Molecular diagnostic findings of lysosomal storage diseases in children and adults suspected to have inborn errors of metabolism

Yuan-Yuan Ho, Laura Murillo, Daniel Beltran, Rachel Hart, Hannah White, Stacey Wong, Michelle Fox, Tom Winder, Britt Johnson

Metabolic Genetics Group, Clinical Genomics, Invitae, San Francisco, CA

BACKGROUND

- Overall prevalence of lysosomal storage diseases (LSDs) is estimated at 1:5000 (0.02%).
- Over 50 genes have been associated with LSDs.
- Aim to determine 1) The attribution of LSDs in 7,581 probands suspected to have inborn errors of metabolism (IEM). 2) Mutation spectrum of LSD genes.

Table 1. List of 52 LSD genes tested in this study

AGA	ARSA	ARSB	ASAH1	ATP13A2	CLN2 (TPP1)	CLN3	CLN5
CLN6	CLN8	CTNS	CTSA	CTSD	CTSF	CTSK	DNAJC5
FUCA1	GAA	GALC	GALNS	GLA	GLB1	GM2A	GNPTAB
GNPTG	GNS	GRN	GUSB	HEXA	HEXB	HGSNAT	HYAL1
IDS	IDUA	KCTD7	LAMP2	LIPA	MAN2B1	MANBA	MCOLN1
MFSD8	NAGA	NAGLU	NEU1	NPC1	NPC2	PPT1	PSAP
SGSH	SLC17A5	SMPD1	SUMF1				

METHODS

- Genomic DNA obtained from whole blood or saliva of the tested individuals was enriched for targeted regions by using a hybridization-based protocol, and then sequenced with Illumina technology.
- Next-generation sequencing (NGS) and validated custom-built algorithms that use depth-of-coverage information and split-read detection were employed to identify copy number variants,[1] small and large indels, and single nucleotide changes.
- Targeted regions were sequenced with ≥50x depth or were supplemented with additional analysis. Reads were aligned to a reference sequence (GRCh37), and sequence changes were called in the context of clinically relevant transcripts of 52 genes known to be associated with LSDs.
- Enrichment and analysis focused on the coding sequence of the referenced transcripts, 10 bp of flanking intronic sequence, and other specific genomic regions demonstrated to be causative of disease at the time of assay design.
- Promoters, untranslated regions, and other non-coding regions were not otherwise interrogated.
- Invitae uses a proprietary, validated algorithm to detect deletions and duplications using NGS. The algorithm calls exonic deletions and duplications by calculating the statistical likelihood of each copy number state by comparing depth of sequence coverage at targeted exons to depth measured from a set of baseline samples.
- Sequence changes were interpreted in the context of referenced transcripts in accordance with ACMG guidelines.[2]
- Likely pathogenic or pathogenic variants identified were confirmed by Sanger sequencing (single nucleotide changes) or aCGH analysis (CNV changes).

RESULTS

- The proportional distribution of P/LP/VUS variants is illustrated in **Figure 1**. Overall 310 unique variants in LSD genes classified as P (pathogenic), LP (likely pathogenic), or VUS (variant of uncertain significance) were identified in 412 of the 7,581 probands.
- Distribution of P/LP/VUS variants by age group is presented in **Figure 2**.

Figure 1. Spectrum and distribution of P/LP/VUS variants identified in LSD genes

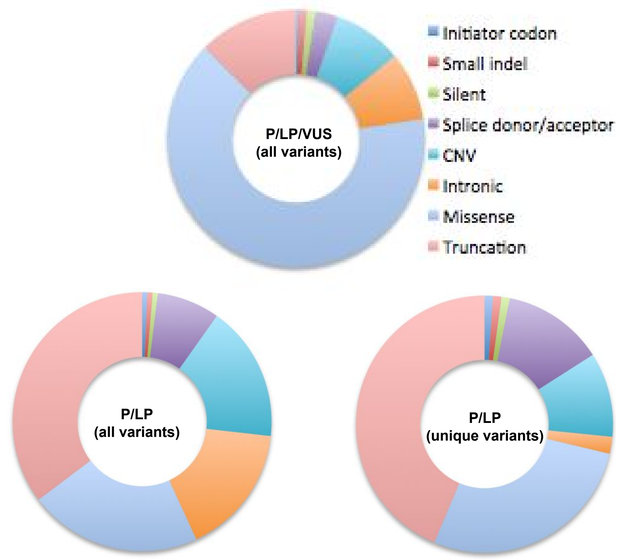
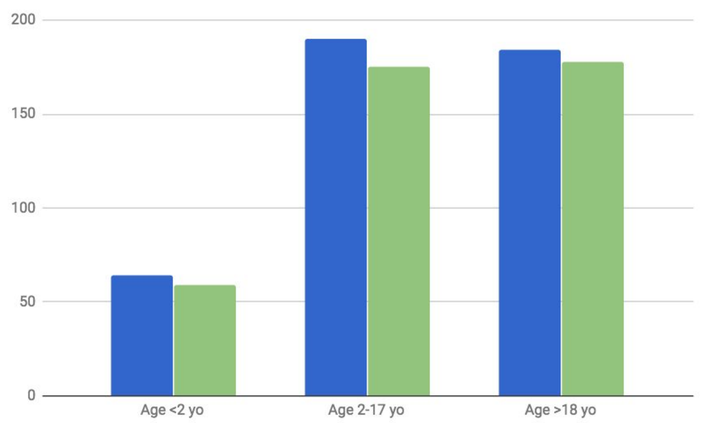


