# Title: Regional expression profiles of risk genes for depression are associated with brain activation patterns in emotion and reward tasks

**Authors**: Arkadiusz Komorowski[1], Ramon Vidal[2], Aditya Singh[3], Matej Murgaš[1], Tonatiuh Pena-Centeno[4], Gregor Gryglewski[1], Siegfried Kasper[5], Jens Wiltfang[6, 7, 8], Rupert Lanzenberger[1], Roberto Goya-Maldonado[3]*

**Affiliations:**

[1]Department of Psychiatry and Psychotherapy, Division of General Psychiatry, Medical University of Vienna, Vienna, Austria.

[2]Max Delbrück Center for Molecular Medicine, Berlin, Germany.

[3]Laboratory of Systems Neuroscience and Imaging in Psychiatry (SNIPLab), Department of Psychiatry and Psychotherapy, University Medical Center Goettingen (UMG), Georg-August University, Von-Siebold-Str. 5, 37075 Goettingen, Germany.

[4]Bioinformatics Unit, Department for Epigenetics and Systems Medicine in Neurodegenerative Diseases, German Center for Neurodegenerative Diseases (DZNE), Von-Siebold-Str. 3a, 37075, Goettingen, Germany.

[5]Center for Brain Research, Medical University of Vienna, Vienna, Austria.

[6]Department of Psychiatry and Psychotherapy, University Medical Center Goettingen (UMG), Georg-August University, Von-Siebold-Str. 5, 37075 Goettingen, Germany.

[7]German Center for Neurodegenerative Diseases (DZNE), Von-Siebold-Str. 3a, 37075, Goettingen, Germany.

[8]Neurosciences and Signalling Group, Institute of Biomedicine (iBiMED), Department of Medical Sciences, University of Aveiro, Aveiro, Portugal.

**One Sentence Summary:** Analysis of the spatial association between whole-brain human gene expression and in-vivo brain activation patterns during emotion and reward processing identified TCF4 and MEF2C as master regulatory genes associated with depressive disorders.

***Correspondence to:**
Dr. Roberto Goya-Maldonado
Labor für Systemische Neurowissenschaften und Bildgebung in der Psychiatrie
Universitätsmedizin Göttingen
Von-Siebold-Str. 5, 37075 Goettingen, Germany
Tel: +49-551-39-22244
Email: roberto.goya@med.uni-goettingen.de

**Abstract**

The exploration of the relationship between gene expression profiles and neural response patterns known to be altered in major depressive disorder provides a unique opportunity to identify novel targets for diagnosis and therapy. Here, we estimated the spatial association between genome-wide transcriptome maps and brain activation patterns from functional magnetic resonance imaging (fMRI) with two established paradigms of great relevance for mood disorders. While task-specific neural responses during emotion processing were primarily associated with expression patterns of genes involved in cellular transport, reward processing was related to neuronal development, synapse regulation, as well as gene transcription. Multimodal integration of single-site and meta-analytic imaging data with risk genes associated with depression revealed a regional susceptibility of functional activity, modulated by master regulators TCF4 and MEF2C. The identification of multiple subordinate genes correlated with fMRI maps and their corresponding regulators presumably will reshape the prospects for neuroimaging genetics.

## Introduction

Over the last decades, genetic and neuroimaging studies have significantly contributed to current knowledge about human neural functions. While single neuroscientific methods contributed to the comprehension of physiological processes as well as pathological alterations in psychiatric disorders, a multimodal integration of large-scale data has proven to be even more conducive for in-depth understanding, especially on the molecular scale. In particular, post-mortem gene expression data from the Allen Human Brain Atlas (AHBA) was repeatedly applied to investigate the relationship between the transcriptome and protein distribution (*1*), brain connectivity (*2*), or morphology (*3*). These studies add up to numerous wide-ranging research findings combining mRNA expression with neuroimaging parameters, partially offering toolboxes for an integrative data analysis (*4*). Different methods offer designated benefits and disadvantages, e.g. the precise temporal resolution of electroencephalography contrasts to its low spatial resolution. In case of positron emission tomography, molecular specificity allows quantification of protein distributions in vivo, but availability of specific radioligands is limited. In general, indirectly measuring brain activation during execution of specific paradigms by means of functional magnetic resonance imaging (fMRI) has evolved into one of the most popular neuroimaging techniques. Although a promising approach to contrast topological brain activation and gene expression patterns making use of the meta-analytic Neurosynth database was presented by Fox and colleagues (*5*), their study lacked disease-related aspects. Other studies, by contrast, have successfully integrated post-mortem data and in vivo imaging findings, to assess the influence of regional gene expression on psychiatric disorders (*6*). While fMRI can depict neural correlates of specific psychological processes with high spatial resolution, measured signal patterns also appear susceptible to genotype variations (*7*). However, the role of specific master regulators (MRs) modulating differentially expressed genes that might influence regional blood oxygenation level dependent (BOLD) signaling has not been resolved yet.

From all psychiatric disorders, major depressive disorder (MDD) is now the leading cause of disability worldwide and a strong contributor to the overall global burden of disease, with increasing prevalence over the years (*8*). Granted that additive genetic effects attribute to approximately 9 % of the variation in liability of MDD (*9*), regional variations in gene expression profiles may determine brain function within a continuum, spanning from a physiological to a more critical pathological state. In this regard, paradigms examining core depressive symptoms like impaired affect modulation or loss of interest and pleasure in common experiences are amongst the best-established within the realms of MDD. Further, alterations of BOLD reactivity during processing of negatively valenced information as well as motivation- or incentive-based learning emphasize the characterization of neuropsychiatric disorders in terms of functions rather than diagnoses (*10, 11*). Positive valence systems and systems for social processes thereby comprise prominent behavioral elements and justify the application of emotion and reward processing paradigms to investigate major domains of human functioning that are currently part of the Research Domain Criteria (RDoC) framework.

In this study, genes with expression patterns correlated with task-specific brain activation were evaluated in regard to biological processes according to the Gene Ontology (GO) database (*12*) as well as their relationship with risk genes for MDD obtained from conventional Genome-wide association studies (GWAS) (*9*). Potential influences of master regulating genes on neuroimaging parameters were evaluated and initial results based on single-site fMRI data subsequently replicated applying meta-analytic activation maps from the Neurosynth online framework (*13*). Ultimately, we aimed to identify individual MRs modulating topological mRNA expression that was associated with brain activation during emotion and reward processing.

## Results

*Topological specificity of emotion and reward processing*

In conjunction with known abnormalities in social interaction and reward responsiveness of patients with depressive disorders, fMRI activation within the social processes and positive valence systems domains of the RDoC framework was evaluated. We minimized unspecific signal variations related to visual, auditory, attentional and executive processing by contrasting brain activation elicited by the experimental condition with task control conditions in each participant. Sequentially, second level analyses provided topological activation elicited by emotion and reward processing (full acquisition and analysis pipeline described in Material and Methods section). The applied emotional face recognition task didn't focus attention on the emotional content of presented faces, but requested gender discrimination instead, provoking rather implicit emotional processing in limbic as well as non-limbic areas such as prefrontal cortices. When testing reward responsiveness, significant activations in dopaminergic brain regions were observed after acceptance of priorly conditioned stimuli, particularly in the mesolimbic reward system (table S1).

To expand the scope of results generated with single-site data, we obtained corresponding uniformity maps from the Neurosynth database for complementary analyses. Detected brain activation clusters were thereby replicated using matching data derived from 91 and 246 studies associated with the terms "fearful faces" and "rewards", respectively. Both meta-analytic maps were in conformity with measured imaging data and can be obtained online from the Neurosynth framework.

*Associations between transcriptome maps and functional brain activation*

For each paradigm, subsequent correlation analyses showed marked associations between functional brain activation and mRNA expression for a large number of individual genes (rankings of genes as well as corresponding Spearman's correlation coefficients for both

datasets are reported in table S2). Due to marked gene expression differences, correlations were performed and ranked separately in cortical and subcortical regions. Findings were highly specific for each paradigm, accounted for by the weak overlap of compiled correlation lists between emotion and reward processing (subcortex: $rho_{RRHO}$ = - 0.282; cortex: $rho_{RRHO}$ = 0.205) (fig. S1). In contrast, high agreement of ranked mRNA-fMRI correlations between two different brain parcellation atlases was observed both for emotional face recognition (subcortex: $rho_{RRHO}$ = 0.697; cortex: $rho_{RRHO}$ = 0.822) and acceptance of monetary rewards (subcortex: $rho_{RRHO}$ = 0.748; cortex: $rho_{RRHO}$ = 0.918) (fig. 1).

Regarding single-site data, functional brain activity evoked by emotion processing correlated positively as well as negatively with whole-brain transcriptome maps. Region-wise correlations yielded similar results compared to voxel-wise analyses, ranging from rho = -0.739 to rho = 0.865 for subcortical and from rho = -0.449 to rho = 0.431 for cortical regions (fig. S2, A and B). Out of all resulting associations between gene expression and brain function the 10 highest positive correlating genes are listed in table 1. In subcortical regions, MALL showed the strongest voxel-wise correlation (rho = 0.633), while C10orf125 showed the highest region-wise correlation (rho = 0.865). In the cortex, SPDYA yielded strongest voxel-wise (rho = 0.328) and FOXN4 strongest region-wise (rho = 0.431) correlation. FOXN4 also showed a high cortical ranking applying voxel-wise analysis (rho = 0.285, $5^{th}$ rank) (fig. 2, A and B).

Analogous to the emotion task, compiled ranked lists included genes strongly correlated with measured functional brain activity related to the reward system, whereby the 10 highest positive correlating genes are listed in table 2. Region-wise analyses yielded higher correlation coefficients than the voxel-wise approach with less prominent associations in the cortex (rho = -0.639 to rho = 0.698), compared to subcortical regions (rho = -0.788 to rho =

0.81). In the subcortex, MDK showed the strongest voxel-wise correlation of all genes (rho = 0.488) and also a high region-wise correlation coefficient (rho = 0.803, 3rd rank) (fig. 2, C and D). Comparing strongest voxel-wise vs. region-wise correlations in the cortex, 7 out of 10 genes were congruent (DUSP3, CA10, PIK3CD, HDAC9, LASS6, GRB14, OLFM3), indicating high agreement between both approaches (fig. S3, A and B). Thereby, gene expression of DUSP3 showed strongest cortical correlations with reward processing both in the voxel-wise (rho = 0.548) as well as the region-wise analysis (rho = 0.698) (fig. S4).

