Acta Medica 2019; 50(4): 42 - 47

acta medica

ORIGINAL ARTICLE

Evaluation of cancer related missense mutations in CENPH

Ceren Sucularlı* ,[PhD]
ORCID: 0000-0003-0627-7900

* Corresponding Author : Assist. Prof. Ceren Sucularli, PhD, Department of Bioinformatics, Institute of Health Sciences, Hacettepe University, 06100, Ankara, TURKEY
E-mail: ceren.sucularli@hacettepe.edu.tr
Tel: +90312 3053038
Fax: +90312 3051459

ABSTRACT

Objective: CENPH, centromere protein H, is one of the constitutive kinetochore proteins. High expression of CENPH has been shown in various forms of cancers; however, studies searching the effect of CENPH mutations in cancers are limited. Therefore, the aim of this study is to investigate the potential effects of the missense mutations of CENPH that have been identified in different cancers.

Materials and Methods: Missense CENPH mutations, which have been observed in cancers, were downloaded from the COSMIC v89. The effect of missense mutations was predicted by using PredictSNP1.0. The protein structure of the CENPH protein was generated with I-TASSER and missense mutations were visualized on CENPH protein with UCSF Chimera. Structural effects of selected mutations were assessed with HOPE.

Results: 34 missense mutations were observed in human cancers. Of the 34 missense mutations 18 mutations were predicted as deleterious and 16 mutations were predicted as neutral with ranging expected accuracies. Predicted missense mutations showed a scattered pattern on 3D CENPH protein. Two of the predicted deleterious missense mutations with higher expected accuracy were further analyzed and assessed according to amino acid properties.

Conclusion: This study provided a systematic analysis and evaluation of missense mutations on a CENPH protein that have been observed in different cancers.

Keywords: Kinetochore, Missense mutation, Prediction, Protein structure

Received: 29 July 2019, Accepted: 21 October 2019,
Published online: 31 October 2019

INTRODUCTION

CENPH is an essential component of the centromere and kinetochore complex and one of the constitutive kinetochore proteins [1]. During the cell cycle, CENPH localizes in the inner kinetochore plate with other constitutive kinetochore proteins [1]. CENPH is required for controlled kinetochore-microtubule plus-end dynamics and chromosome oscillations [2]. Loss of CENPH causes serious mitotic phenotypes, such as chromosome misalignment and multi-polar spindles [3].

Previous studies showed that CENPH has an important role in carcinogenesis of different cancers and is a promising prognostic marker. High CENPH expression has been shown in various cancers and CENPH has been proposed as a prognostic marker for patient survival of hepatocellular carcinoma, non-small cell lung cancer, renal cell carcinoma, nasopharyngeal carcinoma, breast cancer, and esophageal carcinoma [4-9]. In hepatocellular carcinoma cells knockdown of CENPH inhibited cell proliferation and induced apoptosis and growth arrest [10]. In renal cell carcinoma cell lines, CENPH knockdown caused a decrease in cell proliferation and an increase in apoptosis also inhibited cell growth, showing the importance of this protein in cancer progression [7].

Although detailed studies of CENPH expression alterations has been reported in human cancers, the effect of CENPH mutations in human cancers has not been studied in detail. In consideration of the importance of CENPH in carcinogenesis, it is necessary to investigate and classify CENPH mutations in

human cancers. Therefore, in this study missense mutations that were identified in different cancers of CENPH were extracted and their potential effect was investigated.

MATERIALS and METHODS

Obtaining missense mutations from Catalogue of Somatic Mutations in Cancer (COSMIC) database

Missense mutations of CENPH that were observed in cancer samples were downloaded from the COSMIC v89 (https://cancer.sanger.ac.uk/cosmic, [11]).

Predicting the effect of missense mutations

PredictSNP1.0 (https://loschmidt.chemi.muni.cz/predictsnp1/, [12]) was used to predict the effect of missense mutations. This tool is a consensus classifier that integrates the predictions of in silico prediction tools. For prediction analysis, protein sequence of CENPH (NP_075060.1) in fasta format from NCBI database (https://www.ncbi.nlm.nih.gov, [13]) and missense mutations from COSMIC v89 [11], were submitted to PredictSNP1.0 [12]. As a result, the PredictSNP prediction and predictions of each integrated tool were reported for missense mutations as neutral or deleterious along with expected accuracies.

