Supplementary Methods

Subjects

Over a 12 months period, 15 patients with ASD were recruited for the study in the CRC Balbuzie center in Rome (Italy). Inclusion criteria for enrollment were: a) patients who did not receive either cognitive and behavioral or medication treatment; b) range age 30-60 months; c) availability of the patient's history from the birth until the time of the diagnosis. The exclusion criteria included: a) presence of neurological defects or neurological deterioration; b) a genetic or metabolic disease. All subjects were selected to participate in the study, considering inclusion criteria; reasons of exclusion were: refusal to participate (n=4), and drop out of the study for noncompliant treatment (n=6). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The final sample consisted of 5 subjects (1 female, 4 males; mean age=39.4 months; SD=2.7; range 32–45 months). Autism diagnosis was done using the Autism Diagnostic Observation Schedule (ADOS-2) (Lord *et al*., 2012), the Vineland Adaptive Behavior Scales Second Edition (VABS-II) (Sparrow *et al*., 2005), the Psychoeducational profile–Third edition (PEP-3) (Davis, 2014) and the Griffiths Mental Developmental Scales-Extended Revised (GMDS-ER). Psychodiagnostic tests were administered before the therapeutic intervention (time 0) and after 12 months of treatment (time 1). Raters were blinded to child treatment status (pre/post intervention).

The therapeutic intervention, according to the treatment model used, was based on (1) 10 weekly hours of cognitive-behavioral therapy (naturalistic developmental behavioral intervention approach, 5 sessions of two hours) (Schreibman *et al*., 2015) integrated with

10 weekly hours of speech (2 sessions of two hours) and psychomotor (3 sessions of two hours) therapies; (2) parent support following diagnosis communication and parent training to implement specific programs in the familiar context and (3) psychoeducational intervention and meetings with teachers to facilitate child integration in the school context. The total number of hours per week was 20. Every child receive treatment by the same operators (speech and psychomotor therapist, psychologist) during the entire study. Educational efforts focusing on autistic symptoms and its management were discussed in encouraging adherence to treatment model. All parents were adherent to prescribed therapies. All parents gave written informed consent for their children. The consent form was approved by the Ethical Committee.

RNA sequencing

Whole blood was collected just before the beginning of the therapeutic intervention and after 12 months of therapy. Samples were collected in the morning between 9 am and 12 am at both time 0 and time 1. In order to preserve RNA integrity, DNA/RNA ShieldTM (Zymo research) was immediately added to the blood tube. The MagMAXTM for stabilized Blood Tubes RNA isolation kit (Ambion, Life Techonologies) was used to purify total RNA. RNA sequencing was conducted at NovoGene Corporation INC (https://en.novogene.com/; Sacramento, CA) using Globin-Zero Gold rRNA Removal Kit & NEB directional library, at 20 million raw reads per sample. After sequencing, reads were aligned to the Human reference genome (GRCh37) using the *Spliced Transcripts Alignment to a Reference* (*STAR*) software v2.5 (Dobin *et al*., 2013). Aligned reads were summarized as gene-level counts using *featureCounts* 1.4.4 (Liao *et al*., 2014). Outliers

detection was conducted through Principal Component Analysis (PCA), using *R* v3.3.1 (R Core Team, 2016). Gene expression differential analysis between post- and pre-treatment was conducted using the R package *DESeq2* v1.14.1 (Love *et al*., 2014) with a paired model including RIN values as covariate. The method implemented in *DESeq2* fits a generalized linear model (GLM) for each gene, modeling reads counts following a negative binomial distribution. The Logarithmic Fold Change (FC), expressed as Log2 Fold Change, is estimated with an Empirical Bayes procedure, whereas the significance is assessed with a Wald Test (Love *et al*., 2014). The p-values were corrected for multiple testing using the False Discovery Rate (FDR) method (Benjamini and Hochberg, 1995), and genes were annotated using the R-package *BioMart* v2.30.0 (Durinck *et al*., 2005). We considered as significant all the genes with adjusted p-value (adj-p) < 0.05. A total of 5 patients were included in the study, and their RNA was sequenced before and after treatment. We sequenced a total of 487 million (M) of reads (median: 47.3 M, Range: 42.6 – 58.1 M), with 89.9% mapping rate. PCA analysis conducted using the top 2 Principal Components did not show any outlier (Figure S3). We excluded the low responder patient before the differential analysis between post- and pre-treatment. We adjusted for RIN values since we noted larger values in the post-treatment group, although not significant ($p < 0.500$). Additionally, we correlated the variation of ADOS-2 measurements with the changes in gene expression, computing a Pearson's correlation. The expression profiling matrix (raw counts) was transformed using the variance stabilizing transformation method as implemented in *DESeq2* (*vst* function), and adjusted regressing out RIN, age and sex values using the *adjustBatchEffect* function as implemented in the R-package *limma* (Ritchie *et al*., 2015). Then, we computed the ADOS-2 and expression profiling difference

from post to pre-preatment, and we estimated the correlation between the expression of each genes identified and the ADOS-2 changes as absolute values using the function *cor.test* in R.

Validation datasets

We compared our differential analysis results with RNA profiling publicly available datasets from whole blood and post-mortem brains. We hypothesized that associated genes with the treatment should be significantly different in our dataset (post vs pre-treatment) but have an opposite direction when comparing patients vs controls.