*Ontological analysis of task-specific biological processes*

Considering both strong positive and negative mRNA-fMRI correlations, multiple genes associated with imaging data overlapped with specific gene sets listed in the GO knowledgebase (presented in detail in table S3 and fig. S5, A to C). Notably, a marked redundancy between GO terms was present within each paradigm, indicating rather task-specific associations of molecular programs with gene expression patterns throughout the human brain (fig. 3). Considering neural responses during emotional face recognition in subcortical regions, significantly associated biological programs were related to cellular transport processes and mainly included genes that showed positive correlations between gene expression and imaging data. In line with the assumed relevance of molecular transduction for emotion regulation processes, the most significant overlap was observed for the GO term peptide transport. In total, 203 genes positively correlated with emotion processing were also present within this gene set ($p_{corr} = 0.013$), representing 7.9 % of genes included within the GO term (GO:0015833). Functional brain activation in cortical regions yielded no significant overlap with biological processes. Analyzing neural responses during acceptance of monetary rewards in subcortical regions, associated GO terms mainly included genes with positive mRNA-fMRI correlations, predominantly comprising transcription processes. Highest significance was present for nucleic acid metabolic process (GO:0090304, $p_{corr} < 0.001$, 536

genes, 8.95 %), RNA metabolic process (GO:0016070, $p_{corr} < 0.001$, 486 genes, 8.89 %), as well as cellular macromolecule biosynthetic process (GO:0034645, $p_{corr} < 0.001$, 485 genes, 8.45 %). GO terms associated with genes negatively correlated with the acceptance of monetary rewards within subcortical regions mainly related to synaptic processes and neuronal development. Thereby, most significant GO terms were chemical synaptic transmission (GO:0007268, $p_{corr} < 0.001$, 54 genes, representing 6.19 % of genes included within this term) and modulation of chemical synaptic transmission (GO:0050804, $p_{corr} < 0.001$, 35 genes, 6.4 %). However, in the cortex only two overlapping GO terms were found for genes positively associated with reward processing (autophagy of mitochondrion, GO:0000422, $p_{corr} < 0.05$, 4 genes, 3.96 % and sensory perception of sound, GO:0007605, $p_{corr} = 0.04$, 4 genes, 2.34 %).

*The role of risk genes associated with major depression*

The relationship between single-site neuronal brain activation and 42 functional genes associated with MDD obtained from a pre-defined gene set was investigated to evaluate the superordinate role of these risk genes on measured BOLD activation patterns. Master regulator analysis of co-regulatory networks based on previously compiled ranked lists including strongest mRNA-fMRI correlations and revealed individual candidate MRs for each paradigm that potentially regulate genes positively and negatively associated with imaging data. We found 4 regulators from the MDD risk gene set for emotion and 3 for reward processing in subcortical regions, with 79 and 89 possible targets, respectively ($p < 0.001$) (table 3). Curiously, PAX6, LHX2, as well as MEF2C were associated with negatively correlated genes for both paradigms, indicating a rather superordinate role of these MRs in MDD, regardless of cognitive system. Alternatively, TCF4 was identified both as a regulator for genes negatively correlated with emotion as well as genes positively correlated with reward processing. Overall, strongest regulation was found for genes negatively correlated

with reward processing, whereby LHX2 and MEF2C coordinated more than half of the possible targets. Considering regulation for genes that showed positive correlations with functional brain activation, two MRs (SOX5 and TCF4) were identified for reward processing, while no significant regulators were found for genes positively correlated with emotion processing. Subcortical single-site results were replicated when applying independent meta-analytical data from the Neurosynth framework. Identically, MEF2C emerged as regulator of genes negatively correlated with emotional face recognition as well as acceptance of monetary rewards (fig. 4A), while TCF4 showed an inversed regulation of genes correlated with the emotion (negative association) and reward (positive association) paradigm (fig. 4B). Single-site and meta-analytical findings from gene set enrichment analysis (GSEA) complemented results from GO as well as master regulator analyses, showing an inversed aggregation of MDD risk genes for emotion and reward processing. Rather than focusing solely on ranks of single genes, by means of GSEA we could assess the role of the whole gene set associated with depressive disorders. Regarding emotion processing, risk genes were predominantly enriched within positively correlated genes, while location of the maximum enrichment score (ES) for reward processing emphasized genes showing negative correlations with imaging parameters in both data sets, albeit not reaching statistical significance. Within the Neurosynth replication sample, the maximum subcortical ES yielded 0.275 ($p = 0.051$) for emotion and -0.243 for reward processing ($p = 0.65$), respectively (fig. 4C). In the cortex, the distribution of MDD risk genes was identical (fig. S6).

## Discussion

Here, we applied a comprehensive and integrative methodological approach to investigate the relationship between regional gene expression patterns and macroscopic BOLD responses elicited by emotional face recognition and the acceptance of monetary rewards, under the assumption that strongly correlated genes would coincide with distinct biological programs and genes implicated in depressive disorders. Unbiased screening for novel associations between mRNA expression and functional brain activation resulted in ranked lists of 18,686 genes positively and negatively correlated with BOLD signaling. Similar distributions of Spearman's correlation coefficients were present for emotion (ranging from rho = -0.739 to rho = 0.865 in subcortical and from rho = -0.449 to rho = 0.431 in cortical areas) and reward processing (from rho = -0.788 to rho = 0.81 in the subcortex and from rho = -0.639 to rho = 0.698 in the cortex). Considering strongest brain activation elicited by emotion processing in ventral striatum, amygdala, ventral tegmental area, fusiform gyri, insula and medial prefrontal cortex, in that order, it seems plausible that higher correlation levels were observed in subcortical regions, compared to the cortex. Exploring the GO knowledgebase, we detected task-specific processes related to cellular transport as well as neuronal development, synapse regulation, and transcription for emotion and reward processing, respectively. Notably, associated ontologies were interrelated solely within each fMRI paradigm, thus indicating unique biological programs for both investigated RDoC domains. Focusing on systems for social processes and positive valence systems, the meta-analytic replication sample comprised of activation maps from over 90 fMRI studies. Based on correlations between imaging parameters and gene expression, we identified master regulators associated with depressive disorders and task-specific functional brain activation, TCF4 and MEF2C, in two independent datasets. While MEF2C showed a congruent regulation with negative associations related to emotion and reward processing, TCF4 appeared simultaneously as a regulator for genes negatively correlated with the emotion task, but positively correlated with reward processing.

Albeit not reaching statistical significance, the supplementary GSEA suggested an inversed distribution of the risk gene set for major depression, showing a rather positive association with imaging data for emotional face recognition and a negative association for the acceptance of monetary rewards.

Although the role of gene expression patterns has been reviewed for several fMRI measures (*2, 3*), individual genetic influences on emotional face recognition or adaptive reward-based decision-making have only been evaluated for the presence of single gene variants of functional proteins (i.e. CREB1), irrespective of topological distribution (*7*). Particularly in MDD, additive genetic effects may attribute to individual differences in the phenotype and highlight the importance of large-scale data in systems medicine to resolve unsettled genetic influences on neuroimaging paradigms. While over 322 million people worldwide suffer from depressive disorders, a number that increased by 18.4 % between 2005 and 2015 (*8*), a significant part of the population is also affected by subthreshold depressive symptoms, potentially originating from different levels of genetic susceptibility in relevant neuronal systems. Our findings strongly support the idea of a dimensional genetic background influencing activated brain regions within emotion and reward systems that continuously progresses from physiological to pathological states, recently highlighted within the much-noticed RDoC framework (*11*). In line with the debilitating symptoms of depressive disorders, we investigated paradigms reflecting principal functions of both the reward responsiveness construct within the positive valence systems domain as well as the social communication construct, which is part of the systems for social processes domain. Hyper- as well as hypoactivations of brain regions involved in the integration of social information, reduced reward sensitivity and decision-making efficiency suggest a polygenic nature of depressive disorders with distinct imaging features (*14*). However, identifying the core set of risk genes

is complicated due to widespread and disease-specific network interactions as well as modulatory MRs.

In this study, regulating genes associated with MDD included protein coding MEF2C (Myocyte Enhancer Factor 2C), which plays a role in neuronal development, as well as hippocampus-dependent learning and memory (*15*). Relevance for synapse regulation arises due to alternatively spliced transcript variants involved in neuronal processes, e.g. activated TLR4 signaling or the cAMP response element-binding protein (CREB) pathway. Besides depressive disorders, associations of this transcription activator with other psychiatric disorders like schizophrenia (*16*) or attention-deficit/hyperactivity disorder (*17*) have been reported. Most notably, significant dependence of striatal neuronal activation was described for an identified risk variant in the TMEM161B-MEF2C gene cluster during a reward task, endorsing well-known deficits of reward processing in MDD, observed as anhedonia (*18*). Besides, we affirm previously reported relationships between neuropsychiatric disorders and mutations of TCF4, which has been implicated not only in depression, but also in schizophrenia and autism (*19*, *20*). The transcription factor is mainly characterized by its regulatory role for the proliferation and differentiation of neuronal and glial progenitor cells. Interestingly, the master regulatory role of TCF4 in schizophrenia was recently endorsed by Torshizi et al., based on the analysis of transcriptional networks in two independent datasets (*21*).

Since gene variants exhibit their effects by uncountable molecular mechanisms, a closer investigation of genes strongly associated with imaging parameters, as well as the MRs reported in this study, will prospectively allow further statements about regional protein biosynthesis and allocation of resulting proteins to cellular compartments. For example, DNA-binding transcription factor FOXN4 (Forkhead Box N4), belonging to the Forkhead-box (FOX) superfamily, showed highest correlation with the emotional face recognition

paradigm in the cortex. FOX transcription factors are involved in regulatory biological processes and mutations in forkhead genes have been linked to developmental disorders in humans due to substitutions or frameshifts that disable or remove the DNA binding domain. The subtype FOXN4 thereby expresses developmental functions in neural and non-neural tissues, particularly during spinal neurogenesis by modulating a specific expression mosaic of other proneural factors. Further relevance for neural development was shown by Chen et al. (*22*), who demonstrated location of FOXN4 on neurons and astrocytes, as well as an increased expression after spinal cord injury lesions. Although associations with depressive or other neuropsychiatric disorders have not been published, the modulatory role of FOXN4 as a key transcriptional regulator during developmental processes demands further research, especially since the full set of its targets in the CNS are not known yet. Accordingly, the C10orf125 gene (Fucose Mutarotase, FUOM), expressed in the brain and other tissues, showed highest correlation with emotional face recognition in the subcortex. The corresponding gene transcript, fucose mutarotase, is an enzyme of the fucose-utilization pathway performing the interconversion between α-L-fucose and β-L-fucose on human cell surfaces. Hereof, besides one animal study demonstrating male-like sexual behavior in FUOM knock-out mice, presumably resulting from reduced fucosylation during neurodevelopment (*23*), further associations with pathological states in mammals have not been published for this gene. Regarding reward processing, the protein coding DUSP3 (Dual-specificity phosphatase 3) gene, member of the dual-specificity protein phosphatase subfamily, showed strongest correlation with measured fMRI data in the cortex. Members of these protein tyrosine phosphatases (PTPs) regulate the phosphorylation of the mitogen-activated protein (MAP) kinase signaling pathway and control cell signaling, especially in regard to cytoskeleton reorganization, apoptosis and RNA metabolism. DUSP3 shows a wide expression in different tissues as an opposing factor of protein tyrosine kinases (PTKs) and acts as a central mediator of cellular proliferation and differentiation. Whereas a role in neoplastic disorders,

pathologies related to immunology, angiogenesis as well as Parkinson's disease have been related to anomalous tyrosine phosphorylation (*24*), associations with psychiatric disorders have not been described. In subcortical regions, MDK (midkine) was among the highest correlating genes for the reward paradigm. MDK transcribes one of two growth factors from the heparin-binding cytokine family and plays a role during the differentiation of neurons. An important role of this gene has been suggested within dopaminergic pathways, particularly facilitating neuroprotective effects in neurodegenerative disorders, drug-induced neurotoxicity in the striatum, or after neural injury (*25*). Further, a disease-related publication suggests an influence of midkine on addictive behaviors (*26*), and more recently, associations of elevated serum levels of the neurotrophic factor have been found in autism spectrum disorder (*27*).

