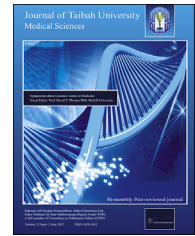Taibah University

Journal of Taibah University Medical Sciences

www.sciencedirect.com

Original Article

Prediction of deleterious single nucleotide polymorphisms and their effect on the sequence and structure of the human *CCND1* gene

Bharath Yoganarasimha, B.E [a], Vivek Chandramohan, M.Sc [b],*,
Thirupathihalli P. Krishna Murthy, M.Tech [a], Bhavya S. Gangadharappa, M.Tech [a],
Gowrishankar B. Siddaiah, Ph.D [b] and Makari Hanumanthappa, M.Sc [c]

[a] *Department of Biotechnology, M. S. Ramaiah Institute of Technology, Bengaluru, Karnataka, India*
[b] *Department of Biotechnology, Siddaganga Institute of Technology, Tumakuru, Karnataka, India*
[c] *Department of Biotechnology, IDSG Government College, Chikkamagaluru, Karnataka, India*

الملخص

أهداف البحث: إن الجين CCND1 المعبر للبروتين G۱/S-السيكلين المحدد، والمنظم لانتقاله في دورة الخلية يثبط أيضا البروتينات في الأورام الشبكية. وزيادة التعبير أو إعادة التنظيم في هذا الجين قد تؤدي إلى أورام مختلفة. تهدف هذه الدراسة لتحديد الأضرار الممكنة من النيوكليوتيد الوحيد المتعدد الأشكال للجين CCND1 باستخدام الطرق الحاسوبية.

طرق البحث: تم استعادة النيوكليوتيد الوحيد المتعدد الأشكال في جينات الإنسان CCND1 من قاعدة معلومات النيوكليوتيدات الوحيدة المتعددة الأشكال. تم فحص هذه النيوكليوتيدات الوحيدة المتعددة الأشكال بواسطة فرز عدم الاحتمال من الاحتمال والتنبؤ بخوادم النيوكليوتيدات الوحيدة المتعددة الأشكال. كما تم بناء الطفرات التي تحوي النيوكليوتيدات الوحيدة المتعددة الأشكال الضارة باستخدام استديو الاكتشاف ٥،٣ وأجريت الدراسات الديناميكية على أصناف محلية وطافرة.

النتائج: تم العثور على ١١٩٤ من النيوكليوتيدات الوحيدة المتعددة الأشكال في الإنسان، منها ٩٤ مغلطة و ٢ لا قيمة لها. كما وُجدت ثلاث نيوكليوتيدات وحيدة متعددة الأشكال ضارة. وأظهرت الديناميكيات الجزيئية وتحليل المسار انحرافا كبيرا في قيم متوسط جذر الانحراف التربيعي في الطفرة N۲۱٦K من قيم البروتين المحلي.

الاستنتاجات: استنادا على هذه الدراسة، نقترح أن النيوكليوتيدات الوحيدة المتعددة الأشكال مع تعريفها G<C.٦٤٨c.:٢.٠٥٣٠٥٦_NM (rs١١٢٥٢٥٠٩٧)) قد تسبب انحرافات في جين CCND1 التي قد تؤدي بالتالي إلى تغيرات في وظائف البروتين G۱/S-السيكلين المحدد. وهذا، بالمقابل يمكن أن يكون السبب في حدوث سرطان الدم النخاعي الحاد.

الكلمات المفتاحية: سرطان الدم النخاعي الحاد؛ الديناميكيات الجزيئية

Abstract

Objective: The *CCND1* gene expresses a protein, G1/S-specific cyclin, that regulates the G1/S transition in the cell cycle and also inhibits retinoblastoma (RB) proteins. Overexpression or rearrangements of this gene can result in various tumours. This study aimed to identify possible deleterious non-synonymous single nucleotide polymorphisms (SNP's) of *CCND1* using computational methods.

Methods: SNPs in the human *CCND1* gene were retrieved from dbSNP. These SNPs were screened by the Sorting Intolerant From Tolerant (SIFT) algorithm and the PredictSNP classification. Mutants with deleterious SNPs were built using Discovery Studio 3.5, and dynamics studies were performed on native and mutant varieties.

Results: In *Homo sapiens*, 1194 SNPs were found, of which 94 were missense and 2 were nonsense SNPs. Three SNPs were found to be deleterious. Molecular dynamics and trajectory analysis showed that there was a significant deviation of the root mean square deviation (RMSD) values in the N216K mutant from the values of the native protein.

Conclusion: Based on this study, we propose that the SNP with SNP ID rs112525097 (NM_053056.2:c.648C>G) might cause aberrations in *CCND1*, which might lead to a change in the function of the G1/S-specific cyclin protein. This, in turn, may lead to the development of acute myeloid leukaemia (AML).

* Corresponding address: Department of Biotechnology, Siddaganga Institute of Technology, Tumakuru 572103, Karnataka, India.
E-mail: vivek@sit.ac.in (V. Chandramohan)
Peer review under responsibility of Taibah University.