As comparison dataset from whole blood we considered the large meta-analysis conducted by Tylee *et al*. (2017), using the complete results of the meta-analysis included in Supplementary Table S3 of their manuscript. In this study the authors combined the results from 7 different *ex vivo* blood or lymphocytes microarray expression profiling studies totally including 626 affected and 447 controls. We considered both the models run including the covariates (non-*sva)*, and the model run with *sva*, with p-values adjusted with the BH method. We considered "validated" genes if: 1) were significant at the unadjusted $p < 0.05$ in one of the two model (non-*sva* and *sva*), 2) with discordant log2 Fold Change direction with our results. Furthermore, we analyzed an additional dataset to investigate whether the DEGs detected in our dataset were related to age changes. The RNA sequencing in our study was conducted when the patients were 3.5 and 4.5 years old (pre and post-treatment, respectively), so potentially some observed changes could be a consequence of the difference in age and not due to the treatment. To further investigate this potential confounding factor, we downloaded the dataset GSE18123 (Kong *et al*.,

2012) including 104 ASD and 184 non-affected characterized with the Affymetrix array HuGene-1_0-st. We selected the sub-dataset GPL6244 because the cohort included both females and males, similarly to our study, and we downloaded the raw data (CEL files) for the ASD patients between 3 and 5 years old (n = 29). Data were normalized using the Robust Multi-array Average (RMA) algorithm (Irizarry *et al*., 2003) with the R-package *oligo* (Carvalho and Irizarry, 2010) and quality controls were conducted using the R-package *arrayqualitymetrics* (Kauffmann *et al*., 2009). We classified and excluded outliers significant in 2 out of 3 metrics computed from the normalized expression values (distances between arrays, Boxplots, and MAplots). Finally, we run a linear model using age as predictor for the gene expression and including sex as covariate. Results were adjusted for multiple testing using the False Discovery Rate method (Benjamini and Hochberg, 1995).

We also compared the results with one publicly available post-mortem brain dataset, including RNA profiling from 13 ASD cases and 39 controls from dorsolateral prefrontal cortex (GSE102741) (Wright *et al*., 2017). FASTQs files were download from the Sequencing Reads Archive (#PRJNA398545), and reads were aligned with *STAR v2.7.5b* (Dobin *et al*., 2013) using *GrCh38* as reference genome. Counts were summarized at the gene level using the function *featureCounts* as implemented in the R-package *Rsubread* (Liao *et al*., 2019). All the downstream analysis were conducted using the R-Package *DEseq2* (Love *et al*., 2014). We removed genes with less than 10 total counts across all samples, and we applied variance-stabilized transformation (*vst* function) before conducting PCA to detect outliers. Finally, we ran a differential expression analysis using the raw counts and adjusting for: pH, RIN, sex, race and age. Results were adjusted for multiple testing with the FDR method (Benjamini and Hochberg, 1995). Statistical

enrichment of our list of DEGs in Tylee *et al*., (2017) and GSE102741 was computed using Fisher's exact test, as implemented in the R-package *bc3net* (de Matos Simoes and Emmert-Streib, 2012), and with a gene set enrichment analysis, using the R-package *fgsea*.

Enrichment analysis

Differentially Expressed genes (DEGs, adj $p < 0.05$) were used for the pathway analysis. We referenced to the REACTOME database (Fabregat *et al*., 2016) adjusting the p-values for multiple testing using the FDR method. We also conducted the Gene Set Enrichment analysis (GSEA). This type of analysis uses as input the entire list of genes (significant and non-significant) ranked by log2 (FC). The goal is to determine whether members of a specific functional gene set tend to occur toward the top (upregulated genes) or bottom (downregulated genes) of the input list (Subramanian *et al*., 2005). The analyses were conducted using the R-package *ReactomePA* (Yu and He, 2016), referencing to the REACTOME database. Finally, we conducted a functional network analysis blood-specific using *HumanBase* web tool (https://hb.flatironinstitute.org/gene).

Supplementary results

Figure S1. Plots reporting the results of the clinical assessment conducted with ADOS-2 at pre and post-treatment time-points.
Treatment responders are indicated in blue (R), and the low responder is indicated in red (LR).

