Functional brain connectome-transcriptional landscape linking to transdiagnostic factors of psychopathological symptoms

Zhiyi Chen (chenzhiyi@email.swu.edu.cn)
Third Military Medical University https://orcid.org/0000-0003-1744-4647

Xuerong Liu
Experimental Research Center for Medical and Psychological Science (ERC-MPS), School of Psychology, Third Military Medical University

Wei Li
Third Military Medical University

Yuanyuan Hu
Chinese Academy of Sciences

Bowen Hu
Faculty of Psychology, Southwest University

Ting Xu
The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, University of Electronic Science and Technology of China

Rong Zhang
Faculty of Psychology, Southwest University

Yancheng Tang
School of Business and Management, Shanghai International Studies University

Lei Xia
Third Military Medical University

Jingxuan Zhang
Third Military Medical University

Zhibing Xiao
Beijing Normal University https://orcid.org/0000-0002-0457-0233

Ji Chen
Department of Psychology and Behavioral Sciences, Zhejiang University https://orcid.org/0000-0002-9476-1737

Zhengzhi Feng

Yuan Zhou
Institute of Psychology, Chinese Academy of Sciences https://orcid.org/0000-0001-6378-1711

Qinghua He

Southwest University https://orcid.org/0000-0001-6396-6273

Jiang Qiu

Southwest University

Xu Lei

Southwest University

Hong Chen

Southwest University

Shaozheng Qin

Beijing Normal University https://orcid.org/0000-0002-1859-2150

Tingyong Feng

Southwest University https://orcid.org/0000-0001-9278-6474

Article

Keywords:

Posted Date: February 13th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3890054/v1

Additional Declarations: There is **NO** Competing Interest.

Research Article

Functional brain connectome-transcriptional landscape linking to transdiagnostic factors of psychopathological symptoms

Running title: Edge-centric transdiagnostic biomarker of depression and anxiety

Authors: Zhiyi Chen[1,2,3]✉[†], PhD, Xuerong Liu[1†], Msc, Wei Li[1†], MBBS, Msc, Yuanyuan Hu[4,5†], Msc, Bowen Hu[6†], Msc, Ting Xu[7,8], Msc, Rong Zhang[2], Msc, Yangcheng Tang[9], Lei Xia[1], Msc, Jing-Xuan Zhang[1], PhD, Zhibing Xiao[6], Msc, Ji Chen[10], PhD, Zhengzhi Feng[1], PhD, MD, Yuan Zhou[4,5]✉, PhD, Qinghua He[2], PhD, Jiang Qiu[2], PhD, Xu Lei[2], PhD, Hong Chen[2], PhD, Shaozheng Qin[6]✉, PhD, Tingyong Feng [2]✉, PhD

Affiliation:

[1] Experimental Research Center for Medical and Psychological Science (ERC-MPS), School of Psychology, Third Military Medical University, Chongqing, 400038, China

[2] School of Psychology, Southwest University, Chongqing, 400415, China

[3] Key Laboratory of Cognition and Personality, Ministry of Education, 400415, China

[4] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China

[5] Department of Psychology, University of Chinese Academy of Sciences, Beijing, 100049, China

[6] State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, 100032, China

[7] The Center of Psychosomatic Medicine, Sichuan Provincial Center for Mental Health, Sichuan Provincial People's Hospital, Chengdu, 611731, China

[8] The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, University of Electronic Science and Technology of China, Chengdu, 611731, China

[9] School of Business and Management, Shanghai International Studies University, Shanghai, 200083, China

[10] Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, 310013 China

[†] **Contributed equally**

✉ **Corresponding at:**

Zhiyi Chen (chenzhiyi@tmmu.edu.cn), Experimental Research Center of Medical and Psychological Science, School of Psychology, Third Military Medical University, Chongqing, 400038 P.R. China; ShaoZheng Qin (szqin@bnu.edu.cn), State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100032 P.R. China; Yuan Zhou (zhouyuan@psych.ac.cn), Institute of Psychological, Chinese Academy of Sciences, Beijing, 100101 P.R. China or Tingyong Feng (fengty0@swu.edu.cn), School of Psychology, Southwest University, Chongqing, 400015 P.R. China

Statements:

Data availability statements:

Imaging data are still pending embargo given this data collection project is ongoing. Resultant data and materials to reproduce the present study are available and accessible at Science Data Bank repository (ScienceDB, https://doi.org/10.57760/sciencedb.13908). Gene expression data that was used for transcriptional analysis can be found in the ABHA database (https://human.brainmap.org/static/download); Data and codes for neurotransmitters can be found at Hansen's atlas https://github.com/netneurolab/hansen_receptors. Codes for estimating gradients and PLS components can be found at Xia's work (https://github.com/mingruixia/MDD_ConnectomeGradient). Data for twin study are available from the Beijing Twins Brain-Behavior Association Project. Brain maps and plots were built by Surf Ice (https://www.nitrc.org/projects/surfice) and R packages.

Funding statements:

This work was supported by National Natural Science Foundation of China (32300907, 32130045), PLA Talent Program Foundation (2022160258), AMU-RD Scholar Foundation (202211001) and the National Key Research and Development Program of China (2022YFC2705201).

Conflict of interest disclosure:

All the authors report no potential competing conflicts.

Ethics approval statement:

All procedures performed in this study were in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study protocol has been approved by the Institutional Review Board (IRB) of the (Faculty of Psychology) Southwest University.

Author contributions: Zhiyi Chen and Tingyong Feng: Conceptualization, Methodology, Software, Visualization, and Writing - Original Draft; Xuerong Liu, Ting Xu, Rongzhang, Wei Li, Yancheng Tang: Data curation, Methodology and Validation; Lei Xia, Jing-Xuan Zhang, Zhibing Xiao, Yuanyuan Hu, Bowen Hu: Data curation; Ji Chen, and Shaozheng Qin: Writing - Review & Editing and Methodology; Yuan Zhou, Qinghua He, Jiang Qiu, Xu Lei, Hong Chen and Zhengzhi Feng: Validation and Data curation; Zhiyi Chen and Tingyong Feng: Project administration and Funding acquisition.

ORCID:

Zhiyi Chen, 0000-0003-1744-4647
Xuerong Liu, 0000-0002-9236-5773
Ting Xu, 0000-0001-9278-6474
Wei Li, 0009-0004-1347-9302
Ji Chen, 0000-0002-9476-1737
Yuanyuan Hu, 0000-0001-9019-6198
Bowen Hu, 0000-0002-5446-5669

Zhengzhi Feng, 0000-0001-6144-5044
Jing-Xuan Zhang, 0000-0002-8979-5107
Qinghua He, 0000-0001-6396-6273
Xu Lei, 0000-0003-2271-1287
Shaozheng Qin, 0000-0002-1859-2150
Hong Chen, 0000-0003-4283-6095
Yuan Zhou, 0000-0001-6378-1711
Tingyong Feng, 0000-0001-9278-6474

Summary:
4,992 Main text (not including title, abstract, acknowledgment, references, tables/figure)
6 of 6 Colored Figures
5 of 5 Extended Data Figures
3 of 3 Supplemental Materials
14 of 14 Supplemental Methods
19 of 19 Supplemental Results
17 of 17 Supplemental Color Figures
43 of 43 Supplemental Tables

Abstract

Transdiagnostic factors are considered promising in elucidating the etiological underpinnings of psychiatric comorbidities, especially in anxiety and depression. However, their symptom-centered neurobiological substrates, encompassing the genetic macro-micro-molecular brain functional landscape, remain elusive. Here, we develop edge-centric functional brain connectome-based predictive models for transdiagnostic factors of anxiety and depression symptoms (sTDF). These factors are estimated from nonlinear Gaussian topological schemes in a nationwide sample and a twin dataset. Edge-centric connectome was found to be reproducible and generalizable neural signatures for the sTDF, showing high sensitivity in neurally representing the sTDF from edge-centric similarity patterns of attention/frontoparietal networks. Such edge-centric signatures were found moderately heritable. Genetic transcriptional analyses further revealed the biological enrichment for gene expression patterns of these edge-centric connectome emerging into vessel systems and metabolism of $CMRO_2$ for sTDF, especially for cerebellar development in late-childhood-to-young-adulthood. Our findings shed lights on the neurobiological architectures of sTDF by clarifying edge-centeric connectome-transcriptional signature.

Introduction

Affective disorders consisting of broad internalizing problems (e.g., anxiety disorder, depressive disorder and bipolar disorder) are of still leading mental health problems in the globe, with strikingly high prevalence of both anxiety and depression disorders and considerable lifetime comorbidities between them (60%)[1,2]. To elucidate the etiological understructure of such comorbidities , cognitive behavioral theories have proposed to move disorder-specific categorical nosology forward into disorder-across "transdiagnostic" components, and have been substantiated by showing common behavioral latent factors (e.g., *p* factor), pharmacological/behavioral treatments and even neurobiological mechanisms, to well-formulate such comorbidites[3-7]. Rather to well-established "paradigm shifting" in the psychiatric diagnosis on patients[8-10], such transdiganostic insights into shared symptoms of anxiety and depression in the "subclinical or at-risk" population are still underdeveloped. Supporting that, the clinical staging models renewed the definition of disease/disorder into a cross-symptom continuum from "preclinical manifestation" to "late-stage disease", rather than a conclusive binary classification[11,12]. Another promising theoretical framework to tackle challenges derived from historically nosologcal structure, that is the Research Domain Criteria (RDoC), also puts forward to decompose psychiatric disorders into preclinical developmental symptom-across "Domains" (e.g., early environment, genetic risks and brain molecular-cellular-circuit vulnerability) and ensuing psychiatric outcome (i.e., diagnosed disorder)[13,14]. Thus, existing evidence may resonate an imperative need to probe the transdiagnostic factors of anxiety and depression symptoms outside clinical patients, especially in their neurobiological architectures.

To capture biologically-explicable markers in the transdiagnosis for psychiatric patients, network neuroscience is remarkably propelling our understandings of intrinsically neural architectures of psychiatric comorbidity from discrete regional dysfunctions into systematic connectome-based perturbations[15,16]. Connectome, a completed component to describe intrinsic region-to-region functional connections (rFC) in the whole brain, has been broadly demonstrated as a fundamental principal of brain functioning[17,18]. Rather mapping the psychiatric categories into specific brain local anomalies or plain connectivity, connectome-based models conceptualized behavioral dysfunctions or neuropsychiatric disorders as resultant distinct phenotype of perturbing brain intrinsic connection profiles[19,20]. Promisingly, cumulative evidence emerged to substantiate that connectome-based features outperformed in predicting depressive/anxiety disorders and even comorbid ones than markers characterized by regional changes or plain neural circuits[21,22]. In this vein, it may resonate that connectome-based prediction paves a feasible way to clarify integrative neurobiological signatures of transdiagnostic architectures for preclinical or "at-risk" population.

Rather to rFC that currently dominates in network neuroscience, the edge-centric FC (eFC) to capture between-rFC communication patterns provided promising insights in featuring brain connectome[23]. The eFC measured the similar patterns of co-fluctuation of spontaneous brain activities at each instant time point[24]. Probing into eFC characterizations of brain connectome surpassed traditional rFC by enabling to track the larger scale and higher dimension of neural network architectures, especially aiding us in delving into the brain-symptom continuous

spectrum[25-27]. As a novel neuromarker, the eFC has been shown reliable and plausible neural representations in classifications of both neuropsychiatric disorder and neurological disease[28-30]. Therefore, extending neural representations of eFC into connectome-based predictive model in characterizing transdiagnostic symptoms could enrich intrinsic neuromarkers of common factor of anxiety and depression.

Notably, brain connectome, as one of the most robust neuromarkers in transpsychiatric conditions, has been demonstrated to be a potent candidate for endophenotype or intermediate phenotype to bridge preclinical/clinical phenotype into genetically molecular understructure[31,32]. Supporting this standpoint, by using the Allen Human Brain Atlas (AHBA) that shared a high-throughout microarray sequencing dataset identifying spatial transcriptome across whole brain, there are robustly empirical evidence to show the genetic transcriptional signatures of connectome-derived phenotype in categorical psychiatric conditions, such as major depressive disorder and schizophrenia[33-36]. By leveraging the ABHA, delving into neuropathological mechanisms of psychiatric conditions could be closely looped from genetically molecular architectures to macroscale connectome-derived changes and ensuing phenotype (symptoms)[37,38]. In addition to phenotypic association for transdiagnosis of anxiety and depression, they share nearly 40% genetic contributions[39]. Thus, it may intensively necessitate a well-established study to clarify neurobiological signatures of such symptom-centered transdiagnostic factors by insights from a genetic connectome-transcriptional landscape into the preclinical populations.

To this end, we aim to clarify edge-centric (eFC) conncetome-based brain signatures of the transdiagnostic factor of anxiety and depression symptoms (sTDF). Furthermore, by examining twin quantitative genetic associations and eFC-specific transcriptional profiles, we further provide systematic landscapes to enrich our understanding of the neurobiological underpinnings of such transdiagnostics into genetic micro-to-molecular architectures **(Fig. 1)**.

Results

Summary of the main analyses.

We recruited a nationwide sample from independent collectors during Nov, 2019 - Feb, 2022, with diverse ethnic groups and balanced socioeconomic conditions **(Fig. 2a, Supplemental Methods 1, Tab. S1-3)**. Main analyses in our present study were categorized into five modules. In the module 1, we capitalized on EBICglasso-derived Gaussian graph-theoretical model for estimating the hubs of symptom-centered connectome in calculating symptom-across sTDF scores **(Fig. 1a)**[5,40]. In the module 2, the edge-centric connectome was built for eligible 1,314 participants in neuroimaging analyses, with each participant for constructing a 12,248,775 x 12,248,775 neural connectome **(Fig. 1b)**. In the module 3, we developed the edge-centric connectome-based predictive model (eCPM)[41,42] to individually predict sTDF score into six relatively independent groups basing on the independent collector (s) or category **(Fig. 1c)**. Rather than individual prediction, we deployed inter-subject representation similarity (IRS) analysis model to probe into edge-centric patterns of sTDF changes in the module 4 **(Fig. 1d)**. In the module 5, we not only probed into the genetic heritability of eFC patterns from a twin dataset (n = 245, 127 Monozygotic twins), but also examined RS-enriched transcriptional

patterns by aligning into the AHBA (n = 6 postmortem brains, 10, 027 genes), with further decodes for their macroscale brain-derived associations and molecular enrichment into biological ontology, cell- and tissue-specific types, and neurodevelopmental periods **(Fig. 1e)**.

