Integrating Computational Approaches to Predict the Effect of Genetic Variants on Protein Stability in Retinal Degenerative Disease

Michelle Grunin, Ellen Palmer, Sarah de Jong, Bowen Jin, David Rinker, Christopher Moth, John A. Capra, Jonathan L. Haines, William S. Bush, and Anneke I. den Hollander

Abstract

Protein function can be impacted by changes in protein structure stability, but determining which change has impact is complex. Stability can be affected by a large change in the tertiary (3D) structure of the protein or due to free-energy changes caused by single amino acid substitutions. Changes in the DNA sequence can have minor or major impact on protein stability, which can lead to disease. Inherited retinal degenerations are generally caused by single mutations which are mostly located in protein-coding regions, while age-

Authors Michelle Grunin, Ellen Palmer, Sarah de Jong, Jonathan L. Haines, William S. Bush, and Anneke I. den Hollander have equally contributed to this chapter.

M. Grunin (✉) · E. Palmer · B. Jin · J. L. Haines · W. S. Bush
Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA
e-mail: mag235@case.edu; elp76@case.edu; bxj139@case.edu; jlh213@case.edu; wsb36@case.edu

S. de Jong (✉)
Department of Ophthalmology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen, the Netherlands
e-mail: sarah.dejong@radboudumc.nl

D. Rinker
Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA
e-mail: david.rinker@vanderbilt.edu

C. Moth
Center for Structural Biology, Vanderbilt University, Nashville, TN, USA
e-mail: chris.moth@vanderbilt.edu

J. A. Capra
Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

Center for Structural Biology, Vanderbilt University, Nashville, TN, USA

Vanderbilt Genetics Institute, Vanderbilt University School of Medicine, Nashville, TN, USA

Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN, USA
e-mail: tony.capra@vanderbilt.edu

A. I. den Hollander
Department of Ophthalmology, Donders Institute for Brain, Cognition, and Behavior, Radboud University Medical Center, Nijmegen, the Netherlands

Department of Human Genetics, Radboud University Medical Center, Nijmegen, the Netherlands
e-mail: Anneke.denHollander@radboudumc.nl

J. D. Ash et al. (eds.), *Retinal Degenerative Diseases XIX*, Advances in Experimental Medicine and Biology 1415, https://doi.org/10.1007/978-3-031-27681-1_24

related macular degeneration (AMD) is a complex disorder that can be influenced by some genetic variants impacting proteins involved in the disease, although not all AMD risk variants lead to amino acid changes. Here, we review ways that proteins may be affected, the identification and understanding of these changes, and how to identify causal changes that can be targeted to develop treatments to alleviate retinal degenerative disease.

Keywords

Proteins · Age-related macular degeneration · Mutations · Variants · Free-energy

1 Part I. Protein Stability: Role and Importance

Multiple efforts are underway to gather information on clinically meaningful mutations in protein coding genes in databases such as the Human Gene Mutation Database (HGMD) [1, 2] or ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) [3]. In 2003, HGMD contained more than 1000 mutations that were directly causative of disease; the number of reported mutations currently in HGMD has grown to 323,661 as of Oct. 2021. ClinVar currently contains 1,159,307 unique disease-contributing variations (Oct. 16, 2021). Many of the mutations found in these databases, but not all, can affect protein stability [4–6]. The deleterious impact of mutations that clearly affect a protein's sequence, fold, and function is sometimes obvious, such as frameshift mutations that alter much of the protein sequence. More often, missense mutations are identified, such as single-point mutations that cause some familial forms of Alzheimer's or Parkinson's disease. Missense mutations are generally more difficult to classify and understand, and newly discovered variants may be given an ambiguous classification (referred to as "variants of unknown significance"; VUS), as these changes may alter protein structure at different, hard to measure levels (e.g., tertiary structure). An excellent example is retinitis pigmentosa, as 60% of disease-causing mutations are still unknown, while this may in part be due to the ambiguous classification of genetic variants [7]. As sequencing-based genetic studies increase the number of identified protein-altering mutations, more work in this area will be critical to understand their role in complex diseases[8] and to discriminate between harmless and more damaging VUS.