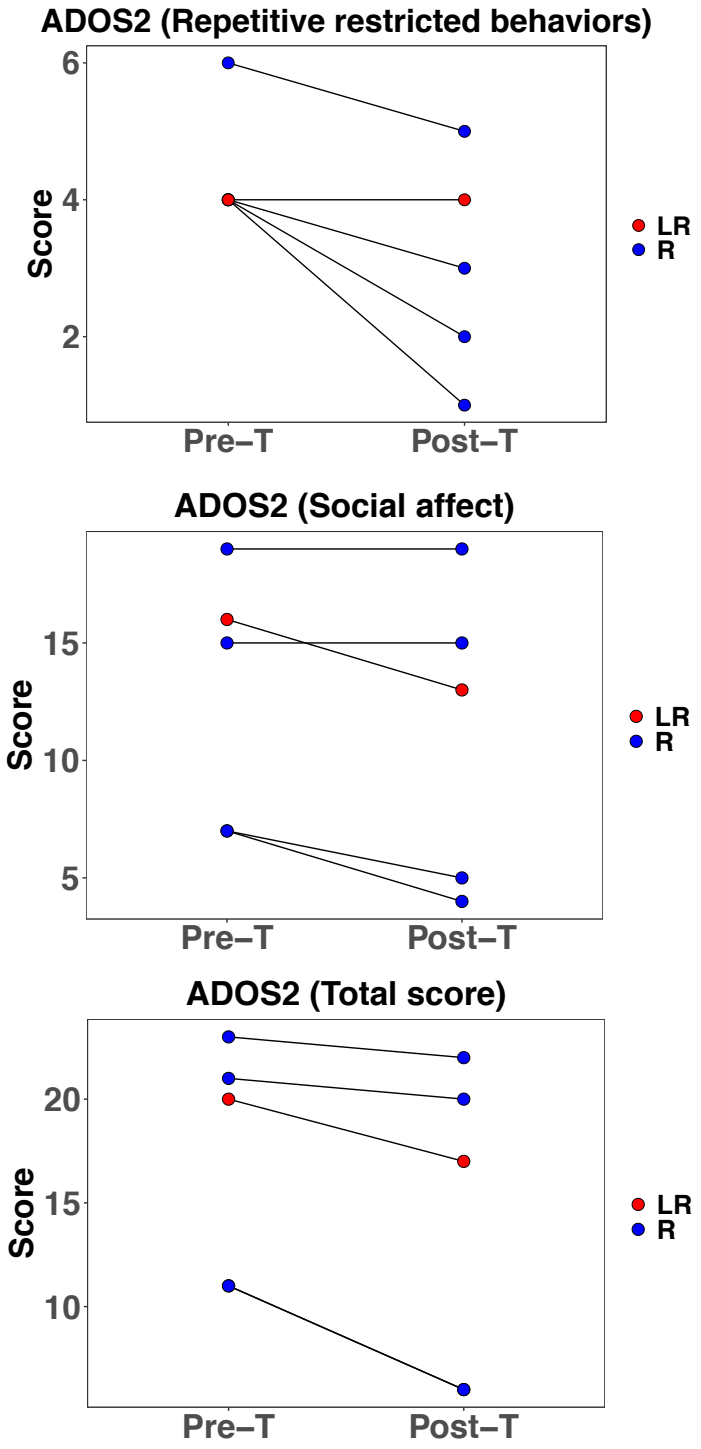


Figure S2. Plots reporting the results of the clinical assessment conducted with GMDS-ER at pre and post-treatment.
(Treatment responders in blue (R) and the low responder is indicated in red (LR).