Despite advances and decreased costs of high-throughput gene expression profiling the necessity for large cohorts in genetic studies call for collaborations and integrative approaches. By including whole-brain expression patterns, our approach can be discriminated from previous neuroimaging studies solely investigating global effects of disease-related SNPs or environmental factors. Still, a range of influencing factors hampers interpretation of spatial associations between gene expression and functional imaging data, like outdated annotation information, inaccurate sample assignment, or presence of locally correlated genes. Further, ubiquitous noise due to expression of genes with a low spatial dependence occurs, which has also been addressed by Gryglewski et al. (*28*). We performed all correlation analyses separately for cortical and subcortical regions, to prevent a possible bias that might arise from expression differences of the transcriptome between these brain areas. While mRNA levels don't necessarily reflect actual in vivo protein densities, signal alterations in fMRI studies are also impaired by confounding variables, such as network-architectures, structural and functional connectivity measures, or non-specific brain activation. To minimize BOLD signaling elicited by superimposed executive functions, control conditions were

implemented for each task, facilitating specificity of performed correlations for both paradigms. Noteworthy, meta-analytic brain activation maps eventually confirmed findings from single-site fMRI measurements. In contrast to a previous study (*5*) that identified gene-cognition associations based on the Neurosynth framework, we also increased spatial resolution and advanced probe selection of gene expression data by application of interpolated mRNA maps with continuous expression estimates throughout the whole brain. Surely, results of this study are limited due to the structure of applied data sets and reflect rather subtle genetic influences on psychological processes, disregarding the dynamic nature of short-term regulatory mechanisms or individual variations due to genetic ancestry and environmental factors. Inter-individual differences regarding effects of sex, age or genotype cannot be considered, when performing integrative analyses on the basis of the AHBA that originally derived expression values from only 6 post-mortem brains (*29*). Also, predictions of in vivo gene expression are limited due to locally and functionally regulated epigenetic and epitranscriptomic modifications that affect actual protein distribution (*30*). Nonetheless, the AHBA still offers the most comprehensive database for the investigation of human whole-brain gene expression, comprising nearly 20,000 genes. Although a whole-brain proteome atlas including relevant genotypes would reflect actual protein expression more reliably, such database has not been published yet. Within the framework of future studies, the vast potential of the AHBA might even be increased by re-assigning available mRNA probes to corresponding genes on the basis of the latest sequencing information to increase the number of specifically annotated genes. Additionally, harmonized data processing pipelines and methodological guidelines, instead of rather unique approaches to data integration and corresponding statistical measures, could increase comparability between studies (*31*).

Since explanations of altered fMRI activations often neglect the spatial distribution of investigated genetic substrates, evaluating the role of gene expression in neuropsychiatric

disorders requires an integration of high-resolution neural information from molecular sources for more precise interventions in the future. Our analyses highlight the advantages of a comprehensive approach to reveal genetic influences on functional brain imaging parameters by integrating multimodal imaging and large-scale genetic data with sufficient power. While insignificant results of GSEA were partly caused by the small number of analyzed risk genes, the low enrichment of the gene set associated with MDD might also be determined by a significant impact of individual MRs that modulate up- and downregulation of multiple subordinate genes, including risk and non-risk genes. In general, the GWAS performed by Wray and colleagues (*9*) is among the largest ever conducted in psychiatric genetics and provides a solid basis for further research about the genetic architecture of MDD. In this study, we successfully applied this GWAS data to explore imaging paradigms that specifically integrate core depressive symptoms in line with current RDoC, like dysfunction in social interaction and inability to properly experience reward.

## Conclusion

While traditional strategies have solely investigated variations of individual genes without topological allocation, whole-transcriptome expression profiles provide superior information for the understanding of neuropsychiatric disorders. Here, biological programs related to cellular transport, neuronal development, synapse regulation, and transcription processes were specifically associated with two paradigms, highlighting the relationship between gene expression and in-vivo imaging parameters for both the emotion and reward system. Further, identification of regulatory genes TCF4 and MEF2C endorses the investigation of commonly altered brain activation during emotional face recognition and the acceptance of monetary rewards in depressed individuals. Overall, our work exemplifies an integrative approach including complementary information from multiscale data, which seems to be increasingly relevant in the big data era.

## Materials and Methods

*Participants*

Healthy subjects were recruited from the university environment and gave written informed consent to study procedures previously approved by the Ethics Committee of the University Medical Center Göttingen. Included participants were aged between 22 and 52 years (M = 40.5, SD = 14.37), of Caucasian European ethnicity and were fluent in German language. Exclusion criteria comprised contraindications to MRI, past or present psychiatric, neurological, or medical disorders, consumption of psychotropic drugs as well as positive family history of psychiatric disorders. In total, a number of 26 men and 22 women completed two fMRI paradigms including reward processing and emotional face recognition. Excessive movement in any of the three translation (> 2 mm) or rotation (> 2°) planes resulted in exclusion of 4 participants.

*Functional brain imaging*

Functional imaging data was acquired using a 3 T scanner (Siemens Magnetom TRIO, Siemens Healthcare, Erlangen, Germany) and a 32-channel head coil with a 2 x 2 x 2 mm voxel size, TR 2500 ms, TE 33 ms, 70° flip angle, 10 % distance factor, FOV 256 mm and 60 slices with multiband factor of 3 for the acquisition of T2*-weighted images. Imaging data analysis was performed using Statistical Parameter Mapping (SPM12; Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) and Matlab R2015b (The Mathworks Inc., Natick, MA, USA). First, echo planar imaging (EPI) images were standardly preprocessed with slice time correction, realignment, and normalization into the MNI space, as well as smoothing with an 8 x 8 x 8 mm FWHM Gaussian kernel.

Two different fMRI paradigms were analyzed for the same group of participants during two states. First, specific activation maps reflected brain activation during performance of the tasks. In contrast, corresponding control conditions represented non-specific hemodynamic

activity inherent to any task performance during fMRI measurement, caused by unspecific physiological activation, e.g. related to visual, auditory, attentional or motor functions. Estimates of neural activity were initially computed with a general linear model (GLM) for each subject individually (first-level analysis) with nuisance movement parameters regressed as covariates-of-no-interest. Later, experimental and control conditions were evaluated in a group level (second-level analysis) and resulting data representing task-specific brain activation used for further analyses.

*Emotional face recognition*

The paradigm of implicit emotional face recognition contains two different contexts: human faces and geometric objects. Pictures of males and females with varying negative face expressions obtained from the Radboud database (*32*) were presented for 17 s, during which participants responded to the gender of the presented person with a button press. Thereby, perception of emotions was rather implicit, which has been shown to enhance the activation of emotional correlates (*33*). For the control condition, participants were instructed to respond analogously to the shape of an object, either an ellipse or a rectangle, positioned in the face area and made from scrambling original face trials. All trials were controlled for brightness, contrast and presented in a very similar composition. The activation patterns representing the experimental and control conditions were computed using first-level (single-subject) contrasts of the trials from emotional faces and object blocks respectively, which were then used for second-level (group) analysis, as standardly performed for random effects model.

*Reward processing*

For this study, a previously established fMRI paradigm was implemented, which has been broadly used to investigate physiological and pathological reward mechanisms (*34*). Briefly, participants performed a modified delayed match to sample task, including two contexts

involving previously conditioned stimuli to monetary rewards: acceptance or rejection of rewards, e.g. pressing a button when green squares are shown. Subjects were instructed they would receive 30 € for their participation, and that they were able to double this amount according to their task performance. As control trials, subjects responded with a button press e.g. to yellow squares as stimuli that required motor performance as well as attentional and memory resources but were not conditioned to monetary reward. To compute the control condition, first-level (single-subject) contrasts of correctly matched sample trials within the same experimental block of reward trials were used. For experimental conditions, first-level experimental contrasts were calculated from brain activation elicited during acceptance of previously conditioned stimuli. At group-level, activations related to experimental and control trials were contrasted to measure activation patterns, as standardly performed for random effects model.

*Meta-analytic functional brain activation*

Besides fMRI data obtained from participants performing two different tasks at our institution, we evaluated large-scale meta-analytical imaging data from the Neurosynth platform (https://neurosynth.org/), which provides probabilistic brain activity mappings computed from an automated synthesis of results from published fMRI studies. The online database combines text-mining and machine-learning techniques to generate meta-analyses of currently 1335 imaging terms from 14371 fMRI studies (*13*). For this study, Neurosynth activation maps related to emotional and reward fMRI data were downloaded in MNI152 2 mm space to validate measured neuronal activation patterns obtained at our institution.

To evaluate statistical overlap between meta-analytical terms and single-site fMRI activation maps, we correlated activation patterns of both datasets. As a first step, activation maps, with z-scores representing how brain regions are related to the chosen term, were generated from the Neurosynth online database. Since the online user interface only provides thresholded

maps, we used the Neurosynth toolbox for python to create unthresholded maps that were further smoothed using 8 x 8 x 8 mm FWHM to match the kernel size of single-site data. Subsequently, we assessed the association of meta-analytic data with measured fMRI maps by means of region-wise Spearman's correlation coefficients in cortical and subcortical regions, whereby imaging paradigms related to recognition of negative faces and the acceptance of monetary rewards matched well with the Neurosynth terms "fearful faces" and "rewards", respectively. Due to low fMRI activity, cerebellum was not considered for the comparison. Within the Neurosynth database, each term includes a uniformity test map that provides the degree to which each voxel is consistently activated in studies that use a given term. Secondly, an association test map is provided that controls for base rate activation differences between brain regions. Association test maps display specific regions, where functional activation related to a distinct term occur more consistently, compared to studies that don't mention that term. In contrast to single-site fMRI data, meta-analytical information comprises rather positive values, due to the sparse reporting of brain regions showing negative neuronal activation in most neuroimaging studies. Hence, when processing unthresholded data from the uniformity as well as association test maps, mainly positive values determined the association analysis between Neurosynth and measured fMRI maps.

*Whole brain gene expression*

The AHBA (www.brain-map.org) consists of microarray assessments from 3702 brain tissue samples collected across 6 human donors (1 female, mean age = 42.5, SD = 13.4) derived from diverse regions of the brain, extensively described in the original publication (*29*). As delineated by other authors (*31*), multiple data processing steps such as gene annotation, data filtering, probe selection or sample assignment need to be considered to facilitate subsequent correlation analyses between expression and neuroimaging data. Also, by using common parcellation schemes expression levels in numerous brain regions devoid of tissue samples

remain indefinite, thus demanding the generation of whole brain transcriptome maps. To sufficiently meet proposed methodological requirements, gene expression data ($log_2$-values) from the AHBA were obtained according to Gryglewski et al. (*28*) to compensate for inter-individual differences between mRNA probes and donor brains as well as sparse anatomical sampling, attributable to high efforts in tissue preparation and processing. Using Gaussian process regression, unbiased whole-brain transcriptome maps comprising expression of 18,686 genes associated with Entrez Gene IDs at all cortical and subcortical structures were created (predicted transcriptome maps are available for download at www.meduniwien.ac.at/neuroimaging/mRNA.html). Nonetheless, for a certain number of remaining genes transcriptome maps have not been predicted sufficiently due to presence of insensitive probes or missing allocations to gene IDs.

