Stem Cell Reports Article Dynamical Electrical Complexity Is Reduced during Neuronal Differentiation in Autism Spectrum Disorder

Debha N Amatya, Sara B Linker, Ana P.D. Mendes, Renata Santos, Galina Erikson, Maxim N Shokhirev, Yuansheng Zhou, Tatyana Sharpee, Fred H Gage, Maria C Marchetto, et al.

Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Dynamical Electrical Complexity Is Reduced during Neuronal Differentiation in Autism Spectrum Disorder

Debha N. Amatya,[1,2] Sara B. Linker,[1] Ana P.D. Mendes,[1] Renata Santos,[1,3] Galina Erikson,[4] Maxim N. Shokhirev,[4] Yuansheng Zhou,[5,6] Tatyana Sharpee,[5] Fred H. Gage,[1,8] Maria C. Marchetto,[1,*] and Yeni Kim[1,7,*]

[1]The Salk Institute, Laboratory of Genetics, La Jolla, CA 92037, USA
[2]University of California San Diego, Department of Neurosciences, La Jolla, CA 92093, USA
[3]Institute of Psychiatry and Neuroscience of Paris (UMR_S1266 INSERM, University of Paris), Laboratory of Dynamic of Neuronal Structure in Health and Disease, Paris, France
[4]The Salk Institute, Integrative Genomics and Bioinformatics Core, La Jolla, CA 92037, USA
[5]The Salk Institute, Computational Neurobiology Laboratory, La Jolla, CA 92037, USA
[6]University of California San Diego, Division of Biological Sciences, La Jolla, CA 92093, USA
[7]Department of Child and Adolescent Psychiatry, National Center for Mental Health, 127 Yongmasanro, Gwangjin-gu, Seoul 04933, South Korea
[8]Senior author
*Correspondence: marchetto@salk.edu (M.C.M.), yenikim@korea.kr (Y.K.)
https://doi.org/10.1016/j.stemcr.2019.08.001

SUMMARY

Neuronal activity can be modeled as a nonlinear dynamical system to yield measures of neuronal state and dysfunction. The electrical recordings of stem cell-derived neurons from individuals with autism spectrum disorder (ASD) and controls were analyzed using minimum embedding dimension (MED) analysis to characterize their dynamical complexity. MED analysis revealed a significant reduction in dynamical complexity in ASD neurons during differentiation, which was correlated to bursting and spike interval measures. MED was associated with clinical endpoints, such as nonverbal intelligence, and was correlated with 53 differentially expressed genes, which were overrepresented with ASD risk genes related to neurodevelopment, cell morphology, and cell migration. Spatiotemporal analysis also showed a prenatal temporal enrichment in cortical and deep brain structures. Together, we present dynamical analysis as a paradigm that can be used to distinguish disease-associated cellular electrophysiological and transcriptional signatures, while taking into account patient variability in neuropsychiatric disorders.

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental condition defined by deficits in communication, social interaction, and repetitive behavior that affects 76 in 10,000 children worldwide (Baxter et al., 2015; Geschwind and State, 2015). Large gene sequencing studies and structural variant analyses have uncovered that rare coding and *de novo* mutations contribute to ASD by targeting molecular processes related to synaptic functioning, chromatin modification, and prenatal cortical development (Geschwind and State, 2015; Parikshak et al., 2013; Sebat et al., 2007; Willsey et al., 2013). Despite this convergence at the gene pathway level, individual mutations are extremely heterogeneous, with none accounting for more than 1% of ASD cases (Devlin and Scherer, 2012). Previous studies have shown a correlation between early brain overgrowth, as measured by head circumference and neuroimaging, and an increased risk for ASD (Courchesne et al., 2003; Hazlett et al., 2011; Shen et al., 2013). In this subset of the ASD population, postmortem analysis has revealed an excess of neurons in the first 3 years of life, and seminal studies using induced pluripotent stem cell (iPSC)-derived neurons have uncovered increased proliferation of neural progenitors and reduced synaptogenesis (Courchesne et al., 2011; Marchetto et al., 2017).

A promising alternative to overcome the challenges of *in vivo* human brain studies are neuronal *in vitro* cultures derived from human iPSCs (Vadodaria et al., 2018). iPSC-derived neurons permit the study of neuronal firing, which may provide insight into the fundamental mechanisms of neuropsychiatric diseases at the cellular level. One of the technologies utilized to study neuronal firing is the multielectrode array (MEA). Multielectrode arrays (MEAs) enable high-throughput, longitudinal recordings of extracellular electrical dynamics from populations of neurons at millisecond resolution, facilitating the analysis of neurons *in vitro* (Kreuz, 2011; Samengo and Elijah, 2013; Taketani and Baudry, 2006). The analysis of electrical recordings typically involves the quantification of spiking-related variables, such as firing rate, spike morphology, and network measures. Electrophysiological analyses of neurons derived from ASD patients have revealed functional defects, such as reductions in neuronal firing, disrupted postsynaptic currents, and imbalanced excitatory/inhibitory tone (Derosa et al., 2018; Liu et al., 2017; Marchetto et al., 2017; Mariani et al., 2015).

The human brain is a complex nonlinear system comprising multiple interacting components at a variety of spatiotemporal scales. This complexity poses a challenge to scientific investigation, but it also provides an opportunity to apply dynamical analysis tools to characterize

neuronal activity (Röschke and Başar, 1988). Minimal embedding dimension (MED) analysis is a mathematical algorithm to determine the dynamical complexity of a time-series recording (Kennel et al., 1992). In this framework, a neuronal network is thought of as a system composed of a finite number of differential equations that govern all possible states of the system. The purpose of MED analysis is to empirically estimate the number of governing equations, which is a measure of complexity, by iteratively forming embeddings on an electrical time-series recording (Kennel et al., 1992; Takens, 1981). Similar techniques have been applied to electroencephalogram (EEG) and magnetoencephalogram recordings in autism, schizophrenia, bipolar disorder, and depression (Akdemir Akar et al., 2015; Bosl et al., 2011; Fernández et al., 2018; Jeong et al., 1998). These studies suggested that dynamical complexity may be increased in mood disorders, such as bipolar disorder and depression, and decreased in developmental conditions, such as schizophrenia and ASD.

Given the evidence of disrupted synapse formation in previous studies (16), we hypothesized that electrical activity during differentiation from the neural progenitor to neuronal stage would be marked by a reduction in dynamical complexity in ASD. We applied MED analysis to the time-series recordings of neurons derived from ASD and neurotypical control subjects. In addition, we aimed to demonstrate that MED analysis of ASD could be used in combination with gene expression analysis to capture disease-relevant genetic mechanisms.

ASD is a highly heterogeneous disorder, which challenges attempts to ascertain disease-related transcriptional changes, given only a binary case status. We show here that the MED supplies a continuous rather than binary variable that is associated with disease severity and can be used to more accurately distinguish disease-associated transcriptional signatures, while taking into account patient variability. Through the application of dynamical analysis techniques to patient-derived neuronal recordings, we hope to introduce this paradigm for the interrogation of disease signatures in neuropsychiatric disorders.

RESULTS

Study Design and MED Analysis Technique

This study was designed to identify a dynamical electrophysiological cellular phenotype associated with idiopathic ASD, as well as the gene expression correlates of this signal. To accomplish this, iPSC lines derived from eight ASD patients with early brain overgrowth and seven neurotypical controls were obtained and differentiated into neurons using a pan-cortical protocol (Marchetto et al., 2010). A schematic of the study design is depicted in Figure 1A, and a description of the samples is given in Table 1. Additional clinical metadata for these samples is presented in Table S1. From the neural progenitor phase, longitudinal electrical recordings were gathered with MEA over the course of 47 days. For each recording, standard spike-related variables, as well as MED, a measure of dynamical complexity, were computed. On day 15 post differentiation, the developing cells were sorted for neuronal identity using the PSA-NCAM marker and then RNA sequenced. The subsequent analysis focused on linking the observed dynamical cellular electrical phenotype with gene expression changes that relate to ASD biology. This was done by using the average MED values on day 15, the same day RNA sequencing was performed, for each subject as a variable in the gene expression models fit to the RNA-sequencing data. Figure 1B depicts how dynamical complexity may vary across samples, even when standard measures such as the firing rate remain constant. In such cases, nonlinear dynamical differences may uncover patterns of disruption in gene expression not identified by binary case-control labels. Further intuition for MED analysis is provided by a simulated example of the Hénon map, a two-dimensional dynamical system, in Figure S1.