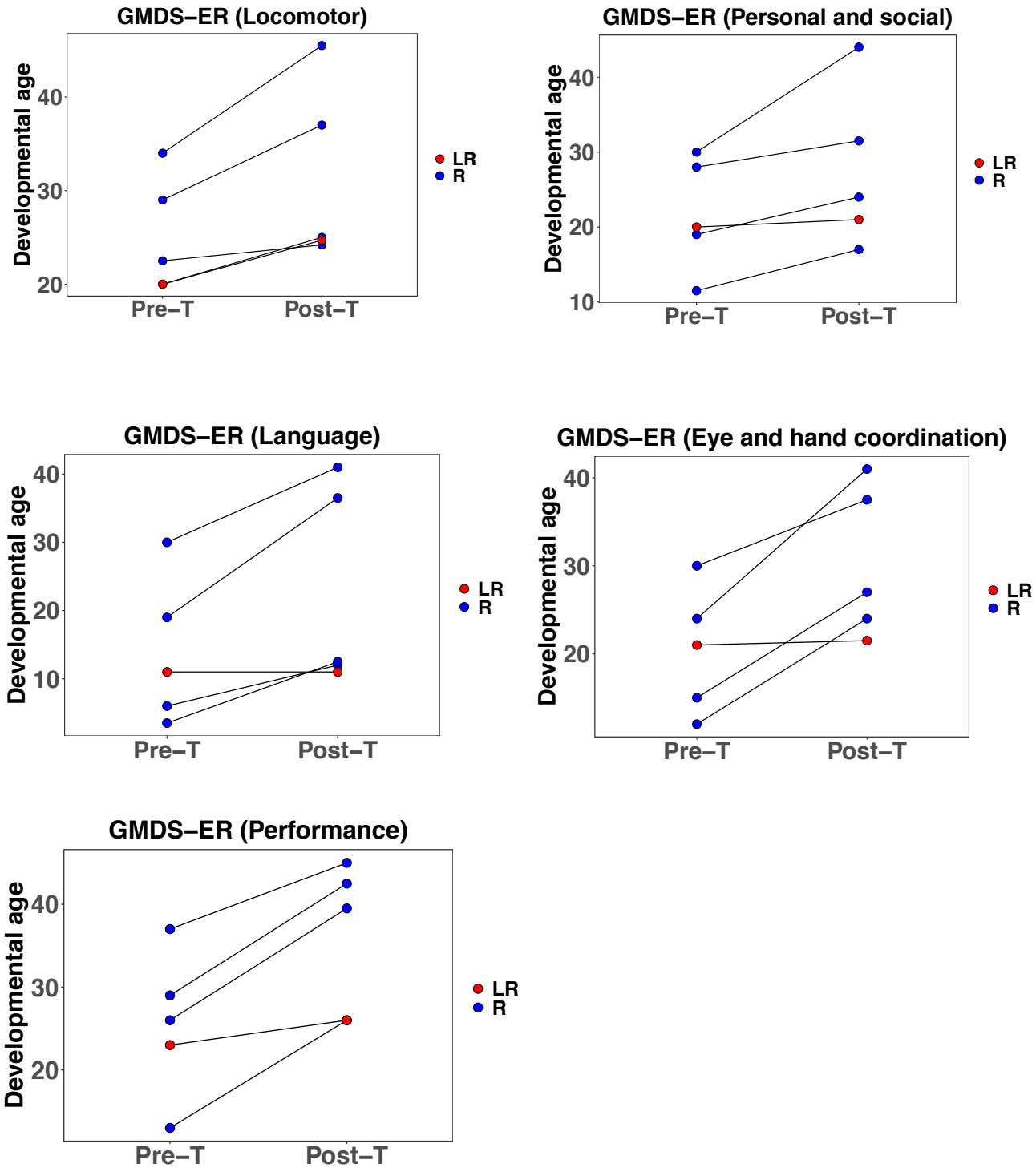


Figure S3. PCA analysis conducted using the components 1 and 2 for all samples (top) and only the responders (bottom). Colors and shape are time-point and patients specific, respectively.

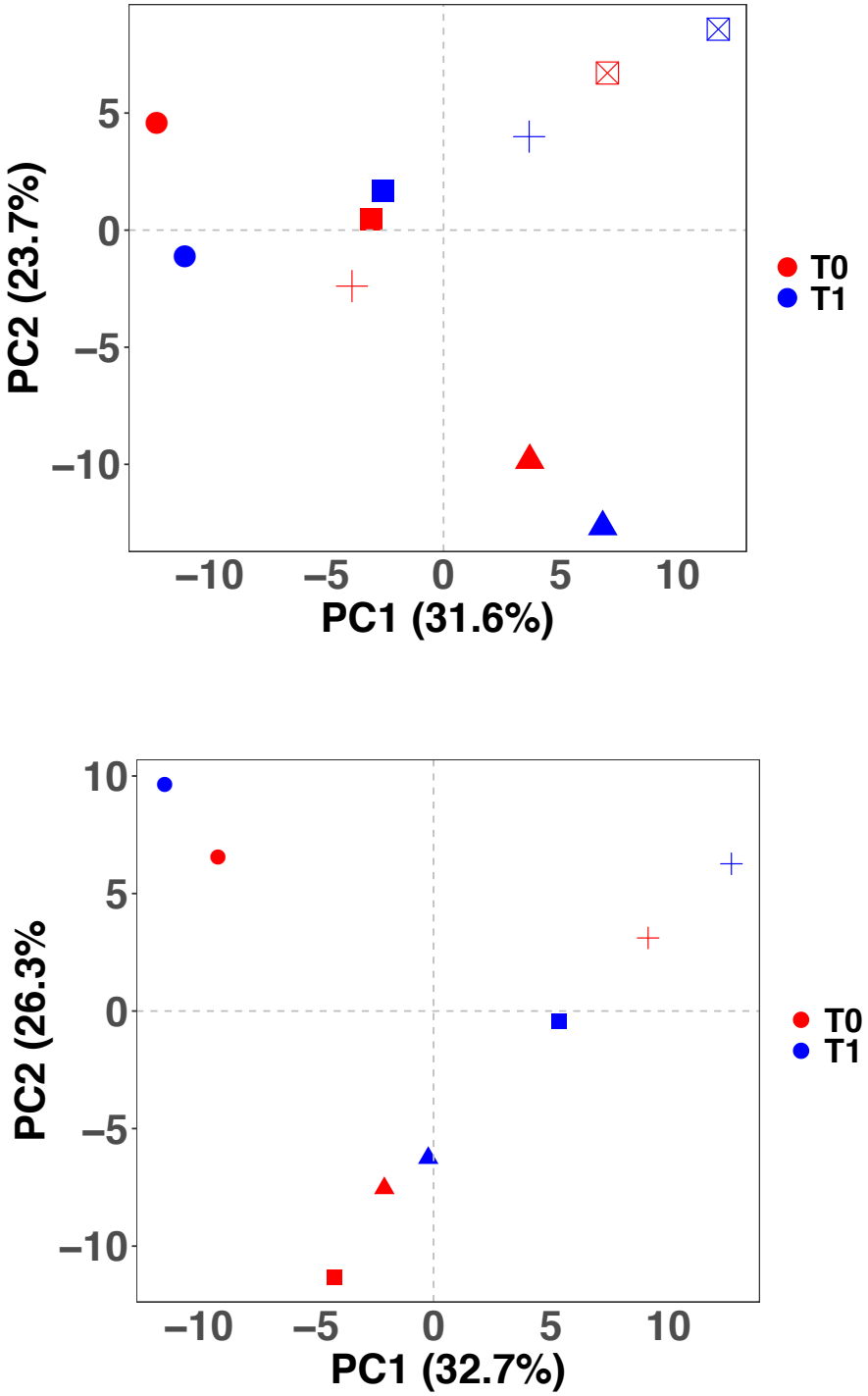


Figure S4. Enrichment of the differentially expressed genes detected in the present study in the dataset GSE102741, including RNA profiling from 13 ASD and 39 controls. We detected a non significant enrichment of our genes (mostly upregulated) among downregulated genes from GSE102714 (p = 0.343; normalized enrichment score = -1.022)

