Genetic Correlations Between NLGN4X Mutations and Autism Spectrum Disorder

Lee Beadle

University of North Alabama
Faculty Mentors: C. Stenger and J. Stupiansky

Scholars Week Presentation
April 23, 2019

Motivation

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 of NLGN4X, can be associated with several cognitive impairments. As a result, I wanted to model NLGN4X and find the best structural representation for analyzing these mutations.

Abstract

Purpose: Find genetic correlations between ASD and mutations of the gene NLGN4X.

Methodology:

- Researching NLGN4X.
- Constructing 3D models.
- Analyzing mutations that are likely damaging.

Results:

- Description of NLGN4X.
- Analysis of the various models.
- Compiled list of relevant variants.
- Identifying links between NLGN4X, other genes, and ASD.

Conclusion: The corresponding results of the research indicate that mutations of NLGN4X increase the risk of developing ASD.

Introduction

- Since 2000, diagnosis rates for ASD have increased from approximately 1 in 150 children to 1 in 59 children [1].
- When considering neurodevelopmental disorders, early intervention is crucial.
- However, early intervention is based on accurate detecting and diagnosis. With neurodevelopmental disorders, this can be problematic due to...
  - Broad range of symptoms.
  - Difficult detection at young ages.
  - Misdiagnosis.
- As a result, early detection can be aided by analysis and examination of the genetic influences.

NLGN4X Background 1

What is NLGN4X?

- Neuroligin 4, X-linked
- Type: Protein coding.
- Group: Neuroligins.

- Responsible for the formation and interactions of synapses between neurons.
- Broadly expressed in the brain.

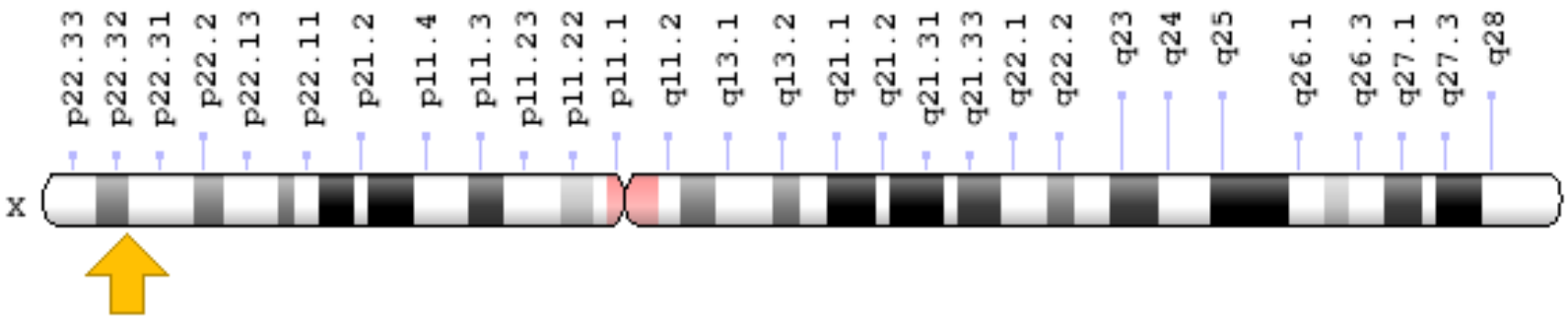

Figure 1: Chromosomal Location

NLGN4X Background 2

Possible conditions resulting from NLGN4X mutations:

- ASD (Significant risk).
- Motor function impairment.
- Communication/language impairments.
- ADHD.
- Schizophrenia.
- Tic disorders.
- Macroencephaly.

Figure 2: ASD

Methodology Overview

1. Research.
2. Constructing models.
3. Choosing the best model.
4. Identifying suspect variants.

Research

Getting Started:

- Compile a list of sources and data for ASD (ClinVar, NCBI, and UniProt).
- Use these resources to understand the function and impact of NLGN4X.
- Find likely damaging mutations of NLGN4X.
- Find links between ASD and the possible impact of NLGN4X mutations.

Constructing Models

Yet Another Scientific Artificial Reality Application (YASARA)

- Fast homology.
- Slow homology.

Iterative Threading ASSEmbly Refinement (I-TASSER)

