Front. Biol. 2014, 9(5): 410–421
DOI 10.1007/s11515-014-1314-2

RESEARCH ARTICLE

Computational screening of disease associated mutations on NPC1 gene and its structural consequence in Niemann-Pick type-C1

Naresh KANDAKATLA (✉)[1], Geetha RAMAKRISHNAN[1], Rajasekhar CHEKKARA[1], Namachivayam BALAKRISHNAN[2]

[1] *Department of Chemistry, Sathayabama University, Jeppiaar Nagar, Chennai-600119, India*
[2] *St. Joseph's College, Bharathidasan University, Tiruchirappalli, 620002, India*

Abstract **Niemann-Pick disease type C1 (NPC1), caused by mutations of NPC1 gene, is an inherited lysosomal lipid storage disorder. Loss of functional NPC1 causes the accumulation of free cholesterol (FC) in endocytic organelles that comprised the characteristics of late endosomes and/or lysosomes. In this study we analyzed the pathogenic effect of 103 nsSNPs reported in NPC1 using computational methods. R1186C, S940L, R958Q and I1061T mutations were predicted as most deleterious and disease associated with NPC1 using SIFT, Polyphen 2.0, PANTHER, PhD-SNP, Pmut and MUTPred tools which were also endorsed with previous *in vivo* experimental studies. To understand the atomic arrangement in 3D space, the native and disease associated mutant (R1186C, S940L, R958Q and I1061T) structures were modeled. Quantitative structural and flexibility analysis was conceded to observe the structural consequence of prioritized disease associated mutations (R1186C, S940L, R958Q and I1061T). Accessible surface area (ASA), free folding energy (FFE) and hydrogen bond (NH bond) showed more flexibility in 3D space in mutant structures. Based on the quantitative assessment and flexibility analysis of NPC1 variants, I1061T showed the most deleterious effect. Our analysis provides a clear clue to wet laboratory scientists to understand the structural and functional effect of NPC1 gene upon mutation.**

Keywords Niemann-Pick disease type C1, SNPs, gene mutation

Introduction

Niemann-Pick disease type C (NPC) (OMIM 257220) is an autosomal recessive genetic disorder that causes an abnormal accumulation of cholesterol and other lipids in many cell types (Morris et al., 1999; Beltroy et al., 2005). The incidence is estimated between 1:120000 live births (Garver and Heidenreich, 2002). NPC is caused by mutations of either the NPC1 (95% of families) or NPC2 genes in humans, although the precise mechanisms of action of these proteins are still under investigation. The identification of mutations within the NPC1 gene is challenging. To date, 252 mutations of NPC1 and 18 mutations of NPC2 gene have been reported at NPC database (http://npc.fzk.de/) (Runz et al., 2008). The NPC1 gene, mapped to chromosome 18q11-q12, spans 56 kb and contains 25 exons (Tamura et al., 2006; Xiong et al., 2012). NPC1 is a multispan membrane protein that is typically associated with late endosomes or lysosomes (Vanier and Suzuki, 1998), degradative organelles which hydrolyzed the cholesteryl esters brought into the cell through lipoproteins (Goldstein and Brown, 1992; Garver and Heidenreich, 2002). NPC1 has a sterol-sensing transmembrane domain which is similar to endoplasmic reticulum proteins that respond to alters in cellular cholesterol (Carstea et al., 1997). The NPC1 protein assists the transbilayer transport of some hydrophobic molecules, but it does not appear to transport cholesterol directly (Puri et al., 1999; Liscum, 2000; Ioannou, 2000; Scott and Ioannou, 2004). Loss of functional NPC1 or NPC2 causes the accumulation of free cholesterol (FC) in endocytic organelles that comprised

Received March 5, 2014; accepted April 28, 2014
Correspondence: Naresh KANDAKATLA
E-mail: nareshkandakatla7@gmail.com

the characteristics of late endosomes and/or lysosomes. These abnormal organelles will be referred to here as lysosome like storage organelles (LSOs) (Pipalia et al., 2006). The LSOs that are associated with NPC are quite similar to the LSOs associated with other hereditary glycosphingolipid storage disorders (often caused by the inability to metabolize a particular lipid). The storage organelles contain multi layered internal whorls of membrane bilayers which contain cholesterol, sphingomyelin, and high amounts of bis-(monoacylglycero)-phosphate (BMP), also known as lysobisphosphatidic acid (LBPA) (Kobayashi et al., 1999; Sun et al., 2001). The clinical features of NPC are vertical supranuclear gaze palsy, cerebellar, ataxia, dysarthria, dysphagia, progressive dementia, cataplexy, seizures and dystonia (Garver et al., 2007, 2010) which severely affect the patient's development and life quality.

Prediction of the disease causing nsSNPs using computational approach has become a eminent methodology. Several research articles state its effectiveness in identifying the deleterious, disease associated mutations, thus predicting the pathogenic phenotypic alleles in correlation to its functional and structural damaging properties (Carvalho et al., 2007, 2009; Goldgar et al., 2004; Karchin, 2009; Kumarand Purohit, 2012a, 2012b, 2012c; Kumar et al., 2012, 2013a, 2013b; Balu and Purohit, 2013; Kamaraj and Purohit, 2013, 2014). An in-depth study of the genetic mutations and their molecular basis of invoking disease related pathways has become an important insight of genomic research (Steward et al., 2003; Mooney, 2005; Ng and Henikoff, 2006; Kumar et al., 2014). The future of SNP analysis significantly lies in the development of personalized medicines that can aid the treatment of genomic variations induced disorders at a higher extent.

The main of objective of this study is to identity the most deleterious and disease-associated nsSNPs in NPC1 gene and their structural effect at molecular level. In this analysis we used SIFT, Polyphen 2.0, PANTHER, PhD-SNP, Pmut and MUTPred tools to ranked the most deleterious and disease associated nsSNPs (non-synonymous) from the available SNP data sets which obtained from dbSNP database. The disease-associated mutations showed high damage in structure and leads to affect the function of the NPC1 protein. Conformational changes in the 3D structure of the protein account for its time dependent physiologic affinities and various biochemical pathway alterations (Purohit and Sethumadhavan, 2009; Purohit et al., 2011a; 2011b; Rajendran et al., 2012; Balu et al., 2013; Keener et al., 2014; Kumar and Purohit, 2014). I-TASSER (Zhang, 2008), a threading based approach was used to model native and mutant NPC1 protein structures. PROCHECK (Laskowski et al., 1996) and PROSA (Wiederstein and Sippl, 2007) program was applied to evaluate the model of native and mutant NPC1 structures. We further applied quantitative assessment and flexibility analysis to observe the structural consequence of NPC1 proteins upon mutation.

Materials and methods

Data set

The data on human NPC1 genes were collected from OMIM (Amberger et al., 2009) and Entrez Gene on National Center for Biological Information (NCBI) Website. The SNP information of NPC1 genes were obtained from dbSNP (http://www.ncbi.nlm.nih.gov/snp/) database (Sherry et al., 2001). The aminoacid sequence of NPC1 protein was retrieved from the Uniprot database (Uniprot ID: O15118).