*Spatial correlation between gene expression and brain activity*

Initially, gene lists ranking correlations between transcriptome maps and functional brain activation patterns elicited by selected paradigms were compiled, whereby the ranking of each gene depended on its correlation strength with the corresponding single-site BOLD activation map. After initial inspection, the cerebellum was excluded from further analysis, due to marginal activation during both fMRI paradigms. On the basis of marked differences in gene expression between broad anatomical regions (*29*) correlation analyses were assessed within cortical and subcortical regions separately. To conduct association analyses, all available transcriptome maps were aligned with group-averaged fMRI activation maps in MNI space for each paradigm. To account for partly non-symmetrical distribution of mRNA data and existence of outliers, Spearman's correlation coefficients were calculated between each gene-imaging pair (mRNA expression vs. fMRI activation). Main findings are reported for region-wise analyses, along with additional results for voxel-wise correlations (total number of voxels was 129,817 in the cortex and 10,863 in subcortex; zero-values outside of the

investigated area were excluded). Thereby, each statistical map was segmented into brain regions according to the Brainnetome atlas, which was selected for primary analyses, because it labels a sufficient number of subcortical (n = 36) and cortical (n = 210) regions-of-interest (ROIs) (*35*). A complementary analysis with fewer brain regions (12 subcortical and 78 cortical ROIs) was done using the automated anatomical labeling (AAL) brain atlas (*36*), in order to evaluate influences of different parcellation methods. Both atlases were aligned with fMRI maps as well as transcriptome maps in MNI space using SPM12, while extraction of ROIs and correlation analyses were performed in MATLAB2018a (www.mathworks.com). Statistical significances of region-wise correlations were assessed by means of randomization tests, including 10,000 iterations. For each sampled permutation, mRNA values were randomly shuffled and correlated with fMRI data. Two-sided p-values were calculated as the proportion of sampled permutations where the absolute value was greater than the true correlation coefficient (non-shuffled data).

*Identification of overlap between analyzed data set*

To compare various sets of mRNA-fMRI correlations we used Rank-Rank Hypergeometric Overlap (RRHO) package (version 1.26.0) in R (https://www.bioconductor.org/packages/release/bioc/html/RRHO.html), which allows statistical testing of the extent of overlap between two ranked lists. RRHO determines the degree of differential expression observed in profiling experiments using the hypergeometric distribution. While originally applied for the comparison of gene expression profiles between different microarray platforms or types of model system, we used RRHO to compare genes that ranked according to relevant measures of differential information, in this case the correlation strength with fMRI data. Thereby, genes of two datasets were ranked according to their names and corresponding ranks tested for statistical overlap. We provide both a graphical representation of the characteristics of analyzed datasets (corresponding p-values)

as well as a statistical measure of overlap ($rho_{RRHO}$). A high overlap implies that positively correlating genes of a given list show high ranks in the second list, while genes with a negative correlation are also negatively associated in the alternative list. Applying this method offered the advantage of using the whole continuum of previously correlated genes without the need to truncate the list by pre-defined thresholds for each combination of used datasets.

*Analysis of biological processes*

Making use of the GO knowledgebase (*12*), a comprehensive resource for computational analysis of large-scale data, we explored enriched biological processes that included genes with expression patterns highly correlated with each fMRI paradigm. Cytoscape plugin "ClueGO" (*37*) was used to compute GO enrichment (default parameters), comprising all listed terms at the time of analysis (18,361 biological process terms). Only findings of specialized biological processes (GO levels higher than 4) were reported, to attain more conclusive information about underlying genetic substrates and their functions within investigated emotion and reward systems. Compiled ranked lists originating from region-wise analyses were included for the investigation of gene ontology, due to more refined results. For each paradigm, the GO analysis was performed separately for positively and negatively correlated genes with cut-off values for correlation coefficients above rho = 0.5 and below rho = -0.5, which excluded non-spatially depending genes showing insufficient associations with fMRI data. Significance level was set to a p-value minimum of 0.05, after applying Bonferroni correction to the enriched GO terms. Considering potential overlaps across imaging paradigms, interrelations between all significant biological processes found in the GO analysis were quantified by means of Spearman's correlation coefficients.

*Association of risk genes for major depression with functional imaging data*

Topological expression patterns of genes associated with major depression and their association with neuroimaging parameters have not yet been investigated. Regarding genetic risks for depressive disorders, recently 44 genetic risk loci for major depression were identified in a genome-wide association meta-analysis by Wray et al., which included 135,458 cases and 344,901 controls (*9*). Based on this meta-analysis, we investigated if the MDD risk gene set contained MRs of previously compiled genes showing strong correlations between gene expression and imaging data. Thereby, the role of all 42 functional risk genes proposed by Wray et al. was evaluated for each fMRI paradigm (table S4).

By using the cytoscape plugin iRegulon (*38*), we performed master regulator analysis separately for all positive and negative mRNA-fMRI correlations above rho = 0.6 and below rho = -0.6, respectively. Rather low thresholds were set to ensure a sufficient number of evaluated genes in subcortical regions. In the cortex, insufficient data availability hampered analyses of co-regulatory networks, due to generally lower correlation coefficients of mRNA-fMRI associations compared to the subcortex. Parameters, such as enrichment score and significances, were used as default; distance from TSS was set to 500bp. We compared the master regulators on each predicted regulon with the genes associated with risk for major depression.

Further, correlations between fMRI data and reported risk genes for MDD were compared with all remaining genes by means of GSEA (*39*), to assess clustering of risk genes for each fMRI paradigm. Regarding applied methodology of GSEA, enrichment score reflects the degree to which the analyzed MDD risk gene set is overrepresented within each ranked list. Thereby, we evaluated whether MDD risk genes were randomly distributed throughout each ranked list or primarily found at the top (showing positive correlations with fMRI data) or bottom (showing negative correlations). We used the GSEA implementation in R available in the package clusterProfiler (*40*). For quantification, the ES was calculated by a stepwise

increase or decrease of the total sum statistic of the ranked list, depending on the ranking of MDD risk genes, as described by Subramanian et al. (*39*). Thereby, the ES, corresponding to a weighted Kolmogorov-Smirnov-like statistic, represents the maximum deviation from zero with positive or negative values indicating enrichment of positive or negative correlations, respectively.

**Supplementary Materials**

Fig. S1. Rank–rank hypergeometric overlap (RRHO) visual representation of measured imaging data for reward vs. emotion processing.

Fig. S2. Voxel-wise vs. region-wise correlation analyses of single-site imaging data during emotion processing.

Fig. S3. Voxel-wise vs. region-wise correlation analyses for single-site imaging data during reward processing.

Fig. S4. Comparison of functional brain activation during reward processing and mRNA expression of DUSP3 in cortical regions.

Fig. S5. Enriched biological programs for emotion and reward processing based on ontological structure.

Fig. S6. Gene Set Enrichment Analysis (GSEA) for emotion and reward processing, including risk genes associated with major depression.

Table S1. Regions showing functional brain activation during emotional face recognition and acceptance of monetary rewards.

Table S2. Associations between functional imaging and transcriptome data.

Table S3. Enriched biological programs for emotion and reward processing based on ontological structure.

Table S4. Previously published risk genes associated with major depression.

**References and Notes**


1. A. Komorowski, G. M. James, C. Philippe, G. Gryglewski, A. Bauer, M. Hienert, M. Spies, A. Kautzky, T. Vanicek, A. Hahn, T. Traub-Weidinger, D. Winkler, W. Wadsak, M. Mitterhauser, M. Hacker, S. Kasper, R. Lanzenberger, Association of Protein Distribution and Gene Expression Revealed by PET and Post-Mortem Quantification in the Serotonergic System of the Human Brain. *Cerebral Cortex* **27**, 117-130 (2017).
2. J. Richiardi, A. Altmann, A.-C. Milazzo, C. Chang, M. M. Chakravarty, T. Banaschewski, G. J. Barker, A. L. W. Bokde, U. Bromberg, C. Büchel, P. Conrod, M. Fauth-Bühler, H. Flor, V. Frouin, J. Gallinat, H. Garavan, P. Gowland, A. Heinz, H. Lemaître, K. F. Mann, J.-L. Martinot, F. Nees, T. Paus, Z. Pausova, M. Rietschel, T. W. Robbins, M. N. Smolka, R. Spanagel, A. Ströhle, G. Schumann, M. Hawrylycz, J.-B. Poline, M. D. Greicius, L. Albrecht, C. Andrew, M. Arroyo, E. Artiges, S. Aydin, C. Bach, T. Banaschewski, A. Barbot, G. Barker, N. Boddaert, A. Bokde, Z. Bricaud, U. Bromberg, R. Bruehl, C. Buchel, A. Cachia, A. Cattrell, P. Conrod, P. Constant, J. Dalley, B. Decideur, S. Desrivieres, T. Fadai, H. Flor, V. Frouin, J. Gallinat, H. Garavan, F. G. Briand, P. Gowland, B. Heinrichs, A. Heinz, N. Heym, T. Hubner, J. Ireland, B. Ittermann, T. Jia, M. Lathrop, D. Lanzerath, C. Lawrence, H. Lemaitre, K. Ludemann, C. Macare, C. Mallik, J.-F. Mangin, K. Mann, J.-L. Martinot, E. Mennigen, F. Mesquita de Carvahlo, X. Mignon, R. Miranda, K. Muller, F. Nees, C. Nymberg, M.-L. Paillere, T. Paus, Z. Pausova, J.-B. Poline, L. Poustka, M. Rapp, G. Robert, J. Reuter, M. Rietschel, S. Ripke, T. Robbins, S. Rodehacke, J. Rogers, A. Romanowski, B. Ruggeri, C. Schmal, D. Schmidt, S. Schneider, M. Schumann, F. Schubert, Y. Schwartz, M. Smolka, W. Sommer, R. Spanagel, C. Speiser, T. Spranger, A. Stedman, S. Steiner, D. Stephens, N. Strache, A. Strohle, M. Struve, N. Subramaniam, L. Topper, R. Whelan, S. Williams, J. Yacubian, M. Zilbovicius, C. P. Wong, S. Lubbe, L. Martinez-Medina, A. Fernandes, A. Tahmasebi; IMAGEN consortium, Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241–1244 (2015).
3. J. Shin, L. French, T. Xu, G. Leonard, M. Perron, G. B. Pike, L. Richer, S. Veillette, Z. Pausova, T. Paus, Cell-Specific Gene-Expression Profiles and Cortical Thickness in the Human Brain. *Cerebral Cortex* **28**, 3267–3277 (2018).
4. G. Rizzo, M. Veronese, P. Expert, F. E. Turkheimer, A. Bertoldo, MENGA: A New Comprehensive Tool for the Integration of Neuroimaging Data and the Allen Human Brain Transcriptome Atlas. *PLoS One* **11**, e0148744 (2016).
5. A. S. Fox, L. J. Chang, K. J. Gorgolewski, T. Yarkoni, Bridging psychology and genetics using large-scale spatial analysis of neuroimaging and neurogenetic data. *bioRxiv* **012310** (2014).
6. I. A. Romme, M. A. de Reus, R. A. Ophoff, R. S. Kahn, M. P. van den Heuvel, Connectome Disconnectivity and Cortical Gene Expression in Patients With Schizophrenia. *Biological psychiatry* **81**, 495-502 (2017).
7. C. Wolf, H. Mohr, E. K. Diekhof, H. Vieker, R. Goya-Maldonado, S. Trost, B. Krämer, M. Keil, E. B. Binder, O. Gruber, CREB1 Genotype Modulates Adaptive Reward-Based Decisions in Humans, *Cerebral Cortex* **26**, 2970-2981 (2015).
8. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**: 1789-1858 (2018).
9. N. R. Wray, S. Ripke, M. Mattheisen, M. Trzaskowski, E. M. Byrne, A. Abdellaoui, M. J. Adams, E. Agerbo, T. M. Air, T. M. F. Andlauer, S.-A. Bacanu, M. Bækvad-

