Journal of Child Neurology
2021, Vol. 36(13–14) 1218–1222

DOI: 10.1177/08830738211018902
journals.sagepub.com/home/jcn
SAGE

Assessing Prevalence and Carrier Frequency of Succinic Semialdehyde Dehydrogenase Deficiency

Kirt Martin, PhD[1], Alice McConnell, BS[2], and Sarah H. Elsea, PhD[3]

Abstract

Pathogenic variants in *ALDH5A1* cause succinic semialdehyde dehydrogenase (SSADH) deficiency, with >180 cases reported worldwide. However, a nonspecific neurologic presentation and inconsistent variant nomenclature have limited diagnoses. In this study, pathogenic variants in *ALDH5A1* were curated and variant prevalence assessed in the Genome Aggregation Database (gnomAD) to determine a minimum carrier frequency and to estimate disease prevalence. Stringent population variant analysis, including 98 reported disease-associated *ALDH5A1* variants, indicates a pan-ethnic carrier frequency of ~1/340, supporting a prevalence of SSADH deficiency of ~1/460 000 worldwide, with highest carrier frequencies observed in East Asian and South Asian populations. Because heterozygous loss of function alleles are rare in gnomAD and >60% of reported disease-causing variants were missense changes that were not present in gnomAD, the pan-ethnic carrier frequency for SSADH deficiency is likely not fully represented in this study. Additional analyses to investigate the potential impact of more common ALDH5A1 variants with reduced but not deficient enzyme activity, including analysis in diverse populations, are needed to fully assess the prevalence of this ultra-rare disease.

Keywords

succinic semialdehyde dehydrogenase deficiency, *ALDH5A1*, neurometabolic disease, succinic semialdehyde dehydrogenase, SSADH, disease prevalence, carrier frequency

Received November 10, 2020. Received revised April 1, 2021. Accepted for publication April 12, 2021.

Succinic semialdehyde dehydrogenase (SSADH) deficiency (OMIM 271980) is a rare autosomal recessive disorder caused by pathogenic variants in *ALDH5A1* (OMIM 610045) which encodes an oxidoreductase involved in the breakdown of γ-aminobutyric acid (GABA). Reduced function of SSADH, which oxidizes succinic semialdehyde to succinate, impairs the degradation of GABA, leading to accumulation of γ-hydroxybutyric acid (GHB). Symptoms include developmental delay, hyperactivity, behavioral abnormalities hypotonia, ataxia, and seizures.[1]

Diagnosis of SSADH deficiency is confirmed by elevated 4-hydroxybutryric acid in urine organic acid analysis and the identification of biallelic pathogenic variants in *ALDH5A1*.[2] Genetic variants identified by next-generation sequencing followed by in silico predictions may not be sufficient to determine pathogenicity, and the private nature of many rare variants adds to this challenge. As a result, functional analysis is often necessary to determine if the variant is pathogenic, which can be assessed by enzyme assay or by the level of GHB. SSADH enzyme analysis is not readily available clinically, and thus, levels of GHB must be determined. GHB can be clinically assessed in whole blood, plasma, or dried blood spots[3,4]; however, a negative correlation with age has been reported for GHB levels in dried blood spots, wherein the GHB concentrations for some patients previously diagnosed with SSADHD, when assessed at older ages, were below the threshold determined for newborn detection.[3,5-7] These findings indicate that some individuals, when assessed at older ages, may be missed with current biomarker cutoffs, increasing the need to better understand variant pathogenicity.[3,5,6]

ALDH5A1 is located at chromosomal locus 6p22.3, with 2 primary isoforms, in addition to other predicted isoforms.[8] Interpretation and reporting of clinical variants for this ultra-rare condition have been challenged by the lack of knowledge

[1] Department of Neurology, Baylor College of Medicine, Houston, TX, USA
[2] Speragen, Austin, TX, USA
[3] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

Corresponding Author:
Sarah H. Elsea, PhD, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, BCM225, Houston, TX 77030, USA.
Email: elsea@bcm.edu

Table 1. *ALDH5A1* Pathogenic and Likely Pathogenic Variants.

