A Comparative Study of Variants in the HDAC8 Gene Linked to Cornelia de Lange Syndrome

Jennifer R. Olszyna

Department of Business, University of North Alabama

Abstract

Cornelia de Lange syndrome (CdLS), is a rare developmental disorder characterized by low birth weight, delayed growth, intellectual disability, behavioral problems, and a distinctive facial appearance (thin, arched eyebrows, low set ears, small teeth, and small nose). This project presents a comparative study of two similar variants both located on the HDAC8 gene thought to cause Cornelia de Lange syndrome. The first variant(R164*) is from a known case of a young girl where her variant in the HDAC8 gene has been linked to her diagnosis of Cornelia de Lange syndrome. The second variant(G140R) is of uncertain significance and has data supporting that it has a high probability that this variant will also cause Cornelia de Lange syndrome. This research project seeks to classify this variant of unknown significance by comparing it to the first.

Introduction

Cornelia de Lange syndrome (CdLS), also known less commonly as Brachmann de Lange syndrome or simply De Lange syndrome, is a rare developmental disorder characterized by low birth weight, delayed growth, malformations beginning in the second trimester of pregnancy, intellectual disability, behavioral problems, and a distinctive facial appearance including thin, arched eyebrows, low set ears, small teeth, and a small nose.[1] About ten percent of people with Cornelia de Lange syndrome have a disease-causing variation in one of four genes such as HDAC8.[3, 6] HDAC8 codes for histone deacetylase and plays a critical role in transcriptional regulation, cell cycle progression, and developmental events by altering chromosome structures and affecting transcription factors' access to DNA.[8] Two similar patients with both a preexisting diagnosis of Cornelia de Lange and a variant on their HDAC8 gene were selected from the ClinVar database for this project. Patient one has a termination at position 164 of their amino acid chain in their HDAC8 gene that was later classified in 2016 as "Pathogenic".[9] Patient two has a missense swap of Glycine for Arginine at position 140 of their amino acid chain in their HDAC8 gene; their variation has yet to receive a classification.[10]

Methodology

Patients' variants were assessed through the previously published sequence-to-structure-to-function workflow,[11] comparing the patients' variants to all ClinVar, COSMIC, and gnomAD missense variants for HDAC8. All variants were assessed with PolyPhen2, Provean, and SIFT. YASARA (Australia) software was used to model changes in the protein. All variants were assessed with PolyPhen2, Provean, and SIFT. A total of 20 nanoseconds of molecular dynamics simulations (mds) were run on a lipid membrane embedded HDAC8 protein model.

Results

Check type	Quality Z-score	Comment
Dihedrals	1.703	Optimal
Packing 1D	-0.615	Good
Packing 3D	-0.273	Good
Overall	-0.120	Good

Z-SCORE from Homology Modeling

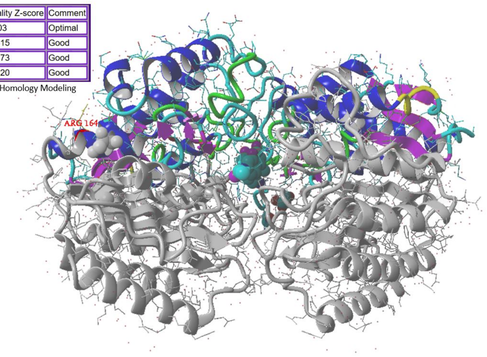

Figure 1:HDAC8 Model showing the effects of a termination of the amino acid chain at position 164

- The HDAC8 protein not fully forming has been linked to patient one's developmental issues

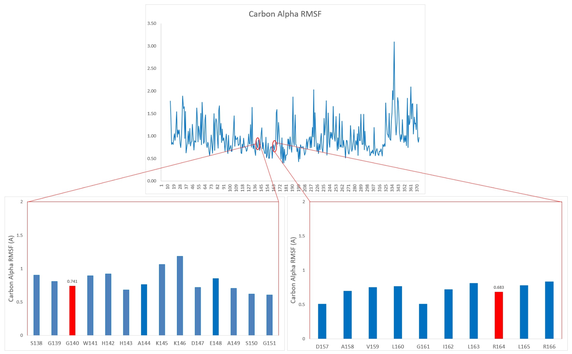

Figure 2:Molecular dynamic simulation of the HDAC8 protein showing the RMSF for each amino acid averaged throughout 20ns of simulation.

- This chart reveals that the site of BOTH mutation is well packed within the structure

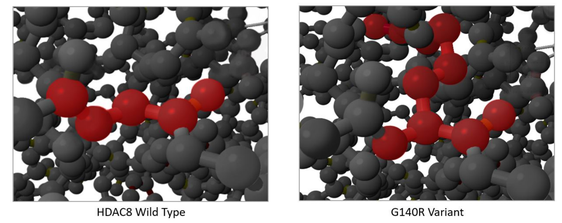

Figure 7:HDAC8 Wild Type and G140R Variant

- G140 mutation changed the shape and length of the amino acid

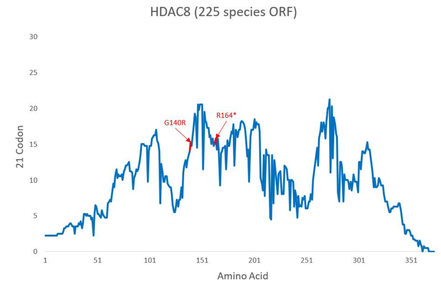

Figure 3:Deep evolutionary analysis using 225 species open reading frame sequences for HDAC8. The plot shows a sliding window calculation for each site (plus ten up and downstream), identifying the most selected and conserved linear motifs within the gene.

