PROTEOMICS

Supporting Information for Proteomics DOI 10.1002/pmic.201500526

Chih-Chiang Tsou, Chia-Feng Tsai, Guo Ci Teo, Yu-Ju Chen
and Alexey I. Nesvizhskii

Untargeted, spectral library-free analysis of data-independent acquisition proteomics data generated using Orbitrap mass spectrometers

Supplementary information for "Untargeted, spectral library-free analysis of data independent acquisition proteomics data generated using Orbitrap mass spectrometers"

Chih-Chiang Tsou, Chia-Feng Tsai, Guoci Teo, Yu-Ju Chen, Alexey I. Nesvizhskii

Supplementary Table 1. Detailed identification results for HEK-293 Q Exactive dataset

Peptide ion IDs (1% Run level FDR): The number of peptide ion identifications determined at 1% individual run level FDR threshold for each run. **Peptide ion IDs (1% Dataset level FDR):** The number of peptide ion identifications at 1% dataset level FDR threshold for each run. For DIA datasets, the numbers include the additional IDs from targeted re-extraction (with a 0.99 probability threshold). **Peptide ion ID coverage (Dataset level):** Percent of peptide ion identifications from the 1% Dataset level FDR peptide ion list that were identified in that particular run. **Protein IDs (1% Run level FDR):** The number of protein identifications at 1% individual run level FDR threshold for each run. **Protein IDs (1% Dataset level FDR):** The number of protein identifications at 1% Dataset level FDR threshold for each run. **Protein ID coverage (Dataset level):** Percent of protein identifications from the 1% Dataset level FDR protein master list identified in that particular run. See Methods for details.

	Peptide ion IDs (1% Run level FDR)	Peptide ion IDs (1% Dataset level FDR)	Peptide ion ID coverage (Dataset level)	Protein IDs (1% Run level FDR)	Protein IDs (1% Dataset level FDR)	Protein ID coverage (Dataset level)
S1_R1_DIA	19,945	24,216	70.2%	2,774	3,359	88.3%
S1_R2_DIA	19,836	24,440	70.9%	2,818	3,365	88.5%
S1_R3_DIA	19,075	23,413	67.9%	2,670	3,320	87.3%
S2_R1_DIA	20,271	24,592	71.3%	2,790	3,402	89.5%
S2_R2_DIA	19,548	24,277	70.4%	2,672	3,329	87.6%
S2_R3_DIA	18,650	23,621	68.5%	2,656	3,284	86.4%
S3_R1_DIA	19,673	23,881	69.3%	2,774	3,346	88.0%
S3_R2_DIA	19,386	24,336	70.6%	2,724	3,402	89.5%
S3_R3_DIA	18,693	24,098	69.9%	2,575	3,322	87.4%
S4_R1_DIA	20,491	24,614	71.4%	2,831	3,413	89.8%
S4_R2_DIA	19,748	24,702	71.7%	2,786	3,415	89.8%
S4_R3_DIA	18,662	23,863	69.2%	2,657	3,288	86.5%
S5_R1_DIA	20,864	24,913	72.3%	2,812	3,409	89.7%
S5_R2_DIA	19,749	24,258	70.4%	2,636	3,332	87.6%
S5_R3_DIA	17,611	24,093	69.9%	2,538	3,344	88.0%
S6_R1_DIA	20,037	23,844	69.2%	2,727	3,349	88.1%
S6_R2_DIA	19,893	24,297	70.5%	2,679	3,373	88.7%
S6_R3_DIA	17,831	23,295	67.6%	2,519	3,253	85.6%
S7_R1_DIA	20,279	24,484	71.0%	2,726	3,351	88.1%
S7_R2_DIA	18,703	23,765	68.9%	2,580	3,308	87.0%
S7_R3_DIA	18,292	23,173	67.2%	2,473	3,229	84.9%
S8_R1_DIA	19,710	23,733	68.8%	2,633	3,283	86.3%
S8_R2_DIA	19,270	23,328	67.7%	2,571	3,265	85.9%
S8_R3_DIA	16,343	21,827	63.3%	2,344	3,118	82.0%
S1_R1_DDA	17,823	18,194	46.2%	2,692	2,930	77.1%

S1_R2_DDA	17,459	17,821	45.3%	2,712	2,931	77.1%
S1_R3_DDA	17,109	17,446	44.3%	2,670	2,870	75.5%
S2_R1_DDA	17,625	17,952	45.6%	2,638	2,912	76.6%
S2_R2_DDA	17,074	17,440	44.3%	2,585	2,885	75.9%
S2_R3_DDA	16,608	16,972	43.1%	2,595	2,834	74.6%
S3_R1_DDA	17,319	17,599	44.7%	2,639	2,885	75.9%
S3_R2_DDA	17,938	18,134	46.1%	2,726	2,945	77.5%
S3_R3_DDA	16,536	17,057	43.3%	2,570	2,868	75.5%
S4_R1_DDA	18,543	18,776	47.7%	2,782	2,996	78.8%
S4_R2_DDA	18,231	18,316	46.5%	2,756	2,932	77.1%
S4_R3_DDA	16,496	16,959	43.1%	2,556	2,805	73.8%
S5_R1_DDA	17,938	18,276	46.4%	2,708	2,930	77.1%
S5_R2_DDA	17,162	17,390	44.2%	2,567	2,785	73.3%
S5_R3_DDA	16,703	16,944	43.0%	2,618	2,801	73.7%
S6_R1_DDA	17,645	18,088	45.9%	2,611	2,905	76.4%
S6_R2_DDA	18,030	18,243	46.3%	2,692	2,865	75.4%
S6_R3_DDA	15,940	16,388	41.6%	2,476	2,747	72.3%
S7_R1_DDA	17,539	17,951	45.6%	2,623	2,882	75.8%
S7_R2_DDA	17,688	17,891	45.4%	2,675	2,879	75.7%
S7_R3_DDA	16,283	16,847	42.8%	2,508	2,793	73.5%
S8_R1_DDA	17,893	18,021	45.8%	2,690	2,879	75.7%
S8_R2_DDA	17,198	17,414	44.2%	2,555	2,820	74.2%
S8_R3_DDA	14,813	15,262	38.8%	2,410	2,658	69.9%

Supplementary Table 2. Detailed identification results of the microtissue Q Exactive dataset

Peptide ion IDs (1% Run level FDR): The number of peptide ion identifications determined at 1% individual run level FDR threshold for each run. **Peptide ion IDs (1% Dataset level FDR):** The number of peptide ion identifications at 1% dataset level FDR threshold for each run. For DIA datasets, the numbers include the additional IDs from targeted re-extraction (with a 0.99 probability threshold). **Peptide ion ID coverage (Dataset level):** Percent of peptide ion identifications from the 1% Dataset level FDR peptide ion list that were identified in that particular run. **Protein IDs (1% Run level FDR):** The number of protein identifications at 1% individual run level FDR threshold for each run. **Protein IDs (1% Dataset level FDR):** The number of protein identifications at 1% Dataset level FDR threshold for each run. **Protein ID coverage (Dataset level):** Percent of protein identifications from the 1% Dataset level FDR protein master list identified in that particular run. See Methods for details.