ALDH5A1 variants[a]	n	Start lost/gene deletion	Missense	Frameshift	In-frame deletions/insertions	Splice site	Stop gain
SSADH deficiency reported variants[b]							
Present in gnomAD	31	0	21	4	0	2	4
Not present in gnomAD[c]	67	1	38	11	3	8	6
Variant present only in gnomAD[d]	59	1	17	16	4	14	7
Total P/LP variants	157	2	76	31	7	24	17

Abbreviations: gnomAD, Genome Aggregation Database; P/LP, pathogenic and likely pathogenic; SSADH, succinic semialdehyde dehydrogenase.
[a]Protein effects of *ALDH5A1* genomic variants identified in the gnomAD cohort.
[b]*ALDH5A1* SSADH deficiency–associated variants reported in the literature.
[c]Population data were not available for variants not present in gnomAD.
[d]Pathogenic variants present in gnomAD cohort but not currently reported in the literature. Likely pathogenic variants that involve the same nucleotide and/or amino acid as a reported pathogenic variant resulting in SSADH deficiency; multiple variants within the same codon may be present.

regarding variant function, resulting in most (>60%) variants reported from clinical diagnostic testing as variants of uncertain significance.[9]

Although more than 180 cases of SSADHD have been reported in the literature, the nonspecific presentation of the disease may contribute to underdiagnoses, and prevalence is not known.[10,11] In this study, we assessed the prevalence of known and predicted pathogenic variants identified in the general population to estimate disease incidence and carrier frequency of this ultra-rare disease.

Methods

ALDH5A1 Variant Curation

Curation of *ALDH5A1* variants utilized a single reference canonical transcript (ENST00000348925.2, NM_170740.1) to account for reporting variability in transcript choice and genome reference changes across the reporting period from 1998-2019 to document specific genomic variants for further population based analysis. *ALDH5A1* pathogenic variants were collated from the historical literature, ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), and an internal SSADH Deficiency Association database (https://www.ssadh.net) to provide the most complete assessment of disease-causing variants in *ALDH5A1*. Variants were excluded if they could not be confirmed and validated in the Allele Registry (http://reg.clinicalgenome.org/redmine/projects/registry/genboree_registry/landing) or in other human genome reference databases, including ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) and ClinVar Miner (https://clinvarminer.genetics.utah.edu). The Allele Registry is a genome integration database supported by the ClinGen Project (https://clinicalgenome.org). Nomenclature for all variants follows HGVS standards.

Carrier Frequency and Disease Prevalence

Population-based exome and whole genome data from the Genome Aggregation Database (gnomAD v2.11, https://gnomad.broadinstitute.org/gene/ENSG00000112294?dataset=gnomad_r2_1) were interrogated for *ALDH5A1* pathogenic variants. Variants included reported disease-associated *ALDH5A1* variants identified in individuals with documented SSADH enzyme deficiency (listed in Supplementary Table S1), in addition to pathogenic/likely pathogenic variants present in the populations represented in gnomAD (Supplementary Table S2) taken together to estimate carrier frequency and disease prevalence.[12] Pathogenic variants not previously associated with SSADH deficiency that were identified in gnomAD were strictly defined either as pathogenic or likely pathogenic per the American College of Medical Genetics/Association for Molecular Pathology (ACMG/AMP) variant analysis guidelines.[13] Variants considered in the analysis included start loss, stop gain (nonsense), frameshift, insertions/deletions (indels), predicted splice site defects that included ± 2 bp from the end of an exon directly impacting splice donor or acceptor sites, and variants affecting the same nucleotide or codon previously associated with SSADH deficiency. Carrier frequency was assessed by determining the minor allele frequency of these variants (Supplementary Tables S1 and S2) across all populations present in the publicly accessible gnomAD database. The minor allele frequency (MAF) of each variant was used to determine overall carrier frequency and disease prevalence, as described previously.[14] Disease prevalence estimates were determined based on cumulative *ALDH5A1* pathogenic/likely pathogenic allele frequency. Data were assumed to be in Hardy-Weinberg equilibrium, and standard Hardy-Weinberg calculations were used to determine carrier frequency and disease prevalence.[14]

Results

In this analysis, we curated 98 reported pathogenic variants associated with SSADH deficiency, representing all forms of mutations typically observed with single gene, exome, or genome sequencing (Supplementary Table S1; see Methods). Single-nucleotide variants (SNVs) were the most common type of mutation in this variant set, resulting in missense changes in $\sim 60\%$ of cases; however, insertion/deletion (indel) variants that resulted in frameshifts or aberrant splicing were also frequently observed, as well as complete gene deletions (Table 1, Supplementary Table S1). Many variants reported in the literature could not be confirmed because of either incorrectly or incompletely reported nucleotide sequence or lack of inclusion of adequate ALDH5A1 isoform or genome information such that the specific variant could not be confidently determined. Thus, some reported variants were not included in this analysis.

