

# Neuroimaging and multiomics reveal cross-scale circuit abnormalities in schizophrenia

Received: 2 November 2022

Accepted: 21 July 2023

Published online: 28 August 2023



Meng Wang<sup>1</sup>, Hao Yan<sup>2,3</sup>, Xiaohan Tian<sup>1</sup>, Weihua Yue<sup>2,3</sup>, Yong Liu<sup>4</sup>, Lingzhong Fan<sup>5,6,7</sup>, Ke Hu<sup>5,6</sup>, Yuqing Sun<sup>1</sup>, Yuxin Zhao<sup>5,6</sup>, Jing Lou<sup>1</sup>, Ming Song<sup>5,6</sup>, Peng Li<sup>2,3</sup>, Jun Chen<sup>8</sup>, Yunchun Chen<sup>9</sup>, Huaning Wang<sup>9</sup>, Wenming Liu<sup>9</sup>, Zhigang Li<sup>10</sup>, Yongfeng Yang<sup>11,12</sup>, Hua Guo<sup>10</sup>, Luxian Lv<sup>11,12</sup>, Jun Yan<sup>2,3</sup>, Huiling Wang<sup>13</sup>, Hongxing Zhang<sup>11,12,14</sup>, Huawang Wu<sup>15</sup>, Yuping Ning<sup>15</sup>, Lin Lu<sup>2,3</sup>, Dai Zhang<sup>2,3,16</sup>, Ang Li<sup>17</sup>✉, Tianzi Jiang<sup>5,6,7,18,19</sup>✉ & Bing Liu<sup>1,20,21</sup>✉

Schizophrenia (SCZ) is a highly heterogeneous disorder with diverse clinical manifestations and macro- and microscale biological variations, usually observed at dissociable levels. Here we propose a cross-scale, circuit-based framework to connect heterogeneous clinical symptoms, large-scale brain circuit dysfunctions, and genetic, molecular and cellular abnormalities in SCZ. Using connectomic and predictive models on three independent neuroimaging datasets ( $n = 1,199$ , including patients with SCZ and healthy controls), we first identified two macroscale dysconnectivity dimensions for corticocortical and corticostriatal circuits, each associated with specific clinical symptoms. We then associated macroscale dysconnectivity with disrupted cellular circuits using extended imaging transcriptomic and genetic analyses on multiomics data. Our findings suggest a two-dimensional cross-scale heterogeneity model of SCZ, which reveals how distinct genetic disruptions affect specific cellular-level deficits, resulting in system-level brain circuit dysconnectivity responsible for the heterogeneous symptoms in SCZ. These findings significantly improve our understanding of cross-scale heterogeneity in SCZ, advancing its pathophysiology and treatment development.

Schizophrenia (SCZ) is a complex and devastating psychiatric disorder with an extremely heterogeneous profile of positive and negative symptoms and cognitive impairments, with distinct levels of prominence across time and individuals<sup>1</sup>. Aside from this clinical heterogeneity, the complex nature of SCZ likewise manifests as significant heterogeneity across different biological systems, from macro- to microscale<sup>2–5</sup>. However, most SCZ research concentrates on certain selective or frequently separate scales, with few attempts to obtain an overall perspective spanning clinical domains and diverse biological systems. Such isolations overlook intrinsically correlated abnormalities across different scales, which may lead to the current absence of consensus regarding the pathophysiology and effective treatment strategies of SCZ.

Considerable multiscale evidence has demonstrated the characteristics of biological heterogeneity in SCZ. At the macroscale level, multifarious non-invasive neuroimaging techniques<sup>6</sup>, especially functional magnetic resonance imaging (fMRI)<sup>7</sup>, coupled with emerging connectomic methodologies<sup>8</sup>, have revealed large-scale corticocortical and cortico-subcortical dysconnectivity in SCZ<sup>9–16</sup> and their potential associations with clinical manifestations and symptoms<sup>15–23</sup>. However, the underlying microscopic variations of brain dysconnectivity characteristics relating to clinical manifestations remain obscure. At the microscale level, several independent lines of evidence have indicated that the molecular and cellular abnormalities in SCZ probably involve genetic risk of common and rare genetic variants<sup>24,25</sup> and disruptions

of neuronal microcircuits. Specifically, the microscopic cellular alterations in SCZ are prominent in terms of cellular composition revealed by postmortem studies<sup>26,27</sup> and a variety of pathologically SCZ-relevant cell types, including excitatory neurons, cortical interneurons, medium spiny neurons, and non-neuronal cells such as oligodendrocytes, astrocytes and so on, were identified by mapping genetic risk loci from SCZ genome-wide association studies (GWAS) onto brain neurons<sup>28,29</sup>, partitioning SCZ heritability with polygenic approaches<sup>30,31</sup>, single-cell dissection of SCZ<sup>32</sup>, and using patient-derived induced pluripotent stem cell (iPSC)<sup>33,34</sup> and brain organoid<sup>35</sup> models. However it remains unknown how the microscale mechanisms can be translated to heterogeneous clinical manifestations in patients with SCZ.

To address the cross-scale heterogeneity of SCZ, here we aim to develop an integrated framework that incorporates individual genetic risk, imaging-based functional circuits, clinical manifestations and group-level multiomics information to disentangle the circuit-based abnormalities of SCZ from a dimensionally resolved perspective. Importantly, neuroimaging has a crucial role in facilitating the crosstalk and integration of clinical and biological information across different scales within this framework. To achieve this, two prerequisites should be met: one is establishing a robust relationship between neuroimaging and heterogeneous clinical phenotypes, preferentially generalizing to different independent cohorts; second, the corresponding macroscale neuroimaging signatures should be plausibly decoded into specific genetic risk, and molecular and cellular mechanisms. Such a framework can integrate SCZ abnormalities at clinical, macroscale and microscale levels and probably provide comprehensive insights into the SCZ complexity.

In this study, we started with macroscale heterogeneity of SCZ measured by neuroimaging techniques and focused on two representative dimensions of large-scale brain circuits (that is, corticocortical and corticostriatal), which were characterized by principal cortical functional gradients<sup>36</sup> and striatum-to-cortex functional connectivity based on resting-state fMRI data, respectively. Using three independent neuroimaging datasets (combined  $n > 1,400$ ) including patients with SCZ and healthy controls, we first identified robust and reproducible macroscale functional dysconnectivity patterns in SCZ. Then we utilized multivariate statistical methods to establish consistent associations between functional dysconnectivity and clinical heterogeneity in patients with SCZ. Finally, we performed comprehensive association analyses by integrating two macroscale brain dysconnectivity with postmortem whole-brain transcriptome, cortical single-cell sequencing, population-based SCZ GWAS and person-specific genome-wide genotyping data to plausibly determine linked dimensions of functional dysconnectivity and cellular abnormalities in SCZ.

## Results

### Two macroscale dysconnectivity dimensions in SCZ

We first characterized macroscale functional dysconnectivity in SCZ and evaluated its reproducibility and robustness by analyzing resting-state fMRI data from three large-sample independent neuroimaging cohorts: SCZ-I, SCZ-II and the Center for Biomedical Research Excellence (COBRE)<sup>37</sup>, including over 1,400 patients with SCZ and normal control (NC) participants altogether. The SCZ-I and SCZ-II datasets were obtained from seven Chinese sites using the same protocol, and images were acquired on 3.0T Siemens or GE MRI scanners, respectively. The COBRE dataset was publicly available and images were acquired on a 3.0T Siemens TimTrio scanner. All participants in both the SCZ-I and SCZ-II datasets were Han Chinese, whereas most participants in the COBRE dataset were either Caucasian or Hispanic. After quality-control screening, there remained 625 participants (318 SCZ and 307 NC) in the SCZ-I dataset, 415 participants (195 SCZ and 220 NC) in the SCZ-II dataset, and 159 participants (72 SCZ and 87 NC) in the COBRE dataset (Methods). There were no significant differences in age and sex between the SCZ and the NC groups in the three datasets (SCZ-I,

two-sample two-sided  $t$ -test for age,  $t = -0.70$ ,  $P = 0.48$ , chi-square test for sex,  $\chi^2 = 0.003$ ,  $P = 0.95$ ; SCZ-II, age,  $t = -1.81$ ,  $P = 0.07$ , sex,  $\chi^2 = 0.23$ ,  $P = 0.62$ ; COBRE, age,  $t = -0.64$ ,  $P = 0.51$ , sex,  $\chi^2 = 1.50$ ,  $P = 0.21$ ; Supplementary Table 1).

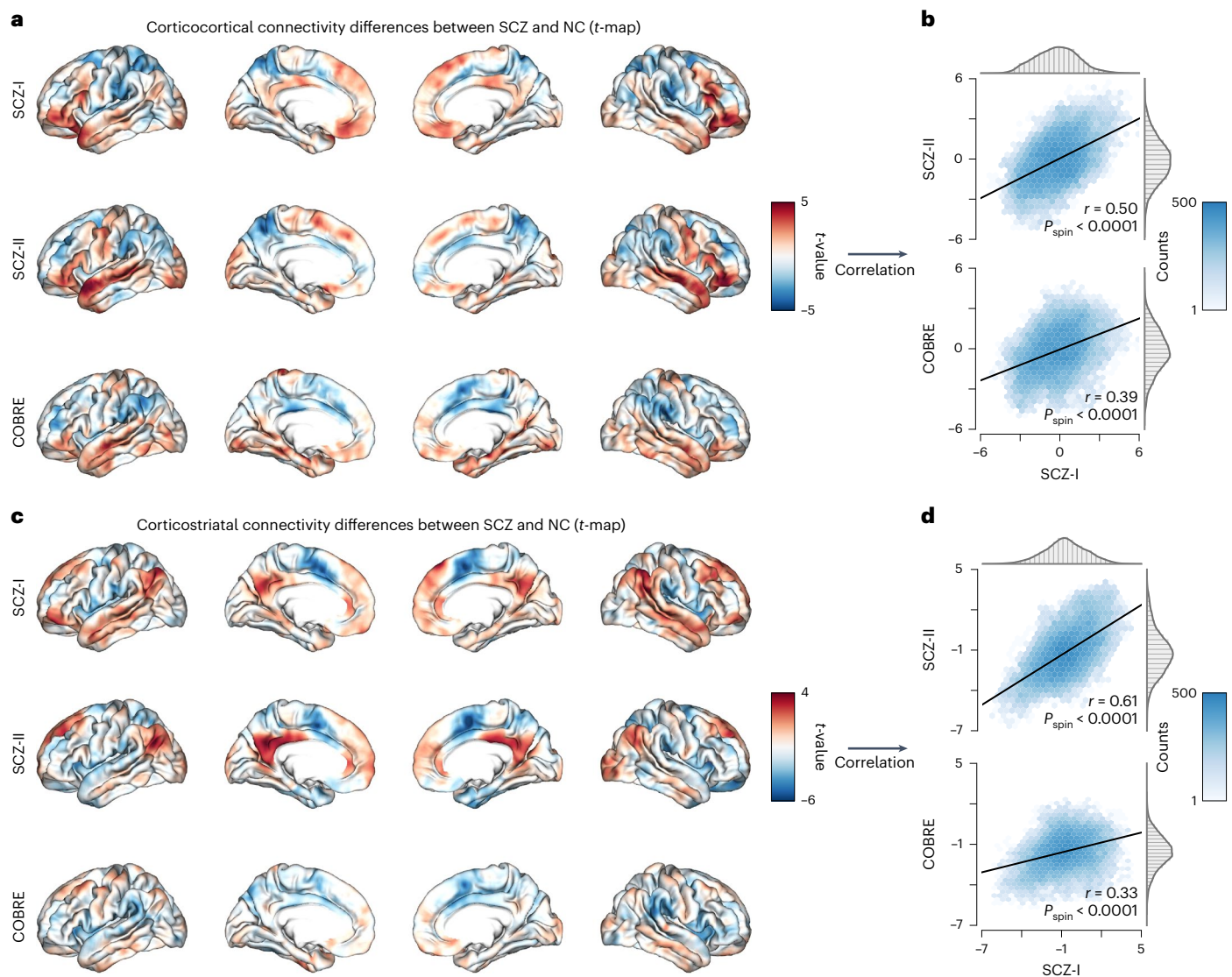
Here we concentrated on two dominant macroscale brain circuits, that is, corticocortical and corticostriatal connectivity, relevant to the pathophysiological characteristics of SCZ. Specifically, the corticocortical connectivity was characterized by the first cortical functional gradient<sup>36</sup>, derived by performing dimensionality reduction on the cortex-wide connectivity matrix with the diffusion maps method<sup>38</sup>. This gradient component was observed to account for the largest amount of connectivity variance (approximately 24%, and similar in SCZ and NC) in the three datasets (Supplementary Fig. 1), suggesting its potential for the low-dimensional representation of corticocortical connectivity. The striatum-to-whole-cortex functional connectivity (corticostriatal connectivity) was also investigated as studies have suggested critical roles of the striatum and its associated circuits to the cortex in SCZ pathology<sup>15,16,39,40</sup>. The averaged patterns of corticocortical and corticostriatal connectivity in SCZ remarkably resembled those in NC across the three datasets (Supplementary Figs. 2 and 3). Furthermore, the variance of corticocortical and corticostriatal connectivity was significantly lower in cross-site images of individual participants compared with single-site images with different participants (Methods) (two-sample  $t$ -test (two-sided), false discovery rate (FDR) correction at 0.05 ( $P_{\text{FDR}} < 0.05$ ); Supplementary Tables 2 and 3), indicating the stability for both connectivity measures.

To identify the abnormalities of the two connectivity circuits in SCZ, we conducted cortical vertex-wise comparisons between the SCZ and NC groups using surface-based linear models while controlling for the effects of age, sex and site. The between-group differences (that is, dysconnectivity patterns indicated with uncorrected  $t$ -maps) showed that the corticocortical connectivity was mostly increased in the inferior frontal and middle temporal cortex and decreased in the superior parietal region in SCZ (Fig. 1a). However, for the corticostriatal connectivity, SCZ predominantly showed an increase in the superior frontal gyrus and a decrease in the supplementary motor area (Fig. 1c). To further evaluate the reproducibility of this result, we calculated spatial correlations between  $t$ -maps across different datasets and concluded that both corticocortical and corticostriatal dysconnectivity patterns were highly reproducible in SCZ (Fig. 1b,d; for corticocortical dysconnectivity, Pearson's  $r = 0.50$  between the SCZ-I and SCZ-II datasets, and 0.39 between the SCZ-I and COBRE datasets; for corticostriatal dysconnectivity,  $r = 0.61$  between the SCZ-I and SCZ-II datasets, and 0.33 between the SCZ-I and COBRE datasets), despite considerable variations in demographics, ethnicity, site and MRI platforms between the cohorts. All these correlations were significant ( $P_{\text{spin}} < 0.0001$ ) after being estimated with spin-based permutation tests, accounting for spatial autocorrelations<sup>41</sup>.

Notably, the dysconnectivity patterns were highly robust after controlling for several potential confounders and clinical factors, including head motions<sup>42</sup> (Supplementary Fig. 4), global signal regression (GSR)<sup>43</sup> in fMRI preprocessing (Supplementary Fig. 5), doses of antipsychotic medications (Supplementary Fig. 6), course of illness (Supplementary Fig. 7) and connectivity matrix thresholding<sup>44</sup> (Supplementary Fig. 8). Besides, the dysconnectivity patterns were also highly consistent when using the ComBat<sup>45</sup> harmonization method to combine multi-site data instead of regressing the site effects in linear models (Supplementary Fig. 9). Taken together, we identified two highly reproducible and robust corticocortical and corticostriatal dysconnectivity dimensions in SCZ, furnishing in vivo neuroimaging evidence for the brain dysconnectivity hypothesis of SCZ<sup>46–50</sup>.

### Link corticocortical dysconnectivity to specific symptoms

Many studies have suggested that large-scale functional dysconnectivity could be responsible for the clinical variations in SCZ<sup>15–21</sup>, albeit with



**Fig. 1 | Highly reproducible corticocortical and corticostriatal dysconnectivity patterns in SCZ.** **a, c.** Cortical renderings show corticocortical (**a**) and corticostriatal (**c**) connectivity differences (that is, uncorrected *t*-maps) between patients with SCZ and NC participants in the SCZ-I (SCZ, *n* = 318; NC, *n* = 307), SCZ-II (SCZ, *n* = 195; NC, *n* = 220) and COBRE (SCZ, *n* = 72; NC, *n* = 87) datasets. Based on rsfMRI data, the corticocortical and corticostriatal connectivity were characterized by the cortical functional gradients and striatum-to-whole-cortex connectivity, respectively. The *t*-maps (that is, dysconnectivity patterns) were derived from cortical vertex-wise between-group comparisons (contrast = SCZ – NC) using surface-based linear models while regressing out effects of the age, sex and site. The warm or cool color indicates

that the connectivity is increased or decreased in SCZ, respectively. **b, d.** Hexbin plots with marginal distributions, and regression lines of best fit show spatial correlations between the corticocortical (**b**) or corticostriatal (**d**) *t*-maps. Both corticocortical and corticostriatal dysconnectivity patterns were highly consistent across three datasets (corticocortical,  $r = 0.50$ ,  $P_{\text{spin}} < 0.0001$  between SCZ-I and SCZ-II datasets and  $r = 0.39$ ,  $P_{\text{spin}} < 0.0001$  between SCZ-I and COBRE datasets; corticostriatal,  $r = 0.61$ ,  $P_{\text{spin}} < 0.0001$  between SCZ-I and SCZ-II datasets and  $r = 0.33$ ,  $P_{\text{spin}} < 0.0001$  between SCZ-I and COBRE datasets). Cortical cross-vertex Pearson's correlation ( $r$ ) (two-sided) was calculated, and the significance ( $P_{\text{spin}}$ ) was determined using the spin-based permutation test.

a lack of consistency. Previous evidence has shown that the individual functional connectivity profile may be used as an intrinsic 'fingerprint' to reliably characterize individual differences in behaviors<sup>51</sup>. Here we separately explored whether our identified two dysconnectivity characteristics could be related to symptom severity in patients with SCZ and overlapped with pertinent functional domains. We first investigated corticocortical dysconnectivity and especially centered on those connectivity features that show significant dysconnectivity to characterize clinical symptoms.

Ten regions of significant between-group differences in the corticocortical connectivity were identified in the SCZ-I dataset (Fig. 2a) after correcting for multiple comparisons at a family-wise error (FWE)

of 0.05, based on the random field theory<sup>52</sup> with a cluster-defining threshold (CDT) of 0.001 ( $P_{\text{FWE}} < 0.05$ , CDT = 0.001), and excluding small survived clusters with less than 30 connected vertices. These ten clusters remained after applying the non-parametric permutation test correction<sup>53,54</sup> ( $P_{\text{perm}} < 0.05$ ; Extended Data Fig. 1a and Methods). When using distinct choices for the CDT (0.005, 0.01), these core clusters were still detected, albeit with potential variations in size and number (Supplementary Fig. 10). Specifically, the corticocortical connectivity in SCZ was significantly increased in four clusters (left middle temporal pole; bilateral inferior frontal gyrus of the orbital part; right inferior frontal gyrus of the triangular part) and decreased in six clusters (bilateral superior parietal gyrus; left precuneus and inferior



parietal gyrus, excluding supramarginal and angular gyri; right insula and supramarginal gyrus; Supplementary Table 4). Furthermore, corticocortical connectivity differences among these clusters showed high reproducibility (Fig. 2b) when performing cluster-wise comparisons with general linear models controlling for the effects of age, sex and site. All ten corticocortical dysconnectivity regions were consistent in the direction of effect size across the three datasets, nine of which were significantly replicated in at least two datasets after FDR correction at 0.05 ( $P_{\text{FDR}} < 0.05$ ; Supplementary Fig. 11).

The above analysis identified a total of ten significant corticocortical dysconnectivity regions with high reproducibility across the three datasets. To examine their potential associations with individual symptom severity, we performed regression analyses to predict clinical assessments of the Positive and Negative Syndrome Scale (PANSS)<sup>55</sup> in patients with SCZ, using mean values of the corticocortical connectivity from each of the ten regions based on generalized additive models (GAMs)<sup>56–58</sup>, a robust and flexible framework for modeling nonlinear relationships. The predictive models were trained and selected via nested tenfold cross-validation<sup>59</sup> (Extended Data Fig. 2) in the SCZ-I dataset (SCZ = 318) to overcome biased estimates and were tested independently in the SCZ-II (SCZ = 195) and COBRE (SCZ = 72) datasets (Methods). The Pearson's correlation and mean absolute error (MAE) between the observed and predicted scores, as well as permutation-based significance ( $P_{\text{perm}}$ ), were calculated to quantify the prediction performance of the model. We found that corticocortical dysconnectivity characteristics could significantly predict PANSS five-factor domains<sup>60</sup> of negative (SCZ-I, Pearson's  $r = 0.28$ ,  $P_{\text{perm}} < 0.0001$ ; SCZ-II,  $r = 0.19$ ,  $P_{\text{perm}} = 0.0048$ ; COBRE,  $r = 0.24$ ,  $P_{\text{perm}} = 0.0165$ ), cognitive (SCZ-I,  $r = 0.38$ ,  $P_{\text{perm}} < 0.0001$ ; SCZ-II,  $r = 0.23$ ,  $P_{\text{perm}} = 0.0015$ ; COBRE,  $r = 0.21$ ,  $P_{\text{perm}} = 0.0288$ ) and excitement (SCZ-I,  $r = 0.30$ ,  $P_{\text{perm}} < 0.0001$ ; SCZ-II,  $r = 0.22$ ,  $P_{\text{perm}} = 0.0009$ ; COBRE,  $r = 0.21$ ,  $P_{\text{perm}} = 0.0226$ ) across the three datasets (Fig. 2e), whereas the effect was not observed for the positive and depression/anxiety domains. Furthermore, significant differences ( $P_{\text{FDR}} < 0.05$ ) were found in four (positive, negative, cognitive and excitement; Supplementary Fig. 12) out of the five-factor scores across the three datasets, indicating that variations in symptom severity distribution had no impact on the predictions.

Moreover, even with slight variations in cluster size and number ( $P_{\text{FWE}} < 0.05$ , CDT = 0.01; 15 clusters, Supplementary Fig. 10), we consistently observed significant predictions for the negative, cognitive and excitement symptoms, while the dimensions of positive and depression/anxiety remained unpredictable (Supplementary Table 5). To examine the influence of different PANSS factor models on prediction, we also used the three-factor division and conducted an identical predictive analysis. The results showed that corticocortical dysconnectivity features consistently predicted negative and general psychopathology subscales across the three datasets ( $P_{\text{perm}} < 0.05$ ), but not the

positive subscale (Supplementary Fig. 13). Given the broad spectrum of symptoms covered by the general psychopathology, which may not directly align with cognitive or excitement symptoms, our predictions showed consistency across distinct PANSS factor models, specifically in terms of the negative and positive domains.

