Supplementary Materials

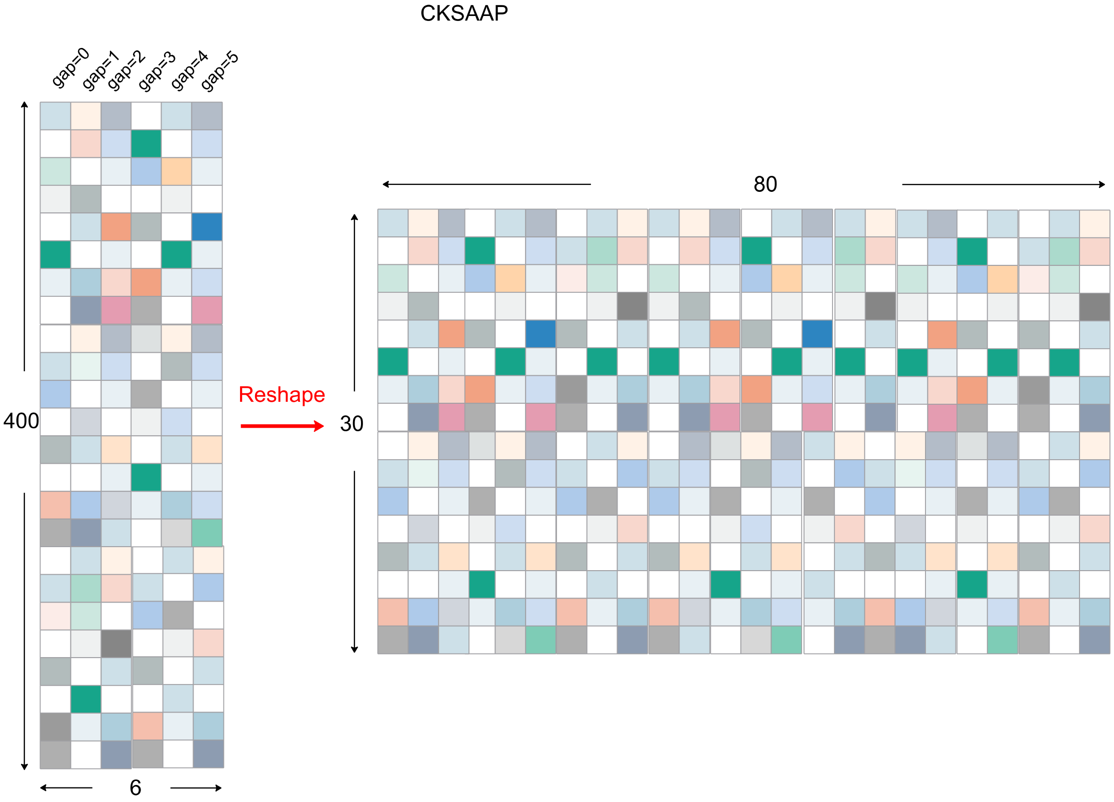
Supplementary Methods: other inputs compared in Section 3.5

CKSAAP

The CKSAAP feature calculates the frequency of amino acid pairs separated by any *k* residues (*k*=0, 1, 2, …, 5). Taking *k*=0 as an example, there are 400 0-spaced residue pairs (AA, AC, AD, …, YY). The feature vector can then be defined as:

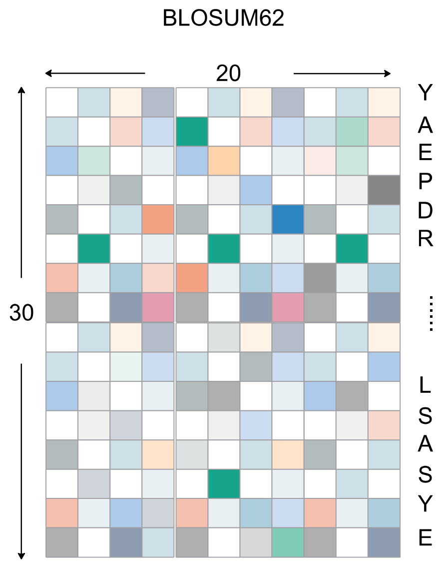
$$(\frac{N_{AA}}{N_{total}},\frac{N_{AC}}{N_{total}},\frac{N_{AD}}{N_{total}},\dots,\frac{N_{YY}}{N_{total}})_{400}$$

where N_{AA} represents the number of times of residue pare AA appears in the protein sequence, and so on. N_{total} is the total number of *k*-spaced residue pairs in the protein. For *k*=0, 1, 2, 3, 4 and 5, $N_{total} = \text{length}(P) - 1, \text{length}(P) - 2, \text{length}(P) - 3, \text{length}(P) - 4, \text{length}(P) - 5, \text{length}(P) - 6$. $\text{length}(P)$ represents the length of protein P. In this study, the CKSAAP encoding was performed using k = 0, 1, 2, 3, 4 and 5. Accordingly, CKSAAP feature was encoded as a 6×400 matrix. For consistency with the other features, we reshaped this matrix into an 80×30 matrix as follows:



BLOSUM62

A BLOSUM62 matrix comprising of *m*×*n* features is used to represent each protein in a training dataset, where *n* denotes the protein length and *m*=20. Thus, in this study, the BLOSUM62 was encoded as a 20×30 matrix as follows.

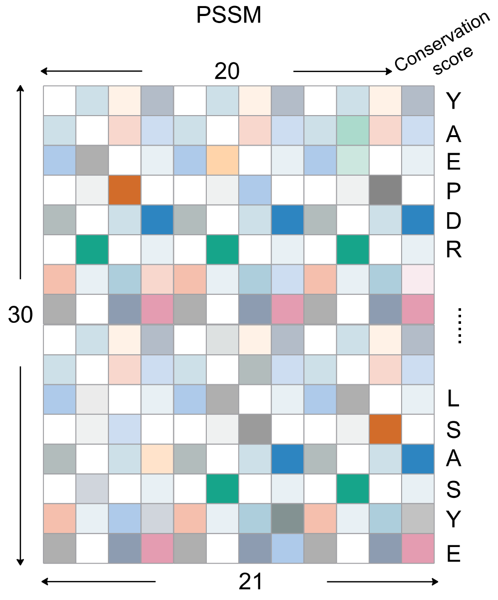


PSSM and conservation scores

PSSM features were extracted from the Position-Specific Scoring Matrix (PSSM) profile. We obtained the PSSM profile by running PSI-BLAST against the UniRef90 database. Each amino acid is represented by a 20-dimensional vector. We also computed conservation scores from the PSSM generated by PSI-BLAST as:

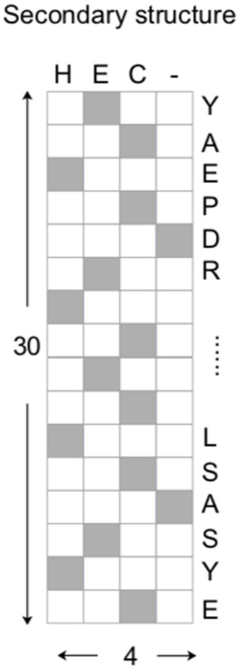
$$Score_i = -\sum_{j=1}^{20} p_{i,j} log_2 p_{i,j}$$

where $p_{i,j}$ is the frequency of amino acid j at position i. In this study, we encoded the PSSM values and the conservation scores into a 21×30 matrix as follow:



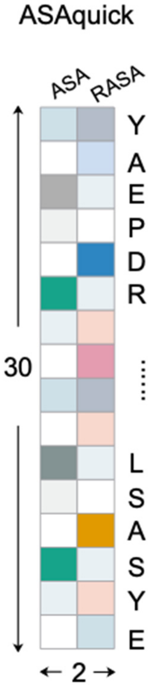
Predicted secondary structure

We employed the PSIPRED to predict the secondary structure. In this study, the predicted secondary structure was encoded into a 4×30 matrix as follow:



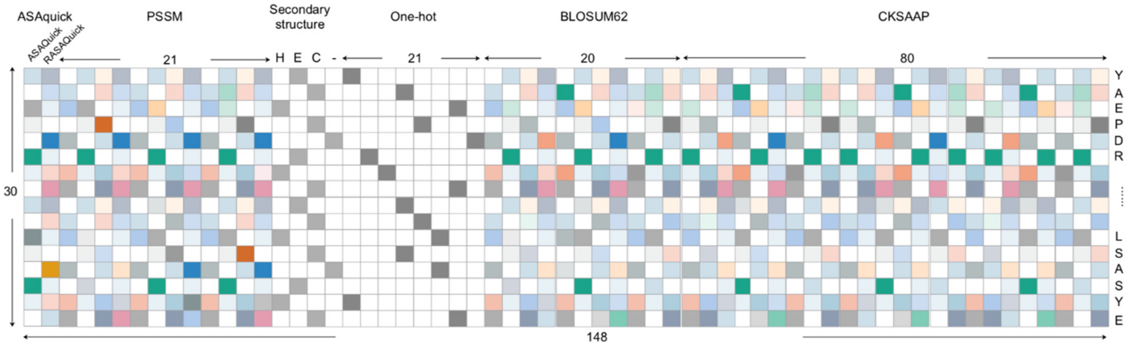
Predicted solvent accessibility

We employed the ASAquick to predict solvent accessibility including raw solvent accessibility (ASA) and relative accessibility (RASA). In this study, the predicted solvent accessibility values were encoded into a 2×30 matrix as follow:



All features

We also combined all features together, including CKSAAP, BLOSUM62, one-hot, secondary structure, PSSM, conservation score, and solvent accessibility, to compare the model performance. The combined input was encoded into a 148×30 matrix as follow:



Supplementary Results

Contents

Figures:

Tables:

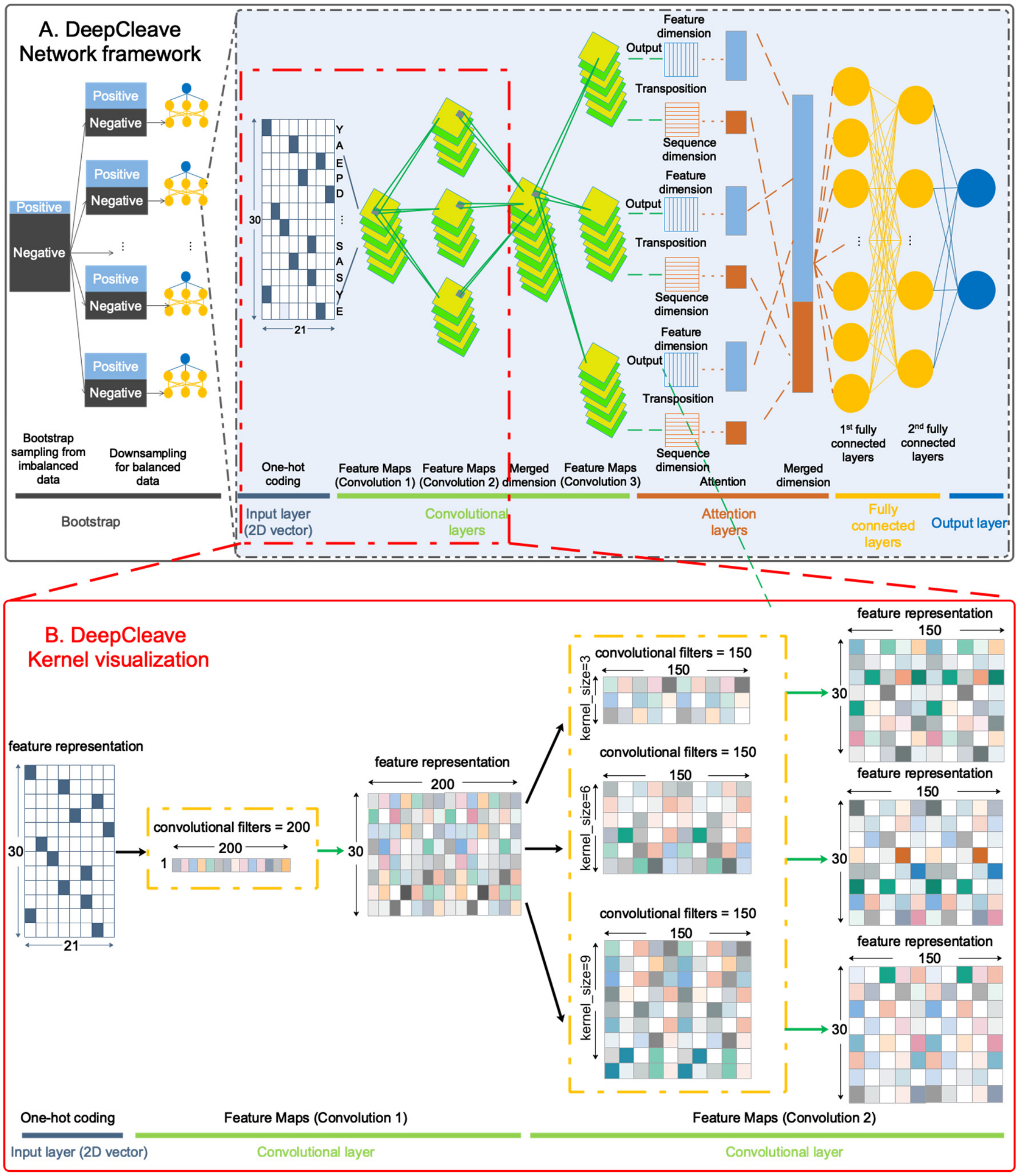


Figure S1. The detailed architecture of DeepCleave. Panel A shows the topology of the deep CNN. Panel B visualizes three different kernel sizes used in the first and the second convolutional layers. We use the first two convolutional layers as an example – note that the architecture of the third layer is similar to the first two layers.

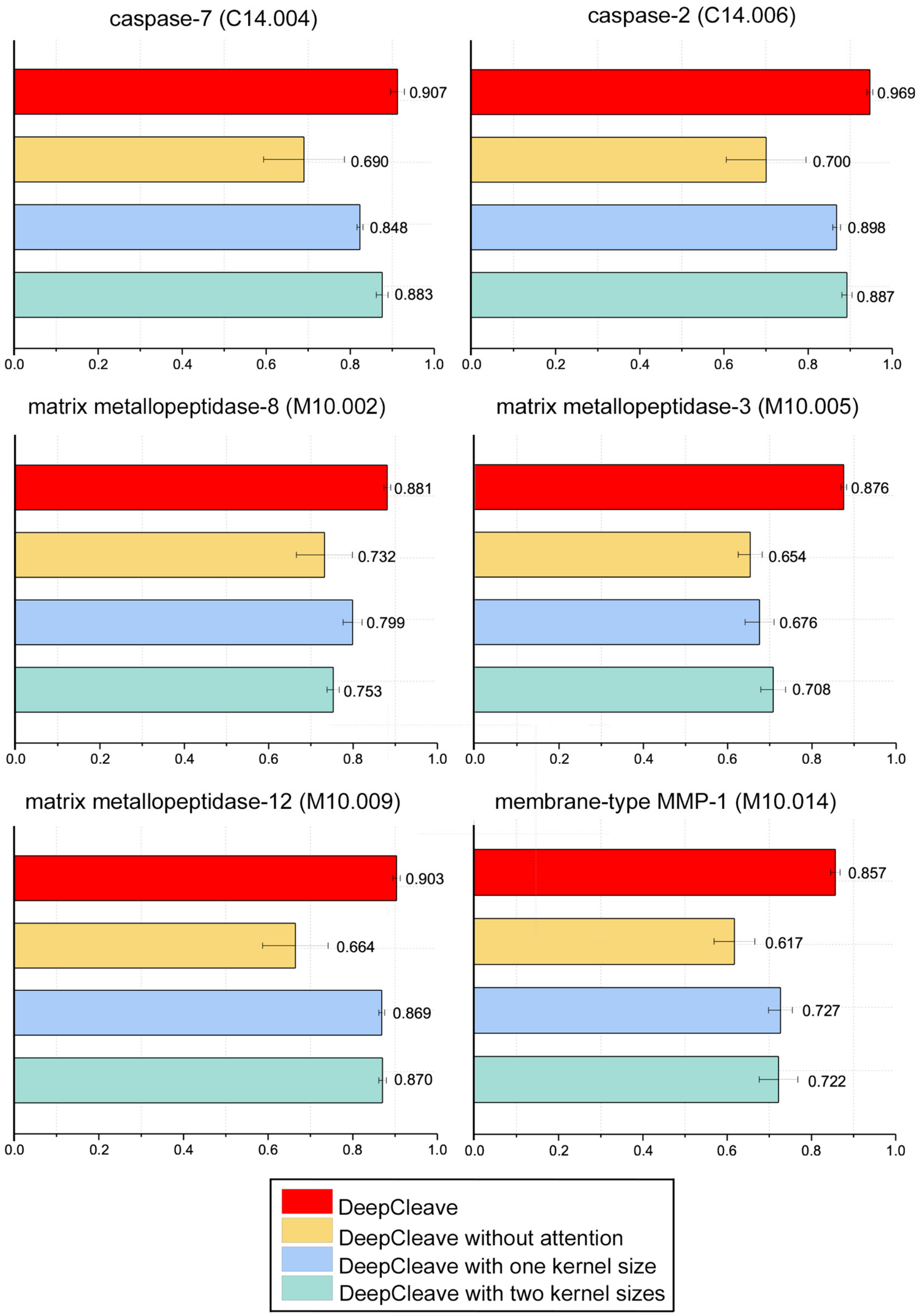


Figure S2. Comparison of the predictive performance for the four models considered in the ablation study on the test dataset using five-fold cross-validation. The results concern six different proteases: caspase-7, caspase-2, MMP-8, MMP-3, MMP-12 and membrane-type MMP-1; the identifier inside the brackets (e.g. "C14.005") is the protease ID in MEROPS. The following models are included: DeepCleave, DeepCleave with-out attention layer, DeepCleave with only one kernel size in the second and the third convolutional layer, and DeepCleave with two kernel sizes in the second and the third convolutional layer.

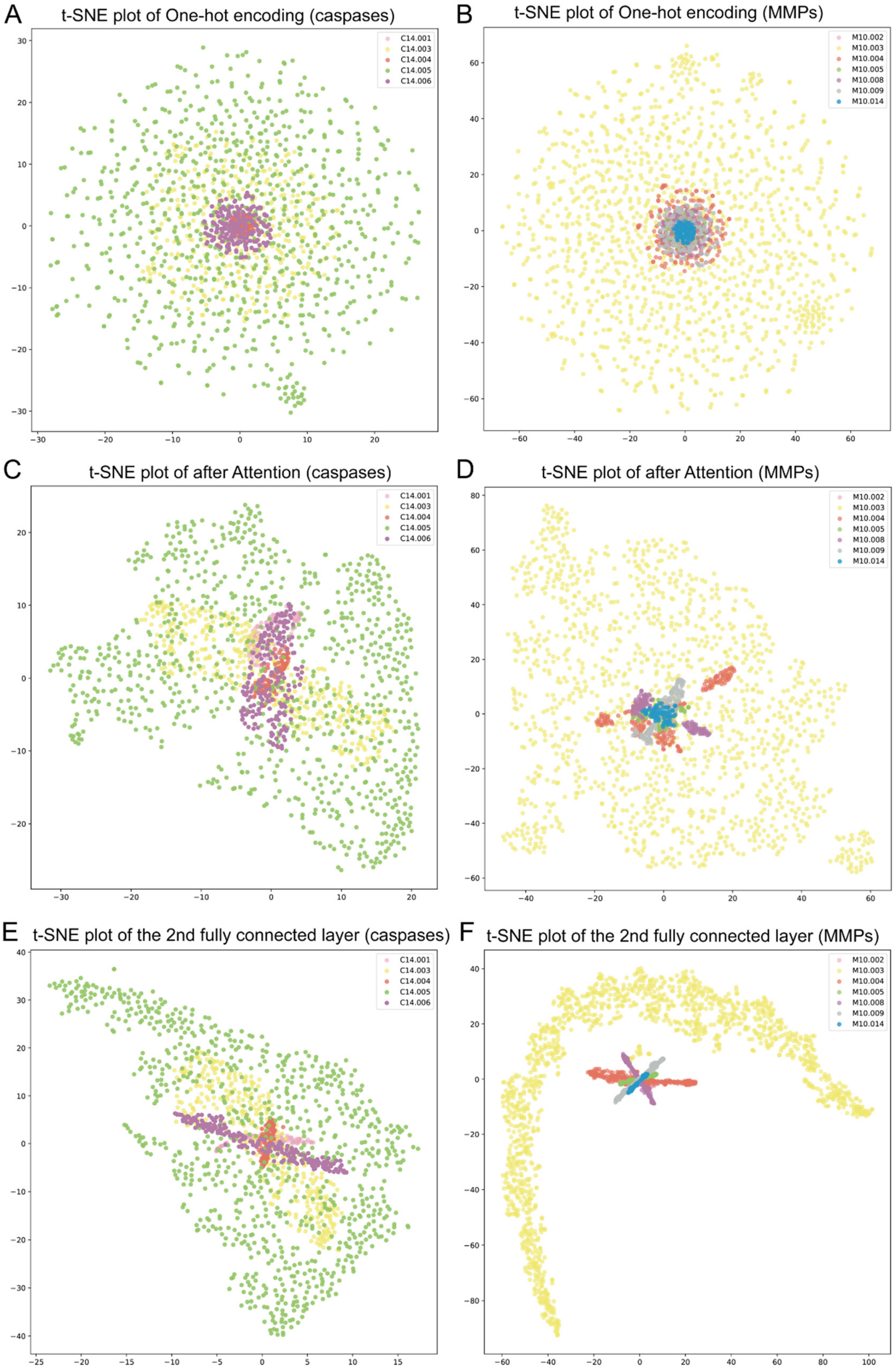


Figure S3. t-SNE plots of the input one-hot encoding (panels A and B), feature representation after the attention layer (panels C and D), and the feature representation of the 2nd fully connected layer (panels E and F) for the models for the caspases (on the left) and the MMPs (on the right).