ELSEVIER | **Production and hosting by Elsevier**

Introduction

Many diseases have their genetic origins in single nucleotide polymorphisms (SNPs). It has been a challenge to identify disease-causing SNPs in genes.[1] Phenotypic variations caused by SNPs also need to be studied.[2] Non-synonymous SNPs (nsSNP) alter DNA, regulate gene expression and also cause changes in the resultant protein.[3] The occurrence of SNPs is roughly once every 1000−2000 bp,[4] and there is an estimated total of 150,482,731 newly released SNPs in the human genome.[5] Non-synonymous SNPs (nsSNPs) are the cause of amino acid substitutions and are major factors that contribute to the functional diversity of proteins. The main aim of pharmacogenomics is to identify deleterious SNPs. The gene that is the focus of our study, *CCND1*, is present on chromosome 11 and has a coding sequence length of 885 nucleotides. Its cytogenetic location is 11q13.3.[6] This gene codes for the protein G1/S-specific cyclin-D1, which is a regulatory subunit of the cyclin D1-CDK4 complex.[7] It inhibits members of the retinoblastoma (RB) protein family, such as RB1, by phosphorylation and also regulates the cell-cycle during the G1/S transition phase. Cyclin D-CDK4 complexes are translocated to the nucleus through interactions with KIP/CIP family members after initially accumulating at the nuclear membrane.[8] Mutations, amplification and overexpression of this gene can contribute to tumourigenesis due to its role in regulating cell cycle progression.[9] When overexpressed, cyclin D1 acts as an oncogene in many human neoplasias. Acute myeloid leukaemia (AML) is a heterogeneous clonal disorder that is characterized by clonal neoplastic cells that express irrepressible proliferation. Blasts with an impaired differentiation program accumulate in the bone marrow.[10] Approximately 80% of all adult leukaemia is AML, which is the most common cause of death from leukaemia. Overall, of all people diagnosed with AML, approximately 20 of 100 people (20%) will survive for 5 years or more after diagnosis. Younger people cope have a better prognosis.[11,12] Metastasis of abnormal white blood cells results in leukaemic infiltration of the bone marrow, resulting in cytopaenia. This leads to reduced red blood cells, which causes fatigue in patients. A decrease in platelets results in internal haemorrhage. Pallor and infection due to a reduced normal white cell count is common. There can be leukaemic infiltration to other organs, which can cause abnormalities in these organs.[13] Untreated AML is usually fatal within a few weeks or months. Mutations in transcription factors that encode for myeloid differentiation can be found in patients with AML.[14] CCND1 binds and activates the G1 cyclin-dependent kinases CDK4 and CDK6. The complex then inhibits members of the retinoblastoma (RB) family of proteins. These processes regulate the G1/S transition in the cell cycle.[15] CCND1 promotes activation of cyclin E/CDK2-containing complexes. It sequesters CDK inhibitors, such as p27 Kip1 and p21Cip1, by a kinase-independent action. CCND1 also inhibits the transcriptional activity and anti-proliferative function of Smad3 by phosphorylation.[16] This gene has also been implicated in many other diseases, such as multiple myeloma, Von Hippel-Lindau syndrome, mantle cell lymphoma, chronic lymphocytic leukaemia, and colorectal cancer, among others.

The aim of this study is to predict non-synonymous SNPs in the *CCND1* gene that might be responsible for causing acute myeloid leukaemia. Mutations in this gene can lead to its overexpression or non-regulation of the retinoblastoma family of proteins (RB). In this study, we performed *in silico* studies to identify deleterious SNPs in the *CCND1* gene. These SNPs might cause changes in the protein structure, which can lead to changes in function. Using this information, we have proposed modelled structures of the mutant proteins. The stabilities of these structures were also checked. These findings could be very helpful in the early diagnosis and treatment of acute myeloid leukaemia. If the SNPs reported in this study are experimentally verified to be a cause of various cancers, fore-knowledge of these mutations would be of great help. Genetic screening can be performed, and people with a particular mutation can be warned about the risks posed by this mutation. These findings might lead to therapeutic solutions, such as effective gene therapy, which might be able to reduce the risk of cancer.

Materials and Methods

The dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) database was used to download SNPs and their related protein sequences for *in silico* studies. Homology-based tools (SIFT and PredictSNP) were used for the study of coding SNPs.

Determination of deleterious SNPs by SIFT and PredictSNP

SIFT determines whether an amino acid substitution has any effect on protein function. The degree of conservation of amino acid residues obtained from the alignment of closely related sequences is essential for the prediction of deleterious SNPs. This technique can be used for predictions in naturally occurring non-synonymous mutations or artificially induced variations.[17] Multiple SNP IDs were collected from the SNP database of the NCBI and submitted as input queries. The threshold value is a tolerance index of ≥ 0.05. The lower the tolerance index, the greater is the functional effect of a particular amino acid substitution. PredictSNP is a consensus classifier that runs a suite of programs that are usually used for the prediction of the effects of a mutation on protein function. It combines the results of PredictSNP, MAPP, PhD-SNP, PolyPhen-1, PolyPhen-2, SIFT, SNAP, PANTHER and the nsSNPs analyser.[18] The output prediction value falls within the continuous interval of $(-1, +1)$. The effects of mutations are said to be neutral if the value falls in the interval $(-1, 0)$. The mutations are

considered to be deleterious if they fall in the interval (0, +1). The greater the distance between the PredictSNP score and zero, the more deleterious the mutations.

