

Mutational profiling of *POT1* gene and its interaction with *TPP1* in cancer- A computational approach

Priyanjali Bhattacharya^a, Pavan Kumar Dhanyamraju^b, Ankita Sarmah^c, Mohana Priya Jay^d, Chris Maria Jose^e, Sayali Dbritto^c, Showmeya Mallavarapu^c, Trupti N. Patel^{a,*}

^a Department of Integrative Biology, Vellore Institute of Technology, Vellore, India

^b Department of Pediatrics, Pennsylvania State University College of Medicine, Hershey, PA, 17033, USA

^c Department of Biomedical Sciences, Vellore Institute of Technology, Vellore, India

^d School of Health Professions, The University of Texas MD Anderson Cancer Center, USA

^e Department of Biotechnology, Vellore Institute of Technology, Vellore, India

ARTICLE INFO

Keywords:

SNPs
Telomeres
POT1 variants
TPP1 gene
POT1-TPP1 interaction

ABSTRACT

Telomeres are specialized structures at the end of eukaryotic chromosomes that maintain genomic stability by preventing chromosomal rearrangements and thereby enabling semi-conservative replication of telomeric DNA. The length of telomeric DNA is retained by telomerase that balances between the processes that lengthen and shorten the telomeres. In human chromosomes, six telomere-associated proteins namely- TRF1, TRF2, POT1, RAP1, TIN2, and TPP1 form the shelterin complex, that is essential for maintenance of telomeric integrity. The human POT1 and TPP1 play a major role in protecting the ssDNA overhangs, formed due to the end replication problem. These proteins along with other repair complexes protect the telomere overhangs from cellular repair complexes. However, shelterin functionality can become compromised due to mutations in any of the six genes and can lead to unwarranted repair of the human telomeres. In cancer and transformed cells, telomerase activation replenishes the telomeres while also, recruiting repair proteins at the telomeres. With an aim to evaluate the functional consequence of non-synonymous single nucleotide polymorphisms (nsSNPs) in *POT1* gene variants, and resulting changes that affect its interactions with *TPP1*, this research was carried out using computational tools. The overall outcomes revealed 16 *POT1* gene mutations that were likely to impact the protein function. Of these 9 mutations, viz., P357S, H437P, V439G, P475L, G534C, P537S, F566C, M587T, and C591W showed that the altered POT1 function impacted its interaction with TPP1 protein. The binding affinity of POT1 with ssDNA overhangs was also changed. A wet-lab follow-up study using site-directed mutagenesis and yeast hybridization techniques can help exploit underlying mechanisms affecting stable association of these two shelterin components.

1. Introduction

Telomeres, as described by Muller and McClintock in 1930, are essential components comprising the chromosomal ends. Human telomeres are composed of hexamer repeats of 'TTAGGG' and provide a protective capping to the chromosomes. These telomeres essentially safeguard the genome from unwanted inter-chromosomal fusion, recombination, nucleolytic degradation and cell death [1]. With every cell division, the length of the telomeres is shortened by 50–100bp, causing cellular senescence in normal cells. However, in cancer cells, the activation of enzyme telomerase causes the telomeres to replenish,

protecting them against senescence and imparting cellular mortality. The six telomere-associated proteins TRF1 (telomeric repeat binding factor 1), TRF2 (telomeric repeat binding factor 2), POT1 (protection of telomeres 1), RAP1 (TERF2 interacting protein), TIN2 (TRF1 interacting nuclear factor 2) and TPP1 (adrenocortical dysplasia protein homolog) forms the shelterin complex that maintains the length of telomeric DNA and enable normal maintenance of linear chromosomal ends in mammalian cells [2,3]. This includes protecting the ssDNA overhang formed due to the end replication problem from the various DNA repair systems. TRF1 and TRF2 bind with double-stranded (ds) telomeric DNA whereas POT1 associates itself with single-stranded (ss) DNA and coats

* Corresponding author,

E-mail addresses: tnpatel@vit.ac.in, Dr.TNPatel@gmail.com (T.N. Patel).

<https://doi.org/10.1016/j.imu.2020.100389>

Received 10 April 2020; Received in revised form 26 June 2020; Accepted 29 June 2020

Available online 4 July 2020

2352-9148/© 2020 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the overhang with its oligonucleotide binding (OB) folds [1,4]. POT1 via an interaction with TPP1 and TIN2 associates itself along with protein pairs, TRF1 and TRF2, to cooperate with telomeric DNA. RAP1, the sixth and most conserved shelterin subunit interacts solely with TRF2 and is known to operate transcriptional regulations and influence NF- κ B signaling [5]. A constitutive and ubiquitous expression of all these subunits makes the shelterin complex highly prolific in shielding the telomeric DNA while it depletes with every cell division.

The POT1-TPP1 heterodimer is critical for regulating telomeres length [6]. Kibe and colleagues (2010) suggested that accumulation of POT1 at telomeres rely on its interaction with TPP1 which in turn cooperates with other shelterin proteins and aid POT1 to bind to the single stranded telomeric DNA through OB-domains [7]. The affinity of binding of POT1 with single stranded telomeric DNA is highly specific to regulate both telomere length and capping. This is achieved by the two N terminal OB folds (OB1 and OB2) with residues ranging from 1 to 300. The C-terminal region, on the other hand, consists of OB3₃₂₀₋₆₃₄ and HJRL₃₉₃₋₅₃₈ (Holiday Junction Like Resolvase) domains, both of which interacts with POT1 binding domain (PBD) of TPP1 (residues 266–320) by forming a tight heterodimer. The packing of OB3-HJRL in POT1-C terminal includes extensive hydrophobic contacts. Presence of zinc ion, stabilizes the orientation between POT1_{OB3} and POT1_{HJRL} [8].

Single nucleotide polymorphisms (SNPs) detected in the *POT1* gene can affect the OB folds of the protein product and lead to telomeric instability. Recently, around 300 SNPs have been identified within the coding region of *POT1* (<https://www.cbioportal.org/>) in various cancers not limited to but including chronic lymphocytic leukaemia (CLL), familial melanoma, familial glioma, and cardiac angiosarcoma [9]. In CLL mutations of *POT1* result in chromosomal aberrations, leading to dominant negative effect on wild type protein [10,11]. Pinzaru et al. (2016) conducted *in vivo* experiments which emphasized that *POT1* mutations particularly promote malignancies of lymphatic system [12]. It was also noted by other group of researchers that *POT1* is a susceptible gene for hereditary cutaneous melanoma [13]. Thus, in case of various cancers, *POT1* gene mutations have a role in tumour development and progression. In cancers with active telomerase, TPP1 controls this enzyme and recruits it to the telomeres in a cell cycle-dependent manner [14–18]. Mutations in *TPP1* that lack the functional OB folds, results in shortening of the telomeres since TPP1 cannot bind to ssDNA independently [19]. Similarly, *POT1* mutants result in complete loss of telomere length control though there is a continuous increment of shelterin load [20]. It is important for POT1-TPP1 interactions to protect the ssDNA telomeres from degradation and repair, however, in an initiated cell it also facilitates extension of telomeres by allowing access of telomerase to ssDNA overhangs [6,20–22].

