

Table S2 The most 50 important CKSAAP encoding mRMR feature scores and their corresponding amino acid pairs are presented.

Ranks	Scores	Pairs	Ranks	Scores	Pairs
1	0.03	E×L	26	0.011	N××××W
2	0.018	A×××Y	27	0.011	G×××G
3	0.017	P×F	28	0.011	M××××G
4	0.017	MF	29	0.011	K×Y
5	0.016	Y××××V	30	0.011	PV
6	0.016	KD	31	0.01	A××××Y
7	0.015	V×××V	32	0.01	K×K
8	0.015	G×××P	33	0.01	R××××F
9	0.014	DL	34	0.01	Q××××E
10	0.014	E××××E	35	0.01	F××G
11	0.014	RI	36	0.01	M××××K
12	0.013	S××××H	37	0.01	TN
13	0.014	M××S	38	0.01	LV
14	0.014	G×××××E	39	0.01	I×××××A
15	0.013	D×A	40	0.01	G×××D
16	0.013	VF	41	0.01	M××××M
17	0.013	Y×××A	42	0.009	RK
18	0.012	P××D	43	0.009	K×××V
19	0.012	LN	44	0.009	K×××D
20	0.012	TT	45	0.009	PE
21	0.012	K××××D	46	0.009	M××L
22	0.011	G×D	47	0.009	Q×N
23	0.011	T××D	48	0.009	D×V
24	0.011	EW	49	0.009	SM
25	0.011	P××××F	50	0.009	N×××T

For example, the feature 'ExL' represents a 1-spaced residue (any amino acid) pair of 'EL', where x stands for any amino acid. And the same representation was applied to other *k*-spaced residue pairs.