To further investigate the specificity, we performed 100 predictive analyses in the SCZ-I dataset and observed significantly lower performance (Pearson's  $r$ ) for positive and depression/anxiety symptoms compared with the other three PANSS factors (two-sample  $t$ -test (two-sided),  $P_{\text{FDR}} < 0.05$ , Supplementary Table 6). In addition, as a supplementary specificity analysis, we formed high- and low-symptom subgroups within the SCZ-I and SCZ-II datasets by selecting patients with SCZ with top-30% and bottom-30% scores for a specific symptom dimension. Additional screening was used to confirm that there were no differences in the other four factors between the subgroups (Extended Data Fig. 3a,d,g). We observed significant spatial positive correlations between the differential maps derived from comparing the corticocortical connectivity profiles of high and low negative subgroups, as well as high and low cognitive subgroups, and the dysconnectivity  $t$ -map (Fig. 1a) (negative, Pearson's  $r = 0.46$ ,  $P_{\text{spin}} < 0.0001$ ; cognitive,  $r = 0.42$ ,  $P_{\text{spin}} < 0.0001$ ; Extended Data Fig. 3b,h). Furthermore, the connectivity strength of the ten clusters (Fig. 2a) significantly differed between the subgroups (negative,  $P = 2.14 \times 10^{-4}$ ; cognitive,  $P = 1.01 \times 10^{-8}$ ), aligning with the direction of differences observed in the SCZ versus NC comparison. However, the corticocortical connectivity differences between the high and low positive subgroups did not spatially correlate with the dysconnectivity pattern ( $r = -0.15$ ,  $P_{\text{spin}} = 0.082$ ), and no changes were observed in the cluster-level connectivity strength ( $P = 0.67$ ) (Extended Data Fig. 3e). These findings demonstrate that corticocortical dysconnectivity is specifically associated with the severity of negative and cognitive symptoms, but not positive symptoms, further supporting the predictive results.

To explore potential functional overlaps with the corticocortical dysconnectivity pattern, we used the Neurosynth<sup>61</sup> platform, which is a well-validated and publicly available resource, to decode  $t$ -maps (Methods). We observed that the term 'face/affective processing' was consistently associated with corticocortical dysconnectivity across the three datasets (Extended Data Fig. 4a and Supplementary Table 7) and may be linked to negative and cognitive symptoms<sup>62–64</sup>, offering partial support for our predictions.

### Link corticostriatal dysconnectivity to specific symptoms

As with corticocortical dysconnectivity, we conducted identical analyses to investigate the associations of corticostriatal dysconnectivity with individual symptom severity of SCZ, and with related functional domains. A total of eight regions indicating significant between-group differences in corticostriatal connectivity were identified in the SCZ-I

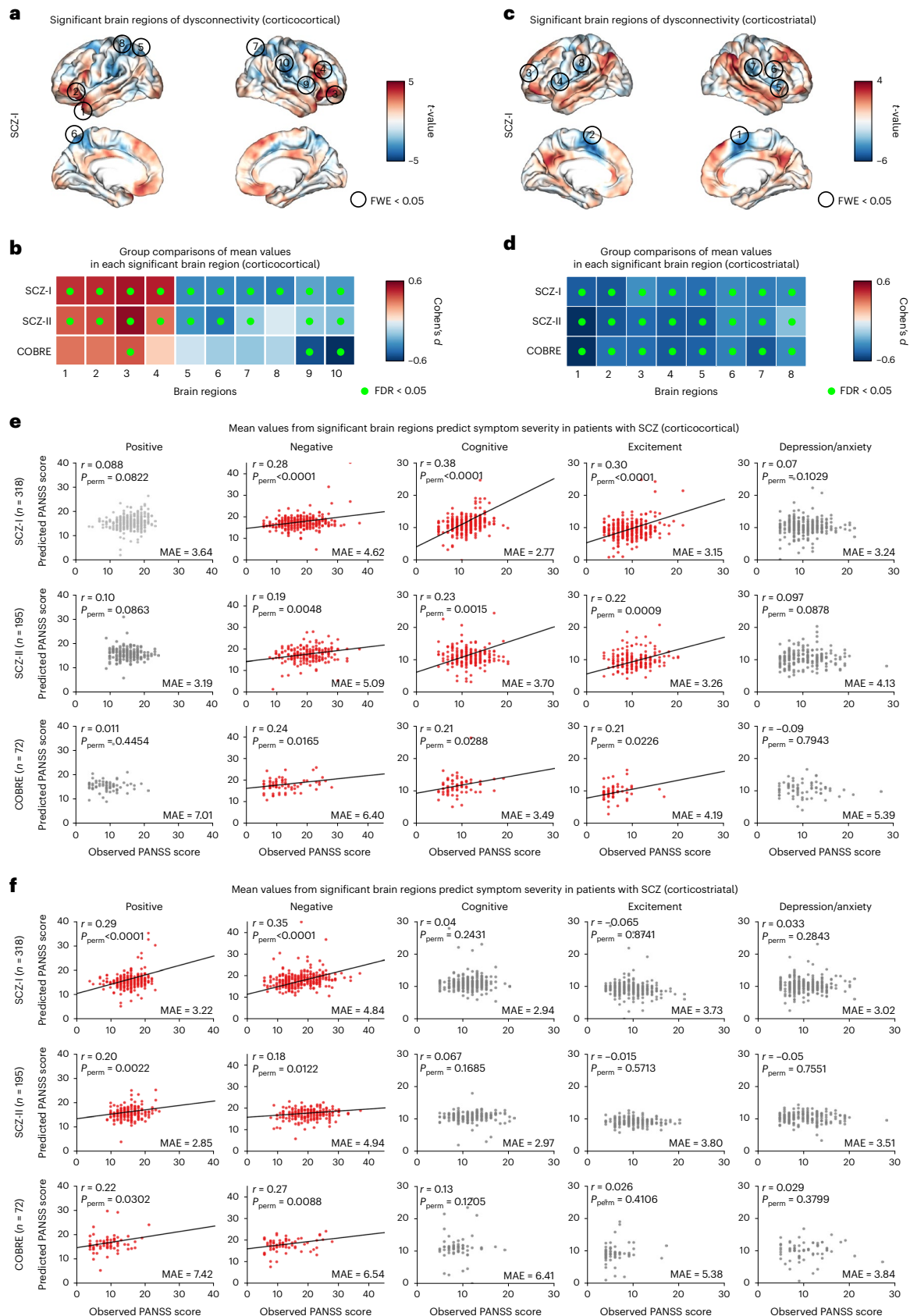
**Fig. 2 | Corticocortical and corticostriatal dysconnectivity associated with clinical symptoms.** **a**, In the SCZ-I dataset, ten regions denoting significant corticocortical dysconnectivity are labeled with circled Arabic numbers (from 1 to 10) on the uncorrected  $t$ -map. These regions were identified after correcting the corticocortical dysconnectivity  $t$ -map (from Fig. 1a, based on a two-sample  $t$ -test (two-sided)) for FWE at 0.05 with a CDT of 0.001 ( $P_{\text{FWE}} < 0.05$ , CDT = 0.001) and excluding small survived clusters (the number of connected vertices was less than 30). Corticocortical connectivity in the regions that are shown in warm (from regions 1 to 4) or cool (from regions 5 to 10) color are significantly increased or decreased in SCZ. **b**, Heatmap showing the between-group comparisons of mean corticocortical connectivity values in each of the ten regions (shown in **a**) in the three datasets, implemented with general linear models controlling for the effects of age, sex and site. The effect size measured with Cohen's  $d$  is given. The FDR method was applied to correct for multiple comparisons. The solid green circle indicates that between-group differences are significant ( $P_{\text{FDR}} < 0.05$ ). All ten significant corticocortical dysconnectivity regions were consistent in the direction of effect size across three datasets, nine of which were significantly replicated in at least two datasets.

**c**, Eight regions that represented significant corticostriatal dysconnectivity were identified in the SCZ-I dataset ( $P_{\text{FWE}} < 0.05$ , CDT = 0.001, the number of connected vertices was greater than 30). In all these regions, corticostriatal connectivity was significantly decreased in SCZ. **d**, Heatmap showing that all eight significant corticostriatal dysconnectivity regions were not only consistent in the direction of effect size but also significantly replicated across the three datasets ( $P_{\text{FDR}} < 0.05$ ). **e**, Mean corticocortical connectivity values extracted from the ten regions (shown in **a**) could significantly predict PANSS five-factor domains of negative, cognitive and excitement across the three datasets, but the effects were not observed in positive and depression/anxiety domains. **f**, Mean corticostriatal connectivity values in the eight regions could significantly predict PANSS positive and negative domains across three datasets, but not cognitive, excitement and depression/anxiety domains. Predictions were implemented with linear GAMs. Model training and selection were performed in the SCZ-I dataset using nested tenfold cross-validation, and further independent tests were completed in the SCZ-II and COBRE datasets. Pearson's  $r$  (two-sided) and MAE between the observed and predicted scores were calculated. The significance ( $P_{\text{perm}}$ ) was determined by permutation tests.



dataset (Fig. 2c;  $P_{\text{FWE}} < 0.05$ , CDT = 0.001; each region had more than 30 connected vertices). After applying a non-parametric permutation test<sup>53,54</sup> for multiple comparison correction, these eight clusters remained significant ( $P_{\text{perm}} < 0.05$ ; Extended Data Fig. 1b and Methods).

These eight clusters were consistently detected across different CDT values (0.005, 0.01), although there were variations in the boundaries and number of clusters (Supplementary Fig. 14). Importantly, the corticostriatal connectivity in SCZ exhibited decreases in all of the



eight regions (bilateral supplementary motor area; left middle frontal and postcentral gyri; bilateral insula; right rolandic operculum and supramarginal gyrus) (Supplementary Table 8), and between-group differences for all these regions were significantly replicated across the three datasets ( $P_{\text{FDR}} < 0.05$ ; Fig. 2d and Supplementary Fig. 15).

Based on GAMs, we observed that mean values of corticostriatal connectivity from the eight regions could significantly predict the PANSS five-factor domains<sup>60</sup> of positive (SCZ-I, Pearson's  $r = 0.29$ ,  $P_{\text{perm}} < 0.0001$ ; SCZ-II,  $r = 0.20$ ,  $P_{\text{perm}} = 0.0022$ ; COBRE,  $r = 0.22$ ,  $P_{\text{perm}} = 0.0302$ ) and negative (SCZ-I,  $r = 0.35$ ,  $P_{\text{perm}} < 0.0001$ ; SCZ-II,  $r = 0.18$ ,  $P_{\text{perm}} = 0.0122$ ; COBRE,  $r = 0.27$ ,  $P_{\text{perm}} = 0.0088$ ) across the three datasets (Fig. 2f), whereas the cognitive, excitement and depression/anxiety symptoms could not be predicted. The predictive outcomes were unaffected by variations in the distribution of the five-factor scores across the three datasets (Supplementary Fig. 12). Despite variations in cluster size and number ( $P_{\text{FWE}} < 0.05$ , CDT = 0.01; 12 clusters, Supplementary Fig. 14), our predictions for positive and negative symptoms remained significant, whereas predictions for the other three factors did not reach significance (Supplementary Table 9). When using the PANSS three-factor division, we found significant predictive capability of corticostriatal connectivity characteristics for both positive and negative symptoms (Supplementary Fig. 16), demonstrating the consistency of our predictions across different PANSS factor models.

In the SCZ-I dataset, we conducted 100 predictions to assess specificity, revealing significantly better predictive performance (Pearson's  $r$ ) for positive and negative symptoms compared with the other three factors (two-sample  $t$ -test (two-sided),  $P_{\text{FDR}} < 0.05$ ; Supplementary Table 10). In our subgroup analysis based on specific symptom dimensions, we found significant positive spatial correlations between the differential maps of high and low positive subgroups, as well as high and low negative subgroups, and the dysconnectivity  $t$ -map (Fig. 1c) (positive, Pearson's  $r = 0.35$ ,  $P_{\text{spin}} = 1.0 \times 10^{-4}$ ; negative,  $r = 0.22$ ,  $P_{\text{spin}} = 5.0 \times 10^{-4}$ ; Extended Data Fig. 3c,f). Moreover, the connectivity strength of the eight clusters (Fig. 2c) showed significant differences between the subgroups, consistent with the direction of differences observed in the SCZ versus NC comparison (positive,  $P = 4.78 \times 10^{-3}$ ; negative,  $P = 1.35 \times 10^{-14}$ ). However, the corticostriatal connectivity differences between the high and low cognitive subgroups did not show spatial correlation with the dysconnectivity pattern ( $r = -0.04$ ,  $P_{\text{spin}} = 0.50$ ), and no alterations were detected in the cluster-level connectivity strength ( $P = 0.12$ ; Extended Data Fig. 3i). These findings indicate that corticostriatal dysconnectivity is primarily linked to the severity of positive and negative symptoms, rather than cognitive symptoms, providing additional support for the predictive results.

Functional decodings of the corticostriatal dysconnectivity  $t$ -map using the Neurosynth database revealed significant and consistent correlations with three broad domains—cued attention, inhibition error and pain—across the three datasets (Extended Data Fig. 4b and Supplementary Table 11). These functional consequences may partially account for the manifestation of positive<sup>65</sup> and negative<sup>66</sup> symptoms, providing indirect support for our predicted results. The two distinct dysconnectivity that were specifically associated with differential clinical symptom profiles probably revealed two different dimensional circuits in the pathophysiology of SCZ.

### Bridge macroscale dysconnectivity to cellular abnormalities

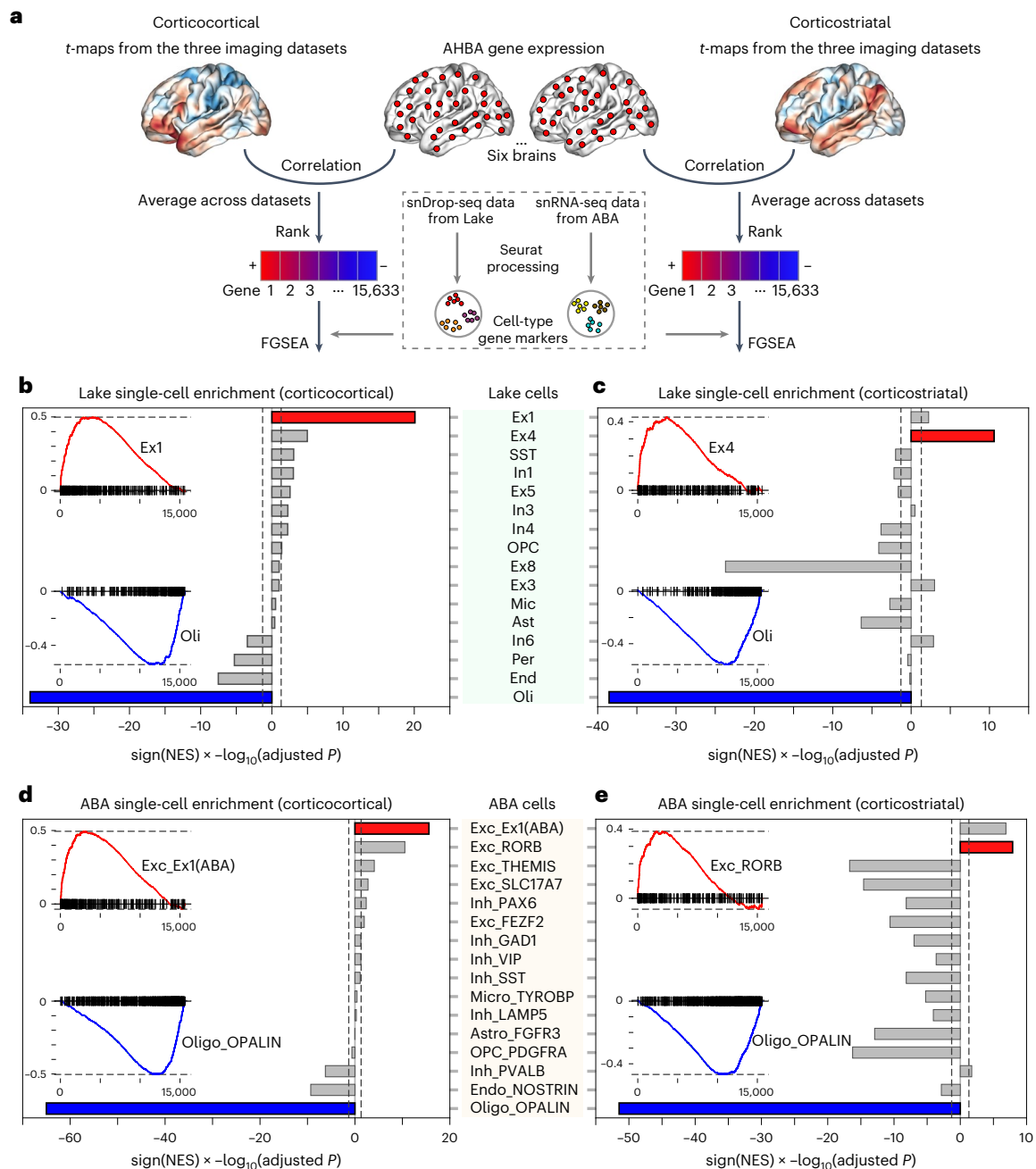
The SCZ pathophysiology implicates microscopic alterations in diverse cell types<sup>3,35</sup>, possibly inducing macroscopic variations of dysconnectivity. Therefore, it is crucial to establish cross-scale correspondence between different pathophysiological facets of SCZ. To this end, we explored the underlying cellular mechanisms of macroscale corticocortical and corticostriatal dysconnectivity in SCZ by integrating dysconnectivity patterns with whole-brain transcriptomics<sup>67</sup> and cortical single-cell sequencing datasets<sup>68,69</sup>.

The schematic diagram of our experimental design illustrates the major multiomics analysis pipeline (Fig. 3a). For the corticocortical or corticostriatal dysconnectivity, we first separately calculated spatial correlations between each dysconnectivity  $t$ -map from the three imaging datasets and cortical gene expression profiles (a total of 15,633 genes after preprocessing) from the Allen Human Brain Atlas (AHBA)<sup>67</sup> (Supplementary Tables 12–15 and 17–19). Gene-wise transcriptional correlations were then averaged across three datasets and ranked in descending order (Supplementary Tables 16 and 20). Finally, we analyzed the sorted gene lists with the fast gene-set enrichment analysis (FGSEA)<sup>70</sup> approach based on the cell-type-specific gene markers, which were derived from two independent single-cell datasets, that is, Lake<sup>68</sup> and ABA<sup>69</sup>. In particular, FGSEA examined whether AHBA expressions of cell gene markers were significantly spatially correlated with global dysconnectivity patterns of SCZ.

From the Lake dataset, 16 cell types with related gene markers were identified (Supplementary Table 21), in agreement with a previous study<sup>71</sup>. We found that corticocortical dysconnectivity correlated genes were positively enriched in the Ex1 excitatory neurons and negatively enriched in the oligodendrocytes, with the most significance (Fig. 3b and Supplementary Table 23). For corticostriatal dysconnectivity, the most significant positively and negatively enriched cells were the Ex4 excitatory neurons and oligodendrocytes, respectively (Fig. 3c and Supplementary Table 25). The positive or negative enrichment herein meant AHBA expressions of cell-type-specific genes were positively or negatively correlated with spatial patterns of dysconnectivity. Furthermore, these results can be well replicated when conducting the same analysis, according to another independent single-cell dataset (ABA) (Fig. 3d,e and Supplementary Tables 22, 24 and 26). It is noted that the ABA cells of Exc\_Ex1(ABA) and Exc\_RORB best matched the Lake cell profiles of Ex1 and Ex4<sup>69</sup>, respectively (Methods). The patterns of cell enrichment revealed by FGSEA were also stable compared with an alternative method, where the AHBA cortical expressions of cell-type-specific genes were averaged and spatially correlated with dysconnectivity. Moreover, we observed strong cell-wise correspondences between these two methods (corticocortical dysconnectivity, Lake cells, Spearman's  $r_s = 0.95$ ,  $P < 7.1 \times 10^{-9}$ , ABA cells,  $r_s = 0.94$ ,  $P < 5.4 \times 10^{-8}$ ; corticostriatal dysconnectivity, Lake cells,  $r_s = 0.93$ ,  $P < 1.0 \times 10^{-8}$ , ABA cells,  $r_s = 0.94$ ,  $P < 2.2 \times 10^{-8}$ ; Supplementary Table 27).

To facilitate statistical analysis, we conducted the entire above analysis at the individual level as well. We first calculated individual dysconnectivity by measuring the deviation of functional connectivity from the NC group (Methods). Each participant, including both patients with SCZ and healthy participants from the SCZ-I, SCZ-II and COBRE datasets (total  $n = 1,199$ ; 585 SCZ and 614 NC), was assigned a dysconnectivity deviation map using this method. Notably, the dysconnectivity map of healthy participants primarily represents deviations from the typical group distribution within the normal range. For individual corticocortical dysconnectivity, Ex1 and Oli were the top two cell types with the highest positive and negative normalized enrichment scores (NESs), respectively (Extended Data Fig. 5a). Meanwhile, Ex1 showed significantly higher NES scores compared with other cell types ( $P_{\text{FDR}} < 0.05$ ), while Oli showed significantly lower NES scores (negative) (Extended Data Fig. 5b). For individual corticostriatal dysconnectivity, Ex4 exhibited the most pronounced positive enrichment, whereas Oli demonstrated the most significant negative enrichment (Extended Data Fig. 5c). In pair-wise comparisons among cell types, Ex4 had the highest NES score, while Oli had the lowest (negative) (Extended Data Fig. 5d). In particular, the individual-level analysis demonstrated consistent replication in the ABA dataset (Extended Data Fig. 5e–h). These findings further strengthen the consistency and specificity of the association between the two dysconnectivity patterns and distinct cell types.

Intriguingly, Ex1 and Ex4 excitatory neurons are mainly located in the cortical upper layer (L2/3) and deeper layer (L5), respectively,



**Fig. 3 | Cell enrichment of dysconnectivity transcriptional correlates.**

**a**, Schematic diagram showing the procedures of single-cell transcriptional enrichment analysis for the corticocortical and corticostriatal dysconnectivity. For each kind of dysconnectivity, *t*-maps obtained from the three imaging datasets (SCZ-I, SCZ-II and COBRE) were separately spatially correlated with all AHBA gene expressions (15,633 genes altogether) using Pearson's correlation measure. The transcriptional correlates were then averaged across the three datasets and ranked in descending order (from positive to negative). Finally, based on single-cell pathways, the FGSEA method was applied to the sorted gene list. Two single-cell datasets Lake<sup>68</sup> and ABA<sup>69</sup> were used separately, following the same preprocessing method using the Seurat toolkit<sup>139</sup>. For each single-cell dataset, 16 cell types with gene markers were identified. **b–e**, Horizontal bars and enrichment plots showing the enrichment results of corticocortical (**b,d**)

and corticostriatal (**c,e**) dysconnectivity using Lake (**b,c**) and ABA (**d,e**) cells. Corticocortical dysconnectivity transcriptional correlated genes were most significantly positively (red bar) enriched in Ex1 or Exc\_Ex1 (ABA) and negatively (blue bar) enriched in Oli or Oligo\_OPALIN. The positively and negatively enriched cells with the most significance for corticostriatal dysconnectivity were Ex4 or Exc\_RORB and Oli or Oligo\_OPALIN, respectively. The bar length was quantified by  $\text{sign}(\text{NES}) \times -\log_{10}(\text{adjusted } P)$ , in which NES and adjusted *P* (two-sided, FDR correction) were given by the FGSEA method.  $\text{sign}(\cdot)$  is the sign function, which equals 1 if  $\text{NES} > 0$  and  $-1$  if  $\text{NES} < 0$ . The two vertical dashed lines indicate the FDR-corrected *P* value at 0.05. Within each panel, two enrichment plots for the most significant positively (red curve) and negatively (blue curve) enriched cells are shown. Each black line on the *x* axis is the position of a cell marker gene. The *y* axis represents the enrichment score.

based on which they are separately categorized into cortical projection neurons and subcortical projection neurons<sup>68,72</sup>. Thus, Ex1 and Ex4 neurons could construct the primary neural circuit projections between cortices and cortex-to-subcortex regions, respectively. The

consistency of these cross-scale findings may suggest a link between the macroscale circuit dysconnectivity and the microscale cellular circuit disruptions in SCZ. Furthermore, our results indicated that such cross-scale circuit association could be SCZ-specific by performing



the same analyses on an additional imaging dataset of major depressive disorder (MDD) (Extended Data Fig. 6, Supplementary Fig. 17 and Supplementary Tables 28–31).

