Regional structural-functional connectome coupling is heritable and associated with age, sex and cognition in adults

Zijin Gu[1], Keith Wakefield Jamison[2], Mert Rory Sabuncu[1,2], and Amy Kuceyeski[2,*]

[1]Electrical and Computer Engineering Department, Cornell University, Ithaca, 14850, USA
[2]Department of Radiology, Weill Cornell Medicine, New York, 10065, USA
[*]amk2012@med.cornell.edu

ABSTRACT

Large scale white matter brain connections quantified via the structural connectome (SC) act as the backbone for the flow of functional activation, which can be represented via the functional connectome (FC). Many studies have used statistical analysis or computational modeling techniques to relate SC and FC at a global, whole-brain level. However, relatively few studies have investigated the relationship between individual cortical and subcortical regions' structural and functional connectivity profiles, here called SC-FC coupling, or how this SC-FC coupling may be heritable or related to age, sex and cognitive abilities. Here, we quantify regional SC-FC coupling in a large group of healthy young adults (22 to 37 years) using diffusion-weighted MRI and resting-state functional MRI data from the Human Connectome Project. We find that while regional SC-FC coupling strengths vary widely across cortical, subcortical and cerebellar regions, they were strongest in highly myelinated visual and somatomotor areas. Additionally, SC-FC coupling displayed a broadly negative association with age and, depending on the region, varied across sexes and with cognitive scores. Specifically, males had higher coupling strength in right supramarginal gyrus and left cerebellar regions while females had higher coupling strength in right visual, right limbic and right cerebellar regions. Furthermore, increased SC-FC coupling in the right lingual gyrus was associated with worse cognitive scores. Finally, we found SC-FC coupling to be highly heritable, particularly in the visual, dorsal attention, and fronto-parietal networks, and, interestingly, more heritable than FC or SC alone. Taken together, these results suggest regional structure-function coupling in young adults decreases with age, varies across sexes in a non-systematic way, is somewhat associated with cognition and is highly heritable.

Introduction

The question of how anatomy and physiology are related is one of the fundamental questions in biology, particularly in neuroscience where studies of form and function have led to fundamental discoveries. In the last few decades, the invention of MRI has enabled *in vivo* investigation of whole-brain, anatomical (white matter) and physiological (functional co-activation) brain networks in human populations. Studies analyzing multi-modal connectivity networks have produced a consensus that, to some extent, alignments exist between the brain's anatomical structural connectome (SC) and its physiological functional connectome (FC)[1–5]. Recent work has focused on implementing computational models, including neural mass models, network diffusion models, graph theoretical or statistical approaches, that formalize the global relationship between SC and FC in both normal and pathological populations[6–9]. Some of the main goals in joint structure-function connectome modeling are to understand how neural populations communicate via the SC backbone[7], how functional activation spreads through the structural connectome[8], to increase the accuracy of noisy connectivity measurements, to identify function-specific subnetworks[10], to predict one modality from the other[1] or to identify multi-modal mechanisms of recovery after injury[11,12]. While useful, these modeling approaches are global in nature and ignore the regional variability in the structure-function relationship that, to date, has not been adequately quantified in adult populations.

Recent publications mapping connectome properties to cognitive abilities have focused on using either FC or SC alone, or concatenating both together to reveal brain-behavior relationships[13–17]. Some recent studies have identified relationships between global, whole-brain SC-FC correlations and cognitive abilities or states of awareness. One such paper showed that stronger global SC-FC correlations were related to worse cognitive function in older adults with cognitive impairment[18]. Another study showed disorders of consciousness patients with fewer signs of consciousness had longer dwell times in dynamic FC states that were most similar to SC[19]. It has also been shown that SC-FC similarity decreases with increasing awareness levels in anesthetized monkeys[20] and, similarly, decreases from deep sleep to wakefulness in humans[21]. Two studies, in severe brain injury and mild traumatic brain injury, revealed that increasing "distance" between SC and FC was related to better

recovery after injury[11,12]. These studies all suggest a weaker coupling of SC and FC is related to better cognitive performance and increasing awareness/consciousness. In contrast, however, a recent study showed increased cognitive flexibility was associated with increased alignment of FC and SC[22]. Therefore, how SC-FC coupling relates to various cognitive functions, awareness or other brain states may vary with the behavioral measure and population in question.

Even fewer studies have explored how the strength of the relationship between SC and FC may vary with age and sex. One such study in a small number of subjects ($N = 14$, 18 months to 18 years of age) showed increasing age was strongly related to higher global correlations between SC and FC ($r = 0.74$, $p < 0.05$)[23]. In one of the few studies to date of regional SC-FC coupling, Baum et. al (2020) studied a large number of developing subjects ($N = 727$, aged $8-23$ years old) and showed that the relationship between age and SC-FC coupling varied across brain regions, with some regions showing positive and fewer regions showing negative relationships. Furthermore, they showed that stronger SC-FC coupling in rostrolateral prefrontal cortex specifically was associated with development-related increases in executive function[24]. Another of regional SC-FC coupling analyzed data from a group of around 100 young adults and showed that, overall, regional SC-FC coupling was stronger in females than in males and that there were sex-specific correlations of SC-FC coupling with cognitive scores[25].

Some recent work has revealed the varying degrees to which the brain's FC is heritable[26–28]. Most studies have focused on FC; however, some recent preliminary work investigated the relationships between gene co-expression, FC, SC and behavior in a developmental cohort[29]. In that pre-print, the authors showed that FC, rather than SC, was more related to genetic co-expression, and, furthermore, that the brain's FC architecture is potentially the mediating factor between genetic variance and cognitive variance across the developing population. However, none of these studies have investigated the heritability of regional SC-FC coupling.

These studies of global, whole-brain SC-FC correlations, while informative, largely ignore regional variability of SC-FC coupling that may provide a more detailed picture of how anatomy and physiology vary with age, sex, genetics and cognitive abilities. There are only two studies to date investigating regional SC-FC coupling. The first used task-based FC in an adolescent population, focused on the cortex and did not assess heritability or sex differences[24] while the second used a data from a moderately sized sample of young adults, did not consider the cerebellum and did not investigate the heritability of SC-FC coupling[25]. In this work, for the first time, we quantify the cortical, subcortical and cerebellar topography of SC-FC coupling at rest in a group of young adults, verify its reproducibility and quantify its association with age, sex and cognition. Moreover, due to the nature of the HCP data, we were also able to assess the patterns of heritability of regional SC-FC coupling. Accurate quantification of the relationship between the brain's structural and functional networks at a regional level is imperative so we can understand how interacting brain circuits give rise to cognition and behavior, and how these relationships can vary with age, sex, cognition and genetics.

Results

We begin by presenting the regional SC-FC coupling values across unrelated young adults and demonstrating this measure's within-subject and out-of-sample reliability. We then map the regional relationships between SC-FC coupling and age, sex and cognition. Finally, we demonstrate the heritability of the SC-FC coupling. Our data is comprised of MRI, demographic, cognitive and familial relationship data from a group of 941 young and healthy adults, curated by the Human Connectome Project[30] (HCP). Individuals from the HCP's S1200 release were included if they had four functional MRI scans, a diffusion MRI scan and a Total Cognition test score. A fine-grained atlas (CC400)[31] was used to partition the brain into 392 spatially contiguous, functionally defined cortical and subcortical regions. Two 392×392 weighted adjacency matrices were then constructed, representing whole brain SC and FC. Here, we calculated FC using a regularized precision approach, which aims to capture only the "direct" connections between brain regions. We chose to use precision-based FC as it was recently shown to result in FC matrices that had stronger correlations with SC than more conventional Pearson correlation-based FC[32]. For completeness and comparison to previous work[24,25], Pearson correlation-based FC results are provided in the Supplemental Information. SC matrices were constructed using anatomically constrained probabilistic tractography; entries in the SC matrices were then a sum of the global filtering weights (SIFT2) of streamlines connecting pairs of regions, divided by the sum of the volumes of the two regions. Once the FC and SC were constructed, the regional SC-FC coupling vector was calculated for each individual in the following way. Each row in the SC matrix, representing a region's SC to the rest of the brain, was correlated with the same region's row in the FC, providing a regional SC-FC coupling vector of length 392 for each subject (Figure 1).

