Genetic Correlations Between NLGN4X Mutations and Autism Spectrum Disorder

Donald Lee Beadle

University of North Alabama

dlbeadle@una.edu

Abstract

Autism spectrum disorder (ASD) is a type of neurodevelopmental disorder. These disorders form a broad group in which the central nervous system is compromised in some significant manner that affects brain development. There are many possible genetic influences for these disorders. One influence, mutations with NLGN4X, can be associated with many cognitive impairments. As a result, the purpose of this research is to find genetic correlations between ASD and mutations with the gene NLGN4X. The methodology for this purpose includes researching the overall function of NLGN4X, building an accurate 3D model of the gene, and analyzing the most likely damaging mutations with respect to the model. The results of this research include the functionality of NLGN4X, an analysis of the various models created, links between NLGN4X and other genes, and a compiled list of relevant variants. The corresponding results of the research conclude that mutations with NLGN4X increase the risk of developing ASD.

Introduction

Since 2000, diagnosis rates for Autism Spectrum Disorder (ASD) have increased from approximately 1 in 150 children to 1 in 59 children [3]. This increase can be attributed to a better understanding of the causes of ASD as well as the implementation of more effective diagnosis methodologies for detecting neurodevelopmental disorders. On average, children are diagnosed with ASD at age four. However, ASD can be effectively diagnosed in children as young as age two [2]. When considering neurodevelopmental disorders, early intervention is a crucial step in developing an effective clinical treatment. However, early intervention is based on the ability to detect ASD both accurately and at a very young age. The symptoms of these neurodevelopmental disorders can vary and be difficult to observe in children across young ages. As a result, early detection of ASD can be aided by analysis and examination of the genetic influences as one of the underlying causes of ASD. One such likely genetic influence is NLGN4X.

Main Objectives

1. Understand the function of NLGN4X.
2. Build 3D models of NLGN4X.
3. Analyze and choose the best model.
4. Identify specific variants with links to neurodevelopmental disorders.

Materials and Methods

- Research: For this project, the first step was to compile a list of sources and data for neurodevelopmental disorders, specifically ASD, and the gene NLGN4X.
- Model Construction: Next, the UniProt database was used to access the current protein data bank (PDB) file and protein sequence (FASTA) file for NLGN4X. Three various models were constructed and analyzed.

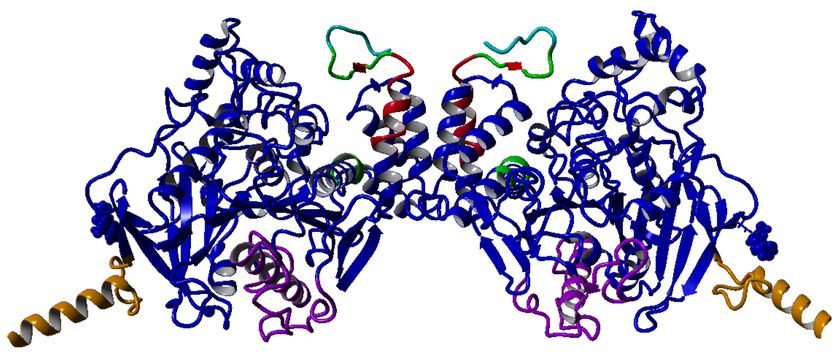

Figure 1: NLGN4X Slow Homology Model

- Suspect Variant Identification: Variants for NLGN4X were compiled to form a database. These variants were then analyzed using resources and tools including Provean, PolyPhen-2 and the GWAS Catalogue.

Results

Gene Function

- NLGN4X is a type I membrane protein-coding gene and belongs to a group of genes known as neuroligins. It encodes a member of the type B carboxylesterase/lipase protein family involved in the formation and remodeling of the central nervous system [4].
- The known molecular functionality includes cell-to-cell interaction, cell adhesion molecule binding, and signaling receptor activity.
- The known biological processes include neuronal cell formation, synapse organization, and social behavior and learning.
- Due to the wide range of involved molecular and biological processes, NLGN4X has a critical impact on the formation and development of the cognitive structures of the brain.

Model Analysis

From Table 1, the results of the overall scores from the three models produced can be observed. With YASARA, the z-scores for the NLGN4X (Q8N0W4) sequence were both below zero. However, because the scores were in a range of [0,-1], they were rated 'Good' by YASARA's scoring standards. The threaded model, by I-TASSER, utilized a confidence, or a C-score, as its scoring method. Out of the three models, the C-score for the threaded model was the lowest.

