




Identification of novel variants in a large cohort of children with Tay–Sachs disease: An initiative of a multicentric task force on lysosomal storage disorders by Government of India

Mehul Mistri¹ · Sanjeev Mehta² · Dhaval Solanki³ · Mahesh Kamate⁴ · Neerja Gupta⁵ · Madhulika Kabra⁵ · Ratna Puri⁶ · Katta Girisha⁷ · Sankar Hariharan⁸ · Sheela Nampoothiri⁹ · Frenny Sheth¹  · Jayesh Sheth¹

Received: 15 October 2018 / Revised: 26 June 2019 / Accepted: 16 July 2019 / Published online: 6 August 2019

© The Author(s), under exclusive licence to The Japan Society of Human Genetics 2019

Abstract

Tay–Sachs disease (TSD) (OMIM) is a neurodegenerative lysosomal storage disorder caused due to mutations in the *HEXA* gene. To date, nearly 190 mutations have been reported in *HEXA* gene. Here, we have characterized 34 enzymatically confirmed TSD families to investigate the presence of novel as well as known variants in *HEXA* gene. Overall study detected 25 variants belonging to 31 affected TSD patients and 3 carrier couples confirmed by enzyme study. Of these 17 patients harbors 15 novel variants, including seven missense variants [p.V206L, p.Y213H, p.R252C, p.F257S, p.C328G, p.G454R, and p.P475R], four nonsense variant [p.S9X, p.E91X, p.W420X, and p.W482X], two splice site variants [c.347–1G>A and c.460–1G>A], and two small deletion [c.1349delC (p.A450VfsX3) and c.52delG (p.G18Dfs*82)]. While remaining 17 patients harbors 10 previously reported variants that includes six missense variants [p.M1T, p.R170Q, p.D322Y, p.D322N, p.E462V, and p.R499C], one nonsense variant [p.Q106X], two splice site variants [c.1073+1G>A and c.459+4A>G] and one 4 bp insertion [c.1278insTATC (p.Y427IfsX5)]. In conclusion, Indian infantile TSD patients provide newer insight into the molecular heterogeneity of the TSD. Combining present study and our earlier studies, we have observed that 67% genotypes found in Indian TSD patients are novel, which are associated with severe infantile phenotypes, while rest 33% genotypes found in our cohort were previously reported in various populations. In addition, higher frequency of the p.E462V and c.1278insTATC mutations in the present study further support and suggest the prevalence of p.E462V mutation in the Indian population.

Introduction

Tay–Sachs disease (TSD) [MIM* 606869] is an autosomal recessive rare neurodegenerative lysosomal storage disorder, marked by the accumulation of GM2 gangliosidosis in the central nervous system (neuronal cell and spinal cell). In TSD, the GM2 gangliosides does not get metabolized into GM3 gangliosides due to deficiency of β -hexosaminidase-A (Hex-A) (*HEXA*; EC: 3.2.1.52) enzyme, characterized by mutation in *HEXA* gene [1, 2]. It is one of the most common lipid storage disorders in India [3].

The clinical phenotype associated with TSD varies from infantile onset presented with progressive neuroregression, seizures, muscle weakness, hyperacusis, and cherry red spot on fundus to adult onset with progressive dementia and gait disturbance [1, 3]. The human *HEXA* gene (14 exons and 13 introns), located on chromosome 15q23–q24 with 35.56 kb spans. As per Human Gene Mutation Database (HGMD), nearly 190 disease causing mutations in *HEXA* gene have

✉ Jayesh Sheth
jshethad1@gmail.com

¹ FRIGE's Institute of Human Genetics, Ahmedabad, Gujarat, India

² Usha Deep Hospital, Naranpura, Ahmedabad, Gujarat, India

³ Mantra Child Neurology & Epilepsy Clinic, Bhavnagar, Gujarat, India

⁴ KLES Prabhakar Kore Hospital, Belgaum, Karnataka, India

⁵ All India Institute of Medical Sciences (AIIMS), New Delhi, India

⁶ Sir Ganga Ram Hospital, New Delhi, India

⁷ Kasturba Medical College, Manipal, Karnataka, India

⁸ SAT Hospital, Thiruvananthapuram, Trivandrum, India

⁹ Amrita Institute of Medical Science & Research Centre, Kochi, Kerala, India

been reported till date. Mutations c.1421+1G>C, c.1274_1277dupTATC, and c.805G>A (p.G269S) are the most commonly prevalent in 98% of Ashkenazi Jews and 35% non-Ashkenazi Jews population [4–6]. Other mutations in the *HEXA* gene are described rarely in various populations across the globe [7]. Among these, some possible common origin or founder mutations have been reported, that includes 7.6 kb deletion commonly originated in ~70% of French-Canadian population [8]; c.459+5A>G mutation in ~27% Argentinean patients [6]; c.571–1G>T mutation in ~80% Japanese patients [9] and p.E462 mutation in ~22% Indian patients [10]. In addition to this, the B1 variant form of TSD, characterized by the presence of a R178H mutation in *HEXA* gene [11] and two pseudodeficiency alleles c.739C>T (p.R247W) and c.745C>T (p.R249W) are also reported in ~35% non-Ashkenazi Jewish population [6].

Present study is a part of National Task Force by Government of India to identify burden of various lysosomal storage disorders, its geographical distribution and prevalence and mutations spectrum in various ethnicity in Indian population. Our earlier studies on 41 unrelated Indian patients carried out for almost 6 years in clinically suspected cases with TSD revealed that this disease though rare, seems to have significant burden in our population but remains unnoticed, under diagnosed due to infantile form [10, 12, 13]. As a result families planning for the next pregnancy often ends up having a similar child due to its autosomal recessive inheritance pattern and putting tremendous social, psychological, and financial burden. From our earlier study we have pointed towards the likely founder mutation in certain community in Gujarat and present study is an attempt to confirm our earlier observations and identify additional novel variants that are prevalent in other Indian population and report their clinical findings.

