bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Title: Machine learning approach to identify resting-state functional connectivity pattern serving as an endophenotype of autism spectrum disorder

Authors

Bun Yamagata[$]

Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan

Postal address: 35 Shinanomachi, 160-8582, Tokyo, Japan

Takashi Itahashi[$]

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Junya Fujino

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Haruhisa Ohta

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Motoaki Nakamura

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Nobumasa Kato

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Masaru Mimura

Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan

Postal address: 35 Shinanomachi, 160-8582, Tokyo, Japan

Ryu-ichiro Hashimoto

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Department of Language Sciences, Graduate School of Humanities, Tokyo Metropolitan University, Tokyo, Japan

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Yuta Aoki*

Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

Postal address: 6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

$ These authors equally contributed to this work.

Corresponding author:

Yuta Aoki, MD, PhD (ORCID: 0000-0001-5617-4948)

Senior Assistant Professor

Medical Institute of Developmental Disabilities Research at Showa University

6-11-11 Kitakarasuyama, 157-8577, Tokyo, Japan

Email: yuaoki-tky@umin.ac.jp

Tel: +81.3.5315.9357

Fax: +81.3.5315.9358

Number of words in Abstract: 191

Number of words in Manuscript: 3720

Number of tables 2

Number of figures 3

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Endophenotype refers to a measurable and heritable component between genetics and diagnosis and exists in both individuals with a diagnosis and their unaffected siblings. We aimed to identify a pattern of endophenotype consisted of multiple connections. We enrolled adult male individuals with autism spectrum disorder (ASD) endophenotype (i.e., individuals with ASD and their unaffected siblings) and individuals without ASD endophenotype (i.e., pairs of typical development (TD) siblings) and utilized a machine learning approach to classify people with and without endophenotypes, based on resting-state functional connections (FCs). A sparse logistic regression successfully classified people as to the endophenotype (area under the curve=0.78, classification accuracy=75%), suggesting the existence of endophenotype pattern. A binomial test identified that nine FCs were consistently selected as inputs for the classifier. The least absolute shrinkage and selection operator with these nine FCs predicted severity of communication impairment among individuals with ASD (r=0.68, p=0.021). In addition, two of the nine FCs were statistically significantly correlated with the severity of communication impairment (r=0.81, p=0.0026 and r=-0.60, p=0.049). The current findings suggest that an ASD endophenotype pattern exists in FCs with a multivariate manner and is associated with clinical ASD phenotype.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Keywords: autism spectrum disorder, endophenotype, machine learning, resting state, unaffected siblings

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Autism spectrum disorder (ASD) is a developmental disorder characterised by deficits in social interaction and repetitive restricted behavior (American Psychiatric Association, 2013). Examples of these social interaction impairments among people with ASD include theory of mind, empathy, and facial emotion recognition (Simon Baron-Cohen, Leslie, & Frith, 1985; S. Baron-Cohen & Wheelwright, 2004; Uljarevic & Hamilton, 2013). Consistent with the findings that such social interactions depend on the social brain neural system (Adolphs, 2009), abnormalities within the social brain have been observed among individuals with ASD (Yuta Aoki, Cortese, & Tansella, 2015; Kana, Keller, Cherkassky, Minshew, & Just, 2009; Kleinhans, Richards, Greenson, Dawson, & Aylward, 2016; Murphy et al., 2017; Pelphrey, Shultz, Hudac, & Vander Wyk, 2011). Besides the anatomical extension of abnormalities, individuals with ASD consistently present atypical functional connections (FCs) (Cherkassky, Kana, Keller, & Just, 2006; Di Martino et al., 2014; Uddin, Supekar, & Menon, 2013). Thus, atypical FCs within the social brain are of interest and may underlie the impairment in social interactions among individuals with ASD.

ASD is a highly heritable condition (Colvert et al., 2015). Indeed, although the prevalence of ASD is about 1% in the general population (Autism & Developmental Disabilities Monitoring Network Surveillance Year Principal, 2014), the risk of developing ASD increases by up to 20%

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

among biological siblings of individuals with ASD (Ozonoff et al., 2011). Despite such obvious heritability, the concordance of clinical diagnoses among monozygotic twins is only about 60% indicates, indicating that ASD is influenced by complex genetic interactions (Hallmayer et al., 2011). Reflecting this complex genetic contribution, autistic traits have a continuous distribution among people with incomplete genetic traits of ASD (Constantino & Todd, 2003). In fact, biological family members of individuals with ASD often have subclinical difficulties in social interactions (Piven & Palmer, 1999).

Endophenotype refers to a measurable and heritable component between genes and disease diagnosis (Gottesman & Gould, 2003). Because of the complex genetic contribution, characterising the ASD endophenotype is particularly important because it may provide an objective intermediate marker and insight into the pathophysiology of ASD. Reflecting its heritability and consistently observed abnormalities among individuals with ASD (Di Martino et al., 2014; Glahn et al., 2010), resting-state functional magnetic resonance imaging (R-fMRI) is one of promising modalities that quantify ASD endophenotype (Khadka et al., 2013).

To identify the ASD endophenotype, prior neuroimaging studies using several modalities have enrolled individuals with ASD, their unaffected siblings, and typical development (TD)

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Barnea-Goraly, Lotspeich, & Reiss, 2010; Jou et al., 2016; Moseley et al., 2015). These studies separately compared each connection across these three groups and showed the shared atypical findings among unaffected siblings of individuals with ASD and individuals with ASD. Most of prior studies recognised such shared abnormality, which is widely scattered throughout the brain, as an endophenotype.

The aim of this study was three-folds. First, we aimed to classify pairs of people as to the ASD endophenotype using a multivariate machine learning approach. Motivation of applying a machine learning approach is that we recognised that ASD endophenotype likely consisted of multiple FCs with different weights (i.e., as a pattern of altered FCs), rather than an unweighted integration of abnormalities shared by individuals with ASD and their unaffected siblings. Second, we examined which FCs were consistently selected when classifying people as to the endophenotype. Finally, to confirm the relation between the selected FCs and the clinical ASD phenotype, we employed the least absolute shrinkage and selection operator (LASSO), to predict the severity of social interaction impairment among people with ASD using the FCs serving as the endophenotype. To do so, we obtained R-fMRI data from 60 participants, consisting of 30 people with ASD endophenotype (15 individuals with ASD and 15 of their unaffected siblings) and 30 people without the ASD endophenotype (15 pairs of

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

TDs). Since ASD is characterized by the impairment of social interactions, we focused on FCs within the social brain. To increase the homogeneity of the participants, we enrolled age- and IQ-matched males.

