New Glycoproteomics Software, *GlycoPep Evaluator*, Generates Decoy Glycopeptides *de novo* and Enables Accurate False Discovery Rate Analysis for Small Data Sets

Zhikai Zhu, Xiaomeng Su, Eden P. Go and Heather Desaire[*]

Department of Chemistry, University of Kansas, Lawrence, Kansas 66047, United States

Supporting Information

* Corresponding author. Phone: 785-864-3015, Fax: 785-864-5396, E-mail: hdesaire@ku.edu.

Name	Composition	Mono m/z	Score
target	NAAGN(1743.6078)FSEPSDSSGAITAR	1199.1564	7.1353
decoy	ARMEARFN(1691.6261)FTEFGAEK	1199.1842	13.9909
decoy	AMGLLHN(1465.2924)VTRVHELAVKNK	1199.1678	8.9555
decoy	NLYRN(1637.6116)RSYFYMQFR	1199.1918	8.7201
decoy	DFHFN(1836.6616)KSKTAYNMR	1199.1714	7.5505
decoy	EYYGEFERN(1654.6437)TSQYVR	1199.1785	7.2385
decoy	QFN(1694.6892)KTEMMRDTWADK	1199.1906	6.8453
decoy	GQTEHVPVQKEPKN(1547.4872)VSDR	1199.1846	6.2557
decoy	QYHANCEMNVN(1677.7382)YSCK	1199.1735	5.7624
decoy	QQFNEWN(1643.5779)MSYAYDKK	1199.1523	4.9396
decoy	LLMPDVPCSSN(1913.6774)SSFK	1199.1615	4.9396
decoy	MPTFSKMN(1887.6664)YTFRGK	1199.1719	4.4829
decoy	N(1755.4980)NSRPRYWFMSLLR	1199.152	4.0451
decoy	NNCQDHRN(1718.6897)YSVSVQR	1199.1837	3.9602
decoy	HEN(1845.6606)VSNEHTMGPNKR	1199.1613	2.965
decoy	SFTRN(1604.4920)ATTLTYGYRSPR	1199.172	2.5651
decoy	ESWVKFHQVNN(1474.5269)HSAPGQR	1199.1929	0
decoy	RVAN(1668.4876)ATGLVTYTRLYTK	1199.1927	0
decoy	DQAMYAACN(1792.7393)ESMKQR	1199.1719	0
decoy	NCMLN(1831.7240)YTVGWFACK	1199.1696	0
decoy	HVGFSSCKN(1790.6631)GSGQLNGR	1199.1762	0

Name	Composition	Mono m/z	Score
target	WEPPLDDGGSEIIN(1419.5022)YTLEK	1199.1871	6.4289
decoy	SGKAAFFEDPHMDN(1297.4379)ESPKYK	1199.1673	17.757
decoy	DMGVVFFDRSTAQNAN(1278.4196)DTNSK	1199.1619	11.0005
decoy	FYEN(1260.2990)KSVRQHMLQLNNWK	1199.1634	8.5615
decoy	NMFESDLVN(1536.6530)DSRKDCTK	1199.1965	7.3229
decoy	YTWNQLLKHQMGQN(1261.4319)DSGWK	1199.185	5.9909
decoy	SHYAHKFNCN(1267.3433)FTMHLRHK	1199.1535	4.5889
decoy	N(1296.5354)STTTDNTDKDLDHSYCRR	1199.1835	4.4139
decoy	WGYDFVN(1361.4380)YTARKVDLVCK	1199.1863	4.397
decoy	NCVKEHVYENREN(1388.5277)VTCR	1199.1835	3.5648
decoy	ACYMVYN(1575.6397)WSARWVEGK	1199.1916	3.1565
decoy	WFTHCN(1580.6961)NTDGLDCCSK	1199.1649	2.8399
decoy	VQNCN(1409.4679)WTAFVSQSTDATTR	1199.1586	2.6244
decoy	QN(1391.4678)WTKNWPNEHNFFTLK	1199.1832	2.5942
decoy	HYLCN(1376.5606)TTEYHTFQEGYR	1199.1788	2.3244
decoy	VGN(1340.3951)YTGTEAPKVWCVVRYR	1199.1829	2.1679
decoy	QNFSRN(1578.6032)HSWKGNGVEEK	1199.1938	0
decoy	KYFVDVGLSN(1518.4152)RTEEYTR	1199.1549	0
decoy	WFFNQTNYDN(1546.5598)NTTEVR	1199.1614	0
decoy	HKMFYFTNWHN(1470.5188)DTGQAK	1199.1681	0
decoy	SLN(1471.4626)WSAKPDMSTATRMTAR	1199.1699	0