SC-FC coupling varies spatially, is consistent over time and is reproducible

The group average SC-FC coupling over 420 unrelated individuals is shown in Figure 2a. We found that, at the group level, regional SC-FC coupling was always positive and varied greatly across cortical and subcortical areas, ranging from $0-0.61$. Visual, and somatomotor areas had significantly higher SC-FC coupling than the other networks (except for dorsal attention network when comparing with somatomotor, see Figure 2b and c, all FDR corrected $p < 0.05$), with average SC-FC coupling values of 0.44 and 0.41, while limbic and subcortical areas had significantly weaker SC-FC coupling than the other networks

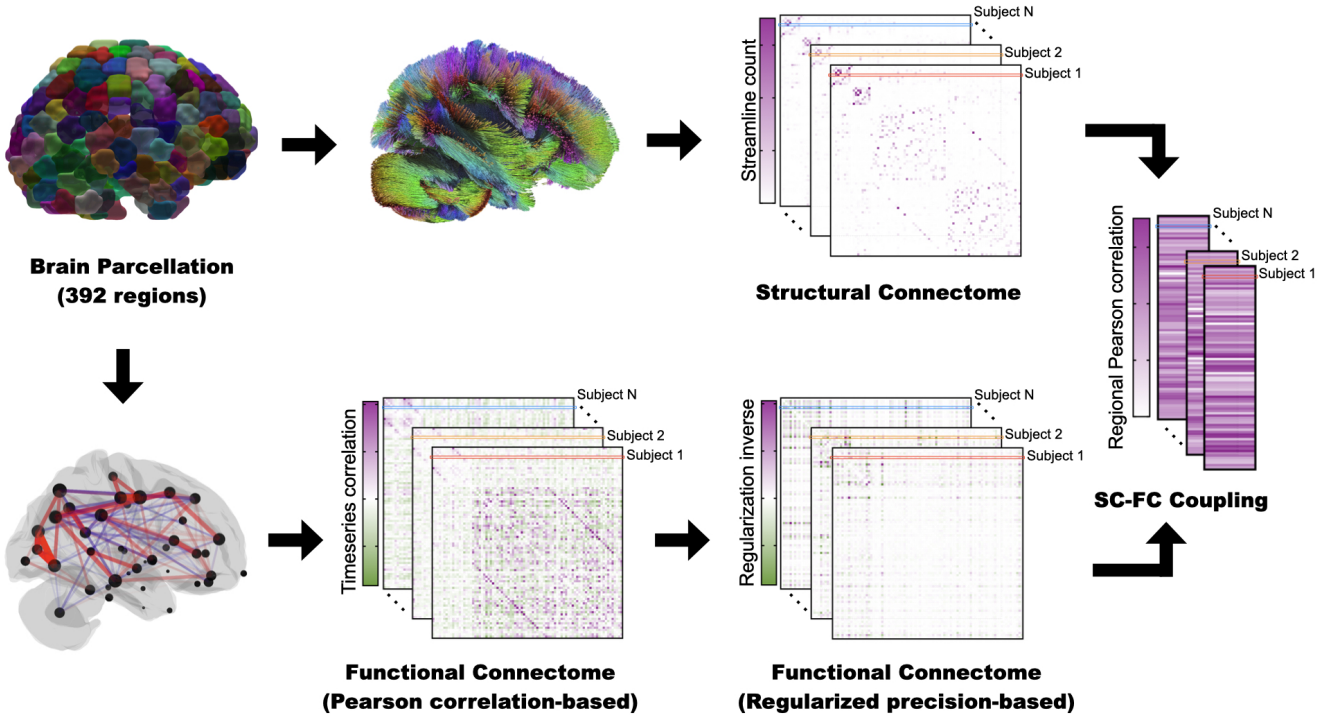

Figure 1. Workflow for quantifying regional SC-FC coupling. The CC400 atlas was used to parcellate the gray matter into 392 cortical and subcortical brain regions[31]. SC matrices were constructed based on probabilistic tractography aimed at reconstructing white matter pathways. FC matrices, representing similarity of functional activation over time, were considered in two ways. The Pearson correlation-based FC matrices were computed by correlating pairwise BOLD time series from the defined regions, while regularized precision-based FC matrices were computed by Tikhonov regularization of the inverse covariance matrix. For each subject, corresponding rows in the SC and FC matrices were correlated to obtain that region's SC-FC coupling value. The result is a vector of regional SC-FC coupling, of length 392, for each individual.

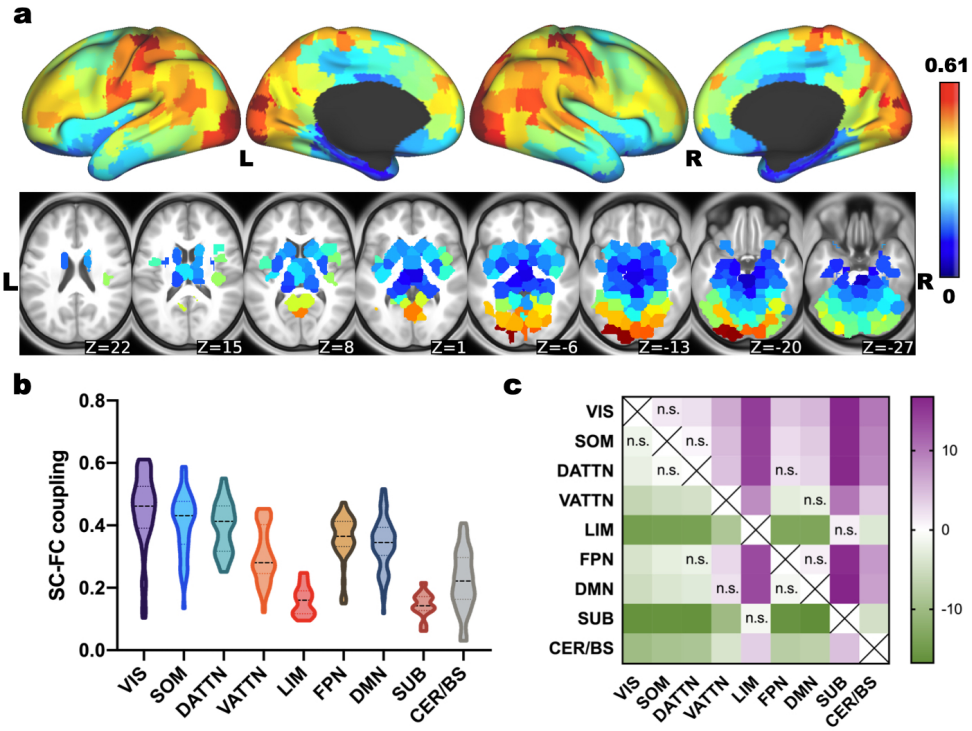

Figure 2. Regional SC-FC coupling varies spatially across the brain. a displays the SC-FC coupling for each cortical and subcortical region in the CC400 atlas. **b** shows the distribution of SC-FC coupling over regions grouped into nine different networks (7 Yeo networks, subcortical and cerebellum/brain stem). **c** shows the t-statistics for all pairwise comparisons of SC-FC coupling across networks, calculated as the network on the y-axis versus the network on the x-axis. Those comparisons with FDR corrected $p > 0.05$ are marked with *n.s.*. Visual, somatomotor and dorsal attention networks have higher SC-FC coupling than other networks while limbic and subcortical areas have weaker SC-FC coupling than other networks. Abbreviations: VIS - visual, SOM - somatomotor, DATTN - dorsal attention, VATTN - ventral attention, LIM - limbic, FPN - frontoparietal, DMN - default mode, SUB - subcortical, CER/BS - cerebellum and brain stem.

(see Figure 2b and c, all FDR corrected $p < 0.05$), with average SC-FC coupling values of 0.16 and 0.14. SC-FC coupling calculated using Pearson correlation-based FC was similar to, but generally weaker than, precision-based SC-FC coupling (Pearson's $r = 0.85$, $p < 1e-109$), see Supplementary Information Figure S1 . All networks, except subcortical, limbic and cerebellum/brain stem, had significantly higher SC-FC coupling when the measure was calculated using the precision-based FC compared to when SC-FC was calculated using Pearson correlation-based FC (FDR corrected $p < 0.05$).

