


RESEARCH NOTE

Open Access



In-silico analysis reveals druggable single nucleotide polymorphisms in angiotensin 1 converting enzyme involved in the onset of blood pressure

Brenda Udosen^{1,2,5}, Opeyemi Soremekun¹, Chinwe Ekenna³, Olaposi Idowu Omotuyi⁴, Tinashe Chikowore^{6,7}, Oyekanmi Nashiru⁵ and Segun Fatumo^{1,5,8*} 

Abstract

Objective: The Angiotensin 1 converting enzyme (*ACE1*) gene plays a critical role in regulating blood pressure and thus, it has become a major therapeutic target of antihypertensives. Single nucleotide polymorphisms (SNPs) occurring within a gene most especially at the functional segment of the genes alter the structure–function relationship of that gene.

Results: Our study revealed that five nsSNPs of the *ACE1* gene were found to be potentially deleterious and damaging and they include rs2229839, rs14507892, rs12709442, and rs4977 at point mutations P351R, R953Q, I1018T, F1051V, and T1187M. The protein stability predictive tools revealed that all the nsSNPs decreased stability of the protein and the ConSurf server which estimates the evolutionary conservation profile of a protein showed that three mutants were in the highly conserved region. In conclusion, this study predicted potential druggable deleterious mutants that can be further explored to understand the pathological basis of cardiovascular disease.

Keywords: SNP informatics, Angiotensin-Converting enzyme 1, Single-nucleotide polymorphisms, Blood Pressure

Introduction

Hypertension is a significant health problem worldwide which accounts for an estimated 7.5 million deaths [1]. The *ACE1* gene is a significant component of the renin-angiotensin system (RAS) [2, 3] which helps to regulate blood pressure and converts the hormone angiotensin I to the active vasoconstrictor angiotensin II [4]. The *ACE1* gene is 21 kb in length on the long arm of chromosome 17 (17q23.3) and is made up of 26 exons and 25 introns. Rigat et al., who first reported the *ACE1* gene polymorphism, proposed that this gene's insertion/deletion

polymorphic form accounts for half (47%) of the phenotypic variance for serum enzyme level.

Considering the implication of *ACE1* Single nucleotide polymorphisms (SNPs) on the phenotypic variance for serum enzyme level and considering the role this gene plays in high blood pressure, it is necessary to study the implications of its SNPs. Therefore, this study aims to identify deleterious and disease-causing nsSNPs in *ACE1* that could serve as molecular and genetic biomarkers to diagnose high blood pressure and are targeted explicitly by inhibitors.

*Correspondence: segun.fatumo@lshtm.ac.uk

¹The African Computational Genomics (TACG) Research Group, MRC/UVRI, and LSHTM, Entebbe, Uganda

Full list of author information is available at the end of the article



Main text

Materials and methods

Single nucleotide polymorphism data retrieval

Non-synonymous SNPs of our target gene were retrieved from the SNPs database (dbSNPs) server of the National Centre for Biotechnology Information (NCBI) [8]. The basis of selection was focused on polymorphisms with small-scale multi-base deletions or insertions and single-base nucleotide substitutions.

Phenotype prediction of deleterious nsSNPs

To identify potential deleterious nsSNPs associated with *ACE1* gene of high blood pressure, we used six different bioinformatics tools implemented in the following web servers: the *Sorting Intolerant from Tolerant* (SIFT) [9], *Protein Variation Effect Analyzer* (PROVEAN) [10], *Polymorphism Phenotyping* (PolyPhen-2) [11], *SNPs&GO* [12], *Predictor of human Deleterious Single Nucleotide Polymorphism* (PhD-SNPs) [13], *PANTHER* [14]. SIFT server classifies nsSNPs based on tolerance index (TI) to be either tolerated ($TI \geq 0.05$) or deleterious (≤ 0.05). PolyPhen-2 classifies nsSNPs as either being possibly/probably (0.00–0.99) damaging or probably benign (≥ 2) by assigning position-specific independent counts (PSIC) score ($0 \leq 2 \leq X$) [15]. The SNP with the highest deleterious prediction by at least five in silico tools were considered the most deleterious nsSNPs for *ACE1* and selected for further investigation.