- Threaded Model.

```
        10         20         30         40         50
MSRPQGLLWL PLLFTPVCVM LNSNVLLWLT ALAIKFTLID SQAQYPVVNT
        60         70         80         90        100
NYGKIRGLRT PLPNEILGPV EQYLGVPYAS PPTGERRFQP PEPPSSWTGI
       110        120        130        140        150
RNTTQFAAVC PQHLDERSLL HDMLPIWFTA NLDTLMTYVQ DQNEDCLYLN
       160        170        180        190        200
IYVPTEDDIH DQNSKKPVMV YIHGGSYMEG TGNMIDGSIL ASYGNVIVIT
       210        220        230        240        250
INYRLGILGF LSTGDQAAKG NYGLLDQIQA LRWIEENVGA FGGDPKRVTI
       260        270        280        290        300
FGSGAGASCV SLLTLSHYSE GLFQKAIIQS GTALSSWAVN YQPAKYTRIL
       310        320        330        340        350
ADKVGCNMLD TTDMVECLRN KNYKELIQQT ITPATYHIAF GPVIDGDVIP
       360        370        380        390        400
DDPQILMEQG EFLNYDIMLG VNQGEGLKFV DGIVDNEDGV TPNDFDFSVS
       410        420        430        440        450
NFVDNLYGYP EGKDTLRETI KFMYTDWADK ENPETRRKTL VALFTDHQWV
       460        470        480        490        500
APAVATADLH AQYGSPTYFY AFYHHCQSEM KPSWADSAHG DEVPYVFGIP
       510        520        530        540        550
MIGPTELFSC NFSKNDVMLS AVVMTYWTNF AKTGDPNQPV PQDTKFIHTK
       560        570        580        590        600
PNRFEEVAWS KYNPKDQLYL HIGLKPRVRD HYRATKVAFW LELVPHLHNL
       610        620        630        640        650
NEIFQYVSTT TKVPPPDMTS FPYGTRRSPA KIWPTTKRPA ITPANNPKHS
       660        670        680        690        700
KDPHKTGPED TTVLIETKRD YSTELSVTIA VGASLLFLNI LAFAALYYKK
       710        720        730        740        750
DKRRHETHRR PSPQRNTTND IAHIQNEEIM SLQMKQLEHD HECESLQAHD
       760        770        780        790        800
TLRLTCPPDY TLTLRRSPDD IPLMTPNTIT MIPNTLTGMQ PLHTFNTFSG
       810
GQNSTNLPHG HSTTRV
```

Figure 3: NLGN4X FASTA

Choosing the Best Model

Base Protein Data Bank (PDB) Comparison:

- PDB vs. Fast homology model.
- PDB vs. Threaded model.
- PDB vs. Slow homology model.

Metrics compared included:

- Predicted secondary structure.
- Predicted profiles.
- Functionality.
- Confidence scores.

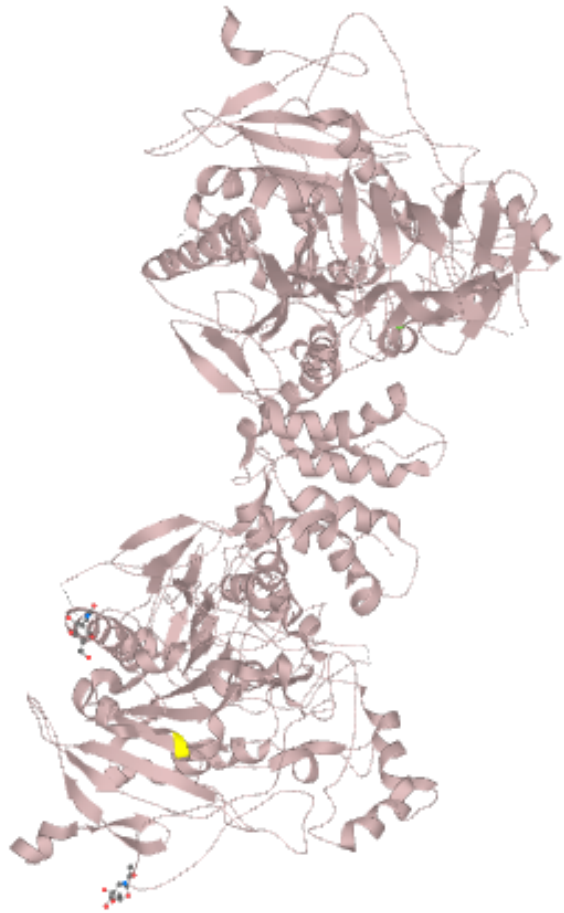

Figure 4: UniProt NLGN4X PDB

Identifying Suspect Variants

Variants for NLGN4X were compiled to form a database.

These variants were then analyzed using resources and tools such as:

- Provean: Protein Variation Effect Analyzer.
- PolyPhen-2: Polymorphism Phenotyping v2 tool.
- GnomAD: Genome aggregation database.
- GWAS Catalogue: Collection of Genome-wide association studies.

Results Overview

1. Gene function.
2. Model analysis.
3. Variant analysis.

NLGN4X Function

Encodes a member of the type B carboxylesterase/lipase protein family involved in the formation and remodeling of the central nervous system [9].

What does this mean?

- Affects the formation of the cognitive structures of the brain.

Molecular functionality:

- Cell-to-cell interaction.
- Cell adhesion molecule binding.
- Signaling receptor activity.

Biological processes:

- Neuronal cell formation.
- Synapse organization.
- Social behavior and learning.

