BRAIN
A JOURNAL OF NEUROLOGY

ORIGINAL ARTICLE

A neurogenetic analysis of female autism

Allison Jack,[1] Catherine A. W. Sullivan,[2] Elizabeth Aylward,[3]
Susan Y. Bookheimer,[4] Mirella Dapretto,[4] Nadine Gaab,[5,6,7] John D. Van Horn,[8,9]
Jeffrey Eilbott,[10] Zachary Jacokes,[9] Carinna M. Torgerson,[11] Raphael A. Bernier,[12,13]
Daniel H. Geschwind,[4,14] James C. McPartland,[10] Charles A. Nelson,[5,6]
Sara J. Webb,[12,13] Kevin A. Pelphrey,[8,15] Abha R. Gupta[2,10,16]* and the
GENDAAR Consortium[†]

[†]Appendix 1.

Females versus males are less frequently diagnosed with autism spectrum disorder (ASD), and while understanding sex differences is critical to delineating the systems biology of the condition, female ASD is understudied.

We integrated functional MRI and genetic data in a sex-balanced sample of ASD and typically developing youth (8–17 years old) to characterize female-specific pathways of ASD risk. Our primary objectives were to: (i) characterize female ASD ($n = 45$) brain response to human motion, relative to matched typically developing female youth ($n = 45$); and (ii) evaluate whether genetic data could provide further insight into the potential relevance of these brain functional differences.

For our first objective we found that ASD females showed markedly reduced response versus typically developing females, particularly in sensorimotor, striatal, and frontal regions. This difference between ASD and typically developing females does not resemble differences between ASD ($n = 47$) and typically developing males ($n = 47$), even though neural response did not significantly differ between female and male ASD.

For our second objective, we found that ASD females ($n = 61$), versus males ($n = 66$), showed larger median size of rare copy number variants containing gene(s) expressed in early life (10 postconceptual weeks to 2 years) in regions implicated by the typically developing female > female functional MRI contrast. *Post hoc* analyses suggested this difference was primarily driven by copy number variants containing gene(s) expressed in striatum. This striatal finding was reproducible among $n = 2075$ probands (291 female) from an independent cohort.

Together, our findings suggest that striatal impacts may contribute to pathways of risk in female ASD and advocate caution in drawing conclusions regarding female ASD based on male-predominant cohorts.

1   Department of Psychology, George Mason University, Fairfax, VA 22030, USA
2   Department of Pediatrics, Yale School of Medicine, New Haven, CT 06510, USA
3   Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA 98101, USA
4   Department of Psychiatry and Biobehavioral Sciences and Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles School of Medicine, Los Angeles, CA 90095, USA
5   Division of Developmental Medicine, Department of Medicine, Boston Children's Hospital, Boston, MA 02115 USA
6   Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA
7   Harvard Graduate School of Education, Cambridge, MA 02138, USA
8   Department of Psychology, University of Virginia, Charlottesville, VA, USA
9   School of Data Science, University of Virginia, Charlottesville, VA, USA
10 Child Study Center, Yale School of Medicine, New Haven, CT 06510, USA
11 Neuroscience Graduate Program, University of Southern California, Los Angeles, CA 90007, USA
12 Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA
13 Center for Child Health, Behavior, and Development, Seattle Children's Research Institute, Seattle, WA 98101, USA

14 Department of Neurology and Center for Neurobehavioral Genetics, University of California Los Angeles School of Medicine, Los Angeles, CA 90095, USA

15 Department of Neurology, Brain Institute, and School of Education and Human Development, University of Virginia, Charlottesville, VA, USA

16 Department of Neuroscience, Yale School of Medicine, New Haven, CT 06510, USA

Correspondence to: Allison Jack
Department of Psychology, George Mason University, 4400 University Drive
3F5, Fairfax, VA 22030, USA
E-mail: ajack@gmu.edu

Correspondence may also be addressed to: Abha R. Gupta
The Anlyan Center, Room S-310, 300 Cedar St.
New Haven, CT 06520, USA
E-mail: abha.gupta@yale.edu

Introduction

Males are diagnosed with autism spectrum disorder (ASD) at rates of around four males to every one female.[1] Although non-biological factors such as ascertainment bias likely contribute to this skewed ratio,[2–4] differential prevalence estimates are likely to remain at ∼3:1 even after accounting for these factors,[5] signalling a robust aetiological role for sexually dimorphic biological processes. Investigating sex differences in ASD brain function, genetics, and gene-brain interactions is thus a critical step in delineating the systems biology of ASD. However, previous work on this question has been hindered by minimal inclusion of ASD females (ASDf).

Much work has focused on characterizing differences in ASD brain function relative to that of typically developing individuals. Emphasis is often placed on the neural bases of social perception, a multifaceted process that includes selectively attending to socially relevant cues, and processing/deriving appropriate information from these stimuli.[6,7] The perception of biological motion—signals conveyed by the human face, eyes, hands, and body—is an important category within social perception.[8] Previous research indicates that typically developing individuals possess specialized neural systems for processing human motion,[9] which appear impacted in ASD.[8,10] Specifically, posterior superior temporal sulcus (pSTS) is selectively recruited during biological motion perception[11] and has shown reduced response to human motion in ASD.[12–14] Members of our team previously characterized the under-response of this region and others (e.g. amygdala, fusiform gyrus) to human motion as constituting a 'neural signature' of ASD.[13] However, thus far, both this and other functional MRI samples used to examine this question have all skewed predominantly male, ranging from ratios of 4:1 to 7:1,[12–14] raising the possibility that this pattern does not best characterize ASDf.

The female protective effect (FPE) hypothesis suggests several reasons why deviations from neurotypical social perception might, themselves, differ for ASDf versus ASD males (ASDm). This hypothesis posits that greater aetiological load is necessary for ASD expression in girls and women due to female-specific protective factors that lend resilience to the development of an autistic phenotype.[15,16] Such factors are presumed to operate regardless of diagnostic status, also influencing typically developing females (TDf) but not typically developing males (TDm). In ASDf, these factors mean that a larger magnitude aetiological 'push' is required for the autism phenotype to manifest. Thus, a core prediction of the FPE hypothesis is that ASDf will demonstrate greater genetic burden than ASDm, supported by evidence of ASDf carrying larger, more frequent, and more deleterious mutations on average than those found in ASDm.[17–19]

While predictions regarding sex-differential genetic load follow in a comparatively straightforward manner from the FPE hypothesis, what can be expected regarding brain functional differences is less clear. If aetiological burden is assumed to translate in relatively direct fashion into later neuroendophenotype, then ASDf might be expected to exhibit higher magnitude deficits than ASDm in terms of brain response to social perceptual stimuli. Alternatively, greater aetiological load might not necessarily lead to more severe impacts than found in ASDm, as female protective factors might act to dampen the impact of these 'hits'. In the latter case, female and male ASD profiles might not appear significantly different.

Objective differences in ASDf versus ASDm social perceptual response are not the only potential reason why profiles of deviation from sex-typical patterns might diverge between female and male ASD. Typically developing females and males might process social stimuli differently; for example, typically developing females might recruit additional mechanisms related to female resilience factors. Initial neuroimaging

work suggests sex-differential brain response among neurotypical adults viewing human motion.[20,21] Applied to autism, even if ASDf and ASDm displayed an underlying similarity in brain response, this pattern might have a different 'meaning' for females versus males, in that it would indicate a deviation from expected function for their sex.

These considerations suggest that examining how ASDf differ from TDf (as well as how ASDf differ from ASDm, and TDf from TDm) will provide new insights that cannot be extrapolated from existing analyses contrasting male-predominant samples of autistic versus typically developing individuals. However, observed differences in brain response at a single time point (particularly in later development) could be due to genetically-mediated processes operating from very early stages of development and/or to experiences over the lifespan, both of which may differ as a function of sex. Combining genetic and neuroimaging information from the same individual, so that data about brain function constrain genetic analyses, could allow for more informed interpretation of neuroimaging results.

To better understand the ASDf neuroendophenotype of social perception, as well as potential genetic contributions to this profile, we over-sampled ASDf to obtain a well-matched cohort of ASD and typically developing youth at a sex ratio of ~1:1 [the 'GENDAAR' (Gender Exploration of Neurogenetics and Development to Advance Autism Research) project]. Youth (8–17 years old) participated in genotyping, neuroimaging, and behavioural phenotyping. We describe results from functional MRI of brain response while viewing point-light displays[13] of coherent (BIO) or scrambled (SCRAM) whole-body human motion in both ASD and sex-, age-, head motion, and IQ-matched typically developing youth from the GENDAAR cohort ($n = 207$, full functional MRI sample; $n = 184$, matched functional MRI sample; $n = 250$, genotyping sample; Table 1). We note that while Kaiser and colleagues[13] used the same paradigm to assess differences between typically developing and ASD subjects, the present study allows for advantages over the previous work in terms of a nearly 4-fold increase in functional MRI sample size among ASD generally and a more than 9-fold increase in ASDf specifically, allowing us to conduct sex difference analyses not possible previously. Sex was defined as sex assigned at birth; no assessment of gender identity was conducted at this time point. We pursued two primary research objectives.

First (Objective 1), we sought to characterize a functional MRI-based profile of ASDf response to socially meaningful motion, defined as brain regions in which TDf displayed a stronger response to the experimental contrast (BIO > SCRAM) than ASDf. We expected that the resulting map would differ from previous reports of typically developing versus ASD differences in male-predominant samples.[12–14] Given little previous research on this question, we aimed to describe the nature of these differences. To contextualize this profile further, we also examined contrasts for ASDf versus ASDm, TDf versus TDm, and TDm versus ASDm.

Second (Objective 2), we integrated genetics and functional MRI data to explore to what degree the TDf > ASDf profile identified in Objective 1 might reflect, at least in part, an aetiological mechanism relevant specifically to ASD development in females (versus exclusively reflecting correlates of differing social experiences). We tested several complementary questions. First (Objective 2A), we predicted that ASDf versus ASDm would exhibit greater mutation load [defined as larger median size of rare genic copy number variants (CNVs)] in genes expressed both (i) in brain regions represented in the functional MRI-defined TDf > ASDf map; and (ii) during prenatal and infant development ('candidate genes'). Second (Objective 2B), given the novelty of the dataset and the large number of regions represented in the TDf > ASDf map, we conducted exploratory tests to determine which brain region(s) seemed to be driving effects observed in Objective 2A. Given previous work demonstrating greater genetic load in ASDf versus ASDm generally,[17,19] we hypothesized (Objective 2C) that potential genetic impacts to brain regions implicated in the TDf > ASDf map, especially those that seemed to be drivers of the ASDf – ASDm CNV size difference (Objective 2B), would contribute to this load differential more so than genetic impacts broadly defined. Finally, due to the exploratory nature of the tests conducted under Objectives 2B and 2C, we examined whether these findings would replicate (Objective 2D) in an independent ASD cohort.

Materials and methods

Sample

We analysed data from Wave 1 of the GENDAAR project, collected across four sites (Yale University, Harvard University/Boston Children's Hospital, University of California Los Angeles, and University of Washington/Seattle Children's Research Institute). Male and female youth (8–17 years old) were recruited for inclusion in the ASD or typically developing group. Unaffected siblings of ASD participants were also recruited but were not included in Objective 1 analyses given insufficient n's for the planned statistics; however, siblings were genotyped and used to filter out CNVs in Objective 2. Parents provided informed written consent; children provided written assent. Procedures were conducted in compliance with the sites' IRBs and the Declaration of Helsinki. Exclusionary criteria for all groups included: full-scale IQ (FSIQ) $\leqslant 70$ as estimated via the Differential Ability Scales, Second Edition[22] General Conceptual Ability Standard Score; twin status; active tic disorder that would interfere with imaging; pregnancy; metal in the body; active seizures within the past year; or current use of any benzodiazepine, barbiturate, or anti-epileptic (Supplementary material). Other medication use must have been stable for over 6 weeks. The Supplementary material describes group-specific exclusionary criteria.

Diagnostic confirmation

ASD diagnoses were confirmed by expert clinicians using the Autism Diagnostic Observation Schedule, Second Edition

Table 1 Phenotypic and data quality characteristics of the samples, by sex and group

	ASD		Sex Diff. P	Typically developing		Sex Diff. P	Omn. Diff. P	Sig. contr.	Pair. Diff. P
	Female	Male		Female	Male				
Full functional MRI sample (n = 207)									
	n = 46	n = 48		n = 54	n = 59				
Age, m	161.21 (30.24)	157.51 (36.80)	0.595	155.62 (37.37)	163.03 (32.92)	0.267	0.666	–	–
FSIQ	101.61 (20.51)	104.12 (18.09)	0.530	111.13 (14.37)	113.69 (16.20)	0.374	**0.001**	TDf > ASDf	**0.033**
								TDm > ASDm	**0.025**
								TDm > ASDf	**0.003**
SRS-2	93.74 (29.43)	97.49 (26.01)	0.530	16.57 (11.81)	15.71 (11.75)	0.705	**<0.001**	ASD > TD	**<0.001**
ADOS	6.63 (1.74)	7.56 (1.84)	**0.013**	–	–	–	–	–	–
ADI-R: A	18.91 (5.57)	19.44 (5.19)	0.638	–	–	–	–	–	–
ADI-R: B	15.78 (3.87)	16.62 (4.48)	0.332	–	–	–	–	–	–
ADI-R: C	5.83 (2.69)	6.60 (2.64)	0.161	–	–	–	–	–	–
Vol. (n)	152.93 (4.76)	152.65 (5.36)	0.783	153.33 (3.85)	154.00 (0.00)	NA	0.314	–	–
Max aRMS	1.21 (0.98)	1.59 (0.99)	*0.067*	1.04 (0.81)	1.09 (0.96)	0.783	**0.016**	ASDm > TDm	**0.033**
								ASDm > TDf	**0.019**
Max rRMS	0.87 (0.76)	1.08 (0.78)	0.203	0.70 (0.68)	0.76 (0.71)	0.633	*0.050*	ASDm > TDf	**0.046**
MO, n	7.96 (5.15)	8.54 (5.99)	0.612	6.85 (5.21)	7.81 (6.42)	0.382	0.515	–	–
Matched functional MRI sample (n = 184)									
	n = 45	n = 47		n = 45	n = 47				
Age, m	161.95 (30.17)	158.77 (36.14)	0.648	158.25 (35.48)	162.63 (32.54)	0.539	0.898	–	–
FSIQ	102.29 (20.21)	104.45 (18.15)	0.592	108.56 (14.26)	110.06 (15.99)	0.634	0.118	–	–
SRS-2	92.83 (29.17)	98.11 (25.97)	0.379	16.69 (11.23)	16.26 (12.51)	0.861	**<0.001**	ASD > TD	**<0.001**
ADOS	6.58 (1.73)	7.55 (1.86)	**0.011**	–	–	–	–	–	–
ADI-R: A	18.80 (5.57)	19.49 (5.23)	0.543	–	–	–	–	–	–
ADI-R: B	15.64 (3.80)	16.66 (4.53)	0.246	–	–	–	–	–	–
ADI-R: C	5.69 (2.56)	6.62 (2.67)	*0.092*	–	–	–	–	–	–
Vol. (n)	152.91 (4.81)	153.02 (4.73)	0.912	154.00 (0.00)	154.00 (0.00)	NA	0.232	–	–
Max aRMS	1.22 (0.99)	1.55 (0.96)	0.110	1.06 (0.81)	1.25 (1.01)	0.307	*0.093*	ASDm > TDf	*0.064*
Max rRMS	0.87 (0.77)	1.05 (0.76)	0.286	0.68 (0.66)	0.88 (0.74)	0.166	0.192	–	–
MO, n	7.98 (5.20)	8.38 (5.95)	0.729	7.00 (5.24)	8.70 (6.76)	0.179	0.533	–	–
CNV sample (n = 250)									
	n = 61	n = 65		n = 65	n = 59				
Age, m	150.77 (32.36)	148.28 (36.38)	0.685	149.62 (37.64)	157.31 (31.87)	0.223	0.492	–	–
FSIQ	98.43 (18.35)	100.58 (16.46)	0.489	111.74 (14.98)	113.80 (15.97)	0.462	**<0.001**	TD > ASD	**<0.001**
SRS-2	93.70 (27.84)	92.89 (29.00)	0.880	17.11 (11.85)	15.68 (13.08)	0.526	**<0.001**	ASD > TD	**<0.001**
ADOS	6.57 (1.75)	7.31 (1.82)	0.023	–	–	–	–	–	–
ADI-R: A	18.79 (5.57)	20.40 (4.79)	*0.085*	–	–	–	–	–	–
ADI-R: B	15.92 (4.02)	16.97 (4.28)	0.158	–	–	–	–	–	–
ADI-R: C	5.87 (2.53)	6.48 (2.74)	0.198	–	–	–	–	–	–

Values are mean (standard deviation, SD) for male/female. Mean differences among all four groups are calculated by *F*-test and, where *P* < 0.10, followed up by *post hoc* pairwise comparisons with Tukey's HSD test; Welch's *t*-tests on the *post hoc* pairwise comparisons can be found in the Supplementary material. Mean differences within diagnostic category are calculated by *t*-test. *P*-values of *P* < 0.05 and *P* < 0.10 are indicated by bold and italics, respectively. ADI-R = Autism Diagnostic Interview-Revised [A = Diagnostic Algorithm A, 'Qualitative Abnormalities in Reciprocal Social Behavior'; B = (verbal), 'Qualitative Impairments in Communication & Language'; C = 'Restricted, Repetitive Behaviors & Interests']; ADOS = Autism Diagnostic Observation Schedule Second Edition Calibrated Severity Score; Diff. = difference; FSIQ = eestimated full-scale IQ; Max a/rRMS = maximum absolute/relative root mean squared motion; m = months; MO = motion outliers; Omn. = omnibus; Pair. = pairwise; Sig. = significant; SRS-2 = Social Responsiveness Scale version 2 total raw score; Vol. = volumes.

(ADOS-2[23]) using either Module 3 or 4 as deemed most appropriate by the assessing clinician and the Autism Diagnostic Interview-Revised[24] (Supplementary material and Supplementary Table 5).

Full and matched functional MRI samples

Table 1 describes characteristics of the full sample (n = 207; 46 ASDf, 48 ASDm, 54 TDf, 59 TDm) meeting phenotyping and MRI data quality inclusion criteria. The Supplementary

material and Supplementary Table 5 describe within-subject magnetic resonance quality assurance. The full sample of TDf and TDm did not differ on any key metrics (age, FSIQ, and MRI quality; Table 1) and consequently this sample is used in tests of TDf versus TDm functional MRI differences. Matching was implemented to resolve head motion and/or FSIQ differences between groups for other planned functional MRI contrasts (Supplementary material and Supplementary Table 5). Table 1 describes the resulting matched functional MRI sample (n = 184). We note that, within each group, no sex differences emerged in scores on the Child Behavior Checklist (CBCL[25]) empirically-based syndrome or DSM-oriented scales, for either the full or matched functional MRI samples, suggesting relatively similar levels of comorbidity; as expected, ASD had higher CBCL scores than typically developing youth on almost every scale (Supplementary Table 6).

Neuroimaging data collection and analysis

Imaging sites

Initially, all sites scanned on a Siemens 3 T Tim Trio magnet. Partway through data collection, two imaging centres (Sites B and C) upgraded to the Siemens 3 T Prisma Fit. Site was included as a nuisance regressor in all group-level functional MRI analyses, with data acquired pre- and post-upgrade considered to come from different sites. See Supplementary material and Supplementary Table 7 for intersite quality assurance details.

Experimental paradigm

Participants viewed blocks (six per condition, 24 s each) of coherent point-light displays of whole-body human motion (BIO) interleaved with scrambled (SCRAM) versions of these displays over a total of 144 2-s volumes, with 10 additional volumes of final fixation. The BIO displays were initially developed and described by Klin et al.[26] and depict child-friendly movements such as pat-a-cake and waving. As refined by Kaiser and colleagues[13] for functional MRI implementation, the BIO > SCRAM contrast allows for assessment of brain activity associated with biological motion perception, controlling for visual motion perception generally (Supplementary Videos 1 and 2). Participants viewed stimuli via VisuaStim Digital MRI compatible video goggles.

Imaging protocol

Whole-brain images were collected on a Tim Trio with 12-channel head-coil (Sites A.Trio, B.Trio, C.Trio, D.Trio, n = 131) or Prisma Fit with 20-channel head-coil (Sites B.Prisma, C.Prisma, n = 76). See Supplementary material for sequence parameters and Supplementary Fig. 3 for functional coverage.

Image analysis

Neuroimaging analysis used FSL v.5.0.8 (fmrib.ox.ac.uk/fsl) tools within the LONI Pipeline v.6.3 environment[27]; additional analysis used R v.3.4.0-1. See Supplementary material for pre-statistics and individual subject level processing.