Contribution of altered telomere length in cancer cell perpetuity is known, but the molecular mechanisms that control this are poorly understood. Therefore, through *in silico* tools, in this study we aim to evaluate the nsSNPs in *POT1* gene and interpret the subsequent changes in POT1-TPP1 protein interactions that lead to telomere length restoration. Since different bioinformatic tools use different algorithms the predictive scores vary. Hence, our focus is to try and infer the final outcome of these mutations on POT1-TPP1 cooperation. Overall, the results of this study may contribute in filtering the significant mutations that can assist the development of diagnostic biomarkers in cancers with active telomerase.

2. Methodology

2.1. Data mining

A total of 338 non-synonymous missense substitutions corresponding to *POT1* gene were mined from COSMIC (Catalogue Of Somatic Mutations In Cancer) database for solid tumors (stomach, thyroid, lung, breast, skin, large intestine) and haematologic malignancies. The amino acid sequence of POT1 [UniProtKB - Q9NUX5 (POTE1_HUMAN)] was

obtained from Swiss-Prot database (<https://www.uniprot.org/>). The crystal structures of POT1 [PDB id: 5H65; 1XJV] and TPP1 [PDB id: 5H65; 2I46] were extracted from RCSB PDB (Protein Data Bank) database and used for non-covalent interaction analyses and docking after altering the mutant residues in the structure using Discovery Studio Visualizer. The nsSNPs were analysed to identify the degree of mutational damage using various predictive tools listed in Table 1.

2.2. Tools used for functional score prediction

PolyPhen-2 and MutPred are sequence and structure based prediction tools; SNPs&GO and PANTHER classify sequences based on support vector machine learning (sequence based) and SIFT, and SNAP, are based on sequence and evolutionary conservation methods. PredictSNP and MAPP are also sequence based tools that provide a percentage score in terms of accuracy of prediction and differentiates physicochemical properties between native and mutant amino acid residues. The selection of tools was based on use of different algorithms to generate the predictive scores. The tools allow the user to identify the nature of a particular substitution i.e. pathogenic or non-pathogenic. Besides, MutPred uniquely helps in prediction of changes in post-translational modification (PTM) features such as-methylation, ubiquitination, phosphorylation, glycosylation along with structural distortion of protein of interest.

2.3. Analysis of protein stability

I-Mutant suite (<http://gpcr.biocomp.unibo.it/cgi/predictor/s/I-Mutant3.0/I-Mutant3.0.cgi>), a support vector machine based web-server, predicts the instability in proteins owing to a point mutation. The protein sequences were submitted along with the position and mutant residue at constant temperature (25 °C) and pH (7). The results displayed, increase or decrease in free energy value (DDG>0, +ve; DDG<0, -ve) in kcal/mol [26].

2.4. HOPE project

Once the scores were retrieved, the damaging mutations were further analysed by HOPE (Have yOur Protein Explained) server (www.cmbi.ru.nl) which predicts the effect of mutation with an insight to study the structural features of the native and mutant protein models [32].

2.5. Analysis of evolutionary conserved residues

Consurf (<https://consurf.tau.ac.il/>) was used to measure the conservation score of an amino acid residue aligned at a given position to determine the significance of that residue in protein structure and function. The score ranges from 1 to 9 which depicts whether a residue is rapidly, moderately or slowly evolving. The lowest scores predict a highly conserved position [33].

2.6. Prediction of interatomic interactions

The web server Arpeggio (<http://biosig.unimelb.edu.au/arpeggio/oweb/>) was used to identify non-covalent interactions viz. van der Waals, ionic, polar, metal complex, carbonyl, and covalent interactions [34]. The hydrogen bonds, aromatic and hydrophobic contacts of POT1 protein which were affected by mutations were also revealed by this tool.

2.7. Prediction of protein secondary structure

The POT1 protein secondary structures were predicted by SOPMA (Self Optimized Prediction Method with Alignment), which explains the distributions of alpha helices, beta bridges and turns, extended strands, bend regions, and random coils (<https://npsa-prabi.ibcp.fr/cgi-bin/nps>

Table 1
Details of the computational tools used in present study.

Tools	Type	Incorporated programme	Type	Scoring system	URL	References
PolyPhen-2 (Polymorphism Phenotyping 2)	Sequence and structure based	-	-	0.000/1.000- benign or damaging	http://genetics.bwh.harvard.edu/pph2/	[23]
Meta-SNP (Meta-predictor of disease causing variants)	Meta Server	SIFT (Sorting Intolerant from Tolerant)	Sequence and evolutionary conservation	>0.05- neutral	http://snps.biofold.org/meta-snp/	[24]
		SNAP (Screening for Non-Acceptable Polymorphisms)	Supervised learning (neural networks)	>0.5- disease		[25]
		PhD-SNP	Supervised-learning (support vector machines)	>0.5- disease		
		PANTHER	Sequence and evolutionary conservation	>0.5- disease		[26]
SNPs&GO	Meta server	PhD-SNP (Predictor of human Deleterious-Single Nucleotide Polymorphisms)	Supervised-learning (support vector machines)	>0.5- disease	http://snps.biofold.org/snps-and-go/pages/method.html	[27]
		PANTHER	Sequence and evolutionary Conservation	>0.5- disease		[26]
		SNPs&GO	Supervised learning (support vector machine)	>0.5- disease		[28]
MutPred	-	-	Structure and sequence based	General score (g) > 0.75	http://MutPred1.mutdb.org/	[29]
PredictSNP	Meta server	PredictSNP	Sequence based	% accuracy prediction	http://loschmidt.chemi.muni.cz/predict-snp/	[30]
		MAPP	Sequence based	differences in physicochemical properties between wild-type and mutant amino acid residues		[31]

a_automat.pl?page=/NPSA/npsa_sopma.html) [35].

2.8. Protein-protein docking

The deleterious mutations in human *POT1* gene were incorporated into the protein structure with the aid of Discovery Studio Visualizer followed by interaction analyses using PatchDock (<https://bioinfo3d.cs.tau.ac.il/PatchDock/>), HDOCK (<http://hdock.phys.hust.edu.cn/>), ZDOCK (<http://zdock.umassmed.edu/>) and ClusPro (<https://cluspro.bu.edu/publications.php>) tools. Four different tools were used to confirm the final outcome without discrepancies. **PatchDock** predicts both wild and mutant type protein-protein interaction with an algorithm using recognition of object and image segmentation techniques through Computer Vision. The protein structures are taken as input and in a three-step process which comprises of-i) Molecular Shape representation, ii) Surface Patch Mixing, Filtering and scoring, and iii) complementary surfaces matching and superimposition; a final output scores is given. The results are quicker to obtain and sorted by shape complementarity [36]. The **HDOCK** server is a highly integrated suite for robust and fast protein-protein/protein-RNA/protein-DNA docking. The interaction, based on the input information of both receptor and ligand (amino acid sequences or protein structures), is carried out through a hybrid algorithm of template based and *ab initio* free docking. The tool returns top 100 predicted complex structures, of which the first is the best-fit [37]. **ZDOCK** utilizes Fast Fourier Transform (FFT) Algorithm to perform rigid body protein-protein docking and generates the predictive scores in terms of shape complementarity and, electrostatic potential [38]; Chen and Weng, 2002). **ClusPro**, the protein-protein docking tool investigates the interaction between two proteins and helps in evaluation of protein structure stability in case of mutant proteins. The docking is performed by considering all known energy parameter sets that generate predictive models which gives rise to a stable structure with low energy values [39].

