**Supplement Table S8**

***1. PolyPhen21 predictions***

PolyPhen-2 (Polymorphism Phenotyping v2) is an online and standalone tool for investigation the AA substitution effect on the structure and function of human proteins. PolyPhen-2 using properties such sequence, phylogenetic and structural to characterize the AA substitution. HumVar is better for Mendelian diseases, when HumDiv should be used for evaluating rare alleles at loci potentially involved in complex phenotypes, dense mapping of regions identified by genome-wide association studies, and analysis of natural selection from sequence data, where even mildly deleterious alleles must be treated as damaging.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Gene** | **cDNA location** | **Aminoacid change** | **PolyPhen2** |
| **HumVar prediction** | **HumVar Score** | **HumDiv prediction** | **HumDiv score** |
| 14 | NIPBL | c.G3431A | p.G1144D | benign | 0.121 | possibly damaging | 0.614 |
| 16 | SPATA7 | c.C769T |  p.R257C | probably damaging | 0.999 | probably damaging | 1 |
| 22 | MYO7A | c.G1955A | p.C652Y | probably damaging | 0.999 | probably damaging | 0.999 |
| MYO7A | c.G5389A | p.D1797N | probably damaging | 0.981 | probably damaging | 0.999 |
| 24 | USH2A | c.G10709T | p.C3570F | probably damaging | 0.997 | probably damaging | 0.999 |
| 26 | MYO7A | c.G3397A | p. G1133R | possibly damaging | 0.455 | possibly damaging | 0.933 |

***2, 3. SIFT2 and PROVEAN3 prediction***

SIFT (Sorting Intolerant from Tolerant) is a sequence homology based tool to predict the effect of a substitution based on AA change. SIFT uses cutoff value of 0.05, which is called tolerance index. Higher tolerance index is better because the less function impact a particular AA substitution is likely to have.

PROVEAN (Protein Variation Effect Analyzer) – is a software tool which predicts whether an amino acid substitution or indel has an impact on the biological function of a protein. If the PROVEAN score is equal to or below a predefined threshold (e.g. -2.5), the protein variant is predicted to have a "deleterious" effect. If the PROVEAN score is above the threshold, the variant is predicted to have a "neutral" effect.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Gene** | **cDNA location** | **Aminoacid change** | **PROVEAN** | **SIFT** |
| **Prediction** | **Score** | **Prediction** | **Score** |
| 14 | NIPBL | c.G3431A | p.G1144D | Neutral | -0.40 | Damaging | 0.010 |
| 16 | SPATA7 | c.C769T |  p.R257C | Deleterious | -5.43 | Damaging | 0 |
| 22 | MYO7A | c.G1955A | p.C652Y | Deleterious | -9.97 | Damaging | 0.001 |
| MYO7A | c.G5389A | p.D1797N | Deleterious | -4.88 | Tolerated | 0.056 |
| 24 | USH2A | c.G10709T | p.C3570F | Deleterious | -7.39 | Damaging | 0 |
| 26 | MYO7A | c.G3397A | p. G1133R | Deleterious | -4.78 | Damaging | 0.017 |

***4. MutPred24 predictions***

A missense mutation having a MutPred score > 0.5 could be considered as “harmful” and a score > 0.75 should be treated as a high conﬁdence “harmful” prediction. Among the six nsSNPs, five was found to be a harmful mutation with a score of > 0.5 and the remaining one nsSNP were found to be normal with the score < 0.5.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Gene** | **cDNA location** | **Aminoacid change** | **MutPred2** |
| **Score** | **Disrupted Mechanism** |
| 14 | NIPBL | c.G3431A | p.G1144D | 0.402 | - |
| 16 | SPATA7 | c.C769T |  p.R257C | 0.501 | Loss of Intrinsic disorder (Pr = 0.42 | P = 0.02); Altered Disordered interface (Pr = 0.38 | P = 5.4e-03); Loss of Phosphorylation at T260 (Pr = 0.29 | P = 0.02); Loss of Proteolytic cleavage at R257 (Pr = 0.15 | P = 0.01); Gain of Sulfation at Y258 (Pr = 0.02 | P = 0.03) |
| 22 | MYO7A | c.G1955A | p.C652Y | 0.975 | Altered Ordered interface (Pr = 0.27 | P = 0.05); Gain of Allosteric site at R657 (Pr = 0.25 | P = 0.01); Loss of Catalytic site at R657 (Pr = 0.20 | P = 0.01); Altered DNA binding (Pr = 0.18 | P = 0.03) |
| MYO7A | c.G5389A | p.D1797N | 0.776 | Gain of Relative solvent accessibility (Pr = 0.24 | P = 0.04); Gain of Allosteric site at F1800 (Pr = 0.23 | P = 0.02); Altered Transmembrane protein (Pr = 0.14 | P = 0.02) |
| 24 | USH2A | c.G10709T | p.C3570F | 0.703 | Altered Transmembrane protein (Pr = 0.25 | P = 1.5e-03); Loss of Disulfide linkage at C3575 (Pr = 0.15 | P = 0.03) |
| 26 | MYO7A | c.G3397A | p. G1133R | 0.528 | Loss of Acetylation at K1128 (Pr = 0.22 | P = 0.03); Altered Transmembrane protein (Pr = 0.16 | P = 0.01) |