Neuronal Electrical Recordings and Dynamical Analysis

To glean knowledge from this *in vitro* model of ASD, both standard analytical tools and nonlinear dynamical methods were applied to the electrical recordings of the neurons. This approach is appropriate, as neuronal activity is a nonlinear dynamical system that arises from the firing of nascent neuronal networks. Representative electrical recordings and spiking events are shown in Figure 2A. Traditional metrics, such as the mean firing rate, fail to show statistical separation between the ASD and control group. However, more complex measures, such as bursting and the variation of interspike intervals (ISI) within bursts, are able to show a difference between groups, suggesting that dynamical analysis may also capture features related to bursting or spike timing that fundamentally distinguish ASD and control electrical activity (Figures 2B and 2C). This is further supported by the strong correlation observed between the MED and number of bursting electrodes ($R = 0.714$, $p \leq 0.001$) and ISI coefficient of variation ($R = 0.492$, $p \leq 0.001$).

MED was computed to quantify the dynamical complexity of the electrical recordings. A previous study provided evidence that iPSC-derived ASD neurons were characterized by diminished synaptogenesis, thus compromising the process of forming connected networks (Marchetto et al., 2017). It was expected that MED would capture this ASD-related deficit of electrophysiological complexity, thereby uncovering a dynamical cellular

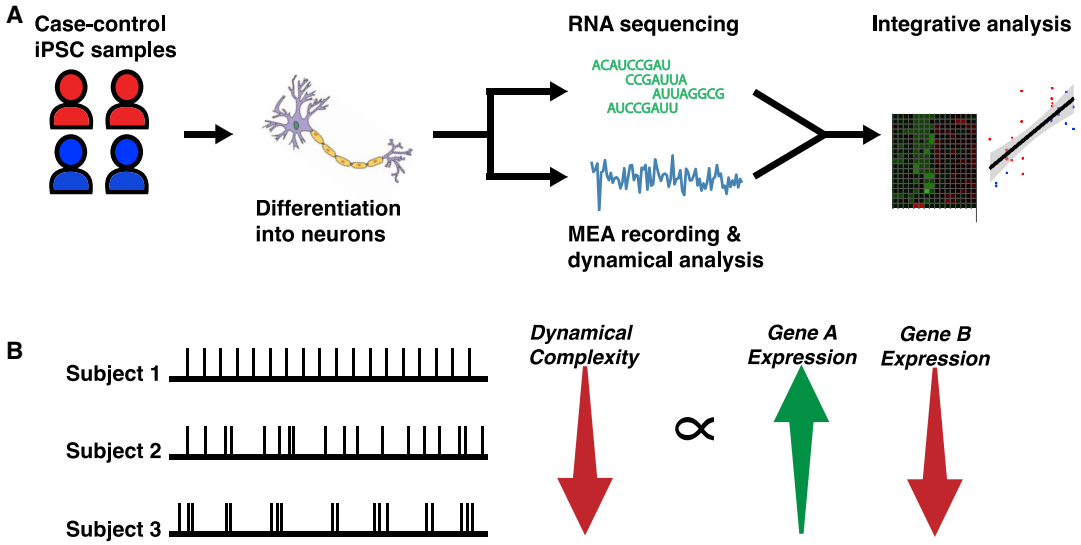

Figure 1. Study Design and Visualizing Dynamical Complexity

(A) Conceptual overview of the study design. iPSCs from seven neurotypical controls and eight ASD cases were differentiated into neurons. Neurons were characterized through RNA sequencing and MEA recordings to identify disease signatures that integrate across both levels of analysis.

(B) Dynamical complexity may vary even when standard spiking variables are constant. For these pedagogical examples, the firing rate is constant at 20 spiking events in the interval. Nevertheless, differences in the organization of the spiking events may result in large differences in complexity, as measured by dynamical analysis. These differences may distinguish groups of interest, such as ASD, and correlate to gene expression changes related to neuronal dysfunction.

See also Figure S1.

electrical phenotype for this disorder. A difference between ASD and control MED was found to emerge after 11 days of differentiation and persist with statistical significance for the next 11 days of recording, as shown in Figure 2D. Correlation analysis between MED and bursting variables revealed a significant association with the number of bursting electrodes and the ISI coefficient of variation, strengthening the relationship between dynamical complexity and coordinated spiking activity and variation. Despite this relationship, the variance associated with MED was less than that of bursting variables, suggesting improved precision in the characterization of electrical activity. Figure S2 details average subject-wise MED boxplots across the key time points of the study, including the day when RNA sequencing was performed.

To assess the behavioral relevance of MED, the correlation between subject MED values and cognitive scores was examined. This revealed a significant correlation in two relevant endpoints, the Vineland ABC Score and Nonverbal IQ, as shown in Figure 2E. This suggests that the cellular dynamic state may have a meaningful impact on cognitive and behavioral states relevant to ASD. Therefore, MED is a dynamical cellular electrophysiological phenotype that is related to both ASD status and some of its relevant clinical features.

Differential Expression Models for ASD and MED

RNA sequencing was performed at 15 days post differentiation in order to explore the gene expression correlates of MED and ASD. Given that MED captures both individual and ASD-related electrical dynamics, we expected that it would be associated with robust and biologically relevant transcriptomic alterations (Figure S2). To test this rationale, gene expression was examined with respect to (1) MED and (2) ASD interacting with MED in separate models. Two subject outliers were excluded based on principal component analysis (Figure S3). As shown in Figure 3, before false discovery rate (FDR) correction, the MED model identified 1,423 differentially expressed genes, and the ASD interaction model was associated with 761 genes. After controlling for multiple comparisons, MED was associated with 53 genes in comparison with 7 for ASD in each respective model. To assess the validity of this result, we permuted the MED values among the subjects and found that the original analysis identified more significant genes than 72% of the permuted trials, and the identified MED genes overlapped with genes that are highly expressed in the brain more strongly than in 91% of permuted trials. The comprehensive lists of differentially expressed genes are given in Tables S2 and S3, and examples of genes differentially expressed with respect to MED value are shown in

Table 1. Characteristics of Subjects from which iPSC Lines Were Generated

	iPSC Lines (N = 15)	
	Control	ASD
Samples	7	8
Brain volume (cm^3)	1,237.2 ± 86.6	1,372.9 ± 87.8
Age (years)	25.0 ± 27.7	13.3 ± 5.6
IQ	118.6 ± 8.6	67.9 ± 14.9
ADOS	–	16.75 ± 2.8
Vineland ABC	99 ± 6.2	57.13 ± 15.1

Age, IQ, ADOS, and Vineland ABC scores were recorded at the time of biopsy. Brain volume was computed via magnetic resonance imaging during early childhood. All samples were derived from males. ASD cases met the diagnostic criteria as defined by ADOS total score cutoffs or guidelines from the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Where relevant, quantities are shown as averages ± one standard deviation. IQ, Wechsler Intelligence Quotient; ADOS, Autism Diagnostic Observation Schedule; Vineland ABC, Vineland Adaptive Behavioral Composite, second edition.

Figure S4. Of note, a model that only included the ASD status without MED was unable to find any differentially expressed genes after FDR correction. This result suggests that cell-derived electrical phenotypes could represent a path forward in studying complex genetic disorders, such as idiopathic ASD. Binary ASD-control labels do not account for the tremendous intersubject heterogeneity present in such samples, and cellular markers may better model both individual and group-related variance for gene expression analyses.

Interrogation of the MED Signature

The gene expression signature of the MED was examined in relation to known ASD genes and pathways, as well as brain regions and stages of development (refer to the Experimental Procedures for details). For these comparisons, the full set of differentially expressed MED and ASD genes, before FDR correction, was considered. Figure 4A shows that both ASD and MED genes are overrepresented in brain-expressed and putative ASD genes. However, the enrichment effect sizes were markedly larger for MED genes. This finding provides evidence that MED is relevant to physiological and disease-related gene lists. Next, biological pathways were analyzed through gene ontology analysis, as shown in Figure 4B. MED genes were enriched in neurodevelopmental, cell migration, junction assembly, and regulatory terms. Additional exploration of cellular components related to the MED gene signature showed a significant overlap with neuronal components, such as the synapse, postsynaptic density, axon, and dendritic spines. Full results for gene ontology analysis testing of biological processes and cellular components are given in Tables S4 and S5, respectively. These findings suggest that MED encodes a measure of neuronal maturation and participation in circuit formation. Finally, to localize the set of MED genes to developmental time points and brain regions, data from the BrainSpan Atlas were used to identify enrichments in specific region-stage pairs (refer to the Experimental Procedures for details). The enrichments were visualized as a matrix in Figure 4C for all region-stage pairs, and regional involvement during the highly enriched, late prenatal time point was mapped onto a representation of the brain in Figure 4D. Table S6 lists the exact p values obtained for the spatiotemporal analysis. This analysis uncovered a strong prenatal temporal enrichment in a mixture of cortical and deep structures (e.g., the visual cortex, inferior temporal cortex, and thalamus), as well as a secondary involvement of late childhood cortical regions, suggesting that MED-related ASD pathology might play out at various stages of development with diverse structural involvement.

