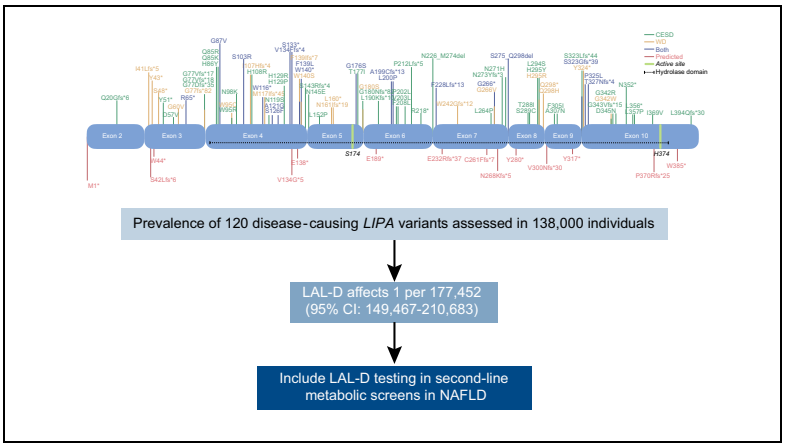
Research Article
Genetic and Metabolic Diseases

JOURNAL
OF HEPATOLOGY

The global prevalence and genetic spectrum of lysosomal acid lipase deficiency: A rare condition that mimics NAFLD

Graphical abstract



Highlights

- There are 98 disease-causing *LIPA* variants associated with LAL-D.
- An additional 22 predicted pathogenic *LIPA* variants have been identified in humans.
- LAL-D has an estimated prevalence of 1 per 177,000.
- Clinicians can be reassured that LAL-D is an ultra-rare mimic of NAFLD.
- Consider LAL-D testing in a second-line metabolic screen in patients with atypical NAFLD.

Authors

Anna Carter, Simon Mark Brackley, Jiali Gao, Jake Peter Mann

Correspondence

jm2032@cam.ac.uk
(J.P. Mann)

Lay summary

Lysosomal Acid Lipase Deficiency (LAL-D) is a rare genetic condition that can cause severe liver disease, but it is difficult to diagnose and sometimes can look like simple fatty liver. It was not clear how common LAL-D was and whether many cases were being missed. To study this, we searched for all genetic mutations that could cause LAL-D, calculated how common those mutations were, and added them up. This let us estimate that LAL-D affects roughly 1 in 175,000 people. We conclude that LAL-D is a very rare condition, but it is treatable so may be included in a 'second-line' of tests for causes of fatty liver.

http://dx.doi.org/10.1016/j.jhep.2018.09.028
© 2018 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved. J. Hepatol. 2019, 70, 142–150

Research Article
Genetic and Metabolic Diseases

JOURNAL OF HEPATOLOGY

The global prevalence and genetic spectrum of lysosomal acid lipase deficiency: A rare condition that mimics NAFLD

Anna Carter[1], Simon Mark Brackley[2], Jiali Gao[2], Jake Peter Mann[3,4,*]

[1]*Manchester University Foundation Trust, Manchester, United Kingdom;* [2]*University of Cambridge, School of Clinical Medicine, Cambridge, United Kingdom;* [3]*University of Cambridge, Department of Paediatrics, Cambridge, United Kingdom;* [4]*University of Cambridge, Institute of Metabolic Science-Metabolic Research Laboratories, Cambridge, United Kingdom*

Background & Aims: Lysosomal acid lipase deficiency (LAL-D) is an autosomal recessive condition that may present in a mild form (cholesteryl ester storage disease [CESD]), which mimics non-alcoholic fatty liver disease (NAFLD). It has been suggested that CESD may affect 1 in 40,000 and is under-diagnosed in NAFLD clinics. Therefore, we aimed to estimate the prevalence of LAL-D using analysis of genetic variation in *LIPA*.

Methods: MEDLINE and EMBASE were systematically searched for previously reported disease variants and prevalence estimates. Previous prevalence estimates were meta-analysed. Disease variants in *LIPA* were annotated with allele frequencies from gnomAD and combined with unreported major functional variants found in humans. Pooled ethnicity-specific prevalences for LAL-D and CESD were calculated using the Hardy-Weinberg equation.

Results: Meta-analysis of existing genetic studies estimated the prevalence of LAL-D as 1 per 160,000 (95% CI 1 per 65,025–761,652) using the allele frequency of c.894G>A in *LIPA*. A total of 98 previously reported disease variants in LIPA were identified, of which 32/98 were present in gnomAD, giving a prevalence of 1 per 307,482 (95% CI 257,672–366,865). Wolman disease was associated with more loss-of-function variants than CESD. When this was combined with 22 previously unreported major functional variants in *LIPA* identified in humans, the pooled prevalence of LAL-D was 1 per 177,452 (95% CI 149,467–210,683) with a carrier frequency of 1 per 421. The prevalence is lowest in those of East Asian, South Asian, and Finnish ancestry.

Conclusion: Using 120 disease variants in *LIPA*, these data can reassure clinicians that LAL-D is an ultra-rare disorder. Given the therapeutic capability of sebelipase alpha, investigation for LAL-D might be included in second-line metabolic screening in NAFLD.

Lay summary: Lysosomal Acid Lipase Deficiency (LAL-D) is a rare genetic condition that can cause severe liver disease, but it is difficult to diagnose and sometimes can look like simple fatty liver. It was not clear how common LAL-D was and whether many cases were being missed. To study this, we searched for all genetic mutations that could cause LAL-D, calculated how common those mutations were, and added them up. This let us estimate that LAL-D affects roughly 1 in 175,000 people. We conclude that LAL-D is a very rare condition, but it is treatable so may be included in a 'second-line' of tests for causes of fatty liver.

© 2018 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

Introduction

Lysosomal acid lipase deficiency (LAL-D) is an autosomal recessive disorder caused by mutations in *LIPA* that manifest as a spectrum of liver disease and dyslipidaemia.[1,2] It is regarded as a rare disorder however recognition of more mild forms of the condition have led to the suggestion that it may represent a significant proportion of patients presenting with non-alcoholic fatty liver disease (NAFLD).[3,4] Emerging data have also reported reduced lysosomal acid lipase (LAL) activity in association with more advanced NAFLD.[5–8] Clarity on the prevalence of LAL-D, and its distribution across ethnicities, is needed to determine whether LAL activity testing should be a routine part of NAFLD clinics.

LIPA encodes the lipase A protein, which functions to catalyse the hydrolysis of cholesteryl esters and triglycerides in the lysosome.[9,10] Loss-of-function mutations are associated with the accumulation of cholesteryl esters in lysosomes in hepatocytes, macrophages, and the spleen. LAL-D describes a spectrum of clinical disease where its most severe form is Wolman disease (WD) (with <1% LIPA activity), presenting with acute liver failure at under 4 months of age, microvesicular steatosis, hepatosplenomegaly, failure to thrive, and adrenal calcification. Without treatment, WD is fatal at under a year, but enzyme replacement therapy significantly improves prognosis.

