

# Diagnostic Classification for Human Autism and Obsessive-Compulsive Disorder Based on Machine Learning From a Primate Genetic Model

Yafeng Zhan, M.S., Jianze Wei, M.S., Jian Liang, Ph.D., Xiu Xu, M.D., Ran He, Ph.D., Trevor W. Robbins, Ph.D., Zheng Wang, Ph.D.

**Objective:** Psychiatric disorders commonly comprise comorbid symptoms, such as autism spectrum disorder (ASD), obsessive-compulsive disorder (OCD), and attention deficit hyperactivity disorder (ADHD), raising controversies over accurate diagnosis and overlap of their neural underpinnings. The authors used noninvasive neuroimaging in humans and nonhuman primates to identify neural markers associated with DSM-5 diagnoses and quantitative measures of symptom severity.

**Methods:** Resting-state functional connectivity data obtained from both wild-type and methyl-CpG binding protein 2 (*MECP2*) transgenic monkeys were used to construct monkey-derived classifiers for diagnostic classification in four human data sets (ASD: Autism Brain Imaging Data Exchange [ABIDE-I], N=1,112; ABIDE-II, N=1,114; ADHD-200 sample: N=776; OCD local institutional database: N=186). Stepwise linear regression models were applied to examine associations between functional connections of monkey-derived classifiers and dimensional symptom severity of psychiatric disorders.

**Results:** Nine core regions prominently distributed in frontal and temporal cortices were identified in monkeys and used as seeds to construct the monkey-derived classifier that informed diagnostic classification in human autism. This same set of core regions was useful for diagnostic classification in the OCD cohort but not the ADHD cohort. Models based on functional connections of the right ventrolateral prefrontal cortex with the left thalamus and right prefrontal polar cortex predicted communication scores of ASD patients and compulsivity scores of OCD patients, respectively.

**Conclusions:** The identified core regions may serve as a basis for building markers for ASD and OCD diagnoses, as well as measures of symptom severity. These findings may inform future development of machine-learning models for psychiatric disorders and may improve the accuracy and speed of clinical assessments.

*Am J Psychiatry* 2021; 178:65–76; doi: 10.1176/appi.ajp.2020.19101091

The pursuit of imaging-based biomarkers for autism spectrum disorder (ASD) has been challenged by a lack of biological accounts (1, 2) and divergent research results (3–5), owing largely to the constraints of ASD's substantial heterogeneity and comorbidity with other diseases, such as obsessive-compulsive disorder (OCD) and attention deficit hyperactivity disorder (ADHD) (6–8). The high frequency of comorbidity could be a result of shared genetic vulnerability and pathophysiology, secondary effects of growing up with autism, shared symptom domains and associated mechanisms, or overlapping diagnostic criteria (9, 10). Emerging evidence from recent trans-diagnostic neuroimaging studies reveals that ASD-ADHD and ASD-OCD may share common features of a dysfunctional brain network, such as regions of high hubness (7, 11). Given the daunting clinical heterogeneity of ASD, a subset of homogeneous samples with a single genetic etiology may help disentangle the complexity of ASD

by establishing a tangible, symptom-related pathway from gene to brain to behavior (12). As extensively demonstrated in both humans (13) and rodent models (14, 15), methyl-CpG binding protein 2 (*MECP2*) is one of few exceptional genes causing autistic features, including intellectual disability, motor dysfunction, anxiety, and social behavior deficits. Mutation of the X-linked *MECP2* gene is found in 90% of patients with Rett syndrome. Duplications of *MECP2*-containing genomic segments cause *MECP2* duplication syndrome, which shares core symptoms with ASD (13, 16). The most recent reports on nonhuman primates have further demonstrated that *MECP2* gain and loss of function in genetically engineered monkeys recapitulate typical phenotypes in autism, such as less active social contact, increased stereotypical behaviors, and elevated anxiety (17, 18). More efficiently, acute viral-based molecular manipulation in the primate amygdala has been successfully used to probe a

See related features: **Editorial** by Dr. van den Heuvel (p. 8), **CME course** (online)

causal role for multiple components of the circuit that may underlie early-life anxiety (19–21). Cumulative evidence from primate genetic models suggests that they have substantial potential to deepen our mechanistic understanding of psychiatric disorders and inform the development of effective therapeutics (12, 17, 18, 22).

Comparative analyses of brain connectomics in closely related species demonstrate a functional resemblance between macaque monkeys and humans under healthy conditions (23–25) and have laid the foundation for exploiting knowledge acquired from nonhuman primate disease models to identify viable diagnostic markers for human brain disorders. Nevertheless, there exists no translational road map that describes the brain-wide mapping of neural circuits between a primate genetic model and human psychiatric patients (26). Thus, we propose a novel cross-species machine-learning framework that leverages connectome-based features learned from a primate genetic model of autism and then builds a classifier for diagnostic utility in humans.

The conceptual design of this cross-species framework originated from considering what features can be learned from the monkey model that capture the underlying neural pathophysiology that is likely shared with ASD patients. Within the network graph setting (25), two basic elements of brain circuitry—nodes (brain regions) and interconnecting edges (connections between pairs of nodes)—are considered (25, 27). Our intuition was that characteristics of nodes are more likely to be homologous between the primate species, given the substantial variations of edges subserving species-specific behavioral and cognitive adaptations (22, 25, 28). Moreover, supposing that only a subset of brain regions (not the entire brain) are particularly relevant to the core neuropathology of autism (termed “core regions” here) (4, 29), we deduced that only these core regions are useful for significantly reducing the complexity of cross-species mapping of this framework (Figure 1A).

We therefore adopted the group lasso feature selection process (30, 31) to determine core regions by treating all edges connecting to one region as a group (i.e., rows in an adjacency matrix). Next, we performed a one-to-one mapping of the identified core regions in the human brain network and used them as seeds to construct classifiers for patient classification. By using whole-brain functional connectivity data sets obtained from wild-type and *MECP2* transgenic monkeys (17) and from four human data sets—the Autism Brain Imaging Data Exchange (ABIDE-I), N=1,112 (3); ABIDE-II, N=1,114 (32); the ADHD-200 sample, N=776 (33); and an OCD local institutional database (N=186) (34)—we aimed to test two specific hypotheses: whether the monkey-derived features (core regions) could inform diagnostic classification in the ASD cohort and whether these monkey-derived features could be generalized to differentiate patients from healthy comparison subjects in the OCD and ADHD cohorts. To further parse the biological accounts of the monkey-derived classifiers, stepwise linear regression models based on features (i.e., functional connections of the monkey-derived classifiers) were built to predict clinical measures of symptom severity in patients.

