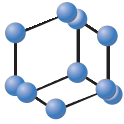


REVIEW ARTICLE


**BENTHAM
SCIENCE**

Nonsynonymous Synonymous Variants Demand for a Paradigm Shift in Genetics


 Mauno Vihinen^{1,*}
¹Department of Experimental Medical Science, Lund University, Lund, BMC B13, Sweden

ARTICLE HISTORY

Received: December 09, 2022

Revised: February 20, 2023

Accepted: March 01, 2023

DOI:
10.2174/1389202924666230417101020

CrossMark

Abstract: Synonymous (also known as silent) variations are by definition not considered to change the coded protein. Still many variations in this category affect either protein abundance or properties. As this situation is confusing, we have recently introduced systematics for synonymous variations and those that may on the surface look like synonymous, but these may affect the coded protein in various ways. A new category, nonsense variation, was introduced to describe variants that do not introduce a stop codon into the variation site, but which lead to different types of changes in the coded protein. Many of these variations lead to mRNA degradation and missing protein. Here, consequences of the systematics are discussed from the perspectives of variation annotation and interpretation, evolutionary calculations, nonsynonymous-to-synonymous substitution rates, phylogenetics and other evolutionary inferences that are based on the principle of (nearly) neutral synonymous variations. It may be necessary to reassess published results. Further, databases for synonymous variations and prediction methods for such variations should consider nonsense variations. Thus, there is a need to evaluate and reflect principles of numerous aspects in genetics, ranging from variation naming and classification to evolutionary calculations.

Keywords: Synonymous variation, silent variation, nonsense variation, phylogenetics, distribution of fitness effects, nonsynonymous-to-synonymous substitution ratio.

1. INTRODUCTION

Nucleotide substitutions in RNA have traditionally been grouped into three categories: synonymous, missense, and nonsense variations; however, there is a need for a fourth category. Synonymous (also called silent) variations are defined as those that do not change the coded amino acid, while missense variants have a nucleotide difference that causes an amino acid substitution in the coded protein. Nonsense variant introduces a premature stop codon, which when early leads usually to mRNA degradation by quality control mechanisms, typically by nonsense-mediated decay (NMD) [1, 2]. It is quite common and wrong to use terms like nonsense and missense to describe changes in DNA or protein level [3, 4].

Many variants annotated as synonymous are not synonymous at all [5-9]. The language for synonymous variants is often confusing and misleading. Therefore, we have recently presented systematic for synonymous variations and introduced a new category of nonsense variation [7]. Nonsense variation is a substitution in the mRNA coding region that affects gene expression, protein, or protein production without introducing a stop codon in the variation site. These variants are not synonymous or silent and indeed have an effect

on the coded protein. The definition has been implemented in Variation Ontology (VariO), which is used for the systematic description of effects, types, consequences, and mechanisms of biological variations [10]. When focusing on the genetic code, a variant may seem synonymous, but it may still affect the protein; thus, calling such variants as synonymous is incorrect.

In the following paper, notation “synonymous” variation is used to indicate cases where nonsense variants have not been separated from synonymous variants.

In this article, the consequences of including nonsense variants in various types of studies are discussed. This exerts an important effect on variation interpretation and genetic disease diagnosis. These variants are typically misannotated [11]. In addition, nonsense variants have to be included in evolutionary inference and in phylogenetic and natural selection predictions. Since synonymous variants are typically considered as neutral or nearly neutral, it is necessary to re-evaluate and sometimes to re-analyze studies based on this assumption. Some authors claim synonymous variants to have a small effect [12], while there is a lot of compelling evidence that many synonymous variants affect the coded protein and have functional and other effects [8, 13, 14]. Databases for “synonymous” variants and predictors for such variations mix different types of variations [15-18]. Thus, a paradigm shift is needed in many genetic studies and approaches.

*Address correspondence to this author at the Department of Experimental Medical Science, Lund University, Lund, BMC B13, Sweden; E-mail: mauno.vihinen@med.lu.se

2. UNSENSE AND SYNONYMOUS VARIATIONS

Systematics has been presented for synonymous and nonsense variants [7]. These variants can have effects on DNA, RNA, and/or protein level (Fig. 1). On the DNA level, synonymous variants can affect transcription factor binding and consequently gene expression, without altering the protein sequence.

