Supporting Information

Integrated platform for decoding hydrophilic peptide fingerprints of hepatocellular carcinoma using artificial intelligence and two-dimensional nanosheets

Zhiyu Li,[a] Bingcun Ma,[b] Shaoxuan Shui,[a] Zunfang Tu,[b] Weili Peng,[c] Yuanyuan Chen,[c] Juan Zhou,*[d] Fang Lan,*[a] Binwu Ying[d] and Yao Wu[a]

[a]National Engineering Research Center for Biomaterials, School of Biomedical Engineering, Sichuan University, Chengdu 610064, China

[b]Sichuan Institute for Drug Control, Chengdu 610097, China

[c]Machine Intelligence Lab, College of Computer Science, Sichuan University, Chengdu 610064, China

[d]Department of Laboratory Medicine, West China Hospital, Sichuan University, Chengdu 610064, China

*Corresponding author: Fang Lan; Juan Zhou

E-mail: fanglan@scu.edu.cn (F. Lan); zhoujuan39@wchscu.cn (J. Zhou)

This part included:

Experimental Section

Fig. S1-S13

Table. S1-S5

Experimental Section.

Chemicals and reagents.

Ti_3AlC_2 (99 wt %) were obtained from 11 technology Co., Ltd. Iron (III) chloride hexahydrate ($FeCl_3 \cdot 6H_2O$), sodium acetate (NaAc), bovine serum albumin (BSA) trypsin, dithiothreitol (DTT), human serum immunoglobulin G (IgG), iodoacetamide (IAA) and 2,5-dihydroxyl benzoic acid (DHB) were purchased from Sigma Aldrich. trifluoroacetic acid (TFA) was obtained from J&K Chemical Ltd (Shanghai, China). Phytic acid 70% (PA) and Lithium fluoride (LiF) (300 mesh) were purchased from Aladdin (Shanghai, China). Ethylene glycol (EG), ammonium bicarbonate, anhydrous ethanol, hydrochloric acid (HCl), sodium chloride (NaCl), 1, 6-hexylenediamine were purchased from Forest Science and Technology Development Co. Ltd (Chengdu, China)

Synthesis of magnetic MXene/PA (MMP) nanosheet.

First, MXene was synthesized using a previously reported method.[1] In the typical process, LiF (1 g) and Ti_3AlC_2 MAX (1 g) were dissolved in 20 mL of HCl (9 M) under magnetic stirring for 24 h, at 35 ℃. The product was washed with deionized water by centrifuging at 3500 rpm until pH ≥ 6. Then, the MXene was collected after centrifuging at the same rpm and proceed to lyophilize for subsequent use. Then, The MXene (90 mg) was dispersed in 60 mL ethylene glycol under sonicated for 30 min. Then the $FeCl_3 \cdot 6H_2O$ (0.27 g), 1, 6-hexylenediamine (0.972 g) and NaAc (1.08 g) were added into above-mentioned suspension under magnetic stirring for 1 h. The mixed liquid was transferred into the reaction kettle and kept at 200 °C for 7 h. The magnetic MXene was obtained by magnetic separation, then washed with ethanol and deionized water several times, and dispersed in deionized water for subsequent use. Finally, The MMP nanosheet was synthesized by physical coating. The magnetic MXene (30 mg) was dispersed in 10 mL PA solution (7%) under sonicated for 30 min. The MMP nanosheet was obtained by magnetic separation, then washed with ethanol and deionized water several times, and dispersed in deionized water for subsequent use.

Sample preparation.

Tryptic digestion of proteins: IgG or BSA (10 mg) was dissolved in 1 mL NH_4HCO_3 solution (50 mM, pH 8.2), and boiled at 100 °C for 15 min. After cooling to room temperature, trypsin (250 μg) was added and incubated at 37 °C for 16 h. Finally, diluted with NH_4HCO_3 solution (50 mM, pH 8.2) to the desired concentration.

Blood samples: All blood samples were drawn into vacutainer tubes by venipuncture and clotted at room temperature within 1 h. Serum samples was collected at 3000 × g for 10 minutes of centrifugation from the blood and immediately stored at -80 ℃ for further analysis.

Selective enrichment of peptides from standard protein tryptic digests and serum.

Selective enrichment of peptides from standard protein tryptic digests: The MMP nanosheet (0.3 mg) were dispersed and equilibrated by 200 μL trypsin digestion solution (IgG or IgG and BSA), incubated at 37 ℃ for 45 min, and collected the peptide-nanosheet complexes by magnetic separation, washed three times with enrichment buffer. Enrichment of HPs from protein tryptic digests diluted to different concentrations was carried out using the same method, with the only difference being the concentration of protein tryptic digests added. Then, the peptides captured were released by eluent (10 μL), shook vigorously at 37 ℃ for 30 min. The eluate was collected by magnetic separation for MALDI-TOF/TOF MS analyses.

Selective enrichment of peptides from serum: The MMP nanosheet (0.3 mg) were dispersed and equilibrated by 200 μL ACN/H_2O/TFA (90: 9.9: 0.1, v/v/v) containing of serum (1%), incubated at 37 ℃ for 45 min, and collected the peptide-nanosheet complexes by magnetic separation, washed three times with enrichment buffer. Then, the peptides captured were released by eluent ACN/H_2O/TFA (1.9: 98: 0.1, v/v/v) (10 μL), shook vigorously at 37 ℃ for 30 min. The eluate was collected by magnetic separation for MALDI-TOF/TOF MS analyses.

Database searching.

The data of mass spectrometry analysis were RAW files, and the database identification and quantitative analysis were performed by MaxQuant (1.6.5.0), the human database downloaded from uniprot (http://www.uniprot.org/).

GO analysis: First mapping all differential proteins to each term in the GO database (http://www.geneontology.org/), calculating the number of proteins in each term. Then, applying hypergeometric test to find out the differential protein's significantly enriched GO entry, and comparing with background all proteins. The formula to calculate:

$$p_{value} = 1 - \sum_{j=0}^{x-1} \frac{\binom{M}{j} \quad \binom{N-M}{n-j}}{\binom{N}{n}}$$

Where N is the number of proteins with GO annotation information in all proteins, n is the number of differential proteins in N, M is the number of proteins annotated to a GO entry in all proteins, x is the number of differential proteins annotated to a GO entry, with $p_{value} \leq 0.05$ as the threshold.