Model	Highest z/C-score
Fast Homology	-0.992
Protein Threading	-1.88
Slow Homology	-0.789

Table 1: Model Results

Interactions with NLGN4X

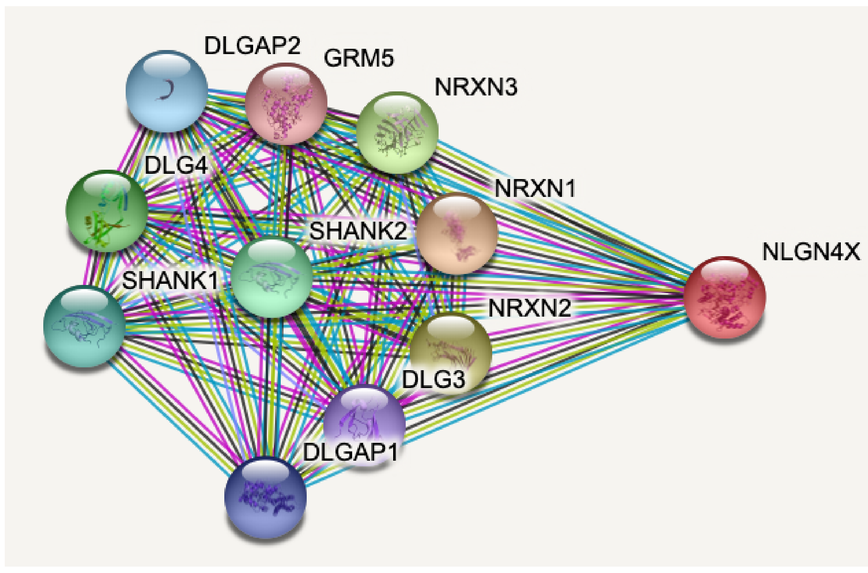

Figure 2: STRING Analysis

Figure 2 shows the results of a STRING analysis of NLGN4X. From the graph edges, it can be observed that NLGN4X shares interactions with genes responsible for presynaptic connections, such as neurexins, and scaffolding proteins, such as the shank family. Both of these groups of genes have mutations associated with ASD.

Variant Analysis

From a compiled variant database, 40 variants of uncertain significance (VUS) were identified for NLGN4X. PolyPhen-2 rated 15 of these as probably damaging and Provean rated 19 as deleterious. Figure 3 shows the 10 most deleterious rated variants from Provean.

Q8N0W4 101 R .	Q8N0W4	101	R	.	-12.28	Deleterious
Q8N0W4 317 C R	Q8N0W4	317	C	R	-11.29	Deleterious
Q8N0W4 94 P L	Q8N0W4	94	P	L	-8.45	Deleterious
Q8N0W4 494 P A	Q8N0W4	494	P	A	-7.35	Deleterious
Q8N0W4 583 R W	Q8N0W4	583	R	W	-7.07	Deleterious
Q8N0W4 84 G R	Q8N0W4	84	G	R	-6.38	Deleterious
Q8N0W4 187 G D	Q8N0W4	187	G	D	-6.07	Deleterious
Q8N0W4 693 F S	Q8N0W4	693	F	S	-5.86	Deleterious
Q8N0W4 710 R C	Q8N0W4	710	R	C	-4.31	Deleterious
Q8N0W4 583 R Q	Q8N0W4	583	R	Q	-3.53	Deleterious

Figure 3: 10 Most Deleterious Variants

Conclusions

- After comparing the three models produced in this project, the YASARA slow homology provided the best 3D representation of NLGN4X.
- Based on the overall function of NLGN4X, deleterious or damaging variants within this gene will have an adverse effect on neurological development.
- After examining a number of predicted deleterious and probably damaging variants, there are still a number of VUS that require analysis.

Forthcoming Research

- Model likely damaging variants of NLGN4X and record/analyze observable structural changes.
- Analyze the various types of mutations to discover the most likely type that results in ASD.

References

[1] Autism Facts and Figures. (n.d.). Retrieved from https://www.autismspeaks.org/autism-facts-and-figures

[2] Data and Statistics on Autism Spectrum Disorder — CDC. (n.d.). Retrieved from https://www.cdc.gov/ncbddd/autism/data.html

[3] Facts and Statistics. (n.d.). Retrieved from http://www.autism-society.org/what-is/facts-and-statistics/

[4] NLGN4X neuroligin 4 X-linked [Homo sapiens (human)] - Gene - NCBI. (n.d.). Retrieved from https://www.ncbi.nlm.nih.gov/gene/57502

Acknowledgements

Faculty Mentors: Dr. Cynthia Stenger and Dr. Jillian Stupiansky. Special thanks to Dr. Jeremey Prokop, Prokop Labs, Michigan State University and David Hinds, HudsonAlpha Institute for Biotechnology.

HudsonAlpha Institute for Biotechnology

CODE

Characterizing Our DNA Exceptions