On the mRNA level, the synonymous and synonymous-like variants can be divided into three major categories (Fig. 1). True synonymous variations are the category that this group of variations is traditionally considered to describe. Although these variants do not affect coded protein sequence, there are variants that have effects on protein regulation, post-translational modifications, protein structure or activity (Fig. 1). The second group is classified as synonymous, but they affect RNA structure and stability. These variants affect the folding and abundance of the coded protein, as shown in the studies presented earlier [5-7, 19].

The third category includes nonsense variants. Three types of nonsense variants are known. They affect either splicing, splicing regulation, or miRNA binding due to exon variations [7]. Nonsense variants are important and apparent-

ly cover a substantial portion of variations, also among those causing diseases. Based on the universal codon table, 23.8% of possible substitutions are for the same amino acid-coding codons [7]. The situation is somewhat different in genes due to different vulnerability of nucleotides, nucleotide composition, gene C+G content, *etc.* [20]. The consequences of variations, whether synonymous or nonsense, depend on many factors, and the context of the variants also plays an important role [7].

Currently, three mechanisms behind nonsense variants are known [7], but there may be more. Splicing-affecting nonsense variants are not synonymous due to aberrant mRNA splicing; they often lead to frameshift alterations, are recognized by NMD machinery, and are degraded. Therefore, the variant is not synonymous and no protein is produced. Those mRNAs that are not degraded code for altered protein due to aberrant splicing [3]. Nonsense variants inactivate exonic splice sites or activate cryptic splice sites [21], alter exonic splice site regulators (ESEs) [22] or exonic splicing silencers (ESSs) [23]), or modify regulatory exonic miRNA binding sites [24-26].

How frequent are nonsense variants? It is not possible to give an exact estimate as it depends on many factors, being