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis: Mapping the information of differentially expressed proteins to the KEGG database, and hypergeometric test was used to identify the pathways that were significantly enriched in differentially expressed genes compared with the entire genome backview.

Protein protein interaction network: Based on string (http://string.embl.de/) database and cytoscape software for interaction network analysis of differential proteins.

Patient characteristics.

A total of 120 serum samples were collected from West China Hospital of Sichuan University. All samples were anonymized, and only the gender, age and cancer-related lab results and pathological diagnosis were recorded. The 120 serum samples including healthy control (HC, 50) and hepatocellular carcinoma patients (HCC, 70) was randomly divided into discovery cohort and test cohort. All patients were verified with medical imaging, biochemical tests and pathological results, and they had no other

major diseases. All healthy controls had normal biochemical profiles (including serum tumor antigens), negative ultrasound/radiological findings and no previous history of any type of cancer.

All the subjects kept an emotional balance without heavy physical exercise and overnight fasting (more than 8 hours) before the blood collection. All relevant ethical regulations were complied with. This study was approved by the Biomedical Ethics Committee of West China Hospital, Sichuan University (Ethic number of 2019-203). All subjects provided informed consent to participate in the study and approved the use of their biological samples for analysis. All experiments were performed following institutional guidelines, in compliance with relevant laws.

MALDI-TOF/TOF MS detection.

In this work, the MALDI-TOF/TOF MS analysis was conducted on the positive ion mode of Autoflex Max mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with both a smartbeam-II laser (355 nm). The signal acquisition mode is set to positive reflection mode with a frequency of 200 Hz and an acceleration voltage of 20 kV. A random walk of 500 shots at raster spot and 50 different spots were measured for each individual sample, therefore, 25000 satisfactory shots were obtained and a range of molecular weights collected from 900 to 3500. The molecular weight of the instrument was calibrated with standard small molecules before each use.

Typically, 1 μL of mixture of the above eluate and DHB was pipetted onto a target plate, after being dried at room temperature. Then, samples were analyzed on the MALDI-TOF/TOF MS. For each sample, five independent experiments were carried out for ensuring the reproducibility and reliability of detection results.

Serum peptide extraction.

The original image was formatted by the software that comes with the MS instrument, and then the pretreatment of peak extraction was then carried out on the raw mass spectra using custom-built code under the same environment (Python 3.10 and PyCharm 2022.1).

Construction of machine learning models.

The FNN model was designed with four layers of neural networks including input and output layers as an overall architecture. The feature dimension in the sample was used as an input node number of the neural network, and the disease label was used as an output node. Two hidden layers were constructed in the middle, and the number of neurons in two hidden layers was 2048 and 256, respectively. The Softmax was used as the activation function in the output layer, which could be used to convert the output value of classification into a probability distribution ranging from 0 to 1, and it was estimated by the following equation:

$$Softmax(x_i) = \frac{exp(x_i)}{\sum_{j=1}^{K} exp(x_j)}$$

x_i is the output value of the ith node, and K is the number of categories classified.

ReLU was used as the activation function in all hidden layers, and it was defined as:

$$ReLU(x_i) = \begin{cases} x_i & if\ x_i \geq 0 \\ 0 & if\ x_i < 0 \end{cases}$$

The BatchNorm1d and Dropout were applied to prevent model overfitting. The cross-entropy loss function was chosen to evaluate the loss via the following formula:

$$Loss = \frac{1}{N}\sum_{i=1}^{N} Loss_i = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{M} y_{ic} log(\tilde{y}_{ic})$$

N is the sample size in the current batch, M is the number of categories, and y was is current input sample label.

Support vector machine (SVM), the basic mathematical model was specified below:

$$max_{w,b} \frac{1}{||w||}$$

$$s.t.\ y_i\left(w^T \cdot x_i + b\right) - 1 \geq 0, i = 1,2,\cdots,n$$

(x_i, y_i) is a data set of sample points, (w, b) are the hyperplane parameters.

Random Forest (RF) categorizes samples through multiple decision trees. For a new sample to be categorized, the likelihood result for the sample in each tree in the forest built from the training cohort was estimated, and the maximum likelihood result was

selected by voting.

The K-Nearest Neighbor (KNN) model was built from a training data set to allocate K instances into a known category based on their similarity. The specific formula was given below:

$$\tilde{y}_i = argmax_{c_j} \sum_{x_i \in N_k(x)} I(y_i = c_j)$$

$$i = 1,2,\cdots,N; j = 1,2,\cdots,K$$

Logistic regression (LR) model was built from estimation of the maximum likelihood and the gradient descent. The specific formula was specified below:

$$h_w(x) = g(w^T x) = \frac{1}{1 + e^{-w^T x}}$$

$$P(y \mid x;w) = (h_w(x))^y (1 - h_w(x))^{1-y}$$

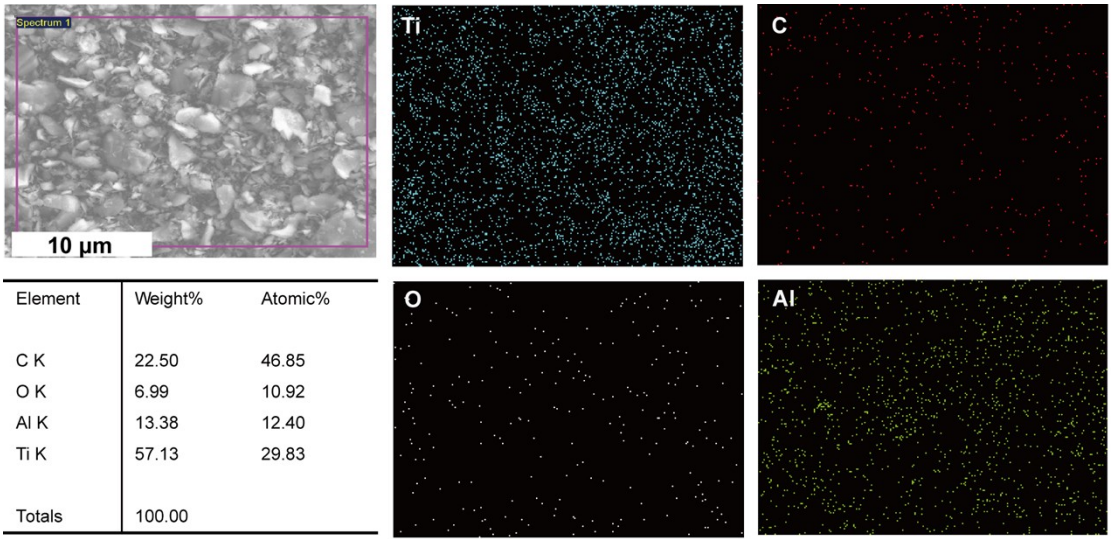
The machine learning models was implemented by pytorch (version 1.13.1) and torchvision (version 0.14.1). Numpy (version 1.23.5), scikit-manifold (version 1.2.1) and pandas (version 1.5.3) were used for data processing, scikit-learn (version 1.2.1) and scipy (version 1.10.0) were used to evaluate machine learning methods. The ROC and AUC values were obtained by scikit-learn (version 1.2.1). The ROC and AUC, and the confusion matrix for train loss & test were plotted by matplotlib (version 3.7.0).

