

Research Article

Comparison of Predictive *In Silico* Tools on Missense Variants in *GJB2*, *GJB6*, and *GJB3* Genes Associated with Autosomal Recessive Deafness 1A (DFNB1A)

Vera G. Pshennikova^{1,2}, Nikolay A. Barashkov^{1,2}, Georgii P. Romanov^{1,2},
Fedor M. Teryutin^{1,2}, Aisen V. Solov'ev^{1,2}, Nyurgun N. Gotovtsev^{1,2},
Alena A. Nikanorova^{1,2}, Sergey S. Nakhodkin², Nikolay N. Sazonov², Igor V. Morozov^{3,4},
Alexander A. Bondar^{1,3}, Lilya U. Dzhemileva^{5,6}, Elza K. Khusnutdinova^{5,7},
Olga L. Posukh^{4,8}, and Sardana A. Fedorova^{1,2}

¹Department of Molecular Genetics, Federal State Budgetary Scientific Institution "Yakut Science Centre of Complex Medical Problems", Yakutsk, Russia

²Laboratory of Molecular Biology, Institute of Natural Sciences, M.K. Ammosov North-Eastern Federal University, Yakutsk, Russia

³Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

⁴Novosibirsk State University, Novosibirsk, Russia

⁵Laboratory of Human Molecular Genetics, Institute of Biochemistry and Genetics, Ufa Scientific Centre, Russian Academy of Sciences, Ufa, Russia

⁶Department of Immunology and Human Reproductive Health, Bashkir State Medical University, Ufa, Russia

⁷Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russia

⁸Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Correspondence should be addressed to Nikolay A. Barashkov; barashkov2004@mail.ru

Received 22 October 2018; Revised 25 January 2019; Accepted 3 February 2019; Published 20 March 2019

Academic Editor: António Amorim

Copyright © 2019 Vera G. Pshennikova et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In silico predictive software allows assessing the effect of amino acid substitutions on the structure or function of a protein without conducting functional studies. The accuracy of *in silico* pathogenicity prediction tools has not been previously assessed for variants associated with autosomal recessive deafness 1A (DFNB1A). Here, we identify *in silico* tools with the most accurate clinical significance predictions for missense variants of the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) connexin genes associated with DFNB1A. To evaluate accuracy of selected *in silico* tools (SIFT, FATHMM, MutationAssessor, PolyPhen-2, CONDEL, MutationTaster, MutPred, Align GVGD, and PROVEAN), we tested nine missense variants with previously confirmed clinical significance in a large cohort of deaf patients and control groups from the Sakha Republic (Eastern Siberia, Russia): Cx26: p.Val27Ile, p.Met34Thr, p.Val37Ile, p.Leu90Pro, p.Glu114Gly, p.Thr123Asn, and p.Val153Ile; Cx30: p.Glu101Lys; Cx31: p.Alal94Thr. We compared the performance of the *in silico* tools (accuracy, sensitivity, and specificity) by using the missense variants in *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes associated with DFNB1A. The correlation coefficient (r) and coefficient of the area under the Receiver Operating Characteristic (ROC) curve as alternative quality indicators of the tested programs were used. The resulting ROC curves demonstrated that the largest coefficient of the area under the curve was provided by three programs: SIFT (AUC = 0.833, p = 0.046), PROVEAN (AUC = 0.833, p = 0.046), and MutationAssessor (AUC = 0.833, p = 0.002). The most accurate predictions were given by two tested programs: SIFT and PROVEAN (Ac = 89%, Se = 67%, Sp = 100%, r = 0.75, AUC = 0.833). The results of this study may be applicable for analysis of novel missense variants of the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) connexin genes.

1. Introduction

The most common form of hereditary nonsyndromic hearing loss is autosomal recessive deafness 1A (DFNB1A, MIM#220290) caused by pathogenic variants in the *GJB2*, *GJB6*, and *GJB3* genes encoding connexin 26 (Cx26), connexin 30 (Cx30), and connexin 31 (Cx31) proteins, respectively. The estimated prevalence of DFNB1A among general human population is 14:100 000, and the main cause of DFNB1A is biallelic recessive pathogenic variants in the *GJB2* gene (MIM#121011) (<http://www.ncbi.nlm.nih.gov/books/NBK1272/>, 2018). Currently, about 400 different pathogenic variations of *GJB2* sequence (more than 70% are missense or nonsense amino acid substitutions) are presented in the Human Gene Mutation Database (HGMD, <http://www.hgmd.cf.ac.uk/ac/all.php>), and this list is regularly updated by novel yet unclassified variants. The majority of nonsense variants are pathogenic since they lead to a premature termination of translation and protein synthesis, while missense variants depending on their location in the amino acid sequence can be neutral, damaging, or partially damaging to the structure and function of protein. As a consequence, pathogenicity of many missense variants is difficult to assess.

Basic information on pathogenic mutations is provided by curated databases such as Online Mendelian Inheritance in Man (OMIM) [1] and the Human Gene Mutation Database (HGMD) [2] collecting data on variants of all genes, mainly from the literature. Disease and gene-specific databases often contain variants that are incorrectly classified including incorrect claims published in peer-reviewed literature since different authors interpret the term “mutation pathogenicity” differently because of the increased complexity of analysis and interpretation of clinical genetic testing. Experimental study of the molecular effects of mutations is laborious, whereas useful and reliable information about the effects of amino acid substitutions can readily be obtained by theoretical methods [3]. A variety of *in silico* tools, both publicly and commercially available, can help in the interpretation of sequence variants without structural or functional studies. However, algorithms used by each tool may differ, but can include determination of the effect of the sequence variant at the nucleotide and amino acid as well as the potential impact of the variant on the protein. The impact of a missense substitution depends on criteria such as the evolutionary conservatism of an amino acid/nucleotide, location, and context within the protein sequence and the biochemical consequence of the amino acid substitution [4].

