connectivity of the true control-disease comparisons is actually not disease-specific, since the “true” or “observed” connectivity trend lies entirely between the permutation bounds. Furthermore, the control-disease comparison show a differential-expression to connectivity relationship that is very similar to the mean relationship of all permutations, indicating that while the real differential-expression to connectivity relationship is non-random, it is not unexpected or disease-specific. This additional test for disease-specificity was not conducted by the only directly comparable study (Lu *et al*, 2007), but indicates that some underlying trend beyond disease effect (which is negated by the permutations) must be creating the special connectivity of differentially expressed genes (see next section).

Baseline expression variability contributes to the differential-expression to connectivity relationship

What could account for the generic relationship of DE genes to connectivity (the red line in Figure 4B)? Meta-analysis comparing baseline variance in gene transcript level to connectivity shows a very strong relationship across all datasets (Figure 5). This indicates that hubs commonly show relatively large swings in transcript level (they are high-variance), with a mean/median correlation of 0.85/0.86 between variability and connectivity across all datasets.

Since hubs show more variable expression levels, this presents a simple explanation for the low connectivity of low p-value (for differential expression) genes since: (1) low variance genes are less connected and (2) low p-values tend are more frequently generated by low-variance genes, therefore low-p-values are associated with low connectivity. Any split of the data (permutation testing) will show this association of low p-values for differential expression with low connectivity, since transcript variance is unaffected by permutation. Note that transcript variability is factored out in the generation of Pearson correlations. So while this connectivity-variability trend exists, it is not circular, but appears to be a natural characteristic of gene networks. Thus, the special connectivity properties of disease genes appear to be fueled by the broader trend of connectivity increasing with variability. Hence taking into consideration the broader context of variability/connectivity with transcriptional programs may lead to uncovering putative disease genes that are closer to the core neuropathology.

DISCUSSION

Resilience of small-world gene network structure to neuropsychiatric diseases

When genetic variants and environmental influences combine to create disease pathology, they utilize and interact with cellular and molecular networks. We showed here that the coexpression networks of brain regions implicated in depression and other neuropsychiatric disorders display small-world and scale-free characteristics. These network architectures are an efficient (low path-length) and well-organized (highly clustered) framework for transcriptional activation. This efficiency comes with a specific weakness – vulnerability attack on key hub nodes (Albert *et al*, 2000), as demonstrated by disease operation in other large-scale networks (Guye *et al*, 2010; Micheloyannis *et al*, 2006; Smit *et al*, 2008; Srinivas *et al*, 2007; Stam *et al*, 2007; van Nas *et al*, 2009). Therefore, we speculated that the connectivity of disease-affected genes could offer a window into pathological mechanisms in neuropsychiatric disorders. However, we found that **the small-world connectivity characteristics of coexpression**

networks are in fact resilient to the effects of depression and other neuropsychiatric disease states, and that the related pathology is not mediated by network disintegration via attack on hub nodes.

Two related studies (Lu *et al*, 2007; Torkamani *et al*, 2010) also observed a broad similarity in the structure of control and schizophrenia gene networks or control and asthmatic networks, indicating that coexpression structure may be unaffected, or only slightly affected in complex disorders. While complete network reconfiguration and targeted destruction of hub connections appears to be rare in postmortem brain networks, it may be more evident in smaller more dedicated local networks that operate on a short time-scale, such as those devoted to metabolism or immune function (Leonardson *et al*, 2010; Reverter *et al*, 2006). The lack of hub-targeting does not preclude existence of differential connectivity shown by more involved methods or if larger samples become available.

As an alternative mechanism of network-pathology interaction, we tested if differentially expressed genes had a characteristic connectivity level within these resilient gene networks. We show that differentially expressed genes in neuropsychiatric disorders tend to have very low connectivity and fall on the edges of the network**. This second form of network resilience to disease (i.e. differential expression of provincial nodes, but not central hubs)** is the opposite of the standard pathological mechanisms in small-world networks, but potentially consistent with the broad range of affected systems in neuropsychiatric disorders. The low connectivity of DE genes is consistent across various brain regions, species, neuropsychiatric diseases and array platforms. Such a diffuse disease signature may be characteristic of complex disorders (Lu *et al*, 2007), but this is unclear since previous studies did not include permutation testing for significance or exploration of the relevance of expression variance.

Why do differentially expressed gene have low connectivity?

Since DE genes in neuropsychiatric disorders have low connectivity, it is natural to ask

(1) What biological and statistical relationships could generate this situation? and (2) How can this knowledge improve selection of disease-associated genes in a network setting? We show that a strong variability-connectivity relationship (Figure 5) creates a situation in which genes detected as differentially expressed are generically low-connected (Figure 4). But there are several biological rationales for why DE genes are located on the edge of the network in these particular datasets. It could be that DE genes follow generic patterns of variation (see Figure 4B for example) due to high false discovery rates associated with depression microarrays. Alternately, if control/disease comparisons have produced an accurate representation of DE genes, they may indeed ride on top of normal patterns of variability, since individual genes have small pro-disease effects in complex diseases.

To determine if the low connectivity of DE genes is specific to complex diseases, a useful future experiment would be to calculate the connectivity of DE gene lists obtained from microarray datasets of severe disorders. To further explore the meaning of connectivity within neuropsychiatric datasets specifically, we checked if various classes of disease genes were associated with lower/higher connectivity (Figure S3). Surprisingly, given the prevalence of the lethality-centrality relationship in other systems, OMIM genes, genes associated with single-mutation disorders, and essential genes showed expected levels of connectivity in coexpression networks. Genes which were frequently differentially expressed across many conditions did show higher connectivity, which fits with our emphasis on intrinsic patterns of variability associating with differential expression.

Inferring mechanisms of pathology from differentially expressed gene connectivity

Regardless of why DE genes are located on the edge of the network, how does this knowledge influence our conceptualization of disease effects on cellular interaction networks? The decentralized nature of DE genes in coexpression networks (Figure 6) may contribute to the illusive nature of depression pathology and the high failure rate of putative anti-depressant drugs – which essentially attempt to influence a vast network from the edge (if they are directly

targeting DE genes). These results are consistent with the multifactorial nature of major depression, bipolar depression and schizophrenia, and, from a coexpression perspective, suggest that single gene modulators will have limited therapeutic effect. It may be discouraging that the disease signal follows generic patterns of network fluctuation, but by understanding patterns of molecular interactions, it may be possible to more effectively track and dismantle disease processes.

Centrality-lethality relationship in coexpression networks

The low connectivity of DE genes calls into question the applicability of the centrality-lethality framework to coexpression networks. The centrality-lethality relationship is exemplified in PPI networks and has gone on to permeate research in small-world and scale-free networks. From a theoretical perspective, it brings a coherent framework to far-flung and complex molecular networks; practically, hubs have been shown to be key components of the modular communities that are centered around them (Horvath *et al*, 2006; Wang *et al*, 2009). The consistent finding that small-world breakdown is a correlate of disease activity (Guye *et al*, 2010; Micheloyannis *et al*, 2006; Smit *et al*, 2008; Srinivas *et al*, 2007; Stam *et al*, 2007; van Nas *et al*, 2009) spurs hope that widely applicable rules for optimal function determine the health of a network, regardless of its scale or composition.

But based on our examination of the position of gene classes in gene coexpression networks (Figure S3), it appears that gene networks do not have the classic centrality-lethality relationship of PPI networks, perhaps because gene networks encompass a variety of biological relationships between molecules (versus physical protein binding). Coexpression relationships can be generated by a several cellular mechanisms, including transcription factor binding sites (Allocco *et al*, 2004; Marco *et al*, 2009), epigenetic regulation (Chen *et al*, 2005), chromosomal gene sequence (Ebisuya *et al*, 2008; Purmann *et al*, 2007), and potentially 3D chromosome configuration (Lieberman-Aiden *et al*, 2009) in addition to fluctuations in cell-type populations

(Oldham *et al*, 2008). Therefore, the definition of what constitutes a network link (physical binding vs correlation) may determine of the lethality-centrality relationship applies. If coexpression networks cannot be evaluated under the same paradigms as protein networks, what other methods can utilize network structure to guide disease-gene selection at the transcriptome/proteomic/metabolomic levels?

Moving forward with gene networks analysis in light of the position of DE genes

The distinctive lack of centrality of DE genes highlights a long-standing challenge in complex diseases: detecting biologically cohesive sets of genes that create a cumulative disease effect. We propose that coexpression links, which encapsulate many cellular relationships, can indicate collective dysfunction. For instance, coexpression links indicated that genes associated with depression in a cross-species analysis of depression were tightly bound together in glial and neuronal-growth related communities (Sibille *et al*, 2009). Specific modules of coexpressed genes may emerge to support specific biological functions, as indicated by a module of neurogenesis-related genes found specifically in the hippocampus (Oldham *et al*, 2008). These diffuse changes may be tied back to specific modulators by searching for transcription factors which link many DE genes, and are themselves dysregulated (Hudson *et al*, 2009). Each of these studies uses networks structure to detect multi-gene cellular functions. But until we understand how disease or environmental influences percolate through the structure of coexpression networks, it is difficult define consistent strategies to predict which genes are critical mediators of disease.

While we examine the association of disease activity with connectivity in numerous datasets, this is a limited representation of transcriptional programs under disease states. In postmortem data, microarray measurements are years removed from potentially key developmental or disease shifts in these networks. What remains in postmortem data is the network steady state – the maladaptive equilibrium of a system sustaining long-term disease

activity. A number of developments could improve detection of key disease modulators, including improved graphs statistics, causal inference, or network-network interactions. It is possible that all information necessary to identify disease genes is contained in current datasets, which simply await application of the optimal graph/network analysis. Indeed new graph statistics that are responsive to physiological changes continue to be developed, such as clique configurations (Volinia *et al*, 2010), differential clustering (Chia and Karuturi, 2010). But calculating higher-order graph statistics does not overcome uncertain causality or the possibility of deficits outside coexpression networks. Understanding how observed coexpression structure responds to known perturbations could inform efforts to trace disease network structure back to unknown deficits. However time-series microarrays from inducible knockout systems, or tightly controlled human populations are extremely rare (Leonardson *et al*, 2010). The transcriptome networks described here are one of many cellular interaction networks, which themselves intersect at multiple contact points through feedback loops that can cloud causal relationships (de la Fuente, 2010). Combining information from different types of biological networks would more accurately reflect the "true" connectivity of genes sets associated with complex disease (Wachi *et al*, 2005), and hopefully further define the structure of pathology (Sharan and Ideker, 2006).

Material and Methods

Study inclusion criteria

We included several postmortem microarray studies of sufficient sample size (see methods) (Aston *et al*, 2005; Iwamoto *et al*, 2004; Iwamoto and Kato, 2006; Sibille *et al*, 2004; Surget *et al*, 2009; Torrey *et al*, 2000). We also performed all analyses on schizophrenia and bipolar datasets found in the main depression studies to see if observed trends were specific to depression or relevant to multiple complex disorders. Mouse data from animals submitted to unpredictable chronic mild stress (UCMS), which develop a depressive-like syndrome, were

also included in the study (Surget *et al*, 2009).

Definition of differential connectivity

To check for evidence of hub-targeting of coexpression links in disease, we use the resampling-based confidence intervals illustrated in Supplemental Figure S2D to identify links that are differentially coexpressed in control and disease networks. If a link is in condition 'A' is greater than the optimized cutoff and the correlation falls below its associated lower 95 % confidence bound in condition 'B' (or the reverse situation for link creation) then we define it as differentially coexpressed. Significance of a particular number of altered links (the p-values in Figure 2) connected to a given node is assessed by permutating the control and disease arrays many times to create sets of networks with no disease effect and then repeating the above check to establish confidence intervals on the expected number of significantly altered links (similar to the process in (Choi et al. 2005)).

Significance of gene connectivity via sampling

The expected mean connectivity of a selected group of genes (for instance low p-value genes in an array experiment) can be accurately estimated through resampling even in degree-heterogeneous scale-free networks. By randomly selecting sets of genes of the same size, null distributions such as those in Figure 3D can provide confidence intervals on expected connectivity. Utilizing these limits it is possible to tell if a group of genes falls outside of the 95% expected range of the resamples, either as highly connected hubs or low-connected "provincial" nodes.

Permutation significance bounds

We repeatedly spliced the data into two pseudo "control" and "depressed" selections, each actually consisting of 50/50 mixture of control/depressed data. Thus any low p-values from

this pseudo-comparison are known to be due to chance or unknown demographic stratification characteristic and not disease effect. Repeated comparisons of degree versus p-value (or fold-change) in such mixed datasets generates an expected null range for the degree associated with each p-value segment. The 95% confidence intervals on expected degree for a given p-value segment are equivalent to the bounds that encompass 95% of the permutation values (see Figure 4B for examples).

Corrections for multiple testing and procedure for meta-analysis of connectivity patterns across datasets

To assess the level of connectivity at a particular DE level across data sets we combined individual p-values for under or over connectivity using the “inverse normal method”. This is more appropriate to this data than the common Fisher's method, as it equally weights high and low values and outputs a consensus p-value as opposed to specifically favoring low p-value results. This combined p-value estimation of under- or over-connectivity we term “meta-connectivity.” Significance of the meta-connectivity values was assessed using the Benjamini-Hochberg FDR. Because it is possible to segment datasets with different bin sizes, correction to the meta-connectivity values due to multiple testing varies based on number of segments into which genes are partitioned (alpha/#bins). Thus, it is possible to reduce the nominal FDR simply by a coarser estimation of the region of p-values with non-random connectivity. However, even with our fine-grained approach, there were clearly defined under- and over-connected regions at 10% FDR, which largely persisted at 1% FDR as well.

Acknowledgements

Supported by NIMH grants MH084060 and MH077159 (ES) and by a fellowship from the University of Pittsburgh Clinical and Translational Science (CG).

Figures and Tables Captions

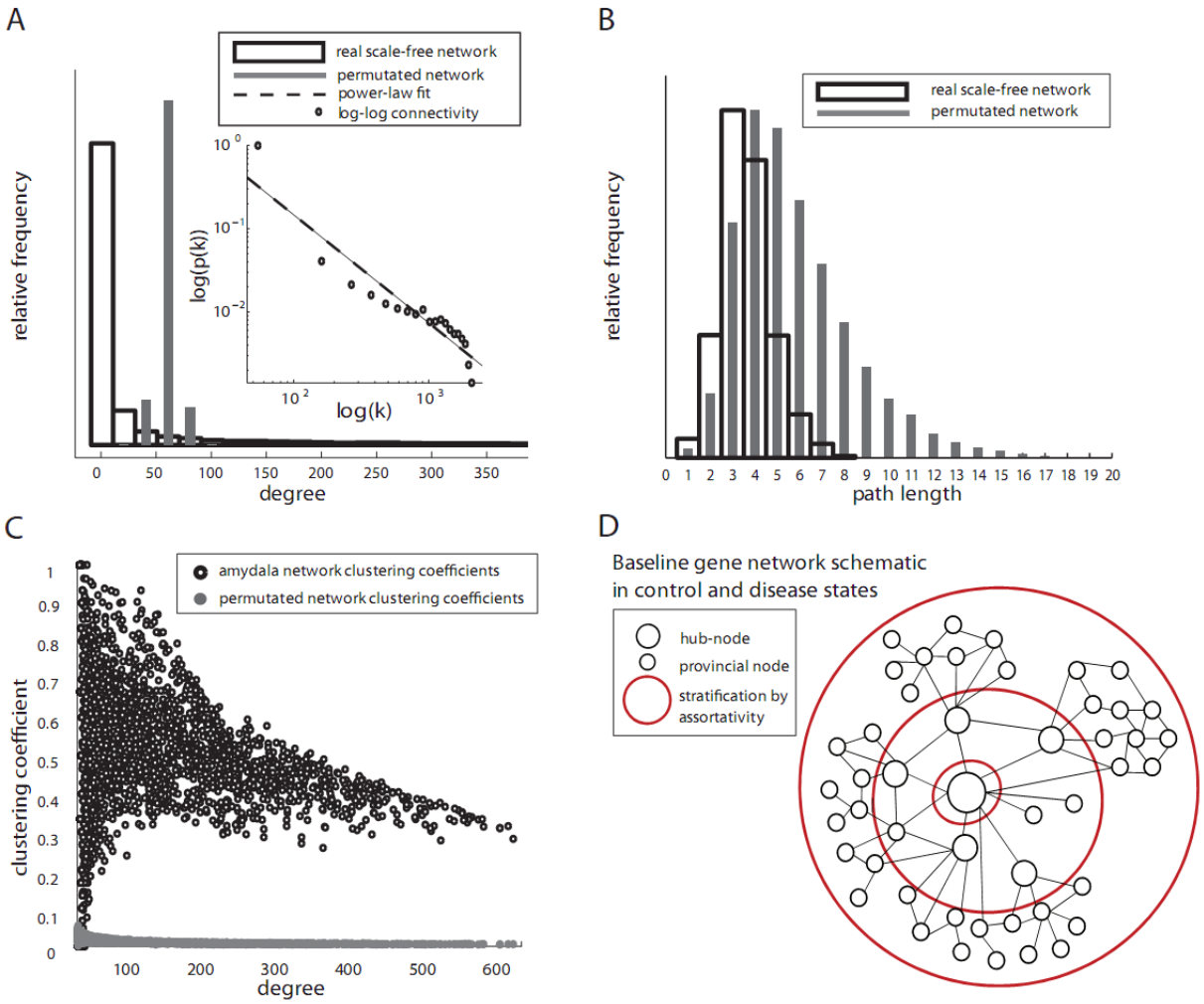


Figure 10 (Figure 1) Scale-free and small-world properties of gene networks common to all human brain datasets – examples from amygdala

(A) Histogram of frequency of connectivity values for exemplar amygdala network and randomized network with equal total number of links (truncated at k=400 for clarity), inset: power-law fit of full connectivity distribution (R=0.82) on log-log scale, indicating distribution is

approximately scale-free. The connectivity distribution does not follow an exact power-law, but regardless, the degree distribution is highly heterogeneous compared to the Gaussian degree distribution expected under random connectivity (gray bars). (B) Path length comparison of actual and randomized network indicates signal pathways through the network are extremely short (since random networks are a common benchmark for low path lengths). (C) Clustering coefficients by degree nodes for a segment of the real amygdala network compared to randomized network with identical degree distribution and number of links. (D) Network schematic of resilient network structure, which persists in both control and disease networks, showing existence of hubs, high clustering and positive assortativity.

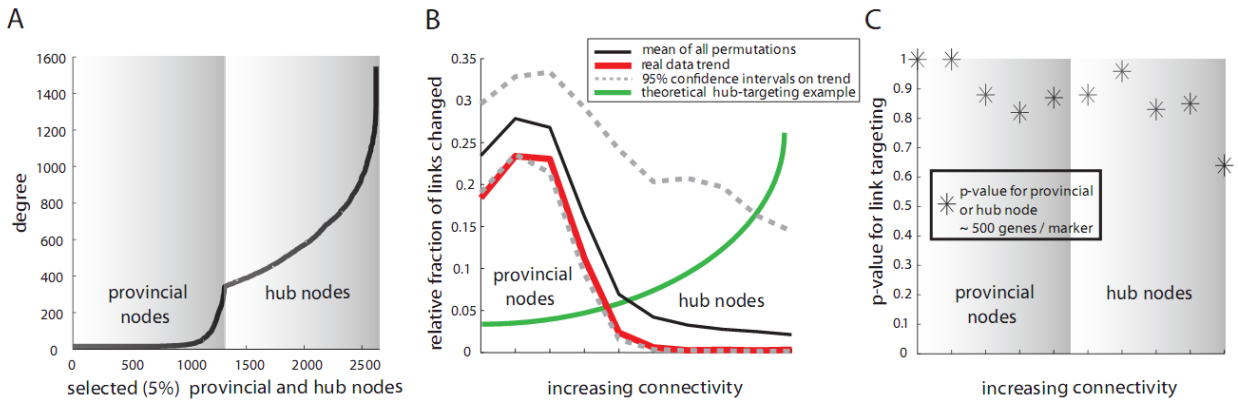


Figure 11 (Figure 2) Small-world network structure is maintained in post-mortem networks in disease states – example from human amygdala

(A) Connectivity of hubs (top 5% connectivity nodes) and equal number of provincial (non-hub) nodes for examination of targeted differential connectivity in disease. This example selection of two different types of genes will be used to illustrate that disease does not target hub connections. *Degree* is the graph theory term for number of network connections of a particular node. (B) For this selection of provincial and hub nodes, we compute the relative fraction of altered links (both created and destroyed) between control and depressed networks.

For comparison the same rewiring statistic is calculated for disease-permutated data (pseudo network comparisons with no disease effect) and these permutations establish the mean and expected confidence bounds on a real effect. While provincial nodes are generically more likely to show differential connectivity vs hub-nodes, this is not a disease effect, but rather due to the greater statistical stability of hub nodes with a large base of connections. Note that “real” network connectivity changes are within the expected bounds of variability. (C) p-values for greater than expected differential connectivity, that further quantify panel 2B, showing that connectivity changes in depression are not greater than expected by chance for both provincial and hub nodes.

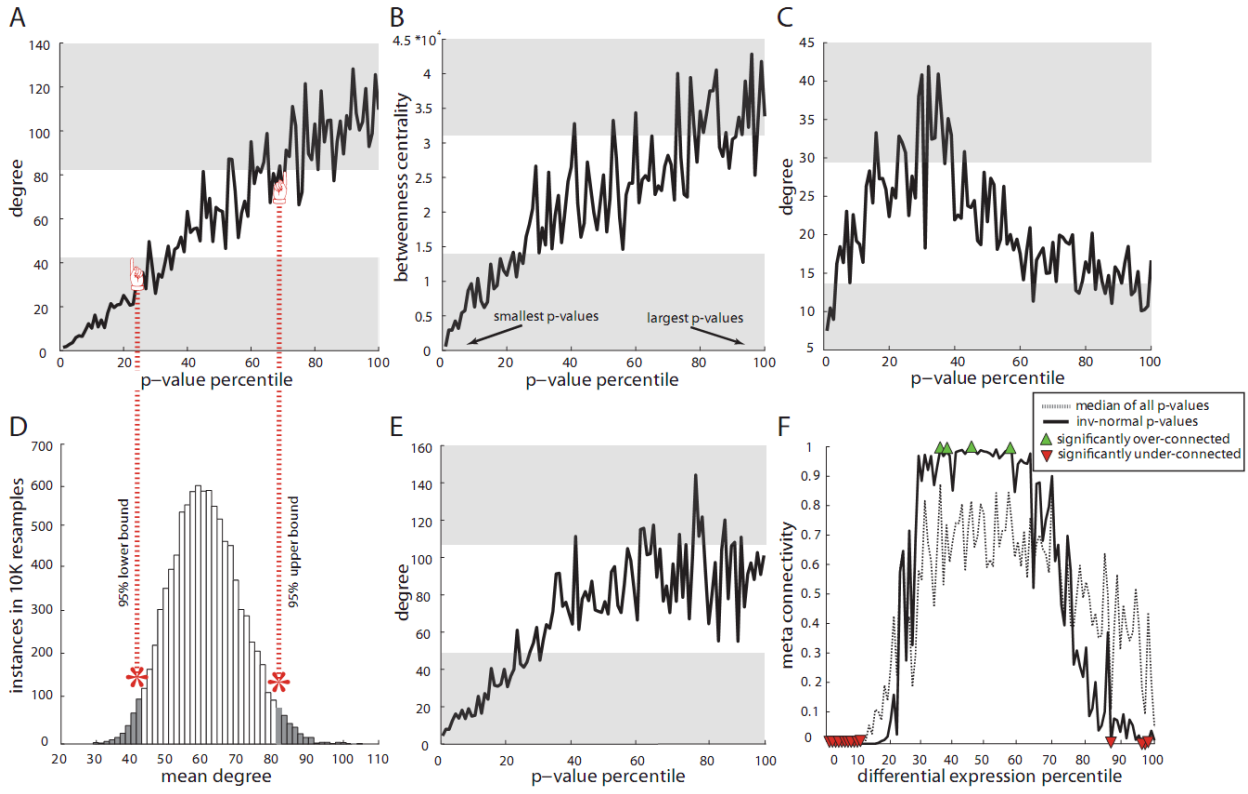


Figure 12 (Figure 3) Examples and meta-analysis of network characteristics stratified by disease effect size, using t-test p-values for differential expression

(A) Stereotypical trends in connectivity by p-value levels in human amygdala data, showing rising connectivity with rising p-values and particularly low connectivity for the most differentially expressed genes (in all figures lowest p-values are to the *left* and transparent gray area shows non-random connectivity values). (B) Betweenness centrality (a measure of how trafficked a particular node is by all shortest network paths) by p-values in amygdala indicates DE genes are not merely low-connected, but on the edge of the network because low p-value genes have the lowest betweenness centrality. (C) Connectivity by p-value levels in largest depression dataset (DLPFC, n=58) also shows stereotypical low p-value/low connectivity trend. (D) Example null connectivity distribution used to estimate expected range of connectivity – each network has its own specific null distribution used to estimate bounds on expected connectivity. (E) Similar low p-value/low connectivity trend as in depression, but in a different area (temporal cortex) and disease (schizophrenia). (F) Combined p-value by degree trends for all datasets (spanning species, disease and array platforms). Meta-connectivity measure (calculated with the “inverse-normal method” close to 0 indicate less connectivity than expected for that percentile of DE genes in all. Meta-connectivity measure close to 1 indicates greater than expected connectivity for that percentile. Note the additional power of meta-analysis is scarcely necessary as mean connectivity is itself highly significant as well. Percentiles with non-random connectivity were estimated at alpha=0.05 and 10% FDR.