Materials and methods

This study was carried out as a part of multicentric task force project approved by Indian Council of Medical Research (ICMR) and Department of Health Research (DHR), Government of India from January 2015 to December 2017. The study protocol and enrollment of the patients was approved by institutional ethics committee and written informed consent was taken from the parents or guardian of all the participating subjects as per Helsinki declaration.

Patients

Thirty-four unrelated affected families by TSD with different geographical background were enrolled in this study

Table 1 Clinical phenotypes and demographic distribution of TSD patients

Age at the time of investigation (years) ($n = 31$) ^a	15.8 ± 5.45
Gender ($n = 31$) ^a	
Male	21 (67.74%)
Female	10 (32.26%)
Demographic distribution ($n = 34$)	
West	18 (52.94%)
North	4 (11.76%)
Central	1 (3.4%)
South	11 (32.35%)
Clinical findings ($n = 34$)	
Regression of physical and cognitive milestones	34 (100%)
Cherry red spot	32 (94.1%)
Generalized hypotonia	29 (85.3%)
Seizures	25 (73.5%)
Hyperacusis	27 (79.4%)
MRI study shows symmetrical abnormal hyperintense signal on T2-weighted images in bilateral basal ganglia with white matter changes and/or demyelination	20 (58.8%)
EEG shows focal abnormalities with spikes and sharp waves	11 (32.35%)

^aThree couples (where proband was not alive) were not included for analysis

from the period of January 2015–December 2017. The study comprises of 21 males and 10 females in the age range of 12–36 months and three carrier parents were investigated (where proband were not alive) having confirmed diagnosis of TSD. The diagnosis in these children was confirmed by the demonstration of reduced or absent activity of Hex-A with normal activity of total-Hex in the leukocytes. All patients included in the study presented with developmental delay since infancy, seizures, exaggerated startle, hyperacusis, muscle weakness, cherry red spot on the fundus and mild or absent hepatomegaly. Table 1 provides the demographic distribution and clinical details of the patients.

Measurement of β -hexosaminidase activity

The enzyme activity of β -hexosaminidase-A (Hex-A) was determined in leukocytes by fluorometric method using sulfated substrate 4-methylumbelliferyl-N-acetyl- β -D-glucosamine-6-sulfate (MUGS), while for carrier analysis Hex A% and total-hexosaminidase activities were measured in leukocytes by a standard heat-inactivation fluorometric method using artificial 4-methylumbelliferyl-N-acetyl- β -D-glucosamine (MUG) substrate [14]. Protein concentrations of the samples were measured using the Lowry method [15].

Mutation analysis of *HEXA* gene

Genomic DNA was isolated by salting out method [16]. The bi-directional sequencing of the *HEXA* gene was carried out by automated sequencer (ABI-3130 genetic analyzer) as previously reported [17]. The mutations identified were then looked up in public databases such as Human Genome Mutation Database (<http://www.hgmd.cf.ac.uk>), Genome aggregation database (<http://gnomad.broadinstitute.org/>), dbSNP database (<http://www.ncbi.nlm.nih.gov/SNP/index.html>), and McGill University database (<http://www.hexdb.mcgill.ca>).

In silico analysis

In silico prediction tools like PolyPhen2, SIFT, PROVEN, FATHMM, FATHMM-MKL, Mutation Taster, and Mutation Assessor to observe the functional effect of novel variants [18].

Structural studies

The structural study of novel missense mutant alleles was carried out using the crystallographic structure of Hex-A (PDB ID: 2GJX) as described in our earlier report [10] and the root mean square deviation (RMSD) of the mutant structures with respect to the wild-type structure was calculated [19].

Results

In the present study the clinical assessment followed by biochemical and molecular investigations confirmed the diagnosis of 34 TSD patients. The mean age at the time of diagnosis was 15.8 (± 5.45) months. Consanguinity was present in 10/34 (29.4%) families. The most common clinical indications observed in these patients were regression of milestones (cognitive and physical) (100%), cherry red spot on fundus (94.1%), generalized hypotonia (85.3%), hyperacusis (79.4%) followed by seizures (73.5%). The MRI study of the brain showed symmetrical abnormal hyperintense signal on T2-weighted images in bilateral basal ganglia with white matter changes ($n = 20$), with demyelination observed in three cases. Details of geographic background, age at diagnosis, and clinical findings are shown in Table 1.

Biochemical investigations revealed significant deficiency of Hex-A with normal total-Hex activity was observed in the leukocytes of all 31 patients and carrier detection was carried out in three couples (where proband was not alive). The deficient Hex-A enzyme activity is consistent with our previous study in infantile TSD patients [10, 12, 13].

Overall analysis by bidirectional Sanger sequencing has identified 25 mutant alleles in 34 affected/carrier (31 affected patients and 3 carrier couples) by TSD. Fifteen of these were novel found to be in 17 TSD patients with damaging effect on the protein function. This includes seven homozygous missense variants [p.V206L, p.Y213H, p.R252C, p.F257S, p.C328G, p.G454R, and p.P475R] one each; three patients were homozygous for novel nonsense variants that include [p.W420X] in two while p.S9X in one patient; three patients were compound heterozygous for novel nonsense variants [p.E91X, p.W420X and p.E482X] with known 4 bp insertion [c.1278insTATC], missense mutation [p.D322Y] and intronic mutation [c.459+4A>G], respectively. Two homozygous small deletions [c.1349delC (p.A450Vfs*3) and c.52delG (p.G18Dfs*82)] were observed in two patients, and two homozygous splice-site variants c.460-1G>A and c.347-1G>A was observed in one each, respectively (Table 2). Chromatograms for all novel variants are shown in Fig. 1.