Different *in silico* tools each have their own strengths and weaknesses depending on the algorithm, and in many cases performance varies depending on the certain gene and protein [5, 6]. Performance of available prediction software is constantly being evaluated by comparing their ability to predict “known” disease-causing variants. As a result, the MutPred performed best for variants of genes associated with the RASopathy and limb-girdle muscular dystrophy (LGMD) [7]; the MAPP and the MAPP + PolyPhen-2.1 provided the best combined model for testing variants of *MLH1*, *MSH2*, *MSH6*, and *PMS2* genes associated with Lynch

syndrome, a hereditary form of colon cancer [8]; the SIFT was well suited for the analysis of variants of the *UGT1A1* gene associated with Crigler-Najjar syndrome (congenital hereditary nonhemolytic unconjugated bilirubinemia) [9]; the Align GVGD *in silico* tool was shown as the best for testing variants of genes associated with cancer (*BRCA1*, *BRCA2*, *MLH1*, and *MLH2*) [10]; *in silico* test of 236 *BRCA1/2* missense variants suggested that SIFT and MutationTaster2 are suitable to predict benignity of variants in these genes [11]. There is also a big class of tools for predicting splice site variations which were tested by comparing the predictions against RNA *in vitro* results for natural splice sites of clinically relevant genes in hereditary breast/ovarian cancer (HBOC) [12]. The analysis revealed that HSF, HSF+SSF-like, or HSF+SSF-like+MES achieved a high performance for predicting the disruption of donor sites, and SSF-like for predicting disruption of acceptor sites [12]. In general, most missense variant prediction algorithms are 65-90% accurate when examining known disease variants.

However, so far the accuracy of *in silico* pathogenicity prediction tools was not assessed for variants of genes associated with autosomal recessive deafness 1A. To date, the only published study was focused on the pathogenicity analysis of 211 missense variants of the *GJB2* gene annotated in the Ensembl and the HGMD databases [13]. Four predictive *in silico* tools, SIFT, PANTHER, PolyPhen-2, and FATHMM, were used but the comparison of their performance was not performed.

The aim of this study is to compare the performance of the *in silico* pathogenicity prediction tools by testing the missense variants in *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes associated with the autosomal recessive deafness 1A.

2. Materials and Methods

2.1. Missense Variants Selection. To assess accuracy of selected *in silico* tools, we tested nine missense variants of the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes found earlier in a large cohort of deaf patients and control groups from the Sakha Republic (Eastern Siberia, Russia): *GJB2* (Cx26): c.79G>A (p.Val27Ile), c.101T>C (p.Met34Thr), c.109G>A (p.Val37Ile), c.269T>C (p.Leu90Pro), c.341A>G (p.Glu114Gly), c.368C>A (p.Thr123Asn), and c.457G>A (p.Val153Ile); *GJB6* (Cx30): c.301G>A (p.Glu101Lys); *GJB3* (Cx31): c.580G>A (p.Alal94Thr) [14–16] (Figure 1). Of these, three variants of the *GJB2* gene, c.269T>C (p.Leu90Pro), c.101T>C (p.Met34Thr), and c.109G>A (p.Val37Ile), are pathogenic variants associated with hearing impairment (DFNB1A); the remaining six variants were interpreted as benign variants of no clinical significance [14, 15]. To assess the clinical relevance of the presented missense variants, we analyzed not only the results of the segregation analysis of genotype-phenotype correlation, but also the data from the databases of annotated variants: OMIM (the Online Mendelian Inheritance in Man, <http://www.omim.org>) [1]; HGMD (the Human Gene Mutation Database, <http://www.hgmd.cf.ac.uk>) [2]; the ClinVar (a public archive with interpretations of clinically