Missense variants comprise over 60% of all known monogenic disease mutations [9], but they disrupt protein structure and function in various, sometimes unclear, ways [10]. Changes to alpha helices or beta sheets are the most clearly understood, as they generally impact protein stability by changing hydrophobicity that affects protein folding, like mutations in TGFb or Pim1 in cancer [9, 11, 12]. Amino acid changes in other locations may affect the stability of a protein or protein complex [6] through changing the tertiary structure of the protein or through changing the free energy needed to fold into a functional form. Depending on the context of the single amino acid change, protein folding (i.e., free-energy change) can be impacted enough to prevent the formation of structural motifs needed for critical function [6]. However, directly measuring free-energy changes resulting from an amino acid substitution is a difficult experimental task because current approaches typically are expensive, are time consuming, are performed on a single mutation at a time, and are hard to scale. In lieu of directly measuring experimental changes in free-energy, computational methods are often employed [13]. These methods typically compare wild-type and mutant protein folding using a solved template protein structure from the protein data bank (PDB, https://www.rcsb.org/) and then changing those structures to model the mutant protein using programs like PoPMusic [14] or Phyre [15]. The free energy before and after folding is calculated for both the wild-type and mutant proteins to produce two respective free-energy change measures (ΔG). These measures are then compared to estimate a change in ΔG due to the amino acid substitution ($\Delta\Delta G$). Both experimentally solved and high-resolution computational models of protein structures like those found in the PDB can be used for accurate

folding estimates [16, 17]. PDB contains 183,386 structures in 2021, which were either determined experimentally or by homology modeling based on similar protein structures [18]. As of 2020, 14,028 solved and proven experimentally structures from the 183,386 were included in the PDB. The sequence based or ab initio approaches for identifying impactful variants are faster but depend on machine learning training sets and do not focus on the ultimate impact on the full post-secondary protein structure, utilizing protein modeling, while the structure-based approaches use multiple scoring methods as well as intensive calculations, like the free-energy perturbation (FEP) and thermodynamic integration (TI) methods [19, 20].

Computationally, protein homology models can be constructed using methods like the Swiss-Model in the ExPASy webserver [21]. Proteins can be graphed using Mol [22]or other programs to visualize exactly where the structure changes and where the variant impacts the protein structure. Variants that impact the free-energy change in binding sites or core regions may be the most damaging, even if the change is not obviously damaging to the tertiary structure of the protein. Mutations that destabilize proteins are described for von Willebrand diseases [23], prion [24], and retinal degenerative diseases that impact rhodopsin [25]. Depending on protein function, some residue substitutions cause disease by increasing protein stability, such as the CLIC2 protein in some mental disorders [26].

2 Part II. Common Computational Methods for Determining Variant Impact

Machine learning algorithms can assess the impact of missense variants using a variety of information including protein sequence and in some cases, structure and folding. Programs like SIFT (http://sift-dna.org) [27], MutPred (http://mutpred.putdb.org) [28], or Polyphen2 (http://genetics.bwh.harvard.ed/pphw) [29] are trained using a set of known deleterious mutations for often severe Mendelian diseases. Some approaches and programs either calculate free-energy changes and take the free-energy changes into account when evaluating protein stability, like I-Mutant2.0 (http://gpcr2.biocomp.unibo.it/%7Eemidio/I-Mutatnt/I-Mutant.htm) [30], SDM (http://www-cryst.bioc.cam.ac.uk/~sdm/sdm.php) [31], CUPSAT (http://cupsat.tu-bs.de/) [32], FoldX (http://fold-x.embl-heidelberg.de) [33], ROSETTA (https://www.rosettacommons.org/) [34], or AUTO-MUTE (http://proteins.gmu.edu/automute) [35]. However, only few programs take into account both information from tertiary structures and the impact of free-energy change on the protein structure. Only PolyPhen, I-Mutant2.0, and HOPE (https://www3.cmbi.umcn.nl/hope/about/) [36] consider both sequence and protein structure changes but include these among many other features in a machine learning prediction based on highly penetrant missense variants[6]. A comparison between the different methods used to estimate the effect of different mutations on protein stability has been described in work of de Groot [37], Thiltgen and Goldstein [38], and Kroncke [39].