## METHODS

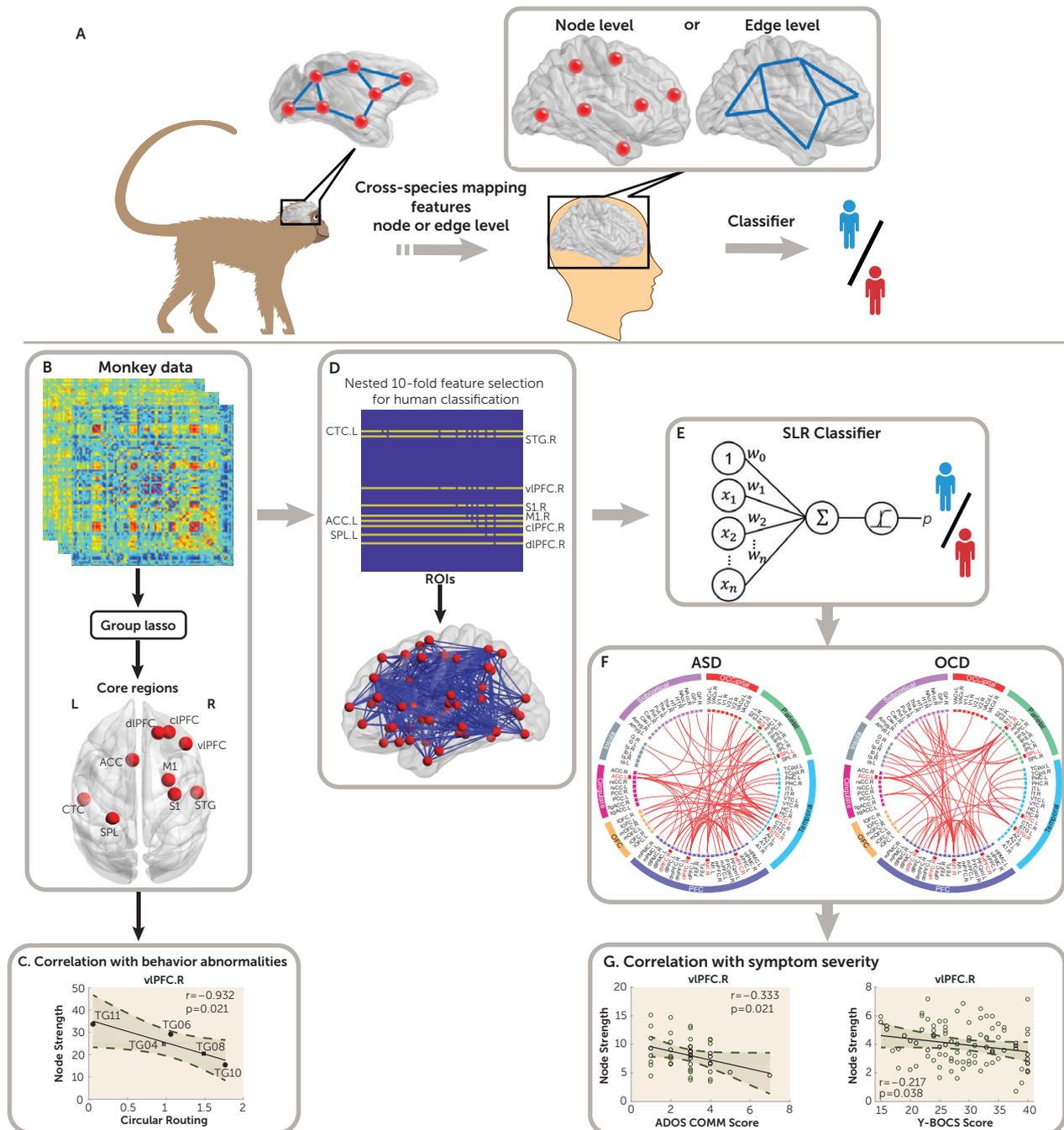
### Data Sets

In this study, we included one monkey data set and four independent human data sets. The monkey data set comprised five *MECP2*-duplication transgenic monkeys (*Macaca fascicularis*: mean age, 4.40 years [SD=0.29], mean weight, 3.26 kg [SD=0.75]; two males and three females) and 11 wild-type monkeys (*Macaca fascicularis*: mean age, 4.68 years [SD=0.46], mean weight, 3.97 kg [SD=1.36]; four males and seven females). Characteristics of the monkey data set are provided in Table S1 in the online supplement. All experimental procedures for animal subjects were approved by the Institutional Animal Care and Use Committee of the Institute of Neuroscience and the Biomedical Research Ethics Committee of the Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, and conformed to the National Institutes of Health guidelines for the humane care and use of laboratory animals. Details of animal preparation are provided in the online supplement.

We analyzed data from the two publicly available ABIDE repositories, ABIDE-I (3) and ABIDE-II (32), one ADHD cohort from the ADHD-200 sample (33), and one OCD cohort from our local institutional database. Similar to other ABIDE and ADHD studies, inclusion and exclusion criteria for participants included right-handedness, a full IQ score >80, known current medication status, and images accepted after quality control and meeting head motion criteria. Additional information on the characteristics of the study participants is provided in the online supplement. Data were available for a total of 336 study subjects from ABIDE-I (ASD group: N=133, male:female ratio, 118:15; healthy control group: N=203, male:female ratio, 167:36) (see also Table S2 in the online supplement), 149 study subjects from ABIDE-II (ASD group: N=60, male:female ratio, 56:4; healthy control group: N=89, male:female ratio, 62:27) (see also Table S3 in the online supplement), and 266 study subjects from the ADHD-200 cohort (ADHD group: N=96, male:female ratio, 79:17; healthy control group: N=170, male:female ratio, 97:73) (see also Table S4 in the online supplement). For the OCD data set, exclusion criteria were translation or rotation in any axis of head motion larger than 3 mm or 3° during scanning; any neurological disorders, psychosurgery, current or past substance abuse or dependence, or pregnancy; and any relevant physical illness, such as brain tumor or brain injury. The OCD cohort comprised 171 participants (OCD group: N=92, male:female ratio, 55:37; healthy control group: N=79, male:female ratio, 51:28) (see also Table S5 in the online supplement).

### Monkey MRI Data Acquisition and Preprocessing

A total of 144 functional MRI (fMRI) data sets from all monkeys were acquired from the Institute of Neuroscience, Chinese Academy of Sciences, on a 3-T whole-body scanner (Trio, Siemens Healthcare, Erlangen, Germany), running with an enhanced gradient coil insert (AC88; 80 mT/m maximum gradient strength, 800 mT/m per second

FIGURE 1. Schematic of a cross-species machine-learning framework<sup>a</sup>

<sup>a</sup> ACC=anterior cingulate cortex; ADOS COMM=Autism Diagnostic Observation Schedule communication subscale; ASD=autism spectrum disorder; cIPFC=centrolateral prefrontal cortex; CTC=central temporal cortex; dIPFC=dorsolateral prefrontal cortex; L=left; M1=primary motor cortex; OCD=obsessive-compulsive disorder; R=right; ROIs=regions of interest; S1=primary somatosensory cortex; SLR=sparse logistic regression; SPL=superior parietal cortex; STG=superior temporal cortex; TG=transgenic monkey; vIPFC=ventrolateral prefrontal cortex; Y-BOCS=Yale-Brown Obsessive Compulsive Scale. Panel A is a diagram of a cross-species circuit-mapping strategy, which was formulated on the basis of the hypothesis that characteristics of the brain circuitry between nonhuman and human primates are evolutionarily conserved at the node or edge level. Panel B illustrates use of a group lasso algorithm to identify core regions in monkeys. Panel C shows associations between characteristics of core regions and behavior abnormalities in transgenic monkeys. Panel D shows how feature engineering in the human data set was initiated by a one-to-one projection of core regions from monkeys to humans and then subjected to a nested 10-fold feature selection for use in patient classification. Panel E shows how optimal features were input into the SLR procedure for patient classification. Panel F shows spatial profiles of the identified functional connections in ASD and OCD classifiers. Panel G shows associations between characteristics of nodes and edges in ASD and OCD classifiers and symptom severity in patients.

maximum slew rate). Whole-brain resting-state fMRI data were collected using a gradient-echo echo-planar sequence, and high-resolution T<sub>1</sub>-weighted anatomical images were acquired using a magnetization prepared rapid gradient echo sequence. The acquisition and preprocessing of monkey MRI data are described elsewhere (35), as well as in the online supplement.

### Human MRI Data Acquisition and Preprocessing

Details of the public data are available online for ABIDE-I/II and the ADHD-200 sample ([http://fcon\\_1000.projects.nitrc.org/indi/abide](http://fcon_1000.projects.nitrc.org/indi/abide) and [http://fcon\\_1000.projects.nitrc.org/indi/adhd200](http://fcon_1000.projects.nitrc.org/indi/adhd200)). For the OCD data set, resting-state fMRI and high-resolution three-dimensional T<sub>1</sub>-weighted data sets were collected with a 12-channel head coil on a Siemens Tim Trio 3-T scanner. Details of the scan protocol are provided in Table S6 in the online supplement. The preprocessing of ABIDE-I was performed by the Preprocessed Connectomes Project (<http://preprocessed-connectomes-project.org/abide/index.html>) using the Data Processing Assistant for Resting-State fMRI toolbox (36). The same preprocessing pipeline was applied to the ABIDE-II, ADHD, and OCD data sets; technical details are summarized in the online supplement.