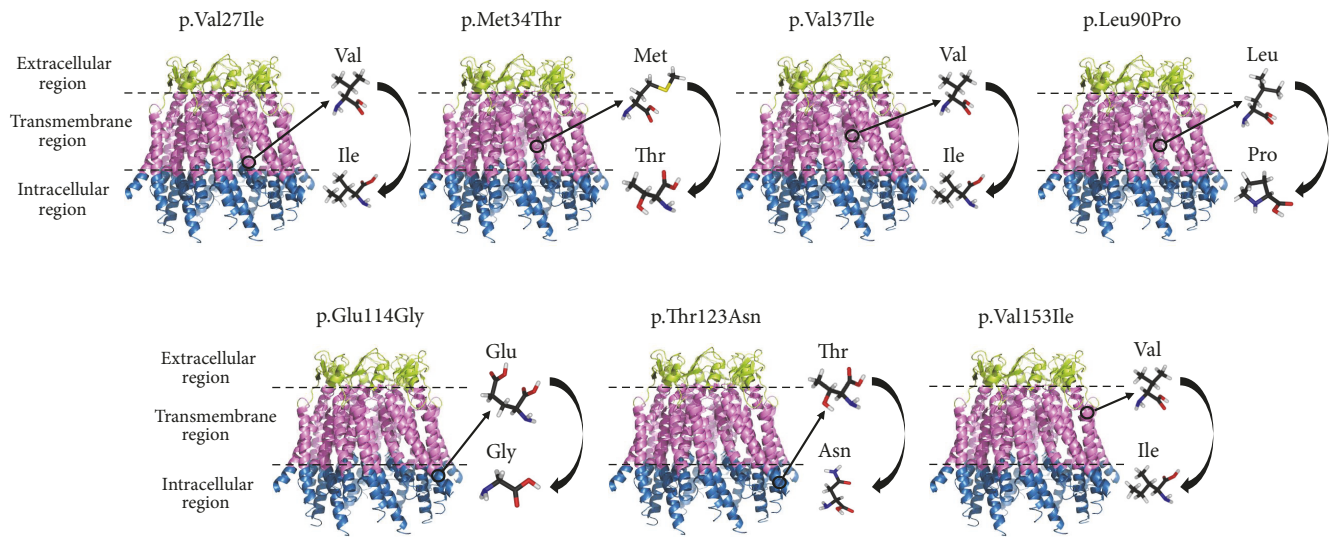


FIGURE 1: Localization of the tested nonsynonymous (missense) amino acid substitutions in the structure of connexin 26. *Note.* The information about the structure Cx26 was obtained from the database of three-dimensional structures of proteins and nucleic acids PDB ID:2ZW3 (<https://www.ncbi.nlm.nih.gov/Structure/pdb/2ZW3>) [22]. Localization of the studied amino acids in structure of Cx26 was obtained using the 3D-structure viewer applet with the protein structure loaded software PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>). Detailed structure models of human Cx30 and Cx31 proteins are currently not defined.

relevant variants, <http://www.ncbi.nlm.nih.gov/clinvar/>) [17, 18]; ExAC (the Exome Aggregation Consortium, <http://exac.broadinstitute.org>) [19]; the 1000 Genomes Project (<http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes>) [20]; dbSNP (the Single Nucleotide Polymorphism database, <http://www.ncbi.nlm.nih.gov/snp/>) [21].

2.2. In Silico Prediction Tools. In this study, 9 predictive computer programs were used to predict pathogenicity: SIFT (Sorting Intolerant From Tolerant) [3, 24–27], FATHMM (Functional Analysis Through Hidden Markov Models) [28–30], MutationAssessor [31, 32], PolyPhen-2 (Polymorphism Phenotyping V-2) [33], CONDEL (Consensus Deleteriousness) [34], MutationTaster [35, 36], MutPred (Mutation Prediction) [37], Align GVG (Align Grantham Variation/Grantham Deviation) [38, 39], and PROVEAN (Protein Variation Effect Analyzer) [40, 41]. Each *in silico* tool uses different parameters for classification of variants which are detailed according to websites listed in Supplementary Materials (see Table S1). The FASTA format and Ensembl sequence identifiers (nucleotide, amino acid, and protein) were used for query in programs (see Table S2).

2.3. Analytical Parameters of In Silico Tools. Analytical parameters of studied tools were calculated according to Fletcher & Fletcher, 2005, and Glantz, 1997 [23, 42]:

Sensitivity (Se) is a proportion of the true-positive results (correct identification of pathogenic variants), according to equation

$$Se = \frac{Tp}{Tp} + FN \times 100\% \quad (1)$$

where Tp denotes true-positive cases and FN denotes false negative cases.

Specificity (Sp) is a proportion of the true negative results (correct identification of benign variants), according to equation

$$Sp = \frac{TN}{TN} + Fp \times 100\% \quad (2)$$

where TN denotes true negative cases and Fp denotes false-positive cases.

Accuracy (Ac) is the ratio of complete correct predictions to the total number of predictions, according to the following equation.

$$Ac = Tp + \frac{TN}{Tp} + TN + Fp + FN \times 100\% \quad (3)$$

Positive predictive values (PPV) are a proportion of positive results that were true-positive (the ratio of true-positive results to all positive results), according to the following equation.

$$PPV = \frac{Tp}{Tp} + Fp \times 100\% \quad (4)$$

Negative predictive values (NPV) are a proportion of negative results that were true negative (the ratio of true negative results to all negative results), according to the following equation.

$$NPV = \frac{TN}{TN} + FN \times 100\% \quad (5)$$

Correlation coefficient (r) is the determination of the relationship between the clinical values of missense variants and predictive evaluation of the program.

ROC curve: the way to express the relationship between sensitivity and specificity for a given test is to construct a curve, called a Receiver Operating Characteristic (ROC) curve [42]. ROC curves are frequently used in the bioinformatic analysis to evaluate classification and prediction models for supporting, diagnosis, and prognosis. To construct a ROC curve, along the Y-axis, the true-positive share (sensitivity) is plotted, along the X-axis, the false-positive share (1 – specificity). The values on the axes ran from probability of 0 to 100% [42]. The quantitative interpretation of ROC is given by AUC (area under ROC curve), the area bounded by the ROC curve and the axis of the share of false-positive cases. The bigger the area under the ROC curve, the better the model. A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system: 0.9-1.0: excellent (A); 0.8-0.9: good (B), 0.7-0.8: fair (C); 0.6-0.7: poor (D); 0.5-0.6: fail (F) (corresponds to random guessing) [43]. The ROC curves were constructed using the MedCalc statistical software for biomedical research (<https://www.medcalc.org>).

3. Results

The predictions for missense variants in the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes by the *in silico* tools in comparison with their established clinical significance are presented in Table 1. Predictions for studied missense variants (3 pathogenic, 6 benign) were different in every analyzed *in silico* tool. Only the c.269T>C (p.Leu90Pro) variant of the *GJB2* gene was evaluated by all programs as a damaging variant.