3. Results

The non-synonymous SNPs in *POT1* gene, selected from COSMIC database, were analysed using nine online computational tools. The highly damaging substitutions were presumed to drive cancer progression by affecting the protein structure and function and by imparting cellular immortality. However the biological mechanisms by which such amino acid variants result in phenotypic change are not completely clear. *In silico* analysis is a powerful tool that facilitates the prediction of effects of non-synonymous SNPs on physicochemical properties of proteins. This information generates a working knowledge of genotype-phenotype correlation in disease biology.

3.1. Prediction of functional scores of nsSNPs by in-silico tools

About 4.73% (16 out of 338) missense substitutions in *POT1* were found to be deleterious using functional score predictive tools. All the 16 nsSNPs were predicted to be damaging with a PSIC (position-specific independent counts) score between 0.577 and 1.000 (Table 2). PANTHER tool predicted that of the 16 mutations, 12 were likely to perturb protein functions of *POT1* since the probability score of deleterious effect was higher than 0.5 (Table 2). The program SIFT predicted 15 of 16 mutations in *POT1* as functionally damaging (score<0.05). SNAP indicated all 16 mutations of *POT1* as highly damaging (score>0.5). The disease-related and neutral mutations as classified by PhD-SNP and SNPs&GO predicted 15 and 10 nsSNPs to be deleterious respectively (Table 2).

The disruption of normal *POT1* functions was predicted by MutPred generating a probability score with g-value greater than 0.5 and p-value less than 0.05; highlighting change in features such as-methylation, ubiquitination, phosphorylation, catalytic activity, solvent accessibility, and secondary structure (Tables 2 and 4). These variants were reported for their role in various solid tumors and haematological cancer (Table 7; Fig. 1).

Table 2Functional scores of deleterious missense substitutions of *POT1* gene mined from COSMIC database.

Mutations	PolyPhen	PANTHER	PhD-SNP	SIFT	SNAP	SNPs&GO	MutPred	PredictSNP	MAPP
L14P	1.000, D (damaging)	NA	0.830, D	0.000, D	0.635, D	0.447, N (neutral)	0.759	87%, D	88%, D
N24S	0.577, D	NA	0.665, D	0.010, D	0.675, D	0.373, N	0.887	66%, D	63%, D
N24T	0.971, D	NA	0.780, D	0.000, D	0.705, D	0.425, N	0.873	87%, D	72%, D
C44Y	1.000, D	NA	0.854, D	0.000, D	0.695, D	0.548, D	0.732	87%, D	81%, D
Y223C	1.000, D	0.854, D	0.864, D	0.000, D	0.755, D	0.758, D	0.877	87%, D	77%, D
G268C	1.000, D	0.874, D	0.878, D	0.000, D	0.690, D	0.823, D	0.716	76%, D	81%, D
G272V	1.000, D	0.799, D	0.879, D	0.000, D	0.685, D	0.845, D	0.406	87%, D	84%, D
P357S	1.000, D	0.660, D	0.344, N	0.020, D	0.615, D	0.352, N	0.579	61%, D	57%, D
H437P	0.999, D	0.781, D	0.865, D	0.090, N	0.655, D	0.840, D	0.663	87%, D	88%, D
V439G	1.000, D	0.687, D	0.703, D	0.000, D	0.735, D	0.558, D	0.740	87%, D	41%, D
P475L	1.000, D	0.742, D	0.669, D	0.000, D	0.600, D	0.571, D	0.821	79%, D	77%, D
G534C	0.999, D	0.704, D	0.845, D	0.020, D	0.625, D	0.692, D	0.586	76%, D	84%, D
P537S	0.999, D	0.580, D	0.637, D	0.000, D	0.640, D	0.432, N	0.698	72%, D	63%, N
F566C	1.000, D	0.763, D	0.735, D	0.000, D	0.770, D	0.571, D	0.760	76%, D	59%, D
M587T	0.998, D	0.418, N	0.561, D	0.000, D	0.730, D	0.387, N	0.470	76%, D	84%, D
C591W	0.999, D	0.754, D	0.875, D	0.010, D	0.705, D	0.621, D	0.618	87%, D	66%, D

*D-damaging, N- neutral, %- expected accuracy.

Table 3Analysis of stability effect of *POT1* variants by I-Mutant Suite.

Mutations	I-Mutant $\Delta\Delta G$ (Kcal/mol)	
	Score	Effect
L14P	-2.37	Destabilizing
N24S	-1.23	Destabilizing
N24T	-0.23	Destabilizing
C44Y	0.13	Stabilizing
Y223C	1.05	Stabilizing
G268C	-0.93	Destabilizing
G272V	-0.17	Destabilizing
P357S	-0.09	Destabilizing
H437P	-1.16	Destabilizing
V439G	-4.78	Destabilizing
P475L	-0.96	Destabilizing
G534C	-0.76	Destabilizing
P537S	-2.4	Destabilizing
F566C	-1.92	Destabilizing
M587T	-0.61	Destabilizing
C591W	-1.17	Destabilizing

3.2. Stability effect analysis

I-Mutant 3.0 Suite predicted changes in free energy viz. $\Delta\Delta G$ classifying 14 of the 16 nsSNPs of *POT1* with decreased stability ($\Delta\Delta G < 0$). The two mutations C44Y and Y223C each in the C-terminus and N-terminus respectively, with increased stability ($\Delta\Delta G > 0$) (Table 3). Thus it might be possible for the pathogenic substitutions C44Y and Y223C not

Table 4Detailed MutPred results of *POT1* variants.

Mutations	Glycosylation	Methylation	Phosphorylation	Catalytic Residue	Ubiquitination sites	MoRF binding	Secondary structure	Solvent accessibility
L14P	Gain	Gain	-	-	Gain	-	Loss of helix	-
N24S	-	-	Gain	-	-	Gain	Loss of sheet/gain of loop	Gain
N24T	-	-	Gain	-	-	-	Loss of sheet/gain of loop	Gain
C44Y	-	Gain	Gain	Loss	Gain	-	-	-
Y223C	-	-	Loss	Loss	-	-	Loss of sheet, gain of loop	-
G268C	-	Loss	-	Loss	-	-	Gain of sheet/loss of loop	-
G272V	-	Loss	Loss	-	-	-	Loss of sheet/loop	-
P357S	Loss	Loss	Gain	-	-	Gain	-	-
H437P	-	Gain	-	-	Gain	Loss	-	-
V439G	-	Loss	-	-	Loss	Loss	-	-
P475L	-	-	Loss	Loss	-	-	Loss of loop	-
G534C	-	-	-	Gain	-	-	Loss of sheet/gain of helix, loop	-
P537S	-	-	Loss	Loss	-	-	Gain of sheet/loss of loop, helix	-
F566C	-	-	-	Loss	Loss	-	Loss of sheet/gain of loop, helix	-
M587T	-	-	Loss	Loss	-	-	Loss of helix	-
C591W	-	-	-	Loss	Gain	-	-	-

to influence the overall stability of *POT1* but disrupt or affect *POT1* dynamics and its *POT1*-TPP1/ssDNA interactive networks.