Disease associated SNP prediction

The single nucleotide polymorphism occurring in the protein coding region may lead to the deleterious and affect its 3D structure, this phenomenon may lead to disease-associated. Here we used SIFT (http://sift.jcvi.org/) (Kumar et al., 2009), PolyPhen 2.0 (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al., 2010), PANTHER (http://www.pantherdb.org/) (Thomas et al., 2003), PhD-SNP (http://snps.biofold.org/phd-snp/) (Capriotti et al., 2006), Pmut (http://mmb2.pcb.ub.es:8080/PMut/) (Ferrer-Costa et al., 2005) and MutPred (http://mutpred.mutdb.org/) (Li et al., 2009) tools in order to examine the disease-associated nsSNP occurring in the NPC1 protein coding region. SIFT uses sequence homology-based approach to classify amino acid substitutions (Kumar et al., 2009). The prediction score < 0.05 is considered to be deleterious. PolyPhen 2.0 is based on combination of sequence and structure based attributes and uses naive Bayesian classifier for the identification of an amino acid substitution and the impact of mutation. The output levels of probably damaging and possibly damaging were classified as functionally significant ($\leqslant 0.5$) and the benign level were classified as tolerated ($\geqslant 0.51$) (Adzhubei et al., 2010). PANTHER program which is a protein family and subfamily database which predicts the frequency of occurrence of amino acid at a particular position in evolutionary related protein sequences. The threshold subPSEC score of -3 has been assigned below which the predictions are considered as deleterious (Thomas et al., 2003). We filtered the nsSNPs that were combinedly predicted to be deleterious and damaging from these three servers. Further we used PhD-SNP, Pmut and MutPred tools to examine the disease-associated nsSNPs. PhD-SNP is SVM based classifier, trained over the million amino acid polymorphism data sets using supervised training algorithm (Capriotti et al., 2006). It predicts whether the given amino acid substitution leads to disease associated or neutral along with the reliability index score (Capriotti et al., 2006). Pmut is a neural network based program which is trained on large database of neutral and pathological mutations (Ferrer-Costa et al., 2005). Pmut uses 3 parameters including mutation descriptors, Solvent accessibility and Residue and sequence properties to calculate the pathogenecity indexes of given input mutation data ranging from 0 to 1.

The mutations with index score > 0.5 is predicted to be pathologically significant (Ferrer-Costa et al., 2005). MutPred is a web based tool, used to predict the molecular cause of disease related amino acid substitution (Li et al., 2009). It uses SIFT, PSI-BLAST and Pfam profiles along with some structural disorder prediction algorithms, including TMHMM, MARCOIL, I Mutant 2.0, B-factor prediction, and DisProt (Li et al., 2009). Functional analysis includes the prediction of DNA binding site, catalytic domains, calmodulin binding targets, and posttranslational modification sites (Li et al., 2009). Combining the scores of all three servers, the accuracy of prediction rises to a greater extent and finally we filtered the most disease associated mutation.

Modeling of NPC1 protein

According to the annotated information available in UNIPROT entry – O15118, the predicted deleterious mutation sites of NPC1 protein was observed in the region of 920–1200. Hence, this protein segment consists of 281 amino acid residues was modeled by I-TASSER server for 3D structure prediction based on threading approach (Zhang, 2008). This program works by combining the folds and secondary structure by profile-profile alignment threading techniques for non-aligned regions. For the submitted sequences, five 3D models were obtained and the best model was selected based on the lowest energy. Further the native structure was mutated with the most deleterious substitution predicted in this study. To build the mutant structures, we made a point mutation in native NPC1 protein at R1186C (arginine to cysteine), S940L (serine to leucine) R958Q (arginine to glutamine) and I1061T (isoleucine to threonine) using SPDB viewer (Kaplan and Littlejohn, 2001). These structures were energetically optimized by GROMACS package 4.5.5 (Hess et al., 2008). During energy minimization both native and mutant structures were solvated in a cubic box with simple point charge (SPC) water molecules at 10Å marginal radius. Initially the solvent molecules were relaxed while all the solute atoms were harmonically restrained to their original positions with a force constant of 100 kcal/mol for 5000 steps. After this, whole molecular system was subjected to energy minimization for 5000 iterations by steepest descent algorithm implementing all atom OPLS force field.

Quantitative assessment and flexibility analysis

VADAR (Volume Area Dihedral Angle Reporter) is a comprehensive web server for quantitative protein structure evaluation. It accepts 3D coordinates of protein as input and calculates, key structural parameters both for individual residues and for the entire protein. These derived parameters can be used to rapidly identify both general and residue-specific problems within newly determined protein structures (Willard et al., 2003). The VADAR web server is accessible at http://redpoll.pharmacy.ualberta.ca/vadar. 3D coordinates of native and mutant NPC1 structures were given as input to the server for quantitative analysis. Normal mode analysis (NMA) is a powerful tool for predicting the possible movements of a given macromolecule. It has been shown recently that half of the known protein movements can be modeled by using two low-frequency normal modes. NMA provides an alternative to molecular dynamics for the study of motions of macromolecules. A quantitative measure of the atomic motions in proteins can be obtained from the mean square fluctuations of the atoms relative to their average positions. Understanding structural dynamics of proteins is essential for gaining greater insights into their biological functions (Kumar and Purohit, 2013; Kumaret al., 2013; Purohit, 2014; Rajendran and Sethumadhavan, 2014). Since protein flexibility is important for protein function and for rational drug design. Therefore flexibility of certain amino acids in protein is useful for various types of interactions which can be analyzed by B factor which are computed from the mean square displacement. We used WEBnm (http://www.bioinfo.no/tools/normalmodes) (Hollup et al., 2005) to calculate the slowest modes of native and mutant NPC1 proteins. 3D coordinates of modeled native and mutants NPC1 structures were given as input to the server.

Results and discussion

Prediction of deleterious nsSNPs using SIFT, PolyPhen 2.0, PANTHER program

Collection of SNP data on the NPC1 gene was retrieved from dbSNP database for our investigation. The SNPs in regulatory regions and coding non-synonymous regions were selected for our studies. Total 103 nsSNPs were computationally analyzed and to identify the deleterious and damaging nsSNPs in NPC1 gene. The SIFT server was used to calculate the tolerance index of all 103 collected nsSNPs by evolutionary conservation analysis. A SIFT score value of < 0.05 was considered to be deleterious. Out of 103 input polymorphic data set, 35 mutations (Y634C, N490T, P471L, L737V, R1186H, R1272C, V780M, F842L, R1183H, A1018T, R1186C, V1141G, S652R, S940L, R161W, R404Q, S1169I, A851T, C113R, R958Q, L1213F, V706A, I1061T, A1054T, P1007A, Q775P, C177Y, P434S, S1200G, R978C, A1035V, Y1088C, N1156S, T1036M and T511M) were predicted to be deleterious with tolerance index ⩽ 0.05 (Table 1). Among these, 18 mutations Y634C, P471L, R1186H, F842L, R1183H, R1186C, V1141G, S652R, R161W, R404Q, S1169I, C113R, R958Q, L1213F, I1061T, Q775P, S1200G, Y1088C, N1156S and T1036M were reported to highly deleterious with SIFT score of 0.00 (Table 1). PolyPhen2.0 server used to predict the possible impact of an amino acid substitution on the structure and function of a protein. Based on polyphen score, 49 nsSNPs were found to be "damaging" (0.5 to 1.000) to protein structure and function

and the remaining 53 nsSNPs were characterized as benign. Among these 49 deleterious nsSNPs, out of which 15 mutations (Y634C, P471L, R1186H, F842L, R1186C, R404Q, S1169I, C113R, R958Q, L1213F, I1061T, Q775P, Y1088C, N1156S and T1036M) were report to be highly deleterious with Polyphen score of 1.000 and all the scores are listed in Table 1. Total of 34 mutations were identified as deleterious in both SIFT and PolyPhen server (Table 1) and it signified that good correlation was observed between these two servers. To further validate these results we carried HMM based statistical prediction method to identify the functionally significant point mutations using PANTHER server. The mutations with subPSEC score less than −3 has been reported to be probably deleterious. 78 mutations with subPSEC score less than or equal to −3 were characterized to be deleterious. Two mutations R958Q and I1061T were predicted as highly deleterious with subPSEC score −8.09327 and −8.53429 respectively. Since the result of PANTHER entirely depends on MSA profile, hence the priority for prediction shall be given to SIFT and PolyPhen scores. We filtered 32 mutations Y634C, N490T, P471L, L737V, R1186H, R1272C, V780M, F842L, R1183H, A1018T, R1186C, V1141G, S652R, S940L, R404Q, S1169I, A851T, C113R, R958Q, L1213F, V706A, I1061T, A1054T, P1007A, Q775P, S1200G, R978C, A1035V, Y1088C, N1156S, T1036M and T511M were commonly predicted to be deleterious mutants and showed good correlations with SIFT, PolyPhen 2.0 and PANTHER servers, which were highlighted as bold in Table 1.