The informative parameters of the compared programs are presented in Table 2. The accuracy of the clinical significance predictions for missense variants among the analyzed nine programs varies from 33% (FATHMM) to 89% (SIFT and PROVEAN). The SIFT and PROVEAN showed high sensitivity and specificity parameters: 67% and 100%, respectively. The programs MutationAssessor, FATHMM, MutationTaster, and CONDEL had 100% sensitivity, but showed a low specificity, between 33% and 67%, and CONDEL showed total absence of specificity. High rates of predictability of positive and negative results were provided by the SIFT and PROVEAN programs (PPV = 100% and NPV = 86% for both programs) while the FATHMM and Align GVGD programs were the most inaccurate, which resulted in a decrease in almost all of the analyzed parameters. However, FATHMM showed 100% sensitivity since all missense variants were classified by this program as equally damaging.

The overall correlation coefficients are presented in Figure 2. The SIFT and PROVEAN programs demonstrate the highest correlation of *in silico* predictions with observed clinical significance of missense substitutions ($r = 0.75$) which corresponds to their analytical parameters (Table 2). The average values of correlation were shown for MutationAssessor ($r = 0.63$), PolyPhen-2 ($r = 0.5$), and CONDEL ($r = 0.5$) which also correspond to their analytical parameters (Table 2). The MutationTaster demonstrated a weak correlation ($r = 0.37$), MutPred showed very weak correlation

($r = 0.18$), and the FATHMM and Align GVGD programs showed no correlation between the observed values ($r = 0$).

The result of ROC curve analysis is shown in Figure 3. The resulting ROC curves demonstrated that the largest coefficient of the area under the curve was shown by three programs: SIFT (AUC = 0.833, $p = 0.046$, 95% CI: 0.45-0.98), PROVEAN (AUC = 0.833, $p = 0.046$, 95% CI: 0.45-0.98), and MutationAssessor (AUC = 0.833, $p = 0.002$, 95% CI: 0.45-0.98). For PolyPhen-2 and CONDEL, the area of the curve was in the range of 0.7-0.8 (AUC = 0.750, $p = 0.175$, 95% CI: 0.37-0.96), and for MutationTaster it was in the range of 0.6-0.7 (AUC = 0.665, $p = 0.114$, 95% CI: 0.29-0.92). Two programs, FATHMM and Align GVGD, showed a complete lack of information in the predictions (AUC = 0.500, $p = 1.000$, 95% CI: 0.17-0.82).

4. Discussion

For the first time, we analyzed the informative parameters of nine predictive *in silico* tools, obtained by predictions of the clinical significance of missense variants of *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) connexin genes associated with hearing impairment. The capabilities of *in silico* prediction tools were demonstrated by testing nine missense variants with confirmed clinical significance of *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes detected earlier in the study of congenital hearing impairment in the Sakha Republic of Russia [14, 15]. The results of this study may be applicable for analysis of novel missense variants of the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes.

We focused on nine programs chosen according to the following criteria: predicting the impact of missense variants on the function or structure of the protein, differing in computational methods and/or tools, popularity (the top programs included in the dbNSFP [44]), and free online access. Parameters such as accuracy, sensitivity, and specificity were chosen to assess their predictive abilities. Without these parameters, it is not possible to fully evaluate the accuracy of a test [42].

As a result, the SIFT and PROVEAN programs showed the highest sensitivity (Se = 67%) and specificity (Sp = 100%). Thus, the requirement for maximum total sensitivity and specificity in our study was 167% (Se + Sp), while the required balance between sensitivity and specificity was 33% (Δ Se - Sp). The accuracy (Ac) of the predictions of the SIFT and PROVEAN programs was 89%. This result can be considered as the best in this study; it can also be compared to accuracy of predictions published earlier in other studies: 80% - 90% [6, 7, 28, 36, 45]. A lower accuracy was shown by MutationAssessor (Ac = 78%), CONDEL (Ac = 67%), and MutationTaster (Ac = 56%) that were highly sensitive (Se = 100%), but not very specific (Sp = 33-67%). These results indicate a low accuracy of predictions for neutral variants. Align GVGD (Ac = 44%) and FATHMM (Ac = 33%) produced a large number of incorrect pathogenicity predictions and thus were unacceptable for testing variants of the studied genes.

TABLE 1: Evaluation of missense variants by predictive *in silico* tools.

