Lifestyle and transcriptional signatures associated with ethnicity/race-related variations in the functional connectome

Ziteng Han[1], Kexin Wang[1], Tiantian Liu[1], Yunxiao Ma[2], Guoyuan Yang[1, 3, *], Tianyi Yan[1, *]

[1]School of Medical Technology, Beijing Institute of Technology, Beijing, China

[2]School of Life Science, Beijing Institute of Technology, Beijing, China

[3]Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing, China

***Corresponding authors**:

Guoyuan Yang, email: yanggy@bit.edu.cn

Tianyi Yan, email: yantianyi@bit.edu.cn

Author contributions: G.Y., Z.H. and T.Y. designed the study; Z.H. and G.Y. analysed the data; Z.H. wrote the first draft of the manuscript; G.Y., Z.H., K.W., T.L., Y.M., and T.Y. edited the manuscript.

Keywords: neuroimaging, trans-ethnicity/race studies, brain-behaviour relationships, precision neuroscience, transcriptome–neuroimaging associations

Competing interest statement: The authors declare that they have no competing interests.

Abstract

The functional connectome architecture of the human brain, which involves ethnicity/race characteristics, has significantly advanced our understanding of brain–behaviour relationships. However, the multifaceted underpinnings of ethnicity/race-related variations in the brain functional connectome remain largely unclear. In this study, we leverage precision individualized functional mapping to examine ethnicity/race-related differences in the brain's intrinsic functional organization, as well as their associations with lifestyle factors and transcriptional signatures. Our findings indicate that variations in network topography and functional connectivity across ethnic/racial groups follow a hierarchical pattern along the sensorimotor–association axis and are constrained by brain morphology. We identify lifestyle factors, primarily education and substance use, that mediate the associations between ethnicity/race and functional connectivity. Via the use of human brain gene expression data, we further demonstrate that cortical gene expression patterns are spatially associated with ethnicity/race-related variations in functional connectivity. Gene ontology analysis of ethnic/racial-associated genes reveals significant enrichment in biological processes, including synaptic signalling pathways and neuronal system development. Collectively, these results reveal the complex relationships between ethnicity/race-related differences in brain function and brain morphology, lifestyle factors, and transcriptomic profiles.

Introduction

As neuroscience has advanced towards greater generalizability, applicability and equity, increasing attention has been given to ethnicity/race-related differences in brain structure and function[1, 2], which are further interwoven with variations in distinct human behaviours and psychiatric systems[3, 4]. For example, evidence from neuroimaging studies indicates that brain anatomical properties, such as cortical thickness[5], curvature[6] and gray matter volume[7, 8], differ significantly across ethnic/racial populations. Moreover, ethnicity/race-related variability has been observed in the variance in intrinsic cortical activity[9] and connectivity profiles[10, 11]. On the basis of precision mapping techniques, ethnicity/race-related variability in the large-scale functional network topography is heterogeneously distributed across the cortical mantle[11, 12]. These spatial patterns of network-level variability also conform with the fundamental properties of brain structure and function[10]. In brain-wide association studies, brain–behaviour relationships have demonstrated limited generalizability across ethnic/racial populations, highlighting model biases introduced by non-representative sampling in neuroimaging cohorts[13-15]. However, given the complex interplay and decoupling of brain structure and function[16, 17], studies focused on exploring the associations between trans-population differences in brain structure and function are lacking. The extent to which ethnicity/race-related differences in the intrinsic functional organization can be explained by differences in cortical morphometric features remains to be systematically investigated.

Ethnicity/race is a social construct intertwined with cultural, environmental, and genetic factors[18]. Focusing solely on superficial ethnicity/race effects may limit neuroscience and lead to potentially unjust treatment of specific populations. Thus, the existing findings emphasize the importance of revealing the root causes of ethnicity/race-related differences[3, 5, 19, 20], which are critical for the discovery of generalizable brain–behaviour relationships. Previous studies have demonstrated that the socioeconomic status is an important driver of ethnicity/race differences in health problems (e.g., mortality) and treatment-seeking behaviours[21-23]. These factors significantly impact the development of brain structures[24] and neurocognitive abilities[25]. Past experiences, such as exposure to social stressors, can contribute to ethnicity/race-related variability in neurology and psychopathology (e.g., ageing and threat response)[5, 26, 27]. Moreover, individual lifestyle choices, such as alcohol dependence and lack of physical activity, are risk factors influencing mental health and occurrence of chronic diseases[28, 29], with the brain fulfilling an important mediating role[30, 31]. Recently, epidemiological evidence has suggested that unhealthy lifestyles can explain ethnicity/race-related health inequities[19]. However, it remains poorly understood whether and how

lifestyle mediates ethnicity/race-related variability in brain function, thus limiting our comprehensive understanding of the neurobiological mechanisms of health inequities across diverse populations.

Understanding the molecular correlates of brain-wide phenotypes in vivo is important in the brain sciences[32, 33]. Via the use of data from the Allen Human Brain Atlas (AHBA), which is a valuable brain-wide atlas for quantifying the transcriptional activity of thousands of genes, prior studies have identified specific transcriptional signatures associated with neurodevelopment[34], cortical functional hierarchy[35], psychiatric disorders[36, 37], and inter-individual variability[38]. These frameworks provide valuable insights into the microscale mechanisms that underlie relevant brain circuits. Recent transcription studies of postmortem brain tissues from African American (AA) individuals have revealed that genetic ancestry affects gene expression in the brain, with both genetic variations and environments contributing to ancestry-associated expression changes[39]. Ancestry-associated differentially expressed genes (DEGs) are primarily enriched in immune response pathways[39, 40] and are linked with the heritability of neurological disorders[39], echoing the observed prevalence differences across populations[41-43]. Despite these advances, it remains largely unknown how ethnicity/race-related multifaceted factors modulate gene expression and the transcriptomic mechanisms underlying ethnicity/race-related variability in brain functional organization.

This study aimed to investigate the multiple mechanisms underlying ethnicity/race effects on brain functional connectome architecture from the perspectives of anatomical constraints, lifestyle and environments, and transcription. We used two age-group datasets from the Human Connectome Project (HCP), namely, the HCP-Young Adult (HCP-YA) and HCP-Development (HCP-D) datasets. These datasets containing multimodal brain imaging measures with ethnic/racial groups were categorized on the basis of self-reported data from the U.S. population. By focusing on U.S. data, we aimed to minimize potential confounding factors resulting from cross-cultural differences (e.g., Western culture vs. East Asian culture). Specifically, we **first** utilized the advanced multisession hierarchical Bayesian model (MS-HBM) to delineate the individual-specific functional topography and functional connectome. We then employed brain-based predictive modelling approaches to demonstrate that ethnicity/race is associated with inter-individual variability in resting-state functional connectivity (RSFC) and applied linear regression models (LRMs) to evaluate ethnicity/race-related variations in brain morphological features. **Second**, we explored the existence of shared predictive network features across ethnicity/race groups and various lifestyle variables. We

also determined that the functional connectivity correlates of ethnicity/race are partially mediated by lifestyle factors through structural equation modelling and mediation analysis. **Finally**, we linked ethnicity/race-relevant functional connectivity features with the AHBA to reveal the transcriptional mechanisms underlying ethnicity/race population diversity. Our work adheres to the latest theoretical framework for neuroimaging research related to ethnicity/race[18] and offers a view explaining the interplay between ethnicity/race–lifestyle and ethnicity/race–gene expression.

Results

We employed two independent large-scale datasets, namely, the HCP-YA and HCP-D datasets, and focused on AA and white American (WA) participants. The quality control process closely followed that in previous work[14]. Afterwards, 101 AA participants and 721 WA participants from the HCP-YA dataset (N = 822; age, 22 to 37 years; 435 females) and 68 AA participants and 404 WA participants from the HCP-D dataset (N = 472; age, 6 to 21 years; 253 females) were used for subsequent analyses. Ethnic/racial groups were categorized on the basis of self-reported information. For each dataset, AA individuals were randomly split into 5 folds with the family structure accounted for, and Hungarian matching was then employed to select matched AA–WA pairs on the basis of age, sex and root-mean-square framewise displacement (RMS)[14]. To prevent sensitivity of the analysis results to the particular data split, the above steps were repeated multiple times until 40 distinct AA splits with matched WA splits emerged.

Ethnicity/race-related variability in brain functional topography and connectome

The MS-HBM, which has been previously shown to capture the functional organization of individuals precisely, was adopted to obtain individual-specific parcellation maps[44]. Briefly, in the MS-HBM, it is assumed that each cortical region exhibits a unique RSFC pattern, and a unified statistical framework is established that accounts for inter-individual RSFC variability, intra-individual RSFC variability and spatial constraints (i.e., encouraging adjacent brain locations with slight changes in RSFC to be assigned the same label) to delineate individualized boundaries between different areas under the guidance of group-level parcellation (Schaefer400)[45]. Compared with several studies primarily focused on examining individual variability in the network topology[46, 47], the MS-HBM takes a further step by revealing the fundamental functional units of these networks within each individual. The resting-state functional magnetic resonance imaging (rs-fMRI) data from all valid sessions for each individual were entered as inputs of the MS-HBM. First,

we randomly selected 40 AA and 40 WA demographically matched individuals to investigate the variability in the group-level functional topography. On the basis of their individual-level parcellations and the maximum probability fusion approach[48], we obtained group-level AA400 and WA400 parcellation maps (**Supplementary Fig. 1**). Similarity analyses quantified with the Dice coefficient revealed that the parcellation maps of the same ethnicity/race group across the different datasets exhibited the highest similarity (**Supplementary Fig. 1**; e.g., AA400 from the HCP-YA dataset compared with AA400 from the HCP-D dataset: 0.844), whereas the parcellation maps spanning both datasets and ethnicity/race groups exhibited the greatest spatial distinctions (e.g., AA400 from the HCP-YA dataset compared with WA400 from the HCP-D dataset: 0.784). Next, we assessed the influence of ethnicity/race on the individual-level functional topography. Specifically, the similarity of individual-specific parcellation maps between all AA–WA pairs (101×721 in the HCP-YA dataset and 68×404 in the HCP-D dataset) was calculated and averaged. We observed a non-uniform distribution along the cortex, with the highest variability in the heteromodal association cortex (**Fig. 1a and Supplementary Fig. 2**; including the control and dorsal attention networks), whereas the unimodal sensorimotor cortex demonstrated the highest similarity (especially in the somatomotor and auditory networks).

The individualized parcellations and group-level Schaefer400 parcellations were then applied to the rs-fMRI data to generate RSFC matrices (400 × 400) for each participant separately. Confounding variables, including age, sex, RMS, intracranial volume (ICV), education and income, were regressed from RSFC[14]. Multivariate kernel ridge regression (KRR), in which RSFC matrices are employed as inputs to distinguish between AA and WA individuals, was utilized[49]. Specifically, we performed nested 5-fold cross validation, with all the training folds determining the optimal value of the regularization parameter and the remaining test fold testing the generalizability of the model on the basis of unseen data. The above 5-fold cross-validation step was repeated 40 times, and the classification accuracy was averaged across 5 folds for each data split. In the HCP-YA dataset, our multivariate approach achieved high classification accuracy, with an accuracy of 85.3% ± 2.3% (mean ± standard deviation) via the individual-specific parcellations and an accuracy of 77.9% ± 2.6% via the Schaefer400 parcellations, both of which significantly exceed chance levels (both $P < 0.001$). In the HCP-D dataset, although the model performance decreased slightly, it remained significantly above chance (both $P < 0.001$), yielding accuracies of 81.7% ± 5.0% with the individual-specific parcellations and 62.5% ± 5.1% with the Schaefer400 parcellations. These results

emphasize the importance of treating individual-specific topographic features as phenotypes and highlight the robust ethnicity/race-related variability in functional connectome architectures.