### Monkey and Human Brain Parcellation and Network Construction

The cortical organizations of both monkeys and humans were parcellated in accordance with the regional map template (37, 38). Because the regional map parcellation does not include subcortical regions, subcortical parcellation for the two species was added on the basis of the INIA19 primate brain atlas (39) and the FreeSurfer template (40), respectively. This generated a whole-brain template with a total of 94 regions of interest for both monkeys and humans. Anatomical regions are listed in Table S7 in the online supplement. Pearson's correlation coefficients between the mean time courses of any pair of regions were calculated to represent their functional connectivity, resulting in a 94×94 connectivity network matrix. Fisher's z transformation was then applied to the connectivity matrix, which was subject to a covariate-regression procedure. Both age (linear and quadratic) and sex were used as covariates for monkey and human data. When information on full IQ score, site, current medication status, and eyes opened/closed condition were available at the time of scanning, they were used as additional covariates for human data.

### Cross-Species Diagnostic Classification for Humans

We adopted a sparse linear regression model based on the group lasso penalty (30, 31), which is designed to eliminate a group of edges simultaneously. Here, we treated all edges connecting to one brain region as a group (i.e., rows in an adjacency matrix): if a specific brain region is irrelevant, then coefficients of all connections should be adjusted to zero. The penalty is deployed to impose a network structure and thereby regularize the problem (i.e., core regions versus

nonrelevant regions) (Figure 1B). It automatically and objectively identified nine core regions (Figure 1B) from the monkey data. Using these nine core regions as seeds, a monkey-derived classifier composed of the functional connections of the human functional connectome was constructed (Figure 1D). Specifically, using the standard lasso method with a 10×10 nested cross-validation (41), we determined a subset of relevant, nonredundant edge features of all 801 functional connections among nine core regions for use in classification. The lasso was applied with L1-norm penalty to achieve a sparse model by excluding the majority of features from the model (41). The sparsity of the model facilitated the optimization of predictors and reduced the model complexity (42). A sparse logistic regression (43) in tandem with leave-one-participant-out cross-validation was then implemented to distinguish patients with ASD from healthy control subjects (for further details, see Figure S1 in the online supplement).

To test the robustness of the monkey-derived classifier, we compared the predictive accuracies of the monkey-derived classifier with that of the randomization classifier and the human-derived classifier, using McNemar's test (44). Specifically, the randomization classifier was constructed on the basis of a randomly selected nine out of 94 regions, and the human-derived classifier was constructed on the basis of core regions identified from the ASD cohort (for further details, see the online supplement). Finally, the monkey-derived classifier was applied to the OCD and ADHD data sets.

### Statistical Analysis and Associations With Symptom Severity

A two-sample t test was applied to evaluate whether a difference in the node strength of nine core regions existed between transgenic and wild-type monkeys. The significance threshold was set at a p value of 0.05 (false-discovery-rate corrected). We quantitatively evaluated the distribution of functional connections for distinguishing patients from healthy comparison subjects in the ASD and OCD cohorts. We sorted functional connections on the basis of lobar location (i.e., prefrontal lobe, orbitofrontal lobe, temporal lobe, parietal lobe, occipital lobe, cingulate cortex, insula, and subcortical region). The spatial profiles of the functional connections were compared by intralobe and interlobe connections (i.e., whether node-to-node functional connections belong to the same lobe or different lobes). One-sided Fisher's exact test was applied to assess shared and disorder-specific functional connections between ASD and OCD. The significance threshold was set at a p value of 0.05 (one-tailed).

To investigate brain-behavior associations, we conducted Pearson's correlation analyses between the node strength of each core region and behavior abnormalities in transgenic monkeys (17) or symptom severity in patients with ASD and OCD ( $p < 0.05$ , subject to false-discovery-rate correction for multiple comparisons). For the core regions that showed significant correlations with symptom severity scores (but did not withstand false-discovery-rate correction), we

conducted a post hoc analysis for core region-based functional connections and dimensional symptom severity using a stepwise linear regression model, which modeled the relationship between the dependent variable (symptom severity score) and independent variables (functional connections to a specific region). For each model, the stepwise regression method iteratively determined a combination of functional connections that were linearly linked to dimensional symptoms. At each iteration, one connection was added or removed from the model for better fitting. The Pearson correlation coefficient and  $r^2$  between the predicted values and its measured values were used to assess the performance of the regression model. The significance threshold of associations of functional connections with dimensional symptoms was set at a  $p$  value of 0.05 (subject to false-discovery-rate correction).

## RESULTS

### Core Region Mapping Between Two Primate Species

The group lasso algorithm identified nine core regions out of 94 nodes of the monkey brain: the left central temporal cortex, right superior temporal cortex (STG), right dorsolateral prefrontal cortex (dlPFC), right primary somatosensory cortex, right primary motor cortex, left anterior cingulate cortex (ACC), right centrolateral prefrontal cortex, left superior parietal cortex, and right ventrolateral prefrontal cortex (vlPFC) (Figure 2A). Compared with the wild-type group, the transgenic group exhibited significantly increased node strength in the right vlPFC ( $p=0.004$ ), left ACC ( $p=0.008$ ), right centrolateral PFC ( $p=0.001$ ), and right dlPFC ( $p=0.004$ ) (Figure 2B). In contrast, the right primary somatosensory cortex ( $p=0.002$ ) and right primary motor cortex ( $p=0.025$ ) showed decreased node strength in the transgenic group (Figure 2B). In addition, we observed a significant negative relationship between the right vlPFC and circular routing ( $r=-0.932$ ,  $p=0.021$ ) and between the left superior parietal cortex and locomotion ( $r=-0.922$ ,  $p=0.026$ ) (see also Figure S2 in the online supplement). Conversely, node strength in the left ACC was positively associated with relative circular routing in the transgenic group ( $r=0.910$ ,  $p=0.032$ ) (see also Figure S2 in the online supplement).

### Highly Accurate Monkey-Derived Classifier for ASD

In the ABIDE-I cohort, the monkey-derived classifier achieved an accuracy of 82.14% (95% CI=77.53, 86.00; permutation test,  $p<0.001$  [see also Figure S3 in the online supplement]), a sensitivity of 79.70% (95% CI=71.66, 85.98), and a specificity of 83.74% (95% CI=77.78, 88.40), corresponding to an area under the receiver operating characteristic curve of 0.884 (Figure 2C; see also Table S8 in the online supplement). In the independent ABIDE-II cohort, the monkey-derived classifier also achieved performance with an accuracy of 75.17% (95% CI=67.30, 81.71; permutation test,  $p=0.014$  [see also Figure S3 in the online supplement]),

a sensitivity of 70.00% (95% CI=56.63, 80.80), and a specificity of 78.65% (95% CI=68.43, 86.35), and the area under the curve was 0.769 (Figure 2C; see also Table S8 in the online supplement).