Gene	Missense variants	Clinical significance	SIFT	FATHMM	Mutation/Assessor	Polyphen-2	CONDEL	Mutation/Taster	MutPred	Align GVGD	PROVEAN
GJB2 (Cx26)	c.79G>A p.Val27Ile rs2274084	Benign	Tolerated score: 0.21	Damaging score: -5.59	Medium FI score: 2.28 VC score: 2.16 VS score: 2.40	Probably damaging HumDiv score: 1.000 HumVar score: 0.998	Deleterious Calculated Condel score: 0.612278613903	Polymorphism score: 29	hypotheses are absent general score: 0.321	Unclassified Class C25 GV 0.00 GD 29.61	Neutral score: -0.660
	c.101T>C p.Met34Thr rs35887622	Pathogenic	Damaging score: 0.01	Damaging score: -5.41	Medium FI score: 2.315 VC score: 2.43 VS score: 2.20	Benign HumDiv score: 0.038 HumVar score: 0.083	Deleterious Calculated Condel score: 0.58786807751	Disease causing score: 81	hypotheses are absent general score: 0.969	Deleterious Class C65 GV 0.00 GD 81.04	Deleterious score: -3.801
	c.109G>A p.Val37Ile rs7247424	Pathogenic	Tolerated score: 0.34	Damaging score: -5.46	Medium FI score: 2.095 VC score: 2.58 VS score: 1.61	Probably damaging HumDiv score: 1.000 HumVar score: 0.996	Deleterious Calculated Condel score: 0.61487213316	Disease causing score: 29	hypotheses are absent general score: 0.902	Unclassified Class C25 GV 0.00 GD 29.61	Neutral score: -0.857
	c.269T>C p.Leu90Pro rs80338945	Pathogenic	Damaging score: 0	Damaging score: -5.64	Medium FI score: 3.33 VC score: 4.26 VS score: 2.40	Probably damaging HumDiv score: 1.000 HumVar score: 0.996	Deleterious Calculated Condel score: 0.676708483818	Disease causing score: 98	Confident hypotheses: Gain of sheet (P = 0.039) general score: 0.915	Deleterious GV 0.00 GD 97.78	Deleterious score: -6.482
	c.341A>G p.Glu114Gly rs2274083	Benign	Tolerated score: 0.16	Damaging score: -4.58	Medium FI score: 2.005 VC score: 2.40 VS score: 1.61	Benign HumDiv score: 0.001 HumVar score: 0.001	Deleterious Calculated Condel score: 0.556433693212	Polymorphism score: 98	hypotheses are absent general score: 0.232	Deleterious Class C65 GV 0.00 GD 97.85	Neutral score: -0.123
	c.368C>A p.Thr123Asn rs11033188	Benign	Tolerated score: 0.59	Damaging score: -4.42	Neutral FI score: -0.305 VC score: -0.61 VS score: -0	Benign HumDiv score: 0.000 HumVar score: 0.000	Neutral Calculated Condel score: 0.513276654484	Disease causing score: 53	hypotheses are absent general score: 0.201	Deleterious Class C55 GV 0.00 GD 64.77	Neutral score: 0.797
	c.457G>A p.Val153Ile rs11033186	Benign	Tolerated score: 1	Damaging score: -3.69	Neutral FI score: -0.305 VC score: -0.43 VS score: -0.18	Benign HumDiv score: 0.003 HumVar score: 0.007	Neutral Calculated Condel score: 0.491937780564	Disease causing score: 29	hypotheses are absent general score: 0.488	Unclassified Class C25 GV 0.00 GD 29.61	Neutral score: 0.138
	c.301G>A p.Glu101Lys rs571454176	Benign	Tolerated score: 0.69	Damaging score: -5.26	Neutral FI score: -0.37 VC score: -0.74 VS score: 0	Benign HumDiv score: 0.193 HumVar score: 0.058	Neutral Calculated Condel score: 0.505405538667	Disease causing Score: 56	Actionable hypotheses: Gain of MoRF binding (P = 0.0064) Gain of ubiquitination at E101 (P = 0.0276) Gain of methylation at E101 (P = 0.0345) general score: 0.506	Deleterious Class C55 GV 0.00 GD 56.87	Neutral score: -1.273
	c.580G>A p.Ala194Thr rs121908852	Benign	Tolerated score: 0.91	Damaging score: -3.67	Low FI score: 1.085 VC score: -0.54 VS score: 2.71	Benign HumDiv score: 0.163 HumVar score: 0.110	Deleterious Calculated Condel score: 0.529626647419	Disease causing Score: 58	hypotheses are absent general score: 0.399	Deleterious Class C55 GV 0.00 GD 58.02	Neutral score: 1.636

Note. The correct results (both “true” positive and “true” negative results) are highlighted by bold font.

TABLE 2: Performance of *in silico* tools.

<i>in silico</i> Tools	Accuracy	Sensitivity	Specificity	PPV	NPV
SIFT	89%	67%	100%	100%	86%
MutationAssessor	78%	100%	67%	60%	100%
FATHMM	33%	100%	0%	33%	0%
Polyphen-2	78%	67%	83%	67%	50%
MutationTaster	56%	100%	33%	43%	33%
PROVEAN	89%	67%	100%	100%	86%
Align GVGD	44%	33%	33%	33%	67%
MutPred	67%	33%	83%	50%	71%
CONDEL	67%	100%	50%	50%	100%

Note. *Accuracy (Ac)* - the proportion of the correct test results (that is the sum of true positive and true negative results) among all the patients examined. In our case, this is the proportion of correct estimates of pathogenic and benign variants; *Sensitivity (Se)* - the ability of the diagnostic method to give the correct result which is defined as the proportion of true positive results among all performed tests. In our case, this is the proportion of true positive results, that is, the correct identification of pathogenic variants; *Specificity (Sp)* - the ability of the diagnostic method not to give false positive results in the absence of disease, which is defined as the proportion of true negative results among healthy individuals in studied group. In our case, this is a share of true negative results, that is, a correct identification of benign variants; *Positive predictive values (PPV)* - prediction of pathogenic variants; *Negative predictive values (NPV)* - prediction of benign variants.

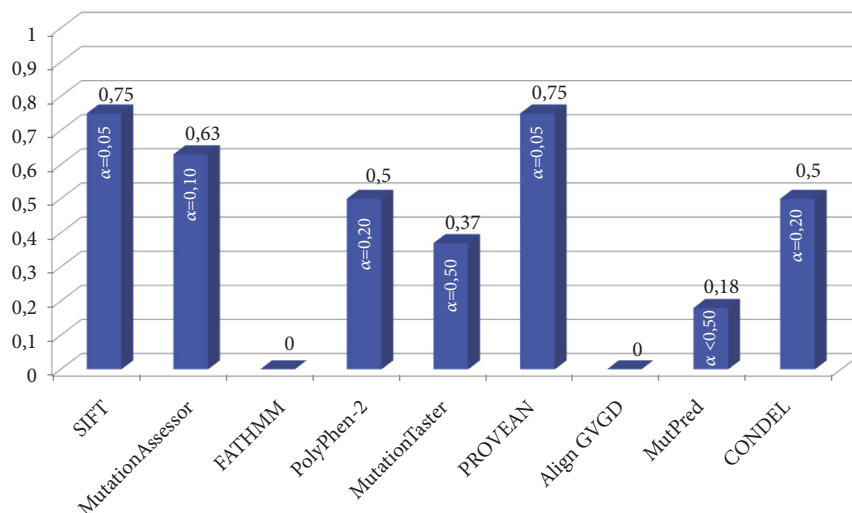


FIGURE 2: The correlation coefficient (r) histogram. Note. r : the relationship between the known clinical significance of missense variants and *in silico* evaluation given by 9 predictive tools; α : the level of significance of the correlation coefficient: the critical value for the significance level and the sample size $n=9$ is 0.933, so the correlation is significant at $p<0.001$ [23].