Next, we tested the reliability and reproducibility of SC-FC coupling by examining its consistency within individuals over time and across different populations of individuals. To test for consistency over time within the same individuals, we used data from a subset of 41 HCP subjects who had a second MRI scan about 6 months after the first. SC-FC coupling was indeed highly consistent across time, with a mean difference of $\mu = -0.004$, limits of agreement $LoA = \mu \pm 0.028$, see Figure 3a, and a test-retest correlation of 0.99 (Pearson's r, $p < 1e-307$). Furthermore, we examined out-of-sample, across population reliability in SC-FC coupling using a subset of 346 unrelated HCP subjects (age, 28.78 ± 3.80 y; 148 males and 198 females), distinct from the initial set of 415 unrelated subjects. It should be noted that, while each set of subjects did not contain relatives within them, there may be some familial relationships across the two sets of subjects which could result in an overestimation of the out-of-sample reliability. Still, out-of-sample reliability was high, with a small mean difference $\mu = 0.005$ and limits of agreement $LoA = \mu \pm 0.017$, see Figure 3b, and high correlation (Pearson's $r = 0.99$, $p < 1e-307$).

Age, sex and cognition have region-specific, significant associations with SC-FC coupling

We used a generalized linear model (GLM) to quantify the association between different characteristics of interest and SC-FC coupling. Specifically, subjects' age, sex, total cognition score, intracranial volume (ICV), in-scanner head motion as well as the two-way interactions terms of age*cognition, sex*cognition and ICV*motion were included in the model. The most prominent relationship observed was a broadly negative association between age and SC-FC coupling, particularly in subcortical structures (mean $\beta = -3.13$), including the caudate, putamen and thalamus, visual areas (mean $\beta = -3.15$) and somatomotor areas

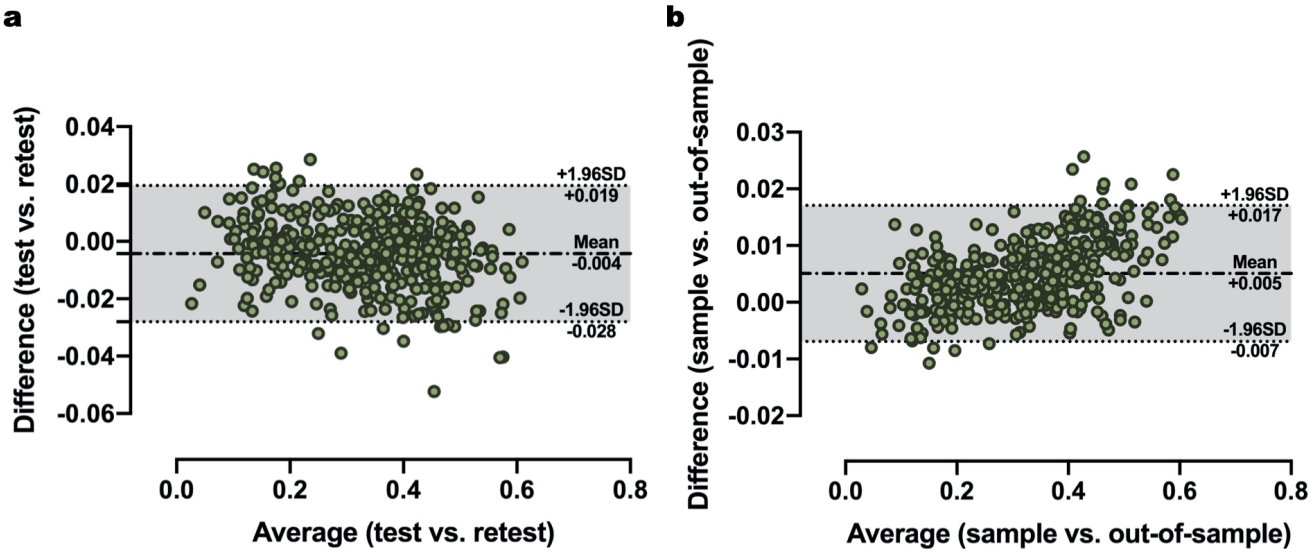

Figure 3. SC-FC coupling is consistent over time and is reproducible. a Bland-Altman plot shows good agreement between the SC-FC coupling calculated in the same set of 41 subjects across two MRI scans taken 6 months apart (mean difference $\mu = -0.004$ and limits of agreement $LoA = \mu \pm 0.028$). **b** Bland-Altman plot shows good agreement between the SC-FC coupling calculated from the original set of 415 subjects and another out-of-sample set of 346 subjects (mean difference $\mu = 0.005$ and limits of agreement $LoA = \mu \pm 0.017$).

(mean $\beta = -3.12$), see Figure 4a,b and c. Males had higher SC-FC coupling in the left cerebellum and right supramarginal gyrus, while females had higher SC-FC coupling in right fusiform gyrus, right cerebellum and right temporal areas (Figure 4d, e and f). The association between cognition and SC-FC coupling was weaker when compared with age and sex. Higher total cognition scores were related to decreased SC-FC coupling in right lingual gyrus areas (Figure 4g, h and i). Similar results were found when using Pearson correlation-based FC to calculate SC-FC coupling, see Figure S2 in Supplementary Information. There were some associations found between SC-FC coupling and both ICV and in-scanner head motion (see Supplementary Information Figure S5 for the precision-based FC results and Supplementary Information Figure S6 for the correlation-based FC results). ICV had more positive than negative associations, while head motion was a mix of both positive and negative associations. For both covariates, most of the coefficients reaching significance were positive, indicating increasing SC-FC coupling with increased head size and motion.

SC-FC coupling is more heritable than FC or SC

Next, we assessed the heritability of SC-FC coupling using a recently developed modeling approach that considers the level of measurement error of the imaging biomarker in question[26]. Specifically, a linear mixed effect (LME) model was designed to independently estimate the inter- and intrasubject variation (representing the unstable, transient component and measurement error) of the total phenotype variability. Heritability was defined as the proportion of intersubject variation attributable to genetics. Overall, SC-FC coupling was highly heritable, particularly in the dorsal attention, visual and fronto-pareital networks (mean heritability 0.56, 0.54 and 0.53, respectively), see Figure 5a and b). SC-FC coupling in limbic and subcortical areas were significantly less heritable (mean heritability 0.16 and 0.18) than the other seven networks (see Figure 5b and c, all FDR corrected $p < 0.05$). For comparison, we calculated the heritability of the node strength ($l1$ norm of each row) of the SC and FC matrices independently, see Figure 5d and g. First, precision-based FC had overall relatively low levels of heritability and was significantly negatively correlated with heritability of SC (Pearson's $r = -0.282$, $p < 1e-7$). Furthermore, SC-FC coupling heritability was not reflective of just SC or FC heritability, being significantly correlated with both (in opposite directions), but was more driven by FC. This is evidenced by the moderate, negative correlation between SC-FC coupling and SC heritability (Pearson's $r = -0.294$, $p < 1e-8$) and the significant, larger positive correlation between SC-FC coupling and FC heritability (Pearson's $r = 0.822$,$p < 1e-96$), see (Figure 5j and k).