DISCUSSION

In this study, we propose a cellular electrical phenotype for idiopathic ASD, the MED, which is derived from the dynamical complexity of neuronal electrical recordings. Previous *in vitro* studies examining ASD neurons have demonstrated both single unit recording abnormalities and alterations in spiking and bursting properties using MEA (Liu et al., 2017; Marchetto et al., 2017; Mariani et al., 2015). Here, we describe ASD-related nonlinear dynamical electrophysiologic alterations using iPSC-derived neurons. Other electrophysiological variables related to bursting also successfully separated the case and control groups, but with less precision than MED. This finding suggests that dynamical complexity may be related to the integrated and synchronized firing of neuronal networks in culture, rather than the isolated behavior of single units. As differentiation from the neural progenitor phase begins, MED increases as a function of neuronal bursting and spike timing properties and peaks after 2 weeks of maturation. The trajectory of this complexity increase is altered in the ASD group, as these samples fail to match the MED found in the control samples.

The divergence in dynamical complexity occurs early on during the first few weeks of differentiation past the neural progenitor phase. A recent study using the same ASD samples revealed that transient pathological priming of expression networks occurs around this time point, supporting the relevance of this early developmental period (Schafer et al., 2019). In our study, MED deficits in the ASD group begin to lessen after 2 weeks of recording, along with a

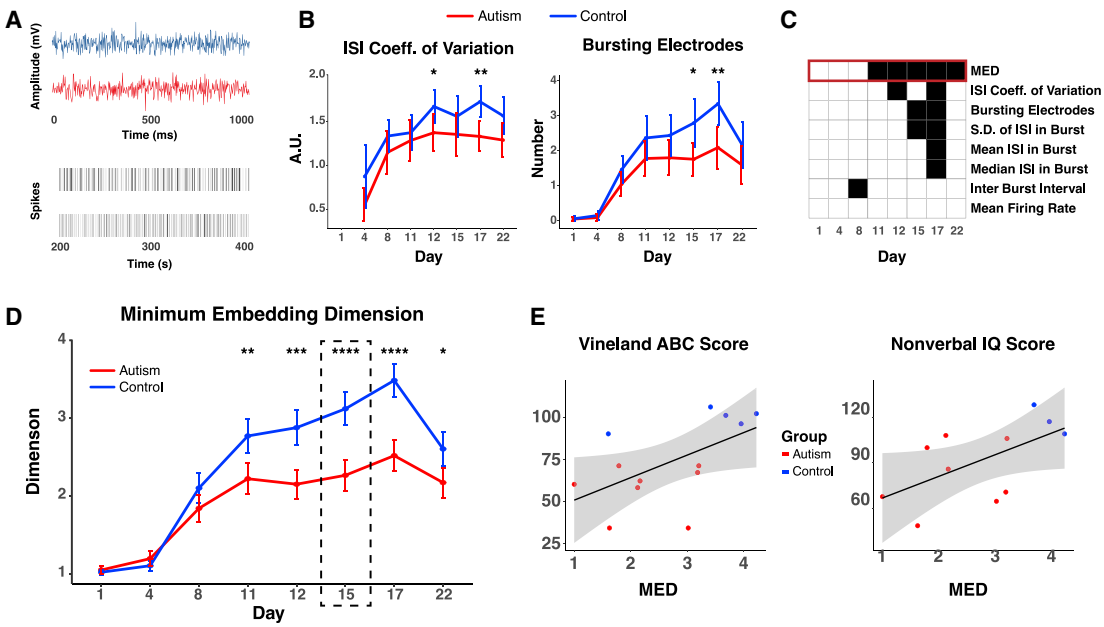

Figure 2. Electrical Analysis of Neuronal Lines, Minimum Embedding Dimension, and Clinical Correlations

(A) Raw MEA data. Raw electrode waveforms (top) are used for spike detection (bottom) and downstream analyses.

(B) Spike interval and bursting variables highlight electrophysiological differences. Group average values with 95% confidence intervals are given for several time points of two variables. Variation in the interspike interval and number of bursting electrodes does show significant group-based differences for day 15 and 17 between control (blue) and ASD (red). For each relevant panel, significance was as tested with Welch's two-sided t test and indicated by asterisks. *$p \leq 5 \times 10^{-2}$, **$p \leq 5 \times 10^{-3}$, ***$p \leq 5 \times 10^{-5}$, and ****$p \leq 5 \times 10^{-7}$.

(C) Overview of group-wise differences across spiking measures and MED. This binary matrix indicates at which time points a measure was able to detect a significant difference between cases and controls (black squares). The MED outperforms all measures in distinguishing groups, but it does overlap with some bursting and spike interval variables, as shown in (B). Relatively common measures, such as the mean firing rate, fail to distinguish control and ASD activity over any day of the recording period.

(D) MED offers a sustained and statistically significant separation of groups. Average MED values and 95% confidence intervals for both groups are depicted for the first eight time points. The control subjects (blue) are associated with higher MED complexity score in comparison with the ASD subjects (red). The dotted rectangle indicates the values of the MED during day 15 of the MEA recordings, when RNA sequencing was performed.

(E) MED is correlated to clinical endpoints of interest. Low MED scores, which are associated with ASD diagnosis, are correlated to lower Vineland adaptive behavior score (left) and lower early nonverbal IQ (right). The gray shading represents a 95% confidence interval around the fitted curve, as estimated by a linear model.

See also Figure S2.

general trend of decreasing activity noted across measures in both groups, perhaps due to the pruning of weakly connected neurons and the establishment of more mature circuits. The convergence of MED at later time points suggests that the electrical complexity gap may be a transient phenomenon in development that nevertheless disrupts the early formation of neural connections and contributes to ASD pathology at later time points. Patient EEG studies have shown that recordings taken from people diagnosed with ASD also exhibit reduced dynamical complexity compared with controls (Bosl et al., 2011). The mechanism by which neuronal complexity deficits may propagate to the circuit and brain-wide levels to drive pathology remains an interesting avenue for further study.

Interestingly, MED was highly correlated with the variability in interspike intervals, indicating that dynamical complexity is sensitive to heterogeneity in the rate of spiking (Figure 2B). As suggested in Figure 1B, neurons firing with a constant rhythm would have a lower MED than neurons firing with an ever-changing rhythm, which would require more embedding dimensions to model. At a molecular level, we found that the MED differentially expressed genes are associated with ASD risk genes and neurodevelopmental pathways, such as neurogenesis, cell migration, and cell morphology. Examination of cellular components implicated in the MED gene expression signature localized to the neuron, implicating the synapse, dendrites, and axons (Table S5). These findings agree with

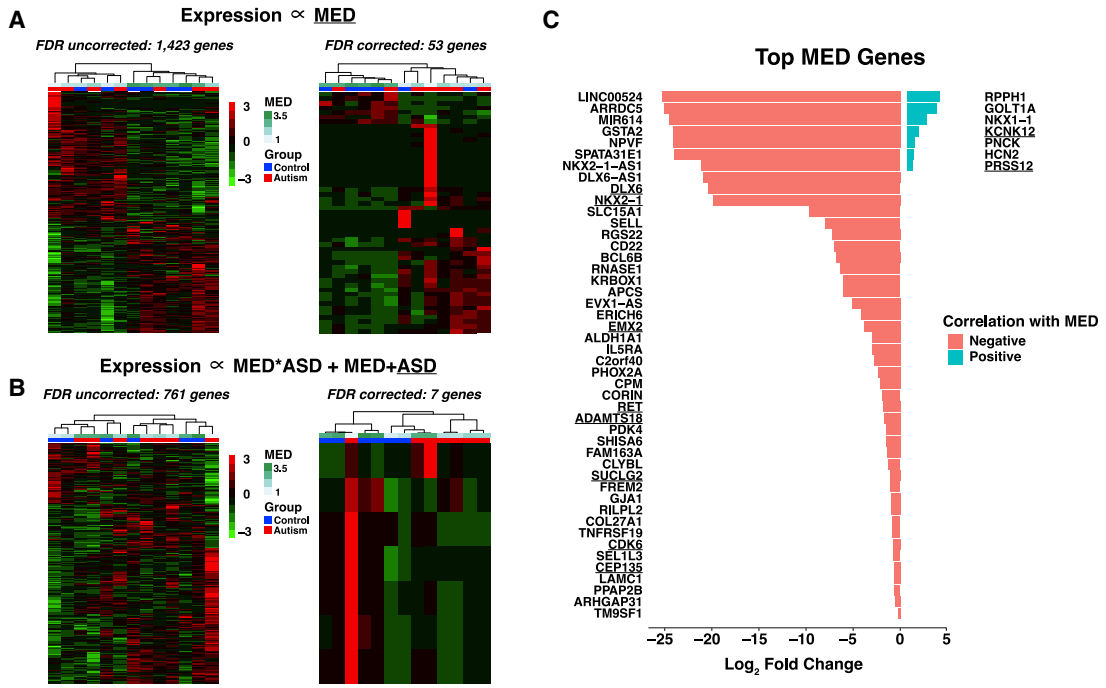

Figure 3. Gene Expression Signatures of MED and ASD

(A) MED is associated with differential expression in 1,423 genes before multiple comparison correction and 53 genes after correction, as depicted in the heatmaps. Hierarchical clustering of the samples based on gene expression separates low- and high-complexity samples, as indicated by the MED color bar.