In addition to the obtained characteristics of accuracy, sensitivity, and specificity, we also used correlation coefficients (r) and areas under the ROC curve (AUC) as alternative indicators of the quality of the tested programs. We compared the values of r and AUC with the quantitative values of the exact predictions of the *in silico* tools under study. For instance, the highest values of $r = 0.75$ were shown by the SIFT and PROVEAN programs that gave the highest number of correct predictions. As is known, the higher the predictive power of the model, the closer the ROC curve to the upper left corner, where the fraction of true-positive cases is 100% (ideal sensitivity) and the share of false-positive cases is zero [42]. The resulting ROC curves demonstrated that the curves of SIFT and PROVEAN were closest to the ideal chart, with the largest area under the curve: AUC = 0.83 (95% confidence interval is 0.45-0.98), which indicates a very

good quality of predictions. The ROC curves of FATHMM and Align GVGD on the diagonal line indicated an absolute lack of informativeness (AUC = 0.500, which corresponds to random guessing); as a result, they had the most erroneous predictions. Our results confirmed that the best programs for bioinformatic analysis of missense variants of the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) connexin genes are SIFT and PROVEAN.

The resulting performance of the PROVEAN and SIFT tools turned out to be fully comparable, as previously described [40, 41]. Note that both programs have the same algorithm of assessing variants by whether they occur in evolutionary conserved region or not, which uses the most popular service, BLASTP (Basic Local Alignment Search Tool) [3, 24, 27, 40, 41]. Thus, we can assume that both tools have the same predictability. However, it should be noted

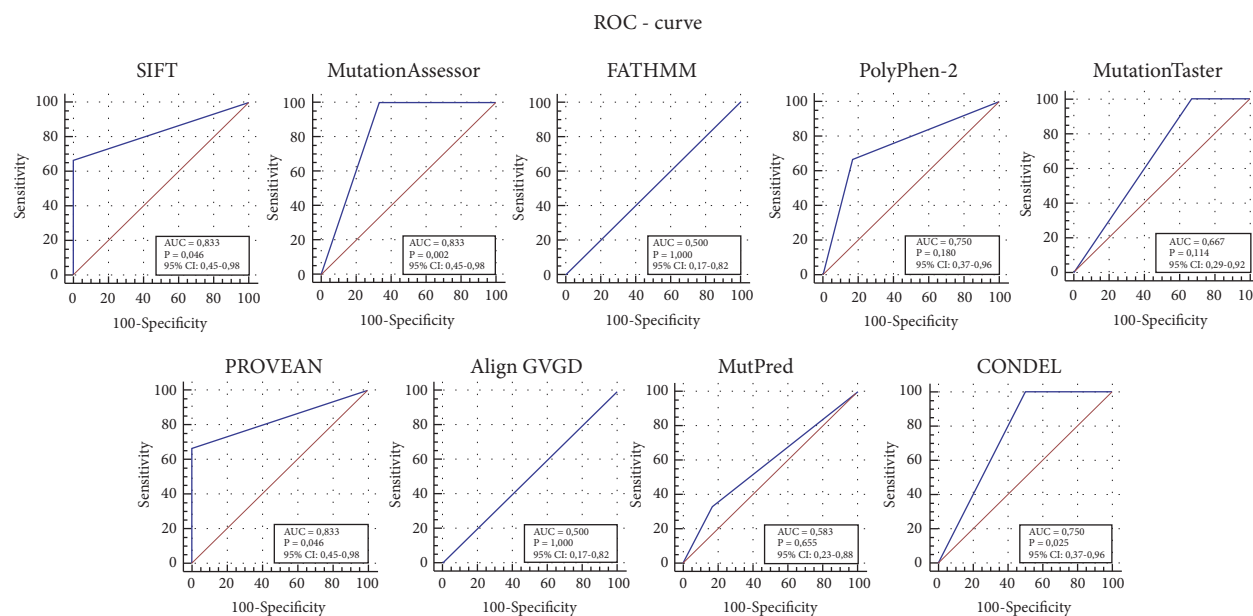


FIGURE 3: ROC curves expressing the relationship of the sensitivity and specificity of the tested programs. These graphs illustrate performance of studied *in silico* tools. The overall accuracy of the tests can be described as the area under the ROC curve (AUC); a higher AUC score indicates a better performance. The diagonal line shows the relationship between true-positive and false-positive values of absolutely uninformative *in silico* tools (FATHMM and Align GVGD). 95% CI indicates 95% confidence interval (Binomial Exact). The ROC curves were constructed using the MedCalc statistical software for biomedical researches (<https://www.medcalc.org>).

that SIFT predicts the effects of all possible substitutions at each position in the protein sequence calculated from a Dirichlet mixture. On the other hand, PROVEAN provides a generalized approach to predict the functional effects of protein sequence variations computed based on BLOSUM62 [40]. The obtained data indicate that, with a wide choice of predictive programs, it is important to consider their methods and tools used for analysis. Also, it should be considered that any computer analysis of biological data is an *in silico* experiment, which has only a more or less reliable prediction that must be verified by other comprehensive structural/functional studies.