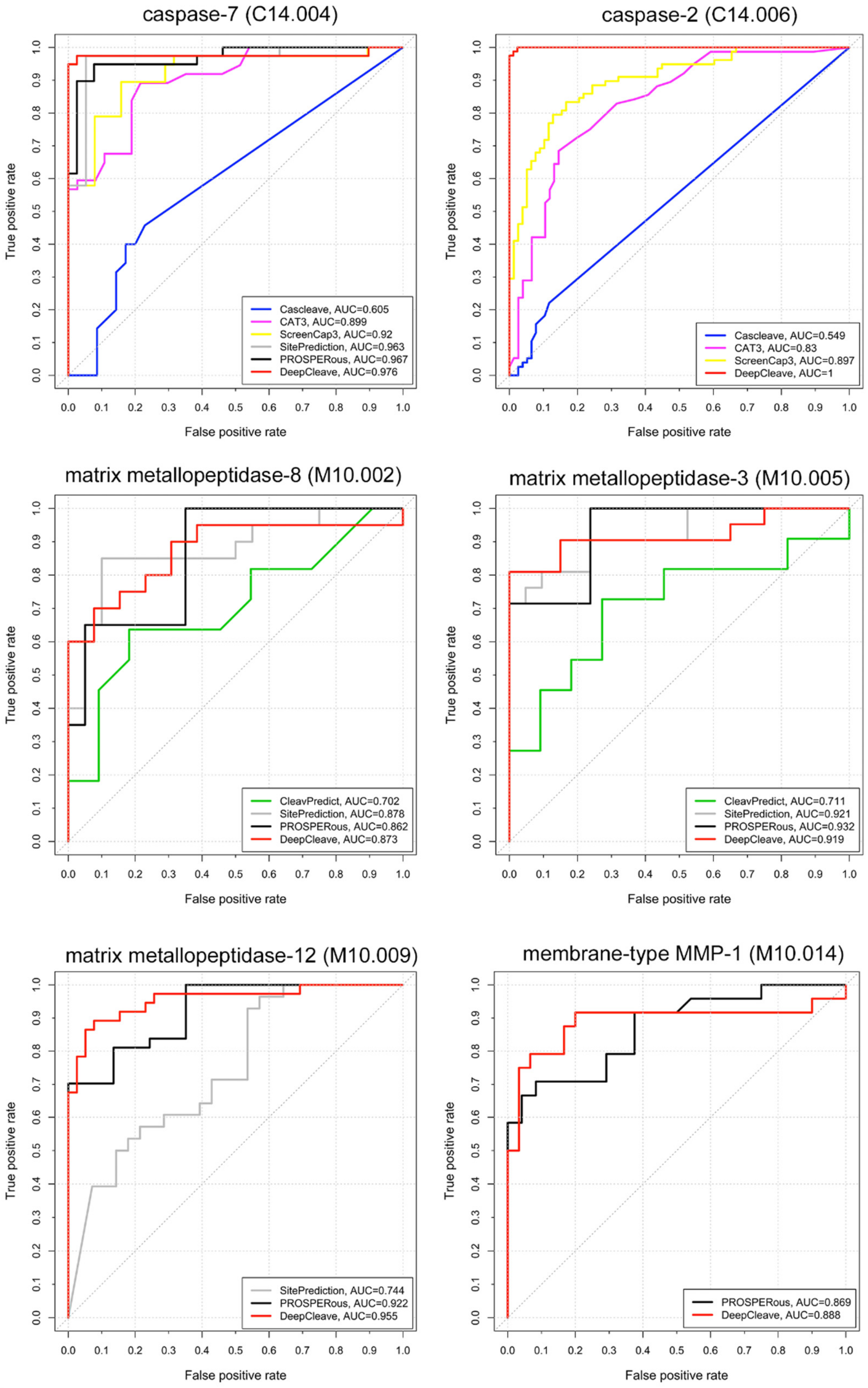
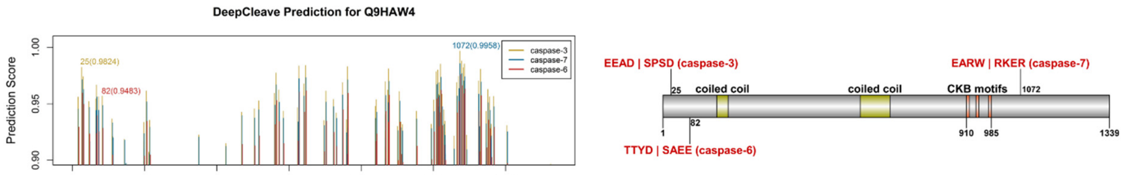
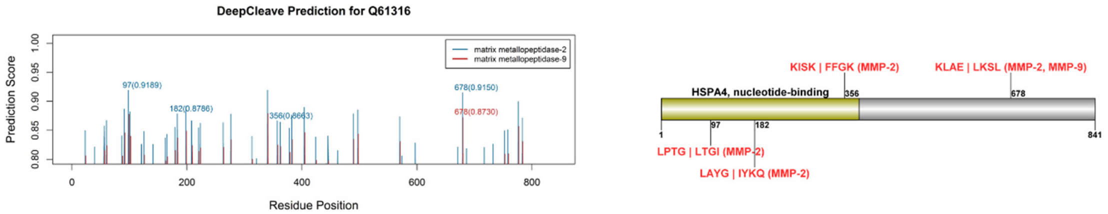


Figure S4. ROC curves and corresponding AUC values generated by seven considered protease cleavage site predictors (DeepCleave, Cascleave, SitePrediction, CleavPredict, CAT3, ScreenCap3, and PROSPERous) for caspase-7, caspase-2, MMP-8, MMP-3, MMP-12, and membrane-type MMP-1; the identifier in the brackets (e.g. "C14.005") is the protease ID in MEROPS.



Claspin



Heat shock 70 kDa protein 4

Figure S5. Illustration of the predictions generated with DeepCleave webserver for two case study proteins collected from the test dataset: Claspin (top two panels) and Heat shock 70 kDa protein 4 (bottom two panel). The panels of the left summarize predictions generated by DeepCleave while the panels on the right show experimental annotations of the cleavage sites.


Caspase-1 (C14.001)
Caspase-3 (C14.003)
Caspase-7 (C14.004)
Caspase-6 (C14.005)
Caspase-2 (C14.006)
MMP-8 (M10.002)
MMP-2 (M10.003)
MMP-9 (M10.004)
MMP-3 (M10.005)
MMP-7 (M10.008)
MMP-12 (M10.009)
membrane-type MMP-1 (M10.014)

Figure S6. Visualization of the enriched Gene Ontology terms for the putative substrates of caspases and MMPs in human.

Table S1. Size and composition of the data used to develop the training and test datasets for the caspase and MMP families. We used the CD-HIT program to cluster the sequences to reduce the sequence similarity between the training and test sets to 20%.

Family	**Protease**	**MEROPS ID**	**Before CD-HIT**		**CD-HIT**	
			Number of substrates	**Number of cleavage sites**	**Number of sequences**	**Number of cleavage sites**
Caspases	caspase-1	C14.001	155	177	134	152
	CED-3 peptidase	C14.002	7	13	7	13
	caspase-3	C14.003	454	629	324	450
	caspase-7	C14.004	150	178	112	133
	caspase-6	C14.005	923	1306	794	1074
	caspase-2	C14.006	244	289	220	262
	caspase-4	C14.007	8	9	8	9
	caspase-5	C14.008	1	1	1	1
	caspase-8	C14.009	47	66	45	64
	caspase-9	C14.010	11	13	10	12
	caspase-10	C14.011	6	9	6	9
	caspase-11	C14.012	3	4	2	3
	caspase-12	C14.013	2	4	2	4
	caspase (insect 1)	C14.015	6	8	6	8
	caspase (insect 2)	C14.016	2	3	2	3
	caspase-14	C14.018	14	14	14	14
	caspase DRONC	C14.019	2	3	2	3
	caspase-15	C14.038	1	2	1	2
	caspase Dredd	C14.040	1	1	1	1
	csp-1 g.p.	C14.A06	3	6	3	6
	Total		2040	2735	1694	2223
MMPs	matrix metallopeptidase-1	M10.001	34	67	27	58
	matrix metallopeptidase-8	M10.002	34	106	17	67
	matrix metallopeptidase-2	M10.003	1555	3373	777	1363
	matrix metallopeptidase-9	M10.004	83	349	60	243
	matrix metallopeptidase-3	M10.005	62	168	32	87
	matrix metallopeptidase-10	M10.006	9	18	7	10
	matrix metallopeptidase-11	M10.007	17	20	15	18
	matrix metallopeptidase-7	M10.008	60	185	41	152

	matrix metallopeptidase-12	M10.009	30	217	19	180
	envelysin	M10.010	2	6	2	6
	matrix metallopeptidase-13	M10.013	21	59	21	59
	membrane-type matrix metallopeptidase-1	M10.014	46	128	34	74
	membrane-type matrix metallopeptidase-2	M10.015	2	14	1	6
	membrane-type matrix metallopeptidase-3	M10.016	5	20	4	14
	membrane-type matrix metallopeptidase-4	M10.017	5	18	5	18
	matrix metallopeptidase-18	M10.018	2	2	2	2
	matrix metallopeptidase-20	M10.019	9	37	6	20
	matrix metallopeptidase-19	M10.021	4	7	4	7
	membrane-type matrix metallopeptidase-5	M10.023	1	1	1	1
	membrane-type matrix metallopeptidase-6	M10.024	7	45	5	25
	matrix metallopeptidase-26	M10.029	8	12	8	12
	At1g59970-type peptidase	M10.038	1	1	1	1
	karilysin	M10.066	1	5	1	5
	Total		1998	4858	1090	2428

Table S2. The number of substrates and cleavage sites for the training dataset.

Family	**Protease**	**MEROPS ID**	**Number of substrates**	**Number of cleavage sites**
Caspases	caspase-1	C14.001	103	115
	CED-3 peptidase	C14.002	7	13
	caspase-3	C14.003	279	378
	caspase-7	C14.004	85	105
	caspase-6	C14.005	641	880
	caspase-2	C14.006	161	195
	caspase-4	C14.007	8	9
	caspase-5	C14.008	1	1
	caspase-8	C14.009	45	64
	caspase-9	C14.010	10	12
	caspase-10	C14.011	6	9
	caspase-11	C14.012	2	3
	caspase-12	C14.013	2	4
	caspase (insect 1)	C14.015	6	8
	caspase (insect 2)	C14.016	2	3
	caspase-14	C14.018	14	14
	caspase DRONC	C14.019	2	3
	caspase-15	C14.038	1	2
	caspase Dredd	C14.040	1	1
	csp-1 g.p.	C14.A06	3	6
	Total		1379	1825
MMPs	matrix metallopeptidase-1	M10.001	27	58
	matrix metallopeptidase-8	M10.002	19	64
	matrix metallopeptidase-2	M10.003	692	1337
	matrix metallopeptidase-9	M10.004	54	229
	matrix metallopeptidase-3	M10.005	38	96
	matrix metallopeptidase-10	M10.006	7	10
	matrix metallopeptidase-11	M10.007	15	18
	matrix metallopeptidase-7	M10.008	30	120
	matrix metallopeptidase-12	M10.009	16	136
	envelysin	M10.010	2	6
	matrix metallopeptidase-13	M10.013	21	59
	membrane-type matrix metallopeptidase-1	M10.014	25	73
	membrane-type matrix metallopeptidase-2	M10.015	1	6
	membrane-type matrix metallopeptidase-3	M10.016	4	14

	membrane-type matrix metallopeptidase-4	M10.017	5	18
	matrix metallopeptidase-18	M10.018	2	2
	matrix metallopeptidase-20	M10.019	6	20
	matrix metallopeptidase-19	M10.021	4	7
	membrane-type matrix metallopeptidase-5	M10.023	1	1
	membrane-type matrix metallopeptidase-6	M10.024	5	25
	matrix metallopeptidase-26	M10.029	8	12
	At1g59970-type peptidase	M10.038	1	1
	karilysin	M10.066	1	5
	Total		984	2317

Table S3. The number of sequences and cleavage sites in the test dataset.