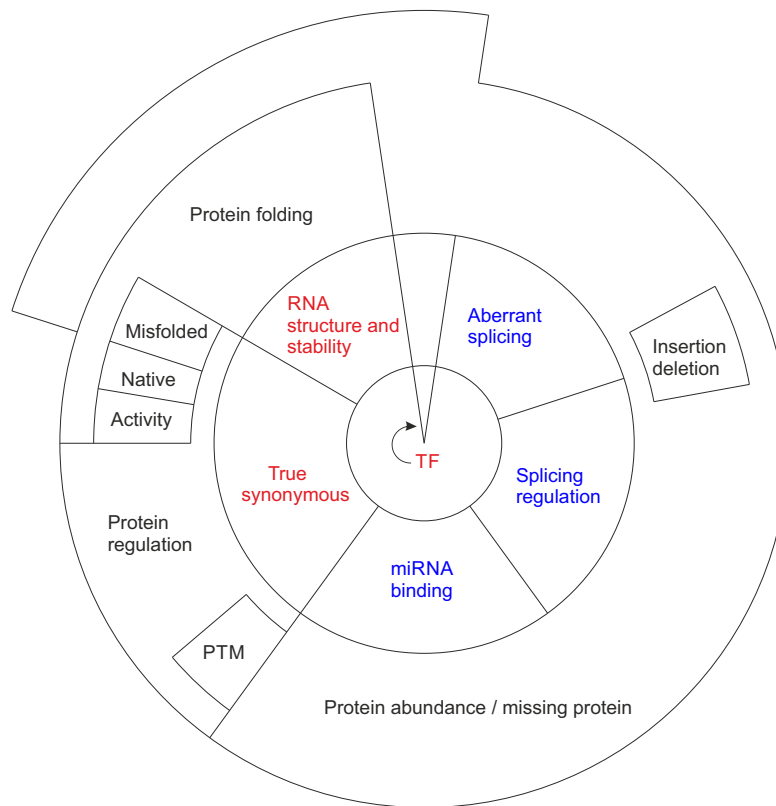


Fig. (1). Classification of synonymous (red) and nonsense (blue) variants and their consequences at DNA (center), RNA (middle), and protein (outer circle) levels. The inner circle indicates DNA variants, which although synonymous can affect transcription factor (TF) binding, and thereby protein production. The middle circle depicts mRNA level alterations, and on the outer ring are shown the protein level effects. At mRNA, synonymous variants are either true synonymous or those that affect RNA structure and stability. Synonymous variants at the protein level have either native or misfolded structure or affect protein activity. Changes to protein abundance are common. PTM indicates post-translational protein modification. The three categories of nonsense variants (blue) affect aberrant splicing, splicing regulation, or miRNA binding. Many of these variants have an effect on protein abundance, and often there is no protein at all. Aberrant splicing and splicing regulation affecting variants can lead also to protein insertion or deletion, if the variant position is close to the C-terminus or the mRNA is not degraded by nonsense-mediated decay. (A higher resolution / colour version of this figure is available in the electronic copy of the article).

different for different genes; however, examples are available in the literature. Of all the possible synonymous variants in exon 7 in the *SMN1* gene, 32 out of 138 variants (23%) decrease exon inclusion [8]. An analysis of 66 out of 67 possible synonymous variations in exon 6 of the *TP53* gene for TP53 protein indicated that nine (13%) variants had a large decrease in splicing [27] due to exon skipping, intron inclusion, or exon truncation. A total of 6.3% of 725 *de novo* coding region variants, which have been identified in autism spectrum disorder families, disrupted splicing [13] and included “synonymous” variants

Recently, Shen and coworkers presented an interesting, systematic study of the fitness effects of thousands of single nucleotide variants on 21 *Saccharomyces cerevisiae* genes [14]. They showed that the majority of synonymous variants had a strong fitness effect, and many of them had an effect on gene expression. In conclusion, we can say that variants that have been classified as synonymous, but which in reality are not synonymous, are frequent, and they often affect splicing and protein abundance. Thus, there is a need for the new classification of variants claimed to be synonymous and for the new term nonsense variant (Fig. 1).

3. PROBLEMS WITH MISCLASSIFICATION OF “SYNONYMOUS” VARIATIONS

Due to a lack of awareness of nonsense variants, they are incorrectly annotated, for *e.g.*, in sequencing projects [11]. They are ignored and lumped together with true synonymous variants by variation annotation tools. For example, ANNOVAR [28], SnpEff [29], and Variant Effect Predictor (VEP) [30] have just one category for synonymous/silent variants. In variation interpretation, these variants are usually ignored, and therefore, disease diagnosis may be prevented or substantially delayed, which may have severe consequences for the patients.

One of the problems with “synonymous variations” was indicated in the title of the News and Views piece describing the work of Shen and others [14]: “Mutations matter even if proteins stay the same” [31]. In the case of “synonymous” variants, the proteins do not always stay the same, and there may not be any protein at all.

4. FITNESS EFFECTS OF “SYNONYMOUS” VARIATIONS

Fitness effects of variations, including “synonymous” variants, have been investigated experimentally in several organisms and genes [9]. These studies have been conducted in viruses, bacteria, and fungi, and widely indicate the variable distribution of fitness effects (DFEs). The DFE scores of synonymous variants can be the same or even lower than those for non-synonymous variants. Thus, many “synonymous” variants are likely not synonymous. The mechanisms are unknown; splicing-related nonsense variants do not occur as viruses and bacteria do not contain introns and have splicing. In the case of yeast, at least some of these observations are likely due to nonsense variants.

We argue that in the extensive study of 21 yeast genes [14], a substantial number of the “synonymous” variants that have non-neutral fitness effects are in fact nonsense vari-

ants. Shen and colleagues investigated, among others, 1866 “synonymous” variants, which showed fitness effects quite similar to missense variants. According to the neutral theory of synonymous variants, these observations cannot be interpreted. To elucidate mechanisms for the effects, relative expression levels (RELs) of variant proteins need to be investigated.