To determine the population frequency of *ALDH5A1* pathogenic alleles, we interrogated the gnomAD data set, which contains genome variant data from $\sim$125 748 exomes and 15 708 whole genomes mapped to the GRCh37/hg19 reference

sequence. Inclusion in this analysis required that variants be interpretable based on the ACMG/AMP variant analysis guidelines.[13] These extensive guidelines provide a methodological process for variant interpretation that is used by clinical diagnostic laboratories for the analysis of genomic variants. We identified a total of 795 variants in *ALDH5A1* in the gnomAD v2.1.1 data set and found ~32% (31/98) of the reported pathogenic variants present in this database (Supplementary Table S1); however, an additional 18 variants in gnomAD involved the same nucleotide and/or amino acid as the reported pathogenic variant; thus, these likely pathogenic variants were also included in this analysis (Supplementary Table S2). In addition, 41 pathogenic/likely pathogenic loss of function variants not previously reported in patients with SSADH deficiency were also identified in the gnomAD population and contributed to this analysis (Supplementary Table S2). In total, although 157 pathogenic/likely variants were identified between gnomAD and reported SSADH deficiency cases, population data were available for only 90 disease-associated and/or pathogenic/likely pathogenic variants (Table 1).

The 90 *ALDH5A1* variants that met the criteria for inclusion in disease frequency analysis included 31 previously reported variants associated with SSADH deficiency (Table 1, Supplementary Table S1), plus 59 additional variants present only in gnomAD (Table 1, Supplementary Table S2). Analyses included 63 SNVs resulting in alterations in translation start/stop or splicing, or in missense changes affecting enzyme function, in addition to 27 intragenic insertions or deletions resulting in frameshift or splicing errors. The inability to clearly assess pathogenicity of missense variants in the gnomAD data set resulted in strict inclusion of only those missense variants reported in individuals with SSADH deficiency or those variants meeting criteria as likely pathogenic by ACMG/AMP variant interpretation standards (summarized in Table 1, Supplementary Tables S1 and S2).

Heterozygous *ALDH5A1* pathogenic/likely pathogenic variants were present in ~0.15% of all chromosomes in the gnomAD pan-ethnic population, with the greatest variety of variants identified in non-Finnish Europeans and a single variant (NM_170740.1: c.858del; p.Asp287IlefsTer27) uniquely present in 0.04% of Latino chromosomes. A single predominant variant each was present in ~0.1% of South Asian chromosomes (NM_170740.1: c.518G>A; p.Arg173His) and East Asian chromosomes (NM_170740.1: c.515G>A; p.Arg172His), with carrier estimates of 1/255 to 1/262, respectively (Table 2). Hardy-Weinberg population analysis of the total MAF of these pathogenic/likely pathogenic variants across the multiethnic populations in gnomAD resulted in an estimated pan-ethnic carrier frequency of ~1/340, with an estimated disease prevalence of ~1/460 000 (Table 2).

Table 2. *ALDH5A1* Carrier Frequency and Disease Prevalence Estimates.

Population	*ALDH5A1* MAF	Carrier frequency	Disease prevalence
Pan-ethnic	1/681	1/341	1/463 761
European (non-Finnish)	1/713	1/357	1/508 623
African	1/653	1/327	1/426 409
East Asian	1/508	1/255	1/258 064
South Asian	1/524	1/262	1/274 232

Abbreviations: gnomAD, Genome Aggregation Database; MAF, minor allele frequency, sum all pathogenic variants.

Discussion

Knowledge of the pathogenicity of genome variants is critical for proper diagnosis. Although the pathogenicity of some genome variants can be clearly determined,[13] many variants are left to prediction programs to assess potential risk. Missense variants assessed by these bioinformatic risk calculations may provide guidance regarding the impact of a nucleotide change on the amino acid sequence; however, programs may conflict in their assessments, and they may or may not accurately predict consequences of enzyme function.[15]

In this study, we curated reported *ALDH5A1* variants associated with SSADH deficiency to better inform genetic testing laboratories and to estimate carrier frequency and disease prevalence. We followed stringent guidelines for variant interpretation and assessment in our attempt to assess all reported cases from the literature through 2019. Unfortunately, not all reported variants in *ALDH5A1* were described and reported within the parameters of these standard guidelines, so not all variants in the literature could be included in this analysis. This stringent definition for variant description and pathogenicity results in the very likely possibility that some pathogenic variants may not be reported as pathogenic in clinical diagnostic testing. A complete analysis of gene deletions as assessed by chromosomal microarray was not performed in this study; thus, the contribution of partial or complete gene deletions to SSADH deficiency is not included in the carrier frequency calculation.