Symptom-centered hubs defined transdiganostic factor of anxiety and depression symptoms. We refined each symptom (item) that was described by both Zung's self-reported depression (SDS) and trait anxiety inventory (TAI) to build the Gaussian graph-theoretical network. We first found significant network-wise correlation between depressive symptoms and anxious symptoms (r = .40, p < .001, Mantel's test), statistically justifying to integrate them into a single transdiagnostic network **(Fig. 2b**, *Supplemental Methods 1-5, Fig. S1-6***)**. In this transdiagnostic network, the high topological centrality was found in the specific symptoms of "perceiving meaningless life", while the clusters of "feeling exhaustion" were captured to bridge core symptoms between depression and anxiety across multifarious topological properties **(Fig. 2c,** *Supplemental Methods 5, Tab. S4-5***)**. By estimating the normalized Shannon's entropy (SE, *Supplemental Methods 6*), the transdiagnostic hubs were found in the cluster of "perceiving meaningless life" and "feeling exhaustion" consisted of 12 symptoms (all SE > 0.8**, Fig. 2d and Tab. S6),** and the common loading integrating these symptoms were estimated as sTDF scores. Stability and statistical powers of estimating these topological properties from this network had been validated well (*Supplemental Methods 7, Fig. S7-10*). Collectively, beyond linear latent component in patients (e.g., p factor), we illuminated hubs of transdiagnostic symptoms of anxiety and depression by insights into the network-wise architecture. By doings so, we extended to provide this new sTDF scalar integrating to depict their topologically symptom-centered properties.

Edge-centric brain functional connectivity (eFC) could be reproducible and generalizable neural signatures of sTDF.

Given the strengths in predicting transpsychiatric conditions from connectome-based insights, we built upon the eCPM to individually predict sTDF from those edge-centric connectome by training support vector regression model, with each connectome for containing 12,248,775 putative eFCs from 4,950 "edge-centric nodes" that defined by Schaefer atlas (100 parcels) across whole brain per participant (*Methods*, *Supplemental Methods 8-10*). By using Fruchterman-Reingold algorithm, we found significantly high network centrality of visual network (VIS), sensorimotor network (SMN) and attention networks (VAN/DAN) in this connectome (normalized degree centrality > .6, p < .001) **(Fig. 3a-b**, *Tab. S7***).** These findings fit with existing studies[23-25,43], showing high validity of constructing these eFCs for neural features.

The eCPM showed statistically significant predictive roles of the eFCs to individual sTDF in the Main sample with 10-fold cross-validation, irrespective of training the regressor from positive eFCs (R^2 = .23, p_{perm} < .01), negative eFCs (R^2 = .26, p_{perm} < .01) or combined one (R^2 = .41, p_{perm} < .01) **(Fig. 3c)**. In the relatively independent Validation sample, we validated high reproducibility of this eCPM by all replicating these findings (p_{perm} < .01) **(Fig. 3d)**. Despite decreased model performance, the generalizability of this eCPM has partially manifested by showing the significantly predictive effects of this trained eCPM in the Generalization sample 1 (p_{perm} < .05) **(Fig. 3d)**. To validate the robustness of generalizability in heterogeneous cohorts,

we further examined prediction of this trained eCPM into additional independent samples. In the Generalization sample 2 encompassing over 30 locally ethical minorities in the China, we still found the significant predictive effects ($p_{perm} < .05$) **(Fig. 3d)**. Such generalizability has been found partly in another Generalization sample 3 that contained all participants in the main ethical group in the China (Han) ($p_{perm} < .05$) **(Fig. 3d)**. Given the potential confounding effects of the COVID-19 pandemic, we capitalized on this trained eCPM for generalizing in the Generalization sample 4 that recruited after COVID-19 pandemic in the China, and demonstrated the statistically significant generalizability ($p_{perm} < .05$) **(Fig. 3d)**. The specificity of this eFC has been also validated by showing optimum model performance for predicting sTPD compared to single-disorder symptoms **(Extended Data Fig. 1)**. On balance, these results highlight that the edge-centric connectome could be reproducible and generalizable brain signatures of transdiagnostic factor of anxiety and depression symptoms.

Edge-centric communication patterns of attention and frontoparietal network contributed to explain changes of such transdiagnostic factors.

To strengthen the neurobiological interpretability of this eCPM, we examined the contributive features of these eFC by extracting edge-centric within- and across-system communication patterns, respectively. We found no apparent outliers in the co-fluctuations across all the time points **(Extended Data Fig. 2a)**, enabling estimations of eFC-sTDF correlations without additional corrections. Based on the feature selections (**Supplemental Methods 10**), we then illustrated these inter-subject eFC-sTDF correlations with $p < .05$, showing uneven distributions into attention and frontoparietal networks **(Extended Data Fig. 2b-c)**. By integrating these eFCs into the brain systems that defined by Yeo-7 atlas, we found prominently high within-system weights and high communication density in the frontoparietal and attention networks (all $p_{perm} < .01$, FDR-corrected) **(Extended Data Fig. 2d, Tab. S8-11)**. In addition to such intra-connection, we demonstrated high between-system communications between ventral/dorsal attention networks and frontoparietal network by showing high normalized Shanno's SE (all $p_{perm} < .01$, FDR-corrected) **(Extended Data Fig. 2e-f)**.

Given the heterogeneous individual-between variants in the eCPM prediction, we furthered our analysis in the group-averaged multivariate representation similarity model (RSA) to sensitively decode brain-symptoms patterns. We found that both neural representation similarity matrix (nRDM) and behavioral RDM showed apparent individual-between variances on Euclidean distance across participants, implying the feasibility to decode neural representations of such brain-behavior changes (***Extended Data Fig. 3a-b***). We found statistically significant higher representation similarity (RS) in eFCs involving into frontal pole, superior frontal cortex, precuneus, and visual areas (all $p_{perm} < .05$, FDR-corrected), these regions that were predominantly assembled into frontoparietal, attention and default mode networks (***Extended Data Fig.3c****, Tab. S13*). Based on brain systems parceled by Yeo-7 atlas, we demonstrated higher RS within the attention networks and limbic networks (all $p_{perm} < .05$, FDR-corrected, ***Extended Data Fig. 3d***, Tab. S14-17). Beyond intra-system patterns, we observed significant RS in the eFCs relating to cross-system communications, such as edge-centric connection of default mode network to limbic network (all $p_{perm} < .05$, FDR-corrected, ***Extended Data Fig. 3e-f***). Thus, we may decode edge-centric communication patterns of such

transdiagnostic factor by showing brain eFC-derived disruptions in attention, limbic and frontoparietal systems.

Brain edge-centric patterns are moderately heritable.

We built upon the univariate quantitative ACE (A, addictive genetic factor; C, common environment; E, unique environment) model to disentangle phenotypic heritability of such eFC features in an independent twin dataset (*Supplemental Methods 11*). A moderate heritability of these eFC patterns (22.9%, 95% CI: 7.4 - 37.2) was found in the clusters of attention network at the optimal best-fitting AE model (*Fig. S11 and Tab. S18-19*). Supporting that, we observed a statistically significant correlation within monozygotic twins (ICC r = .22, p < .0001) for such eFC features, but not yet within the dizygotic ones (ICC r = .06, p = .28). All the statistics had been adjusted for correcting artifacts of sex, age and head-motion parameters. In total, beyond the (intermediate) phenotypic neural signatures, the edge-centric connectome may be a manifestation of outcomes of individual heritable variants, and it thus prompting the probes into molecular genetic associations of these eFC patterns associating to such transdiagnostic factor.

Edge-centric functional connectome of this transdiagnostic factor correlated to specific cortical gene expressions.

Given the heritability of such eFC signatures, we used the normative AHB atlas (http://human.brain-map.org) to delve into the genetically neurobiological understructure of such neural signatures of this transdiagnostic factor by capturing their transcriptional profiles (see *Methods*). By using representation similarity of eFCs into each region as neural phenotype **(Fig. 4a)**, we carried out partial least squares (PLS) regression model to estimate eFC-transcriptional alignments. Both first and second component (s) of PLS (PLS1, PLS2) were found to explain cumulative 32.4% variances in the spatial patterns of eFC from gene expressions, showing the anterior-posterior hierarchy **(Fig. 4a**, *Supplemental Results*). This results indicate significant connectome-transcriptional co-changes of above transdiagnostic symptoms.

These results were further supported by showing the significant correlation between neural signatures (i.e., representation similarity of eFC) and gene expression maps (i.e., PLS weighted scores) in both PLS1 and PLS2 (r_{PLS1} = .31, p_{perm} < .01; r_{PLS2} = .30, p_{perm} < .01) **(Fig. 4a)**. By examining the statistical significance of gene sets in PLS1 and PLS2, we further found 27 (or 231) genes overexpressed (or under-expressed) with increased (or decreased) cortical eFCs (PLS1+, Z > 3.0 or PLS1-, Z < 3.0, p < .005, **Fig. 4b**, *Tab. S20-21*) in the PLS1, as well observed similar associations in the PLS2 **(Fig. 4b**, *Tab. S22-23*). These results were reinforced by showing prominent univariate correlations between spatial gene expressions with top weights and neural phenotype in both PLS components (p < .05, FDR-corrected, **Fig. 4c-d**, *Tab. S24-25*). Collectively, the edge-centric connectome may be a promising candidate as neural endophenotype (or intermediate phenotype) of this transdiagnostic factor of anxiety and depression symptoms, which possibly revealed the genetically molecular mechanisms of vulnerability of such comorbidity.

Gene expression patterns of eFC-derived phenotype were correlated with macro-scale brain

network changes, cortical metabolism and risks of neuropsychiatric diseases.
For annotations of these transcriptional changes, we decoded macro-scale brain correlates of these gene sets (PLS1 and PLS2) depicting gene expression patterns of this transdiagnostic factor by using general linear models in the meta-analytic databases (http://dutchconnectomelab.nl/GAMBA/). These genes were found to spatially expressed across the whole brain (*Extended Data Fig. 4a*), with significant correlates of limbic/visual networks and even whole-brain functional connectivity ($p < .05$, FDR-corrected, *Extended Data Fig. 4b, f, Tab. S26-27*). Furthermore, we found trends of associations of these gene expressions to brain cognitive functions involving into visual and executive functions by being annotated from these cognitive ontology (*Extended Data Fig. 4c, Tab. S28-29*). Supporting that, results of cognitive meta-analytic decoding in the NeuroSynth dataset showed correlates of these gene sets to visual ability, negative affect and emotional performance (*Extended Data Fig. 4d, Tab. S30-31*). Although no significant correlates of cortical expansion and such gene expressions patterns were found (*Extended Data Fig. 4e*), the cortical metabolic rate of oxygen ($CMRO_2$) was observed to prominently correlate with such gene expressions ($p < .05$, FDR-corrected, *Extended Data Fig. 4f, Tab. S32*). In short, based on such decoding analyses, we found the macro-scale brain associations of these gene sets on this transdiagnostic factor, including visual/limbic systems and cortical metabolism to oxygen.

We further probed into the correlates of these gene sets to the risks of neurological and neuropsychiatric diseases in BrainMap datasets. Gene sets have been respectively assembled from the PLS1 (comprising PLS1+ and PLS1-) and PLS2 (comprising PLS2+ and PLS2-) components, showing specific spatial distributions of gene expression across whole brain (*Extended Data Fig. 5a*). PLS1 component was found to be significantly correlate with risks of autism spectrum disorder, Asperger, stroke and dyslexia ($p < .05$, FDR-corrected, *Extended Data Fig. 5b, Tab. S33*). These findings were partially validated by showing the significant correlations of PLS2 gene set to risks of these diseases as well (*Extended Data Fig. 5c, Tab. S34*). Thus, in addition to macro-scale brain associations, such connectome-transcriptional patterns for this transdiagnostic factor may imply risks of comorbidities in other specific neurological/neuropsychiatric diseases, such as ASD, Asperger and dyslexia (*Fig. S12-13*).

The eFC-derived transcriptional patterns were functionally enriched into specific biological pathways, tissue, cell types, diseases and neurodevelopmental periods.
To characterize associations of such transcriptional patterns on multiscale neurobiological processes, we used gene-expression specific enrichment analysis for decoding functional-specific, cell type-specific, disease-specific and neurodevelopment-specific distributions. In the PLS1 component, we firstly capitalized on the Metaspace (https://metascape.org/gp/index.html#/main/step1) platform that embedded with ChatGPT engine (https://openai.com/blog/chatgpt), to examine functional enrichment of these gene sets. We found statistically significant functional enrichment into the biological process (GO) of "regulation of protein kinase activity", "blood vessel development", "alcohol metabolic process" and "sex differentiation" (all $p < 5 \times 10^{-6}$, FDR-corrected) **(Fig. 5a,** *Tab. S35 and Fig. S14*), and illustrated the GO network as well protein-to-protein modules to show their regulations and interactions **(Fig. 5b-c**, *Tab. S36 and Fig. S15*). Full results in the PLS2 can be

found in the **Supplemental Results** (*Tab. S37, Fig. S16-17*).

Given the significant findings in the GO functional enrichment, we moved forward our analyses to probe into this enrichment from tissue-specific, cell type-specific, disease-specific and neurodevelopment-specific schemes. By combining Metascape and SEA (Specific Expression Analysis, http://genetics.wustl.edu/jdlab/csea-tool-2/), the PLS1 was found to be significantly enriched across body tissues, such as brain, heart, muscle, bone marrow ($p < .05$, FDR-corrected, **Fig. 6a, Tab. S38**). We demonstrated prominent cell type-specific enrichment as well, such as smooth muscle and Manno midbrain cells ($p < .05$, FDR-corrected, **Fig. 6b, Tab. S39**). These significantly disease-specific enrichment had also been captured, such as "Down Syndrome" and "Fatigue" ($p < .05$, FDR-corrected, **Fig. 6b, Tab. S40**). In conjunction with BrainSpan atlas (http://www.brainspan.org/), we finally found the specific enrichment of these gene sets linking to sTDF in the cerebellum, especially in the sensitively developmental periods from late-mid childhood to young adulthood **(Fig. 6c)**. Despite failure to reach statistical significance level, this eFC-derived gene set was found to be enriched in cerebellar and cortical regions than other ones at adulthood (specificity index probability, pSI < .001) **(Fig. 6d)**. Full results for the PLS2 can be found in the **Supplemental Results** (*Tab. S41-43*). Together, these results indicate the overarching neurobiological enrichment for the transdiagnostic component of anxiety and depression, linking to eFC-derived brain signatures from phenotypic changes to molecular protein and blood-vessel development as well childhood-to-adulthood cerebellar neurodevelopment.

Discussion

The present study investigates the neurobiological underpinnings of the symptom-centered transdiagnostic factor of anxiety and depression symptoms, by capturing its brain macro-micro-molecular signatures in a non-WEIRD nationwide preclinical cohort. We mainly found that external life meaning perception and internal emotional exhaustion emerge as the hubs of bridging transdiagnostic symptoms of anxiety and depression, and support this new scalar to represent its transdiagnostic factor from symptom-across topological architecture. We further show that the edge-centric communication patterns could be reproducible and generalizable neural signatures predicting this transdiagnostic factor. To improve interpretability of such connectome-based neural phenotype, brain-symptom representation similarities are probed, and found the similarity of intra-communications of frontoparietal and attention networks. Moreover, results derived from Twin-ACE model demonstrated the moderate heritable for these connectome-based eFC signatures. By examining the transcriptional correlates of such representation similarity, we extended our knowledge of neurobiological signatures of this transdiagnostic factor from macroscale brain connectome to genetically micro-molecular architectures, showing specific enrichment into vessel systems, tissues of brain/heart, Manno midbrain cells and childhood-to-adulthood cerebellar neurodevelopments. Together, our findings shed lights on the neurobiological underpinnings of the preclinical transdiagnostic factors of anxiety and depression symptoms by clarifying these genetically molecular-micro-macro brain signatures.

