# Dynamics of the personalities of PCSK9 on missense variants (rs505151 and rs562556) from elderly cohort studies in Brazil

Vitor Galvão Lopes, Victor Fernandes de Oliveira, Livia Mendonça Munhoz Dati, Michel Satya Naslavsky, Glaucio Monteiro Ferreira & Mario Hiroyuki Hirata

View supplementary material

Published online: 03 Apr 2023.

Submit your article to this journal

Article views: 217

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Taylor & Francis
Taylor & Francis Group

Check for updates

# Dynamics of the personalities of PCSK9 on missense variants (rs505151 and rs562556) from elderly cohort studies in Brazil

Vitor Galvão Lopes[a] iD, Victor Fernandes de Oliveira[a] iD, Livia Mendonça Munhoz Dati[a] iD, Michel Satya Naslavsky[b] iD, Glaucio Monteiro Ferreira[a†] iD and Mario Hiroyuki Hirata[a†] iD

[a]Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of Sao Paulo, São Paulo, Brazil; [b]Human Genome and Stem Cell Research Center (HUG-CELL), Biosciences Institute, University of Sao Paulo, São Paulo, Brazil

## ABSTRACT

The Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) promotes the degradation of the low-density lipoprotein receptors (LDLR). Gain-of-function (GOF) variants of PCSK9 significantly affects lipid metabolism leading to coronary artery disease (CAD), owing to the raising the plasma low-density lipoprotein (LDL). Considering the public health matter, large-scale genomic studies have been conducted worldwide to provide the genetic architecture of populations for the implementation of precision medicine actions. Nevertheless, despite the advances in genomic studies, non-European populations are still underrepresented in public genomic data banks. Despite this, we found two high-frequency variants (rs505151 and rs562556) in the ABraOM databank (Brazilian genomic variants) from a cohort SABE study conducted in the largest city of Brazil, São Paulo. Here, we assessed the structural and dynamical features of these variants against WT through a molecular dynamics study. We sought fundamental dynamical interdomain relations through Perturb Response Scanning (PRS) and we found an interesting change of dynamical relation between prodomain and Cysteine-Histidine-Rich-Domain (CHRD) in the variants. The results highlight the pivotal role of prodomain in the PCSK9 dynamic and the implications for the development of new drugs depending on patient group genotype.
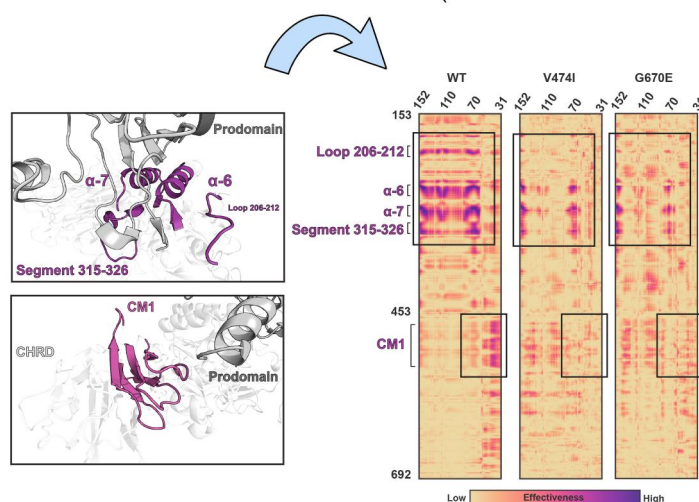
Motions of PCSK9 on missense variants (rs505151 and rs562556)

## Introduction

The Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) negatively regulates the low-density lipoprotein receptors (LDLR) through its lysosomal degradation (Cunningham et al., 2007). Variants that promote gain-of-function (GOF) of PCSK9 affects significantly lipid metabolism by raising the plasma low-density lipoprotein (LDL), which entails the development of coronary artery disease (CAD) (Seidah et al., 2014). The GOF phenotype is linked to familial hypercholesterolemia (FH), a genetically inherited disease that is responsible for early CAD events (Berberich & Hegele, 2019). Meanwhile, the PCSK9 variants have gained clinical interest due to the discovery of

loss-in-function (LOF) variants that were protective against CAD (Berge et al., 2019). As well, that discovery started a wide field for targeting PCSK9 as lipid-lowering therapy for FH patients (Costet et al., 2008). Hence, the PCSK9-variants assessment through whole-genome sequencing (WGS) has pivotal importance for FH understanding and control (Berberich & Hegele, 2019). The populational variants represented by whole-genome sequencing (WGS) orientate functional studies, drug discovery strategies, and precision medicine (Vicini et al., 2016).

Large-scale genomic studies have been conducted worldwide to provide the genetic architecture of populations for the implementation of precision medicine actions (Rocha et al., 2020). Regardless of the noteworthy advances in genomic studies, non-European populations were underrepresented in public genomic data banks (Popejoy & Fullerton, 2016). In the Brazilian context, the ABraOM (Brazilian genomic variants) repository contains WGS data from the SABE study (Naslavsky et al., 2017), a sample of 1171 elder individuals from São Paulo, Brazil's largest city. The occurring variants of this study may represent the Brazilian main frame of variants (Naslavsky et al., 2022) and might contribute to the understanding of more prevalent PCSK9 variants in São Paulo city and Brazil. Additionally, the composition of populational variants might affect the drug binding mode and impair therapy efficacy (Wan et al., 2021). Thereby, the continuous evaluation of variant structural features is required, either to overcome these possible issues or unravel the complex structural features of proteins through their variants.