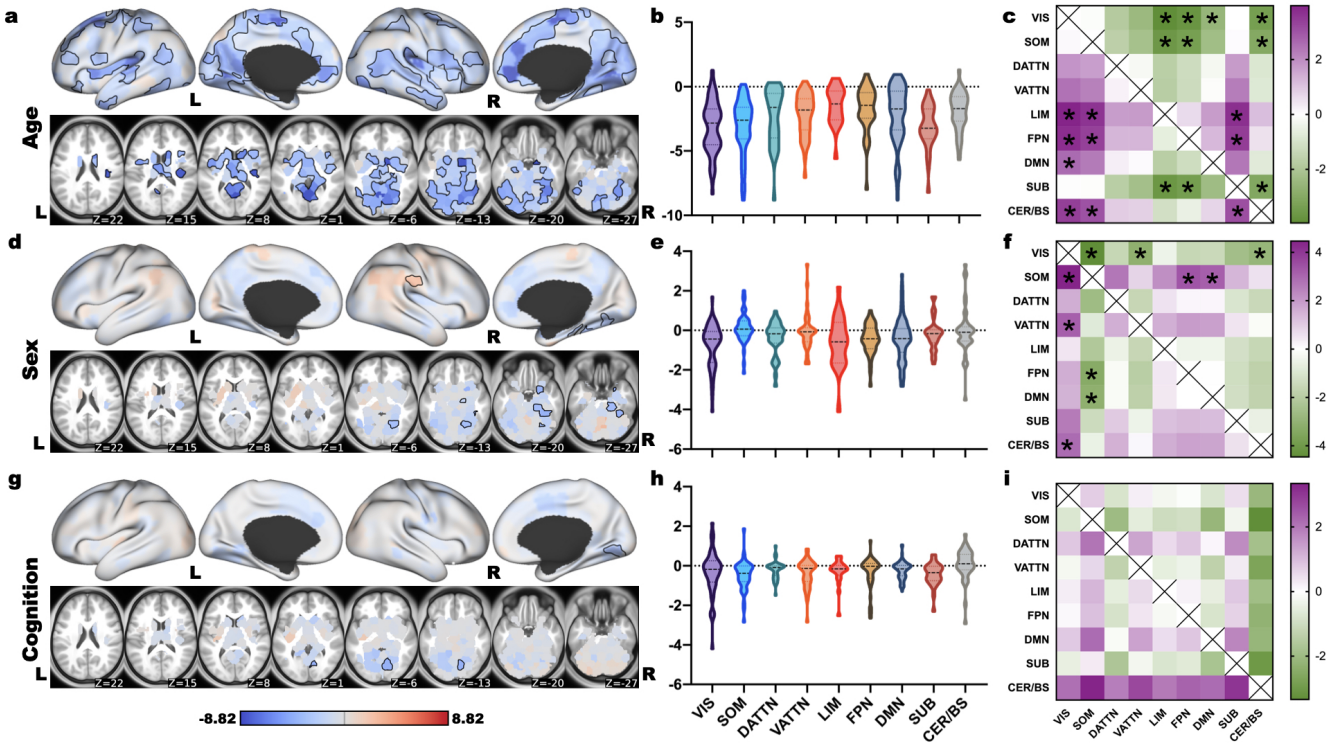

Figure 4. Associations between regional SC-FC coupling and age, sex and total cognition. **a**, **d** and **g** display regional β values from the GLM quantifying associations between SC-FC coupling and age, sex (blue indicates higher SC-FC coupling in females, red higher in males) and total cognition, respectively. Areas with significant β values (after correction) are outlined in black. **b**, **e** and **h** show the network-wise β values for age, sex and total cognition, respectively. **c**, **f** and **i** show the t-statistics for all pairwise comparisons. Those comparisons with FDR corrected $p < 0.05$ are marked with $*$.

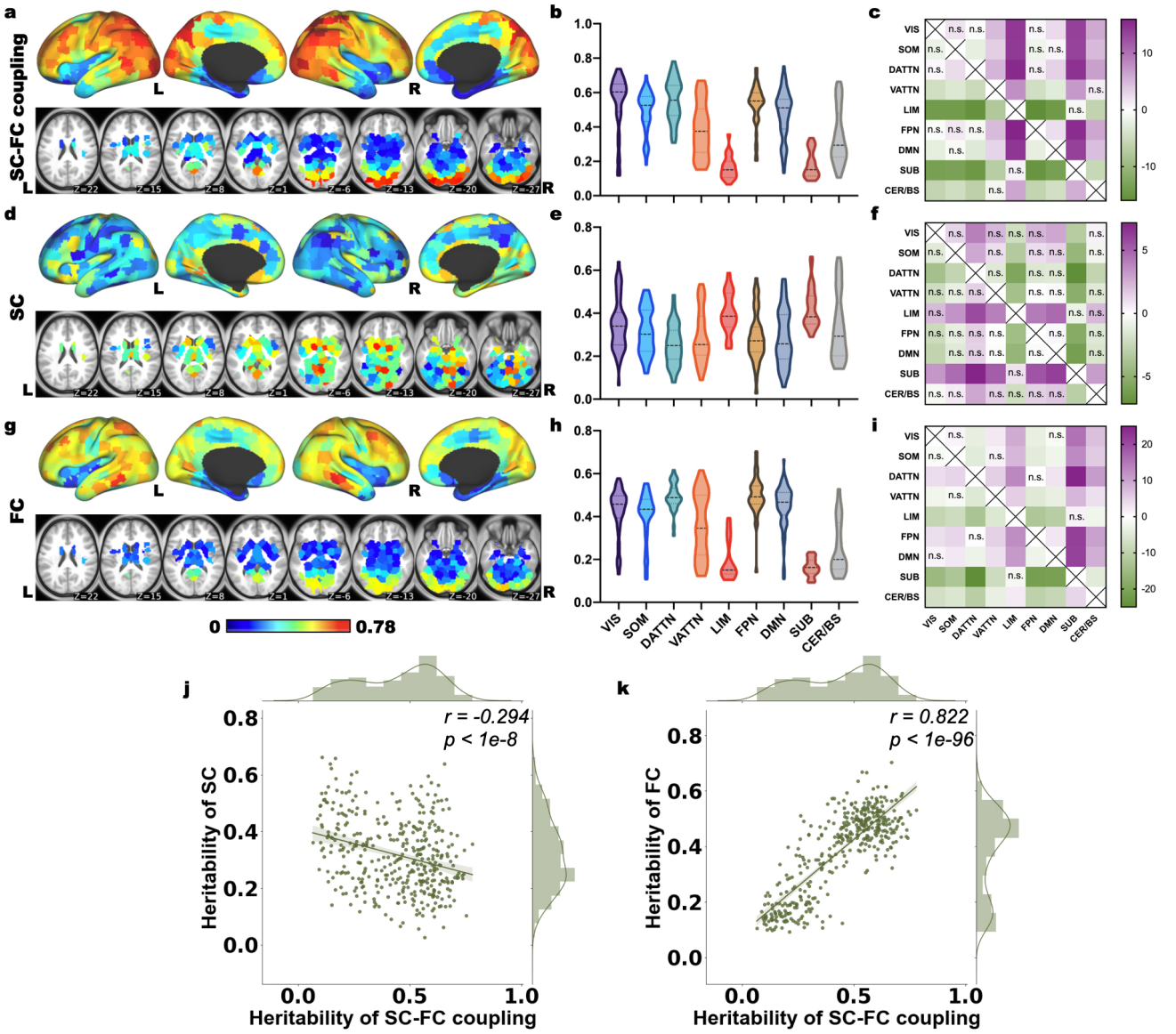

Figure 5. **SC-FC coupling heritability estimates.** **a**, **d** and **g** Regional heritability estimates of SC-FC coupling, SC node strength and precision-based FC node strength. **b**, **e** and **h** Rregional heritability estimates of SC-FC coupling, grouped by functional network, for SC-FC coupling, SC node strength and precision-based FC node strength, respectively. **c**, **f** and **i** Comparisons of heritability values between networks (t-statistics); those with FDR corrected $p > 0.05$ are marked with *n.s.*. **j** and **k** Regional heritability estimates of SC-FC coupling are significantly negatively correlated with regional heritability of SC node strength (Pearson's $r = -0.294$, $p < 1e-8$) and significantly positively correlated with regional heritability of FC node strength (Pearson's $r = 0.882$, $p < 1e-96$).

Discussion

In this paper, we quantified the strength of coupling between the structural and functional connectivity profiles of cortical, subcortical and cerebellar brain regions in a large sample of healthy young adults. We demonstrate that SC-FC coupling is strongest in visual and somatomotor areas, weakest in limbic and subcortical regions and is consistent across time and different sample populations. Furthermore, we show SC-FC coupling has a broadly negative relationship with age, varies across sexes, although not in uniform manner across brain regions, and that stronger SC-FC coupling, particularly in the right lingual gyrus, is related to lower total cognition scores. Finally, we show SC-FC coupling is highly heritable, particularly in the dorsal attention, visual and fronto-parietal control networks, demonstrating stronger values across the brain compared to SC or FC alone.

The ordering of cortical regions into anatomical hierarchies, wherein primary sensory regions are at the bottom and higher-order association areas are at the top, provides a way to organize brain regions. Anatomical hierarchies defined by myelination and white matter connectivity patterns have been shown to reflect functional and transcriptome specialization[33–35]. The cortical SC-FC coupling pattern found in our young adult population, which closely tracks with cortical myelination, further supports the argument that regional SC-FC coupling is reflective of anatomical hierarchies[24]. In fact, the Spearman correlation of the population average SC-FC coupling and regional average myelination from the HCP subjects was 0.42 for precision-based FC ($p < 1e-15$) and 0.53 for Pearson-correlation based FC ($p < 1e-25$). Lower-order areas of high cortical myelination, including primary visual, somatosensory and motor regions, tend to have functional activation patterns that are strongly aligned to their white matter connectivity profiles. Higher-order association areas with lower myelination tend to have complex, dynamic functional profiles that are less anchored to their structural connectivity profiles. Furthermore, we showed relatively low SC-FC coupling in subcortical and limbic structures, which could be reflective of their diverse structural connections and their role as relay stations for functional signals traveling between cerebellar, sensory and other cortical regions. Subcortical and limbic structures also tend to have lower signal-to-noise ratio due to MR imaging artifacts[36] which could also contribute to lower SC-FC coupling.