Although widely studied, probing into the network-wise architectures of pathopsychological

connectome of depression and anxiety was still one of the most promising pathway to understand this transdiagnostic factor[4,44]. In the present study, we found the integrative high centrality of external life meaning perception (e.g., “perceiving meaningless life”) and internal emotional exhaustion (e.g., “feeling exhaustion”) in this symptom-centered network in a preclinical adult population, which merited to characterize early vulnerable symptoms for “at-risk or preclinical” ones[8,45-49]. Thus, these findings complement previous studies by revealing preclinical or “at-risk” trans-symptoms to the anxiety and depression. Importantly, this network analysis conceptualized on a new scalar called “sTDF”, which may forward over theoretical framework of the “p factor” of transdiagnostic psychiatry. Though the “p factor” offered a promising framework to describe common liability or vulnerability across different psychiatric conditions, such “single” dimension may oversimplify between-symptom interactions as critiqued recently[5,50]. In this vein, our findings benefit to partly address this concern by going beyond simplistic linear common associations from this symptom-centered network-wise model[51].

Based on the sTDF describing general transdiagnostic symptom-centered factor of anxiety and depression, we have demonstrated its brain edge-centric connectome-based signatures, with well reproducibility and generalizability. Rather to regional FC, the edge-centric connectome reflects whether the communication patterns of pairs of FC were linked, with mathematical assumptions of such communication for correlations of element-wise co-fluctuations across time series[52,53]. Compared to the nodal FC, it has been broadly manifested that the eFC outperformed in subject-specific identifiability and predictive robustness by depicting community-wise architecture[24,52]. That is, our findings are reinforced by existing evidence to indicate the robustly predictive brain signatures of such symptom-centered phenotype, possibly gaining promising biomarkers of this transdiagnostic factor. One important point to promote this conclusion from previous studies was to replicate and generalize this predictive model originally developed in other independent samples. A robust body of evidence stressed the “generalizability challenges” in tremendously increasing neuroimaging-based machine-learning models[54,55]. As shown in our study, the model performance was overestimated by showing prominently higher predictive accuracy in the internal sample than external samples. Thus, our current eCPM partly addressed this issue by generalizing it into other independent samples involving into ethnic biases and COVID-related confounding factors, highly strengthening the pre-clinical practicability in other broader populations [56,57]. Collectively, our findings derived from this predictive model may imply that the edge-centric connectome could be reproducible and generalizable neural signatures for this transdiagnostic factor of anxiety and depression for preclinical populations.

By probing into the quantitative twin neuroimaging associations and genetically transcriptional patterns of such neural phenotype, our study has bridged a link of brain signatures of this transdiagnostic factor with the underlying gene expression mechanisms. Rather to integrate transdiagnostic symptoms, many studies have demonstrated the genetic associations that depicted by transcriptional profiles for the overlapping of neural anomalies across psychiatric disorders, which was conceptualized as “transdiagnostic neural biomarkers”[58-60]. Supporting that, Xie et al (2023) illuminated transdiagnostic neural substrates

underlying the transpsychiatric conditions by establishing the neuropathopsychological (NP) factor as the endophenotype, showing prominently inheritable characteristics[61]. Furthermore, the polygenic "p factor" has been revealed by showing high risks in heritability (i.e., 57%) across major psychiatric disorders[62,63]. Thus, our findings were supported by these evidences to substantiate genetically neurobiological substrates of edge-centric connectome in predicting transdiagnostic symptom-centered phenotype in the preclinical population[64-66], which may partly indicate the increased genetic/neural risks of "early preclinical symptoms" to "diagnosed disorder". In other words, this finding provided secondary evidence to support the interplay of genetic risks on this preclinical transdiagnostic factor across heterogeneous psychiatric conditions (beyond a commonly general factor) by shaping cortical activities[38,67-69]. One point in this enrichment analysis worthy to be highlighted was to demonstrate the enriched neurodevelopment periods from late-mid childhood to young adulthood for neural phenotype of transdiagnostic factor, especially in cortical and cerebellar regions. There was a consistent conclusion to demonstrate the increasing trends for emotional problems (e.g., depressive symptoms) from childhood to adulthood, but had not yet reached a conclusive interpretation[70-72]. Our findings thus provided indirect evidence to encapsulate such trajectory as results of cortical and cerebellar underdevelopments. From what has been mentioned, the present study may enrich our understandings of multiscale neurobiological underpinnings of transdiagnostic symptom-centered factor across anxiety and depression.

Despite providing multiscale insights, several limitations should be warranted. We recruited a non-WEIRD healthy adult sample to ensure the sociodemographic diversity, but not to consider the contrasts to ones in clinical conditions. Relating to this concern, we didn't included measurements for more internalizing symptoms, thus limiting the "transdiagnostic factor" into anxiety and depression only. Given the cross-sectional design, the present study gained abundant findings from correlative evidences. Thus, we recommended high-quality experimental study to provide robustly causal evidence strengthening the validity of our conclusions. Last aspect of limitations was the moderate strength of evidence. While the pioneering evidence has established an seemingly far-reaching associations between brain macro-micro-molecular signatures and symptoms of anxiety/depressions, specific aspects still require further consolidation through more direct investigations.

In conclusion, we originally established a new network-wise symptom-centered transdiagnostic factor of these anxiety and depression (sTDF) in the preclinical populations. Based on this scalar, we further build upon the machine-learning model to capture its reproducible and generalizable brain edge-centric signatures involving into attention and frontoparital systems. By using twin dataset, AHBA and neurophysiological datasets, the eFC-derived connectome-transcriptional landscapes were disentangled, especially in multifarious biologically functional enrichment associating to vessel systems and childhood-to-adulthood cerebellar neurodevelopments. Thus, these findings paved the pathways to understand neurobiologically plausible understructure of preclinical transdiagnostics of anxiety and depression symptoms by overarching multiscale insights into the genetic macro-micro-molecular framework.

Methods

Participants and neuroimaging data acquisition

We included a representative non-WEIRD nationwide sample consisted of 2, 022 healthy adults ***(Tab. 1 and Supplemental Methods 1)****. Given the COVID-19-derived neurobiological changes, neuroimaging data had been acquired preceding to pandemic, or such data had been collected for participants who were currently free from COVID-19 infections. As behavioral measurements, the Zung's self-report depression scale (SDS) and status-trait anxiety inventory (STAI) were used to describe symptoms* ***(Supplemental Methods 2)****. Data collection protocol and these ensuing analyses had been officially approved by the Institutional Review Board (IRB) of Southwest University. Data acquisition and preprocessing for neuroimaging of these participants were all in line with our previously standardized pipelines to this dataset[73,74]* ***(Supplemental Methods 3)****.*

Gaussian graph-theoretical model and transdiagnostic factor

To capture the hubs of transdiagnostic factor of anxiety and depression, we carried on the EBICglasso Gaussian graph-theoretical model. As guided by didactic framework, the extended Bayesian information criterion glasso (EBICglasso) algorithm was used for regularization of this network[75] with each item as node and with partial correlations of pairs of them as edge ***(Supplemental Methods 4-5)****. Furthermore, we estimated topological hubs of this connectome by using 10 topologically nodal and bridging centrality, with high values of centrality for detecting hubs* ***(Supplemental Methods 6)****.To integrate these 10 multifarious topological properties (i.e., centrality), we used the normalized Shannon's entropy (SE) that described the extent to which this node (item) appeared relatively higher comprehensive centrality values than of others across all the topological properties of centrality* ***(Supplemental Methods 7-8)****. Finally, we calculated the principal common loading scores among these hubs as transdiagnostic factor of this connectome of anxiety and depression symptoms (sTDF) [5].*

Edge-centric functional connectome-based predictive model (eCPM)

Edge-centric brain functional connectome (eFC). *The Schaefer-100 atlas was used to parcel cortical areas into 100 regions, and time series of each region were extracted to be z-scored firstly (i.e., 100 parcels × 240 points). Then, we obtained "edge time series" by estimating the dot products of these time series, thus gaining 4,950 "edge time series". The edge-to-edge connectome was finally built upon for a 12,248,775 x 12,248,775 eFC matrix by correlating each pair of these 4,950 "edge time series" for each participant to be neural features* ***(Supplemental Methods 9)****.*

The eFC conncetome-based predictive model. *In line with the original CPM, we estimated the inter-subject correlations of eFCs to sTDF scores, and retained eFC that reached statistical significance ($p < .05$, uncorrected) as thresholding masks, with separation to positive-correlated and negative-correlated mask. In this vein, individual eFC feature was produced by summing r values of eFC-sTDF correlations in each mask. By using these features, we established the machine-learning model with support vector algorithm to predict sTDF in these independent samples by using MATLAB (MathWorks Inc.)****(Supplemental Methods 10)****.*

Inter-subject representation similarity analysis (IS-RSA)

We deployed the inter-subject representation similarity analysis to capture the multivariate similar patterns of eFC-sTDF correlations, favoring to interpret these ultra-high-volume data. Firstly, each "edge-centric" node had been vectored a 1 × 4,949 "node-specific pattern" by describing all the eFCs relating to this given node. Then, the inter-subject correlations (i.e., r values) of each "node-specific pattern" were calculated, and these 1-r values were used to generated neural representation dissimilarity matrix (RDM). Further, we built upon the behavioral RDM by estimating Euclidean distance of sTDF scores across all the pairs of participant. We vectored all the neural RDMs and one behavioral RDM by using the upper triangular matrix, and iterated to correlate each neural RDM to this behavioral RDM by using Spearman's correlation. Finally, each correlation reflected nodal eFC-sTDF representation similarity (RS), with positive (negative) r value for RS (RDS). Statistical significance for these r values was set to p < .05 with FDR correction.

Edge-centric connectome-transcriptional signatures

Quantitative twin study analysis. *The full model with ACE framework has been established to clarify heritability of eFC features by decomposing variances of addictive genetic effects (A) from latent factor model for 127 monozygotic twins and 118 pairs of dizygotic twins (Beijing Twin Study Dataset,* ***Supplemental Methods 11****). Model performances were further compared to these nested submodels dropping out latent factor (s) (e.g., AE or E), in order to determine the optimal model. Quantitative heritability was finally estimated from this optimal model once the statistical significance of* ΔX^2 *was no longer less than 0.50 (****Supplemental Methods 11****).*

The eFC-derived gene expression patterns to transdiagnostic factor. *Preprocessing of AHBA dataset was all in line with standardized pipeline, and the resultant gene-brain matrix (10,027 genes × 100 parcels) was generated by aligning these gene expression levels into brain spatial map that defined by Schaefer-100 atlas. In addition, the eFC-sTDF RS vector (1 RS × 100 parcels) representing edge-centric neural phenotype of this transdiagnostic factor had been prepared. To capture edge-centric connectome-transcriptional signatures, we carried on the partial least square (PLS) regression model by fitting gene-brain matrix (10,027 × 100, independent variable) into this RS vector (1 × 100, dependent variable)****(Supplemental Methods 12-13)****. The permutation test (at n = 5,000) was used to estimate the statistical significance for each component of this PLS model. Further, the bootstrapping method with n = 5,000 was deployed estimating weights and corresponding statistics (Z values) for these genes. To balance both Type-I and Type-II errors, the statistical threshold was set to Z > 3 (PLS+) or Z < -3 (PLS-).*

Gene sets decoding to brain networks and risks of neurological and psychiatric diseases. *We capitalized on Gene Annotation by Macroscale Brain-imaging Association (GAMBA) dataset for associating brain functions and risks of neurological/psychiatric diseases. This GAMBA dataset integrated resources to link brain changes to gene sets that users provided by online meta-analytic linear regression model, such as macroscale networks, brain cognitive ontology, cognitive annotations, cortical expansion and metabolism* ***(Supplemental Methods 13)****. The statistical significance for each model was corrected by Bonferrioni-Holm algorithm at p < .05. Based on gene sets that we found above, we separately decoded these brain*

correlates by first and second PLS components, with positive weights or reverse ones. To examine the specificity of these gene-neuroimaging associations, we utilized on the Permutation tests with multiple null distributions ***(Supplemental Methods 14)*** [76]*. Here, gene sets had been realigned into orthogonal PLS components to strengthen neuropahtologically plausible interpretability.*

Enrichment analysis. *We deployed the "Metascape" with newest version and "SEA" datasets to delineate functional processes that were enriched from above gene sets (PLS1 and PLS2). The gene set was used as input for this platform, and was further annotated by multiple biological databases* ***(Supplemental Methods 15)****. To show the gene enrichment specificity, we estimating the specificity index probability (pSI) to quantify how this given gene set could enrich higher into specific tissues compared to other ones from different thresholds to include background enrichment term. The statistical significance of such enrichment for this given gene set was estimated, with Benjamini-Hochberg FDR corrections.*