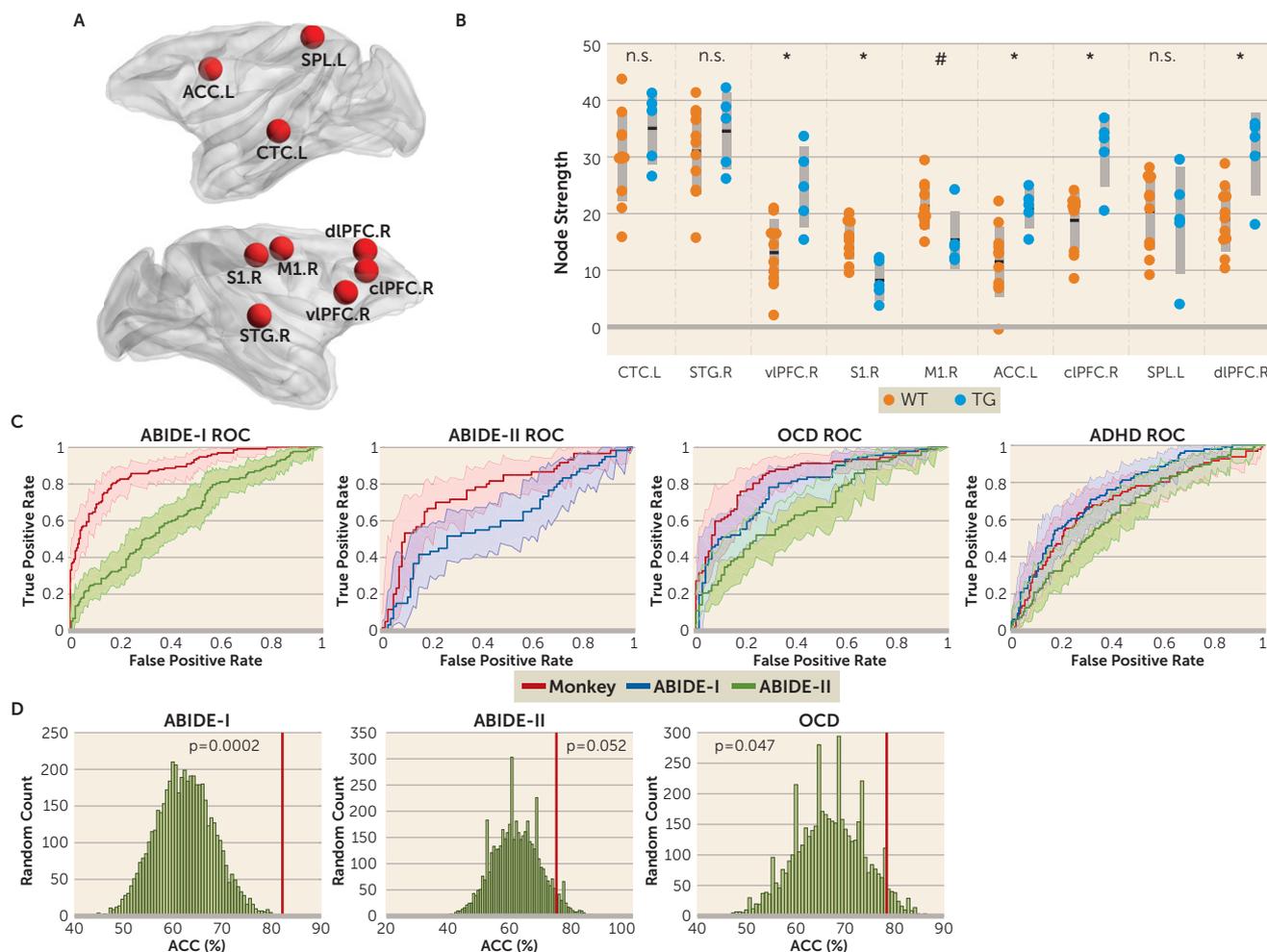
In the ABIDE-II data set, the same group lasso algorithm identified nine core regions: the left STG, left dorsal part of the anterior visual area, left secondary visual cortex, right primary motor cortex, left and right ACC, right centrolateral PFC, right vlPFC, and right globus pallidus (see also Figure S4 in the online supplement). The classifier based on these core regions achieved an accuracy of 61.31% (95% CI=55.85, 66.51), a sensitivity of 56.39% (95% CI=47.53, 64.88), a specificity of 64.53% (95% CI=57.49, 71.02), and an area under the curve of 0.644 in the ABIDE-I cohort (Figure 2C; see also Table S8 in the online supplement), which was significantly lower than that for the monkey-derived classifier ( $p<0.001$ ; see also Table S8 in the online supplement). Similarly, in the ABIDE-I data set, this group lasso algorithm identified four core regions: the left thalamus, right primary visual area, right secondary visual cortex, and right STG (see also Figure S4 in the online supplement). The classifier based on these four core regions achieved an accuracy of 60.40% (95% CI=52.04, 68.21), with a sensitivity of 53.33% (95% CI=40.10, 66.14), a specificity of 65.17% (95% CI=54.26, 74.76), and an area under the curve of 0.611 in the ABIDE-II cohort, which was significantly lower than that for the monkey-derived classifier ( $p=0.003$ ; Figure 2C; see also Table S8 in the online supplement).

Because only the identity information of nine core regions was translated from the monkey model to humans, we questioned whether this particular set of nine core regions would outperform an arbitrary choice (i.e., randomly selecting nine of the whole brain nodes as core regions). We observed that the performance of the monkey-derived classifier was significantly better than chance in the ABIDE-I cohort ( $p=0.0002$ ) and ABIDE-II cohort ( $p=0.052$ ) (Figure 2D). Note that some random selections showed higher accuracy than that for the monkey-derived classifier (Figure 2D). However, on closer examination, we found that even though a random selection of nine regions outperformed the monkey-derived classifier in one cohort, it always failed to achieve reasonable performance in other cohorts, indicating poor generalizability for the random selection of core regions (see also Figure S5 in the online supplement).

### Application of the Monkey-Derived Classifier to Other Human Disorders

Next, we tested the generalizability of the monkey-derived classifier in the OCD and ADHD data sets. In the OCD cohort, the classifier based on the same set of nine core regions from the monkey cohort achieved an accuracy of 78.36% (95% CI=71.29, 84.13; permutation test,  $p=0.002$  [see also Figure S3 in the online supplement]), a sensitivity of 73.91% (95% CI=63.53, 82.26), a specificity of 83.54% (95% CI=73.14, 90.61), and an area under the curve of 0.848 (Figure 2C),

**FIGURE 2. Performance of diagnostic classification for autism spectrum disorder (ASD) and obsessive-compulsive disorder (OCD) cohorts with monkey- and human-derived classifiers<sup>a</sup>**



<sup>a</sup> ACC=anterior cingulate cortex; ABIDE=Autism Brain Imaging Data Exchange; ADHD=attention deficit hyperactivity disorder; cIPFC=centrolateral prefrontal cortex; CTC=central temporal cortex; dIPFC=dorsolateral prefrontal cortex; L=left; M1=primary motor cortex; n.s.=not significant; OCD=obsessive-compulsive disorder; R=right; ROC=receiver operating characteristic; S1=primary somatosensory cortex; SPL=superior parietal cortex; STG=superior temporal cortex; TG=transgenic monkey; vIPFC=ventrolateral prefrontal cortex; WT=wild-type monkey. Panel A is a spatial illustration of the nine core regions identified from the monkey data sets. Panel B shows the node strength of the nine core regions in TG monkeys (N=5) compared with WT monkeys (N=11), using two-sample t tests. Compared with the WT group, the TG group exhibited significantly increased node strength in the right vIPFC, left ACC, right cIPFC, and right dIPFC but decreased node strength in the right S1 and right M1. A black line represents group-averaged node strength, and gray shading represents standard deviation. \* $p < 0.05$  (false-discovery-rate corrected); # $p < 0.05$  (uncorrected). Panel C shows the ROC curves of the monkey-derived classifier plotted with those of the human-derived classifier in the ABIDE-I, ABIDE-II, OCD, and ADHD data sets. Red, blue, and green lines indicate ROC curves of the monkey-derived classifier, ABIDE-I-derived classifier, and ABIDE-II-derived classifier, respectively; shaded areas indicate 95% confidence intervals. Panel D shows the accuracy of histograms of cross-validation using randomly generated core regions to construct the classifier in two ASD data sets and one OCD data set. Vertical red lines represent the classification accuracy of the monkey-derived classifiers, which were significantly higher than chance level at  $p = 0.0002$ ,  $p = 0.052$ , and  $p = 0.047$ .