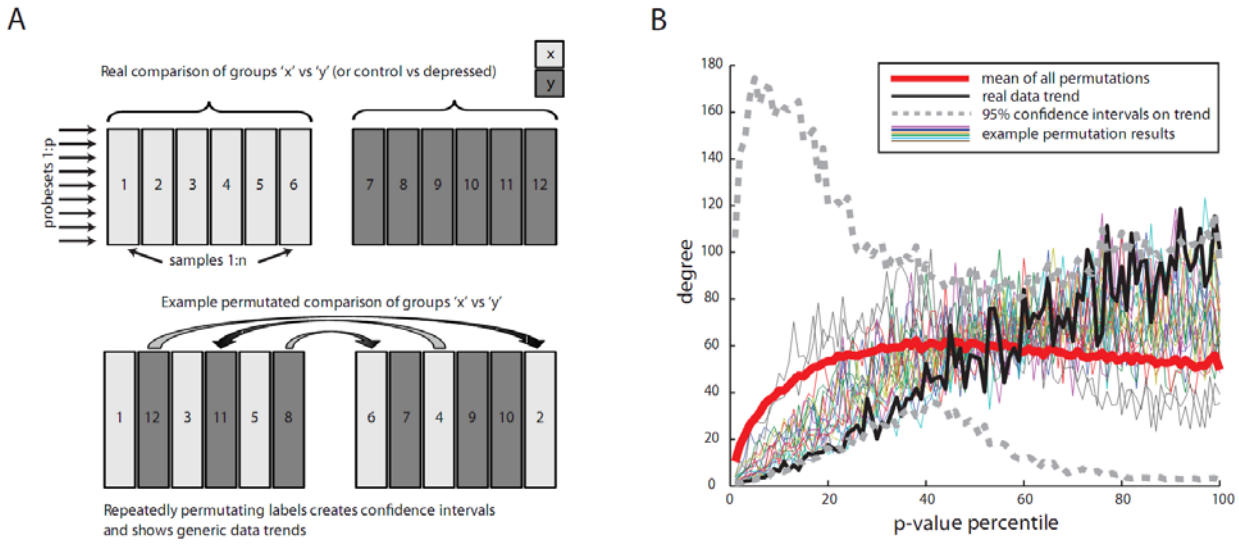


Figure 13 (Figure 4) Disease specificity of differentially expressed gene connectivity patterns using example of DLPFC (n=58) dataset

(A) Schematic of method of data permutation to generate null comparisons, which contain no disease effect since control/disease samples are balanced. (C) Stereotypical and actual trends in connectivity by p-value, showing that non-random connectivity (see Figure 3C) is a generic trend of all comparisons because the permutation bounds encompass the real comparison and because the mean permutation trend is similar to the actual trend.

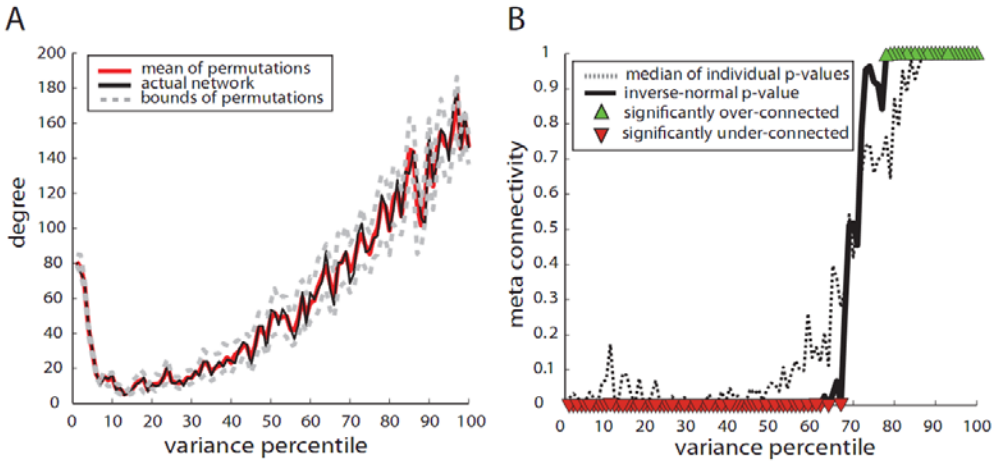


Figure 14 (Figure 5) Underlying network characteristics driving disease-connectivity effects

(A) Variability (transcript expression variance) plotted compared to degree - example from amygdala dataset (B) Meta-analysis of variability-connectivity relationship

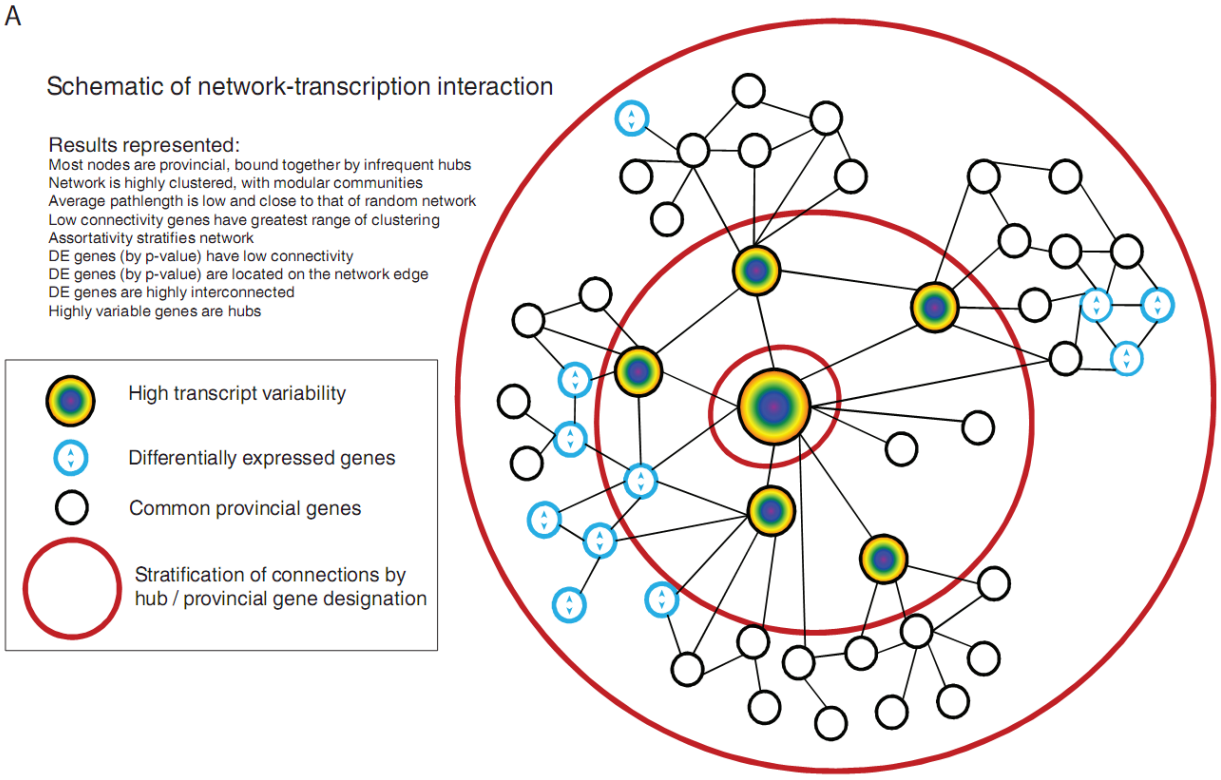


Figure 15 (Figure 6) Schematic of relationship between network structure and differential expression incorporating all results.

5.0 RESEARCH SUMMARY

Rationale for network investigations in major depression Depression is the most common neuropsychiatric disorder with massive human and economic impact that is not effectively managed by current antidepressant drugs (Chapter 1.1). Studies of the relationship of specific genes, neuronal subtypes, brain-regions to depression effects (Chapter 1.2) have a repeating pattern: after the initial finding that some molecular or cellular process is necessary for antidepressant efficacy, it is determined that there is no analogous deficit in naturally occurring depression. Even if some molecule or system is affected by depression, restoring it to normal levels is frequently insufficient to restore healthy behavior. The inadequacy of any single explanation to account for diverse depression pathology and behavioral effects may be a result of depression's multisystem composition. But, perhaps due to career pressure to hyper-specialize, or lack of a computational model of depression, most investigations focus on a single specific aspect of depression, despite the fact that such investigations themselves clearly indicate the insufficiency of single-factor theories. Even traditional microarray analysis of post-mortem depression data, which can sift through many biological processes, has been inconclusive, likely due to multiple pathological mechanisms that only have a faint transcriptome signature. At this point in the field of depression research, it is apparent a linear continuation of standard methodologies and techniques that investigate depression based on the contribution of independent factors is insufficient to deal with the true scope of a multi-system multi-scale disorder.

The novel methods and interdisciplinary questions that characterize studies in chapters 2-4 are driven by the scientific crisis surrounding depression research wherein hundreds of parallel investigations have produced a litany of unconnected disease effects, to the point that pharmaceutical companies are abandoning psychiatric research as a waste of time/investment (Miller, 2010). The inadequate treatment options for depressed patients, the research funding crisis in psychiatry, and the limited progress of single-factor hypotheses of depression all prompted us to directly confront the complexity of biological interactions that have largely prevented coherent understanding of depression pathology. To do this we conducted investigations focused on interactive and collective effects of multiple factors and systems, which is a hallmark of depression pathology. Specifically, we use postmortem transcriptome networks to answer: Is there a biological or molecular process that is consistently dysregulated in the cross-species depressed brain-state (Chapter 2)? What is the basis of altered brain-region communication seen in depression (Chapter 3)? Is there a basic coexpression network structure that mediates differential expressed genes in depression (Chapter 4)?

Philosophical links between studies Of course, these questions are of general interest to depression research; however, we approach them on multiple physical scales in parallel, and use network analysis to access higher-order transcriptome-wide representations of pathology. Thus, what binds these investigations together in their conception (beyond the specific findings, discussed later) is a willingness to go after depression pathology wherever it is found. The studies are not bound by a specific physical scale, but investigate brain dysfunction at the levels of gene interactions, cell-types, and brain regions. This willingness to investigate the full-spectrum of depression deficits in human post-mortem samples (as opposed to through several inaccurate mouse models) is crucial to success in a disease which does not have a characteristic neural process or biomarker. Furthermore, in each of these investigations, we go beyond common protocols of measuring one or more molecular markers and harness the full complexity of gene-gene correlations embedded in microarray data, to see into higher-order

network effects at work in depression. This approach contrasts with the historic trend in depression research to investigate single putative mechanisms in disease models (antidepressants applied to non-depressed mice) that have little to do with the naturally occurring pathology. The result of this unbiased, multi-scale, cross-species, network-centric approach to depression datasets implicates multiple pro-depressive mechanisms are at work simultaneously.

In addition to our specific findings (summarized in Chapters 5.1-5.3), based on the multi-system multi-factor nature of our findings, we advocate a new “floodgate” model of depression that emphasizes the potential for spreading regulatory failure across multiple systems as a pathological mechanism in depression (Chapter 5.4). The potential for cooperation among various biological hypotheses of depression is crucially important in the future success of a building a coherent depression pathology. It indicates that the key elements of understanding the disorder are in place, but what is required is a new integrative philosophy that focuses on the convergence of several mechanisms and how they could dysynergize to create the depressive state through a chain of regulatory/feedback failures.

5.1 PAPER #1 DISCUSSION: HOW DOES REDUCTION IN GLIAL AND NEURONAL FACTORS RELATE TO VARIOUS HYPOTHESES OF DEPRESSION?

Glia are non-neuronal cells of several classes that perform crucial functions related to glucose metabolism, neurotransmitter recycling, synaptic signaling, synaptic plasticity, and immune response (McNally *et al*, 2008). Because glia are involved in maintaining so many brain processes implicated in depression, they are well-situated to simultaneously mediate morphological and neurotransmission deficits hypothesized to underlie depression pathology. Glia cell death may account for the broad trends for prefrontal, orbitalfrontal, and cingulate

cortex hypotrophy seen in depression (Chana *et al*, 2003; Cotter *et al*, 2002; Rajkowska *et al*, 1999) and even though neurons in the amygdala ramify in depression, glial loss is also observed in that region as well (Bowley *et al*, 2002; Hamidi *et al*, 2004). When antidepressants stimulate neurogenesis in the hippocampus they also lead to gliogenesis (Santarelli *et al*, 2003) and thus glia may be partially responsible for antidepressant effects. In a rare instance of agreement among microarray results, oligodendrocyte abnormalities have been detected by different researchers in multiple brain regions (Aston *et al*, 2005; Sibille *et al*, 2009; Sokolov, 2007). The oligodendrocyte deficit may be particularly acute in late-life depression (Khundakar *et al*, 2009), which is characterized by white matter hyperintensities in T2 MR images that indicate altered vasculature and myelination (Nobuhara *et al*, 2006; Thomas *et al*, 2003). Peripheral markers of glial function also indicate that glial down-regulation is likely occurring as S100B is down-regulated in blood and CSF of unmedicated depressed patients (Kronenberg *et al*, 2009; Schroeter *et al*, 2010). (The S100B protein involved in calcium-mediated cellular growth (Santamaria-Kisiel *et al*, 2006) is neurotoxic at high concentrations and released by astrocytes and oligodendrocytes (Pinto *et al*, 2000).)

The sustained HPA activation and inflammation frequently observed in depressed patients offer mechanistic explanation for observed oligodendrocyte deficits. Cortisol has both direct and indirect effects on oligodendroctyes: in the direct mechanism cortisol binds to glucocorticoid receptors in the oligodendrocytes, in particular the NG2 oligodendrocyte precursors and prevents their maturation (Alonso, 2000; Schröter *et al*, 2009). But cortisol and synthetic steroids sometimes used to treat multiple sclerosis may preserve oligodendrocytes (Mann *et al*, 2008; Melcangi *et al*, 2000) by preventing cytokines from altering potassium channel density. Thus stress response and cortisol levels can exert complex direct control over oligodendrocyte populations.

The indirect path from cortisol release to oligodendrocyte death is through interaction with the increased inflammation seen in depression, interaction with other glia, and glutamate

excitoxicity. Cytokines from the periphery or those released in the brain lead to activation of microglia and increased levels of indolamine 2,3dioxygenase, which increases conversion of tryptophan to quinolinic acid, simultaneously decreasing production of serotonin and increasing activation of NMDA receptors. Not only do astrocytes then have more glutamate to scavenge, but quinolinic acid acts on them to decrease rates of glutamate uptake (Tavares *et al*, 2002). These conditions can complete a positive feedback loop wherein the excess glutamate levels lead to further microglia activation and release of TNF-α which sustains high inflammation levels, which are associated with glial death (Mann *et al*, 2008). The astrocyte glutamate transporters EEAT1/2 are down regulated in multiple sclerosis and their absence is most severe around cortical lesions (Vercellino *et al*, 2007). Therefore, glia have a certain tolerance for glutamate scavenging and the high amygdala activation and inflammation seen in depression may push them beyond capacity and lead to glial death.

Several clinical and antidepressant drug observations support the glutamate toxicity hypothesis. Chronic stress in mice leads to a decrease in oligodendrocyte density and using a gliotoxin to decrease glia numbers (in the prefrontal cortex) creates similar depressive-type behaviors (Banasr and Duman, 2008). The glia death from chronic stress appears to be glutamate-mediated as the NMDA antagonist Riluzole prevents loss of glia or generation of depressive behaviors (Banasr *et al*, 2010). Resting state connectivity (default mode connectivity) occurs on a timescale that reflects astrocyte contribution to signaling, and indeed several abnormalities in the anterior cingulate and prefrontal cortex default mode connectivity and task switching responses have been detected in depressed patients (Greicius *et al*, 2007; Sheline *et al*, 2009). Combining resting state connectivity with MRS, abnormalities in glutamate and GABA cycling were observed in depressed patients and severity of the deficit correlated with HAMD scores (Horn *et al*, 2010). Thus, an abundance of mechanisms link glial death, glutamate toxicity and depressive behavior.

Because several mechanisms leading to excitotoxicity could all be recruited by stress response and found in depression, there is a possibility for a perfect glutamate storm in depression. For instance, excess glutamate from prolonged HPA activation, decreased glutamate scavenging due to fewer and less effective astroglia, positive glutamate feedback onto microglia, decreased GABAergic tone and increased NMDA activation from quinolinic acid could all occur simultaneously to increase glutamate levels, which would lead to oligodendrocyte cell death through excitotoxicity (Takahashi *et al*, 2003). While the collusion of all of these mechanisms has not been shown in the context of a single experiment, the independent components are well-replicated by behavioral testing (Banasr *et al*, 2010), MR studies (Horn *et al*, 2010; Price *et al*, 2009), post-mortem microarrays (Choudary *et al*, 2005; Rajkowska *et al*, 1999) and peripheral assays(Petty *et al*, 1981). Therefore in determining the potential culpability of glia/oligodendrocytes in depressive processes, it is not merely the direct effects that are important, but the opportunity glial deficits provide for an excitotoxitic positive feedback loop. Despite all these potential mechanisms behind oligodendrocite death, the actual way in which this contributes to the symptoms of depression, has not yet been specified and probably relates to myelination associated with cell growth (see next).

The second module of “neuronal-related” genes that were differentially expressed across species in depression is likely related to region-specific connectivity changes in depression. The amygdala shows increased dendritic arborization in depressed patients, in contrast to the anterior cingulate cortex, hippocampus, and prefrontal cortex which all show hypotrophy under stress and depression conditions (Morales-Medina *et al*, 2009; Radley *et al*, 2004; Shansky *et al*, 2009). Concordant with amygdala hypertropy we found many differentially affected genes related to calcium regulation and cellular scaffolding, potentially mediating net dendritic growth in the amygdala. Even a single high dose of corticosterone in rats led to amygdala hypertrophy, probably as a function of calcium influx through NMDA and AMPA channels (Duvarci *et al*,

2007; Inglis *et al*, 2002). However, increased calcium influx and resultant neuroplasticity could be created through other mechanisms in naturalistic depression.

While the theory of simultaneous neuronal hypertrophy and glial death in the amygdala are consistent with increased HPA activity, in most other brain regions depression is associated with less ramified dendrites, probably mediated by extrasynaptic NMDA receptors (see below) and if that is the case, then the amygdala hypertrophy has a unique mechanism driving it. While the objective distinctions in plasticity and dendrites in various regions have been cataloged repeatedly, there is frustratingly little research concerning the exact mechanism responsible for these distinctions between cortisol's actions on different regions. It is likely some combination of the structure of feedback circuits and patterns of activation in response to stress, receptor densities, and input from other limbic structures. Again, while our results fit with the approximate story of amygdala reactivity undamped by the frontal cortex, it is far from a mechanistic explanation for exactly how that situation occurs. Currently, since there is no computational model of multi-region interaction and also scant biological justification for how these changes are implemented, we exist in a disturbing situation in which the field does not possess a global framework nor local landmarks that would serve to mechanistically couple these depression effects.

The scenario wherein NMDA is a component in a positive feedback loop of glutamate signaling levels, taken in conjunction with the decreased arborization seen in depression, appears to conflict with experiments showing that NMDA receptors stimulate synaptic plasticity, the antidepressant properties of SNRI's, and the elements of neuronal outgrowth we find upregulated in the amygdala. One possible explanation for the diversity of glutamate effects that is relevant to depression is as follows: low-dose ketamine would be expected to worsen glutamate excitoxicity and inflammation, as opposed to providing immediate relief, as it does in reality (Zarate 2006). However, synaptic and extrasynaptic NMDA channels have different binding affinities and activate different calcium cascades (Hardingham *et al*, 2002). The

extrasynaptic NR2b-containing receptors have higher affinity and are surrounded by lower glutamate concentrations, so ketamine likely acts on those receptors. While the exact mechanisms are still debated (Hardingham and Bading, 2010), extrasynaptic receptors are largely pro-apoptotic and decrease BDNF by decreasing CREB levels, while synaptic receptors are pro-survival. By stimulating synaptic NMDA receptors at a relatively higher level than the extrasynaptic receptors, it seems SNRI's shift the signaling balance towards neuroprotection. However, if extrasynaptic glutamate levels rise due to glial dysfunction then extrasynaptic NMDA receptors are more likely to be activated and decrease proliferation. Therefore, it is at least in theory possible to stimulate dendritic ramification in the amygdala, while generally promoting excitotoxicity in other regions and non-specific glial death. However, there has been no simultaneous examination of glutamate signaling and glial activity in the amygdala and other areas, which would be necessary to move the field beyond speculation about how these region-specific effects might be occurring.

Intrinsic, referred and collective brain region deficits in depression

In light of the core amygdala deficits found in this study, how can reports of depression deficits in other brain regions, or the joint activity of the amygdala with other regions, be interpreted consistently? For instance Chapter 2 shows cross-species oligodendrocyte and neuronal deficits in the amygdala, but not the cingulate cortex; Chapter 3 shows deficits in the communication between those regions; Chapter 4 shows consistent coexpression patterns of differentially expressed genes in both regions. Post-mortem microarrays show a mixture of causal and long-term adaptive changes to environmental and pathological influences, due to the single time-point assessment. Concordantly, expression changes may indicate a brain region generates a particular condition, or they may be activity-driven reactions to changes in other regions. If the changes are reactive, it may be the particular region is targeted by disease-affected regions (convergent or focal input) or selectively vulnerable (lacking some standard compensatory mechanism). Because activity percolates through neural systems, all of these

types of deficits may be present simultaneously, and may or may not be causally-linked. For instance the strong amygdala findings of Chapter 2 are mathematically consistent with the altered AMY-ACC results of Chapter 3, since changes in coordination do not require changes in expression level. However, it is possible that the amygdala is the source of altered coordination - that abnormal amygdala responses to emotional stimuli originate a pattern of communication that results in decorrelated gene expression. Distinguishing between these possibilities, or even the relative likelihood of either scenario, is not possible without a strong modeling framework (developed in Chapter 5.4).

5.2 PAPER #2 DISCUSSION: WHAT IS THE POTENTIAL FOR GENE COORDINATION AS A FUNCTIONAL MARKER OF NEURAL NETWORK ACTIVITY?

Patterns of brain region feedback in depressed patients correspond to the core clinical symptoms of inability to suppress negative thoughts and a tendency toward rumination on negative life events (Cooney *et al*, 2010; Johnson *et al*, 2009). Therefore, a primary research question is how to identify and validate the molecular mechanism behind these effects that encompass multiple brain regions. The influence of specific metabolites on brain function may be assessed through PET or MRS studies, but they require specific ligands and/or can only measure a limited number of metabolites. Alternately, any gene variant can be correlated with fMRI task responses, but those still require *apriori* selection of specific genes. Thus, if there were a gene-based proxy for interregional communication, it could be an effective way of linking abnormalities in brain region activity to potential targets behind those effects, without bias towards a small number of well-studied genes. Gene coordination (Chapter 3) utilizes inter-regional correlations in expression level as a transcriptome-wide filter to relate altered functional connectivity to putative gene markers. Changes to gene coordination in disease are less

specific than, for instance results of a PET study, because while gene coordination is likely driven by related patterns of neural activity, the physical cause of altered coordination is a black box. The list of gene-results from a gene coordination study cannot be traced back to a specific functional deficit or specific fMRI task, because samples come from post-mortem subjects and represent the steady-state relationship of two or more brain regions. But what gene coordination lacks in specificity, it compensates for, by relating combined brain function back to many gene markers in a manner not possible via more specific measures.

Because decreased feedback between the amygdala and anterior cingulate is one of the most replicated fMRI findings in depression (Drevets, 1999; Hooley *et al*, 2009; Matthews *et al*, 2008; Pezawas *et al*, 2005), the genes which also lose synchrony across regions could (A) represent the mechanisms responsible for the abnormal joint activity or (B) represent genes that are dependent on joint activity regimes for synchrony. For future studies it would be useful to clarify the nature of its relationship to regional network activity. This could be done by performing microarrays on many brain regions and then considering if the strongest gene correlation links correspond to strong functional or structural link between brain regions. This would be particularly useful for neuropsychiatric disorders, since interregional connectivity is studied in parallel with expression changes, but systematic understanding of how those two systems couple, is scarce. Regardless of whether gene coordination is causal or correlative with changes to regional activity patterns, it is a robust marker of interregional dysfunction that can easily be linked to specific genes.

Because a large number of genes show altered coordination between AMY and ACC, even at low FDR, we use IPA modulators to represent their concerted function. Two over-arching properties of the IPA-identified modulators behind the observed coordination changes are that (1) they consist of distinct hormonal and circulating factors that affect many brain regions, and (2) they correspond to many existing hypotheses of depression pathogenesis. The association of insulin, beta-estradiol and thyroid hormone with altered coordination suggests

that risk factors of sex and metabolic deficits are reflected in altered gene coordination. Specifically, diabetes increases risk of developing depression (Eaton 1996) and low thyroid hormone may cause depression, which is then directly treatable through artificial thyroid hormone supplements. While the sample in this study was completely male, estradiol's influence on many decoordinated genes could lead to even stronger effects in females.

The interaction of estradiol with many differentially coordinated genes hints that AMY-ACC feedback may be preferentially affected in females, which would fit with higher depression rates in females. There is some evidence that the psychological experience of depression may be different for females, in that they rank social perception as a greater source of depressive symptoms than do males, but exactly what structures and pathways mediate this is unclear (Scheibe 2003). Behavioral tests of depression in mice sometimes show stronger response to stress, but effects are highly test and strain dependent (Dalla 2009). The amygdala and ACC also meet the minimum criteria of response to estradiol levels. (Goldstein 2005). Furthermore, BOLD responses in functionally dimorphic areas including AMY and ACC are most similar to those of males when estrogen levels are lowest, indicating activity is actively regulated and not purely structurally programmed (Holsen 2010).

A separate set of likely partners in gene decoordination are IL-1 and glucocorticoid signaling, as these normally would have negative feedback on each other, but are both elevated in depression (see Chapter 1.2.2 for review of glucocorticoid and inflammatory hypotheses of depression). Since altered glucocorticoid signaling enables many other hypotheses of depression, and it is highly influential among genes with altered coordination, it is both a validation of the technique of gene coordination and an incitement of collective AMY-ACC activity in depression. Thus, the main modulators of altered coordination represent some of the strongest markers of depression, from classic causal factors, to risk factors, to putative mechanisms.