Functional activation flows not only through direct SC but also indirect, multi-synaptic white matter connections, which likely contributes to divergence of SC and FC to varying degrees[37]. Statistical, communication, biophysical and machine learning models have been applied to better align FC and SC[3,7,8,38]. Recent work has also found the strength of global SC-FC correlation depends on how FC is calculated[32]. In particular, this work showed FC calculated using partial correlation (precision), which aims to isolate direct and remove the effect of indirect functional connections, had stronger correlations with SC than standard FC calculated using full (Pearson) correlation. Largely, our results are consistent with the global findings in that regional SC-FC coupling is generally larger when using precision-based FC compared to using full Pearson correlation FC. However the overall intra-areal patterns across the brain (see Supplementary Information Figure S1), heritability (see Figure Supplementary Information Figure S3) and relationships of SC-FC coupling with age, sex, cognition (see Supplementary Information Figure S2) were similar across the FC types.

We showed largely negative associations of SC-FC coupling with age in this young adult population, which we hypothesize could reflect an increase in functional diversity over young adulthood compared against a relatively static myelination pattern. Interestingly, Baum et al. (2020) found mostly age-related increases and some decreases in SC-FC coupling during adolescence which they interpreted as possibly reflecting both functional diversification and increase in myelination in development. We also show sex differences in SC-FC coupling, with males having higher coupling in right supramarginal gyrus and left cerebellar regions and females having higher coupling in right fusiform gyrus, right cerebellum, right parahippocampus/medial temporal structures, and right lingual gyrus. This disagrees somewhat with recent findings in young adults that females had overall greater SC-FC coupling than their male counterparts, particularly in left inferior frontal gyrus, left inferior parietal lobe, right superior frontal gyrus and right superior parietal gyrus[25]. They furthermore found higher SC-FC coupling in males in right insula, left hippocampus and right parahippocampal gyrus[25]. Both studies did agree on males having larger SC-FC coupling in right supramarginal gyrus, but the rest of the results diverge. We hypothesize this may be due to differences in sample size/characteristics or imaging acquisition/preprocessing strategies; particularly important when investigating sex differences in FC is the use of global signal regression which can remove non-neuronal signals like motion[39] and respiration that are known to have sex-specific effects[40]. Our GLM framework additionally controlled for covariates like in-scanner motion and intracranial volume which have known sex differences and a complex relationship with BOLD signals[41,42].

Most previous publications investigating SC-FC relationships and their cognitive implications have explored correlations between impairment or cognition with the strength of the correlation between global, whole-brain SC and FC[19,22,43,44]. Studies in controls have revealed worse cognitive performance in healthy aging was associated with longer latency in dynamic FC states that are more similar to SC[44] and that cognitive flexibility was associated with FC's alignment with SC[22]. Studies in individuals with neurological disorders have shown that SC-FC similarity increases with dementia diagnosis and individuals' performance on memory tasks[43] and that increasing awareness levels in individuals with disorders of consciousness are related to longer latency in dynamic FC states less similar to SC[19]. Regional SC-FC coupling was found to be differently correlated with cognitive function in females and males; specifically, poorer working memory in females was related to weaker SC-FC

coupling in local (non-hub/feeder) connections and better reasoning ability in males was related to stronger SC-FC coupling in rich-club hub connections[25]. In their adolescent population, Baum et al. (2020) found mostly positive correlations between executive function and SC-FC coupling, particularly in lateral frontal and right medial occipital regions; the only region to show the negative associations with cognitive scores was the right primary motor cortex[24]. In the present study, we observe a generally negative association of regional SC-FC coupling across the brain, indicating stronger SC-FC coupling was related to lower total cognition scores. However, SC-FC coupling associations with cognition were generally weaker than associations with age and sex; we hypothesize this is due to the many covariates considered in the model compared to previous work. The only region that achieved significance after all the other covariates were considered was right lingual gyrus in the medial occipital cortex, which has been associated with visual memory and word recognition[45,46]. Interestingly, this region was also one identified in the adolescent study as having an association between SC-FC coupling and executive function, although the association was in the opposite direction[24].

For the first time, we show that regional SC-FC coupling is highly heritable across the brain (with values up to 0.78), particularly in the visual, fronto-parietal control and dorsal attention network. Interestingly, we found regional SC-FC coupling to be more heritable than SC or FC alone, and furthermore, that it was not driven entirely by one modality or the other. Previous studies have shown heritability of FC profiles, with the default mode network having highest heritability (estimates ranging from $0.42-0.8$) and motor and visual areas having lowest heritability estimates ($0.2-0.3$)[26,47]. Both our precision-based and Pearson-based FC results are very similar to these previous findings; however the precision-based FC demonstrates lower levels of heritability than Pearson-based FC ($p < 1e-10$). We hypothesize this could be due to the procedure for calculating the precision matrix. First, the inversion of the covariance matrix is ill-posed so inverting it may introduce noise. Second, the regularization parameter is chosen to minimize the difference between individuals' precision matrices and the population-level mean unregularized precision matrix, which could obscure individual (heritable) characteristics. Furthermore, for the first time, we show regional SC heritability estimates, which are lower than both the heritability of the precision-based FC and the heritability of the Pearson-based FC. One consideration for the SC heritability is that our statistical model uses estimates of between-measure variability based on repeat measurements to account for noise in the heritability estimate. However, we only had one SC per subject so the these estimates could be lower relative to the FC heritability estimates. Interestingly, we found highest SC heritability in limbic and subcortical networks, which were the networks with the lowest heritability in FC and SC-FC coupling. Previous work has suggested different genetic signatures underlying brain anatomy and physiology[47]. However, these areas do tend to have the most noise in fMRI which could also contribute to lower FC heritability estimates. While no other studies have investigated the heritability of SC, one recent preprint quantifying heritability of the size of cortical areas showed unimodal motor/sensory networks had higher heritability (0.44) relative to heteromodal association networks (0.33)[48]. We do show general agreement with their findings in that unimodal visual and motor networks had the highest SC heritability across cortical networks.

Limitations

The results of the analyses in this work are limited by the characteristics of the individuals in the HCP young adult data set. As seen in previous work, SC-FC coupling relationships may vary differently with age across the lifespan, so interpretations of our current findings should be restricted to young adult populations. In addition, we chose to perform global signal regression when processing the fMRI data, as it has been shown that doing so can mitigate systematic non-neuronal shifts in the intensity of the BOLD signal that are not reflective of brain activity[39]. However, a few groups have advocated that performing global signal regression results in anti-correlations that are not straightforwardly interpretable[49]. Finally, tractography algorithms are known to produce streamlines that are not fully reflective of actual anatomical connections[50,51]. Here, we somewhat mitigate this effect by using a global filtering algorithm, which has been shown to result in streamlines that are more reflective of underlying anatomy[52].

Conclusions

Understanding how macroscopic anatomical and physiological connectomes are intertwined and can influence behavior or be influenced by an individual's characteristics or environment is an important, unanswered question in human neuroscience. Here, we use neuroimaging, demographic/familial relationship information and cognitive measures in a large population of young healthy adults to begin to uncover some of these associations. We show that regional structure-function coupling is strongest in highly myelinated visual and somatomotor networks, decreases with age, varies with sex, is related to cognition and is highly heritable. Taken together, these results demonstrate that investigating structure-function relationships at a macroscopic scale can reveal important knowledge in the study of brain form and function.

Methods

Data Description

The data for this study comes from the publicly available HCP database containing high-resolution, preprocessed anatomical, diffusion and resting-state functional MRI data. Specifically, we use WU-Minn HCP minimally processed S1200 release which includes high-resolution 3T MR scans, demographics, behavioral and cognitive scores for a large population of young healthy adults (age 22 to 37 years). For the SC-FC coupling results shown in Figure 2, we used the subset of 420 unrelated subjects that had all four fMRI scans and a complete dMRI scan. For the GLM analyses shown in Figure 4, we selected 415 unrelated subjects from them that had all cognitive scores (age, 28.69 ± 3.69 years; 213 males, 202 females). For the heritability analysis shown in Figure 5, we analyzed 941 subjects (age, 28.67 ± 3.70 years; 441 males, 500 females) from 425 different families. In this set of 941 subjects that had all four fMRI scans and a dMRI scan, there were 116 MZ twin pairs, 61 DZ twin pairs, 455 full siblings and 132 singletons (single-birth individuals without siblings).