Cholesteryl ester storage disease (CESD) is used to refer to a more mild form of the disorder, characterised by 1–5% LIPA activity.[11,12] It presents insidiously with elevated aminotransferases, hepatic steatosis, dyslipidaemia, gradual decline in hepatic function, and premature atherosclerosis. These conditions are difficult to diagnose; it is unclear how many children die of undiagnosed WD and there is currently no consensus on routine testing for *LIPA* activity in adults presenting with NAFLD.

The true prevalence of LAL-D (both CESD and WD) is not known. The most common LIPA variant associated with CESD is a splice-junction mutation in exon 8 (rs116928232: c.894G>A, p.S275_Q298del), referred to as E8SJM,[13] which has

Keywords: Lysosomal storage disease; Dyslipidaemia; Hepatic steatosis; Meta-analysis; Epidemiology.
Received 22 July 2018; received in revised form 13 September 2018; accepted 27 September 2018; available online 11 October 2018
* Corresponding author. Address: University of Cambridge, Department of Paediatrics, Cambridge, United Kingdom. Tel.: +44 1223 763480, fax: +44 1223 336996.
E-mail address: jm2032@cam.ac.uk (J.P. Mann).

been used to derive an estimated prevalence as high as 1 per 40,000. This is much higher than clinical observations, leading to the suggestion that LAL-D is under-diagnosed and may be unmasked in patients presenting with NAFLD.[3,14–16]

It is also not yet established whether heterozygous carriers of pathogenic *LIPA* variants are at increased risk of progressive liver disease in combination with other insults to the hepatic parenchyma. Data are conflicting as to whether the heterozygous carrier state influences circulating lipid profile.[17,18]

Based on these data, we hypothesised that the clinical prevalence of LAL-D would be lower than the estimated genetic prevalence and that there may be a minority of patients with NAFLD who have undiagnosed LAL-D. Therefore, in order to address this and estimate the potential impact of underdiagnosis we aimed to calculate the population prevalence of LAL-D by: i) Meta-analysis of previous prevalence estimates; ii) aggregation of population allele frequencies for previously reported pathogenic *LIPA* variants; iii) Identification of previously unreported major functional *LIPA* variants from next generation sequencing data.

Materials and methods

Systematic review of previous prevalence estimates

NCBI Pubmed/MEDLINE and EMBASE were systematically searched for all articles related to LIPA, WD, CESD, or LAL-D (see Supplementary methods for full search term used). The search was finalised on 12/07/2018. Two reviewers (AC & JPM) independently screened abstracts to assess relevance. Any disagreements were resolved through discussion with a third reviewer. Inclusion criteria were: reporting an original estimate of prevalence or incidence or pathogenic allele frequency for LAL-D, CESD, or WD, determined by either genetically or by identification of clinical cases. Reviews, commentaries and editorials reporting non-original data; and *in vitro* or non-human studies were excluded.

Included articles were first grouped into those reporting the prevalence of LAL-D, CESD, or WD, in the general population and those reporting the prevalence in specific patient cohorts. Genetic studies (reporting prevalence derived from LIPA sequencing or pathogenic allele frequency) and epidemiological studies (reporting identification of clinical cases) were assessed separately. It was also recorded whether studies referred to the prevalence of LAL-D, CESD, and/or WD.

Three prevalence estimates from genetic studies of population cohorts were suitable for meta-analysis using an inverse variance model with a double arcsine transformation.[19] Variance was defined as p[1–p]/POP, with POP representing the sample size for genetic studies. Included studies were assessed for risk of bias using the appraisal tool for cross-sectional studies (AXIS).

Meta-analysis of clinical/epidemiological studies was used to derive a pooled prevalence at birth. Meta-analysis of genetic studies gave an overall mutant allele frequency, which was used to calculate a prevalence of LAL-D (at birth) via the Hardy-Weinberg equation.

Identification of previously reported LIPA variants

A systematic search was performed to identify all previously reported pathogenic variants in LIPA. The same search of Pubmed/MEDLINE and EMBASE (as described above) was used. Two independent reviewers screened abstracts to determine suitability for inclusion and disagreements were resolved through discussion with a third reviewer. Inclusion criteria were describing variants in LIPA in humans associated with a loss-of-function and/or LAL-D disease spectrum. Articles not in English and *in vitro* or non-human studies were excluded. Variants reported as disease-causing, pathogenic, probably pathogenic, or harmful were extracted.

In addition, ClinVar (www.ncbi.nlm.nih.gov/clinvar/, accessed 12/07/2018) was screened for LIPA variants with published reports of pathogenicity and these were added to our list. Mutations derived from personal communications and without a linked publication were excluded.

Variants were annotated as being associated with WD, CESD, or LAL-D (if both, or not otherwise stated).

Variants were classified on a case by case basis into 'pathogenic' and 'likely pathogenic' according to ACMG criteria.[20] Variants classified as 'benign', 'likely benign', or of 'uncertain significance' were excluded. (Likely) Pathogenic variants for WD were those found in patients with: hepatic steatosis, raised aminotransferases, reduced LAL activity, and fatal under age 1 (except for cases in a trial of sebelipase alpha); and for CESD as: the presence of hepatic steatosis with dyslipidaemia with reduced LAL activity, with onset over 1 year of age. In addition to ACMG criteria, variants were classified as 'likely pathogenic' if there was no evidence of reduced LAL activity in the affected patients or when there were conflicting reports regarding individual variants.

Identification of previously unreported major functional variants in LIPA

Ensembl (http://www.ensembl.org/, accessed 12/07/2018) for LIPA transcript ENST00000336233.9 was used to search for all *LIPA* variants identified in humans via the Exome Aggregation Consortium (ExAC),[21] Exome Sequencing Project (ESP), or 1000 Genomes Project (1000G).[22] Those with predicted major consequences on protein structure and function were extracted, including: frameshift, premature stop codon, initiator codon, splice donor, splice acceptor, and start loss variants. The Genome Aggregation Database (gnomAD, gnomad.broadinstitute.org/, accessed 12/07/2018) was also searched and all duplicates were removed. Coverage was assessed in gnomAD variants using specific site quality metrics. Non-PASS filter variants were excluded. This list of major functional variants was compared with previously reported *LIPA* variants found in patients to identify unreported variants predicted to affect LAL activity. Finally, variants underwent frequency filtering.[23]

Annotation of variants with allele frequencies and functional predictions

Coding sequence nucleotide changes for each identified variant were converted to HGVS format (hg38) using Mutalyzer (https://mutalyzer.nl/). The Ensembl Variant Effect Predictor (http://www.ensembl.org/Tools/VEP), was used to annotate variants with SIFT/PolyPhen-2 *in silico* predictions of pathogenicity and population allele frequencies from the 1000G, ESP, ExAC, and gnomAD data sets. Where variants had been identified in more than one dataset the larger cohort was used.