(B) Examination of ASD genes in a combined MED and ASD interaction model identifies 761 differentially expressed genes before correction and 7 differentially expressed genes after multiple comparison correction.

(C) The $\log_2$ fold changes of the 53 differentially expressed genes after multiple comparison correction for MED are shown. The red bars correspond to genes that are negatively correlated to MED, and the blue bars correspond to genes that are positively correlated to MED. Underlined genes are also implicated in ASD risk or brain function in previous studies. Most differentially expressed genes are negatively correlated to MED, and positively correlated genes have less variability in their log fold changes.

See also Figure S4.

sequencing studies that have shown ASD affects a variety of neurodevelopmental processes, such as neuronal proliferation, migration, and synaptogenesis (Courchesne et al., 2018; De Rubeis et al., 2014; Krumm et al., 2014; Marchetto et al., 2017; Mariani et al., 2015; Parikshak et al., 2013; Pinto et al., 2014; Voineagu et al., 2011; Willsey et al., 2013). While it is unclear how spiking dynamics unfold in the developing human brain, our finding of decreased complexity of ASD neurons in culture indicates that early neuronal physiology is disrupted. The clear and sustained reduction in complexity found in the ASD group likely arises from deficiencies in synaptogenesis, which were previously reported in this cohort (Marchetto et al., 2017). Difficulty in forming synaptic connections *in vitro* diminishes the underlying complexity of their electrical outputs, as measured by MED.

The detection of molecular changes related to idiopathic ASD is complicated by the immense genetic heterogeneity associated with the disorder (De Rubeis et al., 2014; Devlin and Scherer, 2012). Although the sample size in this study is much smaller than those typically recruited for population genomics analyses, this iPSC-based approach offers a major advantage over other methods in that cellular variables relevant to disease may be directly recorded and studied in relation to gene expression changes (Vadodaria et al., 2018). In this manner, we are able to find molecular changes associated with idiopathic ASD by tapping into the fundamental electrical dynamics that lie at the core of the disease. Notably, MED was associated with a broader differential gene expression signature than ASD status, and these genes were associated with fetal cortical development, supporting the growing body of literature that suggests ASD disrupts the normal formation of brain circuits during the early stages of life (Parikshak et al., 2013; Willsey et al., 2013). Therefore, the utility of MED as a cellular electrical phenotype for ASD is demonstrated both directly, through electrical recordings, and indirectly, through the examination of gene expression trends related to MED.

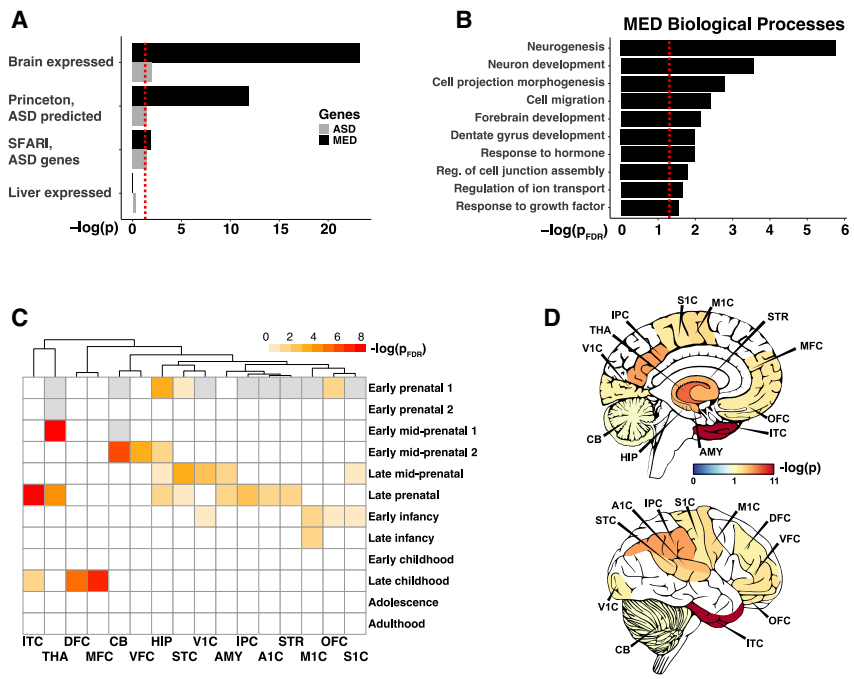

Figure 4. Enrichment of MED Gene Expression Signature with ASD-Related Genes, Biological Pathways, and Spatiotemporal Analysis

To interpret the broad biological relevance of MED and ASD expression signatures, differentially expressed genes before FDR correction were further analyzed.

(A) The MED (black) and ASD (gray) differentially expressed genes were tested for enrichment in five lists: (1) highly brain-expressed genes, (2) putative ASD risk genes, (3) gold standard ASD risk genes, and (4) highly liver-expressed genes (negative control). Significance was tested using the binomial test with a cutoff of p = 0.05 (red line). Both ASD and MED genes were enriched in brain, Princeton, and SFARI lists, although the effect size of the MED signature was notably stronger. Neither set overlapped with liver-expressed genes, a negative control.

(B) Biological processes implicated with MED. The gene expression signature of MED affects neurodevelopmental, cell migration, and growth-related ontologies. Significance of enrichment was calculated using Fisher's exact test with false discovery rate control and a cutoff of $p_{FDR} = 0.05$ (red line).

(C) Gene lists from 16 neuroanatomical regions and 13 developmental stages were tested for enrichment with the MED genes. The neuroanatomical regions include the inferior temporal cortex (ITC), thalamus (THA), dorsal frontal cortex (DFC), medial frontal cortex (MFC), cerebellar cortex (CB), ventral frontal cortex (VFC), hippocampus (HIP), superior temporal cortex (STC), primary visual cortex (V1C), amygdala (AMY), inferior parietal cortex (IPC), primary auditory cortex (A1C), striatum (STR), primary motor cortex (M1C), olfactory cortex (OFC), and primary somatosensory cortex (S1C). Fisher's exact test was performed for each region-stage pair and plotted on a heatmap after false discovery rate correction. A variety of cortical and deeper structures are implicated in MED, primarily at prenatal time points. Strong enrichment of the DFC and MFC is also indicated during late childhood. The grayed-out regions represent structures that are not present in the early prenatal brain.

(D) Visualization of cortical and interior MED-associated regions during late prenatal development (25–28 post conceptual weeks). Raw enrichment p values (p_{raw}) were plotted to show the range of structural involvement at this time point. Cortical regions such as the A1C, S1C, and ITC are enriched for the MED signature; however, deeper lying structures such as the thalamus and striatum are also revealed.

In summary, we demonstrate that nonlinear dynamical analysis can be applied to electrical recordings from patient-derived neurons to reveal differences in complexity between ASD and control subjects. The MED is associated with gene expression changes that both validate existing genetic and mechanistic studies and present less well-described findings, such as the implication of fetal basal brain regions and late childhood cortical development in ASD mechanisms. iPSC-derived models are a rich platform to study neurodevelopmental disorders, because they allow for the examination of cellular disease mechanisms in otherwise inaccessible tissue. Further study of MED as a cellular electrical phenotype for ASD will shed light on the utility of this approach. We believe that this integration of methods is a good strategy for tackling the immense heterogeneity associated with idiopathic ASD and hope that it facilitates future studies in autism and related disorders.

EXPERIMENTAL PROCEDURES

iPSC Samples and Neuronal Differentiation

iPSC lines were derived from two sources. First, 13 samples from Marchetto et al. (2017) comprised the bulk of the lines (Marchetto et al., 2017). These included eight lines derived from ASD patients with an early brain overgrowth phenotype and five age-matched neurotypical control lines. Second, two additional control lines from neurotypical adults were added to balance the groups. The samples from the first cohort included standardized clinical and functional assessments, including the Autism Diagnostic Observation Schedule, Wechsler Intelligence Scale, and Vineland Adaptive Behavior Scale. Sample metadata and clinical scores are given in Table S1. Samples were obtained with approval from the internal review board of the Salk Institute for Biological Studies and informed consent of all subjects. Neuronal differentiation was accomplished through differentiation of the iPSCs into neural progenitor cells, followed by the removal of FGF2 to drive maturation

into neurons, as described in the literature (Marchetto et al., 2017, 2010).