Materials and methods

Participants

We analysed data from 60 adult males consisting of 30 pairs of biological siblings. Thirty people had the ASD endophenotype. Specifically, 15 pairs of participants were discordant for the diagnosis of ASD: namely, one of the siblings was affected with ASD, and the other was unaffected. Another 30 people did not have the endophenotype and consisted of 15 pairs of TD siblings. None of the TD siblings had a family member who had been diagnosed as having ASD. All the participants with ASD were diagnosed by experienced psychiatrists based on the DSM-IV-TR (American Psychiatric, 2000). The diagnosis was further supported by the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2001). In addition, to confirm the absence of a diagnosis of ASD in the unaffected sibling, the parents of the siblings discordant for the diagnosis of ASD were interviewed using the Autism Diagnostic Interview-Revised (ADI-R) (Tsuchiya et al., 2013). We used the Edinburgh Handedness Inventory to evaluate handedness (Oldfield, 1971). The intelligence quotient (IQ) of each participant was assessed

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

using either the Wechsler Adult Intelligence Scale-Third Edition or the WAIS-Revised (Wechsler, 1997; Wechsler & De Lemos, 1981). All the participants completed the Japanese version of the Autism-Spectrum Quotient (AQ) (Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006). Psychiatric comorbidities were observed in two participants with ASD: one had attention-deficit/hyperactivity disorder, and the other had learning disabilities. Five of the participants were taking medication at the time of scanning: benzodiazepine (n = 3), anti-depressants (n = 3), and psychostimulant (n = 1). For inclusion in the TD group, we confirmed the absence of an ASD diagnosis among the family member and the absence of an Axis I diagnosis per the DSM-IV-TR using the Mini-International Neuropsychiatric Interview (Hergueta, Baker, & Dunbar, 1998). The absence of a history of psychotropic medication use was also required. The exclusion criteria for all the participants were known genetic diseases, an estimated full intelligence quotient (IQ) of 80 or below, or the total AQ score of 33 or above. Written informed consent was obtained from all the participants, after they had received a complete explanation of the study. The Ethics Committee of Showa University approved the study protocol. The study was prepared in accordance with the ethical standards of the Declaration of Helsinki.

MRI acquisition

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

All MRI data were acquired using a 3.0 T MRI scanner (MAGNETON Verio, Siemens Medical Systems, Erlangen, Germany) with a 12ch head coil. Functional images were acquired with an echo planar imaging sequence (repetition time [TR]: 2500 ms, echo time [TE]: 30 ms, flip angle: 80°, field of view [FOV]: 212 mm, matrix size: 64 × 64, slice thickness: 3.2 mm with a 0.8-mm gap, 40 axial slices) at rest for 10 min 10 s (244 volumes). During the resting-state scans, participants were asked to gaze at a cross-hair displayed at the center of the screen, not to think about specific things, and to stay awake. To correct for the distortion of the functional image, gradient echo field mapping images were acquired immediately after the resting-state scans (TR: 488 ms, short TE: 4.92 ms, long TE: 7.38 ms, flip angle: 60°, FOV: 212 mm, matrix size: 64 × 64, slice thickness: 3.2 mm with a 0.8-mm gap, 40 axial slices). For normalisation purpose, a T1-weighted image was acquired using an MPRAGE sequence (TR: 2.3 s, TE: 2.98 ms, flip angle: 9°, FOV: 256 mm, matrix size: 256 × 256, slice thickness: 1mm, 240 sagittal slices, voxel size: 1 × 1× 1 mm).

R-fMRI data preprocessing

All the functional images were preprocessed using the Statistical Parametric Mapping (SPM12; Wellcome Department of Cognitive Neurology, London, UK) and functions implemented in FMRIB's Software Library (FSL; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/).

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Preprocessing was performed as follows: 1) the first four volumes were discarded to allow for T1 equilibration; 2) slice timing correction; 3) head motion correction using mcflirt implemented in the FSL (Jenkinson, Bannister, Brady, & Smith, 2002), 4) distortion correction using the FUGUE implemented in the FSL, 5) co-registration of functional images to an anatomical image, 6) spatial normalisation and resampling to a resolution of 2 × 2 × 2 mm, and 7) spatial smoothing (6-mm full-width at half-maximum).

To remove the effects of subtle head motions during the scans (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012), ICA-AROMA was applied (Pruim, Mennes, Buitelaar, & Beckmann, 2015; Pruim, Mennes, van Rooij, et al., 2015). After ICA-AROMA was performed, nuisance regression was further performed. Nuisance signals consisted of signals averaged over white matter, cerebrospinal fluid, and grey matter, respectively (Parkes, Fulcher, Yu Cel, & Fornitod, 2017; Yahata et al., 2016). A band-pass filter (0.008 – 0.1 Hz) was then applied to the residual time-series in a voxel-wise manner. For each participant, the mean frame-wise displacement (FD) was calculated from head motion parameters to quantify the amount of head motion during scans (Jenkinson et al., 2002). The mean FD of each group is shown in Table 1.

The social brain connectome atlas was used for network construction (Alcala-Lopez et al.,

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2017). This atlas consists of 36 brain regions associated with social functioning, such as theory of mind, empathy, and facial emotion recognition. For each participant, the mean time-series was extracted from each region of interest (ROI) and the correlation coefficients between all possible pairs of ROIs were calculated, resulting in a 36 × 36 correlation matrix for each participant. Finally, Fisher's *r*-to-*z* transformation was applied to each correlation coefficient. This procedure generated 630 features for each participant.

Statistical analysis

Identification of endophenotype pattern of FCs

Since our aim was to identify a pattern of FCs serving as an endophenotype from a given data, we formulated the problem in the following manner. Based on previous findings that individuals with ASD and their unaffected siblings shared a pattern of alterations when compared to TDs (Moseley et al., 2015), we recognised that endophenotype was a measurable component satisfying the following conditions:

1. individuals with ASD have higher values than TDs (i.e., ASD > TD);
2. unaffected siblings of individuals with ASD also have higher values than TDs (i.e., unaffected sibling > TD);
3. TD sibling pairs have similar values (i.e., TD ≈ TD sibling).

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

According to the conditions (1) and (2), our aim could be achieved by solving a classification problem, in which we intended to identify FCs that could discriminate persons with a high endophenotype value from persons with a low endophenotype value. Individuals with ASD and their unaffected siblings were regarded as having the ASD endophenotype, while TD and TD siblings were regarded as not having the endophenotype throughout this manuscript.

To examine whether the FCs serve as an endophenotype, we employed sparse logistic regression (SLR) (Yamashita, Sato, Yoshioka, Tong, & Kamitani, 2008). SLR can train a logistic regression model while automatically selecting endophenotype-related FCs. Briefly, SLR relies on a hierarchical Bayesian estimation, in which the prior distribution of each element of the parameter vector is represented as a Gaussian distribution. Based on the automatic relevance determination, irrelevant features are not used in the classification because the respective Gaussian prior distributions have a sharp peak at zero. Such efficient feature elimination method implemented in SLR can mitigate the problems of over-fitting caused by a small sample size. To evaluate the performance of the classifier, a leave-one-pair-out cross-validation (LOPOCV) was performed. Of note, the term 'pair' stands for sibling pairs. In each fold, all-but-one pair was used to train the SLR classifier, while the remaining pair was used for the evaluation.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To further examine the statistical significance of the classification accuracy, a permutation test was performed. At each iteration, a permuted dataset was generated by shuffling the endophenotype label while keeping the pair information. Then, LOPOCV was performed to calculate the classification accuracy for the permuted dataset. This procedure was repeated 5,000 times to construct a null distribution. Statistical significance was set at $P < 0.05$.

Binomial test

SLR selects a small number of relevant FCs from a given data. To confirm that FCs selected by the classifier were not randomly selected, the statistical significance of the selection counts was examined using a binomial test. The classifier selected 8.13 ± 0.94 (mean ± standard deviation [SD]) out of 630 FCs in each of the 30 validation folds (see Results). Thus, we assumed a binomial distribution, $Bi(n, p)$, where n stands for the number of validation folds (i.e., $n = 30$) and p stands for the probability of being selected from the set of FCs (i.e., $p = 8/630$).

