

Table S1. The predictive performance of balanced and unbalanced dataset

Positive dataset : Negative dataset	Sn (%)	Sp (%)	Acc (%)	MCC
1:1	80.83	78.17	79.50	0.5902
1:2	60.47	88.94	79.45	0.5215
1:3	47.79	93.81	82.30	0.4841
1:4	40.71	94.69	83.89	0.4298
1:5	34.22	96.70	86.28	0.4141

Table S2. Euclidean distances between the matrix of complete dataset and 60 subsets

Datasets	1	2	3	4	5	6	7	8
Distances	0.0377	0.0319	0.0306	0.0307	0.0306	0.0260	0.0337	0.0362
Datasets	9	10	11	12	13	14	15	16
Distances	0.0377	0.0399	0.0320	0.0354	0.0279	0.0275	0.0308	0.0318
Datasets	17	18	19	20	21	22	23	24
Distances	0.0325	0.0274	0.0282	0.0366	0.0279	0.0301	0.0346	0.0367
Datasets	25	26	27	28	29	30	31	32
Distances	0.0296	0.0267	0.0316	0.0336	0.0355	0.0287	0.0321	0.0277
Datasets	33	34	35	36	37	38	39	40
Distances	0.0319	0.0336	0.0267	0.0276	0.0257	0.0283	0.0304	0.0281
Datasets	41	42	43	44	45	46	47	48
Distances	0.0281	0.0372	0.0345	0.0299	0.0338	0.0297	0.0328	0.0297
Datasets	49	50	51	52	53	54	55	56
Distances	0.0305	0.0354	0.0293	0.0385	0.0386	0.0322	0.0362	0.0319
Datasets	57	58	59	60				
Distances	0.0293	0.0343	0.0341	0.0336				

Table S3. Predictive performances of O-GlcNAcylated sites based on different local window sizes using BPB features on jackknife test

Window size	Sn (%)	Sp (%)	Acc (%)
11	67.26	74.34	70.80
13	68.44	73.16	70.80
15	70.80	73.45	72.12
17	74.04	71.98	73.01
19	74.63	74.04	74.34
21	73.45	73.16	73.30
23	73.45	73.45	73.45
25	74.63	74.63	74.63
27	74.63	74.04	74.34

Table S4. Predictive performances of O-GlcNAcylated sites based on different local window sizes using BRABSB features on jackknife test

Window size	Sn (%)	Sp (%)	Acc (%)
11	74.04	75.52	74.78
13	71.39	74.04	72.71
15	73.45	72.86	73.16
17	75.81	73.45	74.63
19	75.22	75.22	75.22
21	74.93	74.04	74.48
23	77.29	76.40	76.84
25	76.40	77.58	76.99
27	75.81	77.29	76.55

Table S5. Predictive performances of O-GlcNAcylated sites based on different local window sizes using ANBPB features on jackknife test

Window size	Sn (%)	Sp (%)	Acc (%)
11	70.80	73.16	71.98
13	72.86	74.63	73.75
15	74.93	73.75	74.34
17	77.29	73.75	75.52
19	76.40	75.22	75.81
21	78.47	73.45	75.96
23	79.94	77.58	78.76
25	79.35	76.99	78.17
27	77.58	76.70	77.14

Table S6. Predictive performances of O-GlcNAcylated sites based on different local window sizes using relative RANS features on jackknife test

Window size	Sn (%)	Sp (%)	Acc (%)
11	64.60	72.57	68.58
13	67.55	70.21	68.88
15	68.73	71.98	70.35
17	69.91	72.27	71.09
19	68.44	72.86	70.65
21	68.73	73.16	70.94
23	70.21	71.68	70.94
25	70.50	72.27	71.39
27	69.03	73.16	71.09

Table S7. Comparison of O-GlcNAcPRED with YinOYang and OGlcNAcScan on the independent test dataset

<i>UniProt ID</i>	<i>O-GlcNAcylation site</i>	<i>YinOYang</i>	<i>OGlcNAcScan</i>	<i>O-GlcNAcPRED</i>
Q15750	S395	+	+	+
	S396	-	-	-
	S399	-	-	+
	T400	+	-	-
	S401	+	-	-
Q7Z589	S238	-	-	-
	T535	-	-	+
	S557	-	-	+
P48382	S439	+	+	+
	S442	-	-	-
Q9UHD9	T113	-	-	+
	S116	-	-	-
	T124	+	-	+
Q16186	S213	+	-	-
	S220	+	+	+
	T221	-	+	+
	T222	+	-	-
	T225	+	-	+
Q99996	T2794	-	-	-
	T2797	-	-	-
Q6UN15	T205	-	+	+
	T208	-	-	+
Q15723	T375	+	-	+

	T376	+	-	+
	S400	-	+	+
	S500	+	+	+
	S572	-	-	+
O95487	S315	+	+	+
	T341	+	-	-
	S342	-	-	-
Q32P51	T74	-	-	+
	T75	-	-	-
Q86YP4	T593	-	-	+
Q5JSZ5	S1982	+	+	+
	S1984	-	-	-
Q9P2N6	T860	-	+	+
O00268	S528	+	+	+
	T532	-	-	+
	T533	+	-	-
Q5IRN2	T192	-	-	-
Q9NZZ3	T23	-	-	-
Q9UPN6	T615	-	-	+
Q9NTZ6	T121	-	-	-
	S139	-	-	-
	S143	-	-	-
Q7Z6J0	S519	-	+	+
	T520	-	+	+
Q9H1B7	S179	+	+	+
	T186	-	-	-
P52597	S265	-	-	+

Q9UQ35	S957	-	-	-
B4DKX5	S772	-	-	-
Q9ULU4	T750	+	+	+
Q96JB8	S436	-	-	+
P42167	S265	-	+	-
Q15942	S259	+	-	+
Q8IVW6	T408	+	-	+
P22681	S601	+	+	+
Q9Y520	T2242	-	-	-
Q5JPK0	S411	-	-	-
Q8NEZ4	T1375	-	+	+
O94842	T409	-	-	-
O15014	S1196	-	+	-
P14314	T287	-	-	-
Q9UQR0	S568	-	+	+
Q8N3X1	S797	+	-	+
Q96L91	S3027	-	+	+

Table S8. Performances comparisons among O-GlcNAcPRED, YinOYang and O-GlcNAcscan

Prediction methods	Sn (%)	Sp (%)	Acc (%)	MCC
YinOYang	34.33	89.36	88.85	0.0725
O-GlcNAcscan	31.34	92.45	91.89	0.0847
O-GlcNAcPRED	56.72	64.77	64.70	0.0428