LIPA structural information and modelling

Data on the domain structure and active site positions for LIPA were extracted from UniProtKB (https://uniprot.org) for

P38571.[24] Modelling was extracted from ExPasy SWISS-MODEL (https://swissmodel.expasy.org) for P38571, based upon the gastric triacylglycerol lipase.[10,25,26] Detail on the exon structure of LIPA was also retrieved from neXtProt (https://nextprot.org).[27]

Prevalence estimation for CESD, WD, and LAL-D

We performed a prevalence estimation without allele frequency filtering to identify the maximum genetic prevalence of LAL-D. Total allele frequencies were used to calculate estimations of global population prevalence of homozygous and heterozygous LAL-D genotypes using the Hardy-Weinberg equation. This was performed for: previously reported pathogenic variants (separately for CESD and WD), previously reported pathogenic plus possibly-pathogenic variants, and then previously reported plus unreported major functional variants.

In addition, we estimated the prevalence across 8 different ethnicities. We calculated 95% CIs for the total allele frequencies using the Wilson score method across a binomial distribution where more than 5 alleles had been identified.

Software

Meta-analysis was performed using the Meta-XL add-in for Microsoft Excel (www.epigear.com) and GraphPad Prism version 7.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.

For further details regarding the materials and methods used, please refer to the CTAT table and supplementary information.

Results

Meta-analysis of previous prevalence estimates

A total of 1,926 abstracts were identified by systematic search, of which 194 studies were included (Fig. S1). Nine reported the prevalence of WD or CESD in the general population (Table 1), of which 6/9 used clinically identified cases and 3/9 used genetic sequencing of *LIPA*. Studies were generally of high quality, with a low risk of bias (Table 1, Fig. S2).

Epidemiological studies were retrospective in nature and predominantly identified patients with WD from national or regional referral centres for inherited metabolic disease. Prevalences ranged from 1 per 165,530 to 1 per 909,091 (Fig. S3). Meta-analysis of prevalence estimates using random effects model gave a prevalence of 1 per 451,116 (95% CI 282,022–840,092) with a carrier frequency of 1 per 336 (95% CI 266–458).

All genetic studies used sequencing for c.894G>A in *LIPA* to generate a pathogenic allele frequency. Disease prevalence estimates were then calculated using the proportion of pathogenic *LIPA* variants comprised by c.894G>A (estimated 50–60%, where Scott *et al.* calculated 95% CI 51–69%).[4,34]

Meta-analysis of c.894G>A genetic studies using a random effects model gave a pooled allele frequency of 0.0015 (95% CI 0.001–0.002, Fig. 1). Assuming that c.894G>A accounts for 60% (95% CI 51%–69%)[34] of all CESD mutations, this corresponds to a prevalence for CESD of 1 per 160,000 (95% CI 1 per 65,025–761,652) and a carrier frequency of 1 per 400 (95% CI 1 per 255–873).

We also identified 8 studies reporting the prevalence of CESD in specific patient populations: 5 in cohorts of patients with dyslipidaemia (n = 4,193 patients) and 3 in cohorts with liver disease (n = 430, Table S1). Of these studies, 2/8 (25%) used *LIPA* sequencing and 6/8 assessed LAL activity. These data were not suitable for meta-analysis because of significant patient and methodological heterogeneity.

In light of these wide confidence intervals and assumption of pathogenic contribution we then proceeded to attempt to obtain a more accurate estimate of global LAL-D prevalence, using publicly available sequencing data.

Identification and analysis of previously reported LIPA variants

After removal of duplicates and benign variants (Fig. S4), 98 disease-causing variants in *LIPA* were identified from publications (Table S2). Of these variants, 23/98 (23%) had been associated with WD only, 55/98 (56%) with CESD only, and the remainder with either both or not stated. Additionally, 36/98 (37%) of variants were classified as pathogenic and the remaining 62/98 (63%) likely pathogenic. The exonic distribution along *LIPA* of protein-coding variants and relationship to the active site is shown (Fig. 2). The most common mutation consequence was missense, accounting for 43% (42/98) of all disease variants (Fig. 3A). However, this was not evenly distributed across the LAL-D spectrum: WD was associated with a higher proportion of frameshift mutations and premature stop variants (Fig. 3C), compared to CESD (Fig. 3B).

Population databases enabled identification of 32/98 (33%) of disease variants (Table 2); all variants found in ESP, ExAC, or 1000G were also found in gnomAD and therefore only gnomAD allele frequencies are reported. The most frequently reported allele (E8SJM, c.894G>A) had the highest allele frequency of 0.0089 overall in the gnomAD data set and was notably not found in any individuals of East Asian ancestry. The 66 variants not identified in gnomAD were in regions with >60× mean depth coverage per base (Fig. S5).

Identification and analysis of previously unreported LIPA variants

In order to account for genetic variants that have not yet been identified in patients with LAL-D, we examined all variants that been identified in humans that were predicted to have a significant functional impact on LAL activity (Fig. S6). We identified 22 previously unreported major functional LIPA variants (Table 3, Fig. 2, Table S3). The most common functional classification was splice donor or acceptor in 8/22 (36%).

c.894G>A accounted for 49% of the pooled allele frequencies for all 98 previously reported disease-causing variants and 38% of the pooled allele frequencies when including all 120 variants. The top 4 most frequent LIPA variants (c.894G>A, c.891C>T, c.676-23 T>C, and c.966+3A>T) accounted for 81% of the total allele frequencies for previously reported variants and 61%, when including unreported variants.

Estimation of population prevalence of WD, CESD, and LAL-D

Pooling of the allele frequencies for all previously reported disease variants gave a global mutant allele frequency of 0.0018 (95% CI 0.0017–0.002) for LAL-D (Table 4, Fig. 4), which is equivalent to a prevalence at birth of LAL-D of 1 per 307,482 (95% CI 257,672–366,865). When combined with unreported major functional variants, the pooled allele frequencies increased to 0.0024 (95% CI 0.0022–0.0026), giving a prevalence at birth of LAL-D of 1 per 177,452 (95% CI 149,467–210,683) and a heterozygous carrier rate of 1 per 421 (95% CI 387–459).

Table 1. Epidemiological and genetic studies estimating the prevalence of LAL-D.