3 Part III. New Methodologies: PathProx and POKEMON

PathProx [16, 17] employs an alternative approach: it specifically focuses on the spatial context and stability of the full three-dimensional protein. PathProx performs two types of analyses: (1) PathProx evaluates the spatial proximity of input variants to known pathogenic variants (from resources like ClinVar) and to presumed neutral variants within the protein; (2) PathProx calculates differences in the free-energy and stability of the protein utilizing ROSETTA and examines those variants in relationship to other known pathogenic and presumably neutral variants that may be present in the protein structure. The spatial proximity portion of the algorithm uses Euclidean distance to determine if the input variants are nearer to pathogenic risk variants (or any variant of interest), as compared to "normal" or neutral variants, such as those found in data-

bases like GnomAD, which draws on over 141,456 samples from the controls of dozens of disease studies (https://gnomad.broadinstitute.org/) [40]. Variants that are found closer to known pathogenic variants and that are enriched in cases (when case-control data is used) could indicate that a particular protein region plays a significant role in disease pathogenesis, whereas variants that are found near neutral variants and are more often found in controls or are evenly distributed between cases and controls are likely neutral or perhaps "protective" variants. The ROSETTA-based portion of the method uses free-energy calculations to identify variants that are predicted to impact the structure of a protein associated with a given disease. Together, these two approaches within PathProx create a candidate variant list for further case enrichment, gene-based testing, and functional testing at the bench. Paired together, these approaches allow for the identification of two different patterns of protein-altering relationships.

POKEMON [41] adapts gene-based testing methods employing kernel functions (such as SKAT) to include information about missense variant proximity within the 3D structure of a protein. Using this kernel, a statistical test can be conducted to determine if the spatial proximity of variants within the protein is related to case status. For example, a pocket of missense variants found more frequently in cases within one particular area of a protein may point to a segment of the protein that plays a role in disease. That area may be important as a binding site, or impact protein stability at the weakest point, contributing to disease pathogenesis. Results from POKEMON and the spatial-clustering analyses of PathProx have been found to be concordant in a study of Alzheimer's disease and provide orthogonal methods of identifying disease-associated protein regions (unpublished data, 2021). We recently utilized these two methods to identify variants that can have a functional impact on protein stability and expression of complement factors in AMD (Grunin, Palmer, de Jong et al., unpublished).

4 Part IV. Retinal Diseases, AMD, and Protein Stability

Multiple retinal degenerative diseases are caused by mutations that affect protein stability [42]. One of the best known are rhodopsin mutations, the most common cause of autosomal dominant retinitis pigmentosa [43].The first identified mutations inhibit protein stability that either reduces its export out of the endoplasmic reticulum (ER) (Class II mutations) [44] or increases accumulation inside the cell (Class III mutations) [45]. The majority of these mutations lead to folding defects of the protein [45]. Class II mutations also impact tertiary folding stability [46]. These mutations are common among the G protein-coupled receptors (GPCRs), but how their thermodynamics are impacted by folding stability is still not fully understood. Therefore, computational methods to understand the effect of these mutations is the avenue that has been most explored [43].

Investigation of VUS has been a major focus in genetic research, and the combination of both computational and functional biology approaches has been useful in identifying the impact of VUSs. If computational approaches can differentiate variants of significance from among the VUSs, specifically those that have functional effect, this would significantly improve variant interpretation in diagnostic testing. For rhodopsin, variant interpretation has been performed using a combination of computational and experimental approaches: gain of function mutations have systemically been analyzed using a computational approach, in addition to a full-scale experimental screen to evaluate rhodopsin expression in cells [47]. Two-stage approaches using computational methods before moving to functional testing have been applied, for example, in Best disease, to determine destabilizing mutations using I-mutant and then performing functional testing on all variants [48]. Many tests have shown that reliable first-stage computational testing of missense mutations can be done through I-Mutant, Dmutant, and FoldX [49].

However, a two-stage approach can be used to first predict the effect of a variant computationally and then focus experimental work on variants that are predicted to affect protein stability or function, which would narrow the testing pool, time, and costs [50]. In AMD, one of the most studied proteins is complement factor H (CFH). Rare protein-coding AMD risk variants have been mapped on the protein with corresponding issues of protein unfolding or incorrect folding, with 70% of mutations showing a destabilizing effect. Recently, 105 variants in CFH were classified according to their pathogenicity and effect on function utilizing functional assays [51].