Protein stability analysis of predicted *ACE1* nsSNPs

To have a higher prediction accuracy of protein stability changes upon single AA mutation, we used the istable 2.0 server [16] to exploit in-built sequence-based tools like the MUpPro [17], interpretable decision tree method iPTREE-STAB [18], I-Mutant 2.0 [19], and also the impact of non-synonymous variations on Protein Stability (INPS) [20]. The I-Mutant tool uses a reliability index (RI) score ranging from 0 through 10 for prediction. DDG is a parameter used by various tools to evaluate the stability of the protein upon mutation at pH 7.0 and 25°C temperatures. A decrease of free energy change ($\Delta\Delta G$) in the value is encoded as 0, and an increase of free energy change (DDG) is encoded as 1.

Protein conservation analysis

To identify putative functional and structural amino acids and estimate their evolutionary conservation profile, we used ConSurf [21], a web server tool that uses the Bayesian approach to analyze the evolutionary pattern of

the amino acid. The conservation grades were mapped onto the query structure and specified using the ConSurf color-code, with cyan-through-purple corresponding to grades 1 (most evolving) through 9 (most evolutionary conserved).

Protein modeling of wild and mutant type *ACE1* and structural difference

To predict our protein's three-dimensional (3D) structure and further analyze the difference between the mutant and wild type of *ACE1* protein, we used the homology modeling tool in Robetta [22], the resultant structure was then viewed using Chimera 1.11[23]. Validation of the predicted protein structure was assessed using ERRAT [24], Verify-3D [25], and PROCHECK [26] programs available from the structural and verification analysis server SAVE (<http://nihserver.mbi.ucla.edu/SAVES>). A TM-align algorithm was then used to compare the wild and the mutant type protein structure of the *ACE1* gene [27].

Molecular docking was carried out using AutoDock vina tools in-built in Chimera (ref). A grid box with coordinate (Centre: $x = 17.588$, $y = 60.433$, $z = 40.395$ and size: $x = 23$, $y = 34$, $z = 33$) was set around the binding site of *ACE1* protein to accommodate Benazepril. Benazepril is an inhibitor drug that is used to treat high blood pressure by lowering blood pressure through inhibiting the formation of angiotensin. The 2D structure of Benazepril in mol2 format was retrieved from DrugBank [29], the structure was further optimized using the GAFF force-field and steepest descent in Avogadro [30].

Results

Non-synonymous single nucleotide polymorphism retrieved from dbSNPs database

The identification of disease-causing SNPs is vital to understand the role a protein plays in disease. A total of 80 nsSNPs were retrieved from the National Centre for biotechnology informatics dbSNPs database server [8] (Additional file 1: Table S1). The retrieval process was filtered to retrieve only nsSNPs that had a clinical consequence reported by Clinvar and implicated in High blood pressure [31].

Identification of deleterious nsSNPs in *ACE1*

All eighty nsSNPs retrieved were subjected to six tools which predicted fifty-nine nsSNPs to be potentially deleterious (Additional file 1: Table S2a). SIFT tool predicted twenty-three nsSNPs to be deleterious (Additional file 1: Table S2b). SNPs&GO tool predicted thirty-seven nsSNPs to be diseased (Additional file 1: Table S2c) while PhD-SNP predicted forty nsSNPs to be diseased (Additional file 1: Table S2c). PANTHER revealed eleven

nsSNPs to be diseased (Additional file 1: Table S2c). According to PROVEAN results, thirty-nine nsSNPs were predicted to be deleterious (Additional file 1: Table S2d). Unlike other tools, PolyPhen-2 had the least deleterious prediction of only six nsSNPs (Additional file 1: Table S2e). Fifty-nine SNPs identified to be deleterious across the six tools were shortlisted to five, which included SNPs that were common in at least five of the predictive tools so that only the highly deleterious SNPs will be used for downstream analyses (Table 1).

Stability profile prediction of nsSNPs protein in ACE1

To predict the protein stability changes, we used the istable algorithm tool, which incorporated results from three sequence-based tools; MUpro, iPTREE-STAB, I-Mutant2.0, and an additional tool, INPS. The five highly deleterious nsSNPs identified from the previous analysis were used to predict protein stability. The five predictive tools revealed that all these nsSNPs decrease stability of ACE1 protein (Additional file 1: Table S3a–d).

