

Supplementary materials for:

# Machine Learning-Based Blood RNA Signature for Diagnosis of Autism Spectrum Disorder

Irena Voinsky <sup>1</sup>, Oleg Y. Fridland <sup>2</sup>, Adi Aran <sup>3,4</sup>, Richard E. Frye <sup>5,6,\*</sup> and David Gurwitz <sup>1,7,\*</sup>

<sup>1</sup> Department of Human Molecular Genetics and Biochemistry, Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel; irenavoinsky@mail.tau.ac.il

<sup>2</sup> Independent Researcher, Tel Aviv 69978, Israel; olegfrid@post.bgu.ac.il

<sup>3</sup> Shaare Zedek Medical Center, Jerusalem 91031, Israel; aaran@szmc.org.il

<sup>4</sup> Obesity and Metabolism Laboratory, Institute for Drug Research, School of Pharmacy, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem 91240, Israel

<sup>5</sup> Autism Discovery and Treatment Foundation, Phoenix, AZ 85050, USA

<sup>6</sup> Rossignol Medical Center, Phoenix, AZ 85050, USA

<sup>7</sup> Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel

\* Correspondence: drfrye@rossignolmedicalcenter.com or drfrye@autismdiscovery.org (R.E.F.); gurwitz@tauex.tau.ac.il (D.G.)

**Supplementary Table S1: Pearson r correlation of studied genes expression ( $2^{-\Delta Ct}$ ).** The number of XY Pairs=101. Correlation in bold represent chosen genes combinations for ML testing ( $r>0.3$ ,  $*p\leq 0.05$ ).  
\* $P\leq 0.05$ , \*\* $P\leq 0.01$ , \*\*\* $P\leq 0.001$ , \*\*\*\* $P\leq 0.0001$ .

Gene	<i>BATF2</i>	<i>MT2A</i>	<i>ISG15</i>	<i>LY6E</i>
<i>SERPING1</i>	<b><math>r=0.3030</math></b> <b>**<math>p=0.0021</math></b>	<b><math>r=0.4525</math></b> <b>****<math>p&lt;0.0001</math></b>	$r=0.2948$ <b>**<math>p=0.0028</math></b>	$r=-0.1348$ $p=0.1791$
<i>BATF2</i>		<b><math>r=0.3303</math></b> <b>***<math>p=0.0007</math></b>	$r=0.09726$ $p=0.3332$	$r=0.2832$ <b>**<math>p=0.0041</math></b>
<i>EFHC2</i>	$r=-0.006877$ $p=0.9456$	$r=-0.1668$ $p=0.0956$	$r=0.007879$ $p=0.9377$	$r=-0.1816$ $p=0.0692$
<i>FCGR1A</i>	$r=0.2766$ <b>**<math>p=0.0051</math></b>	$r=0.1139$ $p=0.2567$	$r=0.2377$ <b>*<math>p=0.0167</math></b>	$r=-0.1737$ $p=0.0824$
<i>CDC20</i>	$r=0.2495$ <b>*<math>p=0.0119</math></b>	$r=0.02850$ $p=0.7772$	$r=0.09243$ $p=0.3579$	$r=-0.1396$ $p=0.1639$
<i>MT2A</i>	<b><math>r=0.3303</math></b> <b>***<math>p=0.0007</math></b>		<b><math>r=0.3126</math></b> <b>**<math>p=0.0015</math></b>	$r=0.009802$ $p=0.9225$
<i>ISG15</i>	$r=-0.006877$ $p=0.9456$	<b><math>r=0.3126</math></b> <b>**<math>p=0.0015</math></b>		$r=0.1379$ $p=0.1692$
<i>FBXO6</i>	<b><math>r=0.3134</math></b> <b>**<math>p=0.0014</math></b>	<b><math>r=0.3884</math></b> <b>****<math>p&lt;0.0001</math></b>	<b><math>r=0.3618</math></b> <b>***<math>p=0.0002</math></b>	<b><math>r=0.4343</math></b> <b>****<math>p&lt;0.0001</math></b>
<i>LINC00869</i>	$r=-0.03090$ $p=0.7590$	$r=0.1456$ $p=0.1463$	$r=0.1852$ $p=0.0637$	$r=0.006298$ $p=0.9502$
<i>LY6E</i>	$r=0.2832$	$r=0.009802$	$r=0.1379$	

	**p=0.0041	p=0.9225	p=0.1692	
--	------------	----------	----------	--

**Supplementary Figure S1: Feature importance using the Random Forest Classifier.** The bars represent the feature importance of the “forest”, and standard deviation error bars represent the inter-trees variability. See Methods 2.3 for further details.