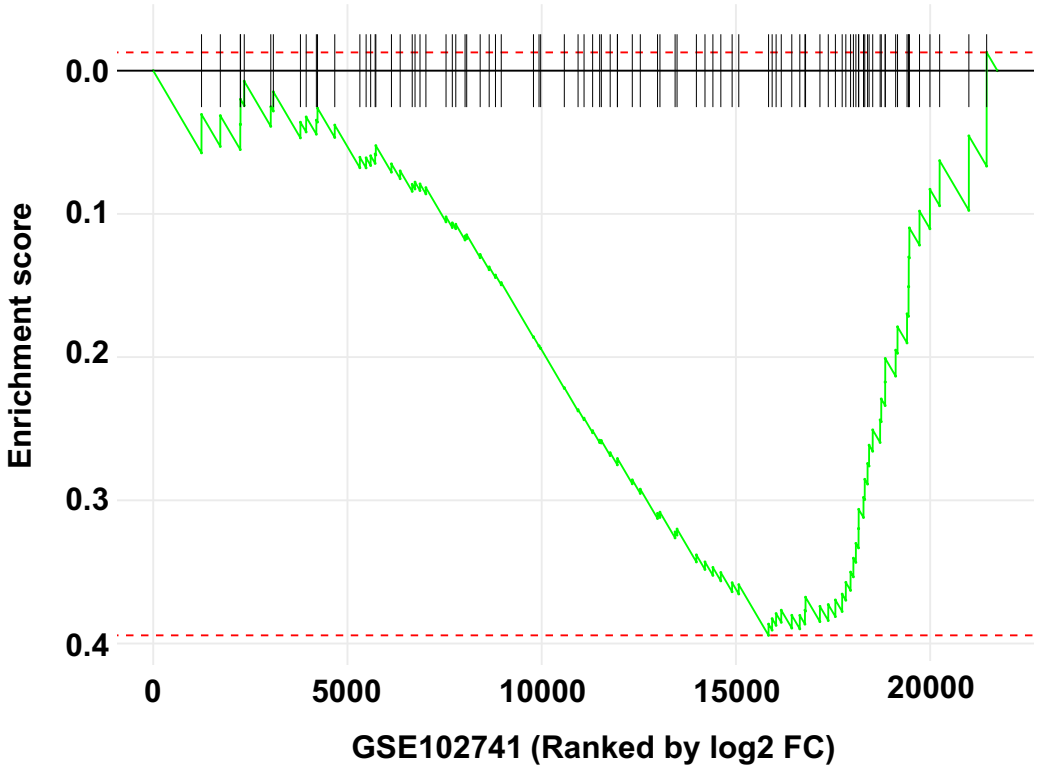


Figure S5. Correlation between ADOS-2 and gene expression changes between post and pre-treatment using all the 5 patients, including the non-responder.

(A) Barplot showind the distribution of *R* correlation strengh classes(Wuensch and Evans, 1996) found for both the DEGs (adj $p < 0.05$) and non-DEGs (adj $p \geq 0.05$). The distribution between the two groups was statistically significant ($p < 1.0E-06$).

(B) Scatterplot showing the correlation between log2FC (differential analysis high-responders) and correlation coefficient ADOS-2/gene expression (all patients). DEGs are reported in red (R = 0.718; p < 2.2E-16).

