Phenome-transcriptome correlation unravels anxiety and depression related pathways

Philipp Gormanns a,1, Nikola S. Mueller b,1, Claudia Ditzen a, Simone Wolf a, Florian Holsboer a, Christoph W. Turck a,∗

a Max Planck Institute of Psychiatry, Proteomics and Biomarkers, 80804 Munich, Germany
b Max Planck Institute of Biochemistry, Independent Junior Research Group Cellular Dynamics and Cell Patterning, 82152 Martinsried, Germany

1 Both authors contributed equally to this work.

∗ Corresponding author. Max Planck Institute of Psychiatry, Proteomics and Biomarkers, Kraepelinstr. 2, D-80804 Munich, Germany. Tel.: +49 89 30622177; fax: +49 89 30622010.
E-mail address: turck@mpipsykl.mpg.de (C.W. Turck).

Keywords:
Pathway analysis
Systematic correlation
Transcriptome
Anxiety
Depression
Molecular biology of disease

ARTICLE INFO

Article history:
Received 8 September 2010
Received in revised form 1 December 2010
Accepted 21 December 2010

ABSTRACT

The identification of pathways pertinent to human diseases is critical for gaining a better understanding of their pathophysiology. Pathway knowledge in turn can provide disease marker information required for diagnosis, drug development and improved patient treatment. Psychiatric disorders including anxiety and depression are complex diseases and are caused by a combination of multiple genetic and environmental factors affecting certain brain circuits. Here we used a systems biology approach to identify molecular pathways that affect anxiety- and depression-like phenotypes. For this purpose we screened pathways for stable enrichment in a great number of publicly available transcriptome data from the Gene Expression Omnibus related to anxiety- and depression-like phenotypes. In case of anxiety our analysis implicates a dysregulation of carbohydrate metabolism, tight junction and the phosphatidylinositol signaling system, whereas for depression gap junction, gonadotropin-releasing hormone signaling and ubiquitin-mediated proteolysis pathways are affected. Furthermore, both anxiety and depression show a dysregulation of VEGF signaling, long term potentiation and the glycolysis pathway. Molecular entities that are part of the identified pathways can serve as biomarkers and potential therapeutic targets for diagnosis and treatment of depression and anxiety disorders.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

At present the diagnosis of depression and anxiety disorders is solely based on clinical classification schemes such as the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2000). These brain disorders represent complex phenotypes and thus far no well-defined biochemical disease markers are at hand. In order to improve this situation an advancement of the understanding of their complex pathophysiology is mandatory. In this regard great efforts are being made in the identification of biomarkers for improved patient stratification and monitoring therapy response. The common mechanisms involved in the pathophysiology of psychiatric disorders, however, remain elusive. To capture the complex pathophysiology of anxiety and depression, we systematically evaluated publicly available molecular data for these phenotypes.

A variety of —omics technologies have been applied to the analysis of psychiatric disorders in order to gain better insights into the characteristics of their pathophysiology. Unlike the genome the transcriptome is highly dynamic in nature and thus has the potential to reveal molecular biomarker information. A great number of high-throughput transcriptome data are centrally collected and freely available in public databases, such as the Gene Expression Omnibus (GEO), which contains data entries from over 380,000 samples representing more than 14,000 experiments (Barrett et al., 2007). Each experiment embodies molecular characterizations of defined phenotypic aspects of a disease. Consequently, a combination of multiple biological snapshots for each phenotype allows the systematic elucidation of the involved pathways. For a comprehensive interpretation of high-throughput —omics data the development of system approaches is mandatory. Butte and Kohane (2006) were the first to correlate structured phenotype information with transcriptome data and later showed that disease signatures are robust across multiple experiments (Dudley et al., 2009). In subsequent studies large-scale correlation approaches have integrated phenotypes with genome and interactome data (Ghosh and Poisson, 2009; Hajduk and Greer, 2007) and have been successfully applied to advance the understanding of a tumor suppressor mechanism.
The present study is based on the concept that transcriptional expression level alterations can reveal pathways dysregulated in a phenotype and thus can provide critical information on the pathophysiology involved.