The modulators associated with altered coordination are not part of a single pathway, but involved in multiple hormonal systems. It could be that depressed sub-populations are defined by a single dysregulated modulator and collectively our data sample implicated these several modulators, however the small samples size here and requirements for statistical robustness, make the separate-and-equal modulator hypothesis unlikely. A more likely conclusion is that these results indicate a combination of disrupted circulating factors is necessary for depression. Mechanistic connections between the hypotheses of depression would certainly support the interoperability of multiple mechanisms in contributing to cortico-limbic dysregulation. As described in Chapter 1.2.5, one possible chain linking all of these hormones would be if lowered sensitivity of glucocorticoid signaling could lead to insufficient suppression of inflammation, which in turn could lead to higher glucocorticoid levels or increased reactivity to emotional stimuli. Increased inflammation and cortisol would lead to amygdala hypertrophy and prefrontal/cingulate hypotrophy, and the accompanying decrease in feedback on the amygdala would exacerbate emotional reactivity (Dantzer *et al*, 2008; McNally *et al*, 2008). However, this concerted activation of several pathways implicated in depression has not been validated or explored in a unified single-organism setting, because depression research is segmented into research cliques. While results from altered gene coordination support the idea of multiple regulatory dysfunctions leading to depression (further developed in Chapter 5.4 as the “floodgate model of depression”) this larger concept requires additional direct tests of combined multi-system influence.

Because gene coordination is a new marker of regional communication, that has not yet been artificially manipulated, there are several aspects of the measure that cloud mechanistic interpretation of results of AMY-ACC communication. For instance, what is the source of gene coordination? Since there is both gain and loss of synchrony among sets of genes in depression, it is possible that either (1) interregional AMY-ACC activity selects sets of genes to be coordinated or else (2) cellular conditions in each region interact with input from the other

region to determine which genes are coordinated. Therefore another source of ambiguity in results interpretation is that it may not be altered feedback activity itself that specifically creates altered coordination, but perhaps independent dysfunctions in each region which then lead to altered coordination. In this case the decoordinated genes would reflect the individual dysfunctions to the extent that they alter communication between the regions. Indeed, based on the numerous morphological changes to AMY and ACC, it appears unlikely that in depression the AMY and ACC have perfect internal operation and merely altered feedback(Drevets, 2003; Rajkowska and Miguel-Hidalgo, 2007; Sheline *et al*, 2001). However, altered feedback between depression-implicated brain regions (Cooney *et al*, 2010; Matthews *et al*, 2008), excess HPA activation (Pariante *et al*, 2008), metabolic defects(Marcus *et al*, 1992; Weber *et al*, 2000), and unsupressed inflammation (Dantzer *et al*, 2008; Smith, 1991) could interact and reinforce specific morphological or cellular deficits. Therefore, altered gene coordination reflects within-area deficits which result in, or result from, altered communication across regions. Thus there are several potential mechanisms behind alterations in gene coordination in depression. While the biological mechanism behind these changes is unclear, they do offer a transcriptome-wide representation of brain-region communications, so the method is well-suited to detecting the combinatorial brain-region dysfunction thought to characterize depressive states.

5.3 PAPER #3 DISCUSSION: UNDERSTANDING BIOLOGICAL FUNCTION THROUGH NETWORK STRUCTURE

The centrality-lethality “rule” permeates research in small-world and scale-free networks. From a theoretical perspective, it brings a coherent framework to far-flung and complex molecular networks; practically, hubs have been shown to be key components of the modular communities that are centered around them (Horvath *et al*, 2006; Wang *et al*, 2009). The

consistent finding that small-world breakdown is a correlate of disease activity (Guye *et al*, 2010; Micheloyannis *et al*, 2006; Smit *et al*, 2008; Srinivas *et al*, 2007; Stam *et al*, 2007; van Nas *et al*, 2009) spurs hope that widely applicable rules for optimal function determine the health of a network, regardless of its scale or contents. The need for coherence and understanding of concerted action of hundreds of genes and multiple neural systems is acute in neuropsychiatric research, because unlike cancer research, there is not even a consistent list of differentially expressed genes and the significance of individual SNPs is quite low and debatable (Bosker *et al*). Thus, network analysis in neuropsychiatry faces the dual challenge of identifying core dysregulated processes, and understanding how subtle far-flung changes in expression relate to disease phenotypes.

To understand how coexpression networks structure may reflect the impact of neuropsychiatric disorders, we measure the centrality of genes that are differentially expressed and differentially coexpressed in depression. Since post-mortem sample sizes are small, we compiled results across all available high-quality depression datasets. These include data from multiple regions, so the results are indicative of any general network pathology mechanisms spanning or targeting multiple regions. Since we include data from multiple regions in a mouse model of depression (Surget *et al*, 2009), these results benefit from the constant genetic background and identical depression induction, in contrast to humans which have unique combinations of genetic background and environmental influences that precipitate depression, which also manifests with different physical symptoms. Furthermore, we also include samples from schizophrenia and bipolar populations, to improve the applicability of results to a range of neuropsychiatric disorders. As described in detail in Chapter 4, we observe a robust trend across all datasets for differentially expressed genes to be much less connected than expected at random. We find that is probably fueled by variability trends in the data, which was not clear in the only previous study (Lu *et al*, 2007), but the fact remains that genes thought to be key mediators of disease are in fact very peripheral in the coexpression network. From a pure

methodological standpoint, this indicates that data from psychiatric disorders pose more substantial challenges to network analysis than have the previous cancer-based datasets, because the typical method of searching for disease-associated hubs may not apply, or may need to be adapted in this situation.

Simultaneous investigations of differential expression and network connectivity are relatively rare (Fuller *et al*, 2007; Lu *et al*, 2007; van Nas *et al*, 2009). It is however, these studies do find specific modules of genes that show consistent up or down-regulation in response to disease, but the general connectivity level of differentially expressed gene is not specified. With exceptions, (van Nas *et al*, 2009) usually the relationship is described with a scatter plot with little or no quantification or significant calculation on the trends. Because there are few examples of this type of analysis, it is difficult to estimate the specificity of this peripheral impact to neuropsychiatric disorders. There does appear to be a correlation between the severity of genetic impact of a condition and the centrality of the impact: datasets with major effects (cancer/sex) show differential expression in well-connected genes, while complex disease datasets (depression/schizophrenia/asthma) show differential expression among low-connected genes.

Thus our findings are not anti-modularity or anti-hub in philosophy, but may represent the flipside of the centrality-lethality relationship. Because deficits in high centrality nodes are very deleterious, and since depression is non-lethal and does not affect the potential of individuals to perform survival tasks, expression changes are concordantly on the edge of the network. A contributing factor to the results may be that post-mortem depression microarrays are inconsistent with each other and near the effect of noise (positive interpretation) or potentially littered with false positives (negative interpretation). However, asthma microarrays lack the highly contentious and discordant nature of depression microarrays, and differentially expressed genes in those networks are also low connected (Lu *et al*, 2007), so it seems that

disease properties, and not simply connectivity-variability relationships are driving the position of differentially expressed genes.

Regardless of how differentially expressed genes come to be positioned on the edge of gene networks, this presents multiple challenges to the routine of selecting hubs which connect many differentially expressed genes. Similarly, if there is differential connectivity between control and depressed states, it is not occurring as relatively easy-to-detect changes in hub connectivity, but rather as fluctuations in connectivity among many low-connected genes. In addition to the technical questions prompted by the position of differentially expressed genes, there is also the larger biological question about how these expression patterns relate to difficulty in finding consistently dysregulated molecular functions in depression.

Simply based on the volume of cases in which disease states correspond to targeted attack on small-world/scale-free network hubs, depression microarrays might be thought to continue in this vein, illustrating a universal disease mechanism. Because results indicate the opposite of this process is occurring – that changes are largely peripheral to the influential coexpression network hubs - this may be the network manifestation of a decentralized multi-system/multi-module dysfunction in depression. The diffuse coexpression impact could be both a symptom and a cause of disagreement over the specific genes and pathways detected by array in depression. If indeed there is no unitary core component of depression which is either differentially connected or differentially expressed, then the frequent disagreement over the microarray correlates of depression have a basis in the disease impact, which is very faint, highly distributed, and skirts around network hubs. From the perspective of what biological processes may be involved in depression, these results support our contention that depression cannot be studied as the result of a single mechanism but likely reflects several simultaneous mechanisms that may be collectively destabilized (Chapters 1.2.8 and 5.4). Thus, in addition to answering specific questions about the connectivity of differentially expressed genes, based on results from hundreds of arrays in multiple species, brain areas and neuropsychiatric disorders,

this study supports a more collective integrative, multi-mechanism approach to research on depression pathology.

If the network representation of disease changes is highly distributed along the edge of the network, practically how then should coexpression studies of depression be conducted? As indicated in Chapter 2, cross-species analysis may select a set of differentially expressed genes of sufficiently high statistical quality, that they do form modular functions in coexpression networks. Overlaying additional sources of biological information into coexpression networks, such as transcription factor networks, or GWAS candidate genes may also link several differentially expressed genes into detectable disease-related communities. If robust differentially connectivity is found in larger datasets, even if it is not targeted at hub nodes, linking differential connectivity to differential expression could prioritize selection of genes closer to primary depression deficits, which is currently difficult to do in the acausal coexpression framework. These possibilities would all benefit from a more mechanistic understanding of how coexpression communities arise from transcription dynamics, and a catalog of specific examples of how different types of disruptions at the levels of DNA sequence, and posttranslational/epigenetic modifications can affect coexpression relationships. However, sample-sizes in knock-out experiments large enough to infer networks are rare, and the distribution of KO genes do not have the systematic coverage needed to build understanding of how expression changes percolate in coexpression networks. Given that depression likely consists of many slightly altered expression levels, and many array datasets have very high false discovery rates, it is unsurprising that multiple datasets and techniques will be necessary to detect robustly affected genes.

5.4 COLLECTIVE IMPLICATIONS OF THESE STUDIES FOR DEPRESSION RESEARCH

These studies reiterate many suspected pathways and mediators of depression pathology and highlight network-based methods to extract the transcriptome impact of major depression at the levels of genes, cells and brain regions. But what new predictions do they collectively make about multi-system multi-scale interactions among putative depressive mechanisms? What, if any, grand organization do they implicate behind depression pathology? In Chapters 2-3 we observe evidence for several hypotheses of depression, ironically excepting the canonical monoamine hypothesis. The observed deficits are highly distributed in coexpression networks (Chapter 4) and only detectable based on interaction between multiple brain regions, or cross-species analysis (Chapter 3). These results could be taken as a marker of contention in depression research over the primacy of different mechanisms and the continued lack of a single mechanism which is capable of accounting for the depressed state. However, we show all of these deficits in the overlapping or related datasets, indicating that they are simultaneously present in the same individuals. For instance, altered gene coordination between AMY and ACC is found in the same individuals who have oligodendrocyte and dendritic abnormalities, who are the same individuals with diffusely distributed expression level changes. Since these deficits interact mechanistically as repeatedly reviewed in Chapter 1.2, and we observe evidence that they are occurring simultaneously, we accept the *prima facie* conclusion that these dysfunctions co-occur in depression. As opposed to selecting a single primary deficit as the core mechanism of depression, we propose that depression is a collective multi-scale multi-system disorder. If this is the case, what conceptual framework is there to understand these depression deficits, so that the disorder does not merely become a syndrome of frequently associated mechanisms and symptoms?

The floodgate model of major depression The standard model for complex disease induction is a threshold model, wherein a sufficient number of deficits (driven by genetics or

environments) eventually becomes sufficient to create disease symptoms. Based on results in Chapters 2-4, and in line with the approaches outlined in Chapter 1.2.8, we propose a "floodgate" model of depression, so named because it proposes that depression occurs as an uninhibited chain of events propagating among several related systems which are no longer able to compensate for, or contain external pressure. Consider the directed nature of the mechanistic links shown in Figure 1: It is not the case that genetic and environmental damage merely occurs statically and is purely confined to a local network. Rather, deficits occur and alter the baseline responsively of the systems and overflow into related systems. Therefore, the "floodgate" model is a macro behavioral-level reiteration of the small cascades of depression-related events which are constantly occurring at the molecular level. This is a distinct mechanism of depression induction from existing theories, which emphasize the severity of a particular deficit or the non-specific total contribution of random deficits that surpass a given threshold (under the allostatic load model).

The emphasis on dysynergy in the floodgate model represents a shift in the philosophy of depression research: instead of asking "what is depression?" – and generating a laundry list of unrelated deficits, the central question in the floodgate model is: "how do combinations of factors all lead to a common depressed brain-state?". This new paradigm focuses on how brain structure and function can unify the various hypotheses of depression. For much of the time-course of depression research, the brain has been considered almost incidental to depression - it was merely the setting in which deficits happened to occur. But focusing on how the deficits relate to each other hinges on finding convergent neural mechanisms for hundreds of discrete findings that may appear unconnected. Currently, much research unsuccessfully attempts to skip from anti-depressants to clinical symptoms, almost as if playing a game that involves matching molecules to symptoms, while the brain stands on the sidelines. To make progress in the floodgate model, all an experiment needs to show is what depression is not, i.e. depression is NOT independently a deficit in factor X and factor Y, but both X and Y activate some pathway

Z. (This process essentially corresponds to dimensionality reduction applied to the field of depression research – looking for colinearity among experimental assays). Ironically, by accepting a complex multi-factor basis for depression, it might be possible for traditional single-mediator-style experiments to begin constructing a systematic understanding of depression, which is more akin to the progress made in other areas of neuroscience, as opposed to the current hodge-podge of discordant findings.

A concrete example of how the floodgate model could be used to evaluate pro-depressive impact may be in the case of severe HPA activation: even if the patient has an oligodendrocyte deficit, as long as there are no simultaneously pro-depressive factors at work i.e. inflammation is low, so that astrocyte glutamate scavenging is adequate and there are not prestanding brain region connectivity changes (stemming from long-term stress or childhood abuse for instance) then the system will likely adapt to the severe HPA activation without creating a depressive state. If depression were to occur, the flood-gate model would predict concurring mechanistically-linked deficits in multiple systems that allow stress-inputs to have largely unregulated impact across several systems. The reason the several linked systems ultimately do collapse is probably related to repeated activation from stress combined with genetic deficits. Thus regulatory systems in each individual may be primed for destruction in different combinations in separate individuals. Therefore, the floodgate model does not emphasize any specific set of priming events (which could vary by individual), but rather the reinforcing effects that occur when regulatory mechanisms fail simultaneously across multiple systems. The floodgate analogy is stretched by the complex regulation between depression-affected systems that control percolation among them. Due to multiple regulatory loops between systems, it may be difficult to detect those sets of deficits that lead to unregulated sequences of adaptation (see Chapter 5.4.1 on multiscale modeling). However, the traditional threshold model provides no framework for how deficits cohere to produce depression, making it even less testable.

To lend support to the floodgate's emphasis on the collusion of distributed deficits, it would be necessary to show that specific combinations of deficits in functionally linked systems have a greater effect than randomly distributed deficits that do not form sequential regulatory links across systems. Treatment with interferon (for hepatitis C) could provide a platform to test the floodgate model vs the traditional threshold model for complex disorders. Because this treatment is a known extreme stress on the immune system, under the floodgate model, the difference between those patients who do/do not develop depression will be relatively lower functioning in ALL depression-related systems, whereas any single measure should be less predictive, even if it is severely affected. This could be tested by regressing depression status at the end of interferon treatment against either a combination of PET, DSI, DMT tests, fMRI stress responsivity and peripheral glial markers (to address functionality of several depression-related systems) or some normalized minimum of those scores. These could indicate if depression is more likely to occur in individuals with some distribution of low-grade deficits (floodgate model) or simply occurs in those individuals with the greatest total deficits. There is some supporting evidence that this proposed experiment could distinguish between depression models, as poor sleep patterns prior to interferon treatment predicts depression during treatment (Franzen *et al*, 2010) as do elevated HAM-D scores (Lotrich *et al*, 2007).

Assuming that all results from depression research are in fact correct, these deficits ultimately contribute towards a brain-state which is prone to stress reactivity and focused on negative personal events. In this framework, the question becomes how exactly deficits combine to create a brain-state characterized by depression. For instance, the floodgate model would predict certain deficits that are particularly detrimental in combination. Even if the floodgate model is validated experimentally, moving depression research into this paradigm of studying the breakdown of robustness vs single depression mediator will require a framework that binds together the hundreds of depression-related deficits. Exploring the concurrent influences of depression deficits will entail dynamical systems models to account for the specific

contribution and interaction of various far-flung depression influences, within a unified setting. However, building such models is technically challenging. Part of what has prevented an integrated model of depression is the challenge of mechanistically uniting deficits that range from complex second messenger systems, to neurotransmitters, to synaptic configurations to the temporal evolution of brain region communication during specific tasks. Modeling all of these influences simultaneously requires “multi-scale” models that search for convergent dynamic patterns of brain activity that are generated by both molecular and systems-level effects – just as they are in the actual brain.

5.4.1 Overview of multi-scale modeling

A pubmed search for all variants of “multi-scale” modeling currently (12-25-2010) yields 12 results, mainly related to large-scale cardiac models. However, I predict multi-scale models will become a key multi-center investigative modeling technique. What are multi-scale models? Multi-scale models mathematically combine the activities of biological elements that operate on different physical or temporal “scales”. Compared to “normal” single-scale models, they more closely reflect the nested complexity of real biological systems (see figure 14). For example, gene transcription is influenced by DNA sequences, post-translational modification, feedback from protein-protein networks, cellular/molecular activity, and ultimately the activity of the organism. While these systems are sometimes modeled in isolation (single-scale), a multi-scale model mathematically couples activity on two or more of these levels, which literally can scale-up the relevance of findings on lower levels. In short, the end point of most experiments – the “limitations” section of confounding factors and unanswered questions is where multi-scale models begin.

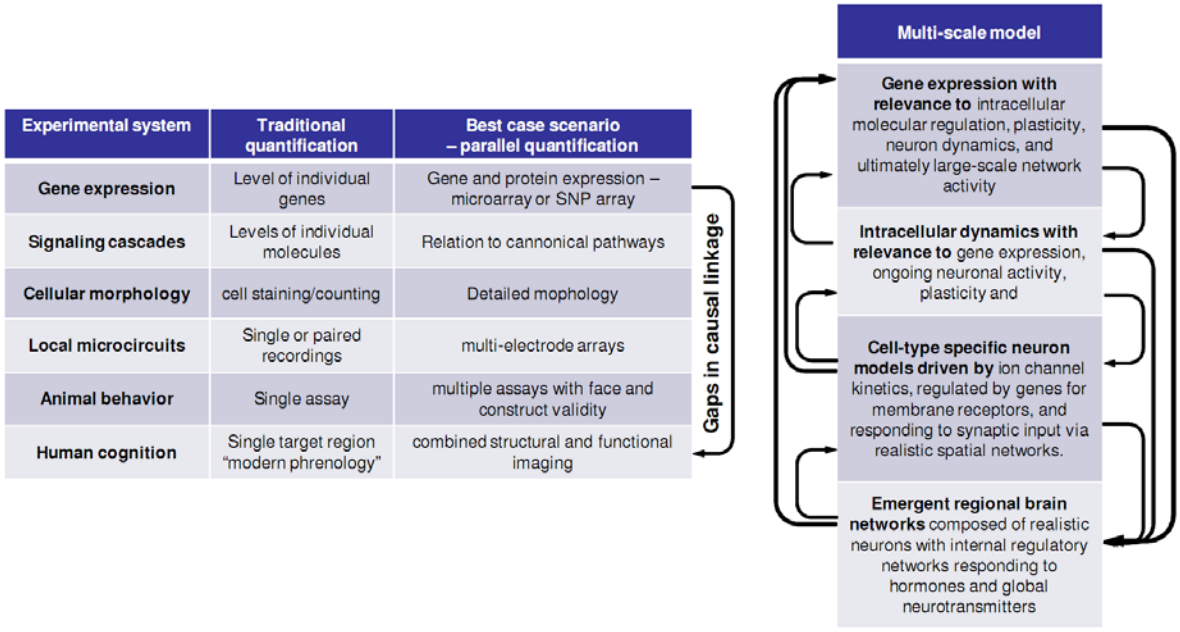


Figure 16 How is a multi-scale model different from other models?

Here we contrast how single-scale vs multi-scale models would be formulated for multi-scale disorders such as depression. Multi-scale models form a unified framework, wherein the results of one simulation feed into the next layer of complexity. This nested organization is useful in preventing the tentative conclusions, isolated results, system-specific results, and parallel but unconnected investigations that characterize depression research. However, multi-scale modeling does require an extensive computational framework between mechanisms that would have to be gleaned from hundreds of papers. Currently such a framework relating the actions of multiple modulators and biochemical links between key brain structures does not exist, even in primitive form, for depresion.

High-throughput technologies indicate that depression stems from the combination and interaction of many causal factors whose individual contribution is small. This distributed pathology impinges on cellular and molecular networks which adapt and interact with each other, simultaneously creating the cognitive symptoms of depression and disguising their cellular origins. Thus, the process of fully testing how or if putative antidepressant agents can

reverse the cognitive symptoms of depression in humans is long, expensive and error-prone. Combining knowledge about depression pathology at the levels of genes, neurons, and brain regions into a unified multi-scale model should allow a more coherent understanding of interactions among causal factors and how they collectively maintain the depressive state.

The advantage of multi-scale models for depression is that multiple polymorphisms that affect specific components of neurotransmission can be modeled in a realistic parallel fashion to understand their collective effect on network activity. A practical example in the case of depression would be to create a detailed simulation of the amygdala, anterior cingulate cortex, and DLPFC as a brain region network of mood regulation with outputs to a virtual hypothalamus and reticular activating system. Each of these regions will have a realistic three-dimensional synaptic coupling between neurons, which each have different receptor subtypes that determine their firing patterns. The activity of neurons as well as the level of virtual circulating hormones and neuromodulators will then interact (via differential equations) with the regulations of genes. Genes for trophic factors, receptor insertion and other processes will regulate neuron firing rates and network structure, to ultimately influence inter-regional communication. Thus a multi-scale depression model would form a mechanistic explanation for signal transduction and biological adaptation from the level of drugable molecules all the way up to brain region communication associated with mood regulation and perceptual states.

This simulation environment would enable faster antidepressant compound profiling with fewer false positives, because it incorporates the relevant biological complexity. Because the model is mathematical and can be “frozen” at any moment, the complete cascade of effects from antidepressants (which interact with virtual cellular machinery) is fully dissectible, and may be utilized in understanding how to extend and improve the activity of promising compounds. Moreover, multi-scale models could provide new insight into the mechanisms most responsible for sustaining the depressive state, and thus a novel basis for new classes of antidepressants. Conversely, multiscale models can be used to test specific compounds or several compounds

with particular properties that are most successful at switching the network from depressed to normal activity.

Integrating lists of differentially expressed genes into a multi-scale model

Gene transcript products can affect intracellular signaling, cell-structures, and cell-cell communication - the full gamut of cellular functions. Therefore implementing differential equations that describe the effects of dozens of genes on a cell (for instance, the effects of a disease-associated gene set) can quickly snowball into a project of simulating every known cellular system. There are at least two complementary strategies to estimate the contribution of numerous genes to multi-scale dynamical models. These enable more realistic simulation of disease complexity (effects of multiple genes and systems) in the framework of incomplete information on the cellular function of genes and gene products.

Strategy 1: Collapsing gene lists into modular affected systems

Under this strategy, microarray results are interpreted as a proxy for subsystem-specific dysregulation: the degree of dysfunction corresponds to the number of affected genes found in a given system, or potentially the coexpression connectivity of those genes. The main assumption of this approach is that the disease relevance of a set of genes is proportional to their cumulative effects on major cellular systems that ultimately affect spiking activity. For instance, if microarrays from the prefrontal cortex showed differential expression of enzymes related to GABA, under this strategy, the first step would be to construct a computational system that encapsulates major components of GABAergic neurotransmission, such as synthesis, vesicle release probability, receptor subtypes and reuptake. Then the probable effect of the differentially expressed genes could be simulated by varying parameters related to the efficiency of the core system components. There are multiple advantages to constructing the simulation around well-characterized elements of a particular subsystem, rather than around specific differentially expressed genes: data may actually exist to realistically model these core systems and modifying the core system may more accurately represents how disease

processes interact with prestanding systems. Furthermore, since microarray measurements present a mixture of causal and (mal)adaptive changes, it may be more appropriate to implement systems-level changes than gene-level changes.

Strategy #1 essentially subsumes genes into parameters (or sometimes variables) that affect neuronal activity. This strategy will still permit multi-scale feedback loops between neuronal activity and gene expression levels, however those "genes" will be functionally implemented as elements in signaling cascades related to neuronal function. This nested complexity is essential in multi-scale models, because it permits examination of the high-level examination of low-level deficits. But how can we increase the richness of gene-gene interactions, at the lowest level of the model, when coexpression networks are acausal and direct interaction pathways are incomplete?

Strategy #2: Modular interactions to determine virtual transcription

Creating a transcriptional activation system for modular cell functions could provide more realistic feedback interactions at the gene level of the model. The "gene" markers of modular cellular functions (which are essentially tied to sets of neuronal parameters) would follow simple activation and repression relationships thought to exist between the systems (likely composed of the major hypothesized systems involved in depression). This is a useful approximation of transcriptional programming since depression is a heterogeneous disorder (different cellular systems are affected in different patients), and because transcriptional programs are regulated in many ways like a classic hidden-layer neural network: classifying a range of inputs into transcriptional programs (Babu *et al*, 2008; Shmulevich *et al*, 2005). While there would be a limited number of these transcriptional markers, they could have fairly detailed combinatorial interactions in response to a set of environmental and endogenous conditions present at a given time in the model. For instance, markers of low thyroid hormone+high stress could activate a very different final selection of genes/parameters than low thyroid+inflammation. The interaction of these modular systems is largely informed by relevant literature, and so may seem *adhoc* in

comparison to data-driven coexpression networks. However, in the absence of causal data from time-series microarrays, it represents an acceptable approximation of multi-gene interactions. Furthermore, these interactions exist mainly to bring further realism to temporal evolution of neuronal parameters, which themselves have a much more refined set of parameters.

