Molecular Neurobiology (2019) 56:6426–6435
https://doi.org/10.1007/s12035-019-1528-z

Niemann-Pick Disease Type C: Mutation Spectrum and Novel Sequence Variations in the Human *NPC1* Gene

Márcia Polese-Bonatto[1,2] · Hugo Bock[1,2] · Ana Carolina S. Farias[1] · Rafaella Mergener[1] · Maria Cristina Matte[1] · Mirela S. Gil[1] · Felipe Nepomuceno[3] · Fernanda T. S. Souza[4,5] · Rejane Gus[4] · Roberto Giugliani[1,4,5,6] · Maria Luiza Saraiva-Pereira[1,2,4,6,7]

Received: 24 September 2018 / Accepted: 15 February 2019 / Published online: 28 February 2019

Abstract

Niemann-Pick type C (NP-C) is a rare autosomal recessive disorder characterized by storage of unesterified glycolipids and cholesterol in lysosome and/or late endosome due to mutations in either *NPC1* or *NPC2* gene. This study aims to identify the spectrum of sequence alterations associated to NP-C in individuals with clinical suspicion of this disease. The entire coding region and flanking sequences of both genes associated to NP-C were evaluated in a total of 265 individuals that were referred to our laboratory. Clinical and/or biochemical suspicion of NP-C was confirmed by molecular analysis in 54 subjects. In this cohort, 33 different sequence alterations were identified in *NPC1* and one in *NPC2*. Among those, 5 novel alterations in *NPC1* gene were identified as follows: one deletion (p.Lys38_Tyr40del), one frameshift (p.Asn195Lysfs*2), and three missense mutations (p.Cys238Arg, p.Ser365Pro and, p.Val694Met) that are likely to be pathogenic through different approaches, including in silico tools as well as multiple sequence alignment throughout different species. We have also reported main clinical symptoms of patients with novel alterations and distribution of frequent symptoms in the cohort. Findings reported here contribute to the knowledge of mutation spectrum of NP-C, defining frequent mutations as well as novel sequence alterations associated to the disease.

Keywords Niemann-Pick type C disease · *NPC1* gene · *NPC2* gene · Mutation spectrum · Novel variation

Introduction

Niemann-Pick type C disease (NP-C disease; OMIM #257220) is a rare autosomal recessive neurodegenerative disorder characterized by storage of unesterified glycolipids and cholesterol in lysosome and/or late endosome (LE/L) due to mutations in either *NPC1* or *NPC2* genes [1]. This disorder causes premature death, and subjects from different ethnic groups can be affected [2–4]. NP-C prevalence is approximately 1/100,000 live births, but incidence can vary among different countries [4]. Hepatosplenomegaly, vertical supranuclear ophthalmoplegia, progressive ataxia, dystonia, and dementia are among symptoms characterized as the "classic" phenotype [5–7]. Mutations in genes coding for the large transmembrane endosomal NPC1 and a small soluble lysosomal NPC2 proteins result in intracellular sterol

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12035-019-1528-z) contains supplementary material, which is available to authorized users.

✉ Maria Luiza Saraiva-Pereira
mlpereira@hcpa.edu.br

[1] Laboratório de Identificação Genética, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil

[2] Programa de Pós-Graduação em Ciências Biológicas: Bioquímica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[3] Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[4] Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil

[5] Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[6] INAGEMP—Instituto Nacional de Genética Médica Populacional, Porto Alegre, RS, Brazil

[7] Departamento de Bioquímica, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

cycling alterations [8]. Great majority of NP-C cases is due to mutations in *NPC1* gene (95%) whereas the remaining are caused by mutations in *NPC2* gene [4, 9]. The human *NPC1* gene is located at *locus* 18q11, spans more than 47 kb, and is organized into 25 exons. The transcript is 4.9 kb long encoding a protein with 1278 amino acids [10]. NPC1 protein has 13 transmembrane domains, 3 large and 4 small luminal loops, 6 small cytoplasmic loops, and a cytoplasmic tail [11]. High homology was observed among NPC1 protein and other NP-C orthologs, such as mouse, *Saccharomyces cerevisiae*, and *Caenorhabditis elegans* [12]. High sequence homology between NPC1 and other proteins that are involved in cholesterol metabolism was also observed [13, 14]. The human *NPC2* gene is located at *locus* 14q24.3, spans more than 13 kb, and is organized into 5 exons. The transcript of 0.9 kb produces a small soluble glycoprotein that contains 131 amino acid residues [5, 15]. *NPC1* and *NPC*2 genes have many mammalian orthologs with highly conserved primary sequences [16].

Diagnosis of NP-C requires biochemical evaluation, such as Filipin staining test in fibroblasts or plasma oxysterols evaluation, and/or molecular analysis of *NPC1* and *NPC2* genes [5, 9]. To date, more than 460 different sequence alterations have been reported to be associated to NP-C.

This study describes the mutation spectrum of a broad genetic analysis in a cohort of patients with NP-C, including five novel sequence variants and rare mutations.

Material and Methods

Patients

In this study, we have included biological samples from 265 individuals that were sent to our laboratory from different regions of Brazil, from 2011 to 2017, through the NPC Network. Inclusion criteria were positive or inconclusive result in the Filipin staining test or a strong clinical suspicion of NP-C, regardless the biochemical evaluation outcome. This study was approved by our local Institutional Review Board (project number 05168).

DNA Isolation and Amplification of *NPC1* and *NPC2* Genes

Genomic DNA was isolated from peripheral white blood cells using standard protocols and stored at − 20 °C. Polymerase chain reaction (PCR) was used to selectively amplify specific fragments of *NPC1* (NG_012795.1) and *NPC2* (NG_007117.1) genes. Primer sequences can be found in Supplementary Table S1. Coding sequences and flanking regions (exons 1 to 25 of *NPC1* gene and exons 1 to 5 of *NPC2* gene) were amplified by PCR using genomic DNA as template. The whole coding region of *NPC1* was divided into 24 different amplicons (exons 15 and 16 were amplified as one fragment). *NPC2* coding region was divided into 5 different amplicons. Amplification reaction was performed in final volumes of 25 μL containing 25 ng genomic DNA, 200 mM of each dNTP, 2.5 μM of each primer (forward and reverse), 2.5 mM of $MgCl_2$, 200 mM of Tris-HCl (pH 8.4), 50 mM of KCl, and 1.25 U of *Taq* DNA Polymerase (Invitrogen™, Carlsbad, CA, USA). Cycling conditions were initial denaturation at 95 °C for 5 min, followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 30 s, and extension at 72 °C for 1 min, followed by final extension at 72 °C for 10 min. Each PCR product was verified by electrophoresis on a 1.5% (*w*/*v*) agarose gel and visualization under UV light.