Present study also identifies 10 previously known mutations in 17 TSD patients, that includes homozygous 4 bp insertion [c.1278insTATC (p.Y427IfsX5)] identified in five patients. Four patients harbor homozygous [p.E462V] mutation. Six patients carried homozygous missense mutations [p.M1T, p.R170Q, D322Y, p.D322N, and p.R499C] one each. One patient was observed having compound heterozygous mutation p.E462V and c.1278insTATC while one patient was compound heterozygous for nonsense variant p.Q106X with splice site variant c.1073+1G>A.

In silico prediction analyses of all novel variants of *HEXA* gene was found to be probably damaging with a deleterious effect on protein function. The predictions of the PolyPhen2, SIFT, PROVEN, FATHMM, FATHMM-MKL, Mutation Taster, and Mutation Assessor were in concordance with the observed pathogenicity in patients with all novel variants and the prediction scores are presented in Table 3. The root mean square deviation (RMSD) values for the modeled mutants were also found to be significant to prove the pathogenicity of all missense mutations. In addition to this, Genome aggregation database allele frequency in general, as well as in south Asian populations are also mentioned in Table 3. Functional study was not conducted due to unavailability of patients after identification of the novel variants. However, we have analysed 500 normal healthy unrelated individuals to identify the frequency of novel variants and none were found to be positive.

Discussion

Present study reports large series of children with TSD that were presented with infantile form of the disease, which is

Table 2 Biochemical and Molecular analysis in patients with Tay-Sachs disease

Patient ID	Hex-A activity (MUGS) (nmol/hr/mg) = (x)	Total Hex activity (MUG) = (y)	Hex A % = (x/y) × 100	Genotypes	Zygoty	Genome aggregation database frequency (allele frequency in South Asian population)	dbSNP reference sequence (HGMD accession number)	Reference
T1	6.1	2451	0.25	Ex12:c.1385A>T (p.E462V)	Homozygous	0.00000406 (0.00003251)	rs863225434 (CMI25495)	Mistri et al. [10] Sheth et al. [3, 12, 13]
T2	4.7	2049.7	0.23	Ex12:c.1385A>T (p.E462V)	Homozygous	0.00000406 (0.00003251)		
T3	1.39	ND	-	Ex12:c.1385A>T (p.E462V)	Homozygous	0.00000406 (0.00003251)		
T26	1.95	-	-	Ex12:c.1385A>T (p.E462V)	Homozygous	0.00000406 (0.00003251)		
T29	1.46	-	-	Ex12:c.1385A>T (p.E462V)	Compound Heterozygous	0.00000406 (0.00003251)	rs863225434 (CMI25495)	Mistri et al. [10]
T4	2.5	862	0.29	Ex11:c.1278insTATC (p.Y427IfsX5)	Homozygous	0.0008041 (0.0001299)	rs387906309 (CI880091)	Myerowitz and Costigan [4]
T19	0.59 ^a	NA	-	Ex11:c.1278insTATC (p.Y427IfsX5)	Homozygous	0.0008041 (0.0001299)		
T21	1.3 ^a	NA	-	Ex11:c.1278insTATC (p.Y427IfsX5)	Homozygous	0.0008041 (0.0001299)		
T34	1.47	-	-	Ex11:c.1278insTATC (p.Y427IfsX5)	Homozygous	0.0008041 (0.0001299)/-		
T25	1.0	-	-	NA	NA	0.0008041 (0.0001299)		
				Mother- Ex11: c.1278insTATC (p.Y427IfsX5)	Mother: Heterozygous			
				Father-Ex11: c.1278insTATC (p.Y427IfsX5)	Father: Heterozygous			
T23	1.43	432.0	-	Ex11:c.1278insTATC (p.Y427IfsX5)	Compound Heterozygous	0.0008041 (0.0001299)	rs387906309	Present study
T31	0.4	709.2	-	Ex2:c.271G>T (p.E91X)	Homozygous	NA	-	
				Ex11:c.1259G>A (p.W420X)	Homozygous	-	rs1057519465	
T18	1.7	424	0.4	Ex11:c.1259G>A (p.W420X)	Homozygous	-		
T5	6.7	1525	0.44	Ex8:c.964G>T (p.D322N)	Homozygous	0.00000406 (0.00003249)	rs772180415 (CMI25496)	Mistri et al. [10] Sheth et al. [3, 12]
T6	8.1	1301.5	0.62	In4:c.459+4A>C	Homozygous	-	rs1057519459 (CS1313566)	Sheth et al. [3, 12]
T7	9.0	1807	0.50	In4:c.460-1G>A	Homozygous	-	rs764343937	Present study
T8	12.5	2499.9	0.50	In9:c.1073+1G>A	Compound Heterozygous	0.0002114 (0.0000)	rs76173977 (CS910444)	Akli et al. [27]
				Ex2:c.316C>T (p.Q106X)	Heterozygous	0.00000406 (0.00003249)	rs773446161 (CM970715)	Akerman et al. [25]
T9	5.2	1000	0.52	Ex13:c.1495C>T (p.R499C)	Homozygous	0.00000406 (0.0000)	rs121907966 (CM920341)	Mules et al. [34]
T10	9.6	1682.2	0.57	Ex6:c.616G>C (p.V206L)	Homozygous	-	rs543071358	Present study
T11	7.5	1958.8	0.38	Ex6:c.637T>C (p.Y213H)	Homozygous	-	rs1057519461	

Table 2 (continued)