Characterization methods.

A vibrating sample magnetometer (VSM, Model PPMS, Quantum Design Company, USA) was employed to measure the magnetization of the samples with field strength varying from 0 to 20000 Oe at 300 K. The mass loss of the samples was analyzed at temperatures ranging from 35 to 500 ℃ at the heating rate of 10 K min^{-1} by simultaneous thermal analysis (STA449 C Jupiter, NETZSCH). Fourier transform infrared spectra were obtained by a spectrometer (FTIR, PE spectrometer) with wavenumber in the range of 400–4000 cm^{-1}. The zeta potential and size distribution of samples were obtained by dynamic light scattering (DLS, Zetasizer NanoZS90,

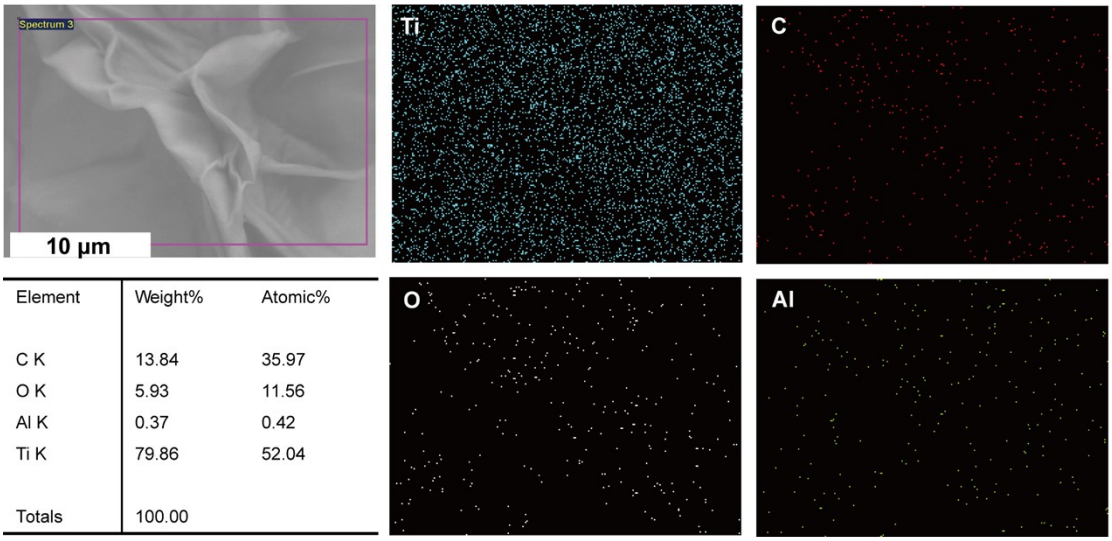
Malvern Instruments Ltd, UK). The morphologies of the samples were observed by scanning electron microscopy (SEM) with an energy disperse spectroscopy (EDS) (Hitachi S-4800, Japan), atomic force microscope (AFM) (Bruker, Icon, Germany) and transmission electron microscopy (TEM, JEOL, JEM-100CX, Japan). The structure of the products was characterized by X-ray diffractometer (XRD) with Cu-Kα radiation (PANalytical, Empyrean, Netherlands) from 3 ° to 80 °.

Figure



Element	Weight%	Atomic%
C K	22.50	46.85
O K	6.99	10.92
Al K	13.38	12.40
Ti K	57.13	29.83
Totals	100.00	

Fig. S1. SEM EDS of $Ti_3Al_2C_x$ MAX, a precursor of MXene. The middle layer Al in $Ti_3Al_2C_x$ MAX was corroded to obtain the MXene nanosheet.



Element	Weight%	Atomic%
C K	13.84	35.97
O K	5.93	11.56
Al K	0.37	0.42
Ti K	79.86	52.04
Totals	100.00	

Fig. S2. SEM EDS of MXene. The Al content decreased from 13.38% in the original $Ti_3Al_2C_x$ MAX sample to 0.37%, suggesting that the Al intermediate layer was well etched, and the MXene nanosheet was successfully exfoliated.