File	Peptide ion IDs (1% Run level FDR)	Peptide ion IDs (1% Dataset level FDR)	Peptide ion ID coverage (Dataset level)	Protein IDs (1% Run level FDR)	Protein IDs (1% Dataset level FDR)	Protein ID coverage (Dataset level)
S1_DIA_R1	16,678	20,060	74.9%	1,889	2,341	88.6%
S1_DIA_R2	17,254	20,160	75.3%	1,921	2,333	88.3%
S1_DIA_R3	17,339	19,994	74.7%	1,921	2,355	89.2%
S3_DIA_R1	16,550	20,408	76.2%	1,828	2,341	88.6%
S3_DIA_R2	16,945	20,612	77.0%	1,891	2,368	89.7%
S3_DIA_R3	16,791	20,191	75.4%	1,881	2,332	88.3%
S4_DIA_R1	16,639	20,030	74.8%	1,818	2,293	86.8%
S4_DIA_R2	17,561	21,038	78.6%	1,893	2,393	90.6%
S4_DIA_R3	17,633	20,644	77.1%	1,900	2,369	89.7%
S7_DIA_R1	17,841	21,264	79.4%	1,970	2,396	90.7%
S7_DIA_R2	18,093	21,227	79.3%	1,996	2,412	91.3%
S7_DIA_R3	17,778	20,574	76.9%	1,926	2,375	89.9%
S9_DIA_R1	17,068	19,810	74.0%	1,896	2,318	87.8%
S9_DIA_R2	17,507	20,227	75.6%	1,896	2,365	89.5%
S9_DIA_R3	17,380	20,307	75.9%	1,969	2,356	89.2%
pool_DDA_R1	16,514	16,607	53.0%	2,156	2,253	81.1%
pool_DDA_R2	17,027	16,979	54.2%	2,150	2,255	81.2%
S1_DDA	12,529	13,195	42.1%	1,787	2,014	72.5%
S3_DDA	15,966	16,034	51.2%	2,115	2,206	79.4%
S7_DDA	16,846	16,857	53.8%	2,187	2,258	81.3%
S9_DDA	15,941	16,093	51.4%	2,121	2,229	80.2%

Supplementary Table 3. Detailed identification results for the Orbitrap Fusion dataset

Peptide ion IDs (1% Run level FDR): The number of peptide ion identifications determined at 1% individual run level FDR threshold for each run. **Peptide ion IDs (1% Dataset level FDR):** The number of peptide ion identifications at 1% dataset level FDR threshold for each run. For DIA datasets, the numbers include the additional IDs from targeted re-extraction (with a 0.99 probability threshold). **Peptide ion ID coverage (Dataset level):** Percent of peptide ion identifications from the 1% Dataset level FDR peptide ion list that were identified in that particular run. **Protein IDs (1% Run level FDR):** The number of protein identifications at 1% individual run level FDR threshold for each run. **Protein IDs (1% Dataset level FDR):** The number of protein identifications at 1% Dataset level FDR threshold for each run. **Protein ID coverage (Dataset level):** Percent of protein identifications from the 1% Dataset level FDR protein master list identified in that particular run. See Methods for details.

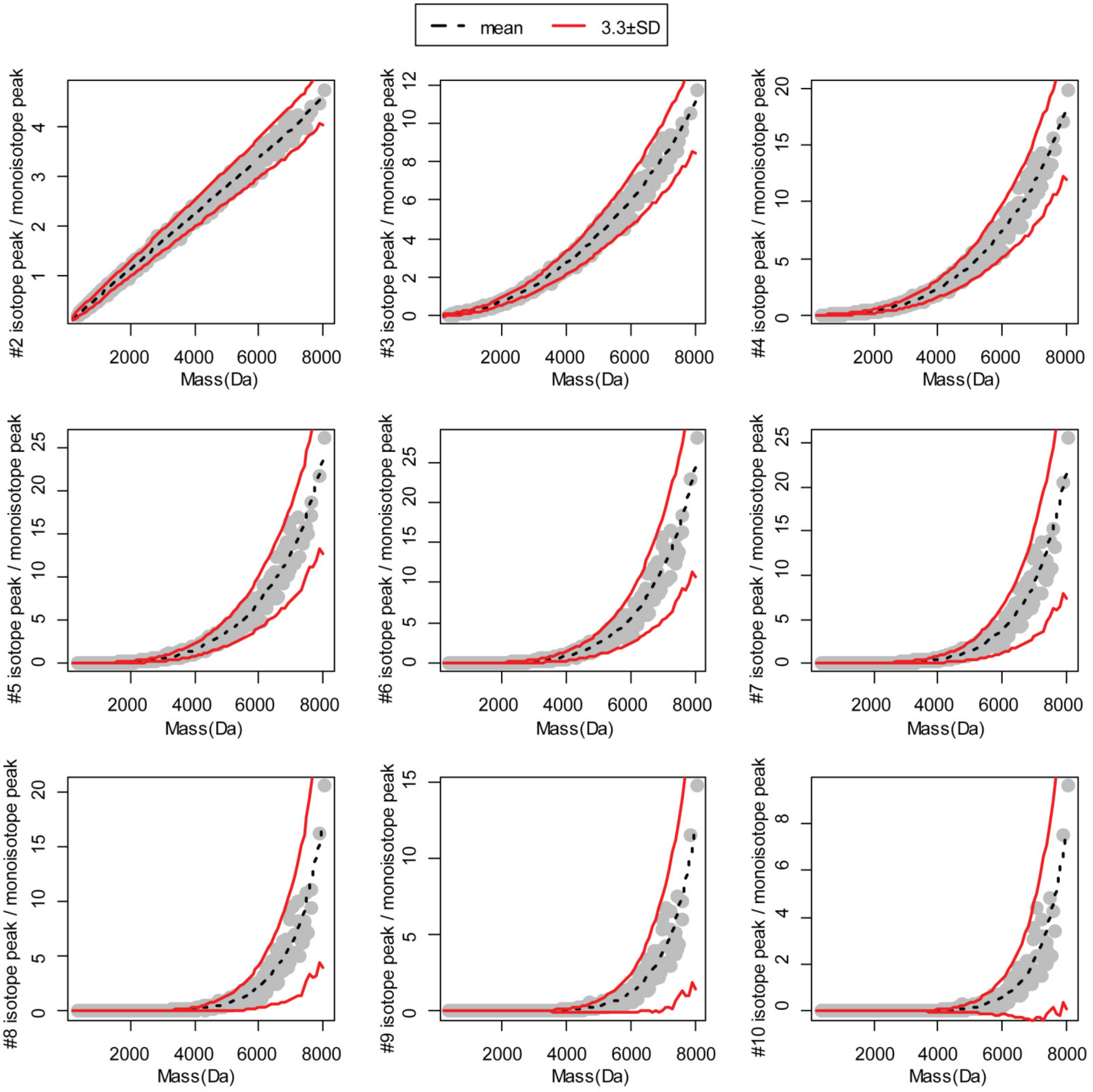
File	Peptide ion IDs (1% Run level FDR)	Peptide ion IDs (1% Dataset level FDR)	Peptide ion ID coverage (Dataset level)	Protein IDs (1% Run level FDR)	Protein IDs (1% Dataset level FDR)	Protein ID coverage (Dataset level)
DIA 5Da R1	28,719	30,336	76.6%	3,846	4,066	92.3%
DIA 5Da R2	29,434	31,014	78.3%	3,854	4,101	93.1%
DIA 5Da R3	29,341	30,604	77.3%	3,858	4,101	93.1%
DIA 10Da R1	31,941	34,117	82.6%	3,691	4,082	93.2%
DIA 10Da R2	33,159	34,946	84.6%	4,009	4,220	96.3%
DIA 10Da R3	33,449	34,818	84.3%	3,962	4,190	95.6%
DIA 15Da R1	29,862	31,419	86.2%	3,545	3,788	95.7%
DIA 15Da R2	29,953	31,494	86.4%	3,598	3,797	95.9%
DIA 15Da R3	29,783	31,514	86.5%	3,616	3,818	96.4%
DIA 20Da R1	26,964	28,606	86.0%	3,342	3,547	96.1%
DIA 20Da R2	26,605	28,419	85.5%	3,348	3,530	95.6%
DIA 20Da R3	26,739	28,605	86.0%	3,330	3,532	95.7%
DIA 25Da R1	23,924	25,926	85.9%	3,125	3,373	96.0%
DIA 25Da R2	23,880	25,956	86.0%	3,052	3,367	95.8%
DIA 25Da R3	24,199	26,033	86.3%	3,101	3,385	96.3%
DDA1 R1	31,851	32,011	79.2%	4,256	4,378	91.6%
DDA1 R2	31,732	31,944	79.0%	4,257	4,409	92.3%
DDA1 R3	32,003	32,143	79.5%	4,314	4,413	92.3%
DDA2 R1	29,623	30,075	71.8%	4,102	4,284	90.6%
DDA2 R2	29,813	30,186	72.1%	4,123	4,310	91.1%
DDA2 R3	30,813	30,847	73.7%	4,227	4,319	91.3%