In silico prediction of CENPH protein structure

The complete protein structure of the CENPH protein was predicted using I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/, [14-16]). The model with highest C-score was selected for CENPH model. CENPH protein and predicted deleterious mutations were visualized by using UCSF Chimera (http://www.rbvi.ucsf.edu/chimera, [17]). Dunbrack rotamer library [18] was used to display rotamers for selected missense mutations in UCSF Chimera.

Structural effects of selected missense mutations on CENPH protein

The structural effects of p.R105W and p.V237F on CENPH protein were analyzed by using HOPE server (http://www.cmbi.ru.nl/hope/, [19]).

RESULTS

Mutations of CENPH from COSMIC database

Cancer associated CENPH mutations were investigated by using COSMIC database v89. The information for total 49 mutations; 34 missense, 7 coding silent, 1 nonsense substitutions and 3 frameshift deletions, 1 frameshift insertion and 3 unknown mutations, which had no information at the amino acid level were obtained for CENPH. The 34 missense mutations, which cause a change in amino acid of the protein, were selected for further analysis.

Prediction of deleterious mutations of CENPH

In order to assess the effect of 34 missense mutations on CENPH, PredictSNP1.0 tool was used. Of the 34 missense mutations, 18 mutations were predicted as deleterious, while 16 mutations were predicted as neutral by PredictSNP1.0 (Table 1).

Table 1. The PredictSNP prediction and expected accuracy results of 34 CENPH missense mutations.

CDS Mutation	AA Mutation	Mutation ID (COSM)	PredictSNP prediction	Expected accuracy
c.4G>A	p.E2K	33151	DELETERIOUS	0.51
c.29C>T	p.A10V	6393548	NEUTRAL	0.83
c.62G>T	p.R21L	4638766	NEUTRAL	0.74
c.95A>G	p.Q32R	6398012	NEUTRAL	0.65
c.116G>T	p.R39L	6321194	NEUTRAL	0.65
c.140G>C	p.R47T	7099599	DELETERIOUS	0.55
c.180G>A	p.M60I	3856099	NEUTRAL	0.83
c.196G>A	p.E66K	6864852	DELETERIOUS	0.65
c.235G>A	p.E79K	1438372	DELETERIOUS	0.65

c.239C>T	p.A80V	3856100	NEUTRAL	0.83
c.251A>T	p.D84V	6171577	NEUTRAL	0.83
c.262G>C	p.E88Q	6104363	DELETERIOUS	0.65
c.268G>A	p.E90K	3072066	DELETERIOUS	0.61
c.313A>T	p.R105W	6104362	DELETERIOUS	0.87
c.314G>A	p.R105K	6900614	DELETERIOUS	0.51
c.314G>T	p.R105M	1438373	DELETERIOUS	0.72
c.368C>T	p.S123F	3920291	NEUTRAL	0.63
c.403G>C	p.E135Q	7120230	NEUTRAL	0.83
c.414A>C	p.K138N	328568	NEUTRAL	0.63
c.430C>A	p.Q144K	6426334	DELETERIOUS	0.55
c.472A>G	p.R158G	738326	DELETERIOUS	0.65
c.477G>T	p.K159N	3429596	DELETERIOUS	0.72
c.553G>T	p.D185Y	6398013	DELETERIOUS	0.65
c.568G>A	p.E190K	7179206	NEUTRAL	0.83
c.570A>C	p.E190D	1200752	NEUTRAL	0.74
c.581G>A	p.R194K	4895150	NEUTRAL	0.83
c.587A>G	p.K196R	3366136	NEUTRAL	0.65
c.588G>T	p.K196N	1069685	DELETERIOUS	0.61
c.686A>G	p.E229G	3366137	DELETERIOUS	0.61
c.691C>A	p.P231T	7186878	DELETERIOUS	0.61
c.695C>A	p.A232D	3669472	NEUTRAL	0.74
c.709G>T	p.V237F	7314207	DELETERIOUS	0.87
c.730G>T	p.V244F	3366138	DELETERIOUS	0.52
c.738G>T	p.M246I	6171576	NEUTRAL	0.74

Visualization of the predicted deleterious mutations on CENPH protein structure

In order to evaluate predicted deleterious missense mutations on CENPH protein, the protein structure of CENPH was predicted by using I-TASSER and visualized by using UCSF Chimera. The localizations of the residues, which carry predicted deleterious missense mutations, were shown on 3D CENPH protein model (Figure 1). According to distribution pattern the predicted missense mutations were scattered along the whole protein structure (Figure 1).