Group-level mixed-effects analyses were conducted on the BIO > SCRAM contrast of parameter estimates images (COPEs). The full sample was used to calculate the mean BIO > SCRAM effect within each group (ASDf, ASDm, TDf, TDm; Supplementary material). Group difference analyses tested

pairwise comparisons within sex (e.g. TDf > ASDf) and within group (e.g. ASDf > ASDm). We considered testing a traditional Sex × Diagnosis interaction effect in an analysis of all individuals in the matched functional MRI sample (Table 1). However, because the TDf and ASDm groups demonstrated significant differences in head motion even after matching (Supplementary material), we were concerned that effects detected in such an analysis might prove partially attributable to lower head motion in TDf. We therefore decided that the most conservative option was to ensure that the TDf and ASDm groups were never included in the same model.

The full sample was used for TDf versus TDm comparisons and the matched sample for other comparisons. As recommended by Eklund et al.,[28] these analyses used non-parametric permutation inference, via FSL's Randomise[29] with threshold-free cluster-enhancement (TFCE),[30] 10 000 permutations, corrected $P < 0.05$. Analyses were limited to regions for which the 'greater than' group (e.g. for TDf > TDm, the TDf group) demonstrated a mean group BIO > SCRAM z-value > 0, as we were not interested in relative differences in deactivation. Age (grand-mean centred), estimated FSIQ (grand-mean centred), Social Responsiveness Scale-2 (SRS-2) total raw score (group-mean centred) and site (factor-effect dummy-coded) were included as covariates. For ASDf ⩽ ASDm analyses, ADOS-2 Calibrated Severity Score (grand-mean centred) was included as an additional covariate (Supplementary material).

Genetics analyses

Genotyping

One hundred and twenty-seven ASD individuals (61 female) were genotyped using the HumanOmni2.5M BeadChip (Illumina). All DNA samples were hybridized and scanned on the Illumina iScan to minimize batch effects and variation. All subjects had a genotyping call rate > 95%. Genotyping data were analysed by PLINK v1.07[31] using the forward strand and confirmed the reported sex and sibling relationships of all subjects. One ASDm was removed for failing to pass quality metrics of the CNV detection algorithms, leaving 126 probands. One hundred and thirty typically developing subjects (67 female) were also genotyped. Four TDm genotyping chips failed; two TDf failed to pass CNV quality metrics, leaving 124 typically developing subjects (see Table 1 and Supplementary Table 6 for genotyping sample characteristics). ASDf and ASDm were balanced on race/ethnicity as were TDf and TFm (Supplementary material and Supplementary Fig. 4).

Copy number variant detection

CNV detection (duplications and deletions) was performed using three algorithms: PennCNV v1.0.4, QuantiSNP v1.1, and GNOSIS, as previously described.[32] Analysis and merging of CNV predictions used CNVision.[32] All rare genic CNVs (⩾50% of CNV at ⩽1% frequency in the Database of Genomic Variants[32]; hereafter, CNVs) predicted by at least PennCNV and QuantiSNP and having a CNVision pCNV of ⩽0.001—i.e. those considered high-quality predictions[33]—were obtained for further analysis. For the 37 ASD subjects who had unaffected sibling(s) with available genotyping data (n = 40 siblings), we further specified that rare CNVs in a proband should only include those not shared by their sibling(s). Subsequent analyses combined duplications and deletions to maximize the

number of CNVs available for examination; separating the two types led to low numbers and no significant effects.

Gene expression levels

Gene-level human brain expression data (Platform GPL5175, Affymetrix GeneChip Human Exon 1.0 ST Array), generated as part of the BrainSpan project (hbatlas.org[34]) were downloaded from the NCBI GEO database (accession number GSE25219) as $\log_2$-transformed signal intensity values. Values $\geqslant 6$ in at least one sample are considered positive brain expression.[34] Affymetrix[35] uses background probes with matching guanine-cytosine content for background correction for all array probes.

Creation of spatiotemporal regions of interest and identification of candidate genes

We expected that hyporesponsivity to human motion in some regions represented in the functional MRI-defined TDf > ASDf map might relate not simply to differing social experiences over time, but could instead (or in addition) relate to a possible aetiological mechanism (i.e. mutation load) operating during the developmental window during which ASD is thought to onset, that is, from the prenatal period up to $\sim$2 years.[36] To test this prediction, we used the BrainSpan exon-array transcriptome dataset[34] to categorize CNVs in our sample based on the spatiotemporal expression of gene(s) they contained during human brain development. Related approaches using this resource have been used by a number of groups.[37,38] To create a candidate gene set, we selected genes expressed during early development [10 postconceptual weeks (pcw) to 2 years] in brain regions ($n = 11$) that were significant in the TDf > ASDf functional MRI contrast and for which BrainSpan data were available (Fig. 2A). These 11 regions included: right/left (R/L) dorsolateral prefrontal cortex (DFC), primary motor cortex (M1C), primary somatosensory cortex (S1C), striatum (STR), ventrolateral prefrontal cortex (VFC), and right superior temporal cortex (R-STC) (Supplementary material).

We chose the summed total size of CNVs per individual to represent mutation load since CNV size has been associated with deleteriousness in previous ASD genetics studies,[18,39] and this parameter had the widest data distribution in our cohort. The distribution of number of CNVs or number of genes contained within CNVs per individual was limited in this dataset and did not yield significant differences between ASDf and ASDm (Supplementary Table 3). We did not restrict our analysis to CNVs containing genes previously associated with ASD to avoid removing potentially pathogenic CNVs not yet identified. The Supplementary material describe ASD-associated genes contained by CNVs in GENDAAR and replication cohorts.

Permutation testing strategy

Our sample contained subjects with CNV sizes that were statistically outlying; we retained these subjects for analysis given that these large CNVs might be aetiologically meaningful. We used the median rather than the mean as our summary statistic for CNV size given its greater robustness to outlier influence. The female minus male (F − M) difference (ASDf − ASDm or TDf − TDm) in median summed total CNV size was calculated by first obtaining the median CNV size for each group and then taking the difference between the two medians. All tests were conducted with custom Perl (v.5.24.1) scripts that used 10 000

permutations and $P \leqslant 0.05$, with P-values based on the number of iterations that yielded the actual F − M difference in median CNV size or greater. Fewer occurrences of such a difference led to lower P-values.

Permutation testing was conducted both by permuting the sex label of 'person-sets' of CNVs found within individuals and by permuting the sex label of CNVs independently of individuals. We defined a person-set of CNVs as all CNVs within an individual meeting certain criteria, the sizes of which are summed to give the total CNV size for that individual; person-set criteria varied by test (Supplementary material). We did not have parental information (*de novo* versus inherited) to help determine which of multiple CNVs within an individual may contribute risk to ASD. Therefore, we prioritized permuting by person-sets (versus by CNVs independent of individuals) as this would constitute a more conservative set of tests; we primarily rely on these results in drawing inferences. A given person-set of CNVs was selected only once within an iteration of a permutation test.

We also conducted control analyses to test the significance of the sex difference in median size of CNVs containing gene(s) not expressed in a region of interest (Supplementary material).

Replication in the Simons Simplex Collection

Rare, genic proband CNVs identified in the Simons Simplex Collection (SSC)[33] were analysed and filtered using the same criteria as for our original dataset (Supplementary Table 4C and F). SSC CNVs were included regardless of inheritance status, and we again specified that CNVs in a proband should only include those not shared by their unaffected sibling(s) ($n = 2282$ siblings). Permutation analyses run in the SSC cohort largely paralleled those conducted in the GENDAAR cohort. However, in SSC, the sex ratio is skewed (SSC M: F n's $\sim$3:1; CNVs $\sim$6:1; GENDAAR both $\sim$1:1). Therefore, we performed an additional permutation test controlling for dataset size.

Data availability

Raw and processed data, scripts, and workflows are available for this project through the National Database for Autism Research at ndar.nih.gov, DOI : 10.15154/1478424; NeuroVault (neurovault.org/collections/BZOJTKVI/); and the Open Science Framework (https://osf.io/mgzny) (Supplementary material).

Results

Objective 1: functional MRI

TDf $\leqslant$ TDm, ASDf $\leqslant$ ASDm, TDm $\leqslant$ ASDm, and TDf $\leqslant$ ASDf contrasts were tested. Significant group differences in BIO > SCRAM response were observed for TDf > TDm and TDf > ASDf. All other contrasts were non-significant. For mean within-group effects, see Supplementary material, Supplementary Fig. 1 and Supplementary Table 1.

Typically developing females > ASD females

TDf showed stronger response than ASDf in primarily parietal, posterior temporal, and posterior frontal regions, with
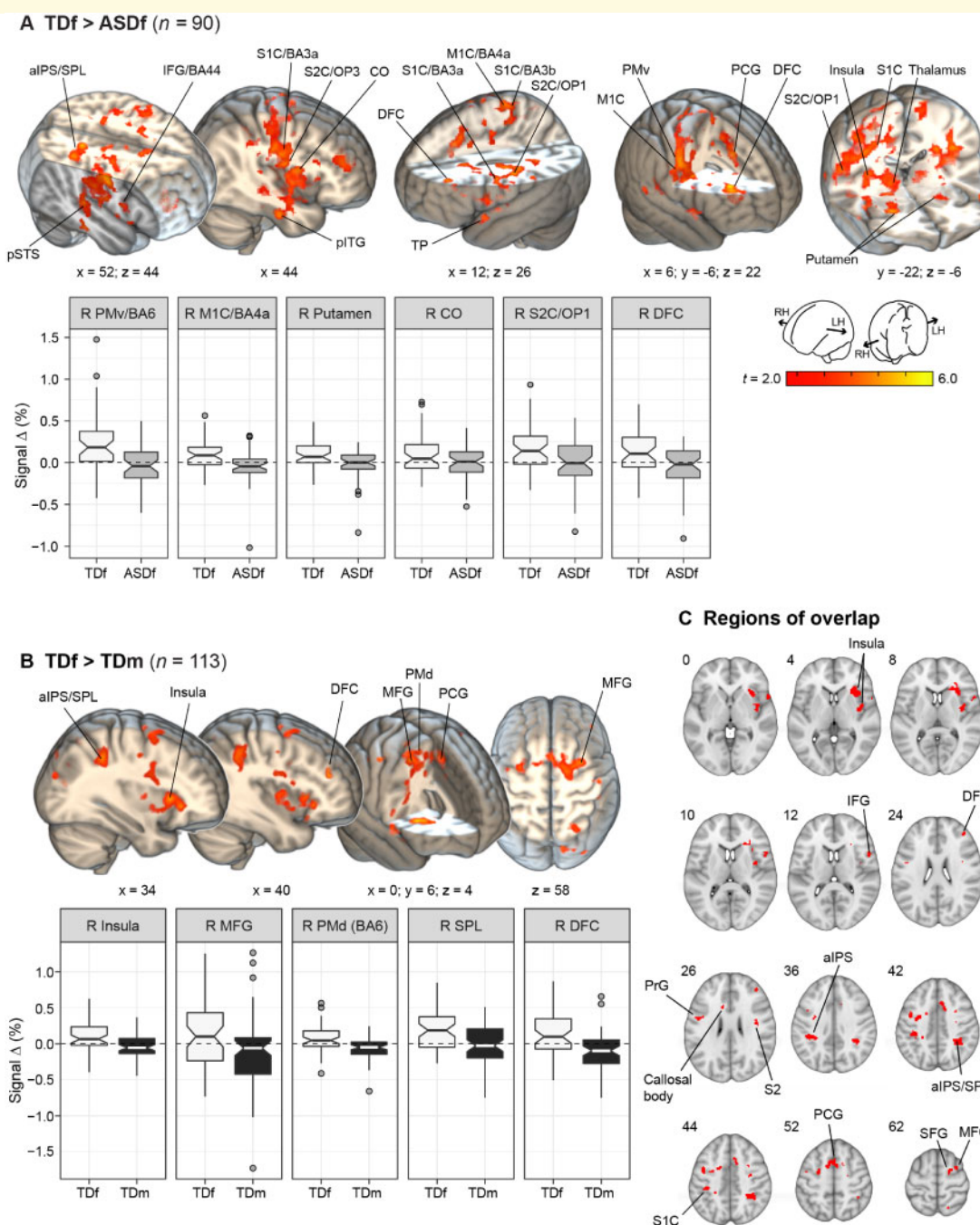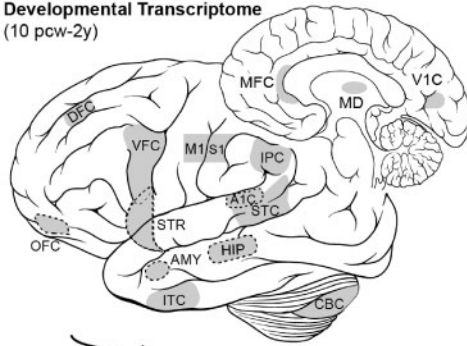
Figure 1 Group differences in response to the BIO > SCRAM contrast, controlling for FSIQ, age, site, and SRS total raw score. *T*-statistic images of regions showing a significantly [permutation-based $P < 0.05$ (corrected) with TFCE, $n_{permutations} = 10\,000$] greater BIO > SCRAM response for one group than another (**A**: TDf > ASDf; **B**: TDf > TDm) are displayed on 3D renderings of the MNI standard brain in neurological convention, with MNI coordinates provided. Box plots depict the distribution of subject-specific per cent signal change (% signal Δ) for the BIO > SCRAM contrast, and are clustered by subject group and panelled by brain region. (**C**) Binary mask indicating regions of overlap between **A** and **B**, displayed as axial slices labelled by MNI *z*-coordinate at *top left*. Box plot elements: centre line = median; notches = confidence interval [±1.58 × IQR/√(n)] around the median; box limits = interquartile range (IQR; 25th to 75th percentile); whiskers = extend to largest/smallest value no further than 1.5 × IQR from box limit; points = outliers. aIPS = anterior intraparietal sulcus; BA = Brodmann area; CO = central operculum; IFG = inferior frontal gyrus; MFG = middle frontal gyrus; PCG = paracingulate; pITG = posterior inferior temporal gyrus; PMd = dorsal premotor cortex; PMv = ventral premotor cortex; PrG = precentral gyrus; pSTS = posterior superior temporal sulcus; R = right; S1C = primary somatosensory cortex; S2C = secondary somatosensory cortex; SFG = superior frontal gyrus; SPL = superior parietal lobule; TP = temporal pole.

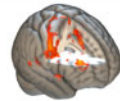**Figure 2 Size of rare CNVs containing candidate genes.** (**A**) Creation of candidate gene-set involved identifying regions that had both been characterized in the BrainSpan Developmental Transcriptome (*left*) and showed a significant TDf > ASDf effect (Fig. 1A); genes were then identified that showed positive expression in these overlapping regions (*right*) between 10 postconceptual weeks (pcw) through to 2 years (y). (**B**) *Left*: Bar plot with standard errors demonstrating the sex difference (ASDf – ASDm) in median total CNV size, in base pairs (bp), within the candidate gene-set. Plots of the CNV size difference are provided for each of the 11 regions of interest. *Right*: Distribution of size of rare CNVs containing gene(s) expressed in right (R) and left (L) striatum (STR) from 10 pcw to 2 years (*right* panels, cyan) by group and sex. For comparison, 'non-STR (R + L)' includes rare CNVs containing gene(s) characterized in BrainSpan that were not positively expressed in R/L-STR from 10 pcw to 2 years (*left* panels, grey). Of note, rare CNVs often contained gene(s) that were expressed in multiple brain regions [e.g. STR plus additional region of interest(s)]. Violin plots depict a Gaussian kernel density estimate, and are overlaid with Tukey-style box plots. (Brain art adapted from illustrations created by Patrick Lynch and C. Carl Jaffe, commons.wikipedia.org/wiki/File:Brain_human_lateral_view.svg and Brain_human_sagittal_section.svg, licensed under a Creative Commons Attribution 2.5 Generic License, 2006).

largest statistical effect sizes (Cohen's $d > 0.90$) observed in right-lateralized ventral premotor cortex (PMv), M1C, parietal white matter, and putamen, and in bilateral DFC and central operculum. Additional regions with a relatively high COPE value (>20) but a medium statistical effect size included right-lateralized secondary somatosensory cortex, superior parietal lobule (SPL), and left temporal pole (Fig. 1A and Table 2).

Table 2 Cluster peaks and local maxima of sites with a significant TDf > ASDf difference in response to BIO > SCRAM

Macroanatomical site	Cyt	Hem	x	y	z	t	d	COPE	P	k
Right fronto-temporo-parietal cluster										12562
PMv/PreCG	BA6	R	54	−6	46	4.34	0.91	26.29	0.005	
M1C/PreCG	BA4a	R	54	−6	40	4.34	0.91	21.79	0.005	
Putamen		R	26	14	−6	4.31	0.91	12.55	0.007	
Parietal WM/SPLA		R	18	−44	44	4.65	0.98	11.86	0.012	
Frontal pole/Frontal WM/DFC		R	34	34	22	4.44	0.94	13.92	0.010	
COper		R	44	2	12	4.32	0.91	11.36	0.009	
PMd/PreCG	BA6	R	30	−18	58	4.20	0.89	12.63	0.011	
PoCG/S1C/Corticospinal tract	BA3b	R	12	−34	70	3.73	0.79	16.59	0.012	
Posterior temporal fusiform cortex		R	40	−14	−28	4.02	0.85	14.66	0.019	
S2C/POper	OP1	R	60	−22	22	3.56	0.75	23.05	0.010	
M1C/PreCG	BA4a	L	−14	−32	64	3.44	0.73	13.87	0.019	
Callosal body		R	16	−30	26	3.85	0.81	10.98	0.017	
Ventroposterior thalamus		R	14	−22	2	3.27	0.69	9.43	0.019	
pSTG/Temporal WM		R	52	−30	8	3.61	0.76	12.44	0.019	
SPL/PoCG/S1C	5L/BA3b	L	−16	−42	64	3.63	0.77	11.15	0.020	
SPL/sLOC	7A	R	16	−64	62	3.26	0.69	35.43	0.026	
Planum polare/Optic radiation		R	40	−4	−18	3.36	0.71	13.43	0.016	
pITG		R	50	−30	−22	3.73	0.79	9.06	0.019	
Left fronto-parietal cluster										8453
M1C/PreCG	BA4p	L	−48	−2	26	4.11	0.87	12.96	0.011	
Anterior corona radiata		L	−26	12	20	4.02	0.85	7.40	0.013	
PMd/SFG	BA6	L	−6	14	54	3.82	0.81	16.79	0.017	
POper/Parietal WM		L	−36	−40	18	4.23	0.89	7.70	0.012	
Paracingulate gyrus		R	6	18	40	4.24	0.89	17.68	0.016	
COper/Broca's area/SLF	BA44	L	−38	6	16	4.34	0.91	10.76	0.014	
Callosal body		L	−12	14	28	3.68	0.78	9.95	0.019	
Insula		L	−38	−10	6	3.86	0.81	10.86	0.016	
Callosal body		L	−24	−52	28	3.79	0.80	7.75	0.016	
Frontal pole/DFC		L	−36	40	24	4.68	0.99	17.15	0.033	
COper/S2C/POper	OP1	L	−40	−22	16	3.48	0.73	13.97	0.016	
Putamen		L	−22	6	−10	3.73	0.79	14.98	0.021	
Anterior supramarginal gyrus/aIPS	hIP2	L	−42	−34	36	3.47	0.73	15.76	0.019	
Broca's area/PreCG	BA44	L	−56	8	10	3.52	0.74	19.79	0.021	
Temporal Pole		L	−58	12	−8	3.35	0.71	25.57	0.030	
SLF		L	−22	−38	42	3.49	0.74	7.15	0.019	
IFG, pars opercularis/Frontal WM		R	40	16	20	2.77	0.58	11.65	0.047	34
Optic radiation		R	40	−44	2	3.37	0.71	6.37	0.046	33
IFG, pars triangularis/Frontal WM		L	−40	28	12	2.69	0.57	9.12	0.049	10
Callosal body		L	−10	−38	18	2.98	0.63	8.38	0.048	4

Coordinates are in MNI space. Statistical inference used 10 000 permutations and TFCE, corrected *P* = 0.05. For table creation purposes only, for clusters k > 1000 voxels, the cluster was further thresholded at *t* > 3.00 and a minimum extent of k = 10 prior to coordinate extraction and atlas query in order to yield more anatomically distinct subclusters. Where this procedure resulted in further clusters where k > 1000, additional local maxima were identified as necessary to describe the anatomical extent of activation. Macroanatomical labels are assigned from the Harvard-Oxford Cortical/Subcortical Structural Atlases[40] and supplemented with macroanatomical and cytoarchitectonic labels from the Juelich Histological Atlas[41]; additional white matter labels are provided by the Mars Parietal connectivity-based parcellation atlas,[42] and the JHU White-Matter Tractography and ICBM-DTI-81 White-Matter Labels Atlases.[43–45] The dorsal/ventral premotor boundary is set at z = 48 per Tomassini *et al.*[46] aIPS = anterior intraparietal sulcus; BA = Brodmann area; COPE = contrast of parameter estimates; COper = central operculum; Cyt = cytoarchitectonic; *d* = Cohen's *d*; IFG = inferior frontal gyrus; Hem = hemisphere; k = voxel extent; L = left; pITG = posterior inferior temporal gyrus; PMd = dorsal premotor cortex; PMv = ventral premotor cortex; PoCG = postcentral gyrus; POper = parietal operculum; PreCG = precentral gyrus; pSTG = posterior superior temporal gyrus; R = right; S1C = primary somatosensory cortex; S2C = secondary somatosensory cortex SLF = superior longitudinal fasciculus; sLOC = superior lateral occipital cortex SPL = superior parietal lobule; *t* = *t*-statistic; WM = white matter. SPLA is a white matter division from the Mars Parietal connectivity-based atlas.

