Deep Learning Identification of Autism Mutations: Using Molecular and Computational Methods to Characterize a Novel Prioritized Non-Coding Variant in FEZF1

Zhang, Vanessa (School: Manhasset High School)

Autism spectrum disorder is a heritable, developmental disorder affecting 1 in 68 children in the United States. Development of a more definitive diagnostic method is especially critical for early intervention and therapy. DeepSEA, a form of artificial intelligence, is a deep convolutional neural network that learns information on chromatin features to rank de novo mutations on predicted impact. The purpose of this study was to use interdisciplinary methods to uncover and investigate novel autism non-coding mutations. The impact of top-ranked variants on expression was tested via a Luciferase Reporter Assay. An Independent T-test ($p<0.01$) revealed that the FEZF1 variant led to significant differences between expression driven by the reference and alternate alleles. This suggests that deep learning methods can effectively and accurately refine extensive lists of potential mutations. An in vitro binding assay comparing regulation at the chromatin level revealed differential binding of the two FEZF1 variant sequences to nuclear proteins. In order to characterize expression driven by the variants, an in vivo expression assay was performed in the proto-vertebrate Ciona intestinalis. It was found that expression driven by the FEZF1 variant overlapped with a CNS marker, indicating autism-relevant expression. Based on the results and analysis of chromatin marks, the FEZF1 variant may potentially alter regulation and expression of the FEZF1 gene, which is involved in important neuronal stem cell fate decisions. This study identified a novel variant near the FEZF1 gene, and identified a connection between autism and a proposed gene interaction pathway associated with development.