[BIOINFO 2025]

AI-Driven Framework Modeling Perturbation in Brain Organoids Reveals Candidate Genes for Autism

In Gyeong Koh[1,2], Eunhyong Chang[1,2], Young Seok Choi[1,2], Soo-Whee Kim[1,2], Yujin Kim[1,2], Hyeji Lee[1,2], Gaeun Byun[1,2], Yebin Ryu[1,2], Seoyeon Kim[1,2], Jihae Lee[3], Hayoon Park[3], Hwanseok Shim[3], Yeojin Ryu[1,2], Woojong Shim[4], Jiho Lee[4], Nelson B. Salazar[5,6], Marla Mendes de Aquino[6,7], Worrawat Engchuan[6,7], Xiaopu Zhou[6,7], Il Bin Kim[8], Jae Hyun Han[9,10], Donna M. Werling[11], So Hyun Kim[12], Miae Oh[13], Min-Sik Kim[14], Daekee Lee[15], Jaesang Kim[15], Yong-Seok Lee[16], Woong Sun[17], Eunjoon Kim[18,19], Stephen W. Scherer[5,6,7,20], Min Ji Jeon[21,*], Hee Jeong Yoo[9,10,*], Joon-Yong An[1,2,3,23,*]

[1] L-HOPE Program for Community-Based Total Learning Health Systems, Korea University, Seoul, Republic of Korea

[2] Department of Integrated Biomedical and Life Science, Korea University, Seoul, Republic of Korea

[3] School of Biosystem and Biomedical Science, College of Health Science, Korea University, Seoul, Republic of Korea

[4] Department of Biomedical Sciences, Korea University College of Medicine, Seoul, Republic of Korea

[5] Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

[6] Genetics and Genome Biology Program, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada

[7] The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada

[8] Department of Psychiatry, CHA Gangnam Medical Center, CHA University School of Medicine, Seoul, Republic of Korea

[9] Department of Psychiatry, Seoul National University Bundang Hospital, Seongnam,

Republic of Korea

[10] Department of Psychiatry, Seoul National University College of Medicine, Seoul, Republic of Korea

[11] Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI, USA

[12] School of Psychology, Korea University, Seoul, Republic of Korea

[13] Department of Psychiatry, College of Medicine, Kyung Hee University, Seoul, Republic of Korea

[14] Department of New Biology, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Republic of Korea

[15] Department of Life Science, Ewha Womans University, Seoul, Republic Korea

[16] Department of Physiology, Seoul National University College of Medicine, Seoul, Republic of Korea

[17] Department of Anatomy, College of Medicine, Korea University, Seoul, Republic of Korea

[18] Center for Synaptic Brain Dysfunctions, Institute for Basic Science (IBS), Daejeon, Republic of Korea

[19] Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

[20] McLaughlin Centre, Toronto, ON M5G 0A4, Canada

[21] Department of Biomedical Informatics, Korea University College of Medicine, Seoul, Republic of Korea

* *Correspondence*: joonan30@korea.ac.kr

Traditional gene discovery for genetically heterogeneous autism spectrum disorder (ASD) requires large cohorts, which has limited the identification of the full spectrum of susceptibility genes. We present an integrative framework that combines a foundation model for predicting gene perturbation responses with orthogonal genomic validation using family-based genome sequence data (n=88,263) to prioritize ASD-linked genes. By training the model on the Neural Organoid Cell Atlas with Perturbations (NOCAP), comprising 3.6 million cells from normal, perturbed, and disease-model organoids, we built a foundation model that predicts transcriptional responses to ASD gene perturbations and enables in silico neurodevelopmental screens. Our model identified two biologically distinct gene groups: one enriched in synaptic pathways (e.g., *NBEA*, *FYN*, *ERC1*) during corticogenesis, and another active in early radial glia, involved in ubiquitin signaling (e.g., *RNF216, TNKS2, YWHAE*). Furthermore, we identified 167 additional candidate ASD genes from these two groups. Our findings demonstrate the value of perturbation-informed foundation models in ASD gene discovery and implicate early radial glia dysfunction in ASD etiology.