Family	Protease	Merops ID	Number of substrates	Number of cleavage sites
Caspases	caspase-1	C14.001	38	44
	caspase-3	C14.003	93	141
	caspase-7	C14.004	36	39
	caspase-6	C14.005	228	300
	caspase-2	C14.006	72	81
	Total		467	605
MMPs	matrix metallopeptidase-8	M10.002	6	20
	matrix metallopeptidase-2	M10.003	242	463
	matrix metallopeptidase-9	M10.004	21	54
	matrix metallopeptidase-3	M10.005	9	21
	matrix metallopeptidase-7	M10.008	19	44
	matrix metallopeptidase-12	M10.009	9	59
	membrane-type matrix metallopeptidase-1	M10.014	15	24
	Total		321	685

Table S4. Predictive performance for the prediction of the substrate cleavage sites for 5 proteases from caspase family for the DeepCleave trained without the transfer learning (original) and DeepCleave trained with transfer learning (transfer): C14.001 (caspase-1), C14.003 (caspase-3), C14.004 (caspase-7), C14.005 (caspase-6) and C14.006 (caspase-2). The results were evaluated in terms of MCC, ACC (Accuracy), Sensitivity (Sn), Specificity (Sp), Precision, and AUC on the test dataset.

Merops ID		**Models**	**MCC**	**ACC**	**Sensitivity**	**Specificity**	**Precision**	**AUC**
C14.001	original	1	0.9199	0.9583	0.9167	1.0000	1.0000	1.0000
		2	0.9199	0.9583	0.9167	1.0000	1.0000	1.0000
		3	0.9199	0.9583	0.9167	1.0000	1.0000	1.0000
		4	0.9199	0.9583	0.9167	1.0000	1.0000	1.0000
		5	0.9199	0.9583	0.9167	1.0000	1.0000	1.0000
	transfer learning	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		2	0.9199	0.9583	0.9167	1.0000	1.0000	0.9931
		3	0.9199	0.9583	0.9167	1.0000	1.0000	0.9931
		4	0.9199	0.9583	0.9167	1.0000	1.0000	0.9931
		5	0.9199	0.9583	0.9167	1.0000	1.0000	0.9931
		6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		16	0.9199	0.9583	0.9167	1.0000	1.0000	0.9583
		17	0.9199	0.9583	0.9167	1.0000	1.0000	0.9583
		18	0.9199	0.9583	0.9167	1.0000	1.0000	0.9583
		19	0.9199	0.9583	0.9167	1.0000	1.0000	0.9583
		20	0.9199	0.9583	0.9167	1.0000	1.0000	0.9583
		21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C14.003	original	1	0.9740	0.9868	1.0000	0.9737	0.9744	0.9972
		2	0.9740	0.9868	1.0000	0.9737	0.9744	0.9986
		3	0.9474	0.9737	0.9737	0.9737	0.9737	0.9882
		4	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979

		5	0.9740	0.9868	1.0000	0.9737	0.9744	0.9958
	transfer learning	1	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		2	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		3	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		4	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		5	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		6	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		7	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		8	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		9	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		10	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		11	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		12	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		13	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		14	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		15	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		16	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		17	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		18	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		19	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		20	0.9740	0.9868	1.0000	0.9737	0.9744	0.9979
		21	0.9740	0.9868	1.0000	0.9737	0.9744	0.9972
		22	0.9740	0.9868	1.0000	0.9737	0.9744	0.9972
		23	0.9740	0.9868	1.0000	0.9737	0.9744	0.9972
		24	0.9740	0.9868	1.0000	0.9737	0.9744	0.9972
		25	0.9740	0.9868	1.0000	0.9737	0.9744	0.9972
C14.004	original	1	0.9129	0.9545	1.0000	0.9091	0.9167	0.9917
		2	0.9129	0.9545	0.9091	1.0000	1.0000	0.9752
		3	0.9129	0.9545	0.9091	1.0000	1.0000	0.9835
		4	0.9129	0.9545	0.9091	1.0000	1.0000	0.9587
		5	0.9129	0.9545	0.9091	1.0000	1.0000	0.9256
	transfer learning	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

		11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
C14.005	original	1	0.9660	0.9830	0.9773	0.9886	0.9885	0.9963
		2	0.9665	0.9830	1.0000	0.9659	0.9670	0.9944
		3	0.9548	0.9773	0.9886	0.9659	0.9667	0.9948
		4	0.9548	0.9773	0.9886	0.9659	0.9667	0.9943
		5	0.9555	0.9773	1.0000	0.9545	0.9565	0.9929
	transfer learning	1	0.9545	0.9773	0.9773	0.9773	0.9773	0.9919
		2	0.9545	0.9773	0.9773	0.9773	0.9773	0.9919
		3	0.9545	0.9773	0.9773	0.9773	0.9773	0.9919
		4	0.9545	0.9773	0.9773	0.9773	0.9773	0.9917
		5	0.9545	0.9773	0.9773	0.9773	0.9773	0.9919
		6	0.9555	0.9773	1.0000	0.9545	0.9565	0.9884
		7	0.9555	0.9773	1.0000	0.9545	0.9565	0.9881
		8	0.9555	0.9773	1.0000	0.9545	0.9565	0.9880
		9	0.9555	0.9773	1.0000	0.9545	0.9565	0.9881
		10	0.9555	0.9773	1.0000	0.9545	0.9565	0.9880
		11	0.9555	0.9773	1.0000	0.9545	0.9565	0.9952
		12	0.9555	0.9773	1.0000	0.9545	0.9565	0.9951
		13	0.9555	0.9773	1.0000	0.9545	0.9565	0.9952
		14	0.9555	0.9773	1.0000	0.9545	0.9565	0.9951
		15	0.9555	0.9773	1.0000	0.9545	0.9565	0.9951
		16	0.9548	0.9773	0.9886	0.9659	0.9667	0.9921
		17	0.9548	0.9773	0.9886	0.9659	0.9667	0.9923
		18	0.9548	0.9773	0.9886	0.9659	0.9667	0.9920
		19	0.9548	0.9773	0.9886	0.9659	0.9667	0.9919
		20	0.9548	0.9773	0.9886	0.9659	0.9667	0.9925
		21	0.9548	0.9773	0.9886	0.9659	0.9667	0.9879

		22	0.9548	0.9773	0.9886	0.9659	0.9667	0.9876
		23	0.9548	0.9773	0.9886	0.9659	0.9667	0.9879
		24	0.9548	0.9773	0.9886	0.9659	0.9667	0.9876
		25	0.9548	0.9773	0.9886	0.9659	0.9667	0.9876
C14.006	original	1	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		2	0.9512	0.9750	1.0000	0.9500	0.9524	0.9925
		3	0.9512	0.9750	1.0000	0.9500	0.9524	0.9900
		4	0.9000	0.9500	0.9500	0.9500	0.9500	0.9775
		5	0.9512	0.9750	1.0000	0.9500	0.9524	0.9925
	transfer learning	1	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		2	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		3	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		4	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		5	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		6	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		7	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		8	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		9	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		10	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		21	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		22	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		23	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		24	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975
		25	0.9512	0.9750	1.0000	0.9500	0.9524	0.9975

Table S5. Predictive performance for the prediction of the substrate cleavage sites for 7 proteases from MMP family for the DeepCleave trained without the transfer learning (original) and DeepCleave trained with transfer learning (transfer): M10.002(matrix metallopeptidase-8), M10.003 (matrix metallopeptidase-2), M10.004 (matrix metallopeptidase-9), M10.005 (matrix metallopeptidase-3), M10.008 (matrix metallopeptidase-7), M10.009 (matrix metallopeptidase-12) and M10.014 (membrane-type matrix metallopeptidase-1). The results were evaluated in terms of MCC, ACC (Accuracy), Sensitivity (Sn), Specificity (Sp), Precision, and AUC on the test dataset.

Merops ID		Models	MCC	ACC	Sensitivity	Specificity	Precision	AUC
M10.002	original	1	0.4082	0.6429	0.2857	1.0000	1.0000	0.9184
		2	0.4082	0.6429	0.2857	1.0000	1.0000	1.0000
		3	0.6325	0.7857	1.0000	0.5714	0.7000	1.0000
		4	0.5222	0.7143	0.4286	1.0000	1.0000	1.0000
		5	0.2774	0.5714	0.1429	1.0000	1.0000	1.0000
	transfer learning	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		16	0.8660	0.9286	1.0000	0.8571	0.8750	0.9796
		17	0.8660	0.9286	1.0000	0.8571	0.8750	0.9796
		18	0.8660	0.9286	1.0000	0.8571	0.8750	0.9796
		19	0.8660	0.9286	1.0000	0.8571	0.8750	0.9796
		20	0.8660	0.9286	1.0000	0.8571	0.8750	0.9796
		21	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		22	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		23	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		24	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		25	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
M10.003	original	1	-	-	-	-	-	0.5
		2	-	-	-	-	-	0.5
		3	-	-	-	-	-	0.5

		4	-	-	-	-	-	0.5
		5	-	-	-	-	-	0.5
	transfer learning	1	0.8362	0.9179	0.9030	0.9328	0.9308	0.9681
		2	0.8439	0.9216	0.9030	0.9403	0.9380	0.9684
		3	0.8362	0.9179	0.9030	0.9328	0.9308	0.9681
		4	0.8362	0.9179	0.9030	0.9328	0.9308	0.9678
		5	0.8439	0.9216	0.9030	0.9403	0.9380	0.9685
		6	0.8209	0.9104	0.9104	0.9104	0.9104	0.9622
		7	0.8209	0.9104	0.9104	0.9104	0.9104	0.9603
		8	0.8209	0.9104	0.9104	0.9104	0.9104	0.9622
		9	0.8373	0.9179	0.8881	0.9478	0.9444	0.9611
		10	0.8209	0.9104	0.9104	0.9104	0.9104	0.9623
		11	0.7836	0.8918	0.8955	0.8881	0.8889	0.9607
		12	0.7996	0.8993	0.8731	0.9254	0.9213	0.9607
		13	0.7987	0.8993	0.8881	0.9104	0.9084	0.9615
		14	0.7996	0.8993	0.8731	0.9254	0.9213	0.9615
		15	0.7996	0.8993	0.8731	0.9254	0.9213	0.9616
		16	0.8433	0.9216	0.9254	0.9179	0.9185	0.9628
		17	0.8508	0.9254	0.9328	0.9179	0.9191	0.9676
		18	0.8433	0.9216	0.9254	0.9179	0.9185	0.9639
		19	0.8507	0.9254	0.9254	0.9254	0.9254	0.9684
		20	0.8507	0.9254	0.9254	0.9254	0.9254	0.9680
		21	0.8284	0.9142	0.9179	0.9104	0.9111	0.9661
		22	0.8359	0.9179	0.9254	0.9104	0.9118	0.9662
		23	0.8359	0.9179	0.9254	0.9104	0.9118	0.9669
		24	0.8359	0.9179	0.9254	0.9104	0.9118	0.9669
		25	0.8359	0.9179	0.9254	0.9104	0.9118	0.9663
M10.004	original	1	0.4564	0.7174	0.5652	0.8696	0.8125	0.6786
		2	0.0510	0.5217	0.2609	0.7826	0.5455	0.3516
		3	0.0436	0.5217	0.5652	0.4783	0.5200	0.4858
		4	-0.1325	0.4348	0.5217	0.3478	0.4444	0.3327
		5	0.0464	0.5217	0.3478	0.6957	0.5333	0.4216
	transfer learning	1	0.6623	0.8261	0.9130	0.7391	0.7778	0.8110
		2	0.6093	0.8043	0.8261	0.7826	0.7917	0.8110
		3	0.6093	0.8043	0.8261	0.7826	0.7917	0.8110
		4	0.6093	0.8043	0.8261	0.7826	0.7917	0.8129
		5	0.6093	0.8043	0.8261	0.7826	0.7917	0.8110
		6	0.6547	0.8261	0.8696	0.7826	0.8000	0.8488
		7	0.6547	0.8261	0.8696	0.7826	0.8000	0.8507
		8	0.6547	0.8261	0.8696	0.7826	0.8000	0.8488
		9	0.6547	0.8261	0.8696	0.7826	0.8000	0.8507