The PCSK9 presents a self-inhibited proconvertase structure (Hampton et al., 2007), which compromises a prodomain (**PD**: 31-152) containing a segment 31-50 that is a transient alpha helix (THA), a catalytic domain (**CD**: 153-404), and a 3-module C-terminal domain (**CM1**: 457-530, **CM2**: 534-601, and **CM3**: 608-692) linked by an exposed hinge region (405–452) (Cunningham et al., 2007). The catalytic domain is responsible for binding the epidermal growth factor A (**EGF-A**) from LDLR (Sundararaman et al., 2021). Some of the critical mutations that lead to the FH occur in this domain, promoting increased affinity and posterior degradation in the late lysosome due to the lower pH (Cunningham et al., 2007). On the other hand, the C-terminal does not interact directly with EGF-A, but a set of variants in that domain would increase LDLR degradation (Sánchez-Hernández et al., 2019) and interfere with the PCSK9 secretion (Saavedra et al., 2012). There are mainly two clusters of variants occurring in the C-terminal structure: The GOF variants that lie on the CM1 loops and beta-sheets (R469W, R499H (Sánchez-Hernández et al., 2019), E482G, A514T, F515L, and A522T); The LOF CM3 variants (P616L, S668R, and C679X) (Saavedra et al., 2012). Nonetheless, the CM2 is unique among the C-terminal modules that are required for PCSK9 extracellular activity (Saavedra et al., 2012). Connecting the CD to the C-terminal, the hinge domain plays a critical role in PCSK9 folding and secretion (Deng et al., 2020). Previously reported, the hinge domain variant R434W results in a misfolded PCSK9 with ~70% decreased extracellular pathway activity (Dubuc et al., 2010). An extended observation showed that R424W leads to a lack of intracellular activity on the LDLR (Saavedra

et al., 2012), suggesting that the hinge domain is critical for PCSK9 active fold both intracellular and extracellular.

As mentioned before, the public genomic databanks underrepresent the Brazilian population due to the singular admixed characteristics (Rodrigues de Moura et al., 2015). In this way, the assessment of the tangle PCSK9 mechanism by its domain synergies is the paramount importance to getting insights into the genetic disease and guiding new high-cover drug discovery approaches. Finally, we aimed to select the more frequent PCSK9 variants from the ABraOM databank and explore their structural behavior through molecular dynamics simulations, precisely, the molecular collective motions, domain interchange, and interactions between PCSK9 and EGF-A domain.

## Material and methods

### Variant selection

The search for the most prevalent PCSK9 variants was done on the web server ABraOM (https://abraom.ib.usp.br/). We searched for the key gene 'PCSK9' in the Genome reference SABE-WGS-1171 (hg38) using a quality filter for GATK pass and Exonic or Splicing/CDS variants only. The frequencies were obtained as a standard output of the ABraOM request, representing the ratio between allelic number and allelic count. Therefore, the higher frequency variants were selected for molecular dynamics analysis.

### Homology modeling

The PCSK9 structure was constructed using a template, three tridimensional structures deposited in the Protein Data Bank (https://www.rcsb.org/) (PDBs: 3bps (Kwon et al., 2008), 6mvu, 2p4e (Cunningham et al., 2007)). The human PCSK9 protein sequence was downloaded from UniprotKB: (ID: Q8NBP7) and the modeling was done by software Modeller 2021 (Webb & Sali, 2016). The structure validation was assessed by the Ramachandran plot (Anderson et al., 2005), overall quality Z-score (Wiederstein & Sippl, 2007), and local model quality (Sippl, 1993) (Figure S1).

### Structure preparation

The variants rs505151 and rs562556 were created by Maestro edition tool. The WT and variants structures were prepared by adding hydrogen atoms and fixing missing side chains using the Biopolymer tool of Sybyl-X 2.0 (TRIPOS Inc.) and PROPKA (Bas et al., 2008) code (Github code: https://github.com/jensengroup/propka).

### Molecular dynamics simulations

The prepared PCSK9 variants and WT structures were simulated through Molecular Dynamics (MD). The engine used was the software Desmond (Bowers et al., 2006) with the OPLS-2005 force-field (Banks et al., 2005). The system was constructed with a protein-ligand complex, a predefined solvent water model (TIP3P (Jorgensen et al., 1983)), charge
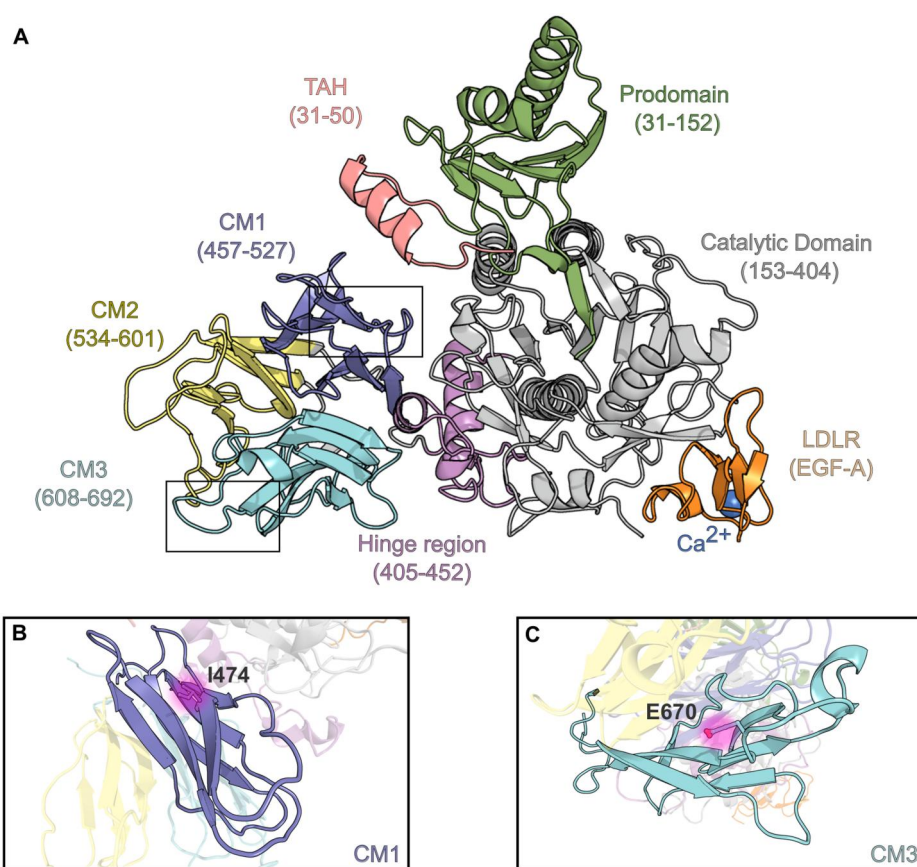
**Figure 1.** Overview of PCSK9, variants, and EGF(A) structure separated by domains: (**A**) Prodomain (green) and the transient alpha helix (TAH, salmon), Catalytic (grey), connecting Hinge domain (violet), and Cysteine-Histidine-Rich-Domain (CHRD) divided in 3 modules (CM1 (purple), CM2 (yellow) and CM3 (cyan)). The EGF-A domain from LDLR is shown in orange. (**B, C**) The CM structures with a pink highlight over the respective variants p.(V474I) and p.(G670E).

