Table 1: 26 most deleterious functional coding SNPs out of 140 predicted by different prediction tools for CLDN-3

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **S.No** | **nsSNP ID** | **Amino acid change** | **Sift Score** | **Sift Prediction** | **PolyPhen2 score** | **PolyPhen2 Prediction** | **SNP &GO Score** | **SNP &GO Prediction** |
| 1 | rs11549498 | R30C, R>S | 0 | Deleterious | 1 | Damaging | 4 | Disease |
| 2 | rs139191328 | P134Q,P>L | 0 | Deleterious | 1 | Damaging | 2 | Disease |
| 3 | rs141257286 | I143N | 0 | Deleterious | 0.988 | Damaging | 5 | Disease |
| 4 | rs201650771 | D75H | 0.04 | Deleterious | 1.000 | Damaging | 0 | Disease |
| 5 | rs781999702 | L15P | 0 | Deleterious | 0.994 | Damaging | 4 | Disease |
| 6 | rs782211762 | C24G | 0.03 | Deleterious | 1 | Damaging | 7 | Disease |
| 7 | rs782235875 | L129R | 0 | Deleterious | 0.515 | Damaging | 2 | Disease |
| 8 | rs782350219 | Q62R,Q>P | 0.02 | Deleterious | 0.084 | Damaging | 6 | Disease |
| 9 | rs782368222 | V107E | 0.02 | Deleterious | 0.969 | Damaging | 2 | Disease |
| 10 | rs782664344 | V119E | 0.02 | Deleterious | 0.435 | Damaging | 2 | Disease |
| 11 | rs1168202234 | G93R | 0.01 | Deleterious | 0.021 | Damaging | 3 | Disease |
| 12 | rs1199877723 | A128P | 0.01 | Deleterious | 1 | Damaging | 3 | Disease |
| 13 | rs1243569776 | A78V | 0.01 | Deleterious | 0.590 | Damaging | 1 | Disease |
| 14 | rs1375975812 | K64E | 0 | Deleterious | 0 | Damaging | 5 | Disease |
| 15 | rs1397705319 | R144W | 0 | Deleterious | 0.998 | Damaging | 2 | Disease |
| 16 | rs1401007590 | C181W | 0 | Deleterious | 0 | Damaging | 5 | Disease |
| 17 | rs1554626603 | Y219S | 0 | Deleterious | 1 | Damaging | 4 | Disease |
| 18 | rs1554626657 | V151G | 0 | Deleterious | 0.509 | Damaging | 1 | Disease |
| 19 | rs1554626663 | P149L | 0 | Deleterious | 0.003 | Damaging | 4 | Disease |
| 20 | rs1554626665 | Y147C | 0 | Deleterious | 0.999 | Damaging | 2 | Disease |
| 21 | rs1554626675 | N140I | 0 | Deleterious | 1.000 | Damaging | 3 | Disease |
| 22 | rs1554626735 | S68L | 0 | Deleterious | 0.021 | Damaging | 5 | Disease |
| 23 | rs1554626749 | A33T | 0.01 | Deleterious | 0.001 | Damaging | 2 | Disease |
| 24 | rs1554626750 | P27R | 0 | Deleterious | 0.984 | Damaging | 6 | Disease |
| 25 | rs1554626752 | P27S | 0 | Deleterious | 1 | Damaging | 4 | Disease |
| 26 | rs1554626763 | W17C | 0 | Deleterious | 1 | Damaging | 4 | Disease |