DNA Sequencing

Amplicons were purified using 2.5 U of Exonuclease I (USB, Cleveland, OH, USA) and 0.25 U of Shrimp Alkaline Phosphatase (USB, Cleveland, OH, USA). DNA sequencing was performed using BigDye® Terminator Cycle Sequencing kit v. 3.1 (Applied Biosystems, Foster City, CA, USA) from universal M13 (− 20) forward and reverse primers, following the manufacturer's instructions. Sequences were analyzed with DNA Sequencing Analysis software v. 5.2 (Applied Biosystems) in an ABI PRISM® 3130*xl* Genetic Analyzer. All identified sequence variations were confirmed by sequencing an independent sample from both forward and reverse primers. Sequence variations were compared to data available in the NP-C database in the Human Gene Mutation Database (HGMD® - http://www.hgmd.cf.ac.uk), the Exome Aggregation Consortium (ExAC) browser (http://exac.broadinstitute.org/), the Genome Aggregation Database (gnomAD) browser (http://gnomad.broadinstitute.org/), and 1000genomes (http://www.internationalgenome.org/home).

Evaluation of Novel Mutations

Amino acid sequences of *NPC1* from 10 different species were compared by multiple alignment in order to determine whether changes identified in their amino acid sequences were associated to conserved residues. Sequences were searched for using the protein database from the National Center for Biotechnology Information (NCBI - https://www.ncbi.nlm.nih.gov/). Amino acid sequences were aligned with Clustal Omega using FASTA format (https://www.ebi.ac.uk/Tools/msa/clustalo/). In order to assess their potential pathogenicity, novel sequence variations in the *NPC1* coding region were analyzed using eight web-based tools. Those tools were PolyPhen-2 (Polymorphism Phenotyping v2, http://genetics.bwh.harvard.edu/pph2/) [17], SNPs3D (http://www.snps3d.org/) [18], Align GVGD (http://agvgd.iarc.fr/) [19], Mutation Taster (http://www.mutationtaster.org/) [20], Mendelian Clinically Applicable Pathogenicity (M-CAP) Score (http://bejerano.

Table 1 Alleles defined by this study. Novel sequence variants are shown in italics

Mutation	cDNA nucleotide substitution	Exon	# of alleles	Allelic frequence
NPC1 gene				
p.Lys38_40Tyrdel	*c.114-122del9*	*2*	*1*	*0.011*
p.Gln117*	c.349C > T	4	1	0.011
p.Cys177Tyr	c.530G > A	5	2	0.023
p.Ala183Thr + *p.Ser365Pro*	c.547G > A + *c.1093 T> C*	5 and 8	1	0.011
p.Ser151Phefs*70	c.451_452delAG	4	1	0.011
*p.Asn195Lysfs*2*	*c.584dupA*	*5*	*1*	*0.011*
p.Cys238Arg	*c.712 T> C*	*6*	*1*	*0.011*
p.Cys247Tyr	c.740G > A	6	1	0.011
p.Arg372Trp	c.1114C > T	8	1	0.011
p.Arg615His	c.1844G > A	12	1	0.011
p.Val664Met	c.1990G > A	13	2	0.023
p.Ser667Leu	c.2000C > T	13	1	0.011
p.Val694Met	*c.2080G> A*	*13*	*1*	*0.011*
p.Gly710Alafs*19	c.2129delA	13	2	0.023
p.Pro733Serfs*10	c.2196_2197insT	14	1	0.011
p.Ala764Val	c.2291C > T	15	1	0.011
p.Ser865Leu	c.2594C > T	17	1	0.011
p.Ala926Thr	c.2776G > A	18	1	0.011
p.Trp942Cys	c.2826G > T	19	1	0.011
p.Asp945Asn	c.2833G > A	19	1	0.011
p.Cys957Tyr	c.2870G > A	19	1	0.011
p.Gly992Arg	c.2974G > C	20	1	0.011
p.Pro1007Ala	c.3019C > G	20	15	0.170
p.Ala1035Val	c.3104C > T	21	24	0.273
p.Ile1061Thr	c.3182 T > C	21	4	0.045
p.Gly1140Val	c.3419G > T	22	1	0.011
p.Asn1156Ile	c.3467A > T	22	1	0.011
p.Leu1157Pro	c.3470 T > C	22	1	0.011
p.Val1165Met	c.3493G > A	23	1	0.011
p.Glu1166Lys	c.3496G > A	23	1	0.011
p.Arg1186His	c.3557G > A	23	2	0.023
p.Phe1221Serfs*20	c.3662delT	24	13	0.148
Total			88	
NPC2 gene				
p.Glu20*	c.58G > T	2	2	1.000
Total			2	

stanford.edu/mcap/) [21], Combined Annotation Dependent Depletion (CADD) (http://cadd.gs.washington.edu/snv) [22], Rare Exome Variant Ensemble level (REVEL) (https://sites.google.com/site/revelgenomics/) [23], and Variant Effect Scoring Tool (VEST3) (http://hg19.cravat.us/CRAVAT/) [24]. Mutalyzer 2.0 was used as a reference for naming novel sequence variations (https://mutalyzer.nl/) [25]. Model structure of *NPC1* was generated by PyMOL 2.0 software (https://pymol.org/2/), and mutant models by Modeler 9.1 software [26], using PDB ID code 3JD8.

Results

Genotype of NP-C patients was defined through identification of novel and/or rare sequence alterations in *NPC1* or *NPC2* genes. Mutant alleles were confirmed in 54 out of 265 individuals, being 29 (53.7%) females and 25 (46.3%) males. Within NP-C confirmed patients, 18 (33.3%) were from consanguineous marriage. Among NP-C cases identified in this work, 52 (96.3%) patients have mutations in *NPC1* gene while the remaining 2 (3.7%) patients carry mutations in *NPC2*. In