		10	0.6547	0.8261	0.8696	0.7826	0.8000	0.8488
		11	0.5262	0.7609	0.8261	0.6957	0.7308	0.8072
		12	0.5262	0.7609	0.8261	0.6957	0.7308	0.8015
		13	0.5222	0.7609	0.7826	0.7391	0.7500	0.8091
		14	0.5222	0.7609	0.7826	0.7391	0.7500	0.8072
		15	0.5262	0.7609	0.8261	0.6957	0.7308	0.8072
		16	0.6547	0.8261	0.7826	0.8696	0.8571	0.7902
		17	0.6093	0.8043	0.7826	0.8261	0.8182	0.7807
		18	0.6139	0.8043	0.7391	0.8696	0.8500	0.7883
		19	0.6093	0.8043	0.7826	0.8261	0.8182	0.7826
		20	0.6093	0.8043	0.7826	0.8261	0.8182	0.7826
		21	0.5652	0.7826	0.7826	0.7826	0.7826	0.8110
		22	0.5652	0.7826	0.7826	0.7826	0.7826	0.8185
		23	0.6093	0.8043	0.8261	0.7826	0.7917	0.8223
		24	0.5652	0.7826	0.7826	0.7826	0.7826	0.8129
		25	0.6093	0.8043	0.8261	0.7826	0.7917	0.8242
M10.005	original	1	0.7035	0.8500	0.9000	0.8000	0.8182	0.8900
		2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		5	0.8165	0.9000	1.0000	0.8000	0.8333	0.9200
	transfer learning	1	0.6547	0.8000	0.6000	1.0000	1.0000	0.7700
		2	0.6547	0.8000	0.6000	1.0000	1.0000	0.7700
		3	0.6547	0.8000	0.6000	1.0000	1.0000	0.7700
		4	0.6547	0.8000	0.6000	1.0000	1.0000	0.7500
		5	0.6547	0.8000	0.6000	1.0000	1.0000	0.7600
		6	0.7338	0.8500	0.7000	1.0000	1.0000	0.8900
		7	0.7338	0.8500	0.7000	1.0000	1.0000	0.8900
		8	0.7338	0.8500	0.7000	1.0000	1.0000	0.8900
		9	0.7338	0.8500	0.7000	1.0000	1.0000	0.8900
		10	0.7338	0.8500	0.7000	1.0000	1.0000	0.8900
		11	0.4000	0.7000	0.7000	0.7000	0.7000	0.6800
		12	0.3015	0.6500	0.6000	0.7000	0.6667	0.6700
		13	0.3015	0.6500	0.6000	0.7000	0.6667	0.6700
		14	0.3015	0.6500	0.6000	0.7000	0.6667	0.6700
		15	0.4000	0.7000	0.7000	0.7000	0.7000	0.6800
		16	0.7338	0.8500	0.7000	1.0000	1.0000	0.8100
		17	0.7338	0.8500	0.7000	1.0000	1.0000	0.7900
		18	0.7338	0.8500	0.7000	1.0000	1.0000	0.8100
		19	0.7338	0.8500	0.7000	1.0000	1.0000	0.7900
		20	0.7338	0.8500	0.7000	1.0000	1.0000	0.7900

		21	0.5025	0.7500	0.7000	0.8000	0.7778	0.7300
		22	0.5025	0.7500	0.7000	0.8000	0.7778	0.7400
		23	0.5025	0.7500	0.7000	0.8000	0.7778	0.7300
		24	0.5025	0.7500	0.7000	0.8000	0.7778	0.7400
		25	0.5025	0.7500	0.7000	0.8000	0.7778	0.7300
M10.008	original	1	0.8333	0.9167	0.9167	0.9167	0.9167	0.9722
		2	0.7526	0.8750	0.9167	0.8333	0.8462	0.9375
		3	0.8452	0.9167	0.8333	1.0000	1.0000	0.9722
		4	0.5000	0.7500	0.7500	0.7500	0.7500	0.8472
		5	0.7526	0.8750	0.9167	0.8333	0.8462	0.9167
	transfer learning	1	0.7526	0.8750	0.8333	0.9167	0.9091	0.8750
		2	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
		3	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
		4	0.7526	0.8750	0.8333	0.9167	0.9091	0.8750
		5	0.7526	0.8750	0.8333	0.9167	0.9091	0.8750
		6	0.7526	0.8750	0.9167	0.8333	0.8462	0.9306
		7	0.7526	0.8750	0.9167	0.8333	0.8462	0.9306
		8	0.7526	0.8750	0.9167	0.8333	0.8462	0.9375
		9	0.7526	0.8750	0.9167	0.8333	0.8462	0.9375
		10	0.7526	0.8750	0.9167	0.8333	0.8462	0.9167
		11	0.5854	0.7917	0.8333	0.7500	0.7692	0.8472
		12	0.5854	0.7917	0.8333	0.7500	0.7692	0.8264
		13	0.5854	0.7917	0.8333	0.7500	0.7692	0.8403
		14	0.5854	0.7917	0.8333	0.7500	0.7692	0.8264
		15	0.5854	0.7917	0.8333	0.7500	0.7692	0.8403
		16	0.6025	0.7917	0.9167	0.6667	0.7333	0.8611
		17	0.6025	0.7917	0.9167	0.6667	0.7333	0.8611
		18	0.5854	0.7917	0.7500	0.8333	0.8182	0.8681
		19	0.5854	0.7917	0.7500	0.8333	0.8182	0.8750
		20	0.6025	0.7917	0.9167	0.6667	0.7333	0.8611
		21	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
		22	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
		23	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
		24	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
		25	0.6667	0.8333	0.8333	0.8333	0.8333	0.8681
M10.009	original	1	0.9309	0.9643	0.9286	1.0000	1.0000	0.9337
		2	0.7454	0.8571	0.7143	1.0000	1.0000	0.8469
		3	0.8571	0.9286	0.9286	0.9286	0.9286	0.9286
		4	0.9309	0.9643	0.9286	1.0000	1.0000	0.9337
		5	0.7454	0.8571	0.7143	1.0000	1.0000	0.7908
	transfer	1	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694

	learning	2	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		3	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		4	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		5	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		6	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		7	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		8	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		9	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		10	0.8571	0.9286	0.9286	0.9286	0.9286	0.9694
		11	0.8571	0.9286	0.9286	0.9286	0.9286	0.9490
		12	0.8571	0.9286	0.9286	0.9286	0.9286	0.9490
		13	0.8571	0.9286	0.9286	0.9286	0.9286	0.9490
		14	0.8571	0.9286	0.9286	0.9286	0.9286	0.9490
		15	0.8571	0.9286	0.9286	0.9286	0.9286	0.9490
		16	0.7877	0.8929	0.9286	0.8571	0.8667	0.9286
		17	0.7877	0.8929	0.9286	0.8571	0.8667	0.9286
		18	0.7877	0.8929	0.9286	0.8571	0.8667	0.9184
		19	0.7877	0.8929	0.9286	0.8571	0.8667	0.9235
		20	0.7877	0.8929	0.9286	0.8571	0.8667	0.9286
		21	0.8571	0.9286	0.9286	0.9286	0.9286	0.9439
		22	0.8571	0.9286	0.9286	0.9286	0.9286	0.9439
		23	0.8571	0.9286	0.9286	0.9286	0.9286	0.9337
		24	0.8571	0.9286	0.9286	0.9286	0.9286	0.9490
		25	0.8571	0.9286	0.9286	0.9286	0.9286	0.9337
M10.014	original	1	0.8819	0.9375	0.8750	1.0000	1.0000	0.8906
		2	0.7746	0.8750	0.7500	1.0000	1.0000	0.8906
		3	0.8819	0.9375	0.8750	1.0000	1.0000	0.8906
		4	0.8819	0.9375	0.8750	1.0000	1.0000	0.8906
		5	0.8819	0.9375	0.8750	1.0000	1.0000	0.8906
	transfer learning	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0.8819	0.9375	1.0000	0.8750	0.8889	0.9375
17	0.8819	0.9375	1.0000	0.8750	0.8889	0.9375
18	0.8819	0.9375	1.0000	0.8750	0.8889	0.9375
19	0.8819	0.9375	1.0000	0.8750	0.8889	0.9375
20	0.8819	0.9375	1.0000	0.8750	0.8889	0.9375
21	0.8819	0.9375	0.8750	1.0000	1.0000	0.9844
22	0.8819	0.9375	1.0000	0.8750	0.8889	0.9844
23	0.8819	0.9375	1.0000	0.8750	0.8889	0.9844
24	0.8819	0.9375	1.0000	0.8750	0.8889	0.9844
25	0.8819	0.9375	0.8750	1.0000	1.0000	0.9844

Table S6. Comparison of the predictive performance for Cascleave, CAT3, CleavPredict, ScreenCap3, SitePrediction, PROSPERous, DeepCleave (original; before transfer learning) and DeepCleave (transfer; after transfer learning) for the prediction of the substrate cleavage sites of 12 proteases from the caspase and MMP families: C14.001 (caspase-1), C14.003 (caspase-3), C14.004 (caspase-7), C14.005 (caspase-6), C14.006 (caspase-2), M10.002 (matrix metallopeptidase-8), M10.003 (matrix metallopeptidase-2), M10.004 (matrix metallopeptidase-9), M10.005 (matrix metallopeptidase-3), M10.008 (matrix metallopeptidase-7), M10.009 (matrix metallopeptidase-12) and M10.014 (membrane-type matrix metallopeptidase-1). The results were evaluated in terms of MCC, ACC (Accuracy), Sensitivity (Sn), Specificity (Sp), Precision and AUC on the test dataset. NA (not applicable) means that a given tool does not make this prediction. Methods with the best results (AUC) for a given protease are shown in bold font.