Prediction of disease-associated nsSNPs

We applied Support Vector machine based PhD-SNP tool to further classify the predicted deleterious nsSNP's as disease associated. Total 32 nsSNPs were commonly predicted in SIFT, PolyPhen 2.0 and PANTHER tools (32.96%) were used in PhD-SNP server. Prediction carried out by PhD-SNP depends on intensive supervised training for over million amino acid polymorphism data sets and hence the prediction efficiency is extremely higher.

Out of 32 nsSNPs, 25 mutations Y634C, N490T, P471L, R1186H, R1272C, F842L, R1183H, A1018T, R1186C, V1141G, S652R, S940L, R404Q, S1169I, C113R, R958Q, L1213F, I1061T, A1054T, P1007A, Q775P, A1035V, Y1088C, N1156S and T1036M were predicted to be disease related (84%) and listed in Table 2. To verify this prediction, we further employed artificial neural network (ANN) based Pmut tool and these 32 nsSNPs were also submitted as input to Pmut tool. Out of 32 nsSNPs, 21 mutations were showed pathogenecity and remaining 11 nsSNPs showed as neutral and listed in Table 2.

Seven mutations (V1141G, S652R, R958Q, I1061T, R978C and Y1088C) were showed high pathological phenotype with pathogenicity index greater than 0.8. Two mutations R958Q and I1061T were predicted as highly disease associated with pathogenecity index 0.8995 and 0.9486 respectively (Table 2). Mutpred tool was used to predict the SNP disease-association probability and probable change in the molecular mechanism in the mutant. We found 12 disease associated mutations (16%) of which 4 mutations R1186C, S940L, 958Q and I1061T were commonly showed by the three (PhD-SNP, Pmut and MUTPred) servers (Table 2). Mutations R1186C, S940L, R958Q and I1061 were found to be highly deleterious with actionable, confident and highly confident hypothesis respectively. Finally we screened four most disease associated mutations R1186C, S940L, R958Q and I1061T which were highlighted as bold in Table 2. Our prediction also endorsed with experimental evidences (Greer et al., 1999; Millat et al., 1999; Yamamoto et al., 2000; Millat et al., 2001; Ribeiro et al., 2001; Sun et al., 2001; Bauer et al., 2002; Tarugi et al., 2002; Fernandez-Valero et al., 2005; Millat et al., 2005).

Molecular modeling

The human NPC1 protein (domain region between 920–1200) was modeled by I-TASSER program based on threading approach. I-TASSER used more than ten templates to model the protein. The top most template (PDB ID: 1b3u) was covered 97% of the NPC1 protein query sequence. The best structure with high confidence score was collected and used for further analysis. The deleterious mutations R1186C, S940L, R958Q and I1061T can possibly alter the native conformation of the NPC1 protein. Hence we made a point mutation in native NPC1 protein at the position of R1186C (arginine to cysteine), S940L (serine to leucine), 958 (arginine to glutamine) and 1061 (isoleucine to threonine) to build the mutant structures. The quality of the modeled structure of native and mutant NPC1 protein was evaluated independently by the PROCHECK (Millat et al., 1996) and PROSA (Wiederstein and Sippl, 2007) programs, which shows good stereo chemical properties of the modeled proteins. In native NPC1, 94.7% of residues in most favored and allowed regions and z-score value of −2.04. Mutant R1186C showed 92.4% of residues in most favored and allowed region, and z-score value of −1.94. Mutant S940L showed 91.6% of residues in most favored and allowed region and z-score value of −1.81. Mutant R958Q showed 87.6% of residues in most favored and allowed region z-score value of −1.96. Mutant I1061T showed 97.6% of residues in most favored ans allowed region and z-score value of −6.2. The overall G-factors of native and mutant NPC1 structures (acceptable between 0 and 0.5) were produced by PROCHECK in the range of 0.09–0.38. These scores implicate high confidence level and the structures were selected for further quantitative and flexibility analysis.

Quantitative assessment and flexibility analysis

VADAR (Willard et al., 2003) was used to evaluate different

Table 1 Summary of nsSNPs that were analyzed by three computational methods SIFT (Tolerated/Deleterious), PolyPhen 2.0 (Benign/Damaging) and PANTHER (Tolerated/Deleterious).

SNP ID	Mutation	SIFT		PolyPhen 2.0		PANTHER	
		Tolerance index	Prediction	PSIC	prediction	subPSEC	Prediction
rs202140203	Y634C	0.00	Deleterious	1.000	Damaging	−6.31284	Deleterious
rs202046984	I787V	0.71	Tolerated	0.000	Benign	−3.66577	Deleterious
rs201956601	D232G	0.46	Tolerated	0.000	Benign	NA	NA
rs201791992	V517I	0.52	Tolerated	0.003	Benign	−3.72368	Deleterious
rs201351377	N490T	0.04	Deleterious	0.866	Damaging	−5.18241	Deleterious
rs201236716	N598S	0.16	Tolerated	0.877	Damaging	−3.59379	Deleterious
rs201226297	P471L	0.00	Deleterious	1.000	Damaging	−8.04132	Deleterious
rs201156397	A558S	0.18	Tolerated	0.990	Damaging	−5.01162	Deleterious
rs201100763	L737V	0.01	Deleterious	0.995	Damaging	−3.39472	Deleterious
rs201021988	D182N	0.22	Tolerated	0.000	Benign	NA	NA
rs200444084	R1186H	0.00	Deleterious	1.000	Damaging	−4.25642	Deleterious
rs200323346	V148L	0.34	Tolerated	0.000	Benign	NA	NA
rs200291759	N70S	0.45	Tolerated	0.001	Benign	−2.68708	Tolerated
rs200264267	R1272C	0.04	Deleterious	0.999	Damaging	−6.61915	Deleterious
rs200243024	S365L	0.06	Tolerated	0.334	Benign	−3.67158	Deleterious
rs199963560	A181T	0.07	Tolerated	0.500	Damaging	NA	NA
rs199812609	V494M	0.11	Tolerated	0.008	Benign	−4.24951	Deleterious
rs199693280	C334S	0.36	Tolerated	0.000	Benign	−3.37473	Deleterious
rs193182840	V780M	0.01	Deleterious	0.961	Damaging	−6.53794	Deleterious
rs192963719	A183V	0.28	Tolerated	0.000	Benign	NA	NA
rs191876836	K1010Q	0.36	Tolerated	0.007	Benign	−3.84354	Deleterious
rs191537721	D502E	1.00	Tolerated	0.000	Benign	−3.52121	Deleterious
rs190298665	F842L	0.00	Deleterious	1.000	Damaging	−4.95077	Deleterious
rs151305963	R1274Q	0.20	Tolerated	0.003	Benign	−4.17639	Deleterious
rs150334966	S1004L	0.23	Tolerated	0.837	Damaging	−5.30274	Deleterious
rs150154006	E301K	0.50	Tolerated	0.109	Benign	NA	NA
rs150053420	R372Q	0.57	Tolerated	0.001	Benign	−3.57866	Deleterious
rs149074243	M156V	0.42	Tolerated	0.019	Benign	NA	NA
rs149020783	C352F	0.46	Tolerated	0.000	Benign	−4.21514	Deleterious
rs148571882	W1122R	0.46	Tolerated	0.013	Benign	−4.31283	Deleterious
rs148237665	Y932S	0.42	Tolerated	0.077	Benign	−3.74345	Deleterious
rs148035987	R1183H	0.00	Deleterious	0.976	Damaging	−6.7934	Deleterious
rs147615070	M156T	0.67	Tolerated	0.014	Benign	NA	NA
rs147021046	N589S	0.28	Tolerated	0.001	Benign	−4.34832	Deleterious
rs146874573	V753M	0.07	Tolerated	0.643	Damaging	−5.2153	Deleterious
rs146666146	A1018T	0.03	Deleterious	0.883	Damaging	−4.6953	Deleterious
rs145666943	Q60H	0.09	Tolerated	0.656	Damaging	−4.24272	Deleterious
rs145362908	V810L	0.30	Tolerated	0.000	Benign	−3.88266	Deleterious
rs145297180	R1186C	0.00	Deleterious	1.000	Damaging	−6.79621	Deleterious
rs144725473	V1141G	0.00	Deleterious	0.984	Damaging	−4.77998	Deleterious
rs144687654	S652R	0.00	Deleterious	0.999	Damaging	−5.59673	Deleterious
rs143908453	T759A	0.62	Tolerated	0.340	Benign	−3.43971	Deleterious
rs143797098	P424A	0.98	Tolerated	0.014	Benign	−4.17737	Deleterious
rs143205855	G149R	0.33	Tolerated	0.248	Benign	NA	NA
rs143124972	S940L	0.01	Deleterious	0.999	Damaging	−5.68616	Deleterious
rs141892620	I450V	0.45	Tolerated	0.007	Benign	−3.9303	Deleterious
rs141440861	G1073S	0.77	Tolerated	0.069	Benign	−2.81269	Tolerated
rs141361998	V327I	0.13	Tolerated	0.002	Benign	NA	NA
rs141243713	R161W	0.00	Deleterious	0.982	Damaging	NA	NA