Study	Cohort assessed	Diagnostic method	Population (n)	Cases (n)	Reported prevalence	Comments	Quality (AXIS tool)
Epidemiological/clinical studies							
Giugliani *et al.*, 2017[28] [Brazil]	General population	Clinical diagnosis lysosomal acid lipase deficiency	41,719,041 births (part of study period)	10 patients	0.011 per 100,000	Birth data only available for part of the study period. May be an under–estimation.	High quality, low RoB
Moammar *et al.*, 2010[29] [Saudi Arabia]	General population	Clinical diagnosis Wolman disease	165,530	1 patient	Not stated	Only 1 case identified. No calculation of prevalence.	High quality, low RoB
Poupětová *et al.*, 2010[30] [Czech Republic]	General population	Clinical diagnosis Wolman disease or CESD	5,480,885 births	15 patients	1 per 365,392 hom affected; 1 per 302 het carrier	Majority of patients had juvenile CESD. No confidence intervals calculated for CESD.	High quality, low RoB
Dionisi–Vici *et al.*, 2002[31] [Italy]	General population	Clinical diagnosis Wolman disease / CESD	7,13,959 births	9 patients	Not stated	No calculation of prevalence.	High quality, low RoB
Applegarth *et al.*, 2000[32] [Canada]	General population	Clinical diagnosis Wolman disease	1,035,816 births	6 patients	Not stated	No calculation of prevalence. Stringent diagnostic criteria used.	High quality, low RoB
Meikle *et al.*, 1999[33] [Australia]	General population	Clinical diagnosis Wolman disease	4,222,323 births	8 patients	1 per 528,000 hom affected; 1 per 363 het carrier	WD only, not CESD/LAL–D. Australian population ancestry.	High quality, very low RoB
Genetic studies							
Scott *et al.*, 2013[30] [United States]	General population	*LIPA* sequencing for c.894G>A	33,156 alleles	53 alleles	0.8 per 100,000	Estimated c.894G>A accounted for 60% of LAL–D. Multiple genetic ancestries included.	High quality, low RoB
Stitziel *et al.*, 2013[18] [United States]	i. General population with lipid profile data ii. General population with CAD data	*LIPA* sequencing for c.894G>A	i. 26,388 alleles ii. 54,944 alleles	i. 42 het ii. 60 het	Not stated	No association with lipid profile or CAD in heterozygous state. European ancestry population.	High quality, low RoB
Muntoni *et al.*, 2007[4] [Germany]	General population	*LIPA* sequencing for c.894G>A	4,046 alleles	10 alleles	1 per 40,000 hom affected; 1 in 200 het carrier	Estimated c.894G>A accounted for 50% of LAL–D. European ancestry population.	High quality, low RoB

AXIS, appraisal tool for cross-sectional studies; CAD, coronary artery disease; CESD, cholesteryl ester storage disease; Het, heterozygote carrier of *LIPA* mutation; hom, homozygous patient affected with LAL-D; LAL-D, lysosomal acid lipase deficiency; RoB, risk of bias; WD, Wolman disease.

JOURNAL OF HEPATOLOGY

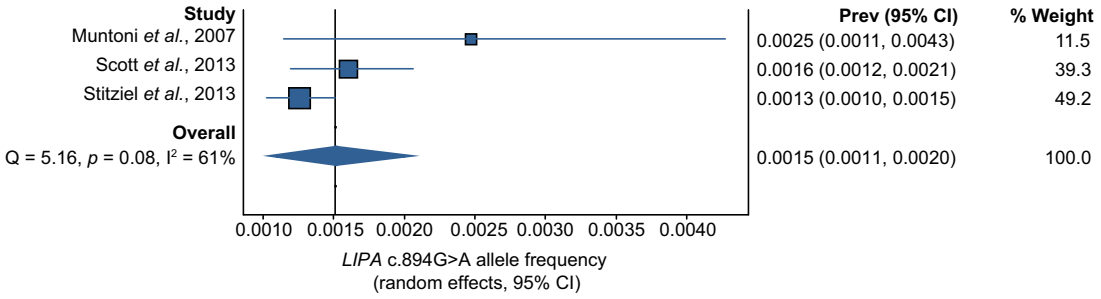


Fig. 1. Random effects meta-analysis of genetic studies estimating the prevalence of LAL-D using c.894G>A allele frequency. AF, allele frequency; LAL-D, lysosomal acid lipase deficiency.

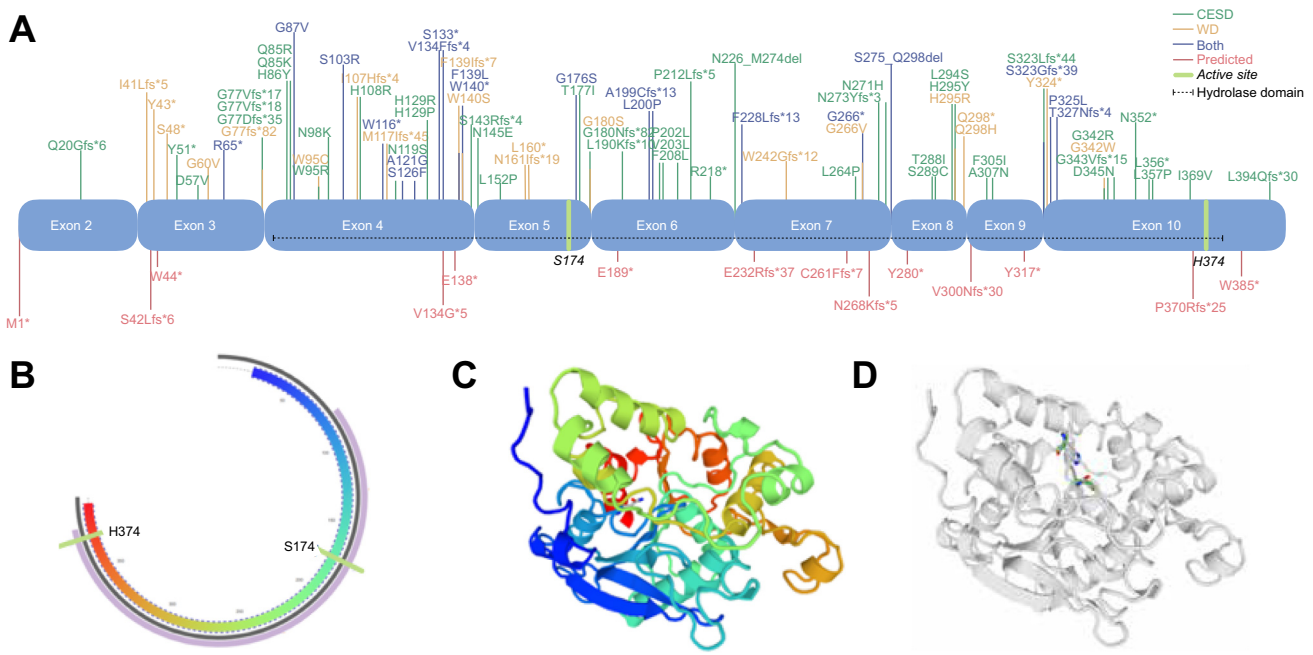


Fig. 2. The modelled structure of LIPA and its active site, with the exonic distribution of disease-causing mutations. (A) The exonic distribution of previously reported protein-coding disease mutations and unreported major functional variants in *LIPA*. (B,C) A modelled structure of lysosomal acid lipase with its (D) active site highlighted. (This figure appears in colour on the web.)

Analysis by ethnicity demonstrated a higher prevalence in those of non-Finnish European ancestry: 1 per 103,286 (95% CI 83,142–128, 321), whilst the lowest prevalence of LAL-D was in those of East Asian, Finnish, South Asian, and Ashkenazi Jewish ancestry (Fig. 4, Table S4).