Family	Merops ID	Tools	MCC	ACC	Sensitivity	Specificity	Precision	AUC
Caspase	C14.001	Cascleave	0.0311	0.5119	0.1905	0.8333	0.5333	0.5088
		CAT3	0.3608	0.6786	0.7500	0.6071	0.6563	0.6875
		CleavPredict	NA					
		ScreenCap3	0.4743	0.7368	0.7105	0.7632	0.7500	0.7479
		SitePrediction	0.7175	0.8571	0.8095	0.9048	0.8947	0.8926
		PROSPERous	0.7526	0.8750	0.8333	0.9167	0.9091	0.9141
		DeepCleave (original)	0.6641	0.8315	0.7955	0.8667	0.8537	0.8833
		DeepCleave (transfer)	**0.8464**	**0.9205**	**0.8636**	**0.9773**	**0.9744**	**0.9556**
	C14.003	Cascleave	0.2023	0.5920	0.3840	0.8000	0.6575	0.5789
		CAT3	0.5466	0.7731	0.7923	0.7538	0.7630	0.8561
		CleavPredict	NA					
		ScreenCap3	0.7492	0.8741	0.8489	0.8993	0.8939	0.9275
		SitePrediction	0.8559	0.9275	0.9058	0.9493	0.9470	0.9814
		PROSPERous	0.9317	0.9656	0.9813	0.9500	0.9515	0.9823
		DeepCleave (original)	0.9530	0.9759	1.0000	0.9539	0.9521	0.9912
		DeepCleave (transfer)	**0.9713**	**0.9856**	**0.9928**	**0.9784**	**0.9787**	**0.9959**
	C14.004	Cascleave	0.2408	0.6143	0.4571	0.7714	0.6667	0.6049
		CAT3	0.6797	0.8378	0.8919	0.7838	0.8049	0.8992
		CleavPredict	NA					
		ScreenCap3	0.7379	0.8684	0.8947	0.8421	0.8500	0.9197
		SitePrediction	0.9213	0.9605	0.9737	0.9474	0.9487	0.9626
		PROSPERous	0.8721	0.9359	0.9487	0.9231	0.9250	0.9671
		DeepCleave (original)	0.9252	0.9625	0.9487	0.9756	0.9737	0.9518
		DeepCleave (transfer)	**0.9487**	**0.9744**	**0.9744**	**0.9744**	**0.9744**	**0.9763**
	C14.005	Cascleave	0.0710	0.5286	0.2328	0.8244	0.5701	0.5261
		CAT3	0.3206	0.6583	0.7374	0.5791	0.6366	0.6906
		CleavPredict	NA					
		ScreenCap3	0.5604	0.7802	0.7785	0.7819	0.7811	0.8541
		SitePrediction	0.7577	0.8767	0.9300	0.8233	0.8404	0.9325
		PROSPERous	0.9701	0.9850	0.9933	0.9767	0.9770	0.9968
		DeepCleave (original)	0.9870	0.9935	0.9933	0.9936	0.9933	0.9995
		DeepCleave (transfer)	**0.9803**	**0.9901**	**0.9967**	**0.9837**	**0.9836**	**0.9994**
	C14.006	Cascleave	0.1387	0.5519	0.2208	0.8831	0.6538	0.5491
		CAT3	0.5280	0.7632	0.7237	0.8026	0.7857	0.8301
		CleavPredict	NA					
		ScreenCap3	0.6667	0.8333	0.8333	0.8333	0.8333	0.8971
		SitePrediction	NA					
		PROSPERous	NA					
		DeepCleave (original)	0.9759	0.9878	0.9753	1.0000	1.0000	0.9993
		DeepCleave (transfer)	**0.9760**	**0.9880**	**0.9877**	**0.9884**	**0.9877**	**0.9996**

MMP	M10.002	Cascleave	NA					
		CAT3	NA					
		CleavPredict	0.4623	0.7273	0.6364	0.8182	0.7778	0.7025
		ScreenCap3	NA					
		SitePrediction	**0.7509**	**0.8750**	**0.8500**	**0.9000**	**0.8947**	**0.8775**
		PROSPERous	0.6939	0.825	1.0000	0.6500	0.7407	0.8625
		DeepCleave (original)	0.2277	0.6250	0.6500	0.5833	0.7222	0.5667
		DeepCleave (transfer)	0.5829	0.7879	0.7500	0.8462	0.8824	0.8731
	M10.003	Cascleave	NA					
		CAT3	NA					
		CleavPredict	0.2099	0.6032	0.5119	0.6944	0.6262	0.6103
		ScreenCap3	NA					
		SitePrediction	0.4603	0.7133	0.9012	0.5253	0.6550	0.7809
		PROSPERous	**0.7471**	**0.8736**	**0.8736**	**0.8736**	**0.8736**	**0.9384**
		DeepCleave (original)	NA					
		DeepCleave (transfer)	0.7799	0.8899	0.8991	0.8814	0.8750	0.9401
	M10.004	Cascleave	NA					
		CAT3	NA					
		CleavPredict	0.2169	0.6081	0.5676	0.6486	0.6176	0.5720
		ScreenCap3	NA					
		SitePrediction	0.6757	0.8378	0.8378	0.8378	0.8378	0.8543
		PROSPERous	0.6158	0.8077	0.7885	0.8269	0.8200	0.9059
		DeepCleave (original)	0.2835	0.5734	0.8654	0.4066	0.4545	0.6103
		DeepCleave (transfer)	**0.7067**	**0.8613**	**0.8269**	**0.8824**	**0.8113**	**0.9100**
	M10.005	Cascleave	NA					
		CAT3	NA					
		CleavPredict	0.4545	0.7273	0.7273	0.7273	0.7273	0.7107
		ScreenCap3	NA					
		SitePrediction	0.7175	0.8571	0.8095	0.9047	0.8947	0.9206
		PROSPERous	**0.7845**	**0.8810**	**1.0000**	**0.7619**	**0.8077**	**0.9320**
		DeepCleave (original)	0.3861	0.7027	0.8095	0.5625	0.7083	0.6250
		DeepCleave (transfer)	0.7566	0.8780	0.9048	0.8500	0.8636	0.9190
	M10.008	Cascleave	NA					
		CAT3	NA					
		CleavPredict	NA					
		ScreenCap3	NA					
		SitePrediction	0.4824	0.7375	0.6500	0.8250	0.7879	0.7744
		PROSPERous	0.7091	0.8523	0.9091	0.7955	0.8163	0.9424
		DeepCleave (original)	0.2551	0.6264	0.6591	0.5957	0.6042	0.6533
		DeepCleave (transfer)	**0.8636**	**0.9318**	**0.9318**	**0.9318**	**0.9318**	**0.9551**
	M10.009	Cascleave	NA					
		CAT3	NA					
		CleavPredict	NA					
		ScreenCap3	NA					
		SitePrediction	0.3656	0.6786	0.5714	0.7857	0.7273	0.7436
		PROSPERous	0.6766	0.8378	0.8108	0.8649	0.8571	0.9218
		DeepCleave (original)	0.2629	0.6316	0.5405	0.7179	0.6452	0.6341
		DeepCleave (transfer)	**0.8158**	**0.9079**	**0.8919**	**0.9231**	**0.9167**	**0.9550**
	M10.014	Cascleave	NA					
		CAT3	NA					
		CleavPredict	NA					
		ScreenCap3	NA					
		SitePrediction	NA					
		PROSPERous	0.6390	0.8125	0.7083	0.9167	0.8947	0.8689
		DeepCleave (original)	0.2988	0.6275	0.4444	0.8333	0.7500	0.5448
		DeepCleave (transfer)	**0.7044**	**0.8519**	**0.8750**	**0.8333**	**0.8077**	**0.8875**

Table S7. Comparison of the predictive performance of models trained using the CKSAAP, putative secondary structure (SS), PSSM with conservation scores (PSSM+CS), putative solvent accessibility (SAS), one-hot encoding, and all inputs combined (all features) for the substrate cleavage site prediction of 12 proteases. The results were evaluated in terms of MCC, ACC (Accuracy), Sensitivity (Sn), Specificity (Sp), Precision and AUC on the training dataset. The best results for a given protease are marked in bold.

Family	**Merops ID**	**Encoding**	**MCC**	**ACC**	**Sensitivity**	**Specificity**	**Precision**	**AUC**
Caspase	C14.001	CKSAAP	0.8169±0.0346	0.9000±0.0204	0.8000±0.0408	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
		BLOSUM62	0.7746±0.0000	0.8750±0.0000	0.7500±0.0000	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
		SS	0.8508±0.0427	0.9200±0.0233	0.8433±0.0429	0.9967±0.0163	0.9964±0.0178	0.9553±0.0402
		PSSM+CS	0.7611±0.0270	0.8667±0.0167	0.7333±0.0333	**1.0000±0.0000**	**1.0000±0.0000**	0.9522±0.0301
		SAS	0.7510±0.0421	0.8650±0.0178	0.7500±0.0000	0.9800±0.0356	0.9760±0.0427	0.9142±0.0397
		One-hot	**0.9712±0.0385**	**0.9850±0.0200**	**0.9700±0.0400**	**1.0000±0.0000**	**1.0000±0.0000**	0.9906±0.0163
		All features	0.9625±0.0602	0.9800±0.0321	0.9657±0.0610	0.9943±0.0280	0.9950±0.0245	**1.0000±0.0000**
	C14.003	CKSAAP	0.8625±0.0345	0.9263±0.0197	0.8526±0.0394	1.0000±0.0000	1.0000±0.0000	0.9669±0.0123
		BLOSUM62	0.9474±0.0000	0.9737±0.0000	0.9737±0.0000	0.9737±0.0000	0.9737±0.0000	0.9976±0.0006
		SS	0.0847±0.0881	0.5405±0.0409	0.6253±0.1072	0.4558±0.1256	0.5380±0.0355	0.6200±0.0242
		PSSM+CS	0.9094±0.0119	0.9526±0.0064	1.0000±0.0000	0.9053±0.0129	0.9136±0.0108	0.9787±0.0090
		SAS	0.8698±0.0380	0.9332±0.0204	0.8926±0.0407	0.9737±0.0000	0.9713±0.0013	0.9700±0.0020
		One-hot	**0.9740±0.0000**	**0.9868±0.0000**	**1.0000±0.0000**	**0.9737±0.0000**	**0.9744±0.0000**	**0.9978±0.0003**
		All features	**0.9740±0.0000**	**0.9868±0.0000**	**1.0000±0.0000**	**0.9737±0.0000**	**0.9744±0.0000**	0.9992±0.0005
	C14.004	CKSAAP	0.9861±0.0319	0.9927±0.0167	0.9891±0.0295	0.9964±0.0178	0.9967±0.0163	1.0000±0.0000
		BLOSUM62	1.0000±0.0000	1.0000±0.0000	1.0000±0.0000	1.0000±0.0000	1.0000±0.0000	1.0000±0.0000
		SS	-0.0328±0.3012	0.5055±0.1222	0.7491±0.1505	0.2618±0.2991	0.5240±0.0998	0.4785±0.1965
		PSSM+CS	0.9129±0.0000	0.9545±0.0000	1.0000±0.0000	0.9091±0.0000	0.9167±0.0000	1.0000±0.0000
		SAS	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
		One-hot	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
		All features	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
	C14.005	CKSAAP	0.7963±0.0436	0.8923±0.0252	0.9745±0.0049	0.8100±0.0535	0.8387±0.0383	0.9428±0.0155