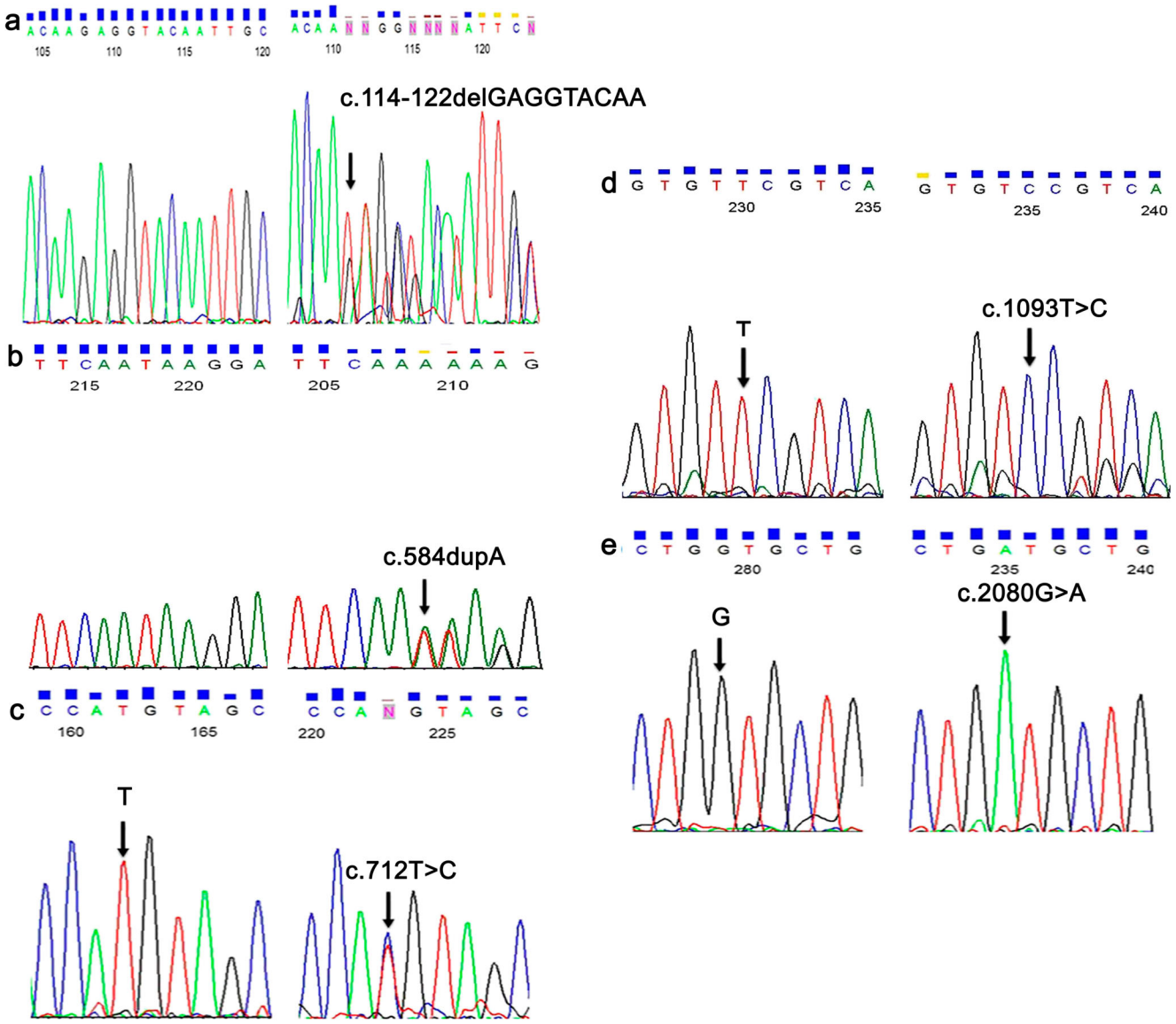

Fig. 1 DNA sequencing of novel sequence variations identified in *NPC1* gene. (a) Direct sequencing of exon 2 from the forward primer. Arrow indicates the beggiging of the deletion in the p.Lys38_Tyr40del (c.114_122delGAGGTACAA) variation. Patient is heterozygous for this variation; therefore, after sequence variation, two different profiles can be seen in the figure: one from wild-type allele and the other from mutant allele. (b) Direct sequencing of exon 5 from the forward primer. The arrow shows nucleotide duplication that characterizes the p.Asn195Lysfs*2 (c.584dupA) variation. Patient is heterozygous for this variation; therefore, after sequence variation, two different profiles can be seen in the figure: one from wild-type allele and the other from mutant allele. (c) Direct sequencing of exon 6 from the forward primer. The arrow indicates T to C substitution in the p.Cys238Arg (c.712 T > C) variation. Patient is heterozygous for this variation. (d) Direct sequencing of part of exon 8 from the forward primer. The arrow indicates T to C substitution in the p.Ser365Pro (c.1093 T > C) variation. Patient is homozygous for this variation. (e) Direct sequencing of part of exon 13 from the forward primer. The arrow indicates G to A substitution in the p.Val694Met (c.2080G > A) variation. Patient is homozygous for this variation

total, 34 different sequence alterations were identified, including 5 novel variations in *NPC1*, and a detailed distribution of mutations is shown in Table 1. Frequency of variants was estimated using unrelated chromosomes only; i.e., we have just considered one allele from homozygous patients of consanguineous marriages, giving a total of 90 alleles. The most frequent mutation was p.Ala1035Val (27.0%), followed by p.Pro1007Ala (16.9%), and p.Phe1221Serfs*20 (14.6%).

All novel variants described here are located in *NPC1* gene, and distributed as follows: one small deletion (p.Lys38_Tyr40del), one frameshift (p.Asn195Lysfs*2), and 3 missense mutations (p.Cys238Arg, p.Ser365Pro, and p.Val694Met). Sequencing profile of novel mutations is in Fig. 1. These novel changes were not found among 400 alleles from normal individuals. All 5 patients carrying novel mutations were previously evaluated by Filipin staining test and results were positive. A brief clinical description of patients

with novel sequence variations can be found in Table 2. Variations cited in this work as novel were not found in the Human Gene Mutation Database (HGMD®), ExaC, gnomAD, and 1000genomes. Findings regarding these novel changes are described below.

The p.Lys38_Tyr40del (c.114-122del9) mutation is a deletion of nine nucleotides within exon 2 (Fig. 1(a)) that leads to a protein with three missing amino acids (Lys, Arg, and Tyr). This mutation was found in a compound heterozygote carrying p.Phe1221Serfs*20 in the other allele. This male patient was diagnosed when he was 1 year old and clinical symptoms include dysphagia, cognitive decline, and developmental delay (Table 2).

Variation p.Asn195Lysfs*2 is due to a duplication of one nucleotide (adenine) in exon 5 of *NPC1* gene (Fig. 1(b)). This frameshift was identified in a female patient *in trans* with p.Phe1221Serfs*20. She was diagnosed at 6 months of age, and main symptoms were hepatosplenomegaly, hypotonia, and developmental delay. This mutation produces a truncated protein that is expected to have a defective function.

Regarding missense mutations, p.Cys238Arg is located at exon 6, due to T to C change (Fig. 1(c)), which was found in a female patient diagnosed at 2 years of age. Her clinical symptoms included cerebellar ataxia, developmental delay, and cataplexy. The p.Cys238Arg was found *in trans* with p.F1221Sfs*20 (patient was a compound heterozygote), and variants were confirmed in her parents.