Multiple sequence alignment to determine conservation/ variation of the deleterious residues

FASTA sequences of the CCND1 protein of various organisms were retrieved from the UniProt database. The organisms were *Homo sapiens*, *Mus musculus*, *Rattus norvegicus* and *Pongo abelii*. The UniProt identifiers are P24385, P25322, P39948 and Q5R6J5, respectively. Multiple sequence alignment was performed using CLC Genomics Workbench software.[19] The conservation/variation was checked at the 130th and 216th residues.

Preparation of protein, stability check and molecular dynamics

The protein, a crystal structure of CDK4 in complex with a D-type cyclin with the PDB ID 2W96, was retrieved from the PDB database.[20] Because only the A-chain containing CCND1 was required for this study, the B-chain containing CDK4 was deleted. Other extra ligand molecules were also deleted to stabilize the protein structure.[21] The resulting structure was considered to be the wild-type variety of the protein. Mutational analysis was performed based on the results obtained from the various tools that predicted deleterious mutations. The original amino acid residues were replaced with mutant variants (N130D, N130S and N216K).[22] The mutated structures were optimized using CHARMm force fields.[23] They were subjected to a standard dynamics cascade with a convergence energy of 0.001 kcal/mol.[24] The optimized structures were then used for further molecular simulation studies.[25] In the mutant variant NP_444284.1:p.Asn130Asp (N130D), A changes to G at the 597th position in the mRNA (cDNA variation: NM_053056.2:c.388A>G). Similarly, A changes to G at the 598th mRNA (c DNA variation: NM_053056.2:c.389A>G) position in the mutant NP_444284.1:p.Asn130Ser (N130S). In the NP_444284.1:p.Asn216Lys (N216K) variant, C changes to G at the 857th mRNA position (cDNA variation: NM_053056.2:c.648C>G). These polymorphisms result in a single amino acid substitution in the protein, which might result in changes in structure and function. The stability of these mutations was tested by DUET and SDM servers.[26] The molecular dynamics simulation protocol in Discovery Studio 3.5 was applied to the wild-type and mutants derived from mutational studies. The trajectory was analysed for a period of 6 ns with respect to backbone atoms. The root mean square deviation is the average inter-atomic distance of superimposed proteins. Six thousand steps were performed in which every frame was compared to the first frame and RMSD deviations were calculated.[27]

Results

Retrieval of SNPs from dbSNP

Investigations of the gene *CCND1* were conducted, and the variants of the gene were obtained from the dbSNP database. A total of 1194 SNPs were found in *Homo sapiens*. Of these, 89 encoding synonymous SNPs and 605 intron SNPs were found. Two stop-gain and 2 frameshift SNPs were found (Figure 1). The non-synonymous SNPs consisting of 94 missense SNPs and 2 nonsense SNPs were selected for investigation. A large number of SNPs were in intron regions.

Obtaining deleterious SNPs using SIFT and PredictSNP

The homology-based tool SIFT was used to determine the conservation range of a particular amino acid. Ninety-six nsSNPs were submitted to the SIFT tool. Their tolerance indices were obtained. The lower the tolerance index, the greater the effect of the amino acid substitution on the function of the protein.[17] The G1/S-specific cyclin-D1 protein sequence was downloaded from the UniProt database (P24385).[28] ProtParam[29] analysis showed that the protein has 295 amino acids. Its molecular weight is 33729.1 Da and its theoretical PI is 4.97. The protein has 47 negatively charged residues and 34 positively charged residues and an instability index of 57.71, which classifies the protein as unstable. The protein has an estimated half-life of 30 h in humans and mammals. Among the 96 SNPs, 3 were predicted to be deleterious, with a score of less than or equal to 0.05. The results are presented below in Table 1. The SNPs of SNPids rs1050971, rs1131439, and rs112525097 were found to be damaging by the SIFT server. An SNP with a score of ≤0.05 is predicted to be damaging. These mutations can cause conformational changes in the protein, which is why
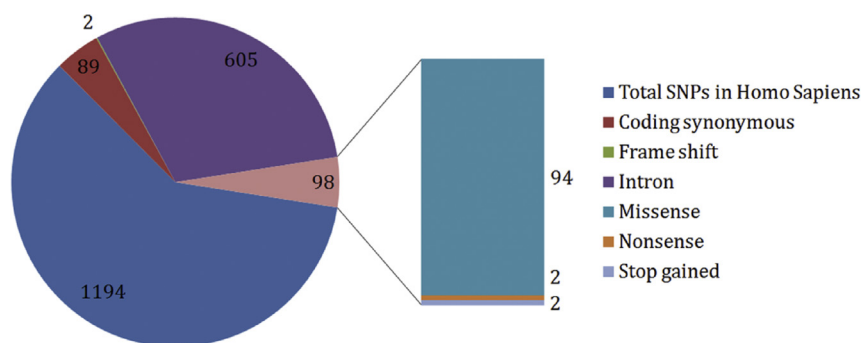

Figure 1: Distribution of nsSNPs: frame shift, coding synonymous and intron SNPs.

Table 1: Deleterious SNPs predicted by SIFT.