Relation between endophenotype-related FCs and clinical phenotype

Once the endophenotype-related FCs had been identified by the classifier with binomial tests,

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

we predicted the severity of the clinical symptoms measured by the ADOS. The score of each individual was predicted using LASSO with nine FCs consistently selected in the classifier (Tibshirani, 1996). While an optimal regularization parameter in the LASSO was determined by an internal 10-fold CV, the weight parameters were determined through a leave-one-subject-out CV (LOSOCV), in which the scores of all-but-one participants were linearly regressed using the nine FCs as explanatory variables. Given the notion that age-related differences might have an impact on the severity of clinical symptoms (Lee, Park, James, Kim, & Park, 2017), we added age as an additional explanatory variable. To evaluate how predicted and actual scores matched, the correlation coefficient was calculated. We further examined the extent to which the FCs were significantly correlated with the severity of clinical symptoms by calculating correlation coefficient. Because of the explorative of this analysis nature, we used a liberal statistical threshold of $P < 0.05$.

Results

Identification of endophenotype pattern of FCs

We discriminated participants with the ASD endophenotype (i.e., individuals with ASD and their unaffected siblings) from TDs using an SLR with the LOPOCV. This classifier separated participants with endophenotype from TDs with 75% accuracy (sensitivity = 76.67% and

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

specificity = 73.33%) and an area under the curve (AUC) of 0.78 (permutation test with 5,000 iterations, $p < 0.001$; Figures 1a and 1b), suggesting that FCs selected by this classifier captured the endophenotype-related features.

Post-hoc paired *t*-tests were performed to examine whether the conditions (see Methods) were satisfied. As shown in Figure 1c, paired *t*-tests demonstrated that the weighted linear summation (WLS) of FCs selected in the classifier was not statistically significantly different between individuals with ASD and their unaffected siblings (*t*-value = -0.03, $df = 14$, $p = 0.98$). In contrast, a two-sample *t*-test showed statistically significant differences between people with endophenotype (i.e., individuals with ASD and their unaffected siblings) and those without endophenotype (*t*-value = 10.28, $df = 58$, $p < 0.001$). These results indicate that the WLS of FCs selected by the classifier satisfied the set of conditions regarding endophenotype, but not the clinical diagnosis. Furthermore, the analysis did not show any significance difference between TD sibling pairs (*t*-value = 1.14, $df = 14$, $p = 0.27$).

To rule out the possibility that nuisance covariates (i.e., head motion and age) have potential to be important features for classification, we repeated the same classification analysis while adding the two nuisance variables as additional explanatory variables (i.e., 630 FCs + 2

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

nuisance variables). If the nuisance variables could explain the endophenotype label over the previously selected FCs, then the SLR would automatically select these nuisance variables as input for the logistic function. We confirmed that the classification accuracy and the AUC were not changed by adding these nuisance variables (accuracy = 75% and AUC = 0.78). Furthermore, in each fold, the SLR never selected the two nuisance variables as inputs for the logistic function (i.e., the weighted parameters for these variables were zero).

Binomial test

Furthermore, we investigated which FCs were stably selected by the SLR across the LOPOCV. In each fold, the classifier selected 8.13 ± 0.94 (mean ± SD) out of the 630 FCs across 30 validation folds. We counted how many times each FC was selected. Under the null hypothesis that eight FCs were randomly selected from 630 FCs, a binomial test was applied to examine the probability of the selection count. We found that nine FCs were selected at a significant frequency ($p < 0.05$, Bonferroni corrected for 630 connections; Figure 2 and Table 2). The selection count for these nine FCs was 22.78 ± 7.61 while that for the remaining FCs was 0.0627 ± 0.33. This result indicates that the nine FCs were consistently selected across the validation folds.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Relation between endophenotype-related FCs and clinical phenotype

We predicted the severity of the clinical symptoms measured by the ADOS using the nine FCs identified in the classifier for endophenotype. LASSO with LOSOCV was individually applied to determine the weights of the nine FCs so that their weighted linear summation was used as a predictor for the corresponding severity score. We found that the communication domain of the ADOS was well predicted from the nine FCs with a statistically significant correlation (r = 0.68, p = 0.021; Figure 3a). Furthermore, we found that two of nine FCs were significantly correlated with the severity of impaired communication measured by ADOS: one was FC between the right temporo-parietal junction and right inferior frontal gyrus (r = 0.81, p = 0.0026; Figure 3b), and the other was FC between the right nucleus accumbens and anterior middle cingulate cortex (r = -0.60, p = 0.049; Figure 3c).

Discussion

Our novel framework was motivated by 1) the observation that neural correlates of ASD emerge as a pattern of FCs and 2) the lack of any previous study attempting to identify the endophenotype as a pattern of FCs. Thus, we applied the machine learning approach to classify people as to the ASD endophenotype. Subsequent binomial tests identified nine FCs serving as the endophenotype. The LASSO further revealed that the selected FCs were

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

associated with the clinical phenotype among individuals with ASD. This high-performance proof-of-concept study shows the importance and relevance of using the machine learning approach to identify the endophenotype as a pattern of FCs.

We utilised the SLR to classify participants as to the ASD endophenotype. Most prior neuroimaging studies identified abnormalities shared by individuals with ASD and their unaffected siblings or abnormalities in the siblings that were intermediate between ASD and TD as an endophenotype (Barnea-Goraly et al., 2010; Jou et al., 2016; Moseley et al., 2015). In these studies, each variable (i.e., connectivity index) was separately examined to determine whether it belonged to the endophenotype. In contrast, we applied a multivariate and data-driven approach to classify individuals with or without the ASD endophenotype using all the available variables (i.e., 630 FCs within the social brain). This approach inherently assumes that the endophenotype emerges as a pattern of altered FCs and allows all the variables to contribute to the endophenotype to different extents. In the case of ASD, the genetic abnormalities vary across individuals (Geschwind & Levitt, 2007; Miles, 2011). In addition, multiple FCs are altered (Di Martino et al., 2014; Uddin et al., 2013). Thus, the genetic influence is not identical across FCs (Ameis & Szatmari, 2012; Meyer-Lindenberg & Weinberger, 2006; Zhan et al., 2014), and the endophenotype should emerge as multiple FCs

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with different extents. The consistency with the definition of the endophenotype might explain the good performance of classification.

The current study identified an ASD endophenotype among the FCs within the social brain. Since not all individuals with the endophenotype develop ASD clinical diagnosis (Ozonoff et al., 2011), the neural correlates for the endophenotype and the diagnosis may differ. Although we did not investigate the neural correlates for the diagnosis because of the small sample size, neural correlates for the clinical diagnosis of ASD might be located within the social brain (Yuta Aoki et al., 2015; Kana et al., 2009; Kleinhans et al., 2016; Murphy et al., 2017; Pelphrey et al., 2011). Given that there was no significant difference in the WLS of the selected FCs between individuals with ASD and their unaffected siblings, the selected FCs represent the ASD endophenotype, rather than the ASD diagnosis. Although it is beyond the scope of this proof-of-concept study, future research involving a large sample size is expected to dissociate the neural correlates of the ASD clinical diagnosis and the ASD endophenotype.