Supplementary Table 4. DIA-Umpire v2 computation time and size of generated pseudo MS/MS spectra

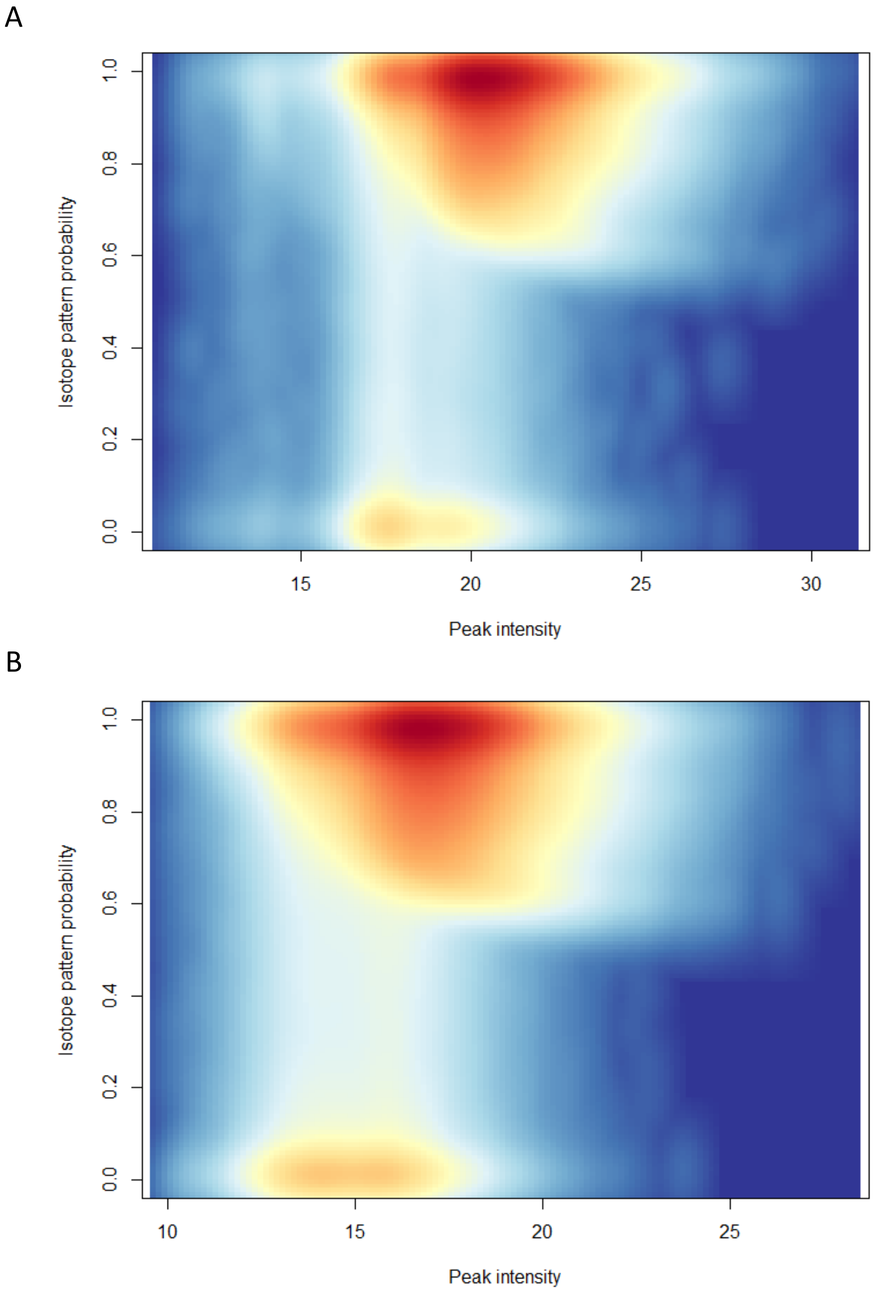
DIA run	Isotope pattern probability threshold	Fraction mass filter applied	Processing time (hours)	mgf file size (MB)			No. of pseudo MS/MS spectra		
				Q1	Q2	Q3	Q1	Q2	Q3
Hela1ug_DIA_10Da_150226_01	0	FALSE	2.74	349.11	546.14	110.27	83,232	142,860	29,246
	0	TRUE	2.36	309.73	359.89	102.74	75,433	97,196	27,938
	0.3	TRUE	2.21	296.81	330.08	90.85	72,183	88,739	24,573
	0.6	TRUE	2.15	282.68	243.18	73.31	68,677	65,991	19,570
	0.9	TRUE	1.89	201.61	132.21	42.79	48,871	35,685	11,291
B_D140314_SGSDSsample1_R01	0	FALSE	7.64	591.9	1187.96	459.01	146,277	293,648	105,208
	0	TRUE	5.7	516.65	765.8	397.71	120,282	179,786	92,798
	0.3	TRUE	5.59	479.47	677.1	351	111,023	157,739	82,169
	0.6	TRUE	4.53	440.52	452.84	266.06	99,582	103,652	62,348
	0.9	TRUE	3.68	286.19	228.08	148.69	63,966	51,981	35,008
Computer hardware specification and operating system: Intel Xeon E5645 CPU, 7 GB ram and single thread used in Java execution, x86_64 GNU/Linux operating system									

Supplementary Table 5. DIA-Umpire v2 identification performance for the original DIA-Umpire published AB Sciex 5600 datasets (*E. coli* and Human)

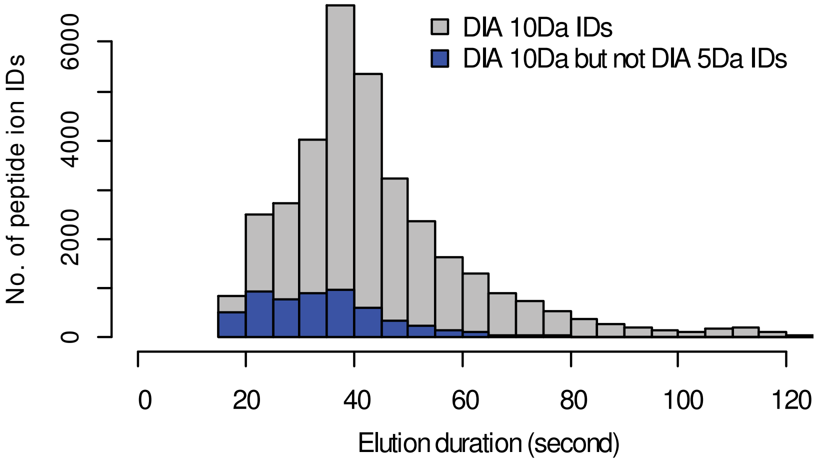
Dataset	File name	No. of protein IDs	No. of peptide ion IDs
E. coli	18484_REP3_1ug_Ecoli_NewStock2_SWATH_1	894	6692
	18486_REP3_1ug_Ecoli_NewStock2_SWATH_2	909	6806
Human	18300_REP2_500ng_HumanLysate_SWATH_1	1428	8927
	18302_REP2_500ng_HumanLysate_SWATH_2	1474	9429



Supplementary Figure 1. Theoretical intensity ratios of i^{th} isotope peak over monoisotope peak. Grey dots represent isotope peak intensity ratio between i^{th} isotope peak vs. monoisotope peak for tryptic peptides generated from human proteome sequences. In each plot, the grey dots were partitioned into 100 Da mass bins and mean and standard deviation (SD) for each bin were calculated. The black dash lines are the mean values of each 100 Da mass bin, and red solid lines represent the boundary for each bin calculated by mean ± 3.3 standard deviations.