Table 2: Protein stability prediction for CLDN-3 most deleterious nsSNPs

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **S.No.** | **CLDN-3 (SNP-ID)** | **Amino acid mutation** | **MuPro** | **MuPro Prediction** | **I mutant** | **I mutant Prediction** |
| 1 | rs11549498 | R30C,R>S | -0.080766614 | Decrease | 1 | **Increase** |
| 2 | rs139191328 | P134Q,P>L | -0.22876211 | Decrease | 9 | Decrease |
| 3 | rs141257286 | I143N | -1.5428041 | Decrease | 4 | Decrease |
| 4 | rs201650771 | D75H | -0.94299293 | Decrease | 9 | Decrease |
| 5 | rs781999702 | L15P | -1.6197015 | Decrease | 1 | **Increase** |
| 6 | rs782211762 | C24G | -1.5552914 | Decrease | 8 | Decrease |
| 7 | rs782235875 | L129R | -2.3250581 | Decrease | 7 | Decrease |
| 8 | rs782350219 | Q62R,Q>P | -0.33846057 | Decrease | 9 | Decrease |
| 9 | rs782368222 | V107E | -1.2419602 | Decrease | 5 | Decrease |
| 10 | rs782664344 | V119E | -1.5076056 | Decrease | 6 | Decrease |
| 11 | rs1168202234 | G93R | -0.48443254 | Decrease | 1 | Decrease |
| 12 | rs1199877723 | A128P | -2.0104554 | Decrease | 1 | Decrease |
| 13 | rs1243569776 | A78V | 0.29228699 | **Increase** | 6 | Decrease |
| 14 | rs1375975812 | K64E | -0.087413489 | Decrease | 8 | Decrease |
| 15 | rs1397705319 | R144W | -1.2739943 | Decrease | 4 | Decrease |
| 16 | rs1401007590 | C181W | -0.48487663 | Decrease | 6 | Decrease |
| 17 | rs1554626603 | Y219S | -1.3271995 | Decrease | 3 | Decrease |
| 18 | rs1554626657 | V151G | -1.8146445 | Decrease | 9 | Decrease |
| 19 | rs1554626663 | P149L | -0.52235807 | Decrease | 5 | Decrease |
| 20 | rs1554626665 | Y147C | -1.5671437 | Decrease | 3 | Decrease |
| 21 | rs1554626675 | N140I | -0.53933944 | Decrease | 4 | **Increase** |
| 22 | rs1554626735 | S68L | -0.14259964 | Decrease | 5 | Decrease |
| 23 | rs1554626749 | A33T | -1.7059897 | Decrease | 5 | Decrease |
| 24 | rs1554626750 | P27R | -0.46505872 | Decrease | 3 | Decrease |
| 25 | rs1554626752 | P27S | -0.75381712 | Decrease | 7 | Decrease |
| 26 | rs1554626763 | W17C | -0.92688127 | Decrease | 1 | Decrease |

Table 3: Effect of CLDN-3 core SNPs on its function predicted by Mutpred2



Table 4: Structural analysis of CLDN-3 core SNPs

|  |  |  |  |
| --- | --- | --- | --- |
| **Amino acid mutation** | **TM align value** | **RMSD value** | **Residue in core region (Procheck)** |
| R30C | 1.00000 | 0.020 | 96.69% |
| L15P | 0.99983 | 0.121 | 96.69% |
| A78V | 1.00000 | 0.009 | 96.69% |
| N140I | 1.00000 | 0.012 | 96.69% |

Table 5: Chemical analysis of CLDN-3 core SNPs by BIOVIA discovery studio visualizer

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **S.N.** | **nsSNPs ID** | **Amino acid position** | | **Residue** | **Hydrophobicity** | **Secondary structure** | **H- Bond** |
| 1 | rs11549498 | 30 | Native | Arginine | -4.5 | Sheet | 5 |
| Mutant | Cysteine | 2.5 | Sheet | 2 |
| 2 | rs781999702 | 15 | Native | Leucine | 3.8 | Helix | 3 |
| Mutant | Proline | -1.6 | Helix | 2 |
| 3 | rs1243569776 | 78 | Native | Alanine | 1.8 | Helix | 2 |
| Mutant | Valine | 4.2 | Helix | 3 |
| 4 | rs1554626675 | 140 | Native | Asparagine | -3.5 | Coil | 3 |
| Mutant | Isoleucine | 4.5 | Coil | 3 |