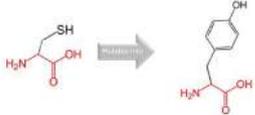
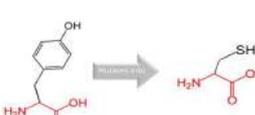
3.3. HOPE prediction

Project HOPE determined the structural changes of protein upon mutation with changes in size, charge, hydrophobicity, salt-bridge and hydrogen bond formation (Table 5). For most of the substitutions differences in size and hydrophobicity between native and mutant amino acid residues were directly affecting the hydrogen bond formation. Similarly, the differences in charge were seen to alter the ionic interactions.

3.4. Analysis of conserved residues of *POT1*

The ConSurf results predicted that phylogenetically, residues L14, N24, C44, Y223, G268, G272, H437, V439, P357, P475, G534, F566, M587 and C591, were highly conserved and P537 was moderately conserved (Fig. 2). The highly conserved amino acid residues are not prone to frequent mutational events, however, in case there is a mutation at such a residue, it can cause a severe impact on protein structure, function and interaction. On the contrary low or moderately conserved residues might show no significant impact in changing the protein upon mutations. In this study all except one mutation had some adverse effect on the protein (see Fig. 3).

Table 5
HOPE prediction of selected deleterious nsSNPs.

Variants	Structural changes	Altered properties
C44Y		The mutant residue is bigger than the wild-type residue. The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit. The hydrophobicity of the wild-type and mutant residue differs. The mutation will cause loss of hydrophobic interactions in the core of the protein.
L14P		The mutant residue is smaller than the wild-type residue. The mutation will cause an empty space in the core of the protein.
N24S		The mutant residue is smaller than the wild-type residue. The mutation will cause an empty space in the core of the protein. The hydrophobicity of the wild-type and mutant residue differs. The mutation will cause loss of hydrogen bonds in the core of the protein and as a result disturb correct folding
N24T		The mutant residue is smaller than the wild-type residue. The mutation will cause an empty space in the core of the protein. The hydrophobicity of the wild-type and mutant residue differs. The mutation will cause loss of hydrogen bonds in the core of the protein and as a result disturb correct folding
G272V		The mutant residue is bigger than the wild-type residue. The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit. The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure
G268C		The mutant residue is bigger than the wild-type residue. The residue is located on the surface of the protein, mutation of this residue can disturb interactions with other molecules or other parts of the protein. The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure
Y223C		The wild-type residue forms a hydrogen bond with: Histidine at position 264. The size difference between wild-type and mutant residue makes that the new residue is not in the correct position to make the same hydrogen bond as the original wild-type residue did. The difference in hydrophobicity will affect hydrogen bond formation. The mutant residue is smaller than the wild-type residue. This will cause a possible loss of external interactions. The hydrophobicity of the wild-type and mutant residue differs
P357S, P537S		Prolines are known to be very rigid and therefore induce a special backbone conformation which might be required at this position. The mutation can disturb this special conformation. The mutant residue is smaller than the wild-type residue. This will cause a possible loss of external interactions. The wild-type residue is more hydrophobic than the mutant residue; Sometimes, hydrophobicity is important for multimerisation and therefore this mutation could affect the multimer contacts.
H437P		The mutant residue is smaller and more hydrophobic than the wild-type residue. The wild-type residue forms a hydrogen bond with: Serine at position 421. The mutation will cause an empty space in the core of the protein and loss of hydrogen bonds in the core of the protein and as a result disturb correct folding.
V439G		The mutant residue is smaller than the wild-type residue. The wild-type residue is more hydrophobic than the mutant residue. The mutation introduces a glycine at this position. Glycines are very flexible and can disturb the required rigidity of the protein at this position. The mutation will cause an empty space in the core of the protein. The mutation will cause loss of hydrophobic interactions in the core of the protein.
P475L		The mutant residue is bigger than the wild-type residue. The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit.
G534C		The wild-type residue is a glycine, the most flexible of all residues. This flexibility might be necessary for the protein's function. Mutation of this glycine can abolish this function. The mutant residue is bigger than the wild-type residue. The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit. The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure.
F566C		The mutant residue is smaller than the wild-type residue; resulting an empty space in core of the protein. The wild-type residue is predicted to be located in its preferred β -strand. The mutant residue prefers to be in another secondary structure, therefore the local conformation will be slightly destabilized.
M587T		The mutant residue is smaller and more hydrophobic than the mutant residue. The mutation will cause loss of hydrophobic interactions in the core of the protein.
C591W		The mutant residue is bigger than the wild-type residue. The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit.

3.5. Analysis of interatomic interactions

The atomic interactions of the various variants of *POT1* were categorically demarcated by the Arpeggio web server. Accordingly the variants were characterized as per their interatomic interactions between the residues in normal and mutant. The results of the mutant residues were deviated from the normal in-i) Hydrogen bonds (L14P, G268C, H437P, G534C, C591W), ii) hydrophobic (L14P, G272V, H437P, V439G, P475L, G534C, P537S, F566C, M587T, C591W, N24 S/T, C44Y, Y223C), iii) ionic bonds (H437P), iv) van der Waals interactions (L14P,

N24 S/T, C44Y, G272V, H437P, F566C, M587T, C591W), v) polarity (L14P, Y223C, G268V, H437P, M587T, C591W), vi) aromatic contacts (G272V, H437P, C44Y, Y223C, C591W) contributing to the instability (Table 6).

3.6. Analysis of protein secondary structure

Among 16 highly deleterious nsSNPs, SOPMA predicted altered secondary structures that lead to 25% residues to be in α -helices (L14, V439, P537, M587), 31.2% in random coils (C44, G272, G268, P357,

C591) and extended strands (N24, Y223, H437, P475, F566), and 6.2% in β -turns (G534).(Fig. 2).

3.7. Protein-protein docking

Through the present study we try to establish wild type and mutant POT1-TTP1 interactions. POT1 C-terminal variants in the OB3 domain and POT1 binding domain (PBD) of TPP1 were docked. Four comparative tools - PatchDock, HDOCK, ZDOCK and ClusPro provided a quantitative assessment on binding efficacy between native and mutant protein models. The native POT1-TTP1 structure was found to have a lowest energy value, indicating towards a 'stable and well docked complex' as compared with the mutant-native (POT1-TTP1) protein interactive pairs (Table 8A). Here, in HDOCK the mutant-native (POT1-TTP1) scores deviated considerably from the wild type due to the algorithm which allows inbuilt modelling of the proteins from the sequence. The other three docking tools showed minute deviations in the scores between mutant-native and native-native docking.

Further, POT1 N-terminus mutant residues were docked to the single stranded telomeric DNA in native and mutant configurations. The scores were generated using HDOCK that has unique algorithm permitting the protein-nucleic acid docking. The mutants gave a score deviated from the normal docking.

4. Discussion and conclusion

Identification and characterization of non-synonymous SNPs is one of the central objectives in molecular biology. This has helped in discovery of diagnostic markers and to study molecular targets in pharmacotherapy by providing insights into cancer biology [40]. The nsSNPs in coding regions changes the amino acids which further deviates the protein structure and function that may be accountable for disease pathogenicity. Many experiments have established the importance of correlating nsSNPs with protein expression, stability, folding, interactions and drug response [41–43]. With an exponential increase in high throughput data and discovery of nsSNPs in cancers, it has become increasingly difficult to explore individual biological significance by wet-lab experiments. *In silico* platform has made it easier to identify and predict the deleteriousness of selective SNPs from the pool of mutations recorded in various databases. In the present study, we make an effort to systematically analyse damaging nsSNPs of *POT1* gene using computational tools for various solid tumors and haematologic malignancies. We also predict the changes that affect protein-protein/protein-DNA interactions of POT1-TTP1/POT1-ssDNA due to mutations in the *POT1* gene. The missense mutations screened from COSMIC database were

Table 6

Arpeggio prediction of interatomic interactions of native and mutant POT1.