SNP ID	Amino acid change	Tolerance Index	Prediction
rs1050971	N130D	0.01	**Damaging**
rs1131439	N130S	0.00	**Damaging**
rs112525097	N216K	0.05	**Damaging**

Note: Values less than or equal to a threshold of 0.05 are considered to be damaging.

they were deemed to be deleterious by SIFT. The confidence levels of structural alterations were obtained by using the PredictSNP program. The protein sequences, along with the mutations specified in the above SNPs, were submitted as inputs, and the results obtained are depicted in Table 2. PredictSNP is a consensus classifier. The mutations were classified into two classes by all of the incorporated tools: neutral or deleterious. An extra class of 'possibly deleterious' was given by PolyPhen1. A confidence score equal to 0.5 enabled this class to be considered a deleterious class. The SNP p.Asn130Asp (N130D) was found to be deleterious by MAPP and SIFT, with a confidence score of 56 and 45 percent, respectively. The SNP p.Asn130Ser (N130S) was found to be deleterious by SIFT due to a confidence score of 53 percent. The SNP p.Asn216Lys (N216K) was found to be deleterious by PhD-SNP server due to a confidence score of 59 percent. It can be observed that all of these confidence scores are below the 60 percent mark. PredictSNP gives a higher score because it combines the results of all of the analysis tools. The p.Asn130Ser and p.Asn216Lys variations had similar scores. The effects of their mutations were determined by molecular dynamics studies.

Multiple sequence alignment to determine conservation/ variation of the deleterious residues

Multiple sequence alignment was performed to check amino acid conservation/variation across species. In the output, the rows correspond to the input sequences. The columns correspond to the amino acid residue in the respective positions. Gaps or positions with '-' indicate insertions or deletions at that position. In this study, the sequences of human, mouse, rat and orangutan were aligned. We observed that asparagine (N) was conserved at the 130th and 216th positions in all 4 species, as shown in (Figure 2). Therefore, we can conclude that this gene is an essential gene that regulates key functions of the body.

Mutational studies of the protein and stability check

The wild-type protein was used, and mutants were built based on the results predicted by SIFT and PredictSNP. For

the mutant variant p.Asn130Asp, A was substituted for G at the 597th mRNA position. The amino acid asparagine was changed to aspartic acid. Similarly, A was changed to G at the 598th mRNA position in the mutant p.Asn130Ser. Here, asparagine was changed to serine. In the p.Asn216Lys variant, C was replaced with G at the 857th mRNA position. Asparagine was changed to lysine in this variant (Figure 3). All 3 variants were subjected to a stability check by DUET and SDM. The stability change is depicted in terms of the variation in Gibb's free energy. A negative energy variation is said to be destabilizing, and a positive energy variation is said to be stabilizing. The p.Asn130Ser variant had a variation of −2.42 kcal/mol and −0.439 kcal/mol, as predicted by SDM and DUET, respectively. The p.Asn130Asp variant had a variation of −1.51 kcal/mol and −0.659 kcal/mol, as predicted by SDM and DUET, respectively. The p.Asn216Lys mutation had a variation of −0.23 kcal/mol and −0.502 kcal/mol, as predicted by SDM and DUET, respectively. All of the energy variations were negative. Consequently, we can say that all 3 variations are destabilizing.

Molecular dynamics simulation of the wild and mutant types

A comparative study of molecular dynamic properties was conducted using the molecular dynamics simulation protocol in Discovery Studio 3.5. This was conducted on the 3 mutant variants and the native wild-type protein. The trajectory was analysed for 6 ns. Different physico-chemical parameters, such as the potential energy, electrostatic and root mean square deviation, were calculated. The backbone RMSD values were calculated from trajectory analysis. A graph of RMSD versus trajectory time was prepared. The wild-type protein was stable after a jump from 1.16 Å to 1.5 Å at 2 ns. The mutant p.Asn130Ser was generally stable in the 1.4−1.6 Å range after some distortion until reaching the 0.26 ns time period. The variant p.Asn130Asp was distorted until 3.4 ns and then stabilized. The mutant p.Asn216Lys showed a maximum deviation from the wild-type protein and had a stable RMSD range of 1.6−1.8 Å (Figure 4). The potential energy values were also calculated. It can be observed that the p.Asn216Lys variant had the least final potential energy of −14330.1 kcal/mol, whereas the wild-type had a final potential energy of −13830.7 kcal/mol. The p.Asn130Ser model had a final potential energy of −13772.7 kcal/mol, and the p.Asn130Asp model had a final potential energy of −14030 kcal/mol (Figure 5). There was a significant deviation between the RMSD values of the wild-type and the p.Asn216Lys model. Therefore, it can be inferred that in the p.Asn216Lys variant, the protein undergoes significant structural changes relative to the wild-type protein. This might result in a change of function in the mutated protein. This is also substantiated by the potential energy studies, which show that the p.Asn216Lys

Table 2: nsSNPs predicted to be functionally significant by PredictSNP, in percentages.