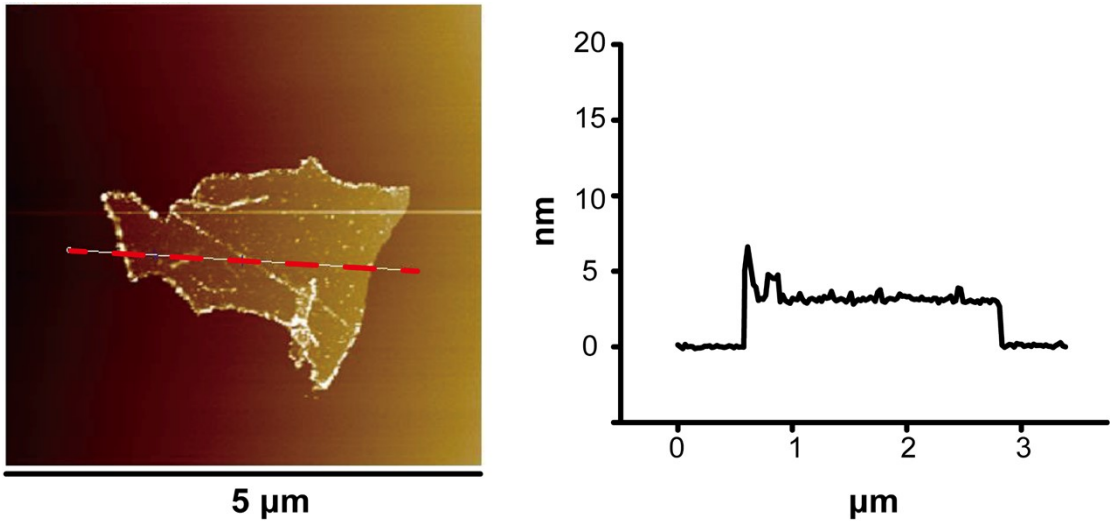
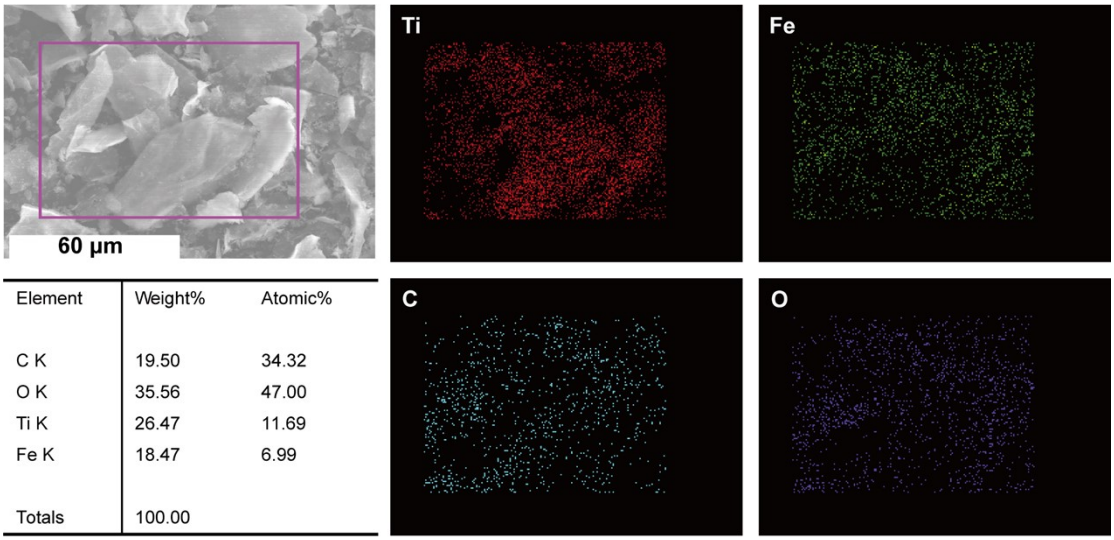


Fig. S3. AFM of MXene. The thickness of a single piece of MXene was about 3-4 nm.



Element	Weight%	Atomic%
C K	19.50	34.32
O K	35.56	47.00
Ti K	26.47	11.69
Fe K	18.47	6.99
Totals	100.00	

Fig. S4. SEM EDS of MM. The Fe content is 18.47%, suggesting that iron oxide was successfully introduced to MXene.

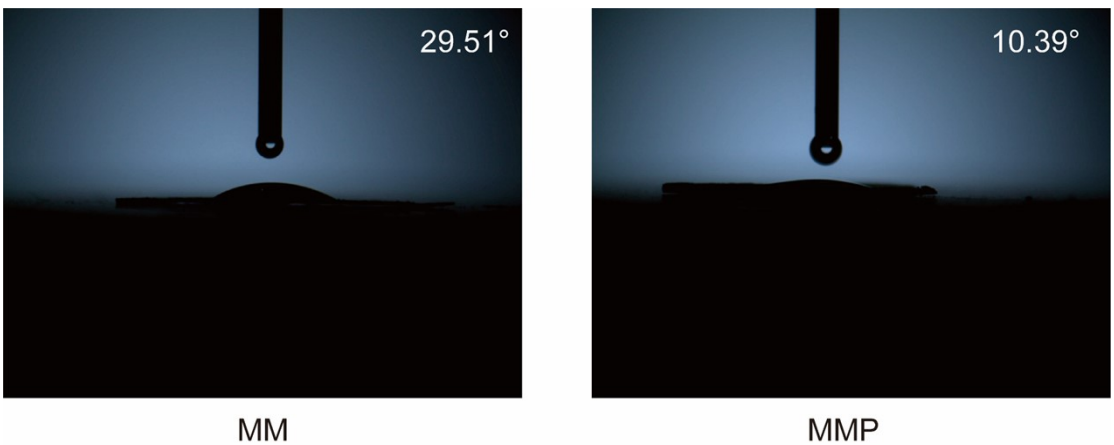


Fig. S5. Water contact angle of MM and MMP.

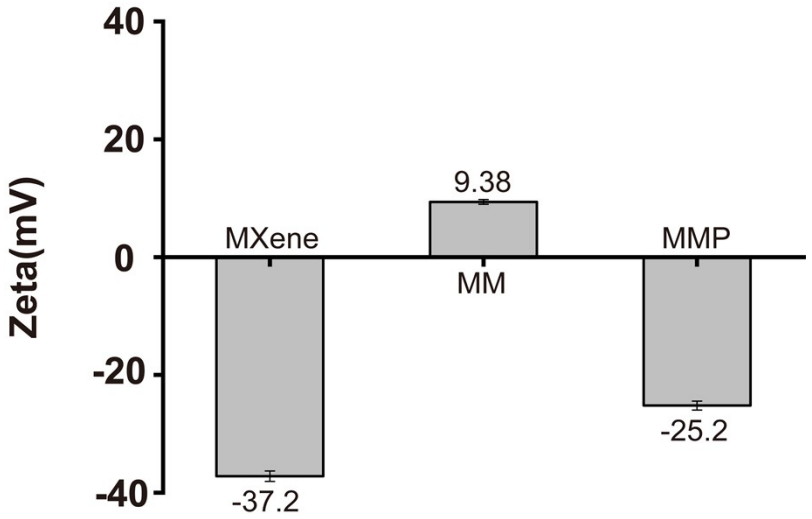


Fig. S6. Zeta potentials of MXene, MM and MMP.