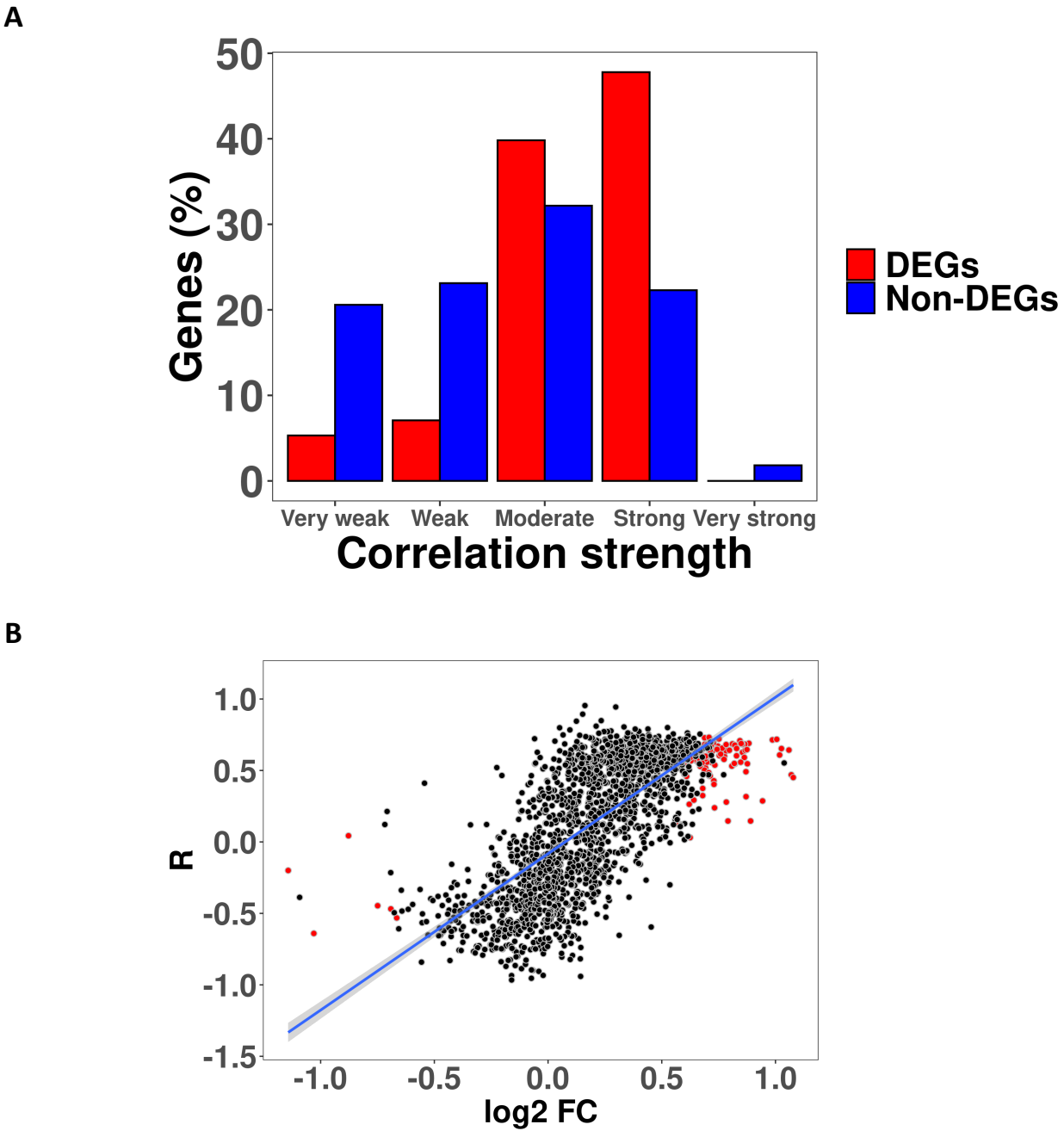


Figure S6. Functional blood-specific network obtained using the 109 DEGs. A total of 5 functional modules were identified (M1-M5).

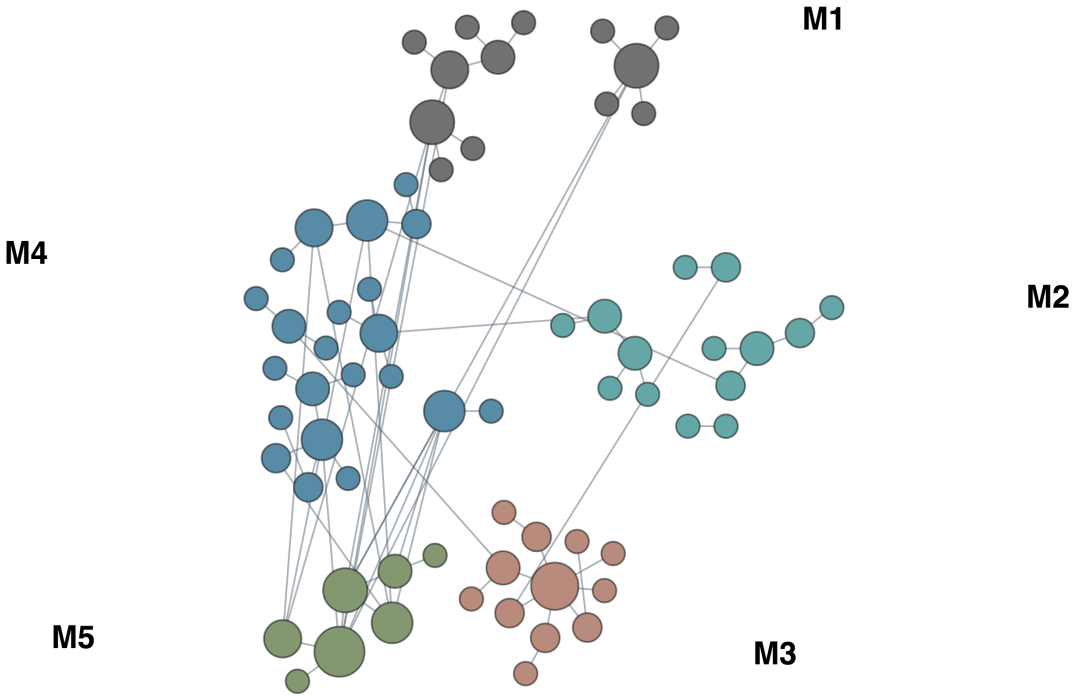


Figure S7. Heatmap representing the expression of the 22 candidate genes in whole blood and different brain regions in GTEx. Complete data are reported in Table S5.