Typically developing females > typically developing males

TDf exhibited increased BIO > SCRAM activation versus TDm in a variety of bilateral parietal and frontal regions.

The largest statistical effect sizes were observed in right-lateralized anterior insula (anterior insula; Cohen's *d* = 0.79) and DFC (Cohen's *d* = 0.84); additional regions with a relatively high COPE value (>20) but a medium statistical effect size

Table 3 Cluster peaks and local maxima of sites with a significant TDf > TDm difference in response to BIO > SCRAM

Macroanatomical site	Cyt	Hem	x	y	z	t	d	COPE	P	k
Bilateral superior frontal cluster										1476
Superior frontal WM		R	18	8	50	3.82	0.72	8.46	0.027	
MFG		R	32	6	58	3.66	0.69	21.65	0.027	
PMd	BA6	R	16	−4	58	3.51	0.66	8.52	0.028	
Corticospinal tract		L	−30	−6	38	3.50	0.66	6.25	0.031	
PMd/MFG	BA6	L	−30	−2	60	3.33	0.63	22.97	0.032	
SLF/S1C	BA3a	R	40	−8	28	3.30	0.62	6.29	0.038	
PMd/SFG	BA6	R	20	6	64	3.20	0.60	14.91	0.032	
Insula		R	34	20	4	4.20	0.79	13.89	0.021	893
Superior parietal WM/SPLC/SPLD		R	16	−60	44	3.63	0.68	10.97	0.036	509
aIPS/SPL	hIP3	R	34	−44	42	3.86	0.73	20.92	0.028	291
aIPS/SLF	hIP1	L	−30	−38	36	3.61	0.68	7.39	0.037	229
PreCG/Corticospinal tract		L	−48	−2	28	3.70	0.70	9.44	0.040	106
Frontal Pole/DFC		R	40	42	24	4.44	0.84	20.39	0.043	29
S1C/PoCG	BA2	R	48	−30	52	3.36	0.63	22.07	0.047	26
M1C/PreCG	BA4a	R	44	−8	48	3.25	0.61	12.97	0.047	17
PoCG		R	52	−10	58	3.66	0.69	10.86	0.046	16
PMv/Corticospinal tract		R	32	−10	44	3.70	0.70	7.75	0.045	13
SPL/Superior parietal WM		L	−24	−50	46	3.40	0.64	13.65	0.048	13
Callosal body		L	−10	16	26	3.65	0.69	8.17	0.047	12
PMd/SFG	BA6	R	26	−4	70	3.00	0.56	17.60	0.048	12
Optic radiation		R	24	−60	32	2.96	0.56	8.82	0.048	8

Coordinates are in MNI space. Statistical inference used 10 000 permutations and TFCE, corrected $P = 0.05$. See notes for Table 2 for details on table creation and atlases used for label assignment. aIPS = anterior intraparietal sulcus; BA = Brodmann area; COPE = contrast of parameter estimates; Cyt = cytoarchitectonic; d = Cohen's d; Hem = hemisphere; k = voxel extent; L = left; MFG = middle frontal gyrus; PMd = dorsal premotor cortex; PMv = ventral premotor cortex; PoCG = postcentral gyrus; PreCG = precentral gyrus; R = right; S1C = primary somatosensory cortex; SFG = superior frontal gyrus; SLF = superior longitudinal fasciculus; SPL = superior parietal lobule; t = t-statistic; WM = white matter; SPLC and SPLD are white matter divisions from the Mars Parietal connectivity-based atlas.

included bilateral middle frontal gyrus (MFG) and right-lateralized anterior intraparietal sulcus (aIPS)/SPL, dorsal premotor cortex (PMd), and S1C (Fig. 1B and Table 3).

Post hoc analyses

Overlap between TDf > ASDf and TDf > TDm

As an exploratory step, we examined overlap between significant regions of TDf > ASDf and TDf > TDm to determine regions that might relate to TDf resilience in social perception. Overlap occurred in frontal and parietal regions including right anterior insula, DFC, inferior frontal gyrus (IFG), and MFG, and in bilateral aIPS and paracingulate (Fig. 1C and Supplementary Table 2).

Overlap between TDf > ASDf and TDm > ASDm

Contrary to expectations, TDm > ASDm did not produce significant results. We considered whether TDf > ASDf and TDm > ASDm might in fact be similar, with a lack of statistical power in the TDm > ASDm contrast preventing us from detecting this similarity. At trend level (corrected $P < 0.10$), the TDm > ASDm contrast generated two small clusters in right corticospinal tract (Supplementary Fig. 2A); overlap between this map and that of TDf > ASDf was

restricted to $k = 10$ voxels (Supplementary Fig. 2B). We also ran both the TDm > ASDm and TDf > ASDf contrasts again using less stringent thresholding procedures commonly used in previous literature (cluster-wise inference with $z > 2.3$ and corrected $P < 0.05$, i.e. FSL's default thresholding prior to April 2017). Using these parameters, TDm > ASDm yielded a single cluster ($k = 386$; MNI peak: $x = 48$, $y = −60$, $z = 8$; Supplementary Fig. 2C) in right posterior caudal STS/middle temporal gyrus. This cluster was located posterior to right pSTS involvement in the TDf > ASDf contrast, with no overlap (Supplementary Fig. 2D).

Objective 2: integration of genetics and functional MRI data

Descriptive characteristics of rare genic CNV size in ASD sample

Table 1 summarizes the genotyping sample. As shown in Table 4, out of 61 ASDf, 116 CNVs in 53 unique individuals were identified, with a median summed total CNV size of 144 378 base pairs (bp) per individual (hereafter, 'median size'); out of 65 ASDm, 114 CNVs in 48 unique individuals were identified, with a median size of 106 740 bp and an

Table 4 Analysis of rare genic CNVs in GENDAAR cohort

ASD sample characteristics	Female		Male		Total individuals with CNVs (n)	F–M difference in median total CNV size (bp)
	Individuals with CNVs (n)	CNVs (n)	Individuals with CNVs (n)	CNVs (n)		
CNVs containing any gene(s)	53	116	48	114	101	37 638
CNVs containing BrainSpan gene(s)	51	88	44	87	95	4350
CNVs containing gene(s) expressed in R/L-STR	40	62	34	48	74	89 999
CNVs containing gene(s) not expressed in R/L-STR	22	26	28	39	50	1002

Permutation tests for ASD sample	Female		Male		Compared to (bp)	P
	Person-sets labelled female (n)	Person-sets sampled from (n)	Person-sets labelled male (n)	Person-sets sampled from (n)		
Sex label shuffling of person-sets of CNVs containing gene(s) expressed in R/L-STR	40	74	34	74	89 999	**0.0293**
Person-sets of all CNVs: sex labels randomized	40	101	34	101	89 999	**0.0178**
Person-sets of all CNVs: sex labels preserved	40	53	34	48	89 999	**0.0172**
Person-sets of CNVs containing BrainSpan gene(s): sex labels randomized	40	95	34	95	89 999	**0.0048**
Person-sets of CNVs containing BrainSpan gene(s): sex labels preserved	40	51	34	44	89 999	**0.0000**
Sex label shuffling of person-sets of CNVs containing gene(s) not expressed in R/L-STR	22	50	28	50	1002	0.4314
Person-sets of CNVs containing gene(s) expressed in R/L-STR: sex labels randomized, sample size reduced to non-R/L-STR size	22	74	28	74	89 999	0.0533[a]
Person-sets of CNVs containing gene(s) expressed in R/L-STR: sex labels preserved, sample size reduced to non-R/L-STR size	22	40	28	34	89 999	0.4955[b]

Typically developing sample characteristics	Female		Male		Total individuals with CNVs (n)	F–M difference in median total CNV size (bp)
	Individuals with CNVs (n)	CNVs (n)	Individuals with CNVs (n)	CNVs (n)		
CNVs containing gene(s) expressed in R/L-STR	35	55	36	70	71	−58 136
CNVs containing gene(s) not expressed in R/L-STR	28	41	25	33	53	1595

Permutation tests for typically developing sample	Female		Male		Compared to (bp)	P
	Person-sets labelled female (n)	Person-sets sampled from (n)	Person-sets labelled male (n)	Person-sets sampled from (n)		
Sex label shuffling of person-sets of CNVs containing gene(s) expressed in R/L-STR	35	71	36	71	−58 136	**0.0779**
Sex label shuffling of person-sets of CNVs containing gene(s) not expressed in R/L-STR	28	53	25	53	1595	0.4131

Tests use 10 000 permutations. P-values indicate how many iterations yielded the actual female minus male (F−M) difference in median total CNV size (value under 'compared to') or greater; $P \leqslant 0.05$ (bold) is a significant difference. CNV = copy number variant; R/L-STR = right and left striatum.
[a,b]Tests intended to determine likelihood that results in full sample are driven by outlier individual(s).
[a]Randomly assigning sex labels of person-sets of CNVs containing R/L-STR, it is still unlikely (5.3%) to obtain a difference ⩾89 999 bp.
[b]Preserving sex labels of person-sets in the reduced sample, a difference ⩾89 999 bp is still likely (49.5%).

overall ASDf – ASDm difference of 37 638 bp. Table 4 summarizes the distribution of subcategories of CNVs among ASDf and ASDm, including CNVs containing gene(s) with available BrainSpan data ('BrainSpan genes', Fig. 2A).

Objective 2A: do females with ASD versus males with ASD show larger median size of CNVs containing candidate genes?

Figure 2B shows the ASDf – ASDm difference in median size of CNVs containing candidate genes. For all 11 brain regions of interest combined, the difference in median size was 81 597 bp (Fig. 2B, solid red line). Shuffling the sex label of person-sets of CNVs containing gene(s) expressed in the combination of all 11 regions of interest 10 000 times produced $P < 0.05$, indicating that this ASDf – ASDm difference in median CNV size was significant (Supplementary Table 4A). By contrast, when comparing TDf to TDm, there was a trend toward greater median CNV size among TDm (TDf – TDm: –63 128 bp; Supplementary Table 4B).

Objective 2B: which regions of interest appear to drive the ASDf – ASDm difference?

As depicted in Fig. 2B, the ASDf – ASDm difference in median size of CNVs containing candidate genes appeared to be driven by two regions of interest in particular: R-STR (83 529 bp) and L-STR (84 184 bp). L-VFC (78 733 bp) and L-M1C (75 922 bp) also showed notable peaks for difference in median size but, unlike R/L-STR, did not exceed the all regions of interest size difference (Fig. 2B, left). Combining data for R/L-STR (two regions of interest) yielded the largest difference (89 999 bp; Table 4). Figure 2B (right) shows the distribution of size of CNVs containing gene(s) expressed in R/L-STR for ASDf and ASDm.

Sex-label shuffling of person-sets of CNVs containing gene(s) expressed in each of the 11 regions of interest produced $P < 0.05$ only for L-M1C, R-STR, L-STR, and the combination of R/L-STR (Table 4 and Supplementary Table 4A). While sex-label shuffling alone suggested that CNVs containing genes expressed in motor or striatal cortex might be driving the effect, further control analyses (eight total, described below under Objective 2C) were only passed unanimously by R/L-STR, whereas L-M1C failed one of the control tests; therefore we focus our attention in the remainder of the text on R/L-STR results. Full results of all control tests on L-M1C (and other regions of interest) are available in Supplementary Table 4.

As for the combination of all 11 regions of interest, there was a trend among TDm (versus TDf) toward greater median CNV size for R/L-STR (TDf – TDm: –58 136 bp) (Table 4, Fig. 2B and Supplementary Table 4B).

Objective 2C: are results attributable to increased CNV size in female ASD more generally?

Sixty-two ASDf and 48 ASDm CNVs (in 40 ASDf and 34 ASDm individuals) contained at least one gene expressed in R- or L-STR. CNVs often contained genes that were expressed in multiple brain regions, such that genes expressed in R/L-STR were not exclusively expressed in these two regions of interest, and some genes were expressed in brain regions other than these two regions of interest. The ASDf – ASDm difference in median size of CNVs containing gene(s) expressed in R/L-STR (89 999 bp) was significant when 40 person-sets of CNVs were labelled female and 34 male 10 000 times after being selected from the set of (i) all person-sets of CNVs containing any gene(s), sex labels randomized; or (ii) preserved; and from (iii) all person-sets of CNVs containing BrainSpan gene(s), sex labels randomized; or (iv) preserved (Table 4). Tests in which CNVs were permuted independently of individuals (rather than as person-sets) were also significant, as were all of the above tests for the combination of all 11 regions of interest (Supplementary Table 4).

To further test the specificity of our findings, we calculated the ASDf – ASDm difference in median size of CNVs containing BrainSpan gene(s) that were not expressed in R/L-STR from 10 pcw to 2 years (non-R/L-STR; Table 4). Twenty-six non-R/L-STR CNVs with a median size of 34 708 bp were identified in 22 ASDf, while 39 non-R/L-STR CNVs with a median size of 33 706 bp were identified in 28 ASDm. The ASDf – ASDm difference (1002 bp) in median size for these CNVs was not significant (Table 4). Additional tests indicated our findings were unlikely to be driven by outlier individuals (Table 4).

Objective 2D: does larger ASDf versus ASDm size of CNVs containing genes expressed during early striatal development replicate in an independent ASD cohort?

We determined the replicability of our results in an independent ASD cohort, the Simons Simplex Collection (SSC[33]) Table 5 shows n's of SSC ASDf and ASDm carrying CNVs containing any gene(s), BrainSpan gene(s), and gene(s) expressed or not expressed in R/L-STR. In the SSC cohort, the ASDf – ASDm difference in median size of CNVs containing gene(s) expressed in R/L-STR was 45 133 bp. Shuffling the sex labels of person-sets of CNVs containing gene(s) expressed in R/L-STR 10 000 times produced $P = 0.0041$, indicating that the ASDf – ASDm difference in median size of CNVs containing genes expressed in R/L-STR was significant in SSC (Table 5); similar control tests to those run in the GENDAAR cohort for Objectives 2A–C were also largely significant (Table 5 and Supplementary Table 4C). As in the GENDAAR cohort, the ASDf – ASDm difference in median size of CNVs containing gene(s) not expressed in R/L-STR (2641 bp) was not significant ($P = 0.2595$; Table 5).

To mimic the ~1:1 sex ratio of CNVs in our cohort (given that the ratio of ASDm to ASDf CNVs in SSC is ~6:1), we randomly selected 40 person-sets of ASDf CNVs and 34 ASDm CNVs 10 000 times from the set of SSC CNVs containing gene(s) expressed in R/L-STR and compared the ASDf – ASDm size difference generated in these permutations to the size difference in the GENDAAR cohort (89 999 bp). Permutation testing yielded $P = 0.1235$, indicating that,

Table 5 Analysis of rare genic CNVs in SSC cohort

ASD sample characteristics	Female		Male		Total individuals with CNVs (n)	F–M difference in median total CNV size (bp)
	Individuals with CNVs (n)	CNVs (n)	Individuals with CNVs (n)	CNVs (n)		
CNVs containing any gene(s)	291	658	1784	3664	2075	27 824
CNVs containing BrainSpan gene(s)	261	498	1573	2807	1834	19 703
CNVs containing gene(s) expressed in R/L-STR	195	313	1212	1784	1407	45 133
CNVs containing gene(s) not expressed in R/L-STR	137	183	819	1023	953	2641

Permutation tests for ASD sample	Female		Male		Compared to (bp)	P
	Person-sets labelled female (n)	Person-sets sampled from (n)	Person-sets labelled male (n)	Person-sets sampled from (n)		
Sex label shuffling of person-sets of CNVs containing gene(s) expressed in R/L-STR	195	1407	1212	1407	45 133	**0.0041**
Person-sets of all CNVs: sex labels randomized	195	2075	1212	2075	45 133	**0.0013**
Person-sets of all CNVs: sex labels preserved	195	291	1212	1784	45 133	**0.1020**
Person-sets of CNVs containing BrainSpan gene(s): sex labels randomized	195	1834	1212	1834	45 133	**0.0013**
Person-sets of CNVs containing BrainSpan gene(s): sex labels preserved	195	261	1212	1573	45 133	**0.0211**
GENDAAR n's and comparison value for SSC person-sets of CNVs containing gene(s) expressed in R/L-STR, sex labels preserved	40	195	34	1212	89 999	0.1235[a]
Sex label shuffling of person-sets of CNVs containing gene(s) not expressed in R/L-STR	137	953	816	953	2641	0.2595

Tests use 10 000 permutations. P-values indicate how many iterations yielded the actual female minus male (F–M) difference in median total CNV size (value under 'compared to') or greater; $P \leqslant 0.05$ (bold) is a significant difference. CNV = copy number variant; R/L-STR = right and left striatum.

[a]Test indicates that when SSC sample size is limited to that of GENDAAR sample, we obtain an ASDf – ASDm difference in median size of CNVs containing gene(s) expressed in R/L-STR equivalent or greater than that found in the GENDAAR sample 12.4% of the time.

in the SSC cohort, the ASDf – ASDm difference in median size of CNVs containing gene(s) expressed in R/L-STR was not significantly different from that of GENDAAR (Table 5).

Discussion

This ASDf-enriched sample has yielded a number of novel insights into female neuro-endophenotypes of social motion perception and potential contributors to female risk for ASD. While functional MRI highlights widespread functional differences between ASDf and TDf viewing human motion, analysis of the size of rare CNVs containing genes expressed in these functional MRI-identified brain regions suggests that potential impacts to striatum may be related to a sex-differential process of risk in early development. These larger ASDf CNVs support the FPE model prediction of greater genetic load in ASDf versus ASDm. Below, we discuss findings related to our major research objectives: (i) characterization of a functional MRI-based profile of ASDf (versus TDf) response to socially meaningful motion; and (ii) integration of functional MRI and genetics data.

First, we observed that the ASDf brain response during human action observation is characterized by less recruitment of parietal and posterior frontal cortex relative to TDf, particularly right somatosensory cortex, motor/premotor areas, and the putaminal region of striatum. This is distinct both from the ASD neural response associated with this paradigm in previous ASDm-predominant literature,[13,14] and from trend-level TDm > ASDm results in this sample, which exhibit minimal overlap with TDf > ASDf. One prominent peak of TDf > ASDf occurred in right PMv, a region putatively associated with 'mirroring' properties,[47,48] and which some suggest may help observers 'fill in' information missing from point-light human motion displays.[49] Somatosensory regions detected in TDf > ASDf also display putative mirroring properties.[50] Thus, greater recruitment of these regions by TDf might imply stronger engagement of such processes. PMv was not represented in BrainSpan, and was thus excluded from our Objective 2 analyses.

To contextualize our TDf > ASDf results, we also analysed differences in response between TDf and TDm, TDm and ASDm, and between ASDf and ASDm. TDf showed increased response to BIO > SCRAM relative to TDm in a variety of frontal and parietal regions. As in the sample of typically developing adults from the study by Anderson and colleagues[20], TDf versus TDm demonstrated greater BIO > SCRAM activation within right DFC, although other regions demonstrating typically developing child (e.g. ventromedial prefrontal cortex) or typically developing adult (e.g. amygdala) sex differences in their cohort did not replicate in our sample, possibly due to differences in the age ranges of our samples. Many of the regions that emerged from our TDf > TDm contrast overlapped with those represented in the TDf > ASDf map, including right-lateralized anterior insula, IFG, DFC, MFG, and bilateral aIPS and paracingulate.

Together, these regions resemble the salience and central executive brain networks. The salience network contains bilateral fronto-insular cortex and dorsal anterior cingulate, and contributes to monitoring and detection of salient stimuli.[51] The central executive network is correlated with right fronto-insular activity and includes DFC, supplementary motor area, and lateral parietal cortices; these systems together play a role in attention, working memory, and cognitive control.[52] The executive and salience sites recruited more strongly by TDf could play a number of roles potentially contributory to resilience in social perception. Right anterior insula contributes to detection of novel salient stimuli[51] and switching between the task-negative (default) and task-positive central executive network[53]; activity in right anterior insula, IFG, and MFG/DFC can indicate renewed attention to a stimulus.[54] These functions suggest more robust attentional reorienting among TDf to the human stimulus after a scrambled block, and/or greater attribution of salience to BIO displays by TDf than either TDm or ASDf.

In previous work examining resting state functional connectivity in our GENDAAR cohort, we found that typically developing youth demonstrated sex differences in functional connectivity of the salience but not the central executive network, while ASD youth showed the opposite pattern, with sex differences in the central executive, but not the salience network.[55] Given our previous results, and the role of the salience network in managing switching to the central executive network,[53] the TDf > ASDf differences we observed in response to social stimuli within nodes of these two networks could be driven by intrinsic neurotypical sex differences in the salience network that are not evident in ASD. Unfortunately, while our present results, and those of our previous resting state work, suggest that anterior insula and aIPS might have relevance to TDf resilience in social perception, these regions were not characterized in BrainSpan, and thus could not be assessed in our Objective 2 analyses.