(*Continued*)

SNP ID	Mutation	SIFT		PolyPhen 2.0		PANTHER	
		Tolerance index	Prediction	PSIC	prediction	subPSEC	Prediction
rs140952850	R116Q	0.16	Tolerated	0.000	Benign	−3.01094	Deleterious
rs140827681	F1207L	0.47	Tolerated	0.715	Damaging	−3.80895	Deleterious
rs140786703	A659V	0.41	Tolerated	0.033	Benign	−4.31882	Deleterious
rs140527006	E1271Q	0.24	Tolerated	0.604	Damaging	−4.43776	Deleterious
rs140211089	H1016R	0.23	Tolerated	0.243	Benign	−2.86351	Tolerated
rs140149624	S425L	0.14	Tolerated	0.006	Benign	−3.56317	Deleterious
rs139751448	R404Q	0.00	Deleterious	1.000	Damaging	−5.56223	Deleterious
rs139612110	S1169I	0.00	Deleterious	1.000	Damaging	−3.97501	Deleterious
rs139485263	N185D	0.17	Tolerated	0.005	Benign	NA	NA
rs139297968	A851T	0.02	Deleterious	0.994	Damaging	−5.57919	Deleterious
rs138277307	E43K	0.93	Tolerated	0.001	Benign	−2.86264	Tolerated
rs138184115	A521S	0.52	Tolerated	0.003	Benign	−3.62631	Deleterious
rs138151007	W291C	0.17	Tolerated	0.957	Damaging	NA	NA
rs138079168	A321V	0.29	Tolerated	0.001	Benign	NA	NA
rs120074136	C113R	0.00	Deleterious	1.000	Damaging	−3.53647	Deleterious
rs120074135	V950M	0.28	Tolerated	0.007	Benign	−4.28583	Deleterious
rs120074134	V378A	0.09	Tolerated	0.969	Damaging	−5.35292	Deleterious
rs120074132	R958Q	0.05	Deleterious	1.000	Damaging	−8.09327	Deleterious
rs120074131	L1213F	0	Deleterious	1.000	Damaging	−3.49078	Deleterious
rs120074130	V889M	0.12	Tolerated	0.995	Damaging	−4.2027	Deleterious
rs113371321	A1187V	0.12	Tolerated	0.999	Damaging	−3.41161	Deleterious
rs112387560	R646H	0.28	Tolerated	0.010	Benign	−5.13502	Deleterious
rs112101747	V706A	0.03	Deleterious	0.994	Damaging	−4.37216	Deleterious
rs111256741	A183T	0.36	Tolerated	0.000	Benign	NA	NA
rs80358259	I1061T	0	Deleterious	1.000	Damaging	−8.53429	Deleterious
rs80358258	A1054T	0.01	Deleterious	0.875	Damaging	−5.70035	Deleterious
rs80358257	P1007A	0.01	Deleterious	0.995	Damaging	−5.56021	Deleterious
rs80358254	G992R	0.50	Tolerated	0.789	Damaging	−3.84824	Deleterious
rs80358253	Q775P	0.00	Deleterious	1.000	Damaging	−7.73353	Deleterious
rs80358252	C177Y	0.02	Deleterious	0.999	Damaging	NA	NA
rs80358251	P237S	0.38	Tolerated	0.003	Benign	NA	NA
rs77815278	Y420S	0.17	Tolerated	0.742	Damaging	−3.46767	Deleterious
rs77080672	R411Q	0.58	Tolerated	0.000	Benign	−4.66248	Deleterious
rs76615690	V624I	1.00	Tolerated	0.005	Benign	−3.83876	Deleterious
rs61731969	M1179V	0.48	Tolerated	0.000	Benign	−4.50621	Deleterious
rs61731962	P434S	0.02	Deleterious	0.001	Benign	−4.94008	Deleterious
rs55680026	N222S	0.74	Tolerated	0.003	Benign	NA	NA
rs35248744	S1200G	0.00	Deleterious	0.998	Damaging	−6.50452	Deleterious
rs34302553	G911S	0.66	Tolerated	0.904	Damaging	−4.16641	Deleterious
rs34226296	V1115F	0.07	Tolerated	0.001	Benign	−4.57463	Deleterious
rs34084984	N961S	0.78	Tolerated	0.034	Benign	−3.56842	Deleterious
rs28942108	R978C	0.05	Deleterious	0.955	Damaging	−3.84198	Deleterious
rs28942107	A1035V	0.01	Deleterious	0.996	Damaging	−5.95932	Deleterious
rs28942106	Y1088C	0.00	Deleterious	1.000	Damaging	−7.52745	Deleterious
rs28942105	N1156S	0.00	Deleterious	1.000	Damaging	−4.42918	Deleterious
rs28942104	T1036M	0.00	Deleterious	1.000	Damaging	−4.29841	Deleterious
rs28940897	Q928P	0.09	Tolerated	0.082	Benign	NA	NA
rs17855819	S151G	0.12	Tolerated	0.091	Benign	NA	NA
rs13381670	T511M	0.01	Deleterious	0.978	Damaging	−4.23833	Deleterious
rs1805084	R1266Q	0.54	Tolerated	0.003	Benign	−3.3839	Deleterious

(*Continued*)

SNP ID	Mutation	SIFT		PolyPhen 2.0		PANTHER	
		Tolerance index	Prediction	PSIC	prediction	subPSEC	Prediction
rs1805082	I858V	0.58	Tolerated	0.047	Benign	−3.71112	Deleterious
rs1805081	H215R	0.53	Tolerated	0.000	Benign	NA	NA
rs1788799	M642I	1.00	Tolerated	0.000	Benign	−3.60971	Deleterious
rs1140458	N931K	0.12	Tolerated	0.034	Benign	−3.89891	Deleterious

AA – Amino Acid; NA – not available (i.e. it indicates that, position does not align to the HMM).