Novel variation p.Ser365Pro is due to T to C change in exon 8 of *NPC1* (Fig. 1(d)) and it was found *in cis* with p.Ala183Thr. This complex allele was found in a homozygous male patient from a consanguineous marriage, and both mutations were confirmed in his parents.

The other novel missense mutation, p.Val694Met, is due to G to A change in exon 13 (Fig. 1(e)), and this alteration was detected in a female patient from consanguineous marriage. She was 14 years old at diagnosis and clinical symptoms include cerebellar ataxia, and vertical supranuclear gaze palsy. Patient was homozygous for this alteration, and variant was also confirmed in her parents.

All novel sequence variations were evaluated through alignment of amino acid sequences from 10 different organisms, and alterations are located within conserved residues, suggesting an effect on protein function or structure (Fig. 2). Pathogenicity of novel missense variations was evaluated using different web-based tools (supplementary Table S2), following the guidelines by the American College of Medical Genetics and Genomics (ACMG) to the interpretation of sequence variants [27]. Two missense mutations (p.Cys238Arg and p.Val694Met) were defined as being pathogenic by all different tools. p.Ser365Pro was also classified as pathogenic by great majority of tools except by one (CADD software). However, considering that this serine residue is conserved among species (Fig. 2), and that this substitution introduces a novel imino group in the protein, a pathogenic effect might be also expected in this case.

Finally, distribution of age at diagnosis ranged from 2 months to 46 years with an average of 11.3 years, and 32 cases (59.3%) were diagnosed in patients of up to 10 years of age. Patients included in the analysis were from different regions of Brazil, and more detailed description of symptoms was available from 33 confirmed cases. Therefore, more frequent symptoms, based on cases with complete clinical description, were splenomegaly, hepatomegaly, cerebellar ataxia, and vertical supranuclear gaze palsy, and distribution of symptoms according to age group is shown in supplementary Fig. S1.

Discussion

We identified mutations in 54 NP-C patients, being 96.3% in *NPC1* gene and 3.7% in *NPC2* gene. Distribution of mutations shown here is similar to the described in the literature from studies worldwide, where mutations in *NPC1* gene occur in 95% of the NP-C patients [2, 12, 15, 28].

The most frequent mutation in our sample population was p.Ala1035Val (27.0%) that is different from cohorts reported in North hemisphere, where p.Ile1061Thr is described as the most frequent one [29–38]. As previously reported, high frequency of p.Ile1061Thr in Hispanic patients suggest a founder effect originated in Western Europe [32, 39]. A prevalence of a different mutation in this studied cohort suggests a different

Table 2 Brief clinical description of patients that carry novel sequence variations identified by this study

Mutation	Genotype	Age at diagnosis	Gender	Clinical symptoms
p.Lys38_Tyr40del	p.Lys38_Tyr40del/p.Phe1221Serfs*20	1 year	Male	Dysphagia, cognitive decline, developmental delay
p.Ser365Pro	[p.Ser365Pro + p.Ala183Thr]/ [p.Ser365Pro + p.Ala183Thr]	27 years	Male	Neurological regression
p.Asn195Lysfs*2	p.N195Kfs*2/p.F1221Sfs*20	6 months	Female	Hepatomegaly, splenomegaly, hypotonia, developmental delay
p.Cys238Arg	p.Cys238Arg/p.Phe1221Serfs*20	2 years	Female	Cerebellar ataxia, developmental delay, cataplexy
p.Val694Met	p.Val694Met/p.Val694Met	14 years	Female	Cerebellar ataxia, vertical supranuclear gaze palsy