We did not detect significant differences between ASDf and ASDm in their functional MRI neural response to biological motion. Moreover, contrary to extant literature, ASDm did not differ from TDm on this task. In exploratory follow-up analyses, we considered whether the TDm > ASDm pattern might be similar to that of TDf > ASDf, but below our threshold for statistical detection. Under a more lenient method for statistical inference, ASDm versus TDm displayed right pSTS hypoactivation similar to that found in previous work,[13,14] suggesting that modern methods of functional MRI statistical inference may reduce our power to detect this effect in exchange for greater type I error control. TDf > ASDf did not overlap with TDm > ASDm under this more lenient method. Thus, while ASDf and ASDm response to human motion did not significantly differ, at the same time what distinguishes ASDf from TDf does not appear similar to what distinguishes ASDm from TDm.

While ASDf and ASDm functional brain response did not differ, genetic analyses demonstrated significant differences between these groups. Specifically, ASDf (versus ASDm) exhibited larger size of rare CNVs containing genes

expressed during early development of striatum. This finding, accompanied by ASDf (versus TDf) hypoactivation of putamen (a component of the striatum) during social perception, suggests that potential impacts to striatum may be an element of developmental risk for ASD trajectories in girls. Previously, putaminal disruptions in ASD versus typically developing individuals have been documented,[56–61] albeit largely in ASDm-exclusive or ASDm-predominant samples. We interpret our findings as suggesting that striatal involvement, while not unique to ASDf, may have a particularly important role in ASDf aetiologies. The putamen, historically attributed a primarily motoric role, also appears involved in social and language functions.[62] Among typically developing individuals, the putamen receives projections from motor/premotor (primarily terminating in dorsolateral/central putamen), and prefrontal cortex (primarily terminating in anterior putamen), and appears to serve as an interface between information about motivational value and voluntary behaviour.[63,64] Recent work using resting state functional MRI data suggests that while TDm (females not assessed) demonstrate distinct functional segregation of putamen into anterior and posterior segments, putamen in ASDm appears as one functional unit.[56] In the present investigation, we observed the peak coordinate of TDf > ASDf striatal response in a region of right anterior putamen characterized as having structural connectivity primarily to executive prefrontal regions (including MFG and DFC[65]) It also may be notable that in addition to reduced ASDf response in M1C, we observed larger size of CNVs containing genes expressed in M1C in many (though not all) of our control tests. Taken together, this pattern of results could indicate disturbances to the striatomotor-cortical system more broadly and, thus, processes of linking information about motivational value to action. Differential putaminal recruitment during social perception might reflect differing organization of functional connectivity, in which the region is linked to the central executive network and, perhaps, associated protective functions for TDf but not ASDf. Genetic disruptions specifically impacting striatal cortex during development may underlie such functional atypicalities, and have greater impact via disruption of female protective mechanisms. The general lack of female characterization in the literature on ASD putaminal disruptions, however, makes it difficult to draw strong conclusions along these lines. Future work should analyse ASDf and TDf patterns of functional connectivity and gene co-expression among these regions to clarify this possibility.

When considering together our findings of robust TDf > ASDf and TDf > TDm differences in brain function, lack of ASD sex differences in brain response, and greater ASDf versus ASDm size of CNVs containing genes expressed in early striatal development, the overall picture presented is complex but not inconsistent with an FPE model. While the FPE predicts that ASDf should have greater genetic load than ASDm—a prediction supported by our findings—this does not necessarily equate to greater symptomaticity or disruption of brain function. While some ASDf may lack resilience factors typically found in TDf, other ASDf may retain aspects of female protection that make their phenotype less severe than it might otherwise have been given their greater aetiological load. Moreover, female resilience factors may also have sociocultural aspects (e.g. more emotion-oriented talk to daughters versus sons[66]); the different socialization experiences that an ASDf might encounter could lead, by adolescence, to a brain profile that does not significantly differ from ASDm despite greater genetic load.

In sum, our findings provide new insights into ASDf brain response during social perception, reveal a potential substrate of female risk for ASD trajectories, and illuminate unique qualities of TDf response to human motion relative to TDm. In addition to the basic systems for processing social motion engaged by both sexes, TDf (unique from TDm or ASDf) recruit additional salience and central executive systems. Further, relative to TDf, ASDf show reduced recruitment of striatum during this perceptual task. Compared to ASDm, ASDf (both in our cohort and an independent sample) demonstrate larger size of rare CNVs containing genes expressed in early striatal development, suggesting that, for ASDf, potential impacts to striatum may be particularly relevant. Our results demonstrate the risk of drawing conclusions regarding ASDf based on work comprised of ASDm-predominant samples, and argue for continued attention to the unique characteristics of ASDf.

Acknowledgements

Funding

Competing interests

J.C.M. consults with Blackthorn Therapeutics and has received research funding from Janssen Research and Development; he receives royalties from Guilford Press, Lambert, and Springer.

Supplementary material

Supplementary material is available at *Brain* online.

Appendix I

GENDAAR Consortium Members (Wave 1)

Raphael A. Bernier, PhD; James C. McPartland, PhD; Pamela Ventola, PhD; Anna Kresse; Emily Neuhaus, PhD; Sarah Corrigan; Julie Wolf, PhD; Nicole McDonald, PhD; Katy Ankenman, MS; Sara J. Webb, PhD; Shafali Jeste, PhD; Charles A. Nelson, PhD; Adam Naples, PhD; Erin Libsack; Kevin A. Pelphrey, PhD; Elizabeth Aylward, PhD; Susan Y. Bookheimer, PhD; Nadine Gaab, PhD; Mirella Dapretto, PhD; John D. Van Horn, PhD; Allison Jack, PhD; Desiree Guilford; Carinna Torgerson; Olivia Welker; Daniel H. Geschwind, MD, PhD; Abha R. Gupta, MD, PhD; Catherine A. W. Sullivan, MS; Jennifer K. Lowe, PhD; Zachary Jacokes; Erin MacDonnell; Heidi Tsapelas; Dianna Depedro-Mercier; Cara M. Keifer; Pamela Ventola, PhD.

Roles and affiliations within the GENDAAR consortium, and a description of named author contributions, are described in this project's Open Science Framework site (https://osf.io/mgzny/wiki/0.%20Authors%20%26%20Consortium%20Members/).

References

1. Baio J, Wiggins L, Christensen DL, et al. Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR Surveill Summ*. 2018;67:1–23.

2. Duvekot J, van der Ende J, Verhulst FC, et al. Factors influencing the probability of a diagnosis of autism spectrum disorder in girls versus boys. *Autism*. 2017;21:646–658.

3. Lai M-C, Lombardo MV, Auyeung B, Chakrabarti B, Baron-Cohen S. Sex/gender differences and autism: Setting the scene for future research. *J Am Acad Child Adolesc Psychiatry*. 2015;54:11–24.

4. Lai M-C, Lombardo MV, Ruigrok AV, et al.; MRC AIMS Consortium. Quantifying and exploring camouflaging in men and women with autism. *Autism*. 2017;21:690–702.

5. Loomes R, Hull L, Mandy WPL. What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry*. 2017;56:466–474.

6. Adolphs R. The social brain: Neural basis of social knowledge. *Annu Rev Psychol*. 2009;60:693–716.

7. Klein JT, Shepherd SV, Platt ML. Social attention and the brain. *Curr Biol*. 2009;19:R958–R962.

8. Pavlova MA. Biological motion processing as a hallmark of social cognition. *Cereb Cortex*. 2012;22:981–995.

9. Pelphrey KA, Morris JP. Brain mechanisms for interpreting the actions of others from biological-motion cues. *Curr Dir Psychol Sci*. 2006;15:136–140.

10. Pelphrey KA, Carter EJ. Brain mechanisms for social perception: Lessons from autism and typical development. *Ann N Y Acad Sci*. 2008;1145:283–299.

11. Allison T, Puce A, McCarthy G. Social perception from visual cues: Role of the STS region. *Trends Cogn Sci*. 2000;4:267–278.

12. Freitag CM, Konrad C, Häberlen M, et al. Perception of biological motion in autism spectrum disorders. *Neuropsychologia*. 2008;46:1480–1494.

13. Kaiser MD, Hudac CM, Shultz S, et al. Neural signatures of autism. *Proc Natl Acad Sci U S A*. 2010;107:21223–21228.

14. Koldewyn K, Whitney D, Rivera SM. Neural correlates of coherent and biological motion perception in autism. *Dev Sci*. 2011;14:1075–1088.

15. Robinson EB, Lichtenstein P, Anckarsäter H, Happé F, Ronald A. Examining and interpreting the female protective effect against autistic behavior. *Proc Natl Acad Sci U S A*. 2013;110:5258–5262.

16. Werling DM, Geschwind DH. Sex differences in autism spectrum disorders. *Curr Opin Neurol*. 2013;26:146–153.

17. Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D. Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron*. 2011;70:898–907.

18. Jacquemont S, Coe BP, Hersch M, et al. A higher mutational burden in females supports a 'female protective model' in neurodevelopmental disorders. *Am J Hum Genet*. 2014;94:415–425.

19. Levy D, Ronemus M, Yamrom B, et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*. 2011;70:886–897.

20. Anderson LC, Bolling DZ, Schelinski S, Coffman MC, Pelphrey KA, Kaiser MD. Sex differences in the development of brain mechanisms for processing biological motion. *Neuroimage*. 2013;83:751–760.

21. Pavlova MA, Sokolov AN, Bidet-Ildei C. Sex differences in the neuromagnetic cortical response to biological motion. *Cereb Cortex*. 2015;25:3468–3474.

22. Elliott CD. *DAS introductory and technical handbook*. San Antonio, TX: The Psychological Corporation; 1990.

23. Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop S. *Autism diagnostic observation schedule, second edition (ADOS-2) manual (part I): Modules 1–4*. Torrance, CA: Western Psychological Services; 2012.

24. Lord C, Rutter M, Le Couteur A. Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*. 1994;24:659–685.

25. Achenbach TM, Rescorla LA. *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2001.

26. Klin A, Lin DJ, Gorrindo P, Ramsay G, Jones W. Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*. 2009;459:257–261.

27. Dinov ID, Horn JD, Van Lozev KM, et al. Efficient, distributed and interactive neuroimaging data analysis using the LONI pipeline. *Front Neuroinform*. 2009;3: 22.

28. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A*. 2016;113:7900–7905.

29. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *Neuroimage*. 2014;92:381–397.

30. Smith SM, Nichols TE. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*. 2009;44:83–98.

31. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.

32. Sanders SJ, Ercan-Sencicek AG, Hus V, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*. 2011;70:863–885.

33. Sanders SJ, He X, Willsey AJ, et al.; Autism Sequencing Consortium. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron*. 2015;87:1215–1233.

34. Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. *Nature*. 2011;478:483–489.

35. Affymetrix. Exon Array Background Correction [Internet]. Affymetrix GeneChip®Exon Array Whitepaper Collection; 2005. https://assets.thermofisher.com/TFS-Assets/LSG/brochures/exon_background_correction_whitepaper.pdf. (25 June 2020, date last accessed)

36. Landa RJ, Holman KC, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Arch Gen Psychiatry.* 2007;64: 853–864.

37. Lin GN, Corominas R, Lemmens I, et al. Spatiotemporal 16p11.2 protein network implicates cortical late mid-fetal brain development and KCTD13-Cul3-RhoA pathway in psychiatric diseases. *Neuron.* 2015;85:742–754.

38. Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell.* 2013;155:997–1007.

39. Pinto D, Delaby E, Merico D, et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet.* 2014;94:677–694.

40. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage.* 2006;31: 968–980.

41. Eickhoff SB, Stephan KE, Mohlberg H, et al. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage.* 2005;25:1325–1335.

42. Mars RB, Jbabdi S, Sallet J, et al. Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *J Neurosci.* 2011;31:4087–4100.

43. Hua K, Zhang J, Wakana S, et al. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *Neuroimage.* 2008;39:336–347.

44. Mori S, Crain BJ. *MRI atlas of human white matter.* Amsterdam, The Netherlands: Elsevier; 2005.

45. Wakana S, Caprihan A, Panzenboeck MM, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimage.* 2007;36:630–644.

46. Tomassini V, Jbabdi S, Klein JC, et al. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *J Neurosci.* 2007;27:10259–10269.

47. Molenberghs P, Cunnington R, Mattingley JB. Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neurosci Biobehav Rev.* 2012;36:341–349.

48. Rizzolatti G, Fadiga L, Gallese V, Fogassi L. Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res.* 1996;3:131–141.

49. Saygin AP, Wilson SM, Hagler DJ, Bates E, Sereno MI. Point-light biological motion perception activates human premotor cortex. *J Neurosci.* 2004;24:6181–6188.

50. Keysers C, Kaas JH, Gazzola V. Somatosensation in social perception. *Nat Rev Neurosci.* 2010;11:417–428.

51. Menon V, Uddin LQ. Saliency, switching, attention and control: A network model of insula function. *Brain Struct Funct.* 2010;214:655–667.

52. Seeley WW, Menon V, Schatzberg AF, et al. Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci.* 2007;27:2349–2356.

53. Goulden N, Khusnulina A, Davis NJ, et al. The salience network is responsible for switching between the default mode network and the central executive network: Replication from DCM. *Neuroimage.* 2014;99:180–190.

54. Weissman DH, Roberts KC, Visscher KM, Woldorff MG. The neural bases of momentary lapses in attention. *Nat Neurosci.* 2006;9:971–978.

55. Lawrence KE, Hernandez LM, Bowman HC, et al. Sex differences in functional connectivity of the salience, default mode, and central executive networks in youth with ASD. *Cereb Cortex.* 2020;30: 5107–5120.

56. Balsters JH, Mantini D, Wenderoth N. Connectivity-based parcellation reveals distinct cortico-striatal connectivity fingerprints in autism spectrum disorder. *Neuroimage.* 2018;170:412–423.

57. Damiano CR, Cockrell DC, Dunlap K, et al. Neural mechanisms of negative reinforcement in children and adolescents with autism spectrum disorders. *J Neurodev Disord.* 2015;7: 12.

58. Doyle-Thomas KAR, Card D, Soorya LV, Ting Wang A, Fan J, Anagnostou E. Metabolic mapping of deep brain structures and associations with symptomatology in autism spectrum disorders. *Res Autism Spectr Disord.* 2014;8:44–51.

59. Di Martino A, Kelly C, Grzadzinski R, et al. Aberrant striatal functional connectivity in children with autism. *Biol Psychiatry.* 2011;69:847–856.

60. van Rooij D, Anagnostou E, Arango C, et al. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the ENIGMA ASD working group. *Am J Psychiatry.* 2018;175:359–369.

61. Wegiel J, Flory M, Kuchna I, et al. Stereological study of the neuronal number and volume of 38 brain subdivisions of subjects diagnosed with autism reveals significant alterations restricted to the striatum, amygdala and cerebellum. *Acta Neuropathol Commun.* 2014;2:141.

62. Pauli WM, O'Reilly RC, Yarkoni T, Wager TD. Regional specialization within the human striatum for diverse psychological functions. *Proc Natl Acad Sci U S A.* 2016;113:1907–1912.

63. Báez-Mendoza R, Schultz W. The role of the striatum in social behavior. *Front Neurosci.* 2013;7: 233.

64. Haber SN. Corticostriatal circuitry. *Dialogues Clin Neurosci.* 2016;18:7–21.

65. Tziortzi AC, Haber SN, Searle GE, et al. Connectivity-based functional analysis of dopamine release in the striatum using diffusion-weighted MRI and positron emission tomography. *Cereb Cortex.* 2014;24:1165–1177.

66. Aznar A, Tenenbaum HR. Gender and age differences in parent-child emotion talk. *Br J Dev Psychol.* 2015;33:148–155.

Supplementary Information for

A Neurogenetic analysis of female autism

Allison Jack, Catherine A.W. Sullivan, Elizabeth Aylward, Susan Y. Bookheimer, Mirella Dapretto, Nadine Gaab, John D. Van Horn, Jeffrey Eilbott, Zachary Jacokes, Carinna M. Torgerson, Raphael A. Bernier, Daniel H. Geschwind, James C. McPartland, Charles A. Nelson, Sara J. Webb, Kevin A. Pelphrey, Abha R. Gupta, and the GENDAAR Consortium.

Corresponding Authors:

Allison Jack
Email: ajack@gmu.edu

Abha R. Gupta
Email: abha.gupta@yale.edu

Supplementary Methods

Sample

1.1 List of excluded medications

Benzodiazepines
- a. Alprazolam (e.g., Xanax)
- b. Chlordiazepoxide (e.g., Librium)
- c. Clonazepam (e.g., Klonopin)
- d. Clorazepate (e.g., Tranxene)
- e. Diazepam (e.g., Valium)
- f. Estazolam (e.g., Prosom)
- g. Flurazepam (e.g., Dalmane)
- h. Lorazepam (e.g., Ativan)
- i. Midazolam
- j. Oxazepam (e.g., Serax)
- k. Quazepam (e.g., Doral)
- l. Temazepam (e.g., Restoril)
- m. Triazolam (e.g., Halcion)

Barbituates
- n. Allobarbital
- o. Alphenal
- p. Amobarbital (e.g., Amytal)
- q. Aprobarbital
- r. Barbital
- s. Brallobarbital
- t. Butabarbital (e.g., Butisol)
- u. Pentobarbital (e.g. Nembutal)
- v. Phenobarbital (e.g., Luminal)
- w. Primidone
- x. Secobarbital (e.g., Seconal)

Anti-epileptics
- a. Carbamazepine
- b. Valproic acid

1.2 Group-specific exclusionary criteria

Typically Developing (TD). Additional exclusionary criteria specific to the TD group were as follows: diagnosed, referred, or suspected ASD, schizophrenia, intellectual disability, learning disability, or other developmental or psychiatric disorder; a first- or second-degree relative with ASD; a total *t*-score > 60 on the Social Responsiveness Scale, Second Edition (Constantino and Gruber, 2012); a raw score > 11 on the Lifetime version of the Social Communication Questionnaire (Rutter *et al.*, 2003); clinical impression suggesting ASD, other developmental delay or disorder, broader autism phenotype, or significant psychiatric disorder.

Autism Spectrum Disorder (ASD). Additional exclusionary criteria specific to the ASD group were as follows: known single gene disorder related to ASD or syndromic form of ASD (e.g., Fragile X); medical conditions likely to be etiological (e.g. focal epilepsy or infantile spasms); any neurological disorder involving pathology above the brainstem, other than uncomplicated non-focal epilepsy; history of significant pre- or perinatal injury, i.e. birth at < 36 weeks *and* weight < 2000 grams, *or* neonatal intensive care unit hospital stay > 3 days; history of neonatal brain damage; any known environmental circumstances that might account for the picture of ASD in the proband (e.g., severe nutritional or psychological deprivation); clinically significant visual or auditory impairment after correction; or any sensorimotor difficulties that would preclude valid use of the diagnostic instruments.

1.3 Diagnostic confirmation

Children and adolescents who received the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Module 3 (for which revised algorithms were available from the outset of the study) were required to achieve a Calibrated Severity Score (CSS) of ≥ 4; adolescents who received Module 4 (for which a revised algorithm became available during the course of data collection (Hus and Lord, 2014)) were required to meet ASD criteria according to *either* the updated algorithm (CSS ≥ 4) *or* the version of the algorithm published with the ADOS-2 (Communication + Social Interaction Total ≥ 7). Scores on the Autism Diagnostic Interview-Revised (ADI-R), a standardized parent interview designed to obtain ASD symptom information, were required to meet the diagnostic algorithm within one point.

FMRI Analyses

1.4 Within-subject MR quality assurance

Participants who met phenotyping criteria were then passed to within-subject MR quality assurance (QA). MRI and fMRI data from participants who met inclusion criteria was reviewed to determine that head motion was within acceptable parameters, that functional slice prescription was correct, and that the anatomical scan was of adequate quality for accurate coregistration to be performed. With regards to head motion during the functional scan, exclusionary criteria were absolute or relative root mean squared (RMS) motion > 4 mm and/or ≥ 20% of timepoints identified as outliers according to Power and colleagues' (Power *et al.*, 2012) DVARS metric. If motion above the RMS threshold(s) occurred within the final 20% of volumes (i.e., between volumes 124 and 154), the tail of the dataset was trimmed to retain the maximal amount of data possible while adhering to motion criteria; otherwise, the dataset was excluded. See Supplementary Table 5, stages 3-5. Visual inspection of the head motion distribution revealed that only three participants (2 ASDm, 1 TDm) displayed both absolute and relative RMS motion > 3mm; they were removed from the dataset as outliers. All participants who passed this stage of QA (*N* = 207) were included in analyses of mean within-group effects and are referred to as the "full fMRI sample." A summary of head motion statistics and the number of volumes retained in the full fMRI sample can be found in Table 1A.