Native vs. Variants	Hydrogen bonds	Hydrophobic contacts	van der Waals interactions	Ionic interactions	Polar contacts	Aromatic contacts
1XJV ^a	271	751	147	28	374	58
L14P	270	744	146	28	372	58
N24S	271	749	147	28	374	58
N24T	271	752	148	28	374	58
C44Y	271	765	149	28	374	65
Y223C	271	745	147	28	373	56
G268C	274	751	147	28	377	58
G272V	271	762	148	28	374	28
5H65 ^a	253	675	158	31	366	8
P357S	253	675	158	31	366	8
H437P	250	678	159	28	363	5
V439G	253	661	158	31	366	8
P475L	253	691	158	31	366	8
G534C	254	677	158	31	368	8
P537S	253	671	158	31	366	8
F566C	253	650	157	31	366	5
M587T	253	663	157	31	367	8
C591W	254	685	160	31	365	14

^a - native *POT1* protein structure.

subjected to computational tools that used algorithms based on 'evolutionary, structural and computational methods'. Thus we could obtain a single consensus accurate prediction through a combinatorial approach relying on diverse algorithms. The initial screening with different computational algorithms helped in identification of 16 highly damaging mutations corresponding to *POT1* gene (Table 2).

The HOPE server gave a more precise understanding of the variants in structural and functional context. Each amino acid, upon mutation, exhibited an altered size, charge and hydrophobicity (Table 5). If the mutant residue misfolds to a larger protein as seen in substitutions C44Y, G272V, G268C, P475L, G534C, and C591W, there is a possibility that it might not fit in the core of protein, leading to bumps and incorrect protein-protein interactions with loss of hydrogen bond. On the contrary, a smaller sized mutant variant such as L14P, N24S, N24T, Y223C, P357S, P537S, H437P, V439G, F566C, and M587T might result in an empty space in core of protein leading to loss of external interaction. In both cases, it will affect the multimeric interactions. The deleterious mutations observed in our study were found to disrupt the native hydrogen bond which resulted in disruption of local structure required for allosteric regulation followed by exposure of hydrophobic core to water molecule. Hydrogen atoms form the essential components contributing towards the atomic structure of macromolecules. Polarity of hydrogen atoms containing partial positive charge is largely responsible for the formation of hydrogen bonds; the network of which participates in numerous biological functions. The substitution of amino acid to tyrosine(Y), serine (S), threonine (T), cysteine (C), and proline (P) showed changes in H-bond formation impacting the interaction with atomic moieties. Likewise, substitutions can bring about a change in the charge of the protein and can affect the ionic interactions. Besides, conformational changes in protein resulted in gain of solvent

Table 7

Diseases associated with selected deleterious ns SNPs of *POT1*.

Mutation	Disease	Mutation	Disease
L14P	Adenocarcinoma	H437P	Malignant melanoma
N24S	Breast carcinoma	V439G	Malignant melanoma
N24T	Squamous cell carcinoma	P475L	Squamous cell carcinoma
C44Y	Mantle cell lymphoma	G534C	adenocarcinoma
Y223C	Chronic lymphocytic leukaemia	P537S	Squamous cell carcinoma
G268C	adenocarcinoma	F566C	Malignant melanoma
G272V	Chronic lymphocytic leukaemia	M587T	Tripple negative breast carcinoma
P357S	Thyroid carcinoma	C591W	Chronic lymphocytic leukaemia

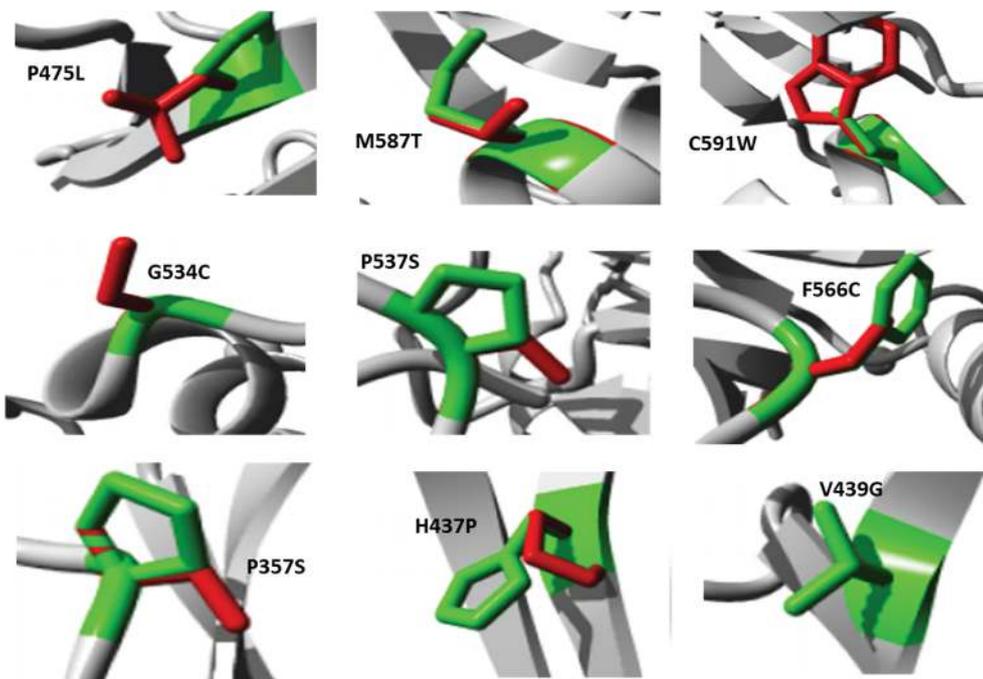


Fig. 1. POT1 variants that interacts with TPP1.