Conservation prediction of deleterious nsSNPs in ACE1

To further explore the possible effect of the five nsSNPs, Consurf was used to reveal the essential functional and structural regions by analyzing the evolutionary pattern of the five nsSNPs protein. The results predicted I1018T and F1051V as structural residues making them highly conserved and buried, while P351R, R953Q, and T1187M were indicated as functional residues making them highly conserved and exposed (Additional file 1: Table S4).

Comparative ACE1 modeling of wild and mutant type and structural characterization

The protein sequence of ACE1 (1300 AA residues) with accession number P12821 was retrieved from the UniProt database [32]. The sequence was then inputted into Robetta [22] and was used as a template to model the 3D structure of ACE1 (Fig. 1a). Following the validation of the modeled structure in PROCHECK, ERRAT, and Verify-3D [24–26], the output result from Verify-3D revealed that 89.73% of the residues have an average 3D-1D score of ≥ 0.2 [33] (Fig. 1b). The quality of the

3D protein structure was further assessed through the Ramachandran plot available in PROCHECK. The plot from the predictive model showed that 92.6%, 6.8%, 0.5%, and 0.1% residues be in favored, allowed, outlier regions, and disallowed, respectively, confirming that the protein structure is of good quality (Fig. 1c). ERRAT showed a quality factor of 96.395 (Fig. 1d). Generally, results obtained from the above tools suggested that our modeled protein structure is of good quality and thus could be used for downstream analysis.

ACE1 Mutant type as a potential drug target

To predict the bound conformations and the binding affinity of the molecule with the protein, we performed molecular docking using the AutoDock tool in Chimera [33]. The binding affinity and interacting amino acids residues within the ACE1 binding pocket of each mutant protein upon binding to Benazepril are highlighted in (Table 2). Some of the interactions exhibited by Benazepril include hydrogen bond, pi-alkyl, pi-pi stacked, etc. (Additional file 2: Figs. S1 and S2).

Discussion

We found a total of fifty-nine deleterious nsSNPs, five of which were consistent across the six tools used at mutation points in P351R, R953Q, T1187M, I1018T, and F1051V. The protein stability analysis highlighted five nsSNPs to decreased stability. According to previous studies, decreased protein stability leads to misfolding and degradation of the protein [34]. In addition, using Consurf for conservation analysis showed that three of the amino acid mutation point were highly conserved with a conservation score of 9.

Amino acids involved in protein–protein interaction located at the enzymatic sites are known to be more conserved than others and are also known to be involved in various cellular processes in a biological system including the stability of the genome [35]. For this reason, highly conserved nsSNPs located in the conserved region are more deleterious than nsSNPs located in the variable regions because they destabilize the protein structure and function.

Table 1 Clinically significant information of Deleterious predicted SNPs

SNP ID	Chr17 (GRCh37) location	Nucleotide change	Protein ID	Amino acid change	Functional Consequence
rs2229839	61559033	C > G	ENSP00000290866	P351R	Missense variant
rs143507892	61568688	G > A	ENSP00000290866	R953Q	Coding sequence variant
rs4976	61570937	T > C	ENSP00000290866	I1018T	Coding sequence variant
rs4977	61571297	T > G	ENSP00000397593	F1051V	Coding sequence variant
rs12709442	61574215	C > T	ENSP00000290866	T1187M	Coding sequence variant