Table 2 The disease-associated SNPs are predicted from PHDSnp server results, pathogenecity index obtained from Pmut server and g score, P score, molecular variations and prediction reliability calculated from MutPred server. Here the most deleterious nsSNPs are displayed in bold.

SNP IDs	Amino acid change	PHDsnp results	Pmut		MUTPred			
			Score	Prediction	g score	P score	Molecular variation	Prediction reliability
rs202140203	Y634C	Disease	0.6874	Pathological	0.910	0.1836	Gain of catalytic residue	Low confidence
rs201351377	N490T	Disease	0.2170	Neutral	0.669	0.132	Gain of helix	Low confidence
rs201226297	P471L	Disease	0.2581	Neutral	0.572	0.062	Gain of helix	Low confidence
rs201100763	L737V	Neutral	0.2533	Neutral	0.464	0.2612	Loss of stability	Low confidence
rs200444084	R1186H	Disease	0.5891	Pathological	0.900	0.0647	Loss of MoRF binding	Low confidence
rs200264267	R1272C	Disease	0.7847	Pathological	0.448	0.0125	Gain of sheet	Low confidence
rs193182840	V780M	Neutral	0.2847	Neutral	0.706	0.3997	Loss of catalytic residue	Low confidence
rs190298665	F842L	Disease	0.3521	Neutral	0.745	0.1532	Loss of stability	Low confidence
rs148035987	R1183H	Disease	0.5679	Pathological	0.443	0.0269	Loss of MoRF binding	Low confidence
rs146666146	A1018T	Disease	0.6817	Pathological	0.777	0.0737	Gain of phosphorylation	Low confidence
rs145297180	R1186C	Disease	0.6215	Pathological	0.804	0.0127	Loss of disorder	Actionable hypothesis
rs144725473	V1141G	Disease	0.8750	Pathological	0.664	0.0428	Loss of stability	Actionable hyphothesis
rs144687654	S652R	Disease	0.8608	Pathological	0.888	0.0516	Gain of MoRF binding	Low confidence
rs143124972	S940L	Disease	0.7018	Pathological	0.874	0.0284	Gain of catalytic residue	Actionable hypothesis
rs139751448	R404Q	Disease	0.4922	Neutral	0.969	0.0457	Loss of sheet	Actionable hypothesis
rs139612110	S1169I	Disease	0.7754	Pathological	0.787	0.0827	Gain of sheet	Low confidence
rs139297968	A851T	Neutral	0.5439	Pathological	0.661	0.079	Loss of helix	Low confidence
rs120074136	C113R	Neutral	0.7311	Pathological	0.974	0.0396	Gain of disorder	Actionable hypothesis
rs120074132	R958Q	Disease	0.8995	Pathological	0.815	0.0281	Loss of catalytic residue	Confident hypothesis
rs120074131	L1213F	Disease	0.9486	Pathological	0.968	0.0777	Loss of ubiquitination	Low confidence
rs112101747	V706A	Neutral	0.2515	Neutral	0.643	0.0375	Gain of catalytic residue	Actionable hypothesis
rs80358259	I1061T	Disease	0.7847	Pathological	0.950	6e-04	Gain of disorder	Very Confident hypothesis
rs80358258	A1054T	Disease	0.6808	Pathological	0.863	0.141	Loss of ubiquitination	Low confidence
rs80358257	P1007A	Disease	0.3212	Neutral	0.907	0.0151	Loss of stability	Actionable hypothesis
rs80358253	Q775P	Disease	0.7092	Pathological	0.898	0.1853	Gain of catalytic residue	Low confidence
rs35248744	S1200G	Neutral	0.6112	Pathological	0.671	0.0144	Loss of glycosylation	Actionable hypothesis
rs28942108	R978C	Neutral	0.8950	Pathological	0.850	0.0374	Loss of loop	Actionable hypothesis
rs28942107	A1035V	Disease	0.4889	Neutral	0.873	0.039	Gain of sheet	Actionable hypothesis
rs28942106	Y1088C	Disease	0.8843	Pathological	0.897	0.0558	Loss of helix	Low confidence
rs28942105	N1156S	Disease	0.4500	Neutral	0.936	0.2294	Gain of helix	Low confidence
rs28942104	T1036M	Disease	0.6654	Pathological	0.842	0.0986	Loss of loop	Low confidence

quantitative measures with default parameters. Modeled and validated native and mutant NPC1 coordinate files were submitted as input to the server. We calculate the numbers of hydrogen bonds, total accessible surface area (ASA) and free folding energy of native and mutant NPC1 structures and shown in Table 3. Native NPC1 showed 19809.7 $Å^2$ of ASA, while mutant structures (R1186C, S940L, R958Q and I1061T) showed 19879.1, 19813.4, 19805.7 and 19973.8 $Å^2$ of ASA respectively (Table 3). Accessible surface area (ASA) is the exposed surface area of the protein that a water molecule could access or touch. It is believed that ASA is good measure of structure geometry of protein. Expanded

structure has more ASA as compared to shrink structure. In our analysis native showed minimum ASA when compared to mutant structures. It indicates that due to mutation in NPC1 protein acquired more expanded structural geometry in turns achieve higher value of ASA. Mutant I1061T showed highest value of ASA than other structures.

Native structure contains free folding energy of – 204.52 kcal/mol, while mutant I1061T contains – 181.33 kcal/mol. Other mutants contain intermediate level of free folding energy (FFE) (Table 3). Native NPC1 protein structure showed least free folding energy as compared to mutant structures. Lowest FFE indicates correct folding which is essential for NPC1 protein function, while higher FFE leads toward misfolding and functionally inactive. It is clear that while mutation in NPC1 acquires higher FFE, it lost its activity. Compare to native and other mutations in NPC1, mutant I1061T showed highest FFE (–181.33 kcal/mol) and it's responsible for destabilizing the NPC1 structure. Hence, it produces most deleterious effect on NPC1. It indicates the mutation I1061T produced high structural deleterious effect as compared to other mutations.

During the superimposition analysis mutant R1186C, S940L, R958Q and I1061 showed structure deviation of 0.077 Å, 0.082 Å, 0.076 Å and 0.088 Å respectively, when compared to native structure. Superimposed images of native and mutants are shown in Fig. 1(A)–(D). Superimposition analysis with reference to native and mutant structure shows significant deviation at c-alpha carbon atom. Mutant T373K showed higher structural deviations as compared to other analyzed mutation. Estimation of the contribution of hydrogen bonding to protein stability has been made by a combination of experiments on model compounds and site-directed mutational studies (Harpaz eet al., 1994). The change in amino acid residues is often accompanied with the alterations in interaction pattern, specially the H bond formation of the corresponding protein (Sunyaev et al., 2000; Chasman and Adams, 2001). Number of hydrogen bonds was found significantly altered upon mutation. Formations of hydrogen bonds in protein structure affect its conformational flexibility (Purohit and Sethumadhavan, 2009; Purohit et al., 2011a, 2011b; Rajendran et al., 2012; Balu et al., 2013; Kumar et al., 2014; Kumar and Purohit, 2014). More hydrogen bonds lead to rigid structure and vice versa. Mutant I1061T structure acquired 215 hydrogen bonds, while native structure has 221 hydrogen bonds as depicted in Table 3. All mutant structures showed significantly decreased number of hydrogen bonds compared to native structure. Due to mutation the structure got expanded and it was well supported by ASA analysis. We conducted flexibility analysis to confirm the above observations. We analyzed the flexibility behavior of native and mutant NPC1 protein by WEBnm approach. Normal mode analysis (NMA) has become the method of choice to investigate the slowest motions in macromolecular systems. NMA relies on the hypothesis that the vibrational normal modes having the lowest frequencies (also named soft modes) describe the largest movements in a protein and are the ones that are functionally relevant. In WEBnm analysis we observed mode 7 which has lower deformation energy than other mode. Native and mutant fluctuation in mode 7 is shown in Fig. 2(A)–(E). The mutant structures showed more motions and flexibility behavior than native structure. It confirms that due to mutation, NPC1 structure became more flexible in nature and because of this structural flexibility it may lose the correct function and leads