Hansen, A. F. T. Beekman, T. B. Bigdeli, E. B. Binder, D. R. H. Blackwood, J. Bryois, H. N. Buttenschøn, J. Bybjerg-Grauholm, N. Cai, E. Castelao, J. H. Christensen, T.-K. Clarke, J. I. R. Coleman, L. Colodro-Conde, B. Couvy-Duchesne, N. Craddock, G. E. Crawford, C. A. Crowley, H. S. Dashti, G. Davies, I. J. Deary, F. Degenhardt, E. M. Derks, N. Direk, C. V. Dolan, E. C. Dunn, T. C. Eley, N. Eriksson, V. Escott-Price, F. H. F. Kiadeh, H. K. Finucane, A. J. Forstner, J. Frank, H. A. Gaspar, M. Gill, P. Giusti-Rodríguez, F. S. Goes, S. D. Gordon, J. Grove, L. S. Hall, E. Hannon, C. S. Hansen, T. F. Hansen, S. Herms, I. B. Hickie, P. Hoffmann, G. Homuth, C. Horn, J.-J. Hottenga, D. M. Hougaard, M. Hu, C. L. Hyde, M. Ising, R. Jansen, F. Jin, E. Jorgenson, J. A. Knowles, I. S. Kohane, J. Kraft, W. W. Kretzschmar, J. Krogh, Z. Kutalik, J. M. Lane, Y. Li, Y. Li, P. A. Lind, X. Liu, L. Lu, D. J. MacIntyre, D. F. MacKinnon, R. M. Maier, W. Maier, J. Marchini, H. Mbarek, P. McGrath, P. McGuffin, S. E. Medland, D. Mehta, C. M. Middeldorp, E. Mihailov, Y. Milaneschi, L. Milani, J. Mill, F. M. Mondimore, G. W. Montgomery, S. Mostafavi, N. Mullins, M. Nauck, B. Ng, M. G. Nivard, D. R. Nyholt, P. F. O'Reilly, H. Oskarsson, M. J. Owen, J. N. Painter, C. B. Pedersen, M. G. Pedersen, R. E. Peterson, E. Pettersson, W. J. Peyrot, G. Pistis, D. Posthuma, S. M. Purcell, J. A. Quiroz, P. Qvist, J. P. Rice, B. P. Riley, M. Rivera, S. Saeed Mirza, R. Saxena, R. Schoevers, E. C. Schulte, L. Shen, J. Shi, S. I. Shyn, E. Sigurdsson, G. B. C. Sinnamon, J. H. Smit, D. J. Smith, H. Stefansson, S. Steinberg, C. A. Stockmeier, F. Streit, J. Strohmaier, K. E. Tansey, H. Teismann, A. Teumer, W. Thompson, P. A. Thomson, T. E. Thorgeirsson, C. Tian, M. Traylor, J. Treutlein, V. Trubetskoy, A. G. Uitterlinden, D. Umbricht, S. Van der Auwera, A. M. van Hemert, A. Viktorin, P. M. Visscher, Y. Wang, B. T. Webb, S. M. Weinsheimer, J. Wellmann, G. Willemsen, S. H. Witt, Y. Wu, H. S. Xi, J. Yang, F. Zhang, V. Arolt, B. T. Baune, K. Berger, D. I. Boomsma, S. Cichon, U. Dannlowski, E. C. J. de Geus, J. R. DePaulo, E. Domenici, K. Domschke, T. Esko, H. J. Grabe, S. P. Hamilton, C. Hayward, A. C. Heath, D. A. Hinds, K. S. Kendler, S. Kloiber, G. Lewis, Q. S. Li, S. Lucae, P. F. A. Madden, P. K. Magnusson, N. G. Martin, A. M. McIntosh, A. Metspalu, O. Mors, P. B. Mortensen, B. Müller-Myhsok, M. Nordentoft, M. M. Nöthen, M. C. O'Donovan, S. A. Paciga, N. L. Pedersen, B. W. J. H. Penninx, R. H. Perlis, D. J. Porteous, J. B. Potash, M. Preisig, M. Rietschel, C. Schaefer, T. G. Schulze, J. W. Smoller, K. Stefansson, H. Tiemeier, R. Uher, H. Völzke, M. M. Weissman, T. Werge, A. R. Winslow, C. M. Lewis, D. F. Levinson, G. Breen, A. D. Børglum, P. F. Sullivan; eQTLGen; 23andMe; Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics* **50**, 668-681 (2018).

10. L. C. Foland-Ross, I. H. Gotlib, Cognitive and neural aspects of information processing in major depressive disorder: an integrative perspective. *Front Psychol* **3**, 489 (2012).
11. C. A. Sanislow, M. Ferrante, J. Pacheco, M. V. Rudorfer, S. E. Morris, Advancing translational research using nimh research domain criteria and computational methods. *Neuron* **101**, 779-782 (2019).
12. The Gene Ontology Consortium, The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic acids research* **47**, D330-D338 (2018).
13. T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, T. D. Wager, Large-scale automated synthesis of human functional neuroimaging data. *Nature methods* **8**, 665-670 (2011).
14. B. Knutson, J. P. Bhanji, R. E. Cooney, L. Y. Atlas, I. H. Gotlib, Neural responses to monetary incentives in major depression. *Biological psychiatry* **63**, 686-692 (2008).

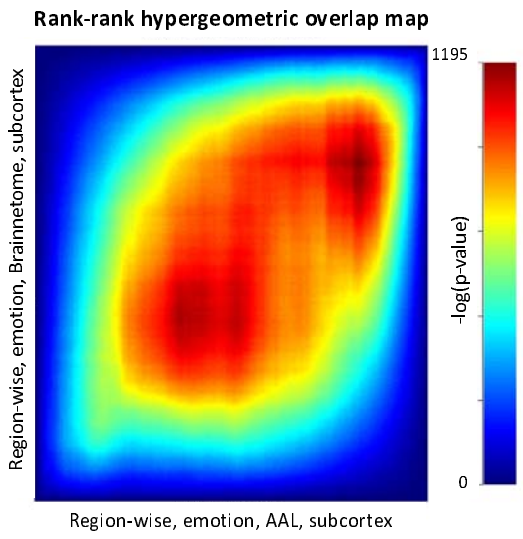
15. A. C. Barbosa, M. S. Kim, M. Ertunc, M. Adachi, E. D. Nelson, J. McAnally, J. A. Richardson, E. T. Kavalali, L. M. Monteggia, R. Bassel-Duby, E. N. Olson, MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function. *Proc Natl Acad Sci USA* **105**, 9391-9396 (2008).
16. A. C. Mitchell, B. Javidfar, V. Pothula, D. Ibi, E. Y. Shen, C. J. Peter, L. K. Bicks, T. Fehr, Y. Jiang, K. J. Brennand, R. L. Neve, J. Gonzalez-Maeso, S. Akbarian, MEF2C transcription factor is associated with the genetic and epigenetic risk architecture of schizophrenia and improves cognition in mice. *Molecular psychiatry* **23**, 123-132 (2018).
17. A. A. Shadrin, O. B. Smeland, T. Zayats, A. J. Schork, O. Frei, F. Bettella, A. Witoelar, W. Li, J. A. Eriksen, F. Krull, S. Djurovic, S. V. Faraone, T. Reichborn-Kjennerud, W. K. Thompson, S. Johansson, J. Haavik, A. M. Dale, Y. Wang, O. A. Andreassen, Novel Loci Associated With Attention-Deficit/Hyperactivity Disorder Are Revealed by Leveraging Polygenic Overlap With Educational Attainment. *J Am Acad Child Adolesc Psychiatry* **57**, 86-95 (2018).
18. C. Muench, M. Schwandt, J. Jung, C. R. Cortes, R. Momenan, F. W. Lohoff, The major depressive disorder GWAS-supported variant rs10514299 in TMEM161B-MEF2C predicts putamen activation during reward processing in alcohol dependence. *Translational psychiatry* **8**, 131 (2018).
19. H. Li, Y. Zhu, Y. M. Morozov, X. Chen, S. C. Page, M. D. Rannals, B. J. Maher, P. Rakic, Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Molecular psychiatry* **24**, 1235-1246 (2019).
20. A. T. Amare, A. Vaez, Y. H. Hsu, N. Direk, Z. Kamali, D. M. Howard, A. M. McIntosh, H. Tiemeier, U. Bültmann, H. Snieder, C. A. Hartman, Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. *Molecular psychiatry*, 1 (2019).
21. A. Doostparast Torshizi, C. Armoskus, H. Zhang, M. P. Forrest, S. Zhang, T. Souaiaia, O. V. Evgrafov, J. A. Knowles, J. Duan, K. Wang, Deconvolution of transcriptional networks identifies TCF4 as a master regulator in schizophrenia. *Science advances* **5**, eaau4139 (2019).
22. X. Chen, Y. Yao, J. Guan, X. Chen, F. Zhang, Up-regulation of FoxN4 expression in adult spinal cord after injury. *J Mol Neurosci* **52**, 403-409 (2014).
23. D. Park, D. Choi, J. Lee, D. S. Lim, C. Park, Male-like sexual behavior of female mouse lacking fucose mutarotase. *BMC Genet* **11**, 62 (2010).
24. P. Cohen, D. R. Alessi, Kinase drug discovery–what's next in the field?. *ACS Chem Biol* **8**, 96-104 (2012).
25. Y. Yoshida, H. Sakakima, F. Matsuda, M. Ikutomo, Midkine in repair of the injured nervous system. *Br J Pharmacol* **171**, 924-930 (2014).
26. E. Gramage, C. Pérez-García, M. Vicente-Rodríguez, S. Bollen, L. Rojo, G. Herradón, Regulation of extinction of cocaine-induced place preference by midkine is related to a differential phosphorylation of peroxiredoxin 6 in dorsal striatum. *Behav Brain Res* **253**, 223-231 (2013).
27. E. Esnafoglu, S. Cirrik, Increased serum midkine levels in autism spectrum disorder patients. *Int J Neurosci* **128**, 677-681 (2018).
28. G. Gryglewski, R. Seiger, G. M. James, G. M. Godbersen, A. Komorowski, J. Unterholzner, P. Michenthaler, A. Hahn, W. Wadsak, M. Mitterhauser, S. Kasper, R. Lanzenberger, Spatial analysis and high resolution mapping of the human whole-brain transcriptome for integrative analysis in neuroimaging. *NeuroImage* **176**, 259-267 (2018).