neutralizing counterions (Na$^+$ or Cl$^-$), and periodic boundary conditions (PBC). The system box was constructed considering at least the distance of 13 Å from the protein atoms to the box edges. Short-range coulombic interactions were set at a cut-off value of 9.0 Å and the integration time step calculated every 2 fs, while long-range coulombic interactions were estimated using the Smooth Particle Mesh Ewald (PME) method (Darden et al., 1993). Each system was subjected to at least 1 μs and all trajectory and interaction data are available on the Zenodo repository (DOI: 10.5281/zenodo.6612263).

Atomic interactions, distances, and secondary structure elements were determined using the Simulation Event Analysis pipeline as implemented in Maestro 2019v.4 (Schrödinger LCC). The criteria for protein-ligand H-bond are 2.5 Å distance between the donor and acceptor atoms (D—H⋯A); ≥120° angle between the donor-hydrogen-acceptor atoms (D—H⋯A); and ≥90° angle between the hydrogen-acceptor-bonded atoms (H⋯A—X). The protein-water and water-ligand H-bonds were considered when 2.8 Å (D—H⋯A); ≥110° (D—H⋯A); and ≥90° (H⋯A—X). Non-specific hydrophobic interactions are defined by the presence of a hydrophobic side chain within 3.6 Å of the ligand's aromatic or aliphatic carbons. π-π interactions are recorded when two aromatic groups are stacked face-to-face or face-to-edge and within 4.5 Å of distance. The MD trajectories were visualized in PyMol v.2.2.3 (Schrödinger LCC, New York, NY, USA). The distance and angles analyses were used to assess the movement of structures. The distance of selected regions was calculated by the center of mass distance using the script (trj_asl_distance.py) available on Schrodinger package 2019v.4.

### Principal component analysis and essential dynamics

The Principal Component Analysis (PCA) was used to study the main features of WT PCSK9 against its mutations. The backbone of domains was extracted and centered to the initial frame of WT simulation (trj_selection_dl.py and trj_center.py) from Schrodinger package 2019v.4. The simulations were aligned and merged using gmx_trjconv from GROMACS [22] 2022. The covariance matrix was calculated by gmx_covar and the eigenvectors were analyzed through gmx_anaeig.

The essential dynamics analysis (Amadei et al., 1993) was generated by the script trj_essential_dynamics.py from the Schrödinger pipeline. For the calculations, the trajectories were aligned by trj_align.py. The modevectors were carried out by PyMol v.2.2.3 (Schrödinger LCC, New York, NY, USA). PCA Plotting information is available at Github: https://github.com/vitor-lopes17/PCSK9-mutations_PCA-DATA/blob/main/Mutations-PCSK9-PCA_analysis.ipynb

### Perturb response scanning (PRS) analysis

The PRS analysis was carried out by the Python package ProDy v.2.0 (Zhang et al., 2021). The PRS matrix was generated

by the function calcPerturbResponse (Atilgan & Atilgan, 2009). The function was fed using the covariance matrix output from essential dynamics analysis. The PRS analysis was generated to PCSK9 WT and variants individually.

## Results

### Examining the PCSK9 structure and its mutational landscape

Two high-frequency nonsynonymous (SNV) missense variants were found at the ABraOM databank (Naslavsky et al., 2017): the p.(G670E) (Mohamed et al., 2021) (dbSNP: rs505151) and p.(V474I) (Mohamed et al., 2021) (dbSNP: rs562556), presenting 91.93% and 83.26% of frequency in the SABE-WGS-1171 cohort, respectively. In this way, tridimensional models of WT PCSK9-EGF(A) and variants G670E/V474I-EGF(A) were built by homology modelling strategy and subsequently submitted to molecular dynamics studies. p.(G670E) lies on a two-beta-sheet connecting loop of CM3 (Figure 1).

Given the importance of EGF-A affinity in the GOF phenotype (Bottomley et al., 2009), we investigated the interactions between PCSK9 and EGF(A) domain in our simulations. Our results show that PCSK9 interacts with the EGF(A) domain more frequently in variant missense simulations than in WT simulations (Figure S3A). We also recorded the total number of hydrogen bonds for each frame (Figure S3B) and observed that the variants had higher fluctuations in the number of hydrogen bonds compared to WT. We further investigated these fluctuations by analyzing the center of mass distance between EGF(A) and PCSK9 (Figure S3C) and the Root Mean Square Fluctuation of EGF(A) domain (Figure S3D).

PCA revealed that PC1 on prodomain captured 58.16% of the total variance of simulations (Figure 2A). This amount of variation suggested an extensive conformational variation of the WT prodomain than p.(V474I) and p.(G670E) (Figure 2A, cyan and purple, respectively). Further, PC2 showed a conformational space overlap among simulations represented by 20.95% of total variation (Figure 2B).

We found that the segments 31-50 from the WT prodomain showed more than 70% secondary structure prevalence when compared to the mutated PCSK9 (Figure 2C). Interestingly, the TAH segment was more labile in the WT simulations (Figure 2D). To explore the structural features involved, representative frames of the simulations were selected from PCA clusters (Figure 2E–G). In all those frames we notice interdomain contacts (prodomain-CHRD). The most important interaction was seen for R469 in WT and R476 in both variants (Figure 2H–J). Hence, the panel of interactions between prodomain and CHRD was shown for hydrogen bonding (Figure 2K) and water-bridge bonding (Figure 2L) across the systems. The variants showed interdomain interactions in higher extension when compared to WT.