5. Conclusion

In summary, the analysis of all obtained informative parameters (accuracy, sensitivity, and specificity) of the nine *in silico* tools along with the correlation coefficient and the area under the ROC curve showed that SIFT and PROVEAN were the tools with the best pathogenicity prediction power; MutationAssessor, PolyPhen-2, and CONDEL performed at an average level; MutationTaster and MutPred were below average; and Align GVGD and FATHMM were uninformative. The results of this study may be applicable for analysis of novel missense variants of the *GJB2* (Cx26), *GJB6* (Cx30), and *GJB3* (Cx31) genes.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Ministry of Education and Science of the Russian Federation (Grant #6.1766.2017), the NEFU in Yakutsk (project: "Genetic features of the population of Sakha Republic: gene pool structure, cold adaptation, psychogenetic characteristics, prevalence of certain genetic and infectious diseases"), the Russian Foundation for Basic Research (Grants #17-29-06-016_ofi_m, #18-015-00212_A, #18-013-00738_A, #18-05-600035_Arctica, and #18-34-00439_mol_a), and the Program for Support of the Bioresource Collections of FASO of Russia "Genome of Sakha Republic", YSC CMP (BRK 0556-2017-0003).

Supplementary Materials

Table S1: description of the *in silico* tools; Table S2: sequence identifiers. (*Supplementary Materials*)

References

- [1] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, "Online mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders," *Nucleic Acids Research*, vol. 33, pp. D514–D517, 2005.
- [2] P. D. Stenson, M. Mort, E. V. Ball et al., "The human gene mutation database: 2008 update," *Genome Medicine*, vol. 1, no. 13, 2009.

- [3] P. C. Ng and S. Henikoff, "Predicting the effects of amino acid substitutions on protein function," *Annual Review of Genomics and Human Genetics*, vol. 7, pp. 61–80, 2006.
- [4] S. Richards, N. Aziz, S. Bale et al., "ACMG laboratory quality assurance committee. 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," *Genetics in Medicine*, vol. 17, no. 5, pp. 405–423, 2015.
- [5] J. Thusberg and M. Vihinen, "Pathogenic or not? and if so, then how? Studying the effects of missense mutations using bioinformatics methods," *Human Mutation*, vol. 30, no. 5, pp. 703–714, 2009.
- [6] J. Thusberg, A. Olatubosun, and M. Vihinen, "Performance of mutation pathogenicity prediction methods on missense variants," *Human Mutation*, vol. 32, no. 4, pp. 358–368, 2011.
- [7] L. C. Walters-Sen, S. Hashimoto, D. L. Thrush et al., "Variability in pathogenicity prediction programs: impact on clinical diagnostics," *Molecular Genetics & Genomic Medicine*, vol. 3, no. 2, pp. 99–110, 2015.
- [8] B. A. Thompson, M. S. Greenblatt, M. P. Vallee et al., "Calibration of multiple in silico tools for predicting pathogenicity of mismatch repair gene missense substitutions," *Human Mutation*, vol. 34, no. 1, pp. 255–265, 2013.
- [9] H. Golehddari, N. Saki, J. Mohammadi-asl, and F. Rahim, "Meta-analysis diagnostic accuracy of SNP-based pathogenicity detection tools: a case of UTG1A1 gene mutations," *International Journal of Molecular Epidemiology and Genetics*, vol. 4, no. 2, pp. 77–85, 2013.
- [10] I. D. Kerr, H. C. Cox, K. Moyes et al., "Assessment of in silico protein sequence analysis in the clinical classification of variants in cancer risk genes," *Journal of Community Genetics*, vol. 8, no. 2, pp. 87–95, 2017.
- [11] C. Ernst, E. Hahnen, C. Engel et al., "Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics," *BMC Medical Genomics*, vol. 11, no. 35, 2018.
- [12] A. Moles-Fernández, L. Duran-Lozano, G. Montalban et al., "Computational tools for splicing defect prediction in breast/ovarian cancer genes: how efficient are they at predicting RNA alterations?" *Frontiers in Genetics*, vol. 9, no. 366, 2018.
- [13] A. Yilmaz, "Bioinformatic analysis of GJB2 gene missense mutations," *Cell Biochemistry and Biophysics*, vol. 71, no. 3, pp. 1623–1642, 2015.
- [14] N. A. Barashkov, V. G. Pshennikova, O. L. Posukh et al., "Spectrum and frequency of the GJB2 gene pathogenic variants in a large cohort of patients with hearing impairment living in a subarctic region of Russia (the Sakha Republic)," *PLoS ONE*, vol. 11, no. 5, Article ID e0156300, 2016.
- [15] V. G. Pshennikova, N. A. Barashkov, A. V. Solovyev et al., "Analysis of GJB6 (Cx30) and GJB3 (Cx31) genes in deaf patients with monoallelic mutations in GJB2 (Cx26) gene in the Sakha Republic (Yakutia)," *Russian Journal of Genetics*, vol. 53, no. 6, pp. 705–715, 2017.
- [16] F. M. Teryutin, N. A. Barashkov, N. L. Kunelskaya et al., "Variability of auditory threshold at deaf patients with splice site c.-23+1G>A mutation in GJB2 gene (Konneksin 26)," *Yakut Medical Journal*, vol. 2, no. 50, pp. 167–172, 2015.
- [17] M. J. Landrum, J. M. Lee, G. R. Riley et al., "ClinVar: public archive of relationships among sequence variation and human phenotype," *Nucleic Acids Research*, vol. 42, no. 1, pp. 980–985, 2014.
- [18] M. J. Landrum, J. M. Lee, M. Benson et al., "ClinVar: public archive of interpretations of clinically relevant variants," *Nucleic Acids Research*, vol. 44, no. D1, pp. 862–868, 2016.
- [19] M. Lek, K. J. Karczewski, E. V. Minikel et al., "Analysis of protein-coding genetic variation in 60,706 humans," *Nature*, vol. 536, no. 7616, pp. 285–291, 2016.
- [20] A. Auton, L. D. Brooks, R. M. Durbin et al., "A global reference for human genetic variation," *Nature*, vol. 526, pp. 68–74, 2015.
- [21] P. C. Ng and S. Henikoff, "Predicting deleterious amino acid substitutions," *Genome Research*, vol. 11, no. 5, pp. 863–874, 2001.
- [22] S. Maeda, S. Nakagawa, M. Suga et al., "Structure of the connexin 26 gap junction channel at 3.5 Å resolution," *Nature*, vol. 458, no. 7238, pp. 597–602, 2009.
- [23] S. A. Glantz, *Primer of Biostatistics*, McGraw-Hill, Health Professions Division, 1997.
- [24] P. C. Ng and S. Henikoff, "Accounting for human polymorphisms predicted to affect protein function," *Genome Research*, vol. 12, no. 3, pp. 436–446, 2002.
- [25] P. C. Ng and S. Henikoff, "SIFT: predicting amino acid changes that affect protein function," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3812–3814, 2003.
- [26] P. Kumar, S. Henikoff, and P. C. Ng, "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm," *Nature Protocols*, vol. 4, no. 7, pp. 1073–1082, 2009.
- [27] H. A. Shihab, J. Gough, D. N. Cooper et al., "Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models," *Human Mutation*, vol. 34, no. 1, pp. 57–65, 2013.
- [28] H. A. Shihab, J. Gough, D. N. Cooper, I. N. M. Day, and T. R. Gaunt, "Predicting the functional consequences of cancer-associated amino acid substitutions," *Bioinformatics*, vol. 29, no. 12, pp. 1504–1510, 2013.
- [29] H. A. Shihab, J. Gough, M. Mort, D. N. Cooper, I. N. M. Day, and T. R. Gaunt, "Ranking non-synonymous single nucleotide polymorphisms based on disease concepts," *Human Genomics*, vol. 8, no. 1, 2014.
- [30] B. Reva, Y. Antipin, and C. Sander, "Determinants of protein function revealed by combinatorial entropy optimization," *Genome Biology*, vol. 8, no. 11, p. 232, 2007.
- [31] B. Reva, Y. Antipin, and C. Sander, "Predicting the functional impact of protein mutations: application to cancer genomics," *Nucleic Acids Research*, vol. 39, no. 17, p. e118, 2011.
- [32] I. A. Adzhubei, S. Schmidt, L. Peshkin et al., "A method and server for predicting damaging missense mutations," *Nature Methods*, vol. 7, no. 4, pp. 248–249, 2010.
- [33] A. González-Pérez and N. López-Bigas, "Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel," *American Journal of Human Genetics*, vol. 88, no. 4, pp. 440–449, 2011.
- [34] J. M. Schwarz, D. N. Cooper, M. Schuelke, and D. Seelow, "MutationTaster2: mutation prediction for the deep-sequencing age," *Nature Methods*, vol. 11, no. 4, pp. 361–362, 2014.
- [35] J. M. Schwarz, C. Rödelsperger, M. Schuelke, and D. Seelow, "MutationTaster evaluates disease-causing potential of sequence alterations," *Nature Methods*, vol. 7, no. 8, pp. 575–576, 2010.
- [36] S. T. Sherry, M. Ward, and K. Sirotkin, "dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation," *Genome Research*, vol. 9, no. 8, pp. 677–679, 1999.