Conclusion of multi-scale modeling and relevance to the floodgate model Thus, the floodgate model is not a purely descriptive model of depression, but may be useful in understanding how slight changes across multiple systems do more than accumulate – they can occur in specific combinations that are particularly deleterious due to linked dysregulation. The floodgate model pushes depression into an individualized medicine paradigm in that each patient may have a different set of predisposing factors which are most likely to be mediating depression. Rather than simply dissolving the etiology of depression into many subcomponents, the floodgate hypothesis relates broad trends in multi-system depression involvement that accounts for its irregular presentation and self-sustaining pathology.

5.5 CONCLUSIONS AND FUTURE DIRECTIONS

Challenges and recommendations for neuropsychiatry The past 60 years of depression research have been marked by blossoming understanding of the effects of depression on the brain, a diversified series of antidepressant drugs, but scant change in disease remission rates. What are the key factors that have prevented molecular research from improving patient response rates? Is the lack of some specific technologies or type of study holding back patient treatment? Results from Chapter 2-4 confirm several main suspects in depression pathology. But if the field has indeed identified the key mediators of the depressive state, why do we still lack a coherent understanding of depression symptoms to guide antidepressant development and patient treatment? Our inability to control the outcome of cases of major depression

suggests we do not understand how key components of depression pathology interact. As indicated by Figure 1 and reiterated in the floodgate model of depression, there are known mechanistic interactions between various hypotheses of depression that offer hints to common mechanisms behind the observed pathology. But the vast majority of depression research describes isolated effects, and hence a forest of vaguely related results has grown up around the amalgam of symptoms that make up depression. Comparatively few studies are devoted to understanding how distinct depression effects converge into brain states that produce a depressive phenotype. Because depression consists of a combination of deficits, it is utterly important to understand how they interact though common neural pathways. Those common pathways could form the basis for logical drug development, whereas research on the currently disjoint hypotheses of depression mechanisms has spurred branching subfields.

How can we understand the common, unique and combined contributions of different mechanisms to the depressed state? Capturing the shifting combinatorial structure of depression requires revamped organization of research focused on the relationships between specific experimental results. If study results are structured relative to each other, in the same way that depression effects exist relative to each other, a coherent pathology may emerge as depression models align with actual brain mechanisms. At the very least, this entails a discussion section in each paper linking specific findings to larger mechanisms that are phenomenologically closer to depression symptoms. However, there is a limit to the power of arrow diagrams and verbal references: the multitude of depression effects necessitate rigorous implementation in a unified setting. The multi-scale model necessary to capture these effects could be one of the most involved modeling projects to date, because known depression deficits span several physical scales and multiple systems within scales. However, the alternative to commencing model construction is to allow the present cacophony of results to rattle on without an overarching framework. Based on the current level of scientific understanding in depression, objectively quantified by pharmaceutical exodus and low remission rates in patients, we suggest

the organization of depression research should be revamped so the field can capitalize on key findings. Specifically we suggest configuring research results to mimic the interconnected structure of depression mechanisms by focusing on interactions between depressive mechanisms. In short, the span of depression mechanisms in the brain demands a corresponding bridge between studies in research.

Global summary We directly confront the dominant obstacles in the field of depression research - multifactorial and multi-scale pathology – by constructing transcriptome-wide networks that comment on depression mechanism at the level of genes, cells, and brain regions. Each of these studies detects a set of biological and network effects that indicate depression persists through multi-component multi-system failures. For instance, in meta-analysis of postmortem depression microarray datasets, we investigated if differentially expressed genes interact with coexpression network structure – potentially using the small-world structure to leverage pathology. We show that the expression and connectivity changes in depression and other neuropsychiatric disorders are confined to the fringes of coexpression networks. Thus the contested genes of interest in depression may create a collective pro-depressive effect without any single gene predominating.

Because the false discovery rates are high in depression microarray studies, we use cross-species analysis to refine a set of differentially expressed genes and then use coexpression networks to show that these results represent two communities of glial- and neuronal growth-related genes. This oligodendrocyte/glial deficit is one of the few results confirmed across microarray studies and fits with the gene markers of change in neuronal structure. So once again, depression deficits do not occur in isolation, but appear simultaneously in meaningful sets of effects that mirror their codependent function in the brain.

We establish “gene coordination” as a new gene-based synchrony measure that appears to reflect function communication between brain regions. Genes with altered coordination implicate the action of glucocorticoids, thyroid hormone, estradiol, insulin and

cytokines in mediating depression effects on cortico-limbic regulatory circuits. These suspect pathologies operate simultaneously on brain regions which are crucial in determining mood and emotional responses. Thus a set of circulating factors with molecular interactions is related to a set of brain regions with functional interactions. While such network-network interactions are complex, they may produce meaningful disease characterizations, because this conception of depression is closer to reality, evidenced by the well-validated results of this study.

Chapters 2-4 already comprise the most network-centric molecular studies of depression to date, yet they point toward the need to integrate more types of cellular networks, to better detect and conceptualize the widespread impact of depression. Embedding these effects into realistic cellular networks would be helpful in understanding the confluence of multi-system interactions that characterize depression. This is an extremely challenging computational modeling task, but we outline how a multi-scale model could satisfy requirements for a realistic simulation environment and could break depression research out of the strangely parallel worlds of competing theories. We also propose a non-competitive framework for depression research - the floodgate model of depression – designed to meld the relationships between hypotheses of depression into a more coherent theory that will develop, rather than fragment, under pressure from new experimental results.

From the perspective of a depression researcher, there has never been a more diverse cadre of potential depression mechanisms. From the perspective of patients with unremitting depression, their vitality is on hold until superior new treatment options arrive. Fortunately there are many insightful experiments currently underway to determine the precise causes and effects of depression. Thus the primary elements necessary for antidepressant drug development - and substantial improvement in the human experience - already exist. But it may be possible to catalyze meaningful progress in depression research through new conceptual approaches. Based on the results of Chapters 2-4 we propose that many forms of network analysis will be

useful in gathering diverse components of pathology and then projecting a coherent representation of depression.

APPENDIX A– Supplementary information for paper #1

Supplemental information

A Molecular Signature of Depression in the Amygdala

Etienne Sibille,[1,2,*] Yingjie Wang,[1] Jennifer Joeyen-Waldorf,[1] Chris Gaiteri,[1,2] Alexandre Surget,[4] Sunghee Oh,[3] Catherine Belzung,[4] George Tseng,[3] and David A Lewis[1,2]

[1]Department of Psychiatry, [2]Center for Neuroscience, [3]Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA 15213, USA.

[4]U930 FRE CNRS 2448, INSERM and Université François Rabelais, 37200 Tours, France.

*Address correspondence to Etienne Sibille, Department of Psychiatry, University of Pittsburgh, 3811 O'Hara Street, BST W1643, Pittsburgh, PA 15213; sibilleel@upmc.edu (E-mail).

1. Detailed Material and Methods

Subjects

Brain samples were obtained during autopsies conducted at the Allegheny County Medical Examiner's Office following consent from the surviving next-of-kin. After careful examination of demographic, clinical and technical parameters, we selected a cohort of male depressed subjects and matched control samples.

For all subjects, consensus DSM-IV diagnoses of MDD were made by an independent committee of experienced clinical research scientists at a case conference utilizing information obtained from clinical records, toxicology exam and a standardized psychological autopsy

(Glantz and Lewis, 1997). This latter incorporates a structured interview, conducted by a licensed clinical psychologist with family members of the index subject, to assess diagnosis, psychopathology, medical, social and family histories, as well as history of substance abuse. A symptom score was calculated based on the presence at time of death (1= unequivocal yes; 0.5= unsure or subthreshold; 0= unequivocal no) of nine major depressive episode symptoms: depressed mood, anhedonia, appetite disturbance, sleep disturbance, psychomotor change, anergia, self-recrimination, diminished ability to concentrate or make decision, and suicidality.

We further focused on patients with familial depression, as these subjects in general display earlier onset of symptoms, more recurring episodes, shorter inter-episode duration, and suffer from more severe and incapacitating episodes than non-familial depressed patients(Fava *et al*, 2000). To determine familial MDD, the next-of-kin was asked about each 1st-degree family member and about the psychiatric history of other family members. This approach has the advantage of being prompt, efficient and appropriate for postmortem studies, but it also underestimates the presence of psychiatric illness in 1st-degree relatives (Andreasen *et al*, 1977). All MDD subjects had at least one 1st-degree relative with a history of MDD. The increased disease severity was supported by a longer average duration of illness in the familial depressed cohort compared to non-familial subjects collected under the same conditions in the same brain donation program (9±2 years versus 3±1 years; Mean±sem; t-Test, p=0.01).

Cases who did not commit suicide, died from natural causes, thus ruling out the possibility of accidental death as masked suicide. MDD subjects with co-morbid psychiatric disorders were excluded. Antidepressant drug exposure was assessed by clinical data from structured interviews, review of records and toxicology studies. Control subjects were paired to each case as closely as possible on age and freezer storage time. Control subjects did not have an Axis I psychiatric disorder, were antidepressant drug-free and died from natural or accidental causes other than suicide. The family histories of MDD in controls included one positive, nine negatives and six unknowns. Subjects with advanced disease stages (i.e., cancer, neurodegenerative

disorders) were excluded. All cases and controls were white Caucasian and were selected for rapid modes of death and short agonal phases, to limit the influence of agonal factors on RNA quality and pH (Tomita *et al*, 2004). All selected brains were analyzed for adequate brain pH (>6.4) and RNA integrity by optical density (OD_{260} ≥ 1.6) and Agilent bioanalyzer analysis (Agilent Technologies, Palo Alto, CA; RIN expert scoring system ≥ 7) as previously described (Eggan *et al*, 2008). Two pairs did not pass quality control in AMY, leaving 16 pairs in ACC and 14 pairs in AMY for the final analysis (Table 1). Rates of death by suicide, disease recurrence, evidence for antidepressant treatment at time of death, and alcohol dependence in MDD subjects are described in Table 1. Toxicological screens on peripheral fluids identified the presence of at least one antidepressant in 5 subjects, including four different tricyclics, one selective serotonin reuptake inhibitor and one weak dopamine reuptake inhibitor. Importantly, all antidepressant-treated subjects were currently depressed at time of death, suggesting either a lack of efficacy, suboptimal treatment or treatment-resistance in these subjects. All procedures were approved by the University of Pittsburgh's Institutional Review Board and Committee for Oversight of Research Involving the Dead.

Brain samples

Upon collection, coronal blocks through the rostral to caudal extent of the brain were cut in ~2 cm blocks and stored at -80C. The AMY is located ~2-3 cm caudal to the temporal pole. Tissue samples were dissected from 20µm section in the cryostat and stored in Trizol (Invitrogen, Carlsbad, CA). Sampling was adapted from (Hamidi *et al*, 2004) (Figure 1A in manuscript). The lateral and ventral borders were delimited by the white matter surrounding the AMY. The medial border was defined by the deep layer of the cortex along the medial edge of the temporal lobe. Finally, the dorsal border of the AMY was drawn along the lateral, basolateral and basomedian nuclei. In view of the heterogeneity of the AMY structure, we performed a pilot study to determine appropriate protocols for reliable and consistent dissection, and to assess the sample-to-sample variation in transcript levels within the rostral part of the AMY compared

to more caudal samples. Rostral samples were reliably sub-dissected and resulted in samples enriched in lateral, basolateral and basomedian nuclei tissue (De Olmos, 2004), while avoiding tissue dilution from the cortical and transitional amygdaloid nuclei that are more prominent in middle and caudal AMY. Rostral, middle and caudal AMY samples were processed on arrays. Results indicated that rostral sub-dissected samples displayed low intra-variability of RNA levels compared to more caudal samples (Figure 1A in manuscript). Accordingly, sampling proceeded on sub-dissected samples corresponding to rostral sections 1 and 2 in Figure 1A. This protocol enriches samples in nuclei of interest and increases the probability that signal differences will reflect subject differences rather than AMY rostral-caudal variability.

ACC samples containing all six cortical layers were harvested from coronal sections at the anatomical level corresponding to subgenual ACC (Brodman area 25), located in the third prefrontal cortex block along the rostral-caudal axis of the brain. A similar microarray pilot study revealed that, within a subject, very little variability in transcript levels was observed for most genes along 10mm of the rostral-caudal axis of the subgenual ACC (Figure 1B in manuscript), in agreement with the more homogeneous anatomical structure of this brain area compared to the AMY. Accordingly, sampling on all cases and controls occurred in the rostral part of the subgenual ACC, immediately caudal to the genu of the corpus callosum. Replicate samples were processed for 4 pairs in the AMY at 3-4 months interval from different RNA extractions obtained from the same subjects. A few white matter (WM) samples were obtained for analysis of cellular origin of transcripts (Sibille *et al*, 2008). These samples were collected adjacent to the grey matter (GM) samples in ACC (n=7) and as an easily recognizable thin band located between the lateral and ventral borders of the amygdala (n=4).

WM/GM analysis

While subgroups of genes are expressed in cell type-specific manners, the majority of gene transcripts display relative enrichments across cell types, including neurons and glia. Here, we used array data from adjacent white matter (WM) samples to generate WM/GM ratios

that are specific for each gene and brain region. We have shown that these ratios represent valid estimates of relative gene transcript enrichment from glia (WM/GM>1.5), neurons (WM/GM<-1.5) or both cellular population (-1.5<WM/GM<1.5) and that incorporating the use of these ratios into transcriptome analysis can provide wider views of overall patterns relating to glial and neuronal functions (Erraji-Benchekroun *et al*, 2005; Sibille *et al*, 2008). Here, WM/GM ratios generated in control samples were used, although ratios generated in psychiatric subjects or treated mice were essentially identical, as Pearson correlation factors between control and all-samples ratios were greater than 0.99 (Sibille *et al*, 2008).

Real-time quantitative real-time PCR (qPCR)

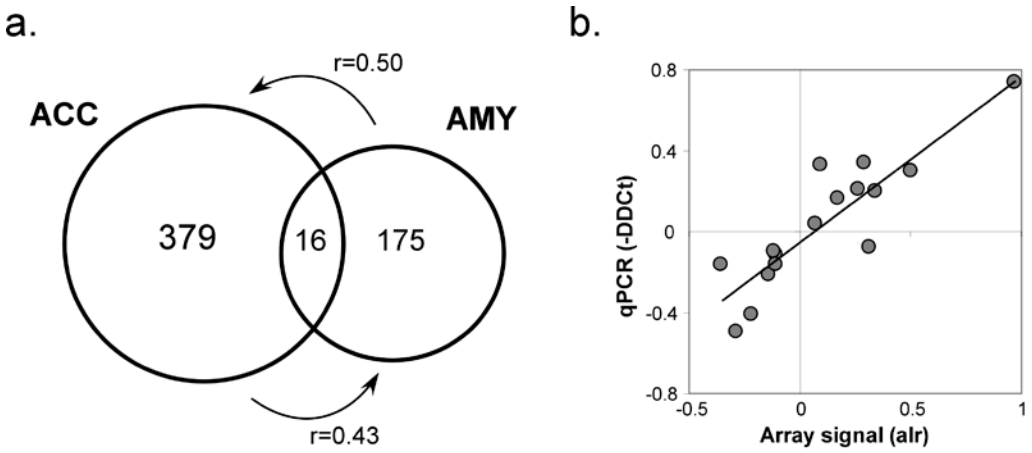
qPCR was performed as previously described (Erraji-Benchekroun *et al*, 2005; Galfalvy *et al*, 2003) In brief, small PCR products (80-120 base-pairs) were amplified in quadruplets on an Opticon real-time PCR machine (Bio-Rad, Hercules, CA), using universal PCR conditions [65C to 59C touch-down, followed by 35 cycles (15" at 95C, 10" at 59C and 10" at 72C)]. 150 pg of cDNA was amplified in 20µl reactions [0.3X Sybr-green, 3mM $MgCl_2$, 200µM dNTPs, 200µM primers, 0.5 unit Platinum Taq DNA polymerase (Invitrogen, Carlsbad, CA)]. Primer-dimers were assessed by amplifying primers without cDNA. Primers were retained if they produced no primer-dimers or non-specific signal only after 35 cycles. Results were calculated as the geometric mean of the relative intensities compared to three internal control genes (actin, glyceraldehyde-3-phosphate dehydrogenase and cyclophilin).

Western Blot analysis

Proteins were isolated from phenol-ethanol supernatant obtained during the RNA isolation for array samples and re-suspended in urea/SDS buffer. 5µg of protein samples were resolved by SDS PAGE in 10% Tris/glycine gels and transferred to PVDF membrane. After 1 hour in Odyssey blocking buffer (LI-COR Biosciences, Lincoln, NE), the blots were incubated with primary antibodies [Rabbit anti-actin 1:10,000, Sigma #A2066, and mouse anti-cyclic nucleotide 3' phosphodiesterase (CNP); SMI-91R from Covance, Denver, PA] in Odyssey

blocking buffer, followed by incubation with secondary antibodies (IRDye® 800 anti-rabbit and 680 anti-mouse; LI-COR Biosciences). After extensive washing, the signals were simultaneously detected using the LI-COR Odyssey® Infrared imaging system. To compare protein content between different samples and to correct for any experimental variations that occur during sample processing on SDS PAGE and Western blots, CNP protein content was expressed relative to the actin content in the same sample. Paired samples were processed in quadruplicate on the same gel.

Gene coexpression networks *See Section #5*



2. Figure 17 (Figure S1) AMY-ACC altered gene expression in MDD and qPCR validation

(a) Venn diagram of altered gene expression. 45% of these genes were upregulated and 55% downregulated in AMY, while ACC results displayed 60% and 40% up- and downregulated genes, respectively. Arrows indicate directional correlations between changes in transcript levels for genes identified in one area (origin of arrow) and changes for the same genes in the other area (end of arrow); $p<e^{-6}$ in both directions. Although the overlap in gene selection was limited, transcript changes in AMY and ACC significantly predicted similar trends for the same

transcripts in the other area, suggesting coordinated changes across areas, despite variability in statistical thresholds.

(b) Technical validation of array results by independent qPCR measurements. Alr, Average Log_2 of (MDD/Control) expression ratio. (-DDCt) represent differences in PCR cycle thresholds between MDD and control samples, which are equivalent to Log_2 values of ratios (See also Table 3). Upregulated: *GRIN2B, DGKG, GABRA2, KCTD12, CALB1, DUSP4, GPNMB, ASPH, RAB27B*; Downregulated: *MOBP, CNP, EGR1, MBP, ENPP2, MAPK1, RPH3A*; Unchanged (*RAB27B*). For all but one (*GRIN2B*), qPCR and array results correlated highly (All genes, R=0.88, $p<5.e^{-6}$; y slope=1.07). Line indicates linear fit.

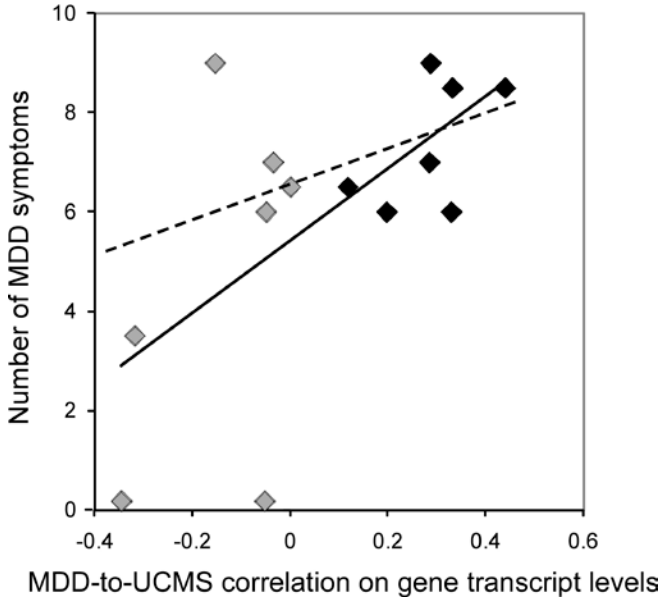


Figure 18 (Figure S2) Positive correlation between the number of MDD symptoms and UCMS/MDD correspondence

See the following section in the manuscript: "AMY cross-species correlations of depression-related molecular changes identified a subgroup of human MDD subjects". The solid line indicates the linear fit for all 14 pairs in AMY (r=0.62; p=0.02). The dashed line indicates the

linear fit after removing two subjects with 0 symptoms (r=0.51, p=0.09).Black squares indicate MDD^{UCMS} subjects.

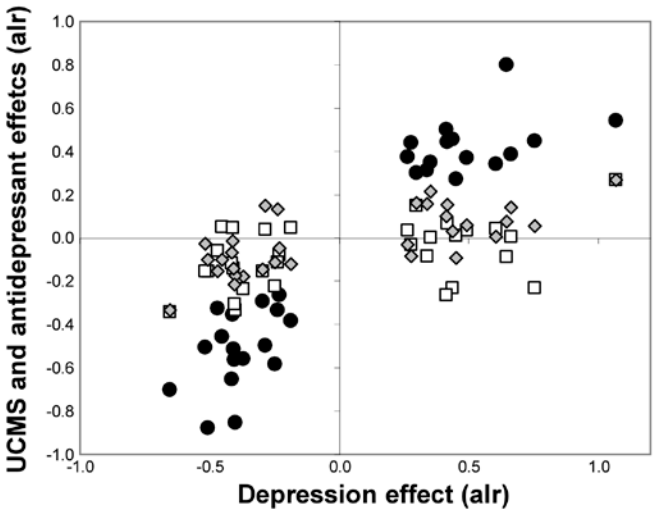


Figure 19 (Figure S3) Antidepressant reversal of MDD conserved changes in the mouse UCMS model

32 genes displayed significant and highly similar changes in transcript levels in human depression and mouse UCMS (black dots). All changes were reversed in UCMS-exposed mice after treatments with an effective (fluoxetine; white squares) or a putative (Crf1R antagonist; grey diamonds) antidepressant treatments in UCMS-exposed mice. Values are from (Surget *et al*, 2009) and are described in Table 3

5. GENE COEXPRESSION NETWORKS (Methods & supporting findings)

Gene networks based on coexpression (i.e., correlated patterns of expression) appear to represent intrinsic attributes of cellular and neural systems that are helpful in identifying functionally-related genes (Lee *et al*, 2004). On the genomic scale, these gene interactions networks are clustered into functional modules (Zhang *et al*, 2005) embedded within a generic scale-free structure (Agrawal, 2002) and this organization persists across species (Bergmann *et al*, 2004). Coexpression networks of genes built through Pearson correlation are broad yet

reliable representation of gene interactions (Eisen *et al*, 1998; Lee *et al*, 2004) and have successfully dissected canonical datasets into functional modules (Carlson *et al*, 2006). We use the Pearson product-moment correlation coefficient to estimate pair-wise coexpression of our set of 32 genes in both mouse and humans. All analyses were performed using the boost graph library and custom MATLAB code. Network visualizations were created using Cytoscape.

Validity and robustness of experimentally derived networks

The sample sizes (n=6/group in mouse; 4 groups: Control, UCMS, UCMS+ fluoxetine, UCMS + CRF1R antagonist; n=14 pairs in human) ruled out the network creation through non-linear or information theoretic measures, which commonly require sample sizes that are an order of magnitude larger. Thus, to ensure that our coexpression links were reliable markers of gene interactions, we used clustering coefficient analysis (Step 1) and jackknife correlation (Step 2) to optimize our cut-off selection. Our goal was to use these techniques to optimize the biologically valid information in the network and to ensure an independent unbiased perspective on glial/neuronal gene regulation in depression.

Network Authentication Step 1: Network Validity

Clustering coefficients estimate the density of local connections in a network. They are calculated for each node in the network as the number of connections between neighboring nodes, divided by the total possible number of connections between all neighboring nodes. Clustering coefficients are a fundamental measure of network structure with wide applicability in brain networks (Sporns *et al*, 2004). We used clustering coefficient as a signature of structural information in the network that was generated by a particular cutoff point, in a method based on Elo (2007). Maxima in the plot of clustering coefficient vs. cutoff represent an optimal ratio of the biological structure vs. noisy or spurious connections. In each case, we compared the clustering coefficients to degree-matched randomly selected networks (see representative plots in Figures S3-4). These plots showed maxima in the region of Pearson correlations of ~0.6-0.8 for all

conditions in both species, indicating that using a Pearson correlation cutoff in this region would maximize the number of links which are representative of biological structure (Elo *et al*, 2007).

Network Authentication Step 2: Network Robustness

Pearson correlation is susceptible to outlying values. To prevent such occurrences from generating links in our networks, we used jackknife correlation in combination with an optimized cutoff threshold (Step 1) to optimally prune the network and maximize the biological signal. Figure S5 shows the common bimodal distributions of potential links that clearly segregated as either robust (i.e., right columns in graphs) or spurious (left columns in graphs) links in the human and mouse datasets. Links in our network were gathered from the most robust groups in jackknife histogram. Within each bin, links were selected in order of Pearson correlation value, beginning with the highest values, until the required network size was filled. Because of the hybrid link selection technique, actual Pearson correlation values of included links are substantially higher than cutoff values. A network size of 100 links will have: *1)* a cutoff value in the suggested range 0.65 (human) or 0.75 (mouse), *2)* will only be composed of robust links, and *3)* will be representative of glial-glial, neuronal-neuronal and neuronal-glial connectedness (See Fig 5). Results for networks with 100 bidirectional links are summarized in table S1. This rigorous criterion generated networks which are highly valid representation of underlying biological interactions.

	Human		Mouse			
	Control	**MDD**	**Control**	**UCMS**	**Fluoxetine**	**Crf1R atg**
Pearson cutoff in jackknife correlation	0.65	0.65	0.75	0.75	0.75	0.75
Average Pearson correlation value of links	0.77	0.76	0.93	0.93	0.88	0.9
% increase of clustering coefficient vs Random network	60	110	95	41	76	77

Table 5 (Table S1) Pearson cutoff values are selected by monitoring increased clustering coefficients of experimental networks vs. controls.

Genes above the cutoff value are then selected in order of their robustness in jackknife correlation. Combining these techniques naturally leads to the selection of genes with a higher correlation than the baseline cutoff.