To capture the complex pathophysiology of anxiety and depression we conducted a comprehensive pathway analysis taking into account all relevant transcriptome data currently deposited in the public domain. Although we aimed to identify dysregulated pathways of human disease phenotypes, we also included transcript data from model organisms that represent certain aspects of the phenotypes, e.g. mouse models of trait anxiety or depression-like behavior. For this purpose we have developed an approach that systematically elucidates pathways globally dysregulated in anxiety and depression based on the correlation of structured phenotypes and GEO transcriptome data. Because of the limited number of public datasets available for the human phenotypes of interest and to capture as many datasets as possible related to anxiety and depression phenotypes, we consulted the structured Unified Medical Language System (UMLS) (Bodenreider, 2004) to identify disease-related GEO datasets (GDS). In addition, the unification of transcriptionally dysregulated orthologs allowed us to identify pathways that are significantly enriched across several experiments and thus are strongly related to the investigated phenotypes. The results of our analysis reveal novel insights into the pathophysiology of anxiety and depression that is not apparent from individual experimental data.

2. Materials and methods

2.1. Phenome definition

The phenomes of anxiety and depression was defined as a generalized level of disease classifications. Keywords describing phenotypic states of anxiety and depression were selected and grouped into generalized disease phenotypes of anxiety and depression (Supplementary Method), respectively. Furthermore, the UMLS was utilized as systematic phenome description source (Bodenreider, 2004). Based on this we introduced another layer that we termed ‘generalized phenotype’ for depression and anxiety, respectively. It includes all shades for each disease which are not necessarily observed all at once in the clinic. The definition of this artificial phenotype is critical for a successful statistical pathway enrichment analysis since the phenotype should include as many datasets as possible. Furthermore, this definition allowed us to include data sets for Bipolar Disorder, a phenotype that consists of depressive as well as manic episodes. In this regard it is believed that depressive and manic episodes affect the same molecular pathways. We submit that by considering the combined up- and downregulated transcript information the generalization approach has the ability to cover depression related pathways that are part of the bipolar disorder phenotype.

2.2. Phenome to transcriptome correlation

The transcriptome was built from experimental microarray data. The GEO is a public repository providing high-throughput raw experimental data of a great number of microarray-based transcriptome experiments (Barrett et al., 2007). Each dataset provides unstructured free-text for the utilized sample, experiment and platform. For all GEO samples, series, platforms and datasets the annotations were extracted. Publicly available scripts (Butte and Kohane, 2006) were utilized and extended to extract the annotations from GEO data repository.

Phenome and transcriptome data were systematically correlated by two approaches to guarantee a comprehensive analysis method (Fig. 1A). In the first step the anxiety and depression concepts were searched in the transcriptome annotations. The second step detected weak correlations of the phenome and the transcriptome by using the MMTx library mapping of the free-text transcriptome annotations to the UMLS concepts of anxiety and depression (Butte and Kohane, 2006). Since the automatic text mining of MMTx was not always correct (Lage et al., 2007), a manual validation of the correlation results was necessary. Resulting GEO entries were linked back to a list of GEO datasets (GDS) containing the microarray raw data. Each expression data set had already been processed and normalized in various (partly unknown) ways. Thus, the median

![Fig. 1. Disease-related pathways. A. Workflow of phenome-to-transcriptome correlation resulting in GDS related to anxiety and depression. GEO annotations were first correlated with the generalized disease concepts of Anxiety and Depression and then correlated to structured disease concepts of UMLS to detect all related GDS. B. The workflow shows the identification process for disease-related pathways. For pathway analysis that included data from different organisms the significantly regulated genes were mapped through KEGG orthologs. C. Pathways that were significantly enriched at least once on the GDS level and on the unified level resulted in global level pathways.](image-url)
absolute deviation (MAD) (Kim et al., 2007; Smyth and Speed, 2003) was applied to each GDS, to ensure that all datasets were equally processed in our analysis. Every sample from each GDS was manually assigned to a disease or control group (Supplementary Methods). Those GDS where disease and control group could not be distinguished were not considered for further analysis.