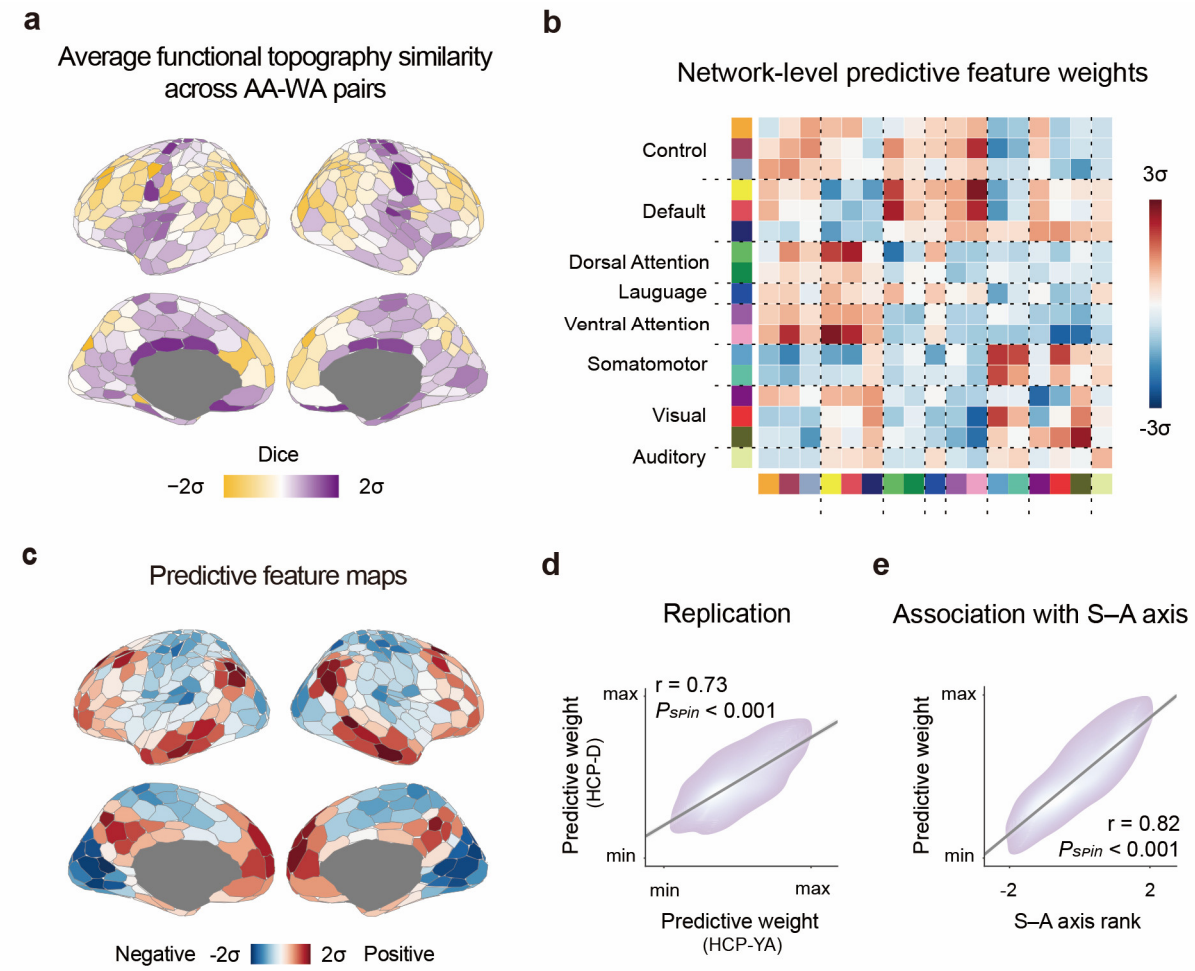

Fig. 1 Functional topography and functional connectivity are associated with ethnicity/race. a, The average similarity of individual-specific parcellations across all the AA–WA pairs revealed that the ethnicity/race-related topographic variability was greatest in the heteromodal association cortices and lowest in the unimodal sensorimotor cortices. The obtained brain maps were visualized using the Schaefer400 parcellations[45] and created with the R package ggseg[50]. **b**, We trained kernel ridge regression (KRR) models using the resting-state functional connectivity (RSFC) as a feature while controlling for age, sex, root-mean-square framewise displacement (RMS), intracranial volume (ICV), education and income to predict individuals' self-reported ethnicity/race. Haufe-transformed regional pairwise feature weights were averaged to the network level via the assignment of cortical regions to one of the Yeo networks. Connections between the default and attention networks, within the default networks, and within the sensorimotor networks largely drive the prediction for both datasets. For visualization, the actual values were subjected to *Z* score standardization. **c,** Haufe-transformed weights of each brain region were averaged to generate ethnicity/race-predictive weight maps. **d,** We performed RSFC-based ethnicity/race prediction in both the discovery HCP-YA dataset and the replication HCP-D dataset. The predictive weights were highly reproducible across the two independent datasets ($r = 0.73$, $P_{spin} < 0.001$). **e,** In the cortex, the ethnicity/race-predictive weight distributions are significantly related to S–A axis ranks ($r = 0.82$, $P_{spin} < 0.001$).

To better understand the network features learned by the model, the regression models were inverted on the basis of Haufe's transform approach to increase interpretability, yielding prediction feature matrices[51]. **Figs. 1b and Supplementary Fig. 3** show the averaged pairwise region-level feature weights at the network level, suggesting that the functional connections associated with ethnicity/race constitute a distributed set of large-scale network circuits, with the most prominent links observed between the default and attentional networks, within the default networks, and within the sensorimotor networks for both datasets. Moreover, the average feature weights of the brain regions were calculated to measure the contribution of each region to the regression model (**Fig. 1c and Supplementary Fig. 3**). We found that the discriminating features derived from the HCP-YA and HCP-D datasets exhibited high generalizability and significant spatial correlations (**Fig. 1d;** r = 0.73, $P_{spin} < 0.001$). In both datasets, the contribution weights showed a hierarchical distribution across the cortex, with opposite contributions between the heteromodal cortices (including the default and frontoparietal control networks) and unimodal cortices responsible for sensory and motor processing. Recently, this global hierarchical layout, which is evident in the spatial pattern of diverse neurobiological properties, has been conceptualized as a cortical axis, referred to as the sensorimotor–association (S–A) axis[52]. Via the use of spatial permutation testing, we found that the contribution of cortical regions to the ethnicity/race prediction model aligns with the S–A axis (**Fig. 1e**; r = 0.82, $P_{spin} < 0.001$), revealing opposing predictive effects for regions located at the sensorimotor and association poles of the S–A axis.

Ethnicity/race-related variability in brain functional connectome is constrained by the fundamental properties of brain organization

The spatial and physical patterns of brain anatomy are known to constrain neural dynamics[16]. Thus, we next aimed to examine whether ethnicity/race-related variability in brain functional connectivity could be explained by the corresponding variability in cortical morphology. We first constructed a morphometric similarity network (MSN) using Pearson's correlation for each individual, which integrates multiple structural MRI features (Z scores) into a biologically meaningful and behaviourally relevant connectome (400×400)[53]. The correlation weights were averaged across rows and within-group individuals to generate group-level regional MSN values (**Fig. 2a**). To investigate the ethnicity/race-related variability in brain morphology, LRMs were employed to regress the covariates, including age, sex, RMS, ICV, education level and income[14]. We observed notable ethnicity/race-related variations in morphometric similarity patterns across the

cerebral cortex (**Fig. 2b and Supplementary Fig. 4**), particularly in regions located in the insula, prefrontal cortex, cuneus and lateral occipital cortex. Importantly, this variability exhibited a high degree of reproducibility across the datasets (**Fig. 2c**; $r = 0.59$, $P_{spin} < 0.001$). We further observed that the *t*-map of the MSN is positively associated with the RSFC-based ethnicity/race predictive weights (**Figs. 2d**; $r = 0.41$, $P_{spin} = 0.012$). These results suggest that brain regions contributing more to RSFC-based ethnicity/race prediction generally exhibit greater ethnicity/race-related variability in the cortical morphology.

In addition to the MSN, we considered other fundamental properties of brain organization. Notably, myelination of the brain, a microstructural marker of neuronal fibres, affects the efficiency of information processing in the nervous system[54, 55]. Intracortical myelination can be estimated using an individual's T1-weighted-to-T2-weighted ratio, whereby higher values reflect greater intracortical myelination. Across the whole cortex, we found that the predictive weight maps were inversely related to the cortical myelin content (**Supplementary Fig. 5**; $r = -0.57$, $P_{spin} < 0.001$). Overall, our results demonstrated that ethnicity/race-related variability in brain functional connectivity is associated with the fundamental properties of brain organization.