The RELs of altogether 53.8% of the “synonymous variants” deviated significantly from 1 [14], the score for normal gene expression. It is likely that the majority of these instances are not synonymous at all, but affect splicing or regulation, and are thus nonsense variants. It would be interesting to sequence the mRNAs to study splicing aberrations for those variants that have residual mRNA. As the performances of prediction methods for consequences of exonic variants beyond the immediate exon-intron boundary are rather poor, these methods would likely not be applicable here. Therefore, a pragmatic way to investigate the data of Shen *et al.* [14] would be to classify the “synonymous” variants with significant REL deviation from 1 as nonsense variants and repeat the analyses for the four variant classes.

The fact that more than 50% of “synonymous” variants can behave against the assumption of the neutral theory of synonymous variants indicates that the variant naming is not correct and a new classification is needed. Further, various predictions based on the assumption have to be re-assessed as the foundations do not hold.

5. NONSYNONYMOUS TO SYNONYMOUS SUBSTITUTION RATIO

One area based on the assumption that synonymous variants are (nearly) neutral is a calculation of nonsynonymous-to-synonymous substitutions ratios [32] (marked as ω , dN/dS or Ka/Ks). This score has been used as the most common measure of the strength and the mode of natural selection of genes. Several algorithms with codon models and additional properties and assumptions have been implemented [33, 34]. These widely used scores are calculated from multiple sequence alignments of related sequences and are prone to confounding effects, for *e.g.*, because of the choice of sequences, their similarities/identities, codon frequencies, how different nucleotide models are handled, *etc.*

Problems with “synonymous” variations in these scores have been known and discussed [35] and remedies have been suggested. However, the actual reason, the heterogeneity of “synonymous” variations, has not been fully considered. It is now evident that these kinds of calculations include nonsense variants and have to apply more complex and more realistic models. Therefore, it is necessary to evaluate and, when necessary, reassess published predictions of codon substitution model-based estimates of natural selection.

6. PHYLOGENETICS AND EVOLUTIONARY MODELING

Substitution models are used in evolutionary biology to describe alterations during time, *i.e.*, rate of change of variations. These models are at the core of phylogenetic inference and other evolutionary biology applications, including calculations of loss and gain of genes (gene turnover) [36]

and nonsynonymous-to-synonymous substitution ratios. Substitution models can be applied to nucleotide or amino acid sequence data. General time reversible (GTR) family of nested models is widely used in maximum likelihood algorithms [33]. In addition to individual rates for the variations and nucleotide frequencies, additional details of invariable sites, variation across sites, neighbor interactions, *etc.*, are used. Model selection is a critical step in evolutionary inference.

As nonsense variations have not been included in the substitution models, it is necessary to evaluate their contribution to the models as well as generate predictions, such as phylogenies. “Synonymous” variations account for theoretically over 20% of single nucleotide substitutions, and as a substantial portion of these is nonsense cases, they are an important variant category.

7. DATABASES AND PREDICTORS

Nonsense variants are misclassified in databases for “synonymous” variants. For example, the Database of Deleterious Synonymous Mutations (dbDSM) [18] mainly contains nonsense variants, not synonymous ones. This database has also another problem. It contains a large number of markers used in genome-wide association studies (GWASs). Even if the markers are synonymous, it is not relevant for the property as the markers, or tags, hardly ever are related to the associated property, they are just markers for the haplotypes that contain the associated variation.

Regarding cancers, two resources contain massive amounts of “synonymous” variation information, SynMicDb [37] and DMSN [17]. Even these resources do not differ between nonsense and synonymous variants. The availability of these data facilitates further studies of some cases. Several prediction methods have been released for synonymous variants; however, the cases used for training and developing these tools are mainly nonsense variants.

Several methods claim to predict the outcome of synonymous variants, including DDIG-IN [38], EnDSM [39], IDSV, an ensemble approach [40], prDSM [41], Silva [42], Syntool [43], regSNPs-splicing [44], and Transcript-inferred Pathogenicity (TraP) [45]. The cases used to train and develop these methods are mainly for nonsense variants and affect splicing.

Tools dedicated to true synonymous variants are missing and those trained with nonsense variants are not optimal for these cases as the effects are not considered, and variant distribution is biased.

CONCLUSION

The traditional category of “synonymous” variants also contains nonsense variants; therefore, it is necessary to re-evaluate the relevance of the results of some prior studies. In forthcoming investigations, nonsense variants should be included.