533 **Reference**

534 1 Substance Abuse and Mental Health Services Administration. *Key substance use and mental*
535 *health indicators in the United States: Results from the 2017 National Survey on Drug Use and*
536 *Health (HHS Publication No. SMA 18-5068, NSDUH Series H-53)* (Center for Behavioral Health
537 Statistics and Quality, Substance Abuse and Mental Health Services Administration, 2018).
538 2 Kalin, N. H. The Critical Relationship Between Anxiety and Depression. *Am J Psychiatry* **177**,
539 365-367 (2020).
540 3 Clark, D. A. Cognitive behavioral therapy for anxiety and depression: possibilities and
541 limitations of a transdiagnostic perspective. *Cogn Behav Ther* **38 Suppl 1**, 29-34 (2009).
542 4 Fusar-Poli, P. *et al.* Transdiagnostic psychiatry: a systematic review. *World psychiatry* **18**, 192-
543 207 (2019).
544 5 Caspi, A. *et al.* The p Factor: One General Psychopathology Factor in the Structure of Psychiatric
545 Disorders? *Clin Psychol Sci* **2**, 119-137 (2014).
546 6 Gosmann, N. P. *et al.* Selective serotonin reuptake inhibitors, and serotonin and
547 norepinephrine reuptake inhibitors for anxiety, obsessive-compulsive, and stress disorders: A
548 3-level network meta-analysis. *PLoS med* **18**, e1003664 (2021).
549 7 Newby, J. M., McKinnon, A., Kuyken, W., Gilbody, S. & Dalgleish, T. Systematic review and meta-
550 analysis of transdiagnostic psychological treatments for anxiety and depressive disorders in
551 adulthood. *Clin Psychol Rev* **40**, 91-110 (2015).
552 8 Kaiser, T., Herzog, P., Voderholzer, U. & Brakemeier, E. L. Unraveling the comorbidity of
553 depression and anxiety in a large inpatient sample: Network analysis to examine bridge
554 symptoms. *Depress Anxiety* **38**, 307-317 (2021).
555 9 Groen, R. N. *et al.* Comorbidity between depression and anxiety: assessing the role of bridge
556 mental states in dynamic psychological networks. *BMC med* **18**, 308 (2020).
557 10 Ter Meulen, W. G. *et al.* Depressive and anxiety disorders in concert-A synthesis of findings on
558 comorbidity in the NESDA study. *J Affect Disord* **284**, 85-97 (2021).
559 11 Scott, J. *et al.* Clinical staging in psychiatry: a cross-cutting model of diagnosis with heuristic
560 and practical value. *Br J Psychiatry* **202**, 243-245 (2013).
561 12 Henry, C. & Scott, J. Clinical staging models: From general medicine to mental disorders.
562 *BJPsych Advances* **23**, 292-299 (2017).
563 13 Cuthbert, B. N. Research Domain Criteria (RDoC): Progress and Potential. *Curr Dir Psychol Sci*
564 **31**, 107-114 (2022).
565 14 Conradt, E., Crowell, S. E. & Cicchetti, D. Using Development and Psychopathology Principles
566 to Inform the Research Domain Criteria (RDoC) Framework. *Dev Psychopathol* **33**, 1521-1525
567 (2021).
568 15 Bassett, D. S., Zurn, P. & Gold, J. I. On the nature and use of models in network neuroscience.
569 *Nat Rev Neurosci* **19**, 566-578 (2018).
570 16 Bassett, D. S., Xia, C. H. & Satterthwaite, T. D. Understanding the Emergence of
571 Neuropsychiatric Disorders With Network Neuroscience. *Biol Psychiatry Cogn Neurosci*
572 *Neuroimaging* **3**, 742-753 (2018).
573 17 Kelly, C., Biswal, B. B., Craddock, R. C., Castellanos, F. X. & Milham, M. P. Characterizing variation
574 in the functional connectome: promise and pitfalls. *Trends Cogn Sci* **16**, 181-188 (2012).
575 18 Finn, E. S. *et al.* Functional connectome fingerprinting: identifying individuals using patterns of
576 brain connectivity. *Nat Neurosci* **18**, 1664-1671 (2015).

19 de Lange, S. C. *et al.* Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nat Hum Behav* **3**, 988-998 (2019).

20 Van Essen, D. C. & Barch, D. M. The human connectome in health and psychopathology. *World psychiatry* **14**, 154-157 (2015).

21 Kaufmann, T. *et al.* Delayed stabilization and individualization in connectome development are related to psychiatric disorders. *Nat Neurosci* **20**, 513-515 (2017).

22 Yang, Z. *et al.* Understanding complex functional wiring patterns in major depressive disorder through brain functional connectome. *Transl Psychiatry* **11**, 526 (2021).

23 Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O. & Betzel, R. F. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nat Neurosci* **23**, 1644-1654 (2020).

24 Jo, Y., Faskowitz, J., Esfahlani, F. Z., Sporns, O. & Betzel, R. F. Subject identification using edge-centric functional connectivity. *NeuroImage* **238**, 118204 (2021).

25 Rodriguez, R. X., Noble, S., Tejavibulya, L. & Scheinost, D. Leveraging edge-centric networks complements existing network-level inference for functional connectomes. *NeuroImage* **264**, 119742 (2022).

26 Faskowitz, J., Betzel, R. F. & Sporns, O. Edges in brain networks: Contributions to models of structure and function. *Netw Neurosci* **6**, 1-28 (2022).

27 Jo, Y. *et al.* The diversity and multiplexity of edge communities within and between brain systems. *Cell Rep* **37**, 110032 (2021).

28 Yang, B. *et al.* Edge-centric functional network analyses reveal disrupted network configuration in autism spectrum disorder. *Affect Disord* **336**, 74-80 (2023).

29 Idesis, S. *et al.* Edge-centric analysis of stroke patients: An alternative approach for biomarkers of lesion recovery. *Neuroimage Clin* **35**, 103055 (2022).

30 Zamani Esfahlani, F. *et al.* Edge-centric analysis of time-varying functional brain networks with applications in autism spectrum disorder. *NeuroImage* **263**, 119591 (2022).

31 Moreau, C. A. *et al.* Brain functional connectivity mirrors genetic pleiotropy in psychiatric conditions. *Brain* **146**, 1686-1696 (2023).

32 Taquet, M. *et al.* A structural brain network of genetic vulnerability to psychiatric illness. *Mol Psychiatry* **26**, 2089-2100 (2021).

33 Romme, I. A., de Reus, M. A., Ophoff, R. A., Kahn, R. S. & van den Heuvel, M. P. Connectome Disconnectivity and Cortical Gene Expression in Patients With Schizophrenia. *Biol Psychiatry* **81**, 495-502 (2017).

34 Xia, M. *et al.* Connectome gradient dysfunction in major depression and its association with gene expression profiles and treatment outcomes. *Mol Psychiatry* **27**, 1384-1393 (2022).

35 Fornito, A., Arnatkevičiūtė, A. & Fulcher, B. D. Bridging the Gap between Connectome and Transcriptome. *Trends Cogn Sci* **23**, 34-50 (2019).

36 Arnatkeviciute, A., Fulcher, B. D., Bellgrove, M. A. & Fornito, A. Imaging Transcriptomics of Brain Disorders. *Biol Psychiatry Glob Open Sci* **2**, 319-331 (2022).

37 Arnatkeviciute, A. *et al.* Genetic influences on hub connectivity of the human connectome. *Nat Commun* **12**, 4237 (2021).

38 Li, J. *et al.* Cortical structural differences in major depressive disorder correlate with cell type-specific transcriptional signatures. *Nat Commun* **12**, 1647 (2021).

39 Hettema, J. M. What is the genetic relationship between anxiety and depression? *Am J Med*

Genet Part C Semin Med Genet **148C**, 140–147 (2008).

40 Fried, E. I., Greene, A. L. & Eaton, N. R. The p factor is the sum of its parts, for now. *World psychiatry* **20**, 69-70 (2021).

41 Shen, X. *et al.* Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc* **12**, 506-518 (2017).

42 Yarkoni, T. & Westfall, J. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect Psychol Sci* **12**, 1100-1122 (2017).

43 Wang, W. *et al.* Edge-centric functional network reveals new spatiotemporal biomarkers of early mild cognitive impairment. *Brain-X* **1**, e35 (2023).

44 Contreras, A., Nieto, I., Valiente, C., Espinosa, R. & Vazquez, C. The study of psychopathology from the network analysis perspective: A systematic review. *Psychother Psychosom* **88**, 71-83 (2019).

45 Beard, C. *et al.* Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychol Med* **46**, 3359-3369 (2016).

46 Grassie, H. L., Kennedy, S. M., Halliday, E. R., Bainter, S. A. & Ehrenreich-May, J. Symptom-level networks of youth- and parent-reported depression and anxiety in a transdiagnostic clinical sample. *Depress Anxiety* **39**, 211-219 (2022).

47 Abdul Karim, M., Ouanes, S., Reagu, S. M. & Alabdulla, M. Network analysis of anxiety and depressive symptoms among quarantined individuals: cross-sectional study. *BJPsych open* **7**, e222 (2021).

48 Jin, Y. *et al.* Network analysis of comorbid depression and anxiety and their associations with quality of life among clinicians in public hospitals during the late stage of the COVID-19 pandemic in China. *J Affect Disord* **314**, 193-200 (2022).

49 McElroy, E., Fearon, P., Belsky, J., Fonagy, P. & Patalay, P. Networks of Depression and Anxiety Symptoms Across Development. *J Am Acad Child Adolesc Psychiatry* **57**, 964-973 (2018).

50 Krueger, R. F. & Eaton, N. R. Transdiagnostic factors of mental disorders. *World psychiatry* **14**, 27-29 (2015).

51 McNally, R. J. Network Analysis of Psychopathology: Controversies and Challenges. *Annu Rev Clin Psychol* **17**, 31-53 (2021).

52 de Reus, M. A., Saenger, V. M., Kahn, R. S. & van den Heuvel, M. P. An edge-centric perspective on the human connectome: link communities in the brain. *Philos Trans R Soc Lond B Biol Sci* **369** (2014).

53 Novelli, L. & Razi, A. A mathematical perspective on edge-centric brain functional connectivity. *Nat Commun* **13**, 2693 (2022).

54 Chen, Z. *et al.* Sampling inequalities affect generalization of neuroimaging-based diagnostic classifiers in psychiatry. *BMC Med* **21**, 241 (2023).

55 Chen, Z. *et al.* Evaluation of Risk of Bias in Neuroimaging-Based Artificial Intelligence Models for Psychiatric Diagnosis: A Systematic Review. *JAMA Netw Open* **6**, e231671 (2023).

56 Mihalik, A. *et al.* Multiple Holdouts With Stability: Improving the Generalizability of Machine Learning Analyses of Brain-Behavior Relationships. *Biol Psychiatry* **87**, 368-376 2020).

57 Dhamala, E., Yeo, B. T. T. & Holmes, A. J. One Size Does Not Fit All: Methodological Considerations for Brain-Based Predictive Modeling in Psychiatry. *Biol Psychiatry* **93**, 717-728 (2023).

58 Ji, G.-J. *et al.* White matter dysfunction in psychiatric disorders is associated with

neurotransmitter and genetic profiles. *Nat Ment Health* **1**, 655-666 (2023).

59 Ardesch, D. J., Libedinsky, I., Scholtens, L. H., Wei, Y. & van den Heuvel, M. P. Convergence of Brain Transcriptomic and Neuroimaging Patterns in Schizophrenia, Bipolar Disorder, Autism Spectrum Disorder, and Major Depressive Disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging* **8**, 630-639 (2023).

60 Hettwer, M. D. *et al.* Coordinated cortical thickness alterations across six neurodevelopmental and psychiatric disorders. *Nat Commun* **13**, 6851 (2022).

61 Xie, C. *et al.* A shared neural basis underlying psychiatric comorbidity. *Nature med* **29**, 1232-1242 (2023).

62 Selzam, S., Coleman, J. R. I., Caspi, A., Moffitt, T. E. & Plomin, R. A polygenic p factor for major psychiatric disorders. *Transl Psychiatry* **8**, 205 (2018).

63 Allegrini, A. G. *et al.* The p factor: genetic analyses support a general dimension of psychopathology in childhood and adolescence. *J Child Psychol Psychiatry* **61**, 30-39 (2020).

64 Buch, A. M. & Liston, C. Dissecting diagnostic heterogeneity in depression by integrating neuroimaging and genetics. *Neuropsychopharmacology* **46**, 156-175 (2021).

65 Darby, R. R., Joutsa, J. & Fox, M. D. Network localization of heterogeneous neuroimaging findings. *Brain* **142**, 70-79 (2019).

66 Verdi, S., Marquand, A. F., Schott, J. M. & Cole, J. H. Beyond the average patient: how neuroimaging models can address heterogeneity in dementia. *Brain* **144**, 2946-2953 (2021).

67 Diez, I. *et al.* Early-life trauma endophenotypes and brain circuit-gene expression relationships in functional neurological (conversion) disorder. *Mol Psychiatry* **26**, 3817-3828 (2021).

68 Berg, L. M. *et al.* The neuroanatomical substrates of autism and ADHD and their link to putative genomic underpinnings. *Mol Autism* **14**, 36 (2023).

69 Xie, Y. *et al.* Brain mRNA Expression Associated with Cortical Volume Alterations in Autism Spectrum Disorder. *Cell Rep* **32**, 108137 (2020).

70 Uljarević, M. *et al.* Anxiety and Depression from Adolescence to Old Age in Autism Spectrum Disorder. *J Autism Dev Disord* **50**, 3155-3165 (2020).

71 Hankin, B. L. *et al.* Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *J Abnorm Child Psychol* **107**, 128-140 (1998).

72 Hankin, B. L. & Abela, J. R. Development of psychopathology: A vulnerability-stress perspective. *Depression from childhood through adolescence and adulthood* 245-288 (2005).

73 Zhang, R., Chen, Z. & Feng, T. The triple psychological and neural bases underlying procrastination: Evidence based on a two-year longitudinal study. *NeuroImage* **283**, 120443 (2023).

74 He, L. *et al.* Functional Connectome Prediction of Anxiety Related to the COVID-19 Pandemic. *Am J Psychiatry* **178**, 530-540 (2021).

75 Borsboom, D. & Cramer, A. O. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol* **9**, 91-121 (2013).

76 Arnatkeviciute, A., Markello, R. D., Fulcher, B. D., Misic, B. & Fornito, A. Toward Best Practices for Imaging Transcriptomics of the Human Brain. *Biol Psychiatry* **93**, 391-404 (2023).

Captions

Fig. 1 Technical and Research Workflow. The main steps to conduct this study included five module. **a,** Module 1, we used each item as node across depression and anxiety scales, and estimate correlations of all the pairs of them to be edge for establishing connectome. The transdiagnostic factor of anxiety and depression symptoms (sTDF) scores were estimated by the common factor loading of symptoms with integrative high centrality in these connectome. **b,** Module 2, The edge-centric functional connectivity (eFC) was estimated by correlating each pair of "edge time series" (240 points) derived from co-fluctuations from Schaefer-100 atlas. **c,** Module 3, we established edge-centric predictive connectome model by using eFC to predict mFC from support vector regression algorithm, and validated this model in several independent samples. **d,** Module 4, the inter-subject representation similarity analysis (IS-RSA) was conducted to decode sTDF-eFC patterns by correlating representation dissimilarity matrix (RDM) of sTDF scores to eFC. **g,** Module 5, we used the sTDF-eFC representation similarity as neural phenotype to align AHBA with 10, 027 gene candidates for capturing neuroimaging-transcriptional associations. The gene-expression macroscale brain decoding and gene enrichment analyses were conducted to annotate micro-cellular functions.

Fig 2. Sociodemograprahic Characteristics and Gaussian Graphic Model of Symptom-Centered Network. a, The geospatial and socioeconomic statistics of this nationwide sample (GGBBP sample recruited from 2019 to 2022). The scale indicated the number of included subjects after log transformation. **b,** Mantel's test plot was illustrated here, and each point into the lower triangle indicated the mean values of corresponding items. **c,** We illustrated the centrality of each symptom (item) from network model by descending order, with the "D" for indicating "depressive symptom" and with the number of this label for indicating the item in this scale. **d,** This showed density with Gaussian kernel function for each symptom by descending order, with each circuit (gray) for indicating the high integrative centrality.