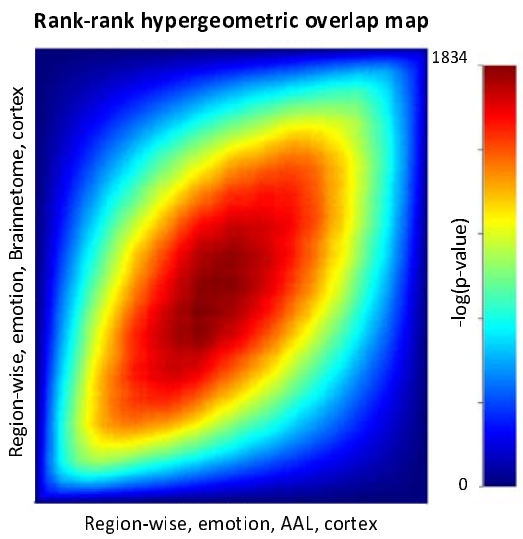
29. M. J. Hawrylycz, E. S. Lein, A. L. Guillozet-Bongaarts, E. H. Shen, L. Ng, J. A. Miller, L. N. van de Lagemaat, K. A. Smith, A. Ebbert, Z. L. Riley, C. Abajian, C. F. Beckmann, A. Bernard, D. Bertagnolli, A. F. Boe, P. M. Cartagena, M. M. Chakravarty, M. Chapin, J. Chong, R. A. Dalley, B. D. Daly, C. Dang, S. Datta, N. Dee, T. A. Dolbeare, V. Faber, D. Feng, D. R. Fowler, J. Goldy, B. W. Gregor, Z. Haradon, D. R. Haynor, J. G. Hohmann, S. Horvath, R. E. Howard, A. Jeromin, J. M. Jochim, M. Kinnunen, C. Lau, E. T. Lazarz, C. Lee, T. A. Lemon, L. Li, Y. Li, J. A. Morris, C. C. Overly, P. D. Parker, S. E. Parry, M. Reding, J. J. Royall, J. Schulkin, P. A. Sequeira, C. R. Slaughterbeck, S. C. Smith, A. J. Sodt, S. M. Sunkin, B. E. Swanson, M. P. Vawter, D. Williams, P. Wohnoutka, H. R. Zielke, D. H. Geschwind, P. R. Hof, S. M. Smith, C. Koch, S. G. Grant, A. R. Jones, An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391-399 (2012).
30. T. Maier, M. Güell, L. Serrano, Correlation of mRNA and protein in complex biological samples. *FEBS lett*, **583**, 3966-3973 (2009).
31. A. Arnatkevic□iūtė, B. D. Fulcher, A. Fornito, A practical guide to linking brain-wide gene expression and neuroimaging data. *NeuroImage* **189**, 353-367 (2019).
32. O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk, A. D. Van Knippenberg, Presentation and validation of the Radboud Faces Database. *Cognition and emotion* **24**, 1377-1388 (2010).
33. M. L. Keightley, G. Winocur, S. J. Graham, H. S. Mayberg, S. J. Hevenor, C. Grady, An fMRI study investigating cognitive modulation of brain regions associated with emotional processing of visual stimuli. *Neuropsychologia* **41**, 585-596 (2003).
34. R. Goya-Maldonado, K. Weber, S. Trost, E. Diekhof, M. Keil, P. Dechent, O. Gruber, Dissociating pathomechanisms of depression with fMRI: bottom-up or top-down dysfunctions of the reward system. *Eur Arch Psychiatry Clin Neurosci* **265**, 57-66 (2015).
35. L. Fan, H. Li, J. Zhuo, Y. Zhang, J. Wang, L. Chen, Z. Yang, C. Chu, S. Xie, A. R. Laird, P. T. Fox, S. B. Eickhoff, C. Yu, T. Jiang, The Human Brainnetome Atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex* **26**, 3508-3526 (2016).
36. N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, M. Joliot, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273-289 (2002).
37. G. Bindea, B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.-H. Fridman, F. Pagès, Z. Trajanoski, J. Galon, ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091-1093 (2009).
38. R. Janky, A. Verfaillie, H. Imrichová, B. Van de Sande, L. Standaert, V. Christiaens, G. Hulselmans, K. Herten, M. Naval Sanchez, D. Potier, D. Svetlichnyy, Z. Kalender Atak, M. Fiers, J.-C. Marine, S. Aerts, iRegulon: From a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol* **10**, e1003731 (2014).
39. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**, 15545-15550 (2005).
40. G. Yu, L. G. Wang, Y. Han, Q. Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16**, 284-287 (2012).

**Acknowledgements:** We thank the diploma students of the Neuroimaging Labs (NIL) for medical support. Parts of this study have been presented by AK at the 32nd ECNP Congress, 07-10 September 2019, Copenhagen, Denmark. **Funding:** With relevance to this work there is no conflict of interest to declare. AS and RGM are funded by the German Federal Ministry of Education and Research (Bundesministerium fuer Bildung und Forschung, BMBF: 01 ZX 1507, ‘‘PreNeSt - e:Med’’). MM is funded by the Austrian Science Fund FWF DOC 33-B27. GG was recipient of a DOC-fellowship of the Austrian Academy of Sciences at the Department of Psychiatry and Psychotherapy, MUV. JW is supported by an Ilídio Pinho professorship, iBiMED (UID/BIM/04501/2013) and FCT project PTDC/DTP-PIC/5587/2014 at the University of Aveiro, Portugal. **Author contributions:** RGM and RL conceived the study; AK and RGM designed the study, primarily interpreted results, and drafted the article; RGM, RV, AS, MM, TPC, and GG made significant contributions to data collection, including quality control, data processing and statistical analysis; SK and JW revised the study and contributed to the intellectual content. All authors have critically revised the article and approved it for publication. **Competing interests:** SK received grants/research support, consulting fees and/or honoraria within the last three years from Angelini, AOP Orphan Pharmaceuticals AG, Celegne GmbH, Eli Lilly, Janssen-Cilag Pharma GmbH, KRKA-Pharma, Lundbeck A/S, Mundipharma, Neuraxpharm, Pfizer, Sanofi, Schwabe, Servier, Shire, Sumitomo Dainippon Pharma Co. Ltd. and Takeda. RL received travel grants and/or conference speaker honoraria within the last three years from Bruker BioSpin MR, Heel, and support from Siemens Healthcare regarding clinical research using PET/MR; he is shareholder of BM Health GmbH since 2019. The remaining authors declare no competing interests. **Data and materials availability:** The functional imaging data generated during the current study are available from the corresponding author upon reasonable request. The correlation lists generated in this study, the associated biological processes, as well as the analyzed risk genes are included in the supplementary information files. All meta-analytical

imaging datasets analyzed are available in the Neurosynth repository, https://neurosynth.org/, whole-brain transcriptome maps are available from the Neuroimaging Labs (NIL), www.meduniwien.ac.at/neuroimaging/mRNA.html, and gene ontology data are available from the Gene Ontology Consortium, http://geneontology.org/.

**Figures:**


Rank-rank hypergeometric overlap map
1195
0
-log(p-value)
Region-wise, emotion, Brainnetome, subcortex
Region-wise, emotion, AAL, subcortex


Rank-rank hypergeometric overlap map
1834
0
-log(p-value)
Region-wise, emotion, Brainnetome, cortex
Region-wise, emotion, AAL, cortex

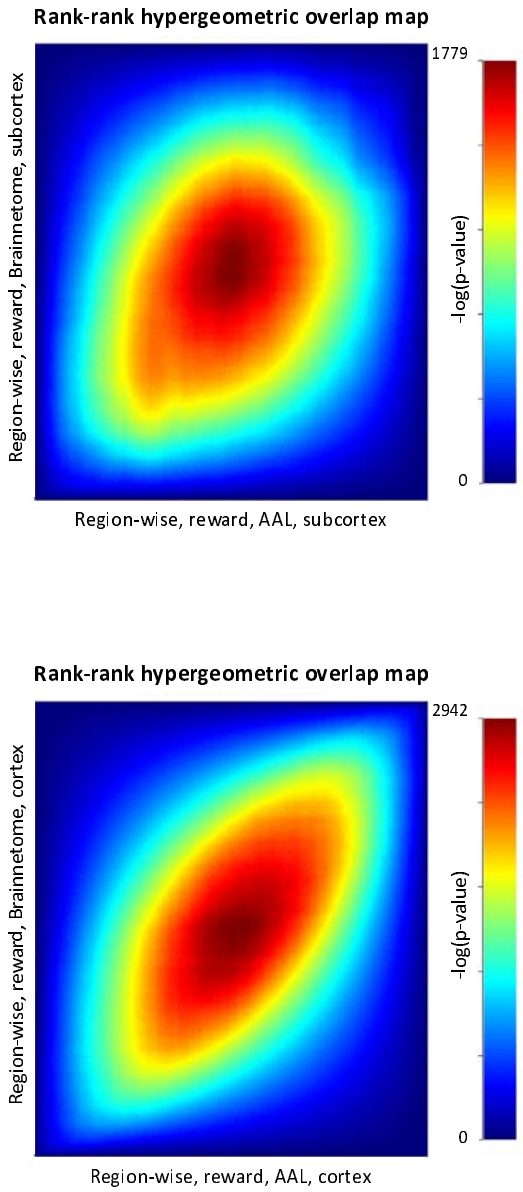
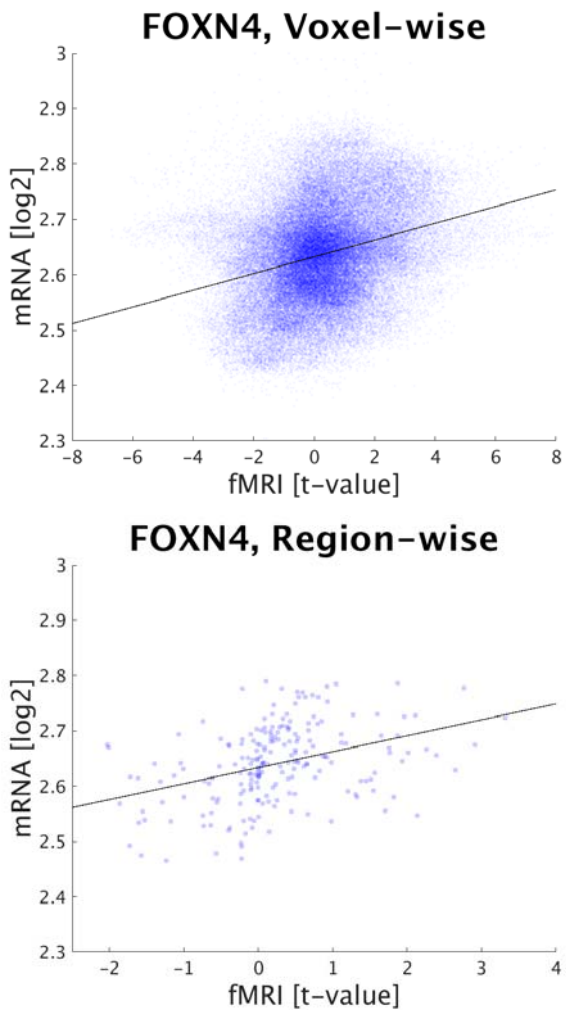