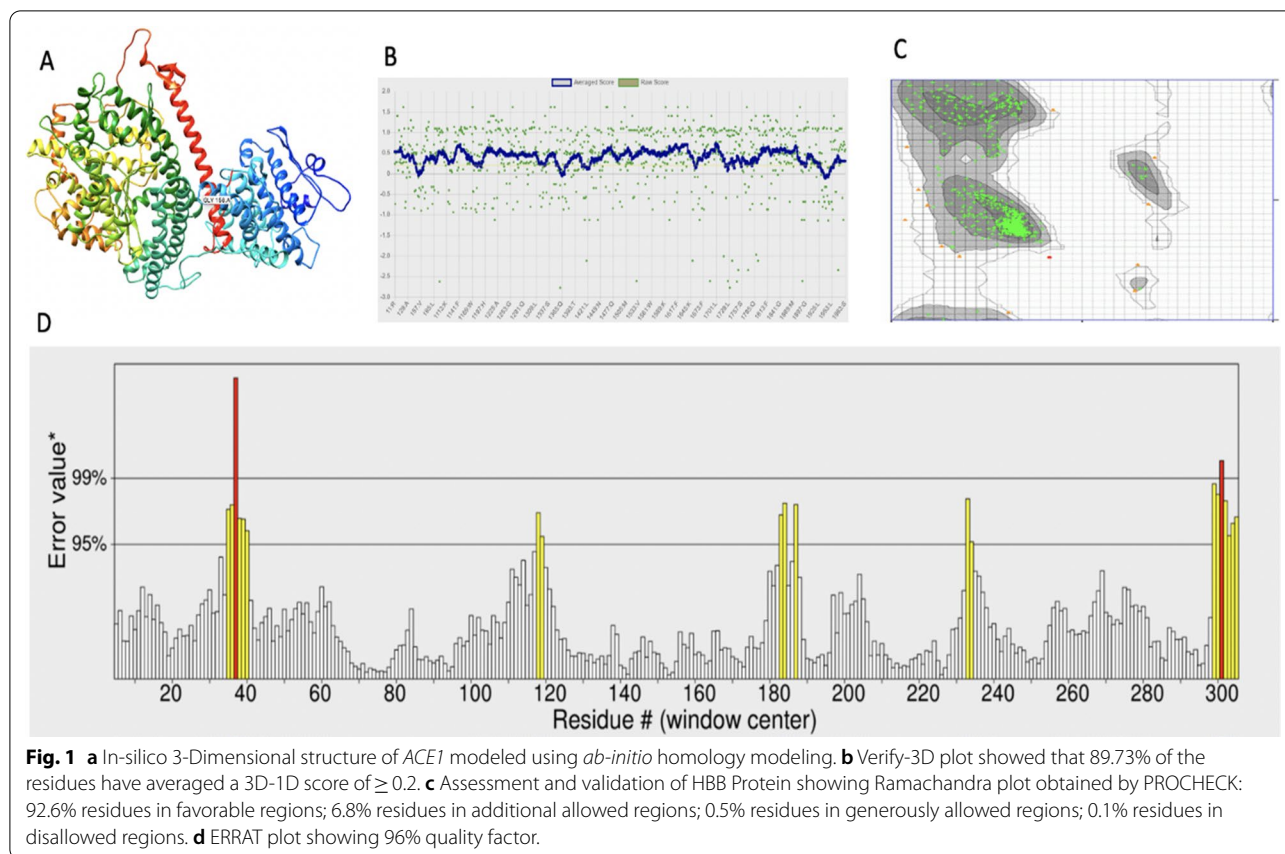


Table 2 Molecular docking of amino acid change against Benazepril

Amino acid change	Interacting amino acids	Binding score (Kcal/mol)
P351R	Tyr805, Cys652, Ala636, Glu658, His635	- 8.7
R953Q	Tyr344, Asn352, Tyr642, Ala345, His692, Ala638, Val800, Trp639	- 9.1
I1018T	Asn348, Asn352, Tyr642, Ala345, Ala638, His692, Val800, Trp639	- 9.0
F1051V	Tyr344, Asn352, Tyr642, Ala345, His692, Val800, Trp639	- 9.2
T1187M	Asn352, Tyr642, Ala345, His692, Ala638, Ser637, Val800, Trp639	- 9.2

Although the use of a single computational tool cannot generate a substantial predictive result of the functional protein region [36], the implementation of several tools could help provide further insight into the impact of nsSNPs on a protein function. Hence, this current study applied multiple in silico tools including INSP, SIFT, Polyphen2, PhD-SNP, PROVEAN, I-Mutant2.0, MUpro, SNPs&GO, Verify-3D, PROCHECK, ERRAT, and ConSurf to identify and evaluate deleterious nsSNPs in the *ACE1* gene. Each of these tools implemented different machine learning approaches such as neural network (NN), decision tree, support vector machine (SVM), and the Bayesian Network to identify nsSNPs that could be used as a drug target for the treatment of high blood pressure [18, 37–39].

Conclusions

Our results showed that SNPs identified through in-silico analysis can alter the structure and function of the *ACE1* gene protein. The five nsSNPs analyzed in this study occur in the functional region of the *ACE1* gene and may therefore alter the functionality of *ACE1*. This can subsequently be used as a basis for enhancing effective drug discovery and pathogenesis targeting *ACE1*. One major strength of this study is that it engages different tools which leverage different algorithms for prediction. This further increases the reliability of the predictions.