Using the LASSO, the selected FCs successfully predicted the severity of communication deficits among individuals with ASD. This relationship corroborates the assumption that the selected FCs are associated with the pathophysiology of ASD. Since we lacked any ADOS scores

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

from the unaffected siblings of individuals with ASD, this relationship was observed only among the individuals with ASD. However, even if the ADOS scores were available from the unaffected siblings, the scores might not reflect the distribution of behaviour well, since the ADOS aims to measure clinical symptoms, not subclinical ones observed among people with the ASD endophenotype (Toth, Dawson, Meltzoff, Greenson, & Fein, 2007). Future studies enrolling unaffected siblings should obtain both MRI data and psychological evaluations that reflect the differences in ASD traits among subclinical individuals.

Several limitations should be considered in this study. First, although we successfully added the condition that the FCs should not differ between TD siblings (see Method), which was not possible in prior studies with only one TD group, the current study had a small sample size because of practical difficulties in recruitment. Thus, because of the lack of statistical power, we could not address brain regions outside of the social brain. However, brain regions outside the social brain are also likely to be involved in the endophenotype of ASD. In addition, although the sparseness of FC selection by the SLR mitigates over-fitting to the current sample, the current sample size was relatively small, compared with the number of FCs. A future study with a sample size large enough to avoid over-fitting is needed to examine the whole brain. Second, although the machine learning approach successfully classified individuals with or

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

without the ASD endophenotype, the endophenotype may not be categorical, as with ASD diagnosis (Y. Aoki et al., 2017). Research using a large sample is encouraged to stratify individuals with an ASD endophenotype. Third, to increase the homogeneity of participants, we recruited only adult males in the current study. Given that the ASD brain shows an atypical developmental trajectory (Y. Aoki, Kasai, & Yamasue, 2012), the current finding may not be generalised to other age ranges. In addition, atypical sex differences have been observed among individuals with ASD (Lai et al., 2017). Thus, further studies that include only female ASD participants are expected.

Conclusion

Using the machine learning approach, we demonstrated that the ASD endophenotype emerges as a pattern of FCs within the social brain. The FCs selected as inputs for the classifier were associated with the clinical ASD severity.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Funding:

This study is the result of "Development of BMI Technologies for Clinical Application" carried out under the Strategic Research Program for Brain Sciences by Japan Agency for Medical Research and Development (AMED). This work is partly supported by the grant by The Japan Foundation for Pediatric Research (to YA).

Conflict of Interest:

All authors report no conflict of interest.

Ethical approval :

The study was prepared in accordance with the ethical standards of the Declaration of Helsinki.

Informed consent:

Written informed consent was obtained from all the participants, after they had received a complete explanation of the study.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annu Rev Psychol, 60*, 693-716.

Alcala-Lopez, D., Smallwood, J., Jefferies, E., Van Overwalle, F., Vogeley, K., Mars, R. B., et al. (2017). Computing the Social Brain Connectome Across Systems and States. *Cereb Cortex*(1460-2199 (Electronic)), 1-26.

Ameis, S. H., & Szatmari, P. (2012). Imaging-genetics in autism spectrum disorder: advances, translational impact, and future directions. *Front Psychiatry, 3*, 46.

American Psychiatric, A. (2000). Diagnostic and statistical manual, 4th edn, Text Revision (DSM-IV-TR). *American Psychiatric Association, Washington*.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Arlington, VA: American Psychiatric Association.

Aoki, Y., Cortese, S., & Tansella, M. (2015). Neural bases of atypical emotional face processing in autism: a meta-analysis of fMRI studies. *The World Journal of Biological Psychiatry, 16*(5), 291-300.

Aoki, Y., Kasai, K., & Yamasue, H. (2012). Age-related change in brain metabolite abnormalities in autism: a meta-analysis of proton magnetic resonance spectroscopy studies. *Translational Psychiatry, 2*, e69.

Aoki, Y., Yoncheva, Y. N., Chen, B., Nath, T., Sharp, D., Lazar, M., et al. (2017). Association of white matter structure with autism spectrum disorder and attention-deficit/hyperactivity disorder. *JAMA Psychiatry, 74*(11), 1120-1128.

Autism, & Developmental Disabilities Monitoring Network Surveillance Year Principal, I. (2014). Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Morbidity and Mortality Weekly Report: Surveillance Summaries, 63*(2), 1-21.

Barnea-Goraly, N., Lotspeich, L. J., & Reiss, A. L. (2010). Similar white matter aberrations in children with autism and their unaffected siblings: a diffusion tensor imaging study using tract-based spatial statistics. *Archives of General Psychiatry, 67*(10), 1052-1060.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J Autism Dev Disord, 34*(2), 163-175.

Cherkassky, V. L., Kana, R. K., Keller, T. A., & Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport, 17*(16), 1687-1690.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., et al. (2015). Heritability of autism spectrum disorder in a UK population-based twin sample. *JAMA Psychiatry, 72*(5), 415-423.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Arch Gen Psychiatry, 60*(5), 524-530.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry, 19*(6), 659-667.

Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: developmental disconnection syndromes. *Current Opinion in Neurobiology, 17*(1), 103-111.

Glahn, D. C., Winkler, A. M., Kochunov, P., Almasy, L., Duggirala, R., Carless, M. A., et al. (2010). Genetic control over the resting brain. *Proceedings of the National Academy of Sciences of the United States of America, 107*(3), 1223-1228.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry, 160*(4), 636-645.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., et al. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry, 68*(11), 1095-1102.

Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IVand ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl 20), 2233.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*(2), 825-841.

Jou, R. J., Reed, H. E., Kaiser, M. D., Voos, A. C., Volkmar, F. R., & Pelphrey, K. A. (2016). White matter abnormalities in autism and unaffected siblings. *Journal of Neuropsychiatry & Clinical Neurosciences, 28*(1), 49-55.

Kana, R. K., Keller, T. A., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Social Neuroscience, 4*(2), 135-152.

Khadka, S., Meda, S. A., Stevens, M. C., Glahn, D. C., Calhoun, V. D., Sweeney, J. A., et al. (2013). Is aberrant functional connectivity a psychosis endophenotype? A resting state functional magnetic resonance imaging study. *Biological Psychiatry, 74*(6), 458-466.

Kleinhans, N. M., Richards, T., Greenson, J., Dawson, G., & Aylward, E. (2016). Altered dynamics of the fMRI response to faces in individuals with autism. *Journal of Autism and Developmental Disorders, 46*(1), 232-241.

Lai, M. C., Lerch, J. P., Floris, D. L., Ruigrok, A. N. V., Pohl, A., Lombardo, M. V., et al. (2017). Imaging sex/gender and autism in the brain: Etiological implications. *Journal of Neuroscience Research, 95*(1-2), 380-397.

Lee, Y., Park, B. Y., James, O., Kim, S. G., & Park, H. (2017). Autism spectrum disorder related functional connectivity changes in the language network in children, adolescents and adults. *Frontiers in Human Neuroscience, 11*(1662-5161 (Print)), 418.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2001). *Autism diagnostic observation schedule (ADOS): Manual*: WPS.