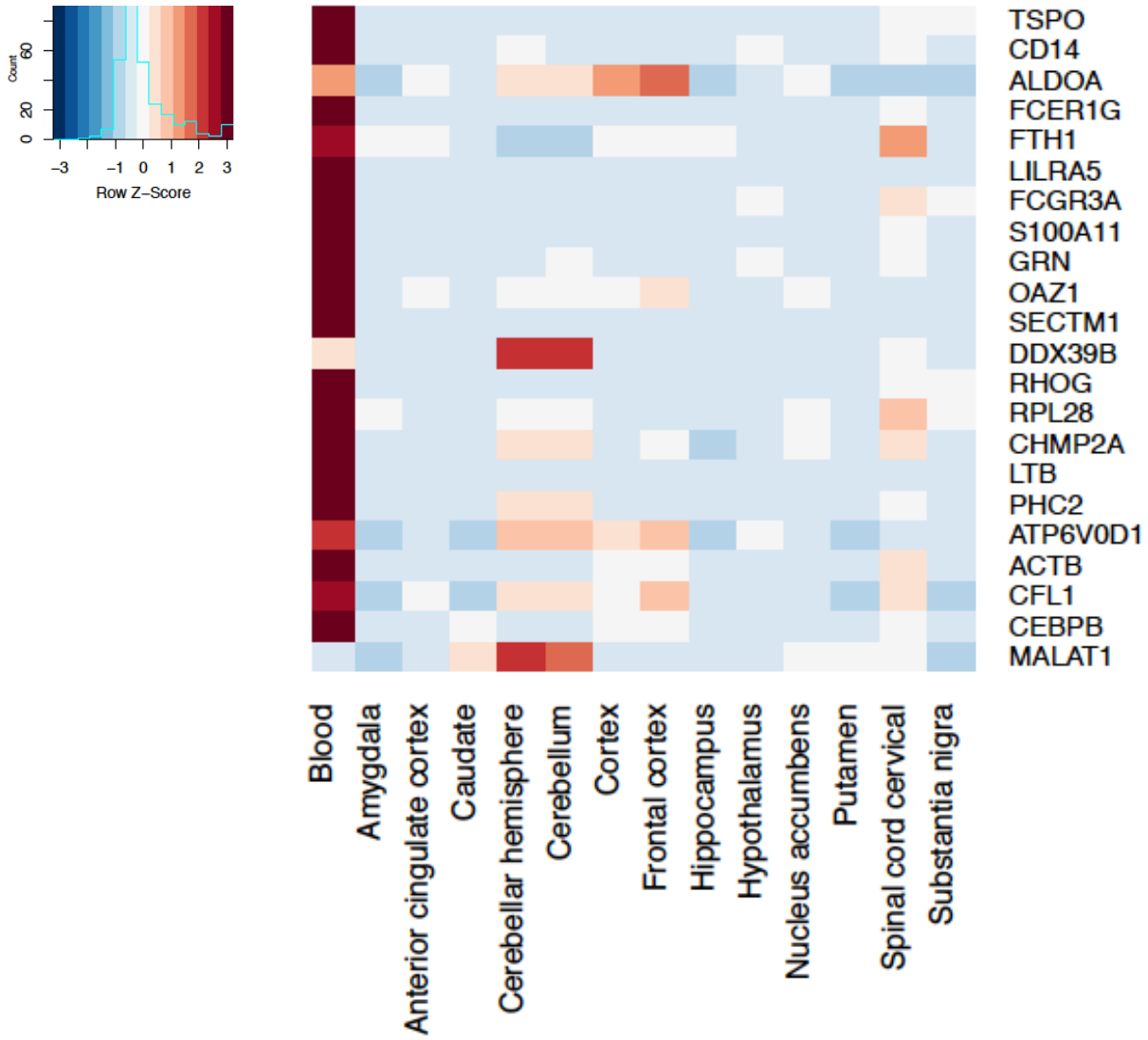
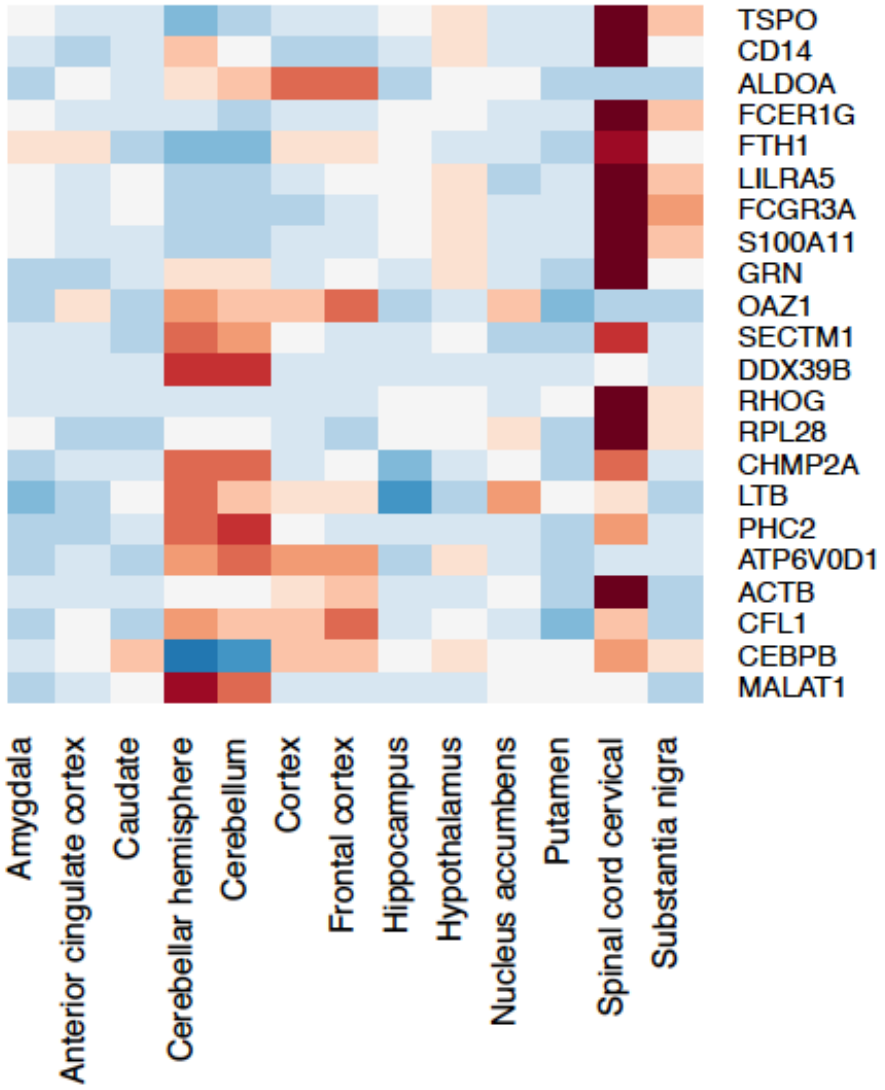