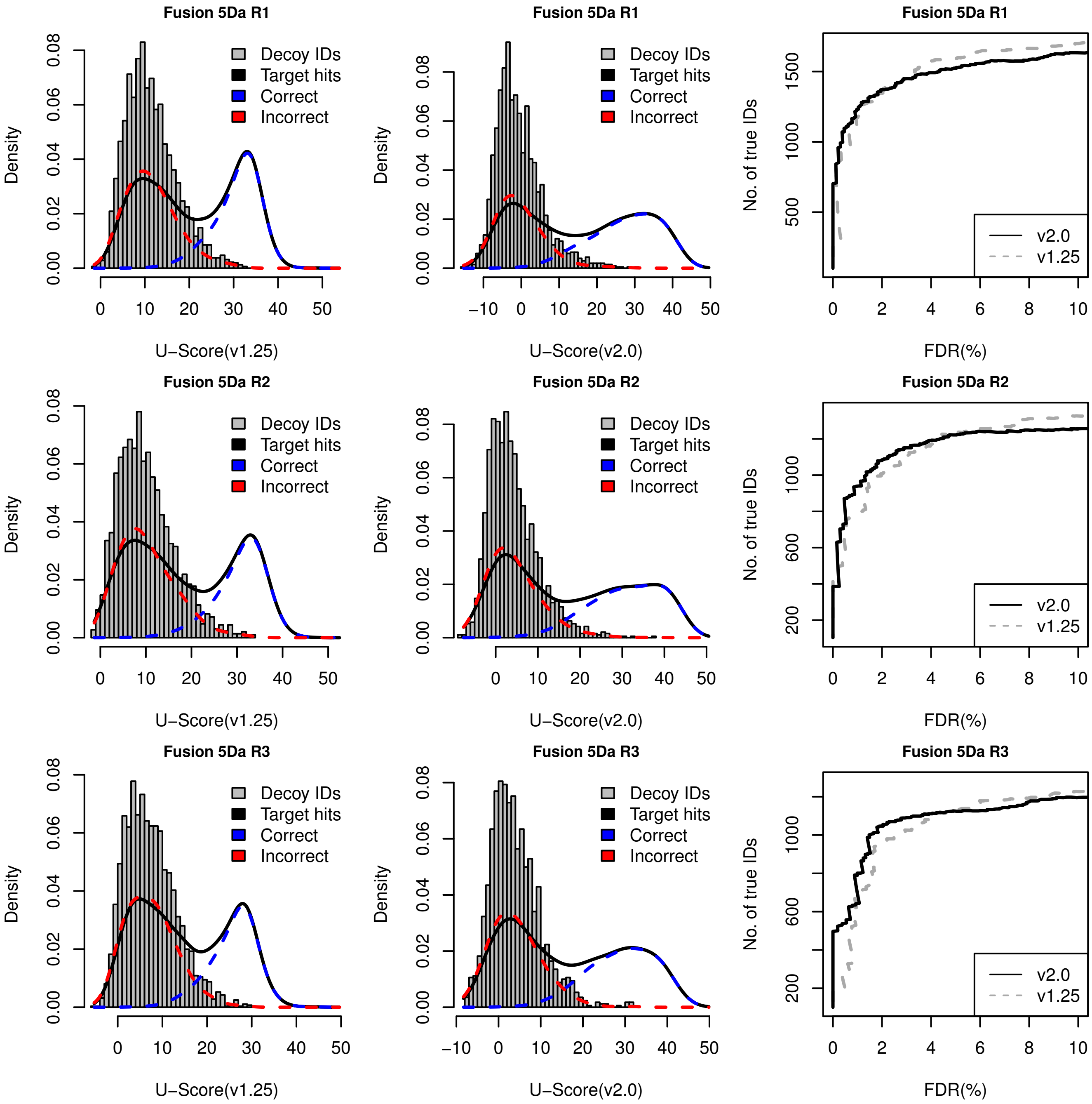


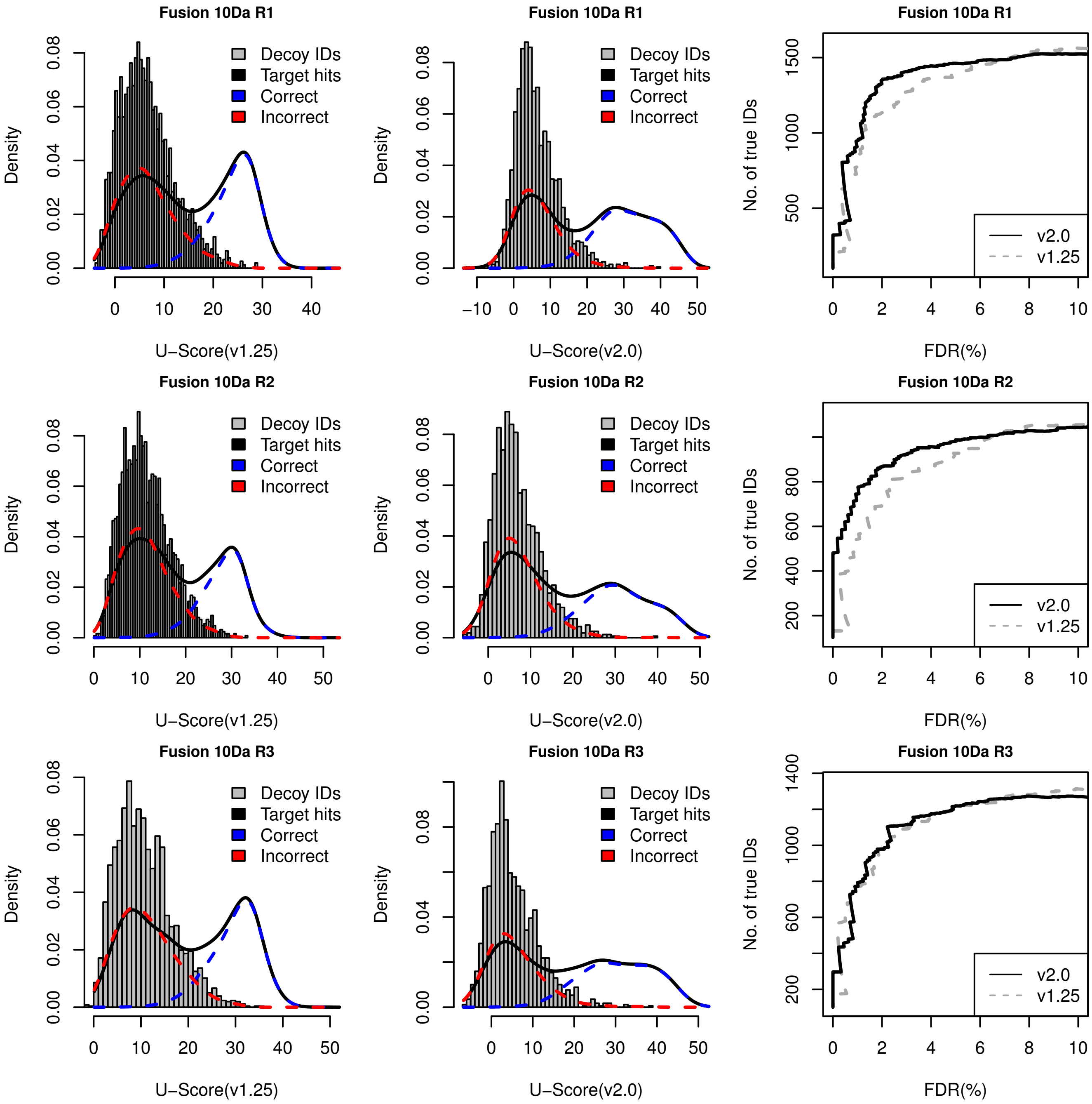
Supplementary Figure 2. Isotope pattern probabilities for all detected peak features plotted against monoisotope peak intensities. The color code indicates the number of detected features in a region of specific peak intensity and isotope pattern probability. (A) The result from the first replicate of the Orbitrap Fusion DIA 10 Da dataset. (B) Same as (A), the result for the first replicate of HEK-293 Q Exactive dataset.