Table 6: Docking analysis of CLDN-3 with ZO-1 along with its core-nsSNPs

|  |  |
| --- | --- |
| **CLDN-3 Wild-type with ZO-1 PDZ1** | |
| HADDOCK score | -80.6 +/- 6.9 |
| Cluster size | 11 |
| RMSD from the overall lowest-energy structure | 2.6 +/- 2.1 |
| Van der Waals energy | -30.7 +/- 3.6 |
| Electrostatic energy | -259.2 +/- 40.0 |
| Desolvation energy | 0.9 +/- 1.0 |
| Restraints violation energy | 10.8 +/- 12.5 |
| Buried Surface Area | 1322.8 +/- 102.4 |
| Z-Score | -2 |
|  |  |
| **A78V Mutant with ZO-1 PDZ1** | |
| HADDOCK score | -78.1 +/- 2.6 |
| Cluster size | 11 |
| RMSD from the overall lowest-energy structure | 10.9 +/- 0.1 |
| Van der Waals energy | -30.7 +/- 2.4 |
| Electrostatic energy | -269.4 +/- 51.8 |
| Desolvation energy | 4.5 +/- 6.5 |
| Restraints violation energy | 19.0 +/- 11.9 |
| Buried Surface Area | 1259.4 +/- 69.0 |
| Z-Score | -1.9 |
|  |  |
| **N140I Mutant with ZO-1 PDZ1** | |
| HADDOCK score | -80 +/- 4.9 |
| Cluster size | 15 |
| RMSD from the overall lowest-energy structure | 1.1 +/- 0.7 |
| Van der Waals energy | -42.5 +/- 2.8 |
| Electrostatic energy | -214.9 +/- 25.7 |
| Desolvation energy | 3.2 +/- 1.2 |
| Restraints violation energy | 16.9 +/- 14.4 |
| Buried Surface Area | 1287.1 +/- 64.3 |
| Z-Score | -2.2 |
|  |  |
| **L15P Mutant with ZO-1 PDZ1** | |
| HADDOCK score | -80.2 +/- 1.5 |
| Cluster size | 8 |
| RMSD from the overall lowest-energy structure | 11.0 +/- 0.1 |
| Van der Waals energy | -37.3 +/- 3.0 |
| Electrostatic energy | -231.7 +/- 40.7 |
| Desolvation energy | 2.3 +/- 5.6 |
| Restraints violation energy | 10.3 +/- 12.9 |
| Buried Surface Area | 1216.8 +/- 40.5 |
| Z-Score | -2 |
|  |  |
| **R30C Mutant with ZO-1 PDZ1** | |
| HADDOCK score | -77.9 +/- 6.7 |
| Cluster size | 10 |
| RMSD from the overall lowest-energy structure | 0.8 +/- 0.5 |
| Van der Waals energy | -31.5 +/- 6.0 |
| Electrostatic energy | -254.2 +/- 56.0 |
| Desolvation energy | 2.6 +/- 3.4 |
| Restraints violation energy | 17.8 +/- 15.1 |
| Buried Surface Area | 1280.0 +/- 100.8 |
| Z-Score | -1.7 |

Table 7: List of Regulome DB Annotated SNPs

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S. N.** | **SNPs ID** | **Allele** | **Probability** | **Ranking** | **Types** |
|  | rs7084 | G/A/C/T | 0.66703 | 1f | 5' UTR |
|  | rs1554626807 | G/C | 0.61749 | 2b | 5' UTR |
|  | rs941413281 | C/A | 0.76166 | 2b | 5' UTR |
|  | rs1258409881 | G/A/T | 0.76166 | 2b | 5' UTR |
|  | rs1338626682 | A/G | 0.61749 | 2b | 5' UTR |
|  | rs1554626813 | G/T | 0.76166 | 2b | 5' UTR |
|  | rs1286176519 | G/A/C | 0.63284 | 2b | 5' UTR |
|  | rs145265326 | G/A | 0.55436 | 1f | 5' UTR |
|  | rs6460054 | T/C | 0.55436 | 1f | 5' UTR |
|  | rs1291822873 | C/T | 0.91667 | 2a | 5' UTR |
|  | rs11549497 | G/A/C | 1 | 2a | 5' UTR |
|  | rs1312953714 | C/A | 0.78848 | 2b | 5' UTR |
|  | rs1449093158 | T/C | 0.82541 | 2b | 5' UTR |
|  | rs1563619234 | G/A/T | 0.93104 | 2b | 3'UTR |
|  | rs1282648052 | T/C | 0.7889 | 2b | 3'UTR |
|  | rs1444708497 | G/A | 0.44059 | 2b | 3'UTR |
|  | rs1464483280 | G/A/C | 0.81166 | 2b | 3'UTR |
|  | rs376199383 | T/A/G | 0.81166 | 2b | 3'UTR |
|  | rs782676405 | G/T | 0.52283 | 2b | 3'UTR |
|  | rs1238223132 | G/A | 0.81166 | 2b | 3'UTR |
|  | rs1177503662 | G/T | 0.81166 | 2b | 3'UTR |
|  | rs371634289 | G/A | 0.81166 | 2b | 3'UTR |
|  | rs1554626585 | G/T | 0.81166 | 2b | 3'UTR |
|  | rs373358033 | G/A | 0.81166 | 2b | 3'UTR |
|  | rs1554626591 | G/A | 0.52283 | 2b | 3'UTR |
|  | rs1274992648 | T/A/G | 0.81166 | 2b | 3'UTR |
|  | rs1221946749 | G/T | 0.81166 | 2b | 3'UTR |
|  | rs1054700451 | G/T | 0.42482 | 2b | 3'UTR |