MEA Recordings and Spike Analysis

Ninety-six-well MEA plates from Axion Biosystems (San Francisco, CA, USA) were used to record electrical activity of neurons derived from all 15 subjects. Each subject's cells were plated in replicates of six and seeded with 10,000 neural progenitor cells that were induced into neuronal differentiation the next day. Wells were coated with poly-ornithine and laminin before cell seeding. Cells were fed every other day and measurements were taken twice a week before feeding. The Maestro MEA system and AxIS software (Axion Biosystems) were used to record neuronal electrical activity from the plates. Voltages were recorded at a frequency of 12.5 kHz and bandpass filtered between 10 Hz and 2.5 kHz. Spike detection was performed using an adaptive threshold set to 5.5 standard deviations above the mean activity of each electrode. Following 5 min of plate rest time, recordings were performed for 10 min. A total of 15 separate recordings were performed over a time span of 47 days for each subject.

Multielectrode data analysis was performed using the Axion Biosystems Neural Metrics Tool, which calculated standard spike-related variables. Bursts were detected with an adaptive Poisson algorithm for high spiking activity that occurred on a single electrode. Variables were averaged across subject replicates and plotted by group for each day of the recordings. The 95% confidence interval around the mean was also calculated and plotted for each day to aid in the discovery of significant trends.

False Nearest Neighbors and Embedding Dimension Analysis

The false nearest neighbor (FNN) method was proposed by Kennel et al. (1992) to find the MED for time-series dynamic systems (Kennel et al., 1992). In this framework, a time series arises from a dynamical system composed of n differential equations. All possible states of the system can be geometrically represented by an attractor, which is a regular shape formed when the n state variables (e.g., x, y for a two-variable system) are plotted against one another.

When only a single variable is accessible, we can instead use d delayed embeddings of this variable to reconstruct the attractor. For example, for $y(n)$ from time series $x(t)$: $y(n) = (x(n), x(n+T), \ldots, x(n+(d-1)T))$ where T is the time delay and the trajectory evolves by $y(n)$, $y(n+1)$ If the embedding dimension d is lower than the native embedding dimension, the system is not fully unfolded, and the reconstruction fails to represent the dynamics of the system, which leads to FNNs in the geometric space. When iteratively increasing the embedding dimension from d to $d+1$, the FNN percentage decreases during the unfolding process until the MED is reached and the state space is fully represented. Therefore, plotting the FNN percentage against the embedding dimension reveals the MED, a quantitative indicator of the intrinsic complexity of the system. Figure S1 visually demonstrates these concepts in an intuitive manner.

Two criteria were used to define the FNNs, as described in the literature (Kennel et al., 1992). The first criterion is the tolerance threshold parameter, R_{tol}, which measures how the neighbor distances change relative to previous distances when increasing d to $d+1$. If the neighbors are false, adding another dimension would largely increase the neighbor distances during the unfolding process. The second criterion models noise in the data. We introduced a threshold parameter A_{tol} to compare the neighbor distances in dimension $d+1$with the size of the attractors in dimension d. By combining these two criteria, the MED for time-series data can be effectively found, even when the data are noisy.

We set $T = 2$, and we used the same threshold parameters as in Kennel's original study $R_{tol} = 15$ and $A_{tol} = 2$ (Kennel et al., 1992). FNN was implemented in MATLAB R2017b (Natick, MA, USA) and performed on both real MEA data and simulated data from the Hénon map (Hénon, 1976; MATLAB, 2017). An online code repository for the MED analysis is freely accessible to others who wish to perform a similar dynamical complexity analysis of electrical recording data. The GitHub repository is available at this link: https://github.com/gyrheart/FNN.

RNA Sequencing

At 15 days post differentiation, the neurons positive for PSA-NCAM (Anti PSA-NCAM Antibody, Miltenyi Biotec) were sorted using flow cytometry to isolate neuronal fate-committed cells. RNA was extracted after cell sorting on TRIzol LS reagent (Invitrogen) from all 15 neuronal samples. Total RNA was extracted using a DNA Free RNA Kit (Zymo Research) according to the manufacturer's instructions. RNA quality was assayed using an Agilent Technologies 2200 TapeStation, and samples with integrity superior to RIN 8.5 were used for library preparation. Stranded mRNA sequencing libraries were prepared using the Illumina TruSeq Stranded mRNA Library Prep Kit according to the manufacturer's instructions. RNA with a poly A tail was isolated using magnetic beads conjugated to polyT oligos. mRNA was then fragmented and reverse transcribed into cDNA. dUTPs were incorporated, followed by second-strand cDNA synthesis. The dUTP-incorporated second strand was not amplified. cDNA was then end repaired, index adapter ligated, and PCR amplified. AMPure XP beads (Beckman Coulter) were used to purify nucleic acid after each step of the library preparation. All sequencing libraries were then quantified, pooled, and sequenced at single-end 50 bp on an Illumina HiSeq 2500 at the Salk Institute Next Generation Sequencing Core.

Then, 1,000 ng of RNA was used for library preparation with the Illumina TruSeq RNA Sample Preparation Kit. The RNAs were sequenced on an Illumina HiSeq 2000 with 50 bp paired-end reads, generating 50 million high-quality sequencing fragments per sample on average. Sequenced reads were quality tested using FASTQC (Andrews, 2010) and aligned to the hg19 (Lander et al., 2001) human genome using the STAR aligner version 2.4.0k (Dobin et al., 2013). Mapping was carried out using default parameters (up to ten mismatches per read, and up to nine multi-mapping locations per read). The genome index was constructed using the gene annotation supplied with the hg19 Illumina iGenomes collection (Illumina, 2015) and sjdbOverhang value of 100. Raw gene expression was quantified across all gene exons (RNA sequencing) using the top-expressed isoform as proxy for gene expression. Transcripts per million (TPM) values were calculated for each sample:

$$TPM_i = \frac{X_i}{l_i} \cdot \left(\frac{1}{\sum \frac{X_j}{l_j}} \right) \times 10^6,$$

where X_i is the count estimate for gene i and l_i is the length of the transcript as determined by querying the Ensembl mart with dataset = "hsapiens_gene_ensembl", using biomaRt in R version 3.5.1. TPM values were $\log_2 + 1$ transformed for downstream analyses.

Two subject outliers were detected by visual inspection of a principal component analysis plot of the normalized gene counts matrix (Figure S3). Thirteen samples were retained for further analysis, including seven ASD lines and six control lines.

Differential Expression Analysis

Gene-based read counts were analyzed for differential expression using the R DeSeq2 package, which uses variance stabilization techniques and the negative binomial distribution to detect expression changes across experimental conditions (Boston, MA, USA) (Love et al., 2014). For this study, two conditions were analyzed for differential expression, using two different model matrices. First, differential expression with respect to MED was analyzed (expression $\propto$ MED). Second, autism-associated genes were gleaned by testing a model that included autism and autism-MED interactions (expression $\propto$ MED + MED × ASD + ASD). A total of 24,162 genes were examined for differential expression, and genes were deemed significant if p_{FDR} was less than 0.05. An expanded set of genes with $p_{raw} \leq 0.05$ was used to assess broad trends in the data with follow-up analyses. Plotting of individual gene expression results, top gene summaries, and heatmaps were performed using TPM normalized counts that were corrected for covariates with the R ComBat package (Boston, MA, USA). All gene expression data are deposited at Gene Expression Omnibus (GEO: GSE125020).

Gene Ontology Analysis

Gene ontology (GO) analysis was performed to determine which biological pathways were associated with differentially expressed genes. The statistical overrepresentation tool from PantherDB was used to perform Fisher's exact tests for a list of uncorrected differentially expressed genes and each GO term in the database (Mi et al., 2017; Thomas et al., 2006). p values were adjusted for multiple comparisons using FDR correction. Biological processes with a p_{FDR} less than 0.05 were reported and plotted in this study.

Statistical Tests for Gene List Overlap and Clinical Correlations

Statistical analyses for gene list overlap and clinical correlations were performed using R version 3.51 (Vienna, Austria) (R Core Team, 2016). Clinical correlation was examined between MED day 15 average values and subject clinical scores. To further explore the differential expression results, overlap was measured with ASD genes, brain-expressed genes, and a negative control of liver-expressed genes (Abrahams et al., 2013; Krishnan et al., 2016; Uhlén et al., 2015). Tissue-specific gene sets were obtained from the Human Protein Atlas (available from www.proteinatlas.org). For each list of interest, overlap was tested using the binomial test, and significance was assigned to p values less than 0.05. Clinical correlations between MED and various clinical endpoints were calculated using the Pearson correlation test. Presented correlations are significant at p values less than 0.05.