Discussion

LAL-D (MIM #278000) is recognised as a treatable[35] monogenic condition that may masquerade as NAFLD or present with cryptogenic cirrhosis. Recent data has led to suggestions that it may be under-diagnosed. In this study we sought to estimate the prevalence of LAL-D using accurate population sequencing data and deepen our understanding of the genetic variation of *LIPA*. Using next generation sequencing data, we found LAL-D to affect 1 per 177,452 across the global population, which is significantly lower than previous estimates that have been up to 1 per 40,000.

The main strength of this study is the use of 120 *LIPA* mutants from a population of >150,000 for our prevalence estimate, whereas previous studies have focussed on c.894G>A (E8SJM).[4,34] In doing so, we have also produced a curated list of disease variants in *LIPA* that may prove useful for further *in vitro* studies or targeted sequencing in patients. The majority of these are extremely rare, with 66 of 98 not being identified in any population sequencing cohort, giving an allele frequency of less than 6.7×10^{-6}.

Using existing data, we have produced prevalence estimates for CESD and WD separately from LAL-D, however there is conflicting data over the consistency of genotype-phenotype correlations.[36–38] These estimates are limited by the difficulty in establishing and reporting clinical diagnoses. The key differentiating factor is residual activity in LAL, which reliably separates the 2 conditions.[12] Although, data in this study supports the notion that WD is more associated with major functional variants (stop gained or splice-site variants). Our estimate of WD prevalence from genetic variants (1 per 393,630) is slightly higher than the prevalence derived from meta-analysis of previous studies (1 per 451,116, (95% CI 282,022–840,092)). Given the challenge in diagnosing LAL-D, the disparity between these clinical epidemiological data and the genetic prevalence are most likely to represent under-diagnosis. The uncertainty

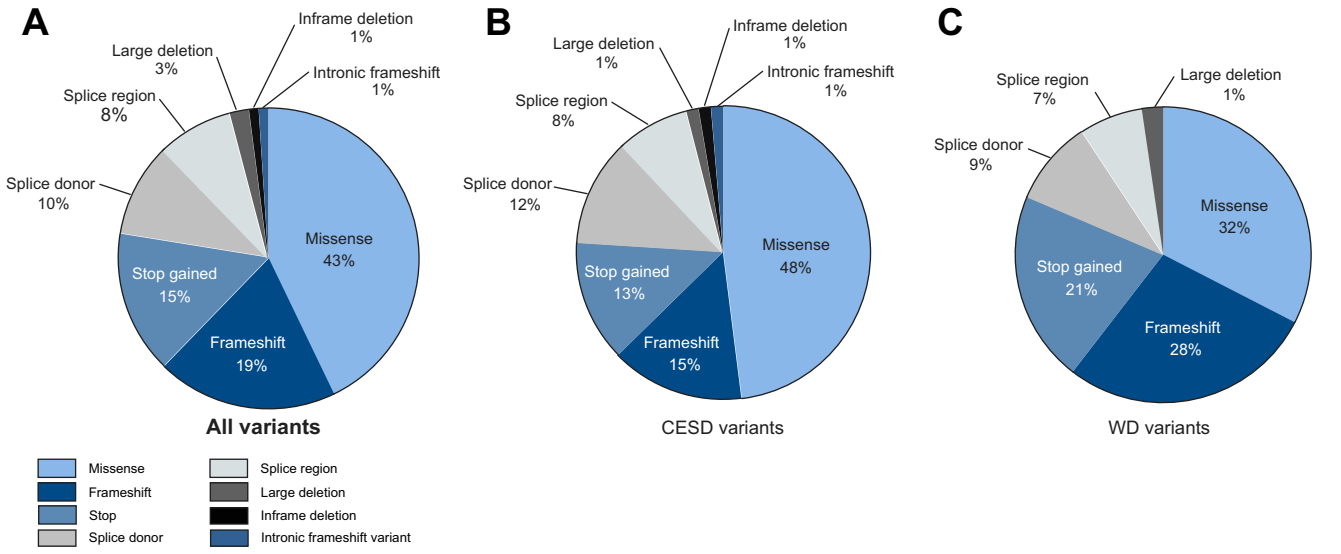


Fig. 3. Functional classification of disease variants in LIPA. (A) All variants; (B) those associated with only CESD; (C) those associated with only with WD. CESD, cholesteryl ester storage disease; Wolman disease.

Table 2. Previously identified LIPA disease variants present in gnomAD with global allele frequencies and combined allele frequencies for each clinical condition.

Genomic position	Nucleotide change	Amino acid change	Consequence	rsID	Clinical presentation	gnomAD AF
10:89214921–89214921	c.1107A>C	p.Ile369Val	Missense variant	rs768436255	CESD	0.000004
10:89214958–89214958	c.1070T>C	p.Leu357Pro	Missense variant	rs772684869	CESD	0.000040
10:89215000–89215000	c.1028delG	p.Gly343Valfs*15	Frameshift variant	rs750301834	CESD	0.000007
10:89215004–89215004	c.1024G>A	p.Gly342Arg	Missense variant	rs776472526	CESD	0.000011
10:89215054–89215054	c.974C>T	p.Pro325Leu	Missense variant		LAL–D	0.000008
10:89215936–89215936	c.966+3A>T	p.(=)	Splice donor variant	rs201242614	CESD	0.000141
10:89215984–89215984	c.920C>A	p.Ala307Asp	Missense variant	rs754964952	CESD	0.000020
10:89222511–89222511	c.894G>A	p.Ser275_Gln298del	Splice region inframe deletion	rs116928232	WD	0.000892
10:89222511–89222511	c.894G>C	p.Gln298His	Splice region missense variant	rs116928232	LAL–D	0.000012
10:89222514–89222514	c.891C>T	p.(=)	Splice region synonymous variant	rs145066614	CESD	0.000271
10:89222524–89222524	c.881T>C	p.Leu294Ser	Missense variant	rs756310979	CESD	0.000004
10:89222542–89222542	c.863C>T	p.Thr288Ile	Missense variant		CESD	0.000004
10:89223683–89223683	c.822 + 1G>A	p.?	Splice donor variant		CESD	0.000032
10:89223710–89223710	c.796G>T	p.Gly266*	Stop gained	rs267607218	LAL–D	0.000007
10:89223822–89223822	c.684delT	p.Phe228Leufs*13	Frameshift variant	rs770074196	LAL–D	0.000012
10:89225093–89225093	c.676–23T>C	p.?	Splice region variant	rs140488274	CESD	0.000148
10:89225115–89225115	c.652C>T	p.Arg218*	Stop gained	rs771330022	CESD	0.000004
10:89225145–89225145	c.622T>C	p.Phe208Leu	Missense variant	rs148713974	CESD	0.000022
10:89225168–89225168	c.599T>C	p.Leu200Pro	Missense variant	rs121965086	LAL–D	0.000012
10:89225173–89225174	c.594dupT	p.Ala199Cysfs*13	Frameshift variant	rs780495201	LAL–D	0.000014
10:89226895–89226895	c.538+6T>C	p.?	Splice donor variant	rs772236690	CESD	0.000016
10:89226978–89226978	c.455T>C	p.Leu152Pro	Missense variant	rs748267444	CESD	0.000004
10:89228230–89228230	c.398delC	p.Ser133*	Frameshift variant	rs756016704	WD	0.000020
10:89228242–89228242	c.386A>G	p.His129Arg	Missense variant		CESD	0.000016
10:89228277–89228278	c.350_351insCC	p.Met117Ilefs*45	Frameshift variant	rs753796180	WD	0.000007
10:89228319–89228319	c.309C>A	p.Ser103Arg	Missense variant	rs766364179	LAL–D	0.000004
10:89228328–89228328	c.229+3A>C	p.?	Splice region variant	rs750405436	CESD	0.000008
10:89228334–89228334	c.294C>G	p.Asn98Lys	Missense variant	rs767688436	CESD	0.000012
10:89228368–89228368	c.260G>T	p.Gly87Val	Missense variant	rs587778878	LAL–D	0.000008
10:89228372–89228372	c.256C>T	p.His86Tyr	Missense variant	rs749180806	CESD	0.000011
10:89228375–89228375	c.253C>A	p.Gln85Lys	Missense variant	rs797045094	CESD	0.000004
10:89245712–89245712	c.193C>T	p.Arg65*	Stop gained	rs779712562	LAL–D	0.000025
				LAL–D (WD + CESD) Combined AF		0.00180
				95% CI		(0.00165–0.00197)
				WD Combined AF		0.00102
				95% CI		(0.00091–0.0012)
				CESD Combined AF		0.00176
				95% CI		(0.0016–0.0019)