Table 3 List of ASA value, Free Folding energy, RMSD and No. of Hydrogen bonds in native and mutant NPC1 proteins

Protein type	Total ASA Value (Å^2)	Free folding energy (Kcal/mol)	RMSD (Å^2)	No. of hydrogen bonds
NPC	19809.7	–204.52	0	221
NPC^{R1186C}	19879.1	–184.29	0.077	218
NPC^{S940L}	19813.4	–184.57	0.082	217
NPC^{R958Q}	19805.7	–182.55	0.076	218
NPC^{I1061T}	19973.8	–181.33	0.088	215

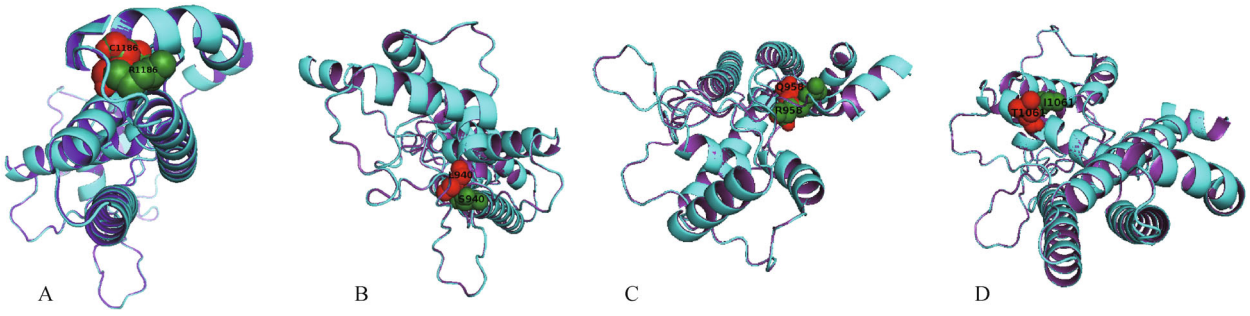

Figure 1 **(A)–(D)** Superimposition of native and mutant (R1186C, S940L, R958Q and I1061T) modeled structures of NPC1.

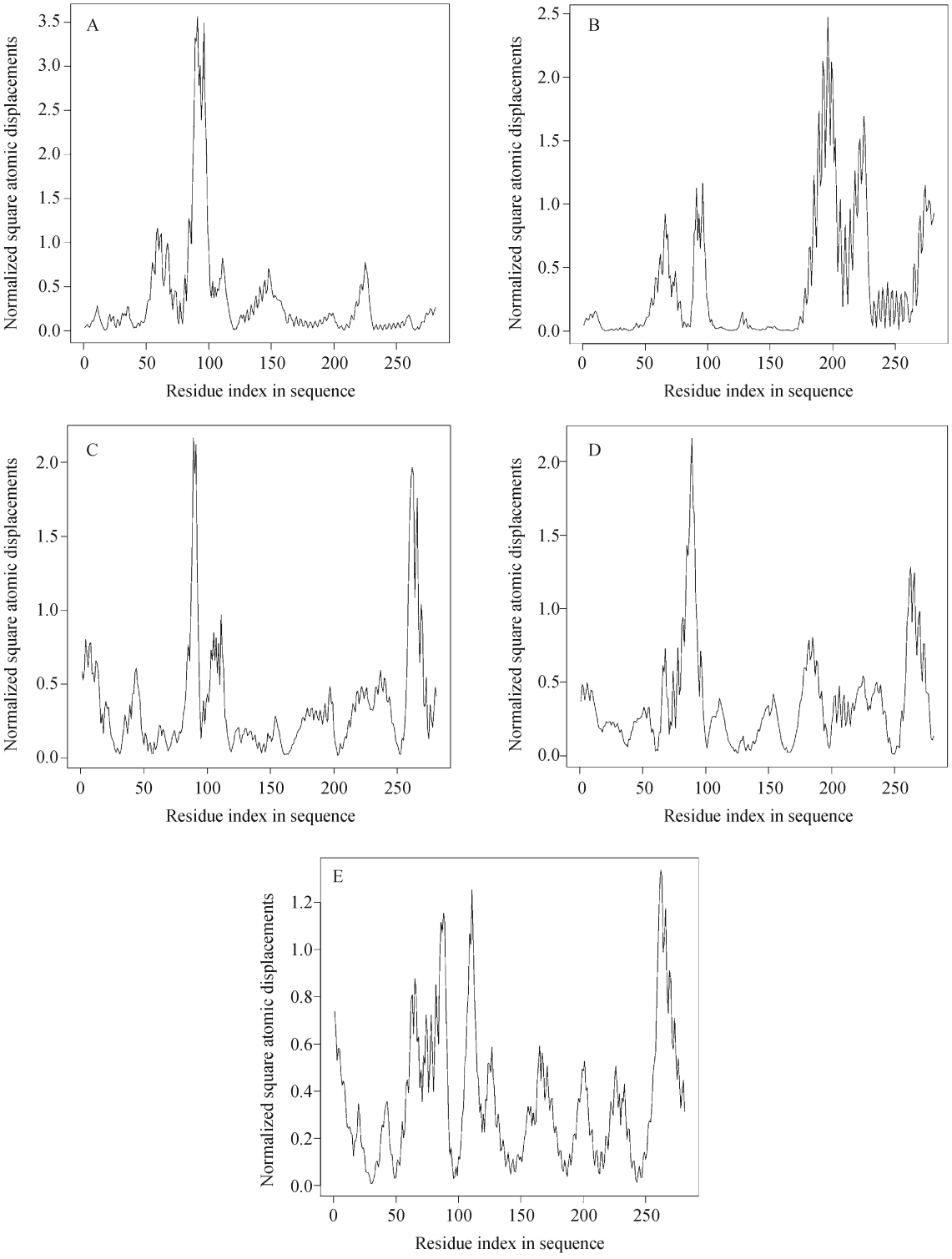

Figure 2 Normal mode analysis of c-alpha carbon atom of native and mutant NPC1. (A) Native, (B) R1186C, (C) S940L, (D) R958Q, (E) I1061T.

to cause Niemann-Pick disease type C1. Our analysis suggests the most disease-associated mutations of NPC1 gene and its structural consequence NPC1 protein.

Conclusions

Computational investigation has now become a roadmap to

characterize a standard disease specific SNP at molecular level. In this study we screened four most disease associated mutations (R1186C, S940L, R958Q and I1061T) which are related to Niemann-Pick disease type C1. Threading based approach was used to model the native and mutant NPC1 proteins. Quantitative and structural approaches have also been extensively used to report the structural consequences of the deleterious predicted point mutations. Due to mutation the structure became more flexible in nature and Mutation I1061T showed most deleterious effect than other mutants in NPC1 which was well supported by ASA, FFE and NHbonds. This may affect the structural and functional behavior of NPC1 protein. Our analysis provides a way to detect the Niemann-Pick disease type C1 associated SNPs from the large SNP data set. Also provides a clear cut clue to researchers about how disease-related SNPs, R1186C, S940L, R958Q and R1186C in NPC1 gene affect the structure of NPC1 protein.