Spatiotemporal Enrichment Analysis

Spatiotemporal enrichment of the MED differentially expressed genes was assessed using gene expression data from the BrainSpan Atlas of the Developing Human Brain (Sunkin et al., 2013). Normalized gene transcript counts were acquired for region-stage pairs. Twelve developmental stages and 16 discrete brain structures were included. In a process similar to one described by Krishnan et al. (2016), representative genes sets were chosen for each region-stage pair by calculating the modified Z score of a given gene in the distribution of counts for all region-stage pairs (Krishnan et al., 2016). The modified Z score was calculated using the median, and the median absolute deviation (MAD) of each gene across all pairs was calculated for the ith gene and jth region-stage pair:

$$z_{i,j} = 0.645 \times \frac{\text{count}_{i,j} - \text{median}_i}{\text{MAD}_i}.$$

For the jth region-stage, all genes for which $z_{i,j} \geq 2$ were selected as representative genes.

Enrichment was calculated by performing the Fisher's exact test using the list of differentially expressed MED genes and each representative region-stage set of genes. Enrichment scores were corrected for multiple comparisons using FDR controlling, and p_{FDR} values are plotted in the heatmap in Figure 4. Significance was assigned to region-stage pairs with p_{FDR} less than 0.05. All spatiotemporal enrichments are listed in Table S6.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2019.08.001.

AUTHOR CONTRIBUTIONS

D.N.A., S.B.L., M.C.M., F.H.G., and Y.K. designed and performed the experiments described in this study and prepared this manuscript for publication. A.P.D.M. cultured the control and ASD cell lines used in this study, in addition to isolating and purifying RNA for high-throughput sequencing. R.S. performed cell sorting for the selection of PSA-NCAM-positive neurons from ASD and controls. G.E. and M.S. assisted with the bioinformatics analysis of the RNA-sequencing data. T.S. and Y.Z. assisted with the MED analysis of neuronal electrical recordings.

ACKNOWLEDGMENTS

Y.K. is supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare (HI18C1077), and a grant from the National Research Foundation (2018R1A2B003640), Republic of Korea. M.C.M. would like to acknowledge the support from the National Institutes of Health grant number R03 MH115426-01A1. D.N.A. would like to acknowledge support from the National Institutes of Health F30 MH115584. F.H.G would like to thank The Leona M. and Harry B. Helmsley Charitable Trust, The Robert and Mary Jane Engman Foundation, and JPB Foundation. Salk core facilities are supported by the Salk Cancer Center with a grant from the NCI (P30

CA014195). The authors would like to thank M.L. Gage for editorial comments.

Received: February 26, 2019
Revised: August 3, 2019
Accepted: August 4, 2019
Published: August 29, 2019

REFERENCES

Abrahams, B.S., Arking, D.E., Campbell, D.B., Mefford, H.C., Morrow, E.M., Weiss, L.A., Menashe, I., Wadkins, T., Banerjee-Basu, S., and Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol. Autism *4*, 36.

Akdemir Akar, S., Kara, S., Agambayev, S., and Bilgiç, V. (2015). Nonlinear analysis of EEGs of patients with major depression during different emotional states. Comput. Biol. Med. *67*, 49–60.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Baxter, A.J., Brugha, T.S., Erskine, H.E., Scheurer, R.W., Vos, T., and Scott, J.G. (2015). The epidemiology and global burden of autism spectrum disorders. Psychol. Med. *45*, 601–613.

Bosl, W., Tierney, A., Tager-Flusberg, H., and Nelson, C. (2011). EEG complexity as a biomarker for autism spectrum disorder risk. BMC Med. *9*, 18.

Courchesne, E., Carper, R., and Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. JAMA *290*, 337–344.

Courchesne, E., Mouton, P.R., Calhoun, M.E., Semendeferi, K., Ahrens-Barbeau, C., Hallet, M.J., Barnes, C.C., and Pierce, K. (2011). Neuron number and size in prefrontal cortex of children with autism. JAMA *306*, 2001–2010.

Courchesne, E., Pramparo, T., Gazestani, V.H., Lombardo, M.V., Pierce, K., and Lewis, N.E. (2018). The ASD Living Biology: from cell proliferation to clinical phenotype. Mol. Psychiatry *24*, 88–107.

Derosa, B.A., El Hokayem, J., Artimovich, E., Garcia-Serje, C., Phillips, A.W., Van Booven, D., Nestor, J.E., Wang, L., Cuccaro, M.L., Vance, J.M., et al. (2018). Convergent pathways in idiopathic autism revealed by time course transcriptomic analysis of patient-derived neurons. Sci. Rep. *8*, 8423.

Devlin, B., and Scherer, S.W. (2012). Genetic architecture in autism spectrum disorder. Curr. Opin. Genet. Dev. *22*, 229–237.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Fernández, A., Al-Timemy, A.H., Ferre, F., Rubio, G., and Escudero, J. (2018). Complexity analysis of spontaneous brain activity in mood disorders: a magnetoencephalography study of bipolar disorder and major depression. Compr. Psychiatry *84*, 112–117.

Geschwind, D.H., and State, M.W. (2015). Gene hunting in autism spectrum disorder: on the path to precision medicine. Lancet Neurol. *14*, 1109–1120.

Hazlett, H.C., Poe, M.D., Gerig, G., Styner, M., Chappell, C., Smith, R.G., Vachet, C., and Piven, J. (2011). Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. Arch. Gen. Psychiatry *68*, 467–476.

Hénon, M. (1976). A two-dimensional mapping with a strange attractor. In The Theory of Chaotic Attractors, P. Cvitanovic, ed. (Routledge), pp. 94–102.

Illumina. (2015). iGenomes online. https://support.illumina.com/sequencing/sequencing_software/igenome.html.

Jeong, J., Kim, D.J., Chae, J.H., Kim, S.Y., Ko, H.J., and Paik, I.H. (1998). Nonlinear analysis of the EEG of schizophrenics with optimal embedding dimension. Med. Eng. Phys. *20*, 669–676.

Kennel, M.B., Brown, R., and Abarbanel, H.D.I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A *45*, 3403.

Kreuz, T. (2011). Measures of neuronal signal synchrony. Scholarpedia *6*, 11922.

Krishnan, A., Zhang, R., Yao, V., Theesfeld, C.L., Wong, A.K., Tadych, A., Volfovsky, N., Packer, A., Lash, A., and Troyanskaya, O.G. (2016). Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. Nat. Neurosci. *19*, 1454.

Krumm, N., O'Roak, B.J., Shendure, J., and Eichler, E.E. (2014). A de novo convergence of autism genetics and molecular neuroscience. Trends Neurosci. *37*, 95–105.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860.

Liu, X., Campanac, E., Cheung, H.H., Ziats, M.N., Canterel-Thouennon, L., Raygada, M., Baxendale, V., Pang, A.L.Y., Yang, L., Swedo, S., et al. (2017). Idiopathic autism: cellular and molecular phenotypes in pluripotent stem cell-derived neurons. Mol. Neurobiol. *54*, 4507–4523.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Marchetto, M.C.N., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell *143*, 527–539.

Marchetto, M.C., Belinson, H., Tian, Y., Freitas, B.C., Fu, C., Vadodaria, K.C., Beltrao-Braga, P.C., Trujillo, C.A., Mendes, A.P.D., Padmanabhan, K., et al. (2017). Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. Mol. Psychiatry *22*, 820.

Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., Amenduni, M., Szekely, A., Palejev, D., Wilson, M., et al. (2015). FOXG1-Dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. Cell *162*, 375–390.

MATLAB. (2017). MATLAB and Statistics Toolbox Release 2017 (The MathWorks, Inc).

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P.D. (2017). PANTHER version 11: expanded

annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. *45*, D183–D189.

Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell *155*, 1008–1021.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z., et al. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am. J. Hum. Genet. *94*, 677–694.

Röschke, J., and Başar, E. (1988). The EEG is not a simple noise: strange attractors in intracranial structures. In Dynamics of Sensory and Cognitive Processing by the Brain, E. Başar, ed. (Springer), pp. 203–216.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209.

R Core Team (2016). R: A language and environment for statistical computing (R Foundation for Statistical Computing). http://www.R-project.org/.

Samengo, I., and Elijah, D.M.M. (2013). Spike-train analysis. In Principles of Neural Coding, R. Quiroga and S. Panzeri, eds. (CRC Press), pp. 75–96.

Schafer, S.T., Paquola, A.C.M., Stern, S., Gosselin, D., Ku, M., Pena, M., Kuret, T.J.M., Liyanage, M., Mansour, A.A.F., Jaeger, B.N., et al. (2019). Pathological priming causes developmental gene network heterochronicity in autistic subject-derived neurons. Nat. Neurosci. *22*, 243–255.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., et al. (2007). Strong association of de novo copy number mutations with autism. Science *316*, 445–449.

Shen, M.D., Nordahl, C.W., Young, G.S., Wootton-Gorges, S.L., Lee, A., Liston, S.E., Harrington, K.R., Ozonoff, S., and Amaral, D.G. (2013). Early brain enlargement and elevated extra-axial fluid in infants who develop autism spectrum disorder. Brain *136*, 2825–2835.

Sunkin, S.M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T.L., Thompson, C.L., Hawrylycz, M., and Dang, C. (2013). Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Res. *41*, D996–D1008.