### PRS analysis showed the effectiveness of mutated residues and the overall residues

To address the dynamical interchange of prodomain and CHRD, a multidomain essential dynamics analysis was generated to

perform the PRS analysis. The PRS is based on a linear response theory where the perturb forces applied via covariance matrix could determine the effectors (by effectiveness) in allosteric signal transduction (Dutta et al., 2015). Therefore, the PRS is useful to determine how single residues or regions can affect protein conformational changes. Here, we applied that analysis to assess qualitatively how the variant residues interfere with the overall dynamics of PCSK9. Firstly, we sought the PRS effectiveness caused by residues 474 and 670 among the systems (Figure 3A–D). Given this, the perturb of WT residue position 474 (WT: V474) acts as a prodomain THA segment effector (Figure 3A) while the perturb of variant 474 (Variant: I474) does not (Figure 3B). Furthermore, the perturb of WT residue position 670 (WT: G670) (Figure 3C) presented more effectiveness on prodomain residues than variant 670 (Variant: E670) (Figure 3D). These two variants showed less perturb in an important region of prodomain when compared to WT residues. Further, the overall PRS analysis of the prodomain (chain P) against the catalytic and CHR domains (chain A) identified a chain A region that acts as a direct effector for the prodomain (Figure 3E). The effectiveness of these regions was stronger in the WT simulations, suggesting that these CHRD mutations could disperse and diminish the allosteric control of the prodomain through chainA-chainP interface regions.

The crucial regions of chain A were identified as: I. Two helices (α6, α7) and two loop segments (206–212, 315–316) from the catalytic domain (Figure 3E and F), and II. The beta sheets from CM1 (Figure 3E and G). Region I of WT showed high effectiveness to the prodomain residues mainly in the segment 50-152, while region I of variants showed lower effectiveness. Region II of WT affected mainly the TAH segment, while region II of variants presented a low effect on TAH residues. Despite this, region II from variants presented a diffuse effectiveness behavior across prodomain residues, indicating that the variants could perturb the allosteric regulation of CHRD.

## Discussion

### Unraveling the affinity of EGF(a), interdomain interactions, and dynamics: Insights into the relationship between variants and pathogenicity

The ABraOM databank variants V474I and G670E were reported previously among other populations (Evans & Beil, 2006; Ferreira et al., 2020; Gai et al., 2021; Li et al., 2020; Mohamed et al., 2021). The V474I variant is associated with raised circulating LDL, blood glucose, body mass index (BMI), mean platelet volume (MPV), and red blood cell distribution width (RDW), suggesting pleiotropic effects more than affecting lipid metabolism (Gai et al., 2021). Therefore, the p.(V474I) variant is in the second beta-sheet of CM1, which is connected to the hinge loop (Figure 1). The CM1 is a module that was suggested to control allosterically the hinge domain (Deng et al., 2020). As mentioned before, the CM1 region compromises a well-described set of GOF variants.

A meta-analysis study has revealed that the G670E variant is associated with increased circulating levels of triglycerides and LDL, as well as a higher risk of cardiovascular disease
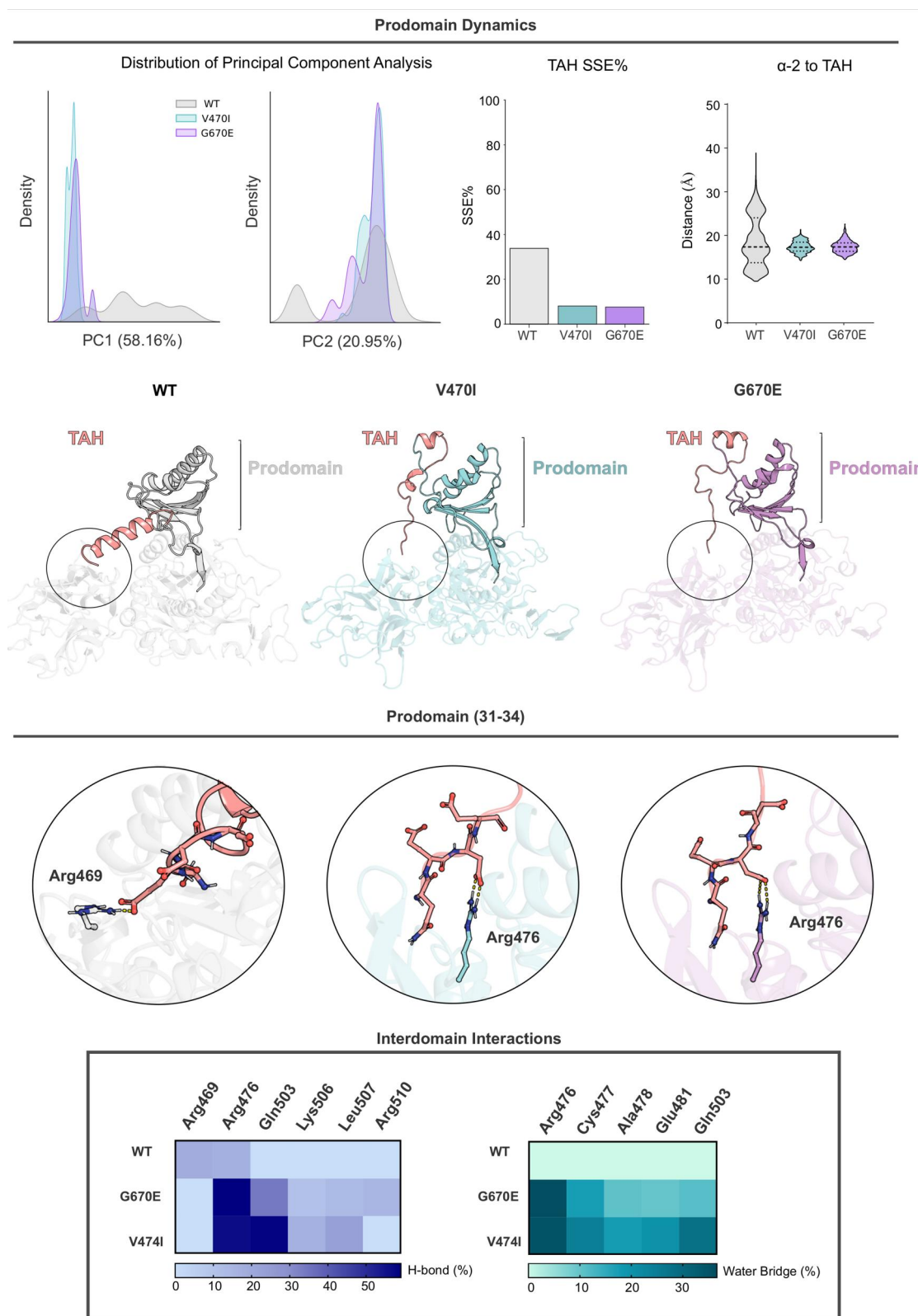
**Figure 2.** PCSK9 prodomain dynamics and interchange with CHRD. (**A, B**) represent the density of displacement along eigenvectors 1 and 2. The complete data of PCA is shown in Figure S2 (**C**) The secondary structure elements (SSE%) along with the simulation for the segment TAH. (**D**) The distances of TAH from prodomain α-2 calculate by the Center of Mass. (**E, F, G**) Representative structure of PCSK9 and its variants selected by PCA clusters of prodomain. A highlight of the TAH structure shows the interdomain contact. (**H, I, J**) Show the principal interactions between prodomain and CHRD that match with figures E, F, and G, respectively. The prevalence of interdomain interactions is shown in Figure S3.