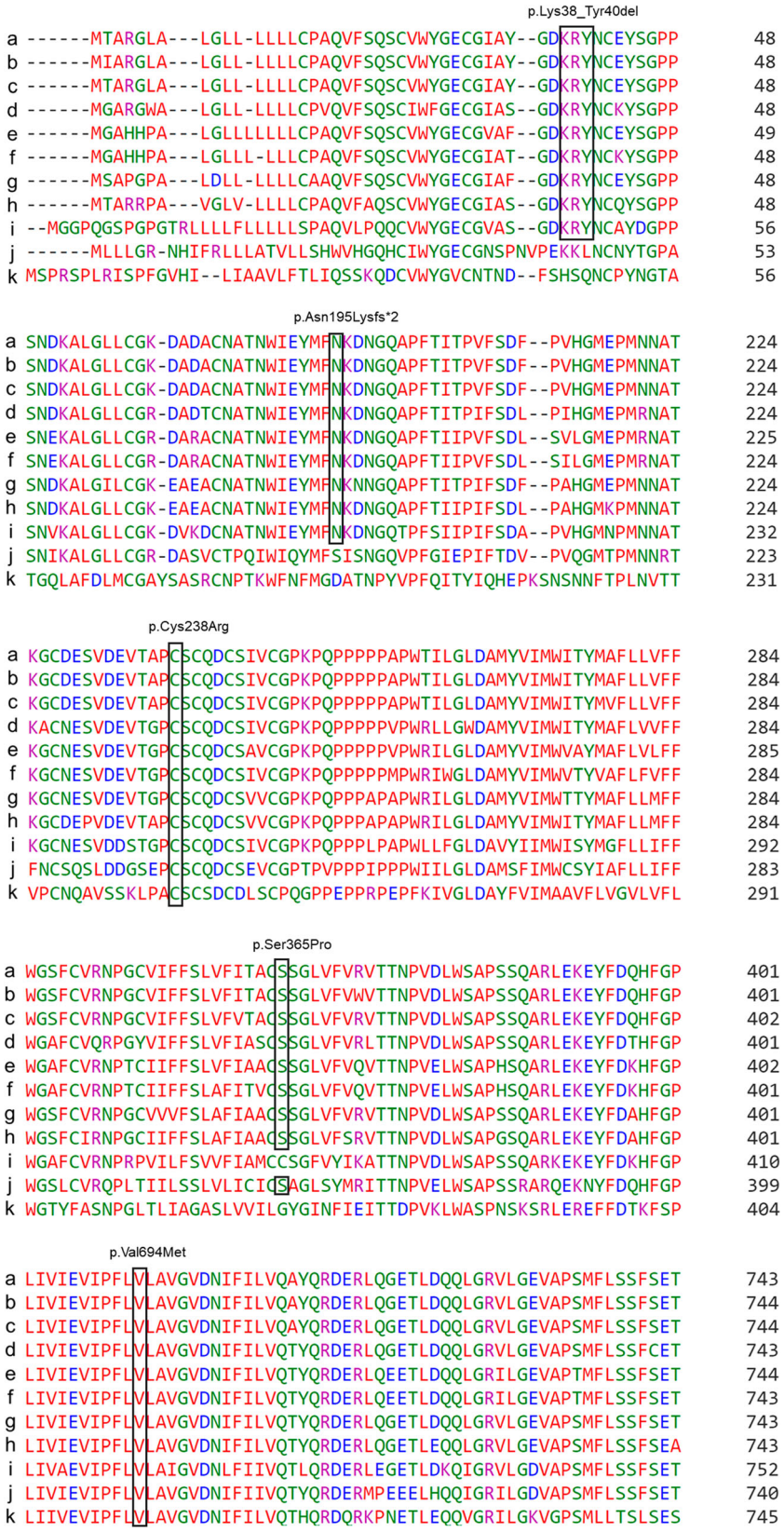

Fig. 2 Alignment of amino acid sequences of several organisms. Specific regions flanking mutations are represented. (a) Human (NP_000262.2); (b) Chimpanzee (XP_016788941.1); (c) Orangutan (XP_009250570.1); (d) Domestic guinea pig (XP_003474080.2); (e) Rat (NP_705888.2); (f) Mouse (NP_032746.2); (g) Horse (XP_005612821.1); (h) Dog (NP_001003107.1); (i) Chicken (XP_419162.3); (j) Zebrafish (NP_001230804.1); (k) Fruit fly (NP_001188769.1). Boxes indicate conserved residues associated to novel mutations. Designation of each mutation is given above the appropriate box

ethnic background of NP-C patients in Brazil. The second most common mutation in our study was p.Pro1007Ala (16.9%), and this alteration was also reported as being frequent in different European countries [5]. Frequency data of

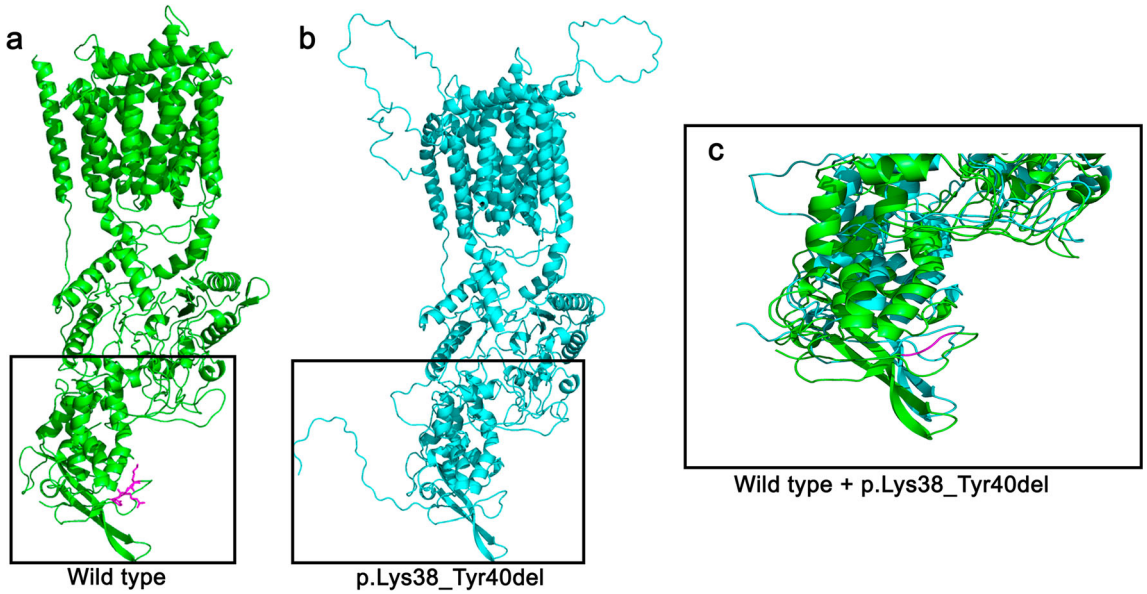

Fig. 3 Location of p.Lys38_40Tyr in the NPC1 protein. (a) Image represents the wild-type NPC1 protein; region of three amino acids (lysine, arginine, and tyrosine) involved in the deletion is represented in pink. (b) Image represents mutant NPC1 protein p.Lys38_40Tyr. (c) Close-up and superposed view of wild-type (in green) and mutant (in blue) NPC1 proteins. Figures were generated by PyMOL 2.0

this mutation reported in Portuguese, British, and German patients ranged from 15 to 20% [37].

We have observed higher frequency of some mutations in specific regions as follows: p.Ala1035Val was present in 32.0% of mutant alleles identified in São Paulo state, p.Pro1007Ala was found in 53.3% of mutant alleles from Paraná state, and p.Phe1221Serfs*20 was identified in 75.0% of alleles from Pernambuco state. Although preliminary, these data indicate that regional variation of ethnic background in a huge country as Brazil might determine higher frequencies of mutations in specific places. Several studies have reported specific disease-causing mutations among different populations and ethnic groups associated to NP-C [30, 31, 38–40]. Additional analyses are required to further investigate this issue.

Distribution of confirmed cases among geographical regions of Brazil was as follows: 46.3% of cases in Southeast, followed by 25.9% of cases from Northeast, 24.1% of cases from South, and 3.7% from West Central. This higher frequency of cases in Southeast might be a combination of highly populated region as well as a more facilitated access to health system.

Clinical presentation of NP-C can be heterogeneous and nonspecific, which makes more difficult to reach a correct diagnosis [9]. Symptoms of pediatric patients ($\leq$4 years) described here are in agreement to a previous report that more discriminatory signs for NP-C in pediatric patients are splenomegaly, hepatomegaly, dysphagia, cognitive decline, delayed neuro-psychomotor, ataxia, and cataplexy [41]. Clinical findings reported here in adult patients, such as seizures, neurological regression, splenomegaly, cognitive decline, cerebellar ataxia, and vertical

