Tracking Genomic Database Queries to Estimate the Prevalence of Rare Diseases: *ASAH1* in the VarSome Database

N. Confer[1], K. Crosby[1], A. Joyner[2], A. Solyom[1]

[1]Aceragen, Durham, North Carolina, [2]Saphetor, Lausanne, Switzerland

ASHG 2022 - PB3213

Introduction

- Pathogenic variants in *ASAH1* cause acid ceramidase deficiency.
- Farber disease and spinal muscular atrophy with progressive myoclonic epilepsy (SMA-PME) result from ceramide accumulation.

ASAH1 gene mutations

↓

Acid ceramidase enzyme deficiency (lysosomal)

↓

Ceramide buildup

↓

Macrophage-driven inflammation / Cellular dysfunction

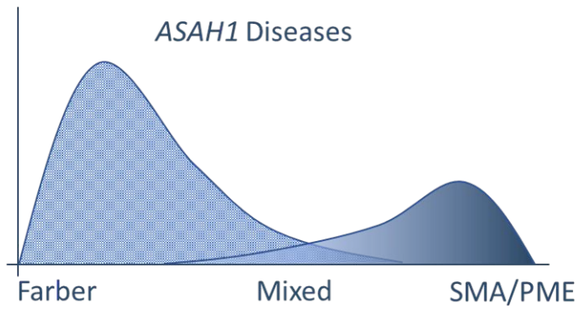
↓

Symptoms of disease

- Fewer than 250 cases of *ASAH1*-related disease are reported in the literature suggesting low prevalence, with frequency not well established.
- Seasonal flu and other conditions utilize aggregated search volumes to aid in surveillance and prevalence estimation[1] .
- A review of recent VarSome query volumes suggests the prevalence of disease associated with this gene is higher than previously estimated.

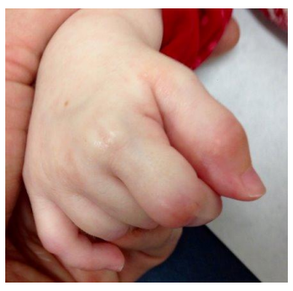
ASAH1-Related Diseases

ASAH1-related disorders present as a phenotypic spectrum.

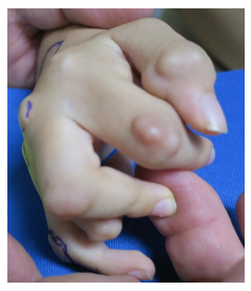


Farber disease
- **Joint disease** (arthritis and/or contractures)
- **Subcutaneous nodules** (lipogranulomas)
- **Hoarse voice** (due to laryngeal nodules)

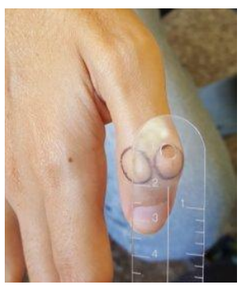
SMA-PME
- **Lower motor neuro disease** (proximal muscle weakness)
- **Progressive myoclonic epilepsy** (myoclonic seizures, myoclonus)



Rapid progression
1 year old



Moderate progression
4 years old



Slow progression
28 years old

Images are from Natural History study showing subcutaneous nodules of the hands of patients with Farber disease.

Methods

Previously published cases of *ASAH1*-releated disease were collected with countries of patient origin plotted as a global heatmap. Where applicable, data from the Observational and Cross-Sectional Cohort Study of the Natural History and Phenotypic Spectrum of Farber Disease (NCT03233841) were included in the analysis.

Records of genetic variant searches for the *ASAH1* gene in the VarSome search engine were used for this analysis. VarSome is a suite of data-driven bioinformatics solutions for clinicians and researchers created by Saphetor SA. Saphetor is a global precision-medicine company dedicated to large-scale identification and interpretation of human genetic variants. VarSome.com is a professional community and search engine that is freely accessible, featuring a community-driven knowledge base that enables flexible queries across more than 140 genetic and genomic data resources.

Data from *ASAH1* gene queries in the past two years (Sept 2020 thru Aug 2022) were collected using the VarSome search engine. These data were compared to published cases to determine if a difference existed in countries represented.