To further gain mechanistic insights into different biological processes that may drive cell-type–dysconnectivity associations, we conducted Gene Ontology (GO) enrichment analyses for the top and bottom deciles of ranked genes (each decile included 1,563 genes) that were correlated with different dysconnectivity dimensions. The top genes for corticocortical and corticostriatal dysconnectivity were linked to synapse-related and cellular modification processes, respectively, while the bottom genes were primarily enriched in axon development processes (Extended Data Fig. 7 and Supplementary Tables 32–35). These biological processes may correspond to the properties of specific cell types and contribute to macroscale dysconnectivity.

### Genomic validation for dysconnectivity-related cell types

Given that SCZ is highly heritable<sup>73,74</sup>, its genetic risk or interaction with environmental factors<sup>75</sup> may contribute to neuronal dysfunctions and subsequently influence the SCZ-related macroscopic alterations. From a genetic perspective, we sought to validate the identified cellular mechanisms underlying macroscale dysconnectivity by analyzing multiple population-based SCZ GWAS datasets and person-specific data, which captured genetic variations linked to different cell types.

We performed MAGMA (Multi-marker Analysis of Genomic Annotation) gene-set property analysis<sup>76,77</sup> to scrutinize the GWAS data and determined which genes of specific cell types were preferentially enriched for the polygenic risk of SCZ (Methods). We analyzed three GWAS datasets: one from the Psychiatric Genomics Consortium (PGC) with East Asian (EAS) and European (EUR) populations<sup>78</sup>, and two datasets derived from our own case–control association analysis on combined SCZ-I and SCZ-II datasets (SCZ = 471, NC = 475), as well as the UK Biobank (UKB) dataset (SCZ = 886, NC = 274,508) respectively. Based on the three GWAS datasets and the inclusion of Lake cells (from the dorsal frontal cortex (DFC) and visual cortex (VIS)) and ABA cells (from the middle temporal gyrus (MTG)), we carried out a total of six MAGMA enrichment analyses. We consistently observed significant enrichment exclusively in Ex1 (Exc\_Exc1(ABA)) and Ex4 (Exc\_RORB) excitatory neurons across both Lake and ABA cells, as well as across multiple GWAS datasets (Fig. 4a, DFC, Ex1  $P = 0.018$ , Ex4  $P = 0.044$ ; VIS, Ex1  $P = 0.009$ , Ex4  $P = 0.031$ ; Fig. 4b, MTG, Exc\_Exc1(ABA)  $P = 0.004$ , Exc\_RORB  $P = 0.029$ ; Fig. 4c, DFC, Ex1  $P = 0.011$ , Ex4  $P = 0.023$ ; VIS, Ex1  $P = 0.006$ , Ex4  $P = 0.038$ ; Fig. 4d, MTG, Exc\_Exc1(ABA)  $P = 0.001$ , Exc\_RORB  $P = 0.049$ ; Fig. 4e, DFC, Ex1  $P = 0.0007$ , Ex4  $P = 0.013$ ; VIS, Ex1  $P = 0.002$ , Ex4  $P = 0.019$ ; Fig. 4f, MTG, Exc\_Exc1(ABA)  $P = 0.001$ , Exc\_RORB  $P = 0.021$ ; Supplementary Tables 36–38). When using an alternative partitioned linkage disequilibrium score regression (LDSC) method<sup>79</sup>, Ex1 (Exc\_Exc1(ABA)) and Ex4 (Exc\_RORB) remained primarily enriched (Extended Data Fig. 8; DFC, Ex1  $P = 6.3 \times 10^{-5}$ , Ex4  $P = 0.016$ ; VIS, Ex1  $P = 1.7 \times 10^{-4}$ , Ex4  $P = 0.002$ ; MTG, Exc\_Exc1(ABA)  $P = 0.008$ , Exc\_RORB

$P = 0.054$  (with marginal significance)), confirming the mapping between GWAS results and cell types identified in our MAGMA analysis.

Based on the person-specific genomic data from the SCZ-I and SCZ-II datasets (SCZ = 471, NC = 475), we calculated the cell-type-specific polygenic risk score (PRS) for each individual to investigate which cell types presented a higher aggregated genetic risk in patients with SCZ compared with that of NC participants. For each individual, the PRSs of both the Lake and ABA cells were calculated (Fig. 4g). Seven cell types showed a significantly high PRS in SCZ relative to the NC group ( $P_{\text{FDR}} < 0.05$ ) when calculating the PRS at a  $P$  threshold of 0.01 (Fig. 4h). Among these cell types, Ex1 was the most significant in terms of between-group differences in cell-type-specific PRSs, while Ex4 and Oli were among the first five significant cell types. In particular, the PRS differences in Ex1 and Ex4 neurons and Oli remained significant when evaluating additional four common  $P$  thresholds (0.001, 0.005, 0.05, 0.1) in PRS calculations (Supplementary Table 39). The PRS analysis of ABA cells replicated the findings, revealing significant between-group differences in Exc\_Exc1(ABA), Exc\_RORB neurons and Oligo\_OPALIN, which remained consistent across various  $P$  thresholds (Fig. 4i and Supplementary Table 40). In addition, we also performed an independent validation using the UKB dataset (SCZ = 886, NC = 274,508). To account for the substantial case–control imbalance, we partitioned the NC group into 310 subsets, each of which was matched in size to the SCZ group. We then conducted PRS comparisons between the SCZ group and each subset. We observed significant between-group differences in Ex1, Ex4 and Oli in over 80% of the comparisons (Fig. 4j and Supplementary Table 41), and these results were also replicated in ABA cells (Fig. 4k and Supplementary Table 42). Taken together, both population-based and individual-specific analyses endorse the genetic involvement of excitatory neurons in the pathogenesis of SCZ. However, the results also indicate that the genetic influence on oligodendrocytes is particularly evident in individual-level analyses, but not significant in population-based analyses (see ‘Discussion’). These observations lend further support to the implication of cell-type-specific abnormalities in SCZ.

### Individual-level investigation of cell-type specificity

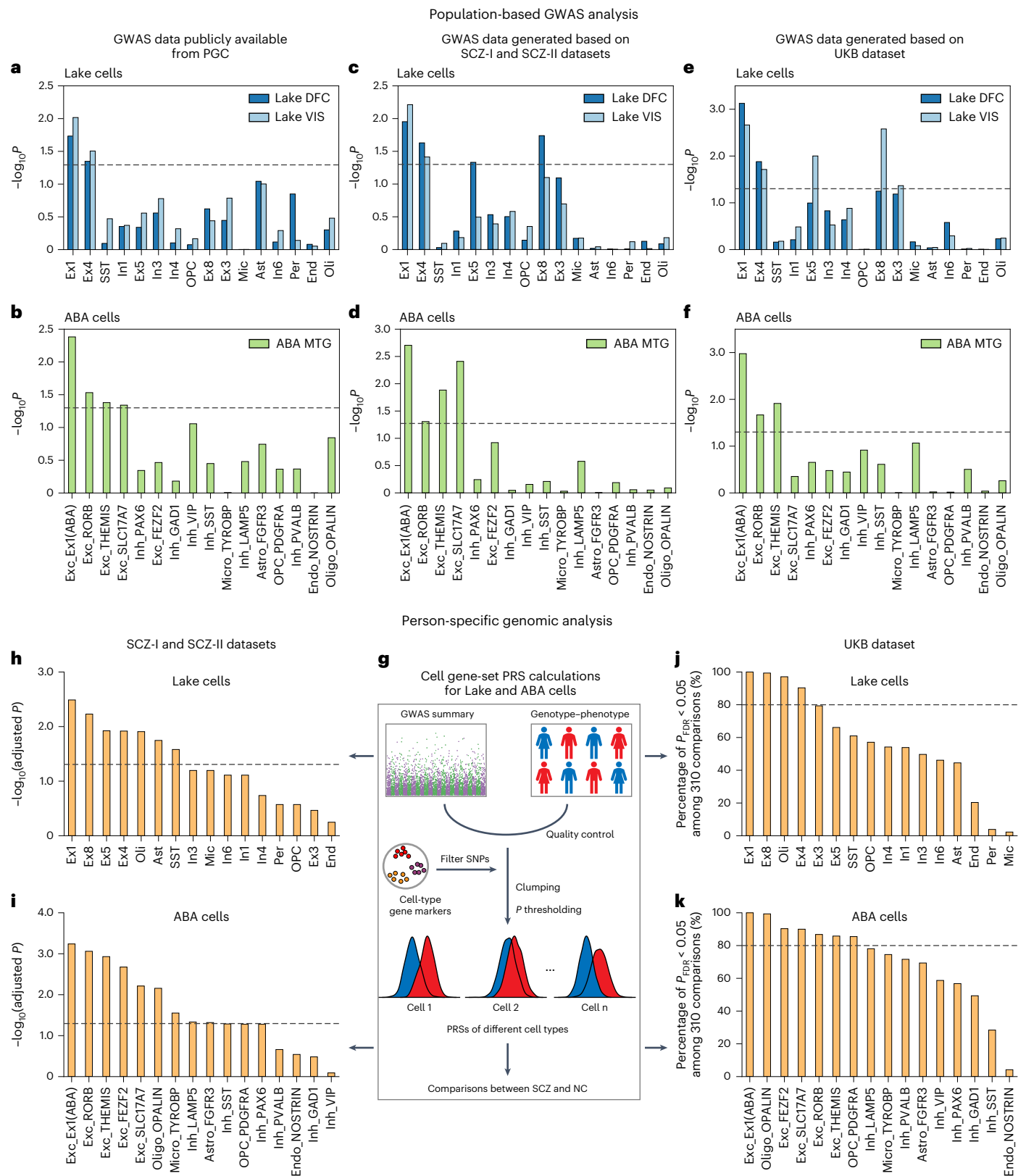
To examine the association between cell-type-specific PRS and cell-type enrichment of individual dysconnectivity, we conducted representational similarity analysis (RSA)<sup>80</sup>. For each cell type, we calculated pair-wise similarity matrix among participants ( $n = 818$ ; 381 SCZ and 437 NC) using the cell-type-specific PRS and the enrichment score (NES) of individual dysconnectivity for that cell type. Rather than calculating the cell-type-specific NES from imaging transcriptome analyses based on the group-level dysconnectivity pattern, here we derive the NES of individual dysconnectivity by evaluating the deviation of each individual's dysconnectivity pattern from a normative model across the population. The two matrices of the same cell type were compared using Spearman's correlation (Fig. 5a). For corticocortical dysconnectivity, a significant positive correlation was observed between PRS and NES similarity in

**Fig. 4 | Population-based and person-specific genomic analyses reveal the cell types implicated in the genetic risk of SCZ.** **a–f**, Vertical bar plots showing the enrichment of Lake cells (**a,c,e**) and ABA cells (**b,d,f**) for polygenic risk in SCZ across different GWAS datasets, utilizing MAGMA gene-set property analysis. The  $P$  value was derived from a one-sided positive test and was not adjusted for multiple statistical tests. The GWAS summary statistics utilized in **a** and **b** were obtained from the PGC. Panels **c–f** use GWAS datasets that were derived from conducting case/control association analysis on the combined data of SCZ-I and SCZ-II (SCZ/NC = 471/475) (**c,d**), as well as the UKB data (SCZ/NC = 886/274,508) (**e,f**). Cell-specific genes were defined using Lake data from the DFC and VIS (**a,c,e**) and ABA data from the MTG (**b,d,f**). The genome-wide risk for SCZ was primarily enriched in Lake cells of Ex1 and Ex4, as well as in ABA cells of Exc\_Exc1(ABA) and Exc\_RORB. **g**, Schematic diagram showing the procedures for calculating cell-type-specific gene-set PRS. **h,i**, Vertical bar plots showing the

two-sample  $t$ -test (two-sided) differences in cell-type-specific PRSs for Lake (**h**) and ABA (**i**) cells, conducted in the combined dataset of SCZ-I and SCZ-II (SCZ/NC = 471/475). The cells were ranked in descending order based on the significance of the between-group differences in PRSs. The PRSs were calculated at a  $P$  threshold of 0.01. The dashed line indicates FDR corrected  $P$  value at 0.05. **j,k**, The UKB data (SCZ/NC = 886/274,508) were used to perform the identical analysis as in **h** and **i**. The healthy participants from the UKB dataset were divided into 310 groups, with each group having nearly an equal number of participants as the SCZ group. Each of the 310 subgroups was compared with the SCZ group. The vertical bar plots show the percentage of significant between-group differences (two-sample  $t$ -test (two-sided) and  $P_{\text{FDR}} < 0.05$ ) in PRS for each cell type among the 310 comparisons. The dashed line represents that 80% of the comparisons yield significant results.

the Ex1 and Oli cell types (Lake cells), as well as the Exc\_Ex1(ABA) and Oligo\_OPALIN cell types (ABA cells) ( $P_{\text{perm}} < 0.05/16$ ; Fig. 5b). For corticostriatal dysconnectivity, PRS demonstrated a significant positive correlation with NES similarity in the Ex4 and Oli cell types, as well as the Exc\_RORB and Oligo\_OPALIN cell types ( $P_{\text{perm}} < 0.05/16$ ; Fig. 5c). Furthermore, the associated patterns of Ex1 and Oli PRS with individual

corticocortical dysconnectivity deviation were significantly positively correlated with the group-level corticocortical dysconnectivity  $t$ -map derived from SCZ and NC comparisons (Ex1,  $r = 0.36$ ,  $P_{\text{spin}} < 0.0001$ ; Oli,  $r = 0.39$ ,  $P_{\text{spin}} < 0.0001$ ; Fig. 5d). Similarly, we found a significant positive correlation between the patterns of Ex4 and Oli PRS associated with individual corticostriatal dysconnectivity deviation and the group-level



corticostriatal dysconnectivity *t*-map (Ex4,  $r = 0.25$ ,  $P_{\text{spin}} = 0.0004$ ; Oli,  $r = 0.24$ ,  $P_{\text{spin}} = 0.0011$ ; Fig. 5e). These associative effects were also observed in ABA cells (Supplementary Fig. 18). Our findings provide individual-level validation of the specific relationships between two different dysconnectivity patterns and cell types.

## Discussion

Owing to the highly heterogeneous nature of SCZ across multiple clinical domains and distinct biological systems from macro- to microscale levels, it is imperative to establish a unifying cross-scale framework along different dimensions to gain a comprehensive insight into the pathophysiological mechanisms of SCZ and develop precision interventions. Here we constructed a circuit-based cross-scale framework and systematically validated its robustness and reliability by integrating a wealth of data, including multiple independent large-sample and multi-site neuroimaging datasets, clinical information of patients with SCZ, functional annotation dataset, postmortem whole-brain transcriptomics, distinct cortical single-cell sequencing, population-based GWAS, and person-specific genome-wide genotyping datasets. Under this framework, we proposed a hypothetical model in which clinical symptoms, macroscale brain functional dysconnectivity, microscale cellular abnormalities and genetic risk related to SCZ were tied together along two different circuit-based dimensions (Fig. 6). Notably, disruptions of the intracortical projection circuits, consisting of upper-layer (L2/3) excitatory neurons and oligodendrocytes associated with genetic risk, may drive corticocortical dysconnectivity, which might be relevant to the negative, cognitive and excitement symptoms of SCZ. Conversely, disruptions of the cortico-subcortical projection circuits that are composed of deeper-layer (L5) excitatory neurons as well as oligodendrocytes associated with genetic risk could give rise to corticostriatal dysconnectivity, which may be responsible for both positive and negative symptoms of SCZ. Our findings, which systematically integrated SCZ abnormalities across clinical, macroscale and microscale levels, represent a significant advance in unraveling the complex pathophysiology of SCZ by sketching out how individual genetic variants affect different cellular-level deficits contributing to distinct system-level brain circuit dysconnectivity that are responsible for the heterogeneous symptoms in SCZ.

Our neuroimaging analyses highlighted robust and consistent alterations of the corticocortical and corticostriatal functional connectivity in SCZ, suggesting dysfunctions in the integration and/or segregation of information among multiple cortical areas and striatum in SCZ<sup>9–13,15,16,19</sup>, which was in line with the hypothesis of dysconnectivity and previous evidences<sup>46–50</sup>. The observed macroscale dysconnectivity in SCZ could probably be indicative of cumulative disruptions of brain microcircuits<sup>81</sup>, given that the fMRI technique can roughly capture underlying variations of local neural activities<sup>82</sup>. Our cross-scale computational integration of neuroimaging and cortical multiomics data reliably showed a unified association between the large-scale brain functional dysconnectivity and disrupted cellular circuits, consisting of cortical layer-specific excitatory neurons and oligodendrocytes

along two distinct circuit-based dimensions. As SCZ is typically a disorder of brain circuits<sup>81,83,84</sup>, our integrated hypothetical model established a plausible connection across multiple well-known facets of SCZ pathophysiology, from person-specific genomic variations<sup>73</sup> and cell-type-specific circuit disruptions<sup>28–35</sup> to brain dysconnectivity of systemic-level neural circuits<sup>9–16</sup>, which may mediate the clinical complexity in SCZ. Disruptions of the corticostriatal circuit in SCZ, microscopically characterized by the composition of cortical deeper-layer (L5) excitatory neurons and oligodendrocytes, and macroscopically measured by striatum-to-cortex functional connectivity, have been supported by multiple independent lines of evidence<sup>16,39,40,85</sup>, and have gained intensive attention in current clinical practices. Our study indicates that the corticocortical neural circuits consisting of upper-layer excitatory neurons are associated with cognitive and negative symptoms that merit further therapeutic consideration. In addition, our genomic analyses offered partial support for the specificity of cellular abnormalities in excitatory neurons and oligodendrocytes in SCZ<sup>86</sup>. Notably, the reliably observed inconsistency of the presence of oligodendrocytes between population-based and person-specific analyses might be attributed to the distinct goals and analytical procedures of the two methods. The population-based approach seeks to discern the cell-type specificity of GWAS results, while the individual-based method evaluates the overall genetic contribution of each cell type. Collectively, our findings uncovered specific brain circuit abnormalities in SCZ at different scales, which could especially be valuable for a comprehensive understanding of the SCZ pathogenesis, although further biological experiments are needed to determine the definitive patterns of interactions between the macro- and microscale brain circuits.

Previous studies have demonstrated pathological alterations or deficits of cortical pyramidal excitatory neurons and oligodendrocytes in SCZ separately<sup>87–89</sup>. Our findings particularly pointed out that the abnormal development and potential interactions between the excitatory neurons from different cortical layers and oligodendrocytes might contribute to the disruptions of the local neuronal circuit and long-range connectivity, which in turn could lead to the macroscopic brain functional dysconnectivity and underlie diverse clinical manifestations in SCZ. The cortical excitatory pyramidal neurons are responsible for the construction of major cortical projection circuits<sup>72,90,91</sup>, whose efficiency is primarily attributed to the oligodendrocyte maturation and myelination<sup>92,93</sup>, supporting the interactions of the two cell types<sup>89</sup>. The interaction of oligodendrocytes with excitatory pyramidal neurons is shown to influence the structural and functional plasticity of neural networks. For example, oligodendrocytes enable the modulation and synchronization of neuronal plasticity and activity<sup>94</sup>, which may conversely impact the dynamic process of myelination by oligodendrocytes<sup>95</sup>. Furthermore, oligodendrocyte dysfunction in SCZ<sup>96–98</sup> could result in abnormal cortical oscillation and broad-spectrum brain dysconnectivity, accounting for the complex clinical characteristics in SCZ. For example, the cortical upper-layer excitatory neurons from which most corticocortical connectivity originates<sup>99</sup>, combined

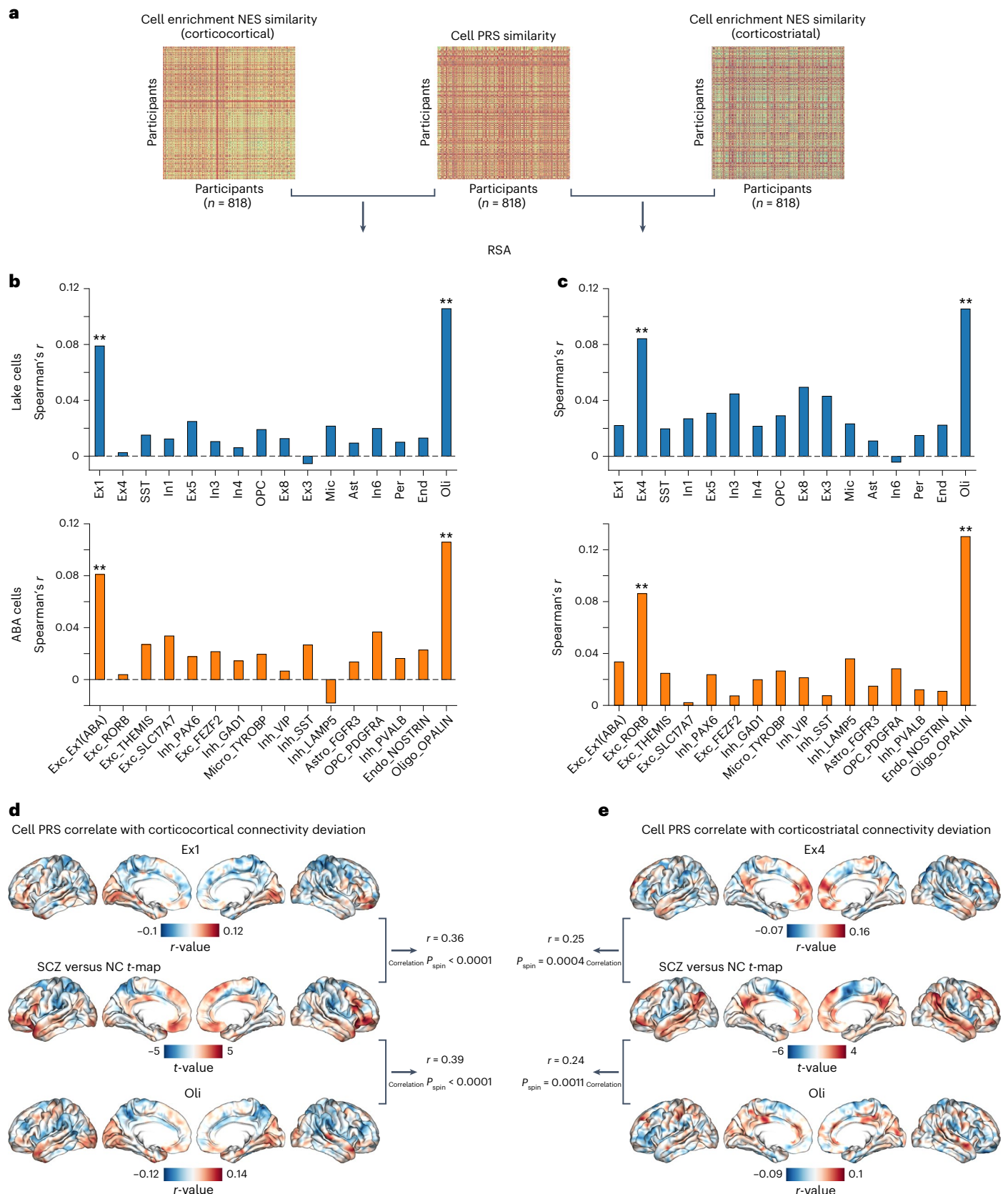
**Fig. 5 | Cell-type-specific PRS linked to dysconnectivity.** **a**, RSA. Individual-level dysconnectivity was obtained by calculating the deviation of each participant's functional connectivity from a normal distribution (Methods). This dysconnectivity measure was then used to conduct the single-cell transcriptional enrichment analysis, as described in Fig. 3, yielding the enrichment score (that is, NES) for each cell. The similarity matrix of cell-type-specific PRS was compared with the similarity based on the enrichment scores of the corresponding cells. **b,c**, Vertical bar plots showing the results of RSA for PRS and corticocortical dysconnectivity (**b**), as well as for PRS and corticostriatal dysconnectivity (**c**). For corticocortical dysconnectivity, there was a significant positive correlation between PRS and NES similarity for cell types of Ex1 (two-sided Spearman's  $r_s = 0.078$ ,  $P_{\text{perm}} < 0.0001$ ) and Oli ( $r_s = 0.105$ ,  $P_{\text{perm}} < 0.0001$ ) (Lake cells), as well as Exc\_Ex1 (ABA) ( $r_s = 0.081$ ,  $P_{\text{perm}} < 0.0001$ ) and Oligo\_OPALIN ( $r_s = 0.106$ ,