Patient ID	Hex-A activity (nmol/hr/mg) = (x)	Total Hex activity (MUGS) (MUG) = (y)	Hex A % = (x/y) × 100	Genotypes	Zygosity	Genome aggregation database frequency (allele frequency in South Asian population)	dbSNP reference sequence (HGMD accession number)	Reference
T12	5.6	1280.1	0.44	Ex1:c.2T>C (p.M1T)	Homozygous	–	rs786204721 (CM930389)	Harmon et al. [21]
T13	1.2	–	–	NA	NA	0.00003231 (0.0000)	rs121907957 (CM920330)	Nakano et al. [22]
	Hex-A% Mother: 59% Father: 62%			Mother-Ex5:c.509G>A (p.R170Q) Father-Ex5:c.509G>A (p.R170Q)	Mother: Heterozygous Father: Heterozygous			
T14	1.5	1502	0.10	Ex12:c.1360G>C (p.G454R) In2:c.347-1G>A	Homozygous	–	rs121907978	Present study
T15	2.3	–	–	NA	Homozygous	–	–	
T16	24.0 ^a	1481.0	1.6	NA	NA	–	rs1057519466	
				Mother –Ex12:c.1349delC (p. A450VfsX3) Father –Ex12:c.1349delC (p. A450VfsX3)	Mother: Heterozygous Father: Heterozygous			
T17	3.5	801.3	0.43	Ex8:c.964G>T (p.D322Y) Ex11:c.1259G>A (p.W420X) Ex13:c.1424C>G (p.P475R)	Compound Heterozygous	0.000008123 (0.00006497) NA	rs772180415 (CM125497) rs1057519465	Mistri et al. [10] Present study
T20	1.53	NA	–	Ex13:c.1424C>G (p.P475R)	Homozygous	0.000004063 (0.000)	rs758829206	
T22	2.5	752.4	0.33	Ex8:c.982T>G (p.C328G)	Homozygous	–	–	
T24	2.6	–	–	Ex7:c.770T>C (p.F257S)	Homozygous	–	–	
T27	0.5 ^a	–	–	Ex13 :c.1444G>T (p.E482X)	Homozygous	–	rs121907952	
T28	2.0	466	–	Ex7:c.754C>T (p.R252C)	Homozygous	0.00001218 (0.00006497)	rs566580738	
T30	1.0 ^a	–	–	Ex1:c.53delG (p.G18DfsX82)	Homozygous	–	–	
T32	0.74 ^a	–	–	Ex8:c.964G>T (p.D322Y)	Homozygous	0.000008123 (0.00006497)	rs772180415 (CM125497)	Mistri et al. [10]
T33	3.8	394.0	–	Ex1:c.26C>A (p.S9X)	Homozygous	–	rs769370282	Present study

^aEnzyme study carried out at other centers.

Normal range for Hex-A - 80 to 410 nmol/hr/mg and for total-Hex - 723 to 2700 nmol/hr/mg

in concordance with the well-defined clinical presentation of the disease. The Hex-A enzyme activity observed in all children was in the range of 0–1.6% with normal total-Hex activity which is consistent with previously observed enzyme activity in infantile TSD cases [20]. In three cases where proband was not available for analysis, parental study has shown the enzyme activity in the carrier range and subsequent molecular study has confirmed the diagnosis as infantile TSD in the index case.

In the present study 25 variants were identified that includes 15 (60%) novel and 10 (40%) previously known mutations. The missense mutations were observed in 18 (52.9%) followed by nonsense mutation in 7 (20.58%), splicing in 4 (11.76%), and frameshift mutation in 9 (26.47%) patients. The previously known deleterious Indian mutations (p.E462V, p.D322Y, p.D322N, and c.1278insTATC) were observed in 41.2% of patients in the present study.

Among 25 variants identified here, 10 have been earlier reported in diverse populations that include missense mutations p.D322Y, p.D322N, and p.E462V from the Indian patients [10, 12, 13]. Other mutations were; p.M1T from non-Jewish English patient [21], p.R170Q from Japanese [22], and Moroccan patients [23], p.R499C from Irish, Italian, and Japanese patients [7], four base pair insertion c.1278insTATC from Indian, Ashkenazi Jews, and non-Jewish patients [7, 10, 24], nonsense mutation p.Q106X from Hispanic-German-Jewish patient [24, 25], splice site mutation c.459+4A>G from Indian patients and c.1073+1G>A from Cajun, non-Jewish, and w. European patients [26–28]. The p.E462V mutation is again found in 5/34 (14.7%) patients that reconfirm the founder effect of this mutation in Indian Gujarati patients. In addition, c.1278insTATC mutation is also more commonly seen 7/34 (20.58%) in Indian patients with TSD.

In present study, we have also found seven novel missense variants that include (p.V206L, p.Y213H, p.R252C, p.F257S, p.C328G, p.G454R, and p.P475R). All these mutations were found to be affecting highly conserved domains and replace the amino acid residues among α -subunit of Hex-A [29]. In addition to this majority of novel variants detected in the present study were neither reported in Genome aggregation database nor 1000 Genome. Moreover, a structural protein analysis was performed using crystallographic structure of *HEXA* (PDB ID: 2 GJX) as template [30]. The analysis showed that variant allele p.V206L disrupts the β -sheet and over packing of residues due to an amino acid change from a relatively medium size chain group. Here, uncharged amino acid residue Valine is replaced by non-polar, large size chain group amino acid Leucine at codon 206 in exon-6 (Fig. 2a). Similarly, in p.Y213H variant, an aromatic, cyclic uncharged residues Tyrosine is replaced by a

positively charged Histidine residue and disrupts the β -sheet (Fig. 2b). Furthermore, all the variants identified so far in residues proximal to Val206 and Tyr213 (V200M, W203G, H204R, D207E, S210R, F211S) have been shown to cause severe TSD phenotypes [17, 28]. Novel missense variant p.R252C is highly likely to be disease-causing mutation because a different mutation in the same codon (R252H and R252L) were previously reported in a Portuguese [31] and Japanese patients [32], respectively, and the bioinformatics predictions scores support pathogenicity further. A novel variant p.R252C disrupts the β -helix with a loss of hydrogen bonding (Fig. 2c) while p.F257S also seems to be disrupting β -sheet and lost hydrogen bonding (Fig. 2d).