2.3. Disease-related pathways

To combine the pathways with genes differentially expressed between disease phenotypes and control, we performed differential expression analysis by ANOVA for every gene of each GDS which was present in at least one pathway (Fig. 1B). A differentially expressed gene ($p < 0.05$) was further parameterized by fold change to determine the regulation direction (Tusher et al., 2001). In addition, pathway dysregulations were analyzed on a “global level” where all GDS datasets were merged (“unified dataset”) into a unified version of each pathway (Fig. 1C). For this purpose the gene lists were first mapped to protein names using Pathway Studio (Ariadne Genomics, Rockville, MD), resulting in one protein list for each GDS. Since the proteins were still organism-specific, they were converted to orthologs with the help of the “Kyoto Encyclopedia of Genes and Genomes” (KEGG), Version Sep. 3, 2009 (Kanehisa et al., 2008). To determine disease-related pathways, we utilized KEGG-standardized reference pathways. We only considered KEGG pathways covered by orthologs derived from human, mouse and rat (Table S4). In addition, we performed an unspecific filtering. For example, we removed pathways like “Metabolic pathways”, which subsume other pathways in the list. Moreover, we discarded pathways like “Progesterone-mediated oocyte maturation” since we did not distinguish between genders in our analysis.

2.4. Pathway enrichment

To identify pathways dysregulated on the GDS level as well as the anxiety and depression unified dataset, pathway enrichment analyses were conducted. A pathway was considered enriched when the number of regulated orthologs was statistically significant ($p < 0.01$) compared to the number of all possible ortholog annotations for this pathway. The number of possible ortholog annotations for each pathway was determined with respect to the data origins; for each GDS either human, mouse or rat and for the two unified sets all three organisms were considered. The hypergeometric distribution was calculated with R (Version 2.9.0) (R Development Core Team, 2009) resulting in one $p$-value for each pathway. In addition, we calculated the FDR (Strimmer, 2008) for the unified datasets. These values confirmed the $p$-value threshold of 0.01 as appropriate. The number of possible ortholog annotations were 3358, 3319, 3238, and 3382 for human, mouse, rat, and all three species combined, respectively. The total number of differentially expressed orthologs mapping to anxiety and depression were 1147 and 1606, respectively. The sample size was set to the number of orthologs, which had at least one pathway entry in KEGG.

3. Results

To find out which pathways were commonly regulated in anxiety- and depression-like phenotypes, KEGG pathways were scanned for transcriptional regulation in the GEO repository (Fig. 1). The GEO contained six and five GDS matching our generalized anxiety- and depression-like phenotypes (Table 1), respectively. The disease-related transcriptomes were used to identify significantly regulated pathways for both phenotypes. Globally regulated pathways, enriched in at least one GDS as well as the unified dataset, are of greatest significance for two reasons. First, the pathway perturbation was strong enough to be detected in a single experiment. Second, the same pathway also showed strong perturbations on the unified level, thus almost the entire pathway as a functional unit is affected. Note that pathways enriched in the unified dataset are derived from a large amount of data compared to the sets originating from a single GDS. As a consequence, a pathway is only significantly enriched if its ortholog coverage is high. In our study we excluded pathways, which were only enriched in the unified dataset or only enriched in a single GDS (Fig. 2 for overview). These pathways may represent biological states not commonly and consistently dysregulated with regard to the generalized disease concepts of anxiety and depression.

3.1. Pathways common in anxiety and depression

Pathways enriched in both anxiety- and depression-like phenotypes are of special interest since it is assumed that anxiety and depression are of high comorbidity and may share pathobiological aspects (Hettema, 2008; Pollack, 2005). We identified three interesting pathways out of five significantly enriched in both phenotypes. One pathway was synaptic Long-Term Potentiation (LTP) that results in a long-lasting increase in synaptic efficacy. LTP was found to be a global level pathway for both the anxiety phenotype ($p = 0.00509$ for unified dataset; pathway coverage: 54.7%) and the depression phenotype ($p = 7.71E-06$ for unified dataset; pathway coverage:

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Gene Expression Omnibus' Depression and Anxiety related datasets.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDS</td>
<td>Organism/Tissue</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Anxiety phenotype</td>
<td></td>
</tr>
<tr>
<td>1GDS346</td>
<td>Mus musculus/Somatosensory cortex</td>
</tr>
<tr>
<td>2GDS1406</td>
<td>Mus musculus/Multiple brain regions</td>
</tr>
<tr>
<td>2GDS1921</td>
<td>Mus musculus/Amynagala and Hippocampus</td>
</tr>
<tr>
<td>2GDS2226</td>
<td>Mus musculus/Hippocampus</td>
</tr>
<tr>
<td>2GDS2309</td>
<td>Mus musculus/Brain</td>
</tr>
<tr>
<td>2GDS526</td>
<td>Rattus norvegicus/Neurophils</td>
</tr>
<tr>
<td>Depression phenotype</td>
<td></td>
</tr>
<tr>
<td>1GDS2190</td>
<td>Homo sapiens/dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>1GDS2191</td>
<td>Homo sapiens/orbitofrontal cortex</td>
</tr>
<tr>
<td>1GDS2779</td>
<td>Homo sapiens/Lymphohoblastoid cell lines</td>
</tr>
<tr>
<td>GDS346</td>
<td>Mus musculus/Somatosensory cortex</td>
</tr>
<tr>
<td>GDS3111</td>
<td>Mus musculus/Hippocampus</td>
</tr>
</tbody>
</table>

1 Number of orthologs with a pathway entry.
2 Identified by direct match to GEO annotations.
80.9%). Closer inspection revealed a tendency of downregulation in the case of depression (Fig. 3). The ratio of up-to-downregulated orthologs was 8:10 and 8:17 for anxiety and depression, respectively.

The orthologs which showed an opposite regulation with respect to both phenotypes were considered characteristic for one or the other generalized phenotype (Fig. S1). Another pathway detected for both phenotypes was glycolysis. Especially for anxiety ($p = 1.25E-05$ for unified dataset; pathway coverage: 75%) a very strong statistical enrichment could be detected and a similar coverage was observed for depression ($p = 0.0028$ for unified dataset; pathway coverage: 75%). Regulation tendency showed a difference between the two phenotypes. The anxiety-like phenotype showed a strong tendency to be upregulated while the depression-like phenotype had a slight tendency for downregulation. The ratio of up-to-downregulation orthologs was 14:5 and 7:10 for anxiety and depression, respectively.

Furthermore, vascular endothelial growth factor (VEGF) signaling was revealed to be enriched in anxiety ($p = 0.00054$; pathway coverage: 63.6%) and depression ($p = 0.00086$; pathway coverage: 75.7%). The up-to-down regulation ortholog ratio (anxiety: 6:8; depression: 7:9) for both phenotypes showed no phenotype specific correlation but rather a general tendency towards downregulation in both, anxiety- and depression-like phenotypes. The VEGF pathway is essential for gap junction permeability control in endothelial cells (Suarez and Ballmer-Hofer, 2001) as well as for tight junction assembly by monolayer permeability control (Wang et al., 2001).

### 3.2. Anxiety-related pathways

The analysis of all anxiety-related GDS yielded 10 pathways enriched on the global level (Table S1). Closer inspection of anxiety-specific pathways revealed a connection to carbohydrate metabolism. Other energy metabolism-related pathways including glycolysis and the tricarboxylic acid (TCA) cycle ($p = 1.85E-04$) also showed a significant dysregulation based on 68.1% of its orthologs. Both carbohydrate metabolism-related pathways qualified as global level pathways. The carbohydrate metabolism module ($p = 9.215E-08$) represented the most prominent pathway system affected in the generalized anxiety phenotype. Fig. 4A depicts the different orthologs according to their regulation for glycolysis and the TCA cycle. A predominant upregulation of genes was observed throughout the heterogeneous GDS (Fig. 4B) including the gene for the rate-limiting enzyme phosphofructokinase.
Another high ortholog dysregulation was observed in the tight junction pathway ($p = 0.0004$; pathway coverage: 55.5%). Tight junctions are key players in substance transport between cells and through tissues and are a part of the blood-brain-barrier (BBB), an important target of psychiatric research as it represents the main barrier for drug entry into the brain. Finally, phosphatidylinositol signaling was identified as a global level pathway ($p = 0.0010$ for unified dataset). No predominant regulation direction was observed in the 63% detected dysregulated orthologs.