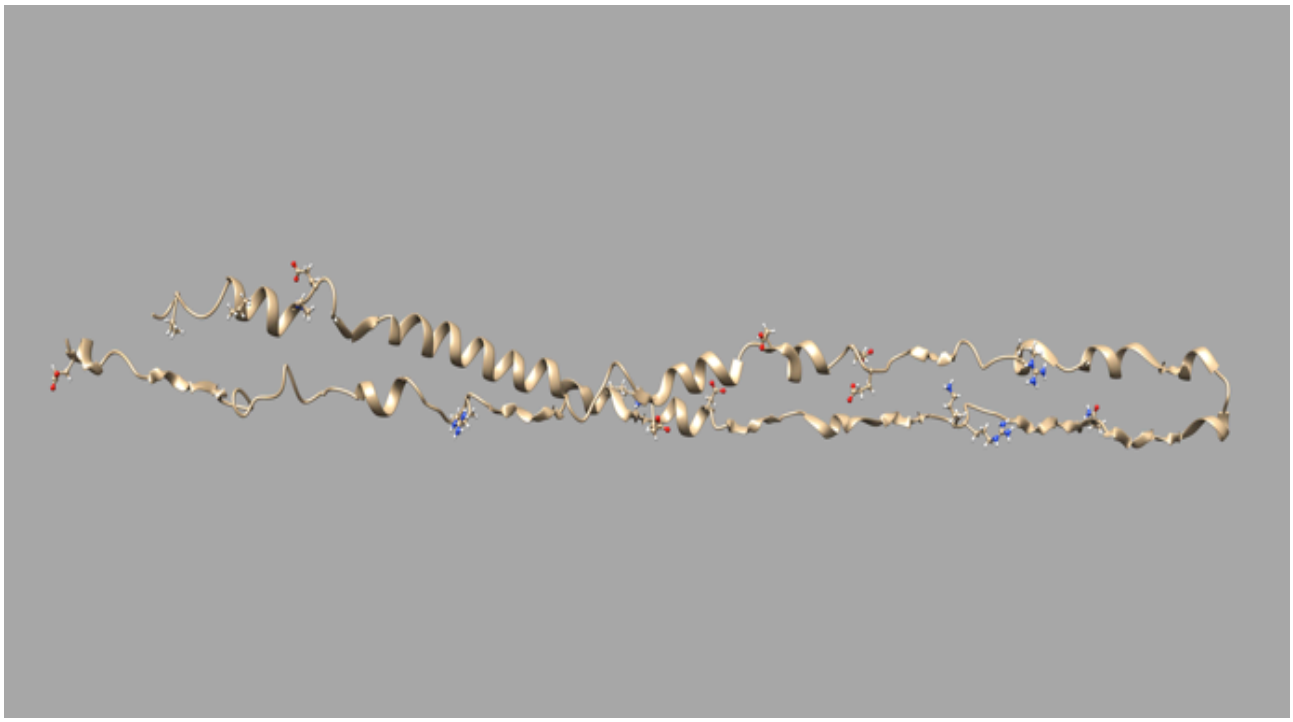

Figure 1. The localization of mutated residues, which carry predicted deleterious mutations, were shown in ball and stick representation on CENPH protein. The protein model of CENPH was generated by using I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/, [14-16]) and visualized by using UCSF Chimera (http://www.rbvi.ucsf.edu/chimera, [17]).

Visualization of p.R105W and p.V237F on CENPH protein

p.R105W and p.V237F were the two missense substitutions that were predicted as deleterious by PredictSNP with high expected accuracy; 0.87 (Table 1). p.R105W was the missense substitution of arginine at position 105 to tryptophan (Figure 2).

