# Supplementary data

Table 1 Comparison of clinical data scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr** | **Peptide Name** | **Sequence** | **HemoNet Scores** | **HemoPred Score** | **HemoPI Score** |
|  | DRAMP18057|MSI-78(Pexiganan) | GIGKFLKKAKKFGKAFVKILKK | 0.614882 | 1 | 0.62 |
|  | DRAMP18058|MSI-78(Pexiganan) | GIGKFLKKAKKFGKAFVKILKK | 0.617765 | 1 | 0.62 |
|  | DRAMP18059|Iseganan(I-367) | RGGLCYCRGRFCVCVGR | 0.200481 | 1 | 0.78 |
|  | DRAMP18060|Iseganan(I-367) | RGGLCYCRGRFCVCVGR | 0.200794 | 1 | 0.78 |
|  | DRAMP18063|P113 | AKRHHGYKRKFH | 0.194891 | 0 | 0.39 |
|  | DRAMP18064|P113D | AKRHHGYKRKFH | 0.192939 | 0 | 0.39 |
|  | DRAMP18079|Mersacidin | CTFTLPGGGGVCTLTSECIC | 0.000794852 | 1 | 0.73 |
|  | DRAMP18080|Plectasin | GFGCNGPWDEDDMQCHNHCKSIKGYKGGYCAKGGFVCKCY | 0.0019304 | 0 | 0.24 |
|  | DRAMP18081|PAC113 | AKRHHGYKRKFH | 0.189225 | 0 | 0.39 |
|  | DRAMP18090|LL-37 | LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES | 0.224533 | 1 | 0.88 |
|  | DRAMP18152|IM942 | KSRIVPAIPVSLL | 0.000329406 | 1 | 0.48 |
|  | DRAMP18159|Pexiganan acetate [MSI-78] | GIGKFLKKAKKFGKAFVKILKK | 0.622829 | 1 | 0.62 |
|  | DRAMP18160|miganan (MI-226/M-226/CLS001) | ILRWPWWPWRRK | 0.409575 | -- | 0.83 |
|  | DRAMP18161|P-145 (24-mer peptide) | IGKEFKRIVERIKRFLRELVRPLR | 0.0810671 | 0 | 0.56 |
|  | DRAMP18172|acteriocin R-7 | KTYYGTNGVHCTKNSLWGKVRLKNMKYDQNTTYMGRLQDILLGWATGAFGKTFH | 0.317911 | 1 | 0.8 |
|  | DRAMP18173|Pediocin PA-1 (acteriocin) | KYYGNGVTCGKHSCSVDWGKATTCIINNGAMAWATGGHQGNHKC | 0.145417 | 1 | 0.45 |
|  | DRAMP18174|Nisin A (Type A lantibiotic) | ITSISLCTPGCKTGALMGCNMKTATCHCSIHVSK | 0.999908 | 1 | 0.58 |
|  | DRAMP18175|ac8c | RIWVIWRR | 0.399976 | 0 | 0.48 |
|  | DRAMP18176|Temporin10a | FLPLASLFSRLL | 5.2754e-10 | 1 | 0.45 |
|  | DRAMP18177|Syphaxin(SP1-22) | GVLDILKGAAKDLAGHVATKVINKI | 0.194938 | 1 | 0.7 |
|  | DRAMP18178|IDR-1 | KSRIVPAIPVSLL | 0.00060035 | 1 | 0.48 |
|  | DRAMP18179|IDR-1002 | VQRWLIVWRIRK | 0.200285 | 1 | 0.37 |
|  | DRAMP18180|uforin II | TRSSRAGLQWPVGRVHRLLRK | 2.70963e-12 | 0 | 0.17 |
|  | DRAMP18181|DP178 (T20, Enfuvirtide & Fuzeon) | YTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWF | 1.17248e-13 | 0 | 0.33 |
|  | DRAMP18182|Sifuvirtide (SFT) | AcSWETWEREIENYTRQIYRILEESQ EQQDRNERDLLE | 4.73229e-07 | 0 | 0.4 |
|  | DRAMP18183|S006 (M6) | QKKIRVRLSA | 0.116088 | 0 | 0.16 |
|  | DRAMP20760|C16G2 | TFFRLFNRSFTQALGKGGGKNLRIIRKGIHIIKKY | 0.463401 | 1 | 0.95 |
|  | DRAMP20773|Dusquetide(SG942) | RIVPA | 5.48762e-08 | 1 | 0.43 |