Construction of the Structural Connectomes

HCP subjects were scanned on a customized Siemens 3T "Connectome Skyra" housed at Washington University in St. Louis. The HCP diffusion data (1.25mm isotropic voxels, TR/TE = 5520/89.5ms, 3x multiband acceleration, b=1000,2000,3000, 90 directions/shell, collected with both left-right and right-left phase encoding) were first minimally preprocessed to correct for motion, EPI and eddy-current distortion, and registered to each subject's T1 anatomical scan[53]. A multi-shell, multi-tissue constrained spherical deconvolution (CSD) model was computed in MRtrix3 to estimate the orientation distribution function[54]. We used a probabilistic (iFOD2[55]), anatomically constrained (ACT[56]) tractography algorithm with dynamic seeding to create individual, whole-brain tractograms containing 5 million streamlines. To better match the whole brain tractogram to diffusion properties of the observed data, we also computed streamline weights that are designed to reduce known biases in tractography data (SIFT2[52]). Finally, the tractograms were used to estimate SC weights for the CC400[31] atlas. The SC between any two regions was the SIFT2-weighted sum of streamlines connecting those regions divided by the sum of the gray matter volume of those regions. The result was an ROI-volume normalized pairwise SC matrix for each subject.

Construction of the Functional Connectomes

There were four gradient-echo EPI resting-state fMRI runs (2.0mm isotropic voxels, TR/TE = 720/33.1ms, 8x multiband acceleration, FoV = 208×180 mm^2, FA = 52°, 72 slices) of approximately 15 minutes each, with two runs in one session and two in a second session, where each session included both right-left and left-right phase encoding. There were 1200 volumes for each run and a total of 4800 volumes (1200 volumes $\times$ 4 runs) for each subject. The data were minimally preprocessed[53] and ICA+FIX[57,58] denoised by the HCP consortium[59]. In scanner motion for each individual was quantified by averaging the overall frame-wise displacement for each of the four fMRI scans. We further regressed out the effect of global gray matter signal and its temporal derivative[60]. To calculate the FC matrices, we first variance-normalized and concatenated the four fMRI runs and calculated the Pearson correlation between each region-pair's average time series in the CC400 atlas[31]; the result was a single Pearson correlation-based FC matrix Σ for each subject. To compute precision-based FC, we first computed the unregularized inverse of the correlation matrix for each individual, and averaged them over the population to obtain the population-level precision matrix. We then calculated the individuals' precision matrices using Tikhonov regularization, which adds a full-rank regularization term (scaled identity) to the correlation matrix before inversion[32]:

$$P_{reg} = (\Sigma + \lambda \cdot I)^{-1}$$

where I is the identity matrix and λ is the regularization parameter. The regularization parameter $\lambda \in [0,1]$ was chosen via grid search to be the value that minimized the sum of the Frobenius norms between the regularized subject precision matrices and the group-averaged unregularized precision matrix, resulting in $\lambda_{opt} = 0.3$. For heritability analysis, the process outlined above was repeated for each of the individual's 4 scans independently, as the LME model uses between-measurement variability in its estimates of heritability[26]. For consistency, we used the same $\lambda_{opt} = 0.3$ for individual scans. For the Pearson correlation-based FC results in the Supplemental Materials, FC matrices were calculated for each of the 4 scans independently and then the average FC over those 4 scans was taken.

Calculation of SC-FC Coupling

SC-FC coupling was constructed by calculating the Pearson correlation between a row of the SC matrix, representing the connectivity fingerprint of that region to every other region in the brain, with the corresponding row of the FC matrix (excluding the self-connection). The result of this step in the analysis is, for each individual, a vector of length 392 that represents the regional SC-FC coupling strength, or similarity of a region's structural and functional connectivity fingerprints, for each of the 392 regions in the atlas.

Quantifying relationships between SC-FC coupling, age, sex and cognition

There are several different covariates that we hypothesized may have significant relationships with SC-FC coupling, namely, age, sex, total cognition, intracranial volume (ICV) and in-scanner head motion. The Total Cognition score, measured using the tests in the NIH toolbox, is the average of the crystallized score (including Picture Vocabulary and Reading Recognition measures) and fluid score (including Dimensional Change Card Sort, Flanker Inhibitory Control and Attention, Picture Sequence Memory, List Sorting, and Pattern Comparison measures). To calculate in-scanner head motion for each subject, we averaged the frame-wise displacement over each volume in the fMRI time series, and then took the average across the four fMRI scans. Finally, using a generalized linear model (GLM) approach, we assessed regional associations between SC-FC coupling and in-scanner motion, demographics and cognitive scores, plus three interaction terms (age*cognition, sex*cognition and ICV*motion). The three interaction terms we included in the GLM were those pairs of variables that we hypothesized may have non-negligible interactions.

$$y_k = \beta_0 + \sum_{i=1}^{8} \beta_i x_i$$

where y_k is the SC-FC coupling of length n (number of subjects) for region $k = 1, 2, ...392$, β_0 is the intercept and β_i are the coefficients for each covariate x_i, also a vector of length n. SC-FC coupling values were Fisher r-to-z transformed for improving normality. All p values for the regression coefficients were FDR corrected for multiple corrections and analyzed for significance at a level of $\alpha = 0.05$.

Quantifying the heritability of SC-FC coupling

LME models were developed to disentangle inter- versus intra-subject variation[61,62]. This LME approach was recently adapted for and applied to HCP data to quantify heritability of functional connectome fingerprints with respect to the inter-subject component, while removing the effect of transient changes across observations of a single subject[26]. This approach allows examination of the association between the genetic relationship and phenotypic similarity, while accounting for shared environment of siblings. Specifically, we write the following:

$$y_{ij} = x_{ij}\beta + \gamma_i + \varepsilon_{ij}$$

where $i = 1, 2, ..., n$ and $j = 1, 2, ...m_i$. m_i is the total number of repeated measures for subject i. The variable y_{ij} is the phenotype measurement for subject i for measurement j, x_{ij} contains all the q covariates while the vector β, also of length q, contains the unknown fixed population-level effects. The scalar γ_i donates the subject-specific deviation from the population mean and ε_{ij} describes denotes the intra-subject measurement error (transient component) of y_{ij} and is assumed to be independent of the random effects and independent between repeated measurements. Stacking all subjects and all repeated observations into a single vector, we have

$$\mathbf{y} = \mathbf{x}^{\mathbf{T}}\beta + \mathbf{T}\gamma + \varepsilon,$$

where $\mathbf{y}$ is the phenotype vector of length $n_{total} = \sum_{i=1}^{n} m_i$, $\mathbf{x}$ is the covariate matrix of dimension $q \times n_{total}$, T is a block diagonal matrix of dimension $n_{total} \times n_{subj}$, γ is a vector of length n_{subj} and ε is a vector of length n_{total}. We consider γ to be the sum of three different effects: additive genetic effect $\mathbf{g} \sim N(0, \sigma_A^2\mathbf{K})$, shared (common) environmental effect $\mathbf{c} \sim N(0, \sigma_C^2\Lambda)$ and unique (subject-specific) environmental effect $\mathbf{e} \sim N(0, \sigma_E^2\mathbf{I}_{\mathbf{n_{total}}})$. Here, σ_A^2, σ_C^2 and σ_E^2 are the additive genetic variance, common environmental variance and unique environmental variance, respectively. The matrix $\mathbf{K}$ is the $m \times m$ genetic similarity matrix derived from the pedigree information where K_{ij} is 1 for monozygotic twins, $1/2$ for dizygotic twins and full siblings and 0 for unrelated individuals. The matrix Λ is an $n_{subj} \times n_{subj}$ matrix indicating shared environment, that is, if the two subjects i and j have the same parents then Λ_{ij} is set to 1, otherwise it is set to 0. Finally, the matrix $\mathbf{I}_{\mathbf{n_{total}}}$ is the identity matrix of size $n_{subj} \times n_{subj}$. Intra-subject variation is assumed to follow a Gaussian distribution, $\varepsilon \sim N(0, \sigma_M^2\mathbf{I}_{\mathbf{n_{total}}})$. Thus, the covariance matrix of $\mathbf{y}$ is

$$\mathbf{cov}[\mathbf{y}] = \sigma_A^2\mathbf{TKT}^T + \sigma_C^2\mathbf{T}\Lambda\mathbf{T}^T + \sigma_E^2\mathbf{TT}^T + \sigma_M^2\mathbf{I}_{\mathbf{n_{total}}}.$$

Finally, we can define the non-transient heritability of a given trait as the proportion of stable, non-transient inter-subject variation that can be explained by genetic variation in the population as

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_C^2 + \sigma_E^2}$$

Unbiased estimates of the variance components σ_A^2, σ_C^2, σ_E^2 and σ_M^2 were obtained using the ReML algorithm[63]. We estimated the nontransient heritability of regional SC-FC coupling (4 measurements per subject), SC node strength as calculated via the sum of rows, excluding the diagonal (1 measurement per subject) and FC node strength as calculated via the sum of absolute value of rows, excluding the diagonal (4 measurements per subject). SC-FC coupling, FC node degree and SC node degree were standardized before calculating heritability. Age, sex and handedness were taken as fixed-effect covariates in each of the heritability models.