The p.C328G variant that likely disturbs the charge distribution near the active site and alters its configuration (Fig. 2e). The p.G454R disrupts the backbone and β -sheet causing conformational changes in the protein (Fig. 2f). This variant occurred at near proximal to the active residue α Glu462, hence likely to be altering the function of this active site. Other substitutions at the same codon (G454S and G454D) have been previously reported in infantile TSD from non-Ashkenazi [27] and Turkish [33] patients, respectively. Another variant p.P475R most likely changes the enzyme interaction properties by increasing the hydrophobic potential on the surface and also destabilizes the protein structure (Fig. 2g).

The two novel splice site variants (c.347-1G>A and c.460-1G>A) seems to be affecting mRNA splicing. The homozygous variants c.347-1G>A and c.460-1G>A found in one patient each were in concordance with the earlier report in American Black patient having mutation at the same position c.460-1G>T, causing infantile TSD [33, 34]. Both the novel splice site variants were predicted to be disease causing or damaging by mutation taster and FATHMM-MKL (Table 3). Nonetheless, numerous homozygote or compound heterozygotes splicing mutations have been reported in *HEXA* gene, resulting in a very low or undetectable mRNA levels during functional analysis. Both the splicing variants found in this study were likely associated with severe infantile phenotypes of the disease, as it has been previously reported [2, 7].

The novel nonsense variants p.R9X, p.E91X, p.W420X, and p.E482X that were detected in a homozygous or compound heterozygous state are highly likely to cause premature termination of the protein and mRNA reduction with a deleterious effect on protein function.

A novel 1 bp deletion (c.1349delC) was observed in exon-12 in an infantile TSD patient that causes the frameshift beginning at the 450th amino acid in the amino acid sequence of the resultant protein, which introduces a downstream stop codon at 453 amino acid (p.A450Vfs*3) which truncates the protein from its original length of 529