Additional VarSome queries over the past 12 months (Oct 2021 thru Sep 2022) were collected for lysosomal disorders having both an approved therapy and an established prevalence. Additional disorders were included based upon assumed historical prevalence comparability to *ASAH1*-releated disease (*ARG1* deficiency) and a recent therapy approval but with poorly established prevalence (MOCDA (*MOCS1)*). VarSome search counts involving pathogenic and likely pathogenic (P/LP) variants were graphed in relation to established prevalence for each disorder with trendline providing estimation of prevalence for disorders with an unknown prevalence.

References

1. Carneiro HA, Mylonakis E. Clin Infect Dis. 2009 Nov 15;49(10).
2. Park KS. Mol Genet Metab Rep. 2021 Feb 26;27.
3. D'Avanzo F, *et. al*. Int J Mol Sci. 2021 Dec 15;22(24).
4. Del Angel G, *et. al*. J. Hum Mutat. 2019 Nov;40(11).
5. Gahl WA, *et. al.* N Engl J Med. 2002 Jul 11;347(2).
6. Giugliani R, *et. al*. Orphanet J Rare Dis. 2021 May 22;16(1).
7. Catsburg C, *et. al*. Orphanet J Rare Dis. 2022 Mar 2;17(1).
8. Yu FPS, *et. al.* Orphanet J Rare Dis. 2018 Jul 20;13(1).
9. Spiegel R, *et. al.* J Inherit Metab Dis. 2022 May;45(3).

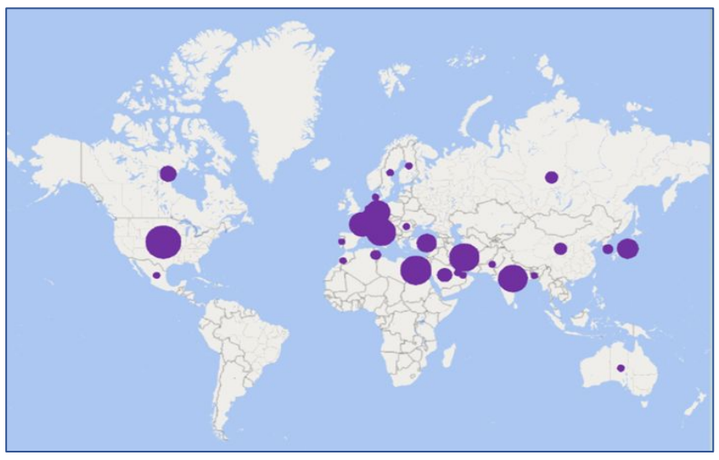
Disclosures: N. Confer and K. Crosby are employees of Aceragen, Inc. A. Joyner is an employee of Saphetor SA.
Contact information: nconfer@aceragen.com

Results

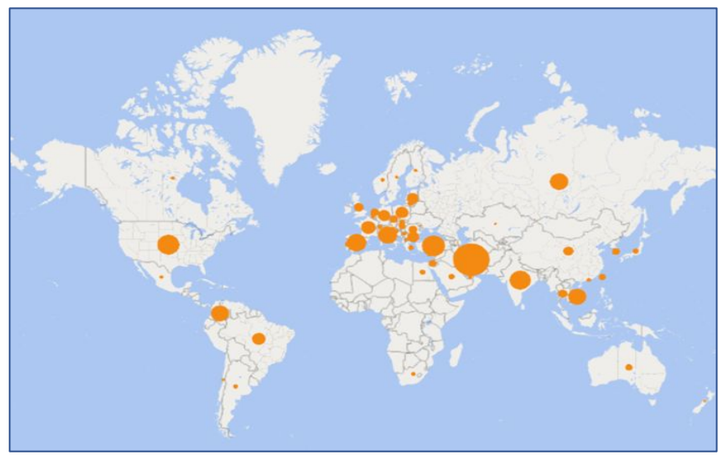
Approximately 250 cases of *ASAH1*-related disease have been published from 32 countries. The countries most represented in the literature differ from the countries with the most queries in the VarSome database.

Published (Country)	Cases (%)	VarSome (Country)	Queries (%)
Egypt	27 (14)	Iran	125 (16)
United States	25 (13)	Turkey	87 (11)
Iran	16 (9)	Spain	63 (8)
France	16 (9)	India	47 (6)
Italy	15 (8)	Russia	39 (5)

Comparison of most represented countries in published literature vs. VarSome database queries. 187 cases from the published literature and natural history study of known origin were included. 787 unique VarSome queries from Sept 2020 thru Aug 2022 were included.