Fig 3. The eFC Line-Graph Connectome and Model Performance of eCPM. a, We used the Gephi (https://gephi.org/) software to visualize edge-centric connectome, with 4,950 nodes and 24,502,500 eFCs. To ensure readability, this connectome density has been threshold to 0.1, and was adjusted by using Fruchterman-Reingold layout. **b,** It showed the edge-centric connectome and brain systems parceled by Yeo-7 network atlas. **c,** To show the model performance, we provided scatter plots for the correlation between true sTDF scores and predicted ones (z-scored) within the sample. The Taylor diagram was drawn to comprehensively evaluate model performance. d, We further figured edge-centric connectome, along with scatter plots and the Taylor diagram to show the external generalizability in validation sample and four generalizability samples, respectively.

Fig 4. Transcriptional Profiles of eFC-derived Representation Similarity to Transdiagnostic Factor. a, We used used partial least squares (PLS) regression model to predict eFC-derived representation similarity by aligning AHBA normative data into Schaefer-100 space (upper panel), and showed the weights for first and second components (PLS1 and PLS2, bottom panel). Further, the scatter plot was provided to show the linear association of PLS scores (weights) to RS. **b,** The colored table detailed the gene expression patterns for PLS1 (upper

panel) and PLS2 (bottom panel), with threshold for the Z value > (<) 3.0. The bar plots in the right panel indicated proportion of the number of genes reaching this statistical threshold from all the candidates. By using this statistical boundary, 27 genes (142 genes) survived from 4772 gene (5823 gene) sets in PLS1, and 44 genes (112 genes) survived from 3994 genes (5989 genes) in PLS2. **c,** We extracted gene expression level for these selected genes from PLS components, and illustrated scatter plots for each PLS component that showing the largest correlation strengths between this given gene and RS coefficient (z-scored). **d,** The univariate correlations for expression levels of all the genes and RS coefficient (z-scored) were calculated, and were presented in this chart with descending order.

Fig 5. Enrichment of Biological Processes/Pathways and Protein-to-Protein Interaction. a, By using the Metaspace tool (amplified by ChatGPT), we presented the top 20 biological processes/pathways that were enriched from the PLS1 gene set at $q < 0.01$ after Benjamini-Hochberg FDR corrections, with the cumulative hypergeometric distribution for estimating corresponding p values. **b,** Circos plot was illustrated by visualizing the term-to-term connectivity, with edges for showing between-term similarity > 0.3. This plot was generated by Cytoscape embodied into the Metascape tool. c, We provided protein-to-protein interaction connectome in this chart, and recolored these proteins that enriched from this gene list by independent modules detected from the Molecular Complex Detection (MCODE) algorithm. Details for each MCODE can be found in the Supplementary Materials.

Fig 6. Specific Enrichment of such sTDF-eFC Gene Set (PLS1). a, We showed the tissue-specific enrichment of this gene set (PLS1) by using both Metascape tool and Specific Expression Analysis (SEA) database. * indicated the $p < .05$ (Benjamin-Hochberg FDR) that found in the Metascape database in the left panel, while the colors of circles indicated the q values (Benjamin-Hochberg FDR) in the right panel. The size of these bullseye plots represented the proportion of the numbers of genes on specific tissues at a given specificity index probability (pSI), which evaluated the levels of enrichment specificity of given genes compared to other ones, with permutation tests. **b,** It had been showed for cell type-specific enrichment (left panel) and disease-specific enrichment (right panel) of this gene set. **c,** Bullseye plots, along with q values, have been illustrated to show the enrichment into the neurodevelopmental periods at different brain areas. **d,** Bullseye plots to show the enrichment of SEA brain regions have been provided though no one reached the statistical significance.

783

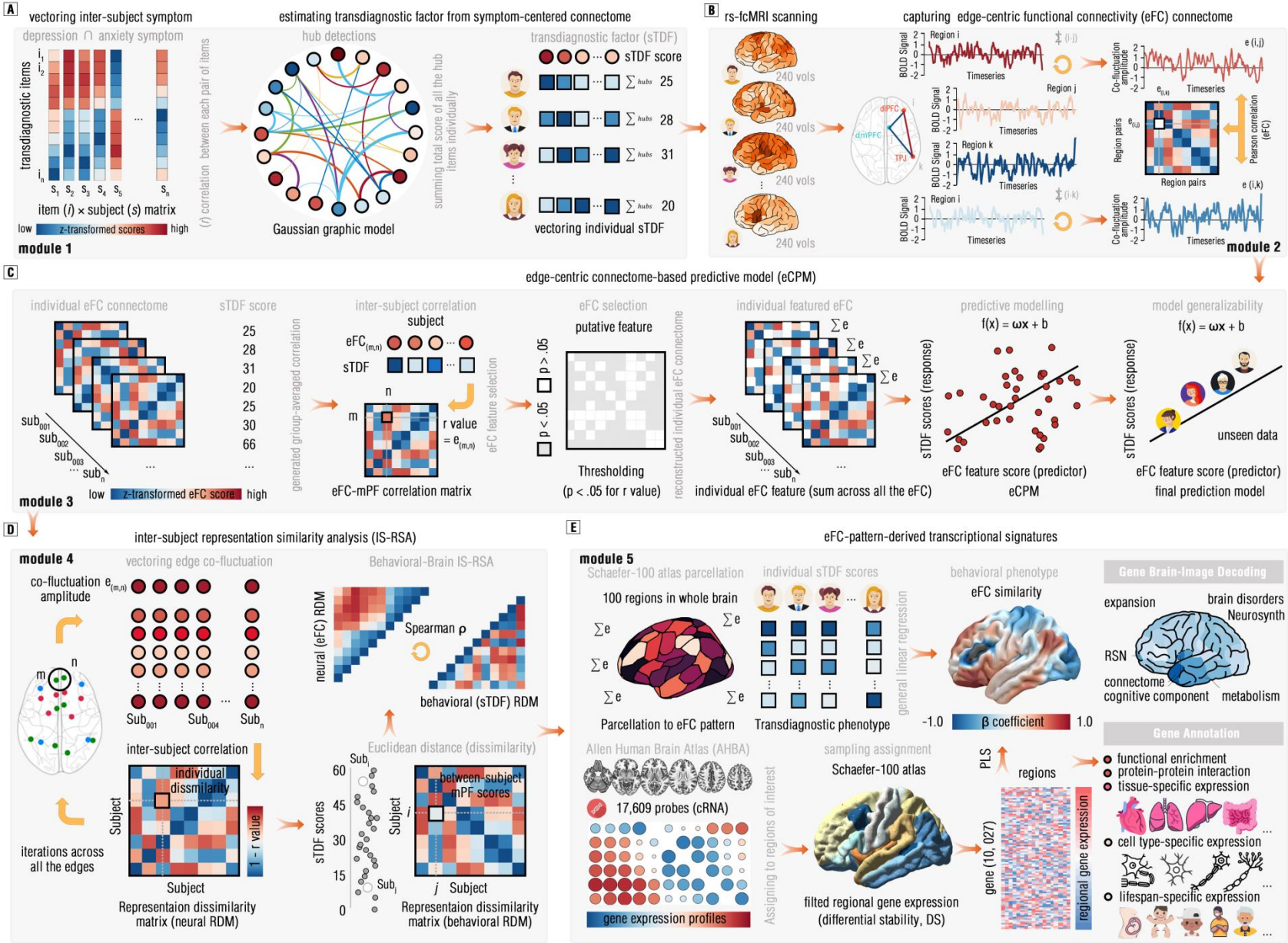

A
vectoring inter-subject symptom
estimating transdiagnostic factor from symptom-centered connectome
depression ∩ anxiety symptom
transdiagnostic items
item (i) × subject (s) matrix
low z-transformed scores high
module 1
(r) correlation between each pair of items
hub detections
Gaussian graphic model
summing total score of all the hub items individually
transdiagnostic factor (sTDF)
sTDF score
vectoring individual sTDF
B
rs-fcMRI scanning
capturing edge-centric functional connectivity (eFC) connectome
240 vols
BOLD Signal
Region i
Region j
Region k
Timeseries
dlPFC
dmPFC
TPJ
Co-fluctuation amplitude
e (i,j)
e (i,k)
Region pairs
Pearson correlation (eFC)
module 2
C
edge-centric connectome-based predictive model (eCPM)
individual eFC connectome
sTDF score
generated group-averaged correlation
inter-subject correlation
subject
eFC(m,n)
sTDF
r value = e(m,n)
eFC-mPF correlation matrix
eFC feature selection
eFC selection
putative feature
p > .05
p < .05
Thresholding
(p < .05 for r value)
reconstructed individual eFC connectome
individual featured eFC
individual eFC feature (sum across all the eFC)
predictive modelling
f(x) = ωx + b
sTDF scores (response)
eFC feature score (predictor)
eCPM
model generalizability
unseen data
final prediction model
low z-transformed eFC score high
module 3
D
inter-subject representation similarity analysis (IS-RSA)
module 4
vectoring edge co-fluctuation
co-fluctuation amplitude
inter-subject correlation
individual dissimilarity
Subject
1 - r value
Representaion dissimilarity matrix (neural RDM)
iterations across all the edges
Behavioral-Brain IS-RSA
neural (eFC) RDM
Spearman ρ
behavioral (sTDF) RDM
Euclidean distance (dissimilarity)
sTDF scores
between-subject mPF scores
Representaion dissimilarity matrix (behavioral RDM)
E
eFC-pattern-derived transcriptional signatures
module 5
Schaefer-100 atlas parcellation
100 regions in whole brain
Parcellation to eFC pattern
individual sTDF scores
Transdiagnostic phenotype
general linear regression
behavioral phenotype
eFC similarity
-1.0 β coefficient 1.0
Gene Brain-Image Decoding
expansion
brain disorders
Neurosynth
RSN
connectome
cognitive component
metabolism
Gene Annotation
functional enrichment
protein-protein interaction
tissue-specific expression
cell type-specific expression
lifespan-specific expression
Allen Human Brain Atlas (AHBA)
17,609 probes (cRNA)
gene expression profiles
Assigning to regions of interest
sampling assignment
Schaefer-100 atlas
filted regional gene expression
(differential stability, DS)
PLS
regions
gene (10, 027)
regional gene expression

Fig. 1 Technical and Research Workflow.

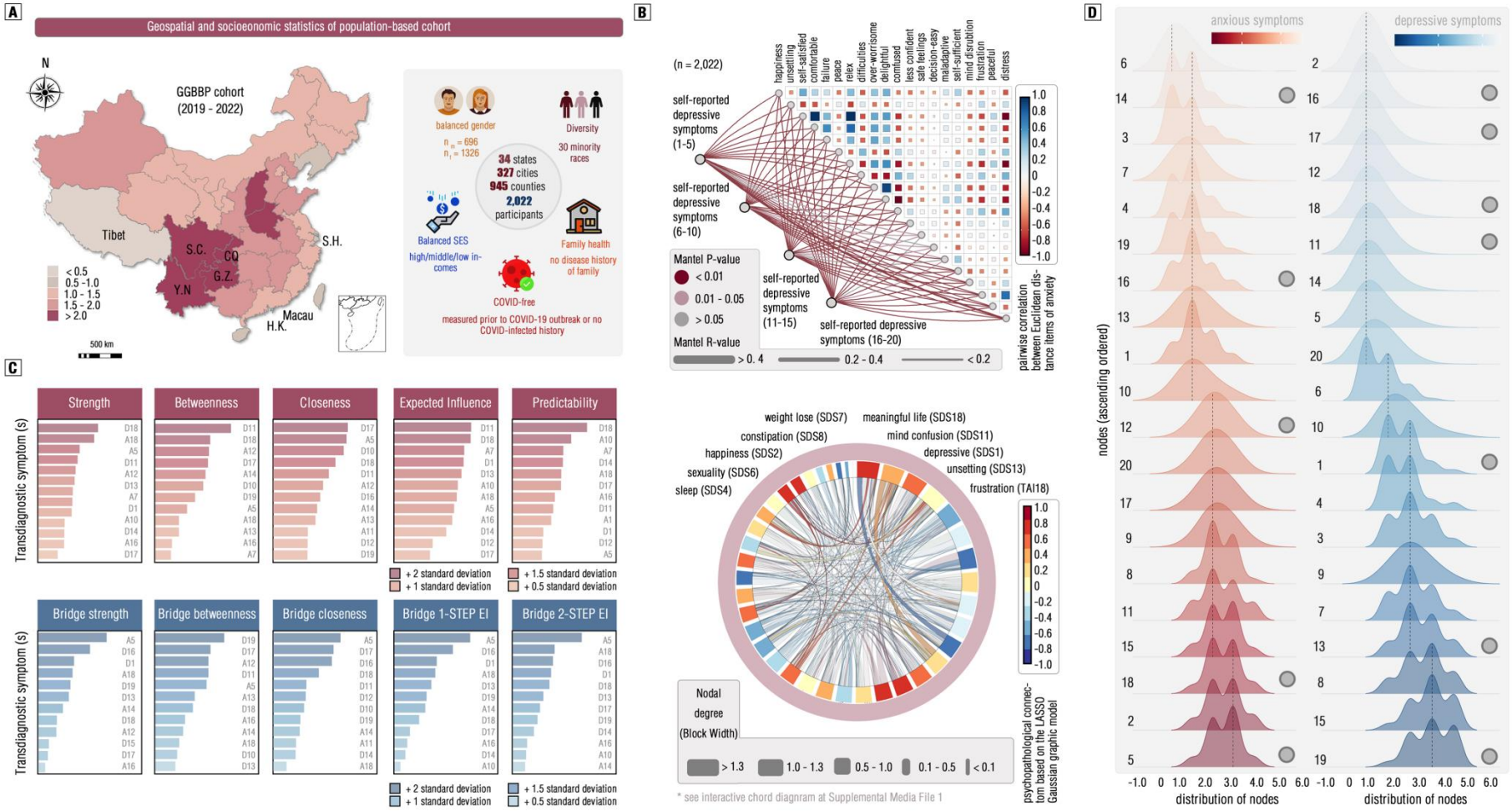


Fig 2. Sociodemograprahic Characteristics and Gaussian Graphic Model of Symptom-Centered Network.