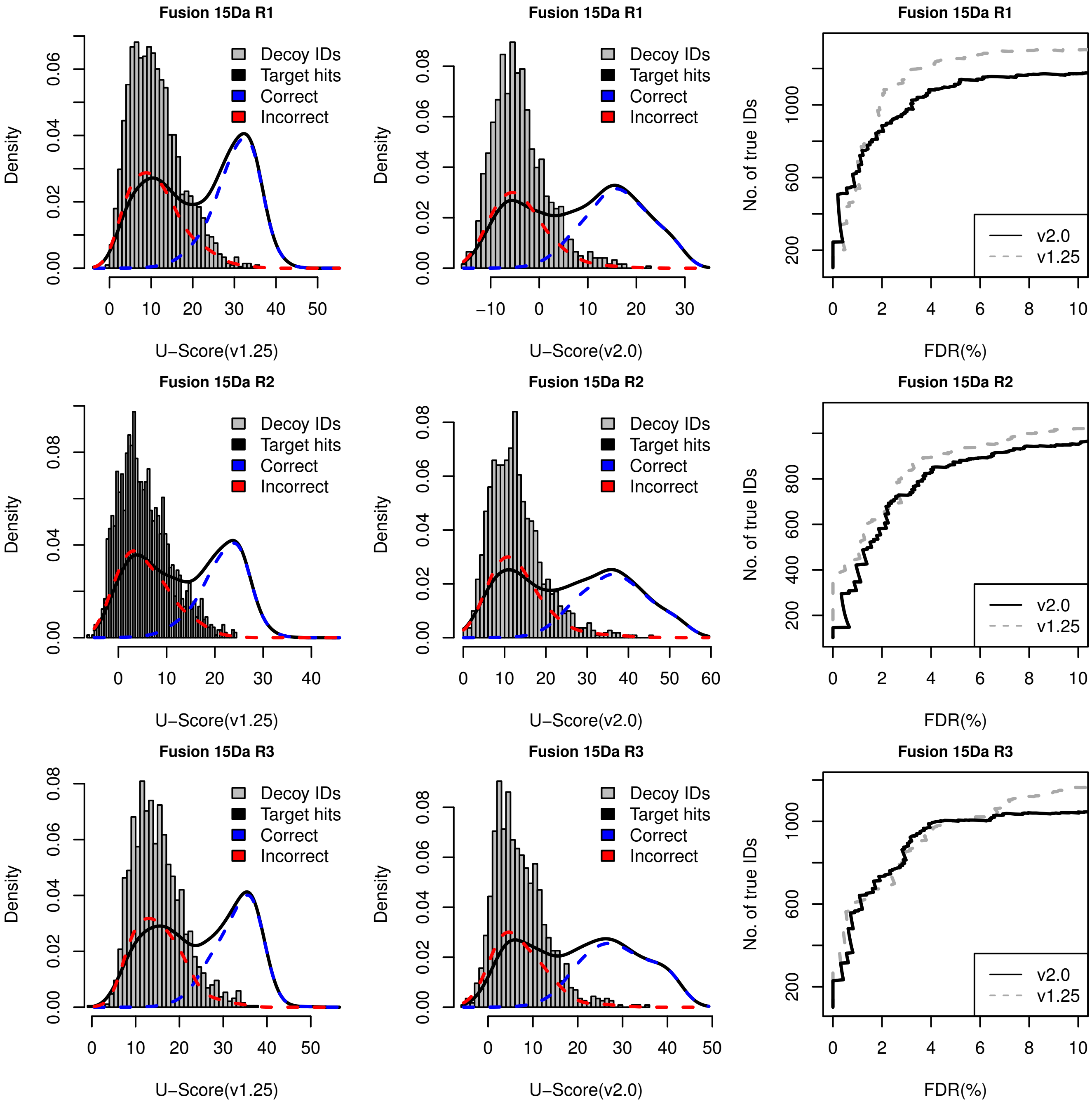


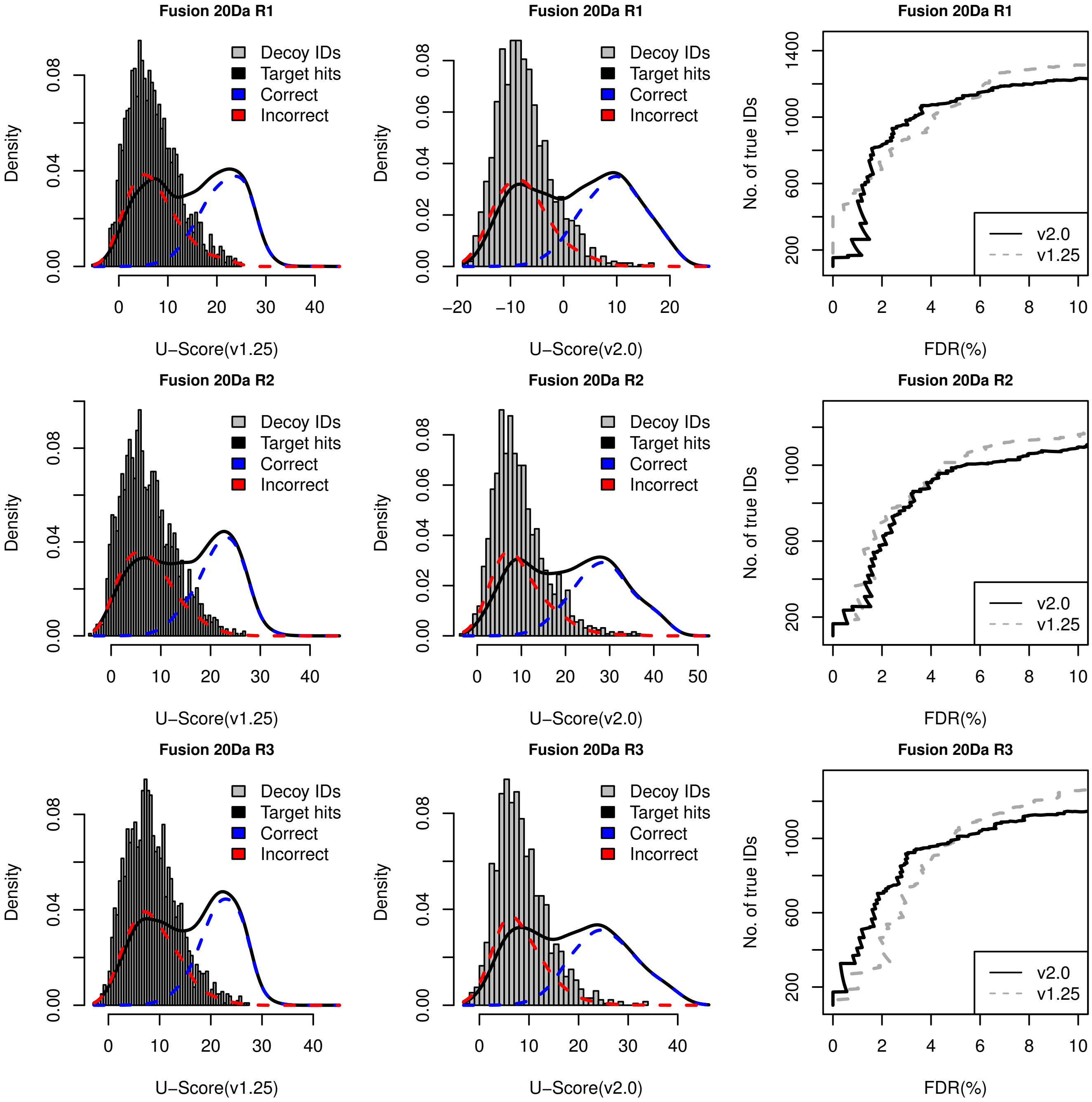
Supplementary Figure 3. Elution time duration of peptide ions in the first replicate of DIA 10 Da Orbitrap Fusion dataset. **Grey**: Histogram of identified peptide ion elution durations in the DIA run. **Dark Blue**: Histogram of the peptide ion elution durations which were identified in the replicate of DIA 10 Da dataset but not identified in any of DIA 5 Da replicates.