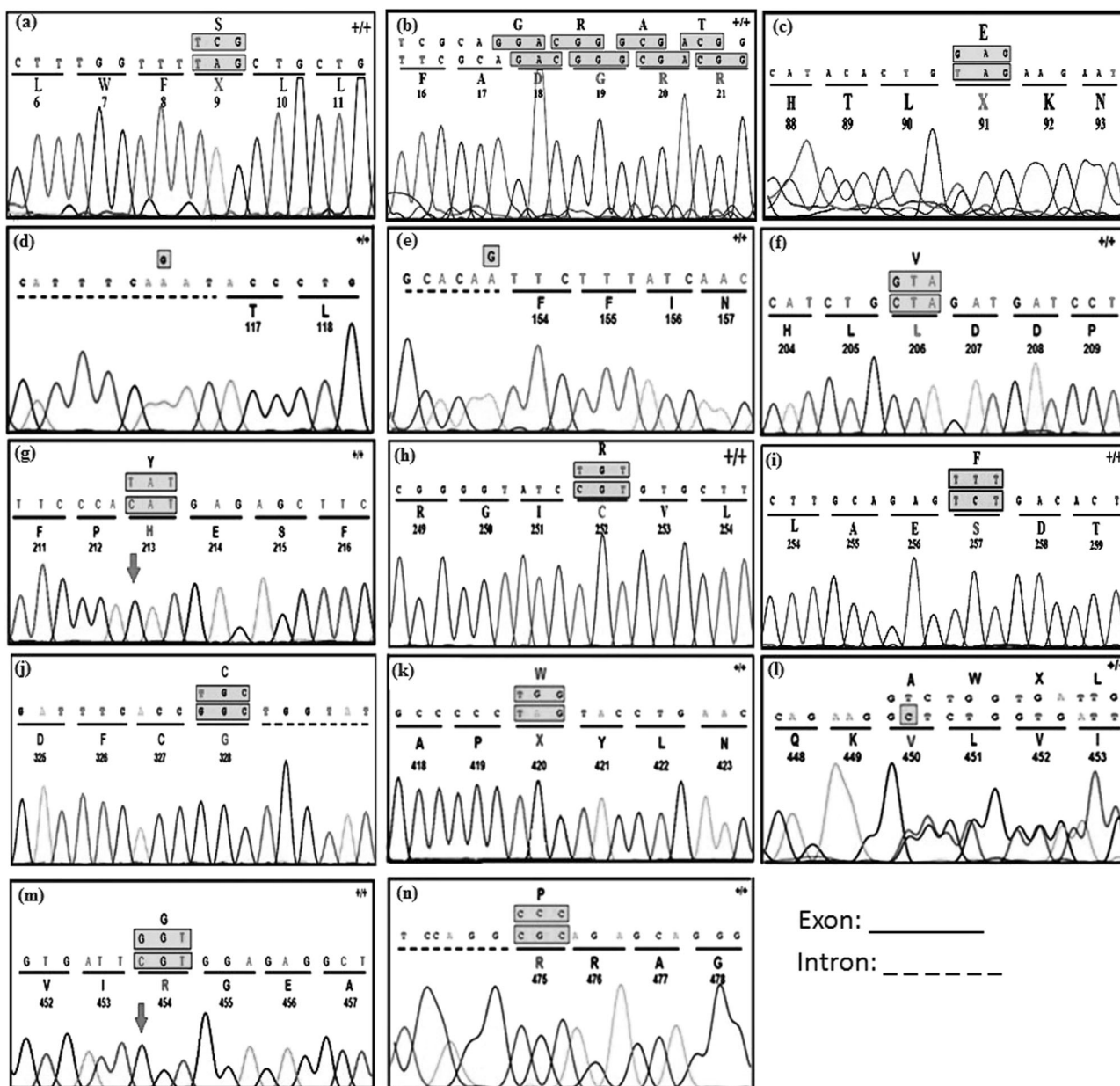


Fig. 1 Sequence chromatogram of novel variants identified in *HEXA* gene. **a** c.26C>A (p.S9X) (Homozygous). **b** c.52delG (p.G18Dfs*82) (homozygous). **c** c.271G>T (p.E91X) (heterozygous). **d** c.347-G>A (homozygous). **e** c.460-1G>A (homozygous). **f** c.616G>C (p.V206L) (homozygous). **g** c.637T>C (p.Y213H) (homozygous). **h** c.754C>T (p.R252C) (homozygous). **i** c.770T>C (p.F257S) (homozygous). **j** c.982T>G (p.C328G) (homozygous). **k** c.1259G>A (p.W420X) (homozygous). **l** c.1349delC (p.A450VfsX3) (homozygous). **m** c.1360G>C (p.G454R) (homozygous). **n** c.1424C>G (p.P475R) (homozygous)

amino acids. Another, one bp deletion c.52delG was observed in Exon-2 that changes the reading frame at amino acid 82 and introduces a downstream stop codon at amino acid 100 (p.G18Dfs*82) also leading to a protein truncation from its original protein length. Both novel 1 bp deletions were also evaluated for pathogenicity using Mutation Taster and was found to be disease causing as shown in Table 3. However, further functional study was not conducted due to unavailability of patients after identification of the novel variants. But all patients with novel variants had

significantly reduced activity of Hexosaminidase-A enzyme in the leukocytes which is likely to reflect the functional effect of the genotype.

A large number of mutations have been previously reported in the *HEXA* gene. Among them, the majority were found in the Hex α -subunit sequence that is generally found in the severe infantile TSD. We have not found any genotype of late-onset phenotype, B1 phenotype and pseudo-deficiency phenotype in the present study, as well as our earlier series reported from India. Nonetheless, there are

Table 3 In-silico predictions for novel variants detected in HEXA gene

Amino acid change	Mutation Taster score	Mutation assessor ^a	SIFT score	Polyphen2 Score (sensitivity, specificity)	PROVEN	FATHMM/FATHMM-MKL	RMSD between native and mutant structures	Structural effect
c.616G>C (p.Val206Leu)	Disease causing	High	0.02 (IT)	0.999 (0.14, 0.99) (PD)	-2.92 (D)	-3.76 (D)/0.9937 (D)	0.385 (DC)	Over packing of residues, likely affect active site and disrupt the beta sheet
c.657T>C (p.Tyr213His)	Disease causing	High	0.03 (IT)	1.00 (0,1) (PD)	-4.81 (D)	-3.76 (D)/0.9943 (D)	0.175 (DC)	Disruption of beta sheet and affect active site
c.754C>T (p.Arg252Cys)	Disease causing	High	0.00 (IT)	1.00 (0,1) (PD)	-7.8 (D)	-4.24 (D)/0.9356 (D)	0.155 (DC)	Disrupt the beta sheet and lost hydrogen bonding
c.770T>C (p.Phe257Ser)	Disease causing	High	0.00 (IT)	1.00 (0,1) (PD)	-7.8 (D)	-3.79 (D)/0.9948 (D)	0.193 (DC)	Disrupt the β-helix and lost hydrogen bonding
c.982T>G (p.Cys328Gly)	Disease causing	High	0.04 (IT)	1.00 (0,1) (PD)	-11.72 (D)	-3.78 (D)/0.9961 (D)	0.285 (DC)	Disrupt the charge distribution near the active site and altering its configuration
c.1360G>C (p.Gly454Arg)	Disease causing	High	0.00 (IT)	1.00 (0,1) (PD)	-7.57 (D)	-7.3 (D)/0.9943 (D)	0.377 (DC)	Over packing, backbone strain and disruption of beta sheet
c.1424C>G (p.Pro475Arg)	Disease causing	High	0.00 (IT)	1.00 (0,1) (PD)	-8.36 (D)	-4.32 (D)/0.9932 (D)	0.282 (DC)	Increase hydrophobic potential on the surface which alters the enzyme interaction properties
c.26C>A (p.S9X)	Disease causing	NA	NA	NA	NA	NA	NA	Protein truncation
c.271G>T (p.E91X)	Disease causing	NA	NA	NA	NA	NA	NA	Protein truncation
c.1259G>A (p.W420X)	Disease causing	NA	NA	NA	NA	NA	NA	Protein truncation
c.1444G>T (p.W482X)	Disease causing	NA	NA	NA	NA	NA	NA	Protein truncation
c.347-1G>A	Disease causing	NA	NA	NA	NA	NA/0.9833 (D)	NA	NA
c.460-1G>A	Disease causing	NA	NA	NA	NA	NA/0.9969 (D)	NA	NA
c.1349delC p.A450Vfs*3	Disease causing	NA	NA	NA	NA	NA	NA	Protein truncation
c.52delG p.G18Dfs*82	Disease causing	NA	NA	NA	NA	NA	NA	Protein truncation

FATHMM the functional analysis through hidden Markov models, SIFT the sorting intolerant from tolerant, PolyPhen-2 polymorphism phenotyping version 2, PROVEN protein variation effect analyzer, RMSD root-mean-square deviation, DC disease causing, IT in tolerance, PD probably damaging, D damaging, NA not applicable

^aMutation Assessor suggests impact of amino acid substitution on protein function