Meyer-Lindenberg, A., & Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

psychiatric disorders. *Nature Reviews Neuroscience, 7*(10), 818-827.

Miles, J. H. (2011). Autism spectrum disorders—a genetics review. *Genetics in Medicine, 13*(4), 278.

Moseley, R. L., Ypma, R. J., Holt, R. J., Floris, D., Chura, L. R., Spencer, M. D., et al. (2015). Whole-brain functional hypoconnectivity as an endophenotype of autism in adolescents. *Neuroimage Clinical, 9*, 140-152.

Murphy, C. M., Christakou, A., Giampietro, V., Brammer, M., Daly, E. M., Ecker, C., et al. (2017). Abnormal functional activation and maturation of ventromedial prefrontal cortex and cerebellum during temporal discounting in autism spectrum disorder. *Human Brain Mapping, 38*(11), 5343-5355.

Okada, N., Kasai, K., Takahashi, T., Suzuki, M., Hashimoto, R., & Kawakami, N. (2014). Brief rating scale of socioeconomic status for biological psychiatry research among Japanese people: A scaling based on an educational history. *Japanese Journal of Biological Psychiatry, 25*, 115-117.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (2011). Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatrics, 128*(3), e488-e495.

Parkes, L., Fulcher, B., Yu Cel, M., & Fornitod, A. (2017). An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *Neuroimage*.

Pelphrey, K. A., Shultz, S., Hudac, C. M., & Vander Wyk, B. C. (2011). Research review: Constraining heterogeneity: the social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 52*(6), 631-644.

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: evidence from a family study of multiple-incidence autism families. *American Journal of Psychiatry, 156*(4), 557-563.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage, 59*(3), 2142-2154.

Pruim, R. H., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2015). Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *Neuroimage, 112*, 278-287.

Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage, 112*, 267-277.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.

Toth, K., Dawson, G., Meltzoff, A. N., Greenson, J., & Fein, D. (2007). Early social, imitation, play, and language abilities of young non-autistic siblings of children with autism. *Journal of Autism and Developmental Disorders, 37*(1), 145-157.

Tsuchiya, K. J., Matsumoto, K., Yagi, A., Inada, N., Kuroda, M., Inokuchi, E., et al. (2013). Reliability and validity of autism diagnostic interview-revised, Japanese version. *Journal of Autism and Developmental Disorders, 43*(3), 643-662.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Uddin, L. Q., Supekar, K., & Menon, V. (2013). Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Frontiers in Human Neuroscience, 7*, 458.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: a formal meta-analysis. *J Autism Dev Disord, 43*(7), 1517-1526.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: a cross-cultural comparison. *Journal of autism and developmental disorders, 36*(2), 263-270.

Wechsler, D. (1997). *WAIS-III: Wechsler adult intelligence scale*: Psychological Corporation.

Wechsler, D., & De Lemos, M. M. (1981). *Wechsler adult intelligence scale-revised*: Harcourt Brace Jovanovich.

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., et al. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nature Communications, 7*.

Yamashita, O., Sato, M. A., Yoshioka, T., Tong, F., & Kamitani, Y. (2008). Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage, 42*(4), 1414-1429.

Zhan, Y., Paolicelli, R. C., Sforazzini, F., Weinhard, L., Bolasco, G., Pagani, F., et al. (2014). Deficient neuron-microglia signaling results in impaired functional brain connectivity and social behavior. *Nature Neuroscience, 17*(3), 400-406.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure Legends

Figure 1. Classification results and post-hoc tests. (a) Receiver operating characteristic curve. A sparse logistic regression with leave-one-pair-out cross validation exhibited an area under the curve of 0.78. (b) Null distribution of classification accuracy derived by a permutation test with 5,000 iterations. The permutation test demonstrated that the observed classification accuracy (75%) was significantly higher than that obtained by the permuted classifier ($p < 0.001$). (c) Post-hoc paired *t*-tests demonstrated that there were no statistically significant differences between individuals with autism spectrum disorder (ASD) and their unaffected siblings in the weighted linear summation of selected FCs (*t*-value = -0.03, $df = 14$, $p = 0.98$) as well as between TD pairs (*t*-value = 1.14, $df = 14$, $p = 0.27$). On the other hand, persons with endophenotype exhibited significantly higher values than those without endophenotype (*t*-value = 10.28, $df = 58$, $p < 0.001$). ***: $P < 0.001$.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2. Results of identified functional connections (FCs) stably selected for the classification.

The left panel shows the results of binomial tests. The red circle indicates the statistical significance of the selected count after multiple comparison correction. The right panel displays the distribution of nine FCs exhibiting statistical significance after multiple comparison corrections. The colors of nodes represent the sub-systems within the social brain atlas defined in (Alcala-Lopez et al.) Red: visual-sensory system, Yellow: limbic system, Green: intermediate-level system, Cyan: higher-level system. **Abbreviations:** AI: anterior insula, aMCC: anterior middle cingulate cortex, Amy: amygdala, Cb: cerebellum, dmPFC: dorsomedial prefrontal cortex, FFA: fusiform face area, FP: frontal pole, HC: hippocampal cortex, IFG: inferior frontal gyrus, MT/V5: middle temporal V5 area, MTG: middle temporal gyrus, NAcc: nucleus accumbens, PCC: posterior cingulate cortex, PCu: precuneus, pSTS: posterior superior temporal sulcus, SMA: supplementary motor area, SMG: supramarginal gyrus, TPJ: temporo-parietal junction, vmPFC: ventromedial prefrontal cortex

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 3: Associations between the severity of clinical symptoms and functional connectivity (FCs) selected by the classifier. (a) Prediction of the severity of communication measured by the Autism Diagnostic Observation Schedule (ADOS) using the least absolute shrinkage and selection operator with leave-one-subject-out cross-validation. The correlation coefficient between the predicted and measured scores exhibited a significant positive correlation (r = 0.68, p = 0.021). (b) The connectivity strength between the right temporo-parietal junction and right inferior frontal gyrus was significantly correlated with the severity of impaired communication as measured by the ADOS-A (r = 0.81, p = 0.0026). (c) The connectivity strength between the right nucleus accumbens and anterior middle cingulate cortex was significantly correlated with the severity of communication impairment as measured by the ADOS-A (r = -0.60, p = 0.049).