- Site of both mutations are shown to fall in highly conserved regions.

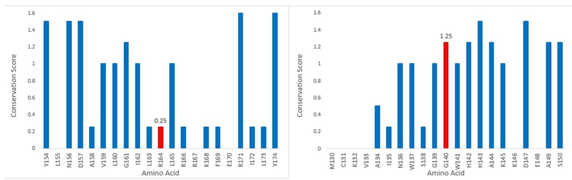

Figure 4:Enlarged view of conservation scores for amino acid 140 and 164 (red) linear motifs. The numbers above represent the percent of sequences with synonymous / non-synonymous variants throughout evolution.

- This chart shows the site of 164 is under very low selection, while 140 is the opposite.

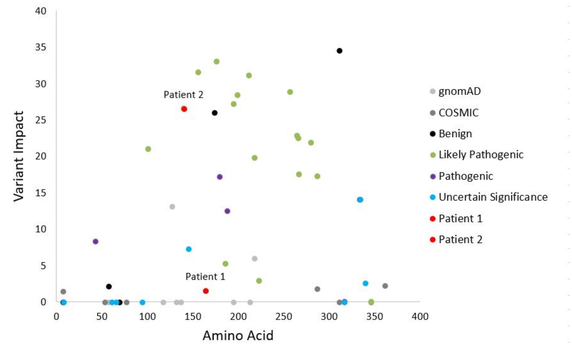

Figure 5:Variant impact scoring for HDAC8 variants according to all files from ClinVar and patients using scores from PROVEAN and SIFT.

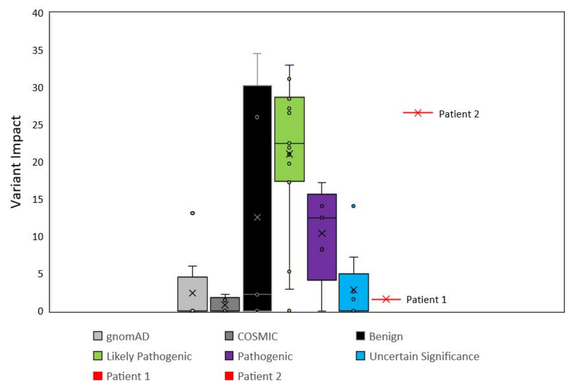

Figure 6:Box and whisker plot for showing the variants combined impact of each group plotted in the previous graph.

Conclusion

Due to the findings so far I believe that this gene will also be a variant significance in causing CdLS due to both patients having similarities in their symptoms and variants. Both of their variations would alter gene regulation, potentially causing many of the developmental problems characteristic of the condition.

References

[1] NORD (National Organization for Rare Disorders), Antonie Kline, Cornelia de Lange Syndrome (CdLS) Foundation, Director of Pediatric Genetics, Harvey Institute of Human Genetics, Greater Baltimore Medical Center. (2019). Cornelia de Lange Syndrome. Retrieved February 2020, from https://rarediseases.org/rare-diseases/cornelia-de-lange-syndrome/

[2] Cornelia de Lange Syndrome: National Library of Medicine (US). Genetics Home Reference [Internet]. Bethesda (MD): The Library; 2015 Sep [cited 2020 Mar 10]. Available from: https://ghr.nlm.nih.gov/condition/cornelia-de-lange-syndromegenes

[3] Genetic and Rare Diseases Information Center (2018). Cornelia de Lange syndrome. Retrieved Febuary 2020, from https://rarediseases.info.nih.gov/diseases/10109/cornelia-de-lange-syndrome.

[4] HDAC8 gene: National Library of Medicine (US). Genetics Home Reference [Internet]. Bethesda (MD): The Library; 2015 Sep [cited 2020 Feb 29]. Available from: https://ghr.nlm.nih.gov/gene/HDAC8.

[5] Deardorff MA, Noon SE, Krantz ID. Cornelia de Lange Syndrome. 2005 Sep 16 [Updated 2016 Jan 28]. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1104/

[6] GeneCards Database. (n.d.). HDAC8 Gene (Protein Coding). Retrieved from https://www.genecards.org/cgi-bin/carddisp.pl?gene=HDAC8keywords=HDAC8

[7] The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 47: D506-515 (2019)

[8] Sebastian Köhler, Leigh Carmody, Nicole Vasilevsky, Julius O B Jacobsen, et al.Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. Nucleic Acids Research. (2018) doi: 10.1093/nar/gky1105

[9] National Center for Biotechnology Information. ClinVar; [VCV000039710.2], https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000039710.2 (accessed April 14, 2020).

[10] National Center for Biotechnology Information. ClinVar; [VCV000523219.1], https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000523219.1 (accessed April 14, 2020).

[11] Prokop JW, Lazar J, Crapitto G, Smith DC, Worthey EA, Jacob HJ. 2017. Molecular modelingin the age of clinical genomics, the enterprise of the next generation. J Mol Model 23: 75.

Acknowledgements

I would would like to thank faculty mentor Dr. Cynthia L. Stenger(UNA) for her insight and guidance as well as Dr. Jeremy W. Prokop(MSU), Dr. Benjie Blair(JSU), and Dr. Jared Painter(UNA) for their additional scientific and technical support. I would also like to thank the UNA Communications Department for their time and help in filming the presentation of this research.

College of BUSINESS
University of NORTH ALABAMA

HUDSONALPHA INSTITUTE FOR BIOTECHNOLOGY
CODE
Characterizing Our DNA Exceptions