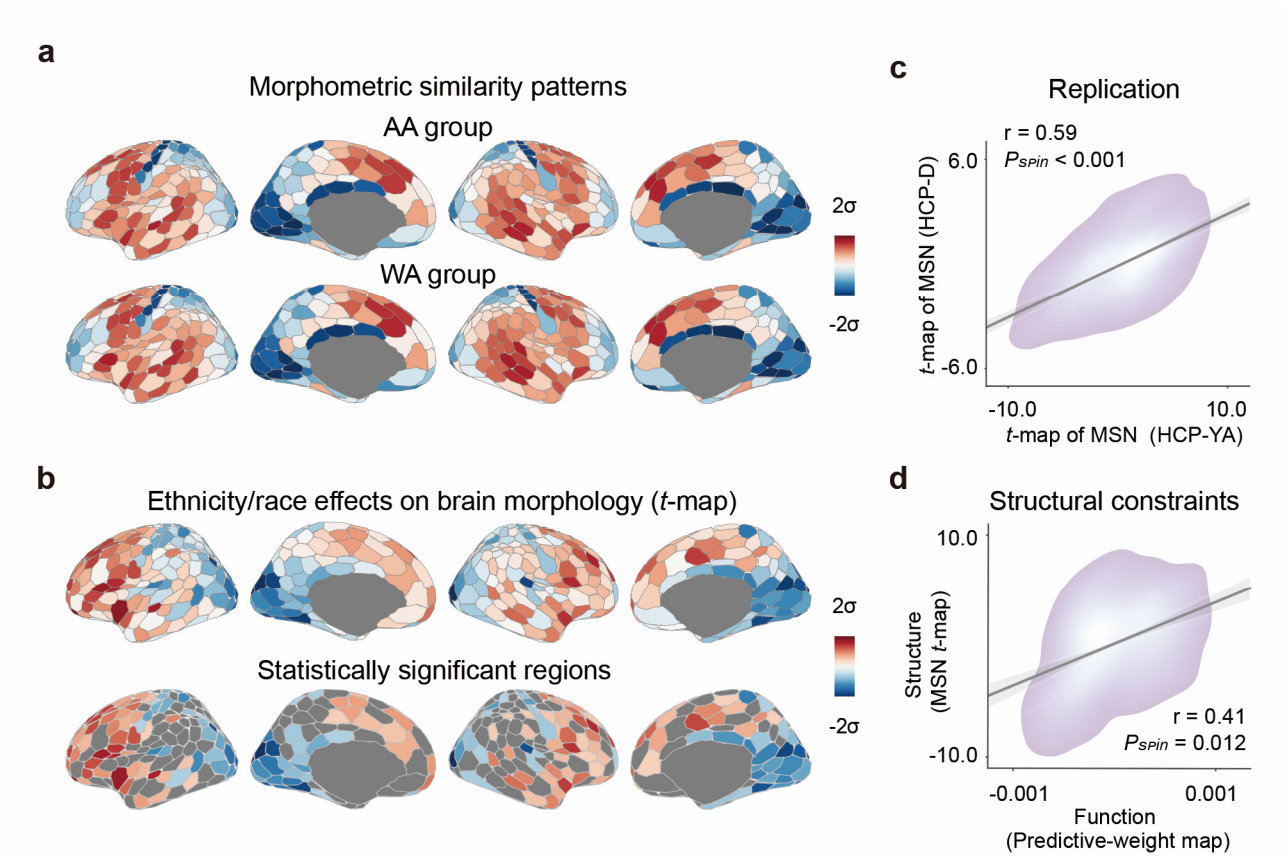

Fig. 2 Ethnicity/race-related variability in functional connectivity profiles conforms with ethnicity/race-related variations in cortical morphometry. a, Cortical distribution pattern of the mean regional morphometric similarity network (MSN) values, which describe the inter-regional similarity across multiple MRI morphometric features (i.e., cortical thickness, gray matter volume, surface area, Gaussian curvature and mean curvature). The obtained brain maps were visualized using the Schaefer400 parcellations[45] and created with the R package ggseg[50]. **b,** Comparison (*t*-map) across ethnic/racial groups of the regional MSNs. Linear regression models (LRMs) were fitted to estimate the ethnicity/race-related variations in the MSN, with age, sex, RMS, ICV, education and income employed as covariates. False discovery rate (FDR) correction with $q < 0.05$ for multiple comparisons across 400 regions revealed statistically significant ethnicity/race-related effects on cortical morphometry. **c,** The *t*-map for regional MSN differences is highly reproducible across the two datasets ($r = 0.59$, $P_{spin} < 0.001$). **d,** Ethnicity/race-related variability in the cortical morphometric similarity can explain parts of the variance in the patterns of ethnicity/race-relevant variability in functional connectivity profiles ($r = 0.41$, $P_{spin} = 0.012$).

Lifestyle influences ethnicity/race-related variability in brain functional connectome

Lifestyle factors, including education, alcohol consumption, smoking, sleep, physical activity, and social connection, have been shown to influence brain structure and function[30, 31, 56, 57]. After revealing the ethnicity/race-related functional topography, RSFC architecture, and MSN variability, we next aimed to identify the shared brain-network features underlying ethnicity/race and various lifestyle factors. We selected 6 lifestyle factors (education, income, substance use, sleep, physical activity, and social relationships) from the HCP-YA dataset commonly associated with health problems. We did not conduct this analysis for the HCP-D dataset because the majority of lifestyle data were only available for a subset of participants (e.g., those under age 17), which would significantly reduce the selected AA sample size. Specifically, we found that education ($r = -0.15$, $P < 0.001$), income ($r = -0.23$, $P < 0.001$), substance use ($r = -0.08$, $P < 0.05$), sleep ($r = 0.11$, $P < 0.01$), physical activity ($r = -0.18$, $P < 0.001$) and social relationships ($r = -0.08$, $P < 0.05$) are significantly associated with ethnicity/race categories in the HCP-YA dataset. We then used KRR models to predict each lifestyle measure on the basis of RSFC matrices. All lifestyle factors could be predicted significantly better than chance, with prediction accuracies of 0.19 ± 0.01 for education, 0.11 ± 0.01 for income, 0.19 ± 0.01 for substance use, 0.09 ± 0.01 for sleep, 0.08 ± 0.01 for physical activity and 0.07 ± 0.01 for social relationships (**Fig. 3a;** Pearson's coefficient was employed as a prediction performance metric). Moreover, we quantified the similarity of Haufe-transformed network features underlying the prediction of lifestyle factors and ethnicity/race via 2-D Pearson correlation coefficients. Notably, the predictive features underlying ethnicity/race-related variability were moderately correlated with those for education ($r = -0.55$), income ($r = -0.57$), substance use (r

= -0.34), sleep (r = 0.36) and physical activity (r = -0.38) but exhibited no associations with those for social relationships (r = 0.05; **Fig. 3b**). Haufe-transformed predictive feature matrices for all the lifestyle factors are shown in **Supplementary Fig. 6.** Furthermore, leveraging data capturing the impacts of neighbourhood-level socioeconomic environments on intrinsic brain activity reported by Sydnor et al.[52], we demonstrated that the observed ethnicity/race-related variability conforms with brain-environment associations across the cortex mantle (**Fig. 3c**; r = 0.57, P_{spin} < 0.001 in the HCP-YA dataset; r = 0.41, P_{spin} = 0.002 in the HCP-D dataset), suggesting potential ethnicity/race-by-environment interaction effects on brain functional organization.

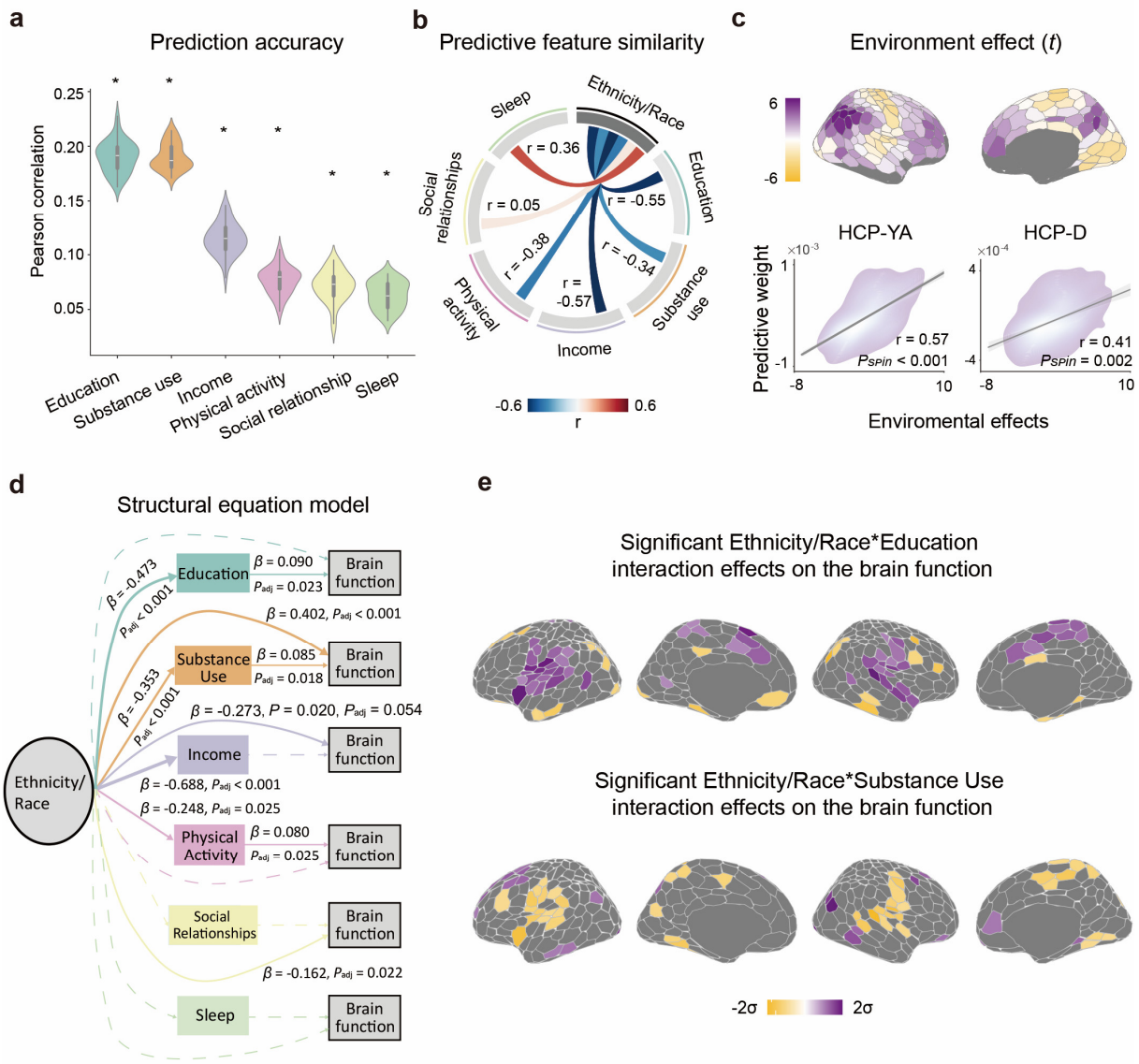

Fig. 3 Lifestyle factors mediate the associations between ethnicity/race and brain functional connectivity. a, Mean prediction accuracies for lifestyle factors. The asterisks indicate above-chance predictions (P < 0.05), determined by comparing the true accuracy with that of 1,000 permuted null models. The violin plots show accuracy distributions across 40 replications. For each violin plot, the width represents the density of the accuracy values, and the central box plot shows the interquartile range and median. **b,**

Correlations of predictive feature matrices between ethnicity/race and lifestyle factors, indicating the shared brain-network features of ethnicity/race and lifestyle factor predictions. **c,** The functional connectivity-based ethnicity/race predictive features are significantly associated with the effects of neighbourhood socioeconomic conditions on the intrinsic activity of the brain (HCP-YA dataset: $r = 0.57$, $P_{spin} < 0.001$; HCP-D dataset: $r = 0.41$, $P_{spin} = 0.002$). **d,** A structural equation model was fitted for each lifestyle factor, thereby revealing the relationships between ethnicity/race, lifestyle and the principal functional gradient (with the top 15 regions associated with the lifestyle factors selected), while controlling for age, sex, RMS, ICV, education and income. The β values represent the regression coefficients in the SEM, whereas the edge thickness scales with the magnitude of the absolute value of these coefficients. Significant links ($P < 0.05$, two-sided) are represented by solid lines, whereas non-significant links are represented by dashed lines. All the significant paths passed the FDR correction ($q < 0.05$). The mediation analysis results revealed that education (absolute mediation proportion = 14.57%) and substance use (absolute mediation proportion = 8.07%) significantly mediated the relationship between ethnicity/race and brain function, with FDR-corrected P values of 0.030 and 0.023, respectively. **e,** Threshold β-maps of the linear regression models (LRMs) showing statistically significant ($P_{FDR} < 0.05$) ethnicity/race–education and ethnicity/race–substance use interaction effects on the principal functional gradient.

Next, we fitted a structural equation model (SEM) for each lifestyle factor to analyse the relationships between ethnicity/race group, lifestyle and the principal functional gradient, thereby controlling age, sex, RMS, ICV, education and income as covariates (**Fig. 3d**). The principal functional gradient captures the dominant patterns of cortical variability in functional connectivity, thus providing a simplified, low-dimensional representation of the functional organization of the brain in a hierarchical, graded manner[58]. The regression coefficients in the converged SEM suggested that ethnicity/race category was a significant predictor for education ($\beta = -0.473$, $P_{FDR} < 0.001$), income ($\beta = -0.688$, $P_{FDR} < 0.001$), substance use ($\beta = -0.353$, $P_{FDR} < 0.001$) and physical activity ($\beta = -0.248$, $P_{FDR} = 0.025$). Education ($\beta = 0.090$, $P_{FDR} = 0.023$), substance use ($\beta = 0.085$, $P_{FDR} = 0.018$) and physical activity ($\beta = 0.080$, $P_{FDR} = 0.025$) were significant predictors for brain function (the first principal gradient of the original RSFC profiles). Subsequent mediation analyses revealed that education (mediation proportion = 14.57%, $P_{FDR} = 0.030$) and substance use (mediation proportion = -8.07%, $P_{FDR} = 0.023$) significantly mediated the associations between ethnicity/race category and brain function. Other mediation effects, including income ($P_{FDR} = 0.646$), physical activity ($P_{FDR} = 0.210$), sleep ($P_{FDR} = 0.279$) and social relationships ($P_{FDR} = 0.486$), were not statistically significant after false discovery rate (FDR) correction.