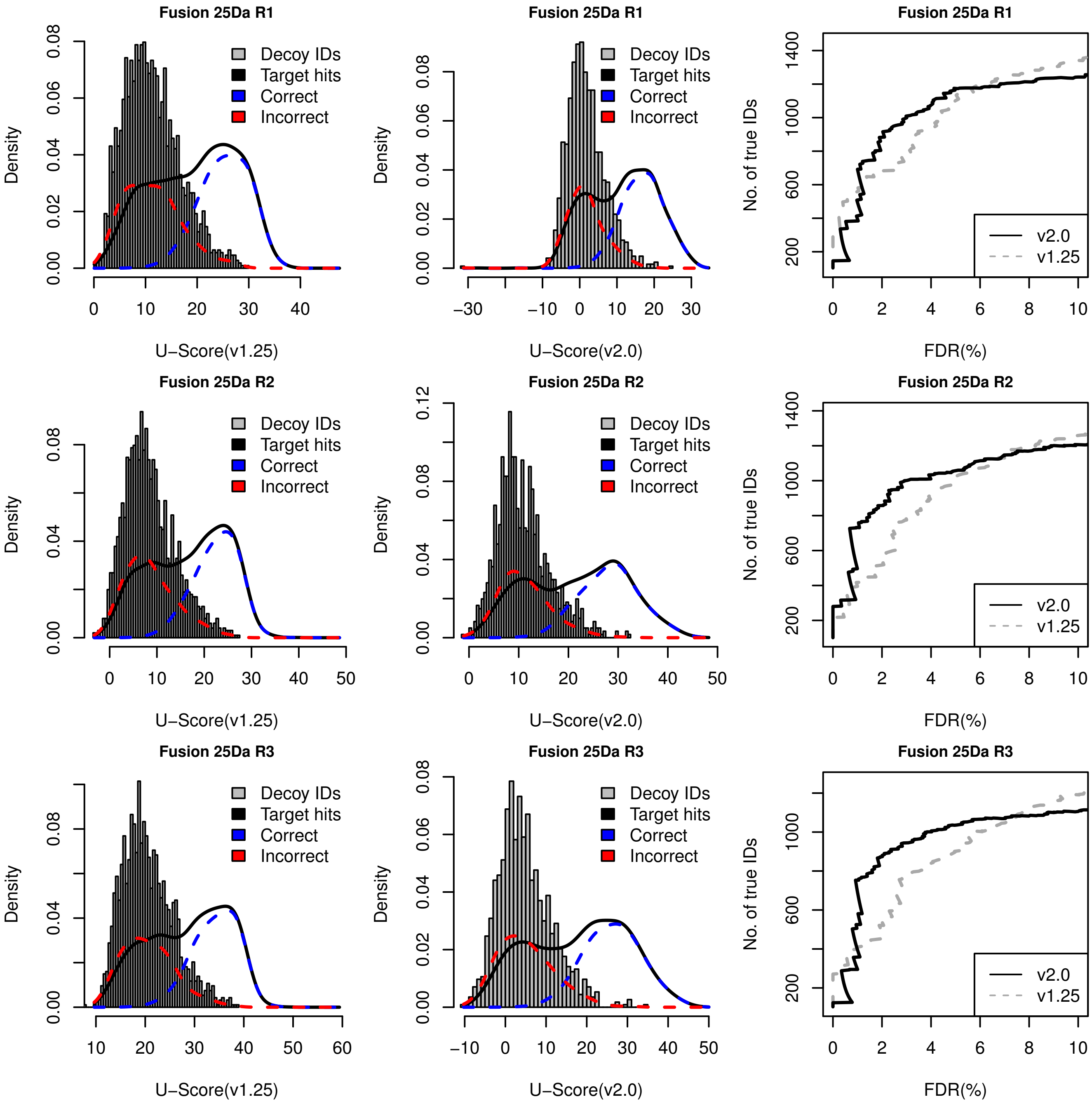
Supplementary Figure 4. The figures shown in the following pages are the comparisons between targeted re-extraction algorithms between DIA-Umpire v1.25 and v2. Each row shows the result for a DIA file. **Left**: Score histograms and parametric Gaussian mixture modeling result obtained from DIA-Umpire v1.25; **Middle**: Score histograms and semi-parametric mixture modeling result obtained from DIA-Umpire v2; **Right**: The numbers of targeted re-extraction identifications as a function of FDR obtained using DIA-Umpire v 1.25 and v2.