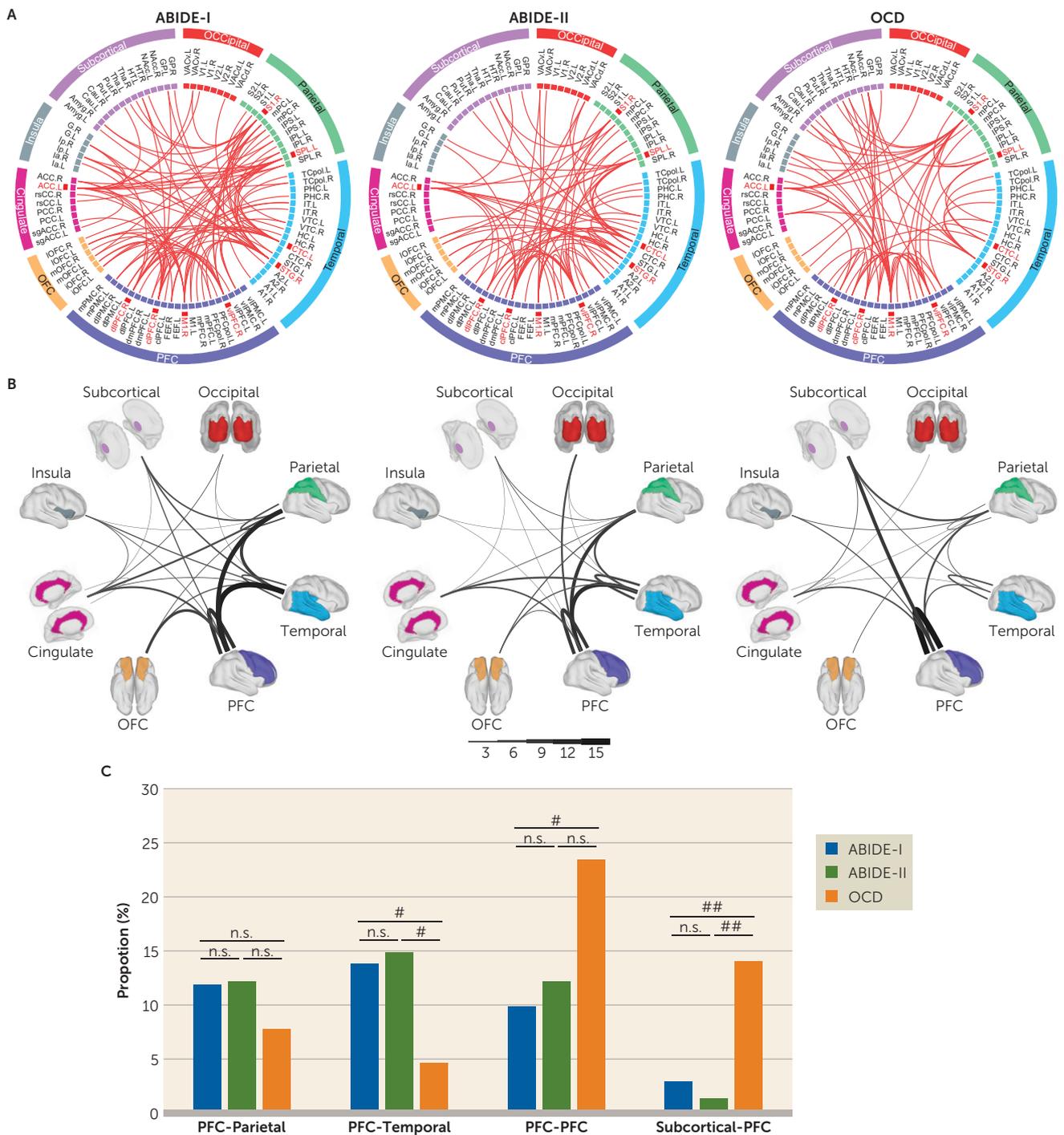
which outperformed two other human-derived classifiers (ABIDE-I,  $p = 0.044$ ; ABIDE-II,  $p < 0.001$ ; see also Table S8 in the online supplement), as well as the randomization classifiers ( $p = 0.047$ ) (Figure 2D). However, in the ADHD cohort, the monkey-derived classifier performed with an accuracy of 68.80% (95% CI=62.80, 74.24), a sensitivity of 56.25% (95% CI=45.76, 66.23), a specificity of 75.88% (95% CI=68.61, 81.96), and an area under the curve of 0.700, which yielded no significant difference compared with other human-derived classifiers (Figure 2C; see also Table S8 in the online supplement). Still, core regions identified in the monkey model showed better generalizability than randomly generated ones

in the ASD and OCD cohorts (see also Figure S5 in the online supplement).

### Characteristics of Nodes and Edges in the Monkey-Derived Classifier for ASD and OCD

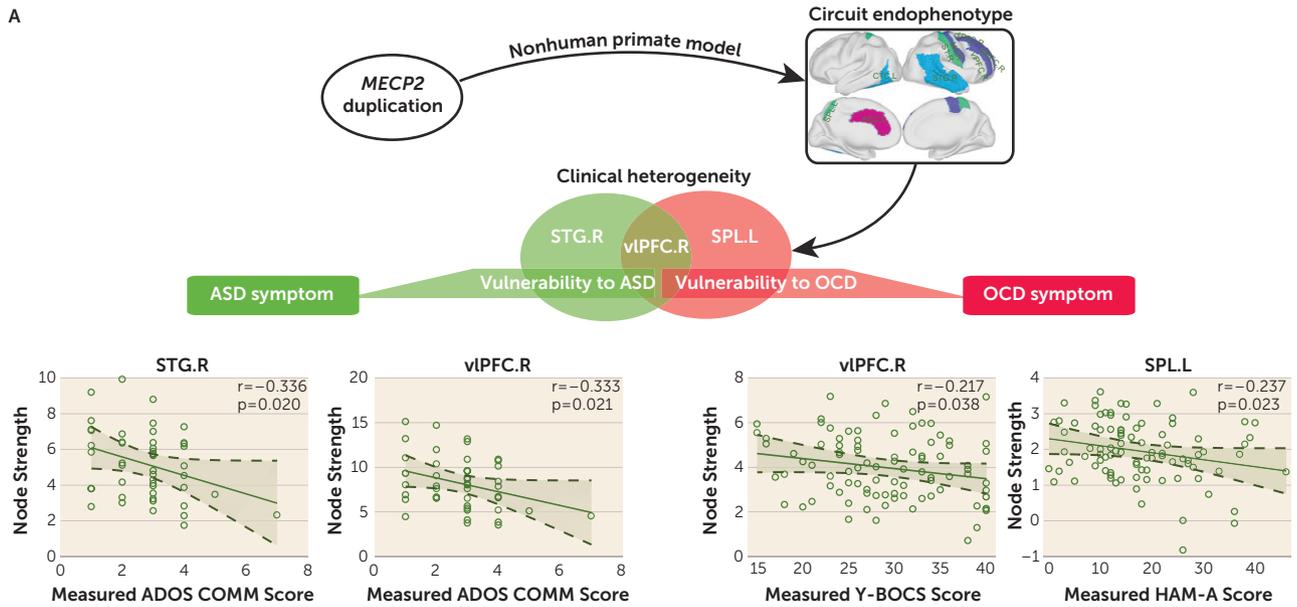
The sparse logistic regression algorithm automatically identified 101, 74, and 64 functional connections from the ABIDE-I, ABIDE-II, and OCD data sets, respectively, for reliable classification of patients and healthy control subjects (Figure 3A). We observed a biased distribution of intralobe functional connections within the prefrontal lobe in both the ASD and OCD cohorts (ABIDE-I:  $N = 10/101$ , ABIDE-II:  $N = 9/74$ ,

**FIGURE 3. Characteristics of the identified functional connections for autism spectrum disorder (ASD) and obsessive-compulsive disorder (OCD)<sup>a</sup>**