Given the mediating effects of education and substance use, we further investigated how the interactions between these factors and individuals' ethnicity/race affect brain functional connectome across the cortical mantle via the application of LRMs (**Fig. 3e**). We identified 91 cortical regions significantly modulated by ethnicity/race–education interactions, primarily located in the insula,

superior temporal, inferior parietal and superior frontal regions. Additionally, we identified 78 cortical regions significantly influenced by ethnicity/race–substance use interactions, including regions in the insula, superior temporal, rostral middle frontal, and superior frontal cortices.

Cortical gene expression related to ethnicity/race-induced variability in brain functional connectome

After demonstrating the influence of lifestyle factors on ethnicity/race-related variability in RSFC profiles, we next aimed to identify the transcriptional signatures associated with this variability. We obtained brain-wide gene expression data from the AHBA (http://human.brain-map.org)[59, 60] and divided them from the original 3,702 distinct microarray samples into 400 region-level parcellations (400 regions × 15,632 gene expression data). Via the use of partial least squares (PLS) regression[61], an advanced approach for determining the fundamental relationships between two high-dimensional matrices, we identified weighted cortical gene expression patterns associated with the spatial distribution of ethnicity/race-predictive features. The first component (PLS1), which captures the most significant pattern linking gene expression data (predictors) to predictive feature weights (responses), explained 40.3% ($P_{perm} < 0.001$) of the total variance in the response variables in the HCP-YA dataset and 22.9% ($P_{perm} < 0.001$) in the HCP-D dataset.

We observed that the PLS1-weighted maps, reflecting an anterior-to-posterior hierarchy along the cortex, were remarkably similar across the two datasets ($r = 0.99$, $P_{spin} < 0.001$) and were spatially correlated with the ethnicity/race predictive weight maps (**Fig. 4a**; $r = 0.65$, $P_{spin} < 0.001$). These positive correlations indicate that genes positively (or negatively) weighted on PLS1 are overexpressed in regions with positive (negative) ethnicity/race-predictive weights. We then ranked the genes on the basis of their normalized PLS1 weights and selected those with the top 2,000 positive weights ($Z > 6.7$ in the HCP-YA dataset and $Z > 4.5$ in the HCP-D dataset), denoted as the PLS+ gene sets, as well as the top 2,000 negative weights ($Z < -6.5$ in the HCP-YA dataset and $Z < -4.6$ in the HCP-D dataset), denoted as the PLS- gene sets (**Fig. 4b and Supplementary Fig. 7**).

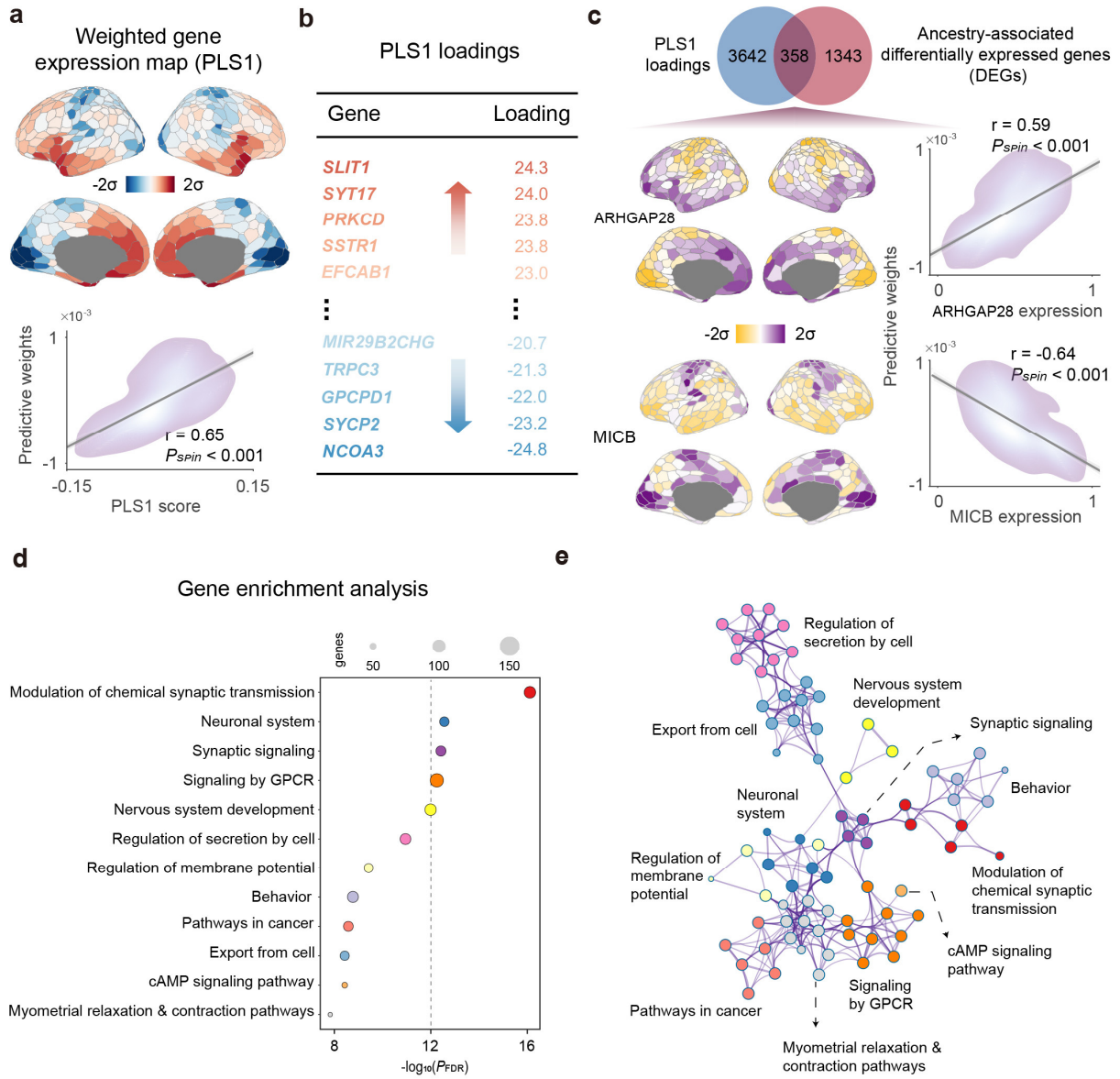

Fig. 4 Gene expression profiles associated with ethnicity/race-related variability in functional connectivity. a, A weighted cortical gene expression map of the regional PLS1 scores (top panel). Scatterplot showing the relationship between the weighted cortical gene expression map (PLS1) and the Haufe-transformed predictive weight map (bottom panel, $r = 0.65$, $P_{spin} < 0.001$). The obtained brain maps were visualized using the Schaefer400 parcellations[45] and created with the R package ggseg[50]. **b,** Genes were ranked on the basis of their loadings in PLS1. **c,** The PLS+ gene list overlapped with the ancestry-associated differentially expressed genes (DEGs) reported by Benjamin et al[39]. **d,** Representative enriched terms of the PLS1+ genes, highlighting GO biological processes related to synaptic signalling and Reactome gene sets associated with the nervous system. The size of the circle denotes the number of genes within the given ontology term, and the multi-test FDR-adjusted *P* values are plotted as log10-transformed values. **e,** Metascape network plot of the enriched terms, capturing the intra-cluster and inter-cluster similarity relationships between them. Each term is represented by a circle, coloured by its cluster identity and scaled proportional to the number of genes involved.

We next conducted an external validation by investigating the relationships between the PLS1 genes and DEGs identified from RNA sequencing data for postmortem brain tissue. Benjamin et al. reported 1,701 DEGs whose expression levels are influenced by genetic ancestry[39]. We noted that among these DEGs, 358 overlapped with the PLS1 genes in the HCP-YA dataset, 358 overlapped with the PLS1 genes in the HCP-D dataset, and 309 overlapped with the PLS1 genes in both datasets (**Fig. 4c** and **Supplementary Fig. 7**), suggesting specificity in the contribution to transcriptional changes across ethnicity/race and genetic ancestry. Among these overlapping genes, ARHGAP28, which encodes the Rho GTPase-activating protein[62], exhibited the strongest positive correlation with the ethnicity/race-predictive weights (**Fig. 4c**; $r = 0.59$, $P_{spin} < 0.001$). In contrast, MICB, which encodes ligands that interact with the natural killer cell-activating receptor (i.e., NKG2D)[63], exhibited the strongest negative correlation (**Fig. 4c**; $r = -0.64$, $P_{spin} < 0.001$).

To interpret the PLS- and PLS+ gene sets, we conducted gene enrichment analysis to characterize the enriched biological functions and pathways within these gene lists via Metascape[64]. After employing the Benjamini–Hochberg method to correct for multiple hypothesis testing and discarding discrete enriched clusters, the remaining significantly enriched terms of the PLS+ genes were rendered into a network (**Figs. 4d–e** and **Supplementary Figs. 7**), including gene ontology (GO) biological processes, such as “modulation of chemical synaptic transmission” and “synaptic signalling”, as well as Reactome gene sets, such as “neuronal system”, “signalling by GPCR” and “nervous system development”. The PLS1- genes were enriched for GO biological processes, such as “metal ion transport” and “potassium ion transmembrane transport”, as well as Reactome pathways, including “neuronal system” and “transport of small molecules” (**Supplementary Fig. 8**). These results highlight the distinct functional roles of genes positively and negatively weighted on PLS1.

We further assessed the reproducibility of the transcriptional enrichment of ethnicity/race-related functional connectivity features. Specifically, we conducted a multi-gene list meta-analysis of the PLS1 genes (combining the PLS+ and PLS- genes) between the HCP-YA and HCP-D datasets. We found that the gene lists in the two independent datasets highly overlapped, with an odds ratio of 271.88 (**Supplementary Fig. 9,** $P < 0.001$). After statistical correction, the surviving ontological terms were highly shared between the two independent gene lists, and the resulting network comprising these terms demonstrated significant overlap with the discovery enrichment analysis results (**Supplementary Figs. 9**).

Discussion

In this study, we comprehensively investigated ethnicity/race-related variability in brain functional organization and its associations with individual lifestyle factors and transcriptomics on the basis of multimodal data from the HCP-YA and HCP-D datasets. We employed a precise brain parcellation approach to construct personalized functional networks and delineate ethnicity/race-related variations in the functional topography. By leveraging a machine learning technique, we determined that multivariate patterns of functional connectivity from sensorimotor and association cortices contributed the opposite to ethnicity/race prediction, and that these patterns were spatially correlated with the morphological variability induced by ethnicity/race. Furthermore, we observed that the shared brain-network features served as predictors for both ethnicity/race and various lifestyle factors, with education and substance use substantially mediating the associations between ethnicity/race and brain function. Finally, we characterized the transcriptomic architectures underlying the observed ethnicity/race-related variability, thereby providing biological microscale insights into brain-wide connectivity properties. Notably, the functional connectivity variability and its associations with transcription were highly reproducible across two large sample datasets for young adults (HCP-YA) and children (HCP-D), which fully demonstrates the advantages of the methodology adopted in our study and the robustness of our results, and echoes recent calls for reproducible neuroscience research. Overall, our findings promote an integrated understanding of ethnicity/race-related variations in functional topography and connectivity, as well as their associations with lifestyle factors and transcriptional signatures.