AF, allele frequency; CESD, cholesteryl ester storage disease; LAL-D, lysosomal acid lipase deficiency; WD, Wolman disease.

Table 3. Previously unreported variants in LIPA predicted to have major functional effects found in gnomAD.

Genomic position	Nucleotide change	Amino acid change	Consequence	rsID	gnomAD AF
10:89214874–89214874	c.1154G>A	p.Trp385*	Stop gained		0.000004
10:89214919–89214919	c.1109delC	p.Pro370Argfs*25	Frameshift variant	rs762074434	0.000004
10:89215953–89215953	c.951T>A	p.Tyr317*	Stop gained	rs760413481	0.000004
10:89216003–89216006	c.898_901delGTTA	p.Val300Asnfs*30	Frameshift variant	rs754498326	0.000004
10:89216010–89216010	c.895–1G>A	p.?	Splice acceptor		0.000032
10:89222565–89222565	c.840T>G	p.Tyr280*	Stop gained	rs139691556	0.000077
10:89223702–89223702	c.804delT	p.Asn268Lysfs*5	Frameshift variant		0.000004
10:89223725–89223726	c.780_781delCT	p.Cys261Phefs*7	Frameshift variant		0.000004
10:89223813–89223813	c.693dupA	p.Glu232Argfs*37	Frameshift variant		0.000004
10:89223832–89223832	c.676–2A>T	p.?	Splice acceptor	rs747508159	0.000004
10:89225202–89225202	c.565G>T	p.Glu189*	Stop gained	rs749421449	0.000004
10:89228216–89228216	c.412G>T	p.Glu138*	Stop gained		0.000004
10:89228228–89228228	c.400dupG	p.Val134Glyfs*5	Frameshift variant	rs760472104	0.000004
10:89245755–89245756	c.149_150delTG	p.?	Splice donor		0.000032
10:89245773–89245773	c.132G>A	p.Trp44*	Stop gained	rs181646633	0.000008
10:89245779–89245780	c.125_126delCT	p.Ser42Leufs*6	Frameshift variant	rs759603689	0.000008
10:89247537–89247537	c.111+1G>A	p.?	Splice donor	rs762960877	0.000008
10:89247597–89247597	c.52T>C	p.?	Splice donor		0.000097
10:89247598–89247598	c.51G>A	p.?	Splice donor		0.000032
10:89247648–89247648	c.1A>C	p.Met1?	Start lost	rs767039444	0.000004
10:89251735–89251735	c.–2+2T>A	p.?	Splice donor		0.000065
10:89251736–89251736	c.–2+1G>A	p.?	Splice donor		0.000162
				Combined AF	0.00057
				95% CI	(0.0004–0.0008)

AF, allele frequency.

Table 4. Prevalence of LAL-D, WD, and CESD estimated from pooled allele frequencies of previously reported variants and combined with unreported variants.

	WD		CESD		LAL–D (WD + CESD)	
	Previously reported variants	Previously reported + unreported variants	Previously reported variants	Previously reported + unreported variants	Previously reported variants	Previously reported + unreported variants
Combined AF	0.00102	0.001593882	0.0018	0.0023	0.001803393	0.002374
95% CI	(0.0009–0.0012)	(0.0014–0.0018)	(0.0016–0.0019)	(0.0021–0.0025)	(0.0017–0.002)	(0.0022–0.0026)
Het carrier (1 per ...)	977	627	556	435	555	421
95% CI	(869–1099)	(561–701)	(519–620)	(393–467)	(508–606)	(387–459)
Hom affected (1 per ...)	954,812	393,630	321,489	183,543	307,482	177,452
95% CI	(754,660–1,208,088)	(314,937–492,010)	(269,021–384,354)	(154,440–218,135)	(257,672–366,865)	(149,467–210,683)

CESD, cholesteryl ester storage disease; Het, heterozygote carrier of *LIPA* mutation; hom, homozygous patient affected with LAL-D; LAL-D, lysosomal acid lipase deficiency; WD, Wolman disease.

associated with these calculations are reflected in their relatively wide confidence intervals.

It is not possible to state whether CESD is under-diagnosed as there are no clinical studies assessing the prevalence of CESD in the general population. However, targeted testing of LAL activity in >1,800 patients with dyslipidaemia and raised aminotransferases found no affected patients.[39]

LIPA is a highly conserved gene and its hydrolase domain is similar to other lipases.[10] Protein-coding disease variants were found to cluster particularly in exons 4 and 10, in addition to splice junctions. The active site codons were not found to be affected directly by any point mutations however previous modelling has demonstrated how adjacent variants disrupt LAL activity.[40,41]

Though this methodology has been used to estimate the prevalence of other monogenic disorders,[42] it is based upon the assumptions of the Hardy-Weinberg principle and there is no *in vitro* data to support many of the variants. These limitations would be likely to result in an over-estimation of the prevalence of LAL-D in these results. Our overall calculated prevalence is also limited by an inability to differentiate between WD and CESD in newly discovered variants.

LAL-D is associated with a high proportion of (intronic) splice-site variants. gnomAD has deep coverage at splice sites (Fig. S6), however this highlights the need for deep exome or whole genome sequencing to diagnose affected patients. This may be a limitation to incorporating genetic diagnosis into clinical practice.