Mutation	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP	PANTHER
N130D	65	56*	68	67	68	45*	61	64
N130S	77	77	68	67	63	53*	61	57
N216K	78	78	59*	67	79	84	67	69

Note: Scores less than 60 (indicated by *) have been flagged as 'deleterious' by the respective software.
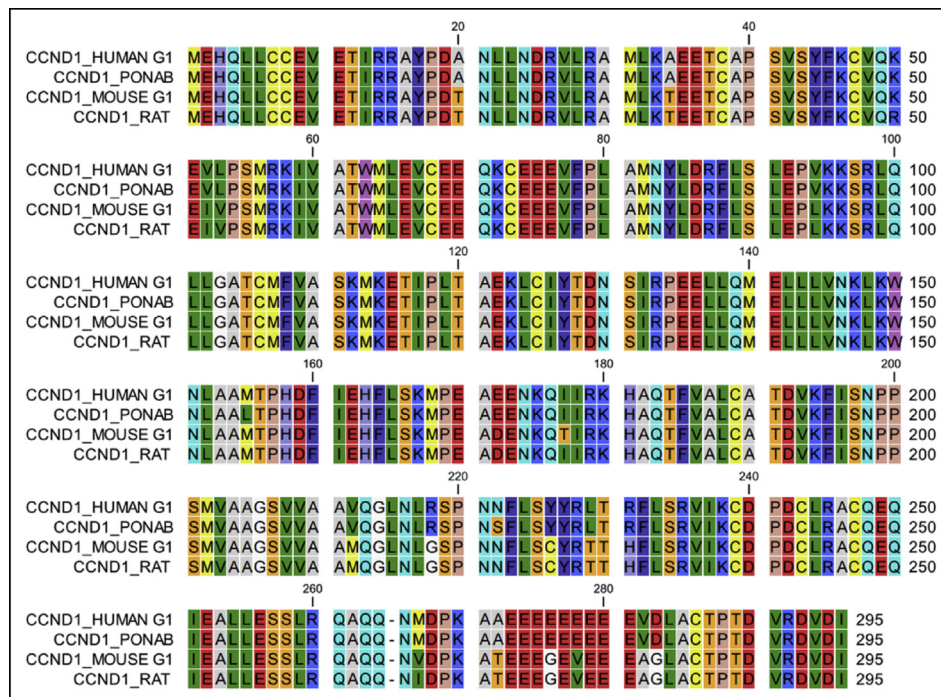
Figure 2: Multiple sequence alignment to determine conservation/variation of the deleterious residues.

model had the least potential energy, a difference of 499.4 kcal/mol from the potential energy of the wild-type. This suggests that this mutated protein is highly stable and is able to exist in natural *in vivo* conditions. This lends support to the inference that this mutation might have the ability to cause a change in protein function and is stable enough to be present for a long period of time.

Discussion

The focus of this study was to analyse the *CCND1* gene, which codes for the G1/S-specific cyclin-D1 protein. This protein is a regulatory subunit of the cyclin D1-CDK4 complex. It inhibits members of the retinoblastoma (RB) protein family, such as RB1, by phosphorylation and also regulates the cell-cycle during the G1/S transition phase. The protein G1/S-specific cyclin-D1 belongs to the cyclin-C terminal family (PF02984). During the cell cycle, the concentrations of these types of proteins vary in a cyclical manner. The fluctuations in the expression of cyclin genes and destruction by the proteasome complex, which is mediated by ubiquitin targeting, drive the cell cycle by inducing fluctuations in the activity of CDK proteins. Activation of CDK is initiated by forming a complex with a cyclin protein, but complete activation is achieved by phosphorylation. The CDK active site is activated by complex formation. Cyclins incorporate binding sites for some substrates. They have no enzymatic properties, but lead the CDKs to particular subcellular locations via targeting. DNA replication is induced by a complex of S cyclin and CDK. High levels of S cyclins are maintained throughout S phase, G2 phase and the early mitosis stages. This helps to induce and promote the initial activities of mitosis. Mutations or SNPs in *CCND1* might be responsible for causing acute myeloid leukaemia or leading to non-regulation of the retinoblastoma family of proteins

(RB).[30,31] DUET and SDM predict that the mutation is destabilizing. We speculate that this might destabilize the cyclinD1/CDK complex. This complex usually interacts with Rb protein, a tumour suppressor. The mutated CCND1 protein might result in a change in tumour suppression, leading to the development of various tumours.

Here, we attempted to identify deleterious and destabilizing SNPs in the *CCND1* gene. We used an *in silico* approach that collates the results of SIFT and PredictSNP with those of Discovery Studio 3.5. In total, we identified 3 deleterious mutants. Further analysis confirmed that 2 mutants were relatively stable. Briefly, investigations of the *CCND1* gene were conducted, and the analysis shows that these SNPs do not affect protein function because they are probably spliced out during mRNA generation.

Among the 96 nsSNPs submitted to SIFT and Predict SNP, 3 SNPs (rs1050971, rs1131439, rs112525097) were found to be deleterious based on their confidence score (≤ 0.05). Further molecular dynamics simulation of the wild-type and mutants were carried out for 3 mutant variants, p.Asn130Asp (N130D), p.Asn130Ser (N130S) and p.Asn216Lys (N216K), and the native wild-type protein. This analysis showed that there is a significant deviation between the RMSD values of the wild-type and the p.Asn216Lys mutant model. This suggests that deviations in the protein might be due to significant structural changes in the mutants compared to the wild-type. Significant structural modifications due to deviations in the RMSD values might result in a change of function of the mutated protein.
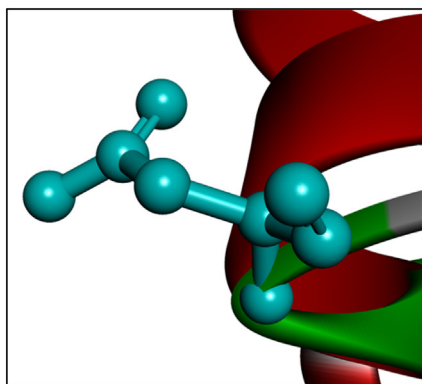
Conclusion