Both the spatial configuration of functional regions and the coupling strength between regions are important drivers of variability in functional organization, which is closely linked to non-imaging measures of behaviour and lifestyle[65]. Here, we observed that both the functional topography and connectivity are associated with individuals' ethnicity/race. First, the spatial variation in functional topography across ethnicity/race groups was greater in the heteromodal association cortices than in the unimodal sensorimotor cortices. This finding extends the results of previous studies showing that association networks exhibit the greatest variations in functional topography across age windows[66], sex[67], and cultural populations[12]. Second, multivariate patterns of brain-ethnicity/race relationships were inversely driven by the connectivity profiles of higher-order association regions and modality-specific regions. This pattern indicates that ethnicity/race-related sociodemographic factors exert widespread and opposite directional effects on the functional connectivity of the two poles of the S–A axis[52]. Additionally, we observed that the ethnicity/race-

related differences in functional connectivity profiles conform with the variations in morphometric similarity across the different ethnic/racial groups. This result suggests that the functional connectivity differences induced by ethnicity/race are associated with brain morphological properties, conforming with the well-established concept of cortical structural constraints on intrinsic neuronal activity[16, 68] and supporting the general homophilic principle of brain organization. According to this principle, if two brain regions are similar in terms of a particular biological or functional characteristic, they typically exhibit similarities across other dimensions and are more likely to be connected by white matter tracts[69].

Given the multidimensional properties of ethnicity/race, a comprehensive perspective is needed to establish a connection between ethnic/racial diversity and the hierarchical functional organization of the brain[14]. Recent compelling evidence has demonstrated that the effects of children's neighbourhood environments on the fluctuation amplitude of the intrinsic activity of the brain are opposite in the sensorimotor and association cortices[52]. Socioeconomically disadvantageous environments contribute to diminished differentiation of the S–A hierarchy[52]. In this study, we demonstrated that the environmental effects are spatially aligned with ethnicity/race-related variability, suggesting overlap of ethnicity/race and various socioenvironmental factors. Moreover, prior studies have suggested that multiple environmental and lifestyle factors are reflected in the brain structure and function, further influencing mental health[30, 31, 70]. Here, we provided evidence that shared functional network features conform with lifestyle factors and ethnicity/race categories. We found that the control, default and somatomotor networks were closely associated with the education level and ethnicity/race of individuals. The connectivity of the control–somatomotor networks, default–ventral attention networks, and somatomotor networks was highly predictive of both the number of years of education and individuals' ethnicity/race. In addition, the connectivity within the default networks, between the default and dorsal attention networks, and between the default and ventral attention networks was highly associated with both the substance use experiences and ethnicity/race of individuals. Indeed, behaviour prediction models may capture broad behaviour within a given behavioural domain rather than the specific behaviours they are trained on, such as the *g* factor in the case of cognition and the *p* factor in the case of mental health[49]. Thus, the observed similarity in the predictive network features suggests that ethnicity/race and lifestyle factors may be supported by shared underlying meta-factors[71].

Moreover, we found that individual-level lifestyle factors (i.e., education level and substance use) significantly mediated the relationships between the functional connectivity profiles and ethnic/racial categories. The principal functional gradient of regions located in the insula, ventral part of the sensory/motor cortex and prefrontal cortex was significantly influenced by the interactions between ethnicity/race and education, as well as by the interactions between ethnicity/race and substance use. These regions serve as crucial control hubs of top-down information processing[72] and partially overlap with the threat neurocircuitry, which fulfils an important role in emotional responses[73, 74]. Notably, both education and substance use are important components in the pathways linking the environment and health status[30]. As a dimension of the socioeconomic position, prior studies have revealed ethnicity/race-related differences in education[75, 76], and our results similarly indicate a linear association between ethnicity/race and education in the HCP-YA dataset. Educational background is often linked to cognitive reserves[77], and higher levels of education may enhance the flexibility and adaptability of the brain during task processing and decision-making[78, 79]. Moreover, a higher education level is associated with greater integrity of brain connectivity in elderly individuals[80, 81], thus providing protection against cognitive decline[82] and the risk of neurogenerative diseases[83, 84]. Substance use rates also differ across ethnic/racial populations, especially marijuana and alcohol use[19, 85]. Differences in exposure to these substances and barriers to assessing substance use care may be shaped in part by sociohistorical factors. Excessive substance use alters the neurobiology of reward processing, control dynamics and emotion regulation[86-88] and is associated with psychiatric conditions, particularly depression symptoms, and reduced quality of life[89-92]. Therefore, fully accounting for socioenvironmental disadvantages may lead to greater understanding of the differences in neuroimaging-derived features across populations and the formulation of more generalizable treatment and intervention strategies. Notably, since many developmental processes, such as those governing plasticity-restricting features such as myelination, are sensitive to exogenous environmental stimuli (e.g., negative experiences)[93, 94], the observed global pattern of ethnicity/race-related variability in functional organization may reflect a close interaction between neurodevelopmental malleability and environmental influences. Finally, the lifestyle factors in our study were relatively coarse measures of individual-level living environments, and the mediation effects accounted for approximately one-quarter of the overall ethnicity/race-related variability. Future work is needed to combine more comprehensive lifestyle (e.g., access to healthcare) and neighbourhood-level environment measures (e.g., population density or percentage of residents living in poverty) to provide a detailed explanation for ethnicity/race-

480 related variability in brain–behaviour interactions and promote the equitable translation of
481 neuroscience findings.

482 After linking functional network properties to transcriptomic data, we identified a set of genes
483 (PLS1) whose weighted expression patterns across the cortex were significantly correlated with
484 ethnicity/race-related variability in connectivity. The PLS1-weighted gene expression map revealed
485 high expression in the primary visual cortex, anterior cingulate cortex, and insular cortex, following
486 the global cortical hierarchy pattern of the brain. Gene-to-function annotation analysis revealed that
487 these genes formed a topologically clustered gene network enriched for several GO biological
488 processes and Reactome pathways in both the HCP-YA and HCP-D cohorts. The identified GO
489 biological processes in both datasets were related primarily to synaptic signalling and the modulation
490 of chemical synaptic transmission, which form the foundation of inter-neuronal communication in
491 the nervous system[95]. Synaptic signalling pathway enrichment is not unexpected: notably, a
492 previous study revealed population differences in gene expression associated with axon guidance and
493 synaptic signalling pathways[40]. The identified Reactome pathways include the neuronal system
494 and its development, which encompass key pathways involving the dynamically regulated processes
495 of neuronal progenitor cell proliferation, migration, differentiation, synaptogenesis, and synaptic
496 pruning that collectively contribute to the establishment of a mature and functionally integrated
497 system[96, 97]. Neurodevelopment-related pathways are also widely documented to fulfil critical
498 roles in epigenetic mechanisms influenced by environmental and lifestyle factors, such as
499 socioeconomic status[98], alcohol consumption[99], early-life stress[100] and physical activity[101,
500 102]. Another important Reactome pathway is G protein-coupled receptor signalling, which mediates
501 most cellular responses to hormones, neurotransmitters, and external environmental stimulants[103].
502 Overall, our results bridge the gap between ethnicity/race-related variability in brain connectivity
503 patterns and transcriptional signatures.

504 We further found that the PLS1 genes slightly overlapped with the ethnicity/race-associated DEGs
505 from postmortem brain tissues reported by Benjamin et al[39]. Among these overlapping genes,
506 ARHGAP28 and MICB—ARHGAP family genes involved in the regulation of cellular
507 signalling[104] and MICB involved in immune function[63, 105]—are mostly spatially coupled with
508 ethnicity/race-related variability in brain connectivity patterns. Since genetic variations (explaining
509 60% of the changes) rather than environments (explaining 15% of the changes) drive genetic
510 ancestry expression changes, no substantial overlap between the MRI-derived PLS1 genes and

511 ancestry DEGs is thus expected[39]. Therefore, our findings emphasize the differences between
512 ethnicity/race social categories and genetic ancestry[18, 106].

513 Our study exhibits several limitations. First, we employed correlation analysis to describe the
514 relationships between functional and structural differences across ethnicity/race groups. It could be
515 meaningful to develop models that can be used explore the causality underlying these differences,
516 which would help clarify the causal structure of the brain. Second, this study is limited to WA and
517 AA participants who self-reported a single ethnic/racial identity but did not account for cultural and
518 geographical factors that may influence the brain. Therefore, it may be challenging to fully
519 generalize the findings across all ethnicity/race groups. Third, considering the multiple differences
520 (e.g., scanner hardware, protocols of the imaging sequences, and length of the rs-fMRI data) between
521 the HCP-YA and HCP-D datasets, we did not investigate the implications of ethnicity/race-related
522 variability changes between children and young adults and across developmental stages. Instead,
523 these two datasets served as independent means to validate our results, and future research can focus
524 on investigating variability in brain functional development across ethnic/racial populations using
525 longitudinal datasets, such as the Adolescent Brain Cognitive Development (ABCD) dataset. Fourth,
526 the AHBA gene expression data were measured from six postmortem brains (one Hispanic, two AA,
527 and three Caucasian specimens), and only two of these donors provided data for the right
528 hemisphere, which limits the effect size and generalizability of the transcriptome–neuroimaging
529 associations reported in our study. Finally, although we demonstrated the impacts of lifestyle and
530 transcriptional signatures on ethnicity/race-related variability in brain connectivity, future work
531 should involve the utilization of individual-level gene expression data, along with more
532 comprehensive lifestyle and environmental characterizations, to examine the proportion of
533 differences that can be explained by each factor. Regardless of these limitations, we provided three
534 perspectives for understanding the impact of ethnicity/race-related factors on brain function, namely,
535 brain structure, lifestyle and the environment, and the transcriptome. In line with recent calls, we
536 argue against the direct causal attribution of the observed differences across populations to
537 ethnicity/race and emphasize the multidimensional attributes of ethnicity/race.

Methods

Datasets. Two publicly available datasets from the HCP were used, namely, the HCP-YA and HCP-D datasets. The HCP-YA participants (N = 822; age: 22-37 years) were part of the HCP S1200 release[107] and matched the HCP subset used by Li et al[14]. Moreover, rs-fMRI data were obtained during two sessions using a multiband sequence (voxel dimension: 2.0 mm isotropic; repetition time: 0.72 s) on a customized Siemens 3-T Skyra at Washington University in St. Louis. Each session comprised two runs with opposite phase-encoding directions, and the duration of each run was approximately 14.4 minutes. Structural images were acquired for each participant with a resolution of 0.7 mm isotropic. Family structures and self-reported ethnicity/race information were carefully considered in subsequent processing. HCP-D participants (N = 472; age: 6-21 years) were obtained from the lifespan HCP release 2.0[108]. For participants 8 years and older, rs-fMRI data were obtained during two sessions using a multiband sequence (voxel dimension: 2.0 mm isotropic) on a 3T Siemens Prisma scanner. Each session encompassed two runs with opposite phase-encoding directions, and the duration of each run was approximately 6.5 minutes. For the youngest participants (5-7 years), the total duration of rs-fMRI scanning was reduced to 21 min. All participants aged 18 years and above provided informed consent themselves, whereas participants younger than 18 years were accompanied by parents who provided informed permission for their participation. Both the HCP-YA and HCP-D datasets were approved by the Institutional Review Board (IRB) at Washington University in St. Louis. Additional details of the data collection process and behavioural measures can be found elsewhere. All procedures in this study were conducted following the applicable guidelines and regulations, with approval from the IRB at the Beijing Institute of Technology.