MMP		BLOSUM62	0.9515±0.0138	0.9751±0.0074	**1.0000±0.0000**	0.9502±0.0148	0.9527±0.0129	0.9929±0.0025
		SS	-0.3814±0.0720	0.3564±0.0364	0.6836±0.0787	0.0291±0.0329	0.4121±0.0242	0.4124±0.0633
		PSSM+CS	0.9447±0.0000	0.9716±0.0000	1.0000±0.0000	0.9432±0.0000	0.9462±0.0000	0.9902±0.0020
		SAS	0.9302±0.0189	0.9650±0.0096	0.9659±0.0000	0.9641±0.0192	0.9645±0.0180	0.9950±0.0014
		One-hot	**0.9550±0.0004**	**0.9773±0.0000**	0.9909±0.0085	**0.9636±0.0085**	**0.9647±0.0077**	0.9910±0.0028
		All features	0.9178±0.0276	0.9570±0.0152	**1.0000±0.0000**	0.9141±0.0303	0.9216±0.0249	**0.9977±0.0009**
	C14.006	CKSAAP	0.8462±0.0345	0.9170±0.0196	0.8340±0.0393	**1.0000±0.0000**	**1.0000±0.0000**	0.9160±0.0136
		BLOSUM62	0.9045±0.0000	0.9500±0.0000	**1.0000±0.0000**	0.9000±0.0000	0.9091±0.0000	0.9980±0.0040
		SS	0.1109±0.2945	0.5680±0.1062	0.9000±0.0316	0.2360±0.2013	0.5484±0.0686	0.5359±0.2613
		PSSM+CS	0.8938±0.0214	0.9450±0.0100	0.9900±0.0200	0.9000±0.0000	0.9082±0.0017	0.9645±0.0066
		SAS	0.8162±0.0309	0.9080±0.0154	0.9100±0.0200	0.9060±0.0162	0.9065±0.0158	0.9826±0.0023
		One-hot	**0.9707±0.0239**	**0.9850±0.0122**	**1.0000±0.0000**	0.9700±0.0245	0.9714±0.0233	0.9985±0.0012
		All features	0.8888±0.0362	0.9410±0.0198	**1.0000±0.0000**	0.8820±0.0397	0.8956±0.0327	**1.0000±0.0000**
	M10.002	CKSAAP	0.6007±0.3090	0.6171±0.1412	**1.0000±0.0000**	0.2343±0.2825	0.5827±0.1024	**1.0000±0.0000**
		BLOSUM62	0.8200±0.1119	0.9086±0.0554	0.9314±0.0714	0.8857±0.0571	0.8914±0.0548	0.9837±0.0082
		SS	0.4038±0.1608	0.6457±0.0979	0.2914±0.1958	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
		PSSM+CS	0.7917±0.4082	0.6514±0.2003	**1.0000±0.0000**	0.3029±0.4007	0.6334±0.1897	0.8955±0.2098
		SAS	0.8370±0.1270	0.9086±0.0743	0.8171±0.1486	**1.0000±0.0000**	**1.0000±0.0000**	**1.0000±0.0000**
		One-hot	**0.9732±0.0536**	**0.9857±0.0286**	**1.0000±0.0000**	0.9714±0.0571	0.9750±0.0500	0.9959±0.0082
		All features	**0.9732±0.0536**	**0.9857±0.0286**	0.9714±0.0571	1.0000±0.0000	1.0000±0.0000	**1.0000±0.0000**
	M10.003	CKSAAP	0.8163±0.0220	0.9058±0.0120	0.9486±0.0215	0.8630±0.0441	0.8759±0.0360	0.9639±0.0100
		BLOSUM62	0.7816±0.0178	0.8884±0.0095	0.9433±0.0060	0.8334±0.0187	0.8502±0.0145	0.9594±0.0058
		SS	0.1816±0.0438	0.5888±0.0195	0.5830±0.0998	0.5946±0.0900	0.5930±0.0265	0.6057±0.0426
		PSSM+CS	0.3835±0.1208	0.6512±0.0587	**0.9704±0.0127**	0.3319±0.1153	0.5949±0.0389	0.7847±0.0532
		SAS	0.5072±0.0661	0.7309±0.0362	0.5418±0.1120	0.9200±0.0714	0.8856±0.0714	0.8563±0.0301
		One-hot	**0.8284±0.0185**	**0.9140±0.0093**	0.9081±0.0178	0.9200±0.0129	**0.9192±0.0118**	**0.9647±0.0030**
		All features	0.8258±0.0986	0.9121±0.0498	0.9030±0.0784	**0.9211±0.0282**	0.9185±0.0307	0.9338±0.0478

M10.004	CKSAAP	0.3239±0.2069	0.6507±0.1095	0.8076±0.1174	0.4939±0.2385	0.6360±0.1244	0.8161±0.0746
	BLOSUM62	0.5901±0.0350	0.7948±0.0175	0.7739±0.0174	0.8157±0.0186	0.8077±0.0188	0.8219±0.0124
	SS	-0.1604±0.0474	0.4200±0.0236	0.4261±0.0174	0.4139±0.0552	0.4220±0.0192	0.4048±0.0269
	PSSM+CS	-0.1906±0.1044	0.4757±0.0270	0.9513±0.0541	0.0000±0.0000	0.4871±0.0147	0.6238±0.1238
	SAS	0.1173±0.2821	0.5583±0.1390	0.5948±0.0918	0.5217±0.2324	0.5785±0.1540	0.5517±0.1183
	One-hot	**0.6003±0.0468**	**0.7991±0.0232**	**0.8191±0.0402**	0.7791±0.0442	0.7892±0.0316	0.8140±0.0212
	All features	0.5837±0.1269	0.7782±0.0612	0.6448±0.0884	**0.9115±0.0994**	**0.8902±0.1000**	**0.8397±0.0419**
M10.005	CKSAAP	-0.0005±0.3275	0.5320±0.1303	**0.8560±0.0496**	0.2080±0.2756	0.5377±0.1041	0.7900±0.0613
	BLOSUM62	0.4952±0.0591	0.7460±0.0280	0.7000±0.0000	0.7920±0.0560	0.7739±0.0483	0.8384±0.0415
	SS	-0.6432±0.1501	0.1900±0.0693	0.2960±0.0196	0.0840±0.1405	0.2485±0.0423	0.1582±0.1442
	PSSM+CS	-0.2336±0.0204	0.4480±0.0098	0.8960±0.0196	0.0000±0.0000	0.4725±0.0057	0.7652±0.0473
	SAS	0.5775±0.1661	0.7737±0.0843	0.6591±0.1354	0.8884±0.1367	0.8857±0.1400	0.7997±0.0825
	One-hot	**0.5931±0.1533**	**0.7840±0.0689**	0.6680±0.0466	**0.9000±0.1265**	**0.8916±0.1366**	0.7720±0.0720
	All features	0.5917±0.0789	0.7800±0.0283	0.6840±0.0967	0.8760±0.1450	0.8852±0.1319	**0.9052±0.0892**
M10.008	CKSAAP	0.6625±0.0884	0.8274±0.0396	0.8011±0.0327	0.8537±0.1025	0.8589±0.0984	0.9057±0.0656
	BLOSUM62	0.6074±0.0994	0.8033±0.0493	0.8000±0.0408	0.8067±0.0655	0.8075±0.0612	0.8803±0.0115
	SS	-0.0131±0.4094	0.4933±0.2037	0.5100±0.1948	0.4767±0.2218	0.4995±0.1999	0.5206±0.2474
	PSSM+CS	-0.0450±0.1680	0.4933±0.0631	0.8100±0.0764	0.1767±0.1381	0.4985±0.0420	0.6953±0.0532
	SAS	0.4585±0.2125	0.6517±0.0897	0.3300±0.1787	**0.9733±0.0564**	**0.9346±0.3876**	0.6075±0.2423
	One-hot	**0.6637±0.0686**	**0.8300±0.0352**	**0.8533±0.0488**	0.8067±0.0696	0.8190±0.0505	0.8744±0.0314
	All features	0.6595±0.0898	0.8259±0.0402	0.8004±0.0322	0.8515±0.1037	0.8571±0.0994	**0.9068±0.0647**
M10.009	CKSAAP	0.1730±0.2106	0.5514±0.1021	**0.9543±0.0343**	0.1486±0.2175	0.5390±0.0843	0.9051±0.0316
	BLOSUM62	0.8101±0.0375	0.9043±0.0195	0.9114±0.0366	0.8971±0.0355	0.9000±0.0296	0.9663±0.0241
	SS	-0.1679±0.1338	0.4314±0.0588	0.4029±0.2276	0.4600±0.2635	0.4262±0.0599	0.4335±0.2059
	PSSM+CS	0.2712±0.2294	0.6000±0.1125	0.8514±0.0967	0.3486±0.2629	0.5835±0.0955	0.8029±0.0961
	SAS	0.4015±0.1518	0.6986±0.0770	0.6400±0.1134	0.7571±0.0808	0.7243±0.0760	0.7394±0.1007
	One-hot	**0.8433±0.0278**	**0.9214±0.0143**	0.9286±0.0000	**0.9143±0.0286**	**0.9162±0.0248**	0.9508±0.0172

	All features	0.7879±0.1074	0.8814±0.0652	0.9257±0.0795	0.8371±0.1835	0.8796±0.1286	**0.9657±0.0238**
M10.014	CKSAAP	-0.2807±0.1398	0.4525±0.0406	0.9050±0.0812	0.0000±0.0000	0.4741±0.0228	0.4719±0.0924
	BLOSUM62	0.8341±0.0591	0.9125±0.0306	0.9500±0.0612	0.8750±0.0791	0.8906±0.0639	**0.9875±0.0063**
	SS	-0.4804±0.0000	0.3125±0.0000	0.6250±0.0000	0.0000±0.0000	0.3846±0.0000	0.1125±0.0063
	PSSM+CS	NA	0.5000±0.0000	1.0000±0.0000	0.0000±0.0000	0.5000±0.0000	0.7375±0.0932
	SAS	0.4621±0.3028	0.7025±0.1524	0.9200±0.0696	0.4850±0.3227	0.6797±0.1667	0.8856±0.0771
	One-hot	**0.9528±0.0578**	**0.9750±0.0306**	**0.9900±0.0339**	**0.9600±0.0583**	**0.9644±0.0518**	0.9844±0.0242
	All features	0.5809±0.2370	0.7750±0.1118	0.6800±0.0620	0.8700±0.2487	0.8995±0.1759	0.8725±0.0889

Table S8. Comparison of the predictive performance of models trained using the CKSAAP, BLOSUM62, putative secondary structure (SS), PSSM with conservation scores (PSSM+CS), putative solvent accessibility (SAS), one-hot encoding, and all inputs combined (all features) for the substrate cleavage site prediction of 12 proteases. The results were evaluated in terms of MCC, ACC (Accuracy), Sensitivity (Sn), Specificity (Sp), Precision and AUC on the independent test dataset. The best results for a given protease are shown in bold font.