The *CCND1* gene was investigated using *in silico* methods. A total of 1194 SNPs were found in humans. Of these, 89 were synonymous and 96 were non-synonymous. The non-
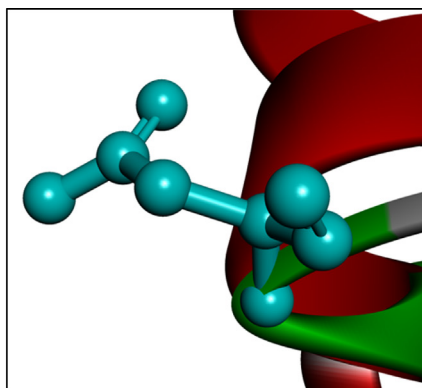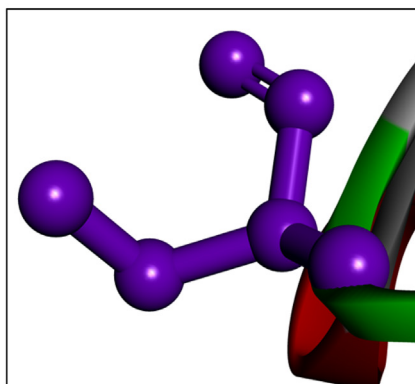
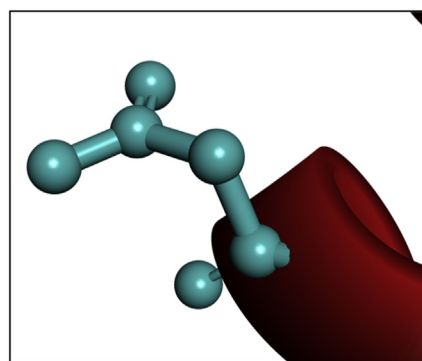*Wild-type protein with asparagine (N) amino acid at the 130th position (blue colour)*
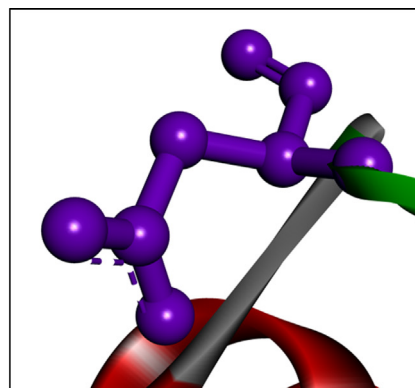
Mutant protein with aspartic acid (D) amino acid at the 130th position (violet colour)
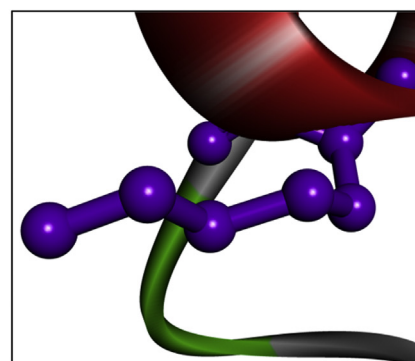

Wild-type protein with asparagine (N) amino acid at the 130th position (blue colour)

Mutant protein with serine (S) amino acid at the 130th position (violet colour)

Wild-type protein with asparagine (N) amino acid at the 216th position (blue colour)


*Mutant protein with lysine (K) amino acid at the 216th position (violet colour)*

Figure 3: Three experimentally validated nsSNPs [N130D, N130S and N216K].