Takens, F. (1981). Detecting Strange Attractors in Turbulence in Dynamical Systems and Turbulence (Springer).

Taketani, M., and Baudry, M. (2006). Advances in Network Electrophysiology: Using Multi-Electrode Arrays (Springer).

Thomas, P.D., Kejariwal, A., Guo, N., Mi, H., Campbell, M.J., Muruganujan, A., and Lazareva-Ulitsky, B. (2006). Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic Acids Res. *34*, W645–W650.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. Science *347*, 1260419.

Vadodaria, K.C., Amatya, D.N., Marchetto, M.C., and Gage, F.H. (2018). Modeling psychiatric disorders using patient stem cell-derived neurons: a way forward. Genome Med. *10*, 1.

Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J., and Geschwind, D.H. (2011). Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature *474*, 380.

Willsey, A.J., Sanders, S.J., Li, M., Dong, S., Tebbenkamp, A.T., Muhle, R.A., Reilly, S.K., Lin, L., Fertuzinhos, S., Miller, J.A., et al. (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell *155*, 997–1007.

Stem Cell Reports, Volume *13*

Supplemental Information

Dynamical Electrical Complexity Is Reduced during Neuronal Differentiation in Autism Spectrum Disorder

Debha N. Amatya, Sara B. Linker, Ana P.D. Mendes, Renata Santos, Galina Erikson, Maxim N. Shokhirev, Yuansheng Zhou, Tatyana Sharpee, Fred H. Gage, Maria C. Marchetto, and Yeni Kim

Supplemental Figure 1

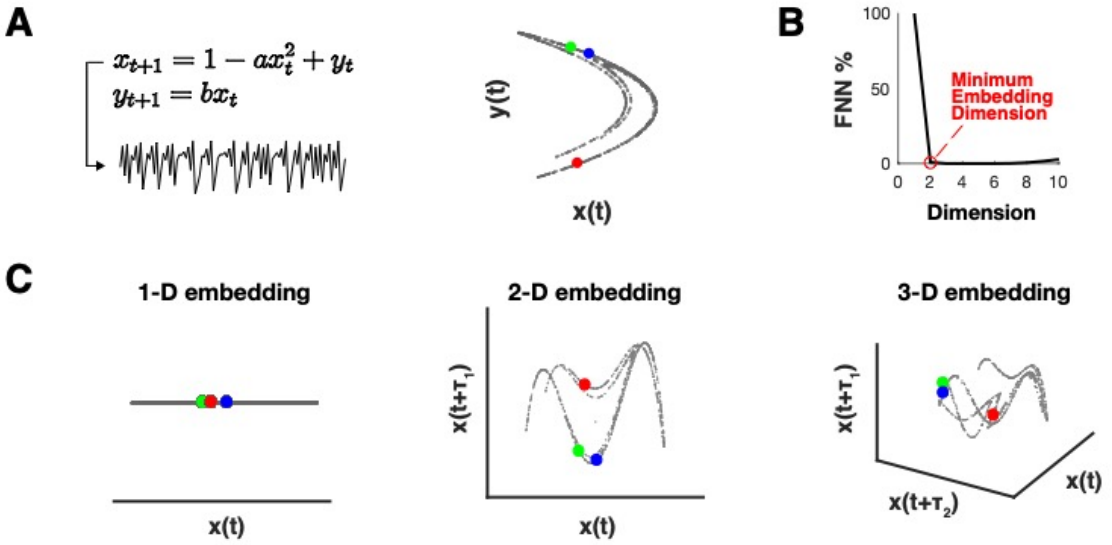

Figure S1: Minimum embedding dimension simulation.
The goal of the MED technique is to estimate the dimensional complexity of the system using only a time-series recording from a single variable in the system. This can be accomplished by using an algorithm to minimize the number of false nearest neighbors (FNNs) identified by iteratively plotting the time series over embedded versions of itself. A. The Hénon map. The Hénon map is a 2-variable (x, y) nonlinear dynamical system that exhibits chaotic behavior over time, simulated to recorded electrical activity (left). A state space reconstruction of this system reveals a smooth and regular attractor (right), in which some (x, y) coordinate pairs are closer together in space (blue and green) relative to other pairs (red). This attractor is a geometric representation of the 2-variable system, which can be summarized as having a dimensionality, or complexity, of 2, because it is fully unfolded in the (x, y) coordinate space. B. False nearest neighbors (FNN) to find the minimum embedding dimension (MED). Percentage of FNN is minimized when a time series is embedded in its native dimensionality. For the Hénon system, the MED occurs at 2, because it is a 2-variable system. C. Demonstration of the FNN procedure to find MED. The recording of a single variable from a dynamical system, such as x in the Hénon system or electrical activity in a system of neurons, can be embedded with delayed versions of the time series to reconstruct the original attractor of the entire system. This allows for the iterative testing of embedding dimensions and measurement of conserved neighboring points until a minimum is reached. In the example of the Hénon system, 1-dimensional embedding erroneously clusters all 3 points together. A 2-dimensional embedding correctly separates the red point from the blue and green points. Finally, a 3-dimensional embedding fails to create a further separation of these points, signifying that sufficient complexity is captured by a 2-dimensional embedding for the Hénon system. In the case of MEA recordings, MED provides a measurement of dynamical complexity of neural activity, potentially providing insights into the neurobiology of development and ASD.

Supplemental Figure 2

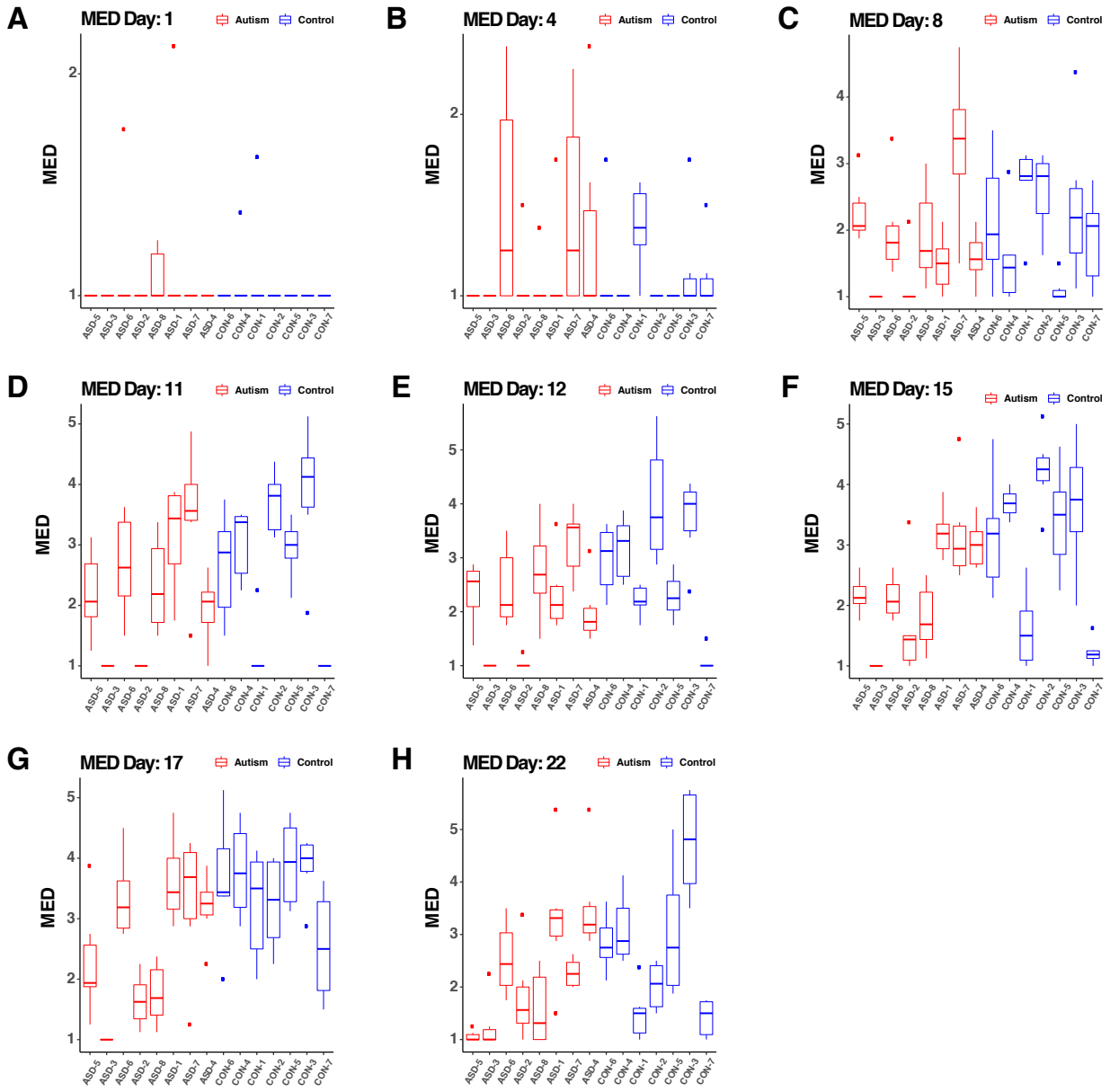

Figure S2: MED intersubject variability.
A-H show boxplots for 7 control (blue) and 8 ASD (red) subject MED values averaged over six replicates across the timepoints of MEA recording. The diminished dynamical complexity of the ASD subjects becomes apparent in the second week timepoints, through the course of the remaining days. Despite the group differences, these panels reveal the individual subject variation in the measurement. F. Of note, this panel shows the MED subject values for day 15 of the recordings, which was the day that cells were extracted for RNAseq. The average subject MED values on this day were used as the variable of interest for gene expression analysis.