Limitation

A major limitation of this study like other in-silico studies is that all the steps taken to predict the impact of the nsSNPs are computer-based, hence, there is a need to explore more robust in-vitro and in-vivo investigations to confirm these result.

Abbreviations

ACE1: Angiotensin-Converting enzyme 1; AA: Amino acid; dbSNPs: Database single-nucleotide polymorphisms; DDG: Delta Delta G; INPS: Impact of non-synonymous variation on protein stability; NCBI: National Centre for Biotechnology Information; nsSNPs: Non-synonymous single nucleotide polymorphisms; PhD-SNP: Predictor of human deleterious single nucleotide polymorphism; Polyphen: Polymorphism phenotyping; PROVEAN: Protein Variation Effect analyzer; SIFT: Sorting Intolerant from Tolerant; SNPs: Single nucleotide polymorphisms; RAS: Renin-angiotensin system; RMSD: Root mean square deviation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13104-021-05879-z>.

Additional file 1: Table S1. Nonsynonymous Single Nucleotide Polymorphism of ACE1 gene retrieved from NCBI dbSNP database. **Table S2. a** Deleterious or disease-causing SNPs across the six in silico tools (SNPs&GO, PROVEAN, PhD-SNP, Polyphen-2, PANTHER and SIFT). **b** SIFT predictive result of deleterious of nsSNPs. **c** SNPs&GO deleterious nsSNPs prediction. **d** PROVEAN predictive result of deleterious or disease-causing nsSNPs. **e** Polyphen-2 predictive result of deleterious or disease-causing nsSNPs. **Table S3. a** INSP protein stability predictions for nsSNPs in ACE1. **b** MUpro stability prediction result for MUpro_NN and MUpro_SVM. **c** I-Mutant prediction result for ACE1 protein stability. **d** istable server protein stability prediction using MUpro_SVM, MUpro_NN, iPTREE-STAB and I-Mutant. **Table S4.** ConSurf conservation prediction analysis.

Additional file 2: Fig. S1. Workflow chart depicting the step-by-step process of insilico analysis of ACE1 gene. **Fig. S2.** ACE1- Benazepril interactions highlighting interacting residues and interaction types for P351R (A), R953Q (B), I1018T (C), F1051V (D), and T1187M (E).

Acknowledgements

SF is an international intermediate fellow funded by the Wellcome Trust grant (220740/Z/20/Z) at the MRC/UVRI and LSHTM. TC is an international training fellow supported by the Wellcome Trust grant (214205/Z/18/Z). The national institutes of Health Common Fund to the H3ABionet Project grant number (5U24HG006941-09).

Authors' contributions

SF and OS conceptualized the study. BU and OS led the main analyses. BU wrote the first draft of the manuscript. CE, OI, TC, and SF reviewed the first draft. SF supervised the project. All authors read and approved the final manuscript.

Funding

Not applicable. Wellcome trust, 220740/Z/20/Z, Segun Fatumo.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. ACE UniProt accession number P1282. Sorting Intolerant from Tolerant (SIFT): <https://sift.bii.a-star.edu.sg/>. Protein Variation Effect Analyzer (PROVEAN): http://provean.jcvi.org/seq_submit.php. Polymorphism Phenotyping (PolyPhen-2): <http://genetics.bwh.harvard.edu/pph2/>. SNPs&GO: <https://snps-and-go.biocomp.unibo.it/snps-and-go/>. Predictor of human Deleterious Single Nucleotide Polymorphism (PhD-SNPs): <https://snps.biofold.org/phd-snp/phd-snp.html>. PANTHER: <http://www.pantherdb.org/tools/csnpscoreForm.jsp>, interpretable decision tree method iPTREE-STAB: <http://bioinformatics.myweb.hinet.net/iptree.htm>. I-Mutant 2.0: <https://folding.biofold.org/i-mutant/i-mutant2.0.html>, non-synonymous variations on Protein Stability (INPS): <https://inpsmd.biocomp.unibo.it/inpsSuite>. ConSurf: <https://consurf.tau.ac.il/>. Robetta: <https://robetta.bakerlab.org/>. SAVES: <https://saves.mbi.ucla.edu/>

[pantherdb.org/tools/csnpscoreForm.jsp](http://www.pantherdb.org/tools/csnpscoreForm.jsp), interpretable decision tree method iPTREE-STAB: <http://bioinformatics.myweb.hinet.net/iptree.htm>. I-Mutant 2.0: <https://folding.biofold.org/i-mutant/i-mutant2.0.html>, non-synonymous variations on Protein Stability (INPS): <https://inpsmd.biocomp.unibo.it/inpsSuite>. ConSurf: <https://consurf.tau.ac.il/>. Robetta: <https://robetta.bakerlab.org/>. SAVES: <https://saves.mbi.ucla.edu/>

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Other than the grants listed in the acknowledgment section, the authors declare that they have no other competing interests.