Compliance with ethics guidelines

Naresh Kandakatla, Rajasekhar Chekkara, Namachivayam Balakrishnan and Geetha Ramakrishnan declare that they have no conflict of interest. This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Adzhubei I A, Schmidt S, Peshkin L, Ramensky V E, Gerasimova A, Bork P, Kondrashov A S, Sunyaev S R (2010). A method and server for predicting damaging missense mutations. Nat Methods, 7(4): 248–249

Amberger J, Bocchini C A, Scott A F, Hamosh A (2009). McKusick's online Mendelian inheritance in man (OMIM®). Nucleic Acids Res, 37(Database issue suppl 1): D793–D796

Balu K, Purohit R (2013). Mutational analysis of TYR gene and its structural consequences in OCA1A. Gene, 513(1): 184–195

Balu K, Rajendran V, Sethumadhavan R, PurohitR (2013). Investigation of binding phenomenon of NSP3 and p130Cas mutants and their effect on cell signalling. Cell Biochem Biophys, 67(2): 623–633

Bauer P, Knoblich R, Bauer C, Finckh U, Hufen A, Kropp J, Braun S, Kustermann-Kuhn B, Schmidt D, Harzer K, Rolfs A (2002). NPC1: Complete genomic sequence, mutation analysis, and characterization of haplotypes. Hum Mutat, 19(1): 30–38

Beltroy E P, Richardson J A, Horton J D, Turley S D, Dietschy J M (2005). Cholesterol accumulation and liver cell death in mice with Niemann-Pick type C disease. Hepatology, 42(4): 886–893

Capriotti E, Calabrese R, Casadio R (2006). Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics, 22(22): 2729–2734

Carstea E D, Morris J A, Coleman K G, Loftus S K, Zhang D, Cummings C, Gu J, Rosenfeld M A, Pavan W J, Krizman D B, Nagle J, Polymeropoulos M H, Sturley S L, Ioannou Y A, Higgins M E, Comly M, Cooney A, Brown A, Kaneski C R, Blanchette-Mackie E J, Dwyer N K, Neufeld E B, Chang T Y, Liscum L, Strauss J F 3rd, Ohno K, Zeigler M, Carmi R, Sokol J, Markie D, O'Neill R R, van Diggelen O P, Elleder M, Patterson M C, Brady R O, Vanier M T, Pentchev P G, Tagle D A (1997). Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. Science, 277 (5323): 228–231

Carvalho M A, Marsillac S M, Karchin R, Manoukian S, Grist S, Swaby R F, Urmenyi T P, Rondinelli E, Silva R, Gayol L, Baumbach L, Sutphen R, Pickard-Brzosowicz J L, Nathanson K L, Sali A, Goldgar D, Couch F J, Radice P, Monteiro A N A (2007). Determination of cancer risk associated with germ line BRCA1 missense variants by functional analysis. Cancer Res, 67(4): 1494–1501

Carvalho M, Pino M A, Karchin R, Beddor J, Godinho-Netto M, Mesquita R D, Rodarte R S, Vaz D C, Monteiro V A, Manoukian S, Colombo M, Ripamonti C B, Graeme SuthersR R, Borg A, Radice P, Grist S A, Monteiro A N A, Billack B (2009). Analysis of a set of missense, frameshift, and in-frame deletion variants of BRCA1. Mutat Res, 660(1): 1–11

Chasman D, Adams R M (2001). Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. J Mol Biol, 307(2): 683–706

Fernandez-Valero E M, Ballart A, Iturriaga C, Lluch M, Macias J, Vanier M T, Pineda M, Coll M J (2005). Identification of 25 new mutations in 40 unrelated Spanish Niemann-Pick type C patients: genotype-phenotype correlations. Clin Genet, 68(3): 245–254

Ferrer-Costa C, Gelpí J L, Zamakola L, Parraga I, de la Cruz X, Orozco M (2005). PMUT: a web-based tool for the annotation of pathological mutations on proteins. Bioinformatics, 21(14): 3176–3178

Garver W S, Francis G A, Jelinek D, Shepherd G, Flynn J, Castro G, Walsh Vockley C, Coppock D L, Pettit K M, Heidenreich R A, Meaney F J (2007). The National Niemann-Pick C1 disease database: report of clinical features and health problems. Am J Med Genet A, 143A(11): 1204–1211

Garver W S, Heidenreich R A (2002). The Niemann-Pick C proteins and trafficking of cholesterol through the late endosomal/lysosomal system. Curr Mol Med, 2(5): 485–505

Garver W S, Jelinek D, Meaney F J, Flynn J, Pettit K M, Shepherd G, Heidenreich R A, Vockley C M W, Castro G, Francis G A (2010). The National Niemann-Pick Type C1 Disease Database: correlation of lipid profiles, mutations, and biochemical phenotypes. J Lipid Res, 51(2): 406–415

Goldgar D E, Easton D F, Deffenbaugh A M, Monteiro A N, Tavtigian S V, Couch F J, the Breast Cancer Information Core (BIC) Steering Committee (2004). Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. Am J Hum Genet, 75(4): 535–544

Goldstein J L, Brown M S (1992). Lipoprotein receptors and the control of plasma LDL cholesterol levels. Eur Heart J, 13(Suppl B): 34–36

Greer W L, Dobson M J, Girouard G S, Byers D M, Riddell D C, Neumann P E (1999). Mutations in *NPC1* highlight a conserved NPC1-specific cysteine-rich domain, Am J Hum Genet, 65(5): 1252–1260

Harpaz Y, Gerstein M, Chothia C (1994). Volume changes on protein folding. Structure, 2(7): 641–649

Hess B, Kutzner C, Van Der Spoel D, Lindahl E (2008). GROMACS 4:

Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput, 4(3): 435–447

Hollup S M, Salensminde G, Reuter N (2005). WEBnm@: a web application for normal mode analyses of proteins. BMC Bioinformatics, 6(1): 52

Ioannou Y A (2000). The structure and function of the Niemann-Pick C1 protein. Mol Genet Metab, 71(1-2): 175–181

Kamaraj B, Purohit R (2014). Computational screening of disease-associated mutations in OCA2 gene. Cell Biochem Biophys, 68(1): 97–109

Kamaraj B, Purohit R (2013). *In silico* screening and molecular dynamics simulation of disease-associated nssnp in tyrp1 gene and its structural consequences in OCA3. BioMed Res Int, 2013: Article ID 697051

Kaplan W, Littlejohn T G (2001). Swiss-PDB viewer (deep view). Brief Bioinform, 2(2): 195–197

Karchin R (2009). Next generation tools for the annotation of human SNPs. Brief Bioinform, 10(1): 35–52

Kobayashi T, Beuchat M H, Lindsay M, Frias S, Palmiter R D, Sakuraba H, Parton R G, Gruenberg J (1999). Late endosomal membranes rich in lysobisphosphatidic acid regulate cholesterol transport. Nat Cell Biol, 1(2): 113–118

Kumar A, Purohit R (2012a). Computational centrosomics: an approach to understand the dynamic behaviour of centrosome. Gene, 511(1): 125–126

Kumar A, Purohit R (2012c). Computational investigation of pathogenic nsSNPs in CEP63 protein. Gene, 503(1): 75–82

Kumar A, Purohit R (2013). Cancer associated E17K mutation causes rapid conformational drift in AKT1 pleckstrin homology (PH) domain. PLoS One, 8(5): e64364

Kumar A, Purohit R (2014). Use of Long Term Molecular Dynamics Simulation in Predicting Cancer Associated SNPs. PLoS Comput Biol, 10(4): e1003318

Kumar A, Rajendran V, Sethumadhavan R, Purohit R (2013). Molecular dynamic simulation reveals damaging impact of RAC1 F28L mutation in the switch I region. PLoS One, 8(10): e77453

Kumar A, Rajendran V, Sethumadhavan R, Purohit R (2013a). Computational investigation of cancer-associated molecular mechanism in Aurora A (S155R) mutation. Cell Biochem Biophys, 66(3): 787–796

Kumar A, Rajendran V, Sethumadhavan R, Purohit R (2013b). Evidence of colorectal cancer-associated mutation in MCAK: a computational report. Cell Biochem Biophys, 67(3): 837–851