Data availability

HCP data are publicly available at `www.humanconnectome.org`. Certain HCP data are restricted to protect subject privacy, such as genetic, medical, and neuropsychiatric information.

Code availability

Python code to reproduce the main results of this paper is publicly available at `https://github.com/zijin-gu/scfc-coupling`. Preprocessing code is available upon request.

References

1. Honey, C. J. *et al.* Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci.* **106**, 2035–2040 (2009).
2. Shen, K. *et al.* Information processing architecture of functionally defined clusters in the macaque cortex. *J. Neurosci.* **32**, 17465–17476, DOI: 10.1523/JNEUROSCI.2709-12.2012 (2012). https://www.jneurosci.org/content/32/48/17465.full.pdf.
3. Mišić, B. *et al.* Network-Level Structure-Function Relationships in Human Neocortex. *Cereb. cortex (New York, N.Y. : 1991)* **26**, 3285–3296, DOI: 10.1093/cercor/bhw089 (2016).
4. Hermundstad, A. M. *et al.* Structural foundations of resting-state and task-based functional connectivity in the human brain. *Proc. Natl. Acad. Sci.* **110**, 6169–6174, DOI: 10.1073/pnas.1219562110 (2013). https://www.pnas.org/content/110/15/6169.full.pdf.
5. Uddin, L. Q. Complex relationships between structural and functional brain connectivity. *Trends Cogn. Sci.* **17**, 600 – 602, DOI: https://doi.org/10.1016/j.tics.2013.09.011 (2013). Special Issue: The Connectome.
6. Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**, 43–56, DOI: 10.1038/nrn2961 (2011).
7. Sanz Leon, P. *et al.* The virtual brain: a simulator of primate brain network dynamics. *Front. Neuroinformatics* **7**, 10, DOI: 10.3389/fninf.2013.00010 (2013).
8. Abdelnour, F., Voss, H. U. & Raj, A. Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *NeuroImage* **90**, 335–47, DOI: 10.1016/j.neuroimage.2013.12.039 (2014).
9. Vázquez-Rodríguez, B. *et al.* Gradients of structure–function tethering across neocortex. *Proc. Natl. Acad. Sci.* **116**, 21219–21227, DOI: 10.1073/pnas.1903403116 (2019). https://www.pnas.org/content/116/42/21219.full.pdf.
10. Chu, S.-H., Parhi, K. K. & Lenglet, C. Function-specific and Enhanced Brain Structural Connectivity Mapping via Joint Modeling of Diffusion and Functional MRI. *Sci. Reports* **8**, 4741, DOI: 10.1038/s41598-018-23051-9 (2018).
11. Kuceyeski, A. *et al.* The application of a mathematical model linking structural and functional connectomes in severe brain injury. *NeuroImage: Clin.* **11**, 635–647, DOI: 10.1016/j.nicl.2016.04.006 (2016).
12. Kuceyeski, A. F., Jamison, K. W., Owen, J. P., Raj, A. & Mukherjee, P. Longitudinal increases in structural connectome segregation and functional connectome integration are associated with better recovery after mild tbi. *Hum. Brain Mapp.* **40**, 4441–4456, DOI: 10.1002/hbm.24713 (2019). https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbm.24713.
13. Amico, E. & Goñi, J. Mapping hybrid functional-structural connectivity traits in the human connectome. *Netw. Neurosci.* **2**, 306–322, DOI: 10.1162/netn_a_00049 (2018).
14. Zimmermann, J., Griffiths, J. D. & McIntosh, A. R. Unique mapping of structural and functional connectivity on cognition. *J. Neurosci.* **38**, 9658–9667, DOI: 10.1523/JNEUROSCI.0900-18.2018 (2018). https://www.jneurosci.org/content/38/45/9658.full.pdf.
15. Liégeois, R. *et al.* Resting brain dynamics at different timescales capture distinct aspects of human behavior. *Nat. Commun.* **10**, 2317, DOI: 10.1038/s41467-019-10317-7 (2019).

16. Rykhlevskaia, E., Gratton, G. & Fabiani, M. Combining structural and functional neuroimaging data for studying brain connectivity: A review. *Psychophysiology* **45**, 173–187, DOI: https://doi.org/10.1111/j.1469-8986.2007.00621.x (2008). https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8986.2007.00621.x.

17. Kesler, S. R. *et al.* Disrupted brain network functional dynamics and hyper-correlation of structural and functional connectome topology in patients with breast cancer prior to treatment. *Brain Behav.* **7**, e00643, DOI: https://doi.org/10.1002/brb3.643 (2017). https://onlinelibrary.wiley.com/doi/pdf/10.1002/brb3.643.

18. Wang, J. *et al.* Alterations in Brain Network Topology and Structural-Functional Connectome Coupling Relate to Cognitive Impairment. *Front. Aging Neurosci.* **10**, 404, DOI: 10.3389/fnagi.2018.00404 (2018).

19. Demertzi, A. *et al.* Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci. Adv.* **5**, DOI: 10.1126/sciadv.aat7603 (2019). https://advances.sciencemag.org/content/5/2/eaat7603.full.pdf.

20. Barttfeld, P. *et al.* Signature of consciousness in the dynamics of resting-state brain activity. *Proc. Natl. Acad. Sci.* **112**, 887–892, DOI: 10.1073/pnas.1418031112 (2015).

21. Tagliazucchi, E., Crossley, N., Bullmore, E. T. & Laufs, H. Deep sleep divides the cortex into opposite modes of anatomical–functional coupling. *Brain Struct. Funct.* **221**, 4221–4234, DOI: 10.1007/s00429-015-1162-0 (2016).

22. Medaglia, J. D. *et al.* Functional alignment with anatomical networks is associated with cognitive flexibility. *Nat. human behaviour* **2**, 156–164 (2018).

23. Hagmann, P. *et al.* White matter maturation reshapes structural connectivity in the late developing human brain. *Proc. Natl. Acad. Sci.* **107**, 19067–19072 (2010).

24. Baum, G. L. *et al.* Development of structure–function coupling in human brain networks during youth. *Proc. Natl. Acad. Sci.* **117**, 771–778 (2020).

25. Zhao, S. *et al.* Sex Differences in Anatomical Rich-Club and Structural–Functional Coupling in the Human Brain Network. *Cereb. Cortex* DOI: 10.1093/cercor/bhaa335 (2020). Bhaa335, https://academic.oup.com/cercor/advance-article-pdf/doi/10.1093/cercor/bhaa335/34469467/bhaa335.pdf.

26. Ge, T., Holmes, A. J., Buckner, R. L., Smoller, J. W. & Sabuncu, M. R. Heritability analysis with repeat measurements and its application to resting-state functional connectivity. *Proc. Natl. Acad. Sci.* **114**, 5521–5526, DOI: 10.1073/pnas.1700765114 (2017). https://www.pnas.org/content/114/21/5521.full.pdf.

27. Sinclair, B. *et al.* Heritability of the Network Architecture of Intrinsic Brain Functional Connectivity. *NeuroImage* **121**, 243, DOI: 10.1016/J.NEUROIMAGE.2015.07.048 (2015).

28. Miranda-Dominguez, O. *et al.* Heritability of the human connectome: A connectotyping study. *Netw. Neurosci.* **2**, 175, DOI: 10.1162/NETN_A_00029 (2018).

29. Bertolero, M. A. *et al.* The network architecture of the human brain is modularly encoded in the genome. *arXiv* (2019). 1905.07606.