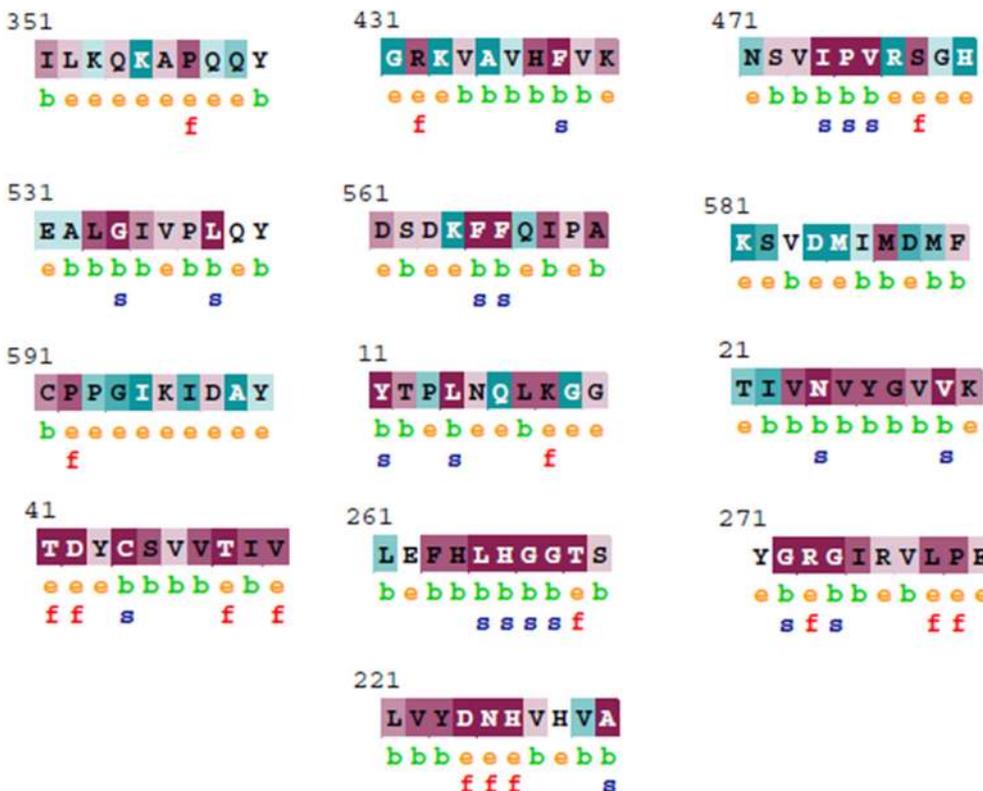


Fig. 2. Consurf analysis of POT1 variants.

accessibility (N24S, N24T) which in turn can affect the overall entropy of the system in bound state further disturbing the free energy of binding [44].

The adverse effects of structural deformities along with altered post-translational modifications (PTMs) triggered by each missense substitutions using MutPred was interpreted. Unsurprisingly, almost all of these substitutions were found to contribute toward a gain or loss of

secondary structures as well as changes in the PTMs like methylation, glycosylation, phosphorylation, ubiquitination. MoRF binding, solvent accessibility and catalytic residues were also altered. These changes cumulatively contributed towards fluctuations in gene expression, cellular differentiation, protein folding, normal protein function, signaling, protein degradation, and ligand-binding [44,45]. Consequently, such susceptible coding variants narrow the gap between

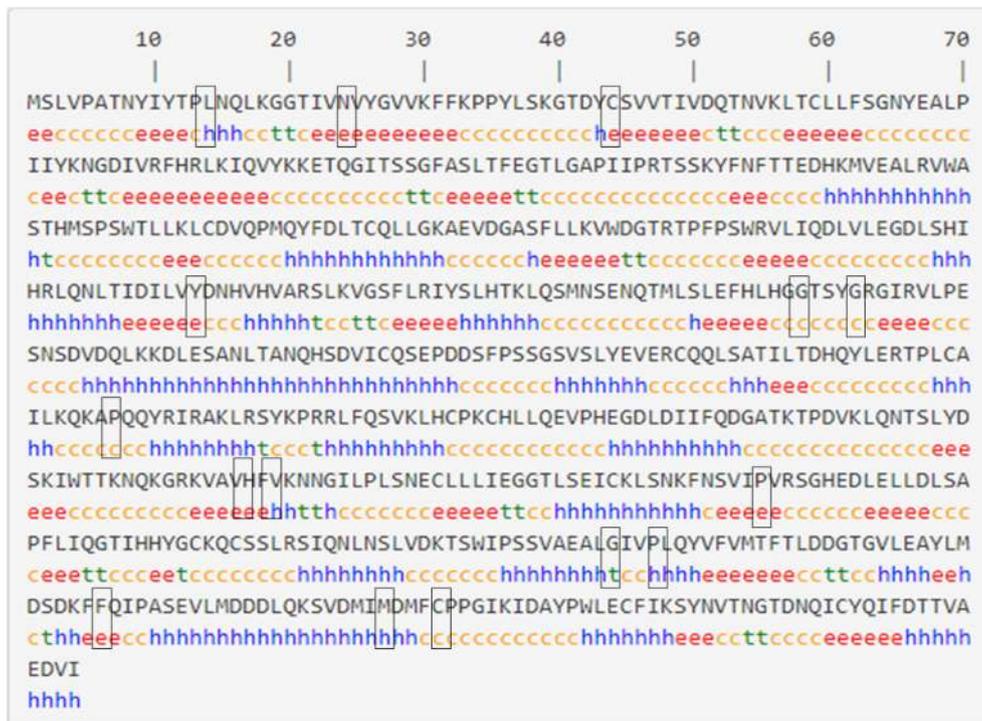


Fig. 3. POT1 protein secondary structure prediction by SOPMA.

Table 8A

Protein-protein interaction results of C-terminal OB3 domain of POT1 with POT1 binding domain of TPP1.

Variants	PatchDock		ClusPro		HDOCK	ZDOCK
	Score	ACE	Centre energy	Lowest energy		
POT1 _w /TPP1 _w	25,164	-462.65	-1695.2	-2268.0	-902.88	2547.740
P357S/TPP1 _w	25,164	-462.65	-1695.2	-2268.0	-882.32	2547.820
H437P/TPP1 _w	25,164	-462.65	-1695.5	-2268.1	-878.28	2547.695
V439G/TPP1 _w	25,164	-462.65	-1695.2	-2268.0	-862.23	2547.739
P475L/TPP1 _w	25,164	-462.65	-1695.1	-2267.9	-909.14	2547.748
G534C/TPP1 _w	20,748	-363.47	-1695.2	-2268.0	-861.01	2547.746
P537S/TPP1 _w	20,748	-363.47	-1695.1	-2268.0	-920.49	2547.737
F566C/TPP1 _w	25,164	-462.65	-1695.3	-2268.0	-857.11	2540.715
M587T/TPP1 _w	20,748	-363.47	-1695.2	-2267.9	-861.60	2547.747
C591W/TPP1 _w	25,164	-462.65	-1695.3	-2268.1	-870.02	2547.753

*w-wild/native.

protein changes and disease outcomes. The mutations in *POT1* plausibly affected the protein function by compromising telomeres integrity and activating cellular repair mechanisms in telomeres. This in turn can lead to cell immortality and halt cellular senescence causing progression of cancer.

Gain or loss of phosphorylation sites are the major underlying mechanisms in deregulation of signal transduction mediated by altered kinase or phosphatase function with a direct effect on gene expression. Substitutions of N24S, N24T, C44Y, and P357S were found to be enriched in gain of phosphorylation sites which holds true for most of the cancer-associated substitutions. This can have impact on enzyme catalysis through a systematic deregulation of normal protein function [45]. The amino acids, directly involved in the process of catalysis influence the reactions and any substitution in these residues can shift the protein kinetics [46]. A gain of catalytic residue was observed in G534C of POT1 Holiday Junction Like Resolvase (HJRL) domain which may change the rate of reaction and binding affinity of the protein with TPP1. The conservation analysis showed 14 out of 16 residues to be highly conserved. Usually, mutations of highly conserved residue is intolerable for protein stability. The majority of disease causing variants are seen to be located in α -helices and random coils [47,48]. In our study, SOPMA

predicted approximately 25% and, 31.2% residues of POT1 protein to be located in α -helices, and extended strands-random coils respectively.