Image acquisition and processing. ICA-FIX denoised rs-fMRI data on top of the minimal preprocessing steps of the HCP were employed[109]. To increase the behavioural prediction performance, additional motion censoring and global signal regression were performed following the procedure described by Li et al[14]. Specifically, volumes with an RMS exceeding 0.2 mm were marked as censored frames. The frames before and two frames after these volumes, as well as any previously uncensored segments shorter than five volumes, were also marked as censored frames. Rs-fMRI runs were discarded if more than half of the frames were marked as censored. Next, the signal of the remaining uncensored frames was averaged across the cortical vertices as the global signal. The global signal and its first temporal derivative were regressed out from the original time courses.

570 **Defining individual-specific cortical parcellations and the functional connectome.** The gradient
571 MS-HBM approach was employed to define individual-specific cortical parcellations for each
572 participant
573 (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Kong2022_
574 ArealMSHBM)[44]. The steps of the MS-HBM approach are as follows: (i) The Connectome
575 Workbench (https://github.com/Washington-University/workbench) is utilized to compute each
576 individual's diffusion embedding matrices for gradients; (ii) connectivity profiles for each seed
577 region, initialization parameters and a probability spatial mask are generated; (iii) group-level priors
578 are estimated for each ethnicity/race group on the basis of the Schaefer400 parcellations[45] and 40
579 independent participants; and (iv) Bayesian model estimation is performed for the generation of
580 individual-level parcellations. The mathematical details of the MS-HBM approach are described
581 elsewhere[44]. We used individualized and Schaefer400 parcellations to compute the RSFC. The
582 vertex-wise time series, excluding the censored frames, were averaged across all the regions of
583 interest (ROIs). For each run, the RSFC matrix was obtained across 400 regions using Pearson's
584 correlation analysis. The RSFC matrices were then Fisher *z*-transformed and averaged across runs,
585 yielding two final 400 × 400 RSFC matrices for each participant.

586 **Predictive modelling.** AA participants from the HCP-YA and HCP-D datasets were randomly split
587 into 5 folds, and participants from the same family were kept together in the same fold. Hungarian
588 matching was applied to select demographically matched AA–WA pairs on the basis of age, sex and
589 RMS within each fold[12, 14]. The procedures for data splitting and matching were performed 40
590 times to obtain generalized results.

591 We applied 5-fold cross-validated KRR models to predict ethnicity/race and various lifestyle factors
592 (**Supplementary Table 1**). To facilitate comparisons of functional brain correlates across
593 ethnicity/race and lifestyle factors, the linear regression framework was implemented for
594 ethnicity/race predictions instead of a classification approach. KRR models aim to predict
595 behavioural measures of the test participants by leveraging the similarity of their RSFC patterns to
596 those of the training participants. Specifically, the prediction is obtained as a weighted average of the
597 behavioural scores for all training participants, where the weights are determined by the value of
598 Pearson's correlation between the vectorized RSFC matrices of the test participants and each training
599 participant. For each test fold, kernel regression parameters were estimated using data from the four
600 training folds. The optimal *l2* regularization parameter (λ) was selected through 5-fold cross

601 validation of the training data. All the determined parameters were subsequently utilized for the
602 predictions of the participants within the test set. To mitigate the potential sensitivity of a single 5-
603 fold cross-validation step to data splitting variations, the entire cross-validation process was repeated
604 40 times. Additionally, for each predicted lifestyle factor, all related measures were subjected to *Z*-
605 normalization across participants and then averaged to obtain a composite score for each individual
606 before being entered into the model. When variables other than education and income were
607 predicted, age, sex, RMS, ICV, education, and income were regressed out from the RSFC. In
608 contrast, when predicting education and income, only age, sex, RMS, and ICV were regressed out
609 from the RSFC.

610 For the ethnicity/race prediction models, the accuracy of each test set was calculated as the
611 proportion of correctly predicted samples to the total number of samples. For the lifestyle factor
612 prediction models, the accuracy of each test set was determined via Pearson's correlation analysis.
613 The accuracy measures were averaged across 5 folds for each data split, yielding 40 prediction
614 accuracy values. To evaluate the prediction significance, a corresponding set of null models was
615 generated for each predictive model. Specifically, the output variable was permuted 1,000 times. For
616 each permutation, a null model was trained and tested following the same procedures applied to the
617 original model. The *P* value of the significance of each model is defined as the proportion of null
618 models whose prediction accuracies exceed those of the true models.

619 **Predictive feature weights.** Haufe transformation was used to interpret the brain–behaviour
620 relationships learned by the prediction models[51]. For each training fold, Haufe-transformed feature
621 weights were calculated as the covariance between the input features ($RSFC_{\text{train}}$) and the output
622 predicted scores (y'_{train}). The feature weight matrices were averaged across the 40 splits to obtain the
623 final importance measures (400 × 400), reflecting the contribution of each edge to the prediction. We
624 subsequently summarized the pairwise Haufe-transformed weights to the network level (17 × 17). In
625 addition, we calculated their averages at the regional level (400 × 1). In this study, we did not obtain
626 the absolute value of the feature weights and preserved their positive or negative signs[110].

627 **Morphometric similarity analysis.** The Scharfer400 parcellation was transformed into the T1w
628 surface space of each participant, and five morphometric features were extracted from the
629 participant's T1w images for each region, including cortical thickness, gray matter volume, surface
630 area, Gaussian curvature and mean curvature[37]. For each participant, each morphometric feature
631 was subjected to *Z*-normalization across regions, and a feature vector (5 × 1) was created for each

632 region. Pearson's correlation analysis was then employed to estimate the morphometric similarity
633 between cortical regions, yielding a pairwise MSN (400 × 400) for each participant. The full MSN
634 matrices were then averaged across rows to generate a regional MSN map.

635 To visualize the spatial pattern of the MSN across the cortex, the regional MSN maps were averaged
636 across participants within each ethnicity/race group. To examine the ethnicity/race-related variations
637 in the MSN, an LRM was fitted for each region, thereby employing the regional MSN value as the
638 response variable and age, sex, RMS, ICV, education and income as covariates. Weights were
639 applied to account for differences in group sizes, with each participant assigned a weight equal to the
640 reciprocal of the size of their corresponding group. Two-sided *t* statistics derived from the LRMs
641 were extracted, and significance was determined at $P < 0.05$ with the application of FDR correction
642 for multiple comparisons across 400 regions to control for Type I errors.

643 **Structural equation modelling.** The *lavaan* package (version 0.6.18) in R was used to fit an SEM
644 for each lifestyle factor and estimate regression coefficients of each path while controlling for age,
645 sex, RMS and ICV[30, 31]. The complete model equations are provided in the Supplementary
646 Materials. The latent variable for brain function was derived from the participants' principal
647 functional gradient. In this study, we applied principal component analysis (PCA) to the RSFC
648 matrices of the participants for dimensionality reduction, and the first principal component was
649 identified as the principal functional gradient. The *Z*-normalized gradient values of the first 15
650 cortical regions, which were significantly correlated with each lifestyle factor, were represented as
651 the latent variable for brain function[31]. Before being entered into the model, we subjected the
652 RMS, ICV and lifestyle factors to *Z*-normalization to avoid scale effects and facilitate coefficient
653 comparisons across models. The *P* value for each path was calculated by dividing the estimated
654 parameter by its robust standard error (SE) and then testing the null hypothesis that the parameter
655 equals zero. The *P* values of all paths within each SEM were subjected to FDR correction to account
656 for multiple comparisons. We further calculated the mediation effects (a*b) of each lifestyle factor
657 and the proportion of ethnicity/race-related variability mediated by the lifestyle factors
658 (a*b/(a*b+c)). Within each SEM, the FDR was also controlled at $q < 0.05$ across the mediation
659 effects, direct effects and total effects.

660 **Interaction effects modelling.** To examine how the ethnicity/race-related variability in brain
661 connectivity is moderated by the education level and substance use status of participants, we fitted
662 two LRMs for each region. These models controlled for age, sex, RMS and ICV and included a term

for the interaction between ethnicity/race and education or substance use[111]. The response terms of these models were the regional principal functional gradient values, namely, the principal components derived from RSFC matrix dimensionality reduction using PCA. Moreover, two-sided *t* statistics of the interaction terms within the LRMs were extracted, and significance was determined at $P < 0.05$ with the application of FDR correction for multiple comparisons across 400 regions to prevent Type I errors. The complete model equations are provided in the Supplementary Materials.

Transcriptomic analysis. We utilized gene expression data from the AHBA dataset, which were derived from six postmortem brains (age: 24 to 57 years; 1 female) with 3702 spatially distinct samples. The preprocessing steps were performed according to the approach of Markello et al.[112] These steps are as follows: (i) the probe-to-gene annotations were updated, thereby performing intensity-based filtering with a threshold of 0.5, selecting probes with the most consistent patterns of regional variation across donors, and aggregating probes to genes; (ii) gene expression samples were mapped to the Schaefer400 parcellation, in which a bidirectional mirroring approach across the left and right hemisphere was applied, and centroid-based imputation was employed in cases of missing data; (iii) the expression values of each sample were normalized across the genes for each donor, and the expression values of each gene were normalized across the samples for each donor; and (iv) the samples were aggregated across donors to generate expression values for a given region. Thus, 15,632 genes remained, and a gene expression matrix (400 × 15,632) was used for subsequent analyses[113].

PLS regression was employed to relate the Haufe-transformed weights of ethnicity/race prediction and the transcriptomic measurements of all 15,632 genes[37]. In the PLS model, *Z*-normalized gene expression data were adopted as the predictor variable, and *Z*-normalized predictive weights were employed as the response variables. PLS1 provides the optimal low-dimensional representation of gene expression profiles that are mostly correlated with ethnicity/race-related variability in brain functional connectivity. The statistical significance of the variance explained by PLS1 was assessed by spatially permuting the response variables 1,000 times. Finally, bootstrapping was applied to quantify the estimation errors in the PLS1 weight of each gene, and *Z* scores were calculated by dividing the weight of each gene by its bootstrap SE, enabling the ranking of genes on the basis of their PLS1 contributions.

In both the HCP-YA and HCP-D datasets, the top 2,000 positive genes and 2,000 negative genes were selected on the basis of their Z values to form the PLS1+ and PLS1- gene sets, respectively.

694 Each gene set was then submitted separately to the Metascape website, an automated functional
695 pathway enrichment tool that provides biological insights into the given genes. To validate the
696 enrichment results, the PLS genes from the HCP-YA and HCP-D datasets were then entered into
697 Metascape together to conduct a multi-gene list meta-analysis. This approach facilitates the
698 understanding of pathways that are shared across gene lists or selectively ascribed to specific gene
699 lists. All identified pathways were considered significant at $P < 0.05$ after FDR correction ($q < 0.05$)
700 for multiple comparisons.

701 **Statistical analysis.** In this study, cortical data were related to each other across parcels using two-
702 sided Pearson's correlation analysis. To prevent overestimation of the significance of correlations,
703 1,000 spin test permutations were performed to generate a null distribution for each cortical feature
704 by randomly rotating the feature maps, thereby controlling for distance-dependent spatial
705 autocorrelation. Permutation *P* values were obtained by counting the number of times that the null
706 correlation exceeded (for positive correlations) or fell below (for negative correlations) the observed
707 correlation, divided by the total number of spatial rotations. Spin tests were implemented using the
708 Neuromaps toolbox (https://github.com/netneurolab/neuromaps).