The full spectrum of the SSADH deficiency phenotype is not known, further complicating accurate and efficient diagnosis. Phenotypic variability and its association with the level of enzyme activity has not been well defined, with reported enzyme activity in affected individuals ranging from <1% to ~20%,[16,17] Common variants in ALDH5A1 have been reported with reduced enzyme activity[2,16,18]; however, the potential contributions of those variants to disease have not been described. Although parents who are carriers of the pathogenic variants present in their children do not have symptoms or features of SSADH deficiency, carriers do have reduced enzyme activity,[18,19] and modifiers of enzyme function that may impact the presence or severity of any features have not been described. Furthermore, one common variant, c.709G>T;p.Ala237Ser (rs62621664), was present in ~1% of all chromosomes across all populations in gnomAD. Although c.709G>A;p.Ala237Thr was identified as a pathogenic variant (Supplementary Table S1), c.709G>T was excluded from this population analysis because of the presence of homozygous individuals across all

populations in the gnomAD data set. However, the contribution of this variant to disease cannot be fully excluded and may have implications in the heterozygous state when in combination with a null variant.

The current analysis indicates that SSADH deficiency is likely more common in Asian populations; however, this preliminary result comes from limited population data and, thus, is insufficient to support a firm conclusion at this time. Genomic data are limited from populations other than non-Finnish Europeans and, as such, these analyses may be over- or under-estimated. That said, these calculations are likely an underestimate of the pan-ethnic carrier frequency and disease prevalence because of the presence of only about one-third of the reported pathogenic variants in gnomAD. Further, curation of variant pathogenicity is limited because of the lack of knowledge related to function of missense variants in ALDH5A1. The nature of the private genetic variants and their absence in gnomAD complicates the analysis of pathogenicity for the variants present in this pan-ethnic population. Interestingly, the general population databases also show $\sim 50\%$ fewer pathogenic loss of function variants than expected for *ALDH5A1*, suggesting that loss of function/null variants are not well tolerated in the general (healthy) adult population; thus, additional studies are required to better understand genome variation and the impact on enzyme function, as well as any potential health concerns in carriers of *ALDH5A1* null alleles, particularly with aging.

Data in this study support a pan-ethnic carrier frequency of $\sim 1/340$, suggesting the prevalence of SSADH deficiency is $\sim 1/460\,000$ worldwide, with disease prevalence higher in the East Asian and South Asian populations (Table 2). These data indicate that the prevalence of SSADH deficiency, though similar to other ultrarare autosomal recessive developmental epileptic encephalopathies, including GABA transaminase deficiency, adenylosuccinate lyase deficiency, and the citrate transporter deficiencies, may be slightly more common. This stringent analysis could be improved with a more rigorous assessment of the missense variants present in the general population. Further, additional studies are required to assess the potential impact of more common ALDH5A1 variants with reduced, but not deficient, activity and to assess more broadly *ALDH5A1* variation in diverse populations.

Author Contributions

SHE designed the study; AM contributed data; KM participated in the acquisition and analysis of data; SHE analyzed the data and drafted the body of the manuscript; and KM, AM, and SHE edited and approved the final manuscript. All authors contributed to the acquisition, analysis, and interpretation of data, gave final approval to the manuscript, and agreed to be accountable for all aspects of the work in ensuring that questions relating to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: AM is employed by Speragen. Remaining authors declare no conflicts of interest.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding to support this work was provided, in part, by the Speragen, Inc, the SSADH Association, and private donors.

ORCID iD

Sarah H. Elsea, PhD https://orcid.org/0000-0002-1400-8519

Supplemental Material

Supplemental material for this article is available online.

Ethical Approval

Variant data assessed in this study were curated from the literature, from publicly accessible databases, or from anonymized variant data voluntarily submitted to the SSADH Association variant database. All studies conformed to the Declaration of Helsinki (as revised in 2013) and with the Declaration of Taipei on Ethical Considerations regarding Health Databases and Biobanks in 2016.