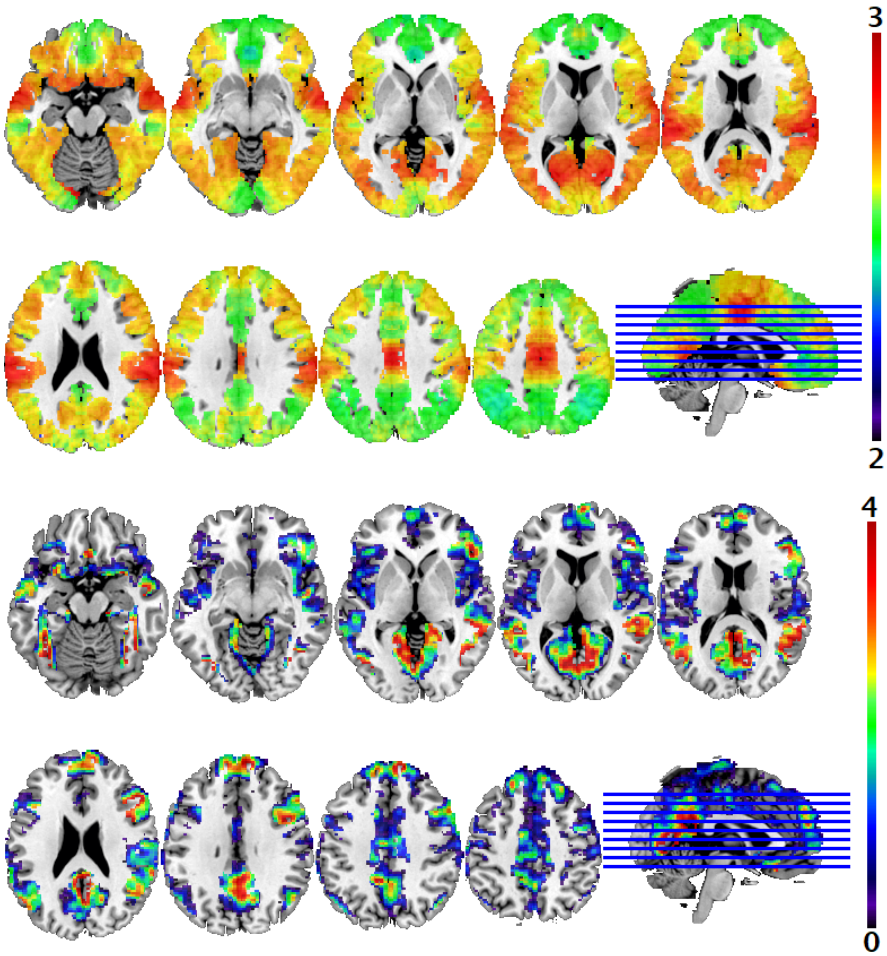

Fig. 1. Rank–rank hypergeometric overlap (RRHO) visual representation of measured imaging data applying automated anatomical labeling (AAL) vs. the Brainnetome atlas. Analyzing single-site data, region-wise RRHO comparing ranked lists including 18,686 genes indicated high agreement between the atlases. Genes with congruent correlation coefficients (either positive or negative) show higher statistical significance in the bottom left and top right corner. Comparisons of both parcellation methods were performed for emotional face

recognition in the subcortex ($rho_{RRHO}$ = 0.697) and cortex ($rho_{RRHO}$ = 0.822), as well as for reward processing in subcortical ($rho_{RRHO}$ = 0.748) and cortical regions ($rho_{RRHO}$ = 0.918).

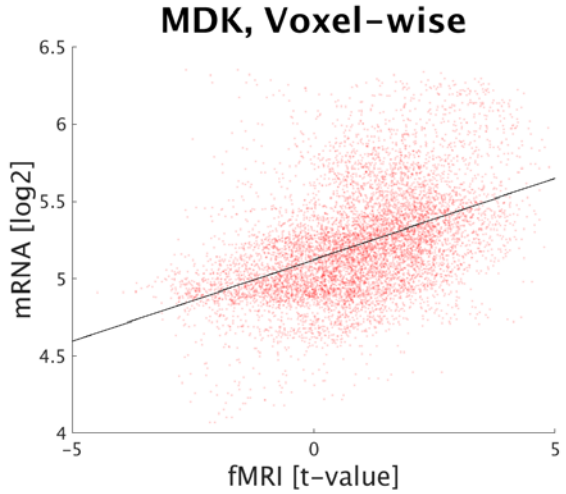
**(A)**

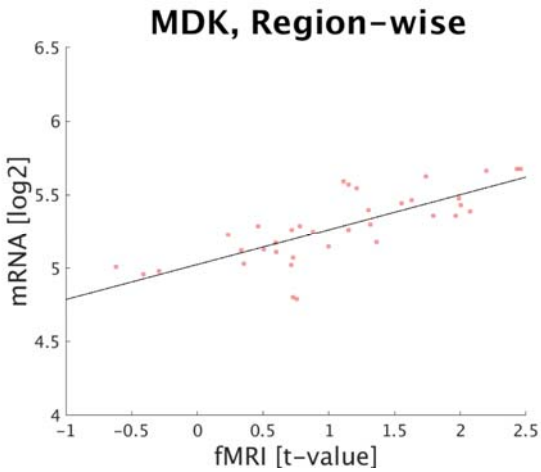


**(B)**




**(C)**






**(D)**

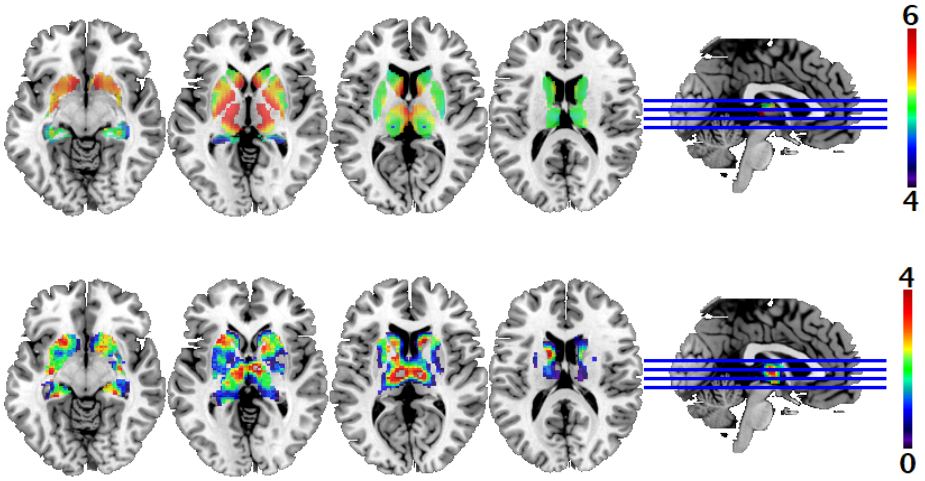



Fig. 2. Comparison of task-specific functional brain activation and mRNA expression of high-correlating genes in cortical and subcortical regions. **(A)** The scatter plots depict correlations between cortical mRNA levels of FOXN4 and single-site imaging data (emotional face recognition) for voxel-wise (rho = 0.285; 129,817 voxels) and region-wise (rho = 0.431; 210 regions, $p < 0.0001$) analyses. Each dot represents expression values and corresponding imaging parameters at target coordinates or within anatomical regions, respectively. **(B)** Cortical FOXN4 gene expression and brain activation patterns during emotion processing are visualized in MNI space based on whole-brain transcriptome maps ($\log_2$) and functional magnetic resonance imaging data (t-value). **(C)** Associations between subcortical mRNA levels of MDK and single-site imaging data (acceptance of monetary rewards) for voxel-wise

(rho = 0.488; 10,863 voxels) and region-wise (rho = 0.803; 36 regions, $p < 0.0001$) analyses. **(D)** Subcortical MDK gene expression ($log_2$) and brain activation patterns during reward processing (t-value).

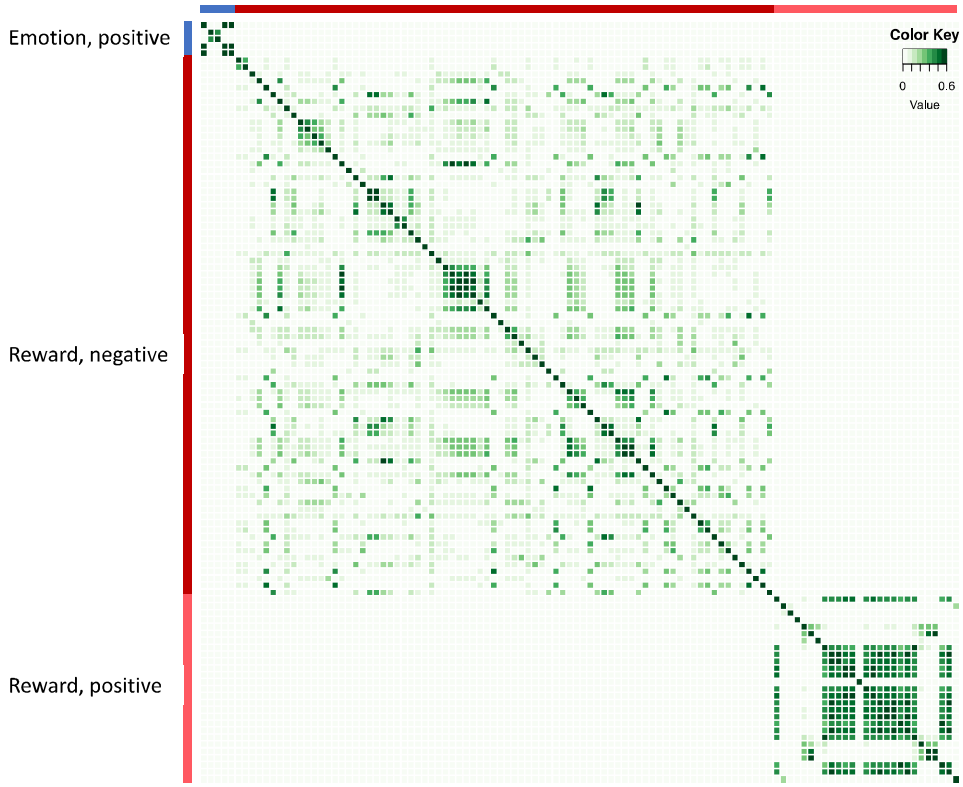
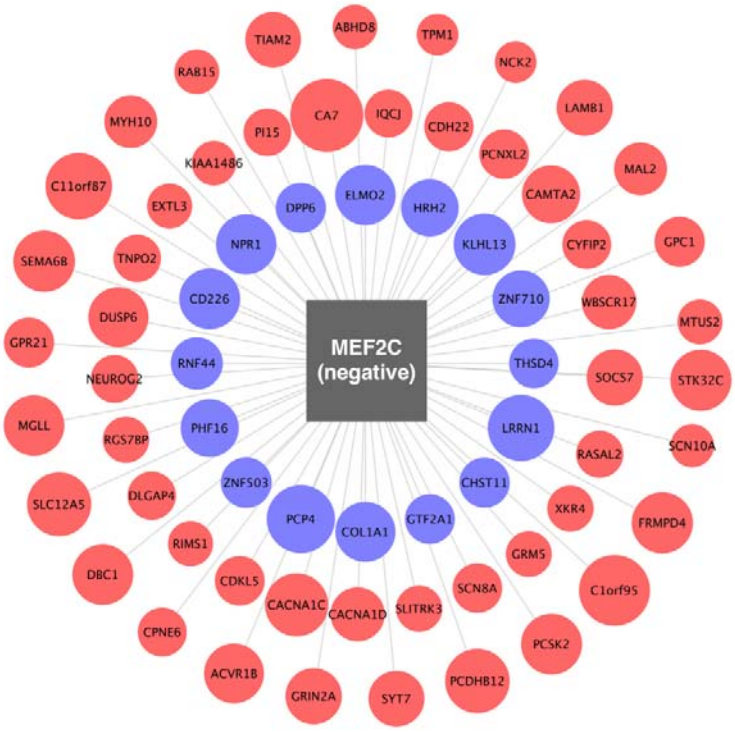


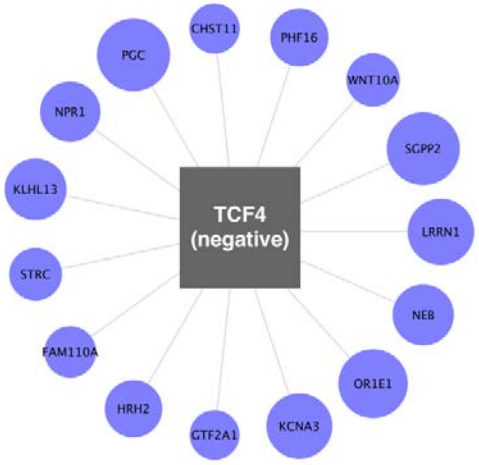

Fig. 3. Heatmap depicting interrelations of associated biological processes listed within the gene ontology (GO) knowledgebase for emotion and reward processing. Overall, 5 (emotion, positive), 78 (reward, negative), and 27 (reward, positive) GO terms that overlapped with genes strongly correlated with single-site imaging data were compared in subcortical regions (GO terms of analyzed biological programs are provided in supplementary table 3). Results yielded marked redundancies for both paradigms (emotional face recognition: blue; acceptance of monetary rewards: red). Considering the multitude of possible associations of 18,686 genes with all listed GO terms, analyses were restricted to higher ontological