1.5 FMRI sample matching

Statistically significant or trend-level ASD versus TD differences in key metrics (age, cognitive ability, and MRI data quality) were detected in the full fMRI sample. These differences were driven by greater root mean squared (RMS) head motion among ASDm than TDm (or TDf) and higher estimated FSIQ among TD than ASD across both sexes (Table 1A). Therefore, we implemented sample matching wherein controls were selected to match to probands on variables demonstrating significant differences in the full sample, within sex.

Sample matching used MatchIt v. 3.0.1 (Ho *et al.*, 2007) within R v.3.4.1. This procedure involved optimal matching (Hansen, 2004) using the propensity score from a logistic regression, where controls were selected to match to probands on the variables demonstrating significant differences in the full sample, within sex. Prior to matching, Welch's *t*-test indicated greater head motion among ASDm than TDm (max absolute RMS: $t(99.44) = 2.64$, $p = .010$; max relative RMS: $t(96.45) = 2.19$, $p = .031$) and higher estimated FSIQ among TD than ASD across both sexes (female: $t(78.79) = -2.64$, $p = .010$; male: $t(95.42) = -2.85$, $p = .005$). Thus, for girls, matching was based on estimated FSIQ; for boys, matching was based on estimated FSIQ, and maximum relative and absolute RMS values. One ASDm and one ASDf participant could not be adequately matched and were discarded. See Supplementary Table 5:vi-viii.

After matching, differences between ASDm and TDm on maximum absolute ($t(91.83) = 1.46$, $p = .147$) and relative ($t(91.95) = 1.06$, $p = .294$) RMS values were no longer present. ASD vs. TD differences in estimated FSIQ were no longer significant for boys ($t(90.57) = -1.59$, $p = .115$) or girls ($t(79.09) = -1.70$, $p = .093$). We note that while the *p*-value approaches significance for girls,

the mean estimated FSIQ score for both groups was within the average range ("Average" classification: 90-109; ASDf M = 102.29; TDf M = 108.56); further, estimated FSIQ scores were included as a covariate in all group-level analyses.

In the full sample, ASDf and ASDm exhibited a trend-level difference in maximum absolute RMS motion that was resolved in the only slightly smaller matched sample, leading us to conduct contrasts between ASDm and ASDf on the matched sample.

Within the matched sample (Table 1B), ASDf and ASDm did not differ in terms of parent-reported symptom history of social behavioral or communication/language symptoms as measured via the ADI-R; they also did not differ in terms of SRS-2 raw scores. However, on average, ASDm scored approximately one point higher than did ASDf on clinician assessment of current symptom severity as calculated via the Calibrated Severity Score (CSS) of the ADOS-2 (Table 1B). Due to this difference, which could not be resolved through matching without too great a loss of sample size, ADOS-2 CSS score was included as a covariate in all fMRI analyses exclusive to ASDf/m. We note that this variable was uncorrelated with SRS-2 raw score (r(84) = 0.04, p = .682), which was also included as a regressor in all fMRI analyses in order to control for individual variability in autistic-like characteristics.

Finally, we note that in the matched sample, significant head motion differences did remain between ASDm and TDf (max absolute RMS: t(88.46) = 2.66, p = .009; max relative RMS: t(89.14) = 2.48, p = .015); thus these groups were not included together in any fMRI statistical model.

1.6 Inter-site MR quality assurance

Previous work on multisite neuroimaging design indicates that the blood oxygen level dependent (BOLD) signal obtained from fMRI experimental paradigms generally replicates well, showing greater variability across persons than sites, with the risk incurred by aggregating multisite data generally restricted to a modest reduction of statistical power (Gee *et al.*, 2015). See Supplementary Table 7A for n's scanned at each scanner-site, by diagnosis and sex.

To assess inter-site differences in MR scan quality, the full imaging protocol was run on both a spherical agar-filled phantom shipped to all sites and a human phantom who traveled to each site (with the exception of Site C.Prisma due to a change in health status precluding MRI). An additional human phantom was scanned at two sites (Sites A & D). Phantoms were scanned at Sites B & C both before and after the Prisma upgrade. To assess whether functional scan quality varied among the six sites, temporal signal to noise ratio (tSNR) values were calculated for fMRI scans of the phantoms using the automated LONI Neuroimaging Quality Control System (qc.loni.usc.edu), which implements tSNR calculations via the AFNI (afni.nimh.nih.gov) tool afni_restproc.py. A summary of tSNR values by site and phantom type can be found in Supplementary Table 7B. These values were entered as the dependent variable in a linear mixed model (fitted using nlme (Pinheiro *et al.*, 2017) v. 3.1-128 within R) with site as a fixed effect and subject (human 1, human 2, or spherical phantom) as a random effect within site. An ANOVA on this model indicated that tSNR values differed significantly among the sites (F(5,48) = 3.11, p = .016); post-hoc testing using Tukey's HSD (see Supplementary Table 7C) indicated that Site A and Site D significantly differed in their tSNR values, with an estimated Site D-Site A mean difference in tSNR of 22.48 (SE = 6.34), p = .005. Given the inter-site difference in tSNR, site was included as a nuisance regressor in all group-level fMRI analyses.

1.7 Imaging parameters

We acquired whole-brain T1-weighted anatomical images using the Siemens MPRAGE (magnetization-prepared rapid gradient-echo sequence); whole-brain T2-weighted in-plane structural images (for improved coregistration between BOLD functional and T1-weighted anatomical images) using a Siemens segmented spin echo EPI (echo planar imaging) sequence; and whole-brain functional images using a Siemens single-shot gradient echo EPI sequence sensitive to the BOLD contrast, as follows:

T1-weighted anatomical image: TR = 2530 ms; $TE_{Tim\ Trio}$ = 3.31 ms/$TE_{Prisma\ Fit}$ = 3.34 ms; flip angle = 7°; FOV = 256 mm; image matrix = 256 mm^2; voxel size = 1 mm^3; 176 slices; NEX = 1.

T2-weighted in-plane structural image: TR = 5000 ms; TE$_{Tim Trio}$ = 34 ms/TE$_{Prisma Fit}$ = 35 ms; flip angle = 90°, FOV = 192 mm; image matrix = 128 mm^2; voxel size = 1.5 mm × 1.5 mm × 4 mm; 34 slices, NEX = 4.

BOLD image: TR = 2000 ms; TE = 30 ms; flip angle = 90°; FOV = 192 mm; image matrix = 64 mm^2; voxel size = 3 mm × 3 mm × 4 mm; 34 slices; NEX = 1.

Four functional volumes were discarded prior to initiating the experimental paradigm to allow for magnetic saturation effects; thereafter 154 volumes were acquired.

Slice prescription for both the in-plane structural and the BOLD functional images was identical and used an oblique prescription to capture the maximal amount of cerebrum and cerebellum while reducing artifact contamination in inferior structures. See Supplementary Figure 3 for a summary image of functional coverage.

1.8 Preprocessing & first-level analysis

Pre-statistics processing. The following pre-statistics processing of functional data was applied: motion correction using MCFLIRT (Jenkinson *et al.*, 2002); slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5mm; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; high pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=50.0s).

Registration. Linear registration of the functional image to the T2-weighted in-plane structural image used FLIRT with 6 degrees of freedom. Thereafter, registration to the high-resolution T1-weighted anatomical image used boundary-based registration (Greve and Fischl, 2009). Registration to the Montreal Neurological Institute (MNI) standard space image was carried out using FLIRT (Jenkinson and Smith, 2001; Jenkinson *et al.*, 2002) and then refined using FNIRT nonlinear registration (Andersson *et al.*, 2007) with 12 degrees of freedom and a warp resolution of 10 mm. Registration results were visually inspected with a focus on the goodness of fit of the raw functional image to the standard space template and, when necessary, individual registration parameters were adjusted to achieve optimal results.

Individual subject level analysis. At the individual subject level, preprocessed data were submitted to fixed-effects analyses to generate subject-specific estimates of whole-brain response to the BIO > SCRAM contrast. The analysis was conducted using FSL's FEAT v.6.00, with time-series statistical analysis carried out using FILM with local autocorrelation correction (Woolrich *et al.*, 2001). BIO and SCRAM time courses were included as regressors in this analysis, convolved with a gamma hemodynamic response function (HRF), with temporal filtering applied and a temporal derivative added. Nuisance regressors at this level were included to account for subject motion: a confound matrix identifying outlier timepoints according to the DVARS metric (Power *et al.*, 2012) as well as six standard motion parameters. The contrast of interest was BIO > SCRAM.

1.9 Methods for within-group analysis of BIO > SCRAM

Analyses were carried out at the whole-brain level within FEAT, using FLAME (Beckmann *et al.*, 2003; Woolrich *et al.*, 2004; Woolrich, 2008) stages 1 and 2 with automatic outlier detection and deweighting (Woolrich, 2008), with cluster-corrected thresholding using traditional Gaussian Random Field Theory (RFT; Worsley, 2007) at $z > 3.10$ and $p = .05$. While inflated familywise error (FWE) rates are a risk with clusterwise inference, nonparametric permutation-based approaches to inference can exhibit invalid behavior in the case of the one-sample *t*-test (Eklund *et al.*, 2016); further, Kessler and colleagues (2017) have demonstrated that a cluster-determining threshold of $p = .001$ (i.e., $z = 3.10$) with RFT-FWE correction identifies clusters that, in nearly all cases, are also categorized as significant when using a nonparametric false discovery rate benchmark of α = .05. Given the risks associated with a sign-flipping permutation test, we elected to use cluster-based thresholding with the recommended cluster-determining threshold (Kessler *et al.*, 2017) in this case. Age, estimated FSIQ, and SRS-2 total raw score (all group-mean centered) were included as covariates in each analysis along with site (factor-effect dummy-coded with $n = 5$ regressors); for ASDf and ASDm, group-mean centered ADOS-2 CSS score was also added as a covariate.

Genetic Analyses

1.10 Ancestry mapping

Eigensoft (github.com/DReichLab/EIG) was used to compare SNP genotypes of probands in our sample to individuals of known ancestry in HapMap3 (broadinstitute.org/medical-and-population-genetics/hapmap-3). 117,471 SNPs from the HumanOmni2.5M BeadChip were included by meeting the following criteria: 1) minor allele frequency > 5%, 2) not in significant linkage disequilibrium with other SNPs analyzed, 3) satisfied Hardy-Weinberg equilibrium ($p <$ .001), and 4) maintained 95% genotype calling in all samples. Smartpca and Eigenstrat within Eigensoft were utilized to perform principal component analysis (PCA) with the HapMap3 individuals and subjects in our sample. Eigenvalues of the first two principal components, which contributed the greatest amount of variation relative to the other principal components, were plotted against each other (Supplementary Fig. 4B). PCA correctly distinguished the ancestry groups of HapMap3 and confirmed the self-reported race and ethnicity of the subjects in this study. ASDf and ASDm were balanced in terms of race/ethnicity as were TDf and TDm (Supplementary Fig. 4A). Of note, 4 ASDf and 6 ASDm from the Simons Simplex Collection (SSC) participated in the GENDAAR study and were not included in the genetics analysis of SSC subjects described below. Their genotyping data were not available to us to include in ancestry mapping, but ASDf and ASDm remained balanced when accounting for their self-reported race/ethnicity.

1.11 Creation of spatio-temporal ROIs and identification of candidate genes.

Custom Perl (v.5.24.1) scripts were used to identify CNVs containing candidate genes. We defined candidate genes based on both their temporal and regional patterns of expression (Fig 2A). *Temporally*, we used genes expressed during the developmental time period that ASD is thought to onset. Specifically, we used BrainSpan (Kang *et al.*, 2011) developmental time periods three ("Early Fetal," 10-13 pcw; the earliest BrainSpan time period during which our brain ROIs were sufficiently differentiated to be analyzed) through 2y within period 10 ("Early Childhood", 1-6y; greater than 2y excluded). *Regionally*, we used genes expressed during these periods within 11 brain regions that were both represented in the TDf > ASDf fMRI contrast and for which transcriptome data was available from BrainSpan (Fig. 2A). These included the following BrainSpan regions: R/L-DFC, R/L-M1C, R/L-S1C, R/L-STR, R/L-VFC, and R-STC. Overlap was determined by Kang et al.'s description of their anatomical definition of these regions, provided in the supplement to their report (2011). We note in particular the extent of tissue sampled in BrainSpan ROIs M1C and STR (which yielded significant or trending ASDf-ASDm results): for R/L-M1C, the sampled area comprised ventrolateral precentral gyrus (BA 4); for R/L-STR, the sampled area comprised anterior striatum including head of caudate and putamen, internal capsule, and nucleus accumbens(Kang *et al.*, 2011). In addition, BrainSpan R/L-VFC is equivalent to inferior frontal gyrus pars opercularis and pars triangularis (BA 44/45; Kang *et al.*, 2011).

1.12 Explanation of all permutation tests

Sex label shuffling of person-sets of CNVs containing gene(s) expressed in the ROI. See Supplementary Table 4A-C, column 3. In this analysis, person-sets of CNVs containing gene(s) expressed in the ROI have their sex labels randomly shuffled. These person-sets consist *only* of CNVs containing gene(s) expressed in the ROI; they do not include any other CNVs that individual may carry. Specifically, to conduct this analysis, the number of individuals carrying at least one CNV containing at least one gene expressed in the ROI is obtained for each sex (e.g., 40 ASDf and 34 ASDm carrying at least one CNV containing at least one gene expressed in R or L-STR). Then, for each permutation, from all person-sets of CNVs containing gene(s) expressed within the ROI, a number of person-sets equivalent to the original female *n* (e.g., 40) are randomly selected and labeled as female, and likewise a number of person-sets equivalent to the original male *n* (e.g.,

34) are randomly selected and labeled as male. The new F-M difference in median CNV size is then calculated.

Person-sets of all CNVs: sex labels randomized. See Supplementary Table 4A-C, column 4. In this test, permutations are conducted in which *all* person-sets of CNVs containing *any* gene(s) have their sex labels randomly assigned in ratios equivalent to the original sex ratio for the ROI. Specifically, from the set of *all* individuals of either sex carrying *any* genic CNV(s), a number of individuals equivalent to the female *n* carrying CNVs containing gene(s) expressed in the ROI are randomly selected and labeled as female, and likewise a number of individuals equivalent to the male *n* carrying CNVs containing gene(s) expressed in the ROI are randomly selected and labeled as male. The new F-M difference in median CNV size is then calculated.

Person-sets of all CNVs: sex labels preserved. See Supplementary Table 4A-C, column 5. In this test, we sample from *all* person-sets of CNVs containing *any* gene(s), using ratios equivalent to the original sex ratio for the ROI, but we preserve the original sex label of each person-set. Specifically, for each permutation, from all female person-sets with CNVs containing *any* gene(s), a number of person-sets is selected equivalent to the female *n* carrying CNVs containing gene(s) expressed in the ROI. Likewise, from all male person-sets with CNVs containing *any* gene(s), a number of person-sets is selected equivalent to the male *n* carrying CNVs containing gene(s) expressed in the ROI. The new F-M difference in median CNV size is then calculated.

Person-sets of CNVs containing BrainSpan gene(s): sex labels randomized. See Supplementary Table 4A-C, column 6. In this analysis, we sample from person-sets of CNVs that contain gene(s) characterized in BrainSpan ("BrainSpan genes"), and sex labels are randomly assigned to these person-sets in ratios equivalent to the original sex ratio for the ROI. Specifically, for a given individual, CNVs containing at least one BrainSpan gene are included in their person-set, while CNVs containing only gene(s) not characterized in BrainSpan are excluded. For each permutation, a number of person-sets equivalent to the female *n* carrying CNVs containing gene(s) expressed in the ROI are randomly selected and labeled as female, and likewise, a number of person-sets equivalent to the male *n* carrying CNVs containing gene(s) expressed in the ROI are randomly selected and labeled as male. The new F-M difference in median CNV size is then calculated.

Person-sets of CNVs containing BrainSpan gene(s): sex labels preserved. See Supplementary Table 4A-C, column 7. In this analysis, we sample from person-sets of CNVs that contain BrainSpan gene(s) in ratios equivalent to the original sex ratio for the ROI, preserving the original sex labels of those person-sets. Specifically, for each permutation, from all female person-sets of CNVs containing BrainSpan gene(s), a number of person-sets equivalent to the female *n* carrying CNVs containing gene(s) expressed in the ROI are randomly selected, and likewise, from all male person-sets of CNVs containing BrainSpan gene(s), a number of person-sets equivalent to the male *n* carrying CNVs containing gene(s) expressed in the ROI is selected. The new F-M difference in median CNV size is then calculated.

Testing F-M difference in median CNV size containing gene(s) expressed in the ROI: permuting by CNVs independently of individuals. We also performed parallel versions of all tests described above, in which we permuted CNVs meeting particular criteria (e.g., CNVs containing any gene[s], CNVs containing BrainSpan gene[s]) independently of individuals rather than in person-sets (Supplementary Table 4).

Sex label shuffling of person-sets of CNVs containing gene(s) not expressed in the ROI. See Supplementary Table 4D-F, column 3. This test is motivated by the expectation that if the F-M difference in median size of CNVs containing gene(s) expressed in the ROI can be explained by a tendency toward larger CNV size *more generally* among ASDf, then we should also detect significantly larger median CNV size in ASDf than ASDm when limiting our analysis to consider only CNVs that contain gene(s) *not* expressed in the ROI. Conversely, if sex differences in median CNV size are primarily driven by CNVs containing gene(s) expressed in the ROI, then these tests should be non-significant. This test examines how often the F-M difference found in median size of CNVs *not* containing ROI-expressed gene(s) can be replicated when the sex labels of these person-sets of CNVs are randomly shuffled. Specifically, for a given individual, CNVs containing BrainSpan gene(s) that are *not* expressed in the ROI are included in their person-set, while CNVs containing at least one BrainSpan gene that *is* expressed in the ROI, as well as CNVs containing genes not characterized in BrainSpan, are excluded. For each permutation, from all person-sets of

CNVs containing gene(s) *not* expressed in the ROI, a number of person-sets equivalent to the original female *n* carrying CNVs containing gene(s) *not* expressed in the ROI are randomly selected and labeled as female, and likewise a number of person-sets equivalent to the original male *n* carrying CNVs containing gene(s) *not* expressed in the ROI are randomly selected and labeled as male. The new F-M difference in median CNV size is then calculated for the permutation and compared to the actual F-M difference in median size of CNVs containing gene(s) *not* expressed in the ROI.

Person-sets of CNVs containing gene(s) expressed in ROI: sex labels random, sample size reduced to non-ROI sample size. See Supplementary Table 4D-F, columns 4 and 6. These permutations allow us to test whether our significant finding for an ROI may be driven by a few outlier individuals; we do so by reducing the sample size for each permutation to be equivalent to the number of individuals carrying CNVs containing gene(s) *not* expressed in the ROI (e.g., 22 ASDf and 28 ASDm for CNVs containing gene(s) not expressed in the R/L-STR). Then, for each permutation, from all person-sets of CNVs containing gene(s) expressed in the ROI (e.g., 74 total ASD for R/L-STR), the equivalent female *n* number of person-sets are randomly selected and labeled female, and the equivalent male *n* number of person-sets are randomly selected and labeled male. The new F-M difference in median CNV size is then calculated for the permutation and compared to the actual F-M difference in median size of CNVs containing gene(s) expressed in the ROI (Supplementary Table 3D-F, column 6), as well as the F-M difference in median size of CNVs *not* containing gene(s) expressed in the ROI (Supplementary Table 4D-F, column 4).

Person-sets of CNVs containing gene(s) expressed in ROI: sex labels preserved, sample size reduced to non-ROI sample size. See Supplementary Table 4D-F, columns 5 and 7. Again, these permutations allow us to test whether our significant finding for the ROI may be driven by a few outlier individuals; we do so by reducing the sample size for each permutation to be equivalent to the number of individuals carrying CNVs containing gene(s) *not* expressed in the ROI (e.g., 22 ASDf and 28 ASDm for CNVs containing gene(s) not expressed in R/L-STR). Then, for each permutation, from all female person-sets of CNVs containing gene(s) expressed in the ROI (e.g., 40 ASDf for R/L-STR), the equivalent *n* number of female person-sets are randomly selected and labeled female. From the set of all male person-sets of CNVs containing gene(s) expressed in the ROI (e.g., 34 ASDm for R/L-STR), the equivalent *n* number of male person-sets are randomly selected and labeled male. The F-M difference in median CNV size is then calculated for the permutation and compared to the actual F-M difference in median size of CNVs containing gene(s) expressed in the ROI (Supplementary Table 4D-F, column 7), and the F-M difference in median size of CNVs *not* containing gene(s) expressed in the ROI (Supplementary Table 4D-F, column 5).