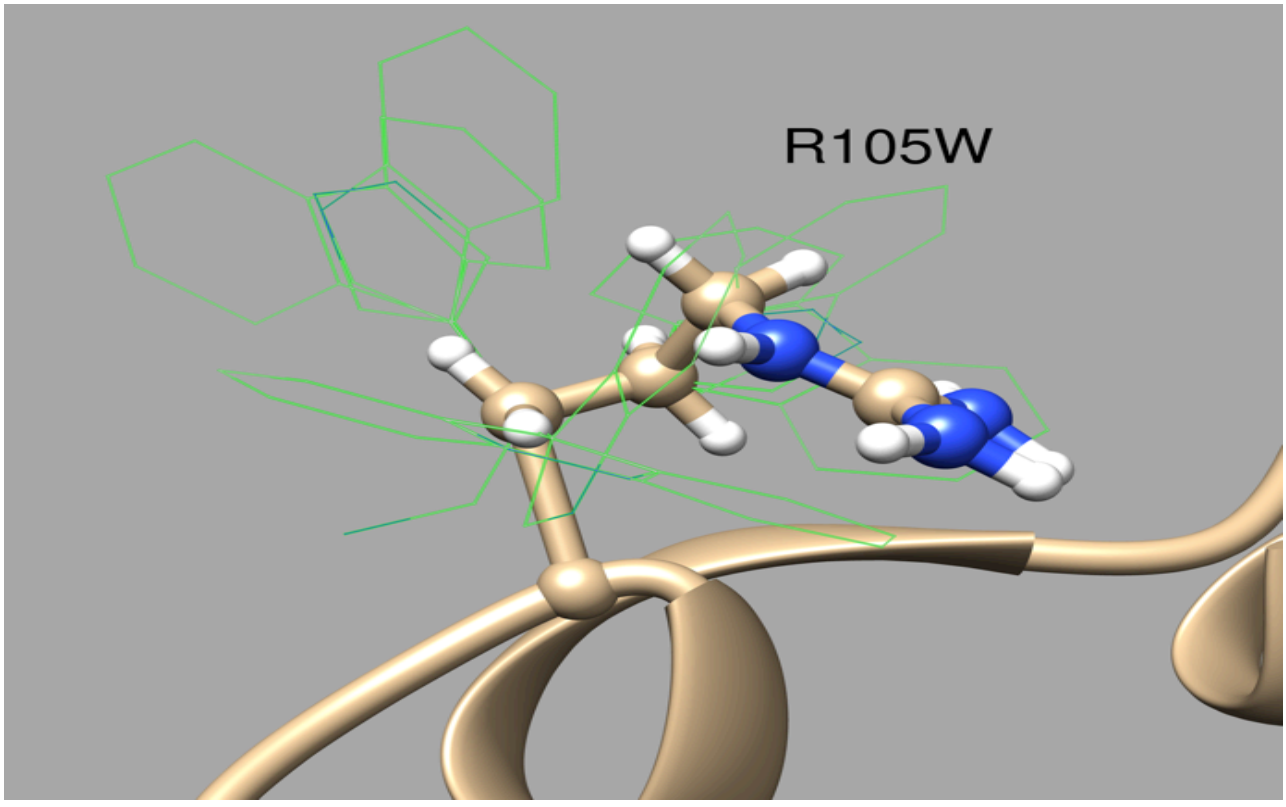

Figure 2. p.R105W mutation on CENPH protein. The original Arginine, R, was shown in ball and stick presentation and rotamers of mutated Tryptophan, W, was shown in wire presentation.

p.V237F was the missense substitution of valine at position 237 to phenylalanine (Figure 3).