30. Van Essen, D. C. *et al.* The wu-minn human connectome project: an overview. *Neuroimage* **80**, 62–79 (2013).

31. Craddock, R. C., James, G. A., Holtzheimer III, P. E., Hu, X. P. & Mayberg, H. S. A whole brain fmri atlas generated via spatially constrained spectral clustering. *Hum. brain mapping* **33**, 1914–1928 (2012).

32. Liégeois, R., Santos, A., Matta, V., Van De Ville, D. & Sayed, A. H. Revisiting correlation-based functional connectivity and its relationship with structural connectivity. *Netw. Neurosci.* **0**, 1–25, DOI: 10.1162/netn_a_00166 (0). https://doi.org/10.1162/netn_a_00166.

33. Burt, J. B. *et al.* Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat. Neurosci.* **21**, 1251–1259, DOI: 10.1038/s41593-018-0195-0 (2018).

34. Barbas, H. & Rempel-Clower, N. Cortical structure predicts the pattern of corticocortical connections. *Cereb. cortex (New York, NY: 1991)* **7**, 635–646 (1997).

35. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci.* **113**, 12574–12579, DOI: 10.1073/pnas.1608282113 (2016). https://www.pnas.org/content/113/44/12574.full.pdf.

36. Marquis, R. *et al.* Spatial resolution and imaging encoding fmri settings for optimal cortical and subcortical motor somatotopy in the human brain. *Front. Neurosci.* **13**, 571, DOI: 10.3389/fnins.2019.00571 (2019).

37. Suárez, L. E., Markello, R. D., Betzel, R. F. & Misic, B. Linking structure and function in macroscale brain networks. *Trends Cogn. Sci.* (2020).

38. Sarwar, T., Tian, Y., Yeo, B., Ramamohanarao, K. & Zalesky, A. Structure-function coupling in the human connectome: A machine learning approach. *NeuroImage* 117609, DOI: 10.1016/j.neuroimage.2020.117609 (2020).

39. Power, J. D., Plitt, M., Laumann, T. O. & Martin, A. Sources and implications of whole-brain fmri signals in humans. *NeuroImage* **146**, 609 – 625, DOI: https://doi.org/10.1016/j.neuroimage.2016.09.038 (2017).

40. Lynch, C. J. *et al.* Prevalent and sex-biased breathing patterns modify functional connectivity MRI in young adults. *Nat. Commun.* **11**, 5290, DOI: 10.1038/s41467-020-18974-9 (2020).

41. Dhamala, E., Jamison, K., Sabuncu, M. & Kuceyeski, A. Sex classification using long-range temporal dependence of resting-state functional MRI time series. *Hum. brain mapping* **in press** (2020).

42. Hodgson, K. *et al.* Shared Genetic Factors Influence Head Motion During MRI and Body Mass Index. *Cereb. cortex (New York, N.Y. : 1991)* **27**, 5539–5546, DOI: 10.1093/cercor/bhw321 (2017).

43. Cao, R. *et al.* Abnormal Anatomical Rich-Club Organization and Structural–Functional Coupling in Mild Cognitive Impairment and Alzheimer's Disease. *Front. Neurol.* **11**, 53, DOI: 10.3389/fneur.2020.00053 (2020).

44. Cabral, J. *et al.* Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Sci. Reports* **7**, 5135, DOI: 10.1038/s41598-017-05425-7 (2017).

45. Mechelh, A., Humphreys, G. W., Mayall, K., Olson, A. & Price, C. J. Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. *Proc. Royal Soc. B: Biol. Sci.* **267**, 1909–1913, DOI: 10.1098/rspb.2000.1229 (2000).

46. Bogousslavsky, J., Miklossy, J., Deruaz, J. P., Assal, G. & Regli, F. Lingual and fusiform gyri in visual processing: A clinico-pathologic study of superior altitudinal hemianopia. *J. Neurol. Neurosurg. Psychiatry* **50**, 607–614, DOI: 10.1136/jnnp.50.5.607 (1987).

47. Glahn, D. C. *et al.* Genetic control over the resting brain. *Proc. Natl. Acad. Sci.* **107**, 1223–1228, DOI: 10.1073/pnas.0909969107 (2010). https://www.pnas.org/content/107/3/1223.full.pdf.

48. Anderson, K. M. *et al.* Heritability of individualized cortical network topography. *bioRxiv* DOI: 10.1101/2020.07.30.229427 (2020). https://www.biorxiv.org/content/early/2020/07/30/2020.07.30.229427.full.pdf.

49. Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B. & Bandettini, P. A. The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *NeuroImage* **44**, 893 – 905, DOI: https://doi.org/10.1016/j.neuroimage.2008.09.036 (2009).

50. Maier-Hein, K. H. *et al.* The challenge of mapping the human connectome based on diffusion tractography. *Nat. Commun.* **8**, 1349, DOI: 10.1038/s41467-017-01285-x (2017).

51. Sarwar, T., Ramamohanarao, K. & Zalesky, A. Mapping connectomes with diffusion mri: deterministic or probabilistic tractography? *Magn. Reson. Medicine* **81**, 1368–1384, DOI: https://doi.org/10.1002/mrm.27471 (2019). https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.27471.

52. Smith, R. E., Tournier, J.-D., Calamante, F. & Connelly, A. Sift2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography. *Neuroimage* **119**, 338–351 (2015).

53. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–124 (2013).

54. Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A. & Sijbers, J. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. *NeuroImage* **103**, 411–426 (2014).

55. Tournier, J. D., Calamante, F. & Connelly, A. Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. In *Proceedings of the international society for magnetic resonance in medicine*, vol. 1670 (Ismrm, 2010).

56. Smith, R. E., Tournier, J.-D., Calamante, F. & Connelly, A. Anatomically-constrained tractography: improved diffusion mri streamlines tractography through effective use of anatomical information. *Neuroimage* **62**, 1924–1938 (2012).

57. Griffanti, L. *et al.* Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. *Neuroimage* **95**, 232–247 (2014).

58. Salimi-Khorshidi, G. *et al.* Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* **90**, 449–468 (2014).

59. Smith, S. M. *et al.* Resting-state fmri in the human connectome project. *Neuroimage* **80**, 144–168 (2013).

60. Li, J. *et al.* Global signal regression strengthens association between resting-state functional connectivity and behavior. *NeuroImage* **196**, 126 – 141, DOI: https://doi.org/10.1016/j.neuroimage.2019.04.016 (2019).

61. Laird, N. M. & Ware, J. H. Random-effects models for longitudinal data. *Biometrics* **38**, 963–974 (1982).

62. Molenberghs, G. & Verbeke, G. *Linear mixed models for longitudinal data* (Springer, 2000).

63. Patterson, D. & Thompson, R. Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**, 545–554 (1971).

64. Mitchell, S. M., Lange, S. & Brus, H. Gendered citation patterns in international relations journals. *Int. Stud. Perspectives* **14**, 485–492 (2013).

65. Maliniak, D., Powers, R. & Walter, B. F. The gender citation gap in international relations. *Int. Organ.* **67**, 889–922 (2013).

66. Caplar, N., Tacchella, S. & Birrer, S. Quantitative evaluation of gender bias in astronomical publications from citation counts. *Nat. Astron.* **1**, 1–5 (2017).

67. Dion, M. L., Sumner, J. L. & Mitchell, S. M. Gendered citation patterns across political science and social science methodology fields. *Polit. Analysis* **26**, 312–327 (2018).

68. Dworkin, J. D. *et al.* The extent and drivers of gender imbalance in neuroscience reference lists. *Nat. Neurosci.* **23**, 918–926 (2020).

Acknowledgements

This work was supported by the following grants: R21 NS104634-01 (AK), R01 NS102646-01A1 (AK), R01 LM012719 (MS), R01 AG053949 (MS), NSF CAREER 1748377 (MS), and NSF NeuroNex Grant 1707312 (MS). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Author contributions statement

A.K. and M.S. conceived the experiments and interpreted the results, Z.G. conducted the experiments, analysed and interpreted the results. K.J. processed the imaging data and interpreted the results. Z.G. and A.K. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Citation gender diversity statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minorities are under-cited relative to the number of such papers in the field[64–68]. Here we sought to proactively consider choosing references that reflect the diversity of the field in thought, form of contribution, gender, and other factors. We obtained predicted gender of the first and last author of each reference by using databases that store the probability of a name being carried by a woman[68]. By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 10.13% woman(first)/woman(last), 11.7% man/woman, 18.06% woman/man, 60.12% man/man. This method is limited in that a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or transgender people. We look forward to future work that could help us to better understand how to support equitable practices in science.