Table 1: Characteristics of the participants

	People with endophenotype (n=30)				Statistics (a)			People without endophenotype (n=30)				Statistics (d)		
	ASD (n=15)		unaffected siblings (n=15)		df	t-value	P-value	TD (b) (n=15)		TD (c) (n=15)		df	t-value	P-value
	Mean	SD	Mean	SD				Mean	SD	Mean	SD			
Age (years)	28.3	6.1	28	7.3	14	0.45	0.657	28.4	6.5	25.1	5.3	14	5.62	<0.001
Full IQ	111.8	16	107.4	13.2	14	0.77	0.452	115.9	15.7	114.9	12.1	14	0.24	0.815
Verbal IQ	106.3	31.2	107.1	16.8	14	0.08	0.94	117.1	15.9	116.7	12.5	14	0.18	0.859
Performance IQ	108.1	16.4	105.3	9.8	14	0.62	0.547	109.4	13.6	107.7	7.9	14	0.42	0.678
Handedness	61.8	68.4	99.3	2.9	14	2.11	0.054	89.2	26.4	80.7	51.4	14	0.55	0.594
ADI-R*														
Social interaction	20.7	6.3	0.8	1.4	13	11.56	<0.001	-	-	-	-	-	-	-
Communication (verbal)	12.6	5	0.4	1.1	13	8.83	<0.001	-	-	-	-	-	-	-
RRB	3.9	2	0	0	13	7.02	<0.001	-	-	-	-	-	-	-
ADOS**														
Communication	4.45	1.13	-	-	-	-	-	-	-	-	-	-	-	-
Social interaction	7.91	2.12	-	-	-	-	-	-	-	-	-	-	-	-
Communication + Social interaction	12.36	2.73	-	-	-	-	-	-	-	-	-	-	-	-
RRB	0.18	0.40	-	-	-	-	-	-	-	-	-	-	-	-
AQ														

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Total	34.2	6.2	19.4	6.9	12	6.27	<0.001	17.6	5.5	14.9	6	14	1.12	0.28
AS	7.4	1.6	3.9	1.8	12	6.3	<0.001	4	2.1	3.7	1.4	14	0.54	0.596
ATD	5.4	1.9	3.2	1.9	12	2.67	0.02	3.4	2.1	3.9	1.6	14	0.88	0.396
COM	7.6	1.6	3.6	2.6	12	4.16	0.001	3.1	2.5	2.6	2.3	14	0.5	0.625
IMG	6	2	4.1	1.4	12	2.31	0.04	3.7	1.7	2.8	1.7	14	1.26	0.229
SS	7.9	2	4.5	2.7	12	4.26	0.001	3.4	2.7	2	2.1	14	1.38	0.191
SES (e	5.5	1.1	5.5	1.1	14	0.2	0.843	5.9	1.2	5.7	1.1	13	0.46	0.655
Mean FD	0.17	0.07	0.14	0.05	14	1.51	0.153	0.16	0.07	0.18	0.10	14	0.687	0.503

(a) Statistics show the results of comparisons between adult males with ASD and their unaffected brothers. (b) Older siblings. (c) Younger siblings. (d) Statistics show the results of comparisons between typical older and younger siblings. (e) A higher score indicates a lower socioeconomic status (Okada et al., 2014). Abbreviations: ADOS-2: Autism Diagnostic Observation Schedule Second Edition, ADI-R: Autism Diagnostic Interview-Revised, AQ: autism-spectrum quotient, AS: attention switching/tolerance of change, ASD: autism spectrum disorder, ATD: attention to detail, COM: communication skills, IMG: imagination, IQ: intelligence quotient, RRB: restricted repetitive behaviors, SD: standard deviation, SES: socioeconomic status, SS: social skills, TD: typical development. *The ADI-R score was missing for one person. **ADOS scores are missing for four people.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/348599; this version posted June 18, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

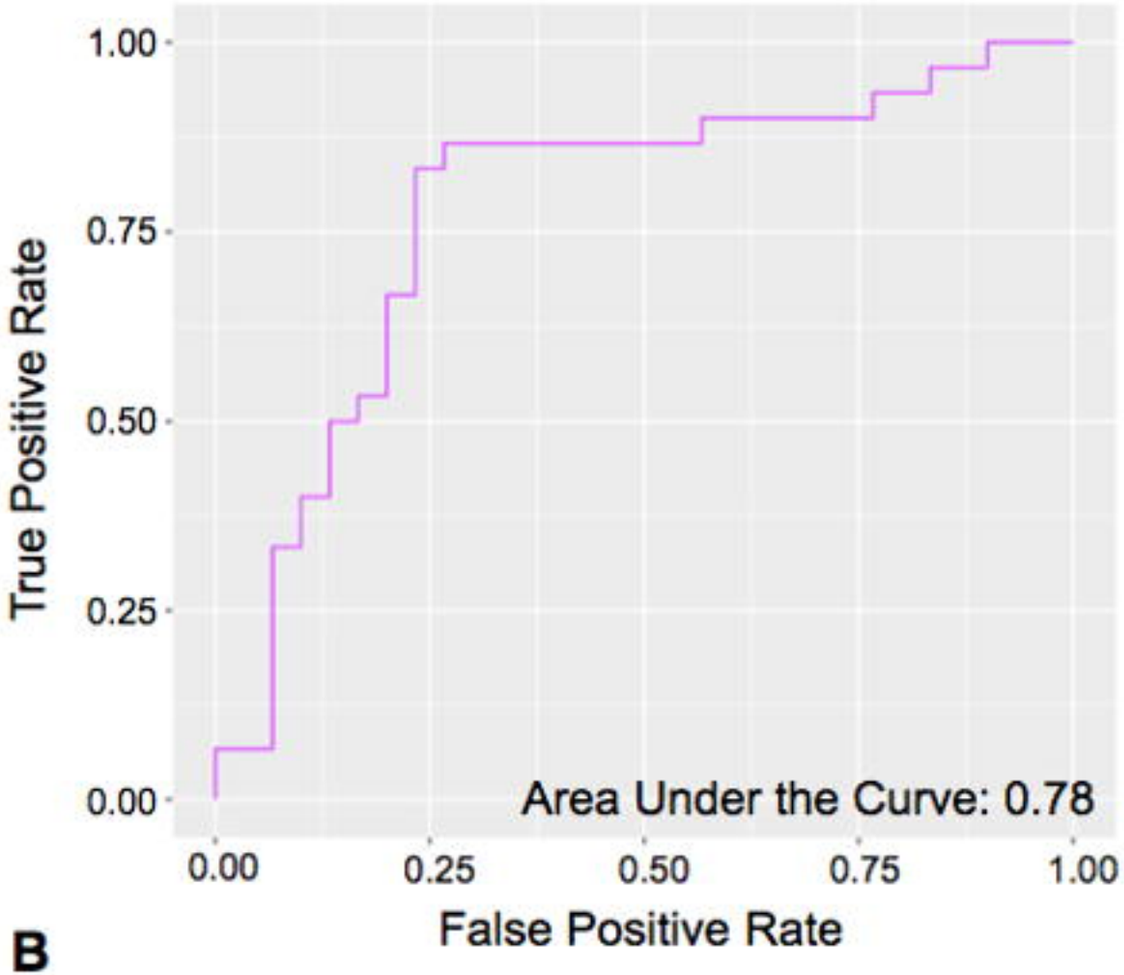
Table 2. The list of functional connections (FCs) stably selected for classification.