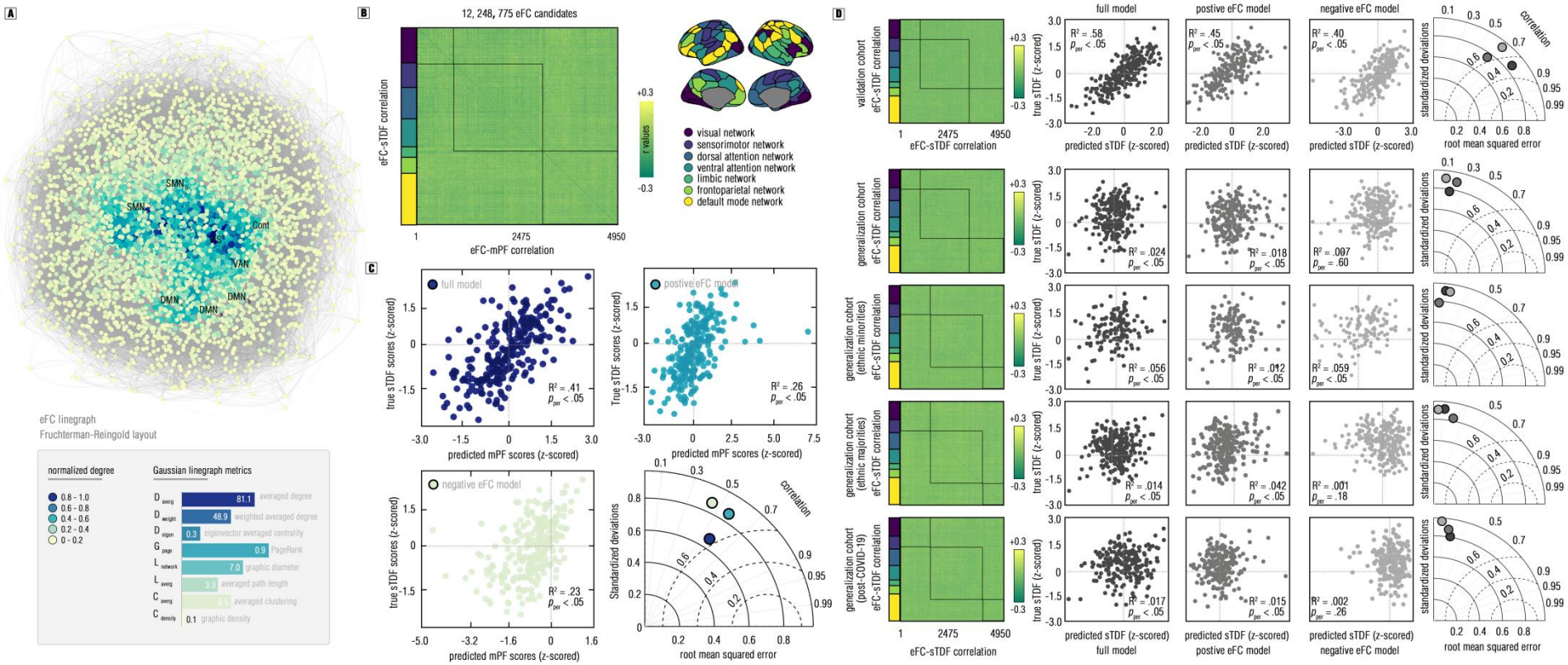
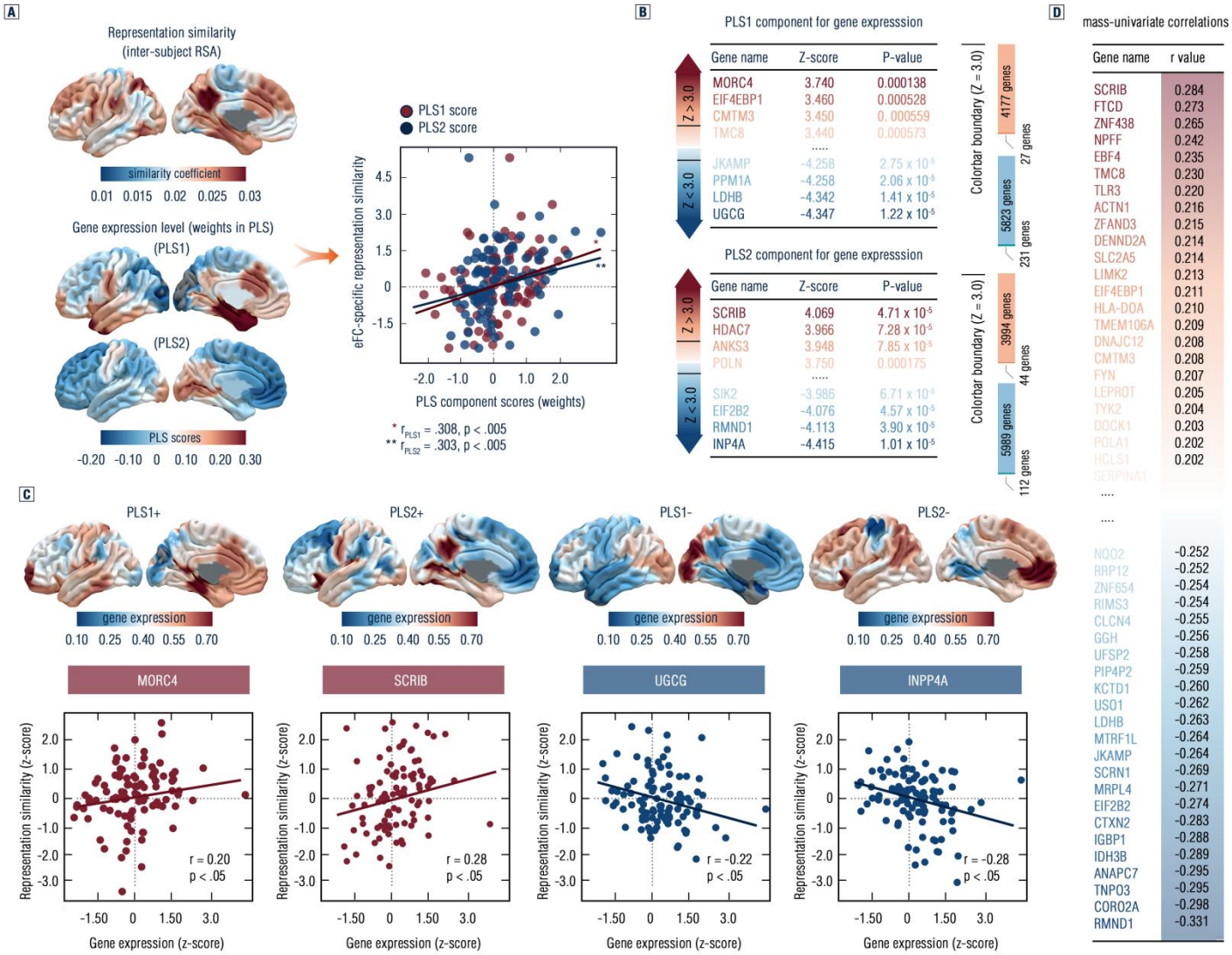
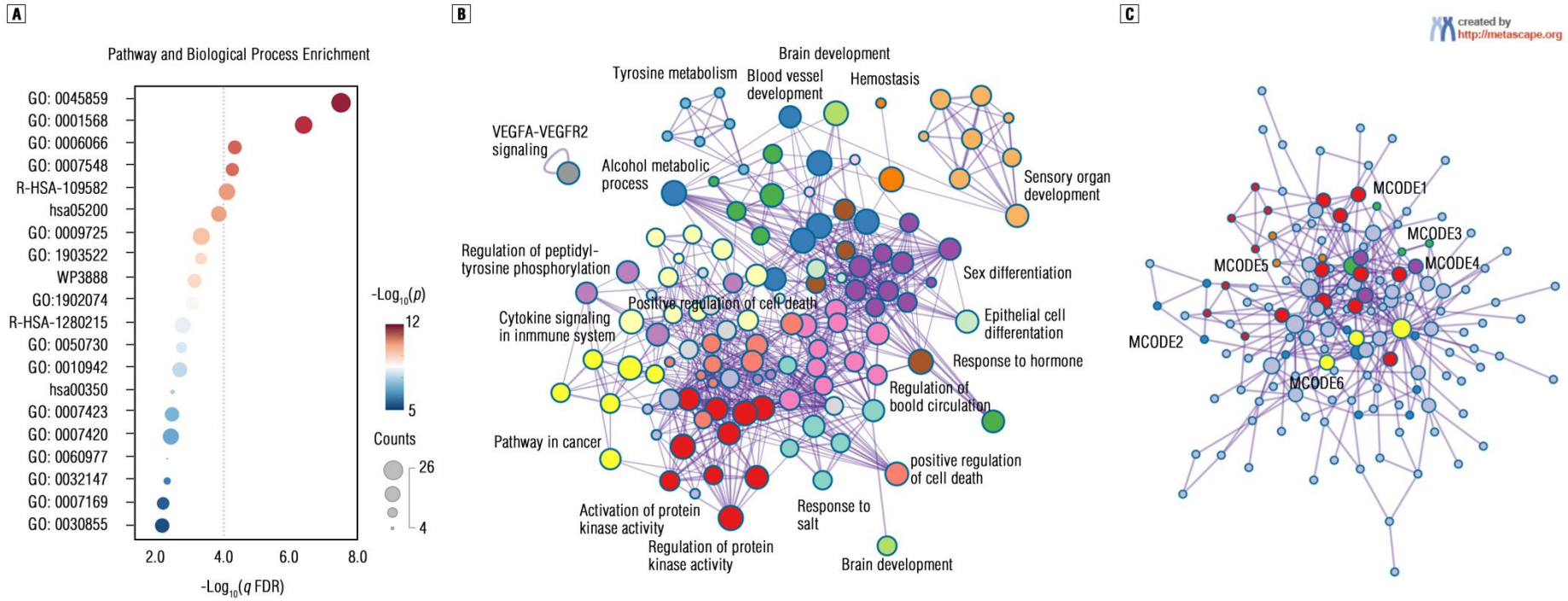


Fig 3. The eFC Line-Graph Connectome and Model Performance of eCPM.