(Qiu et al., 2017). However, in East Asian populations, the G670E variant is associated with extremely low levels of LDL-C, even among individuals with common variants in both APOB and PCSK9 genes (Lee et al., 2017). This suggests that variants in APOB may offer better protection than PCSK9 variants in this scenario. Additionally, in elderly populations in
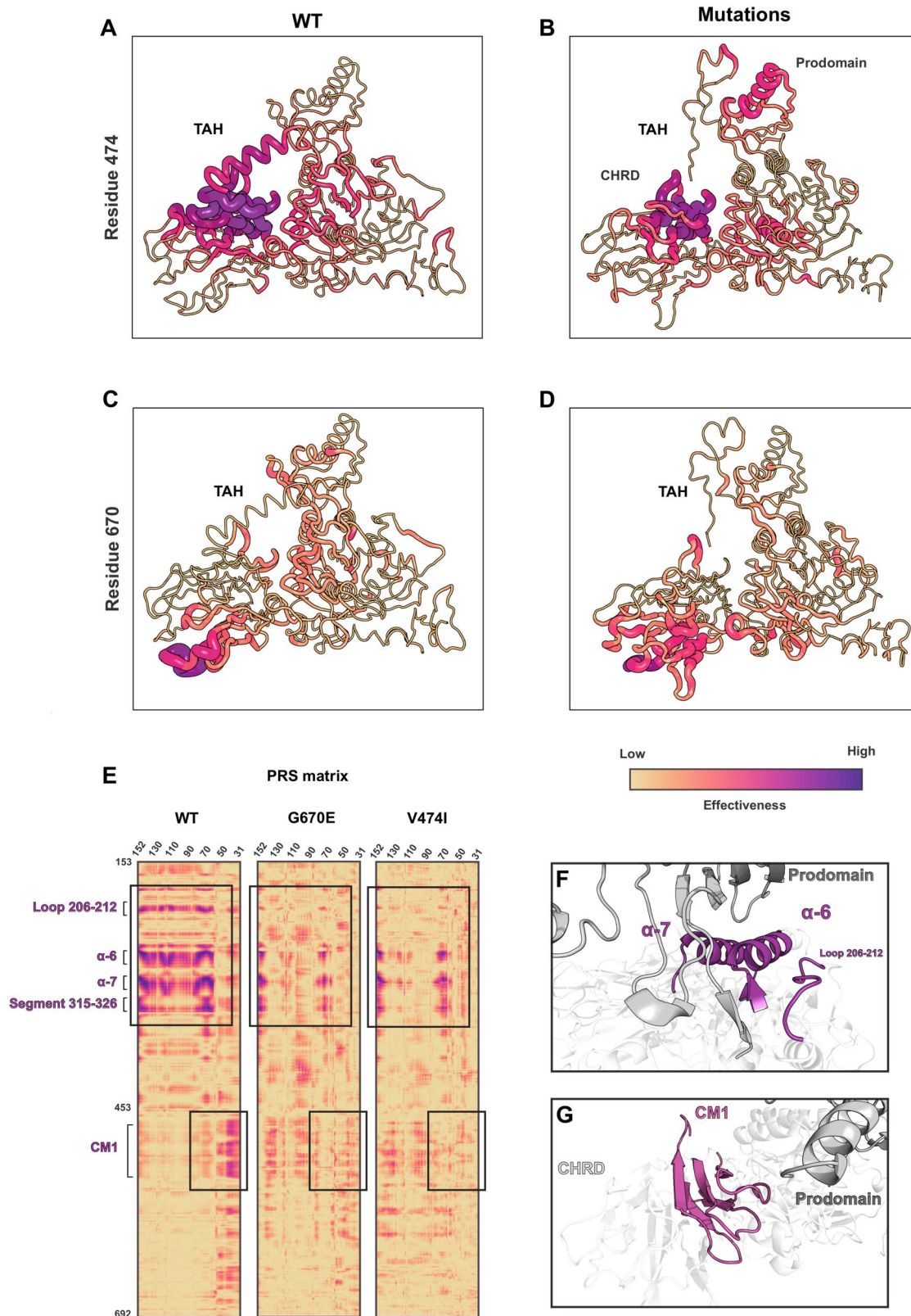
# Effect of residues 474 & 670



**Figure 3.** The effectiveness of mutated residues and the overall effectiveness of the prodomain. (**A, B**) The pair of effectiveness for the WT and V474I variant for position 474, respectively. (**C, D**) The pair of effectiveness for the WT and G670E variant for position 670, respectively. (**E**) The PRS matrix of chain A (catalytic domain and CHRD) against prodomain. The highlights show the regions of more effectiveness in the prodomain and the related structures from chain A. The full PRS matrix is shown in Figure S4. (**F, G**) Represent the structure of the first and second highlights of Figure E, respectively. The secondary structures nomenclature index used in these images is represented in Figure S5.

Brazil, certain characteristics related to miscegenation may have been misunderstood compared to other populations around the world.