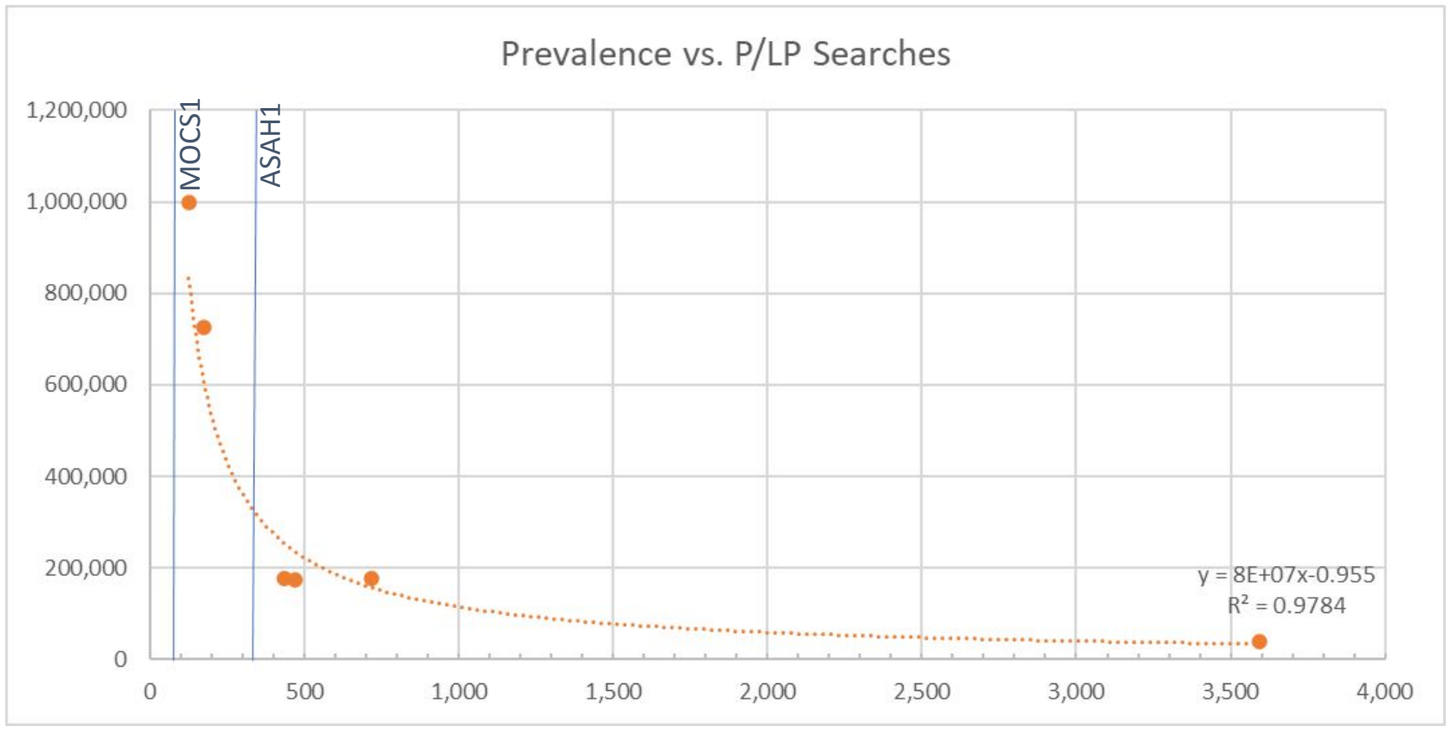
Geographic distribution by country for published cases, with size of dot representing relative volume.