709 Data availability

710 The raw and preprocessed data are available from the HCP
711 (https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release;
712 https://humanconnectome.org/study/hcp-lifespan-development). The subject lists for both the HCP-
713 YA and HCP-D datasets are available on our GitHub page
714 (https://github.com/TianyiYanLab/Ethnicity_Race_Diversity). Human gene expression data are
715 available from the Allen Brain Atlas (https://human.brainmap.org). In our analyses, we also utilized
716 publicly available cortical atlases, including the Schaefer400 atlas
717 (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer20
718 18_LocalGlobal/Parcellations), the S–A axis (https://github.com/PennLINC/S-A_ArchetypalAxis),
719 and environmental effect maps (https://doi.org/10.5281/zenodo.7606653). The PLS1 gene lists and *Z*
720 score weights for both cohorts are provided in Supplementary Data 1. All data supporting the
721 findings of this study are included in the paper and its Supplementary Information, and all additional
722 information is available from the authors upon reasonable request.

Code availability

All the codes used to generate the results are available on GitHub (https://github.com/TianyiYanLab/Ethnicity_Race_Diversity).

Acknowledgements

The authors thank the HCP team and Dr. Zaixu Cui for their support and help. This work was supported by the National Natural Science Foundation of China (grant numbers 62336002, 62406025 and 82302175); the STI 2030-Major Projects (grant number 2022ZD0208500); the Key-Area Research and Development Program of Guangdong Province (grant number 2023B0303030002); the National Science and Technology Innovation 2030 Program (grant number 2021ZD0200500); and the Beijing Nova Program (grant number 20230484465).

References

1. Kopal J, Uddin LQ, Bzdok D. The end game: respecting major sources of population diversity. Nat Meth. 2023;20(8):1122-1128.
2. Ricard J, Parker T, Dhamala E, Kwasa J, Allsop A, Holmes A. Confronting racially exclusionary practices in the acquisition and analyses of neuroimaging data. Nat Neurosci. 2023;26(1):4-11.
3. Harnett NG, Merrill LC, Fani N. Racial and ethnic socioenvironmental inequity and neuroimaging in psychiatry: a brief review of the past and recommendations for the future. Neuropsychopharmacol. 2024:1-13.
4. Dhamala E, Ricard JA, Uddin LQ, Galea LA, Jacobs EG, Yip SW, et al. Considering the interconnected nature of social identities in neuroimaging research. Nat Neurosci. 2024:1-12.
5. Harnett NG, Fani N, Rowland G, Kumar P, Rutherford S, Nickerson LD. Population-level normative models reveal race-and socioeconomic-related variability in cortical thickness of threat neurocircuitry. Commun Biol. 2024;7(1):745.
6. Yang G, Bozek J, Han M, Gao JH. Constructing and evaluating a cortical surface atlas and analyzing cortical sex differences in young Chinese adults. Hum Brain Mapp. 2020;41(9):2495-2513.
7. Dumornay NM, Lebois LA, Ressler KJ, Harnett NG. Racial disparities in adversity during childhood and the false appearance of race-related differences in brain structure. Am J Psychiatry. 2023;180(2):127-138.
8. Yang G, Zhou S, Bozek J, Dong H-M, Han M, Zuo X-N, et al. Sample sizes and population differences in brain template construction. Neuroimage. 2020;206:116318.
9. Fields C, Rosenblatt M, Scheinost D. Racialized Heteroscedasticity in Connectome-Based Predictions of Brain-Behavior Relationships. PsyArXiv. 2024:2025.2001.2010.31234.
10. Harnett NG, Fani N, Carter S, Sanchez LD, Rowland GE, Davie WM, et al. Structural inequities contribute to racial/ethnic differences in neurophysiological tone, but not threat reactivity, after trauma exposure. Mol Psychiatry. 2023;28(7):2975-2984.
11. Ge J, Yang G, Han M, Zhou S, Men W, Qin L, et al. Increasing diversity in connectomics with the Chinese Human Connectome Project. Nat Neurosci. 2023;26(1):163-172.

12. Yang G, Bozek J, Noble S, Han M, Wu X, Xue M, et al. Global diversity in individualized cortical network topography. Cereb Cortex. 2023;33(11):6803-6817.
13. LeWinn KZ, Sheridan MA, Keyes KM, Hamilton A, McLaughlin KA. Sample composition alters associations between age and brain structure. Nat Commun. 2017;8(1):874.
14. Li J, Bzdok D, Chen J, Tam A, Ooi LQR, Holmes AJ, et al. Cross-ethnicity/race generalization failure of behavioral prediction from resting-state functional connectivity. Sci Adv. 2022;8(11):eabj1812.
15. Greene AS, Shen X, Noble S, Horien C, Hahn CA, Arora J, et al. Brain–phenotype models fail for individuals who defy sample stereotypes. Nature. 2022;609(7925):109-118.
16. Pang JC, Aquino KM, Oldehinkel M, Robinson PA, Fulcher BD, Breakspear M, et al. Geometric constraints on human brain function. Nature. 2023;618(7965):566-574.
17. Fotiadis P, Parkes L, Davis KA, Satterthwaite TD, Shinohara RT, Bassett DS. Structure–function coupling in macroscale human brain networks. Nat Rev Neurosci. 2024;25(10):688-704.
18. Cardenas-Iniguez C, Gonzalez MR. Recommendations for the responsible use and communication of race and ethnicity in neuroimaging research. Nat Neurosci. 2024;27(4):615-628.
19. Puka K, Kilian C, Zhu Y, Mulia N, Buckley C, Lasserre AM, et al. Can lifestyle factors explain racial and ethnic inequalities in all-cause mortality among US adults? BMC Public Health. 2023;23(1):1591.
20. Han S, Northoff G. Culture-sensitive neural substrates of human cognition: A transcultural neuroimaging approach. Nat Rev Neurosci. 2008;9(8):646-654.
21. Williams DR, Collins C. US socioeconomic and racial differences in health: patterns and explanations. Annu Rev Sociol. 1995;21(1):349-386.
22. Sudano JJ, Baker DW. Explaining US racial/ethnic disparities in health declines and mortality in late middle age: the roles of socioeconomic status, health behaviors, and health insurance. Soc Sci Med. 2006;62(4):909-922.
23. Roberts AL, Gilman SE, Breslau J, Breslau N, Koenen KC. Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. Psychol Med. 2011;41(1):71-83.
24. Noble KG, Houston SM, Brito NH, Bartsch H, Kan E, Kuperman JM, et al. Family income, parental education and brain structure in children and adolescents. Nat Neurosci. 2015;18(5):773-778.
25. Tomalski P, Moore DG, Ribeiro H, Axelsson EL, Murphy E, Karmiloff‐Smith A, et al. Socioeconomic status and functional brain development – associations in early infancy. Dev Sci. 2013;16(5):676-687.
26. Carter SE, Ong ML, Simons RL, Gibbons FX, Lei MK, Beach SR. The effect of early discrimination on accelerated aging among African Americans. Health Psychol. 2019;38(11):1010.
27. Fani N, Carter SE, Harnett NG, Ressler KJ, Bradley B. Association of racial discrimination with neural response to threat in Black women in the US exposed to trauma. JAMA Psychiatry. 2021;78(9):1005-1012.
28. Firth J, Solmi M, Wootton RE, Vancampfort D, Schuch FB, Hoare E, et al. A meta‐review of "lifestyle psychiatry": the role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. World Psychiatry. 2020;19(3):360-380.
29. Murray CJ, Aravkin AY, Zheng P, Abbafati C, Abbas KM, Abbasi-Kangevari M, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1223-1249.
30. Tian YE, Cole JH, Bullmore ET, Zalesky A. Brain, lifestyle and environmental pathways linking physical and mental health. Nat Ment Health. 2024:1-12.

31. Zhao Y, Yang L, Sahakian BJ, Langley C, Zhang W, Kuo K, et al. The brain structure, immunometabolic and genetic mechanisms underlying the association between lifestyle and depression. Nat Ment Health. 2023;1(10):736-750.
32. Richiardi J, Altmann A, Milazzo A-C, Chang C, Chakravarty MM, Banaschewski T, et al. Correlated gene expression supports synchronous activity in brain networks. Science. 2015;348(6240):1241-1244.
33. Fornito A, Arnatkevičiūtė A, Fulcher BD. Bridging the gap between connectome and transcriptome. Trends Cogn Sci. 2019;23(1):34-50.
34. Feng G, Wang Y, Huang W, Chen H, Cheng J, Shu N. Spatial and temporal pattern of structure–function coupling of human brain connectome with development. eLife. 2024;13:RP93325.
35. Zhang X-H, Anderson KM, Dong H-M, Chopra S, Dhamala E, Emani PS, et al. The cell-type underpinnings of the human functional cortical connectome. Nat Neurosci. 2024:1-11.
36. Morgan SE, Seidlitz J, Whitaker KJ, Romero-Garcia R, Clifton NE, Scarpazza C, et al. Cortical patterning of abnormal morphometric similarity in psychosis is associated with brain expression of schizophrenia-related genes. Proc Natl Acad Sci USA. 2019;116(19):9604-9609.
37. Li J, Seidlitz J, Suckling J, Fan F, Ji G-J, Meng Y, et al. Cortical structural differences in major depressive disorder correlate with cell type-specific transcriptional signatures. Nat Commun. 2021;12(1):1647.
38. Li L, Wei Y, Zhang J, Ma J, Yi Y, Gu Y, et al. Gene expression associated with individual variability in intrinsic functional connectivity. Neuroimage. 2021;245:118743.
39. Benjamin KJ, Chen Q, Eagles NJ, Huuki-Myers LA, Collado-Torres L, Stolz JM, et al. Analysis of gene expression in the postmortem brain of neurotypical Black Americans reveals contributions of genetic ancestry. Nat Neurosci. 2024;27:1067-1074.
40. Foley KG, Adli M, Kim JJ. Single-nuclei sequencing of uterine serous carcinoma reveals racial differences in immune signaling. Proc Natl Acad Sci USA. 2024;121(34):e2402998121.
41. Association As. 2010 Alzheimer's disease facts and figures. Alzheimers Dement. 2010;6(2):158-194.
42. Power MC, Bennett EE, Turner RW, Dowling NM, Ciarleglio A, Glymour MM, et al. Trends in relative incidence and prevalence of dementia across non-Hispanic Black and White individuals in the United States, 2000-2016. JAMA Neurol. 2021;78(3):275-284.
43. KESSLER II. Epidemiologic studies of Parkinson's disease: II. A hospital-based survey. Am J Epidemiol. 1972;95(4):308-318.
44. Kong R, Yang Q, Gordon E, Xue A, Yan X, Orban C, et al. Individual-specific areal-level parcellations improve functional connectivity prediction of behavior. Cereb Cortex. 2021;31(10):4477-4500.
45. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex. 2018;28(9):3095-3114.
46. Wang D, Buckner RL, Fox MD, Holt DJ, Holmes AJ, Stoecklein S, et al. Parcellating cortical functional networks in individuals. Nat Neurosci. 2015;18(12):1853-1860.
47. Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, et al. Precision functional mapping of individual human brains. Neuron. 2017;95(4):791-807. e797.
48. Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex. 2016;26(8):3508-3526.
49. Chen J, Tam A, Kebets V, Orban C, Ooi LQR, Asplund CL, et al. Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. Nat Commun. 2022;13(1):1-17.
50. Mowinckel AM, Vidal-Piñeiro D. Visualization of brain statistics with R packages ggseg and ggseg3d. Adv Methods Pract Psychol Sci. 2020;3(4):466-483.