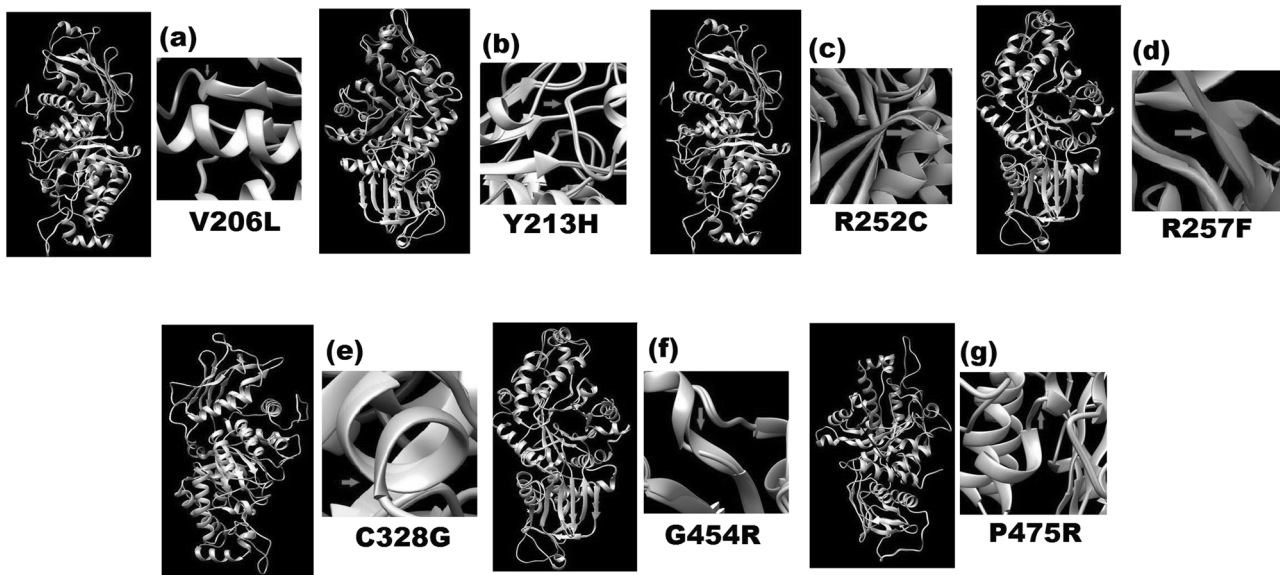


Fig. 2 Superimposed native structures (blue) and mutant structure (brown) of the α subunit produced using UCSF Chimera. **a** p.V206L creates overpacking of residues and disruption of β -sheet. **b** p.Y213H creates overpacking of residues and disruption of β -sheet. **c** p.R252C disrupt the β -helix and lost hydrogen bonding. **d** p.F257S disrupt the β -Sheet and lost hydrogen bonding. **e** p.C328G disrupt the charge distribution near the active site and altering its configuration. **f** p.G454R creates overpacking, backbone and β -sheet disruption and **g** p.P475R alter the enzyme interaction properties

mutations that have been reported in the B1 variants in TSD patients [6].

Conclusion

Present study identifies 15 novel variants that are prevalent in Indian infantile TSD patients which provide newer insight into the molecular pathology of the TSD. Combining present study and our earlier studies, we have observed that 67% genotypes found in Indian TSD patients are novel, which are associated with severe infantile phenotypes, while rest 33% genotypes found in our cohort were previously reported in various populations. Additionally, presence of p.E462V, p.D322Y, and c.1278insTATC variant are the most common hot spot mutations for TSD patients from India, which can be utilized as the first tier molecular screening test for carrier analysis of the large population and to develop therapeutic targets in future.

Acknowledgements We are thankful to the coordinators and members of National Task Force on Lysosomal storage disorders: Dr. Roli Mathur, Dr. Babbanjee, Dr. Seema Kapoor, Dr. Ashwin Dalal, Dr. Subha Phadake and, other clinician: Dr. Chaitanya Datar, Dr. Anil Jalan, Dr. Siddharth Shah, Dr. Raju C. Shah, Dr. Aniket Parashar for referring patients to our center. We are thankful to the patients and their families to participate in this study.

Author contributions Designing of the study, optimization of technical procedure, and preparation of manuscript: MM, FS, and JS. Processing of sample for enzymes study and molecular analysis: MM. Clinical analysis: SM, MK, DS, NG, MK, SN, RP, GK and SH.

Critical revision and approval of the manuscript: JS, FS, and MM. All the authors read and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000 (5).

Informed consent Informed written consent was taken from the parents or guardians of all the participating patients.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Gravel R, Koback M, Proia R, Sandhoff K, Suzuki K, et al. The GM2 gangliosidosis. In: Scriver CR, Beaudet AL, Ly WS, Valle D, editors. The metabolic and molecular basis of inherited diseases. vol. 3, VIIth ed. New York: McGraw-Hill; 2001. p. 3827–76.
2. Sonnino S, Chigorno V. Ganglioside molecular species containing C18- and C20-sphingosine in mammalian nervous tissues and neuronal cell cultures. *Biochim Biophys Acta* 2000;1469:63–77.
3. Sheth J, Mistri M, Sheth F, Shah R, Bavdekar A, Godbole K, et al. Burden of lysosomal storage disorders in India: experience of 387 affected children from a single diagnostic facility. *JIMD Rep* 2014;12:51–63.