Summary: NLGN4X has a critical impact on the formation and development of the cognitive structures of the brain.

Model Analysis 1

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

Table 1: Model Results

Model Analysis 2

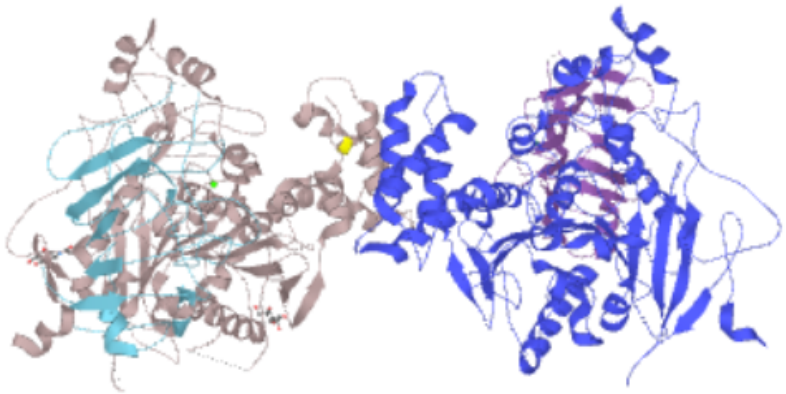

Figure 5: UniProt NLGN4X PDB

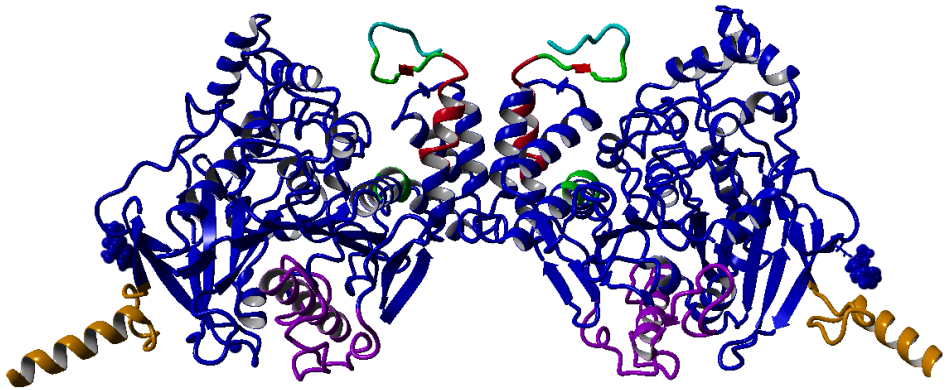

Figure 6: Slow Homology Model

Variant Analysis 1

PolyPhen-2:

- 37 variants.
- 15 probably damaging.

Provean:

- 40 variants.
- 19 deleterious.

Gene Constraint

Category	Exp. no. variants	Obs. no. variants	Constraint metrics		
Synonymous	168	193	Z = -1.51	o/e = 1.15 (1.02 - 1.29)	0 — 1
Missense	367.7	222	Z = 2.7	o/e = 0.6 (0.54 - 0.68)	0 — 1
LoF	19.1	1	pLI = 0.99	o/e = 0.05 (0.02 - 0.25)	0 — 1

exome genome Save plot

Figure 7: GnomAD Results

Variant Analysis 2

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 8: 10 Most Deleterious Variants

Conclusion

- Based on the function of NLGN4X, deleterious variants within this gene will have an adverse effect on neurological development.
- After comparing the three models produced in this project, the YASARA slow homology provided the best 3D representation of NLGN4X.
- After examining a number of predicted deleterious variants, there are still a number of variants of uncertain significance (VUUS) that will have an adverse effect on neurological development.

Thank You

Questions?

References 1

[1] Autism Society
http://www.autism-society.org/what-is/facts-and-statistics/

[2] Autism Speaks
https://www.autismspeaks.org/autism-facts-and-figures

[3] CDC
https://www.cdc.gov/ncbddd/autism/data.html

[4] ClinVar
https://www.ncbi.nlm.nih.gov/gene/57502

[5] Gentics Home Reference
https://ghr.nlm.nih.gov/gene/NLGN4X

References 2

[6] GnomAD
https://gnomad.broadinstitute.org/gene/ENSG00000146938

[7] GWAS Catalog
https://www.ebi.ac.uk/gwas/genes/NLGN4X

[8] I-TASSER
https://zhanglab.ccmb.med.umich.edu/I-TASSER/

[9] NCBI
https://www.ncbi.nlm.nih.gov/gene/57502

[10] PolyPhen-2
http://genetics.bwh.harvard.edu/pph2/

References 3

[11] PROVEAN
http://provean.jcvi.org/index.php

[12] STRING
https://string-db.org/cgi/network.pl?taskId=4WrfKguZfKsg

[13] UniProt (NLGN4X/Q8N0W4)
https://www.uniprot.org/uniprot/Q8N0W4

[14] YASARA
http://www.yasara.org/