3.3. Depression-related pathways

We next analyzed the pathways exclusively enriched in the generalized depression phenotype. In total, 10 pathways were detected to be significantly enriched (Table S2). Extensive analyses revealed three interesting global level pathways with two of them showing a prevalent downregulation in terms of gene expression (Fig. 3C).

The gap junction pathway was found to be significantly enriched in depression ($p = 0.0057$ for unified dataset). This pathway represents several protein kinase cascades specific for adjacent cell communication. Out of 64.9% detected dysregulated orthologs, 45.9% were consistently downregulated in this pathway. No predominant regulation direction was observed in the 63% detected dysregulated orthologs.

Compared to the other depression related pathways the GnRH
signaling pathway shows a slight tendency for upregulation by 38% of its detected orthologs.

4. Discussion

Our study provides a systematic description of dysregulated pathways related to generalized anxiety- and depression-related phenotypes in humans and mouse and rat models that is based on gene transcription level differences. Individual transcriptomic experiments often only capture a fraction of the affected pathways pertinent to a disease phenotype. Accordingly, the unification of multiple datasets enabled us to determine pathway dysregulations in a robust and statistically significant manner. Through the combination of transcriptomic data from a variety of tissues and species we generalized our results to unravel pathways which were either exclusively or commonly dysregulated in anxiety- and depression-like phenotypes.

An altered carbohydrate metabolism revealed by an upregulation of glycolysis and the TCA cycle was shown to be significantly related to the anxiety-like behavioral phenotype. Several key metabolites of the carbohydrate metabolism have previously been implicated in anxiety. Elevated brain lactate levels due to an increased enzymatic activity of phosphofructokinase (PFK) (Esquivel et al., 2009) were found in panic disorder (Esquivel et al., 2009; Maddock et al., 2009). We also found PFK transcription to be upregulated in our study which supports the earlier observations. As the enzyme carries out the rate-limiting step in glycolysis, the resulting imbalance of the metabolites may contribute to the pathophysiology of anxiety. Since panic attacks are inducible by lactate infusions and patients with panic disorder seem to have an increased sensitivity towards higher lactate levels (Maddock et al., 2009), we submit that a dysregulation of the carbohydrate metabolism contributes to a lactate level imbalance and, thus, might mediate anxiety-like phenotypes.

We also found the tight junction pathway dysregulated in anxiety. Tight junctions are critical for the BBB, an important target in neuroscience research due to its ability to block the transfer of specific molecules into the CNS and thus limiting the access of neuropsychiatric drugs. In this regard a SNP in the gene for P-glycoprotein, a BBB transporter protein, was found to influence antidepressant drug uptake (Uhr et al., 2008). Our findings suggest that not only P-glycoprotein but multiple entities that are part of the BBB tight junction pathway are affected in anxiety-like phenotypes.

Also implicated for anxiety-like phenotypes is phosphatidylinositol signaling. Several studies have shown that inositol has anxiolytic activities. Proteins that are part of the phosphatidylinositol signaling pathway could thus represent potential targets for anxiety drug research.

Three pathways were found to be affected only for depression-like phenotypes, with the majority of transcripts downregulated (Fig. 3C). The gap junction pathway affects the exchange of cytoplasmic components between adjacent cells. Although this pathway has not previously been related to depression, its key cell membrane connecting units, the connexons, are frequently discussed as drug targets for psychiatric disorders (Fatemi et al., 2008). The dysregulation of the gap junctions might be the result of an altered vascular endothelial growth factor (VEGF) signaling pathway, which is upstream and reported to be essential for permeability control of gap junctions. In response to VEGF stimulation several kinases are activated leading to disrupted gap junctions in endothelial cells (Suarez and Ballmer-Hofer, 2001). Hence the finding of a downregulated gap junction pathway for depression might ultimately be caused by an altered VEGF signaling pathway.

We also found the ubiquitin-mediated proteolysis to be enriched in depression with a strong downregulation tendency. The proteasome is responsible for protein degradation and is involved in many cellular processes. Ubiquitin-mediated proteolysis is activated in response to cellular stress including oxidative stress. A downregulation of proteolysis has been reported not only for bipolar depression (Konradi et al., 2004; Ryan et al., 2006), but also for schizophrenia (Bousman et al., 2009). With our correlation approach we are able to confirm and statistically back up the hypothesis of ubiquitin-mediated proteolysis involvement in depression.