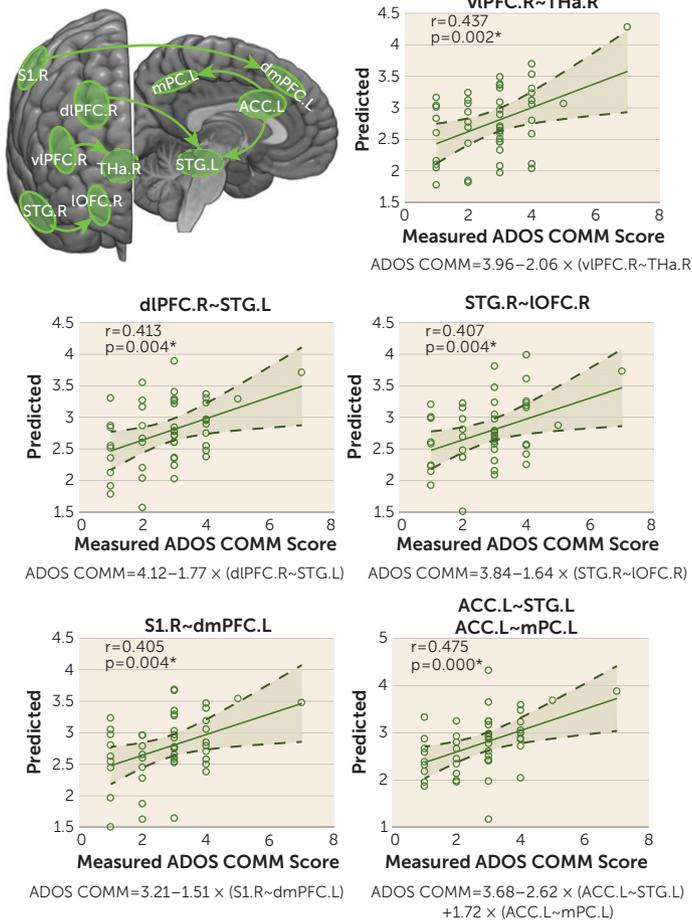


<sup>a</sup> ABIDE=Autism Brain Imaging Data Exchange; n.s.=not significant; OFC=orbitofrontal cortex; PFC=prefrontal cortex. In panel A, the top row shows the identified functional connections and their terminal regions (nodes) for reliable distinction of ASD patients from healthy control subjects and OCD patients from healthy control subjects. The inner layers of the brain nodes (circles) denote 94 brain regions, the interlayer highlights the nine core regions, and the outer layer specifies the lobar division. The abbreviations and parcellation of brain nodes are presented in Table S7 in the online supplement. Panel B shows the sum of the number of intralobe and interlobe connections based on the lobar division for ASD and OCD patients. Panel C shows the comparison of intralobe and interlobe connections with relatively higher functional connections between ASD and OCD patients. # $p < 0.05$ , uncorrected; ## $p < 0.01$ , uncorrected.

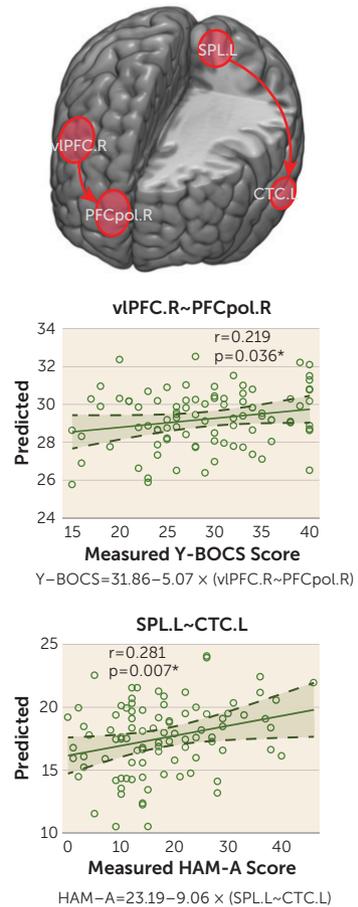
**FIGURE 4. Genetically linked neural circuits associated with distinct symptomatic domains in autism spectrum disorder (ASD) and obsessive-compulsive disorder (OCD)<sup>a</sup>**



**B. Circuits Predict ADOS COMM Score**



**C. Circuits Predict OCD Symptom Score**



<sup>a</sup> ACC=anterior cingulate cortex; ADOS COMM=Autism Diagnostic Observation Schedule communication subscale; ASD=autism spectrum disorder; CTC=central temporal cortex; dIPFC=dorsolateral prefrontal cortex; dmPFC=dorsomedial prefrontal cortex; HAM-A=Hamilton Anxiety Rating scale; L=left; IOFC=lateral orbitofrontal cortex; mPC=medial parietal cortex; OCD=obsessive-compulsive disorder; PFCpol=prefrontal polar cortex; R=right; S1=primary somatosensory cortex; SPL=superior parietal cortex; STG=superior temporal cortex; Tha=thalamus; vIPFC=ventrolateral prefrontal

OCD: N=15/64) (Figure 3B), which exhibited a significant proportional difference between the ABIDE-I and OCD cohorts ( $p=0.017$ ) but not between the ABIDE-II and OCD cohorts ( $p=0.065$ ) (Figure 3C). Meanwhile, interlobe functional connections mainly connecting prefrontal lobes with temporal lobes in patients with ASD compared with OCD patients (ABIDE-I: N=14/101, ABIDE-II: N=11/74, OCD: N=3/64) (Figure 3B) exhibited a significant proportional difference between groups (ABIDE-I compared with OCD,  $p=0.048$ ; ABIDE-II compared with OCD,  $p=0.043$ ) (Figure 3C). In contrast, OCD appeared to have a biased distribution of interlobe functional connections between prefrontal and subcortical areas (ABIDE-I: N=3/101, ABIDE-II: N=1/74, OCD: N=9/64; ABIDE-I compared with OCD,  $p=0.010$ ; ABIDE-II compared with OCD,  $p=0.004$ ) (Figure 3C).

The node strength of the right vIPFC was associated with communication subscores on the Autism Diagnostic Observation Schedule (ADOS) ( $r=-0.333$ ,  $p=0.021$ ), administered by research-reliable personnel, in 48 of 60 ASD patients available in the ABIDE-II cohort and was also associated with Yale-Brown Obsessive Compulsive Scale (Y-BOCS) scores ( $r=-0.217$ ,  $p=0.038$ ) for OCD patients (Figure 4A). In addition, the node strength of the left central temporal cortex, right STG, left ACC, right dlPFC, and right primary somatosensory cortex exhibited significant correlations with the ADOS communication subscore for ASD patients (Figure 4A; see also Table S9 in the online supplement). By contrast, we found a significant association between the Hamilton Anxiety Rating Scale (HAM-A) scores for OCD patients and the node strength of the left superior parietal cortex ( $r=-0.237$ ,  $p=0.023$ ) (Figure 4A). Note that associations between the node strength of nine core regions and symptom severity in patients with ASD and OCD did not survive correction for multiple comparisons.

We next determined whether functional connections of the classifiers were commonly shared or uniquely linked to symptomatic domains of ASD and OCD. Of all the possible right vIPFC-linked connections, the linear model based on the connection of the right vIPFC and the right thalamus achieved an  $r$  correlation of 0.437 ( $r^2=0.191$ ,  $p=0.002$ ) between the predicted and measured ADOS communication subscores for patients with ASD in the ABIDE-II group

(Figure 4B; see also Table S10 in the online supplement). The linear model equation is as follows:

$$\text{ADOS communication subscore} = 3.96 - 2.06 \times (\text{right vIPFC} \sim \text{right thalamus})$$

In fact, the ADOS communication subscore was significantly predicted by connections between the right dlPFC and left STG, the right STG and right orbitolateral prefrontal cortex, and the right primary somatosensory cortex and left dorsomedial prefrontal cortex, as well as between the left ACC and both the left medial parietal cortex and left STG (Figure 4B). By contrast, functional connection between the right vIPFC and right prefrontal polar cortex was predictive of the measured Y-BOCS scores for OCD patients ( $r=0.219$ ,  $r^2=0.048$ ,  $p=0.036$ ; Figure 4C; see also Table S10 in the online supplement). The model equation is as follows:

$$\text{Y-BOC} = 31.86 - 5.07 \times (\text{right vIPFC} \sim \text{right prefrontal polar cortex})$$

Of all left superior parietal cortex-linked connections, the predicted HAM-A scores based on the functional connection between the left superior parietal cortex and left central temporal cortex were significantly correlated with the actual HAM-A scores of these OCD patients ( $r=0.281$ ,  $r^2=0.079$ ,  $p=0.007$ ; Figure 4C; see also Table S10 in the online supplement).

## DISCUSSION

We demonstrated that a set of nine core regions, although identified in a relatively small number of biologically homogeneous transgenic monkeys, is sufficiently powerful to set the stage for cross-species mapping and prioritization of features for reducing clinical heterogeneity. This is probably because predisposition of a single genetic manipulation in animal models can be highly susceptible to specific neural circuits and pathways, revealing the extent of a tangible circuit endophenotype and its manifestation in the behavioral domain (45). These findings speak to the great heuristic value of identifying and investigating biologically homogeneous samples in future clinical studies.