Fig 4. Transcri

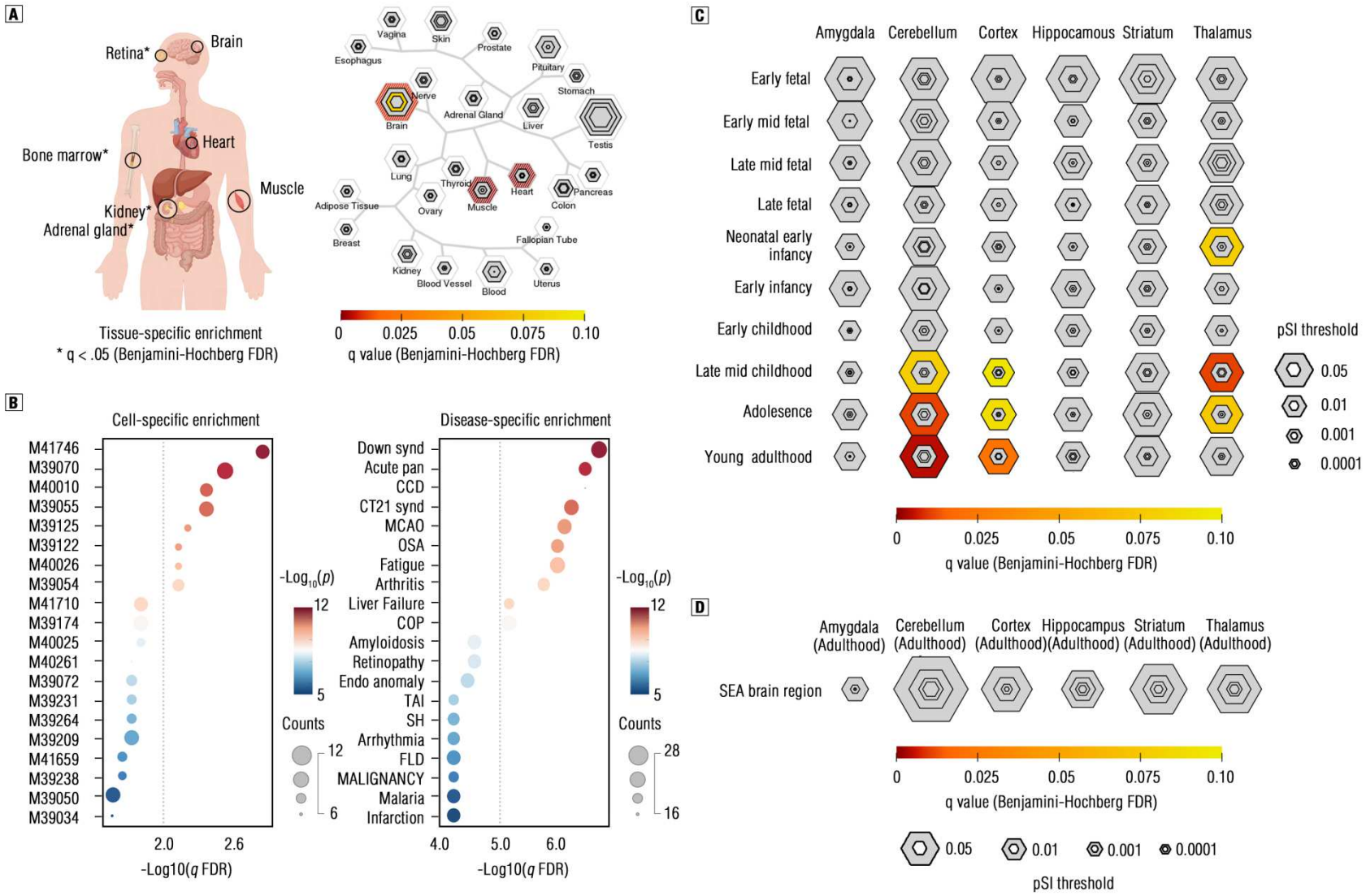


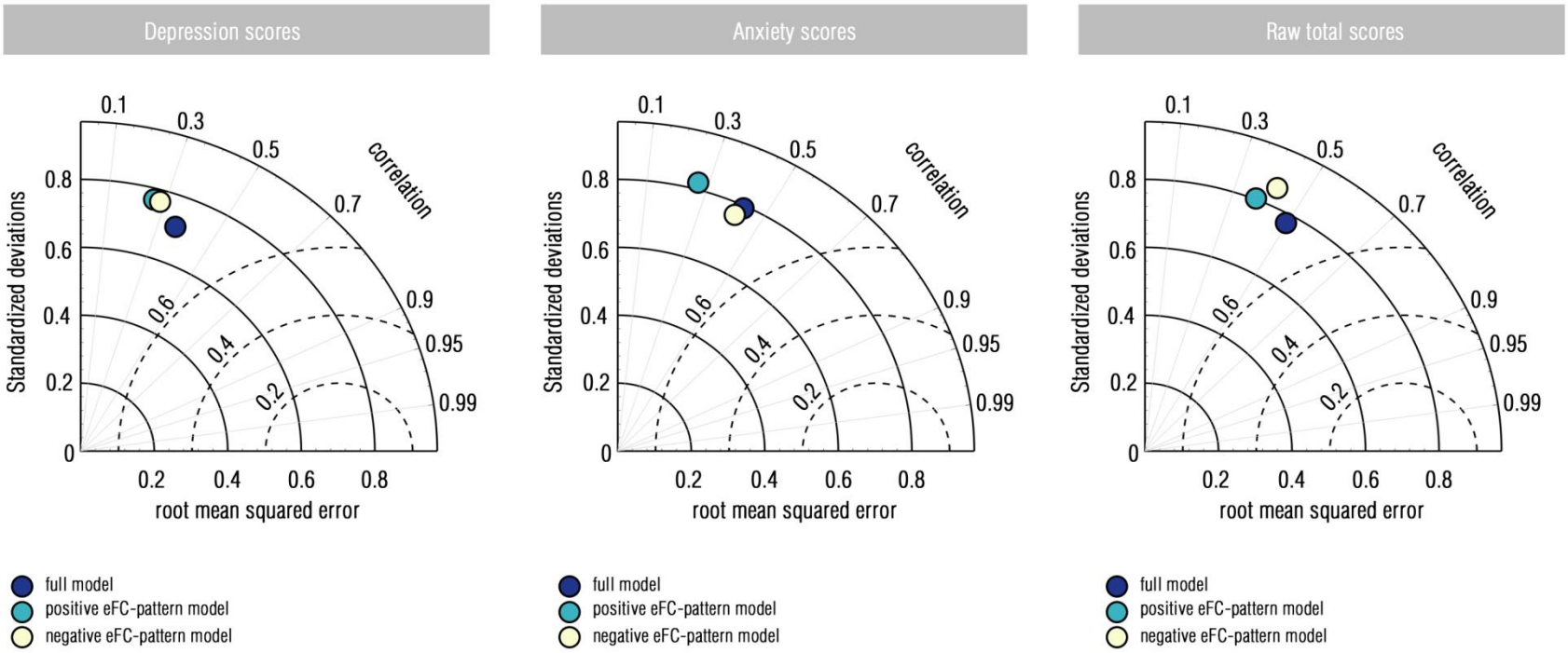
789



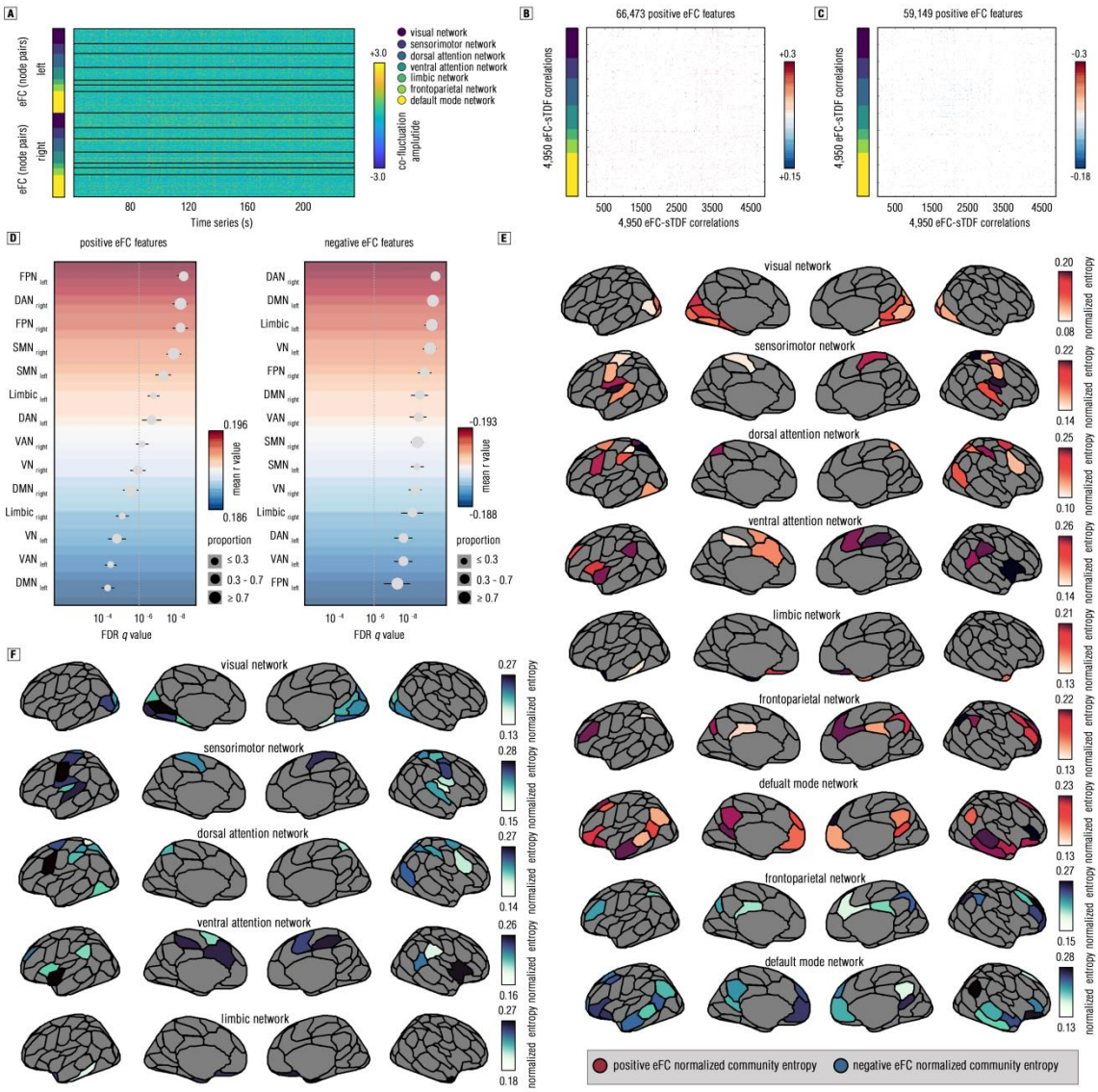
790 **Fig 5. Enrichment of Biological Processes/Pathways and Protein-to-Protein Interaction.**

Fig 6. Sp

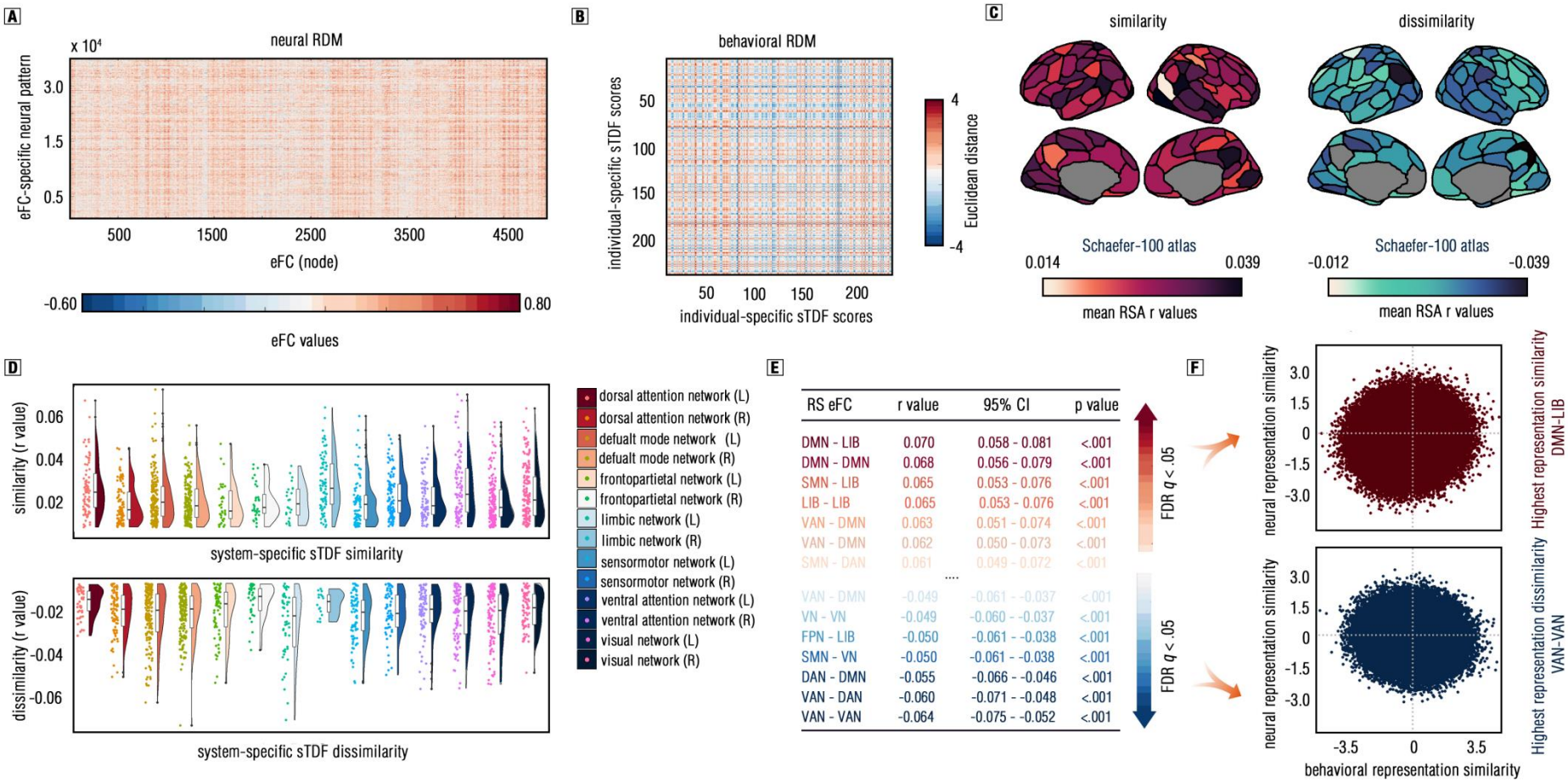

A
Retina*
Brain
Bone marrow*
Heart
Muscle
Kidney*
Adrenal gland*
Vagina
Skin
Esophagus
Prostate
Pituitary
Stomach
Nerve
Brain
Adrenal Gland
Liver
Testis
Lung
Thyroid
Heart
Pancreas
Adipose Tissue
Ovary
Muscle
Colon
Breast
Fallopian Tube
Kidney
Blood Vessel
Blood
Uterus
Tissue-specific enrichment
* q < .05 (Benjamini-Hochberg FDR)
0
0.025
0.05
0.075
0.10
q value (Benjamini-Hochberg FDR)
B
Cell-specific enrichment
M41746
M39070
M40010
M39055
M39125
M39122
M40026
M39054
M41710
M39174
M40025
M40261
M39072
M39231
M39264
M39209
M41659
M39238
M39050
M39034
2.0
2.6
-Log10(q FDR)
-Log10(p)
12
5
Counts
12
6
Disease-specific enrichment
Down synd
Acute pan
CCD
CT21 synd
MCAO
OSA
Fatigue
Arthritis
Liver Failure
COP
Amyloidosis
Retinopathy
Endo anomaly
TAI
SH
Arrhythmia
FLD
MALIGNANCY
Malaria
Infarction
4.0
5.0
6.0
-Log10(q FDR)
-Log10(p)
12
5
Counts
28
16
C
Amygdala
Cerebellum
Cortex
Hippocamous
Striatum
Thalamus
Early fetal
Early mid fetal
Late mid fetal
Late fetal
Neonatal early infancy
Early infancy
Early childhood
Late mid childhood
Adolesence
Young adulthood
pSI threshold
0.05
0.01
0.001
0.0001
0
0.025
0.05
0.075
0.10
q value (Benjamini-Hochberg FDR)
D
Amygdala (Adulthood)
Cerebellum (Adulthood)
Cortex (Adulthood)
Hippocampus (Adulthood)
Striatum (Adulthood)
Thalamus (Adulthood)
SEA brain region
0
0.025
0.05
0.075
0.10
q value (Benjamini-Hochberg FDR)
0.05
0.01
0.001
0.0001
pSI threshold



Extended Data Fig. 1 Model Performance for the Trained eCPM on Single-Disorder Symptoms. By testing this trained eCPM for the single-disorder symptoms (raw total scores), we found the decreased predictability of this model for these single symptoms, irrespective of training from positive (positive eFC-pattern model), negative (negative eFC-pattern model) or the combined eFCs (full model).



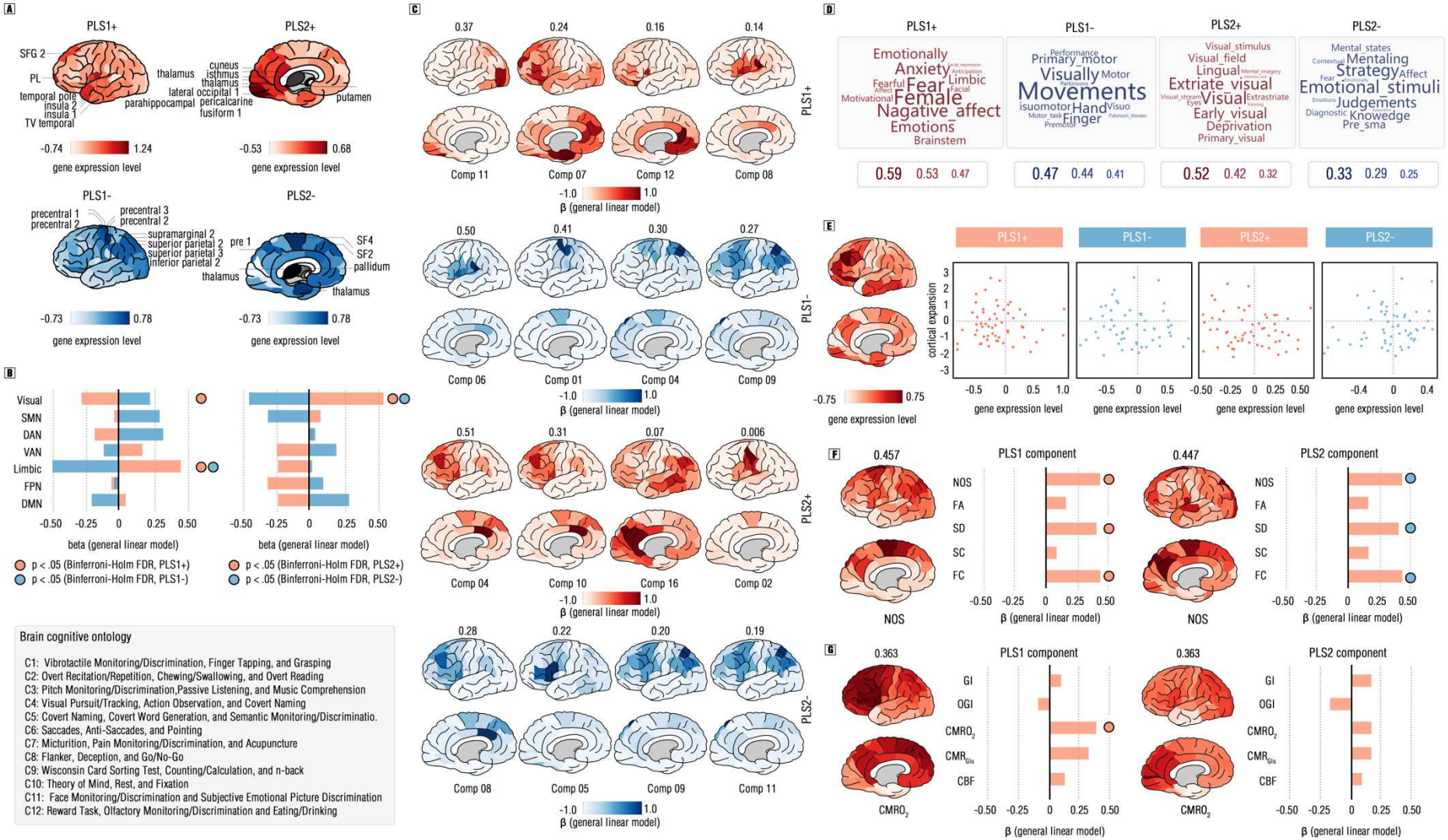
Extended Data Fig 2. Contributive eFC Features of trained eCPM. a, It showed the "edge time series" for each edge-centric "node" by estimating co-fluctuations. The blocks in the left side of matrix indicated corresponding brain system that parceled by Yeo-7 atlas. **b-c,** We have drawn matrix to show "contributive edges" in the eCPM, which were determined by the inter-subject positive (b) and negative (c) correlations between eFCs and sTDF ($p < .05$, uncorrected). **d,** By estimating averaged correlation within each brain system, we showed the mean (95% confidence interval) correlation coefficient for each one, with descending order. The point size in these plots indicated the proportion of the number of included "contributive edges" on the possible maximum number within each brain system. **e-f,** We illustrated normalized entropy from "contributive edges" with positive correlations to sTDF (e) and negative correlations to sTDF (f) into Schaefer-100 atlas, and showed these results by using Yeo 7 brain system, respectively.