An increased affinity to the LDLR's EGF(A) is pointed to as the main factor of GOF phenotype such as seen in D374Y classic mutation (Bottomley et al., 2009; Mousavi et al., 2011; Pandit et al., 2008). However, the frequencies of interactions observed in the N-terminal region of the EGF(A) domain in variant systems alone were insufficient to account for GOF phenotype exhibited by these variants. In summary, the results indicate an increase in the flexibility of the EGF(A) N-terminal region, which has resulted in changes in the distribution of hydrogen bonds, rather than an increase in the number of interactions.

We found an essential dynamics difference between WT and mutated proteins, mainly related to the prodomain and CHRD interchange. The segment THA from the prodomain could be the key to understanding the intricate relationships between PCSK9 domains. In this way, we hypothesized that the formation of alpha helix is a transient event that is triggered by medium content and THA lability. Accordingly, the work of Sarkar et al. (2020) demonstrated the formation of the transient alpha helix could be triggered by the increase of n-dodecyl phosphocholine (DPC) micelles in the medium and it is dependent on the integrity of the THA segment. Additionally, we propose that the lability behavior contributes to the alpha-helix formation and this event is controlled allosterically by the other domains. As seen in the PRS analysis (Figure 3E), the CHRD mutations could interfere with the allosteric regulation of the prodomain, promoting a decrease in the alpha helix formation. To this extent, the TAH formation could be the central key that allows LDL particles to bind on PCSK9 and promote its regulation (Kosenko et al., 2013). Some CHRD mutations could diminish (R469W and F515L) or abolish (R496W) the affinity of PCSK9 by LDL particules (Sarkar et al., 2020). It has been previously reported that more than one-third of circulating human PCSK9 is bound to LDL-C, and this fraction is predominantly present in a less-active monomeric form of PCSK9. Conversely, the unbound fraction can self-associate and generate a multimeric form of PCSK9, which has a higher affinity for LDLR and is more efficient in promoting its catabolism (Fan et al., 2008). In this context, the relationship between the prodomain and LDLR binding may represent the underlying mechanism of CHRD mutations, which desensitize PCSK9 to LDLR binding and favor a multimeric state over a monomeric state. Moreover, these findings have implications for drug discovery, particularly regarding the interdomain interactions. For instance, antibodies targeting the CM1 region have been shown to reduce LDL-C levels (Schiele et al., 2014), and considering the results of this study, this could be attributed to the loss of THA flexibility and immobilization of region II (Figure 3G).

## Conclusion

In conclusion, we used the molecular dynamics simulations and PRS to extract regions that might exhibit unique contributions to prodomain dynamics. PRS analysis showed that the missense variants on PCSK9 reveals the conformational changes that could affect the PCSK9 structure than WT. Hence, it allows for the understanding of the dynamics of important CHRD variants found in Brazilian elderly people and might propose new drug targets for drug discovery based on genotype group. The results highlighted the pivotal role of prodomain in the PCSK9 dynamic and the implications for the development of new drugs depending on patient group genotype.

## ORCID

*Vitor Galvão Lopes* (iD) http://orcid.org/0000-0002-7936-6931
*Victor Fernandes de Oliveira* (iD) http://orcid.org/0000-0001-9219-556X
*Livia Mendonça Munhoz Dati* (iD) http://orcid.org/0000-0002-6303-2144
*Michel Satya Naslavsky* (iD) http://orcid.org/0000-0002-9068-1713
*Glaucio Monteiro Ferreira* (iD) http://orcid.org/0000-0002-1952-9428
*Mario Hiroyuki Hirata* (iD) http://orcid.org/0000-0002-9521-7979

## Author contributions statement

V.G.L, G.M.F and L.M.M.D. design in silico experiments and analysis. V.F.O., M.S.N., M.H.H, and G.M.F helped with the discussion and writing. M.S.N., M.H.H, and G.M.F contributed to resources and overall study supervision. All authors prepared and reviewed the manuscript.

## References

Amadei, A., Linssen, A. B., & Berendsen, H. J. (1993). Essential dynamics of proteins. *Proteins*, *17*(4), 412–425. https://doi.org/10.1002/prot.340170408

Anderson, R. J., Weng, Z., Campbell, R. K., & Jiang, X. (2005). Main-chain conformational tendencies of amino acids. *Proteins*, *60*(4), 679–689. https://doi.org/10.1002/prot.20530

Atilgan, C., & Atilgan, A. R. (2009). Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLOS Computational Biology*, *5*(10), e1000544. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000544.

Banks, J. L., Beard, H. S., Cao, Y., Cho, A. E., Damm, W., Farid, R., Felts, A. K., Halgren, T. A., Mainz, D. T., Maple, J. R., Murphy, R., Philipp, D. M., Repasky, M. P., Zhang, L. Y., Berne, B. J., Friesner, R. A., Gallicchio, E., & Levy, R. M. (2005). Integrated modeling program, applied chemical theory (IMPACT). *Journal of Computational Chemistry*, *26*(16), 1752–1780. https://doi.org/10.1002/jcc.20292

Bas, D. C., Rogers, D. M., & Jensen, J. H. (2008). Very fast prediction and rationalization of pKa values for protein–ligand complexes. *Proteins: Structure, Function, and Bioinformatics*, *73*(3), 765–783. https://pubmed.ncbi.nlm.nih.gov/18498103/.

Berberich, A. J., & Hegele, R. A. (2019). The complex molecular genetics of familial hypercholesterolaemia. *Nature Reviews. Cardiology*, 16(1), 9–20. https://doi.org/10.1038/s41569-018-0052-6

Berge, K. E., Ose, L., & Leren, T. P. (2019). Missense mutations in the PCSK9 gene are associated with hypocholesterolemia and possibly increased response to statin therapy. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 26(5), 1094–1100. https://www.ahajournals.org/doi/10.1161/01.ATV.0000204337.81286.1c. https://doi.org/10.1038/s41569-018-0052-6

Bottomley, M. J., Cirillo, A., Orsatti, L., Ruggeri, L., Fisher, T. S., Santoro, J. C., Cummings, R. T., Cubbon, R. M., Lo Surdo, P., Calzetta, A., Noto, A., Baysarowich, J., Mattu, M., Talamo, F., De Francesco, R., Sparrow, C. P., Sitlani, A., & Carfí, A. (2009). Structural and biochemical characterization of the wild type PCSK9-EGF(AB) complex and natural familial hypercholesterolemia mutants. *The Journal of Biological Chemistry*, 284(2), 1313–1323. https://doi.org/10.1074/jbc.M808363200

Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., & Salmon, J. K. (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. In Proceedings of the 2006 ACM/IEEE Conference on Supercomputing. ACM Press. https://doi.org/10.1145/1188455.1188544

Costet, P., Krempf, M., & Cariou, B. (2008). PCSK9 and LDL cholesterol: Unravelling the target to design the bullet. *Trends in Biochemical Sciences*, 33(9), 426–434. https://doi.org/10.1016/j.tibs.2008.06.005

Cunningham, D., Danley, D. E., Geoghegan, K. F., Griffor, M. C., Hawkins, J. L., Subashi, T. A., Varghese, A. H., Ammirati, M. J., Culp, J. S., Hoth, L. R., Mansour, M. N., McGrath, K. M., Seddon, A. P., Shenolikar, S., Stutzman-Engwall, K. J., Warren, L. C., Xia, D., & Qiu, X. (2007). Structural and biophysical studies of PCSK9 and its mutants linked to familial hypercholesterolemia. *Nature Structural & Molecular Biology*, 14(5), 413–419. https://doi.org/10.1038/nsmb1235

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An *N* ·log (*N*) method for Ewald sums in large systems. *Journal of Chemical Physics*, 98, 10089–10092. https://doi.org/10.1063/1.464397

Deng, S.-J., Shen, Y., Gu, H.-M., Guo, S., Wu, S.-R., & Zhang, D.-W. (2020). The role of the C-terminal domain of PCSK9 and SEC24 isoforms in PCSK9 secretion. *Biochimica et Biophysica Acta. Molecular and Cell Biology of Lipids*, 1865(6), 158660. https://doi.org/10.1016/j.bbalip.2020.158660

Dubuc, G., Tremblay, M., Paré, G., Jacques, H., Hamelin, J., Benjannet, S., Boulet, L., Genest, J., Bernier, L., Seidah, N. G., & Davignon, J. (2010). A new method for measurement of total plasma PCSK9: Clinical applications. *Journal of Lipid Research*, 51(1), 140–149. https://doi.org/10.1194/jlr.M900273-JLR200

Dutta, A., Krieger, J., Lee, J. Y., Garcia-Nafria, J., Greger, I. H., & Bahar, I. (2015). Cooperative dynamics of intact AMPA and NMDA glutamate receptors: Similarities and subfamily-specific differences. *Structure (London, England : 1993)*, 23(9), 1692–1704. https://doi.org/10.1016/j.str.2015.07.002

Evans, D., & Beil, F. U. (2006). The E670G SNP in the PCSK9 gene is associated with polygenic hypercholesterolemia in men but not in women. *BMC Medical Genetics*,7.

Fan, D., Yancey, P. G., Qiu, S., Ding, L., Weeber, E. J., Linton, M. F., & Fazio, S. (2008). Self-association of human PCSK9 correlates with its LDLR-degrading activity. *Biochemistry*, 47(6), 1631–1639. https://doi.org/10.1021/bi7016359

Ferreira, J. P., Xhaard, C., Lamiral, Z., Borges-Canha, M., Neves, J. S., Dandine-Roulland, C., LeFloch, E., Deleuze, J. F., Bacq-Daian, D., Bozec, E., & Girerd, N. (2020). PCSK9 protein and rs562556 polymorphism are associated with arterial plaques in healthy middle-aged population: The STANISLAS cohort. *Journal of the American Heart Association*, 9, e014758.

Gai, M.-T., Adi, D., Chen, X. C., Liu, F., Xie, X., Yang, Y. N., Gao, X. M., Ma, X., Fu, Z. Y., Ma, Y. T., & Chen, B. D. (2021). Polymorphisms of rs2483205 and rs562556 in the PCSK9 gene are associated with coronary artery disease and cardiovascular risk factors. *Scientific Reports*, 11, 11450.

Hampton, E. N., Knuth, M. W., Li, J., Harris, J. L., Lesley, S. A., & Spraggon, G. (2007). The self-inhibited structure of full-length PCSK9 at 1.9 Å reveals structural homology with resistin within the C-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37), 14604–14609. https://doi.org/10.1073/pnas.0703402104

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79, 926–935. https://doi.org/10.1063/1.445869

Kosenko, T., Golder, M., Leblond, G., Weng, W., & Lagace, T. A. (2013). Low density lipoprotein binds to proprotein convertase subtilisin/kexin Type-9 (PCSK9) in human plasma and inhibits PCSK9-mediated low density lipoprotein receptor degradation. *The Journal of Biological Chemistry*, 288(12), 8279–8288. https://doi.org/10.1074/jbc.M112.421370

Kwon, H. J., Lagace, T. A., McNutt, M. C., Horton, J. D., & Deisenhofer, J. (2008). Molecular basis for LDL receptor recognition by PCSK9. *Proceedings of the National Academy of Sciences of the United States of America*, 105(6), 1820–1825. https://doi.org/10.1073/pnas.0712064105

Lee, C. J., Lee, Y., Park, S., Kang, S. M., Jang, Y., Lee, J. H., & Lee, S. H. (2017). Rare and common variants of APOB and PCSK9 in Korean patients with extremely low low-density lipoprotein-cholesterol levels. *Plos One*, 12, e0186446.

Li, Y., Wang, H., Yang, X. X., Geng, H. Y., Gong, G., & Lu, X. Z. (2020). PCSK9 gene E670G polymorphism and coronary artery disease: An updated meta-analysis of 5,484 subjects. *Frontiers in Cardiovascular Medicine*, 7, 582865.