1.13 Data availability

Raw anonymized MR functional and structural DICOM images collected for this project, raw genotyping data from the HumanOmni2.5M BeadChip (Illumina), and accompanying phenotypic data, are available at the National Database for Autism Research (NDAR; Hall *et al.*, 2012) at ndar.nih.gov, DOI:10.15154/1478424. The parents of one TD participant provided consent for their child to participate in the study but did not complete necessary paperwork for NDAR GUID generation and was lost to follow-up. Other measures from the child were complete. Consequently, data from this child are included in the fMRI analysis but are not available in NDAR. The remaining 206 fMRI datasets are available, as are all genotyping datasets. Unthresholded statistical images (*z*-statistic & contrast of parameter estimates [COPE] images) and binary masks for fMRI analyses are available at Neurovault (Gorgolewski *et al.*, 2015) (neurovault.org/collections/BZOJTKVI/). Thresholded statistical images from fMRI analyses and additional resources are provided at this project's Open Science Framework (Foster and Deardorff, 2017) site (https://osf.io/mgzny).

Supplementary Results

2.1 Within-group fMRI response to BIO > SCRAM

Atlas label assignment. For reported results, both in text and in the Supplementary Tables, macroanatomical labels are assigned from the Harvard-Oxford Cortical/Subcortical Structural Atlases (Desikan *et al.*, 2006) and supplemented with macroanatomical and cytoarchitectonic

labels from the Juelich Histological Atlas (Eickhoff *et al.*, 2005); additional white matter labels are provided by the Mars Parietal connectivity-based parcellation atlas (Mars *et al.*, 2011) as well as the JHU White-Matter Tractography and ICBM-DTI-81 White-Matter Labels Atlases (Mori and Crain, 2005; Wakana *et al.*, 2007; Hua *et al.*, 2008). Dorsal/ventral premotor boundary is set at z = 48 based on Tomassini and colleagues (Tomassini *et al.*, 2007).

TDf. Whole-brain analysis of the full sample of TDf demonstrated widespread response to the BIO > SCRAM contrast in temporo-occipital, parietal, and posterior frontal regions (Supplementary Table 1A, Supplementary Fig. 1A). Specifically, bilateral occipital pole activated preferentially to coherent versus scrambled human motion; right posterior temporal regions also showed this pattern, proceeding along a trajectory from temporo-parietal junction (TPJ) and pSTS through temporo-occipital middle temporal gyrus (toMTG) and into inferior lateral occipital cortex (iLOC), temporo-occipital inferior temporal gyrus (toITG), and temporo-occipital fusiform (toFFG). Somewhat less extensive and lower magnitude response was also observed in left hemisphere posterior temporal regions. Bilateral activity was also observed in anterior intra-parietal sulcus (aIPS), inferior parietal lobule (IPL), and central operculum. TDf also showed primarily right-lateralized activity in the pars opercularis of the inferior frontal gyrus (IFGpop), primary (S1) and secondary (S2) somatosensory cortices, superior frontal gyrus, dorsal premotor cortex (PMd), and insula. Left-lateralized activity was observed in precentral gyrus and hemispheric lobules VIIb, VIIIa, and Crus II of the cerebellum.

TDm. Within the full sample of TDm, significant BIO > SCRAM response was observed in bilateral temporo-occipital regions as well as right-lateralized putamen, IPL, and S1 (Supplementary Table 1A, Supplementary Fig. 1B). Temporo-occipital activity in the right hemisphere included recruitment of TPJ, pSTS, toMTG, iLOC, occipital pole, toITG, and toFFG; in the left hemisphere, this activity did not extend into pSTS but included lower magnitude recruitment of more posterior/inferior structures including occipital pole, iLOC, and toMTG.

ASDf. Across the full sample of ASDf, BIO > SCRAM activity was observed in bilateral occipito-temporal regions, including occipital pole and toFFG. Activity in the posterior temporal fusiform cortex and toITG was primarily lateralized to the left hemisphere. In the right hemisphere, a cluster in posterior supramarginal gyrus/IPL was also observed. See Supplementary Table 1A and Supplementary Figure 1C.

ASDm. Among ASDm, no BIO > SCRAM activity surpassed threshold. We note that at a less stringent threshold of $z = 2.3$, corr. $p < .05$, a small cluster ($k = 366$) could be observed in right occipital pole.

2.2 ASD-associated genes contained by CNVs in the GENDAAR cohort

We examined GENDAAR CNVs for genes expressed in R/L-STR which have also been associated with ASD as per the Simons Foundation Autism Research Initiative (SFARI) database, a manually curated list of ASD risk genes (scores of S and 1-3; gene.sfari.org; accessed 03/04/2020). Of the 87 genes contained in ASDf CNVs that were expressed in R/L-STR, 7 (8%) were associated with ASD. Of the 55 genes contained in ASDm CNVs that were expressed in R/L-STR, 0 (0%) were associated with ASD, yielding a significant difference ($p = 0.0430$, Fisher exact test, 2-tailed).

2.3 ASD-associated genes contained by CNVs in the SSC cohort

Of the 626 genes contained in ASDf CNVs that were expressed in R/L-STR, 70 (11%) were associated with ASD. Of the 1882 genes contained in ASDm CNVs that were expressed in R/L-STR, 155 (8%) were associated with ASD, yielding a significant difference ($p = 0.0290$, Fisher exact test, 2-tailed).

Supplementary References

Andersson JLR, Jenkinson M, Smith SM. Non-linear registration, aka spatial normalisation [Internet]. 2007Available from: www.fmrib.ox.ac.uk/analysis/techrep

Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in FMRI. Neuroimage 2003; 20: 1052–63.

Constantino JN, Gruber CP. Social Responsiveness Scale, Second Edition. Los Angeles, CA: Westeren Psychological Services; 2012

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006; 31: 968–80.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, et al. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 2005; 25: 1325–35.

Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci 2016; 113: 7900–5.

Foster ED, Deardorff A. Open Science Framework (OSF). J Med Libr Assoc 2017; 105: 203.

Gee DG, McEwen SC, Forsyth JK, Haut KM, Bearden CE, Addington J, et al. Reliability of an fMRI paradigm for emotional processing in a multisite longitudinal study. Hum Brain Mapp 2015; 36: 2558–79.

Gorgolewski KJ, Varoquaux G, Rivera G, Schwarz Y, Ghosh SS, Maumet C, et al. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Front Neuroinform 2015; 9: 8.

Greve DN, Fischl B. Accurate and Robust Brain Image Alignment using Boundary-based Registration. Neuroimage 2009; 48: 63.

Hall D, Huerta MF, McAuliffe MJ, Farber GK. Sharing heterogeneous data: the national database for autism research. Neuroinformatics 2012; 10: 331–9.

Hansen BB. Full Matching in an Observational Study of Coaching for the SAT. J Am Stat Assoc 2004; 99: 609–18.

Ho DE, Imai K, King G, Stuart EA. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. Polit Anal 2007; 15: 199–236.

Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, et al. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 2008; 39: 336–47.

Hus V, Lord C. The autism diagnostic observation schedule, module 4: revised algorithm and standardized severity scores. J Autism Dev Disord 2014; 44: 1996–2012.

Jenkinson M, Bannister PR, Brady JM, Smith SM. Improved optimization for the robust and

accurate linear registration and motion correction of brain images. Neuroimage 2002; 17: 825–41.

Jenkinson M, Smith SM. A global optimisation method for robust affine registration of brain images. Med Image Anal 2001; 5: 143–56.

Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the human brain. Nature 2011; 478: 483–9.

Kessler D, Angstadt M, Sripada CS. Reevaluating 'cluster failure' in fMRI using nonparametric control of the false discovery rate. Proc Natl Acad Sci U S A 2017; 114: E3372--E3373.

Mars RB, Jbabdi S, Sallet J, O'Reilly JX, Croxson PL, Olivier E, et al. Diffusion-Weighted Imaging Tractography-Based Parcellation of the Human Parietal Cortex and Comparison with Human and Macaque Resting-State Functional Connectivity. J Neurosci 2011; 31: 4087–100.

Mori S, Crain BJ. MRI atlas of human white matter. Elsevier; 2005

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models [Internet]. 2017Available from: https://cran.r-project.org/package=nlme

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE, Li D, et al. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 2012; 59: 2142–54.

Rutter M, Bailey A, Lord C. SCQ: The Social Communication Questionnaire--Manual. Los Angeles, CA: Western Psychological Services; 2003

Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002; 17: 143–55.

Tomassini V, Jbabdi S, Klein JC, Behrens TEJ, Pozzilli C, Matthews PM, et al. Diffusion-Weighted Imaging Tractography-Based Parcellation of the Human Lateral Premotor Cortex Identifies Dorsal and Ventral Subregions with Anatomical and Functional Specializations. J Neurosci 2007; 27

Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 2007; 36: 630–44.

Woolrich MW. Robust group analysis using outlier inference. Neuroimage 2008; 41: 286–301.

Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, Smith SM. Multilevel linear modelling for fMRI group analysis using Bayesian inference. Neuroimage 2004; 21: 1732–47.

Woolrich MW, Ripley BD, Brady JM, Smith SM. Temporal autocorrelation in univariate linear modeling of fMRI data. Neuroimage 2001; 14: 1370–86.

Worsley K. CHAPTER 18 – Random Field Theory. In: Statistical Parametric Mapping. 2007. p. 232–6

A. TDf (n = 54)
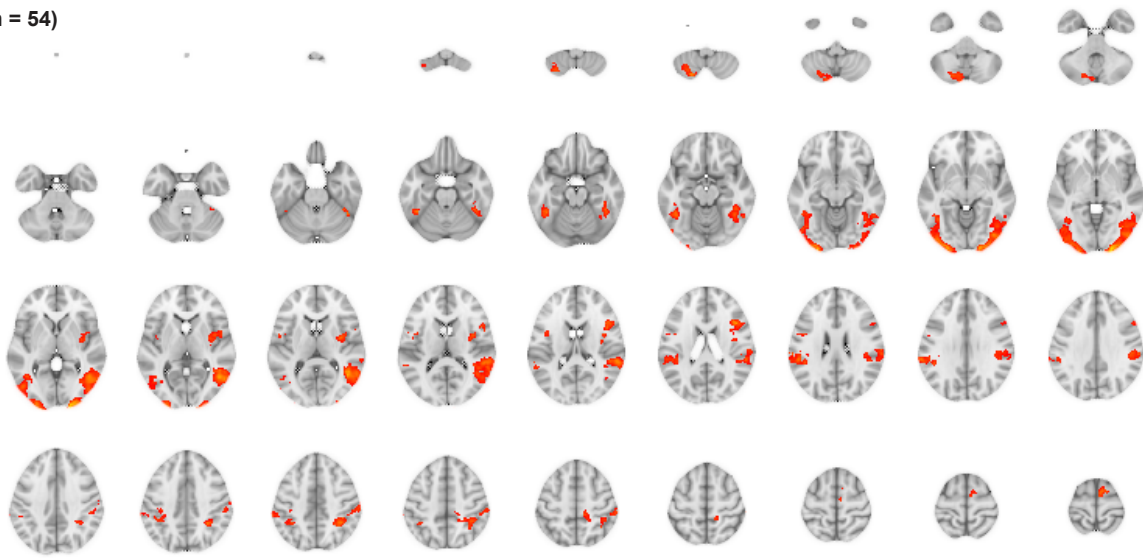
B. TDm (n = 59)
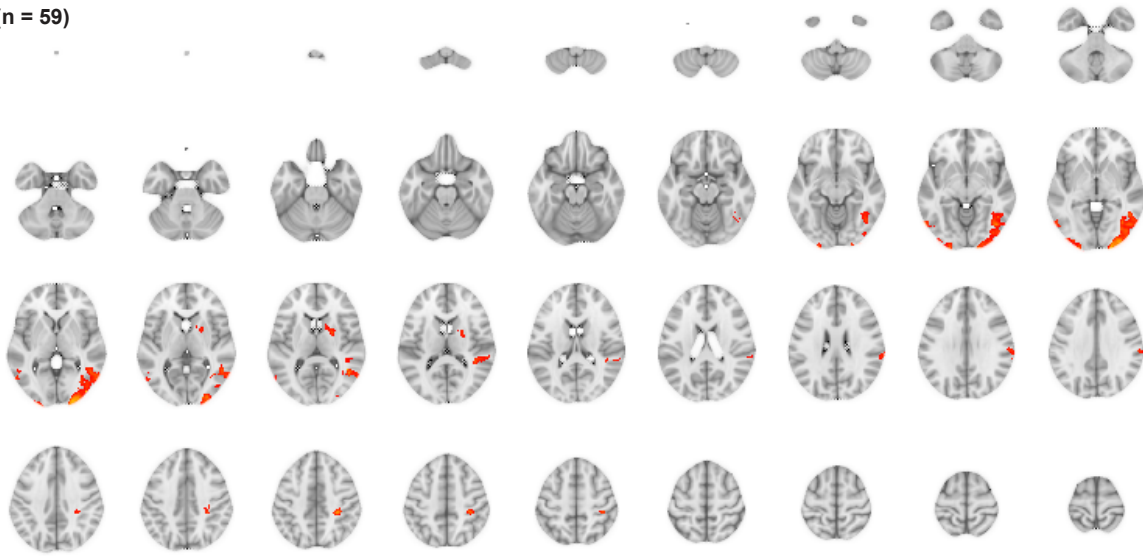
C. ASDf (n = 46)
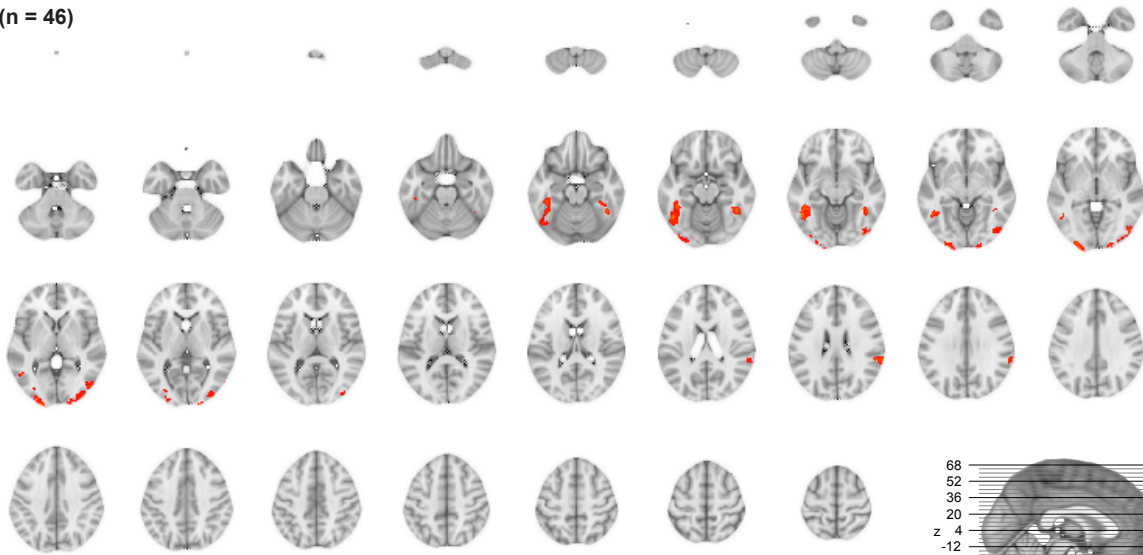
CDT $p < .001$ ($z = 3.1$);
FWE-corrected $p < .05$ $z = 3.1$ 6.0

68
52
36
20
z   4
−12
−28
−44
−60
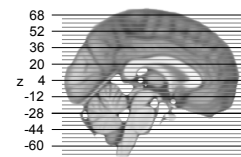
Supplementary Figure 1 | Mean group responses to BIO > SCRAM contrast within the full sample, controlling for FSIQ, age, site, and SRS total raw score. Analysis was conducted at the whole-brain level with a cluster-determining threshold (CDT) of $p < .001$ (i.e., $z = 3.1$) and FWE correction at $p < .05$. Axial z-statistic images for the samples of interest (**A:** TDf; **B:** TDm; **C:** ASDf) are displayed on the MNI standard brain template in neurological convention, with a sagittal reference image indicating slice position in the lower right. No suprath-reshold results were detected for ASDm ($n = 48$). We note that at a less stringent threshold of $z = 2.3$, $p < .05$, a small cluster (k = 366) of ASDm response could be observed in right occipital pole.

Permutation-based TFCE Inference

A. TDm > ASDm (*p* < .10)

x = 28   y = -10   z = 38

t = 2.3 ▮ 5.0

B. Overlap between TDm > ASDm and TDf > ASDf

x = 20   x = 22   x = 24

x = 26   x = 28   x = 30

▮ TDf-ASDf, permutation-based *p* < .05   ▮ Overlap
▮ TDm-ASDm, permutation-based *p* < .10

RFT Clusterwise Inference

C.TDm > ASDm (*z* > 2.3, *p* < .05)

x = 48   y = -60   z = 8

z = 2.3 ▮ 5.0

D. No overlap between TDm > ASDm and TDf > ASDf

x = 56   x = 52   x = 48

▮ TDf-ASDf cluster-wise *z* > 2.3, *p* < .05
▮ TDm-ASDm cluster-wise *z* > 2.3, *p* < .05
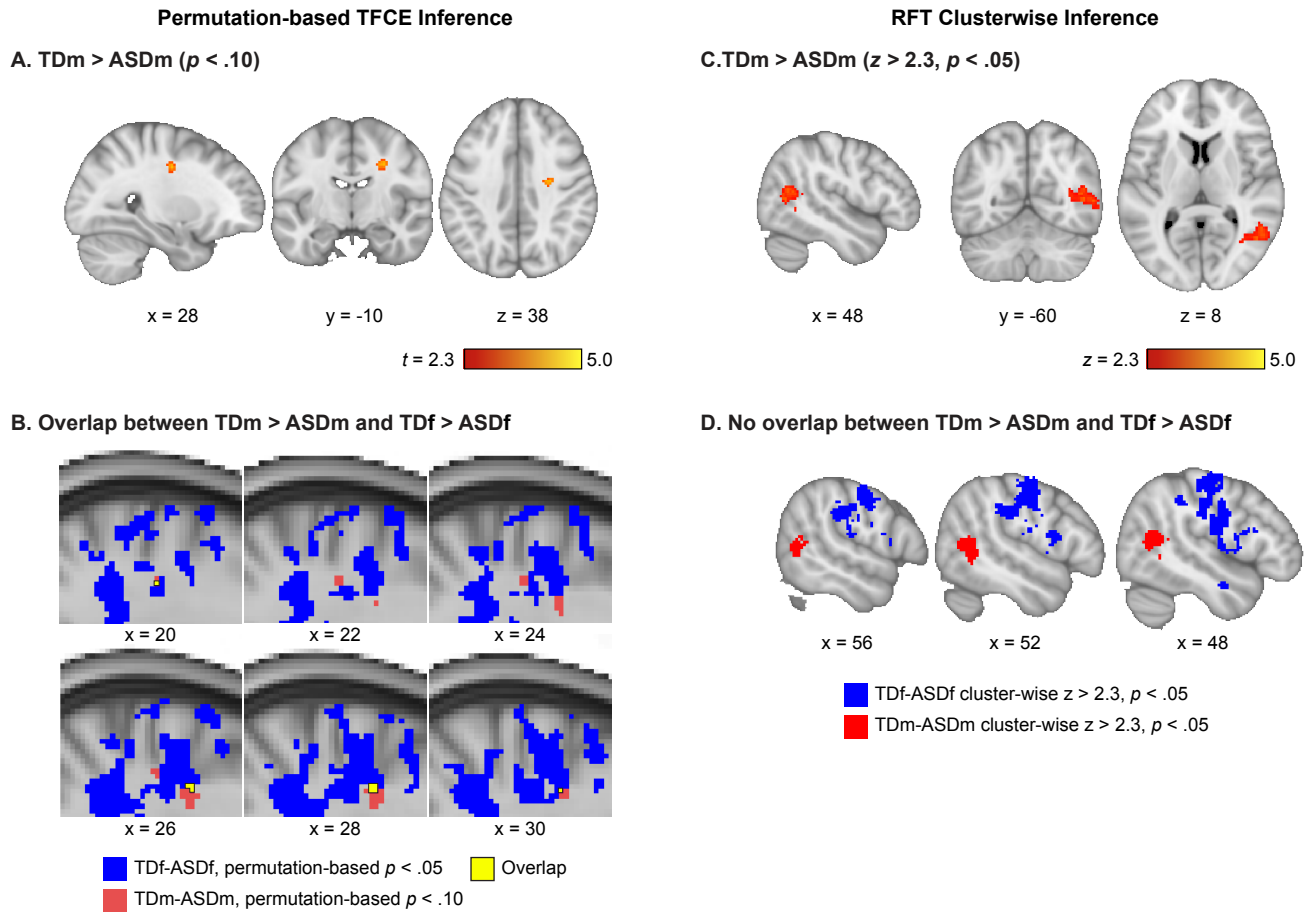
Supplementary Figure 2 | Trend-level TDm > ASDm differences in BIO > SCRAM response within the matched sample, controlling for FSIQ, age, site, and SRS total raw score. **A:** *Z*-statistic image displaying trend level (corrected *p* < .10) results for TDm > ASDm differences in BIO > SCRAM response, with statistical inference based on 10,000 permutations and threshold-free cluster enhancement (TFCE). Two small clusters are detected within the right corticospinal tract: the first with a voxel extent of k = 39 and a peak of *t* = 4.24 at MNI coordinates x = 28, y = 10, z = 38 (shown), and the second with k = 13 and a peak of *t* = 3.62 at x = 24, y = 24, z = 48 (not depicted). **B:** Detail view (sagittal orientation) of the permutation-based TDm > ASDm map at trend level (red), the permutation-based TDf > ASDf map at corrected *p* < .05 (blue), and their overlap (yellow). Overlap was limited to k = 10 voxels within right corticospinal tract and para-premotor white matter. **C:** Z-statistic image displaying results for the TDm > ASDm difference in BIO > SCRAM when using Gaussian random field theory (RFT)-based clusterwise inference with familywise error (FWE) correction at *z* > 2.3, *p* < .05. Depicted is the single cluster of k = 386 detected in right posterior caudal STS/middle temporal gyrus with a peak of *z* = 3.37 at x = 48, y = -60, z = 8. This cluster is located posterior to right pSTS involvement in the TDf > ASDf contrast as determined by permutation-based inference (Fig. 1A), with no overlap. **D:** Sagittal view demonstrating the lack of overlap between TDm > ASDm (red) and TDf > ASDf (blue) results generated using RFT cluster-wise inference at z > 2.3 and FWE corrected *p* < .05. All maps are displayed on the MNI standard brain template in neurological convention, with MNI coordinates of the depicted slice listed below.