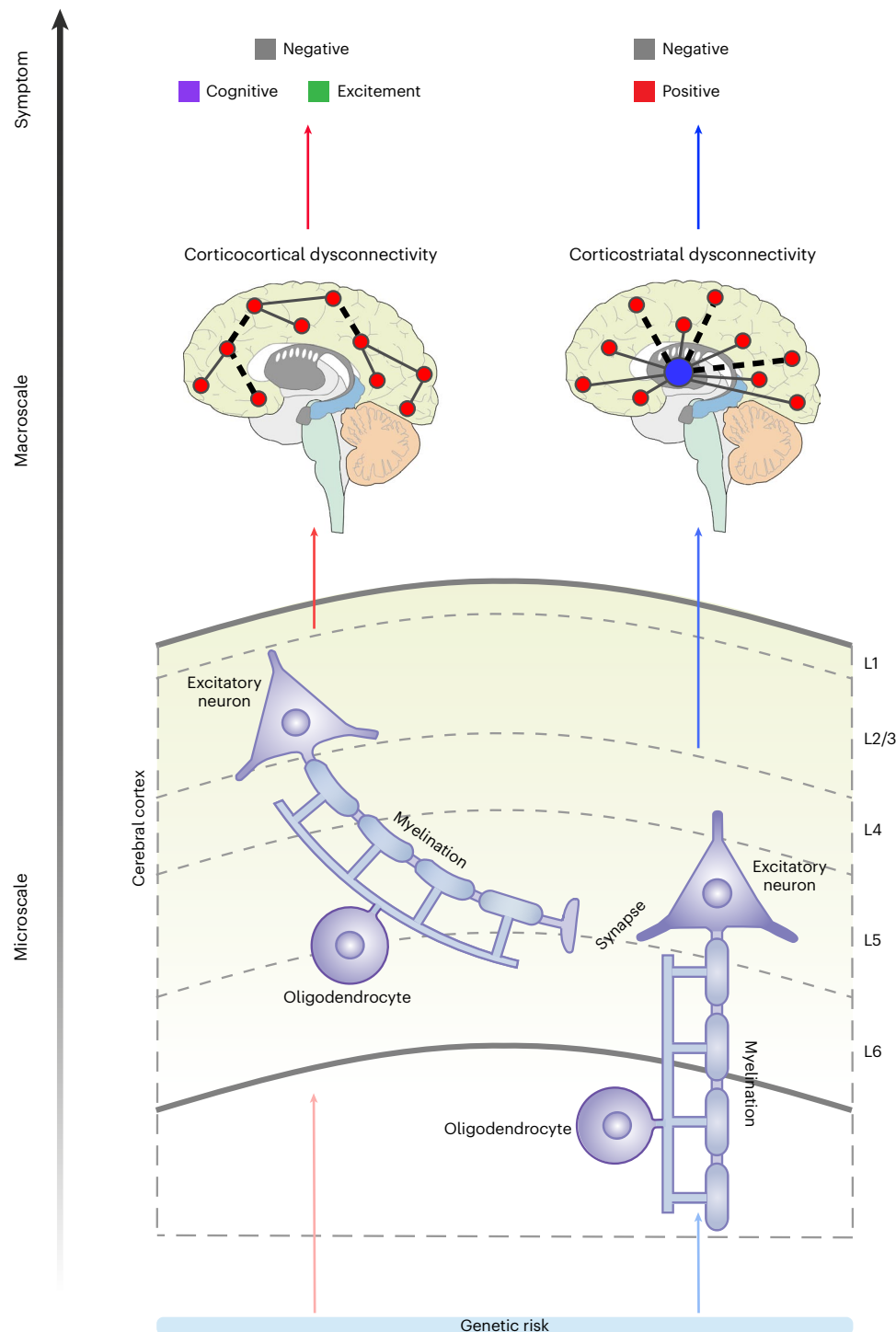
$P_{\text{perm}} < 0.0001$ ) (ABA cells). For corticostriatal dysconnectivity, a significant positive correlation was observed between PRS and NES similarity for Ex4 (two-sided Spearman's  $r_s = 0.084$ ,  $P_{\text{perm}} < 0.0001$ ) and Oli ( $r_s = 0.105$ ,  $P_{\text{perm}} < 0.0001$ ) cells, as well as Exc\_RORB ( $r_s = 0.086$ ,  $P_{\text{perm}} < 0.0001$ ) and Oligo\_OPALIN ( $r_s = 0.13$ ,  $P_{\text{perm}} < 0.0001$ ) cells. The significance was determined through 10,000 permutation tests, where the PRS similarity matrix was shuffled per row in each test. **d**, The associated patterns of Ex1 and Oli PRS with corticocortical connectivity deviation showed significant positive correlations (two-sided Pearson's  $r$ ) with the corticocortical dysconnectivity *t*-map. **e**, The associated patterns of Ex4 and Oli PRS with corticostriatal connectivity deviation were significantly positively correlated (two-sided Pearson's  $r$ ) with the corticostriatal dysconnectivity *t*-map. The significance ( $P_{\text{spin}}$ ) was determined using the spin-based permutation test. \*\* $P_{\text{perm}} < 0.05/16$  (cell types).



with oligodendrocytes responsible for axon myelination enabling fast neurotransmission<sup>93</sup>, may be potential modulators of corticocortical connectivity shaping individualized clinical manifestations<sup>17</sup>. To further understand how the interplay between cellular dysfunctions

in oligodendrocytes and excitatory neurons influences macro-level functional disconnections in SCZ, we recommend future research efforts to clarify the role of neuronal microcircuits in shaping macroscopic functional connectivity. This exploration can greatly benefit





**Fig. 6 | A hypothetical model linking the macroscale dysconnectivity to microscale cellular circuits and clinical symptoms in SCZ.** The corticocortical dysconnectivity is responsible for symptoms of negative, cognitive and excitement and may be associated with perturbed cortical projection circuits consisting of excitatory neurons in the cortical upper layer (L2/3) and

oligodendrocytes associated with specific genetic risk. The corticostriatal dysconnectivity may contribute to both negative and positive symptoms and is related to the disturbed cortico-subcortical projection circuits consisting of excitatory neurons in the cortical deeper layer (L5) and oligodendrocytes associated with specific genetic risk.

from the integration of cell-specific optogenetic stimulation with high-field fMRI in small animal models<sup>100,101</sup>. Moreover, future studies with the advent of novel technologies, such as spatial multiomics co-profiling technologies<sup>102</sup> of brain specimens from individuals with SCZ or the availability of laminar fMRI used to assess layer-specific cortical activity techniques<sup>103</sup>, could further validate our findings and refine the proposed hypothetical model.

Our findings of cross-scale associations along the two abnormal circuit dimensions in SCZ may provide insights for enhancing dysconnectivity-based non-pharmacological interventions for patients with SCZ and developing promising therapeutic strategies based on neurobiological targets for pharmacological modulation, although more clinical and biological validations are desired. Antipsychotic medications currently licensed mainly target the striatal dopamine

circuits and have been found to be less effective in treating negative symptoms, particularly cognitive deficits, in SCZ<sup>104,105</sup>. Our identified cross-scale circuit abnormality may facilitate the development of new therapeutic strategies for effectively treating the heterogeneous clinical spectrum of SCZ. For instance, to enhance the efficacy of improving cognitive and/or negative symptoms, more attention should be paid to the upper-layer corticocortical circuitry, which may provide feasible strategies for promoting connectivity-based non-pharmacological interventions<sup>106,107</sup>. Specifically, biophysically based simulations might be employed to guide the parameter selection for repetitive transcranial magnetic stimulation or transcranial direct current stimulation, with the aim of preferentially stimulating excitatory neurons at varying cortical depths, especially along disrupted corticocortical circuitry<sup>108,109</sup>. There are other prospective treatment strategies that can be explored to address the disrupted corticocortical dysconnectivity and alleviate cognitive or negative symptoms in patients with SCZ. For instance, newly designed drugs that target the gene pathway modifying neuronal transmission of excitatory neurons in the cortical layer 2/3 or new cell-specific gene therapy techniques<sup>110</sup> could be considered. We believe that with future breakthroughs in new clinic or laboratory intervention technologies combining the cross-scale biological mechanistic insights of SCZ, circuit-informed precision treatment formulation of SCZ may move forward.

SCZ remains highly heterogeneous in its causation and such that dissociated observations from different scales are greatly diverse. Our cross-scale, circuit-based study has sketched out a cross-scale association model of circuit-based abnormality in SCZ to help better understand the SCZ pathomechanism. Nevertheless, we have to acknowledge that such cross-scale, data-driven computational integrations can only offer suggestive evidence for associations of possible etio-pathological factors, without additional casual indications, warranting further in-depth neurobiological investigations. In this study, we focused on two circuit-based dimensions of SCZ to propose and validate our framework. However, we acknowledge that we have not considered functional abnormalities in other circuits, such as those related to cerebellar dysconnectivity<sup>20</sup>. This limitation may potentially impact the completeness of information needed for clinical practice. Therefore, exploring the impact of dysfunction in other circuits would indeed be a worthwhile endeavor for future research. In addition, our imaging transcriptome analyses had limitations arising from the utilization of independent samples for imaging and transcriptomic data. The relationship between these data domains was indirectly established through spatial variations. Although previous evidence has showcased the potential of this approach in unraveling the molecular mechanisms underlying different brain disorders<sup>71,111</sup>, more biological validations from postmortem samples or SCZ-derived iPSC are strongly recommended.

## Conclusion

Our study highlighted the cross-scale heterogeneity of SCZ, by illustrating cross-scale associations of genetic risk, microscale cellular circuits, macroscale brain circuits and clinical symptoms through two distinct dysconnectivity dimensions. Our study introduced a robust cross-scale framework that harnesses neuroimaging and multiomics data, leveraging advanced machine learning and statistical techniques. This framework provided a potential avenue for unraveling the intricate pathological mechanisms of psychiatric disorders. Moreover, our findings have significant implications for advancing the understanding of comprehensive cross-scale abnormalities and developing novel treatment strategies for individuals with SCZ.

## Methods

### Participants

In this study, we used three independent SCZ datasets (SCZ-I, SCZ-II and COBRE) involving more than 1,400 participants, one MDD dataset

(MDD-XX) for specificity validation and one genotyping dataset from the UKB, including patients with SCZ and healthy participants, for genomic validation. The participants enrolled in the SCZ-I and SCZ-II datasets were obtained from seven Chinese sites with the same recruitment protocol, which has been reported in our previous study<sup>16</sup>. We divided these participants into two independent datasets according to differences in their MRI scanner types (3.0T Siemens or GE scanners). COBRE<sup>37</sup> is a publicly available dataset in which participants were scanned on a 3.0T Siemens TimTrio scanner, and most individuals are either Caucasians or Hispanics. All the participants included in the SCZ-I, SCZ-II and MDD-XX datasets were Han Chinese descendants.

**SCZ-I.** The SCZ-I dataset ( $n = 736$  participants; 400 SCZ and 336 NC) was developed by acquiring samples from four Chinese hospitals—Peking University Sixth Hospital (PKU), Beijing Huilongguan Hospital (HLG), Xijing Hospital (XJ) and Henan Mental Hospital (HM-Siemens)—where participants were scanned on 3.0T Siemens MRI scanners. All patients with SCZ were diagnosed consensually by two qualified psychiatrists utilizing the Structured Clinical Interview for DSM-IV axis I disorders (SCID-I/P, Patient Edition). None of the patients had any history of other DSM-IV axis I disorders, neurological disorders, cognitive disabilities, serious physical diseases, severe head trauma, substance abuse or dependence and electroconvulsive therapy within the last 6 months. All healthy participants were evaluated clinically using the SCID-I/NP (Non-patient Edition) and reported that they had no family (first- and second-degree relatives) history of any mental disorders. The research protocol was approved by the Medical Research Ethics Committee of participating hospitals. All participants or their guardians provided written informed consent before their enrollment in the study. We first examined data completeness for each participant and excluded those individuals with incomplete clinical assessments, missing modalities of either structural MRI or resting-state fMRI (rsfMRI), and without genotype data. We then performed a quality-control assessment by inspecting the frame-wise displacement<sup>42</sup>, a motion-related parameter, for rsfMRI data and excluded those participants with a mean frame-wise displacement greater than 0.3 mm. By doing this, there were no significant differences in mean frame-wise displacement between the SCZ and NC groups. Furthermore, participants who failed genotype quality control (see details below) were also excluded. In total, 625 individuals (318 SCZ and 307 NC) remained after the screening, and the SCZ and NC groups did not significantly differ in age and sex.

**SCZ-II.** The SCZ-II dataset included 541 participants (279 SCZ and 262 NC) who were scanned on 3.0T GE MRI scanners, and recruited from three hospitals in China: Renmin Hospital of Wuhan University (WHU), Zhumadian Psychiatric Hospital (ZMD) and Henan Mental Hospital (HM-GE). Notably, although participants from Henan Mental Hospital were divided into SCZ-I and SCZ-II datasets, they did not overlap and were scanned on different scanner types. Before their enrollment, written informed consent was obtained from all the participants or their legal guardians. The study protocol was approved by the Medical Research Ethics Committees of the local hospitals. All inclusion criteria and clinical assessments, data gathering and completeness examination and quality-control methods were completely identical to the SCZ-I dataset. Finally, 415 participants (195 SCZ and 220 NC) were included in this dataset, significantly matching their age and sex between the SCZ and NC groups.

**COBRE.** The COBRE dataset contains multimodal MRI imaging and phenotypic data of patients with SCZ and NC participants, which can be openly accessed via the COINS data exchange portal (<https://coins.trendscenter.org/>)<sup>112</sup>. Detailed study information has been provided on the web platform. The SCID-I/P was adopted by trained clinical psychiatrists for the diagnosis of SCZ. All individuals were screened and fully met the inclusion criteria. Specifically, participants were excluded



if they had any history of neurological disorder, mental retardation, substance abuse or dependence within the last 12 months, and severe head trauma with more than 5 minutes loss of consciousness. Written informed consent was obtained from all participants in accordance with the institutional review board protocols of the University of New Mexico. We further screened these participants by applying the same procedures (except for the genotype data) as for the SCZ-I and SCZ-II datasets, which resulted in a total of 159 participants (72 SCZ and 87 NC) who were eligible for inclusion. There were no significant differences in age and sex between the SCZ and NC groups.

**MDD-XX.** The MDD-XX dataset included patients with MDD and NC, which was obtained at the Second Affiliated Hospital of Xinxiang Medical University in China. Patients were diagnosed to meet the criteria of DSM-IV for MDD by trained psychiatrists using simple non-structured interviews. Participants were excluded if they had any other mental disorders; organic causes of depression, including heart, liver and kidney diseases; or presence of surgically implanted electronic devices or metal frames that might hinder MRI scanning. All healthy participants had no history of psychiatric and/or neurological diseases. Written informed consent was obtained from all the participants. This study was approved by the Ethics Committee of the Second Affiliated Hospital of Xinxiang Medical University. Procedures for the MRI data completeness inspection and quality control for rsfMRI data were the same as described for the SCZ datasets. Finally, 149 participants (75 MDD and 74 NC) were included in the study, with statistically matched age and sex between the MDD and NC groups.

**UKB.** The UKB project gathered data from more than 500,000 participants in the United Kingdom between 2006 and 2010, including cognitive assessments, physical measurements, blood samples and touch-screen questionnaires<sup>113</sup>. This study included a subset of participants' genotype data. Unrelated individuals of Caucasian ancestry (field 21,000) were selected for analysis to reduce confounding by population stratification. The SCZ diagnoses were ascertained based on the International Classification of Diseases (ICD)-10 codes for SCZ (F20.0–F20.9) found in hospital records (fields 41,202 and 41,204) or death records (fields 40,001 and 40,002). Individuals diagnosed with other psychoses (ICD-10 codes F21–23, F25, F28, F29) or bipolar disorder (ICD-10 codes F30, F31 and self-reported code 1291 in field 20,002) were excluded from the control group. Imputed genotypes were downloaded and we applied quality-control filters for minor allele frequency ( $MAF \geq 0.001$ ), information score ( $INFO \geq 0.8$ ) and Hardy–Weinberg equilibrium ( $HWE \leq 1 \times 10^{-7}$ ). Following the exclusion, the study comprised 275,394 participants, including 886 individuals with SCZ and 274,508 in the NC group. All participants provided informed consent and ethical approval was given by the North West Multi-centre Research Ethics Committee (<https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics>). The UKB data are accessed through application number 85139.

### Clinical assessment

The psychiatric symptoms of patients with SCZ in the SCZ-I and SCZ-II datasets were evaluated using PANSS<sup>55</sup> by trained clinical psychiatrists after protocol training and consensus evaluation. In particular, only the patients with SCZ with a PANSS total score exceeding 60 and scores greater than 4 on at least 3 of the 7 positive items were included. The PANSS evaluations for patients with SCZ in the COBRE dataset were also obtained.

### MRI data acquisition

Multimodal imaging data for the SCZ-I and SCZ-II datasets were obtained from one multi-site project using the same protocol whenever possible. Briefly, images of the SCZ-I dataset were acquired on 3.0T Siemens scanners: TimTrio in PKU, HLG and XJ; Verio in HM-Siemens. For

the rsfMRI data, two-dimensional echo planar imaging (EPI) was used with the following parameters: repetition time (TR) 2,000 ms; echo time (TE) 30 ms; flip angle (FA) 90°; field of view (FOV)  $220 \times 220 \text{ mm}^2$ ; matrix size  $64 \times 64$ ; voxel size  $3.4375 \times 3.4375 \times 4.6 \text{ mm}^3$ ; 240 volumes; and 33 slices. For the T1-weighted (T1w) images, three-dimensional magnetization-prepared rapid gradient-echo sequence was performed with the following parameters: TR 2,530 ms; TE 3.5 ms (PKU, HLG and XJ) or 2.43 ms (HM-Siemens); FA 7°; inversion time (TI) 1,100 ms; voxel size  $1 \times 1 \times 1 \text{ mm}^3$ ; and matrix size  $256 \times 256 \times 192$ .

Imaging data of the SCZ-II dataset were obtained from 3.0T GE Signa HDxt scanners in WHU, ZMD and HM-GE. For the rsfMRI data, the two-dimensional EPI sequence was used with the parameters: TR 2,000 ms; TE 30 ms; FA 90°; FOV  $220 \times 220 \text{ mm}^2$ ; matrix size  $64 \times 64$ ; voxel size  $3.4375 \times 3.4375 \times 4.6 \text{ mm}^3$ ; 240 (WHU and HM-GE) or 180 (ZMD) volumes; and 32 (WHU) or 33 (ZMD and HM-GE) slices. For the T1w images, the three-dimensional gradient-echo sequence was performed with the parameters: TR 7.8 ms (WHU and HM-GE) or 6.8 ms (ZMD); TE 3 ms (WHU and HM-GE) or 2.5 ms (ZMD); FA 7°; TI 1,100 ms; voxel size  $1 \times 1 \times 1 \text{ mm}^3$ ; and matrix size  $256 \times 256 \times 188$ .

The COBRE dataset was acquired on a 3.0T Siemens TimTrio scanner, and detailed scanning parameters were previously reported<sup>37</sup>. Imaging data of the MDD-XX dataset were obtained from a 3.0T Siemens Verio scanner using the same parameters as HM-Siemens in the SCZ-I dataset.

### MRI data preprocessing

All rsfMRI data were preprocessed with the same and standardized pipeline using BRANT<sup>114</sup>, a MATLAB toolkit for batch preprocessing of fMRI data. Briefly, the following procedures were included: discarding the first ten time points to account for the equilibration effect; slice timing correction to correct for temporal shifts of different slices; within-participant EPI image realignment for estimating and spatially correcting for head motions of different EPI volumes; co-registration from T1w to mean EPI image by rigid-body transformation; spatial normalization of EPI images to standard Montreal Neurological Institute (MNI) space using the segmentations of T1w; resampling the normalized EPI images to  $3 \times 3 \times 3 \text{ mm}^3$ ; removing confounding effects by regressing out linear trends, averaged white matter and cerebrospinal fluid, the first derivatives of white matter and cerebrospinal fluid, and estimated head motion parameters<sup>115</sup>; and temporal band-pass filtering at 0.01–0.08 Hz to suppress low-frequency drifts and physiological noises. We completed the quality control by inspecting the head motion (mean frame-wise displacement less than 0.3 mm) and examining the quality of registration and normalization. After finishing these steps, voxel-based rsfMRI data were obtained in the MNI space, which was additionally interpolated to the cerebral cortex along the mid-thickness surface and resampled to fsaverage5 space<sup>116</sup> (with 10,242 vertices per hemisphere), followed by a surface-based Gaussian kernel smoothing (full-width at half-maximum 2 mm).

### Genotype data acquisition and processing

Ethylene diamine tetraacetic acid anti-coagulated venous blood samples were acquired from all participants in the SCZ-I and SCZ-II datasets. We used the EZgene Blood gDNA Miniprep Kit to extract genomic DNA from whole blood and performed whole-genome genotyping on Illumina Human OmniZhongHua-8 BeadChips, based on the standard Illumina genotyping protocol. The quality control for genotyping was carried out using PLINK (v1.07)<sup>117</sup>, referred to in our previous study<sup>118</sup>. First, participants with missing genotype rates greater than 0.05 were excluded. Then, pair-wise identity-by-descent was estimated to identify pairs of participants who were likely to have relative relationships (with more similar genotypes than those from random samples by chance), and we removed the one from each pair with a greater missing genotype rate. Next, we removed single nucleotide polymorphisms (SNPs), according to the data filtering criteria: missing genotype rates

greater than 0.05; a significant departure from the Hardy–Weinberg equilibrium ( $P < 0.001$ ) and a minor allele frequency of less than 0.01. In addition, we applied the EIGENSTRAT method<sup>119,120</sup> to control for population stratification. Specifically, principal component analysis was carried out on the linkage disequilibrium pruned set of autosomal SNPs, which were obtained by conducting linkage disequilibrium pruning and removing five long-range linkage disequilibrium regions with the HapMap phase 3 reference dataset<sup>121</sup>. We obtained 10 principal components and excluded outlying samples with greater than 6 s.d. values. Finally, imputation was completed for ungenotyped SNPs using SHAPEIT (v2.r790)<sup>122</sup> and IMPUTE2<sup>123</sup> based on the 1,000 Genome Phase 1 reference panel<sup>124</sup>. Autosomal SNPs with imputation quality scores of greater than 0.8 were reserved for further analysis.

### Calculation of functional connectivity for each individual

**Corticocortical connectivity.** We characterized corticocortical connectivity via the notion of functional gradients<sup>36</sup>, which were derived from rsfMRI data by applying a nonlinear dimensionality reduction approach, that is, diffusion maps<sup>38</sup>, to the cortical functional connectivity matrix. This method resulted in several low-dimensional components (also known as gradients), among which we selected the first principal gradient for representation, given that it explains the most amount of variance in the original matrix and reflects an intrinsic cortical organization spanning between the sensorimotor and transmodal areas<sup>36,125</sup>, in both healthy individuals<sup>36,126</sup> and psychiatric patients<sup>44</sup>.