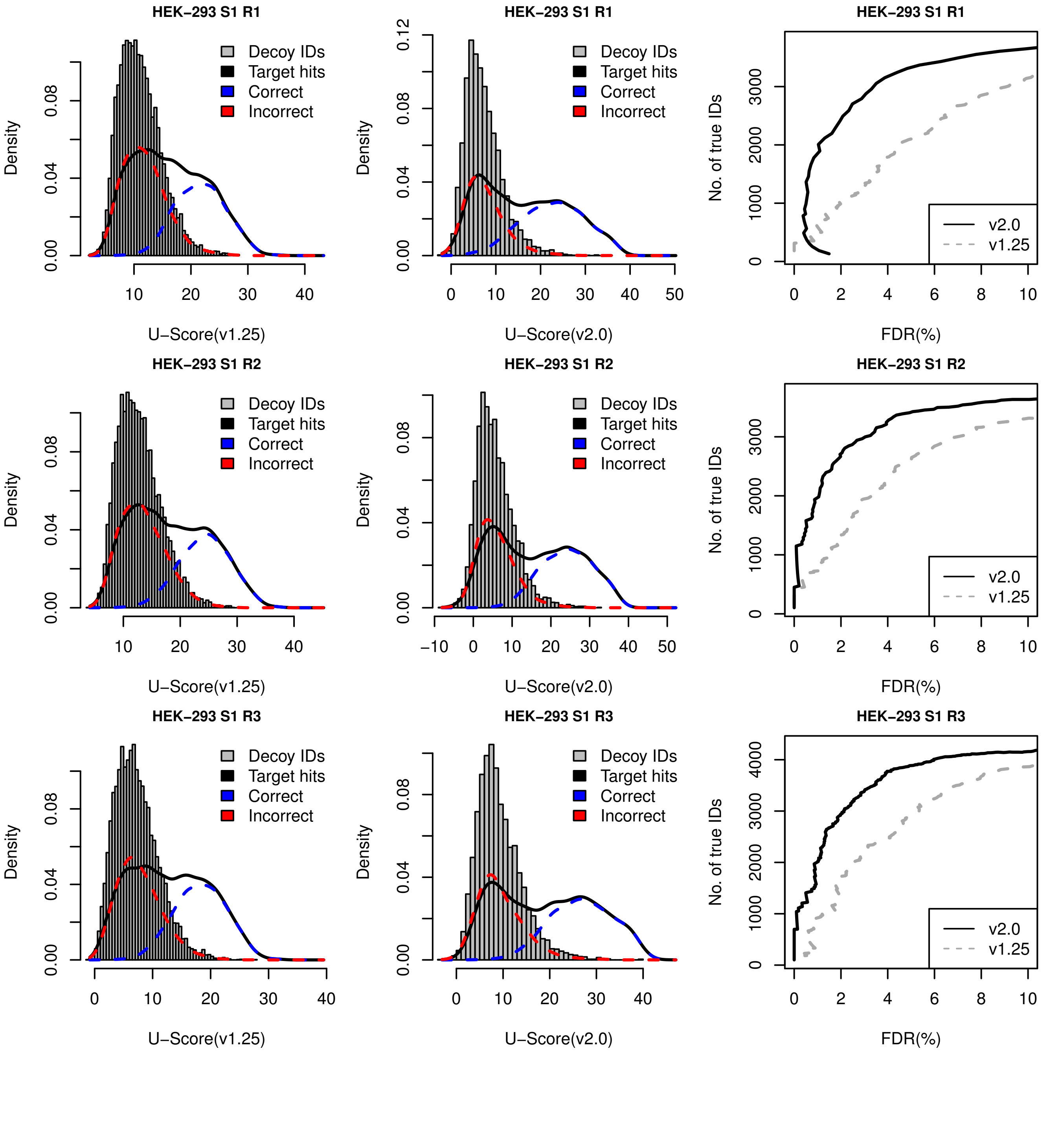

Fusion 5Da R1
Density
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
50
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 5Da R1
Density
−10
0
10
20
30
40
50
U−Score(v2.0)
Fusion 5Da R1
No. of true IDs
500
1000
1500
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
Fusion 5Da R2
U−Score(v1.25)
Fusion 5Da R2
U−Score(v2.0)
Fusion 5Da R2
200
600
1000
FDR(%)
Fusion 5Da R3
U−Score(v1.25)
Fusion 5Da R3
U−Score(v2.0)
Fusion 5Da R3
FDR(%)

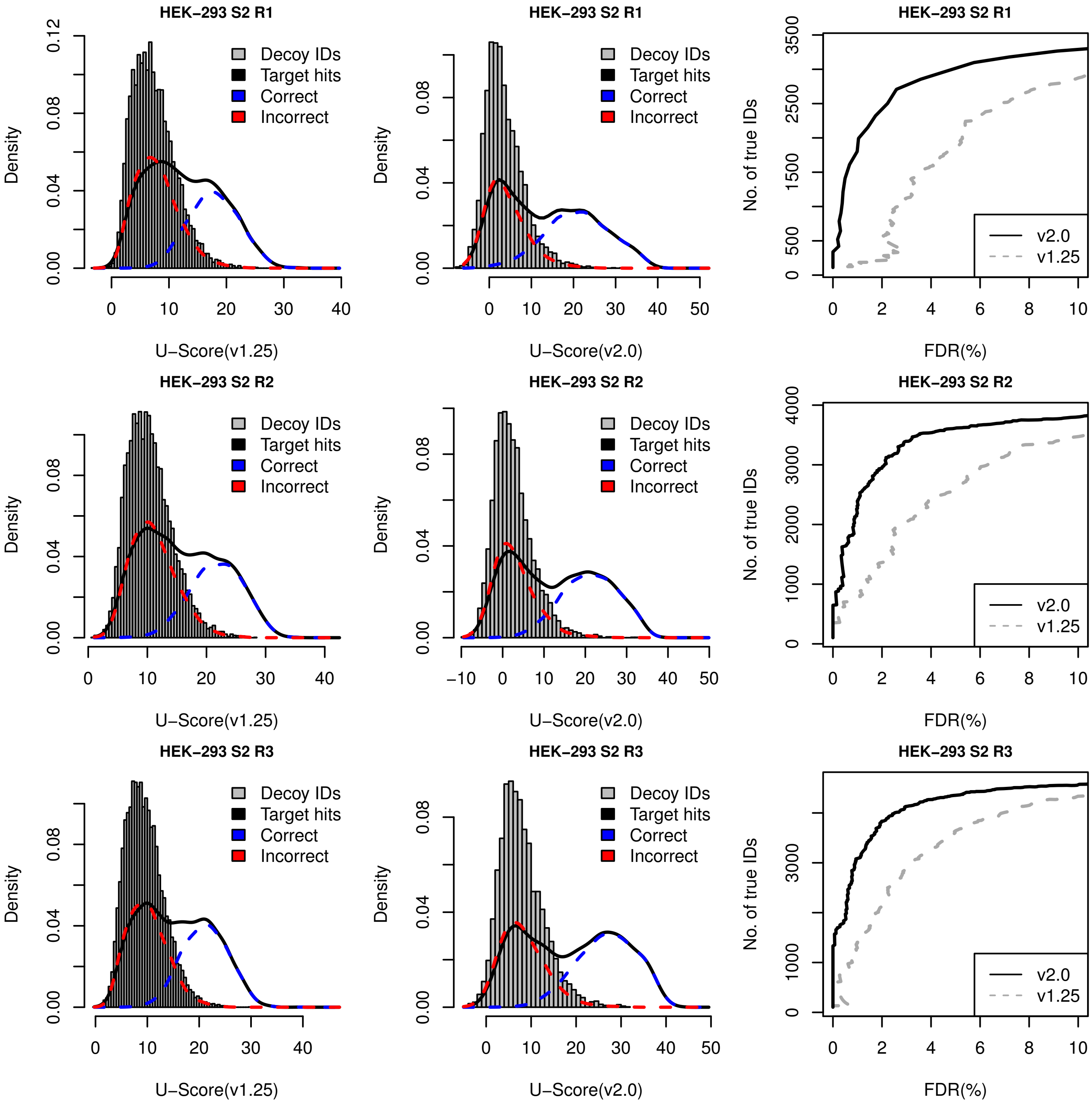

Fusion 10Da R1
Density
U–Score(v1.25)
U–Score(v2.0)
No. of true IDs
FDR(%)
Decoy IDs
Target hits
Correct
Incorrect
v2.0
v1.25
Fusion 10Da R2
Fusion 10Da R3
0.00
0.02
0.04
0.06
0.08
–10
0
10
20
30
40
50
500
1000
1500
200
400
600
800
1400
2
4
6
8