***5. I-Mutant 2.0 5 prediction***

The stability of six nsSNPs was predicted by I-Mutant 2.0 through comparing free energy. I-Mutant predicts the stability of the protein upon amino acid substitution by examining the Gibbs free energy by ∆∆G(DDG) value = ∆G (New protein) - ∆G (Wild type) in kcal/mol, which is calculated at pH 7 and 25°C. Among the six nsSNPs, four of them showing decreasing stability of protein and two shows increasing stability.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Gene** | **cDNA location** | **Aminoacid change** | **I-Mutant 2.0** |
| **DDG** | **Stability** |
| 14 | NIPBL | c.G3431A | p.G1144D | 0.72 | Increase |
| 16 | SPATA7 | c.C769T | p.R257C | -1.27 | Decrease |
| 22 | MYO7A | c.G1955A | p.C652Y | 0.18 | Increase |
| MYO7A | c.G5389A | p.D1797N | -1.61 | Decrease |
| 24 | USH2A | c.G10709T | p.C3570F | -0.64 | Decrease |
| 26 | MYO7A | c.G3397A | p. G1133R | -0.61 | Decrease |

***6. Project HOPE6 results***

Project HOPE 3D is web portal for wide analysis of AA substitution on human protein. HOPE collects structural information, sequence annotations, calculations of 3D structure and prediction of secondary structure from different software and databases. Furthermore, HOPE combines this information to give analyze for AA substitution.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Gene** | **cDNA location** | **Aminoacid change** | **Project HOPE** |
| **Structure** | **Properties** |
|
| 14 | NIPBL | c.G3431A | p.G1144D | https://lh5.googleusercontent.com/-eKj09YBCoXIYnq5LVYUp5QT3rekHL2Enpaz6xLE30qRaZ49yOusaLWG9NDR8YOBP_xLptHdVvI3O1cmkHY0bS5rxcZYkWtg1Rek8VtCLNmdSv3H5etoqawH40jOeS3E7yFT7ci4https://lh4.googleusercontent.com/PIh2iIdQf75ISubm182oxZD7WrQuvmyD5MQExYhhPBfEHIdNK-l2eTmVWqlDipIExaOTbJeMCiCJTtzNXFrEU9rpAp0IkBnsDDsGE9kUCDFdTo2zrCKCPJFU1yzYeu9teMZc-T97https://lh6.googleusercontent.com/o-0lv0p08DVNhIuFSk8IXSnSZf1AnC2Fi6_Tw1FkIM-ODADjABrrl38bZy3bvnzLQYJnVE7ZJpwFoeTvQL43wmJh2s-L2qLHdJgSWyiar-kYCqtbfMSz3E7_gOxVtW-uKGB-0sEo | There is a difference in charge between the wild-type and mutant amino acid.The mutation introduces a charge, this can cause repulsion of ligands or other residues with the same charge. The wild-type and mutant amino acids differ in size. The mutant residue is bigger, this might lead to bumps. The torsion angles for this residue are unusual. only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure. |
| 16 | SPATA7 | c.C769T |  p.R257C | https://lh3.googleusercontent.com/mOh2atu8icZmc2NDxgIrYsGQNL07QgkyFaLDL2ZZJxzKyM7iaJTIb4ACT5TqC41dBKH7aAzd4oZ9hT5LtgY5VafE28qmZW6c4-UivG_kM_7SqLbzj4_P1YQTMHLZZiG_H7ndDPNDhttps://lh4.googleusercontent.com/GyzSFgQE_6vkEdrl7R_ReUznvNId8qoBpiYIB-9JeyGZkzWmCTbh8yWvGfSrKH5GyBo4o-_pHFt_k4cS1IYfUNrFeHbFCw64U_9ArXHJ-yPqmE_fmUOiLgw-rOIhRDwDFHVkherbhttps://lh4.googleusercontent.com/Om9v8tC3Kz9uycNRuSL_WhXe-GQigXF74KBjUQeeOJr_2M48tT_OMhId_W4FCoHmeuOI4af8ceL1nWpsSCQENcMGc0HBOQLooUjqO3Zuuee8JpsfVTkSiCVo7d6L4bYl0rtX6Lia | There is a difference in charge between the wild-type and mutant amino acid.The charge of the wild-type residue will be lost, this can cause loss of interactions with other molecules or residues.The wild-type and mutant amino acids differ in size.The mutant residue is smaller, this might lead to loss of interactions.The hydrophobicity of the wild-type and mutant residue differs.The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding. |