Table 2 External validation

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Peptide** | **Sequence** | **N-Terminal** | **C-Terminal** | **%Hemolysis** | **Percentage Concentration in μM** | **Target** | **HELMO Scores** | **HemoPred Score** | **HemoPI Score** | **Refer- -ence** |
| Anoplin [K4,7k] | GLLKRIKTLL |  | AMD | 10 | >256 | Non-hemolytic | -0.02 | 1 | 0.51 | [5]  [5] |
| GNU7 | RLLRPLLQLLKQKLR |  |  | 2 | 65 | Non-hemolytic | -0.10 | 0 | 0.47 | [6] |
| Serrulin | GFGGGRGGFGGGRGGFGGGGIGGGGFGGGYGGGKIKG |  |  | 0 | 50 | Non-hemolytic | 0.01 | 0 | 0.5 | [7] |
| Flower-specific defensin / NaD1 | RECKTESNTFPGICITKPPCRKACISEKFTDGHCSKILRRCLCTKPC |  |  | 1.5 | 50 | Non-hemolytic | -0.12 | 0 | 0.45 | [8] |
| Odorranin-HP, Odorranain-W-RA1 | GLLRASSVWGRKYYVDLAGCAKA |  |  | 50 | 500 | Non-hemolytic | 0.71 | 0 | 0.44 | [9] |
| Synthetic peptide | XXXXGXXX |  |  | 0.7 | 100 | Hemolytic | -0.33 | - | 0.48 | [10] |
| Synthetic peptide | GIVKXIVKXIVKXI |  | AMD | 25 | 150 | Hemolytic | 0.84 | - | 0.44 | [11] |
| D-Chensinin-1b (3-13) [h4,10r] | vwrrwrrfwrr |  | AMD | 50 | 4000 | Hemolytic | 0.48 | - | 0.44 | [12] |
| Synthetic  peptide | klrsllrtlsrakaaklrtllralsr |  |  | 1.67 | 500 | Hemolytic | 1.16 | 1 | 0.46 | [13] |
| Cathelicidin-6, Antibacterial peptide BMAP-27(1-18), Myeloid antibacterial peptide 27(1-18) | GRFKRFRKKFKKLFKKLS | ACT | AMD | 50 | 4 | Hemolytic | 1.18 | 0 | 0.47 | [14] |
| BMAP-28(1-18) | GGLRSLGRKILRAWKKYG |  | AMD | 50 | >200 | Hemolytic | 0.46 | 0 | 0.53 | [15] |
| BMAP-28(1-18) [G1V, G2L, S5A, G7A, R12L, A13G, W14I] | VLLRALARKILLGIKKYG | C12 | AMD | 60 | 100 | Hemolytic | 0.52 | 1 | 0.44 | [15] |
| Synthetic peptide | XXGXXXXX | C12 | AMD | 92 | 100 | Hemolytic | 0.44 | - | 0.44 | [10] |
| D2D [k1a] | akxfkxkx |  | AMD | 56 | 150 | Hemolytic |  | - | 0.44 | [16] |
| D2D [k5f] | kkxffxkx |  | AMD | 41 | 150 | Hemolytic |  | - | 0.44 | [16] |