APPENDIX B SUPPLEMENTARY INFORMATION FOR PAPER #3

Differentially expressed genes in depression and other neuropsychiatric disorders are distributed on the periphery of resilient gene coexpression networks

Supplementary Experiments and Methods:

A. Selecting optimal coexpression threshold for maximum modularity

B. Connectivity estimates of known gene classes

Supplementary Figures and Tables Captions

Supplementary Experiments and Methods:

A. Selecting optimal coexpression threshold for maximum modularity

Optimally selecting network links to establish a high-confidence gene networks using human postmortem microarray data There are multiple ways to infer gene interaction networks, however all methods rely on synchronous fluctuations in two or more genes across multiple microarrays. Therefore, when inferring biological networks based on microarrays, a primary concern is selecting an optimal level of correlation that ensures core network structure is a product of real biological interactions, not spurious data correlations. To assure networks generated via datasets listed in Table S1 represent biological reality, we first individually optimize the link selection process for each dataset, since each represents a distinct combination of sample size, data quality, and biological structure (see next).

Figure S2A illustrates the importance of large samples sizes in addressing these issues by shrinking the null distribution of expected Pearson correlation values between genes (decreasing noise). The larger number of extreme correlation values of the non-permutated dataset in Figure S2A (shown in red) compared to the permutated data (gray) is thus indicative of biological relationships underlying gene correlations. Using postmortem array data from psychiatric populations, an additional concern is the influence of subpopulations on network generation. Increasing sample size addresses this concern as it refines the set of actual correlations, decreasing false positives (shown in Figure S2B distribution plateauing at n=14 samples). These plots show that false-positive network links are unlikely to occur at high link selection thresholds (i.e. correlations of 0.8 or greater).

To translate raw gene-gene correlation values into gene networks, we apply this optimized threshold to the absolute correlation values and select all higher correlations to be links in the network (greater than 0.85 in Fig.1C). The outcome of this network generation procedure is extremely low false discovery rate (FDR) for network links, commonly under 1%.

Bootstrap estimates of correlation values shows that at the high thresholds used in these analyses, the correlations are very stable and that there is little influence from outlying values, such as would be generated by the presence of subpopulations (Figure S2D). Based on this complete assessment of gene-gene correlations, each dataset is transformed into a high confidence network, in which links represent biologically undefined but highly consistent relationship between connected nodes.

To optimally balance false positives and false negative links in the networks, so that the real biologically-driven correlations predominate, we utilize the stereotypical community structure of biological networks. Because modularity (segregated clusters) in network structure is a hallmark of meaningful network connections, we chose the cutoff for "real" correlations to be the exact point of greatest biological/clustered structure (Figure S2C, specifically example networks shown for different thresholds). We do this by minimizing network synchronizability (λ_1) - a measure which is small when the network is composed of nearly disconnected clusters, indicating the network has been pruned down to its core modular components (Perkins *et al*, 2009).

Assortative mixing in postmortem gene networks Assortativity (the likelihood of connection to nodes of similar degree) is another global characteristic that may direct disease activity in coexpression networks. A commonly cited distinction between technological and biological network compared to social networks is that the former tend to be dissortative while social networks are positively assortative. However, we show here that all gene array based networks in this analysis are strongly assortative (mean/median assortativity .396/.468, see Table S1) meaning that hubs are preferentially connected to hubs, while provincial nodes are preferentially connected to other provincial nodes. See Fig.1D for a graphical representation summarizing the gene network characteristics observed in all tested postmortem gene arrays datasets (from Table S1).

B. Connectivity estimates of known gene classes

Understanding relevance of connectivity through known gene classes Since DE genes have a particularly low connectivity, we were interested in what other classes of genes might share this characteristic and/or whether biological meaning can be assigned to over- or under-connectivity. Examining the connectivity of disease genes in particular allows us to assess if the centrality-lethality relationship found in protein-protein interaction (PPI) networks holds for gene networks. Additionally mapping gene classes onto postmortem networks can provide an external validation for our inferred network structures. As with differentially expressed genes, for each class of genes we test the hypothesis that it is either central or peripheral to the gene network. Many of these classes are hubs in PPI networks; however, given the lack of overlap in gene and protein network links, their role in gene interaction networks may not be similar (Bhardwaj and Lu, 2009; Xulvi-Brunet *et al*, 2009).

Composite results show these gene classes have more non-random types of connectivity, both above and below the expected range, readily demonstrable from the large number of extreme p-values in Table S4. Also, under hierarchical clustering, similar brain regions from different studies showed similar patterns of connectivity across gene categories, (as indicated by adjacency of similar brain regions under hierarchical clustering in Figure S3). For instance, the proximity of anterior cingulate studies, amygdala studies, and of BA8-9-and-10 studies in Figure S3 indicates that the region by gene class connectivity variations are not random, but consistent characteristics of those particular systems, potentially related to their distinct biological functions.

Gene classes with non-random connectivity Concordant results across many included studies were found for connectivity of cancer (n=497) (Futreal *et al*, 2004) and transcriptions factor gene lists (n=1835) (Vaquerizas *et al*, 2009). Surprisingly, given their presumptive roles in

directing large swathes of activity, these classes were both less connected in the investigated datasets than expected at random. Also, those genes that are commonly differentially expressed ("DER" genes - commonly differentially expressed ratio genes, n=400) across a large number of microarray studies (Chen *et al*, 2008b) were more connected than expected. This gene category may correspond to our highly variable/highly connected gene class and indeed shows similarly high connectivity. Expanding on the original centrality-lethality correlation, Barrenas (2009) showed that the protein products of complex genes have lower connective than monogenic diseases. However at the gene network level it appears they above expected connectivity.

Gene classes with expected connectivity In order to check if the lethality-centrality relationship of PPI networks exists in the postmortem gene interaction networks, we include list of genes whose proteins are classified as "essential" (n=118) (Liao and Zhang, 2008). While genes associated with severe diseases are frequently hubs in PPI's, they were not more or less connected than expected at random in the gene network. Similarly longevity-associated proteins (n=261) (de Magalhaes *et al*, 2009) are organizing hubs in PPI networks (Budovsky *et al*, 2007), however their associated genes are not hubs in the gene networks. Disease genes cataloged in OMIM (n=1646) (Hamosh *et al*, 2005) were at first largely monogenic (n=738) (Jimenez-Sanchez *et al*, 2001), but now include an increasing number of genes implicated in complex disorders (n=411) (Barrenas *et al*, 2009). While the complex disorder genes appear to be over-connected, there is no evidence of under or over-connectivity of the complete OMIM catalogue or of monogenic disease gene specifically.

The connectivity of specific gene classes indicates that these classes may operate in brain-region specific ways, which we speculate are related to their under- or over-connectivity in the respective systems. Additionally, these function-specific patterns of connectivity support the biological validity of our inferred networks, since many gene classes do exhibit non-random

connectivity and because categories of genes show similar levels of connectivity in similar brain regions.

Establishing significant connectivity by gene class For each class of genes we generate a null distribution of median connectivity values for a group of randomly selected genes of equal number to the special gene class in question. The percentile in which the actual median connectivity of the gene class falls determines its p-value for hub/non-hub connectivity signature (Table S4). Meta-p-values for connectivity of each class of gene were computed using the inverse normal method.

Supplementary Figures and Tables Captions

Brain Bank	Nominal author or reference	Species	Condition	array chipset	nominal brain region	ctrl samples	disease samples	# samples used in network	#network nodes	Opt. Pearson threshold	~scale-free?	mean clustering coefficient (CC)	mean non-zero CC	assortativity
Columbia	Sibille et al. 2004	Human	Depression	Affy U133a+2	BA 9	15	15	30	22277	0.82	yes	0.136	0.381	0.4916
Columbia	Sibille et al .2004	Human	Depression	Affy U133a+2	BA 47	15	15	30	22277	0.82	yes	0.148	0.419	0.5302
Pittsburgh	Sibille et al. 2009	Human	Depression	Affy U133a+2	basolateral amygdala	*14*	*14*	28	26199	0.85	yes	0.196	0.477	0.264
Pittsburgh	Sibille et al. 2009	Human	Depression	Affy U133a+2	anterior cingulate	*14*	*14*	28	26199	0.85	yes	0.1646	0.505	0.4738
Pittsburgh	Sibille et al. 2010 (forthcoming)	Human	Depression	Affy U133a+2	DLPFC	*29*	*29*	58	29211	0.8	yes	0.158	0.551	0.4639
Pittsburgh	Sibille et al. 2010 (forthcoming)	Human	Depression	Affy U133a+2	ant. cingulate ctx	*28*	*28*	56	29211	0.8	yes	0.143	0.529	0.3546
Stanley	Feinberg et al.	Human	Depression	Affy hgu95av2	cerebellum	*14*	*13*	*50*	12453	0.8	yes	0.15	0.501	0.4724
Stanley		Human	Bipolar	Affy hgu95av2	cerebellum		*12*	*50*			yes			
Stanley		Human	Schizophrenia	Affy hgu95av2	cerebellum		*11*	*50*			yes			
Stanley	Altar (B) et al.	Human	Depression	Agilent	BA 46/10	*10*	*20*	39	12235	0.82	yes	0.136	0.381	0.6056
Stanley		Human	Schizophrenia	Agilent	BA 46/10		*9*	39			yes			
Stanley	Altar (C) et al.	Human	Depression	Affy hgu133a	BA 46	*10*	*11*	44	22383	0.85	yes	0.19	0.503	0.2819
Stanley		Human	Bipolar	Affy hgu133a	BA 46		*12*	44			yes			
Stanley		Human	Schizophrenia	Affy hgu133a	BA 46		*11*	44			yes			
Stanley	Sklar (A) et al.	Human	Depression	Affy hgu95av2	BA 8/9	*12*	*11*	*47*	12453	0.79	yes	0.4383	0.6495	0.0854
Stanley		Human	Bipolar	Affy hgu95av2	BA 8/9		*11*	*47*			yes			
Stanley		Human	Schizophrenia	Affy hgu95av2	BA 8/9		*13*	*47*			yes			
Stanley	Sklar (B) et al.	Human	Depression	Affy hgu95av2	cerebellum	*10*	*13*	*46*	12453	0.84	yes	0.305	0.503	0.1989
Stanley		Human	Bipolar	Affy hgu95av2	cerebellum		*11*	*46*			yes			
Stanley		Human	Schizophrenia	Affy hgu95av2	cerebellum		*12*	*46*			yes			
Stanley	Iwamoto 2004	Human	Depression	Affy hgu95av2	BA 10	*15*	*11*	*50*	12453	0.82	yes	0.176	0.526	0.3846
Stanley		Human	Bipolar	Affy hgu95av2	BA 10		*11*	*50*			yes			
Stanley		Human	Schizophrenia	Affy hgu95av2	BA 10		*13*	*50*			yes			
Stanley	Aston 2004	Human	Depression	Affy hgu95av2	temporal ctx– BA 21	*19*	*13*	*59*	12453	0.79	yes	0.26	0.557	0.104
Stanley		Human	Bipolar	Affy hgu95av2	temporal ctx– BA 21		*12*	*59*			yes			
Stanley		Human	Schizophrenia	Affy hgu95av2	temporal ctx– BA 21		*15*	*59*			yes			
na	Sibille et al. 2009	Mouse	UCMS	MOE 430_2	amygdala	*6*	*6*	*24*	25859	0.83	yes	0.173	0.348	0.4998
na	Sibille et al. 2009	Mouse	UCMS + FLX	MOE 430_2	amygdala		*6*	*24*			yes			
na	Sibille et al. 2009	Mouse	UCMS + CRF1A	MOE 430_2	amygdala		*6*	*24*			yes			
na	Sibille et al. 2009	Mouse	UCMS	MOE 430_2	cingulate cortex	*6*	*6*	*24*	25859	0.78	yes	0.242	0.317	0.4829
na	Sibille et al. 2009	Mouse	UCMS + FLX	MOE 430_2	cingulate cortex		*6*	*24*			yes			
na	Sibille et al. 2009	Mouse	UCMS + CRF1A	MOE 430_2	cingulate cortex		*6*	*24*			yes			
na	Sibille et al. 2009	Mouse	UCMS	MOE 430_2	dentate gyrus	*6*	*6*	*24*	25859	0.87	yes	0.085	0.4316	0.65
na	Sibille et al. 2009	Mouse	UCMS + FLX	MOE 430_2	dentate gyrus		*6*	*24*			yes			
na	Sibille et al. 2009	Mouse	UCMS + CRF1A	MOE 430_2	dentate gyrus		*6*	*24*			yes			

Table 6 (Table S1) Summary of studies included in meta-analysis – array details, brain regions, disorders, and network parameters

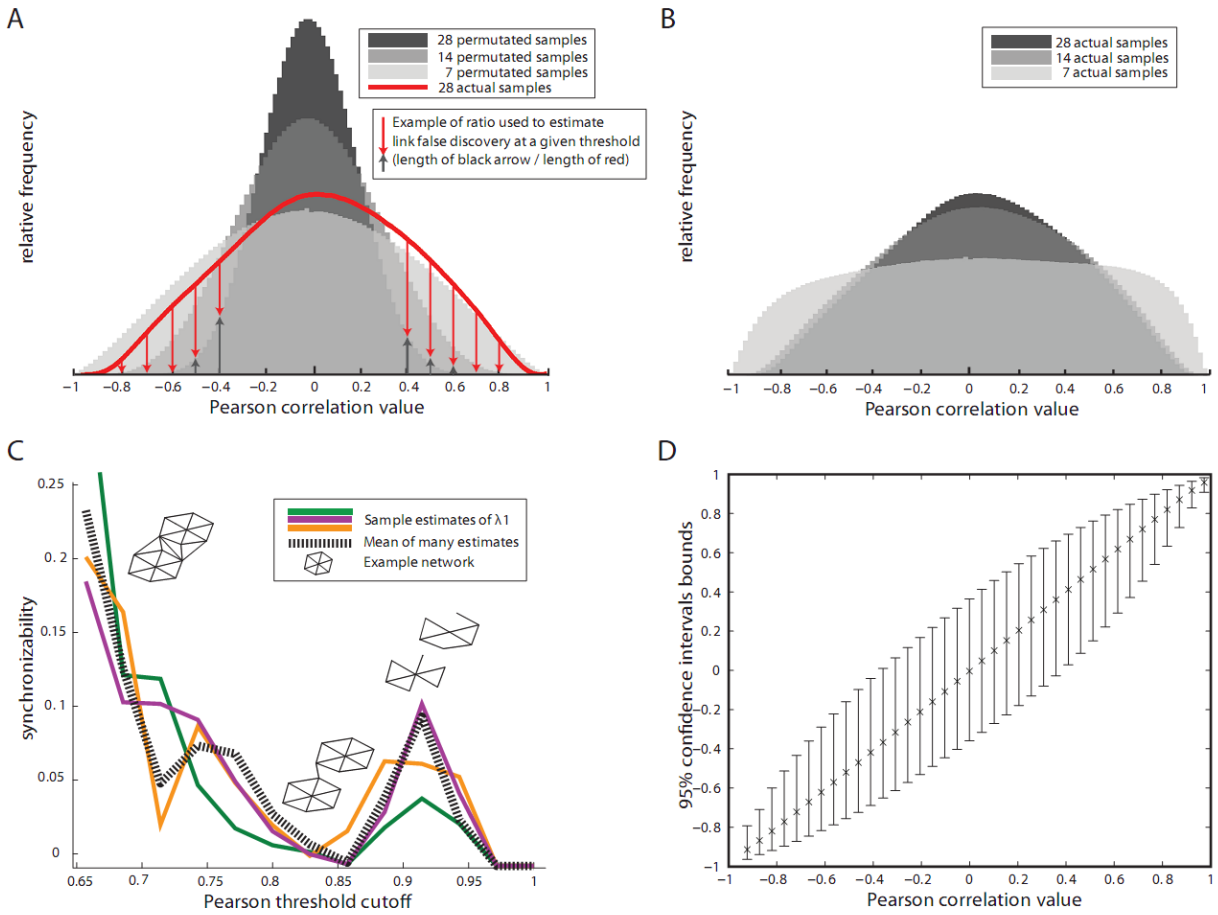


Figure 20 (Figure S2) Gene network validity and optimization – examples from human amygdala dataset

(A) Decreasing spurious network links with increasing sample size, shown by the null correlation distributions shrinking towards zero with increasing sample size, and the greater number of high correlations in real versus permutated data. (B) Decreasing false positive correlations in the actual data set with increasing sample size, shown as decrease in number of extreme correlations when comparing 7 to 14 samples, but which then remains constant between n=14 and n=28 samples. (C) Example estimates of network synchronizability (low synchronizability implies high modularity) at various thresholds in order to optimize correlation cutoff (example estimates based on different subsets of nodes are shown in different colors). Inset: schematic of link pruning and changes in modularity shown for increasing cutoffs. (D) 95% confidence bounds on Pearson correlation shrink for extreme correlation values (estimated

by resampling) indicating selected network links are robust since most optimized cutoffs are 0.8 or greater.

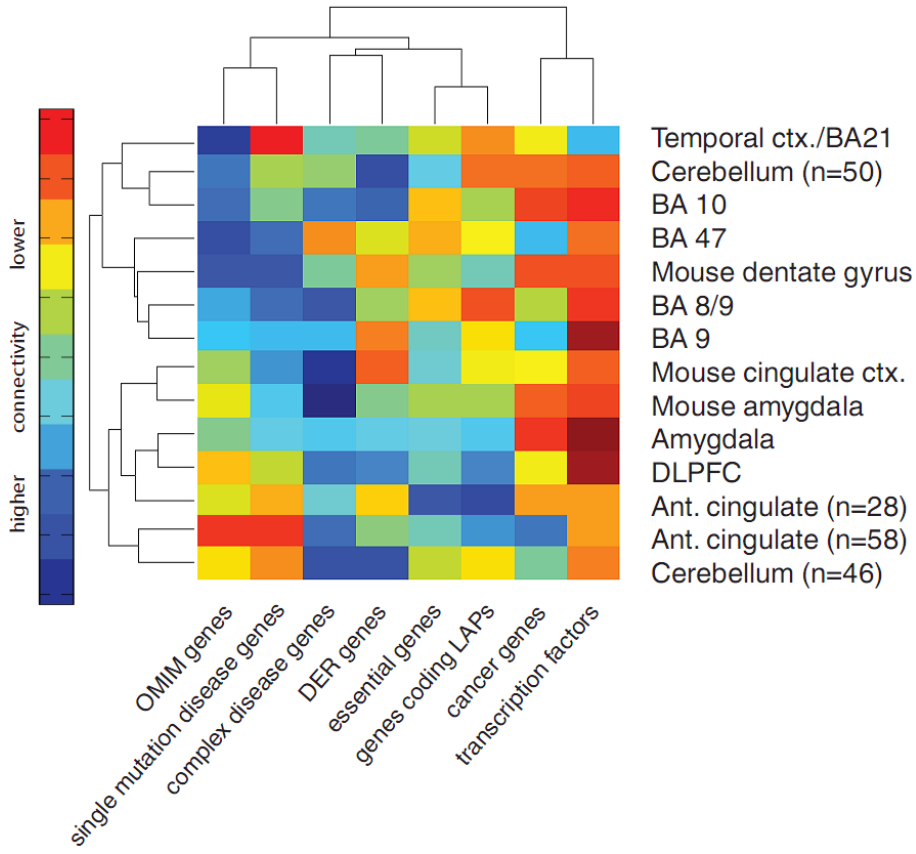


Figure 21 (Figure S3) Hierarchical clustering of brain regions by gene class connectivity

(A) Regions with similar patterns of connectivity across the different gene classes are listed proximally. Heatmap shows under-connected categories as orange and over-connected as blue.

p-values for greater than expected connectivity, 1-pvalues for less than expected connectivity

higher connectivity
lower connectivity

single mutation	OMIM	DER	complex	essential	LAP	cancer associated	transcription factor	Nominal Brain Region
0.135	0.575	0.33	0.35	0.42	0.505	0.44	0.27	temporal ctx / BA21
0.22	0.64	0.58	0	0.37	0.97	0.975	1	cerebellum n=50
0.035	0.415	0.065	0.005	0.765	0.5	0.955	1	BA 46
0.21	0.355	0.96	0.79	0.92	0.85	0.45	0.995	BA 47
0.155	0.17	0.535	0.885	0.62	0.485	1	1	mouse dentate gyrus
0.28	0.19	0.115	0.575	0.8	0.96	0.62	1	BA 8/9
0.005	0	0	0.345	0.075	0.265	0.005	0.52	BA 9
0.725	0.49	0.25	1	0.605	0.845	0.85	1	mouse cingulate ctx
0.755	0.44	0.04	0.585	0.67	0.66	0.965	1	mouse amygdala
0.185	0.025	0	0.015	0.04	0	0.755	1	amygdala
0.605	0.415	0	0.02	0.24	0.005	0.505	1	DLPFC
0.885	0.985	0.7	0.96	0.495	0.42	1	1	anterior cingulate n=28
1	0.985	0	0.42	0.315	0.075	0.02	0.815	anterior cingulate n=56
0.845	0.96	0.01	0.025	0.695	0.835	0.53	1	cerebellum n=46
median p-value	0.25	0.4275	0.09	0.385	0.55	0.5025	0.6875	1	
consensus p-value	0.1696	0.333	0	0.0242	0.4391	0.3764	0.9991	1	

Table 7 (Table S4) p-values for under- or over-connectivity of 8 important biological gene categories in each gene network. Low p-values represent over-connectivity and very high p-values represent under-connectivity.

REFERENCES

Agrawal H (2002) Extreme self-organization in networks constructed from gene expression data. *Phys Rev Lett* **89:** 268702.

Airan RD, Meltzer LA, Roy M, Gong Y, Chen H, Deisseroth K (2007) High-speed imaging reveals neurophysiological links to behavior in an animal model of depression. *Science* **317:** 819-823.

Aisa B, Tordera R, Lasheras B, Río JD, Ramírez MJ (2007) Cognitive impairment associated to HPA axis hyperactivity after maternal separation in rats. *Psychoneuroendocrinology* **32:** 256-266.

Akaike H (1974) A new look at the statistical model identification. *IEEE transat Autoamtic Control* **19:** 716-723.

Albert, Jeong, Barabasi (2000) Error and attack tolerance of complex networks. *Nature* **406:** 378-382.

Alexander RP, Kim PM, Emonet T, Gerstein MB (2009) Understanding modularity in molecular networks requires dynamics. *Sci Signal* **2:** pe44.

Alfonso J, Frasch AC, Flugge G (2005) Chronic stress, depression and antidepressants: effects on gene transcription in the hippocampus. *Rev Neurosci* **16:** 43-56.

Allman J, Hakeem A, Watson K (2002) Two phylogenetic specializations in the human brain. *Neuroscientist* **8:** 335-346.

Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P (2001) The anterior cingulate cortex. The evolution of an interface between emotion and cognition. *Ann N Y Acad Sci* **935:** 107-117.

Allocco DJ, Kohane IS, Butte AJ (2004) Quantifying the relationship between co-expression, co-regulation and gene function. *BMC Bioinformatics* **5:** 18.

Alonso G (2000) Prolonged corticosterone treatment of adult rats inhibits the proliferation of oligodendrocyte progenitors present throughout white and gray matter regions of the brain. *Glia* **31:** 219-231.

Altar CA, Hunt RA, Jurata LW, Webster MJ, Derby E, Gallagher P, Lemire A, Brockman J, Laeng P (2008) Insulin, IGF-1, and muscarinic agonists modulate schizophrenia-associated genes in human neuroblastoma cells. *Biol Psychiatry* **64:** 1077-1087.

Altman RB, Raychaudhuri S (2001) Whole-genome expression analysis: challenges beyond clustering. *Curr Opin Struct Biol* **11:** 340-347.

Andreasen NC, Endicott J, Spitzer RL, Winokur G (1977) The family history method using diagnostic criteria. Reliability and validity. *Arch Gen Psychiatry* **34:** 1229-1235.

Angelucci F, Brenè S, Mathé AA (2005) BDNF in schizophrenia, depression and corresponding animal models. *Mol Psychiatry* **10:** 345-352.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25:** 25-29.

Aston C, Jiang L, Sokolov BP (2005) Transcriptional profiling reveals evidence for signaling and oligodendroglial abnormalities in the temporal cortex from patients with major depressive disorder. *Mol Psychiatry* **10:** 309-322.

Ayuso-Gutiérrezd JL (2005) Depressive subtypes and efficacy of antidepressive pharmacotherapy. *World J Biol Psychiatry* **6 Suppl 2:** 31-37.

Babu MM, Janga SC, de Santiago I, Pombo A (2008) Eukaryotic gene regulation in three dimensions and its impact on genome evolution. *Curr Opin Genet Dev* **18:** 571-582.

Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA (2004) Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14:** 283-291.

Bamberger CM, Schulte HM, Chrousos GP (1996) Molecular determinants of glucocorticoid receptor function and tissue sensitivity to glucocorticoids. *Endocr Rev* **17:** 245-261.

Banasr M, Chowdhury GMI, Terwilliger R, Newton SS, Duman RS, Behar KL, Sanacora G (2010) Glial pathology in an animal model of depression: reversal of stress-induced cellular, metabolic and behavioral deficits by the glutamate-modulating drug riluzole. *Mol Psychiatry* **15:** 501-511.

Banasr M, Duman RS (2007) Regulation of neurogenesis and gliogenesis by stress and antidepressant treatment. *CNS Neurol Disord Drug Targets* **6:** 311-320.

Banasr M, Duman RS (2008) Glial loss in the prefrontal cortex is sufficient to induce depressive-like behaviors. *Biol Psychiatry* **64:** 863-870.

Banks SJ, Eddy KT, Angstadt M, Nathan PJ, Phan KL (2007) Amygdala-frontal connectivity during emotion regulation. *Soc Cogn Affect Neurosci* **2:** 303-312.

Barabasi, Albert (1999) Emergence of scaling in random networks. *Science* **286:** 509-512.

Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* **5:** 101-113.