Family	Merops ID	Encoding strategy	MCC	ACC	Sensitivity	Specificity	Precision	AUC
Caspase	C14.001	CKSAAP	0.6269	0.8068	0.7045	0.9091	0.8857	0.8533
		BLOSUM62	0.7735	0.8864	0.8636	0.9091	0.9048	0.9437
		SS	0.3453	0.6705	0.7500	0.5909	0.6471	0.7020
		PSSM+CS	0.5965	0.7955	0.8636	0.7273	0.7600	0.8585
		SAS	0.7800	0.8864	0.8182	0.9545	0.9474	0.9236
		One-hot	**0.8464**	**0.9205**	0.8636	**0.9773**	**0.9744**	**0.9556**
		All features	0.7957	0.8977	**0.8864**	0.9091	0.9070	0.9411
	C14.003	CKSAAP	0.6088	0.7979	0.6950	0.9007	0.8750	0.9022
		BLOSUM62	0.9583	0.9787	**1.0000**	0.9574	0.9592	0.9932
		SS	0.3411	0.6702	0.7021	0.6383	0.6600	0.7442
		PSSM+CS	0.7885	0.8936	0.9220	0.8652	0.8725	0.9666
		SAS	0.8516	0.9255	0.9433	0.9078	0.9110	0.9787
		One-hot	**0.9713**	**0.9856**	0.9928	**0.9784**	**0.9787**	**0.9959**
		All features	0.9506	0.9752	0.9858	0.9645	0.9653	0.9953
	C14.004	CKSAAP	0.7436	0.8718	0.8718	0.8718	0.8718	0.9145
		BLOSUM62	0.8986	0.9487	0.9744	0.9231	0.9268	0.9717
		SS	0.1604	0.5769	0.7179	0.4359	0.5600	0.5401
		PSSM+CS	0.8473	0.9231	0.9487	0.8974	0.9024	0.9671
		SAS	0.8986	0.9487	0.9231	**0.9744**	0.9730	0.9698
		One-hot	**0.9487**	**0.9744**	0.9744	**0.9744**	**0.9744**	0.9763
		All features	0.9258	0.9615	**1.0000**	0.9231	0.9286	**0.9869**
	C14.005	CKSAAP	0.5465	0.7633	0.6300	0.8967	0.8591	0.8752
		BLOSUM62	0.9576	0.9783	**1.0000**	0.9567	0.9585	0.9987
		SS	0.2275	0.6117	0.7067	0.5167	0.5938	0.6474
		PSSM+CS	0.8241	0.9117	0.9333	0.8900	0.8946	0.9683
		SAS	0.9167	0.9583	0.9533	0.9633	0.9630	0.9887
		One-hot	**0.9803**	**0.9901**	0.9967	**0.9837**	**0.9836**	**0.9994**
		All features	0.9350	0.9667	0.9967	0.9367	0.9403	0.9930
	C14.006	CKSAAP	0.7174	0.8580	0.8272	0.8889	0.8816	0.9079
		BLOSUM62	0.9401	0.9691	**1.0000**	0.9383	0.9419	0.9921
		SS	0.1656	0.5802	0.7037	0.4568	0.5644	0.6322
		PSSM+CS	0.8906	0.9444	0.9753	0.9136	0.9186	0.9893
		SAS	0.8765	0.9383	0.9383	0.9383	0.9383	0.9875
		One-hot	**0.9760**	**0.9880**	0.9877	**0.9884**	**0.9877**	**0.9996**
		All features	0.9401	0.9691	**1.0000**	0.9383	0.9419	0.9974
	M10.002	CKSAAP	0.4201	0.6500	**1.0000**	0.3000	0.5882	0.7475
		BLOSUM62	0.5507	0.7750	0.7500	0.8000	0.7895	0.8300
		SS	0.1502	0.5750	0.6000	0.5500	0.5714	0.6588
		PSSM+CS	0.2801	0.6000	0.9500	0.2500	0.5588	0.6725
		SAS	0.4193	0.7000	0.8500	0.5500	0.6538	0.7400
		One-hot	**0.5829**	**0.7879**	0.7500	**0.8462**	**0.8824**	**0.8731**
		All features	0.4551	0.7250	0.6500	0.8000	0.7647	0.7875
	M10.003	CKSAAP	0.1958	0.5896	0.3888	0.7905	0.6498	0.6376
		BLOSUM62	0.7148	0.8564	0.8942	0.8186	0.8313	0.9353
		SS	0.2031	0.6015	0.5875	0.6156	0.6044	0.6012
		PSSM+CS	0.2225	0.5940	0.8618	0.3261	0.5612	0.6688

	SAS	0.3378	0.6641	0.5464	0.7819	0.7147	0.7302
	One-hot	**0.7799**	**0.8899**	**0.8991**	0.8814	**0.8750**	**0.9401**
	All features	0.2332	0.5680	0.1620	**0.9741**	0.8621	0.7287
M10.004	CKSAAP	0.2809	0.6186	0.3505	0.8866	0.7556	0.5980
	BLOSUM62	0.5571	0.7778	0.8148	0.7407	0.7586	0.8769
	SS	0.0371	0.5185	0.5000	0.5370	0.5192	0.4642
	PSSM+CS	0.3175	0.6574	0.7222	0.5926	0.6393	0.6344
	SAS	0.1809	0.5833	0.3889	0.7778	0.6364	0.6211
	One-hot	**0.7067**	**0.8613**	**0.8269**	0.8824	0.8113	**0.9100**
	All features	0.4961	0.7222	0.5000	**0.9444**	**0.9000**	0.8059
M10.005	CKSAAP	0.3540	0.6750	0.6000	0.7500	0.7059	0.7356
	BLOSUM62	0.4811	0.7381	0.8095	0.6667	0.7083	0.8776
	SS	0.1010	0.5476	0.7143	0.3810	0.5357	0.4773
	PSSM+CS	0.3162	0.6429	0.8571	0.4286	0.6000	0.7574
	SAS	0.1925	0.5952	0.6667	0.5238	0.5833	0.5805
	One-hot	**0.7566**	**0.8780**	**0.9048**	0.8500	0.8636	**0.9190**
	All features	0.4851	0.6905	0.3810	**1.0000**	**1.0000**	0.7392
M10.008	CKSAAP	0.4457	0.6863	0.4118	0.9608	0.9130	0.7532
	BLOSUM62	0.7502	0.8750	0.8864	0.8636	0.8667	0.9303
	SS	0.1820	0.5909	0.6136	0.5682	0.5870	0.6472
	PSSM+CS	0.3093	0.6477	0.5000	0.7955	0.7097	0.7314
	SAS	0.1260	0.5341	0.1136	0.9545	0.7143	0.5842
	One-hot	**0.8636**	**0.9318**	**0.9318**	0.9318	0.9318	**0.9551**
	All features	0.3929	0.6477	0.3182	**0.9773**	**0.9333**	0.7515
M10.009	CKSAAP	0.4529	0.6702	0.3404	**1.0000**	**1.0000**	0.6125
	BLOSUM62	0.6174	0.8051	0.8814	0.7288	0.7647	0.8564
	SS	0.1772	0.5763	0.3220	0.8305	0.6552	0.7070
	PSSM+CS	0.4082	0.7034	0.7458	0.6610	0.6875	0.7297
	SAS	0.3250	0.6610	0.7288	0.5932	0.6418	0.7032
	One-hot	**0.8158**	**0.9079**	**0.8919**	0.9231	0.9167	**0.9550**
	All features	0.4956	0.7373	0.5932	0.8814	0.8333	0.7738
M10.014	CKSAAP	0.1644	0.5806	0.4839	0.6774	0.6000	0.6504
	BLOSUM62	0.5833	0.7917	0.7917	0.7917	0.7917	0.8576
	SS	0.0845	0.5417	0.6250	0.4583	0.5357	0.4427
	PSSM+CS	0.3146	0.6458	0.8333	0.4583	0.6061	0.7292
	SAS	0.4226	0.7083	0.6250	0.7917	0.7500	0.7708
	One-hot	**0.7044**	**0.8519**	**0.8750**	0.8333	0.8077	**0.8875**
	All features	0.5130	0.7083	0.4167	**1.0000**	**1.0000**	0.8767

Table S9. Summary of the results of the proteome-wide substrate cleavage site predictions for the 20425 human sequences and 12 proteases from caspase and MMP families generated by DeepCleave. These predictions utilize high-confidence specificity level (>= 0.99) for each protease to generate high-confidence results.

Family	**Protease**	**MEROPS ID**	**Number of sequences**	**Number of cleavage sites**
Caspases	caspase-1	C14.001	68	286
	caspase-3	C14.003	293	2745
	caspase-7	C14.004	77	4784
	caspase-6	C14.005	717	5518
	caspase-2	C14.006	189	2703
	Total		1344	16036
MMPs	matrix metallopeptidase-8	M10.002	24	357
	matrix metallopeptidase-2	M10.003	577	2518
	matrix metallopeptidase-9	M10.004	22	591
	matrix metallopeptidase-3	M10.005	49	1294
	matrix metallopeptidase-7	M10.008	44	1355
	matrix metallopeptidase-12	M10.009	23	829
	membrane-type matrix metallopeptidase-1	M10.014	38	845
	Total		777	7789

Table S10. The top-5 significantly enriched GO terms for each protease. The GO terms are divided into three types: BP (biological process); MF (molecular function); and CC (cellular component).

Protease	Type	Term	P-value
caspase-1 (C14.001)	BP	mRNA splicing	1.41E-09
	BP	positive regulation of neuroinflammatory response	3.68E-06
	BP	cellular response to interferon-gamma	5.05E-06
	BP	astrocyte activation	1.70E-05
	BP	RNA export from nucleus	5.14E-05
	CC	membrane	2.56E-12
	CC	cytosol	2.45E-10
	CC	catalytic step 2 spliceosome	1.64E-07
	CC	nucleus	2.06E-07
	CC	focal adhesion	3.79E-07
	MF	RNA binding	5.53E-18
	MF	protein kinase binding	1.90E-06
	MF	cadherin binding	1.14E-05
	MF	identical protein binding	2.03E-05
	MF	apolipoprotein binding	6.83E-05
caspase-3 (C14.003)	BP	apoptotic process	7.19E-20
	BP	viral process	2.64E-17
	BP	intrinsic apoptotic signaling pathway in response to DNA damage	5.78E-14
	BP	protein phosphorylation	2.82E-13
	BP	neuron apoptotic process	3.29E-13
	CC	cytosol	2.03E-50

	CC	cytoplasm	1.73E-29
	CC	nucleoplasm	2.26E-29
	CC	nucleus	1.80E-26
	CC	protein-containing complex	2.34E-22
	MF	identical protein binding	1.23E-23
	MF	RNA binding	5.35E-23
	MF	protein kinase binding	2.15E-20
	MF	ATP binding	2.28E-19
	MF	ubiquitin protein ligase binding	4.64E-17
caspase-7 (C14.004)	BP	mesenchyme migration	9.31E-10
	BP	platelet formation	7.39E-09
	BP	ephrin receptor signaling pathway	2.62E-08
	BP	neuron apoptotic process	3.88E-07
	BP	negative regulation of apoptotic process	2.12E-06
	CC	cytosol	3.70E-17
	CC	cytoplasm	3.80E-11
	CC	nucleoplasm	2.71E-10
	CC	extracellular exosome	1.85E-08
	CC	nucleus	1.79E-07
	MF	activating transcription factor binding	3.46E-09
	MF	ATP binding	1.30E-06
	MF	structural constituent of cytoskeleton	1.91E-06
	MF	identical protein binding	3.42E-06
	MF	protein domain specific binding	3.49E-06
caspase-6 (C14.005)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08
caspase-2 (C14.006)	BP	mRNA splicing	1.34E-15
	BP	RNA metabolic process	1.48E-08
	BP	viral process	3.48E-06

	BP	negative regulation of mRNA splicing	3.57E-05
	BP	RNA export from nucleus	5.47E-05
	CC	nucleoplasm	2.18E-18
	CC	cytosol	3.99E-18
	CC	membrane	1.45E-15
	CC	nucleus	2.52E-13
	CC	integral component of membrane	6.64E-11
	MF	RNA binding	8.29E-30
	MF	cadherin binding	3.03E-08
	MF	proteasome-activating ATPase activity	1.26E-07
	MF	mRNA binding	2.10E-06
	MF	RNA strand annealing activity	8.81E-06
MMP-8 (M10.002)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08
MMP-2 (M10.003)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08

MMP-9 (M10.004)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08
MMP-3 (M10.005)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08
MMP-7 (M10.008)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17

	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08
MMP-12 (M10.009)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08
membraine-type-MMP-1 (M10.014)	BP	mRNA splicing	4.92E-22
	BP	RNA splicing	1.15E-11
	BP	viral process	2.46E-10
	BP	mRNA processing	1.83E-09
	BP	regulation of mRNA stability	6.14E-09
	CC	nucleoplasm	2.51E-62
	CC	cytosol	1.27E-47
	CC	nucleus	2.66E-39
	CC	integral component of membrane	1.89E-31
	CC	membrane	1.13E-28
	MF	RNA binding	9.94E-68
	MF	cadherin binding	4.34E-17
	MF	G protein-coupled receptor activity	9.61E-13
	MF	ATP binding	1.77E-11
	MF	mRNA binding	1.02E-08