Table 3 HemoNet for fivefold Hyperparameters

|  |  |  |
| --- | --- | --- |
| **Features** | **Classifier** | **Parameters** |
| SeqVeq, OHE | Neural Network |  |
| SVM | C=4, g=512 |
| XGBoost | e= 135 d=5 |
| RF | e= 140, d=145 |
| 1mer, OHE | Neural Network |  |
| SVM | Rbf 1,64 |
| XGBoost | e= 120 d=11 |
| RF | e= 90, d=20 |
| 2mer, OHE | Neural Network |  |
| SVM | Rbf, 16,16 |
| XGBoost | e= 130 d=15 |
| RF | e= 130, d=50 |
| SeqVeq, Smiles-based-ECFP4 | Neural Network |  |
| SVM | Rbf, 16,16 |
| XGBoost | 0.87 e= 85 d= 5 |
| RF | 0.86 e= 80 d= 13 |
| 1mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 128 g= 128 |
| XGBoost | e= 140 d= 50 |
| RF | e= 150 d= 30 |
| 2mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 32 g= 256 |
| XGBoost | e= 130 d= 40 |
| RF | e= 110 d= 60 |
| SeqVeq, 1mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 8 g= 256 |
| XGBoost | e=90 d=140 |
| RF | 0.8598 e= 200 d= 200 |
| SeqVeq, 2mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 1 g= 512 |
| XGBoost |  |
| RF | 0.8598 e= 200 d= 200 |

Table 4 90 percent non redundancy removal Hyperparameters

|  |  |  |
| --- | --- | --- |
| **Features** | **Classifier** | **parameters** |
| SeqVeq, OHE | Neural Network |  |
| SVM | c= 1 AM= 0.7135 |
| XGBoost | e= 170 d= 5 |
| RF | d=10, e=170 AM= 0.689 |
| 1mer, OHE | Neural Network |  |
| SVM | c= 128 AM = 0.6656 |
| XGBoost | e= 70 d= 200 AM = 0.7787 |
| RF | e= 170 d= 170 AM = 0.659 |
| 2mer, OHE | Neural Network |  |
| SVM | c= 64 AM = 0.64082 |
| XGBoost | e= 170 d= 15 AM = 0.6819 |
| RF | e= 130 d= 50 AM = 0.624 |
| SeqVeq, Smiles-based-ECFP4 | Neural Network |  |
| SVM | 0.70 c= 1 g= 1e-05 |
| XGBoost | e= 170 d= 5 AM= 0.684 |
| RF | e= 170 d= 135 AM = 0.669 |
| 1mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM |  |
| XGBoost | e= 185 d= 10 AM = 0.7709 |
| RF | e= 200 d= 150 AM= 0.768 |
| SeqVeq, 1mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | C=64 |
| XGBoost | e= 200 d= 10 AM = 0.735 |
| RF | e= 200 d= 90 AM = 0.7305 |
| SeqVeq, 2mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 100 g= 1e-05 |
| XGBoost | e= 145 d= 140 AM = 0.7813 |
| RF | e= 200 d= 90 AM = 0.7724 |

Table 5 70 percent non redundancy removal Hyperparameters

|  |  |  |
| --- | --- | --- |
| **Features** | **Classifier** | **parameters** |
| SeqVeq, OHE | Neural Network | Epoch=500, Loss<0.45 |
| SVM | c= 1 g= 1e-05 AM = 0.632 |
| XGBoost | e= 170 d= 1 AM = 0.5977 |
| RF | D=1,e=170, Auc 0.619, |
| 1mer, OHE | Neural Network |  |
| SVM | c= 100 AM = 0.620 |
| XGBoost | e= 170 d= 1 AM = 0.610 |
| RF | e= 170 d= 5 AM = 0.5781 |
| 2mer, OHE | Neural Network |  |
| SVM | c= 100 AM= 0.552 |
| XGBoost | e= 170 d= 1 AM= 0.59 |
| RF | 0.67 e= 170 d= 1 AM= 0.67 |
| SeqVeq, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 1 g= 1e-05 AM= 0.690 |
| XGBoost | Auc=0.76 E=170, d=5 |
| RF | Auc=0.73 E=170, d=5 |
| 1mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | 256,0.001 |
| XGBoost | e= 170 d= 5, AM= 0.718 |
| RF | e= 170 d= 5, AM= 0.768 |
| SeqVeq, 1mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 256 g= 0.001 AM= 0.755 |
| XGBoost | e= 160 d= 25 AM = 0.74204 |
| RF | e= 100 d= 55 AM = 0.758 |
| SeqVeq, 2mer, Smiles-based-ECFP4 | Neural Network |  |
| SVM | c= 16 g= 0.01 AM= 0.7657 |
| XGBoost | e= 200 d= 90 |
| RF | e= 200 d= 90 AM = 0.73 |