synonymous SNPs included 94 missense SNPs and 2 nonsense SNPs. These 96 nsSNPs were analysed using SIFT and PredictSNP. Three SNPs were found to be damaging by SIFT, and this was corroborated by PredictSNP. The variation in structure caused by these mutations was studied by molecular simulation studies for a time period of 6 ns. Of the 3 mutations identified in this study, it was found that the mutation in the native protein at the 216th position, where asparagine changes to lysine, had the largest deviation from the RMSD value of the native structure. The native structure had an average RMSD value of 1.38, whereas the mutant structure of the NP_444284.1:p.Asn216Lys (N216K) model had an average RMSD value of 1.61. This can be interpreted as the N216K model causing significant changes in structure, which in turn might cause a significant change in function. Hence, we can conclude that the nsSNP (rs112525097) associated with this mutation might be an important candidate for the cause of acute myeloid leukaemia (AML) resulting from
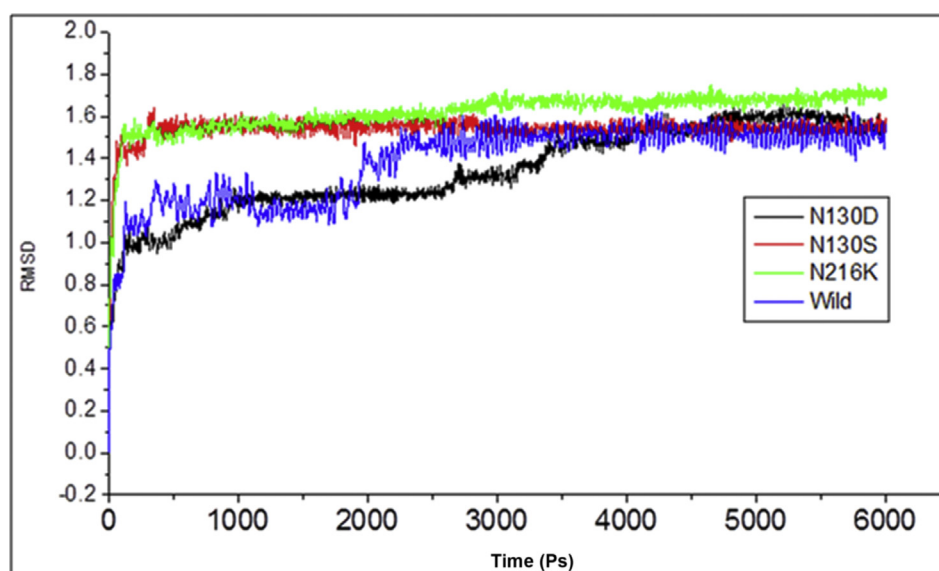
Figure 4: Backbone RMSDs are shown as a function of time for wild-type and mutant CCND1 protein structures over 6 ns.
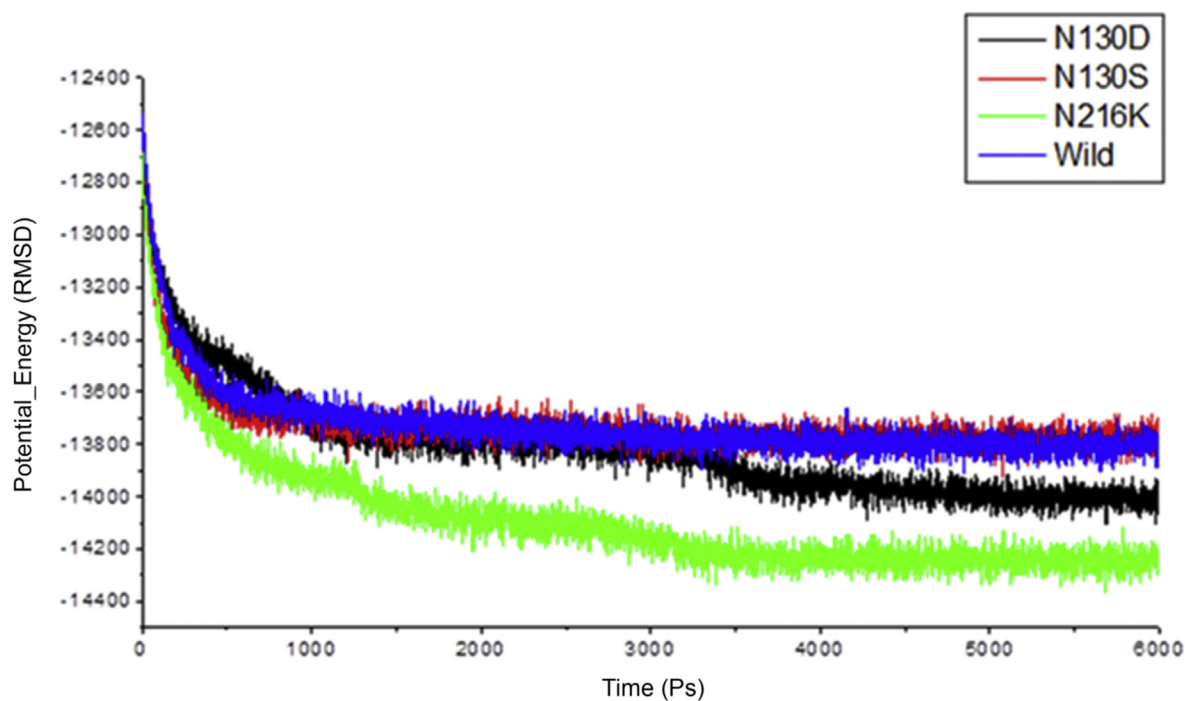
Figure 5: Potential energy is shown as a function of time for wild-type and mutant CCND1 protein structures over 6 ns.

the *CCND1* gene. Further *in vivo* studies are required for confirmation and development of therapeutic treatment.

Authors' contributions

BY and VC designed and conducted the study. KM contributed to data collection and result interpretation. All authors were involved in writing the initial draft of the article and critical review. All authors have critically reviewed and approved the final draft and are responsible for the content and the similarity index of the manuscript.

Conflict of interest

The authors have no conflict of interest to declare.

resources for carrying out this project. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