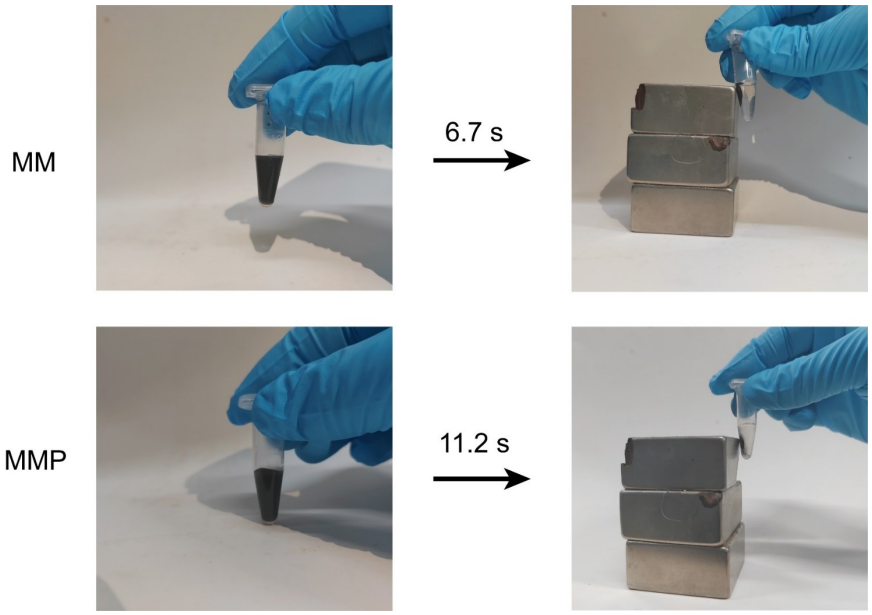


Fig. S7. Magnetic separation time of MM and MMP.

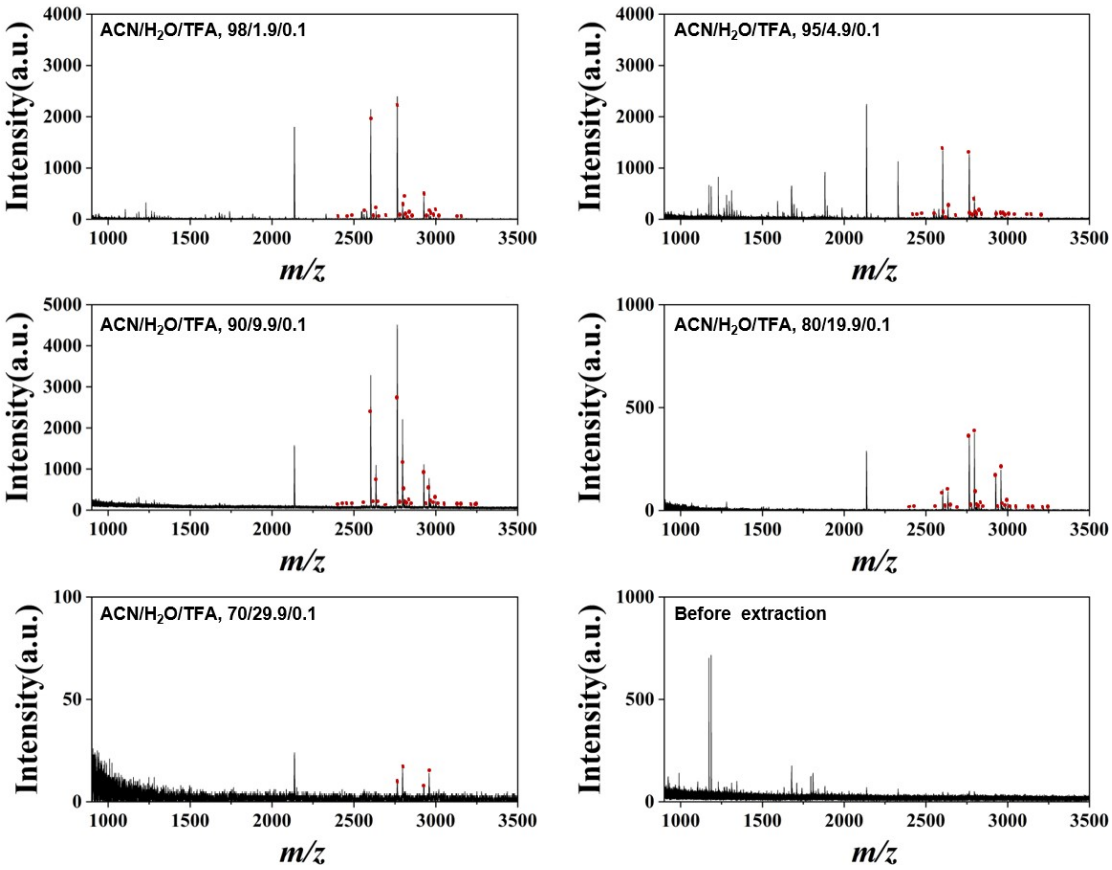


Fig. S8. The enrichment conditions screening of MMP selective adsorption HPs, elution conditions ACN/H_20/TFA:1.9/98/0.1, IgG digestion solution concentration of 100 fmol, HPs were marked by red.