The introduction of the concept of nonsense variants facilitates an understanding of systematic annotation and requires changes in annotation tools [28, 29] and variation interpretation [46]. Existing studies that rely on the neutrali-

ty in various codon indices, as well as those on phylogenetic inference and evolutionary modelling need to be reassessed. Although information and examples of non-neutral synonymous variants have been around for years, new methods are needed with more realistic assumptions and premises, including nonsense variants. Therefore, it is necessary to check the foundations of these studies and include nonsense variations in the models, programs, and algorithms. This is not necessarily an easy task since some nonsense variants may be difficult to predict from sequences and their experimental identification requires more experiments than customary at the moment.

LIST OF ABBREVIATIONS

DFEs	=	Distribution of Fitness Effects
ESEs	=	Exonic Splice Site Enhancers
ESSs	=	Exonic Splicing Silencers
GTR	=	General Time Reversible
GWASs	=	Genome-wide Association Studies
TF	=	Transcription Factor
VariO	=	Variation Ontology
VEP	=	Variant Effect Predictor

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

This work was financially supported by Vetenskapsrådet (Grant number 2019-01403) and the Swedish Cancer Society (Grant number CAN 20 1350).

CONFLICT OF INTEREST

The author declares no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] Kurosaki, T.; Popp, M.W.; Maquat, L.E. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell Biol.*, **2019**, *20*(7), 406-420. <http://dx.doi.org/10.1038/s41580-019-0126-2> PMID: 30992545
- [2] Lindeboom, R.G.H.; Vermeulen, M.; Lehner, B.; Supek, F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat. Genet.*, **2019**, *51*(11), 1645-1651. <http://dx.doi.org/10.1038/s41588-019-0517-5> PMID: 31659324
- [3] Vihinen, M. Systematics for types and effects of RNA variations. *RNA Biol.*, **2021**, *18*(4), 481-498. <http://dx.doi.org/10.1080/15476286.2020.1817266> PMID: 32951567
- [4] Vihinen, M. Muddled genetic terms miss and mess the message. *Trends Genet.*, **2015**, *31*(8), 423-425. <http://dx.doi.org/10.1016/j.tig.2015.05.008> PMID: 26091961
- [5] Sauna, Z.E.; Kimchi-Sarfaty, C. Understanding the contribution of synonymous mutations to human disease. *Nat. Rev. Genet.*, **2011**, *12*(10), 683-691. <http://dx.doi.org/10.1038/nrg3051> PMID: 21878961
- [6] Shabalina, S.A.; Spiridonov, N.A.; Kashina, A. Sounds of silence: Synonymous nucleotides as a key to biological regulation and complexity. *Nucleic Acids Res.*, **2013**, *41*(4), 2073-2094.