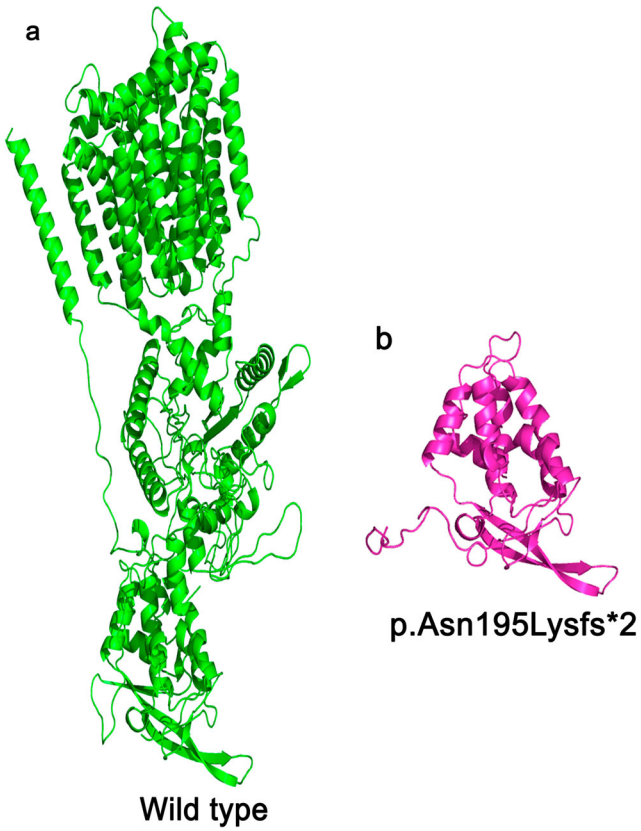

Fig. 4 Images representing wild-type NPC1 protein and NPC1 protein produced in the presence of p.Asn195Lysfs*2. (a) Wild-type NPC1 protein (1278 amino acids). (b) Mutant NPC1 protein produced by p.Asn195Lysfs*2 mutation (197 amino acids). Figures were generated by PyMOL 2.0

supranuclear gaze palsy, were described as more commonly found in older patients (supplementary Fig. S1).

Novel sequence variants appear to be widespread along different regions in the protein: p.Lys38_Tyr40del, p.Asn195Lysfs*2, and p.Cys238Arg are located within lumen A domain, p.Ser365Pro in the transmembrane II (TMII) domain, and p.Val694Met in the transmembrane V (TMV). Position of these variations can be visualized in NPC1 protein topology generated by Protter software [42] (supplementary Fig. S2).

Mutation p.Lys38_Tyr40del is an in-frame deletion that leads to a protein lacking three amino acids (Lys, Arg, and Tyr). This deletion is located within the N-terminal domain (NTD) (supplementary Fig. S2), which is the first luminal domain composed by 240 amino acids [43]. This type of mutation generates a mutant protein with a different tertiary structure (Fig. 3) that will likely affect protein function.

The frameshift variation p.Asn195Lysfs*2 produces a truncated small protein that is expected to have a defective function. The wild-type NPC1 protein has 1278 amino acids, and mutant protein produced by this variation would be expected to have 197 amino acids only (Fig. 4). Therefore, essential domains of NPC1 protein will be missing, and normal function highly impaired.

Missense mutation p.Cys238Arg is also located in a very relevant domain of the protein. The amino acid cysteine at position 238 establishes one of two disulfide bonds (C97–C238 and C227–C243) from Ψ loop. This particular loop has been reported before as being an important interface between N-terminal domain (NTD) and middle luminal domain (MLD) [44]. Therefore, the replacement of this cysteine residue by an arginine prevents the establishment of a disulfide bond, which changes protein conformation as can be observed in Fig. 5(b, c).

p.Ser365Pro variation introduces a novel imino group in the protein (Fig. 5(d, e)). This imino group will be part of a transmembrane domain (supplementary Fig. S2) and can cause a very relevant change in the protein structure by itself. In this particular case, as this mutation is found in a complex allele *in cis* with another mutation, an even stronger impact in protein function is expected due to a combination of effects.

The remaining missense mutation p.Val694Met is also located within a transmembrane region, which is important for NPC1 (Fig. 5(f) and Fig. 5(g)). This region shows high homology to the sterol-sensing domains (SSD) of HMG-Co A

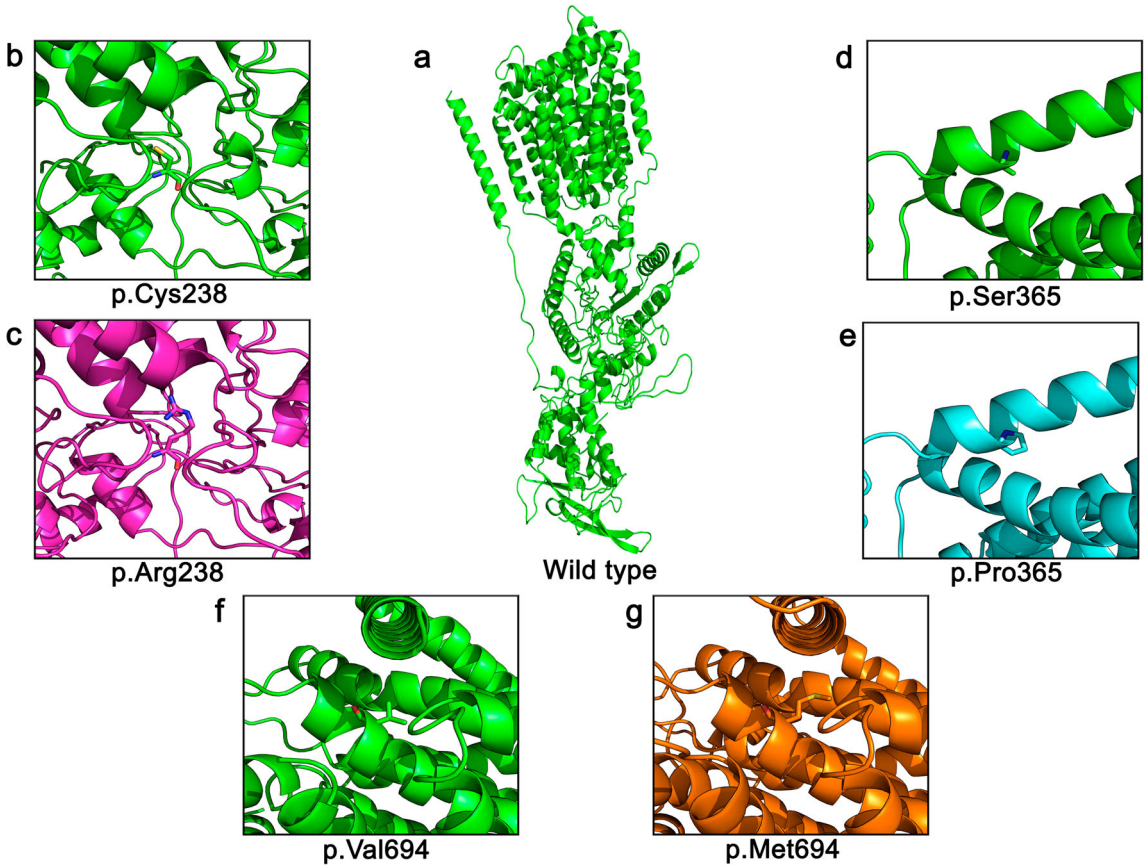

Fig. 5 Missense mutations in NPC1 when compared to the wild-type amino acido residue. Arrows indicate location of wild-type and mutant amino acid residue in each variant. (a) Wild-type NPC1 protein. (b) Wild-type residue Cys (cysteine) compared to (c) mutant residue Arg (arginine). (d) Wild-type residue Ser (serine) compared to (e) mutant residue Pro (proline). (f) Wild-type residue Val (valine) compared to (g) mutant residue Met (methionine). Figures were generated by PyMOL 2.0