Barlev NA, Liu L, Chehab NH, Mansfield K, Harris KG, Halazonetis TD, Berger SL (2001) Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol Cell* **8:** 1243-1254.

Barrenas F, Chavali S, Holme P, Mobini R, Benson M (2009) Network properties of complex human disease genes identified through genome-wide association studies. *PLoS ONE* **4:** e8090.

Baskerville S, Bartel DP (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11:** 241-247.

Batada NN, Urrutia AO, Hurst LD (2007) Chromatin remodelling is a major source of coexpression of linked genes in yeast. *Trends Genet* **23:** 480-484.

Bauer M, Whybrow PC (2003) Thyroid hormone and mood modulation: new insights from functional brain imaging techniques. *Curr Psychiatry Rep* **5:** 163-165.

Bauer M, Whybrow PC, Angst J, Versiani M, Möller H-J, of Societies of Biological Psychiatry Task Force on Treatment Guidelines for Unipolar Depressive Disorders WF (2002) World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for Biological Treatment of Unipolar Depressive Disorders, Part 2: Maintenance treatment of major depressive disorder and treatment of chronic depressive disorders and subthreshold depressions. *World J Biol Psychiatry* **3:** 69-86.

Baxter LR, Schwartz JM, Phelps ME, Mazziotta JC, Guze BH, Selin CE, Gerner RH, Sumida RM (1989) Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Arch Gen Psychiatry* **46:** 243-250.

Belmaker RH, Agam G (2008) Major depressive disorder. *N Engl J Med* **358:** 55-68.

Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* **125:** 279-284.

Bergmann S, Ihmels J, Barkai N (2004) Similarities and differences in genome-wide expression data of six organisms. *PLoS Biol* **2:** E9.

Berto P, D'Ilario D, Ruffo P, Virgilio RD, Rizzo F (2000) Depression: cost-of-illness studies in the international literature, a review. *J Ment Health Policy Econ* **3:** 3-10.

Bet PM, Penninx BWJH, Bochdanovits Z, Uitterlinden AG, Beekman ATF, van Schoor NM, Deeg DJH, Hoogendijk WJG (2009) Glucocorticoid receptor gene polymorphisms and childhood adversity are associated with depression: New evidence for a gene-environment interaction. *Am J Med Genet B Neuropsychiatr Genet* **150B:** 660-669.

Bezchlibnyk YB, Wang JF, McQueen GM, Young LT (2001) Gene expression differences in bipolar disorder revealed by cDNA array analysis of post-mortem frontal cortex. *J Neurochem* **79:** 826-834.

Bhagwagar Z, Wylezinska M, Jezzard P, Evans J, Ashworth F, Sule A, Matthews PM, Cowen PJ (2007) Reduction in occipital cortex gamma-aminobutyric acid concentrations in medication-free recovered unipolar depressed and bipolar subjects. *Biol Psychiatry* **61:** 806-812.

Bhardwaj N, Lu H (2005) Correlation between gene expression profiles and protein-protein interactions within and across genomes. *Bioinformatics* **21:** 2730-2738.

Bhardwaj N, Lu H (2009) Co-expression among constituents of a motif in the protein-protein interaction network. *J Bioinform Comput Biol* **7:** 1-17.

Bikeye S-NN, Colin C, Marie Y, Vampouille R, Ravassard P, Rousseau A, Boisselier B, Idbaih A, Calvo CF, Leuraud P, Lassalle M, Hallani SE, Delattre J-Y, Sanson M (2010) ASPM-associated stem cell proliferation is involved in malignant progression of gliomas and constitutes an attractive therapeutic target. *Cancer Cell Int* **10:** 1.

Borate BR, Chesler EJ, Langston MA, Saxton AM, Voy BH (2009) Comparison of threshold selection methods for microarray gene co-expression matrices. *BMC Res Notes* **2:** 240.

Bosker FJ, Hartman CA, Nolte IM, Prins BP, Terpstra P, Posthuma D, van Veen T, Willemsen G, Derijk RH, de Geus EJ, Hoogendijk WJ, Sullivan PF, Penninx BW, Boomsma DI, Snieder H, Nolen WA (2010) Poor replication of candidate genes for major depressive disorder using genome-wide association data. *Mol Psychiatry*.

Botteron KN, Raichle ME, Drevets WC, Heath AC, Todd RD (2002) Volumetric reduction in left subgenual prefrontal cortex in early onset depression. *Biol Psychiatry* **51:** 342-344.

Boutanaev AM, Kalmykova AI, Shevelyov YY, Nurminsky DI (2002) Large clusters of co-expressed genes in the Drosophila genome. *Nature* **420:** 666-669.

Bowery N (1989) GABAB receptors and their significance in mammalian pharmacology. *Trends Pharmacol Sci* **10:** 401-407.

Bowley MP, Drevets WC, Ongür D, Price JL (2002) Low glial numbers in the amygdala in major depressive disorder. *Biol Psychiatry* **52:** 404-412.

Brazma A, Jonassen I, Vilo J, Ukkonen E (1998) Predicting gene regulatory elements in silico on a genomic scale. *Genome Res* **8:** 1202-1215.

Budovsky A, Abramovich A, Cohen R, Chalifa-Caspi V, Fraifeld V (2007) Longevity network: construction and implications. *Mech Ageing Dev* **128:** 117-124.

Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* **10:** 186-198.

Byrne A, Weirauch M, Wong V, Koeva M, Dixon S, Stuart J, Roy P (2007) A global analysis of genetic interactions in Caenorhabditis elegans. *J Biol* **6:** 8.

Campbell S, MacQueen G (2006) An update on regional brain volume differences associated with mood disorders. *Curr Opin Psychiatry* **19:** 25-33.

Carlson MRJ, Zhang B, Fang Z, Mischel PS, Horvath S, Nelson SF (2006) Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks. *BMC Genomics* **7:** 40.

Caron H, van Schaik B, van der Mee M, Baas F, Riggins G, van Sluis P, Hermus MC, van Asperen R, Boon K, Voûte PA, Heisterkamp S, van Kampen A, Versteeg R (2001) The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291:** 1289-1292.

Carpenter LL, Anderson GM, Pelton GH, Gudin JA, Kirwin PD, Price LH, Heninger GR, McDougle CJ (1998) Tryptophan depletion during continuous CSF sampling in healthy human subjects. *Neuropsychopharmacology* **19:** 26-35.

Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122:** 33-43.

Carrozza MJ, Utley RT, Workman JL, Côté J (2003) The diverse functions of histone acetyltransferase complexes. *Trends Genet* **19:** 321-329.

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301:** 386-389.

Castrén E (2004) Neurotrophic effects of antidepressant drugs. *Curr Opin Pharmacol* **4:** 58-64.

Chana G, Landau S, Beasley C, Everall IP, Cotter D (2003) Two-dimensional assessment of cytoarchitecture in the anterior cingulate cortex in major depressive disorder, bipolar disorder, and schizophrenia: evidence for decreased neuronal somal size and increased neuronal density. *Biol Psychiatry* **53:** 1086-1098.

Chen AC, Shirayama Y, Shin KH, Neve RL, Duman RS (2001a) Expression of the cAMP response element binding protein (CREB) in hippocampus produces an antidepressant effect. *Biol Psychiatry* **49:** 753-762.

Chen B, Dowlatshahi D, MacQueen GM, Wang JF, Young LT (2001b) Increased hippocampal BDNF immunoreactivity in subjects treated with antidepressant medication. *Biol Psychiatry* **50:** 260-265.

Chen L, Lawlor DA, Lewis SJ, Yuan W, Abdollahi MR, Timpson NJ, Day INM, Ebrahim S, Smith GD, Shugart YY (2008a) Genetic association study of BDNF in depression: Finding from two cohort studies and a meta-analysis. *Am J Med Genet B Neuropsychiatr Genet*.

Chen L, Zhao H (2005) Gene expression analysis reveals that histone deacetylation sites may serve as partitions of chromatin gene expression domains. *BMC Genomics* **6:** 44.

Chen R, Morgan AA, Dudley J, Deshpande T, Li L, Kodama K, Chiang AP, Butte AJ (2008b) FitSNPs: highly differentially expressed genes are more likely to have variants associated with disease. *Genome Biol* **9:** R170.

Chen W-H, de Meaux J, Lercher MJ (2010) Co-expression of neighbouring genes in Arabidopsis: separating chromatin effects from direct interactions. *BMC Genomics* **11:** 178.

chi Hsiung S, Adlersberg M, Arango V, Mann JJ, Tamir H, peing Liu K (2003) Attenuated 5-HT1A receptor signaling in brains of suicide victims: involvement of adenylyl cyclase, phosphatidylinositol 3-kinase, Akt and mitogen-activated protein kinase. *J Neurochem* **87:** 182-194.

Chia BKH, Karuturi RKM (2010) Differential co-expression framework to quantify goodness of biclusters and compare biclustering algorithms. *Algorithms Mol Biol* **5:** 23.

Choi JK, Yu U, Yoo OJ, Kim S (2005) Differential coexpression analysis using microarray data and its application to human cancer. *Bioinformatics* **21:** 4348-4355.

Choudary PV, Molnar M, Evans SJ, Tomita H, Li JZ, Vawter MP, Myers RM, Bunney WE, Akil H, Watson SJ, Jones EG (2005) Altered cortical glutamatergic and GABAergic signal transmission with glial involvement in depression. *Proc Natl Acad Sci U S A* **102:** 15653-15658.

Cipriani A, Brambilla P, Furukawa T, Geddes J, Gregis M, Hotopf M, Malvini L, Barbui C (2005) Fluoxetine versus other types of pharmacotherapy for depression. *Cochrane Database Syst Rev***:** CD004185.

Cohen BA, Mitra RD, Hughes JD, Church GM (2000) A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression. *Nat Genet* **26:** 183-186.

Conti B, Maier R, Barr AM, Morale MC, Lu X, Sanna PP, Bilbe G, Hoyer D, Bartfai T (2007) Region-specific transcriptional changes following the three antidepressant treatments electro convulsive therapy, sleep deprivation and fluoxetine. *Mol Psychiatry* **12:** 167-189.

Cooney RE, Joormann J, Eugène F, Dennis EL, Gotlib IH (2010) Neural correlates of rumination in depression. *Cogn Affect Behav Neurosci* **10:** 470-478.

Cotter D, Mackay D, Chana G, Beasley C, Landau S, Everall IP (2002) Reduced neuronal size and glial cell density in area 9 of the dorsolateral prefrontal cortex in subjects with major depressive disorder. *Cereb Cortex* **12:** 386-394.

Cotter D, Mackay D, Landau S, Kerwin R, Everall I (2001) Reduced glial cell density and neuronal size in the anterior cingulate cortex in major depressive disorder. *Arch Gen Psychiatry* **58:** 545-553.

Covington HE, Maze I, LaPlant QC, Vialou VF, Ohnishi YN, Berton O, Fass DM, Renthal W, Rush AJ, Wu EY, Ghose S, Krishnan V, Russo SJ, Tamminga C, Haggarty SJ, Nestler EJ (2009) Antidepressant actions of histone deacetylase inhibitors. *J Neurosci* **29:** 11451-11460.

Covington HE, Vialou V, Nestler EJ (2010) From synapse to nucleus: novel targets for treating depression. *Neuropharmacology* **58:** 683-693.

Croarkin PE, Levinson AJ, Daskalakis ZJ (2010) Evidence for GABAergic inhibitory deficits in major depressive disorder. *Neurosci Biobehav Rev* **35:** 818-825.

Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW (2008) From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat Rev Neurosci* **9:** 46-56.

David DJ, Samuels BA, Rainer Q, Wang J-W, Marsteller D, Mendez I, Drew M, Craig DA, Guiard BP, Guilloux J-P, Artymyshyn RP, Gardier AM, Gerald C, Antonijevic IA, Leonardo ED, Hen R (2009) Neurogenesis-dependent and -independent effects of fluoxetine in an animal model of anxiety/depression. *Neuron* **62:** 479-493.

Day A, Dong J, Funari VA, Harry B, Strom SP, Cohn DH, Nelson SF (2009) Disease gene characterization through large-scale co-expression analysis. *PLoS ONE* **4:** e8491.

Dayan P, Huys QJM (2009) Serotonin in Affective Control. *Annu Rev Neurosci*.

de Kloet ER, Joëls M, Holsboer F (2005) Stress and the brain: from adaptation to disease. *Nat Rev Neurosci* **6:** 463-475.

de la Fuente A (2010) From 'differential expression' to 'differential networking' - identification of dysfunctional regulatory networks in diseases. *Trends Genet* **26:** 326-333.

de Magalhaes JP, Budovsky A, Lehmann G, Costa J, Li Y, Fraifeld V, Church GM (2009) The Human Ageing Genomic Resources: online databases and tools for biogerontologists. *Aging Cell* **8:** 65-72.

De Olmos JS (ed) (2004) *Amygdala*. San Diego, CA: Elsevier, 739-868pp.

de Wit E, Braunschweig U, Greil F, Bussemaker HJ, van Steensel B (2008) Global chromatin domain organization of the Drosophila genome. *PLoS Genet* **4:** e1000045.

Dedovic K, Duchesne A, Andrews J, Engert V, Pruessner JC (2009) The brain and the stress axis: the neural correlates of cortisol regulation in response to stress. *Neuroimage* **47:** 864-871.

Dekker J (2008) Gene regulation in the third dimension. *Science* **319:** 1793-1794.

del Rio G, Koschützki D, Coello G (2009) How to identify essential genes from molecular networks? *BMC Syst Biol* **3:** 102.

Dell'Osso L, Del Debbio A, Veltri A, Bianchi C, Roncaglia I, Carlini M, Massimetti G, Catena Dell'Osso M, Vizzaccaro C, Marazziti D, Piccinni A (2010) Associations between brain-derived neurotrophic factor plasma levels and severity of the illness, recurrence and symptoms in depressed patients. *Neuropsychobiology* **62:** 207-212.

Deng J, Deng L, Su S, Zhang M, Lin X, Wei L, Minai AA, Hassett DJ, Lu LJ (2010a) Investigating the predictability of essential genes across distantly related organisms using an integrative approach. *Nucleic Acids Res*.

Deng Y, Dai X, Xiang Q, Dai Z, He C, Wang J, Feng J (2010b) Genome-wide analysis of the effect of histone modifications on the coexpression of neighboring genes in Saccharomyces cerevisiae. *BMC Genomics* **11:** 550.

Denys D, Mantione M, Figee M, van den Munckhof P, Koerselman F, Westenberg H, Bosch A, Schuurman R (2010) Deep brain stimulation of the nucleus accumbens for treatment-refractory obsessive-compulsive disorder. *Arch Gen Psychiatry* **67:** 1061-1068.

Dobrin R, Zhu J, Molony C, Argman C, Parrish ML, Carlson S, Allan MF, Pomp D, Schadt EE (2009) Multi-tissue coexpression networks reveal unexpected subnetworks associated with disease. *Genome Biol* **10:** R55.

Donohue JM, Pincus HA (2007) Reducing the societal burden of depression: a review of economic costs, quality of care and effects of treatment. *Pharmacoeconomics* **25:** 7-24.

Dorogovtsev SN, Goltsev AV, Mendes JFF (2002) Pseudofractal scale-free web. *Phys Rev E Stat Nonlin Soft Matter Phys* **65:** 066122.

Dowlatshahi D, MacQueen GM, Wang JF, Young LT (1998) Increased temporal cortex CREB concentrations and antidepressant treatment in major depression. *Lancet* **352:** 1754-1755.

Drevets WC (1999) Prefrontal cortical-amygdalar metabolism in major depression. *Ann N Y Acad Sci* **877:** 614-637.

Drevets WC (2003) Neuroimaging abnormalities in the amygdala in mood disorders. *Ann N Y Acad Sci* **985:** 420-444.

Drevets WC, Bogers W, Raichle ME (2002) Functional anatomical correlates of antidepressant drug treatment assessed using PET measures of regional glucose metabolism. *Eur Neuropsychopharmacol* **12:** 527-544.

Drevets WC, Price JL, Simpson JR, Todd RD, Reich T, Vannier M, Raichle ME (1997) Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* **386:** 824-827.

Drigues N, Poltyrev T, Bejar C, Weinstock M, Youdim MBH (2003) cDNA gene expression profile of rat hippocampus after chronic treatment with antidepressant drugs. *J Neural Transm* **110:** 1413-1436.

Duman RS, Heninger GR, Nestler EJ (1997) A molecular and cellular theory of depression. *Arch Gen Psychiatry* **54:** 597-606.

Duvarci S, Paré D (2007) Glucocorticoids enhance the excitability of principal basolateral amygdala neurons. *J Neurosci* **27:** 4482-4491.

Ebisuya M, Yamamoto T, Nakajima M, Nishida E (2008) Ripples from neighbouring transcription. *Nat Cell Biol* **10:** 1023-1024.

Egeland M, Warner-Schmidt J, Greengard P, Svenningsson P (2010) Neurogenic effects of fluoxetine are attenuated in p11 (S100A10) knockout mice. *Biol Psychiatry* **67:** 1048-1056.

Eggan SM, Hashimoto T, Lewis DA (2008) Reduced cortical cannabinoid 1 receptor messenger RNA and protein expression in schizophrenia. *Arch Gen Psychiatry* **65:** 772-784.

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95:** 14863-14868.

Eller T, Vasar V, Shlik J, Maron E (2008) Pro-inflammatory cytokines and treatment response to escitalopram in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* **32:** 445-450.

Elo LL, Järvenpää H, Oresic M, Lahesmaa R, Aittokallio T (2007) Systematic construction of gene coexpression networks with applications to human T helper cell differentiation process. *Bioinformatics* **23:** 2096-2103.

Engel JD, Tanimoto K (2000) Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. *Cell* **100:** 499-502.

Ensel WM, Lin N (2000) Age, the stress process, and physical distress: the role of distal stressors. *J Aging Health* **12:** 139-168.

Erraji-Benchekroun L, Underwood MD, Arango V, Galfalvy H, Pavlidis P, Smyrniotopoulos P, Mann JJ, Sibille E (2005) Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. *Biol Psychiatry* **57:** 549-558.

Fava M, Kendler KS (2000) Major depressive disorder. *Neuron* **28:** 335-341.

Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, Biggs MM, Zisook S, Leuchter A, Howland R, Warden D, Trivedi MH (2008) Difference in Treatment Outcome in Outpatients With Anxious Versus Nonanxious Depression: A STAR*D Report. *Am J Psychiatry* **165:** 342-351.

Feldman I, Rzhetsky A, Vitkup D (2008) Network properties of genes harboring inherited disease mutations. *Proc Natl Acad Sci U S A*.

Floresco SB, Tse MT (2007) Dopaminergic regulation of inhibitory and excitatory transmission in the basolateral amygdala-prefrontal cortical pathway. *J Neurosci* **27:** 2045-2057.

Franzen PL, Buysse DJ, Rabinovitz M, Pollock BG, Lotrich FE (2010) Poor sleep quality predicts onset of either major depression or subsyndromal depression with irritability during interferon-alpha treatment. *Psychiatry Res* **177:** 240-245.

Freemantle N, Mason J (2000) The importance of achieving additional drug benefits at a reasonable cost. A review of the fluoxetine years. *Pharmacoeconomics* **17:** 319-324.

Frenois F, Moreau M, O'Connor J, Lawson M, Micon C, Lestage J, Kelley KW, Dantzer R, Castanon N (2007) Lipopolysaccharide induces delayed FosB/DeltaFosB immunostaining within the mouse extended amygdala, hippocampus and hypothalamus, that parallel the expression of depressive-like behavior. *Psychoneuroendocrinology* **32:** 516-531.

Friedman A, Deri I, Friedman Y, Dremencov E, Goutkin S, Kravchinsky E, Mintz M, Levi D, Overstreet DH, Yadid G (2007) Decoding of dopaminergic mesolimbic activity and depressive behavior. *J Mol Neurosci* **32:** 72-79.

Fuller TF, Ghazalpour A, Aten JE, Drake TA, Lusis AJ, Horvath S (2007) Weighted gene coexpression network analysis strategies applied to mouse weight. *Mamm Genome* **18:** 463-472.

Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR (2004) A census of human cancer genes. *Nat Rev Cancer* **4:** 177-183.

Gaiteri C, Guilloux J-P, Lewis DA, Sibille E (2010) Altered gene synchrony suggests a combined hormone-mediated dysregulated state in major depression. *PLoS ONE* **5:** e9970.

Galfalvy HC, Erraji-Benchekroun L, Smyrniotopoulos P, Pavlidis P, Ellis SP, Mann JJ, Sibille E, Arango V (2003) Sex genes for genomic analysis in human brain: internal controls for comparison of probe level data extraction. *BMC Bioinformatics* **4:** 37.

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* **11:** 4241-4257.

Gaynes BN, Rush AJ, Trivedi MH, Wisniewski SR, Spencer D, Fava M (2008) The STAR*D study: treating depression in the real world. *Cleve Clin J Med* **75:** 57-66.

Ge H, Liu Z, Church GM, Vidal M (2001) Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. *Nat Genet* **29:** 482-486.

Gerner RH, Hare TA (1981) CSF GABA in normal subjects and patients with depression, schizophrenia, mania, and anorexia nervosa. *Am J Psychiatry* **138:** 1098-1101.

Gershon AA, Vishne T, Grunhaus L (2007) Dopamine D2-like receptors and the antidepressant response. *Biol Psychiatry* **61:** 145-153.

Giacobbe P, Mayberg HS, Lozano AM (2009) Treatment resistant depression as a failure of brain homeostatic mechanisms: implications for deep brain stimulation. *Exp Neurol* **219:** 44-52.

Gibson G (2009) Decanalization and the origin of complex disease. *Nat Rev Genet* **10:** 134-140.

Gillis J, Pavlidis P (2009) A methodology for the analysis of differential coexpression across the human lifespan. *BMC Bioinformatics* **10:** 306.

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci U S A* **99:** 7821-7826.

Glantz LA, Lewis DA (1997) Reduction of synaptophysin immunoreactivity in the prefrontal cortex of subjects with schizophrenia. Regional and diagnostic specificity. *Arch Gen Psychiatry* **54:** 943-952.

Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L (2007) The human disease network. *Proc Natl Acad Sci U S A* **104:** 8685-8690.

Goshen I, Yirmiya R (2009) Interleukin-1 (IL-1): a central regulator of stress responses. *Front Neuroendocrinol* **30:** 30-45.

Greenbaum D, Colangelo C, Williams K, Gerstein M (2003) Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol* **4:** 117.

Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, Reiss AL, Schatzberg AF (2007) Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol Psychiatry* **62:** 429-437.

Guelzim N, Bottani S, Bourgine P, Képès F (2002) Topological and causal structure of the yeast transcriptional regulatory network. *Nat Genet* **31:** 60-63.

Gundersen BB, Blendy JA (2009) Effects of the histone deacetylase inhibitor sodium butyrate in models of depression and anxiety. *Neuropharmacology* **57:** 67-74.

Gustin M-P, Paultre CZ, Randon J, Bricca G, Cerutti C (2008) Functional meta-analysis of double connectivity in gene coexpression networks in mammals. *Physiol Genomics* **34:** 34-41.

Guye M, Bettus G, Bartolomei F, Cozzone PJ (2010) Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. *MAGMA*.

Hamidi M, Drevets WC, Price JL (2004) Glial reduction in amygdala in major depressive disorder is due to oligodendrocytes. *Biol Psychiatry* **55:** 563-569.

Hammen C (2005) Stress and depression. *Annu Rev Clin Psychol* **1:** 293-319.

Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA (2005) Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* **33:** D514-517.

Hardingham GE, Bading H (2010) Synaptic versus extrasynaptic NMDA receptor signalling: implications for neurodegenerative disorders. *Nat Rev Neurosci* **11:** 682-696.

Hardingham GE, Fukunaga Y, Bading H (2002) Extrasynaptic NMDARs oppose synaptic NMDARs by triggering CREB shut-off and cell death pathways. *Nat Neurosci* **5:** 405-414.

Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D, Egan MF, Weinberger DR (2002) Serotonin transporter genetic variation and the response of the human amygdala. *Science* **297:** 400-403.

Hartman SE, Bertone P, Nath AK, Royce TE, Gerstein M, Weissman S, Snyder M (2005) Global changes in STAT target selection and transcription regulation upon interferon treatments. *Genes Dev* **19:** 2953-2968.

Hasler G, Drevets WC, Manji HK, Charney DS (2004) Discovering endophenotypes for major depression. *Neuropsychopharmacology* **29:** 1765-1781.

He X, Zhang J (2006) Why do hubs tend to be essential in protein networks? *PLoS Genet* **2:** e88.

Hellsten J, West MJ, Arvidsson A, Ekstrand J, Jansson L, Wennström M, Tingström A (2005) Electroconvulsive seizures induce angiogenesis in adult rat hippocampus. *Biol Psychiatry* **58:** 871-878.

Hinkelmann K, Moritz S, Botzenhardt J, Riedesel K, Wiedemann K, Kellner M, Otte C (2009) Cognitive impairment in major depression: association with salivary cortisol. *Biol Psychiatry* **66:** 879-885.

Holden C (2000) Mental health. Global survey examines impact of depression. *Science* **288:** 39-40.

Holick KA, Lee DC, Hen R, Dulawa SC (2008) Behavioral effects of chronic fluoxetine in BALB/cJ mice do not require adult hippocampal neurogenesis or the serotonin 1A receptor. *Neuropsychopharmacology* **33:** 406-417.

Hollon SD, Shelton RC, Wisniewski S, Warden D, Biggs MM, Friedman ES, Husain M, Kupfer DJ, Nierenberg AA, Petersen TJ, Shores-Wilson K, Rush AJ (2006) Presenting characteristics of depressed outpatients as a function of recurrence: preliminary findings from the STAR*D clinical trial. *J Psychiatr Res* **40:** 59-69.

Holsboer F (2000) The corticosteroid receptor hypothesis of depression. *Neuropsychopharmacology* **23:** 477-501.

Holsboer F (2008) How can we realize the promise of personalized antidepressant medicines? *Nat Rev Neurosci* **9:** 638-646.

Holsboer F, Ising M (2008) Central CRH system in depression and anxiety--evidence from clinical studies with CRH1 receptor antagonists. *Eur J Pharmacol* **583:** 350-357.