Name	Composition	Mono m/z	Score
target	TCILEILN(2303.8409)STK	1199.1827	5.4209
decoy	ALHFAAGN(2280.7746)VTWK	1199.1615	7.5504
decoy	PWWEN(2291.8156)KSTQK	1199.1577	6.5228
decoy	YCSLN(2314.9140)TSGNHK	1199.1658	5.9909
decoy	WQQPN(2341.8339)HTKSK	1199.1624	5.126
decoy	N(2283.9264)TSDQWRYNK	1199.1829	4.2816
decoy	GDSN(2245.8074)RTPGYKVR	1199.1714	4.1285
decoy	RYGN(2168.8991)QSTNACQK	1199.1876	3.8859
decoy	YLN(2283.7518)GSTSKWKK	1199.1573	3.5251
decoy	WN(2262.7918)ASADTKQRR	1199.1611	3.1565
decoy	N(2145.8925)VTAPSSMDEQDR	1199.1782	2.9745
decoy	GRWN(2239.6750)LSVRVLR	1199.1638	0
decoy	PAEN(2293.9808)CSEPNRK	1199.1952	0
decoy	N(2264.8949)WTMNEWHGR	1199.1614	0
decoy	N(2458.0150)GTQEFQWK	1199.1873	0
decoy	GSHLAKGRN(2252.8488)CSR	1199.1797	0
decoy	YVCN(2246.8510)KTYNMR	1199.1598	0
decoy	DYN(2392.8995)STFKAEK	1199.161	0
decoy	YN(2346.7617)STVVNPVKK	1199.157	0
decoy	MN(2457.9191)ATVWFLR	1199.1737	0
decoy	GCN(2379.0307)DTVEYMK	1199.1809	0

Name	Composition	Mono m/z	Score
target	LN(1378.4328)GSAPIQVCWYRDGVLLR	1199.2024	2.0622
decoy	RSVTNLN(1363.3711)KTGTNPGDEYLPR	1199.1742	15.3993
decoy	NWWSCWAN(1182.5144)CTRVEGGASMLK	1199.1975	9.6913
decoy	APRPCCN(1518.5499)NSRCTQNWR	1199.1556	8.881
decoy	QAVVRKLAYYFN(1550.4671)NTAMR	1199.1856	6.6824
decoy	N(1279.4870)DTQEDEPRAPNASVRMNSGK	1199.1883	5.9909
decoy	VVLSLCN(1278.3198)CTRDYVYKSPDK	1199.1553	5.2002
decoy	DCHAACWWDAKVRN(1375.5599)TSDK	1199.1818	4.6675
decoy	GGGEEQMAYVKYN(1305.5171)ESQEENK	1199.1748	4.6356
decoy	EHNFGPVYCTGN(1356.5390)WSMFHR	1199.1711	4.4829
decoy	GDLCMLYRGVN(1529.6517)TTYDCK	1199.1935	4.1652
decoy	ETDNAHNLEEPSCN(1252.4670)LSGNWR	1199.1624	3.5648
decoy	SCDHAFMPFN(1323.6879)CSGWHCMK	1199.1892	3.5251
decoy	N(1230.5094)MSETQFEFWYMMMYNGR	1199.1595	3.1845
decoy	MGN(1535.5851)WTWMKKWQEGYSK	1199.1838	2.3039
decoy	AAKAEN(1402.4966)FSYSQKLSCSHHK	1199.1872	2.1122
decoy	PFN(1338.4314)MTGLENKYNDRHQHR	1199.1713	1.7558
decoy	PN(1379.3610)ASLDPMVVRQEFECPVK	1199.1532	1.7558
decoy	WCFGN(1356.5501)SSNDWRDLRGNVR	1199.1952	0
decoy	VVEWDHRGFDNAGGN(1445.5196)HSPR	1199.1751	0
decoy	EAQEFVMHN(1324.3931)RSQVDLWPGK	1199.1684	0

Supplementary Table 1. The decoy glycopeptides and their scoring results returned by GPE for the four incorrect glycopeptides that were input as target candidates (labeled in red). The four target glycopeptide compositions are NAAG*N*FSEPSDSSGAITAR+ [Hex]7[HexNAc]3 (3+, *m/z* 1199.16), WEPPLDDGGSEII*N*YTLEK+ [Hex]5[HexNAc]3 (3+, *m/z* 1199.19), TCILEIL*N*STK +[Hex]4[HexNAc]6[Fuc]1[Neu5Ac]1(3+, *m/z* 1199.18), and L*N*GSAPIQVCWYRDGVLLR+ [Hex]3[HexNAc]4[SO3]1(3+, *m/z* 1199.20) (*N* is the glycosylation site).