- [37] B. Li, V. G. Krishnan, M. E. Mort et al., "Automated inference of molecular mechanisms of disease from amino acid substitutions," *Bioinformatics*, vol. 25, no. 21, pp. 2744–2750, 2009.
- [38] S. V. Tavtigian, A. M. Deffenbaugh, L. Yin et al., "Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral," *Journal of Medical Genetics*, vol. 43, no. 4, pp. 295–305, 2006.
- [39] E. Mathe, M. Olivier, S. Kato, C. Ishioka, P. Hainaut, and S. V. Tavtigian, "Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods," *Nucleic Acids Research*, vol. 34, no. 5, pp. 1317–1325, 2006.
- [40] Y. Choi, G. E. Sims, S. Murphy, J. R. Miller, and A. P. Chan, "Predicting the functional effect of amino acid substitutions and indels," *PLOS ONE*, vol. 7, no. 10, Article ID 46688, 2012.
- [41] Y. Choi and A. P. Chan, "PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels," *Bioinformatics*, vol. 31, no. 16, pp. 2745–2747, 2015.
- [42] R. H. Fletcher and S. W. Fletcher, *Clinical Epidemiology: The Essentials*, Lippincott Williams & Wilkins, 2005.
- [43] K. H. Zou, "Receiver operating characteristic (ROC) literature research," 2002, <http://splweb.bwh.harvard.edu:8000/pages/ppl/zou/roc.html>.
- [44] X. Liu, C. Wu, C. Li, and E. Boerwinkle, "dbNSFP v3.0: a one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs," *Human Mutation*, vol. 37, no. 3, pp. 235–241, 2016.
- [45] C. Dong, P. Wei, X. Jian et al., "Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies," *Human Molecular Genetics*, vol. 24, no. 8, pp. 2125–2137, 2015.