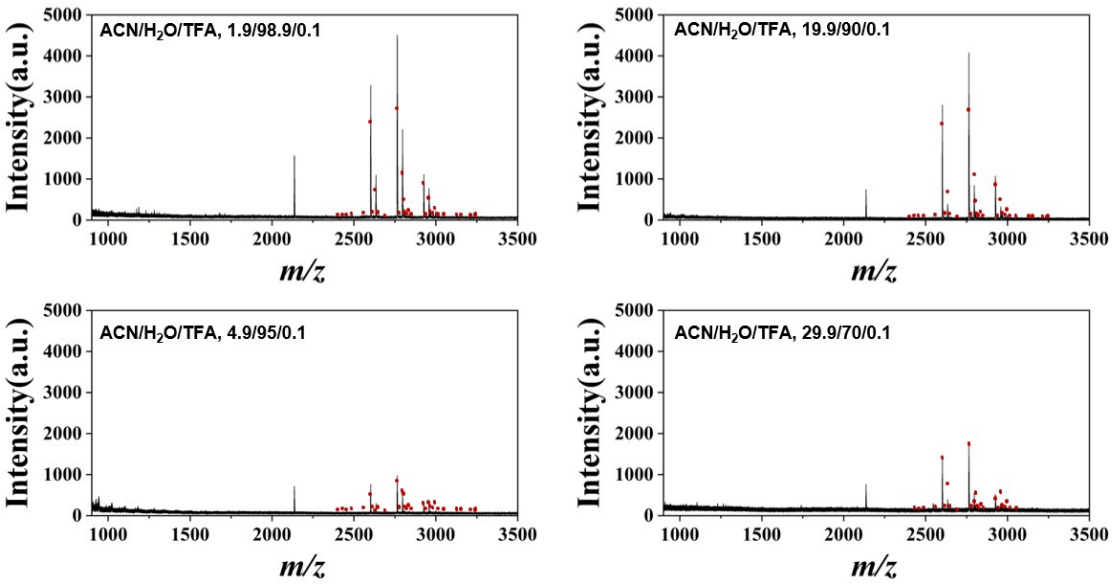


Fig. S9. The MMP selective adsorption HPs elution conditions screening, IgG digestion solution concentration of 100 fmol, HPs were marked by red dots.

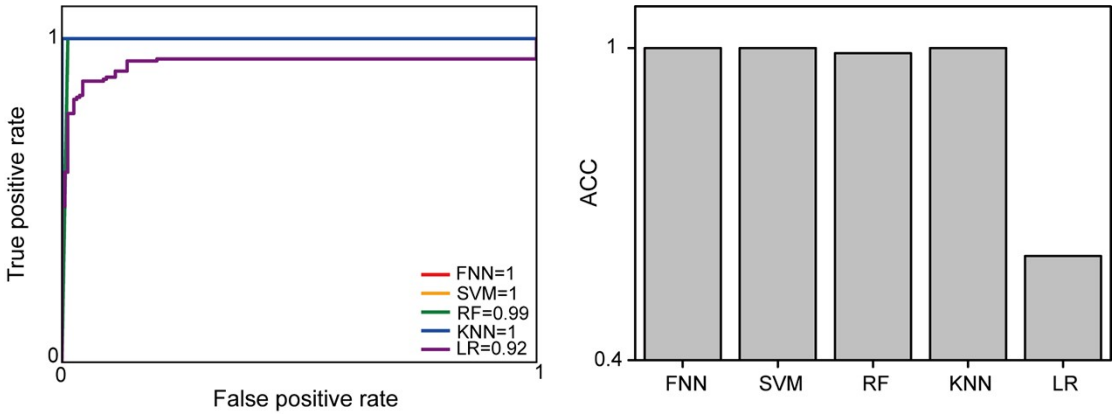


Fig. S10. ROC curves and ACC value of HC/HCC for the training cohort.

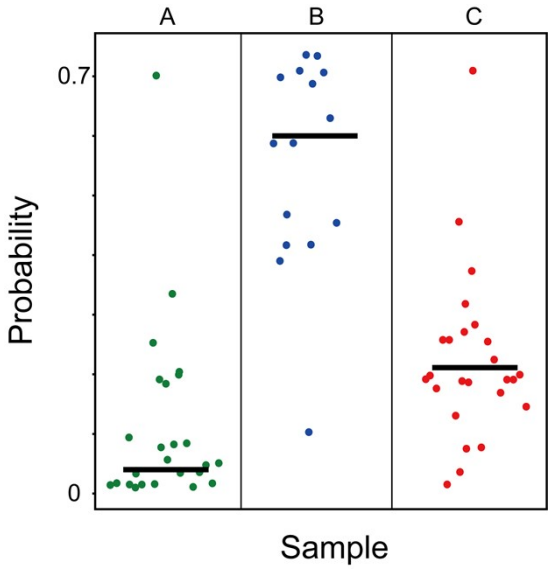


Fig. S11. The probability of A, B and C stages through the machine learning model