- <http://dx.doi.org/10.1093/nar/gks1205> PMID: 23293005
- [7] Vihinen, M. When a synonymous variant is nonsynonymous. *Genes*, **2022**, *13*(8), 1485.
- <http://dx.doi.org/10.3390/genes13081485> PMID: 36011397
- [8] Mueller, W.F.; Larsen, L.S.Z.; Garibaldi, A.; Hatfield, G.W.; Hertel, K.J. The silent sway of splicing by synonymous substitutions. *J. Biol. Chem.*, **2015**, *290*(46), 27700-27711.
- <http://dx.doi.org/10.1074/jbc.M115.684035> PMID: 26424794
- [9] Bailey, S.F.; Alonso Morales, L.A.; Kassen, R. Effects of synonymous mutations beyond codon bias: The evidence for adaptive synonymous substitutions from microbial evolution experiments. *Genome Biol. Evol.*, **2021**, *13*(9), evab141.
- <http://dx.doi.org/10.1093/gbe/evab141> PMID: 34132772
- [10] Vihinen, M. Variation Ontology for annotation of variation effects and mechanisms. *Genome Res.*, **2014**, *24*(2), 356-364.
- <http://dx.doi.org/10.1101/gr.157495.113> PMID: 24162187
- [11] Vihinen, M. Systematic errors in annotations of truncations, loss-of-function and synonymous variants. *Front. Genet.*, **2023**, *14*, 1015017.
- <http://dx.doi.org/10.3389/fgene.2023.1015017> PMID: 36713076
- [12] Dhindsa, R.S.; Wang, Q.; Vitsios, D.; Burren, O.S.; Hu, F.; DiCarlo, J.E.; Kruglyak, L.; MacArthur, D.G.; Hurles, M.E.; Petrovski, S. A minimal role for synonymous variation in human disease. *Am. J. Hum. Genet.*, **2022**, *109*(12), 2105-2109.
- <http://dx.doi.org/10.1016/j.ajhg.2022.10.016> PMID: 36459978
- [13] Rhine, C.L.; Neil, C.; Wang, J.; Maguire, S.; Buerer, L.; Salomon, M.; Meremikwu, I.C.; Kim, J.; Strande, N.T.; Fairbrother, W.G. Massively parallel reporter assays discover de novo exonic splicing mutants in paralogs of Autism genes. *PLoS Genet.*, **2022**, *18*(1), e1009884.
- <http://dx.doi.org/10.1371/journal.pgen.1009884> PMID: 35051175
- [14] Shen, X.; Song, S.; Li, C.; Zhang, J. Synonymous mutations in representative yeast genes are mostly strongly non-neutral. *Nature*, **2022**, *606*(7915), 725-731.
- <http://dx.doi.org/10.1038/s41586-022-04823-w> PMID: 35676473
- [15] Zeng, Z.; Bromberg, Y. Predicting functional effects of synonymous variants: A systematic review and perspectives. *Front. Genet.*, **2019**, *10*, 914.
- <http://dx.doi.org/10.3389/fgene.2019.00914> PMID: 31649718
- [16] Zeng, Z.; Aptekmann, A.A.; Bromberg, Y. Decoding the effects of synonymous variants. *Nucleic Acids Res.*, **2021**, *49*(22), 12673-12691.
- <http://dx.doi.org/10.1093/nar/gkab1159> PMID: 34850938
- [17] Miao, X.; Li, X.; Wang, L.; Zheng, C.; Cai, J. DSMNC: A database of somatic mutations in normal cells. *Nucleic Acids Res.*, **2019**, *47*(D1), D971-D975.
- <http://dx.doi.org/10.1093/nar/gky1045> PMID: 30380071
- [18] Wen, P.; Xiao, P.; Xia, J. dbDSM: A manually curated database for deleterious synonymous mutations. *Bioinformatics*, **2016**, *32*(12), 1914-1916.
- <http://dx.doi.org/10.1093/bioinformatics/btw086> PMID: 27153700
- [19] Bali, V.; Bebok, Z. Decoding mechanisms by which silent codon changes influence protein biogenesis and function. *Int. J. Biochem. Cell Biol.*, **2015**, *64*, 58-74.
- <http://dx.doi.org/10.1016/j.biocel.2015.03.011> PMID: 25817479
- [20] Vihinen, M. Individual genetic heterogeneity. *Genes*, **2022**, *13*(9), 1626.
- <http://dx.doi.org/10.3390/genes13091626> PMID: 36140794
- [21] Kim, Y.J.; Kang, J.; Seymen, F.; Koruyucu, M.; Zhang, H.; Kasimoglu, Y.; Bayram, M.; Tuna-Ince, E.B.; Bayrak, S.; Tuloglu, N.; Hu, J.C.C.; Simmer, J.P.; Kim, J.W. Alteration of exon definition causes amelogenesis imperfecta. *J. Dent. Res.*, **2020**, *99*(4), 410-418.
- <http://dx.doi.org/10.1177/0022034520901708> PMID: 31999931
- [22] Nielsen, K.B.; Sørensen, S.; Cartegni, L.; Corydon, T.J.; Doktor, T.K.; Schroeder, L.D.; Reinert, L.S.; Elpeleg, O.; Krainer, A.R.; Gregersen, N.; Kjems, J.; Andresen, B.S. Seemingly neutral polymorphic variants may confer immunity to splicing-inactivating mutations: A synonymous SNP in exon 5 of MCAD protects from deleterious mutations in a flanking exonic splicing enhancer. *Am. J. Hum. Genet.*, **2007**, *80*(3), 416-432.
- <http://dx.doi.org/10.1086/511992> PMID: 17273963