For each participant, we calculated functional gradients across the whole cerebral cortex, following the same procedures as described elsewhere<sup>36,44,126</sup>. Precisely, we first constructed a cortical vertex-based functional connectivity matrix (rsFC,  $20,484 \times 20,484$  entries) based on the Pearson's correlation. We then performed  $z$  transformation and thresholded the rsFC by reserving the top-10% values of connectivity in each row while zeroing others and the negative connectivity, as well. Next, we created an affinity matrix by calculating cosine distances between all pairs of rows and subtracting them by one. The obtained affinity matrix, with values ranging from zero to one, was symmetric and captured the similarity of connectivity profiles among the cortical vertices. Furthermore, we performed dimensionality reduction on the affinity matrix using the diffusion maps<sup>38</sup> method, which was determined by a single parameter  $\alpha$ . We followed the choice made by previous studies<sup>36,44,126</sup> and assigned  $\alpha$  to 0.5. This analysis eventually estimated several components in low-dimensional embedding space, referred to as functional gradients, and each of which ( $1 \times 20,484$  entries) explained a proportion of connectivity variance. Finally, for each dataset, we constructed an averaged template based on the gradients of all individuals and performed orthonormal alignment from the gradient of each participant to the template through Procrustes analysis<sup>127</sup>.

**Corticostriatal connectivity.** For each participant, we obtained two forms of rsfMRI data after preprocessing, that is, voxel-based in the MNI space and vertex-based in the fsaverage5 space. We first extracted the mean time series around voxels in the striatum region from voxel-based rsfMRI data, then calculated Pearson's correlations between the extracted striatal time series and the time series of each vertex across the whole cerebral cortex from vertex-based rsfMRI data. The striatum mask used here was developed in our previous study<sup>16</sup>, based on the Brainnetome atlas<sup>128</sup>.

### Stability analysis for single-participant functional connectivity pattern

Before the acquisition of the SCZ-I and SCZ-II datasets, a preliminary study was conducted with three healthy traveling participants (participant 01, participant 02 and participant 03) to assess the consistency of images obtained from different scanners. Each participant underwent a single imaging scan at multiple sites: 7 sites (PKU, HLG,

HM-Siemens, XJ, HM-GE, WHU and ZMD) for participant 01, 5 sites (PKU, HLG, HM-Siemens, XJ and HM-GE) for participant 02, and 4 sites (PKU, HM-Siemens, XJ and HM-GE) for participant 03.

To evaluate the stability of single-participant functional connectivity patterns, we analyzed data from the aforementioned three participants, as well as images from healthy participants in the SCZ-I and SCZ-II datasets. Specifically, we examined the variations in corticocortical and corticostriatal connectivity between scans of the same participant across different sites and scans of different participants at single sites. The analysis procedure is as follows, with participant 01 serving as an example for the sake of clarity (participant 02 and participant 03 follow a similar process). First, we separately computed the functional connectivity pattern for each of the 7 cross-site images of participant 01, and then calculated the variance of each vertex across the cortex (18,715 vertices after excluding the medial wall), resulting in a variance map. Second, for each site where participant 01 was scanned, we randomly selected imaging data from 7 different participants (equivalent to the image number of participant 01) and obtained the variance map using the same method. In particular, to minimize bias, we repeated this process 100 times and selected the variance map that had a median value of the averaged cortical variance (mean value across 18,715 vertices) for subsequent analysis. Finally, we performed a two-sample  $t$ -test to compare the variance map of participant 01 with the randomly selected variance maps from each site.

### Identification of dysconnectivity in SCZ

To identify dysconnectivity patterns of SCZ, we separately conducted cortical vertex-wise comparisons for corticocortical and corticostriatal connectivity between the SCZ and NC groups in each dataset, using surface-based linear models implemented in the MATLAB toolbox of SurfStat (<https://www.math.mcgill.ca/keith/surfstat/>). The models included effects of age, sex, site and group, as follows:  $Y = \beta_1 \times \text{age} + \beta_2 \times \text{sex} + \beta_3 \times \text{site} + \beta_4 \times \text{group} + \text{intercept}$ , where  $Y$  was the dependent variable representing connectivity value at each vertex (the medial wall was excluded),  $\beta$  was the corresponding regression coefficient and the site variable was dummied in the SCZ-I and SCZ-II datasets and was not included in the COBRE dataset. For each kind of connectivity in each dataset, an unthresholded  $t$ -map was generated (contrast = SCZ – NC). We calculated spatial Pearson's correlations across the cortex between  $t$ -maps from different datasets to evaluate reproducibility. The significance of correspondence was estimated using permuted spin tests<sup>41</sup>, which preserved the spatial and contralateral dependence. In particular, we performed a total of 10,000 random rotations to spherical representations of the cortical surface to generate a null distribution and estimated the  $P_{\text{spin}}$  value relative to this null model.

To further detect statistically significant dysconnectivity,  $t$ -maps from the SCZ-I dataset were corrected for FWEs at the level of 0.05, based on the random field theory<sup>52</sup> (that is,  $P_{\text{FWE}} < 0.05$ ) with a CDT of 0.001. We identified clusters with significant between-group differences in the corticocortical and corticostriatal connectivity in the SCZ-I dataset, requiring each cluster to have at least 30 connected vertices. To further evaluate the reproducibility of these significant connectivity differences, we extracted the mean connectivity values from each cluster and carried out cluster-wise comparisons between the SCZ and NC groups in the SCZ-II and COBRE datasets utilizing general linear models. Age, sex and site were controlled as covariates. The site variable was handled similarly as described above. The FDR correction method was adopted for multiple comparisons.

### Robustness analysis for dysconnectivity

To demonstrate the robustness of our identified global patterns of the corticocortical and corticostriatal dysconnectivity in SCZ, we performed several replication analyses in the SCZ-I dataset. When considering distinct methodological alternatives, we followed the same

procedures to achieve new dysconnectivity *t*-maps, which were then compared with our main findings. (1) Head motion can cause variations in the rsfMRI connectivity measures<sup>42</sup>, thus we repeated between-group comparisons including mean frame-wise displacement as an additional control covariate. (2) The influence of GSR on the rsfMRI data has not reached a consensus<sup>43</sup>, and our main results were based on the non-GSR data. Nevertheless, we re-preprocessed rsfMRI data using GSR, and repeated the between-group comparisons. (3) To investigate the effect of antipsychotic drugs on the connectivity differences, we repeated the between-group comparisons while additionally controlling doses of the antipsychotics (chlorpromazine equivalents). (4) To examine the influence of disease course on connectivity differences, we repeated the between-group comparisons while regressing the disease course (month). (5) To assess the effects of distinct matrix thresholding on our corticocortical results, we additionally repeated the analysis for corticocortical connectivity at another four thresholds (5%, 15%, 20%, 25%; default was 10%). (6) We also investigated the effects of harmonizing multi-site imaging data using the ComBat<sup>45</sup> approach.

To assess the robustness of significant dysconnectivity clusters, we conducted surface-based corticocortical and corticostriatal connectivity comparisons between patients with SCZ and NC, employing a non-parametric permutation test<sup>54</sup> for correction. Briefly, we randomly shuffled group membership (that is, SCZ and NC) and conducted between-group comparisons following the methodology described in Fig. 1, with a CDT of 0.001. The process was repeated 10,000 times, and the maximum cluster size was aggregated to create an empirical null distribution. The significance was determined by comparing the cluster sizes obtained from the actual between-group comparisons with the permutation distribution.

### Prediction of individual symptom severity based on dysconnectivity

To investigate the underlying relationships between individual dysconnectivity and clinical symptoms in SCZ, we conducted regression analyses to predict different PANSS<sup>55</sup> domains in patients with SCZ, based on the connectivity values in brain regions that showed significant dysconnectivity (derived from between-group comparisons, and survived for  $P_{FWE} < 0.05$  in the SCZ-I dataset). These prediction analyses were separately performed for the corticocortical and corticostriatal connectivity.

We constructed predictive models under the framework of GAMs, which have been applied in brain research to model nonlinear relationships<sup>129,130</sup>, using the pyGAM package (<https://github.com/dswah/pyGAM>). Briefly, linear GAMs were built for simplicity, which extended general linear models by allowing flexible nonlinear functions of predictor variables (known as feature functions), while maintaining the additivity. The feature functions were generally smooth and could capture arbitrary nonlinear relationships between each predictor and response variable. In our analysis, penalized B splines<sup>131</sup> were used as feature functions, which allowed us to control the degree of smoothness by a penalty parameter to prevent overfitting. Our models were as follows:  $Y = f(x_1) + f(x_2) + \dots + f(x_n)$ , where  $Y$  represented PANSS subscale scores,  $f(\cdot)$  was the spline term with 20 basis functions,  $x_i$  was the connectivity value from brain regions with significant dysconnectivity, and  $n$  was the total number of connectivity features. All  $x_i$  were corrected for age, sex and site effects.

The SCZ-I dataset was used as the training set for model training and selection, while the SCZ-II and COBRE datasets served as two independent test sets. To avoid biases in model selection, we applied nested tenfold cross-validation<sup>59</sup>, which contained inner and outer tenfold procedures used for tuning penalty parameters of feature functions, and testing model performances, respectively. To estimate the penalty parameters, we employed a randomized search strategy<sup>132</sup> (100,000 times for each inner round) as the search spaces were of high dimensions, thus limiting the use of grid search. We selected the model with

the best testing performance and further tested that model in the SCZ-II and COBRE datasets after retraining it on the whole SCZ-I dataset. To quantify the prediction performance of the model, Pearson's correlation and MAE between the observed and predicted scores were calculated. We then estimated the significance of predictions (that is,  $P_{perm}$  value) by performing altogether 10,000 permutation tests, in which we randomly shuffled the observed PANSS scores and re-obtained Pearson's *r* values, using the fitted models to create null distributions.

### Functional decode for dysconnectivity using Neurosynth

We applied the Neurosynth database (<https://neurosynth.org/>)<sup>61</sup>, a web-based platform for large-scale meta-analysis of published fMRI studies, to decode corticocortical and corticostriatal dysconnectivity *t*-maps by assessing their similarities to brain activation maps (also known as association test maps in *z*-scores) generated for different terms. In this analysis, we used 50 topic terms (<https://neurosynth.org/analyses/topics/v4-topics-50/>), which capture different aspects of human cognitive functions and are generated from more than 11,406 functional studies by a topic modeling approach<sup>133</sup>. Briefly, this method extracts distinct topics from article abstracts and assigns greater weights to terms that are closely related in meaning (semantically related words). Following that, a meta-analysis is performed to ascertain the neural associations of each topic by examining brain regions that consistently exhibit higher activation in articles strongly associated with each topic compared with those less associated with the topic. We computed cross-cortex spatial correlations (Pearson's *r*) between *t*-maps and 50 topic maps. The correlative signs (that is, positive or negative) were included in the calculations. The *t*-maps of both corticocortical and corticostriatal dysconnectivity from three SCZ datasets were decoded individually. To estimate the significance of each correlation (that is,  $P_{perm}$  value), we conducted 10,000 permutation tests to create a null distribution of correlation coefficients in each dataset. In particular, we first randomly shuffled the diagnostic labels (that is, SCZ or NC) from the beginning, then re-obtained *t*-maps for each permutation, and finally decoded them as described above to achieve arbitrary correlation coefficients. Only correlations reaching the threshold of  $P_{perm} < 0.05$  were reported. Notably, among the 50 topic terms, 'noise' terms that did not capture any coherent cognitive processes were also calculated, but were not shown in the main results. The term index and corresponding labels were referred to in a previous study<sup>134</sup>.

### Imaging transcriptomic analyses

**Cortical transcriptomic data acquisition and processing.** To investigate the underlying molecular mechanisms of the macroscopic dysconnectivity of SCZ, we used the whole-brain transcriptomic data from the AHBA<sup>67</sup>, which is an anatomically comprehensive adult human brain transcriptome dataset and can be freely accessed online (<https://human.brain-map.org/>). It provides gene expression data from six postmortem donors (male/female 5/1; ages  $42.5 \pm 13.38$  years, from 24 to 57 years; 4 donors have left hemisphere only, and 2 donors have both left and right hemispheres). AHBA data were preprocessed using the abagen toolbox<sup>135</sup>, which offers a whole standardizing workflow including the following steps: gene re-annotation, background filtering, probe selection, sample assignment, normalization and gene-set filtering. All these procedures were given default options, according to the recommendations of ref. 136. As our analysis specifically focused on the whole cerebral cortex, we provided a cortical mask for the processing workflow to obtain microarray gene expression data of tissue samples within the mask. Finally, we obtained an expression matrix with 1,733 samples  $\times$  15,633 genes based on the default settings. The MNI coordinates of these samples were also achieved simultaneously.

**Dysconnectivity-associated genes obtained by transcriptional correlation analysis.** We further identified dysconnectivity-associated genes by linking dysconnectivity *t*-maps with the transcriptomic



profiling of each gene. To do so, we included the obtained MNI coordinates of all cortical samples as seeds and extracted mean values of *t*-maps around these seeds within a 2 mm sphere. Then we calculated the spatial correlations (Pearson's *r*) between each gene expression (15,633 genes in total) and the extracted *t*-map values. Finally, genes were sorted in descending order from positive to negative based on the respective Pearson's *r*. The above analyses were performed for both corticocortical and corticostriatal *t*-maps from each SCZ dataset, resulting in six gene lists associated with dysconnectivity patterns of SCZ. Moreover, for each kind of dysconnectivity, a sorted gene list was achieved based on average values of AHBA spatial correlations across the three SCZ datasets. The subsequent analyses were mainly centered on the average sorted gene lists.

**GO enrichment analysis.** We performed GO enrichment analysis using Metascape<sup>137</sup> to identify significant biological processes underlying dysconnectivity-associated genes. For each kind of dysconnectivity, the average sorted gene list obtained from transcriptional correlation analysis was split into deciles, among which the top and bottom deciles of genes were analyzed. Metascape is a web-based platform (<https://metascape.org/gp/index.html#/main/step1>) that combines multiple features (for example, gene annotation, functional enrichment) within one integrated portal. In particular, it can automatically cluster enriched terms into non-redundant groups and represent them as an enrichment network. Minimal overlap and minimal enrichment options were set to 3 and 1.5, respectively. The significance of enriched terms was corrected by FDR at the level of 0.05, and discrete enrichment clusters were discarded. For visualization, enrichment networks were depicted with Cytoscape<sup>138</sup>.

### Identification of cell types associated with dysconnectivity based on single-cell transcriptional enrichment analysis

**Single-cell datasets.** Two cortical single-cell datasets were used in this study. The first was single-nucleus droplet-based-sequencing (snDrop-seq) data from postmortem brain samples covering the frontal cortex (Brodmann area (BA) 6, BA 9 and BA 10), and VIS (BA 17) from 6 different human adults, as reported by ref. 68. Of note, the originally released Lake data also contained brain samples from the lateral cerebellar hemisphere, which were excluded from our analysis since our study simply concentrated on the cerebral cortex. The Lake data are publicly available from the Gene Expression Omnibus (GSE97942). The second single-cell dataset was single-nucleus RNA-sequencing (snRNA-seq) data, which can also be publicly accessed through the Allen Brain Atlas data portal (<https://portal.brain-map.org/atlas-es-and-data/rnaseq>)<sup>69</sup>. The snRNA-seq data (which we called ABA data) were obtained from either human adult postmortem or resected brain samples from patients with epilepsy, including a total of eight donors, and involved the MTG area.

**Single-cell data processing and cell-type assignment.** We used identical processing pipelines for the two single-cell datasets with the Seurat toolkit (v3)<sup>139</sup>, consistent with those applied by ref. 71. First, initial filtering was conducted to remove genes expressed in less than three cells and exclude cells with fewer than 200 expressed genes. Then log normalization ('NormalizeData' function) was performed for gene expressions of each cell, based on total expression values, using a size factor of 10,000 molecules. Finally, we regressed out covariates of the sequencing platform and batch processing, and the residuals were scaled and centered ('ScaleData' function). We used predefined cell classes from each single-cell dataset and identified transcriptional markers (that is, differentially expressed genes) for each cell type based on the Wilcoxon rank-sum test ('FindMarkers' function).

As reported by ref. 71, we identified 16 cell classes for the Lake dataset, where all cell types were defined in both the frontal cortex and VIS, including five interneurons (In1, In3, In4, In6 and SST), five

excitatory neurons (Ex1, Ex3, Ex4, Ex5 and Ex8) and six non-neuronal cells (astrocyte (Ast), oligodendrocyte (Oli), pericyte (Per), endothelial (End), microglia (Mic) and oligodendrocyte precursor cell (OPC)). The ABA data originally defined 75 transcriptomically distinct cell types<sup>69</sup>, including 45 inhibitory neurons, 24 excitatory neurons and 6 non-neuronal cells. Cell types were annotated based on four segments: major cell class, layer enrichment, a subclass marker gene and a cell-specific marker gene. For example, 'Exc L2 LAMP5 LTK' indicated an excitatory neuron, enriched in layer 2, which expressed a subclass marker LAMP5 and a cell-specific marker LTK. In our analysis, we collapsed cell identities into overarching categories based on the major cell class and subclass marker genes. For example, 'Exc L2 LAMP5 LTK' was redesignated as 'Exc LAMP5'. By this means, the number of cell types for the ABA dataset was decreased to 17, a comparative quantity with that of the Lake cell classes. In addition, from the results of mapping ABA cell types onto Lake cell classes<sup>69</sup>, we observed that both 'Exc LAMP5' and 'Exc LINC00507' matched in the largest proportion to Ex1, among the 16 Lake cell types. We thus collapsed 'Exc LAMP5' and 'Exc LINC00507', and named 'Exc Ex1(ABA)' instead. Finally, we identified 16 cell types for the ABA data, including Exc\_Ex1(ABA), Exc\_RORB, Exc\_THEMIS, Exc\_SLC17A7, Exc\_FEZF2, Inh\_PAX6, Inh\_GAD1, Inh\_VIP, Inh\_SST, Inh\_LAMP5, Inh\_PVALB, Micro\_TYROBP, Astro\_FGFR3, OPC\_PDGFRA, Endo\_NOS-TRIN and Oligo\_OPALIN. (Note that the whitespace separation in the original cell names was substituted with an underscore.)

**Identification of dysconnectivity-related cell types by cell enrichment analysis.** We identified cell-type-specific genes, according to positive and significant differential expressions, that survived the Bonferroni correction at the level of 0.05. The positive value indicated a high level of expression of the respective gene in that particular cell type. For the Lake cells, these analyses were accomplished separately for the frontal cortex and VIS samples, and differentially expressed genes were intersected. For each single-cell dataset, we performed cell enrichment analysis for both corticocortical and corticostriatal dysconnectivity-associated genes obtained from the imaging transcriptomic analysis, using the FGSEA<sup>70</sup> method, which can quickly estimate arbitrarily low GSEA *P* values with high accuracy. The NES, indicating the enrichment score normalized to the mean enrichment of random genes with the same size in each cell, and the significance level (FDR adjusted *P*) were given in this analysis. The enrichment score was the same as in broad GSEA implementation<sup>140</sup>, which had both positive and negative values. In particular, we reported  $\text{sign}(\text{NES}) \times -\log_{10}(\text{adjusted } P)$  values, which indicated either a positive or negative enrichment, as well as the significance, simultaneously. If NES was positive, then  $\text{sign}(\text{NES})$  was equal to one, otherwise to minus one.

### Validation for dysconnectivity-related cell types by population-based and person-specific genomic analysis

**Population-based single-cell GWAS enrichment analysis.** SCZ GWAS data<sup>78</sup> were obtained from the PGC, including two independent summary statistics derived from the EAS and EUR populations, respectively. To validate whether polygenic risks for SCZ have enriched in cell-type-specific genes, MAGMA (v 1.08)<sup>76</sup> gene-set property analysis<sup>77</sup> was performed following previous studies<sup>29,71</sup>. Cell-specific gene expression was measured using the R package of EWCE<sup>141</sup> (Expression Weighted Cell type Enrichment). The whole pipeline of MAGMA analysis consisted of three basic steps and was performed for each GWAS dataset (that is, EAS and EUR) in each single-cell dataset (that is, Lake and ABA). First, an annotation procedure was carried out to map SNPs onto genes. For that, we set a 5-kb symmetrical window (that is, 5 kb upstream and 5 kb downstream) to extend the annotation region. Second, we performed gene analysis on SNPs to achieve gene-specific *P* values. The referenced panels (1,000 Genomes East Asian and European Phase 3)<sup>124</sup> used to account for the linkage disequilibrium between SNPs were matched with the GWAS summary statistics. Furthermore, gene meta-analysis

was conducted to integrate the results from both the EAS and EUR cohorts. Finally, gene property analysis was implemented by regression models to obtain gene-set  $P$  values. In particular, for the Lake dataset, the above analysis was separately conducted for the frontal cortex and VIS.

To further validate the results, we performed the above MAGMA enrichment analyses on two additional GWAS datasets. These two GWAS summary statistics were obtained by separately conducting the case-control association analysis on the combined data of SCZ-I and SCZ-II datasets (SCZ = 471, NC = 475), as well as the UKB dataset (SCZ = 886, NC = 274,508) using PLINK<sup>117</sup>. The detailed process of genotyping quality control was described above.

We also utilized the partitioned LDSC method<sup>79</sup> to establish the relationships between GWAS results and cell types, which was accomplished by examining the enrichment of common SNP heritability for SCZ in the most cell-type-specific genes. Specifically, the top-10% of specific genes (measured using EWCE) for each cell type were analyzed, and we chose the one-sided coefficient  $z$ -score  $P$  value to quantify the association.

**Person-specific single-cell PRS analysis.** We calculated cell-type-specific PRS based on person-specific genotype data in the SCZ-I and SCZ-II datasets using PLINK (v1.07)<sup>117</sup> to investigate whether specific cell types could exhibit a higher genetic risk in patients with SCZ relative to healthy participants. As our in-house genomic data were obtained from the Chinese population, we used independent Chinese GWAS summary statistics<sup>142</sup> as a reference to calculate the PRS for each participant. We first mapped GWAS SNPs onto genes and selected genes corresponding to 16 cell-type-specific gene sets achieved from the Lake single-cell dataset. Linkage disequilibrium adjustment via clumping<sup>143</sup> was then performed. The 'score' function provided in the PLINK was used to generate PRS for 5 common  $P$  thresholds: 0.1, 0.05, 0.01, 0.005 and 0.001. For each participant, we obtained cell-type-specific PRSs for 16 cell types  $\times$  5 different  $P$  thresholds. Finally, two-sample  $t$ -tests were calculated to evaluate cell-type-specific PRS differences between the SCZ and NC groups. Multiple comparisons were corrected by the FDR method. The same set of analyses was also carried out based on ABA cells.

In addition, we conducted an independent validation in the UKB dataset. Using the GWAS summary statistics from the EUR populations<sup>78</sup> as a reference, we calculated gene-set PRSs for both Lake and ABA cells under the primary threshold of  $P = 0.01$ . Given the significant imbalance in sample sizes between SCZ and NC groups in the UKB dataset, we divided the NC group into 310 subsets, with each subset having a sample size comparable to that of the SCZ group. The SCZ group was then compared with each NC subgroup using two-sample  $t$ -tests. We calculated the percentage of cell types showing significant differences (FDR < 0.05) among the 310 statistical results.

### Individual-level analysis combining dysconnectivity, cell types and PRS

This part of the analysis was implemented in the SCZ-I and SCZ-II datasets, including participants who had both imaging and genotype data.