All the 16 variants showed an altered intratomic interactions in terms of hydrogen bonds, aromatic and polar contacts, hydrophobic, ionic and van der Waals interactions. Protein-protein docking of native and mutant models by PatchDock, HDOCK, ZDOCK and ClusPro confirmed the interaction of native POT1 OB3 fold with POT1 binding domain of TPP1 to be highly stable as defined by lowest energy values of -462.65 kcal/mol, -902.88 kcal/mol, 2547.740 kcal/mol and -2268.0 kcal/mol respectively (Table 8A). In PatchDock, the docking analysis of interactive damaging variants of POT1- G534C, P537S and M587T showed relatively higher energy values (-363.47 kcal/mol) indicating non-favourable conjugation possibly due to conformational changes of POT1 protein, enforcing the neighboring residues toward an antagonistic interaction. In HDOCK variants P475L (-909.14 kcal/mol) and P357S (-920.49) showed higher binding efficacy as compared to native POT1-TPP1 interactive pairs (-902.88 kcal/mol). However, both in ClusPro and ZDOCK, the mutant-native interactive pairs showed minor drift in energy values. Since the docking tools use different algorithms, the scores obtained cannot be consistent and comparable although the defined outcome may be in synchrony. Our results indicate that the

Table 8B

Protein-ssDNA interaction of POT1 N-terminal OB fold mutants to the telomeric DNA.

Variants	HDOCK Docking score
POT1 _w /ssDNA	-710.36
C44Y/ssDNA	-727.90
L14P/ssDNA	-694.81
G268C/ssDNA	-696.91
G272V/ssDNA	-708.98
N24S/ssDNA	-665.13
N24T/ssDNA	-699.34
Y223C/ssDNA	-648.81

*w-wild/native.

selected mutations are causing low impact damage to the POT1 structure and probably function. The POT1-TPP1 binding may grossly be affected as perceptible at interatomic interactions. In our study, we found residues L14, N24, C44 in POT1 N_{OB1} fold; Y223, G268 and G272 in POT1 N_{OB2} fold; P357, F566, M587 and C591 in POT1 C_{OB3} fold and H437, V439, P475, G534, P537 in HJRL domain. The structural inspection of POT1 revealed that single or multiple SNPs can alter the natural state of POT1-TPP1 interaction with a change in POT1 C-fold which in turn can result in dysfunctional telomeric capping complex resulting in telomeric instability and cancer cell immortality. Experimental evidence showed that POT1C mutants C591W and P475L cause a significant decrease and/or loss in TPP1 binding. The P475 residue of HJRL contributes to fold of this domain whereas the C591 of C terminus (helix α 8) spans the entire length of the side of the OB3 fold and is critical for organization of POT1C_{OB3}. Therefore displacement of the POT1 helix α 8 upon mutation would lead to reorganization of POT1C_{OB3} thus affecting TPP1 binding. Besides both C591W and P475L showed a significant increase (~10%) of fragile telomeres (Rice et al., 2016). Biochemical and structural analyses revealed M587T mutant to be tolerable for POT1 stability, protein-folding, telomere-localization and TPP1-binding [8]. The mechanistic role of C-terminal mutations which display hallmarks of dysfunctional telomeres is unclear but could be due to complete or partial loss of functional OB folds [49–51], loss of interaction with TPP1 or decreased expression of POT1 protein [13,52,53,54,55,57]. The segments 5'-TTAGGG and its downstream TTAG-3' are recognized by OB1 and 2 folds respectively which remain unaffected on dimerization of POT1 and TPP1 [56].

The docking of N-terminal variants of POT1 with ssDNA (5'-TTAGGGTTAG-3') showed an altered binding efficacy; while the interaction of native POT1 with ssDNA was found to have lowest binding energy suggesting best docked complex. This indicates that the mutant variants of POT1 still binding to telomeric DNA can perhaps destabilize the protein-DNA complex which is stimulated by hydrophobic interactions between bases and aromatic side chains of the amino acids (Table 8B) [58]. The majority of cancer related SNPs of POT1 tend to be located on N-terminal OB domains of POT1 [52]. Identified 12 somatic mutations in CLL of which Y223C, and G272V (OB2) belong to the N-terminal of POT1 and were predicted to be deleterious as a result of either complete or partial termination of POT1-telomere interaction interface. Y223C variant was found to affect the binding efficacy of POT1 with terminal guanine bases of telomeric DNA. In our study too these mutations showed deviated scores indicating similar outcomes. Further, as per research conducted by various scientific groups, it is hypothesized that the N-terminal mutations of OB1 and 2 folds could be associated with chromosomal abnormalities such as irregular telomeres length, fragile telomeres and chromosome end-to-end fusions, all of which point toward telomeres uncapping and recruitment of repair proteins at the telomeres.

Since, the DNA binding and the dimerization of POT1-TPP1 is solely dependent upon POT1 protein, recognition and elucidating biological significance of human POT1 mutants can help to decipher their role in cancers.

In summary, we analysed POT1 nsSNPs in cancer to predict the most damaging mutations followed by its interaction with TPP1 and single strand telomeric DNA. The detailed results emphasized the importance of highly deleterious SNPs altering protein structure and function. Experimental research is worth carrying out in the future through techniques like site-directed mutagenesis so as to prioritize SNPs as molecular markers in cancer diagnostics.

Financial support

No.

Ethical statement

We do not have any ethical statement to declare for the aforementioned work since it is a bioinformatics based work and no animals and/or patients are associated with the study.

Declaration of competing interest

I wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgement

This work was not supported financially by any agency. We would like to thank all the authors and VIT, Vellore for their support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.imu.2020.100389>.

References

- [1] Rice C, Shastrula PK, Kossenkov AV, et al. Structural and functional analysis of the human POT1-TPP1 telomeric complex. *Nat Commun* 2017;8:14928.
- [2] Xin H, Liu D, Songyang Z. The telosome/shelterin complex and its functions. *Genome Biol* 2008;9:232.
- [3] Patel TN, Vasani R, Gupta D, et al. Shelterin proteins and cancer. *Asian Pac J Cancer Prev APJCP* 2015;16:3085–90.
- [4] Hockemeyer D, Palm W, Else T, et al. Telomeres protection by mammalian POT1 requires interaction with TPP1. *Nat Struct Mol Biol* 2007;14:754.
- [5] Sfeir A. Telomeres at a glance. *J Cell Sci* 2012;125:4173–8.
- [6] Wang F, Podell ER, Zaug AJ, et al. The POT1-TPP1 telomeres complex is a telomerase processivity factor. *Nature* 2007;445:506–10.
- [7] Kibe T, Osawa GA, Keegan CE, et al. Telomeres protection by TPP1 is mediated by POT1a and POT1b. *Mol Cell Biol* 2010;30:1059–66.
- [8] Chen C, Gu P, Wu J, et al. Structural insights into POT1-TPP1 interaction and POT1 C terminal mutations in human cancer. *Nat Commun* 2017;8:14929.
- [9] Xu M, Kiselar J, Whitted TL, et al. POT1-TPP1 differentially regulates telomerase via POT1 His266 and as a function of single-stranded telomeres DNA length. *Proc Natl Acad Sci Unit States Am* 2019;116:23527–33.
- [10] Chang N. Cancer chromosomes going to POT1. *Nat Genet* 2013;45:473–5.
- [11] Quesada V, Conde L, Villamor N, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* 2011;44:47–52.
- [12] Pinzaru AM, Hom RA, Beal A, et al. Telomere replication stress induced by POT1 inactivation accelerates tumorigenesis. *Cell Rep* 2016;15:2170–84.
- [13] Shi J, Yang XR, Ballew B, et al. Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. *Nat Genet* 2014;46:482–6.
- [14] Nandakumar J, Bell CF, Weidenfeld I, et al. The TEL patch of telomeres protein TPP1 mediates telomerase recruitment and processivity. *Nature* 2012;492:285–9.
- [15] Abreu E, Aritonovska E, Reichenbach P, et al. TIN2-tethered TPP1 recruits human telomerase to telomeres in vivo. *Mol Cell Biol* 2010;30:2971–82.
- [16] Sexton AN, Youmans DT, Collins K. Specificity requirements for human telomeres protein interaction with telomerase holoenzyme. *J Biol Chem* 2012;287:34455–64.
- [17] Zhong FL, Batista LF, Freund A, et al. TPP1 OB-fold domain controls telomeres maintenance by recruiting telomerase to chromosome ends. *Cell* 2012;150:481–94.
- [18] Schmidt JC, Zaug AJ, Cech TR. Live cell imaging reveals the dynamics of telomerase recruitment to telomeres. *Cell* 2016;166:1188–97.
- [19] Martinez P, Blasco MA. Role of shelterin in cancer and aging. *Aging Cell* 2010;9:653–6.