---

cortex; Y-BOCS=Yale-Brown Obsessive Compulsive Scale. Panel A illustrates how the effects of a single genetic mutation in animal models may be limited to specific neural circuits (nine core regions) and associated behavioral manifestations in ASD and OCD. Node strengths of the right vIPFC and right STG exhibited significant correlations with ADOS COMM subscores of ASD patients; the right vIPFC was associated with Y-BOCS scores for OCD patients, while the left SPL was correlated with their HAM-A scores. Panel B illustrates the predictions of ADOS COMM subscores based on linear models, shown at the bottom of each plot for core regions, including the right vIPFC, right dlPFC, right STG, right S1, and left ACC. The connection between the right vIPFC and right thalamus achieved a correlation of  $r=0.437$  ( $p=0.002$ ) between predicted and measured ADOS COMM subscores for ABIDE-II patients. Similarly, the measured ADOS COMM subscore was predicted by connections with the right dlPFC and left STG ( $r=0.413$ ,  $p=0.004$ ), the right STG and right lateral orbitofrontal cortex ( $r=0.407$ ,  $p=0.004$ ), and the right S1 and left dorsomedial prefrontal cortex ( $r=0.405$ ,  $p=0.004$ ), and the connections between the left ACC and both the left medial parietal cortex and left STG ( $r=0.475$ ,  $p<0.001$ ), respectively. Panel C shows the predictions of Y-BOCS and HAM-A scores for OCD patients based on a linear model for the right vIPFC and left SPL (bottom of plot), respectively. The connection between the right vIPFC and right prefrontal polar cortex was sufficiently strong to predict Y-BOCS scores for OCD patients ( $r=0.219$ ,  $p=0.036$ ). HAM-A scores for these OCD patients could be predicted by the functional connection between the left SPL and left CTC ( $r=0.281$ ,  $p=0.007$ ). Shaded areas represent 95% confidence intervals. The abbreviations and parcellation of brain nodes are presented in Table S7 in the online supplement. \* $p<0.05$  (false-discovery-rate corrected).

This machine-learning framework enables the search for some straightforward explanations regarding how a neural circuit informs diagnostic classification for different human cohorts, thus allowing us to gain illuminating insights into the mechanisms by which these features were linked to symptom domains of ASD and OCD. This is illustrated by the results showing that the core regions, mostly distributed in frontal and temporal lobes (including the right vIPFC and right STG), were correlated with social communicative deficits (4, 29, 46). The right vIPFC and left ACC play an important role in the stop circuit (47, 48), which is heavily implicated in relation to restricted, repetitive behaviors and interests in ASD (49). Moreover, the functional connection between the left superior parietal cortex and left central temporal cortex predicted the severity of anxiety, reiterating the relationship between anxiety and attention in OCD (50). Additionally, our findings highlight the crucial role of two core regions—the primary somatosensory cortex and primary motor cortex—indicating their attributes to motor impairment in ASD (51) and impaired social-communicative skill development (51, 52), since deficits in the primary somatosensory cortex may affect the ability to master motor skills (53). Because sensory and motor disturbances have been hypothesized as some of the earliest signs of abnormality in children with ASD (52), studies characterizing the trajectory of sensorimotor disturbance and the downstream effects on social and communicative interaction in later development would be invaluable to addressing its biological basis and developing early therapeutics in autistic children. Intriguingly, the set of nine core regions identified in transgenic monkeys must be treated as one indiscernible group to achieve desirable performance compared with randomly generated ones. Although structural abnormalities of some of these nine regions, including gray matter volume and white matter tracts, have been reported in part (7, 54, 55), our observation suggests an integrated functional role of these nine regions, which calls for further investigation for the underlying neurobiological substrates.

From a circuit perspective, these core regions may simultaneously function as critical hubs for forming differential patterns of functional coupling with the rest of the brain, and which underlie distinct clinical symptom profiles in a variety of diseases. ASD-specific features exhibit apparently marked distribution of fronto-temporal and fronto-parietal connections, prominently involved in cognitive control and social communication (56). By contrast, dysconnectivity in OCD shows a significant bias toward the fronto-subcortical pathways, largely involved in response inhibition and cognitive flexibility (4, 47). Additionally, dissection of two distinct vIPFC-centered circuits from the monkey-derived classifier has revealed dual neuropathology-dependent roles for the vIPFC (the social communication of ASD and the compulsivity of OCD), thus emphasizing both unitary and diverse features contributing to the overlap of pathophysiological circuits between these disorders (8, 47).

This is, to our knowledge, the first study to propose this cross-species machine-learning method for defining circuit endophenotypes across multiple psychiatric disorders, and therefore caution is warranted. Replication to validate the present predictive algorithm with data from other populations and settings will be critical to extend the current use of genetically engineered animal models to broader human application scenarios, which also represents an important criterion for validating a primate genetic model. Considering the complex genetic architecture of ASD and related disorders, further work is essential to fully assess the predictive validity of nonhuman primate models based on single gene manipulation.

## CONCLUSIONS

In this possibly unique study, we explored the homologous nature of primate brain connectomes, showing that specific brain regions or circuits derived from genetically edited animal models may serve as a basis for defining a biologically defined circuit endophenotype shared between ASD and OCD. Such a circuit endophenotype associates genetics with distinct symptomatic domains across multiple current categorical diagnoses, which may represent an alternative route to deconstructing inherent heterogeneity and complex comorbidity. In addition, this translational framework can serve as an innovative general road map for a wide range of genetic animal models that bridge gaps between diagnostic categories of complex brain diseases and genetic and circuit mechanisms.

## AUTHOR AND ARTICLE INFORMATION

Institute of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, Chinese Academy of Sciences, Shanghai (Zhan, Wang); University of Chinese Academy of Sciences, Beijing (Zhan, Wang); School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing (Wei); Institute of Automation, Center for Excellence in Brain Science and Intelligence Technology, National Laboratory of Pattern Recognition, Chinese Academy of Sciences, Beijing (Wei, Liang, He); Department of Child Health Care, Children's Hospital of Fudan University, Shanghai (Xu); Department of Psychology, Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, U.K. (Robbins); Institute of Science and Technology for Brain Inspired Intelligence, Fudan University, Shanghai (Robbins); Shanghai Center for Brain Science and Brain-Inspired Intelligence Technology, Shanghai (Wang); and Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China (Wang).

Send correspondence to Dr. Wang (zheng.wang@ion.ac.cn) and Dr. He (ran.he@ia.ac.cn).

Presented at the 27th Annual Meeting of the International Society of Magnetic Resonance in Medicine, Montreal, May 11–16, 2019.

Mr. Zhan and Mr. Wei share first authorship.

Supported by the National Key R&D Program of China (grants 2017YFC1310400 and 2018YFC1313803 to Dr. Wang), the Strategic Priority Research Program of the Chinese Academy of Sciences (grant XDB32030000 to Dr. Wang), the Shanghai Municipal Science and Technology Major Project (grant 2018SHZDZX05 to Dr. Wang), and the National Natural Science Foundation (grants 81571300, 81527901, and

31771174 to Dr. Wang), as well as the Key Realm R&D Program of Guangdong Province (grant 2019B030335001 to Dr. Wang) and Wellcome Trust (grant 104631/Z/12/Z to Dr. Robbins).

The authors thank Hu Zhang, Zhiwei Wang, Jinqiang Peng, and Wenwen Yu for assistance with monkey data acquisition, Zhiwei Wang and Xiaoyu Chen for assistance with data preprocessing, and Drs. John Gore, Ravi Menon, Valerie Voon, and Tieniu Tan for editorial assistance.

Mr. Zhan and Dr. Wang are co-inventors on submitted patents for improvement in MRI-based classification of psychiatric disorders by use of a monkey-human interspecies machine-learning algorithm. Dr. Robbins has served as a consultant for Cambridge Cognition, Lundbeck, Mundipharma, and Unilever; he receives royalties from Cambridge Cognition and editorial honoraria from Elsevier and Springer-Verlag. The other authors report no financial relationships with commercial interests.

Received October 24, 2019; revisions received February 21 and March 21, 2020; accepted March 30, 2020; published online June 16, 2020.