This updated prevalence estimate has implications for the diagnosis of LAL-D and exclusion of secondary metabolic causes of NAFLD. EASL-EASD-EASO guidelines recommended that Wilson disease and alpha-1-antitrypsin deficiency should be tested

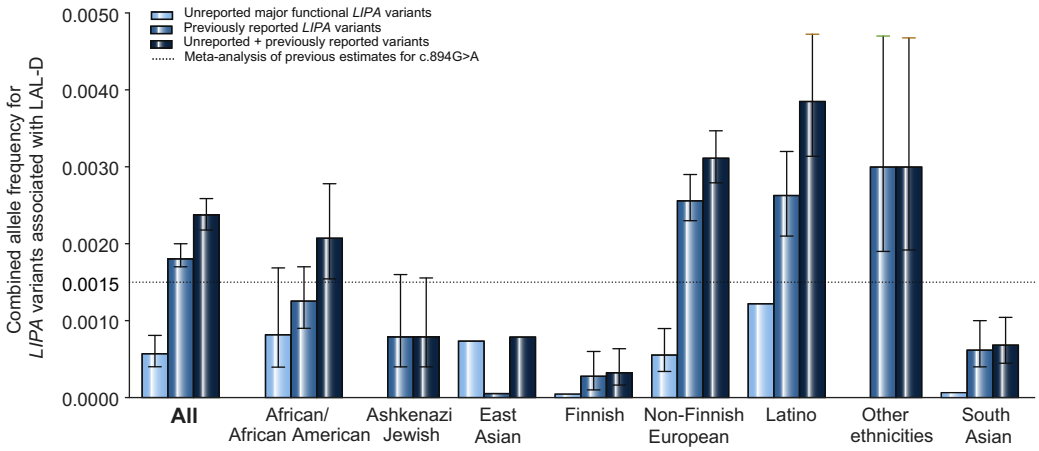


Fig. 4. The pooled allele frequency for disease variants in LIPA from unreported major functional variants, previously reported disease variants, and combined for 8 genetic ancestries. These are compared to the meta-analysed genetic prevalence of c.894G>A (dotted line).

at an 'extended investigation' stage.[43] These data suggest that LAL-D is significantly less common than these 2 conditions and therefore could be included in a 'second-line' of metabolic tests along with other rare mimics of NAFLD (Fig. S7, Table S5). A consensus is needed on whether to include LAL-D screening in patients with atypical features of NAFLD – for example, those who are especially young, with a significant family history, low body mass index, or without insulin resistance.

It remains to be established whether heterozygotes have a clinically manifest phenotype and whether 1–5% LAL activity would accelerate concomitant NAFLD, viral hepatitis, or any other hepatic parenchymal disease. The currently presented data would suggest that the heterozygote carrier rate is much lower than that for other monogenic liver diseases that may influence NAFLD progression, such as HFE hereditary haemochromatosis (1 in 8–10)[44] or alpha-1-antitrypsin deficiency (1 in 15–20).[45]

In conclusion, through comprehensive analysis of genetic variation in *LIPA* we have expanded our recognition of disease-causing mutants to 120 variants. LAL-D is an ultra-rare disease, even in its late-onset form as CESD, and these data can help reassure clinicians that LAL-D is unlikely to represent a significant proportion of patients presenting with NAFLD. A consensus is needed for when testing LAL activity should be performed in patients with dyslipidaemia or hepatic steatosis.

Financial support
The authors received no financial support to produce this manuscript.

Conflict of interest
The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions
AC: data collection, analysis, manuscript drafting, approval of the final manuscript; JG & SB: data collection, analysis, approval of the final manuscript; JPM: concept, data collection, analysis, manuscript drafting, approval of the final manuscript.

Acknowledgements
The authors thank Dr. Matthew J. Armstrong (Birmingham, UK) for his advice throughout this work.

Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2018.09.028.

References
Author names in bold designate shared co-first authorship

[1] Wolman M, Sterk VV, Gatt S, Frenkel M. Primary familial xanthomatosis with involvement and calcification of the adrenals. Report of two more cases in siblings of a previously described infant. Pediatrics 1961;28:742–757.
[2] Bernstein DL, Hülkova H, Bialer MG, Desnick RJ. Cholesteryl ester storage disease: Review of the findings in 135 reported patients with an underdiagnosed disease. J Hepatol 2013;58:1230–1243.
[3] Himes RW, Barlow SE, Bove K, Quintanilla NM, Sheridan R, Kohli R. Lysosomal acid lipase deficiency unmasked in two children with nonalcoholic fatty liver disease. Pediatrics 2016;138, e20160214–e20160214.
[4] Muntoni S, Wiebusch H, Jansen-Rust M, Rust S, Seedorf U, Schulte H, et al. Prevalence of cholesteryl ester storage disease. Arterioscler Thromb Vasc Biol 2007;27:1866–1868.
[5] Tovoli F, Napoli L, Negrini G, D'Addato S, Tozzi G, D'Amico J, et al. A relative deficiency of lysosomal acid lypase activity characterizes non-alcoholic fatty liver disease. Int J Mol Sci 2017;18:1134.
[6] Vespasiani-Gentilucci U, Valentini F, Carotti S, Vorini F, Zingariello M, Francesconi M, et al. The hepatic expression of lysosomal acid lipase (LAL) is reduced in NAFLD patients, and associated with features of genetically determined LAL deficiency and NAFLD activity score. Dig Liver Dis 2017;49:e1–e2.
[7] **Shteyer E**, **Villenchik R**, Mahamid M, Nator N, Safadi R. Low serum lysosomal acid lipase activity correlates with advanced liver disease. Int J Mol Sci 2016;17:312.
[8] Baratta F, Pastori D, Del Ben M, Polimeni L, Labbadia G, Di Santo S, et al. Reduced lysosomal acid lipase activity in adult patients with non-alcoholic fatty liver disease. EBioMedicine 2015;2:750–754.
[9] Anderson RA, Byrum RS, Coates PM, Sando GN. Mutations at the lysosomal acid cholesteryl ester hydrolase gene locus in Wolman disease. Proc Natl Acad Sci U S A 1994;91:2718–2722.
[10] **Roussel A**, **Canaan S**, Egloff M-P, Riviere M, Dupuis L, Verger R, et al. Crystal structure of human gastric lipase and model of lysosomal acid