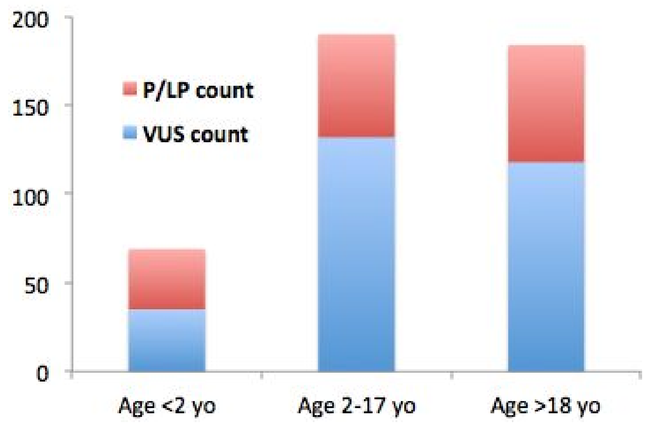
Figure 2. Distribution of P/LP/VUS variants in LSD genes by patient age group. variant count, proband count



RESULTS

- Distribution of P/LP/VUS variants identified in LSD genes varies by age group (**Figure 3**).

Figure 3. Distribution of P/LP and VUS variants in LSD genes by age



- A positive diagnosis was defined as monoallelic LP/P identified for autosomal dominant or X-linked LSDs and biallelic LP/P identified for autosomal recessive LSDs.
- In probands with positive molecular diagnoses, the majority of LP/P variants in infantile patients occurred in the *HEXA* gene. In pediatric patients, LP/P variants clustered in *CLN3*. In adult patients, pathogenic mutations clustered in *GAA*, *GLA*, and *LAMP2* **(Table 2).**

Table 2. LSD genes mutated in probands with positive molecular diagnosis of **LSDs**

Age (positive case count)	Genes mutated in LSD positive cases (LP/P mutant allele count)
< 2 yo (12 cases)	***GALC***(2), ***GAA***(2*), ***IDUA***(2), ***LIPA***(2*), ***NPC1***(4), ***HEXA***(12#) Recurring mutations: *HEXA* p.427Ilefs*5, p.Gly269Ser, Exon 11-13 deletion
2-17 yo (26 cases)	***ARSA***(4), ***ATP13A2***(2), ***CLN3***(8#), ***CLN5***(2), ***CLN6***(2), ***CTNS***(2), ***GAA***(4), ***GLA***(3), ***HEXA***(2), ***IDS***(4), ***LAMP2***(3), ***MFSD8***(2), ***NPC1***(2), ***SMPD1***(2) Recurring mutations: CLN3 Exon 8-9 deletion
> 17 yo (28 cases)	***ARSA***(2), ***CTNS***(2*), ***GAA***(6#), ***GLA***(12#) ***GRN***(2), ***IDS***(1), ***LAMP2***(8#) Recurring mutations: *GAA* c.-32-13T>G; GLA p.Gly325Asp, p.Tyr152*, p.Asn215Ser.

mutation clusters, *homozygous

CONCLUSIONS

- **66 out of 7,581 probands with suspected IEM have positive molecular diagnosis of LSDs.**
- **Different LSD genes are associated with LSDs in infants and adults.**
- **The actual attribution of LSDs to IEM is likely higher in this study cohort because not all probands received comprehensive testing for all 52 LSD genes.**

References: 1) Deletion/duplication analysis (https://www.invitae.com/en/assay/). 2) Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria (PMID: 28492532).

Disclosures: All authors are stockholders and employees of Invitae.