**Calculation of individual-level dysconnectivity.** We first established a normative distribution delineating the typical range of variation in functional connectivity within the general population. For any participant, we then calculated the deviation of their functional connectivity relative to the normative distribution. By doing this, we could define dysconnectivity for patients with SCZ at an individual level. Nevertheless, even healthy participants showed deviations, which we interpreted as a manifestation of normal variation. Specifically, for both the SCZ-I and SCZ-II datasets, we performed analyses of multivariate linear regression using data from the NC group, following the formula:  $Y = \beta_0 + \beta_1 X_{\text{age}} + \beta_2 X_{\text{sex}} + \beta_3 X_{\text{site}} + \varepsilon$ , where  $Y$  was the functional connectivity score at each vertex, and the site variable was dummied. After that,

we calculated the deviation for each participant with the formula:  $Y_d = (Y_{\text{new}} - \beta_0 - \beta_1 W_{\text{age}} - \beta_2 W_{\text{sex}} - \beta_3 W_{\text{site}}) / (\text{s.d. of } \varepsilon)$ , where  $Y_{\text{new}}$  was the functional connectivity value of each vertex for the participant and  $Y_d$  was the deviation.  $W_{\text{age}}$ ,  $W_{\text{sex}}$  and  $W_{\text{site}}$  represented the participant's age, sex and dummied site, respectively. The coefficients of  $\beta$  and residuals of  $\varepsilon$  were obtained from regression models. In particular, the effects of age, sex, and site were considered in the deviation calculation.

**Individually linking genetic risk to dysconnectivity.** In alignment with the approach employed for the dysconnectivity  $t$ -map, we carried out single-cell transcriptional enrichment analyses on each participant to detect cell types associated with their dysconnectivity. Notably, individual-level dysconnectivity was analyzed for both patients with SCZ and healthy participants. To examine the link between cell-type-specific PRSs and dysconnectivity-associated cell types, we performed RSA<sup>80</sup>. Specifically, for each cell type, we first computed the pair-wise similarity between participants for the PRS using the Euclidean distance, which resulted in a similarity matrix. We then repeated this procedure for the dysconnectivity enrichment score (NES). Finally, we compared the PRS similarity matrix to the NES matrix using Spearman's correlation. The permutation test strategy was adopted to assess the significance of correlations between similarity matrices. For each cell type, the PRS matrix was permuted 10,000 times per row to create a null distribution.

### Statistical analysis

Details of the statistical methods, and the software and packages used in each relevant section have been described above, with all remaining analyses performed using Python v3.7. Before conducting any statistical analysis, participants with incomplete essential data or failing quality control, as well as those who did not meet the inclusion criteria, were excluded (see the 'Participants' section). The two-sample two-sided  $t$ -test was employed to assess differences between groups. A two-sided Pearson's correlation coefficient ( $r$ ) was used to measure the strength of correlation. We utilized two permutation test strategies, spin-based and non-spin-based (that is, random shuffle), with 10,000 iterations for each method to construct null distributions. The significance was determined using thresholds of  $P_{\text{spin}} < 0.05$  or  $P_{\text{perm}} < 0.05$ , as well as controlling for FWE and FDR at 0.05 for multiple testing.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The COBRE dataset can be accessed through the COINS data exchange portal<sup>112</sup> (<https://coins.trendscenter.org/>). The UKB data can be requested through a standard protocol (<https://www.ukbiobank.ac.uk/register-apply/>). The Brainnetome atlas can be downloaded at <https://atlas.brainnetome.org/download.html>. The Neurosynth database can be downloaded at <https://github.com/neurosynth/neurosynth-data>. The AHBA is freely available at <https://human.brain-map.org/>. The Lake<sup>68</sup> single-cell data are publicly available at the National Center for Biotechnology Information under the SuperSeries accession code GSE97942. The ABA<sup>69</sup> single-cell data are publicly available at <https://portal.brain-map.org/atlas-and-data/rnaseq>. The SCZ GWAS<sup>78</sup> data are publicly available at <https://pgc.unc.edu/for-researchers/download-results/>. All study data supporting the findings are provided within the paper and in its Supplementary Information. All raw data from the SCZ-I and SCZ-II datasets will be made available upon reasonable request to the corresponding authors.

### Code availability

The preprocessing software for resting-state fMRI data is freely available (BRANT<sup>114</sup> v3.35, <http://brant.brainnetome.org/en/latest/>). The

SurfStat toolbox for surface-wide statistical comparisons is freely available at <https://www.math.mcgill.ca/keith/surfstat/>. Corticocortical connectivity (first functional gradient) was calculated based on open-source codes at [https://github.com/NeuroanatomyAndConnectivity/gradient\\_analysis](https://github.com/NeuroanatomyAndConnectivity/gradient_analysis). The toolbox for performing spatial permutation test (spin test) is freely available at <https://github.com/spin-test/spin-test>. Functional decodings were performed based on the Neurosynth package openly available at <https://github.com/neurosynth/neurosynth>. The pyGAM package for performing symptom prediction analysis is openly available at <https://github.com/dswah/pyGAM>. The abagen toolbox for AHBA data processing is freely available at <https://github.com/rmarkello/abagen>. The pipeline for single-cell data processing is consistent with that performed by ref. 71, and the codes are openly available at [https://github.com/kevmmanderson/2020\\_PNAS\\_Depression](https://github.com/kevmmanderson/2020_PNAS_Depression). FGSEA was performed based on the fgsea package, which is openly available at <https://github.com/ctlab/fgsea>. The MAGMA (v1.08)<sup>76</sup> software and reference data are publicly available at <https://ctg.cncr.nl/software/magma>. The PLINK (v1.07)<sup>17</sup> software is freely available at <https://www.cog-genomics.org/plink/>. GO enrichment analysis was performed using the Metascape<sup>137</sup> platform (<https://metascape.org/gp/index.html#/main/step1>). The LDSC package is available at <https://github.com/bulik/ldsc>. All custom codes used in the analysis are publicly available at [https://github.com/BingLiu-Lab/scz\\_cross-scale\\_abnormalities](https://github.com/BingLiu-Lab/scz_cross-scale_abnormalities).

## References

- Owen, M. J., Sawa, A. & Mortensen, P. B. Schizophrenia. *Lancet* **388**, 86–97 (2016).
- Voineskos, A. N., Jacobs, G. R. & Ameis, S. H. Neuroimaging heterogeneity in psychosis: neurobiological underpinnings and opportunities for prognostic and therapeutic innovation. *Biol. Psychiatry* **88**, 95–102 (2020).
- Price, A. J., Jaffe, A. E. & Weinberger, D. R. Cortical cellular diversity and development in schizophrenia. *Mol. Psychiatry* **26**, 203–217 (2021).
- van den Heuvel, M. P., Scholtens, L. H., de Reus, M. A. & Kahn, R. S. Associated microscale spine density and macroscale connectivity disruptions in schizophrenia. *Biol. Psychiatry* **80**, 293–301 (2016).
- Scholtens, L. H. & van den Heuvel, M. P. Multimodal connectomics in psychiatry: bridging scales from micro to macro. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **3**, 767–776 (2018).
- Keshavan, M. S. et al. Neuroimaging in schizophrenia. *Neuroimaging Clin. N. Am.* **30**, 73–83 (2020).
- Gur, R. E. & Gur, R. C. Functional magnetic resonance imaging in schizophrenia. *Dialogues Clin. Neurosci.* **12**, 333–343 (2010).
- Smith, S. M. et al. Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* **17**, 666–682 (2013).
- Fornito, A., Zalesky, A., Pantelis, C. & Bullmore, E. T. Schizophrenia, neuroimaging and connectomics. *Neuroimage* **62**, 2296–2314 (2012).
- van den Heuvel, M. P. & Fornito, A. Brain networks in schizophrenia. *Neuropsychol. Rev.* **24**, 32–48 (2014).
- Li, T. et al. Brain-wide analysis of functional connectivity in first-episode and chronic stages of schizophrenia. *Schizophr. Bull.* **43**, 436–448 (2017).
- Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P. & Mechelli, A. Dysconnectivity in schizophrenia: where are we now? *Neurosci. Biobehav. Rev.* **35**, 1110–1124 (2011).
- Lynall, M.-E. et al. Functional connectivity and brain networks in schizophrenia. *J. Neurosci.* **30**, 9477–9487 (2010).
- Woodward, N. D., Karbasforoushan, H. & Heckers, S. Thalamocortical dysconnectivity in schizophrenia. *Am. J. Psychiatry* **169**, 1092–1099 (2012).
- Fornito, A. et al. Functional dysconnectivity of corticostriatal circuitry as a risk phenotype for psychosis. *JAMA Psychiatry* **70**, 1143–1151 (2013).
- Li, A. et al. A neuroimaging biomarker for striatal dysfunction in schizophrenia. *Nat. Med.* **26**, 558–565 (2020).
- Sheffield, J. M. & Barch, D. M. Cognition and resting-state functional connectivity in schizophrenia. *Neurosci. Biobehav. Rev.* **61**, 108–120 (2016).
- Rotarska-Jagiela, A. et al. Resting-state functional network correlates of psychotic symptoms in schizophrenia. *Schizophr. Res.* **117**, 21–30 (2010).
- Shukla, D. K. et al. Aberrant frontostriatal connectivity in negative symptoms of schizophrenia. *Schizophr. Bull.* **45**, 1051–1059 (2019).
- Brady, R. O. et al. Cerebellar-prefrontal network connectivity and negative symptoms in schizophrenia. *Am. J. Psychiatry* **176**, 512–520 (2019).
- Chen, J. et al. Intrinsic connectivity patterns of task-defined brain networks allow individual prediction of cognitive symptom dimension of schizophrenia and are linked to molecular architecture. *Biol. Psychiatry* **89**, 308–319 (2021).
- Adhikari, B. M. et al. Functional network connectivity impairments and core cognitive deficits in schizophrenia. *Hum. Brain Mapp.* **40**, 4593–4605 (2019).
- Adhikari, B. M. et al. Effects of ketamine and midazolam on resting state connectivity and comparison with ENIGMA connectivity deficit patterns in schizophrenia. *Hum. Brain Mapp.* **41**, 767–778 (2019).
- Singh, T. et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* **604**, 509–516 (2022).
- Trubetskoy, V. et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
- Harrison, P. J. Postmortem studies in schizophrenia. *Dialogues Clin. Neurosci.* **2**, 349–357 (2000).
- Roeske, M. J., Konradi, C., Heckers, S. & Lewis, A. S. Hippocampal volume and hippocampal neuron density, number and size in schizophrenia: a systematic review and meta-analysis of postmortem studies. *Mol. Psychiatry* **26**, 3524–3535 (2021).
- Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
- Watanabe, K., Umičević Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P. & Posthuma, D. Genetic mapping of cell type specificity for complex traits. *Nat. Commun.* **10**, 3222 (2019).
- Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).
- Calderon, D. et al. Inferring relevant cell types for complex traits by using single-cell gene expression. *Am. J. Hum. Genet.* **101**, 686–699 (2017).
- Ruzicka, W. B. et al. Single-cell dissection of schizophrenia reveals neurodevelopmental-synaptic axis and transcriptional resilience. Preprint at medRxiv <https://doi.org/10.1101/2020.11.06.20225342> (2020).
- Räsänen, N., Tiihonen, J., Koskivi, M., Lehtonen, Š. & Koistinaho, J. The iPSC perspective on schizophrenia. *Trends Neurosci.* **45**, 8–26 (2022).
- Sebastian, R., Song, Y. & Pak, C. Probing the molecular and cellular pathological mechanisms of schizophrenia using human induced pluripotent stem cell models. *Schizophr. Res.* <https://doi.org/10.1016/j.schres.2022.06.028> (2022).
- Notaras, M. et al. Schizophrenia is defined by cell-specific neuropathology and multiple neurodevelopmental mechanisms in patient-derived cerebral organoids. *Mol. Psychiatry* **27**, 1416–1434 (2022).
- Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl Acad. Sci. USA* **113**, 12574–12579 (2016).
- Aine, C. J. et al. Multimodal neuroimaging in schizophrenia: description and dissemination. *Neuroinformatics* **15**, 343–364 (2017).



38. Coifman, R. R. & Lafon, S. Diffusion maps. *Appl. Comput. Harmon. Anal.* **21**, 5–30 (2006).
39. McCutcheon, R. A., Abi-Dargham, A. & Howes, O. D. Schizophrenia, dopamine and the striatum: from biology to symptoms. *Trends Neurosci.* **42**, 205–220 (2019).
40. Shepherd, G. M. G. Corticostriatal connectivity and its role in disease. *Nat. Rev. Neurosci.* **14**, 278–291 (2013).
41. Alexander-Bloch, A. F. et al. On testing for spatial correspondence between maps of human brain structure and function. *Neuroimage* **178**, 540–551 (2018).
42. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142–2154 (2012).
43. Murphy, K. & Fox, M. D. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* **154**, 169–173 (2017).
44. Hong, S.-J. et al. Atypical functional connectome hierarchy in autism. *Nat. Commun.* **10**, 1022 (2019).
45. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
46. Friston, K. J. & Frith, C. D. Schizophrenia: a disconnection syndrome? *Clin. Neurosci.* **3**, 89–97 (1995).
47. Stephan, K. E., Baldeweg, T. & Friston, K. J. Synaptic plasticity and dysconnection in schizophrenia. *Biol. Psychiatry* **59**, 929–939 (2006).
48. Friston, K. J. The disconnection hypothesis. *Schizophr. Res.* **30**, 115–125 (1998).
49. Friston, K., Brown, H. R., Siemerkus, J. & Stephan, K. E. The dysconnection hypothesis (2016). *Schizophr. Res.* **176**, 83–94 (2016).
50. Stephan, K. E., Friston, K. J. & Frith, C. D. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophr. Bull.* **35**, 509–527 (2009).
51. Finn, E. S. et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
52. Worsley, K. J., Andermann, M., Koulis, T., MacDonald, D. & Evans, A. C. Detecting changes in nonisotropic images. *Hum. Brain Mapp.* **8**, 98–101 (1999).
53. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl Acad. Sci. USA* **113**, 7900–7905 (2016).
54. Nichols, T. E. & Holmes, A. P. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* **15**, 1–25 (2002).
55. Kay, S. R., Fiszbein, A. & Opler, L. A. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
56. Hastie, T. & Tibshirani, R. Generalized additive models. *Stat. Sci.* **1**, 297–310 (1986).
57. Hastie, T. & Tibshirani, R. Generalized additive models for medical research. *Stat. Methods Med. Res.* **4**, 187–196 (1995).
58. Hastie, T. & Tibshirani, R. Generalized additive models: some applications. *J. Am. Stat. Assoc.* **82**, 371–386 (1987).
59. Varoquaux, G. et al. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *Neuroimage* **145**, 166–179 (2017).
60. Lindenmayer, J.-P., Bernstein-Hyman, R. & Grochowski, S. A new five factor model of schizophrenia. *Psychiatr. Q.* **65**, 299–322 (1994).
61. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
62. Hooker, C. & Park, S. Emotion processing and its relationship to social functioning in schizophrenia patients. *Psychiatry Res.* **112**, 41–50 (2002).
63. Lincoln, T. M., Mehl, S., Kesting, M.-L. & Rief, W. Negative symptoms and social cognition: identifying targets for psychological interventions. *Schizophr. Bull.* **37**, S23–S32 (2011).
64. Moura, B. M. et al. A network of psychopathological, cognitive, and motor symptoms in schizophrenia spectrum disorders. *Schizophr. Bull.* **47**, 915–926 (2021).
65. Guillem, F., Rinaldi, M., Pampoulova, T. & Stip, E. The complex relationships between executive functions and positive symptoms in schizophrenia. *Psychol. Med.* **38**, 853–860 (2008).
66. Koshiyama, D. et al. Hierarchical pathways from sensory processing to cognitive, clinical, and functional impairments in schizophrenia. *Schizophr. Bull.* **47**, 373–385 (2021).
67. Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
68. Lake, B. B. et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* **36**, 70–80 (2018).
69. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
70. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* <https://doi.org/10.1101/060012> (2021).
71. Anderson, K. M. et al. Convergent molecular, cellular, and cortical neuroimaging signatures of major depressive disorder. *Proc. Natl Acad. Sci. USA* **117**, 25138–25149 (2020).
72. Lake, B. B. et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
73. Trifu, S. C., Kohn, B., Vlasie, A. & Patrichi, B.-E. Genetics of schizophrenia (Review). *Exp. Ther. Med.* **20**, 3462–3468 (2020).
74. Gejman, P. V., Sanders, A. R. & Duan, J. The role of genetics in the etiology of schizophrenia. *Psychiatr. Clin. North Am.* **33**, 35–66 (2010).
75. Krabbendam, L. & van Os, J. Schizophrenia and urbanicity: a major environmental influence—conditional on genetic risk. *Schizophr. Bull.* **31**, 795–799 (2005).
76. Leeuw, C. A., de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
77. de Leeuw, C. A., Neale, B. M., Heskes, T. & Posthuma, D. The statistical properties of gene-set analysis. *Nat. Rev. Genet.* **17**, 353–364 (2016).
78. Lam, M. et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat. Genet.* **51**, 1670–1678 (2019).
79. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
80. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis—connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
81. Nath, M., Wong, T. P. & Srivastava, L. K. Neurodevelopmental insights into circuit dysfunction in schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **104**, 110047 (2021).
82. Heeger, D. J. & Ress, D. What does fMRI tell us about neuronal activity? *Nat. Rev. Neurosci.* **3**, 142–151 (2002).
83. Dienel, S. J., Schoonover, K. E. & Lewis, D. A. Cognitive dysfunction and prefrontal cortical circuit alterations in schizophrenia: developmental trajectories. *Biol. Psychiatry* **92**, 450–459 (2022).
84. Lewis, D. A. & Sweet, R. A. Schizophrenia from a neural circuitry perspective: advancing toward rational pharmacological therapies. *J. Clin. Invest.* **119**, 706–716 (2009).

85. Howes, O. D. et al. Molecular imaging studies of the striatal dopaminergic system in psychosis and predictions for the prodromal phase of psychosis. *Br. J. Psychiatry* **191**, s13–s18 (2007).
86. Romme, I. A. C., de Reus, M. A., Ophoff, R. A., Kahn, R. S. & van den Heuvel, M. P. Connectome disconnectivity and cortical gene expression in patients with schizophrenia. *Biol. Psychiatry* **81**, 495–502 (2017).
87. Lewis, D. A. Neuroplasticity of excitatory and inhibitory cortical circuits in schizophrenia. *Dialogues Clin. Neurosci.* **11**, 269–280 (2009).
88. Tkachev, D. et al. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. *Lancet* **362**, 798–805 (2003).
89. Raabe, F. J. et al. Studying and modulating schizophrenia-associated dysfunctions of oligodendrocytes with patient-specific cell systems. *NPJ Schizophr.* **4**, 23 (2018).
90. Spruston, N. Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* **9**, 206–221 (2008).
91. Han, W. & Šestan, N. Cortical projection neurons: sprung from the same root. *Neuron* **80**, 1103–1105 (2013).
92. Michalski, J.-P. & Kothary, R. Oligodendrocytes in a nutshell. *Front. Cell. Neurosci.* **9**, 340 (2015).
93. Tomassy, G. S. et al. Distinct profiles of myelin distribution along single axons of pyramidal neurons in the neocortex. *Science* **344**, 319–324 (2014).
94. de Hoz, L. & Simons, M. The emerging functions of oligodendrocytes in regulating neuronal network behaviour. *Bioessays* **37**, 60–69 (2015).
95. Gibson, E. M. et al. Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. *Science* **344**, 1252304 (2014).
96. Hof, P. R. et al. Loss and altered spatial distribution of oligodendrocytes in the superior frontal gyrus in schizophrenia. *Biol. Psychiatry* **53**, 1075–1085 (2003).
97. Windrem, M. S. et al. Human iPSC glial mouse chimeras reveal glial contributions to schizophrenia. *Cell Stem Cell* **21**, 195–208.e6 (2017).
98. Aberg, K., Saetre, P., Jareborg, N. & Jazin, E. Human QKI, a potential regulator of mRNA expression of human oligodendrocyte-related genes involved in schizophrenia. *Proc. Natl Acad. Sci. USA* **103**, 7482–7487 (2006).
99. Harris, K. D. & Shepherd, G. M. G. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
100. Lee, J. H., Liu, Q. & Dadgar-Kiani, E. Solving brain circuit function and dysfunction with computational modeling and optogenetic fMRI. *Science* **378**, 493–499 (2022).
101. Grimm, C. et al. Optogenetic activation of striatal D1R and D2R cells differentially engages downstream connected areas beyond the basal ganglia. *Cell Rep.* **37**, 110161 (2021).
102. Zhang, D. et al. Spatial epigenome-transcriptome co-profiling of mammalian tissues. *Nature* **616**, 113–122 (2023).
103. Huber, L. et al. Layer-dependent functional connectivity methods. *Prog. Neurobiol.* **207**, 101835 (2021).
104. Carbon, M. & Correll, C. U. Thinking and acting beyond the positive: the role of the cognitive and negative symptoms in schizophrenia. *CNS Spectr.* **19**, 35–53 (2014).
105. Kaar, S. J., Natesan, S., McCutcheon, R. & Howes, O. D. Antipsychotics: mechanisms underlying clinical response and side-effects and novel treatment approaches based on pathophysiology. *Neuropharmacology* **172**, 107704 (2020).
106. Valiengo, L. et al. Efficacy and safety of transcranial direct current stimulation for treating negative symptoms in schizophrenia. *JAMA Psychiatry* **77**, 121–129 (2020).
107. Kostova, R., Cecere, R., Thut, G. & Uhlhaas, P. J. Targeting cognition in schizophrenia through transcranial direct current stimulation: a systematic review and perspective. *Schizophr. Res.* **220**, 300–310 (2020).
108. Koponen, L. M., Nieminen, J. O. & Ilmoniemi, R. J. Multi-locus transcranial magnetic stimulation-theory and implementation. *Brain Stimul.* **11**, 849–855 (2018).
109. Abera, A. S., Wang, B., Grill, W. M. & Peterchev, A. V. Simulation of transcranial magnetic stimulation in head model with morphologically-realistic cortical neurons. *Brain Stimul.* **13**, 175–189 (2020).
110. Qiu, Y. et al. On-demand cell-autonomous gene therapy for brain circuit disorders. *Science* **378**, 523–532 (2022).
111. Martins, D. et al. Imaging transcriptomics: convergent cellular, transcriptomic, and molecular neuroimaging signatures in the healthy adult human brain. *Cell Rep.* **37**, 110173 (2021).
112. Wood, D. et al. Harnessing modern web application technology to create intuitive and efficient data visualization and sharing tools. *Front. Neuroinformatics* **8**, 71 (2014).
113. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
114. Xu, K., Liu, Y., Zhan, Y., Ren, J. & Jiang, T. BRANT: a versatile and extendable resting-state fMRI toolkit. *Front. Neuroinformatics* **12**, 52 (2018).
115. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. & Turner, R. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* **35**, 346–355 (1996).
116. Fischl, B., Sereno, M. I., Tootell, R. B. H. & Dale, A. M. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* **8**, 272–284 (1999).
117. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
118. Liu, B. et al. Polygenic risk for schizophrenia influences cortical gyrification in 2 independent general populations. *Schizophr. Bull.* **43**, 673–680 (2017).
119. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
120. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
121. Thorisson, G. A., Smith, A. V., Krishnan, L. & Stein, L. D. The International HapMap Project web site. *Genome Res.* **15**, 1592–1593 (2005).
122. Delaneau, O., Marchini, J. & Zagury, J.-F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9**, 179–181 (2011).
123. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
124. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
125. Huntenburg, J. M., Bazin, P.-L. & Margulies, D. S. Large-scale gradients in human cortical organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).
126. Zhao, Y. et al. The development of cortical functional hierarchy is associated with the molecular organization of prenatal/postnatal periods. *Cereb. Cortex* <https://doi.org/10.1093/cercor/bhac340> (2022).
127. Langs, G., Golland, P. & Ghosh, S. S. Predicting activation across individuals with resting-state functional connectivity based multi-atlas label fusion. *Med. Image Comput. Comput. Assist. Interv.* **9350**, 313–320 (2015).
128. Fan, L. et al. The human Brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb. Cortex* **26**, 3508–3526 (2016).
129. Bethlehem, R. A. I. et al. Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).