Kumar A, Rajendran V, Sethumadhavan R, Purohit R (2014). Relationship between a point mutation S97C in CK1δ protein and its affect on ATP-binding affinity. J Biomol Struct Dyn, 32(3): 394–405

Kumar A, Rajendran V, Sethumadhavan R, Shukla P, Tiwari S, Purohit R (2014). Computational SNP analysis: current approaches and future prospects. Cell Biochem Biophys, 68(2): 233–239

Kumar P, Henikoff S, Ng P C (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc, 4(7): 1073–1081

Kumar A, Purohit R (2012b). Computational screening and molecular dynamics simulation of disease associated nsSNPs in CENP-E. Mutat Res, 738–739: 28–37

Kumar A, Rajendran V, Sethumadhavan R, Purohit R (2012). *In silico* prediction of a disease-associated STIL mutant and its affect on the recruitment of centromere protein J (CENPJ). FEBS Open Bio, 2: 285–293

Laskowski R A, Rullmannn J A, MacArthur M W, Kaptein R, Thornton J M (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR, 8 (4): 477–486

Li B, Krishnan V G, Mort M E, Xin F, Kamati K K, Cooper D N, Mooney S D, Radivojac P (2009). Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics, 25(21): 2744–2750

Liscum L (2000). Niemann-Pick type C mutations cause lipid traffic jam. Traffic, 1(3): 218–225

Millat G, Baïlo N, Molinero S, Rodriguez C, Chikh K, Vanier M T (2005). Niemann-Pick C disease: use of denaturing high performance liquid chromatography for the detection of NPC1 and NPC2 genetic variations and impact on management of patients and families. Mol Genet Metab, 86(1-2): 220–232

Millat G, Marçais C, Rafi M A, Yamamoto T, Morris J A, Pentchev P G, Ohno K, Wenger D A, Vanier M T (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(5): 1321–1329

Millat G, Marçais C, Tomasetto C, Chikh K, Fensom A H, Harzer K, Wenger D A, Ohno K, Vanier M T (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(6): 1373–1385

Mooney S (2005). Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis. Brief Bioinform, 6(1): 44–56

Morris J A, Zhang D, Coleman K G, Nagle J, Pentchev P G, Carstea E D (1999). The genomic organization and polymorphism analysis of the human Niemann-Pick C1 gene. Biochem Biophys Res Commun, 261 (2): 493–498

Ng P C, Henikoff S (2006). Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet, 7 (1): 61–80

Pipalia N H, Huang A, Ralph H, Rujoi M, Maxfield F R (2006). Automated microscopy screening for compounds that partially revert cholesterol accumulation in Niemann-Pick C cells. J Lipid Res, 47 (2): 284–301

Puri V, Watanabe R, Dominguez M, Sun X, Wheatley C L, Marks D L, Pagano R E (1999). Cholesterol modulates membrane traffic along the endocytic pathway in sphingolipid-storage diseases. Nat Cell Biol, 1(6): 386–388

Purohit R (2014). Role of ELA region in auto-activation of mutant KIT receptor: a molecular dynamics simulation insight. J Biomol Struct Dyn, 32(7): 1033–1046

Purohit R, Rajendran V, Sethumadhavan R (2011a). Studies on adaptability of binding residues and flap region of TMC-114 resistance HIV-1 protease mutants. J Biomol Struct Dyn, 29(1): 137–152

Purohit R, Rajendran V, Sethumadhavan R (2011b). Relationship between mutation of serine residue at 315th position in M. tuberculosis catalase-peroxidase enzyme and Isoniazid susceptibility: an in silico analysis. J Mol Model, 17(4): 869–877

Purohit R, Sethumadhavan R (2009). Structural basis for the resilience of Darunavir (TMC114) resistance major flap mutations of HIV-1 protease. Interdiscip Sci, 1(4): 320–328

Rajendran V, Purohit R, Sethumadhavan R (2012). In silico investigation of molecular mechanism of laminopathy caused by a point mutation (R482W) in lamin A/C protein. Amino Acids, 43(2): 603–615

Rajendran V, Sethumadhavan R (2014). Drug resistance mechanism of PncA in Mycobacterium tuberculosis. J Biomol Struct Dyn, 32(2): 209–221

Ribeiro I, Marcão A, Amaral O, Sá Miranda M C, Vanier M T, Millat G (2001). Niemann-Pick type C disease: NPC1 mutations associated with severe and mild cellular cholesterol trafficking alterations. Hum Genet, 109(1): 24–32

Runz H, Dolle D, Schlitter A M, Zschocke J (2008). NPC-db, a Niemann-Pick type C disease gene variation database. Hum Mutat, 29(3): 345–350

Scott C, Ioannou Y A(2004). The NPC1 protein: structure implies function. Biochimica et Biophysica Acta (BBA)-Mol Cell Biol L, 1685(1): 8–13

Sherry S T, Ward M H, Kholodov M, Baker J, Phan L, Smigielski E M, Sirotkin K (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res, 29(1): 308–311

Steward R E, MacArthur M W, Laskowski R A, Thornton J M (2003). Molecular basis of inherited diseases: a structural perspective. Trends Genet, 19(9): 505–513

Sun X, Marks D L, Park W D, Wheatley C L, Puri V, O'Brien J F, Kraft D L, Lundquist P A, Patterson M C, Pagano R E, Snow K (2001). Niemann-Pick C variant detection by altered sphingolipid trafficking and correlation with mutations within a specific domain of NPC1. Am J Hum Genet, 68(6): 1361–1372

Sun X, Marks D L, Park W D, Wheatley C L, Puri V, O'Brien J F, Kraft D L, Lundquist P A, Patterson M C, Pagano R E, Snow K (2001). Niemann-Pick C variant detection by altered sphingolipid trafficking and correlation with mutations within a specific domain of NPC1. Am J Hum Genet, 68(6): 1361–1372

Sunyaev S, Ramensky V, Bork P (2000). Towards a structural basis of human non-synonymous single nucleotide polymorphisms. Trends Genet, 16(5): 198–200

Tamura H, Takahashi T, Ban N, Torisu H, Ninomiya H, Takada G, Inagaki N (2006). Niemann-Pick type C disease: novel NPC1 mutations and characterization of the concomitant acid sphingomyelinase deficiency. Mol Genet Metab, 87(2): 113–121

Tarugi P, Ballarini G, Bembi B, Battisti C, Palmeri S, Panzani F, Di Leo E, Martini C, Federico A, Calandra S (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 (11): 1908–1919

Thomas P D, Campbell M J, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A (2003). PANTHER: a library of protein families and subfamilies indexed by function. Genome Res, 13(9): 2129–2141

Vanier M T, Suzuki K (1998). Recent advances in elucidating Niemann-Pick C disease. Brain Pathol, 8(1): 163–174

Wiederstein M, Sippl M J (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res, 35(Web Server issue suppl 2): W407-10

Willard L, Ranjan A, Zhang H, Monzavi H, Boyko R F, Sykes B D, Wishart D S (2003). VADAR: a web server for quantitative evaluation of protein structure quality. Nucleic Acids Res, 31(13): 3316–3319

Xiong H, Higaki K, Wei C J, Bao X H, Zhang Y H, Fu N, Qin J, Adachi K, Kumura Y, Ninomiya H, Nanba E, Wu X R (2012). Genotype/phenotype of 6 Chinese cases with Niemann-Pick disease type C. Gene, 498(2): 332–335

Yamamoto T, Ninomiya H, Matsumoto M, Ohta Y, Nanba E, Tsutsumi Y, Yamakawa K, Millat G, Vanieer M, Pentchev P G, Ohno K (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(9): 707–712

Zhang Y (2008). I-TASSER server for protein 3D structure prediction. BMC Bioinformatics, 9(1): 40