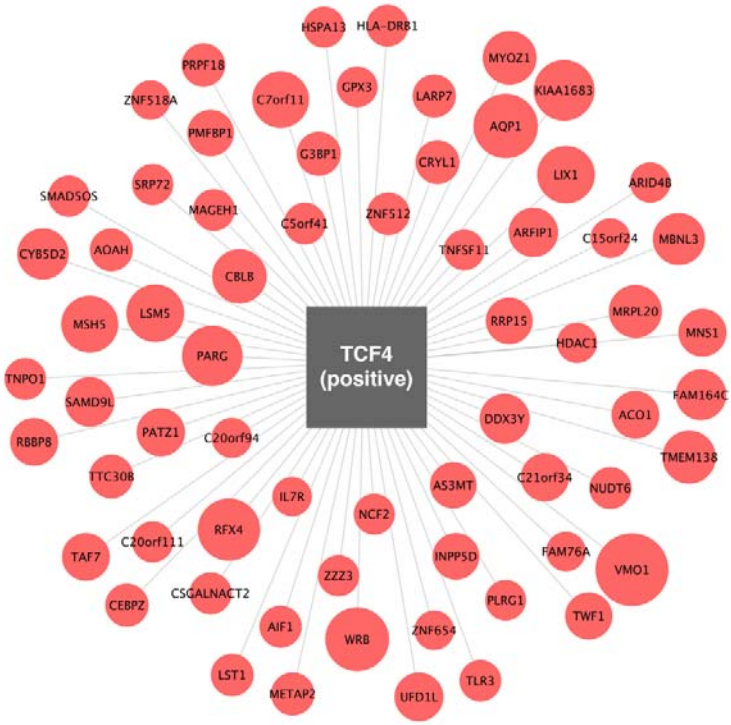
hierarchies (GO levels above 4). Further, only reasonable correlations between imaging and gene expression data were analyzed, with a cut-off value of rho < -0.5 and > 0.5, respectively.

**(A)**




**(B)**




TCF4
(positive)
HSPA13
HLA-DRB1
MYOZ1
KIAA1683
PRPF18
ZNF518A
C7orf11
GPX3
LARP7
AQP1
PMFBP1
G3BP1
CRYL1
LIX1
ARID4B
SMAD5OS
SRP72
MAGEH1
C5orf41
ZNF512
ARFIP1
C15orf24
MBNL3
CYB5D2
AOAH
TNFSF11
CBLB
MRPL20
MSH5
LSM5
RRP15
MNS1
HDAC1
PARG
TNPO1
FAM164C
SAMD9L
ACO1
DDX3Y
RBBP8
PATZ1
C20orf94
TMEM138
TTC30B
C21orf34
NUDT6
AS3MT
IL7R
NCF2
RFX4
C20orf111
FAM76A
TAF7
INPP5D
VMO1
ZZZ3
CSGALNACT2
PLRG1
CEBPZ
TWF1
AIF1
WRB
ZNF654
LST1
TLR3
METAP2
UFD1L

**(C)**

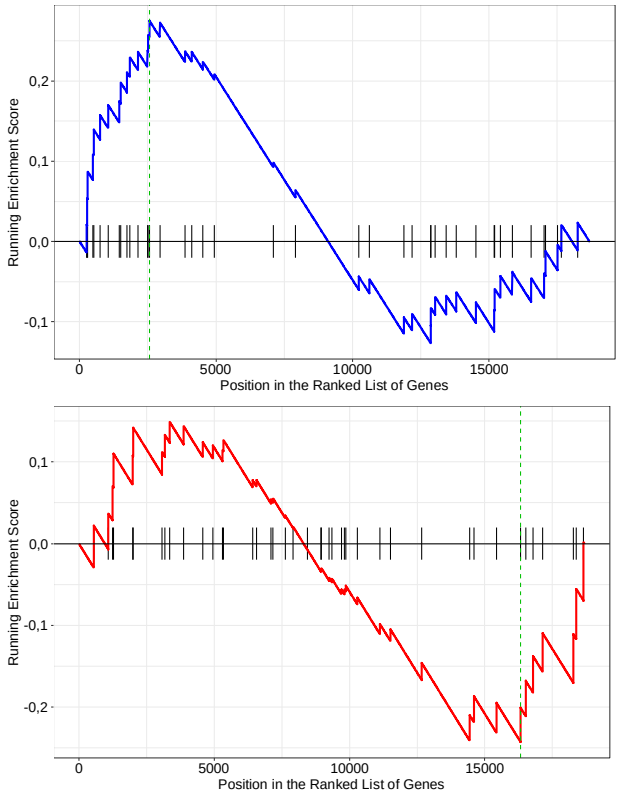

Running Enrichment Score
0,2
0,1
0,0
-0,1
0
5000
10000
15000
Position in the Ranked List of Genes
Running Enrichment Score
0,1
0,0
-0,1
-0,2
0
5000
10000
15000
Position in the Ranked List of Genes

Fig. 4. Relationship between risk genes for depressive disorders and genes correlated with functional brain imaging. **(A)** Two master regulators associated with depressive disorders and corresponding subordinate genes correlated with emotion and reward processing present in single-site and meta-analytical data sets are presented. Graphical visualizations are based on associations between gene expression and single-site imaging data in subcortical regions ($p < 0.001$). The size of each circle corresponds to the absolute value of Spearman's correlation coefficient of the respective gene. MEF2C mainly regulates genes with expression patterns negatively correlated with emotion (16 targets, blue) and reward processing (51 targets, red). **(A)** In contrast, TCF4 inversely regulates genes showing negative associations with emotion (15 targets), but positive associations with reward processing (67 targets). **(C)** Gene set enrichment analysis revealed an inversed relationship between emotion and reward processing in the subcortex. Vertical lines on the x-axis represent positions of 42 functional risk genes within each ranked list including 18,686 genes, dashed lines mark the locations of the maximum enrichment score (ES) which draws a density line depicting the peak enrichment of risk genes. Analyzing data from the Neurosynth framework, ES yielded a positive value for emotion (0.275, $p = 0.051$) and a negative value for reward processing (-0.243, $p = 0.65$).

## Tables

Table 1. Ranking of Spearman's correlation coefficients for genes with expression patterns showing highest positive associations with single-site imaging data (emotional face recognition).

| Subcortex | | | | Cortex | | | |
|---|---|---|---|---|---|---|---|
| Voxel-wise correlations | | Region-wise correlations | | Voxel-wise correlations | | Region-wise correlations | |
| rho | Gene name | rho | Gene name | rho | Gene name | rho | Gene name |
| 0.633 | MALL | 0.865 | C10orf125 | 0.328 | SPDYA | **0.431** | **FOXN4** |
| 0.616 | HRASLS5 | 0.845 | PTRH1 | **0.298** | **CCDC62** | 0.395 | PIK3R6 |
| 0.614 | FAT4 | 0.818 | GHRLOS | 0.296 | PYGO2 | 0.390 | RYBP |
| 0.612 | SCARA5 | 0.817 | SLC24A4 | 0.289 | CPZ | 0.387 | STC1 |
| 0.610 | MESP1 | 0.817 | AC022098.3 | **0.285** | **FOXN4** | **0.384** | **CCDC62** |
| 0.608 | LINC00260 | 0.812 | ZNF280C | 0.283 | FRMD3 | 0.379 | FUBP1 |
| 0.604 | SKAP2 | 0.810 | FUT1 | 0.282 | PHOX2B | 0.378 | SMYD1 |
| 0.602 | SCPEP1 | 0.804 | NLE1 | 0.280 | ATXN10 | 0.356 | MIA2 |
| 0.596 | CRHBP | 0.803 | FBP1 | 0.276 | XAGE3 | 0.356 | TMPRSS4 |
| 0.594 | RAB3GAP1 | 0.803 | C16orf55 | 0.273 | EGFL6 | 0.350 | DLG3 |

Footnote: Bold names correspond to genes that were ranked within the 10 highest positive correlating genes both for voxel-wise and region-wise analyses. All listed region-wise correlation coefficients were significant in permutation tests ($p < 0.0001$).

Table 2. Ranking of Spearman's correlation coefficients for genes with expression patterns showing highest positive associations with single-site imaging data (acceptance of monetary rewards).

| Subcortex | | | | Cortex | | | |
|---|---|---|---|---|---|---|---|
| Voxel-wise correlations | | Region-wise correlations | | Voxel-wise correlations | | Region-wise correlations | |
| rho | Gene name | rho | Gene name | rho | Gene name | rho | Gene name |
| **0.488** | **MDK** | 0.810 | VMO1 | **0.548** | **DUSP3** | **0.698** | **DUSP3** |
| 0.481 | HELLS | 0.805 | OSTM1 | **0.547** | **CA10** | **0.681** | **CA10** |
| 0.475 | RBBP8 | **0.803** | **MDK** | **0.542** | **PIK3CD** | **0.680** | **PIK3CD** |
| 0.453 | ATF1 | **0.798** | **KRT18P19** | **0.533** | **GRB14** | **0.653** | **HDAC9** |
| **0.451** | **KRT18P19** | 0.791 | IMPACT | 0.517 | ASS1 | **0.640** | **LASS6** |
| 0.450 | CD274 | **0.787** | **USP24** | **0.516** | **LASS6** | 0.639 | CCNYL1 |
| 0.446 | C8orf22 | 0.784 | NEK1 | **0.505** | **HDAC9** | **0.637** | **GRB14** |
| 0.442 | SALL4 | 0.782 | RCBTB2 | 0.504 | FBXL2 | **0.627** | **OLFM3** |
| **0.440** | **USP24** | 0.782 | PCBD2 | **0.497** | **OLFM3** | 0.609 | SHC1 |
| 0.439 | SFRP5 | 0.781 | CD99 | 0.492 | TMEM150C | 0.607 | NT5DC2 |

Footnote: Bold names correspond to genes that were ranked within the 10 highest positive correlating genes both for voxel-wise and region-wise analyses. All listed region-wise correlation coefficients were significant in permutation tests ($p < 0.0001$).

Table 3. Master regulator analysis in subcortical regions for emotion and reward processing.

| | **Emotional face recognition** | | **Acceptance of monetary rewards** | |
|---|---|---|---|---|
| | **Negative gene associations** | **Positive gene associations** | **Negative gene associations** | **Positive gene associations** |
| **Single-site imaging data** | TCF4 (15/79)<br>MEF2C (16/79)<br>LHX2 (18/79)<br>PAX6 (18/79) | --- | PAX6 (20/89)<br>LHX2 (43/89)<br>MEF2C (51/89) | SOX5 (57/513)<br>TCF4 (67/513) |
| **Meta-analytical imaging data** | MEF2C (52/1332)<br>TCF4 (649/1332) | --- | MEF2C (21/348) | TCF4 (87/208) |

Footnote: Master regulators were evaluated in single-site as well as meta-analytical imaging datasets by means of co-regulatory networks built with iRegulon software ($p < 0.001$). Ranked genes that showed highest correlation with brain activation maps and 42 functional risk genes associated with major depression were used as input data. Values in parenthesis represent the number of targets of each regulatory gene within the group and total number of possible targets.