Several other proteins of the complement pathway are associated with AMD including C3, C9, CFB, and CFI [52]. Rare protein-coding variants in these complement genes have been associated with differences in protein concentration in patients and controls [53–57]. However, in AMD, a combination of common variants with modest effect and rare variants with large effect collectively contribute to the disease, and how these variants contribute to AMD and whether they have synergistic effects are currently unknown. Analysis of variants in RPE65 and rhodopsin have shown similar results in destabilization of the protein, utilizing an unfolding mutation screen (UMS) [50]. However, these results were not followed up by actual functional testing on those mutations, because the mutations were already known, and thus the method did not show predictive power. We recently applied computational methodology utilizing PathProx and POKEMON to patients with AMD and identified variants that impact proteins in the complement system. These variants were located in unique spatial locations in the protein and lead to distinct free-energy changes. We determined through computational mining which variants cause an in vitro change in complement protein expression. Therefore, the application of computational filtering allows us to identify variants that have a foreseeable functional impact (Grunin, Palmer, deJong, unpublished data, 2021)

In conclusion, utilizing novel methods like PathProx and POKEMON will enable understanding of which variants are likely to affect protein stability and may provide new avenues for identifying treatment-amenable variants without the need to bench test every single mutation. In addition, free-energy changes can be utilized to predict the consequences that these variants might have on the protein of interest. Reliable predictive testing allows the identification of variants of interest in disease more rapidly and allows for targeted functional testing.

References

1. Stenson PD, Ball E V., Mort M, et al. Human Gene Mutation Database (HGMD®): 2003 updateHum Mutat;21. https://doi.org/10.1002/humu.10212. Epub ahead of print 2003.
2. Stenson PD, Mort M, Ball E V., et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet;136. https://doi.org/10.1007/s00439-017-1779-6. Epub ahead of print 2017.
3. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res;46. https://doi.org/10.1093/nar/gkx1153. Epub ahead of print 2018.
4. Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat;17. https://doi.org/10.1002/humu.22. Epub ahead of print 2001.
5. Casadio R, Vassura M, Tiwari S, et al. Correlating disease-related mutations to their effect on protein stability: a large-scale analysis of the human proteome. Hum Mutat;32. https://doi.org/10.1002/humu.21555. Epub ahead of print 2011.
6. Stefl S, Nishi H, Petukh M, et al. Molecular mechanisms of disease-causing missense mutations. J Mol Biol;425. https://doi.org/10.1016/j.jmb.2013.07.014. Epub ahead of print 2013.
7. González-Del Pozo M, Fernández-Suárez E, Martín-Sánchez M, et al. Unmasking Retinitis Pigmentosa complex cases by a whole genome sequencing algorithm based on open-access tools: hidden recessive inheritance and potential oligogenic variants. J Transl Med;18. https://doi.org/10.1186/s12967-020-02258-3. Epub ahead of print 2020.
8. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with support vector machines and evolutionary information. Bioinformatics;22. https://doi.org/10.1093/bioinformatics/btl423. Epub ahead of print 2006.
9. Yue P, Li Z, Moult J. Loss of protein structure stability as a major causative factor in monogenic disease. J Mol Biol;353. https://doi.org/10.1016/j.jmb.2005.08.020. Epub ahead of print 2005.