Supplemental Figure 3

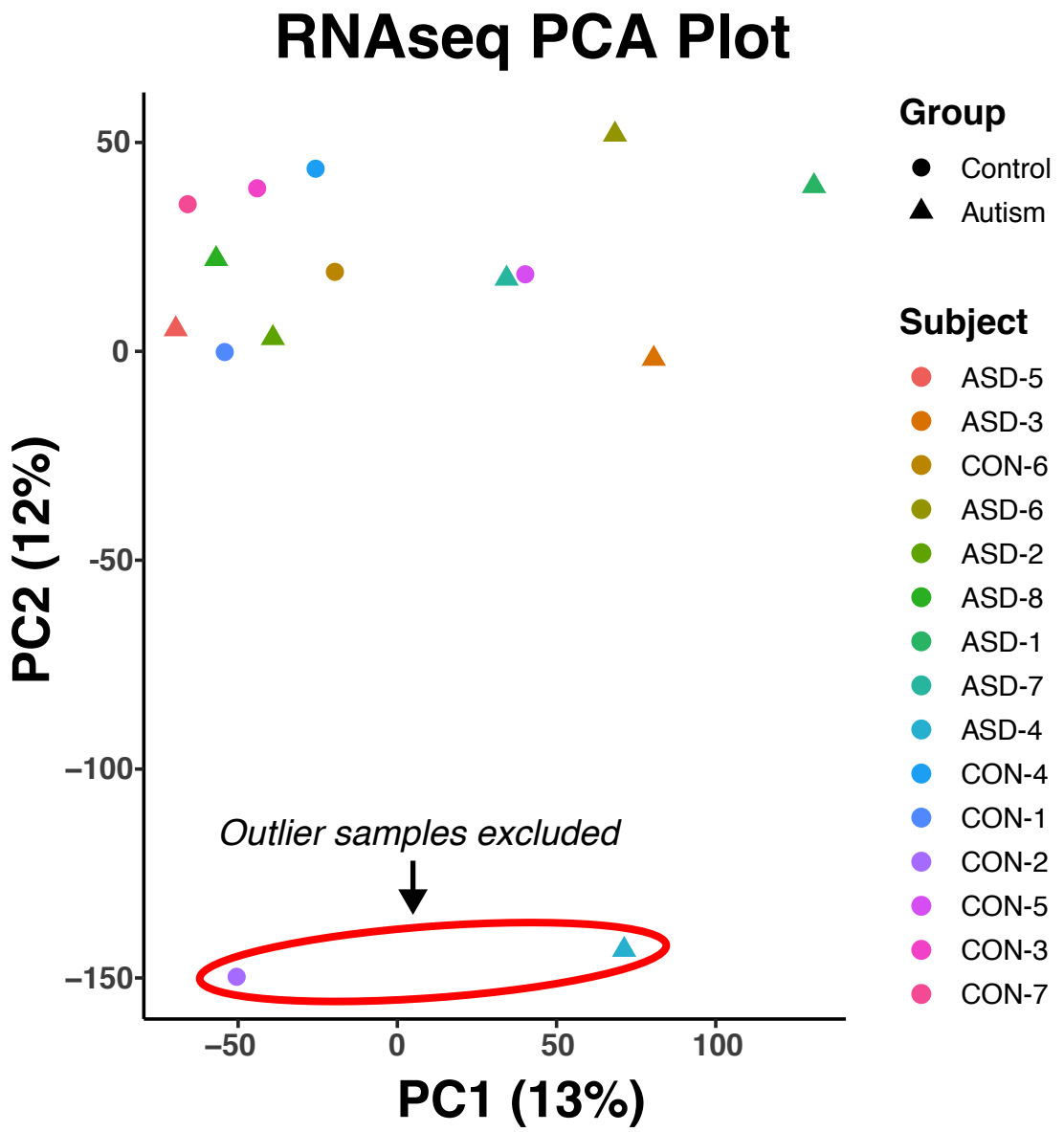

Figure S3: RNA sequencing counts matrix PCA plot.
Principal components analysis plot of the RNA sequencing counts matrix reveals two outlier subjects, CON-2 and ASD-4, which were removed for differential expression analysis. In parentheses, the percentage of variance explained by each principal component is given.

Supplemental Figure 4

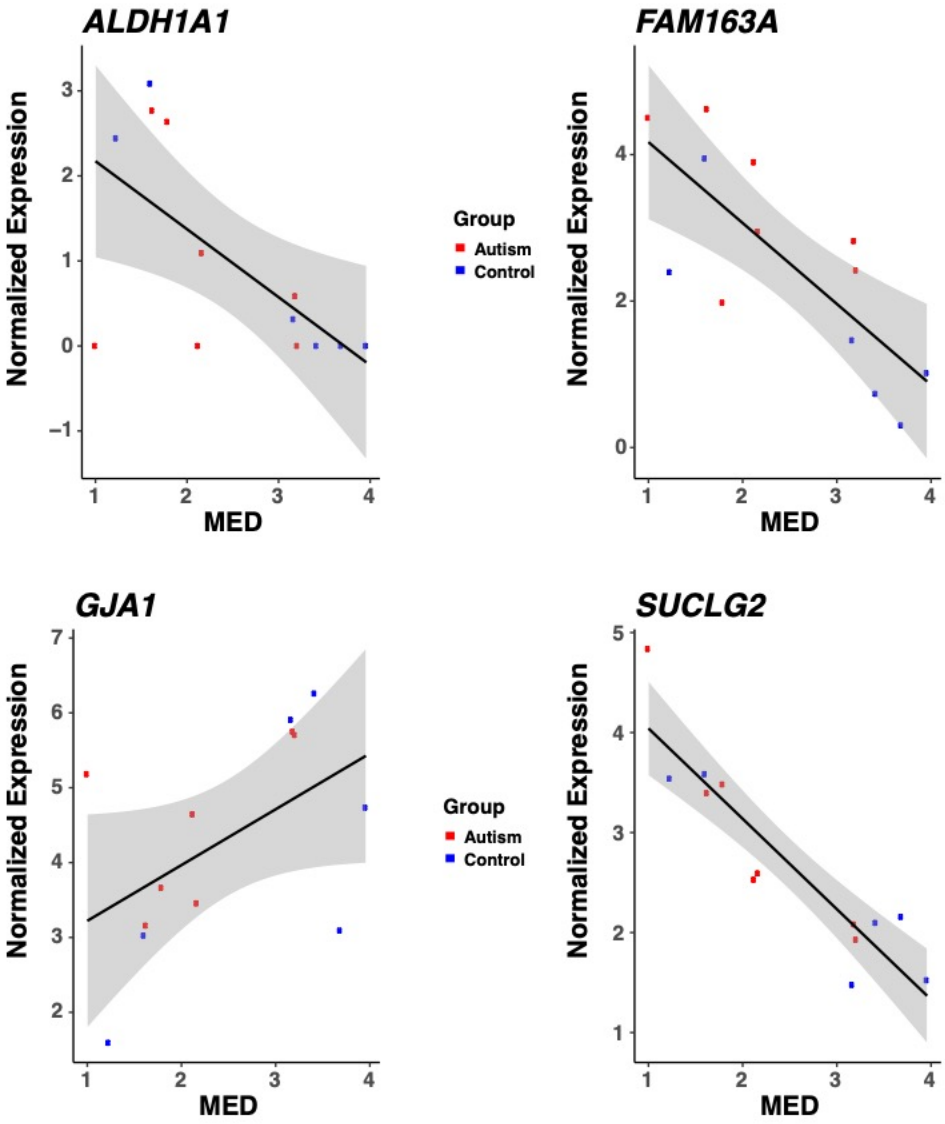

Figure S4: Examples of genes differentially expressed for MED.
Individual gene plots for genes that are differentially expressed for the MED value reveal linear correlations between gene expression and dynamical complexity. These relationships capture transcriptomic variance more comprehensively than ASD-control group labels. For all panels, the grey shading represents a 95% confidence interval around the mean values, as estimated by a linear model.