Hooley JM, Gruber SA, Parker HA, Guillaumot J, Rogowska J, Yurgelun-Todd DA (2009) Cortico-limbic response to personally challenging emotional stimuli after complete recovery from depression. *Psychiatry Res* **171:** 106-119.

Horn DI, Yu C, Steiner J, Buchmann J, Kaufmann J, Osoba A, Eckert U, Zierhut KC, Schiltz K, He H, Biswal B, Bogerts B, Walter M (2010) Glutamatergic and resting-state functional connectivity correlates of severity in major depression - the role of pregenual anterior cingulate cortex and anterior insula. *Front Syst Neurosci* **4**.

Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, Laurance MF, Zhao W, Qi S, Chen Z, Lee Y, Scheck AC, Liau LM, Wu H, Geschwind DH, Febbo PG, Kornblum HI, Cloughesy TF, Nelson SF, Mischel PS (2006) Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci U S A* **103:** 17402-17407.

Hsu DT, Langenecker SA, Kennedy SE, Zubieta J-K, Heitzeg MM (2010) fMRI BOLD responses to negative stimuli in the prefrontal cortex are dependent on levels of recent negative life stress in major depressive disorder. *Psychiatry Res* **183:** 202-208.

Hu TW (2006) Perspectives: an international review of the national cost estimates of mental illness, 1990-2003. *J Ment Health Policy Econ* **9:** 3-13.

Hu W, Zhang M, Czeh B, Flugge G, Zhang W (2010) Stress impairs GABAergic network function in the hippocampus by activating nongenomic glucocorticoid receptors and affecting the integrity of the parvalbumin-expressing neuronal network. *Neuropsychopharmacology* **35:** 1693-1707.

Hudson NJ, Reverter A, Dalrymple BP (2009) A differential wiring analysis of expression data correctly identifies the gene containing the causal mutation. *PLoS Comput Biol* **5:** e1000382.

Huot RL, Plotsky PM, Lenox RH, McNamara RK (2002) Neonatal maternal separation reduces hippocampal mossy fiber density in adult Long Evans rats. *Brain Res* **950:** 52-63.

Ihmels J, Levy R, Barkai N (2004) Principles of transcriptional control in the metabolic network of Saccharomyces cerevisiae. *Nat Biotechnol* **22:** 86-92.

Inglis FM, Crockett R, Korada S, Abraham WC, Hollmann M, Kalb RG (2002) The AMPA receptor subunit GluR1 regulates dendritic architecture of motor neurons. *J Neurosci* **22:** 8042-8051.

Iversen LL (1965) The inhibition of noradrenaline uptake by drugs. *Adv Drug Res* **2:** 1-46.

Iwamoto K, Kakiuchi C, Bundo M, Ikeda K, Kato T (2004) Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol Psychiatry* **9:** 406-416.

Iwamoto K, Kato T (2006) Gene expression profiling in schizophrenia and related mental disorders. *Neuroscientist* **12:** 349-361.

Janga SC, Collado-Vides J, Babu MM (2008) Transcriptional regulation constrains the organization of genes on eukaryotic chromosomes. *Proc Natl Acad Sci U S A* **105:** 15761-15766.

Jeong H, Mason SP, Barabási AL, Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* **411:** 41-42.

Jimenez-Sanchez G, Childs B, Valle D (2001) Human disease genes. *Nature* **409:** 853-855.

Johnson MK, Nolen-Hoeksema S, Mitchell KJ, Levin Y (2009) Medial cortex activity, self-reflection and depression. *Soc Cogn Affect Neurosci* **4:** 313-327.

Jordan IK, Mariño-Ramírez L, Wolf YI, Koonin EV (2004) Conservation and coevolution in the scale-free human gene coexpression network. *Mol Biol Evol* **21:** 2058-2070.

Jothi R, Balaji S, Wuster A, Grochow JA, Gsponer J, Przytycka TM, Aravind L, Babu MM (2009) Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Mol Syst Biol* **5:** 294.

Kang HJ, Adams DH, Simen A, Simen BB, Rajkowska G, Stockmeier CA, Overholser JC, Meltzer HY, Jurjus GJ, Konick LC, Newton SS, Duman RS (2007) Gene expression profiling in postmortem prefrontals cortex of major depressive disorder. *J Neurosci* **27:** 13329-13340.

Kendler KS, Hettema JM, Butera F, Gardner CO, Prescott CA (2003a) Life event dimensions of loss, humiliation, entrapment, and danger in the prediction of onsets of major depression and generalized anxiety. *Arch Gen Psychiatry* **60:** 789-796.

Kendler KS, Liu X-Q, Gardner CO, McCullough ME, Larson D, Prescott CA (2003b) Dimensions of religiosity and their relationship to lifetime psychiatric and substance use disorders. *Am J Psychiatry* **160:** 496-503.

Kent S, Bluthé RM, Kelley KW, Dantzer R (1992) Sickness behavior as a new target for drug development. *Trends Pharmacol Sci* **13:** 24-28.

Kern S, Oakes TR, Stone CK, McAuliff EM, Kirschbaum C, Davidson RJ (2008) Glucose metabolic changes in the prefrontal cortex are associated with HPA axis response to a psychosocial stressor. *Psychoneuroendocrinology* **33:** 517-529.

Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS, Replication NCS (2003) The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* **289:** 3095-3105.

Khanin R, Wit E (2006) How scale-free are biological networks. *J Comput Biol* **13:** 810-818.

Khundakar A, Morris C, Oakley A, McMeekin W, Thomas AJ (2009) Morphometric analysis of neuronal and glial cell pathology in the dorsolateral prefrontal cortex in late-life depression. *Br J Psychiatry* **195:** 163-169.

Kim YK, Lee SW, Kim SH, Shim SH, Han SW, Choi SH, Lee BH (2008) Differences in cytokines between non-suicidal patients and suicidal patients in major depression. *Prog Neuropsychopharmacol Biol Psychiatry* **32:** 356-361.

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT (2008) Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* **5:** e45.

Knorr U, Vinberg M, Kessing LV, Wetterslev J (2010) Salivary cortisol in depressed patients versus control persons: a systematic review and meta-analysis. *Psychoneuroendocrinology* **35:** 1275-1286.

Koo JW, Duman RS (2008) IL-1beta is an essential mediator of the antineurogenic and anhedonic effects of stress. *Proc Natl Acad Sci U S A* **105:** 751-756.

Kostka D, Spang R (2004) Finding disease specific alterations in the co-expression of genes. *Bioinformatics* **20 Suppl 1:** i194-i199.

Kramer MA, Kolaczyk ED, Kirsch HE (2008) Emergent network topology at seizure onset in humans. *Epilepsy Res* **79:** 173-186.

Kranz GS, Kasper S, Lanzenberger R (2010) Reward and the serotonergic system. *Neuroscience* **166:** 1023-1035.

Krishnan V, Nestler EJ (2010) Linking molecules to mood: new insight into the biology of depression. *Am J Psychiatry* **167:** 1305-1320.

Kronenberg G, Kirste I, Inta D, Chourbaji S, Heuser I, Endres M, Gass P (2009) Reduced hippocampal neurogenesis in the GR(+/-) genetic mouse model of depression. *Eur Arch Psychiatry Clin Neurosci* **259:** 499-504.

Ladd CO, Huot RL, Thrivikraman KV, Nemeroff CB, Plotsky PM (2004) Long-term adaptations in glucocorticoid receptor and mineralocorticoid receptor mRNA and negative feedback on the hypothalamo-pituitary-adrenal axis following neonatal maternal separation. *Biol Psychiatry* **55:** 367-375.

Lai Y, Wu B, Chen L, Zhao H (2004) A statistical method for identifying differential gene-gene co-expression patterns. *Bioinformatics* **20:** 3146-3155.

Lee DY, Hayes JJ, Pruss D, Wolffe AP (1993) A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72:** 73-84.

Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P (2004) Coexpression analysis of human genes across many microarray data sets. *Genome Res* **14:** 1085-1094.

Leonardson AS, Zhu J, Chen Y, Wang K, Lamb JR, Reitman M, Emilsson V, Schadt EE (2010) The effect of food intake on gene expression in human peripheral blood. *Hum Mol Genet* **19:** 159-169.

Lercher MJ, Blumenthal T, Hurst LD (2003) Coexpression of neighboring genes in Caenorhabditis elegans is mostly due to operons and duplicate genes. *Genome Res* **13:** 238-243.

Levinson AJ, Fitzgerald PB, Favalli G, Blumberger DM, Daigle M, Daskalakis ZJ (2010) Evidence of cortical inhibitory deficits in major depressive disorder. *Biol Psychiatry* **67:** 458-464.

Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M, Craddock N, Owen MJ, Jones L, Jones I, Korszun A, Aitchison KJ, Shi J, Quinn JP, Mackenzie A, Vollenweider P, Waeber G, Heath S, Lathrop M, Muglia P, Barnes MR, Whittaker JC, Tozzi F, Holsboer F, Preisig M, Farmer AE, Breen G, Craig IW, McGuffin P (2010) Genome-wide association study of major recurrent depression in the U.K. population. *Am J Psychiatry* **167:** 949-957.

Lewis DA, Kathol RG, Sherman BM, Winokur G, Schlesser MA (1983) Differentiation of depressive subtypes by insulin insensitivity in the recovered phase. *Arch Gen Psychiatry* **40:** 167-170.

Li M, Belozerov VE, Cai HN (2010) Modulation of chromatin boundary activities by nucleosome-remodeling activities in Drosophila melanogaster. *Mol Cell Biol* **30:** 1067-1076.

Liao BY, Zhang J (2008) Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proc Natl Acad Sci U S A* **105:** 6987-6992.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326:** 289-293.

Lotrich FE, Rabinovitz M, Gironda P, Pollock BG (2007) Depression following pegylated interferon-alpha: characteristics and vulnerability. *J Psychosom Res* **63:** 131-135.

Lu X, Jain VV, Finn PW, Perkins DL (2007) Hubs in biological interaction networks exhibit low changes in expression in experimental asthma. *Mol Syst Biol* **3:** 98.

Luscher B, Shen Q, Sahir N (2010) The GABAergic deficit hypothesis of major depressive disorder. *Mol Psychiatry*.

Luscombe NM, Babu MM, Yu H, Snyder M, Teichmann SA, Gerstein M (2004) Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431:** 308-312.

Ma H-W, Kumar B, Ditges U, Gunzer F, Buer J, Zeng A-P (2004) An extended transcriptional regulatory network of Escherichia coli and analysis of its hierarchical structure and network motifs. *Nucleic Acids Res* **32:** 6643-6649.

Mann SA, Versmold B, Marx R, Stahlhofen S, Dietzel ID, Heumann R, Berger R (2008) Corticosteroids reverse cytokine-induced block of survival and differentiation of oligodendrocyte progenitor cells from rats. *J Neuroinflammation* **5:** 39.

Marco A, Konikoff C, Karr TL, Kumar S (2009) Relationship between gene co-expression and sharing of transcription factor binding sites in Drosophila melanogaster. *Bioinformatics* **25:** 2473-2477.

Marcus MD, Wing RR, Guare J, Blair EH, Jawad A (1992) Lifetime prevalence of major depression and its effect on treatment outcome in obese type II diabetic patients. *Diabetes Care* **15:** 253-255.

Matthews SC, Strigo IA, Simmons AN, Yang TT, Paulus MP (2008) Decreased functional coupling of the amygdala and supragenual cingulate is related to increased depression in unmedicated individuals with current major depressive disorder. *J Affect Disord* **111:** 13-20.

Mayberg HS, Brannan SK, Mahurin RK, Jerabek PA, Brickman JS, Tekell JL, Silva JA, McGinnis S, Glass TG, Martin CC, Fox PT (1997) Cingulate function in depression: a potential predictor of treatment response. *Neuroreport* **8:** 1057-1061.

Mayberg HS, Liotti M, Brannan SK, McGinnis S, Mahurin RK, Jerabek PA, Silva JA, Tekell JL, Martin CC, Lancaster JL, Fox PT (1999) Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am J Psychiatry* **156:** 675-682.

Mazure CM, Maciejewski PK, Jacobs SC, Bruce ML (2002) Stressful life events interacting with cognitive/personality styles to predict late-onset major depression. *Am J Geriatr Psychiatry* **10:** 297-304.

McDonald EM, Mann AH, Thomas HC (1987) Interferons as mediators of psychiatric morbidity. An investigation in a trial of recombinant alpha-interferon in hepatitis-B carriers. *Lancet* **2:** 1175-1178.

McEwen BS (2003) Mood disorders and allostatic load. *Biol Psychiatry* **54:** 200-207.

McEwen BS (2007) Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiol Rev* **87:** 873-904.

McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonté B, Szyf M, Turecki G, Meaney MJ (2009) Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nat Neurosci* **12:** 342-348.

McNally L, Bhagwagar Z, Hannestad J (2008) Inflammation, glutamate, and glia in depression: a literature review. *CNS Spectr* **13:** 501-510.

Mehta D, Menke A, Binder EB (2010) Gene expression studies in major depression. *Curr Psychiatry Rep* **12:** 135-144.

Melcangi RC, Cavarretta I, Magnaghi V, Ciusani E, Salmaggi A (2000) Corticosteroids protect oligodendrocytes from cytokine-induced cell death. *Neuroreport* **11:** 3969-3972.

Meshi D, Drew MR, Saxe M, Ansorge MS, David D, Santarelli L, Malapani C, Moore H, Hen R (2006) Hippocampal neurogenesis is not required for behavioral effects of environmental enrichment. *Nat Neurosci* **9:** 729-731.

Micheloyannis S, Pachou E, Stam CJ, Breakspear M, Bitsios P, Vourkas M, Erimaki S, Zervakis M (2006) Small-world networks and disturbed functional connectivity in schizophrenia. *Schizophr Res* **87:** 60-66.

Mijalski T, Harder A, Halder T, Kersten M, Horsch M, Strom TM, Liebscher HV, Lottspeich F, de Angelis MH, Beckers J (2005) Identification of coexpressed gene clusters in a comparative analysis of transcriptome and proteome in mouse tissues. *Proc Natl Acad Sci U S A* **102:** 8621-8626.

Miller G (2010) Is pharma running out of brainy ideas? *Science* **329:** 502-504.

Mineur YS, Belzung C, Crusio WE (2007) Functional implications of decreases in neurogenesis following chronic mild stress in mice. *Neuroscience* **150:** 251-259.

Minucci S, Pelicci PG (2006) Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer* **6:** 38-51.

Modell S, Lauer CJ, Schreiber W, Huber J, Krieg JC, Holsboer F (1998) Hormonal response pattern in the combined DEX-CRH test is stable over time in subjects at high familial risk for affective disorders. *Neuropsychopharmacology* **18:** 253-262.

Monteggia LM, Barrot M, Powell CM, Berton O, Galanis V, Gemelli T, Meuth S, Nagy A, Greene RW, Nestler EJ (2004) Essential role of brain-derived neurotrophic factor in adult hippocampal function. *Proc Natl Acad Sci U S A* **101:** 10827-10832.

Monteggia LM, Luikart B, Barrot M, Theobold D, Malkovska I, Nef S, Parada LF, Nestler EJ (2007) Brain-derived neurotrophic factor conditional knockouts show gender differences in depression-related behaviors. *Biol Psychiatry* **61:** 187-197.

Morales-Medina JC, Sanchez F, Flores G, Dumont Y, Quirion R (2009) Morphological reorganization after repeated corticosterone administration in the hippocampus, nucleus accumbens and amygdala in the rat. *J Chem Neuroanat* **38:** 266-272.

Mössner R, Mikova O, Koutsilieri E, Saoud M, Ehlis A-C, Müller N, Fallgatter AJ, Riederer P (2007) Consensus paper of the WFSBP Task Force on Biological Markers: biological markers in depression. *World J Biol Psychiatry* **8:** 141-174.

Müller N, Schwarz MJ (2007) The immune-mediated alteration of serotonin and glutamate: towards an integrated view of depression. *Mol Psychiatry* **12:** 988-1000.

Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, Warshaw M, Maser JD (1999) Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *Am J Psychiatry* **156:** 1000-1006.

Muglia P, Tozzi F, Galwey NW, Francks C, Upmanyu R, Kong XQ, Antoniades A, Domenici E, Perry J, Rothen S, Vandeleur CL, Mooser V, Waeber G, Vollenweider P, Preisig M, Lucae S, Müller-Myhsok B, Holsboer F, Middleton LT, Roses AD (2008) Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. *Mol Psychiatry*.

Nair A, Vaidya VA (2006) Cyclic AMP response element binding protein and brain-derived neurotrophic factor: molecules that modulate our mood? *J Biosci* **31:** 423-434.

Nakanishi M, Niidome T, Matsuda S, Akaike A, Kihara T, Sugimoto H (2007) Microglia-derived interleukin-6 and leukaemia inhibitory factor promote astrocytic differentiation of neural stem/progenitor cells. *Eur J Neurosci* **25:** 649-658.

Nayak RR, Kearns M, Spielman RS, Cheung VG (2009) Coexpression network based on natural variation in human gene expression reveals gene interactions and functions. *Genome Res*.

Nestler EJ, Barrot M, DiLeone RJ, Eisch AJ, Gold SJ, Monteggia LM (2002) Neurobiology of depression. *Neuron* **34:** 13-25.

Netoff TI, Clewley R, Arno S, Keck T, White JA (2004) Epilepsy in small-world networks. *J Neurosci* **24:** 8075-8083.

Newman MEJ (2002) Assortative mixing in networks. *Phys Rev Lett* **89:** 208701.

Newman MEJ (2003) Mixing patterns in networks. *Phys Rev E Stat Nonlin Soft Matter Phys* **67:** 026126.

Newton SS, Thome J, Wallace TL, Shirayama Y, Schlesinger L, Sakai N, Chen J, Neve R, Nestler EJ, Duman RS (2002) Inhibition of cAMP response element-binding protein or dynorphin in the nucleus accumbens produces an antidepressant-like effect. *J Neurosci* **22:** 10883-10890.

Ng YK, Wu W, Zhang L (2009) Positive correlation between gene coexpression and positional clustering in the zebrafish genome. *BMC Genomics* **10:** 42.

Ni L, Bruce C, Hart C, Leigh-Bell J, Gelperin D, Umansky L, Gerstein MB, Snyder M (2009) Dynamic and complex transcription factor binding during an inducible response in yeast. *Genes Dev* **23:** 1351-1363.

Nibuya M, Morinobu S, Duman RS (1995) Regulation of BDNF and trkB mRNA in rat brain by chronic electroconvulsive seizure and antidepressant drug treatments. *J Neurosci* **15:** 7539-7547.

Nikolaus S, Antke C, Müller H-W (2009) In vivo imaging of synaptic function in the central nervous system: II. Mental and affective disorders. *Behav Brain Res* **204:** 32-66.

Nin MS, Salles FB, Azeredo LA, Frazon APG, Gomez R, Barros HMT (2008) Antidepressant effect and changes of GABAA receptor gamma2 subunit mRNA after hippocampal administration of allopregnanolone in rats. *J Psychopharmacol* **22:** 477-485.

Nobuhara K, Okugawa G, Sugimoto T, Minami T, Tamagaki C, Takase K, Saito Y, Sawada S, Kinoshita T (2006) Frontal white matter anisotropy and symptom severity of late-life depression: a magnetic resonance diffusion tensor imaging study. *J Neurol Neurosurg Psychiatry* **77:** 120-122.

Nutt DJ (2006) The role of dopamine and norepinephrine in depression and antidepressant treatment. *J Clin Psychiatry* **67 Suppl 6:** 3-8.

Obayashi T, Kinoshita K (2009) Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. *DNA Res* **16:** 249-260.

Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci U S A* **103:** 17973-17978.

Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, Geschwind DH (2008) Functional organization of the transcriptome in human brain. *Nat Neurosci* **11:** 1271-1282.

Ongür D, Drevets WC, Price JL (1998) Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc Natl Acad Sci U S A* **95:** 13290-13295.

Oquendo MA, Barrera A, Ellis SP, Li S, Burke AK, Grunebaum M, Endicott J, Mann JJ (2004) Instability of symptoms in recurrent major depression: a prospective study. *Am J Psychiatry* **161:** 255-261.

Orchinik M, Carroll SS, Li YH, McEwen BS, Weiland NG (2001) Heterogeneity of hippocampal GABA(A) receptors: regulation by corticosterone. *J Neurosci* **21:** 330-339.

Overington JP, Al-Lazikani B, Hopkins AL (2006) How many drug targets are there? *Nat Rev Drug Discov* **5:** 993-996.

Papakostas GI, Thase ME, Fava M, Nelson JC, Shelton RC (2007) Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in treating major depressive disorder? A meta-analysis of studies of newer agents. *Biol Psychiatry* **62:** 1217-1227.

Pariante CM, Lightman SL (2008) The HPA axis in major depression: classical theories and new developments. *Trends Neurosci* **31:** 464-468.

Pariante CM, Miller AH (2001) Glucocorticoid receptors in major depression: relevance to pathophysiology and treatment. *Biol Psychiatry* **49:** 391-404.

Parsey RV, Olvet DM, Oquendo MA, yu Huang Y, Ogden RT, Mann JJ (2006) Higher 5-HT1A receptor binding potential during a major depressive episode predicts poor treatment response: preliminary data from a naturalistic study. *Neuropsychopharmacology* **31:** 1745-1749.

Paus T (2001) Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci* **2:** 417-424.

Pavlidis P, Qin J, Arango V, Mann JJ, Sibille E (2004) Using the gene ontology for microarray data mining: a comparison of methods and application to age effects in human prefrontal cortex. *Neurochem Res* **29:** 1213-1222.

Perkins AD, Langston MA (2009) Threshold selection in gene co-expression networks using spectral graph theory techniques. *BMC Bioinformatics* **10 Suppl 11:** S4.

Petty F, Schlesser MA (1981) Plasma GABA in affective illness. A preliminary investigation. *J Affect Disord* **3:** 339-343.

Pezawas L, Meyer-Lindenberg A, Drabant EM, Verchinski BA, Munoz KE, Kolachana BS, Egan MF, Mattay VS, Hariri AR, Weinberger DR (2005) 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nat Neurosci* **8:** 828-834.

Pinna G, Costa E, Guidotti A (2006) Fluoxetine and norfluoxetine stereospecifically and selectively increase brain neurosteroid content at doses that are inactive on 5-HT reuptake. *Psychopharmacology (Berl)* **186:** 362-372.

Pinto SS, Gottfried C, Mendez A, Gonçalves D, Karl J, Gonçalves CA, Wofchuk S, Rodnight R (2000) Immunocontent and secretion of S100B in astrocyte cultures from different brain regions in relation to morphology. *FEBS Lett* **486:** 203-207.

Price RB, Shungu DC, Mao X, Nestadt P, Kelly C, Collins KA, Murrough JW, Charney DS, Mathew SJ (2009) Amino acid neurotransmitters assessed by proton magnetic resonance spectroscopy: relationship to treatment resistance in major depressive disorder. *Biol Psychiatry* **65:** 792-800.

Prieto C, Risueño A, Fontanillo C, las Rivas JD (2008) Human gene coexpression landscape: confident network derived from tissue transcriptomic profiles. *PLoS ONE* **3:** e3911.

Pruessner JC, Dedovic K, Khalili-Mahani N, Engert V, Pruessner M, Buss C, Renwick R, Dagher A, Meaney MJ, Lupien S (2008) Deactivation of the limbic system during acute psychosocial stress: evidence from positron emission tomography and functional magnetic resonance imaging studies. *Biol Psychiatry* **63:** 234-240.

Pucak ML, Kaplin AI (2005) Unkind cytokines: current evidence for the potential role of cytokines in immune-mediated depression. *Int Rev Psychiatry* **17:** 477-483.

Purmann A, Toedling J, Schueler M, Carninci P, Lehrach H, Hayashizaki Y, Huber W, Sperling S (2007) Genomic organization of transcriptomes in mammals: Coregulation and cofunctionality. *Genomics* **89:** 580-587.

Qiu P, Gentles AJ, Plevritis SK (2009) Fast calculation of pairwise mutual information for gene regulatory network reconstruction. *Computer methods and programs in biomedicine* **94:** 177-180.

Raab JR, Kamakaka RT (2010) Insulators and promoters: closer than we think. *Nat Rev Genet* **11:** 439-446.

Radley JJ, Sisti HM, Hao J, Rocher AB, McCall T, Hof PR, McEwen BS, Morrison JH (2004) Chronic behavioral stress induces apical dendritic reorganization in pyramidal neurons of the medial prefrontal cortex. *Neuroscience* **125:** 1-6.

Rajkowska G, Miguel-Hidalgo JJ (2007) Gliogenesis and glial pathology in depression. *CNS Neurol Disord Drug Targets* **6:** 219-233.

Rajkowska G, Miguel-Hidalgo JJ, Wei J, Dilley G, Pittman SD, Meltzer HY, Overholser JC, Roth BL, Stockmeier CA (1999) Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol Psychiatry* **45:** 1085-1098.

Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL (2002) Hierarchical organization of modularity in metabolic networks. *Science* **297:** 1551-1555.

Ray M, Zhang W (2010) Analysis of Alzheimer's disease severity across brain regions by topological analysis of gene co-expression networks. *BMC Syst Biol* **4:** 136.

Ren X-Y, Fiers MWEJ, Stiekema WJ, Nap J-P (2005) Local coexpression domains of two to four genes in the genome of Arabidopsis. *Plant Physiol* **138:** 923-934.

Reverter A, Hudson NJ, Nagaraj SH, Pérez-Enciso M, Dalrymple BP (2010) Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data. *Bioinformatics* **26:** 896-904.

Reverter A, Ingham A, Lehnert SA, Tan S-H, Wang Y, Ratnakumar A, Dalrymple BP (2006) Simultaneous identification of differential gene expression and connectivity in inflammation, adipogenesis and cancer. *Bioinformatics* **22:** 2396-2404.

Rhodes RA, Murthy NV, Dresner MA, Selvaraj S, Stavrakakis N, Babar S, Cowen PJ, Grasby PM (2007) Human 5-HT transporter availability predicts amygdala reactivity in vivo. *J Neurosci* **27:** 9233-9237.