10. Calamini B, Lo DC, Kaltenbach LS. Experimental models for identifying modifiers of polyglutamine-induced aggregation and neurodegeneration. Neurotherapeutics;10. https://doi.org/10.1007/s13311-013-0195-4. Epub ahead of print 2013.
11. Kucukkal TG, Alexov E. Structural, dynamical, and energetical consequences of RETT syndrome mutation R133c in MeCP2. Comput Math Methods Med. 2015. https://doi.org/10.1155/2015/746157. Epub ahead of print 2015.
12. Petukh M, Dai L, Alexov E. SAAMBE: webserver to predict the charge of binding free energy caused by amino acids mutations. Int J Mol Sci;17. https://doi.org/10.3390/ijms17040547. Epub ahead of print 2016.
13. Magliery TJ. Protein stability: computation, sequence statistics, and new experimental methods. Curr Opin Struct Biol;33. https://doi.org/10.1016/j.sbi.2015.09.002. Epub ahead of print 2015.
14. Dehouck Y, Kwasigroch JM, Gilis D, et al. PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. BMC Bioinf;12. https://doi.org/10.1186/1471-2105-12-151. Epub ahead of print 2011.
15. Kelley LA, Mezulis S, Yates CM, et al. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10:845–58.
16. Sivley RM, Sheehan JH, Kropski JA, et al. Three-dimensional spatial analysis of missense variants in RTEL1 identifies pathogenic variants in patients with Familial Interstitial Pneumonia. BMC Bioinf;19. https://doi.org/10.1186/s12859-018-2010-z. Epub ahead of print 2018.
17. Sivley RM, Dou X, Meiler J, et al. Comprehensive analysis of constraint on the spatial distribution of missense variants in human protein structures. Am J Hum Genet;102. https://doi.org/10.1016/j.ajhg.2018.01.017. Epub ahead of print 2018.
18. Berman HM, Battistuz T, Bhat TN, et al. The protein data bank Acta Crystallogr Sect D Biol Crystallogr;58. https://doi.org/10.1107/S0907444902003451. Epub ahead of print 2002.
19. Petukh M, Kucukkal TG, Alexov E. On human disease-causing amino acid variants: Statistical study of sequence and structural patterns. Hum Mutat;36. https://doi.org/10.1002/humu.22770. Epub ahead of print 2015.
20. Michel J, Foloppe N, Essex JW. Rigorous free energy calculations in structure-based drug design Mol Inf;29. https://doi.org/10.1002/minf.201000051. Epub ahead of print 2010.
21. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res;46. https://doi.org/10.1093/nar/gky427. Epub ahead of print 2018.
22. Sehnal D, Bittrich S, Deshpande M, et al. Mol∗Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res;49. https://doi.org/10.1093/nar/gkab314. Epub ahead of print 2021.
23. Xu AJ, Springer TA. Mechanisms by which von willebrand disease mutations destabilize the A2 domain. J Biol Chem;288. https://doi.org/10.1074/jbc.M112.422618. Epub ahead of print 2013.
24. Jetha NN, Semenchenko V, Wishart DS, et al. Nanopore analysis of wild-type and mutant prion protein (PrPC): single molecule discrimination and PrPC kinetics. PLoS One;8. https://doi.org/10.1371/journal.pone.0054982. Epub ahead of print 2013.
25. Rakoczy EP, Kiel C, McKeone R, et al. Analysis of disease-linked rhodopsin mutations based on structure, function, and protein stability calculations. J Mol Biol;405. https://doi.org/10.1016/j.jmb.2010.11.003. Epub ahead of print 2011.
26. Witham S, Takano K, Schwartz C, et al. A missense mutation in CLIC2 associated with intellectual disability is predicted by in silico modeling to affect protein stability and dynamics. Proteins Struct Funct Bioinf;79. https://doi.org/10.1002/prot.23065. Epub ahead of print 2011.
27. Sim NL, Kumar P, Hu J, et al. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res;40 https://doi.org/10.1093/nar/gks539. . Epub ahead of print 2012.
28. Pejaver V, Urresti J, Lugo-Martinez J, et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat Commun;11. https://doi.org/10.1038/s41467-020-19669-x. Epub ahead of print 2020.
29. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods;7. https://doi.org/10.1038/nmeth0410-248. Epub ahead of print 2010.
30. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res;33. https://doi.org/10.1093/nar/gki375. Epub ahead of print 2005.
31. Pandurangan AP, Ochoa-Montaño B, Ascher DB, et al. SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res;45. https://doi.org/10.1093/nar/gkx439. Epub ahead of print 2017.
32. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res;34. https://doi.org/10.1093/nar/gkl190. Epub ahead of print 2006.
33. Schymkowitz J, Borg J, Stricher F, et al. The FoldX web server: an online force field. Nucleic Acids Res;33. https://doi.org/10.1093/nar/gki387. Epub ahead of print 2005.
34. Alford RF, Leaver-Fay A, Jeliazkov JR, et al. The Rosetta all-atom energy function for macromolecular modeling and design. J Chem Theory Comput;13. https://doi.org/10.1021/acs.jctc.7b00125. Epub ahead of print 2017.
35. Masso M, Vaisman II. AUTO-MUTE 2.0: a portable framework with enhanced capabilities for predicting protein functional consequences upon mutation. Adv Bioinf. 2014. https://doi.org/10.1155/2014/278385. Epub ahead of print 2014.