Extended Data Fig 3. Contributive Feature with High Representation Similarity (RS). a, The eFC-specific neural patterns for each "edge-wise" node have been illustrated by this 4,950 x 4,949 neural representation dissimilarity matrix (RDM). **b,** We drew the behavioral RDM by showing the Euclidean distance between each pair of sTDF scores. **c,** We rearranged RS *r* values into each parcel from the Schaefer-100 atlas., with the left (right) panel for positively (negatively)

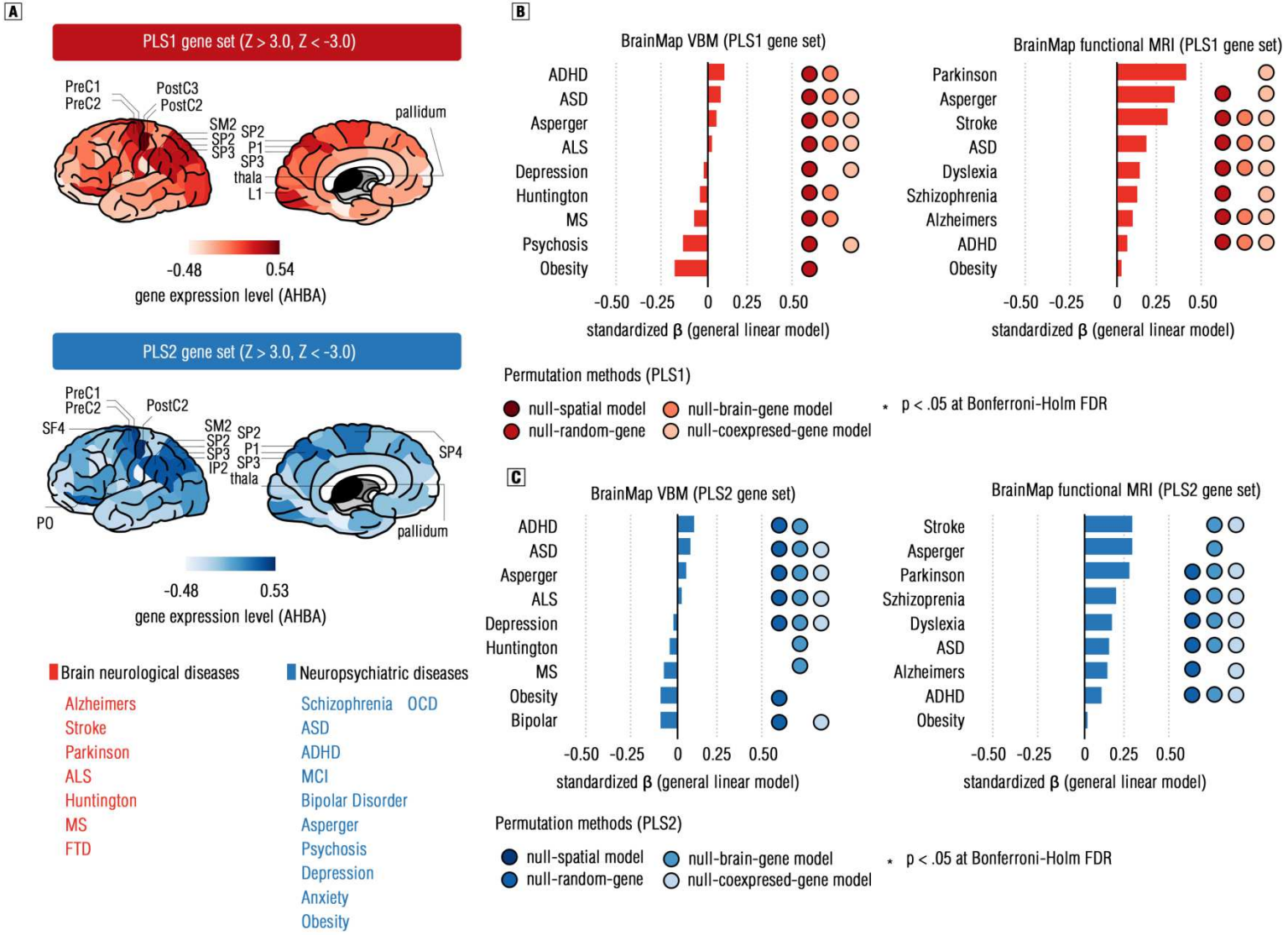
similarity between behavioral and neural RDM. **d,** We further plot the density and distributions of these RS by realigning into intra-communications in the Yeo-7 brain systems that captured by clustering algorithm (Faskowitz et al., 2020). **e,** The inter-connections between these brain systems have been shown with $q < .05$ after Bonferroni-Holm FDR correction. **f,** Scatter plots for the highest positive (top) and negative (bottom) RS have been provided.

838


A
PLS1+
SFG 2
PL
temporal pole
insula 2
insula 1
TV temporal
thalamus
parahippocampal
-0.74
1.24
gene expression level
PLS2+
cuneus
isthmus
thalamus
lateral occipital 1
pericalcarine
fusiform 1
putamen
-0.53
0.68
gene expression level
PLS1-
precentral 1
precentral 2
precentral 3
precentral 2
supramarginal 2
superior parietal 2
superior parietal 3
inferior parietal 2
thalamus
-0.73
0.78
gene expression level
PLS2-
pre 1
SF4
SF2
pallidum
thalamus
-0.73
0.78
gene expression level
B
Visual
SMN
DAN
VAN
Limbic
FPN
DMN
-0.50 -0.25 0 0.25 0.50
beta (general linear model)
p < .05 (Binferroni-Holm FDR, PLS1+)
p < .05 (Binferroni-Holm FDR, PLS1-)
p < .05 (Binferroni-Holm FDR, PLS2+)
p < .05 (Binferroni-Holm FDR, PLS2-)
Brain cognitive ontology
C1: Vibrotactile Monitoring/Discrimination, Finger Tapping, and Grasping
C2: Overt Recitation/Repetition, Chewing/Swallowing, and Overt Reading
C3: Pitch Monitoring/Discrimination,Passive Listening, and Music Comprehension
C4: Visual Pursuit/Tracking, Action Observation, and Covert Naming
C5: Covert Naming, Covert Word Generation, and Semantic Monitoring/Discriminatio.
C6: Saccades, Anti-Saccades, and Pointing
C7: Micturition, Pain Monitoring/Discrimination, and Acupuncture
C8: Flanker, Deception, and Go/No-Go
C9: Wisconsin Card Sorting Test, Counting/Calculation, and n-back
C10: Theory of Mind, Rest, and Fixation
C11: Face Monitoring/Discrimination and Subjective Emotional Picture Discrimination
C12: Reward Task, Olfactory Monitoring/Discrimination and Eating/Drinking
C
PLS1+
0.37 0.24 0.16 0.14
Comp 11 Comp 07 Comp 12 Comp 08
-1.0 1.0
β (general linear model)
PLS1-
0.50 0.41 0.30 0.27
Comp 06 Comp 01 Comp 04 Comp 09
-1.0 1.0
β (general linear model)
PLS2+
0.51 0.31 0.07 0.006
Comp 04 Comp 10 Comp 16 Comp 02
-1.0 1.0
β (general linear model)
PLS2-
0.28 0.22 0.20 0.19
Comp 08 Comp 05 Comp 09 Comp 11
-1.0 1.0
β (general linear model)
D
PLS1+
Emotionally
Anxiety
Fearful
Fear
Limbic
Motivational
Female
Nagative_affect
Emotions
Brainstem
0.59 0.53 0.47
PLS1-
Primary_motor
Visually
Motor
Movements
isuomotor
Hand
Finger
0.47 0.44 0.41
PLS2+
Visual_stimulus
Visual_field
Lingual
Extriate_visual
Visual
Early_visual
Deprivation
Primary_visual
0.52 0.42 0.32
PLS2-
Mental_states
Mentaling
Strategy
Affect
Emotional_stimuli
Judgements
Knowedge
Pre_sma
0.33 0.29 0.25
E
-0.75 0.75
gene expression level
PLS1+ PLS1- PLS2+ PLS2-
cortical expansion
gene expression level
F
0.457
NOS
PLS1 component
NOS FA SD SC FC
-0.50 -0.25 0 0.25 0.50
β (general linear model)
0.447
NOS
PLS2 component
NOS FA SD SC FC
-0.50 -0.25 0 0.25 0.50
β (general linear model)
G
0.363
CMRO2
PLS1 component
GI OGI CMRO2 CMRGlu CBF
-0.50 -0.25 0 0.25 0.50
β (general linear model)
0.363
CMRO2
PLS2 component
GI OGI CMRO2 CMRGlu CBF
-0.50 -0.25 0 0.25 0.50
β (general linear model)

Extended Data Fig 4. Transcriptional Annotation of these Gene Setsby Macroscale Brain-imaging Association (GAMBA) Analysis. **a,** The gene expression distributions for gene sets have been illustrated from PLS1 and PLS2, with separations to positive (PLS1+, PLS2+) and negative weights (PLS1-, PLS2-). The visualization has been implemented by D-K atlas. **b,** By using the GAMBA decoding, The linear regression model was used to fit the expression levels to the network property of resting-state networks (RSNs) by Yeo-7 atlas, and the standardized beta coefficient was presented by bar plots. Dots with light orange (PLS+) and light (PLS-) blue indicated the p value for this beta coefficient reached statistical significance level ($p < .05$ at Bonferronni-Holm FDR correction) for PLS+ and PLS-, respectively. **c,** We decoded brain cognitive ontology by these gene sets (PLS1+, PLS1-, PLS2+, PLS2-), and plot the brain spatial distributions (though no one reached statistical significance), with details for each ontology at the left-bottom panel. **d,** We decoded the cognitive correlates of PLS components by using online meta-analysis at NeuroSynth, respectively, with large font size for high correlation strengths. **e,** It showed the distribution of cortical expansion into the brain model, with scatter plots for these PLS components. No significant correlations were found, but it appeared decreased trends for the PLS+ and such cortical evolution. **f,** Rather the separations to positive and negative weights, we had drawn the plots to show the associates of whole-brain topological properties to the entire PLS component 1 and 2. Nodal strength indices: NOS = Number Of Streamlines, FA = Fractional Anisotropy, SD = Streamline Density, FC = Functional Connectome, SC = Structural Connectome. Dots with light orange and blue represented ones to reach statistical significance ($p < .05$ at Bonferronni-Holm FDR correction) for PLS1 and PLS2, respectively. g, We showed the associates of such gene sets to the cortical metabolism. Label of this dot was in line with previous one. GI = Glycolytic Index, OGI = Oxygen-Glucose Index, $CMRO_2/GMR_{Glu}$ = Cerebral Metabolic Rate of Oxygen/Glucose, CBF, Cerebral Blood Flow.

865


A
PLS1 gene set (Z > 3.0, Z < -3.0)
PreC1
PreC2
PostC3
PostC2
SM2
SP2
SP3
SP2
P1
SP3
thala
L1
pallidum
-0.48
0.54
gene expression level (AHBA)
PLS2 gene set (Z > 3.0, Z < -3.0)
PreC1
PreC2
PostC2
SF4
SM2
SP2
SP3
IP2
SP2
P1
SP3
thala
SP4
PO
pallidum
-0.48
0.53
gene expression level (AHBA)
Brain neurological diseases
Alzheimers
Stroke
Parkinson
ALS
Huntington
MS
FTD
Neuropsychiatric diseases
Schizophrenia
OCD
ASD
ADHD
MCI
Bipolar Disorder
Asperger
Psychosis
Depression
Anxiety
Obesity
B
BrainMap VBM (PLS1 gene set)
ADHD
ASD
Asperger
ALS
Depression
Huntington
MS
Psychosis
Obesity
-0.50 -0.25 0 0.25 0.50
standardized β (general linear model)
BrainMap functional MRI (PLS1 gene set)
Parkinson
Asperger
Stroke
ASD
Dyslexia
Szhizophrenia
Alzheimers
ADHD
Obesity
-0.50 -0.25 0 0.25 0.50
standardized β (general linear model)
Permutation methods (PLS1)
null-spatial model
null-random-gene
null-brain-gene model
null-coexpresed-gene model
* p < .05 at Bonferroni-Holm FDR
C
BrainMap VBM (PLS2 gene set)
ADHD
ASD
Asperger
ALS
Depression
Huntington
MS
Obesity
Bipolar
-0.50 -0.25 0 0.25 0.50
standardized β (general linear model)
BrainMap functional MRI (PLS2 gene set)
Stroke
Asperger
Parkinson
Szhizoprenia
Dyslexia
ASD
Alzheimers
ADHD
Obesity
-0.50 -0.25 0 0.25 0.50
standardized β (general linear model)
Permutation methods (PLS2)
null-spatial model
null-random-gene
null-brain-gene model
null-coexpresed-gene model
* p < .05 at Bonferroni-Holm FDR

Extended Data Fig 5. Risks of Neurological and Psychiatric Diseases of These Gene Sets. a, Spatial distributions of gene expression for PLS1 (top one) and PLS2 (bottom one) sets have been illustrated. Top 20% regions that showed the highest gene expression were labeled. Full name of these abbreviations can be found in the D-K atlas. In the bottom panel, all the neurological psychiatric disorders that these gene sets were involved by examining in the BrainMap database have been detailed. **b-c,** The general linear regression models were conducted to predict the risks of these diseases from PLS1 (*b*) and PLS2 (*c*) gene list, and were visualized by beta coefficients in these bar plots. No associates of diseases reached statistical significance ($p < .05$ at Bonferrioni-Holm correction). To improve gene-expression-correlated specificity, four permutation methods to estimate the statistical significance of these regression models were furthered conducted, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplement1SIFirstSubmissionFinal.docx
- Supplement2.pdf
- Supplement3.xlsx