References

1. Jakobs C, Bojasch M, Mönch E, Rating D, Siemes H, Hanefeld F. Urinary excretion of gamma-hydroxybutyric acid in a patient with neurological abnormalities. The probability of a new inborn error of metabolism. *Clin Chim Acta*. 1981;111(2-3):169-178.
2. Chambliss KL, Hinson DD, Trettel F, et al. Two exon-skipping mutations as the molecular basis of succinic semialdehyde dehydrogenase deficiency (4-hydroxybutyric aciduria). *Am J Hum Genet*. 1998;63(2):399-408.
3. Forni S, Pearl PL, Gibson KM, Yu Y, Sweetman L. Quantitation of gamma-hydroxybutyric acid in dried blood spots: feasibility assessment for newborn screening of succinic semialdehyde dehydrogenase (SSADH) deficiency. *Mol Genet Metab*. 2013; 109(3):255-259.
4. Stove CP, Ingels AS, Lambert WE. Quantitation of gamma-hydroxybutyric acid in dried blood spots in newborn screening. *Mol Genet Metab*. 2013;110(1-2):195.
5. Jansen EE, Vogel KR, Salomons GS, Pearl PL, Roullet JB, Gibson KM. Correlation of blood biomarkers with age informs pathomechanisms in succinic semialdehyde dehydrogenase deficiency (SSADHD), a disorder of GABA metabolism. *J Inherit Metab Dis*. 2016;39(6):795-800.
6. Johansen SS, Wang X, Sejer Pedersen D, et al. Gamma-hydroxybutyrate (GHB) content in hair samples correlates negatively with age in succinic semialdehyde dehydrogenase deficiency. *JIMD Rep*. 2017;36:93-98.
7. Brown M, Ashcraft P, Arning E, Bottiglieri T, Roullet JB, Gibson KM. Gamma-hydroxybutyrate content in dried bloodspots facilitates newborn detection of succinic semialdehyde dehydrogenase deficiency. *Mol Genet Metab*. 2019;128(1-2):109-112.
8. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019;47(D1):D506-d515.
9. Henrie A, Hemphill SE, Ruiz-Schultz N, et al. ClinVar miner: demonstrating utility of a Web-based tool for viewing and filtering ClinVar data. *Hum Mutat*. 2018;39(8):1051-1060.

10. Gibson KMJ, C. Disorders of beta- and alpha-amino acids in free and peptide-linked forms. In: Scriver CR, Beaudet AL, Sly WS, et al, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York, NY: McGraw-Hill; 2001: 2079-2105.
11. Pearl PL, Gibson KM, Acosta MT, et al. Clinical spectrum of succinic semialdehyde dehydrogenase deficiency. *Neurology*. 2003;60(9):1413-1417.
12. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443.
13. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
14. Cappuccio G, Donti TR, Hubert L, Sun Q, Elsea SH. Opening a window on lysosomal acid lipase deficiency: Biochemical, molecular, and epidemiological insights. *J Inherit Metab Dis*. 2019; 42(3):509-518.
15. Walters-Sen LC, Hashimoto S, Thrush DL, et al. Variability in pathogenicity prediction programs: impact on clinical diagnostics. *Mol Genet Genomic Med*. 2015;3(2):99-110.
16. Pattarelli PP, Nyhan WL, Gibson KM. Oxidation of [U-14C]succinic semialdehyde in cultured human lymphoblasts: measurement of residual succinic semialdehyde dehydrogenase activity in 11 patients with 4-hydroxybutyric aciduria. *Pediatr Res*. 1988; 24(4):455-460.
17. Akaboshi S, Hogema BM, Novelletto A, et al. Mutational spectrum of the succinate semialdehyde dehydrogenase (ALDH5A1) gene and functional analysis of 27 novel disease-causing mutations in patients with SSADH deficiency. *Hum Mutat*. 2003;22(6):442-450.
18. Aoshima T, Kajita M, Sekido Y, et al. Mutation analysis in a patient with succinic semialdehyde dehydrogenase deficiency: a compound heterozygote with 103-121del and 1460 T>A of the ALDH5A1 gene. *Hum Hered*. 2002;53(1):42-44.
19. Menduti G, Biamino E, Vittorini R, et al. Succinic semialdehyde dehydrogenase deficiency: the combination of a novel ALDH5A1 gene mutation and a missense SNP strongly affects SSADH enzyme activity and stability. *Mol Genet Metab*. 2018;124(3):210-215.