36. Venselaar H, te Beek TAH, Kuipers RKP, et al. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinf;11. https://doi.org/10.1186/1471-2105-11-548. Epub ahead of print 2010.
37. De Groot BL, Van Aalten DMF, Scheek RM, et al. Prediction of protein conformational freedom from distance constraints. Proteins Struct Funct Genet;29. https://doi.org/10.1002/(SICI)1097-0134(199710)29:2<240::AID-PROT11>3.0.CO;2-O. Epub ahead of print 1997.
38. Thiltgen G, Goldstein RA. Assessing predictors of changes in protein stability upon mutation using self-consistency. PLoS One;7. https://doi.org/10.1371/journal.pone.0046084. Epub ahead of print 2012.
39. Kroncke BM, Duran AM, Mendenhall JL, et al. Documentation of an imperative to improve methods for predicting membrane protein stability. Biochemistry;55. https://doi.org/10.1021/acs.biochem.6b00537. Epub ahead of print 2016.
40. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature;581. https://doi.org/10.1038/s41586-020-2308-7. Epub ahead of print 2020.
41. Jin B, Capra JA, Benchek P, et al. An association test of the spatial distribution of rare missense variants within protein structures improves statistical power of sequencing studies. bioRxiv. https://doi.org/10.1101/2021.08.09.455695
42. Tzekov R, Stein L, Kausha S. Protein misfolding and retinal degeneration Cold Spring Harb Perspect Biol;3. https://doi.org/10.1101/cshperspect.a007492. Epub ahead of print 2011.
43. Marinko JT, Huang H, Penn WD, et al. Folding and misfolding of human membrane proteins in health and disease: from single molecules to cellular proteostasis. Chem Rev;119. https://doi.org/10.1021/acs.chemrev.8b00532. Epub ahead of print 2019.
44. Sung CH, Davenport CM, Hennessey JC, et al. Rhodopsin mutations in autosomal dominant retinitis pigmentosa. Proc Natl Acad Sci U S A;88. https://doi.org/10.1073/pnas.88.15.6481. Epub ahead of print 1991.
45. Kaushal S, Khorana HG. Structure and function in rhodopsin. 7. Point mutations associated with autosomal dominant retinitis pigmentosa. Biochemistry;33. https://doi.org/10.1021/bi00186a011. Epub ahead of print 1994.
46. Rader AJ, Anderson G, Isin B, et al. Identification of core amino acids stabilizing rhodopsin. Proc Natl Acad Sci U S A;101. https://doi.org/10.1073/pnas.0401429101. Epub ahead of print 2004.
47. Wan A, Place E, Pierce EA, et al. Characterizing variants of unknown significance in rhodopsin: A functional genomics approach. Hum Mutat;40. https://doi.org/10.1002/humu.23762. Epub ahead of print 2019.
48. Milenkovic A, Milenkovic VM, Wetzel CH, et al. BEST1 protein stability and degradation pathways differ between autosomal dominant Best disease and autosomal recessive bestrophinopathy accounting for the distinct retinal phenotypes. Hum Mol Genet;27. https://doi.org/10.1093/hmg/ddy070. Epub ahead of print 2018.
49. Khan S, Vihinen M. Performance of protein stability predictors. Hum Mutat;31. https://doi.org/10.1002/humu.21242. Epub ahead of print 2010.
50. McCafferty CL, Sergeev Y V. In silico mapping of protein unfolding mutations for inherited disease. Sci Rep;6. https://doi.org/10.1038/srep37298. Epub ahead of print 2016.
51. Martin Merinero H, Zhang Y, Arjona E, et al. Functional characterization of 105 Factor H variants associated with atypical HUS: lessons for variant classification. Blood. https://doi.org/10.1182/blood.2021012037. Epub ahead of print 2021.
52. Fritsche LG, Igl W, Bailey JNC, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat Genet;48. https://doi.org/10.1038/ng.3448. Epub ahead of print 2016.
53. Hallam TM, Marchbank KJ, Harris CL, et al. Rare genetic variants in complement factor I lead to low FI plasma levels resulting in increased risk of age-related macular degeneration. Investig Ophthalmol Vis Sci;61. https://doi.org/10.1167/IOVS.61.6.18. Epub ahead of print 2020.
54. Kremlitzka M, Geerlings MJ, De Jong S, et al. Functional analyses of rare genetic variants in complement component C9 identified in patients with age-related macular degeneration. Hum Mol Genet;27. https://doi.org/10.1093/hmg/ddy178. Epub ahead of print 2018.
55. Geerlings MJ, Kersten E, Groenewoud JMM, et al. Geographic distribution of rare variants associated with age- related macular degeneration. Mol Vis. 2018;9:75–82.
56. de Jong S, Gagliardi G, Garanto A, et al. Implications of genetic variation in the complement system in age-related macular degeneration. Prog Retin Eye Res;84. https://doi.org/10.1016/j.preteyeres.2021.100952. Epub ahead of print 2021.
57. de Jong S, Volokhina EB, de Breuk A, et al. Effect of rare coding variants in the CFI gene on factor I expression levels. Hum Mol Genet;29. https://doi.org/10.1093/hmg/ddaa114. Epub ahead of print 2020.