4. Myerowitz R, Costigan F. The major defect in Ashkenazi Jews with Tay–Sachs disease is an insertion in the gene for alpha-chain of beta-hexosaminidase. *J Biol Chem* 1988;263:18587–9.
5. Kaback M, Lim-Steele J, Dabholkar D, Brown D, Levy N, Zeiger K. Tay–Sachs disease carrier screening, prenatal diagnosis and the molecular era. An international perspective, 1970–93. The International TSD data collection Network. *JAMA* 1993;270:2307–15.
6. Zampieri S, Montalvo A, Blanco M, Zanin I, Amartino H, Vlahovicek K, et al. Molecular analysis of HEXA gene in Argentinian patients affected with Tay–Sachs disease: possible common origin of the prevalent c.459+5A>G mutation. *Gene* 2012;499:262–5.
7. Mahuran D. Biochemical consequences of mutations causing the GM2 gangliosidosis. *Biochem Biophys Acta* 1999;1455:105–38.
8. Myerowitz R, Hogikyan ND. Different mutations in Ashkenazi Jewish and non-Jewish French Canadians with Tay–Sachs disease. *Science* 1986;232:1646–8.
9. Tanaka A, Fujimaru M, Choeh K, Isshiki G. Novel mutations, including the second most common in Japan, in the beta-hexosaminidase alpha subunit gene, and a simple screening of Japanese patients with Tay–Sachs disease. *J Hum Genet* 1999;44:91–5.
10. Mistri M, Tamhankar P, Sheth F, Sanghavi D, Kondurkar P, Patil S, et al. Identification of novel mutations in HEXA Gene in children affected with Tay–Sachs disease from India. *PLoS One* 2012;7:e39122.
11. Ohno K, Suzuki K. Mutation in GM2-gangliosidosis B1 variant. *J Neurochem* 1988;50:316–8.
12. Sheth J, Mistri M, Datar C, Kalane U, Patil S, Kamate M, et al. Expanding the spectrum of HEXA mutations in Indian patients with Tay–Sachs disease. *Mol Genet Metab Rep* 2014b;1:425–30.
13. Sheth J, Mistri M, Mahadevan L, Mehta S, Solanki D, Kamate M, et al. Identification of deletion–duplication in HEXA gene in five children with Tay–Sachs disease from India. *BMC Med Genet* 2018;19:109.
14. Wendeler M, Sandhoff K. Hexosaminidase assays. *Glycoconj J* 2009;26:945–52.
15. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. *J Biol Chem* 1951;193:265–75.
16. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16:1215.
17. Triggs Raine B, Akerman B, Clarke J, Gravel R. Sequencing of DNA flanking the exons of the HEXA gene, and identification of a mutation in Tay–Sachs disease. *Am J Hum Genet* 1991;49:1041–54.
18. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* 2015;31:761–3.
19. Ohno K, Saito S, Sugawara K, Sakuraba H. Structural consequences of amino acid substitutions causing Tay–Sachs disease. *Mol Genet Metab* 2008;94:462–8.
20. Kaback M, Desnick R. Hexosaminidase A deficiency. In: Pagon R, Bird T, Dolan C, Stephens K, editors. 1999. <http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene&part=tay-sachs>. Accessed 29 July 2019.
21. Harmon DL, Gardner-Medwin D, Stirling JL. Two new mutations in a late infantile Tay–Sachs patient are both in exon 1 of the beta-hexosaminidase alpha subunit gene. *J Med Genet* 1993;30:123–8.
22. Nakano T, Nanba E, Tanaka A, Ohno K, Suzuki Y, Suzuki K. A new point mutation within exon 5 of beta-hexosaminidase alpha gene in a Japanese infant with Tay–Sachs disease. *Ann Neurol* 1990;27:465–73.
23. Drucker L, Proia R, Navon R. Identification and rapid detection of three Tay–Sachs mutations in the Moroccan Jewish population. *Am J Hum Genet* 1992;51:371–7.
24. Tanaka A, Ohno K, Sandhoff K, Maire I, Kolodny EH, Brown A, et al. GM2-gangliosidosis B1 variant: analysis of beta-hexosaminidase alpha gene abnormalities in seven patients. *Am J Hum Genet* 1990;46:329–39.
25. Akerman B, Natowicz M, Kaback M, Loyer M, Campeau E, Gravel R. Novel mutations and DNA-based screening in non-Jewish carriers of Tay–Sachs disease. *Am J Hum Genet* 1997;60:1099–106.
26. McDowell GA, Mules EH, Fabacher P, Shapira E, Blitzer MG. The presence of two different infantile Tay–Sachs disease mutations in a Cajun population. *Am J Hum Genet* 1992;51:1071–7.
27. Akli S, Boue J, Sandhoff K, Kleijer W, Vamos E, Young E, et al. Collaborative study of the molecular epidemiology of Tay–Sachs disease in Europe. *Eur J Hum Genet* 1993;1:229–38.
28. Landels E, Green P, Ellis I, Fensom A, Kaback M, Lim-Steele J, et al. Further investigation of the HEXA gene intron 9 donor splice site mutation frequently found in non-Jewish Tay–Sachs disease patients from the British Isles. *J Med Genet* 1993;30:479–81.
29. Maier T, Strater N, Schuette C, Klingenstein R, Sandhoff K, Saenger W. The X-ray crystal structure of human beta-hexosaminidase B provides new insights into Sandhoff disease. *J Mol Biol* 2003;328:669–81.
30. Lemieux M, Mark B, Cherney M. Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay–Sachs mutations and loss of GM2 ganglioside hydrolysis. *J Mol Biol* 2006;359:913–29.
31. Riberio MG, Sonin T, Pinto RA, Fontes A, Ribeiro H, Pinto E, et al. Clinical, enzymatic, and molecular characterization of a Portuguese family with a chronic form of GM-2 gangliosidosis B1 variant. *J Med Genet* 1996;33:341–3.
32. Tanaka A, Hoang LT, Nishi Y, Maniwa S, Oka M, Yamano T. Different attenuated phenotypes of GM2 gangliosidosis variant B in Japanese patients with HEXA mutations at codon 499, and five novel mutations responsible for infantile acute form. *J Hum Genet* 2003;48:571–4.
33. Ozkara H, Navon R. At least six different mutations in HEXA gene cause Tay–Sachs disease among the Turkish population. *Mol Genet Metab* 1998;65:250–3.
34. Mules EH, Hayflick S, Mille CS, Reynolds LW, Thomas GH. Six novel deleterious and three neutral mutations in the gene encoding the alphasubunit of hexosaminidase A in non-Jewish individuals. *Am J Hum Genet* 1992;50:834–41.