a. TDf

b. TDm

c. ASDf

d. ASDm

Number of participants with coverage
1 [   ] All
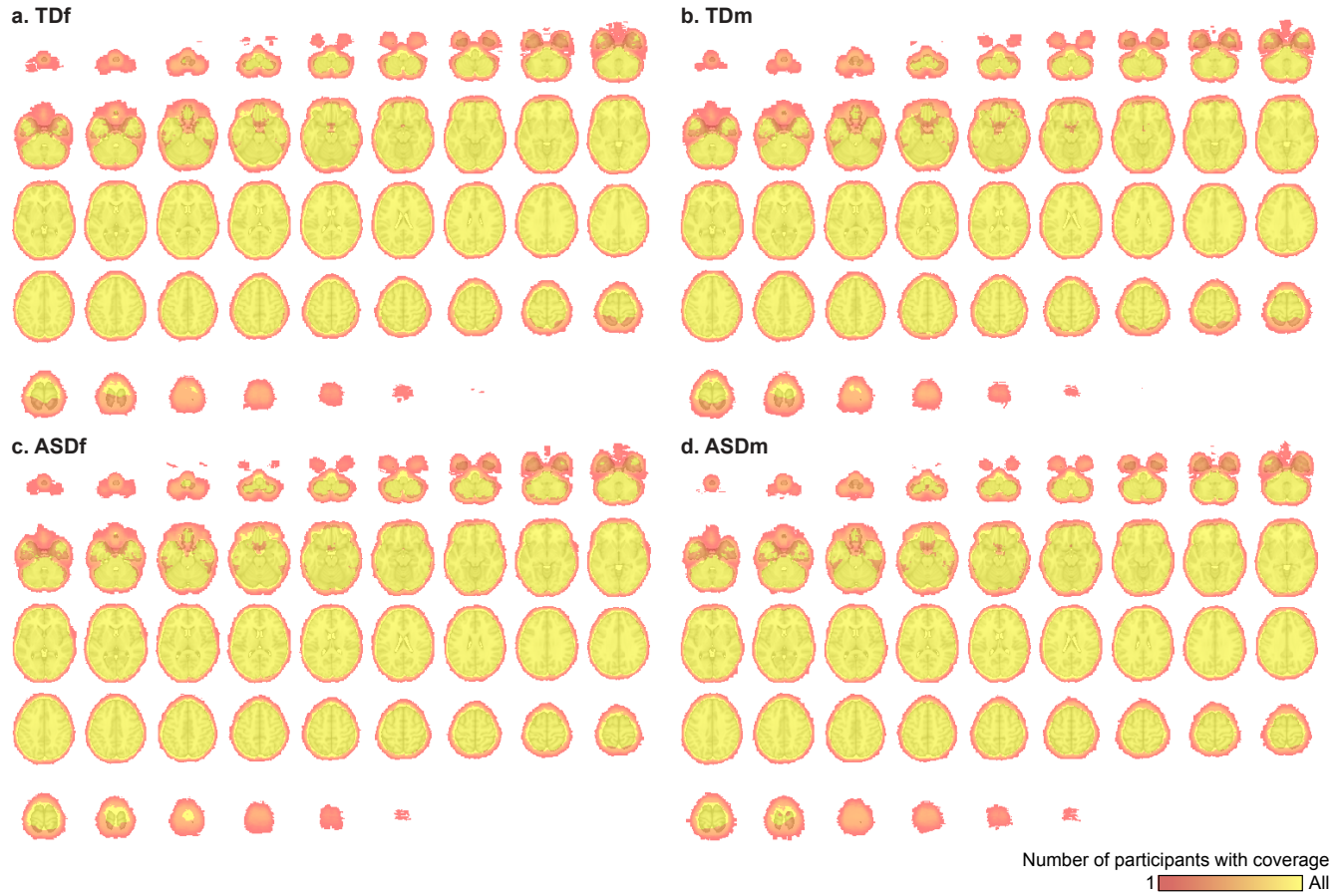
Supplementary Figure 3 | Summary of functional coverage in the full fMRI sample. Axial images for each group (**a:** TDf; **b:** TDm; **c:** ASDf; **d:** ASDm) depict the number of participants with functional coverage over a particular region of the MNI standard space template via a red (coverage in *n* = 1 participant) to yellow (coverage in all participants) color gradient..

A. Self-reported race and ethnicity in the GENDAAR CNV sample.

Race/Ethnicity	ASDf	ASDm	TDf	TDm
White	45 (7)	45 (8)	45 (7)	46 (9)
Black or African American	2 (0)	3 (0)	5 (2)	2 (0)
Asian	2 (0)	1 (0)	5 (0)	4 (0)
American Indian/Native Alaskan	0 (0)	1 (0)	0 (0)	0 (0)
More than one race	6 (1)	10 (2)	8 (2)	5 (1)
Undeclared	2 (0)*	0 (0)	2 (0)*+	2 (0)*

Note: Total n's for each racial group are listed with the subtotal reporting Hispanic or Latino descent in parentheses. *: Found to cluster with European population. +: One individual found to cluster with Asian population.

B. Confirmation of self-reported race and ethnicity in the GENDAAR CNV sample via PCA.
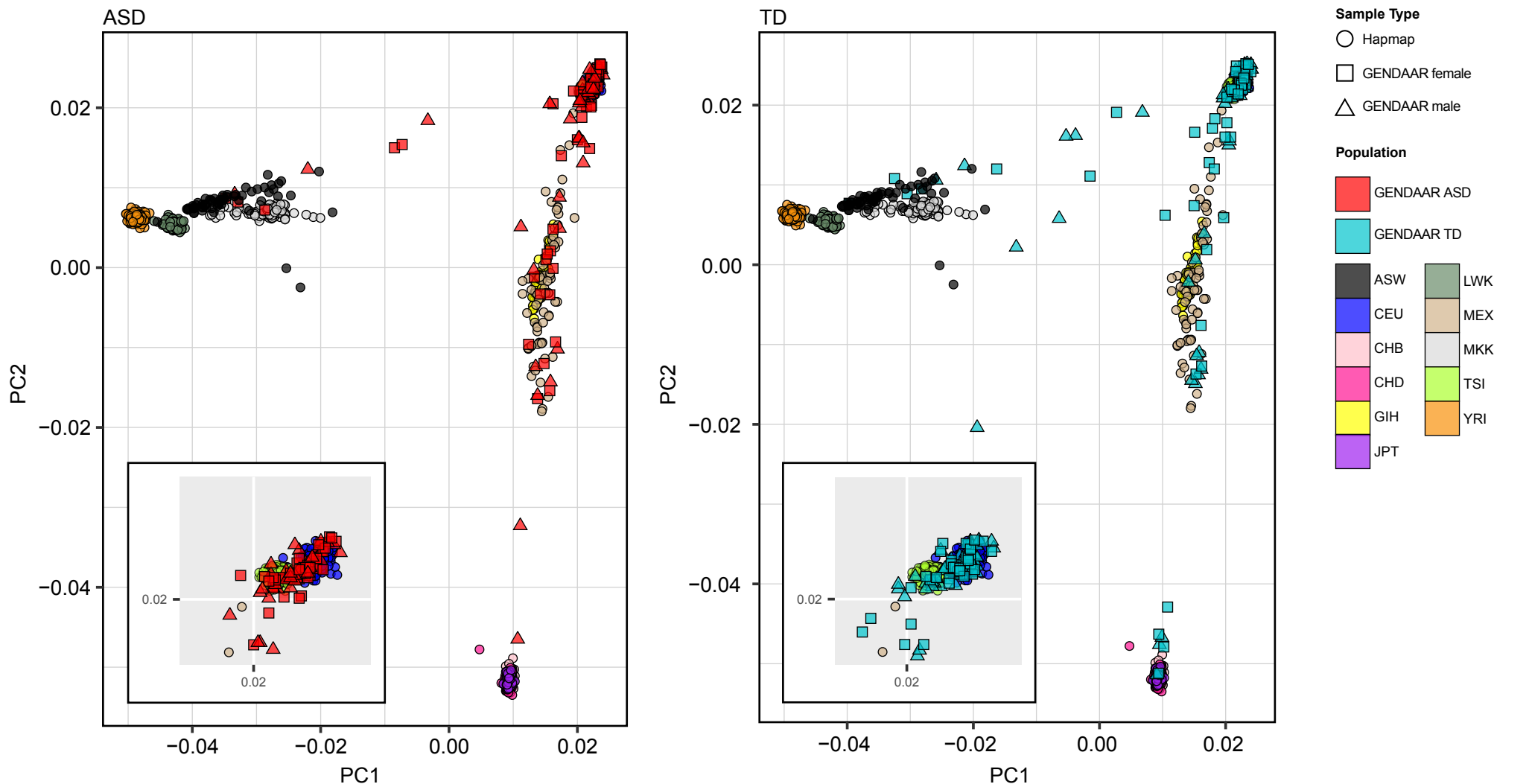
Supplementary Figure 4 | Confirmation of self-reported race and ethnicity of probands in the GENDAAR CNV sample. **A:** Self-reported race and ethnicity among probands. **B:** Population stratification using principal component analysis. Eigensoft was used to compare SNP genotypes of individuals in the GENDAAR CNV sample (ASD: red, left; TD: cyan, right) to individuals of known ancestry in HapMap3. Eigenvalues of the first two principal components (PC1 & PC2), which contributed to the greatest amount of variation relative to the other principal components, are shown plotted against each other. ASW (black): African ancestry in Southwest USA; CEU (blue): residents of Utah with Northern and Western European ancestry; CHB (light pink): Han Chinese in Beijing; CHD (dark pink): Chinese in Denver, CO; GIH (yellow): Gujarati Indians in Houston, TX; JPT (purple): Japanese in Tokyo; LWK (dark green): Luhya in Kenya; MEX (tan): Mexican ancestry in Los Angeles, CA; MKK (light grey): Maasai in Kenya; TSI (light green): Toscans in Italy; YRI (orange): Yoruba in Nigeria.

Supplementary Table 1 | Cluster peaks & local maxima of sites with preferential response to coherent versus scrambled biological motion (BIO > SCRAM) within group.

Macroanatomical site	Cytoarch.	Hem	x	y	z	Z	Cluster p (corr.)	k
ASDf								
Left temporo-occipital cluster							1.07×10⁻⁶	509
Inferior longitudinal fasciculus/Optic radiation		L	-40	-44	-12	4.68		
Temporo-occipital inferior temporal gyrus		L	-46	-54	-14	4.52		
Posterior temporal fusiform cortex		L	-40	-34	-20	4.05		
Right occipital cluster							3.28×10⁻⁶	459
Inferior lateral occipital cortex		R	46	-74	-2	5.97		
Inferior lateral occipital cortex	V4	R	40	-78	-10	4.78		
Left occipital cluster							1.69×10⁻⁵	389
Occipital pole	V3v	L	-22	-98	-4	4.49		
Occipital Pole	V2/BA 18	L	-16	-106	2	4.42		
Right temporo-occipital cluster							5.74×10⁻³	175
Temporal occipital fusiform cortex/Inferior longitudinal fasciculus		R	38	-46	-12	4.84		
Temporal occipital fusiform cortex			44	-48	-18	4.42		
Posterior temporal fusiform cortex			36	-36	-20	3.78		
Right parietal cluster							3.84×10⁻²	118
Posterior supramarginal gyrus/Inferior parietal lobule	PF	R	68	-36	28	4.39		
Posterior supramarginal Gyrus/Inferior parietal lobule	PFcm	R	54	-36	24	3.78		
ASDm								
n.s.								
TDf								
Right temporo-occipital cluster							6.10×10⁻²⁶	3798
Occipital pole	V1/BA 17	R	20	-102	-2	5.77		
Temporo-occipital middle temporal gyrus/Temporal WM		R	48	-52	6	5.25		
Occipital Pole	V3v	R	30	-94	6	4.91		
Inferior lateral occipital cortex	V5	R	54	-62	4	4.78		
Posterior superior temporal sulcus		R	58	-38	14	4.74		
Temporal occipital fusiform cortex		R	42	-42	-18	4.39		
Left temporo-occipital cluster							1.79×10⁻¹⁶	1960
Occipital Pole	V2/BA 18	L	-20	-104	0	5.1		
Occipital Pole	V3v	L	-28	-94	-10	5.06		
Posterior temporal fusiform cortex		L	-40	-44	-22	4.58		
Parietal operculum/Inferior parietal lobule	PFcm	L	-54	-34	22	4.56	2.38×10⁻⁷	625
Right inferior-posterior frontal cluster							4.77×10⁻⁷	
Inferior frontal gyrus (pars opercularis)/Frontal WM	BA 45	R	44	16	20	4.99		589
Insula/S2C/Parietal Operculum	OP3	R	38	-2	10	4.39		
Inferior frontal gyrus (pars opercularis)	BA 44	R	50	12	12	4.24		
Right parietal cluster							1.34×10⁻⁵	
Superior parietal lobule/Anterior intraparietal sulcus/Parietal WM	hIP3	R	30	-48	42	4.78		429
Supramarginal Gyrus/S1C	BA 2	R	38	-38	44	4.55		
Postcentral Gyrus/S1C/Parietal WM	BA 3a	R	14	-34	52	4		
Superior Parietal Lobule/Anterior intraparietal sulcus	hIP3		36	-52	52	3.94		
S1C/SPLA	BA 2		26	-40	48	3.8		
Left cerebellar cluster							2.19×10⁻⁵	407
Hemispheric VIIb		L	-20	-72	-50	4.78		
Hemispheric VIIIa		L	-24	-66	-54	4.27		
Hemispheric Crus II		L	-6	-82	-44	4.03		
Superior frontal gyrus/Dorsal premotor cortex	BA 6	R	8	-2	74	4.54	7.11×10⁻³	180
Left posterior frontal cluster							2.20×10⁻²	143
Precentral Gyrus/Posterior frontal WM		L	-44	0	24	3.89		
Central Operculum		L	-42	0	14	3.77		
Precentral Gyrus/M1C/Posterior frontal WM	BA 4p	L	-50	0	22	3.73		
Central Operculum/S2C/Parietal operculum	OP4	L	-50	0	8	3.72		
Insula		L	-40	-6	4	3.69		
TDm								
Right temporo-occipital cluster							4.80E-16	1636
Inferior lateral occipital cortex		R	52	-60	-6	6.45		
Temporo-occipital fusiform cortex		R	42	-58	-14	5.88		
Inferior lateral occipital cortex	V4	R	44	-82	0	5.43		
Occipital Pole	V1/BA 17	R	24	-96	-2	5.26		
Occipital Pole	V3v	R	32	-92	-2	5.14		
Left occipital cluster							7.24E-05	311
Occipital Pole	V2/BA 18	L	-22	-100	-10	4.87		
Occipital Pole	V3v	L	-28	-98	-6	4.32		
Inferior lateral occipital cortex	V4	L	-32	-88	-4	3.82		
Right temporo-parietal cluster							0.00224	194
Superior longitudinal fasciculus/Parietal WM		R	42	-40	14	4.39		
Supramarginal gyrus/Inferior parietal lobule	PF	R	60	-26	32	3.77		
Posterior superior temporal gyrus			60	-34	12	3.66		
Left temporo-occipital cluster							0.0028	187
Inferior lateral occipital cortex		L	-56	-64	-6	4.31		
Temporo-occipital middle temporal gyrus		L	-56	-60	6	3.9		
Right striatal cluster							0.0271	121
Putamen		R	24	0	10	4.1		
Anterior thalamic radiation/anterior limb of internal capsule		R	16	10	6	3.52		
Caudate/Anterior thalamic radiation		R	14	2	14	3.38		
Right posterior frontal cluster							0.0326	116
Postcentral Gyrus/Parietal WM/SPLA		R	30	-34	44	4.79		
Postcentral Gyrus/S1C	BA 2	R	38	-34	50	4.27		
S1C/Parietal WM/IPLB	BA 3a	R	34	-24	40	3.66		

Note. Cluster peaks and local maxima (in MNI space) of regions showing significant (*z* > 3.10, corr. *p* = .05) BIO > SCRAM response within group. See Supplementary Methods 2.1 for details of atlas label assignment. *BA*: Brodmann's area. *Corr*: Corrected. *Cytoarch*: Cytoarchitectonic. *Hem*: Hemisphere. *k*: Voxel extent. *L*: Left.. *M1C*: Primary motor cortex. *n.s.*: Non-significant. *R*: Right. *S1C*; *S2C*: Primary; secondary somatosensory cortex. *WM*: White matter. *Z*: Z-statistic. *SPLA* & *IPLB* are WM divisions from the Mars Parietal connectivity-based atlas. *V1, V2, V4, V5,* and *V3v* refer to visual areas 1, 2, 4, & 5, and ventral visual area 3.

Supplementary Table 2 | Regions of overlap between TDf > TDm and TDf > ASDf.

Region	k
Left Frontal Lobe	
Precentral gyrus	110
Juxtapositional lobule cortex	34
Superior frontal gyrus	27
Paracingulate	20
Middle frontal gyrus	13
Right Frontal Lobe	
Insula	200
Paracingulate	80
Superior frontal gyrus	75
Juxtapositional lobule cortex	49
Inferior frontal gyrus pars opercularis	45
Middle frontal gyrus	34
Frontal operculum	29
Frontal pole	22
Precentral gyrus	16
Left Parietal Lobe	
Postcentral gyrus	33
Left anterior supramarginal gyrus	26
Right Parietal Lobe	
Superior parietal lobule	108
Posterior supramarginal gyrus	29
Postcentral gyrus	6
Right Temporal Lobe	
Temporal pole	10
Right Occipital Lobe	
Superior lateral occipital cortex	5
White Matter	
Left corticospinal tract	49
Callosal body	12
Right corticospinal tract	9
Left cingulum	2

Note. *k:* Voxel extent.

Supplementary Table 3 | Number of CNVs and genes contained by CNVs in the GENDAAR ASD sample, by sex.

ASDf Median (Range)	ASDm Median (Range)	*p*
Number of CNVs (all CNVs)		
2 (1-6)	2 (1-4)	0.3840
Number of genes contained in CNVs (all genes)		
3 (1-12*)	4 (1-11)	0.8136

Note. *P*-values are calculated from unpaired *t*-tests. *One ASDf participant constituted an outlier with 47 genes contained in CNVs and was not included in median or *p*-value calculation.

Supplementary Table 4 | All permutation analyses probing significance of F-M difference in median total size of CNVs containing gene(s) expressed or *not* expressed in ROIs.