Mohamed, S. H., Hassaan, M. M. M., Ibrahim, B. A., & Sabbah, N. A. (2021). PCSK9 E670G (rs505151) variant and coronary artery disease risk among diabetics. *Genetic Testing and Molecular Biomarkers*, 25(9), 615–623. https://doi.org/10.1089/gtmb.2021.0010

Mousavi, S. A., Berge, K. E., Berg, T., & Leren, T. P. (2011). Affinity and kinetics of proprotein convertase subtilisin/kexin type 9 binding to low-density lipoprotein receptors on HepG2 cells. *The FEBS Journal*, 278(16), 2938–2950. https://doi.org/10.1111/j.1742-4658.2011.08219.x

Naslavsky, M. S., Sclar, M. O., Yamamoto, G. L., Wang, J. Y. T., Zverinova, S., Karp, T., Nunes, K., Ceroni, J. R. M., de Carvalho, D. L., da Silva Simões, C. E., & Bozoklian, D. (2022). Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. *Nature Communications*, 13, 1004.

Naslavsky, M. S., Yamamoto, G. L., de Almeida, T. F., Ezquina, S. A. M., Sunaga, D. Y., Pho, N., Bozoklian, D., Sandberg, T. O. M., Brito, L. A., Lazar, M., Bernardo, D. V., Amaro, E., Duarte, Y. A. O., Lebrão, M. L., Passos-Bueno, M. R., & Zatz, M. (2017). Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Human Mutation*, 38(7), 751–763. https://doi.org/10.1002/humu.23220

Pandit, S., Wisniewski, D., Santoro, J. C., Ha, S., Ramakrishnan, V., Cubbon, R. M., Cummings, R. T., Wright, S. D., Sparrow, C. P., Sitlani, A., & Fisher, T. S. (2008). Functional analysis of sites within PCSK9 responsible for hypercholesterolemias. *Journal of Lipid Research*, 49(6), 1333–1343.

Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature*, 538(7624), 161–164. https://doi.org/10.1038/538161a

Qiu, C., Zeng, P., Li, X., Zhang, Z., Pan, B., Peng, Z. Y., Li, Y., Ma, Y., Leng, Y., & Chen, R. (2017). What is the impact of PCSK9 rs505151 and rs11591147 polymorphisms on serum lipids level and cardiovascular risk: A meta-analysis. *Lipids in Health and Disease*, 16, 111.

Rocha, C. S., Secolin, R., Rodrigues, M. R., Carvalho, B. S., & Lopes-Cendes, I. (2020). The Brazilian Initiative on Precision Medicine (BIPMed): Fostering genomic data-sharing of underrepresented populations. *NPJ Genomic Medicine*, 5, 1–7.

Rodrigues de Moura, R., Coelho, A. V. C., de Queiroz Balbino, V., Crovella, S., & Brandão, L. A. C. (2015). Meta-analysis of Brazilian genetic admixture and comparison with other Latin America countries. *American Journal of Human Biology : The Official Journal of the Human Biology Council*, 27(5), 674–680. https://doi.org/10.1002/ajhb.22714

Saavedra, Y. G. L., Day, R., & Seidah, N. G. (2012). The M2 module of the Cys-His-rich domain (CHRD) of PCSK9 protein is needed for the extracellular low-density lipoprotein receptor (LDLR) degradation pathway. *The Journal of Biological Chemistry*, 287(52), 43492–43501. https://doi.org/10.1074/jbc.M112.394023

Sánchez-Hernández, R. M., Di Taranto, M. D., Benito-Vicente, A., Uribe, K. B., Lamiquiz-Moneo, I., Larrea-Sebal, A., Jebari, S., Galicia-Garcia, U., Nóvoa, F. J., Boronat, M., Wägner, A. M., Civeira, F., Martín, C., & Fortunato, G. (2019). The Arg499His gain-of-function mutation in the C-terminal domain of PCSK9. *Atherosclerosis*, 289, 162–172. https://doi.org/10.1016/j.atherosclerosis.2019.08.020

Sarkar, S. K., Foo, A. C., Matyas, A., Asikhia, I., Kosenko, T., Goto, N. K., Vergara-Jaque, A., & Lagace, T. A. (2020). A transient amphipathic helix in the prodomain of PCSK9 facilitates binding to low-density lipoprotein particles. *Journal of Biological Chemistry*, 295, 2285–2298.

Schiele, F., Park, J., Redemann, N., Luippold, G., & Nar, H. (2014). An antibody against the C-terminal domain of PCSK9 lowers LDL cholesterol levels in vivo. *Journal of Molecular Biology*, 426(4), 843–852. https://doi.org/10.1016/j.jmb.2013.11.011

Seidah, N. G., Awan, Z., Chrétien, M., & Mbikay, M. (2014). PCSK9: A key modulator of cardiovascular health. *Circulation Research*, 114(6), 1022–1036. https://doi.org/10.1161/CIRCRESAHA.114.301621

Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins*, 17(4), 355–362. https://doi.org/10.1002/prot.340170404

Sundararaman, S. S., Döring, Y., & van der Vorst, E. P. (2021). PCSK9: A multi-faceted protein that is involved in cardiovascular biology. *Biomedicines*, 9, 793. https://doi.org/10.3390/biomedicines9070793

Vicini, P., Fields, O., Lai, E., Litwack, E. D., Martin, A.-M., Morgan, T. M., Pacanowski, M. A., Papaluca, M., Perez, O. D., Ringel, M. S., Robson, M., Sakul, H., Vockley, J., Zaks, T., Dolsten, M., & Søgaard, M. (2016). Precision medicine in the age of big data: The present and future role of large-scale unbiased sequencing in drug discovery and development. *Clinical Pharmacology and Therapeutics*, 99(2), 198–207. https://doi.org/10.1002/cpt.293

Wan, S., Kumar, D., Ilyin, V., Al Homsi, U., Sher, G., Knuth, A., & Coveney, P. V. (2021). The effect of protein mutations on drug binding suggests ensuing personalised drug selection. *Scientific Reports*, 11(1), 13452. https://doi.org/10.1038/s41598-021-92785-w

Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma*, 54(1), 5–6.

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(Web Server issue), W407–W410. https://doi.org/10.1093/nar/gkm290

Zhang, S., Krieger, J. M., Zhang, Y., Kaya, C., Kaynak, B., Mikulska-Ruminska, K., Doruker, P., Li, H., & Bahar, I. (2021). ProDy 2.0: Increased scale and scope after 10 years of protein dynamics modelling with Python. *Bioinformatics (Oxford, England)*, 37(20), 3657–3659. https://doi.org/10.1093/bioinformatics/btab187