References

1. Hernandez LM, Blazer DG, editors. *Genes, Behavior, and the Social Environment: Moving beyond the nature/nurture debate.* Washington (DC): The National Academies Collection: Reports funded by National Institutes of Health; 2006.
2. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, Diaz de Stahl T, et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. **Am J Hum Genet 2008**; 82(3): 763−771.
3. Yates CM, Sternberg MJ. The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions. **J Mol Biol 2013**; 425(21): 3949−3963.
4. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. **Nature 2001**; 409(6822): 928−933.
5. NCBI dbSNP Human Build 146. Available from: http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=10065172. Accessed on July 7, 2016.
6. Noorlag R, van Kempen PM, Stegeman I, Koole R, van Es RJ, Willems SM. The diagnostic value of 11q13 amplification and protein expression in the detection of nodal metastasis from oral squamous cell carcinoma: a systematic review and meta-analysis. **Virchows Archiv An Int J Pathology 2015**; 466(4): 363−373.
7. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL. Cyclin D as a therapeutic target in cancer. **Nat Rev Cancer 2011**; 11(8): 558−572.
8. LaBaer J, Garrett MD, Stevenson LF, Slingerland JM, Sandhu C, Chou HS, et al. New functional activities for the p21 family of CDK inhibitors. **Genes & Dev 1997**; 11(7): 847−862.
9. Hosokawa Y, Arnold A. Mechanism of cyclin D1 (CCND1, PRAD1) overexpression in human cancer cells: analysis of allele-specific expression. **Genes, Chromosomes Cancer 1998**; 22(1): 66−71.
10. Kumar CC. Genetic abnormalities and challenges in the treatment of acute myeloid leukemia. **Genes & Cancer 2011**; 2(2): 95−107.
11. Inoue S, Khan I, Mushtaq R, Carson D, Saah E, Onwuzurike N. Postinduction supportive care of pediatric acute myelocytic leukemia: should patients be kept in the hospital? **Leukemia Res Treat 2014**; 2014: 592379.
12. Burke PW, Douer D. Acute lymphoblastic leukemia in adolescents and young adults. **Acta Haematol 2014**; 132(3−4): 264−273.
13. Porcu P, Cripe LD, Ng EW, Bhatia S, Danielson CM, Orazi A, et al. Hyperleukocytic leukemias and leukostasis: a review of pathophysiology, clinical presentation and management. **Leukemia Lymphoma 2000**; 39(1−2): 1−18.
14. Pabst T, Mueller BU, Harakawa N, Schoch C, Haferlach T, Behre G, et al. AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. **Nat Med 2001**; 7(4): 444−451.
15. Quelle DE, Ashmun RA, Shurtleff SA, Kato JY, Bar-Sagi D, Roussel MF, et al. Overexpression of mouse D-type cyclins accelerates G1 phase in rodent fibroblasts. **Genes & Dev 1993**; 7(8): 1559−1571.
16. Matsuura I, Denissova NG, Wang G, He D, Long J, Liu F. Cyclin-dependent kinases regulate the antiproliferative function of Smads. **Nature 2004**; 430(6996): 226−231.
17. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. **Nat Protoc 2009**; 4(7): 1073−1081.
18. 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(1): e1003440.
19. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. **Nat Rev Genet 2011**; 12(6): 443−451.
20. Rose PW, Prlic A, Bi C, Bluhm WF, Christie CH, Dutta S, et al. The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. **Nucleic Acids Res 2015**; 43(Database issue): D345−D356.
21. Huang HJ, Jian YR, Chen CY. Traditional Chinese medicine application in HIV: an in silico study. **J Biomol Struct Dyn 2014**; 32(1): 1−12.
22. Biswas D, Gupta SK, Sindhwani G, Patras A. TLR2 polymorphisms, Arg753Gln and Arg677Trp, are not associated with increased burden of tuberculosis in Indian patients. **BMC Res Notes 2009**; 2: 162.
23. Dammalli M, Chandramohan V, Biradar MI, Nagaraju N, Gangadharappa BS. In silico analysis and identification of novel inhibitor for new H1N1 swine influenza virus. **Asian Pac J Trop Dis 2014**; 4: S635−S640.
24. Morishige K, Shimokawa H, Matsumoto Y, Eto Y, Uwatoku T, Abe K, et al. Overexpression of matrix metalloproteinase-9 promotes intravascular thrombus formation in porcine coronary arteries in vivo. **Cardiovasc Res 2003**; 57(2): 572−585.
25. Chandramohan V, Kaphle A, Chekuri M, Gangarudraiah S, Bychapur Siddaiah G. Evaluating andrographolide as a potent inhibitor of NS3-4A protease and its drug-resistant mutants using in silico approaches. **Adv Virology 2015**: 2015.
26. Pires DE, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. **Nucleic Acids Res 2014**: gku411.
27. Chandramohan V, Nagaraju N, Rathod S, Kaphle A, Muddapur U. Identification of deleterious SNPs and their effects on structural level in CHRNA3 gene. **Biochem Genet 2015**; 53(7−8): 159−168.
28. UniProt C. UniProt: a hub for protein information. **Nucleic Acids Res 2015**; 43(Database issue): D204−D212.
29. Hasan MA, Mazumder MH, Chowdhury AS, Datta A, Khan MA. Molecular-docking study of malaria drug target enzyme transketolase in *Plasmodium falciparum* 3D7 portends the novel approach to its treatment. **Source Code Biol Med 2015**; 10: 7.
30. Tarsitano M, Palmieri S, Ferrara F, Riccardi C, Cavaliere ML, Vicari L. Detection of the t (11; 14)(q13; q32) without CCND1/IGH fusion in a case of acute myeloid leukemia. **Cancer Genet Cytogenet 2009**; 195(2): 164−167.
31. de Oliveira FM, Tone LG, Simoes BP, Falcão RP, Brassesco MS, Sakamoto-Hojo ET, et al. Acute myeloid leukemia (AML-M2) with t (5; 11)(q35; q13) and normal expression of cyclin D1. **Cancer Genet Cytogenet 2007**; 172(2): 154−157.