Figure S8. Heatmap representing the expression of the 22 candidate genes in different brain regions in GTEx. Complete data are reported in Table S5. We removed the whole blood to better highlight the variation across brain regions



References

Benjamini Y, Hochberg Y (1995) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B* **57**, 289–300.

Carvalho BS, Irizarry RA (2010) A framework for oligonucleotide microarray preprocessing. *Bioinformatics*.

Davis H (2014) Psychoeducational Profile-Third Edition (PEP-3). In *Encyclopedia of Special Education*.

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

Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W (2005) BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440.

Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, Jassal B, Jupe S, Korninger F, McKay S, Matthews L, May B, Milacic M, Rothfels K, Shamovsky V, Webber M, Weiser J, Williams M, Wu G, Stein L, Hermjakob H, D'Eustachio P (2016) The reactome pathway knowledgebase. *Nucleic Acids Research* **44**, D481–D487.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264.

Kauffmann A, Gentleman R, Huber W (2009) arrayQualityMetrics - A bioconductor package for quality assessment of microarray data. *Bioinformatics* **25**, 415–416.

Kong SW, Collins CD, Shimizu-Motohashi Y, Holm IA, Campbell MG, Lee IH, Brewster SJ, Hanson E, Harris HK, Lowe KR, Saada A, Mora A, Madison K, Hundley R, Egan J, McCarthy J, Eran A, Galdzicki M, Rappaport L, Kunkel LM, Kohane IS (2012) Characteristics and Predictive Value of Blood Transcriptome Signature in Males with Autism

Spectrum Disorders. *PLoS ONE*.

Liao Y, Smyth GK, Shi W (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930.

Liao Y, Smyth GK, Shi W (2019) The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*.

Lord C, Rutter M, DiLavore P, Risi S, Gotham K, Bishop S (2012) *Autism Diagnostic Observation Schedule, (ADOS-2) Modules 1-4. Los Angeles, California:*

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

de Matos Simoes R, Emmert-Streib F (2012) Bagging statistical network inference from large-scale gene expression data. *PLoS ONE*.

R Core Team (2016) R Development Core Team. *R: A Language and Environment for Statistical Computing* **55**, 275–286.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* **43**, e47.

Schreibman L, Dawson G, Stahmer AC, Landa R, Rogers SJ, McGee GG, Kasari C, Ingersoll B, Kaiser AP, Bruinsma Y, McNerney E, Wetherby A, Halladay A (2015) Naturalistic Developmental Behavioral Interventions: Empirically Validated Treatments for Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*.

Sparrow SS, Cicchetti D V., Balla DA (2005) The Vineland Adaptive Behavior Scales (2nd ed). In *Major psychological assessment instruments*.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*.

Tylee DS, Hess JL, Quinn TP, Barve R, Huang H, Zhang-James Y, Chang J, Stamova BS, Sharp FR, Hertz-Picciotto I, Faraone S V., Kong SW, Glatt SJ (2017) Blood transcriptomic comparison of individuals with and without autism spectrum disorder: A combined-samples mega-analysis. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*.

Wright C, Shin JH, Rajpurohit A, Deep-Soboslay A, Collado-Torres L, Brandon NJ, Hyde TM, Kleinman JE, Jaffe AE, Cross AJ, Weinberger DR (2017) Altered expression of histamine signaling genes in autism spectrum disorder. *Translational Psychiatry*.

Wuensch KL, Evans JD (1996) Straightforward Statistics for the Behavioral Sciences. *Journal of the American Statistical Association*.

Yu G, He QY (2016) ReactomePA: An R/Bioconductor package for reactome pathway analysis and visualization. *Molecular BioSystems*.