Author details

¹The African Computational Genomics (TACG) Research Group, MRC/UVRI, and LSHTM, Entebbe, Uganda. ²The African Center of Excellence in Bioinformatics of Bamako (ACE-B), University of Sciences, Techniques and Technologies of Bamako, Bamako, Mali. ³University at Albany, Albany, NY, USA. ⁴Department of Biochemistry, Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria. ⁵H3Africa Bioinformatics Network (H3ABioNet) Node, Centre for Genomics Research and Innovation, NABDA/FMST, Abuja, Nigeria. ⁶Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. ⁷MRC/Wits Developmental Pathways for Health Research Unit, Department of Pediatrics, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. ⁸Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK.

Received: 5 August 2021 Accepted: 6 December 2021

Published online: 20 December 2021

References

- Mendis S, Global status report on non-communicable diseases 2010. World Health Organisation; 2010. http://www.who.int/nmh/publications/ncd_report2010/en/.
- Krishnan R, Sekar D, Karunanithy S, Subramaniam S. Association of angiotensin converting enzyme gene insertion/deletion polymorphism with essential hypertension in south Indian population. *Genes Dis*. 2016. <https://doi.org/10.1016/j.gendis.2016.03.001>.
- Martínez-Rodríguez N, et al. Single nucleotide polymorphisms of the angiotensin-converting enzyme (ACE) gene are associated with essential hypertension and increased ACE enzyme levels in mexican individuals. *PLoS ONE*. 2013;8(5): e65700. <https://doi.org/10.1371/journal.pone.0065700>.
- J. H. Fountain and S. L. Lappin, "Physiology, Renin Angiotensin System," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2021. [Online]. <http://www.ncbi.nlm.nih.gov/books/NBK470410/>. Accessed 23 Jun 2021.
- Alwi ZB. The Use of SNPs in Pharmacogenomics Studies. *MJMS*. 2005;12(2):4–12.
- Gurdasani D, et al. Uganda genome resource enables insights into population history and genomic discovery in Africa. *Cell*. 2019;179(4):984–1002. e36. <https://doi.org/10.1016/j.cell.2019.10.004>.
- Soremekun OS, Soliman MES. From genomic variation to protein aberration: mutational analysis of single nucleotide polymorphism present in ULBP6 gene and implication in immune response. *Comput Biol Med*. 2019;111: 103354. <https://doi.org/10.1016/j.combiomed.2019.103354>.
- Sherry ST, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29(1):308–11. <https://doi.org/10.1093/nar/29.1.308>.