lipase, two lipolytic enzymes of medical interest. J Biol Chem 1999;274:16995–17002.
[11] Pagani F, Pariyarath R, Garcia R, Stuani C, Burlina AB, Ruotolo G, et al. New lysosomal acid lipase gene mutants explain the phenotype of Wolman disease and cholesteryl ester storage disease. J Lipid Res 1998;39:1382–1388.
[12] Aslanidis C, Ries S, Fehringer P, Büchler C, Klima H, Schmitz G. Genetic and biochemical evidence that CESD and Wolman disease are distinguished by residual lysosomal acid lipase activity. Genomics 1996;33:85–93.
[13] Klima H, Ullrich K, Aslanidis C, Fehringer P, Lackner KJ, Schmitz G. A splice junction mutation causes deletion of a 72-base exon from the mRNA for lysosomal acid lipase in a patient with cholesteryl ester storage disease. J Clin Invest 1993;92:2713–2718.
[14] Bernstein DL, Hulkova H, Bialer MG, Desnick RJ. Cholesteryl ester storage disease: review of the findings in 135 reported patients with an underdiagnosed disease. J Hepatol 2013;58:1230–1243.
[15] Pant M, Oshima K. Cholesteryl ester storage disease: an underdiagnosed cause of cirrhosis in adults. Ann Diagn Pathol 2017;31:66–70.
[16] Elleder M, Chlumská A, Hyánek J, Poupětová H, Ledvinová J, Maas S, et al. Subclinical course of cholesteryl ester storage disease in an adult with hypercholesterolemia, accelerated atherosclerosis, and liver cancer. J Hepatol 2000;32:528–534.
[17] Muntoni S, Wiebusch H, Jansen-Rust M, Rust S, Schulte H, Berger K, et al. Heterozygosity for lysosomal acid lipase E8SJM mutation and serum lipid concentrations. Nutr Metab Cardiovasc Dis 2013;23:732–736.
[18] Stitziel NO, Fouchier SW, Sjouke B, Peloso GM, Moscoso AM, Auer PL, et al. Exome sequencing and directed clinical phenotyping diagnose cholesterol ester storage disease presenting as autosomal recessive hypercholesterolemia. Arterioscler Thromb Vasc Biol 2013;33:2909–2914.
[19] Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. J Epidemiol Community Health 2013;67:974–978.
[20] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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;17:405–424.
[21] **Lek M**, **Karczewski KJ**, **Minikel EV**, **Samocha KE**, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–291.
[22] The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526:68–74.
[23] Whiffin N, Minikel E, Walsh R, O'Donnell-Luria AH, Karczewski K, Ing AY, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. Genet Med 2017;19:1151–1158.
[24] The UniProt Consortium. UniProt: The universal protein knowledgebase. Nucleic Acids Res 2017;45:D158–D169.
[25] **Waterhouse A**, **Bertoni M**, **Bienert S**, **Studer G**, **Tauriello G**, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46:W296–W303.
[26] Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci Rep 2017;7:1–15.
[27] Gaudet P, Michel PA, Zahn-Zabal M, Britan A, Cusin I, Domagalski M, et al. The neXtProt knowledgebase on human proteins: 2017 update. Nucleic Acids Res 2017;45:D177–D182.
[28] Giugliani R, Federhen A, Michelin-Tirelli K, Riegel M, Burin M. Relative frequency and estimated minimal frequency of lysosomal storage diseases in Brazil: Report from a reference laboratory. Genet Mol Biol 2017;40:31–39.
[29] Moammar H, Cheriyan G, Mathew R, Al-Sannaa N. Incidence and patterns of inborn errors of metabolism in the eastern province of Saudi Arabia, 1983–2008. Ann Saudi Med 2010;30:271–277.
[30] Poupětová H, Ledvinová J, Berná L, Dvořáková L, Kožich V, Elleder M. The birth prevalence of lysosomal storage disorders in the Czech Republic: Comparison with data in different populations. J Inherit Metab Dis 2010;33:387–396.
[31] Dionisi-Vici C, Rizzo C, Burlina AB, Caruso U, Sabetta G, Uziel G, et al. Inborn errors of metabolism in the Italian pediatric population: A national retrospective survey. J Pediatr 2002;140:321–327.
[32] Applegarth DA, Toone JR, Lowry RB. Incidence of inborn errors of metabolism in British Columbia, 1969-1996. Pediatrics 2000;105 e10.
[33] Meikle P, Hopwood J, Clague A, Carey W. Prevalence of lysosomal storage disorders. JAMA 1999;281:249–254.
[34] **Scott SA**, **Liu B**, **Nazarenko I**, Martis S, Kozlitina J, Yang Y, et al. Frequency of the cholesteryl ester storage disease common LIPA E8SJM mutation (c.894G>A) in various racial and ethnic groups. Hepatology 2013;58:958–965.
[35] Burton BK, Balwani M, Feillet F, Barić I, Burrow TA, Camarena Grande C, et al. A phase 3 trial of sebelipase alfa in lysosomal acid lipase deficiency. N Engl J Med 2015;373:1010–1020.
[36] Anderson RA, Bryson GM, Parks JS. Lysosomal acid lipase mutations that determine phenotype in Wolman and cholesterol ester storage disease. Mol Genet Metab 1999;68:333–345.
[37] Lohse P, Maas S, Sewell AC, van Diggelen OP, Seidel D. Molecular defects underlying Wolman disease appear to be more heterogeneous than those resulting in cholesteryl ester storage disease. J Lipid Res 1999;40:221–228.
[38] Redonnet-Vernhet I, Chatelut M, Basile JP, Salvayre R, Levade T. Cholesteryl ester storage disease: relationship between molecular defects and in situ activity of lysosomal acid lipase. Biochem Mol Med 1997;62:42–49.
[39] Reynolds TM, Mewies C, Hamilton J, Wierzbicki AS. Identification of rare diseases by screening a population selected on the basis of routine pathology results—the PATHFINDER project: lysosomal acid lipase/cholesteryl ester storage disease substudy. J Clin Pathol 2018;71:608–613.
[40] Zschenker O, Jung N, Rethmeier J, Trautwein S, Hertel S, Zeigler M, et al. Characterization of lysosomal acid lipase mutations in the signal peptide and mature polypeptide region causing Wolman disease. J Lipid Res 2001;42:1033–1040.
[41] Rajamohan F, Reyes AR, Ruangsiriluk W, Hoth LR, Han S, Caspers N, et al. Expression and functional characterization of human lysosomal acid lipase gene (LIPA) mutation responsible for cholesteryl ester storage disease (CESD) phenotype. Protein Expr Purif 2015;110:22–29.
[42] Wallace DF, Subramaniam VN. The global prevalence of HFE and non-HFE hemochromatosis estimated from analysis of next-generation sequencing data. Genet Med 2016;18:618–626.
[43] European Association for the Study of the Liver (EASL)European Association for the Study of Diabetes (EASD)European Association for the Study of Obesity (EASO). EASL – EASD – EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016;64:1388–1402.
[44] **Ye Q**, **Qian BX**, Yin WL, Wang FM, Han T. Association between the HFE C282Y, H63D polymorphisms and the risks of non-alcoholic fatty liver disease, liver cirrhosis and hepatocellular carcinoma: An updated systematic review and meta-analysis of 5,758 cases and 14,741 controls. PLoS ONE 2016;11 e0163423.
[45] **Strnad P**, **Buch S**, **Hamesch K**, Fischer J, Rosendahl J, Schmelz R, et al. Heterozygous carriage of the alpha1-antitrypsin Pi*Z variant increases the risk to develop liver cirrhosis. Gut 2018. https://doi.org/10.1136/gutjnl-2018-316228.