- [23] Tonin, R.; Catarzi, S.; Caciotti, A.; Procopio, E.; Marini, C.; Guerrini, R.; Morrone, A. Progressive myoclonus epilepsy in Gaucher disease due to a new Gly-Gly mutation causing loss of an exonic splicing enhancer. *J. Neurol.*, **2019**, *266*(1), 92-101.
- <http://dx.doi.org/10.1007/s00415-018-9084-4> PMID: 30382391
- [24] Liu, C.; Rennie, W.A.; Carmack, C.S.; Kanoria, S.; Cheng, J.; Lu, J.; Ding, Y. Effects of genetic variations on microRNA: target interactions. *Nucleic Acids Res.*, **2014**, *42*(15), 9543-9552.
- <http://dx.doi.org/10.1093/nar/gku675> PMID: 25081214
- [25] Brest, P.; Lapaquette, P.; Souidi, M.; Lebrigand, K.; Cesaro, A.; Vouret-Craviari, V.; Mari, B.; Barbry, P.; Mosnier, J.F.; Hébuterne, X.; Harel-Bellan, A.; Mograbi, B.; Darfeuille-Michaud, A.; Hofman, P. A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat. Genet.*, **2011**, *43*(3), 242-245.
- <http://dx.doi.org/10.1038/ng.762> PMID: 21278745
- [26] Tay, Y.; Zhang, J.; Thomson, A.M.; Lim, B.; Rigoutsos, I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*, **2008**, *455*(7216), 1124-1128.
- <http://dx.doi.org/10.1038/nature07299> PMID: 18806776
- [27] Bhagavatula, G.; Rich, M.S.; Young, D.L.; Marin, M.; Fields, S. A massively parallel fluorescence assay to characterize the effects of synonymous mutations on TP53 expression. *Mol. Cancer Res.*, **2017**, *15*(10), 1301-1307.
- <http://dx.doi.org/10.1158/1541-7786.MCR-17-0245> PMID: 28652265
- [28] Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.*, **2010**, *38*(16), e164.
- <http://dx.doi.org/10.1093/nar/gkq603> PMID: 20601685
- [29] Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly.*, **2012**, *6*(2), 80-92.
- <http://dx.doi.org/10.4161/fly.19695> PMID: 22728672
- [30] McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The ensembl variant effect predictor. *Genome Biol.*, **2016**, *17*(1), 122.
- <http://dx.doi.org/10.1186/s13059-016-0974-4> PMID: 27268795
- [31] Sharp, N. Mutations matter even if proteins stay the same. *Nature*, **2022**, *606*(7915), 657-659.
- <http://dx.doi.org/10.1038/d41586-022-01091-6> PMID: 35676345
- [32] Goldman, N.; Yang, Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.*, **1994**, *11*(5), 725-736.
- PMID: 7968486
- [33] Arenas, M. Trends in substitution models of molecular evolution. *Front. Genet.*, **2015**, *6*, 319.
- <http://dx.doi.org/10.3389/fgene.2015.00319> PMID: 26579193
- [34] Zhang, Z.; Yu, J. Evaluation of six methods for estimating synonymous and nonsynonymous substitution rates. *Genomics Proteomics Bioinform.*, **2006**, *4*(3), 173-181.
- [http://dx.doi.org/10.1016/S1672-0229\(06\)60030-2](http://dx.doi.org/10.1016/S1672-0229(06)60030-2) PMID: 17127215
- [35] Wisotsky, S.R.; Kosakovsky, S.L.; Shank, S.D.; Muse, S.V. Synonymous site-to-site substitution rate variation dramatically inflates false positive rates of selection analyses: Ignore at your own peril. *Mol. Biol. Evol.*, **2020**, *37*(8), 2430-2439.
- <http://dx.doi.org/10.1093/molbev/msaa037> PMID: 32068869
- [36] Librado, P.; Vieira, F.G.; Sánchez-Gracia, A.; Kolokotronis, S.O.; Rozas, J. Mycobacterial phylogenomics: An enhanced method for gene turnover analysis reveals uneven levels of gene gain and loss among species and gene families. *Genome Biol. Evol.*, **2014**, *6*(6), 1454-1465.
- <http://dx.doi.org/10.1093/gbe/evu117> PMID: 24904011
- [37] Sharma, Y.; Miladi, M.; Dukare, S.; Boulay, K.; Caudron-Herger, M.; Groß, M.; Backofen, R.; Diederichs, S. A pan-cancer analysis of synonymous mutations. *Nat. Commun.*, **2019**, *10*(1), 2569.
- <http://dx.doi.org/10.1038/s41467-019-10489-2> PMID: 31189880
- [38] Livingstone, M.; Folkman, L.; Yang, Y.; Zhang, P.; Mort, M.; Cooper, D.N.; Liu, Y.; Stantic, B.; Zhou, Y. Investigating DNA-, RNA-, and protein-based features as a means to discriminate pathogenic synonymous variants. *Hum. Mutat.*, **2017**, *38*(10), 1336-1347.
- <http://dx.doi.org/10.1002/humu.23283> PMID: 28649752
- [39] Cheng, N.; Wang, H.; Tang, X.; Zhang, T.; Gui, J.; Zheng, C.H.; Xia, J. An ensemble framework for improving the prediction of deleterious synonymous mutation. *IEEE Trans. Circ. Syst. Video Tech.*, **2022**, *32*(5), 2603-2611.