Table 8: miRNAs-SNPs association analysis of 3’ UTR SNPs

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SNPs ID** | **Position** | **Allele** | **Target gain with SNP in 3'UTR** | **Target loss with SNP in 5'UTR** |
| **rs1282648052** | chr7:73769331 | T/C | 2 (hsa-miR-8072, hsa-miR-3960) | 1 (hsa-miR-579-5p) |
| **rs1444708497** | chr7:73769333 | G/A | 3 (hsa-miR-10395-5p, hsa-miR-8069, hsa-miR-598-5p) | 1 (hsa-miR-6786-5p) |
| **rs1464483280** | chr7:73769341 | G/A | 4 (hsa-miR-142-3p, hsa-miR-4255, hsa-miR-4536-5p, hsa-miR-6073) | 1 (hsa-miR-4694-5p) |
| **rs1054700451** | chr7:73769147 | G/T | 4 (hsa-miR-493-3p, hsa-miR-380-5p, hsa-miR-1304-5p, hsa-miR-563 ) | 6 (hsa-miR-7515, hsa-miR-9986, miR-3179, miR-1294, miR-4316, miR-4710) |

Table 9: Minimum free energy (MFE) analysis of wild and mutant structures of mRNAs caused by 3’, 5’ UTR SNPs

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **S. N.** | **SNPs ID** | **Allele** | **Location (Chr 7)** | **Start Region** | **End Region** | **MFE Wild (kcal/mol)** | **MFE Mutated (kcal/mol)** |
| **1** | rs1054700451 | G/T | 73769147 | 73769097 | 73769199 | -32.3 | -30.8 |
| **2** | **rs1563619234** | **G/A/T** | **73769330** | **73769280** | **73769380** | **-33.3** | **-35.00** |
| **3** | **rs1282648052** | **T/C** | **73769331** | **73769281** | **73769381** | **-34.4** | **-38.4** |
| **4** | rs1444708497 | G/A | 73769333 | 73769283 | 73769383 | -30.8 | -30.4 |
| **5** | rs1464483280 | G/A/C | 73769341 | 73769291 | 73769391 | -29.7 | -29.9 |
| **6** | rs376199383 | T/A/G | 73769345 | 73769295 | 73769395 | -32.1 | -30.60 |
| **7** | rs782676405 | G/T | 73769348 | 73769298 | 73769398 | -34.7 | -34.2 |
| **8** | rs1238223132 | G/A | 73769350 | 73769300 | 73769400 | -34.7 | -35.6 |
| **9** | rs1177503662 | G/T | 73769351 | 73769301 | 73769401 | -33.2 | -32.1 |
| **10** | rs371634289 | G/A | 73769356 | 73769306 | 73769406 | -33.3 | -34.6 |
| **11** | rs1554626585 | G/T | 73769357 | 73769307 | 73769407 | -32.9 | -31.4 |
| **12** | rs373358033 | G/A | 73769359 | 73769309 | 73769409 | -31.9 | -32.9 |
| **13** | rs1554626591 | G/A | 73769363 | 73769313 | 73769413 | -32.7 | -31.3 |
| **14** | rs1274992648 | T/A/G | 73769364 | 73769314 | 73769414 | -33.7 | -33.3 |
| **15** | **rs1221946749** | **G/T** | **73769367** | **73769317** | **73769417** | **-22.7** | **-27.2** |
| **16** | **rs7084** | **G/A/C/T** | **73770105** | **73770055** | **73770155** | **-53.4** | **-55.2** |
| **17** | rs1554626807 | G/C | 73770148 | 73770098 | 73770198 | -42.6 | -41.1 |
| **18** | rs941413281 | C/A | 73770150 | 73770100 | 73770200 | -42.5 | -43.3 |
| **19** | rs1258409881 | G/A/T | 73770151 | 73770101 | 73770201 | -42.1 | -43.1 |
| **20** | rs1338626682 | A/G | 73770153 | 73770103 | 73770203 | -42.5 | -47.2 |
| **21** | rs1554626813 | G/T | 73770155 | 73770105 | 73770205 | -44.8 | -44.2 |
| **22** | rs1286176519 | G/A/C | 73770163 | 73770113 | 73770213 | -37.6 | -35.9 |
| **23** | rs145265326 | G/A | 73770186 | 73770136 | 73770236 | -32.5 | -32.2 |
| **24** | rs6460054 | T/C | 73770204 | 73770154 | 73770254 | -35.4 | -35.4 |
| **25** | rs1291822873 | C/T | 73770217 | 73770167 | 73770267 | -43 | -42.1 |
| **26** | **rs11549497** | **G/A/C** | **73770218** | **73770168** | **73770268** | **-43.7** | **-48** |
| **27** | rs1312953714 | C/A | 73770236 | 73770186 | 73770286 | -55.7 | -53.5 |
| **28** | rs1449093158 | T/C | 73770244 | 73770194 | 73770294 | -57 | -57 |