1. CNVs containing gene(s) expressed in ROI	2. F-M difference in median total CNV size (bp)	3. Sex label shuffling (*p*)	4. All CNVs containing *any* gene(s), sex labels random (*p*)	5. All CNVs containing *any* gene(s), sex labels preserved (*p*)	6. All CNVs containing BrainSpan gene(s), sex labels random (*p*)	7. All CNVs containing BrainSpan gene(s), sex labels preserved (*p*)
A. GENDAAR cohort: ASD						
Permuting person-sets of CNVs						
ASDf neurofunctional profile ROIs (*n* = 11)	81,597	0.0487*	0.0274*	0.0285*	0.0112*	0.0006*
R-DFC	48,494	0.1762	0.1519	0.2533	0.1251	0.0832
L-DFC	48,494	0.2365	0.1514	0.2635	0.1277	0.1158
R-VFC	45,087	0.2024	0.1713	0.2930	0.1499	0.1168
L-VFC	78,733	0.0685	0.0364*	0.0446*	0.0146*	0.0022*
R-M1C	53,922	0.1635	0.1177	0.1925	0.0955	0.0739
L-M1C	75,922	0.0364*	0.0449*	0.0584	0.0239*	0.0135*
R-S1C	53,922	0.1430	0.1248	0.1916	0.1033	0.0752
L-S1C	46,169	0.2265	0.1571	0.2568	0.1309	0.0932
R-STR	83,529	0.0486*	0.0271*	0.0303*	0.0122*	0.0004*
L-STR	84,184	0.0379*	0.0300*	0.0371*	0.0122*	0.0023*
R-STC	38,417	0.2624	0.2137	0.3332	0.1803	0.1513
R/L-STR (*n* = 2)	89,999	0.0293*	0.0178*	0.0172*	0.0048*	0.0000*
Permuting CNVs independently of individuals						
ASDf neurofunctional profile ROIs (*n* = 11)	81,597	0.0075*	0.0001*	0.0161*	0.0001*	0.0056*
R-DFC	48,494	0.0559	0.0041*	0.1404	0.0037*	0.0501
L-DFC	48,494	0.0651	0.0048*	0.1086	0.0030*	0.0514
R-VFC	45,087	0.0712	0.0051*	0.1486	0.0049*	0.0571
L-VFC	78,733	0.0189*	0.0002*	0.0183*	0.0003*	0.0100*
R-M1C	53,922	0.0600	0.0025*	0.0820	0.0010*	0.0340*
L-M1C	75,922	0.0099*	0.0000*	0.0178*	0.0003*	0.0124*
R-S1C	53,922	0.0656	0.0032*	0.0755	0.0022*	0.0348*
L-S1C	46,169	0.1030	0.0050*	0.1397	0.0046*	0.0571
R-STR	83,529	0.0088*	0.0000*	0.0144*	0.0000*	0.0059*
L-STR	84,184	0.0020*	0.0001*	0.0130*	0.0001*	0.0064*
R-STC	38,417	0.1550	0.0154*	0.2250	0.0127*	0.0792
R/L-STR (*n* = 2)	89,999	0.0040*	0.0000*	0.0098*	0.0000*	0.0024*
B. GENDAAR cohort: TD						
Permuting person-sets of CNVs						
ASDf neurofunctional profile ROIs (*n* = 11)	-63,128	0.0681	0.1028	0.0121*	0.0593	0.0151*
R/L-STR (n=2)	-58,136	0.0779	0.1252	0.0238*	0.0833	0.0487*
Permuting CNVs independently of individuals						
ASDf neurofunctional profile ROIs (*n* = 11)	-63,128	0.0046*	0.0005*	0.0244*	0.0004*	0.0254*
R/L-STR (n=2)	-58,136	0.0054*	0.0011*	0.0373*	0.0013*	0.0401*
C. SSC cohort: ASD						
Permuting person-sets of CNVs						
ASDf neurofunctional profile ROIs (*n* = 11)	42,183	0.0041*	0.0022*	0.1160	0.0009*	0.0270*
R/L-STR (n=2)	45,133	0.0041*	0.0013*	0.1020	0.0013*	0.0211*
Permuting CNVs independently of individuals						
ASDf neurofunctional profile ROIs (*n* = 11)	42,183	0.0000*	0.0000*	0.0011*	0.0000*	0.0022*
R/L-STR (n=2)	45,133	0.0000*	0.0000*	0.0009*	0.0000*	0.0011*

1. CNVs containing gene(s) *not* expressed in ROI	2. F-M difference in median total CNV size (bp)	3. Sex label shuffling (p)	4. All CNVs containing ROI gene(s), sex labels random, compared to [D2, E2, or F2][a] (p)	5. All CNVs containing ROI gene(s), sex labels preserved, compared to [D2, E2, or F2][a] (p)	6. All CNVs containing ROI gene(s), sex labels random, compared to [A2, B2, or C2][b] (p)	7. All CNVs containing ROI gene(s), sex labels preserved, compared to [A2, B2, or C2][b] (p)
D. GENDAAR cohort: ASD						
Permuting person-sets of CNVs						
ASDf neurofunctional profile ROIs (*n* = 11)	5,822	0.3203	0.4681	0.9343	0.0897	0.5005
R-DFC	6,964	0.2400	0.4674	0.9363	0.2073	0.5838
L-DFC	6,964	0.2363	0.4796	0.9488	0.2329	0.6229
R-VFC	5,163	0.2382	0.4708	0.9749	0.2222	0.5768
L-VFC	5,026	0.2811	0.4825	0.9345	0.1011	0.4413
R-M1C	5,163	0.2287	0.4662	0.9451	0.1864	0.4980
L-M1C	9,552	0.1358	0.4463	0.9552	0.0806	0.5042
R-S1C	5,373	0.1552	0.4786	0.9548	0.1739	0.4824
L-S1C	11,671	0.0713	0.4387	0.8507	0.2355	0.5300
R-STR	1,139	0.3941	0.4928	0.9840	0.0736	0.5060
L-STR	2,082	0.3232	0.4907	0.9973	0.0619	0.4402
R-STC	3,362	0.2247	0.4967	0.9233	0.2819	0.5517
R/L-STR (*n* = 2)	1,002	0.4314	0.4941	0.9922	0.0533	0.4955
Permuting CNVs independently of individuals						
ASDf neurofunctional profile ROIs (*n* = 11)	5,822	0.2600	0.3907	0.6297	0.0206*	0.1413
R-DFC	6,964	0.1131	0.3479	0.5506	0.0591	0.2566
L-DFC	6,964	0.1548	0.3892	0.6830	0.0623	0.3566
R-VFC	5,163	0.2100	0.3920	0.6719	0.0737	0.2698
L-VFC	5,026	0.1692	0.3847	0.6346	0.0208*	0.1586
R-M1C	5,163	0.2010	0.4023	0.6373	0.0650	0.2361
L-M1C	9,523	0.0782	0.3338	0.6481	0.0157*	0.2146
R-S1C	5,373	0.2629	0.3977	0.6716	0.0651	0.2449
L-S1C	11,671	0.0627	0.3327	0.5311	0.0885	0.2676
R-STR	1,139	0.3640	0.4214	0.6946	0.0141*	0.1120
L-STR	2,082	0.3004	0.4257	0.7829	0.0065*	0.1141
R-STC	3,362	0.3069	0.4188	0.6165	0.1284	0.2999
R/L-STR (*n* = 2)	1,002	0.3811	0.4297	0.7482	0.0093*	0.1215
E. GENDAAR cohort: TD						
Permuting person-sets of CNVs						
ASDf neurofunctional profile ROIs (*n* = 11)	-8,345	0.1268	0.4405	0.9857	0.8945	0.5759
R/L- STR (*n* = 2)	1,595	0.4131	0.4979	0.0026*	0.8790	0.4826
Permuting CNVs independently of individuals						
ASDf neurofunctional profile ROIs (*n* = 11)	-8,345	0.2146	0.3214	0.6045	1.0000	0.9275
R/L- STR (n = 2)	1,595	0.4959	0.5064	0.1801	0.9999	0.9269
F. SSC cohort: ASD						
Permuting person-sets of CNVs						
ASDf neurofunctional profile ROIs (*n* = 11)	2,839	0.1913	0.4357	0.9956	0.0184*	0.4967
R/L- STR (n = 2)	2,641	0.2595	0.4428	0.9996	0.0112*	0.4673
Permuting CNVs independently of individuals						
ASDf neurofunctional profile ROIs (*n* = 11)	2,839	0.0128*	0.0631	0.9285	0.0000*	0.2331
R/L- STR (n = 2)	2,641	0.0044*	0.0576	0.9745	0.0000*	0.2512

Note. CNVs passed quality control and were rare; ASD CNVs were not in common with unaffected siblings. All tests use 10k permutations. *P*-values indicate how many iterations yielded the actual female minus male (F-M) difference in median CNV size or greater (value listed under "compared to"); *p* ≤ .05 (*) indicates that the difference is significant; *p* = .0000 indicates that no permutation yielded a difference ≥ than the comparison value. [a]: Compared to the actual F-M difference in median size of CNVs containing gene(s) *not* expressed in the ROI, which can be found in column 2 of section D, E, or F, as appropriate. [b]: Compared to the actual F-M difference in median size of CNVs containing gene(s) expressed in the ROI, which can be found in column 2 of section A, B, or C, as appropriate. **Abbreviations:** *bp:* base pairs. *R-*: Right; *L-*: Left; *DFC*: dorsolateral frontal cortex; *M1C*: primary motor cortex; *S1C*: primary somatosensory cortex; *SSC*: Simons Simplex Collection. *STC*: superior temporal cortex; *STR*: striatum; *VFC*: ventrolateral prefrontal cortex.

Supplementary Table 5 | Numbers of subjects remaining after each stage of data filtering based on subject inclusion/exclusion criteria, MR quality assurance standards, and sample matching.

		ASD		TD	
		Female	**Male**	**Female**	**Male**
i. Received intake assessment	n =	106	135	97	97
ii. Excluded at phenotyping review:	n =	32	38	20	14
Missing required measure(s)	n =	3	10	16	6
FSIQ ≤ 70	n =	5	3	0	0
Neurological disorder or NICU history	n =	0	1	0	0
Did not meet ASD criteria	n =	28	26	NA	NA
Did not meet TD criteria	n =	NA	NA	6	8
iii. Passed to MRI quality check	n =	74	97	77	83
iv. Excluded at quality check:	n =	28	47	23	23
Did not complete biological motion paradigm	n =	9	24	9	12
Excessive motion	n =	17	17	9	8
Insufficient functional coverage	n =	2	4	2	2
Poor quality structural scan	n =	0	1	1	0
Exclusionary incidental finding	n =	0	1	1	1
Signal artefact	n =	0	0	1	0
v. Excluded upon visual inspection for outliers:					
Head motion outliers	n =	0	2	0	1
vi. Passed to matching	n =	46	48	54	59
No adequate match available	n =	1	1	NA	NA
vii. Matched	n =	45	47	45	47
viii. In final sample	n =	45	47	45	47

Note. *FSIQ*: Full scale IQ as estimated from the DAS-2 General Conceptual Ability Standard Score.

Supplementary Table 6 | Child Behavior Checklist (CBCL) *t*-scores, by sex and group.

A. Full fMRI Sample (*N* = 207)

	ASD			TD			Omn. Diff.	Sig. contrast(s)	Pairwise Diff.
	Female *n* = 46	Male *n* = 48	Sex Diff.	Female *n* = 54	Male *n* = 59	Sex Diff.			
	M(SD)	*M(SD)*	*p*	*M(SD)*	*M(SD)*	*p*	*p*		*p*
Emp. syndrome scales									
Anxious/Depressed	63.00(9.80)	60.87(9.59)	.303	52.02(4.15)	51.75(4.60)	.742	<.001	ASD > TD	<.001
Withdrawn/Depressed	66.28(12.54)	63.33(9.36)	.214	52.31(3.83)	51.78(2.98)	.424	<.001	ASD > TD	<.001
Somatic Complaints	57.93(7.05)	58.70(7.98)	.632	53.50(5.01)	52.64(5.39)	.388	<.001	ASD > TD	≤.005
Social Problems	64.14(8.46)	64.28(8.55)	.937	51.12(2.36)	50.73(2.02)	.359	<.001	ASD > TD	<.001
Thought Problems	64.86(9.63)	64.22(8.22)	.736	51.79(4.26)	51.49(2.73)	.668	<.001	ASD > TD	<.001
Attention Problems	66.63(11.01)	65.59(9.27)	.632	51.48(2.70)	51.42(2.14)	.903	<.001	ASD > TD	<.001
Rule Break. Behavior	55.93(4.95)	56.41(6.66)	.698	51.06(2.28)	50.76(1.59)	.437	<.001	ASD > TD	<.001
Aggressive Behavior	58.65(7.55)	59.20(9.22)	.761	50.58(1.47)	50.63(1.54)	.861	<.001	ASD > TD	<.001
DSM-oriented scales									
Depressive problems[a]	61.08(9.07)	62.77(10.53)	.542	52.80(5.43)	51.40(3.27)	.266	<.001	ASD > TD	≤.001
Anxiety problems	60.91(9.03)	59.65(8.27)	.497	50.88(1.86)	51.36(3.19)	.338	<.001	ASD > TD	<.001
Somatic problems	56.14(7.20)	56.22(7.93)	.961	53.83(6.59)	52.93(6.30)	.468	.038	ASDm > TDm	.081
AD/H problems	63.00(8.63)	62.52(7.28)	.779	50.77(1.62)	50.83(1.28)	.827	<.001	ASD > TD	<.001
Opp. defiant problems	57.93(6.98)	59.09(8.70)	.490	51.38(2.64)	51.32(2.36)	.896	<.001	ASD > TD	<.001
Conduct problems	56.49(6.58)	57.24(8.08)	.631	50.69(1.96)	50.63(1.86)	.858	<.001	ASD > TD	<.001

B. Matched fMRI Sample (*N* = 184)

	ASD			TD			Omn. Diff.	Sig. contrast(s)	Pairwise Diff.
	Female *n* = 45	Male *n* = 47	Sex Diff.	Female *n* = 45	Male *n* = 47	Sex Diff.			
	M(SD)	*M(SD)*	*p*	*M(SD)*	*M(SD)*	*p*	*p*		*p*
Emp. syndrome scales									
Anxious/Depressed	63.00(9.92)	61.09(9.58)	.364	51.74(3.67)	52.15(5.08)	.664	<.001	ASD > TD	<.001
Withdrawn/Depressed	66.24(12.69)	63.36(9.47)	.236	52.35(3.80)	51.85(3.13)	.502	<.001	ASD > TD	<.001
Somatic Complaints	58.10(7.05)	58.82(8.03)	.654	53.81(5.35)	53.06(5.85)	.527	<.001	ASD > TD	≤.003
Social Problems	64.26(8.52)	64.42(8.60)	.931	51.05(2.34)	50.87(2.23)	.719	<.001	ASD > TD	<.001
Thought Problems	64.74(9.71)	64.44(8.16)	.879	52.00(4.63)	51.57(2.76)	.602	<.001	ASD > TD	<.001
Attention Problems	66.74(11.12)	65.73(9.32)	.650	51.58(2.92)	51.57(2.32)	.990	<.001	ASD > TD	<.001
Rule Break. Behavior	56.02(4.98)	56.49(6.71)	.713	51.14(2.40)	50.87(1.75)	.551	<.001	ASD > TD	<.001
Aggressive Behavior	58.76(7.61)	59.04(9.27)	.877	50.53(1.50)	50.74(1.70)	.536	<.001	ASD > TD	<.001
DSM-oriented scales									
Depressive problems[b]	51.74(3.67)	52.15(5.08)	.361	52.59(5.29)	51.62(3.46)	.464	<.001	ASD > TD	≤.003
Anxiety problems	52.35(3.80)	51.85(3.13)	.599	50.79(1.86)	51.47(3.52)	.252	<.001	ASD > TD	<.001
Somatic problems	53.81(5.35)	53.06(5.85)	.966	54.28(6.99)	53.34(6.83)	.522	.128		
AD/H problems	51.05(2.34)	50.87(2.23)	.840	50.81(1.74)	50.85(1.37)	.911	<.001	ASD > TD	<.001
Opp. defiant problems	52.00(4.63)	51.57(2.76)	.694	51.33(2.63)	51.47(2.52)	.794	<.001	ASD > TD	<.001
Conduct problems	51.58(2.92)	51.57(2.32)	.612	50.84(2.13)	50.79(2.05)	.910	<.001	ASD > TD	<.001

C. CNV sample (*N* = 250)

	ASD			TD			Omn. Diff.	Sig. contrast(s)	Pairwise Diff.
	Female *n* = 61	Male *n* = 65	Sex Diff.	Female *n* = 65	Male *n* = 59	Sex Diff.			
	M(SD)	*M(SD)*	*p*	*M(SD)*	*M(SD)*	*p*	*p*		*p*
Emp. syndrome scales									
Anxious/Depressed	63.48(10.25)	60.17(9.43)	.074†	52.14(4.16)	51.39(3.63)	.289	<.001	ASD > TD	<.001
Withdrawn/Depressed	65.59(11.91)	61.90(9.43)	.069†	52.44(3.95)	51.72(2.80)	.246	<.001	ASD > TD	<.001
Somatic Complaints	59.27(8.40)	57.73(7.80)	.312	53.35(5.32)	52.65(5.42)	.477	<.001	ASD > TD	≤.003
Social Problems	65.55(9.15)	63.00(9.50)	.145	51.03(2.18)	50.68(1.97)	.362	<.001	ASD > TD	<.001
Thought Problems	66.11(9.57)	64.05(8.22)	.220	51.90(3.82)	50.79(1.57)	.036*	<.001	ASD > TD	<.001
Attention Problems	69.66(11.64)	65.39(9.17)	.032	52.03(3.50)	51.00(1.68)	.039*	<.001	ASD > TD	<.001
Rule Break. Behavior	56.61(5.96)	55.41(6.25)	.294	51.17(2.46)	50.67(1.02)	.137	<.001	ASD > TD	<.001
Aggressive Behavior	60.04(9.41)	58.14(8.58)	.261	50.73(1.94)	50.39(1.03)	.223	<.001	ASD > TD	<.001
DSM-oriented scales									
Depressive problems[c]	63.19(10.25)	61.06(9.62)	.400	52.92(4.72)	51.03(2.34)	.073†	<.001	ASD > TD	<.001
Anxiety problems	61.20(8.82)	59.24(7.86)	.212	50.78(1.64)	51.18(3.02)	.379	<.001	ASD > TD	<.001
Somatic problems	57.12(8.99)	55.73(8.18)	.387	53.30(6.65)	52.75(6.33)	.645	.006	ASDf > TDf ASDf > TDm	.033 .013
AD/H problems	64.59(8.58)	62.37(7.67)	.148	51.32(2.87)	49.79(6.83)	.121	<.001	ASD > TD	<.001
Opp. defiant problems	58.88(7.82)	58.27(8.30)	.689	51.49(2.95)	51.09(1.84)	.365	<.001	ASD > TD	<.001
Conduct problems	57.79(8.08)	55.81(7.17)	.170	50.86(2.28)	50.44(1.10)	.197	<.001	ASD > TD	<.001

Notes. See notes on Table 1. "ASD > TD" under Sig. Contrasts indicates that all pairwise contrasts between TD and ASD were significant (i.e., ASDf > TDf, ASDm > TDm, ASDf > TDm, ASDm > TDf). a-c: One data collection site did not provide Depressive problems scale scores; non-missing *N*'s: a = 106; b = 97*;* c = 118. **Abbreviations:** *Emp.*: Empirically based. *AD/H*: Attention-deficit/hyperactivity. *Opp*: oppositional.

Supplementary Table 7 | Subject distribution and signal to noise ratio across scanning sites.

A. Subject counts by scanner, sex, and group.

	ASD		TD	
	Female	Male	Female	Male
	n	*n*	*n*	*n*
Site A.Trio	4	5	7	6
Site B.Prisma	6	12	16	19
Site B.Trio	10	5	4	4
Site C.Prisma	3	5	8	7
Site C.Trio	10	13	9	12
Site D.Trio	13	8	10	11

B. tSNR across sites, by phantom type.

	A.Trio	B.Prisma	B.Trio	C.Prisma	C.Trio	D.Trio
	M(SD)	*M(SD)*	*M(SD)*	*M(SD)*	*M(SD)*	*M(SD)*
Sphere	122.46(5.82)	178.50(NA)[a]	155.00(8.43)	178.30(NA)[a]	154.02(7.02)	182.56(4.15)
Human 1	145.84(23.57)	153.78(22.12)	162.24(15.90)	NA	148.70(11.41)	156.23(8.44)
Human 2	167.78(11.34)	NA	NA	NA	NA	162.68(11.23)

C. Pairwise comparisons of tSNR values across sites.

Site comparison	*B*	*SE*	*z*	*p*
B.Prisma-A.Trio	15.64	8.52	1.84	0.419
B.Trio-A.Trio	15.85	7.15	2.22	0.211
C.Prisma-A.Trio	34.77	17.82	1.95	0.348
C.Trio-A.Trio	8.59	7.15	1.20	0.822
D.Trio-A.Trio	22.48	6.34	3.55	0.005
B.Trio-B.Prisma	0.22	8.92	0.02	1.000
C.Prisma-B.Prisma	19.13	18.77	1.02	0.903
C.Trio-B.Prisma	-7.04	8.92	-0.79	0.966
D.Trio-B.Prisma	6.85	8.56	0.80	0.964
C.Prisma-B.Trio	18.91	18.01	1.05	0.891
C.Trio-B.Trio	-7.26	7.62	-0.95	0.926
D.Trio-B.Trio	6.63	7.20	0.92	0.935
C.Trio-C.Prisma	-26.17	18.01	-1.45	0.673
D.Trio-C.Prisma	-12.28	17.83	-0.69	0.981
D.Trio-C.Trio	13.89	7.20	1.93	0.359

Note. Mean (*M*) and standard deviation (*SD*) are calculated across the average temporal signal to noise ratio (tSNR) computed from each of the five functional imaging protocols included in the overall multisite project. *p*-values are adjusted using Tukey's HSD. [a]: Only one functional imaging protocol (resting state) was run on the phantom at these sites.