9. 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):1073–81. <https://doi.org/10.1038/nprot.2009.86>.
10. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinforma Oxf Engl*. 2015;31(16):2745–7. <https://doi.org/10.1093/bioinformatics/btv195>.
11. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013. <https://doi.org/10.1002/0471142905.hg0720s76>.
12. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: a web server for predicting the deleterious effect of human protein variants using functional annotation. *BMC Genomics*. 2013;14(Suppl 3):S6. <https://doi.org/10.1186/1471-2164-14-S3-S6>.
13. Capriotti E, Fariselli P. PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants. *Nucleic Acids Res*. 2017. <https://doi.org/10.1093/nar/gkx369>.
14. Thomas PD, et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res*. 2003;13(9):2129–41. <https://doi.org/10.1101/gr.772403>.
15. Xi T, Jones IM, Mohrenweiser HW. Many amino acid substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function. *Genomics*. 2004;83(6):970–9. <https://doi.org/10.1016/j.ygeno.2003.12.016>.
16. Chen C-W, Lin J, Chu Y-W. iStable: off-the-shelf predictor integration for predicting protein stability changes. *BMC Bioinformatics*. 2013;14(Suppl 2):S5. <https://doi.org/10.1186/1471-2105-14-S2-S5>.
17. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res*. 2005. <https://doi.org/10.1093/nar/gki396>.
18. Huang L-T, Gromiha MM, Ho S-Y. iPTREE-STAB: interpretable decision tree based method for predicting protein stability changes upon mutations. *Bioinformatics*. 2007;23(10):1292–3. <https://doi.org/10.1093/bioinformatics/btm100>.
19. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res*. 2005. <https://doi.org/10.1093/nar/gki375>.
20. Fariselli P, Martelli PL, Savojardo C, Casadio R. INPS: predicting the impact of non-synonymous variations on protein stability from sequence. *Bioinformatics*. 2015;31(17):2816–21. <https://doi.org/10.1093/bioinformatics/btv291>.
21. Ashkenazy H, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*. 2016;44(W1):W344–50. <https://doi.org/10.1093/nar/gkw408>.
22. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*. 2004. <https://doi.org/10.1093/nar/gkh468>.
23. Pettersen EF, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605–12. <https://doi.org/10.1002/jcc.20084>.
24. Colovos C, Yeates TO. Verification of protein structures: patterns of non-bonded atomic interactions. *Protein Sci Publ Protein Soc*. 1993;2(9):1511–9. <https://doi.org/10.1002/pro.5560020916>.
25. Kresge CT, Leonowicz ME, Roth WJ, Vartuli JC, Beck JS. Ordered mesoporous molecular sieves synthesized by a liquid-crystal template mechanism. *Nature*. 1992;359(6397):710–2. <https://doi.org/10.1038/359710a0>.
26. R. A. Laskowski, M. W. MacArthur, and J. M. Thornton, PROCHECK: validation of protein-structure coordinates. In: *International Tables for Crystallography*, American Cancer Society, 2012, pp. 684–687. <https://doi.org/10.1107/97809553602060000882>.
27. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res*. 2005;33(7):2302–9. <https://doi.org/10.1093/nar/gki524>.
28. Carugo O, Pongor S. A normalized root-mean-square distance for comparing protein three-dimensional structures. *Protein Sci Publ Protein Soc*. 2001;10(7):1470–3. <https://doi.org/10.1110/ps.690101>.
29. Wishart DS, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res*. 2008. <https://doi.org/10.1093/nar/gkm958>.
30. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminformatics*. 2012;4(1):17. <https://doi.org/10.1186/1758-2946-4-17>.
31. Landrum MJ, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014. <https://doi.org/10.1093/nar/gkt1113>.
32. T. UniProt Consortium et al. UniProt: the universal protein knowledge-base. *Nucleic Acids Res* 2017, doi: <https://doi.org/10.1093/nar/gkw1099>.
33. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multi-threading. *J Comput Chem*. 2010;31(2):455–61. <https://doi.org/10.1002/jcc.21334>.
34. Singh SM, Kongari N, Cabello-Villegas J, Mallela KMG. Missense mutations in dystrophin that trigger muscular dystrophy decrease protein stability and lead to cross-beta aggregates. *Proc Natl Acad Sci U S A*. 2010;107(34):15069–74. <https://doi.org/10.1073/pnas.1008818107>.
35. Williamson K, Schneider V, Jordan RA, Mueller JE, Pozzi MH, Bryk M. Catalytic and functional roles of conserved amino acids in the SET domain of the *S. cerevisiae* lysine methyltransferase set1. *PLoS ONE*. 2013. <https://doi.org/10.1371/journal.pone.0057974>.
36. Gassoum A, Abdelraheem NE, Elsadiq N. Comprehensive analysis of rsSNPs associated with hypertension using in-silico bioinformatics tools. *OALib*. 2016;03(07):1–24. <https://doi.org/10.4236/oalib.1102839>.
37. Capriotti PFE, Casadio R: a neural-network-based method for predicting protein stability changes upon single point mutations. *Bioinformatics*. 2004;20(suppl 1):i63–8.
38. Capriotti E, Fariselli P, Calabrese R, Casadio R. Predicting protein stability changes from sequences using support vector machines. *Bioinforma Oxf Engl*. 2005. <https://doi.org/10.1093/bioinformatics/bti1109>.
39. Mayrose I, Graur D, Ben-Tal N, Pupko T. Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol Biol Evol*. 2004;21(9):1781–91. <https://doi.org/10.1093/molbev/msh194>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