| 22 | MYO7A | c.G1955A | p.C652Y | https://lh3.googleusercontent.com/AQkyEQw7dL53VBh4jOeKyrxEwilWvKW3jcjGHkDG_TN1qsyAoCQNQrV-GMJ1aDv1HfbbK6HH3swTXQrqmLCP4z4HaMr9gmw37hq_frYZ3KqgJFwmwzDwYbkfmfPVBWXHHVclxnByhttps://lh4.googleusercontent.com/7iPHAdS_T_iScqA0Lfuj2pHbR73W_9BHGi1yabUOuV6OKqKaewbP53SgCTvVuaaC0MZG9q3UnxeCIN_LB5tdufiKi7anPfdLTaKCUM8mzsvwK0cnuwezWZzsUekzx2yrx9thUOTihttps://lh5.googleusercontent.com/fRivu_S2mYxayi8Oheyxff8JLlprQCg7byyHcnhLguRS1XeqBpWYowkdB_Z2go_48iB-8FVyNsQ0QBAc77FcJiKs1xsfgZxCVZYE-9tJTuae_Qic1ua0B1j_gb8E9h_LFijW7R3j | The wild-type and mutant amino acids differ in size.The mutant residue is bigger than the wild-type residue.The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit.The hydrophobicity of the wild-type and mutant residue differs.The mutation will cause loss of hydrophobic interactions in the core of the protein. |
| MYO7A | c.G5389A | p.D1797N | https://lh5.googleusercontent.com/8lRBxptuldmtVTkF8bpJ7ZukpfQGNFkjE86dmprUSQmzYsx8hVn3M84777-p8Fxm6ziL7lhbzyjRNh08JbCpuvCwA4rKzPiercnBmaFJn5uvb6iKpc5Q3beU-cqjf1V7HcynT3Kjhttps://lh6.googleusercontent.com/mhIJ3H1nc8_YyoJTJguEAaBBmI9kVvQ3zR35F6hpJqNlOWY_iJlwW5iv0l_GGjyKs-1ryzVx7wPTeedesuZhgP-wC14lrSDzPWRXIr70csf3UKCFD_k3Y8UtjAHKqV6t2ArSje56https://lh6.googleusercontent.com/kpXyo-NifbUU2X3hj4dqe4vnplFd8hGxeX_w5HGTgeca7TO792lErbFGO4GA4DFci0lPTKvSnafI_b8WKmI-LlDQxqbxvl8cOLGQr2Lek80CenVPNI6iDr0qlD76RHM7IDV6njgU | There is a difference in charge between the wild-type and mutant amino acid.The charge of the wild-type residue will be lost, this can cause loss of interactions with other molecules or residues. |
| 24 | USH2A | c.G10709T | p.C3570F | https://lh6.googleusercontent.com/WIY-mbgtCveXMC-gjEmSSF1nhmTzrO--Llsve3HSTcRdPqCh3twA-aUDmvwlk4XNUWhP_EkUXhEI0X2JJGqg9GofeifpciTVmAf8BXXO3A30v-pWW0RrVNIxAdEjT-XGIcHSk9UPhttps://lh3.googleusercontent.com/MXl3MzBxI19Kts0vMGeKnb5M9B17Ng8kfCUfCt9Dn1MCR2NrCFuX-8ieUayZ7phMo9-abvcoKhEbK4Q2fHDEX3cdMvXJzz2ObIbwK8adlEIYqjLX76Wh5Upjv1w4k__VxiKzAmEhhttps://lh5.googleusercontent.com/qJ_v5vBO687-KknCKCgAyo6LQW0ze4NatSF_SBE1qooi5hEedj_M3Pa8LuUdM0mdCZ_goWNk3khJImCaojj4Li-rj0P8QnxcDGLb4AnoS9yRqAgjqzy6b9qGCSJRyQ7bfuEPNV4N | The wild-type and mutant amino acids differ in size.The mutant residue is bigger, this might lead to bumps.The hydrophobicity of the wild-type and mutant residue differs.The mutation introduces a more hydrophobic residue at this position. This can result in loss of hydrogen bonds and/or disturb correct folding. |
| 26 | MYO7A | c.G3397A | p. G1133R | https://lh6.googleusercontent.com/i8iCsTuHzRFRV9TfOZdh9jGUVCHQsKb8DltxRnAwhBEen_sLRIpYipsOvRYZf0S0grkK6M_rwvAzz2MDFwF8WbdEDg0m5L5DvoLinZW4yC-8qDBI9jmB0obkE8OJ7UU91X-0IkKYhttps://lh6.googleusercontent.com/76bOyYDkQyfhujNobVt4FUQMxFORy6AxP3vc5bSru1rd56wIK-MjXQRjWWUHFMKH0tMODaOIzQ-SjH0yMk245xJpRcDXuxiOm5nhtq2bmHJOJNqKHgaK_dVYm1fkCc9ky7abg8Edhttps://lh5.googleusercontent.com/f72aeQPQ1lNULyzcNbkkvlLMUyOvitXzEN8Wo6mO-5Bv8p0CxIqY3hT8bBSsd1ZObiSRlJzr_8JK3aSrE5wSE2iDB41TYzLM5ndRjCKEChztOkLSKpUOSrH86sBHJPKn_KaAHANB | There is a difference in charge between the wild-type and mutant amino acid. The mutation introduces a charge, this can cause repulsion of ligands or other residues with the same charge. The wild-type and mutant amino acids differ in size. The mutant residue is bigger, this might lead to bumps. The torsion angles for this residue are unusual. only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure. |