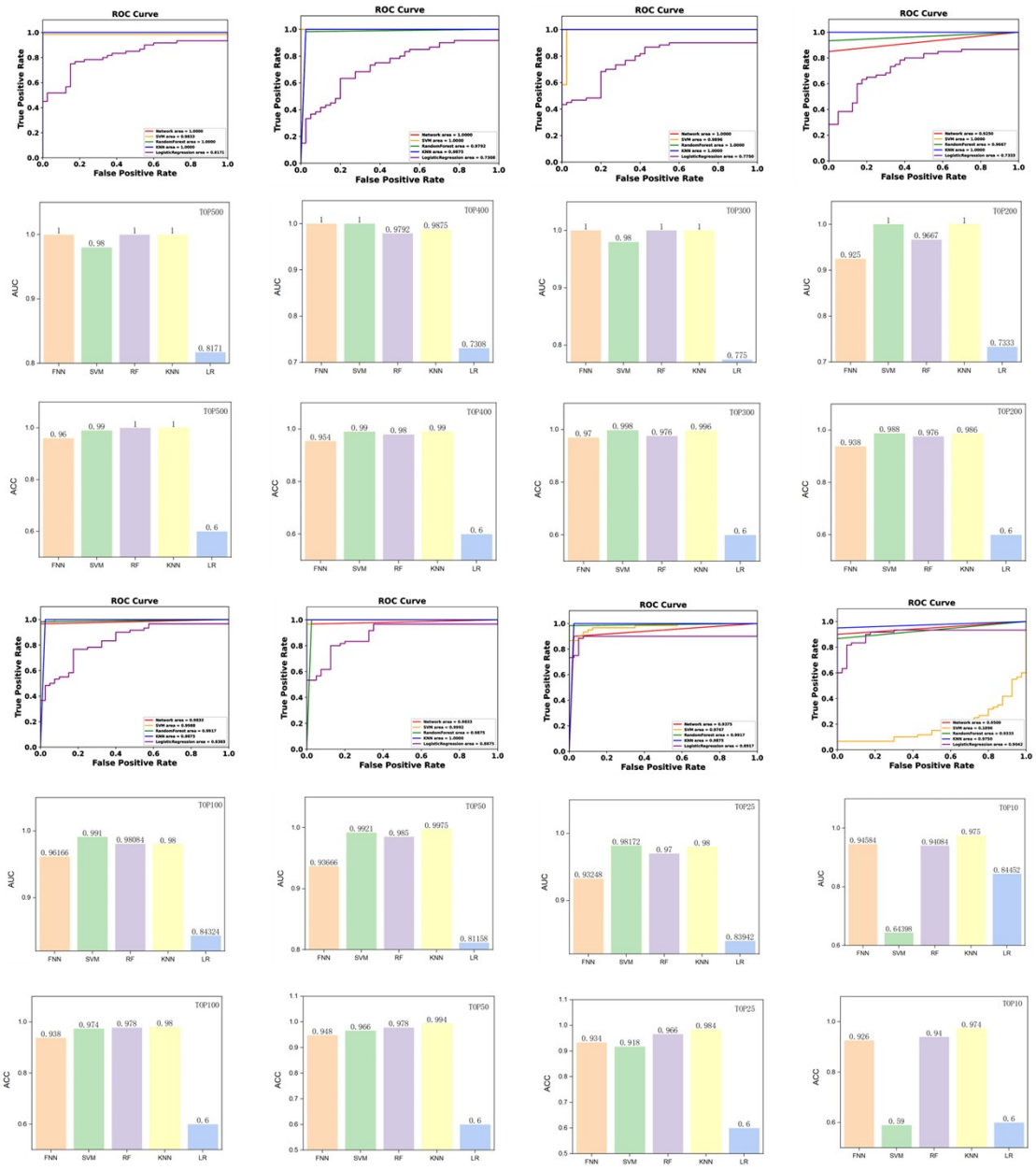


Fig. S12. The ROC curves, AUC value, and ACC value of HC /HCC by the top 500, 400, 300, 200, 100, 50, 25, and 10 weight, respectively.

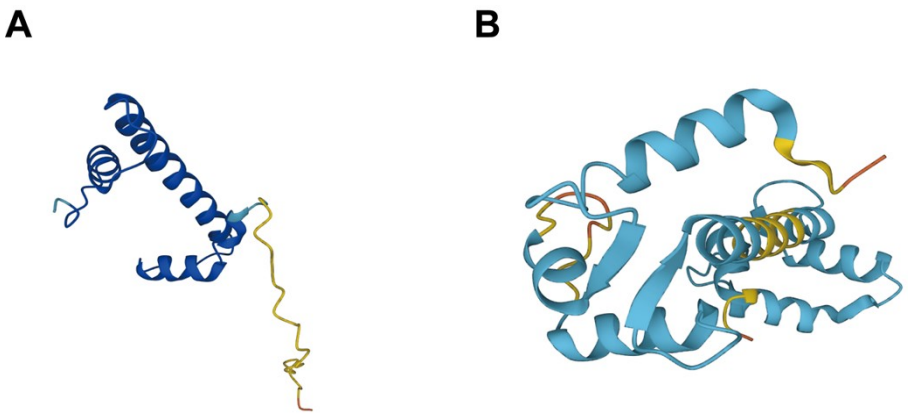


Fig. S13. The structure of (A) Q0VAS5 and (B) Q53H37. (Source: AlphaFold Database)

Table

Table. S1 DES of MMP.

Element	Weight%	Atomic%
C	13.48	24.35
O	40.35	54.73
P	5.07	3.55
Ti	21.71	9.84
Fe	19.39	7.54
Totals	100.00	

Table. S2 Observed molecular masses, proposed glycan compositions and peptide sequences of HPs enriched from IgG tryptic digests by the MMP. Hex, HexNAc, Fuc and NeuAc are the abbreviations of hexose, N-acetylhexosamine, fucose and N-acetylneuraminic acid, respectively. N& denotes the glycosylation sites.

Number	m/z	Glycan composition	Peptide sequence
1	2398.5	[Hex]3[HexNAc]3[Fuc]1	EEQFN&STFR
2	2431.1	[Hex]3[HexNAc]3[Fuc]1	EEQYN&STYR
3	2456.2	[Hex]3[HexNAc]4	EEQFN&STFR
4	2487.1	[Hex]3[HexNAc]4	EEQYN&STYR
5	2561.4	[Hex]4[HexNAc]3[Fuc]1	EEQFN&STFR
6	2602.0	[Hex]3[HexNAc]4[Fuc]1	EEQFN&STFR
7	2616.7	[Hex]3[HexNAc]4[Fuc]1	EEQFN&STFR
8	2633.1	[Hex]3[HexNAc]4[Fuc]1	EEQYN&STYR
9	2649.6	[Hex]4[HexNAc]4	EEQYN&STYR
10	2764.4	[Hex]4[HexNAc]4[Fuc]1	EEQFN&STFR
11	2778.5	[Hex]5[HexNAc]4	EEQFN&STFR
12	2795.7	[Hex]4[HexNAc]4[Fuc]1	EEQYN&STYR
13	2805.0	[Hex]3[HexNAc]5[Fuc]1	EEQFN&STFR
14	2813.0	[Hex]5[HexNAc]4	EEQYN&STYR
15	2822.2	[Hex]4[HexNAc]5	EEQFN&STFR
16	2837.0	[Hex]3[HexNAc]5[Fuc]1	EEQYN&STYR

17	2853.7	[Hex]4[HexNAc]5	EEQYN&STYR
18	2909.5	[Hex]4[HexNAc]4[NeuAc]1	EEQFN&STFR
19	2925.9	[Hex]5[HexNAc]4[Fuc]1	EEQFN&STFR
20	2942.4	[Hex]5[HexNAc]4[Fuc]1	EEQFN&STYR
21	2958.5	[Hex]5[HexNAc]4[Fuc]1	EEQYN&STYR
22	2966.2	[Hex]4[HexNAc]5[Fuc]1	EEQFN&STFR
23	2983.4	[Hex]5[HexNAc]5	EEQFN&STFR
24	3000.2	[Hex]4[HexNAc]5 [Fuc]1	EEQYN&STYR
25	3016.8	[Hex]5[HexNAc]5	EEQYN&STYR
26	3054.3	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQFN&STFR
27	3128.0	[Hex]5[HexNAc]5[Fuc]1	EEQFN&STFR
28	3162.0	[Hex]5[HexNAc]5[Fuc]1	EEQYN&STYR
29	3217.3	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQFN&STFR
30	3247.0	[Hex]4[HexNAc]4[Fuc]1	TKPREEQFN&STFR
31	3250.3	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQYN&STYR

Table. S3 Summary of clinical characteristics for patients and healthy controls.