130. Baum, G. L. et al. Development of structure–function coupling in human brain networks during youth. *Proc. Natl Acad. Sci. USA* **117**, 771–778 (2020).
131. Eilers, P. H. C. & Marx, B. D. Flexible smoothing with B-splines and penalties. *Stat. Sci.* **11**, 89–121 (1996).
132. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13**, 281–305 (2012).
133. Poldrack, R. A. et al. Discovering relations between mind, brain, and mental disorders using topic mapping. *PLoS Comput. Biol.* **8**, e1002707 (2012).
134. Wang, H.-T. et al. Neurocognitive patterns dissociating semantic processing from executive control are linked to more detailed off-task mental time travel. *Sci. Rep.* **10**, 11904 (2020).
135. Markello, R. D. et al. Standardizing workflows in imaging transcriptomics with the abagen toolbox. *eLife* **10**, e72129 (2021).
136. Arnatkevičiūtė, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage* **189**, 353–367 (2019).
137. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
138. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
139. Stuart, T. et al. Comprehensive INtegration of Single-cell Data. *Cell* **177**, 1888–1902.e21 (2019).
140. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
141. Skene, N. G. & Grant, S. G. N. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. *Front. Neurosci.* **10**, 16 (2016).
142. Li, Z. et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat. Genet.* **49**, 1576–1583 (2017).
143. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).

## Acknowledgements

This work was supported by the Startup Funds of Beijing Normal University (to B.L.), the National Key Basic Research and Development Program (973) (grant 2011CB707800 to T.J.), the National Key Research and Development Plan (grant 2016YFC0904300 to B.L.), the Natural Science Foundation of China (grant 81771451 to B.L.; grant

82171543 to A.L.), the Science and Technology Innovation 2030—Brain Science and Brain-Inspired Intelligence Project of China (grant 2021ZD0200200 to T.J.).

## Author contributions

B.L. and T.J. led the project. B.L. and M.W. were responsible for the study concept and the design of the study. H.Y. and W.Y. provided crucial advice for the study. A.L. made substantial contributions to the paper and provided critical comments. M.W. and B.L. analyzed the data, created the figures and wrote the paper. Y.L., L.F., K.H., Y.S., Y.Z., J.L., X.T. and M.S. participated in discussions of the results and the paper. P.L., J.C., Y.C., Huaning Wang, W.L., Z.L., Y.Y., H.G., L. Lv, L. Lu, J.Y., Huiling Wang, H.Z., H. Wu, Y.N. and D.Z. contributed to the data acquisition.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s44220-023-00110-3>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44220-023-00110-3>.

**Correspondence and requests for materials** should be addressed to Ang Li, Tianzi Jiang or Bing Liu.

**Peer review information** *Nature Mental Health* thanks Marta Bosia, Katharina Schmack and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

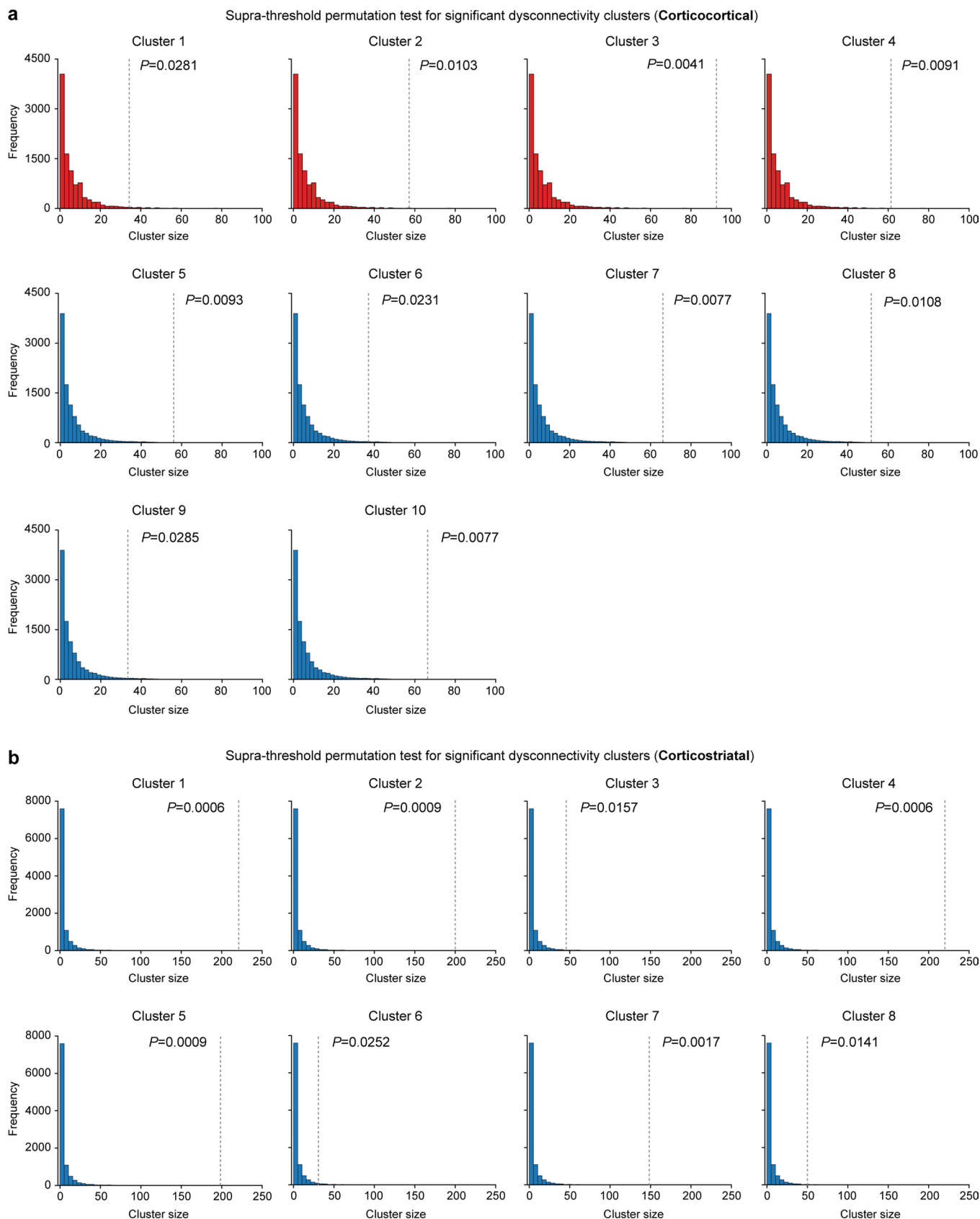
**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2023

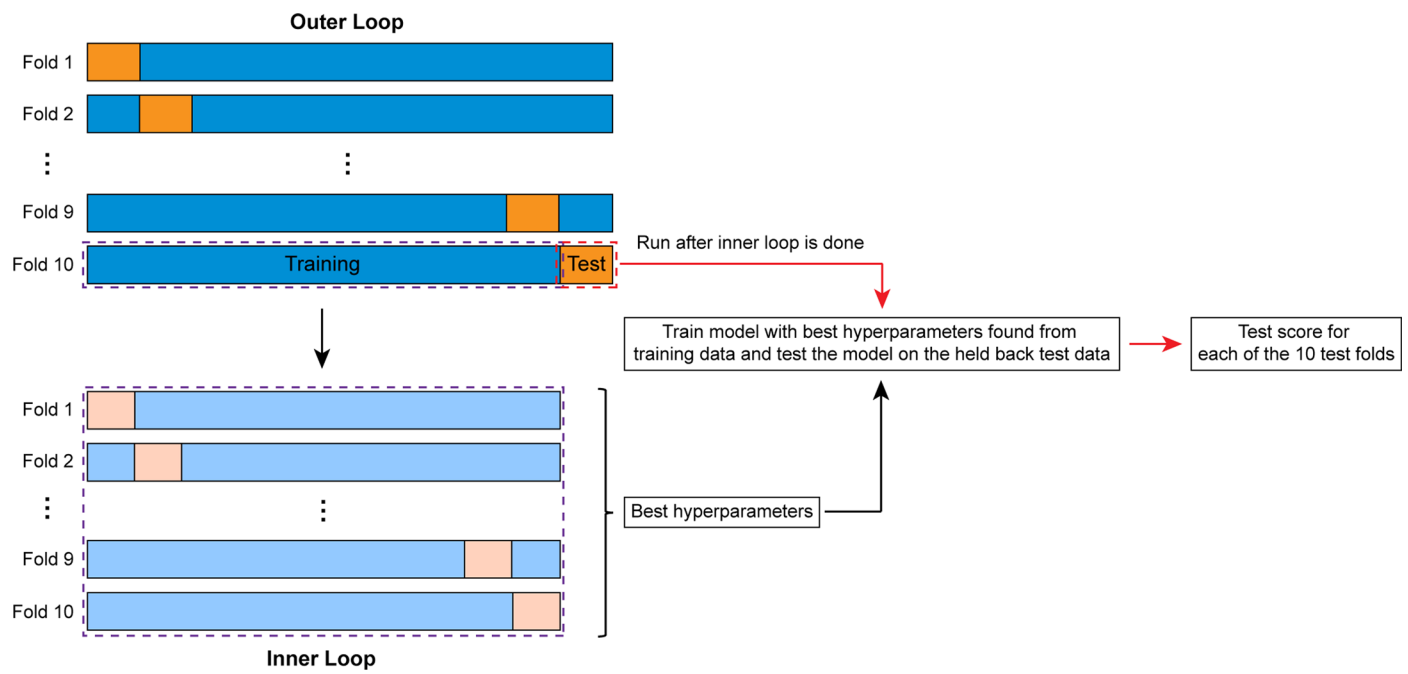
<sup>1</sup>State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China. <sup>2</sup>Peking University Sixth Hospital, Peking University Institute of Mental Health, Beijing, China. <sup>3</sup>NHC Key Laboratory of Mental Health (Peking University) and National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Beijing, China. <sup>4</sup>School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing, China. <sup>5</sup>Brainnetome Center, Institute of Automation, Chinese Academy of Sciences, Beijing, China. <sup>6</sup>School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China. <sup>7</sup>Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing, China. <sup>8</sup>Department of Radiology, Renmin Hospital of Wuhan University, Wuhan, China. <sup>9</sup>Department of Psychiatry, Xijing Hospital, The Fourth Military Medical University, Xi'an, China. <sup>10</sup>Zhumadian Psychiatric Hospital, Zhumadian, China. <sup>11</sup>Department of Psychiatry, Henan Mental Hospital, The Second Affiliated Hospital of Xinxiang Medical University, Xinxiang, China. <sup>12</sup>Henan Key Lab of Biological Psychiatry of Xinxiang Medical University, International Joint Research Laboratory for Psychiatry and Neuroscience of Henan, Xinxiang, China. <sup>13</sup>Department of Psychiatry, Renmin Hospital of Wuhan University, Wuhan, China. <sup>14</sup>Department of Psychology, Xinxiang Medical University, Xinxiang, China. <sup>15</sup>The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou, China. <sup>16</sup>Center for Life Sciences/ PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China. <sup>17</sup>State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. <sup>18</sup>Research Center for Augmented Intelligence, Zhejiang Lab, Hangzhou, China. <sup>19</sup>Innovation Academy for Artificial Intelligence, Chinese Academy of Sciences, Beijing, China. <sup>20</sup>IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. <sup>21</sup>Chinese Institute for Brain Research, Beijing, China. ✉ e-mail: [al@ibp.ac.cn](mailto:al@ibp.ac.cn); [jiangtz@nlpr.ia.ac.cn](mailto:jiangtz@nlpr.ia.ac.cn); [bing.liu@bnu.edu.cn](mailto:bing.liu@bnu.edu.cn)



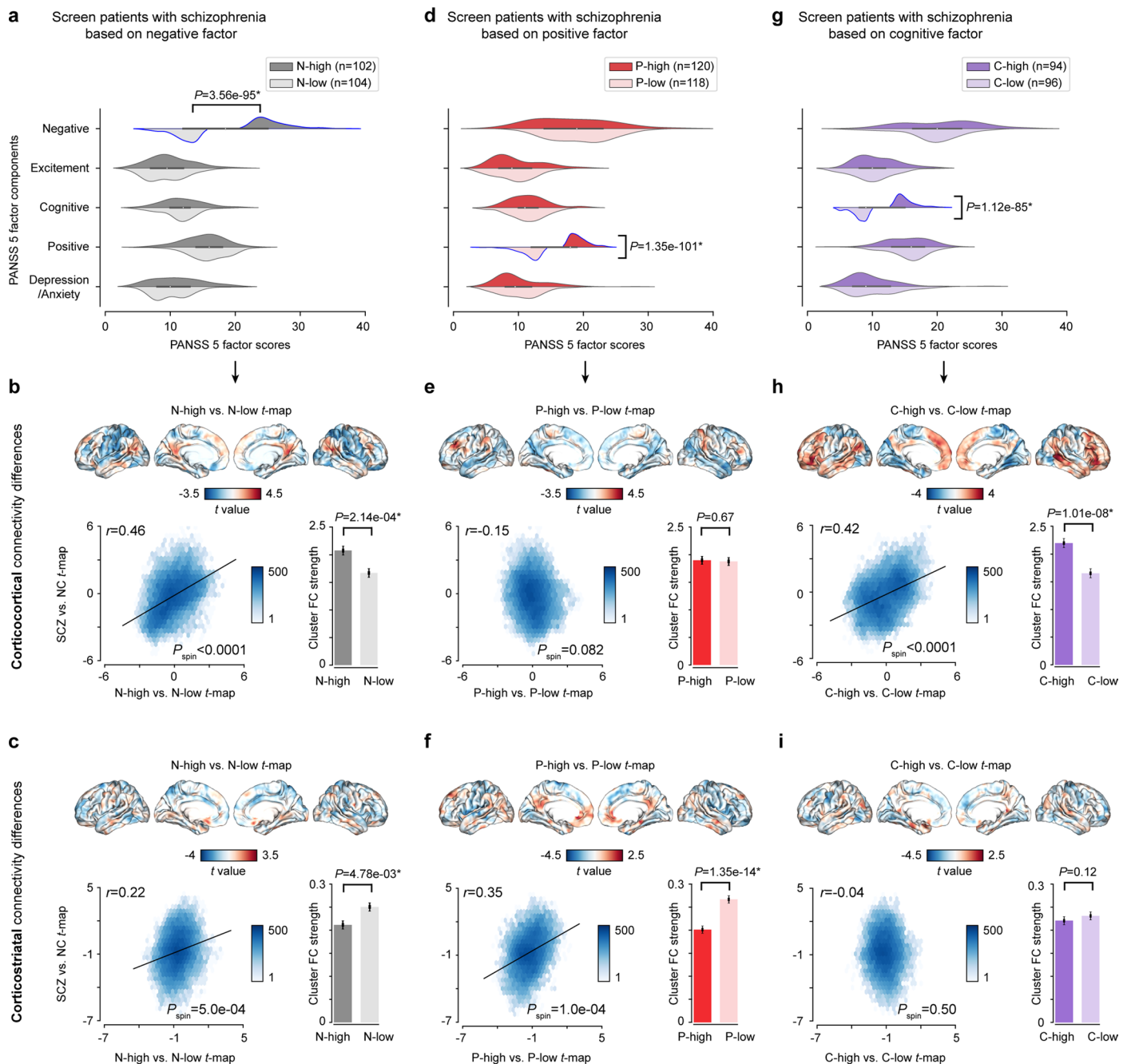


**Extended Data Fig. 1 | Supra-threshold permutation test based on cluster size (10,000 iterations, cluster defining threshold (cdt)=0.001).** Surface-based corticocortical and corticostriatal connectivity comparisons (based on a

two-sample two-sided t-test) between patients with SCZ and NC, with multiple comparisons corrected using permutation-based cluster thresholding instead of random field theory (see Fig. 2a,c).



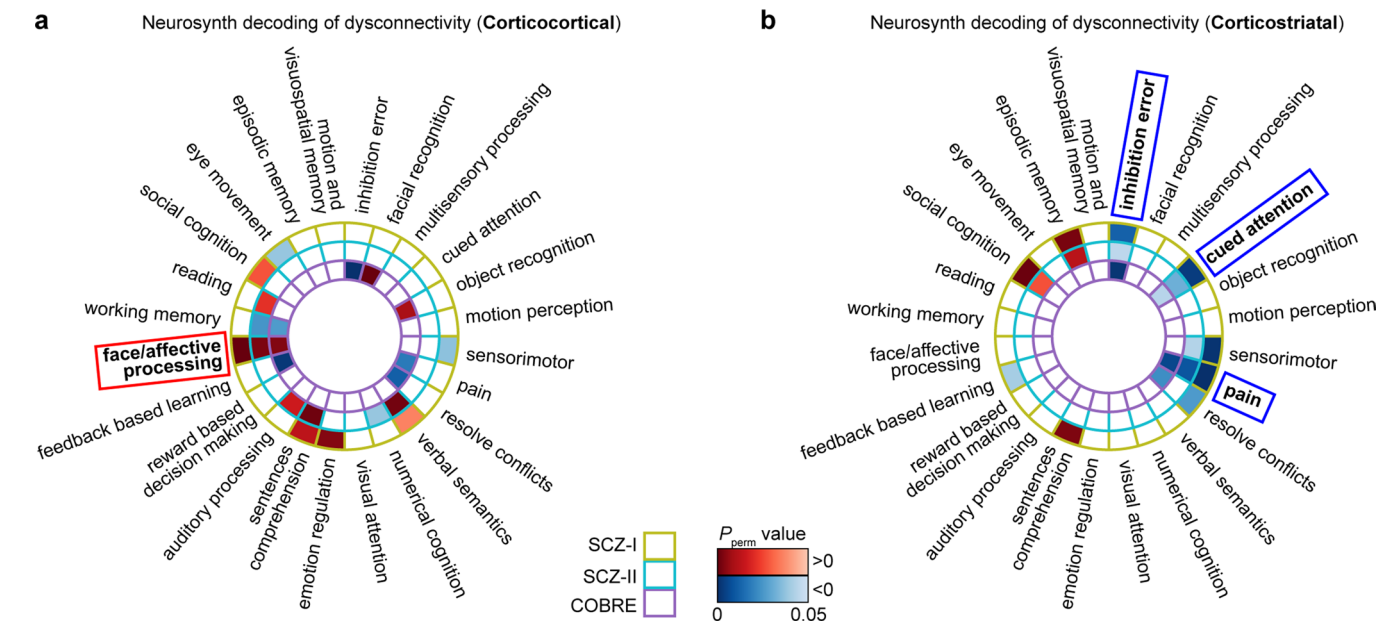
**Extended Data Fig. 2 | Nested 10-fold cross-validation process.** The procedure was applied for model training and selection in the SCZ-I dataset.



**Extended Data Fig. 3 | Comparisons of functional connectivity between high and low PANSS dimension subgroups.** **a.** Based on the negative dimension of the PANSS 5-factor model, the top and bottom 30% of SCZ patients ( $n = 513$ ) from the combined SCZ-I and SCZ-II datasets were selected, forming two subgroups: N-high and N-low. The two subgroups were further screened using a two-sample  $t$ -test (two-sided) to ensure no differences in the other four PANSS dimensions. **b.** Corticocortical connectivity was compared between the N-high and N-low subgroups using the methodology in Fig. 1, resulting in a differential map, termed N-map. The N-map demonstrated a significant positive spatial correlation with the dysconnectivity  $t$ -map from the SCZ and NC group comparison (see Fig. 1). Mean connectivity scores from the ten dysconnectivity clusters (Fig. 2a) were significantly higher in the N-high subgroup. **c.** Corticostriatal connectivity was compared between the N-high and N-low subgroups. The corticostriatal N-map and dysconnectivity  $t$ -map displayed a significant positive spatial correlation. The absolute connectivity values from eight clusters (Fig. 2c) significantly decreased in the N-high subgroup. **d-f.** A similar analysis, like panels **a-c**, compared functional connectivity between high and low positive subgroups. The

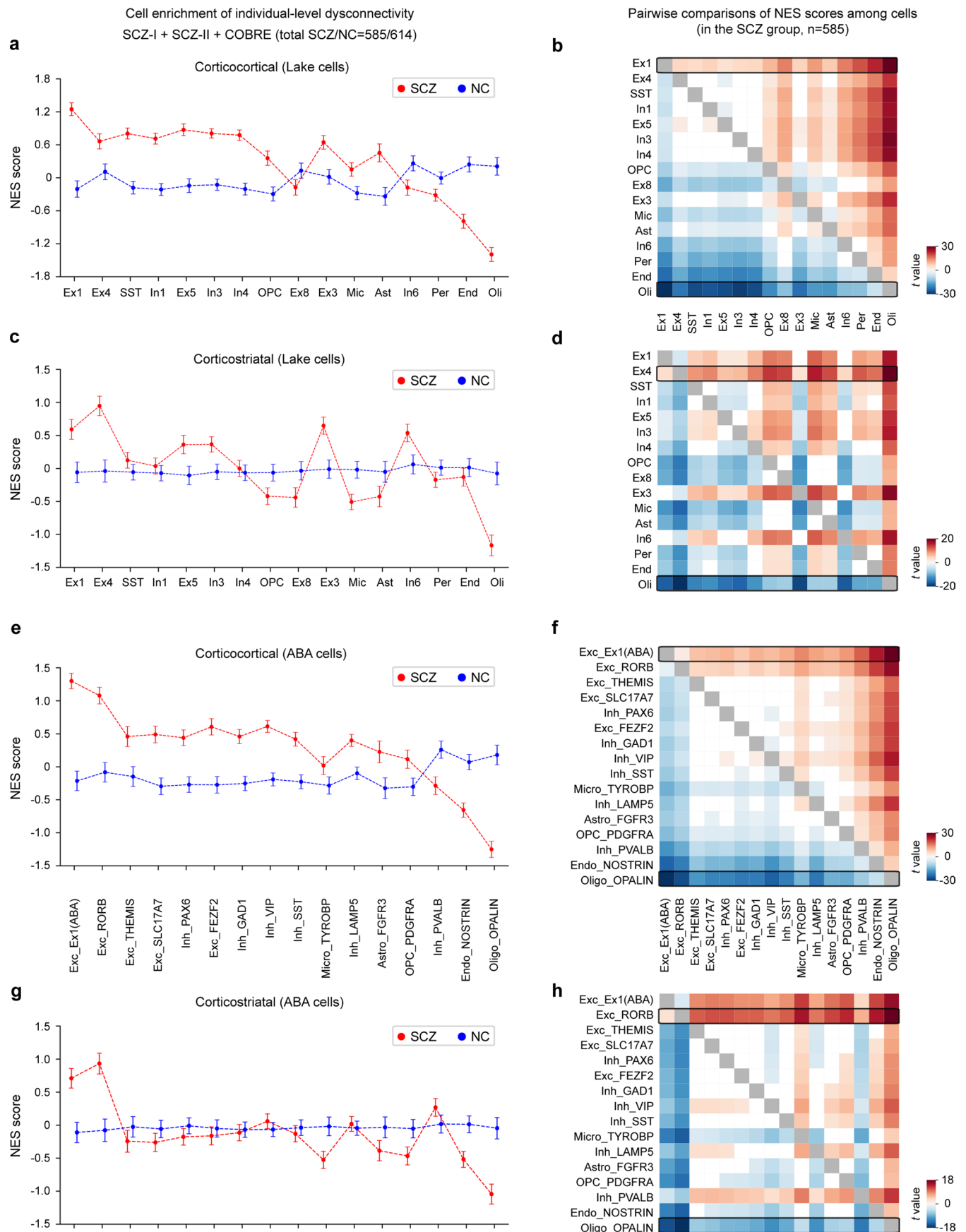
corticocortical P-map and dysconnectivity  $t$ -map showed no correlation, with no significant differences in connectivity strength (absolute value) of the ten clusters between P-high and P-low subgroups (**e**). However, a significant positive correlation was found between the corticostriatal P-map and dysconnectivity  $t$ -map, while the eight clusters' connectivity strength significantly decreased in the P-high subgroup (**f**). **g-i.** Like panels **a-c** and **d-f**, functional connectivity differences were examined between the high and low cognitive subgroups. Significant differences were found in corticocortical connectivity for C-high and C-low subgroups, while corticostriatal connectivity showed no significant differences. Spatial correlations between maps were quantified using two-sided Pearson's  $r$ . The significance ( $P_{\text{spin}}$ ) was determined using the spin-based permutation test. Connectivity strength differences were measured using the two-sample  $t$ -test (two-sided). The error bars represent mean values  $\pm$  95% confidence interval. The box plot presents minimum, 25th percentile, median, 75th percentile, and maximum values (excluding outliers) for the combined high and low groups.