- <http://dx.doi.org/10.1109/TCSVT.2021.3063145>
- [40] Ranganathan, G.S.; Alexov, E. An ensemble approach to predict the pathogenicity of synonymous variants. *Genes*, **2020**, *11*(9), 1102.
- <http://dx.doi.org/10.3390/genes11091102> PMID: 32967157
- [41] Cheng, N.; Li, M.; Zhao, L.; Zhang, B.; Yang, Y.; Zheng, C.H.; Xia, J. Comparison and integration of computational methods for deleterious synonymous mutation prediction. *Brief. Bioinform.*, **2020**, *21*(3), 970-981.
- <http://dx.doi.org/10.1093/bib/bbz047> PMID: 31157880
- [42] Buske, O.J.; Manickaraj, A.; Mital, S.; Ray, P.N.; Brudno, M. Identification of deleterious synonymous variants in human genomes. *Bioinformatics*, **2015**, *31*(5), 799.
- <http://dx.doi.org/10.1093/bioinformatics/btu765> PMID: 25488928
- [43] Zhang, T.; Wu, Y.; Lan, Z.; Shi, Q.; Yang, Y.; Guo, J. Syntool: A novel region-based intolerance score to single nucleotide substitution for synonymous mutations predictions based on 123,136 individuals. *BioMed Res. Int.*, **2017**, *2017*, 1-5.
- <http://dx.doi.org/10.1155/2017/5096208> PMID: 28812016
- [44] Zhang, X.; Li, M.; Lin, H.; Rao, X.; Feng, W.; Yang, Y.; Mort, M.; Cooper, D.N.; Wang, Y.; Wang, Y.; Wells, C.; Zhou, Y.; Liu, Y. regSNPs-splicing: A tool for prioritizing synonymous single-nucleotide substitution. *Hum. Genet.*, **2017**, *136*(9), 1279-1289.
- <http://dx.doi.org/10.1007/s00439-017-1783-x> PMID: 28391525
- [45] Gelfman, S.; Wang, Q.; McSweeney, K.M.; Ren, Z.; La Carpia, F.; Halvorsen, M.; Schoch, K.; Ratzon, F.; Heinzen, E.L.; Boland, M.J.; Petrovski, S.; Goldstein, D.B. Annotating pathogenic non-coding variants in genic regions. *Nat. Commun.*, **2017**, *8*(1), 236.
- <http://dx.doi.org/10.1038/s41467-017-00141-2> PMID: 28794409
- [46] Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; Voelkerding, K.; Rehm, H.L. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet. Med.*, **2015**, *17*(5), 405-424.
- <http://dx.doi.org/10.1038/gim.2015.30> PMID: 25741868