The GnRH signaling pathway showed a significant enrichment and a strong upregulation tendency in depression. The hypothalamic hormone GnRH stimulates the secretion of luteinizing hormone and follicle-stimulating hormone from the pituitary gland. The two hormones in turn regulate ovary and testicular function in mammals and changes in gonadotropin levels are known to impact mood and can contribute to the development of affective disorders. On the other hand the therapeutic use and substitution of sex hormones can improve depressive symptoms in humans (Fleurence et al., 2009).

Pathways enriched for both phenotypes are of special interest since it is assumed that anxiety and depression have a shared pathophysiology (Hettema, 2008; Pollack, 2005). We found LTP dysregulated for both generalized anxiety and depression. This finding is not at all surprising since synaptic plasticity, a key process involved in learning and memory, is widely believed to be affected in anxiety and depression (Cooke and Bliss, 2006).

In the present study we have identified several pathways that showed a dysregulation in both anxiety and depression but with distinct characteristics. Interestingly, the gap junction pathway in depression and the tight junction pathway in anxiety are both affected by VEGF signaling (Suarez and Ballmer-Hofer, 2001; Wang et al., 2001), a pathway significantly enriched for both phenotypes. The other aspect of interest in the case of gap junctions is their presumed function in the tight junction barrier (Nagasawa et al., 2006). A malfunction of the BBB might be caused by a poor permeability control and barrier function of the tight junctions and also by the gap junction pathway components with the net result of altered brain (signaling) metabolite levels. In addition, an abnormal blood–brain communication can lead to a dysregulation of brain synaptic plasticity (Shalev et al., 2009) which our data implicate for both phenotypes.

The frequently discussed oxidative stress pathway in anxiety and depression (Konradi et al., 2004; Ng et al., 2008; Rezин et al., 2009; Tsaluchidu et al., 2008) did not show a significant enrichment in our analyses. Since only a few oxidative stress-related orthologs were dysregulated, we conclude that mitochondrial pathways are not predominantly dysregulated on the transcriptional level. A finding that is not unexpected since mitochondrial activity is mostly dysregulated on the proteome level through posttranslational modifications. Signaling hormones and growth factors regulate mitochondrial remodeling processes through fission and fusion in response to fluctuations in oxygen or nutrients (Carlucci et al., 2008). Closer inspection of the downstream effector pathways of oxidative stress revealed pathways significantly enriched on the transcriptional level. One example is ubiquitin-mediated proteolysis, which is a key mechanism in the regulation of mitochondrial homeostasis and in our analysis strongly enriched for depression. Moreover, in response to low oxygen concentration, growth factors, like VEGF and TGFβ, are activated (Carlucci et al., 2008). Thus, our results indirectly support the notion of mitochondrial dysfunction in anxiety and depression albeit through downstream effector pathways.

The systematic correlation approach enabled us to not only reveal critical pathways involved in anxiety and depression, but also to identify novel pathway relationships. One limitation of our method is the exclusion of transcripts which have no known ortholog and are not associated with any pathway. Also, the fact
that a pathway did not reveal a significant enrichment could simply be due to missing GEO transcriptome data. As a consequence we cannot make a statement about pathways that were not found to be significantly enriched. Despite these limitations our study was able to identify a number of pathways related to anxiety- and depression-like phenotypes, which showed strong and significant enrichments by combining data from multiple experiments deposited in the public domain. The method can be extended to other complex disease phenotypes for which transcription analysis data are available.

**Role of the funding source**

This study was supported by a BMBF QuantPro grant and the Max Planck Society. Both funding sources had no influence on study design.

**Contributors**

PG and NM performed research. CD evaluated data. SW supported script programming. CT and FH contributed to the project idea and design. CT was involved in evaluating the data and writing the manuscript.

**Conflict of interest**

None declared.

**Acknowledgement**

This study was supported by a BMBF QuantPro grant and the Max Planck Society.

**Appendix. Supplementary material**

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.jpsychires.2010.12.010.

**References**


Shalev H, Serlin Y, Friedman A. Breaching the blood-brain barrier as a gate to psychiatric disorder. Cardiovasc Psychiatry Neurol; 2009. 278531.