**Extended Data Fig. 4 | Functional decodings of corticocortical and corticostriatal dysconnectivity.** The spatial correlations (two-sided Pearson's  $r$ ) between  $t$ -maps and 24 predefined topic maps (each is composed of related terms) from the Neurosynth database were calculated. The significance ( $P_{\text{perm}}$ ) was estimated using permutation tests. The warm or cool color indicates a positive or negative correlation, respectively. Only significant correlations ( $P_{\text{perm}} < 0.05$ ) are shown with colors, and darker colors indicate a higher

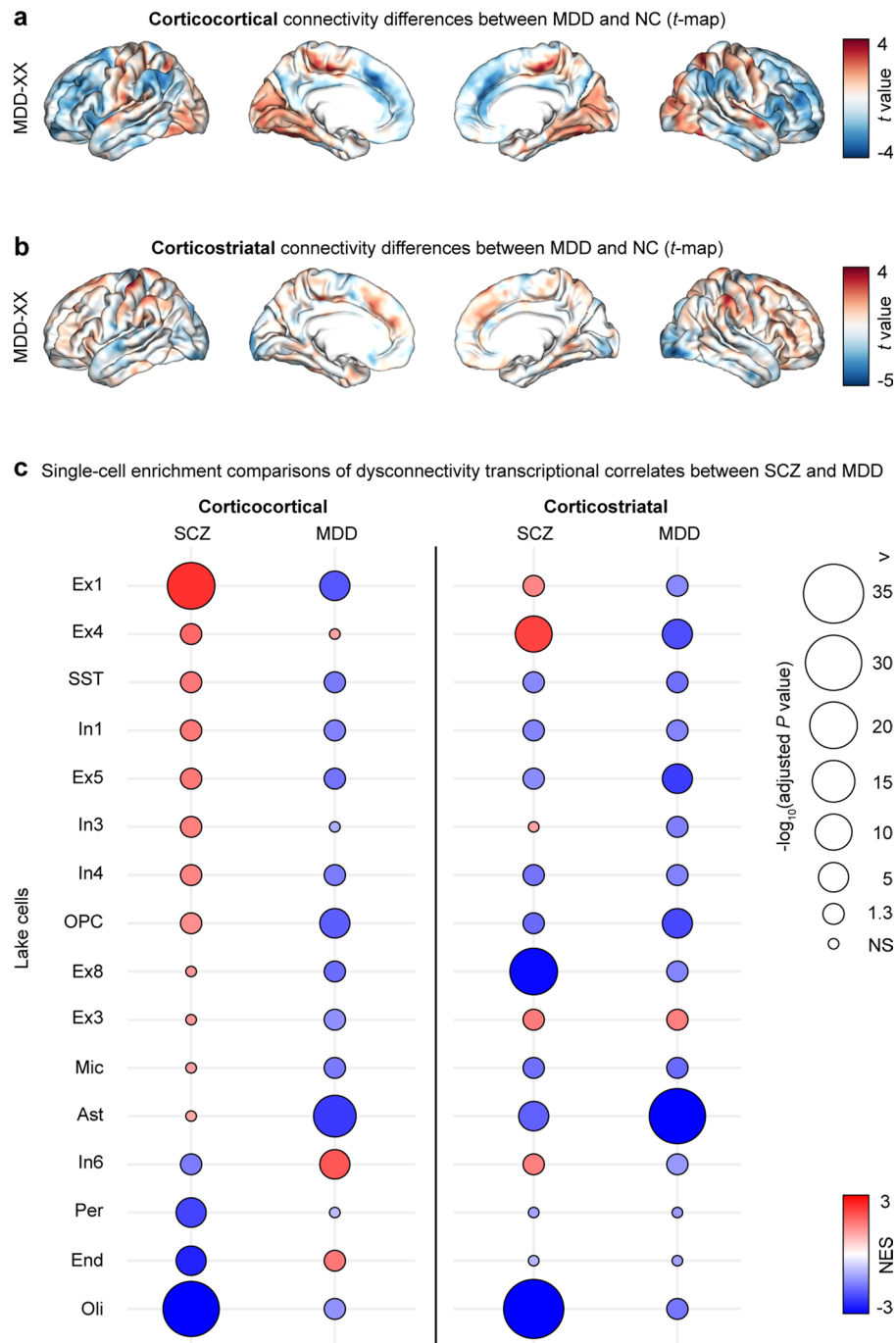
significance. The face/affective processing map (the label is marked in bold and surrounded by a red box) is significantly, positively, and consistently correlated with corticocortical dysconnectivity  $t$ -maps across three datasets (**a**). There were three cognitive maps (pain, cued attention, and inhibition error) that were significantly, negatively, and consistently correlated with corticostriatal dysconnectivity  $t$ -maps across three datasets (**b**).

**Extended Data Fig. 5 | Cell enrichment of individual dysconnectivity.**

**a, c, e, g.** The line graphs illustrate individual-level enrichment scores (NESs) of Lake and ABA cells for corticocortical and corticostriatal dysconnectivity. The participants were derived from the SCZ-I, SCZ-II, and COBRE datasets (total SCZ/

NC=585/614). The error bars represent mean values  $\pm$  95% confidence interval.

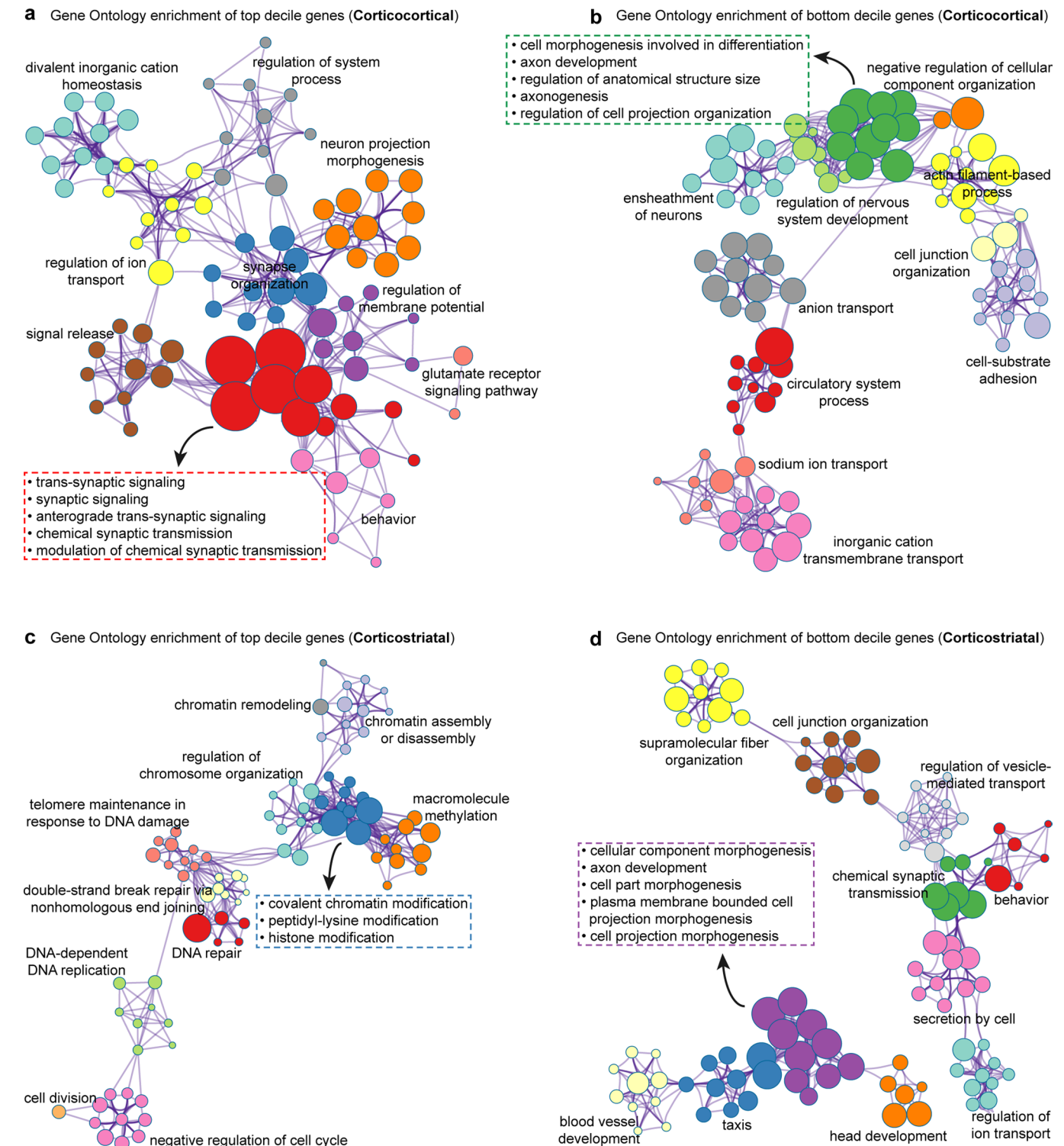
**b, d, f, h.** The heatmaps show pairwise comparisons of individual NES scores within the SCZ group (n = 585), using a two-sample t-test (two-sided). Blank squares indicate no significant differences between cells.



**Extended Data Fig. 6 | Cell enrichment comparisons of dysconnectivity transcriptional correlates between SCZ and MDD. a, b,** Cortical renderings show corticocortical (**a**) and corticostriatal (**b**) connectivity differences (t-maps) between patients with MDD and NC in the MDD-XX dataset (MDD,  $n = 75$ ; NC,  $n = 74$ ). The MDD dysconnectivity t-maps were obtained by applying the same analytical method as in Fig. 1. The warm or cool color indicates connectivity is increased or decreased in MDD. **c.** Cell enrichment of MDD dysconnectivity transcriptional correlates was conducted as with SCZ in Fig. 3. MDD corticocortical dysconnectivity transcriptional correlated genes

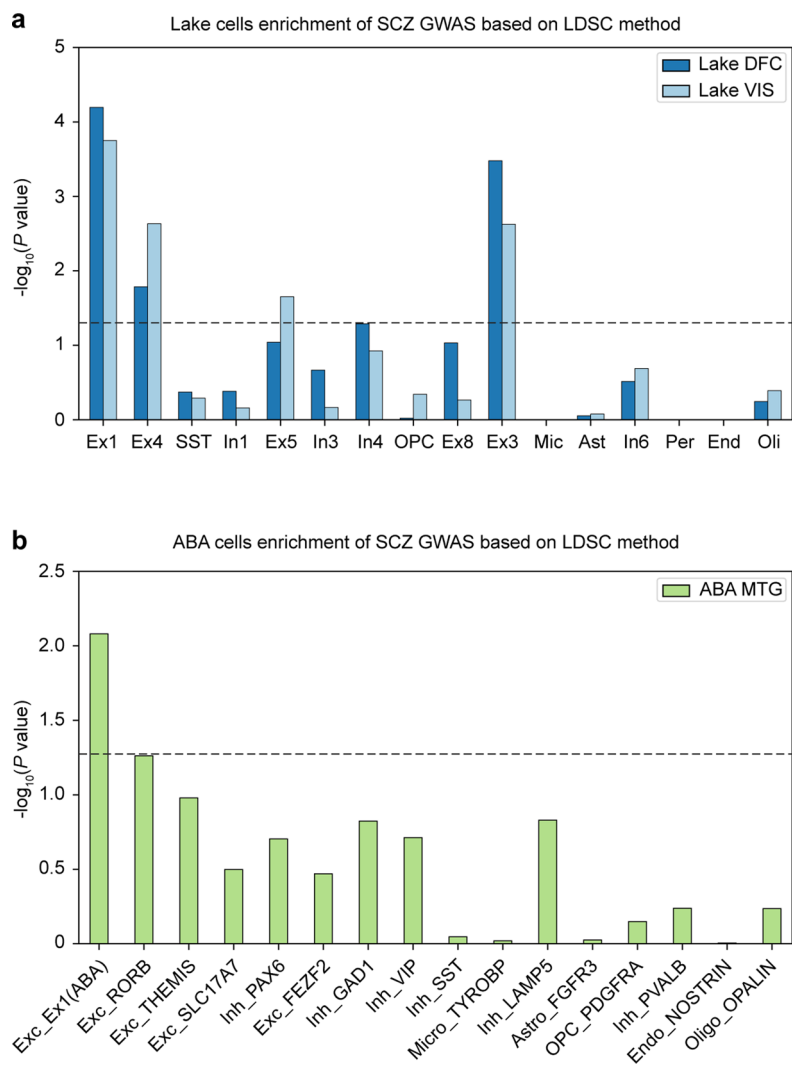
were most significantly positively enriched in In6 and negatively enriched in Ast. The positively and negatively enriched cells with the most significance for MDD corticostriatal dysconnectivity were Ex3 and Ast, respectively. The solid circles with warm or cool colors indicate positive or negative enrichment, determined by NES. Darker colors mean greater NES. The radius is quantified by  $-\log_{10}(\text{adjusted } P \text{ value})$  (two-sided, FDR correction). Larger circles mean more significance of enrichment. NS, non-significant. NES, normalized enrichment score.





**Extended Data Fig. 7 | Gene Ontology enrichment of dysconnectivity transcriptional correlates.** Metascape enrichment networks show inter-cluster and intra-cluster similarities of enriched biological process terms. Each term is characterized by a circle node, where its size represents the significance level of enrichment, and its color denotes cluster identity (nodes with the same color

belong to the same cluster). For each kind of dysconnectivity, a sorted gene list (in descending order) was obtained based on averaged spatial correlations between all AHBA genes and dysconnectivity *t*-maps. The gene list was then split into deciles. Enrichment analyses were separately conducted for the top and bottom gene deciles.



**Extended Data Fig. 8 | Cell enrichment of SCZ GWAS based on LDSC method.** Vertical bar plots show Lake (a) and ABA (b) cells enrichment of polygenic risk in SCZ from GWAS data. Cell-specific genes were defined using Lake data from the dorsal frontal cortex (DFC) and visual cortex (VIS), and ABA data from the middle temporal gyrus (MTG). The dashed line indicates unadjusted  $P < 0.05$ .

Corresponding author(s): Bing Liu

Last updated by author(s): Jul 17, 2023

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection Neuroimaging data: multiple MRI scanners. Genotype data: EZgene Blood gDNA Miniprep Kit. (see Methods)

Data analysis The preprocessing software for resting-state fMRI data is freely available (BRANT v3.35, <http://brant.brainnetome.org/en/latest/>). The SurfStat toolbox for surface-wide statistical comparisons is freely available at <https://www.math.mcgill.ca/keith/surfstat/>. Corticocortical connectivity (first functional gradient) was calculated based on open-source codes at [https://github.com/NeuroanatomyAndConnectivity/gradient\\_analysis](https://github.com/NeuroanatomyAndConnectivity/gradient_analysis). The toolbox for performing spatial permutation test (spin test) is freely available at <https://github.com/spin-test/spin-test>. Functional decodings were performed based on the neurosynth package openly available at <https://github.com/neurosynth/neurosynth>. The abagen pyGAM package (v0.8.0) for performing symptom prediction analysis is openly available at <https://github.com/dswah/pyGAM>. The abagen toolbox (v0.1.1) for AHBA data processing is freely available at <https://github.com/rmarkello/abagen>. The pipeline for single-cell data processing is consistent with that performed by Anderson et al. and the codes are openly available at [https://github.com/kevmanderson/2020\\_PNAS\\_Depression](https://github.com/kevmanderson/2020_PNAS_Depression). Fast gene set enrichment analysis was performed based on the fgsea package, which is openly available at <https://github.com/ctclab/fgsea>. The MAGMA (v1.08) software and reference data are publicly available at <https://ctg.cncr.nl/software/magma>. The PLINK (v1.07) software is freely available at <https://www.cog-genomics.org/plink/>. Gene ontology enrichment analysis was performed using the Metascape platform (<https://metascape.org/gp/index.html#/main/step1>). The LDSC package is available at <https://github.com/bulik/ldsc>. The IMPUTE2 can be downloaded at [https://mathgen.stats.ox.ac.uk/impute/impute\\_v2.html#home](https://mathgen.stats.ox.ac.uk/impute/impute_v2.html#home). The SHAPEIT (v2.r790) can be downloaded at [https://mathgen.stats.ox.ac.uk/genetics\\_software/shapeit/shapeit.html#home](https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html#home). All custom codes used in the analysis are publicly available at [https://github.com/BingLiu-Lab/scz\\_cross-scale\\_abnormalities](https://github.com/BingLiu-Lab/scz_cross-scale_abnormalities).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.



## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The COBRE dataset can be accessed from the COINS data exchange portal (<https://coins.trendscenter.org/>). The UKB data can be requested through a standard protocol (<https://www.ukbiobank.ac.uk/register-apply/>). The Brainnetome atlas can be downloaded at <https://atlas.brainnetome.org/download.html>. The Neurosynth database can be downloaded at <https://github.com/neurosynth/neurosynth-data>. Allen Human Brain Atlas is freely available at <https://human.brain-map.org/>. The Lake et al. single-cell data is publicly available at the National Center for Biotechnology Information under the SuperSeries accession code GSE97942. The ABA single-cell data is publicly available at <https://portal.brain-map.org/atlas-and-data/rnaseq>. The schizophrenia GWAS data is publicly available at <https://pgc.unc.edu/for-researchers/download-results/>. All study data supporting the findings are provided within the paper and its supplementary material. All raw data from the SCZ-I and SCZ-II datasets will be made available upon reasonable request to the corresponding author.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

### Reporting on sex and gender

The term "sex" was utilized to refer to biological attribution. The sex of participants was determined based on self-reporting. Our findings were not related to a specific sex. In our study, sex was regarded as a covariate when performing between-group comparisons. The sex proportions were matched between patients and healthy controls, which were provided in Supplementary Table S1. The SCZ-I dataset included 318 individuals with schizophrenia (male/female=166/152) and 307 healthy subjects (male/female=160/147). The SCZ-II dataset included 195 individuals with schizophrenia (male/female=96/99) and 220 healthy subjects (male/female=102/118). The COBRE dataset included 72 individuals with schizophrenia (male/female=59/13) and 87 healthy subjects (male/female=63/24). For UKB data, we did not include the sex information in the main text as we consider it was irrelevant to our genetic analysis.

### Population characteristics

All participants' information can be found in Supplementary Table S1 and Methods section. When performing between-group comparisons, age and sex are regarded as covariates. We did not provide detailed age and sex information for the UKB data.

### Recruitment

All participants enrolled in the SCZ-I and SCZ-II datasets were obtained from seven Chinese sites with the same recruitment protocol. All participants in the MDD-XX dataset were obtained at the Second Affiliated Hospital of Xinxiang Medical University in China. The third-party COBRE and UKB data have their own distinct recruitment parameters. See the Methods section for details.

### Ethics oversight

The research protocol was approved by the Medical Research Ethics Committee of participating hospitals, including Peking University Sixth Hospital, Beijing Huilongguan Hospital, Xijing Hospital, Henan Mental Hospital, Renmin Hospital of Wuhan University, Zhumadian Psychiatric Hospital, and the Second Affiliated Hospital of Xinxiang Medical University. The COBRE study was approved by the institutional review board (IRB) of the University of New Mexico (UNM). The UKB study was approved by the North West Multi-centre Research Ethics Committee (MREC, <https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us/ethics>).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

### Sample size

The sample size was based on the availability of data, no statistical methods were used to pre-determine sample size.

### Data exclusions

Subjects were excluded if they had incomplete data or failed quality control examinations. Detailed exclusion criteria were provided in the Methods section.

### Replication

The corticocortical and corticostriatal dysconnectivity patterns were highly consistent across three datasets. The individual symptom prediction findings were reproducible across three datasets. The imaging transcriptional analysis results were well replicated across two single-cell datasets. In population-based genetic analysis, the results were well replicated using the public GWAS dataset, our own GWAS

summary derived from SCZ-I and SCZ-II datasets, and GWAS summary generated from UKB data. In person-specific genetic analysis, the results were replicated based on the SCZ-I and SCZ-II datasets, as well as the UKB dataset. In addition, both population-based and person-specific genetic analyses demonstrated high consistency across the Lake and ABA single-cell datasets.

Randomization	Participants were allocated into patient and control groups based on the clinical diagnosis. In our neuroimaging analyses, we included age, gender, and site as covariates when performing between-group comparisons.
Blinding	Researchers were not blinded to patient/control status. Both dysconnectivity pattern analysis (from between-group comparisons) and case-control GWAS analysis necessitate knowledge of the participants' diagnostic status.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

## Magnetic resonance imaging

### Experimental design

Design type	Resting state
Design specifications	A single session was scanned for each participant.
Behavioral performance measures	This is not relevant to our study since we do not have task designs.

### Acquisition

Imaging type(s)	functional and structural
Field strength	3 Tesla
Sequence & imaging parameters	Details were provided in the Methods section.
Area of acquisition	Whole brain scan
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

### Preprocessing

Preprocessing software	BRANT version 3.35 was used for the preprocessing of resting-state fMRI, details of specific parameters were provided in the Methods section.
Normalization	Nonlinear normalization of functional MRI image to the template.
Normalization template	MNI152 template
Noise and artifact removal	We removed confounding effects by regressing out linear trends, averaged white matter (WM) and cerebrospinal fluid (CSF), the first derivatives of WM and CSF, and estimated head motion parameters. We then performed temporal band-pass filtering at 0.01 - 0.08 Hz to suppress low-frequency drifts and physiological noises.
Volume censoring	Artifacts, registration, and normalization quality underwent visual check.

### Statistical modeling & inference

Model type and settings	Surface-based linear models were used to obtain connectivity differences between patients with schizophrenia and healthy controls. The two-sided t-test method was applied to investigate cluster-level between-group connectivity differences.
-------------------------	---

Regression analyses were conducted using multivariate predictive models (linear GAMs). Imaging transcriptional analyses were performed on the dysconnectivity t-maps.

Effect(s) tested

The t-statistic value was used to measure the extent of group-level connectivity differences. Pearson's correlation and mean absolute error were calculated to quantify the performance of individual symptom prediction.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s) The striatum region was investigated which was determined based on the Brainnetome atlas.

Statistic type for inference  
(See [Eklund et al. 2016](#))

cortical vertex-wise

Correction

FWE, FDR, and permutation-based correction methods were applied according to the specific conditions. Detailed descriptions were provided in the Methods section.

## Models & analysis

n/a | Involved in the study

- ☐ ☒ Functional and/or effective connectivity  
☒ ☐ Graph analysis  
☐ ☒ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Pearson's correlation was applied to measure functional connectivity.

Multivariate modeling and predictive analysis

The predictive analysis was performed based on generalized additive models. Pearson's correlation and mean absolute error were calculated to quantify the prediction performance.