nature research

Corresponding author(s): Michael Lombardo, Eric Courchesne

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

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

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Software used for data analysis included Illumina GenomeStudio software (version 1.1.1), R version 3.4.2, RStudio (version 1.1.383), various R libraries available on CRAN, GitHub, and Bioconductor (e.g., easypackages, ggplot2, psych, here, patchwork, CellCODE, limma, qvalue, WGCNA, gplots), MATLAB R2017b, plsgui MATLAB toolbox, SPM8 MATLAB Toolbox, AFNI version 17.3.00, MetaCore GeneGO software version 5.0. Custom code for implementing all analyses can be found at https://github.com/mvlombardo/asdlangoutcomebloodgexfmripls

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

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The raw data that support the findings from this study are publicly available from the NIH National Database for Autism Research (NDAR). Raw blood leukocyte gene expression data is publicly available via Gene Expression Omnibus (GEO) (GSE42133; GSE111175). Song bird area X gene expression data is publicly available on GEO (GSE34819). GTEx data is publicly available at https://gtexportal.org. ASD post-mortem cortical gene expression can be found at https://github.com/dhglab/Genome-wide-changes-in-lncRNA-alternative-splicing-and-cortical-patterning-in-autism.

Field-specific reporting

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

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

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	No statistical methods were used to pre-determine sample sizes, but our sample sizes are currently amongst the largest of any fMRI study to date on ASD at very early ages in toddlerhood. Furthermore, for the primary hypothesis test of the manuscript (i.e. partial least squares analysis), statistical power analyses have not been formally developed for such analyses and thus a priori power calculations could not be done.
Data exclusions	Gene expression data were taken from a larger superset of data available on ASD and TD toddlers. Quality control analysis of this superset was performed to identify and remove 23 outlier samples. Samples were marked as outlier if they showed low signal intensity (average signal two standard deviations lower than the overall mean), deviant pairwise correlations, deviant cumulative distributions, deviant multi-dimensional scaling plots, or poor hierarchical clustering, as described elsewhere (Pramparo et al., 2015, Molecular Systems Biology, 11, 841). From this high-quality superset, we utilized the maximum number of ASD and TD subjects whom also had fMRI data available.
Replication	A replication dataset for the primary PLS analyses was not available. However, for enrichment analyses, when possible, we utilized two independent gene lists from different studies (e.g., human-specific genes, ASD prenatal genes, FMRP or CHD8 targets) in to identify replicable enrichments.
Randomization	Random allocation of participants to groups is not applicable because ASD and TD labels belong to specific groups of participants. No other randomization procedures were implemented as part of the data collection process.
Blinding	Data collection and analyses were not performed blind to the conditions of the experiment.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	A total of n=118 toddlers were scanned with fMRI and had available gene expression data. From these 118 toddlers, n=81 ASD individuals were examined and were split into 2 language outcome subtypes. n=41 individuals with ASD (34 male, 7 female) were classified as 'poor' language outcome (ASD Poor), based on the criteria of having both Mullen EL and RL T-scores more than 1 standard deviation below the norm of 50 (i.e. T<40) at the final testing time-point (mean age at fMRI scan = 29.53 months, SD at fMRI scan = 8.04, range = 12-46 months). Another n=40 individuals with ASD (30 male, 10 female) were classified as 'good' language outcome (ASD Good), based on having either Mullen EL or RL T-scores greater than or equal to 40 (i.e. T ≥ 40) at the final testing time-point (mean age at fMRI scan = 29.73 months, SD at fMRI scan = 8.51, range = 12-45 months). The usage of the term 'Good' here is not used to refer to ability level in absolute terms, but more reflects ability relative to the ASD Poor subgroup. These ASD subtypes were compared to n=37 typically-developing toddlers (21 male, 16 female; mean age at fMRI scan = 26.19 months, SD at fMRI scan = 10.20, range = 12-45 months). ASD subtypes and TD did not statistically differ in age at the time of scanning (F(2,115) = 1.87, p = 0.15).
Recruitment	Identical to the approach used in our earlier studies, toddlers were recruited through two mechanisms: community referrals (e.g., website) or a general population-based screening method called the 1-Year Well-Baby Check-Up Approach that allowed for the prospective study of ASD beginning at 12 months based on a toddler's failure of the CSBS-DP Infant-Toddler Checklist. All toddlers were tracked from an intake assessment around 12 months and followed roughly every 12 months until 3–4 years of age. All toddlers, including normal control subjects, participated in a series of tests collected longitudinally across all visits, including the Autism Diagnostic Observation Schedule (ADOS; Module T, 1, or 2), the Mullen Scales of Early Learning, and the Vineland Adaptive Behavior Scales. All testing occurred at the University of California, San Diego Autism Center of Excellence (ACE). We are not aware of any self-selection or other biases likely to impact the recruitment of our cohort.

Magnetic resonance imaging

Experimental design

Design type	Block design
Design specifications	Number of speech blocks = 9 Length of each block = 20 secs Interval between blocks = 20 secs
Behavioral performance measures	Since toddlers were scanned during natural sleep, no behavioral measures were collected during scanning.

Acquisition

Imaging type(s)	functional
Field strength	1.5T
Sequence & imaging parameters	Imaging data were collected on a 1.5 Tesla General Electric MRI scanner during natural sleep at night; no sedation was used. High-resolution T1-weighted anatomical scans were collected for warping fMRI data into standard atlas space. Blood oxygenation level-dependent (BOLD) signal was measured across the whole brain with echoplanar imaging during the language paradigm (echo time = 30 ms, repetition time = 2,500 ms, flip angle = 90 degrees, bandwidth = 70 kHz, field of view = 25.6 cm, in-plane resolution = 4 x 4 mm, slice thickness = 4 mm, 31 slices).
Area of acquisition	whole-brain coverage

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Preprocessing of functional imaging data was implemented within the Analysis of Functional NeuroImages (AFNI) software package. The preprocessing pipeline was comprised of motion correction, normalization to Talairach space, and smoothing (8mm full-width at half-maximum (FWHM) Gaussian kernel).
Normalization	linear normalization to Talairach space
Normalization template	Talairach
Noise and artifact removal	Motion parameters were included as regressors in the GLMs
Volume censoring	No volume censoring was done

Statistical modeling & inference

Model type and settings	First-level and second-level mass-univariate whole-brain activation analyses were modeled with the general linear model (GLM) in SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). Events in first-level models were modeled using the canonical hemodynamic response function and its temporal derivative. All first-level GLMs included motion parameters as

covariates of no interest. High-pass temporal filtering was applied with a cutoff of 0.0078 Hz (1/128 seconds) in order to remove low frequency drift in the time series.

Effect(s) tested: Speech vs Rest

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

Statistic type for inference (See Eklund et al. 2016): Voxel-wise

Correction: Voxel-wise FDR

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis: Gene co-expression modules were summarized by the module eigengene (first principal component of genes from a module). Module eigengene values were inserted into the PLS analysis as the gene expression dataset, while whole-brain t-maps from the Speech vs Rest contrast were inserted into the PLS analysis as the neuroimaging dataset. A permutation test (10,000 permutations) was done for hypothesis testing of latent-variable pairs and bootstrapping was conducted to compute bootstrap ratios for brain and gene expression variables.