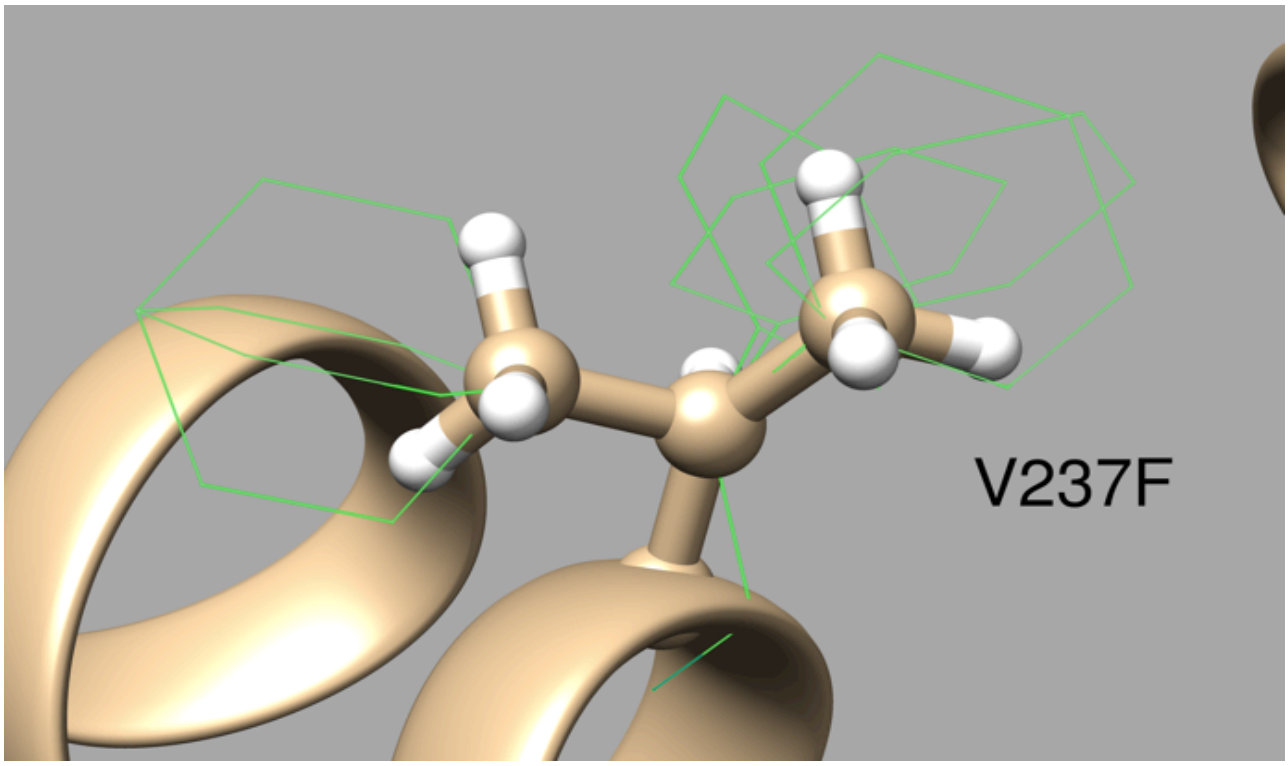

Figure 3. p.V237F mutation on CENPH protein. The original Valine, V, was shown in ball and stick presentation and rotamers of mutated Phenylalanine, F, was shown in wire presentation.

Structural effects of p.R105W and p.V237F on CENPH protein

In order to detect the structural effects of missense substitutions of p.R105W and p.V237F, HOPE server [19] was used.

According to HOPE results, when the amino acid properties of wild type and mutant amino acid was considered, for the p.R105W substitution, the charge of the wild type residue was lost after the substitution. The mutant residue was bigger than the wild type residue. The wild type and mutant residue also differed in hydrophobicity.

For the p.V237F substitution, the wild type and mutant amino acids differed in size.

DISCUSSION

In this study, the missense mutations of CENPH that were observed in different cancers were extracted and the effects of these mutations were investigated.

The kinetochore is a proteinaceous assembly that connects centromeric DNA and spindle microtubules and is vital for chromosome segregation [20]. During mitosis, chromosome segregation problems stimulates aneuploidy, chromosome instability and tumorigenesis [21]. CENPH is one of the constitutive kinetochore proteins and is necessary for kinetochore-microtubule plus-end dynamics and chromosome oscillations [1, 2]. Knockdown of CENPH produces chromosome misalignment and multi-polar spindles [3]. Increased CENPH expression has been shown in different cancers and proposed as a promising prognostic marker [4-9]. Although there were studies in the literature about the aberrant expression of CENPH in cancers, studies searching the effect of CENPH mutations associated with cancers are needed.

According to results of this study, 18 missense mutations were predicted as deleterious for CENPH (Table 1). In order to evaluate the localization of deleterious missense mutations on 3D protein, human CENPH protein was modeled and locations of predicted missense mutations were visualized on CENPH. On the 3D protein structure, the predicted deleterious missense mutations were displayed a scattered pattern instead of accumulating in one region (Figure 1). Interestingly R105 position carried three different substitutions, p.R105W, p.R105K and p.R105M that were all predicted as deleterious (Table 1).

In this study, two missense mutations, p.R105W and p.V237F, were predicted as deleterious with high

expected accuracy; 0.87 (Table 1). p.R105W missense substitution has been detected in lung adenocarcinoma patients [22] and p.V237F has been detected in liver neoplasms according to COSMIC v89.