- [20] de Lange T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Gene Dev* 2005;19:2100–10.
- [21] Nandakumar J, Cech TR. Finding the end: recruitment of telomerase to telomeres. *Nat Rev Mol Cell Biol* 2013;14:69–82.
- [22] Baumann P, Cech TR. Pot1, the putative telomeres end-binding protein in fission yeast and humans. *Science* 2001;292:1171–5.
- [23] Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
- [24] Capriotti E, Altman RB, Bromberg Y. Collective judgement predicts disease-associated single nucleotide variants. *Mutations in proteins. BMC Genomics, Suppl* 2013;3:S2.
- [25] Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003;31:3812–4.
- [26] Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. *Bioinformatics* 2006;22:2729–34.
- [27] Calabrese R, Capriotti E, FariSELLI P, et al. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat* 2009;30:1237–44.
- [28] Mi H, Poudel S, Muruganujan A, et al. Panther version 10: expanded protein families and functions and analysis tools. *Nucleic Acids Res* 2016;44:D336–42.
- [29] Li B, Krishnan VG, Mort ME, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 2009;25:2744–50.
- [30] Bendl J, Stourac J, Salanda O, et al. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS Comput Biol* 2014;10:e1003440.
- [31] Stone EA, Sidow A. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res* 2005;15:978–86.
- [32] Venselaar H, Beek T, Kuipers R, et al. Protein structure analysis of mutations causing inheritable diseases. An e-science approach with life scientist friendly interfaces. *BMC Bioinf* 2010;8:548.
- [33] Glaser F, Pupko T, Paz I, et al. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* 2003;19:163–4.
- [34] Jubb HC, Higuero AP, Ochoa-Montaño B, et al. Arpeggio: a web server for calculating and visualising interatomic interactions in protein structures. *J Mol Biol* 2017;429:365–71.
- [35] Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci* 1995;11:681–4.
- [36] Schneidman-Duhovny D, Inbar Y, Nussinov R, et al. Patchdock and Symmdock: servers for rigid and symmetric docking. *Nucleic Acids Res* 2005;33:W363–367.
- [37] Yan Y, Tao H, He J, et al. The HDOCK server for integrated protein-protein docking. *Nat Protoc* 2020;15:1829–52.
- [38] Pierce BG, Wiehe K, Hwang H, et al. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* 2014;30:1771–3.
- [39] Kaur T, Thakur K, Singh J, et al. Identification of functional SNPs in human *LGALS3* gene by in silico analyses. *The Egyptian Journal of Medical Human Genetics* 2017;18:321–8.
- [40] Bhatnagar R, Dang AS. Comprehensive in-silico prediction of damage associated SNPs in Human Prolidase gene. *Nature Scientific Reports* 2018;8:9430.
- [41] Rotimi SO, Peter O, Oguntade O, et al. In silico analysis of the functional non-synonymous single nucleotide polymorphisms in the human *CYP27B1* gene. *The Egyptian Journal of Medical Human Genetics* 2018;19:367–78.
- [42] Calvete O, Garcia-Pavia P, Domingues F, et al. The wide spectrum of POT1 gene variants correlates with multiple cancer types. *Eur J Hum Genet* 2017;1–4.
- [43] Stefl S, Nishi H, Petukh M, et al. Molecular mechanisms of disease-causing missense mutations. *J Mol Biol* 2013;425:3919–36.
- [44] Yang Y, Peng X, Ying P, et al. AWESOME: a database of SNPs that affect protein post-translational modifications. *Nucleic Acids Res* 2019;47:D874–80.
- [45] Radivojac P, Baenziger PH, Kann MG, et al. Gain and loss of phosphorylation sites in human cancer. *ECCB* 2008;24:i241–7.
- [46] Xin F, Myers S, Li YF, et al. Structure-based kernels for the prediction of catalytic residues and their involvement in human inherited disease. *Bioinformatics* 2010;26:1975–82.
- [47] Kucukkall TG, Petukh M, Lin L, et al. Structural and physico-chemical effects of disease and NonDiseaseSNPs on proteins. *Curr Opin Struct Biol* 2015;32:18–24.
- [48] Zhang M, Huang C, Wang Z, et al. In silico analysis of non-synonymous single nucleotide polymorphisms (nsSNPs) in the human *GJA3* gene associated with congenital cataract. *BMC Molecular and Cell Biology* 2020;21:12.
- [49] Gong Y, Stock AJ, Liu Y. The enigma of excessively long telomeres in cancer: lessons learned from rare human POT1 variants. *Curr Opin Genet Dev* 2020;60:48–55.
- [50] Ye JZ, Hockemeyer D, Krutchinsky AN, et al. POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Gene Dev* 2004;18:1649–54.
- [51] Loayza D, De Lange T. POT1 as a terminal transducer of TRF1 telomere length control. *Nature* 2003;423:1013–8.
- [52] Ramsay AJ, Quesada V, Foronda M, et al. POT1 mutations cause telomeres dysfunction in chronic lymphocytic leukemia. *Nat Genet* 2013;45:526–30.
- [53] Robles-Espinoza CD, Harland M, Ramsay AJ, et al. POT1 loss-of-function variants predispose to familial melanoma. *Nat Genet* 2014;46:478–81.
- [54] Ferrandon S, Saultier P, Carras J, et al. Telomeres profiling: toward glioblastoma personalized medicine. *Mol Neurobiol* 2013;47:64–76.
- [55] Lin TT, Letsolo BT, Jones RE, et al. Telomeres dysfunction and fusion during the progression of chronic lymphocytic leukemia: evidence for a telomeres crisis. *Blood* 2010;116:1899–907.
- [56] Choi KH, Lakamp-Hawley AS, Kolar C, et al. The OB-fold domain 1 of human POT1 recognizes both telomeric and non-telomeric DNA motifs. *Biochimie* 2015;115:17–27.
- [57] Bainbridge MN, Armstrong GN, Gramatges MM, et al. Germline mutations in shelterin complex genes are associated with familial glioma. *J Natl Cancer Inst* 2015;107:384.
- [58] Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein-protein docking. *Nat Protoc* 2017;12:255–78.