## REFERENCES

1. Abi-Dargham A, Horga G: The search for imaging biomarkers in psychiatric disorders. *Nat Med* 2016; 22:1248–1255
2. Sahin M, Sur M: Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. *Science* 2015; 350:350
3. Di Martino A, Yan CG, Li Q, et al: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry* 2014; 19:659–667
4. Ecker C, Bookheimer SY, Murphy DG: Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. *Lancet Neurol* 2015; 14:1121–1134
5. Hahamy A, Behrmann M, Malach R: The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nat Neurosci* 2015; 18:302–309
6. Simonoff E, Pickles A, Charman T, et al: Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *J Am Acad Child Adolesc Psychiatry* 2008; 47:921–929
7. Carlisi CO, Norman LJ, Lukito SS, et al: Comparative multimodal meta-analysis of structural and functional brain abnormalities in autism spectrum disorder and obsessive-compulsive disorder. *Biol Psychiatry* 2017; 82:83–102
8. Robbins TW, Gillan CM, Smith DG, et al: Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci* 2012; 16:81–91
9. Cross-Disorder Group of the Psychiatric Genomics Consortium: Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* 2019; 179:1469–1482 e1411.
10. Lai M-C, Lombardo MV, Baron-Cohen S: Autism. *Lancet* 2014; 383:896–910
11. Lake EMR, Finn ES, Noble SM, et al: The functional brain organization of an individual allows prediction of measures of social abilities transdiagnostically in autism and attention-deficit/hyperactivity disorder. *Biol Psychiatry* 2019; 86:315–326
12. Izpissua Belmonte JC, Callaway EM, Caddick SJ, et al: Brains, genes, and primates. *Neuron* 2015; 86:617–631
13. Ramocki MB, Tavyev YJ, Peters SU: The MECP2 duplication syndrome. *Am J Med Genet A* 2010; 152A:1079–1088
14. Guy J, Hendrich B, Holmes M, et al: A mouse MeCP2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet* 2001; 27:322–326
15. Samaco RC, Mandel-Brehm C, McGraw CM, et al: Crh and Oprm1 mediate anxiety-related behavior and social approach in a mouse model of MECP2 duplication syndrome. *Nat Genet* 2012; 44:206–211
16. Ramocki MB, Peters SU, Tavyev YJ, et al: Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. *Ann Neurol* 2009; 66:771–782
17. Liu Z, Li X, Zhang JT, et al: Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. *Nature* 2016; 530:98–102
18. Chen Y, Yu J, Niu Y, et al: Modeling Rett syndrome using TALEN-edited MECP2 mutant cynomolgus monkeys. *Cell* 2017; 169:945–955.e10, e910
19. Fox AS, Oler JA, Birn RM, et al: Functional connectivity within the primate extended amygdala is heritable and associated with early-life anxious temperament. *J Neurosci* 2018; 38:7611–7621
20. Kalin NH, Fox AS, Kovner R, et al: Overexpressing corticotropin-releasing factor in the primate amygdala increases anxious temperament and alters its neural circuit. *Biol Psychiatry* 2016; 80:345–355
21. Kalin NH: Mechanisms underlying the early risk to develop anxiety and depression: a translational approach. *Eur Neuro-psychopharmacol* 2017; 27:543–553
22. Fox AS, Kalin NH: A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology. *Am J Psychiatry* 2014; 171:1162–1173
23. Miranda-Dominguez O, Mills BD, Grayson D, et al: Bridging the gap between the human and macaque connectome: a quantitative comparison of global interspecies structure-function relationships and network topology. *J Neurosci* 2014; 34:5552–5563
24. Vincent JL, Patel GH, Fox MD, et al: Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 2007; 447:83–86
25. van den Heuvel MP, Bullmore ET, Sporns O: Comparative connectomics. *Trends Cogn Sci* 2016; 20:345–361
26. Pryluk R, Kfir Y, Gelbard-Sagiv H, et al: A tradeoff in the neural code across regions and species. *Cell* 2019; 176:597–609.e18, e518
27. Pu J, Wang J, Yu W, et al: Discriminative structured feature engineering for macroscale brain connectomes. *IEEE Trans Med Imaging* 2015; 34:2333–2342
28. Mars RB, Sotiropoulos SN, Passingham RE, et al: Whole brain comparative anatomy using connectivity blueprints. *eLife* 2018; 7:7
29. Amaral DG, Schumann CM, Nordahl CW: Neuroanatomy of autism. *Trends Neurosci* 2008; 31:137–145
30. Yuan M, Lin Y: Model selection and estimation in regression with grouped variables. *J R Stat Soc B* 2006; 68:49–67
31. Wang J, Wonka P, Ye JP: Lasso screening rules via dual polytope projection. *J Mach Learn Res* 2015; 16:1063–1101
32. Di Martino A, O'Connor D, Chen B, et al: Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci Data* 2017; 4:170010
33. Consortium HD; HD-200 Consortium: The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Front Syst Neurosci* 2012; 6:62
34. Yin D, Zhang C, Lv Q, et al: Dissociable frontostriatal connectivity: mechanism and predictor of the clinical efficacy of capsulotomy in obsessive-compulsive disorder. *Biol Psychiatry* 2018; 84:926–936
35. Lv Q, Yang L, Li G, et al: Large-scale persistent network reconfiguration induced by ketamine in anesthetized monkeys: relevance to mood disorders. *Biol Psychiatry* 2016; 79:765–775
36. Chao-Gan Y, Yu-Feng Z: DPARSF: a MATLAB toolbox for “pipeline” data analysis of resting-state fMRI. *Front Syst Neurosci* 2010; 4:13
37. Kötter R, Wanke E: Mapping brains without coordinates. *Philos Trans R Soc Lond B Biol Sci* 2005; 360:751–766
38. Bezgin G, Vakorin VA, van Opstal AJ, et al: Hundreds of brain maps in one atlas: registering coordinate-independent primate neuroanatomical data to a standard brain. *Neuroimage* 2012; 62:67–76
39. Rohlfing T, Kroenke CD, Sullivan EV, et al: The INIA19 template and NeuroMaps atlas for primate brain image parcellation and spatial normalization. *Front Neuroinform* 2012; 6:27
40. Fischl B, Salat DH, Busa E, et al: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002; 33:341–355
41. Tibshirani R: Regression shrinkage and selection via the Lasso. *J Roy Stat Soc B Met* 1996; 58:267–288

42. Wright J, Ma Y, Mairal J, et al: Sparse representation for computer vision and pattern recognition. *Proc IEEE* 2010; 98: 1031–1044
43. Yamashita O, Sato MA, Yoshioka T, et al: Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage* 2008; 42:1414–1429
44. McNemar Q: Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947; 12: 153–157
45. Krol A, Wimmer RD, Halassa MM, et al: Thalamic reticular dysfunction as a circuit endophenotype in neurodevelopmental disorders. *Neuron* 2018; 98:282–295
46. Eyler LT, Pierce K, Courchesne E: A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain* 2012; 135:949–960
47. Aron AR, Robbins TW, Poldrack RA: Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn Sci* 2014; 18: 177–185
48. Dalley JW, Everitt BJ, Robbins TW: Impulsivity, compulsivity, and top-down cognitive control. *Neuron* 2011; 69:680–694
49. Thakkar KN, Polli FE, Joseph RM, et al: Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD). *Brain* 2008; 131:2464–2478
50. Pacheco-Unguetti AP, Acosta A, Callejas A, et al: Attention and anxiety: different attentional functioning under state and trait anxiety. *Psychol Sci* 2010; 21:298–304
51. Thompson A, Murphy D, Dell'Acqua F, et al: Impaired communication between the motor and somatosensory homunculus is associated with poor manual dexterity in autism spectrum disorder. *Biol Psychiatry* 2017; 81:211–219
52. Robertson CE, Baron-Cohen S: Sensory perception in autism. *Nat Rev Neurosci* 2017; 18:671–684
53. Nebel MB, Eloyan A, Nettles CA, et al: Intrinsic visual-motor synchrony correlates with social Deficits in Autism. *Biol Psychiatry* 2016; 79:633–641
54. Ameis SH, Lerch JP, Taylor MJ, et al: A diffusion tensor imaging study in children with ADHD, autism spectrum disorder, OCD, and matched controls: distinct and non-distinct white matter disruption and dimensional brain-behavior relationships. *Am J Psychiatry* 2016; 173:1213–1222
55. Wolff JJ, Gu H, Gerig G, et al: Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *Am J Psychiatry* 2012; 169:589–600
56. Cheng W, Rolls ET, Gu H, et al: Autism: reduced connectivity between cortical areas involved in face expression, theory of mind, and the sense of self. *Brain* 2015; 138:1382–1393