reductase that is involved in cholesterol synthesis, and to the sterol regulatory element-binding protein (SREBP) cleavage-activating protein (SCAP), which is an activator of a transcription factor in cholesterol biosynthesis [11, 29, 30]. The insertion of cholesterol into the lysosomal membrane involves NPC1 transmembrane domains, including the sterol-sensing motif that has been identified in other proteins as involved in cholesterol homeostasis [45]. The majority of mutations in the SSD region is associated to a severe phenotype [30, 46].

All novel variants were located within conserved regions when multiple alignment analysis was performed with sequences of 10 different species (Fig. 2). These conserved regions imply a requirement of those amino acids for normal protein structure and/or activity throughout species. In addition, all novel variations were tested for hydrophobicity prediction and variations of hydrophobicity levels are expected in each one. This is further evidence related to protein topology susceptibility associated to novel sequences alterations reported here.

In summary, data provided here contribute to the knowledge of worldwide mutation spectrum associated to NP-C. Combination of molecular analyses with in silico tools, as well as molecular modeling, can generate a more comprehensive insight into NP-C associated proteins, with a potential to identify additional targets to the development of novel therapeutics for Niemann-Pick type C.

Acknowledgements We would like to thank the patients and their families for providing biological material for this study. The authors also thank investigators who enrolled patients in the NPC Brazil Network. This study was partially supported by grants from Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq), Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Fundo de Incentivo a Pesquisa e Eventos do HCPA (FIPE-HCPA), and INAGEMP—National Institute of Population Medical Genetics (grant CNPq 573993/2008-4). The NPC Brazil Network is partially funded by an unrestricted grant from Actelion (05168). MPB, HB, and FN were supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); RG and MLSP were supported by CNPq.

References

1. Vance JE, Karten B (2014) Niemann-Pick C disease and mobilization of lysosomal cholesterol by cyclodextrin. J Lipid Res 55:1609–1621. https://doi.org/10.1194/jlr.R047837
2. Vanier MT (2015) Complex lipid trafficking in Niemann-Pick disease type C. J Inherit Metab Dis 38:187–199. https://doi.org/10.1007/s10545-014-9794-4
3. Pandi S, Chandran V, Deshpande A, Kurien A (2014) Niemann-Pick disease type C or Gaucher's disease type 3? A clinical conundrum. BMJ Case Rep 2014:10–13. https://doi.org/10.1136/bcr-2014-203713
4. Vanier MT (2010) Niemann-Pick disease type C. Orphanet J Rare Dis 5:16. https://doi.org/10.1186/1750-1172-5-16
5. Vanier M, Millat G (2003) Niemann-Pick disease type C. Clin Genet 64:269–281. https://doi.org/10.1034/j.1399-0004.2003.00147.x
6. Patterson MC, Vanier MT, Suzuki K et al (2015) Part 16: Lysosomal disorders chapter 145: Niemann-Pick disease type C: a lipid trafficking disorder. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease. Accessed in 10/06/2015. https://doi.org/10.1036/ommbid.175
7. Walterfang M, Fietz M, Fahey M, Sullivan D, Leane P, Lubman DI, Velakoulis D (2006) The neuropsychiatry of Niemann-Pick type C disease in adulthood. J Neuropsychiatry Clin Neurosci 18:158–170. https://doi.org/10.1176/appi.neuropsych.18.2.158
8. Walterfang M, Abel LA, Desmond P, Fahey MC, Bowman EA, Velakoulis D (2013) Cerebellar volume correlates with saccadic gain and ataxia in adult Niemann-Pick type C. Mol Genet Metab 108:85–89. https://doi.org/10.1016/j.ymgme.2012.11.009
9. Patterson MC, Hendriksz CJ, Walterfang M, Sedel F, Vanier MT, Wijburg F, NP-C Guidelines Working Group (2012) Recommendations for the diagnosis and management of Niemann-Pick disease type C: an update. Mol Genet Metab 106:330–344. https://doi.org/10.1016/j.ymgme.2012.03.012
10. Morris JA, Zhang D, Coleman KG, Nagle J, Pentchev PG, Carstea ED (1999) The genomic organization and polymorphism analysis of the human Niemann-Pick C1 gene. Biochem Biophys Res Commun 261:493–498. https://doi.org/10.1006/bbrc.1999.1070
11. Davies JP, Ioannou YA (2000) Topological analysis of Niemann-Pick C1 protein reveals that the membrane orientation of the putative sterol-sensing domain is identical to those of 3-hydroxy-3-methylglutaryl-CoA reductase and sterol regulatory element binding protein cleavage-activating. J Biol Chem 275:24367–24374. https://doi.org/10.1074/jbc.M002184200
12. Carstea ED, Morris JA, Coleman KG, Loftus SK, Zhang D, Cummings C, Gu J, Rosenfeld MA et al (1997) Niemann-Pick C1 isease gene: Homology to mediators of cholesterol homeostasis. Science 277:228–231. https://doi.org/10.1126/science.277.5323.228
13. Hua X, Nohturfft A, Goldstein JL, Brown MS (1996) Sterol resistance in CHO cells traced to point mutation in SREBP cleavage-activating protein. Cell 87:415–426
14. Millard EE, Gale SE, Dudley N, Zhang J, Schaffer JE, Ory DS (2005) The sterol-sensing domain of the Niemann-Pick C1 (NPC1) protein regulates trafficking of low density lipoprotein cholesterol. J Biol Chem 280:28581–28590. https://doi.org/10.1074/jbc.M414024200
15. Naureckiene S, Sleat DE, Lackland H et al (2000) Identification of HE1 as the second gene of Niemann-Pick C disease. Science 290:2298–2301. https://doi.org/10.1126/science.290.5500.2298
16. Vanier MT, Millat G (2004) Structure and function of the NPC2 protein. Biochim Biophys Acta 1685:14–21. https://doi.org/10.1016/j.bbalip.2004.08.007
17. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7:248–249. https://doi.org/10.1038/nmeth0410-248
18. Yue P, Moult J (2006) Identification and analysis of deleterious human SNPs. J Mol Biol 356:1263–1274. https://doi.org/10.1016/j.jmb.2005.12.025
19. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, de Silva D, Zharkikh A et al (2005) Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 43:295–305. https://doi.org/10.1136/jmg.2005.033878
20. Schwarz JM, Cooper DN, Schuelke M, Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11:361–362. https://doi.org/10.1038/nmeth.2890

21. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, Bernstein JA, Bejerano G (2016) M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet 48:1581–1586. https://doi.org/10.1038/ng.3703
22. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46:310–315. https://doi.org/10.1038/ng.2892
23. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q et al (2016) REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet 99:877–885. https://doi.org/10.1016/j.ajhg.2016.08.016
24. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R (2013) Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics 14:S3. https://doi.org/10.1186/1471-2164-14-S3-S3
25. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PEM (2008) Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. Hum Mutat 29:6–13. https://doi.org/10.1002/humu.20654
26. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234(3):779–815. https://doi.org/10.1006/jmbi.1993.1626
27. Richards S, Aziz N, Bale S et al (2015) 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 17:405–423. https://doi.org/10.1038/gim.2015.30
28. Vanier MT, Duthel S, Rodriguez-Lafrasse C et al (1996) Genetic heterogeneity in Niemann-Pick C disease: a study using somatic cell hybridization and linkage analysis. Am J Hum Genet 58:118–125
29. Millat G, Marçais C, Rafi M a, et al (1999) Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant allele in patients of Western European descent and correlates with a classic juvenile phenotype. Am J Hum Genet 65:1321–1329. https://doi.org/10.1086/302626
30. Yamamoto T, Ninomiya H, Matsumoto M, et al (2000) Genotype-phenotype relationship of Niemann-Pick disease type C: a possible correlation between clinical onset and levels of NPC1 protein in isolated skin fibroblasts. J Med Genet 37:707–712
31. Garver WS, Jelinek D, Meaney FJ, Flynn J, Pettit KM, Shepherd G, Heidenreich RA, Vockley CMW et al (2010) The national Niemann-Pick type C1 disease database: correlation of lipid profiles, mutations, and biochemical phenotypes. J Lipid Res 51:406–415. https://doi.org/10.1194/jlr.P000331
32. Park WD, O'Brien JF, Lundquist PA et al (2003) Identification of 58 novel mutations in Niemann-Pick disease type C: correlation with biochemical phenotype and importance of PTC1-like domains in NPC1. Hum Mutat 22:313–325. https://doi.org/10.1002/humu.10255
33. Bauer P, Knoblich R, Bauer C, Finckh U, Hufen A, Kropp J, Braun S, Kustermann-Kuhn B et al (2002) NPC1: Complete genomic sequence, mutation analysis, and characterization of haplotypes. Hum Mutat 19:30–38. https://doi.org/10.1002/humu.10016
34. Tarugi P, Ballarini G, Bembi B, Battisti C, Palmeri S, Panzani F, di Leo E, Martini C et al (2002) Niemann-Pick type C disease: mutations of NPC1 gene and evidence of abnormal expression of some mutant alleles in fibroblasts. J Lipid Res 43:1908–1919. https://doi.org/10.1194/jlr.M200203-JLR200
35. Ribeiro I, Marcão A, Amaral O, Sá Miranda M, Vanier MT, Millat G (2001) Niemann-Pick type C disease: NPC1 mutations associated with severe and mild cellular cholesterol trafficking alterations. Hum Genet 109:24–32
36. Meiner V, Shpitzen S, Mandel H, Klar A, Ben-Neriah Z, Zlotogora J, Sagi M, Lossos A et al (2001) Clinical-biochemical correlation in molecularly characterized patients with Niemann-Pick type C. Genet Med 3:343–348. https://doi.org/10.1097/00125817-200109000-00003
37. Yamamoto T, Nanba E, Ninomiya H, Higaki K, Taniguchi M, Zhang H, Akaboshi S, Watanabe Y et al (1999) NPC1 gene mutations in Japanese patients with Niemann-Pick disease type C. Hum Genet 105:10–16
38. Greer WL, Riddell DC, Murty S, Gillan TL, Girouard GS, Sparrow SM, Tatlidil C, Dobson MJ et al (1999) Linkage disequilibrium mapping of the Nova Scotia variant of Niemann-Pick disease. Clin Genet 55:248–255. https://doi.org/10.1034/j.1399-0004.1999.550406.x
39. Fernandez-Valero EM, Ballart a IC et al (2005) Identification of 25 new mutations in 40 unrelated Spanish Niemann-Pick type C patients: genotype-phenotype correlations. Clin Genet 68:245–254. https://doi.org/10.1111/j.1399-0004.2005.00490.x
40. Millat G, Marçais C, Tomasetto C, Chikh K, Fensom AH, Harzer K, Wenger DA, Ohno K et al (2001) Niemann-Pick C1 disease: correlations between NPC1 mutations, levels of NPC1 protein, and phenotypes emphasize the functional significance of the putative sterol-sensing domain and of the cysteine-rich luminal loop. Am J Hum Genet 68:1373–1385. https://doi.org/10.1086/320606
41. Yang C-C, Su Y-N, Chiou P-C, Fietz MJ, Yu CL, Hwu WL, Lee MJ (2005) Six novel NPC1 mutations in Chinese patients with Niemann-Pick disease type C. J Neurol Neurosurg Psychiatry 76:592–595. https://doi.org/10.1136/jnnp.2004.046045
42. Omasits U, Ahrens CH, Müller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30:884–886. https://doi.org/10.1093/bioinformatics/btt607
43. Kwon HJ, Abi-Mosleh L, Wang ML, et al (2009) Structure of N-Terminal Domain of NPC1 Reveals Distinct Subdomains for Binding and Transfer of Cholesterol. Cell 137:1213–1224. https://doi.org/10.1016/j.cell.2009.03.049
44. Li X, Lu F, Trinh MN, Schmiege P, Seemann J, Wang J, Blobel G (2017) 3.3 Å structure of Niemann–Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol transport. Proc Natl Acad Sci 114:201711716–201719121. https://doi.org/10.1073/pnas.1711716114
45. Vance JE (2010) Transfer of cholesterol by the NPC team. Cell Metab 12:105–106. https://doi.org/10.1016/j.cmet.2010.07.004
46. Millat G, Chikh K, Naureckiene S, et al (2001) Niemann-Pick disease type C: spectrum of HE1 mutations and genotype/phenotype correlations in the NPC2 group. Am J Hum Genet 69:1013–1021. https://doi.org/10.1086/324068