According to HOPE results, for the p.R105W substitution, the loss of charge after the substitution might lead to loss of interactions with other molecules or residues. The bigger mutant residue might lead to bumps. The mutant residue was more hydrophobic than wild type residue, which might cause loss of hydrogen bonds and/or disturb protein folding. For the p.V237F substitution, since the mutant residue was bigger in size, this substitution might lead to bumps.

In this study, 3D CENPH protein was generated by an in silico method. Although, currently in silico protein modeling techniques predict the protein structures with high accuracy, they have also limitations and drawbacks [23]. Therefore, it should be noted that the created protein model of CENPH might not fully reflect the original 3D protein.

In addition, since this study was performed with computational methods, additional experimental methods are needed to validate the pathogenicity and phenotypic associations of predicted deleterious mutations.

CONFLICT OF INTEREST

There is no conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] Sugata N, Li S, Earnshaw WC, Yen TJ, Yoda K, Masumoto H, et al. Human CENP-H multimers colocalize with CENP-A and CENP-C at active centromere--kinetochore complexes. Hum Mol Genet 2000;9:2919-26.

[2] Amaro AC, Samora CP, Holtackers R, Wang E, Kingston IJ, Alonso M, et al. Molecular control of kinetochore-microtubule dynamics and chromosome oscillations. Nat Cell Biol 2010;12:319-29.

[3] Orthaus S, Ohndorf S, Diekmann S. RNAi knockdown of human kinetochore protein CENP-H. Biochem Biophys Res Commun 2006;348:36-46.

[4] Liao WT, Feng Y, Li ML, Liu GL, Li MZ, Zeng MS, et al. Overexpression of centromere protein H is significantly associated with breast cancer progression and overall patient survival. Chin J Cancer 2011;30:627-37.

[5] Guo XZ, Zhang G, Wang JY, Liu WL, Wang F, Dong JQ, et al. Prognostic relevance of Centromere protein H expression in esophageal carcinoma. BMC Cancer 2008;8:233.

[6] Lu G, Shan T, He S, Ren M, Zhu M, Hu Y, et al. Overexpression of CENP-H as a novel prognostic biomarker for human hepatocellular carcinoma progression and patient survival. Oncol Rep 2013;30:2238-44.

[7] Wu X, Lin Y, Shi L, Huang Y, Lai C, Wang Y, et al. Upregulation of centromere protein H is associated with progression of renal cell carcinoma. J Mol Histol 2015;46:377-85.

[8] Liao WT, Song LB, Zhang HZ, Zhang X, Zhang L, Liu WL, et al. Centromere protein H is a novel prognostic marker for nasopharyngeal carcinoma progression and overall patient survival. Clin Cancer Res 2007;13:508-14.

[9] Liao WT, Wang X, Xu LH, Kong QL, Yu CP, Li MZ, et al. Centromere protein H is a novel prognostic marker for human nonsmall cell lung cancer progression and overall patient survival. Cancer 2009;115:1507-17.

[10] Lu G, Hou H, Lu X, Ke X, Wang X, Zhang D, et al. CENP-H regulates the cell growth of human hepatocellular carcinoma cells through the mitochondrial apoptotic pathway. Oncol Rep 2017;37:3484-92.

[11] Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res 2019;47:D941-D7.

[12] Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput Biol 2014;10:e1003440.

[13] O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 2016;44:D733-45.

[14] Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 2010;5:725-38.

[15] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nat Methods 2015;12:7-8.

[16] Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res 2015;43:W174-81.

[17] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605-12.

[18] Dunbrack RL, Jr. Rotamer libraries in the 21st century. Curr Opin Struct Biol 2002;12:431-40.

[19] Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics 2010;11:548.

[20] Perpelescu M, Fukagawa T. The ABCs of CENPs. Chromosoma 2011;120:425-46.

[21] Holland AJ, Cleveland DW. Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. Nat Rev Mol Cell Biol 2009;10:478-87.

[22] Cai W, Zhou D, Wu W, Tan WL, Wang J, Zhou C, et al. MHC class II restricted neoantigen peptides predicted by clonal mutation analysis in lung adenocarcinoma patients: implications on prognostic immunological biomarker and vaccine design. BMC Genomics 2018;19:582.

[23] Gupta CL, Akhtar S, Bajpai P. In silico protein modeling: possibilities and limitations. EXCLI J 2014;13:513-5