51. Haufe S, Meinecke F, Görgen K, Dähne S, Haynes J-D, Blankertz B, et al. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage. 2014;87:96-110.
52. Sydnor VJ, Larsen B, Seidlitz J, Adebimpe A, Alexander-Bloch AF, Bassett DS, et al. Intrinsic activity development unfolds along a sensorimotor–association cortical axis in youth. Nat Neurosci. 2023;26(4):638-649.
53. Seidlitz J, Váša F, Shinn M, Romero-Garcia R, Whitaker KJ, Vértes PE, et al. Morphometric similarity networks detect microscale cortical organization and predict inter-individual cognitive variation. Neuron. 2018;97(1):231-247. e237.
54. Fields RD. White matter in learning, cognition and psychiatric disorders. Trends Neurosci. 2008;31(7):361-370.
55. Glasser MF, Van Essen DC. Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI. J Neurosci. 2011;31(32):11597-11616.
56. Soldan A, Pettigrew C, Zhu Y, Wang M-C, Bilgel M, Hou X, et al. Association of lifestyle activities with functional brain connectivity and relationship to cognitive decline among older adults. Cereb Cortex. 2021;31(12):5637-5651.
57. Rolls ET, Feng R, Feng J. Lifestyle risks associated with brain functional connectivity and structure. Hum Brain Mapp. 2023;44(6):2479-2492.
58. Sydnor VJ, Larsen B, Bassett DS, Alexander-Bloch A, Fair DA, Liston C, et al. Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. Neuron. 2021;109(18):2820-2846.
59. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature. 2012;489(7416):391-399.
60. Sunkin SM, Ng L, Lau C, Dolbeare T, Gilbert TL, Thompson CL, et al. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Res. 2012;41(D1):D996-D1008.
61. Abdi H. Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdisc Rev Comput Stat. 2010;2(1):97-106.
62. Yeung C-YC, Taylor SH, Garva R, Holmes DF, Zeef LA, Soininen R, et al. Arhgap28 is a RhoGAP that inactivates RhoA and downregulates stress fibers. PLoS One. 2014;9(9):e107036.
63. Stephens HA. MICA and MICB genes: can the enigma of their polymorphism be resolved? Trends Immunol. 2001;22(7):378-385.
64. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.
65. Bijsterbosch JD, Woolrich MW, Glasser MF, Robinson EC, Beckmann CF, Van Essen DC, et al. The relationship between spatial configuration and functional connectivity of brain regions. eLife. 2018;7:e32992.
66. Cui Z, Li H, Xia CH, Larsen B, Adebimpe A, Baum GL, et al. Individual variation in functional topography of association networks in youth. Neuron. 2020;106(2):340-353. e348.
67. Shanmugan S, Seidlitz J, Cui Z, Adebimpe A, Bassett DS, Bertolero MA, et al. Sex differences in the functional topography of association networks in youth. Proc Natl Acad Sci USA. 2022;119(33):e2110416119.
68. Valk SL, Xu T, Paquola C, Park B-y, Bethlehem RA, Vos de Wael R, et al. Genetic and phylogenetic uncoupling of structure and function in human transmodal cortex. Nat Commun. 2022;13(1):2341.
69. Sebenius I, Dorfschmidt L, Seidlitz J, Alexander-Bloch A, Morgan SE, Bullmore E. Structural MRI of brain similarity networks. Nat Rev Neurosci. 2024:1-18.

70. Keller AS, Moore TM, Luo A, Visoki E, Gataviņš MM, Shetty A, et al. A general exposome factor explains individual differences in functional brain network topography and cognition in youth. Dev Cogn Neurosci. 2024;66:101370.
71. Poldrack RA. Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci. 2006;10(2):59-63.
72. Gordon EM, Lynch CJ, Gratton C, Laumann TO, Gilmore AW, Greene DJ, et al. Three distinct sets of connector hubs integrate human brain function. Cell Rep. 2018;24(7):1687-1695. e1684.
73. Kim P, Evans GW, Angstadt M, Ho SS, Sripada CS, Swain JE, et al. Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. Proc Natl Acad Sci USA. 2013;110(46):18442-18447.
74. Goodman AM, Harnett NG, Knight DC. Pavlovian conditioned diminution of the neurobehavioral response to threat. Neurosci Biobehav Rev. 2018;84:218-224.
75. Kao G, Thompson JS. Racial and ethnic stratification in educational achievement and attainment. Annu Rev Sociol. 2003;29(1):417-442.
76. Espinosa LL, Turk JM, Taylor M, Chessman HM. Race and ethnicity in higher education: A status report. American Council on Education, 2019.
77. Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. J Int Neuropsychol Soc. 2002;8(3):448-460.
78. Barulli DJ, Rakitin BC, Lemaire P, Stern Y. The influence of cognitive reserve on strategy selection in normal aging. J Int Neuropsychol Soc. 2013;19(7):841-844.
79. Barulli D, Stern Y. Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. Trends Cogn Sci. 2013;17(10):502-509.
80. Teipel SJ, Meindl T, Wagner M, Kohl T, Bürger K, Reiser MF, et al. White matter microstructure in relation to education in aging and Alzheimer's disease. J Alzheimers Dis. 2009;17(3):571-583.
81. Arenaza-Urquijo EM, Landeau B, La Joie R, Mevel K, Mézenge F, Perrotin A, et al. Relationships between years of education and gray matter volume, metabolism and functional connectivity in healthy elders. Neuroimage. 2013;83:450-457.
82. Dufouil C, Alperovitch A, Tzourio C. Influence of education on the relationship between white matter lesions and cognition. Neurology. 2003;60(5):831-836.
83. Bozzali M, Dowling C, Serra L, Spanò B, Torso M, Marra C, et al. The impact of cognitive reserve on brain functional connectivity in Alzheimer's disease. J Alzheimers Dis. 2015;44(1):243-250.
84. Seyedsalehi A, Warrier V, Bethlehem RA, Perry BI, Burgess S, Murray GK. Educational attainment, structural brain reserve and Alzheimer’s disease: a Mendelian randomization analysis. Brain. 2023;146(5):2059-2074.
85. Ali MM, Creedon T, Bagalman E, Bui J, Clemans-Cope L, Winiski E, et al. Substance Use and Substance Use Disorders by Race and Ethnicity, 2015-2019. U.S. Department of Health and Human Services, 2023.
86. Volkow ND, Wang G-J, Fowler JS, Tomasi D, Telang F. Addiction: beyond dopamine reward circuitry. Proc Natl Acad Sci USA. 2011;108(37):15037-15042.
87. Singleton SP, Velidi P, Schilling L, Luppi AI, Jamison K, Parkes L, et al. Altered structural connectivity and functional brain dynamics in individuals with heavy alcohol use elucidated via network control theory. Biol Psychiatry Cogn Neurosci Neuroimaging. 2024.
88. Li C-sR, Sinha R. Inhibitory control and emotional stress regulation: Neuroimaging evidence for frontal–limbic dysfunction in psycho-stimulant addiction. Neurosci Biobehav Rev. 2008;32(3):581-597.
89. Tidey JW, Miller ME. Smoking cessation and reduction in people with chronic mental illness. BMJ. 2015;351.

954 90. Xiang S, Jia T, Xie C, Cheng W, Chaarani B, Banaschewski T, et al. Association between
955 vmPFC gray matter volume and smoking initiation in adolescents. Nat Commun. 2023;14(1):4684.
956 91. Shokri-Kojori E, Tomasi D, Wiers CE, Wang G-J, Volkow ND. Alcohol affects brain functional
957 connectivity and its coupling with behavior: greater effects in male heavy drinkers. Mol Psychiatry.
958 2017;22(8):1185-1195.
959 92. MacKillop J, Agabio R, Feldstein Ewing SW, Heilig M, Kelly JF, Leggio L, et al. Hazardous
960 drinking and alcohol use disorders. Nat Rev Dis Primers. 2022;8(1):80.
961 93. Makinodan M, Rosen KM, Ito S, Corfas G. A critical period for social experience–dependent
962 oligodendrocyte maturation and myelination. Science. 2012;337(6100):1357-1360.
963 94. Forbes TA, Gallo V. All wrapped up: environmental effects on myelination. Trends Neurosci.
964 2017;40(9):572-587.
965 95. Cohen S, Greenberg ME. Communication between the synapse and the nucleus in neuronal
966 development, plasticity, and disease. Annu Rev Cell Dev Biol. 2008;24(1):183-209.
967 96. Tau GZ, Peterson BS. Normal development of brain circuits. Neuropsychopharmacol.
968 2010;35(1):147-168.
969 97. Jiang X, Nardelli J. Cellular and molecular introduction to brain development. Neurobiol Dis.
970 2016;92:3-17.
971 98. Stringhini S, Polidoro S, Sacerdote C, Kelly RS, Van Veldhoven K, Agnoli C, et al. Life-course
972 socioeconomic status and DNA methylation of genes regulating inflammation. Int J Epidemiol.
973 2015;44(4):1320-1330.
974 99. Kapoor M, Wang J-C, Farris SP, Liu Y, McClintick J, Gupta I, et al. Analysis of whole
975 genome-transcriptomic organization in brain to identify genes associated with alcoholism. Transl
976 Psychiatry. 2019;9(1):89.
977 100. Rahman MF, McGowan PO. Cell-type-specific epigenetic effects of early life stress on the
978 brain. Transl Psychiatry. 2022;12(1):326.
979 101. Grazioli E, Dimauro I, Mercatelli N, Wang G, Pitsiladis Y, Di Luigi L, et al. Physical activity
980 in the prevention of human diseases: role of epigenetic modifications. BMC Genomics. 2017;18:111-
981 123.
982 102. Vetr NG, Gay NR, Montgomery SB. The impact of exercise on gene regulation in association
983 with complex trait genetics. Nat Commun. 2024;15(1):3346.
984 103. Rosenbaum DM, Rasmussen SG, Kobilka BK. The structure and function of G-protein-
985 coupled receptors. Nature. 2009;459(7245):356-363.
986 104. Tcherkezian J, Lamarche‐Vane N. Current knowledge of the large RhoGAP family of
987 proteins. Biol Cell. 2007;99(2):67-86.
988 105. Bahram S. MIC genes: from genetics to biology. Adv Immunol. 2001;76:1-60.
989 106. Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, Periyasamy S, et al.
990 Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls,
991 and recommendations. Cell. 2019;179(3):589-603.
992 107. Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, et al. The WU-
993 Minn human connectome project: an overview. Neuroimage. 2013;80:62-79.
994 108. Harms MP, Somerville LH, Ances BM, Andersson J, Barch DM, Bastiani M, et al. Extending
995 the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and
996 Aging projects. Neuroimage. 2018;183:972-984.
997 109. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, et al. The
998 minimal preprocessing pipelines for the Human Connectome Project. Neuroimage. 2013;80:105-124.
999 110. Dhamala E, Ooi LQR, Chen J, Ricard JA, Berkeley E, Chopra S, et al. Brain-based predictions
1000 of psychiatric illness–linked behaviors across the sexes. Biol Psychiatry. 2023;94(6):479-491.

111. Serio B, Hettwer MD, Wiersch L, Bignardi G, Sacher J, Weis S, et al. Sex differences in functional cortical organization reflect differences in network topology rather than cortical morphometry. Nat Commun. 2024;15(1):7714.
112. Markello RD, Arnatkeviciute A, Poline J-B, Fulcher BD, Fornito A, Misic B. Standardizing workflows in imaging transcriptomics with the abagen toolbox. eLife. 2021;10:e72129.
113. Arnatkevičiūtė A, Fulcher BD, Fornito A. A practical guide to linking brain-wide gene expression and neuroimaging data. Neuroimage. 2019;189:353-367.