ID	Terminal region 1			Terminal region 2			Mean
	Name	Hemi.	Sub-system	Label	Hemi.	Sub-system	weight
1	SMA	R	Intermediate	pSTS	L	Visual-sensory	5.75
2	TPJ	R	Higher	IFG	R	Intermediate	2.42
3	NAcc	R	Limbic	NAcc	L	Limbic	-0.81
4	pMCC	L	Intermediate	HC	L	Limbic	-3.02
5	pSTS	R	Visual-sensory	Pcu	M	Higher	-5.71
6	NAcc	R	Limbic	aMCC	M	Intermediate	-6.49
7	Amy	R	Limbic	SMG	R	Intermediate	-6.69
8	SMG	R	Intermediate	IFG	L	Intermediate	-6.84
9	MTG	R	Higher	rACC	M	Limbic	-7.75

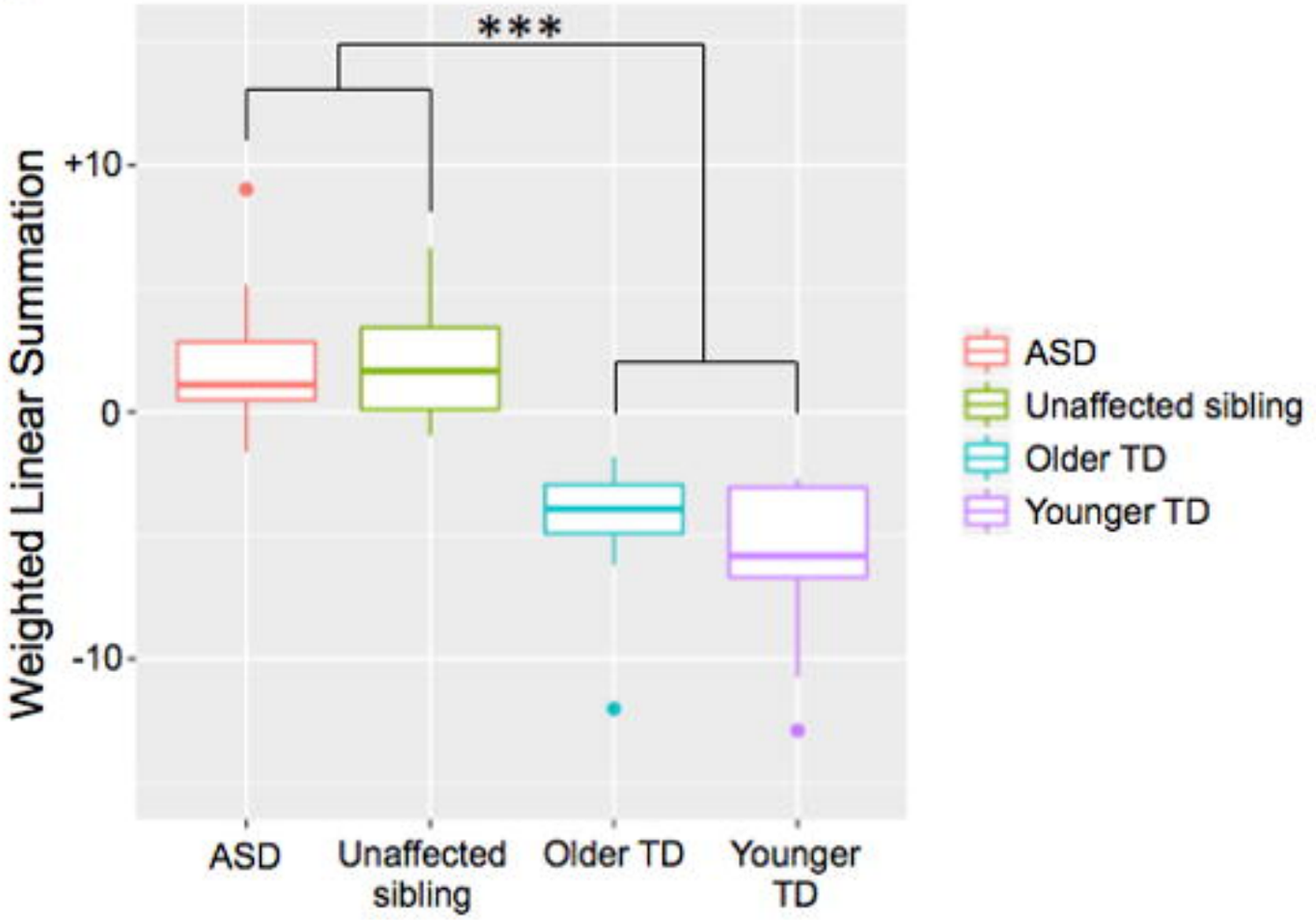
Brain regions and sub-systems were based on (Alcala-Lopez et al.)

Abbreviations: aMCC: anterior middle cingulate cortex, Amy: amygdala, HC: hippocampal cortex, Hemi: hemisphere, IFG: inferior frontal gyrus, L: left, M: middle, MTG: middle temporal gyrus, NAcc: nucleus accumbens, PCu: precuneus, pMCC: posterior medial cingulate cortex, pSTS: posterior superior temporal sulcus, R: right, rACC: right anterior cingulate cortex, SMA: supplementary motor area, SMG: supramarginal gyrus, TPJ: temporo-parietal junction