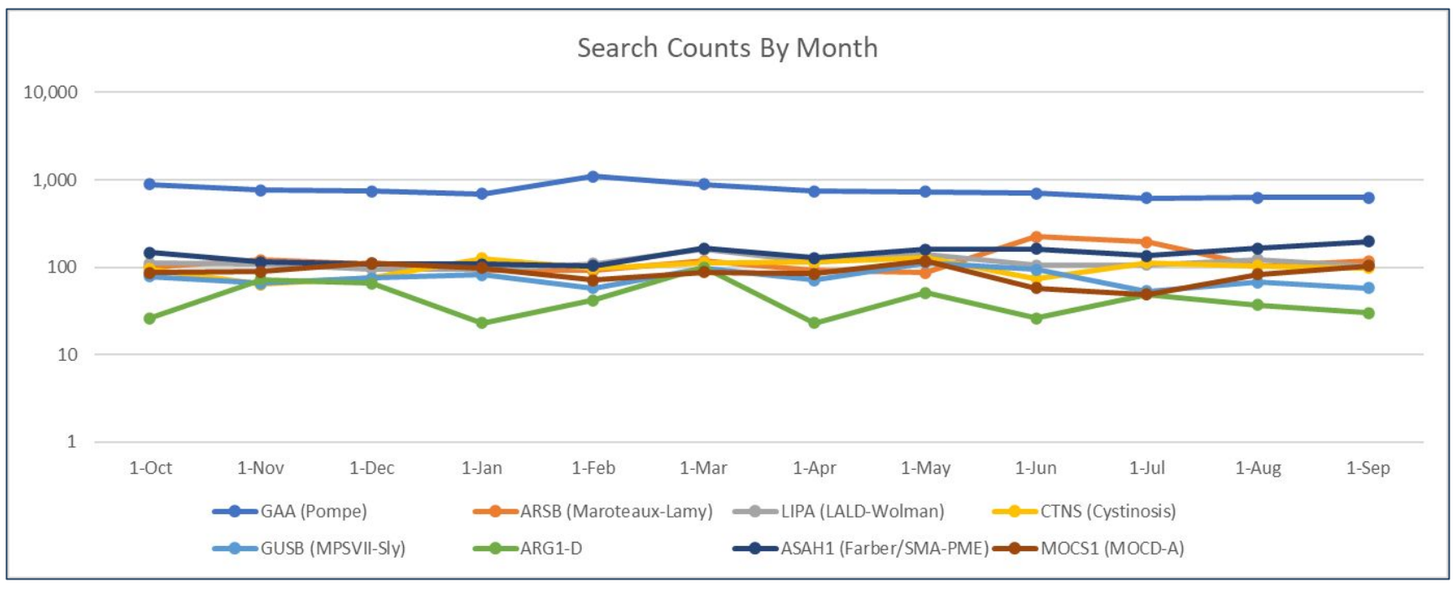
Geographic distribution by country of origin for VarSome queries, with size of dot representing relative volume.

Gene (Disease)	Published / Inferred Prevalence (1/x)	Total VarSome searches	Searches with P/LP	% P/LP	% VUS
GAA (Pompe)	40,000[2]	9,106	3,594	39.4	30.9
ARSB (Maroteaux-Lamy)	177,000[3]	1,449	719	49.6	30.1
LIPA (LALD-Wolman)	177,452[4]	1,369	436	31.8	40.3
CTNS (Cystinosis)	175,000[5]	1,199	468	39.0	39.2
GUSB (MPSVII-Sly)	1,000,000[6]	911	126	13.8	65.3
ARG1 (Arg-1 Def.)	726,000[7]	544	174	32.0	50.4
ASAH1 (Farber/SMA-PME)	1,000,000[8] / 311,000	1,695	321	18.9	43.7
MOCS1 (MOCD-A)	411,187[9] / 1,100,000	1,037	78	7.5	66.9

VarSome queries over the past 12 months for lysosomal disorders with an approved therapy and evidenced prevalence (rows 1-5) and a metabolic disease with a published prevalence similar to *ASAH1*-releated disease (row 6). These search volumes were used to generate an inferred prevalence of *ASAH1*-related diseases, which may be more accurate than the published prevalence. The inferred prevalence for MOCD-A (*MOCS1)* was generated as a second test gene due to the suspected rarity of cases despite the published prevalence of 1:411,187.



Disease prevalence values were graphed in relation to search volume for known P/LP variants for the past 12 months. A trend line was established to estimate disease prevalence for the two unknown diseases, *ASAH1*-related diseases and MOCDA (*MOCS1).*



The VarSome search counts for each involved gene were graphed based upon monthly volume (Oct 2021 – Sept 2022).

Conclusions

- Accurate prevalence estimates for rare diseases are important for resource allocation, drug development, and informing public policy.
- Using published cases to generate disease prevalence estimates may be flawed, as countries less represented in the literature are resultantly under-represented or absent.
- VarSome queries for *ASAH1* variants identified additional geographies with recent cases not represented in the literature, with South America, Eastern Europe, the Middle East, and Southeast Asia notable.
- The inferred prevalence value for *ASAH1*-related disorders was determined to be 1:311,000, while MOCDA (*MOCS1)* was determined to be 1:1,280,000, with both in agreement with recent case volumes.
- VarSome query volume may be a reasonable method to infer disease prevalence for rare monogenic diseases.