Patient type	Number	Gender		Age	BCLC stage		
		Male	Female		A	B	C
HC	50	42	8	62 (53-81)	/	/	/
HCC	70	60	10	59 (41-97)	22	14	24
Total	120	108	18	61 (41-97)	22	14	24

Table. S4 Correlation coefficient of gender, age and the patient disease type.

	Gender	Age	Patient type
Gender	1		
Age	-0.10922	1	
Patient type	0.023669	-0.17018	1

Table. S5 Top 100 weighted HPs identification and protein matching.

MALDI-MS (m/z)	Ms/ms (m/z)	Sequence	Peptide matching to proteins	Weight in AI

1435.72	1433.7052	GNFHAVYRDDLK	P05109	0.054027
1229.94	1229.6616	NILTSNNIDVK	C9JF17	0.051316
1592	1591.773	AGELTPEEEAQYKK	Q53H37	0.051005
1301.46	1301.651	QRQEELCLAR	Q5VY30	0.050633
1324.19	1324.7463	DNIQGITKPAIR	Q0VAS5	0.049638
1082.71	1083.5421	NSVNSHTIGR	A0A494C0J7	0.047987
1055.46	1057.6132	LASLEELKR	P15924	0.04782
1165.95	1165.584	HQTVPQNTGGK	B4E1B2	0.047506
1914.91	1915.972	YGPIVDVYVPLDFYTR	Q5JRI1	0.045471
1319	1319.6721	ELAVQIYEEAR	O00571	0.045449
966.64	965.40589	ECLQTCR	P02760	0.045168
1118.5754	1078.5407	LEYDDLRR	P15924	0.045091
1817.36	1816.7971	EGTCPEAPTDECKPVK	B4E1B2	0.044994
1265.16	1265.6139	TGETNDFELLK	A0A024QYU7	0.044555
1329.41	1329.6605	DEVFEYIIFR	I3L4Q1	0.043897
1389.66	1388.6896	KQELSEAEQATR	V9HWA9	0.043844
901.67	901.4076	VQQPDCR	A0A3B3ISA6	0.043398
1144.37	1144.5183	LCNIEPDER	Q04756	0.043377
2696.23	2696.2619	RGGPPFAFVEFEDPRDAEDAVYGR	Q59FA2	0.043241
1423.12	1422.7355	NPNLPPETVDSLK	C9JF17	0.043178
1005.3	1004.5655	KDLQNFLK	B2R4M6	0.043166
1958.63	1957.9745	GHYTEGAELVDSVLDVVR	P07437	0.043075
1514.11	1513.6719	SLEEEKYDMSGAR	P31944	0.042991
969.67	970.41783	YTACETAR	C9JV37	0.042857
1476.71	1476.682	AEFHHSIMSQYK	Q02413	0.04268
2493.58	2494.2969	FQIGDYLDIAITPPNRAPPPSGR	A0A384ME24	0.042586
1700.44	1700.8985	AVFVDLEPTVIDEVR	Q53GA7	0.042545

1276.99	1275.6248	WEAEPVYVQR	A0A140VK00	0.042346
2408.08	2408.2012	FDGALNVDLTEFQTNLVPYPR	Q53GA7	0.042261
1911.55	1911.9902	SSPVVIDASTAIDAPSNLR	P02751	0.042111
1480.61	1480.7021	CFEGFGIDGPAIAK	A0A8Q3SIA1	0.041978
1039.74	1038.4804	EQYNMLGGK	Q02413	0.041827
972.79	971.51484	ASGQNLNLR	B4DRV1	0.041801
1192.47	1193.5677	QEYVLNESGR	B4DRV1	0.0416
984.38	985.5808	KQLVEIEK	P00738	0.041495
1021.53	1021.5193	ATVVYQGER	B4DPN0	0.041494
1167.82	1168.634	LPEATPTELAK	A0A1B1CYC5	0.041434
1118.4	1118.5568	ESSNVVVTER	Q02413	0.041404
1122.48	1123.5947	SGMYVVIEVK	H7BZK5	0.041282
903.42	904.42502	DAATDVASR	A0A494C0J7	0.04123
1416.06	1415.6834	WYVDGVEVHNAK	A0A5J6KJ24	0.041217
3100.33	3101.4003	FWEVISDEHGIDPTGTYHGDSDLQLDR	P07437	0.041108
1233.96	1233.6982	EHVAHLLFLR	P19652	0.041103
1280.97	1282.5652	WQEEMELYR	A0A024R3E3	0.041081
2172.32	2171.0503	SPVGVQPILNEHTFCAGMSK	P00738	0.041034
1869.23	1868.9844	EVVLTQSPGTLSLSPGER	A0A5C2FUV8	0.040836
1410.01	1410.7831	GALQNIIPASTGAAK	A4UCT1	0.040756
1491.98	1492.6906	FNTANDDNVTQVR	B4DWK8	0.040755
1082.12	1083.4655	DTGDIFCTR	Q9HB00	0.040699
1478.2	1477.7777	ELTSELKENFIR	X6R8F3	0.040621
1280.93	1279.75	VVLEGGIDPILR	P05164	0.040584
1538.87	1538.8417	LLESGGGLVQPGGSLR	A0A5C2GE41	0.040556

1222.28	1221.5044	EGDDDRTVCR	B4DW11	0.040528
1239.66	1239.6136	SDVVYTDWKK	V9HWF6	0.040471
1022.89	1023.5059	FISLGEACK	V9HWA9	0.040394
3675.26	3676.7573	SVEGWILFVTGVHEEATEEDIHDKFAEYGEIK	A0A023T787	0.040299
1445.97	1445.678	FPTDQLTPDQER	P05164	0.040245
1659.27	1659.7563	EHAVEGDCDFQLLK	B7Z556	0.040235
1449.32	1450.6762	AYLEEECPATLR	A0A140VK00	0.040234
1227.84	1227.6836	TVIAQHHVAPR	Q6ZVX7	0.040206
1203.89	1203.603	EGGSMARQLQK	A0A3B3IU58	0.040183

A total of 62 of the top 100 HPs were identified, corresponding to a total of 48 proteins.

References

1. X. Guo, N. Li, C. Wu, X. Dai, R. Qi, T. Qiao, T. Su, D. Lei, N. Liu, J. Du, E. Wang, X. Yang, P. Gao and Q. Dai, *Adv. Mater*, 2022, **34**, 2201120.