***7. For the prediction of unknown structure MUSTER7 web-based application was used.***





**A.** MYO7A, p.C652Y. Both the wildtype and mutant side chain are shown in green and red respectively.**B.** SPATA7, p.R257C, Both the wildtype and mutant side chain are shown in green and red respectively.**C.** MYO7A, p. D1797N, Both the wildtype and mutant side chain are shown in green and red respectively. **D.** MYO7A, G1133R, Both the wildtype and mutant side chain are shown in green and red respectively.**E.** USH2A, p. C3570F, Both the wildtype and mutant side chain are shown in green and red respectively.

**8. Splicing Variants**

For prediction of splicing variant pathogenicity MutationTaster8, CADD9, and Human Splice Finder10 were used. A variant with a CADD score of >15 interpreted as probably pathogenic and <15 as benign, as suggested by the CADD authors.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ID** | **Gene** | **cDNA location** | **rs ID** | **MutationTaster** | **CADD** | **Human Splice Finder(HSF)** |
| **prediction** | **Score** |
| 4 | USH2A | c.8682-9A>G |  rs372347027  | polymorphism | benign | 12.87 | No significant splicing motif alteration detected. |
| 11 | USH2A | c.8682-9A>G |  rs372347027  | polymorphism | benign | 12.87 | No significant splicing motif alteration detected. |
| 13 | MYO7A | c.3109-12G>A | rs782566244 | polymorphism | benign | -0.27 | No significant splicing motif alteration detected. |
| 17 | RPGR | c.247+2T>C | - | Disease Causing | probably pathogenic | 26.3 | Alteration of the WT donor site,most probably affecting splicing. |
| 28 | USH2A | c.11048-2A>G | rs200871041 | Disease Causing | probably pathogenic | 29.4 | Activation of an intronic cryptic acceptor site. Potential alteration of splicing. Alteration of the WT acceptor site, most probably affecting splicing. |

1. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P,
Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense
mutations. Nat Methods. 2010 Apr;7(4):248-9. doi: 10.1038/nmeth0410-248
2. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Research. 2003;31(13):3812-3814.
3. Choi Y, Chan AP (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31(16): 2745-2747.
4. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. MutPred2: inferring the molecular and phenotypic impact of amino acid variants. bioRxiv 134981; doi: https://doi.org/10.1101/134981.
5. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research. 2005;33(Web Server issue):W306-W310. doi:10.1093/nar/gki375.
6. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics. 2010 Nov 8;11(1):548. DOI: 10.1186/1471-2105-11-548.
7. S. Wu, Y. Zhang. MUSTER: Improving protein sequence profile-profile alignments by using multiple sources of structure information. Proteins: Structure, Function, and Bioinformatics 2008; 72: 547-556.
8. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014 Apr;11(4):361-2.
9. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014 Feb 2. doi: 10.1038/ng.2892.
10. Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human
Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009 May;37(9):e67. doi: 10.1093/nar/gkp215