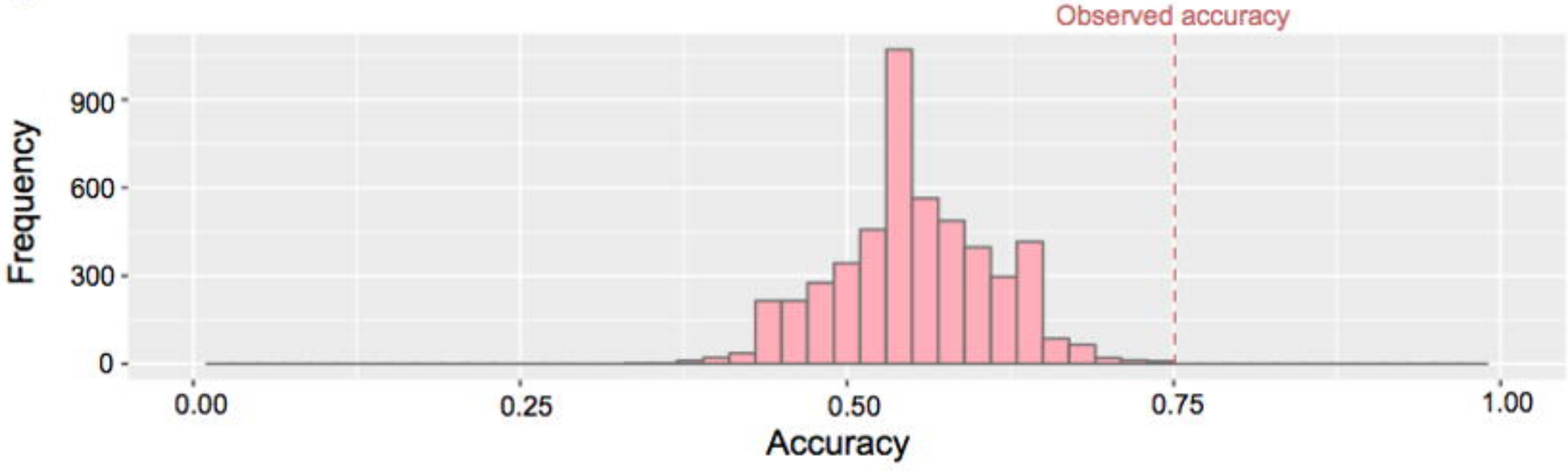
A

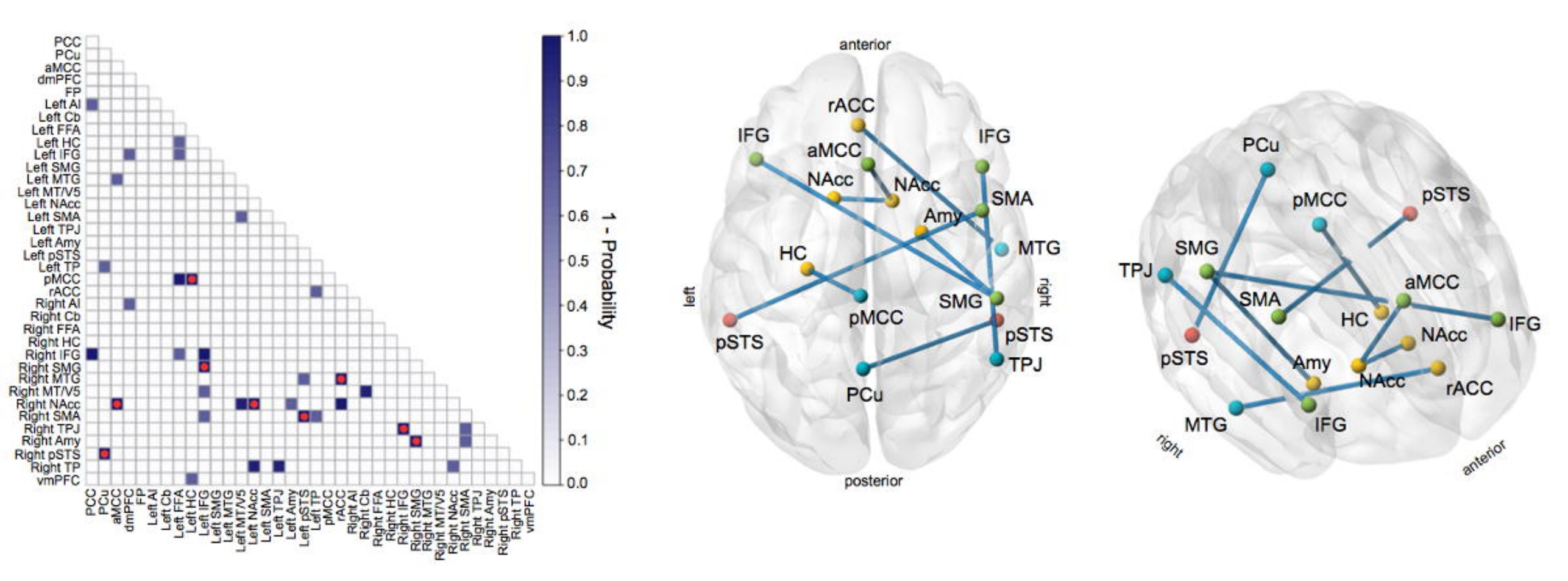

1.00
0.75
0.50
0.25
0.00
True Positive Rate
Area Under the Curve: 0.78
0.00
0.25
0.50
0.75
1.00
False Positive Rate

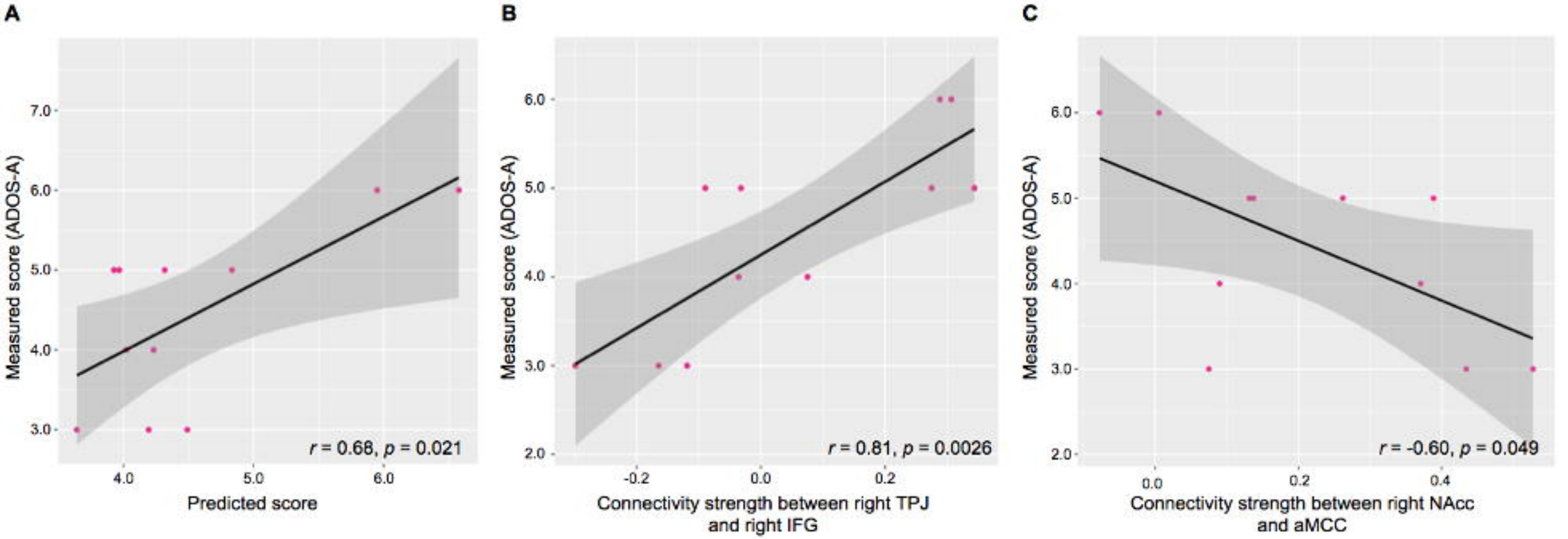
C


***
+10
0
-10
Weighted Linear Summation
ASD
Unaffected sibling
Older TD
Younger TD
ASD
Unaffected sibling
Older TD
Younger TD

B


Observed accuracy
900
600
300
0
Frequency
0.00
0.25
0.50
0.75
1.00
Accuracy


PCC
PCu
aMCC
dmPFC
FP
Left AI
Left Cb
Left FFA
Left HC
Left IFG
Left SMG
Left MTG
Left MT/V5
Left NAcc
Left SMA
Left TPJ
Left Amy
Left pSTS
Left TP
pMCC
rACC
Right AI
Right Cb
Right FFA
Right HC
Right IFG
Right SMG
Right MTG
Right MT/V5
Right NAcc
Right SMA
Right TPJ
Right Amy
Right pSTS
Right TP
vmPFC
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1 - Probability
anterior
posterior
left
right
rACC
IFG
aMCC
NAcc
SMA
Amy
MTG
HC
SMG
pMCC
pSTS
TPJ
PCu
PCu
pMCC
pSTS
SMG
TPJ
aMCC
SMA
HC
IFG
NAcc
pSTS
Amy
rACC
MTG
IFG
right
anterior


A
Measured score (ADOS-A)
7.0
6.0
5.0
4.0
3.0
4.0
5.0
6.0
Predicted score
r = 0.68, p = 0.021
B
Measured score (ADOS-A)
6.0
5.0
4.0
3.0
2.0
-0.2
0.0
0.2
Connectivity strength between right TPJ and right IFG
r = 0.81, p = 0.0026
C
Measured score (ADOS-A)
6.0
5.0
4.0
3.0
2.0
0.0
0.2
0.4
Connectivity strength between right NAcc and aMCC
r = -0.60, p = 0.049