Roberson-Nay R, McClure EB, Monk CS, Nelson EE, Guyer AE, Fromm SJ, Charney DS, Leibenluft E, Blair J, Ernst M, Pine DS (2006) Increased amygdala activity during successful memory encoding in adolescent major depressive disorder: An FMRI study. *Biol Psychiatry* **60:** 966-973.

Robertson B, Wang L, Diaz MT, Aiello M, Gersing K, Beyer J, Mukundan S, McCarthy G, Doraiswamy PM (2007) Effect of bupropion extended release on negative emotion processing in

major depressive disorder: a pilot functional magnetic resonance imaging study. *J Clin Psychiatry* **68:** 261-267.

Rodrigues SM, Doux JEL, Sapolsky RM (2009) The Influence of Stress Hormones on Fear Circuitry. *Annu Rev Neurosci*.

Rosin C, Bates TE, Skaper SD (2004) Excitatory amino acid induced oligodendrocyte cell death in vitro: receptor-dependent and -independent mechanisms. *J Neurochem* **90:** 1173-1185.

Ross SB, Renyi AL (1969) Inhibition of the uptake of tritiated 5-hydroxytryptamine in brain tissue. *Eur J Pharmacol* **7:** 270-277.

Roth TL, Lubin FD, Funk AJ, Sweatt JD (2009) Lasting epigenetic influence of early-life adversity on the BDNF gene. *Biol Psychiatry* **65:** 760-769.

Ruhé HG, Mason NS, Schene AH (2007) Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies. *Mol Psychiatry* **12:** 331-359.

Sahay A, Hen R (2007) Adult hippocampal neurogenesis in depression. *Nat Neurosci* **10:** 1110-1115.

Sanacora G, Mason GF, Rothman DL, Behar KL, Hyder F, Petroff OA, Berman RM, Charney DS, Krystal JH (1999) Reduced cortical gamma-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy. *Arch Gen Psychiatry* **56:** 1043-1047.

Sanacora G, Mason GF, Rothman DL, Krystal JH (2002) Increased occipital cortex GABA concentrations in depressed patients after therapy with selective serotonin reuptake inhibitors. *Am J Psychiatry* **159:** 663-665.

Santamaria-Kisiel L, Rintala-Dempsey AC, Shaw GS (2006) Calcium-dependent and -independent interactions of the S100 protein family. *Biochem J* **396:** 201-214.

Santarelli L, Saxe M, Gross C, Surget A, Battaglia F, Dulawa S, Weisstaub N, Lee J, Duman R, Arancio O, Belzung C, Hen R (2003) Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science* **301:** 805-809.

Schroeter ML, Abdul-Khaliq H, Sacher J, Steiner J, Blasig IE, Mueller K (2010) Mood disorders are glial disorders: evidence from in vivo studies. *Cardiovasc Psychiatry Neurol* **2010:** 780645.

Schröter A, Lustenberger RM, Obermair FJ, Thallmair M (2009) High-dose corticosteroids after spinal cord injury reduce neural progenitor cell proliferation. *Neuroscience* **161:** 753-763.

Schultz SK, Ellingrod VL, Turvey C, Moser DJ, Arndt S (2003) The influence of cognitive impairment and behavioral dysregulation on daily functioning in the nursing home setting. *Am J Psychiatry* **160:** 582-584.

Seminowicz DA, Mayberg HS, McIntosh AR, Goldapple K, Kennedy S, Segal Z, Rafi-Tari S (2004) Limbic-frontal circuitry in major depression: a path modeling metanalysis. *Neuroimage* **22:** 409-418.

Sémon M, Duret L (2006) Evolutionary origin and maintenance of coexpressed gene clusters in mammals. *Mol Biol Evol* **23:** 1715-1723.

Sequeira A, Gwadry FG, Ffrench-Mullen JMH, Canetti L, Gingras Y, Casero RA, Rouleau G, Benkelfat C, Turecki G (2006) Implication of SSAT by gene expression and genetic variation in suicide and major depression. *Arch Gen Psychiatry* **63:** 35-48.

Shansky RM, Hamo C, Hof PR, McEwen BS, Morrison JH (2009) Stress-induced dendritic remodeling in the prefrontal cortex is circuit specific. *Cereb Cortex* **19:** 2479-2484.

Sharan R, Ideker T (2006) Modeling cellular machinery through biological network comparison. *Nat Biotech* **24:** 427-433.

Sheline YI, Barch DM, Donnelly JM, Ollinger JM, Snyder AZ, Mintun MA (2001) Increased amygdala response to masked emotional faces in depressed subjects resolves with antidepressant treatment: an fMRI study. *Biol Psychiatry* **50:** 651-658.

Sheline YI, Barch DM, Price JL, Rundle MM, Vaishnavi SN, Snyder AZ, Mintun MA, Wang S, Coalson RS, Raichle ME (2009) The default mode network and self-referential processes in depression. *Proc Natl Acad Sci U S A* **106:** 1942-1947.

Sheline YI, Gado MH, Price JL (1998) Amygdala core nuclei volumes are decreased in recurrent major depression. *Neuroreport* **9:** 2023-2028.

Shelton RC (2007) The molecular neurobiology of depression. *Psychiatr Clin North Am* **30:** 1-11.

Shen Q, Lal R, Luellen BA, Earnheart JC, Andrews AM, Luscher B (2010) gamma-Aminobutyric acid-type A receptor deficits cause hypothalamic-pituitary-adrenal axis hyperactivity and antidepressant drug sensitivity reminiscent of melancholic forms of depression. *Biol Psychiatry* **68:** 512-520.

Shmulevich I, Kauffman SA, Aldana M (2005) Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proc Natl Acad Sci U S A* **102:** 13439-13444.

Sibille E, Arango V, Galfalvy HC, Pavlidis P, Erraji-Benchekroun L, Ellis SP, Mann JJ (2004) Gene expression profiling of depression and suicide in human prefrontal cortex. *Neuropsychopharmacology* **29:** 351-361.

Sibille E, Arango V, Joeyen-Waldorf J, Wang Y, Leman S, Surget A, Belzung C, Mann JJ, Lewis DA (2008) Large-scale estimates of cellular origins of mRNAs: enhancing the yield of transcriptome analyses. *J Neurosci Methods* **167:** 198-206.

Sibille E, Su J, Leman S, Guisquet AML, Ibarguen-Vargas Y, Joeyen-Waldorf J, Glorioso C, Tseng GC, Pezzone M, Hen R, Belzung C (2007) Lack of serotonin1B receptor expression leads to age-related motor dysfunction, early onset of brain molecular aging and reduced longevity. *Mol Psychiatry* **12:** 1042-1056, 1975.

Sibille E, Wang Y, Joeyen-Waldorf J, Gaiteri C, Surget A, Oh S, Belzung C, Tseng GC, Lewis DA (2009) A molecular signature of depression in the amygdala. *Am J Psychiatry* **166:** 1011-1024.

Siegle GJ, Carter CS, Thase ME (2006) Use of FMRI to predict recovery from unipolar depression with cognitive behavior therapy. *Am J Psychiatry* **163:** 735-738.

Siegle GJ, Steinhauer SR, Thase ME, Stenger VA, Carter CS (2002) Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biol Psychiatry* **51:** 693-707.

Siegle GJ, Thompson W, Carter CS, Steinhauer SR, Thase ME (2007) Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biol Psychiatry* **61:** 198-209.

Sliwerska E, Meng F, Speed TP, Jones EG, Bunney WE, Akil H, Watson SJ, Burmeister M (2007) SNPs on chips: the hidden genetic code in expression arrays. *Biol Psychiatry* **61:** 13-16.

Smit DJA, Stam CJ, Posthuma D, Boomsma DI, de Geus EJC (2008) Heritability of "small-world" networks in the brain: a graph theoretical analysis of resting-state EEG functional connectivity. *Hum Brain Mapp* **29:** 1368-1378.

Smith DF, Jakobsen S (2009) Molecular tools for assessing human depression by positron emission tomography. *Eur Neuropsychopharmacol* **19:** 611-628.

Smith RS (1991) The macrophage theory of depression. *Med Hypotheses* **35:** 298-306.

Sokolov BP (2007) Oligodendroglial abnormalities in schizophrenia, mood disorders and substance abuse. Comorbidity, shared traits, or molecular phenocopies? *Int J Neuropsychopharmacol* **10:** 547-555.

Southworth LK, Owen AB, Kim SK (2009) Aging mice show a decreasing correlation of gene expression within genetic modules. *PLoS Genet* **5:** e1000776.

Spijker AT, van Rossum EFC (2009) Glucocorticoid receptor polymorphisms in major depression. Focus on glucocorticoid sensitivity and neurocognitive functioning. *Ann N Y Acad Sci* **1179:** 199-215.

Sporns O, Chialvo DR, Kaiser M, Hilgetag CC (2004) Organization, development and function of complex brain networks. *Trends Cogn Sci* **8:** 418-425.

Sproul D, Gilbert N, Bickmore WA (2005) The role of chromatin structure in regulating the expression of clustered genes. *Nat Rev Genet* **6:** 775-781.

Srinivas KV, Jain R, Saurav S, Sikdar SK (2007) Small-world network topology of hippocampal neuronal network is lost, in an in vitro glutamate injury model of epilepsy. *Eur J Neurosci* **25:** 3276-3286.

Stam CJ, Jones BF, Nolte G, Breakspear M, Scheltens P (2007) Small-world networks and functional connectivity in Alzheimer's disease. *Cereb Cortex* **17:** 92-99.

Stein JL, Wiedholz LM, Bassett DS, Weinberger DR, Zink CF, Mattay VS, Meyer-Lindenberg A (2007) A validated network of effective amygdala connectivity. *Neuroimage* **36:** 736-745.

Steuer R, Kurths J, Daub CO, Weise J, Selbig J (2002) The mutual information: detecting and evaluating dependencies between variables. *Bioinformatics* **18 Suppl 2:** S231-S240.

Strahl BD, Allis CD (2000) The language of covalent histone modifications. *Nature* **403:** 41-45.

Surget A, Saxe M, Leman S, Ibarguen-Vargas Y, Chalon S, Griebel G, Hen R, Belzung C (2008a) Drug-dependent requirement of hippocampal neurogenesis in a model of depression and of antidepressant reversal. *Biol Psychiatry* **64:** 293-301.

Surget A, Wang Y, Leman S, Ibarguen-Vargas Y, Edgar N, Griebel G, Belzung C, Sibille E (2008b) Brain-region specific and state-dependent transcriptome changes in a rodent model of depression and of antidepressant reversal. *Neuropsychopharmacology (pending)*.

Surget A, Wang Y, Leman S, Ibarguen-Vargas Y, Edgar N, Griebel G, Belzung C, Sibille E (2009) Corticolimbic transcriptome changes are state-dependent and region-specific in a rodent model of depression and of antidepressant reversal. *Neuropsychopharmacology* **34:** 1363-1380.

Sutcigil L, Oktenli C, Musabak U, Bozkurt A, Cansever A, Uzun O, Sanisoglu SY, Yesilova Z, Ozmenler N, Ozsahin A, Sengul A (2007) Pro- and anti-inflammatory cytokine balance in major depression: effect of sertraline therapy. *Clin Dev Immunol* **2007:** 76396.

Sutherland H, Bickmore WA (2009) Transcription factories: gene expression in unions? *Nat Rev Genet* **10:** 457-466.

Takahashi A, Shimamoto A, Boyson CO, DeBold JF, Miczek KA (2010) GABA(B) receptor modulation of serotonin neurons in the dorsal raphé nucleus and escalation of aggression in mice. *J Neurosci* **30:** 11771-11780.

Takahashi JL, Giuliani F, Power C, Imai Y, Yong VW (2003) Interleukin-1β promotes oligodendrocyte death through glutamate excitotoxicity. *Annals of Neurology* **53:** 588-595.

Takahashi Y, Washiyama K, Kobayashi T, Hayashi S (2006) Gene expression in the brain from fluoxetine-injected mouse using DNA microarray. *Ann N Y Acad Sci* **1074:** 42-51.

Tavares RG, Tasca CI, Santos CES, Alves LB, Porciúncula LO, Emanuelli T, Souza DO (2002) Quinolinic acid stimulates synaptosomal glutamate release and inhibits glutamate uptake into astrocytes. *Neurochem Int* **40:** 621-627.

Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM (1999) Systematic determination of genetic network architecture. *Nat Genet* **22:** 281-285.

Tesson BM, Breitling R, Jansen RC (2010) DiffCoEx: a simple and sensitive method to find differentially coexpressed gene modules. *BMC Bioinformatics* **11:** 497.

Tfilin M, Sudai E, Merenlender A, Gispan I, Yadid G, Turgeman G (2010) Mesenchymal stem cells increase hippocampal neurogenesis and counteract depressive-like behavior. *Mol Psychiatry* **15:** 1164-1175.

Thomas AJ, O'Brien JT, Barber R, McMeekin W, Perry R (2003) A neuropathological study of periventricular white matter hyperintensities in major depression. *J Affect Disord* **76:** 49-54.

Thompson Ray M, Weickert CS, Wyatt E, Webster MJ (2010) Decreased BDNF, trkB-TK+ and GAD(67) mRNA expression in the hippocampus of individuals with schizophrenia and mood disorders. *J Psychiatry Neurosci* **36:** 100048.

Tomita H, Vawter MP, Walsh DM, Evans SJ, Choudary PV, Li J, Overman KM, Atz ME, Myers RM, Jones EG, Watson SJ, Akil H, Bunney WE, Jr. (2004) Effect of agonal and postmortem factors on gene expression profile: quality control in microarray analyses of postmortem human brain. *Biol Psychiatry* **55:** 346-352.

Tompa M, Li N, Bailey TL, Church GM, Moor BD, Eskin E, Favorov AV, Frith MC, Fu Y, Kent WJ, Makeev VJ, Mironov AA, Noble WS, Pavesi G, Pesole G, Régnier M, Simonis N, Sinha S, Thijs G, van Helden J, Vandenbogaert M, Weng Z, Workman C, Ye C, Zhu Z (2005) Assessing computational tools for the discovery of transcription factor binding sites. *Nat Biotechnol* **23:** 137-144.

Torkamani A, Dean B, Schork NJ, Thomas EA (2010) Coexpression network analysis of neural tissue reveals perturbations in developmental processes in schizophrenia. *Genome Res* **20:** 403-412.

Torrey EF, Webster M, Knable M, Johnston N, Yolken RH (2000) The stanley foundation brain collection and neuropathology consortium. *Schizophr Res* **44:** 151-155.

Trinklein ND, Aldred SF, Hartman SJ, Schroeder DI, Otillar RP, Myers RM (2004) An abundance of bidirectional promoters in the human genome. *Genome Res* **14:** 62-66.

Trivedi MH, Lin EHB, Katon WJ (2007) Consensus recommendations for improving adherence, self-management, and outcomes in patients with depression. *CNS Spectr* **12:** 1-27.

Tsankova NM, Berton O, Renthal W, Kumar A, Neve RL, Nestler EJ (2006) Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nat Neurosci* **9:** 519-525.

Tsaparas P, Mariño-Ramírez L, Bodenreider O, Koonin EV, Jordan IK (2006) Global similarity and local divergence in human and mouse gene co-expression networks. *BMC Evol Biol* **6:** 70.

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* **358:** 252-260.

Tyring S, Gottlieb A, Papp K, Gordon K, Leonardi C, Wang A, Lalla D, Woolley M, Jahreis A, Zitnik R, Cella D, Krishnan R (2006) Etanercept and clinical outcomes, fatigue, and depression in psoriasis: double-blind placebo-controlled randomised phase III trial. *Lancet* **367:** 29-35.

Uher R, Perroud N, Ng MYM, Hauser J, Henigsberg N, Maier W, Mors O, Placentino A, Rietschel M, Souery D, Zagar T, Czerski PM, Jerman B, Larsen ER, Schulze TG, Zobel A, Cohen-Woods S, Pirlo K, Butler AW, Muglia P, Barnes MR, Lathrop M, Farmer A, Breen G, Aitchison KJ, Craig I, Lewis CM, McGuffin P (2010) Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. *Am J Psychiatry* **167:** 555-564.

Uranova NA, Vostrikov VM, Orlovskaya DD, Rachmanova VI (2004) Oligodendroglial density in the prefrontal cortex in schizophrenia and mood disorders: a study from the Stanley Neuropathology Consortium. *Schizophr Res* **67:** 269-275.

Uzunova V, Sampson L, Uzunov DP (2006) Relevance of endogenous 3alpha-reduced neurosteroids to depression and antidepressant action. *Psychopharmacology (Berl)* **186:** 351-361.

Vaidya VA, Duman RS (2001) Depresssion--emerging insights from neurobiology. *Br Med Bull* **57:** 61-79.

Valenzuela L, Kamakaka RT (2006) Chromatin insulators. *Annu Rev Genet* **40:** 107-138.

van Nas A, Guhathakurta D, Wang SS, Yehya N, Horvath S, Zhang B, Ingram-Drake L, Chaudhuri G, Schadt EE, Drake TA, Arnold AP, Lusis AJ (2009) Elucidating the role of gonadal hormones in sexually dimorphic gene coexpression networks. *Endocrinology* **150:** 1235-1249.

van Noort V, Snel B, Huynen MA (2004) The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model. *EMBO Rep* **5:** 280-284.

van Rossum EFC, Binder EB, Majer M, Koper JW, Ising M, Modell S, Salyakina D, Lamberts SWJ, Holsboer F (2006) Polymorphisms of the glucocorticoid receptor gene and major depression. *Biol Psychiatry* **59:** 681-688.

Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM (2009) A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* **10:** 252-263.

Venkataram S, Fay JC (2010) Is transcription factor binding site turnover a sufficient explanation for cis-regulatory sequence divergence? *Genome Biol Evol* **2:** 851-858.

Vercellino M, Merola A, Piacentino C, Votta B, Capello E, Mancardi GL, Mutani R, Giordana MT, Cavalla P (2007) Altered glutamate reuptake in relapsing-remitting and secondary progressive multiple sclerosis cortex: correlation with microglia infiltration, demyelination, and neuronal and synaptic damage. *J Neuropathol Exp Neurol* **66:** 732-739.

Vogel JH, von Heydebreck A, Purmann A, Sperling S (2005) Chromosomal clustering of a human transcriptome reveals regulatory background. *BMC Bioinformatics* **6:** 230.

Voineskos AN, Wilson AA, Boovariwala A, Sagrati S, Houle S, Rusjan P, Sokolov S, Spencer EP, Ginovart N, Meyer JH (2007) Serotonin transporter occupancy of high-dose selective serotonin reuptake inhibitors during major depressive disorder measured with [11C]DASB positron emission tomography. *Psychopharmacology (Berl)* **193:** 539-545.

Volinia S, Galasso M, Costinean S, Tagliavini L, Gamberoni G, Drusco A, Marchesini J, Mascellani N, Sana ME, Jarour RA, Desponts C, Teitell M, Baffa R, Aqeilan R, Iorio MV, Taccioli C, Garzon R, Leva GD, Fabbri M, Catozzi M, Previati M, Ambs S, Palumbo T, Garofalo M, Veronese A, Bottoni A, Gasparini P, Harris CC, Visone R, Pekarsky Y, de la Chapelle A, Bloomston M, Dillhoff M, Rassenti LZ, Kipps TJ, Huebner K, Pichiorri F, Lenze D, Cairo S, Buendia M-A, Pineau P, Dejean A, Zanesi N, Rossi S, Calin GA, Liu C-G, Palatini J, Negrini M, Vecchione A, Rosenberg A, Croce CM (2010) Reprogramming of miRNA networks in cancer and leukemia. *Genome Res* **20:** 589-599.

Vyas A, Mitra R, Rao BSS, Chattarji S (2002) Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. *J Neurosci* **22:** 6810-6818.

Wachi S, Yoneda K, Wu R (2005) Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics* **21:** 4205-4208.

Walf AA, Frye CA (2006) A review and update of mechanisms of estrogen in the hippocampus and amygdala for anxiety and depression behavior. *Neuropsychopharmacology* **31:** 1097-1111.

Wang L, Tang H, Thayanithy V, Subramanian S, Oberg AL, Cunningham JM, Cerhan JR, Steer CJ, Thibodeau SN (2009) Gene networks and microRNAs implicated in aggressive prostate cancer. *Cancer Res* **69:** 9490-9497.

Warden D, Rush AJ, Trivedi MH, Fava M, Wisniewski SR (2007) The STAR*D Project results: a comprehensive review of findings. *Curr Psychiatry Rep* **9:** 449-459.

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* **393:** 440-442.

Weaver ICG (2007) Epigenetic programming by maternal behavior and pharmacological intervention. Nature versus nurture: let's call the whole thing off. *Epigenetics* **2:** 22-28.

Weaver ICG, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ (2004) Epigenetic programming by maternal behavior. *Nat Neurosci* **7:** 847-854.

Weaver ICG, D'Alessio AC, Brown SE, Hellstrom IC, Dymov S, Sharma S, Szyf M, Meaney MJ (2007) The transcription factor nerve growth factor-inducible protein a mediates epigenetic programming: altering epigenetic marks by immediate-early genes. *J Neurosci* **27:** 1756-1768.

Weber B, Schweiger U, Deuschle M, Heuser I (2000) Major depression and impaired glucose tolerance. *Exp Clin Endocrinol Diabetes* **108:** 187-190.

Weissman MM, Bland R, Joyce PR, Newman S, Wells JE, Wittchen HU (1993) Sex differences in rates of depression: cross-national perspectives. *J Affect Disord* **29:** 77-84.

Wellman CL (2001) Dendritic reorganization in pyramidal neurons in medial prefrontal cortex after chronic corticosterone administration. *J Neurobiol* **49:** 245-253.

Wennström M, Hellsten J, Ekstrand J, Lindgren H, Tingström A (2006) Corticosterone-induced inhibition of gliogenesis in rat hippocampus is counteracted by electroconvulsive seizures. *Biol Psychiatry* **59:** 178-186.

Whitney NP, Eidem TM, Peng H, Huang Y, Zheng JC (2009) Inflammation mediates varying effects in neurogenesis: relevance to the pathogenesis of brain injury and neurodegenerative disorders. *J Neurochem* **108:** 1343-1359.

Wilber AA, Wellman CL (2009) Neonatal maternal separation alters the development of glucocorticoid receptor expression in the interpositus nucleus of the cerebellum. *Int J Dev Neurosci* **27:** 649-654.

Willner P (2005) Chronic mild stress (CMS) revisited: consistency and behavioural-neurobiological concordance in the effects of CMS. *Neuropsychobiology* **52:** 90-110.

Willner P, Mitchell PJ (2002) The validity of animal models of predisposition to depression. *Behav Pharmacol* **13:** 169-188.

Winden KD, Oldham MC, Mirnics K, Ebert PJ, Swan CH, Levitt P, Rubenstein JL, Horvath S, Geschwind DH (2009) The organization of the transcriptional network in specific neuronal classes. *Mol Syst Biol* **5:** 291.

Winter C, von Rumohr A, Mundt A, Petrus D, Klein J, Lee T, Morgenstern R, Kupsch A, Juckel G (2007) Lesions of dopaminergic neurons in the substantia nigra pars compacta and in the ventral tegmental area enhance depressive-like behavior in rats. *Behav Brain Res* **184:** 133-141.

Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, Bakkoury ME, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Véronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW (1999) Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science* **285:** 901-906.

Wong K-F, Huk AC (2008) Temporal Dynamics Underlying Perceptual Decision Making: Insights from the Interplay between an Attractor Model and Parietal Neurophysiology. *Front Neurosci* **2:** 245-254.

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457:** 1033-1037.

Xulvi-Brunet R, Li H (2009) Co-expression Networks: Graph Properties and Topological Comparisons. *Bioinformatics*.

Yanai I, Hunter CP (2009) Comparison of diverse developmental transcriptomes reveals that coexpression of gene neighbors is not evolutionarily conserved. *Genome Res* **19:** 2214-2220.

Yanashima R, Kitagawa N, Matsubara Y, Weatheritt R, Oka K, Kikuchi S, Tomita M, Ishizaki S (2009) Network Features and Pathway Analyses of a Signal Transduction Cascade. *Front Neuroinformatics* **3:** article 13.

Yang MQ, Koehly LM, Elnitski LL (2007) Comprehensive annotation of bidirectional promoters identifies co-regulation among breast and ovarian cancer genes. *PLoS Comput Biol* **3:** e72.

Yu H, Luscombe NM, Qian J, Gerstein M (2003) Genomic analysis of gene expression relationships in transcriptional regulatory networks. *Trends Genet* **19:** 422-427.

Zampieri M, Soranzo N, Altafini C (2008) Discerning static and causal interactions in genome-wide reverse engineering problems. *Bioinformatics* **24:** 1510-1515.

Zhan S, Horrocks J, Lukens LN (2006) Islands of co-expressed neighbouring genes in Arabidopsis thaliana suggest higher-order chromosome domains. *Plant J* **45:** 347-357.

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* **4:** Article17.

Zhu C-B, Blakely RD, Hewlett WA (2006) The proinflammatory cytokines interleukin-1beta and tumor necrosis factor-alpha activate serotonin transporters. *Neuropsychopharmacology* **31:** 2121-2131.

Zhu X, Gerstein M, Snyder M (2007) Getting connected: analysis and principles of biological networks. *Genes Dev* **21:** 1010-1024.

Zobel AW, Nickel T, Sonntag A, Uhr M, Holsboer F, Ising M (2001) Cortisol response in the combined dexamethasone/CRH test as predictor of relapse in patients with remitted depression. a prospective study. *J Psychiatr Res* **35:** 83-94.

Zotenko E, Mestre J, O'Leary DP, Przytycka TM (2008) Why do hubs in the yeast protein interaction network tend to be essential: reexamining the connection between the network topology and essentiality. *PLoS Comput Biol* **4:** e1000140.

Zunszain PA, Anacker C, Cattaneo A, Carvalho LA, Pariante CM (2010) Glucocorticoids, cytokines and brain abnormalities in depression. *Prog Neuropsychopharmacol Biol Psychiatry*.