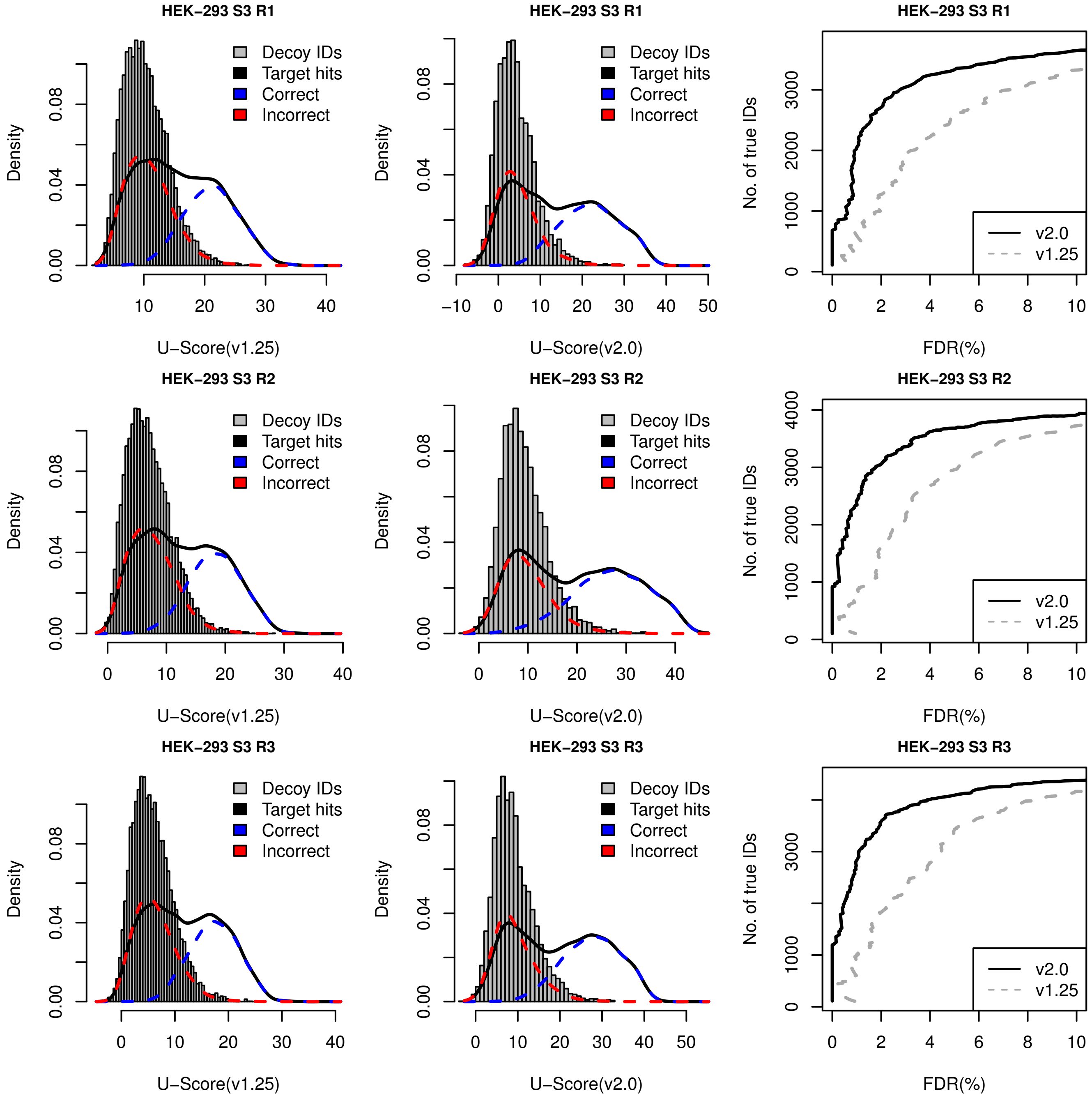

Fusion 15Da R1
Density
0.00
0.02
0.04
0.06
0
10
20
30
40
50
U−Score(v1.25)
▨ Decoy IDs
■ Target hits
🟦 Correct
🟥 Incorrect
Fusion 15Da R1
Density
0.00
0.02
0.04
0.06
0.08
−10
0
10
20
30
U−Score(v2.0)
▨ Decoy IDs
■ Target hits
🟦 Correct
🟥 Incorrect
Fusion 15Da R1
No. of true IDs
200
600
1000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
Fusion 15Da R2
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
▨ Decoy IDs
■ Target hits
🟦 Correct
🟥 Incorrect
Fusion 15Da R2
Density
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
50
60
U−Score(v2.0)
▨ Decoy IDs
■ Target hits
🟦 Correct
🟥 Incorrect
Fusion 15Da R2
No. of true IDs
200
400
600
800
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
Fusion 15Da R3
Density
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
50
U−Score(v1.25)
▨ Decoy IDs
■ Target hits
🟦 Correct
🟥 Incorrect
Fusion 15Da R3
Density
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
50
U−Score(v2.0)
▨ Decoy IDs
■ Target hits
🟦 Correct
🟥 Incorrect
Fusion 15Da R3
No. of true IDs
200
600
1000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25

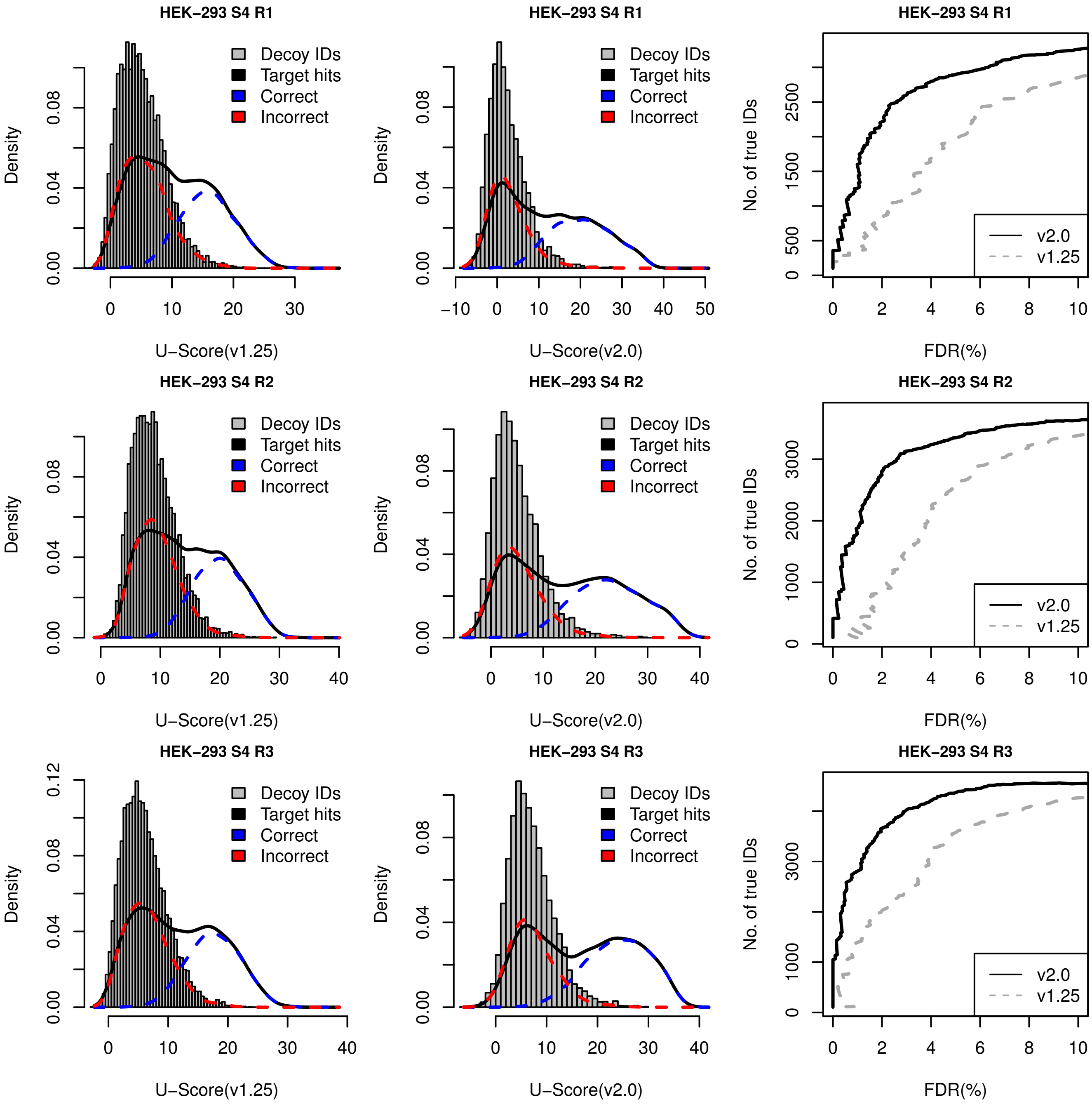

Fusion 20Da R1
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v1.25)
Fusion 20Da R1
Density
0.00
0.02
0.04
0.06
0.08
Decoy IDs
Target hits
Correct
Incorrect
−20
−10
0
10
20
U−Score(v2.0)
Fusion 20Da R1
No. of true IDs
200
600
1000
1400
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
Fusion 20Da R2
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v1.25)
Fusion 20Da R2
Density
0.00
0.02
0.04
0.06
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
50
U−Score(v2.0)
Fusion 20Da R2
No. of true IDs
200
600
1000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
Fusion 20Da R3
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v1.25)
Fusion 20Da R3
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v2.0)
Fusion 20Da R3
No. of true IDs
200
600
1000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)

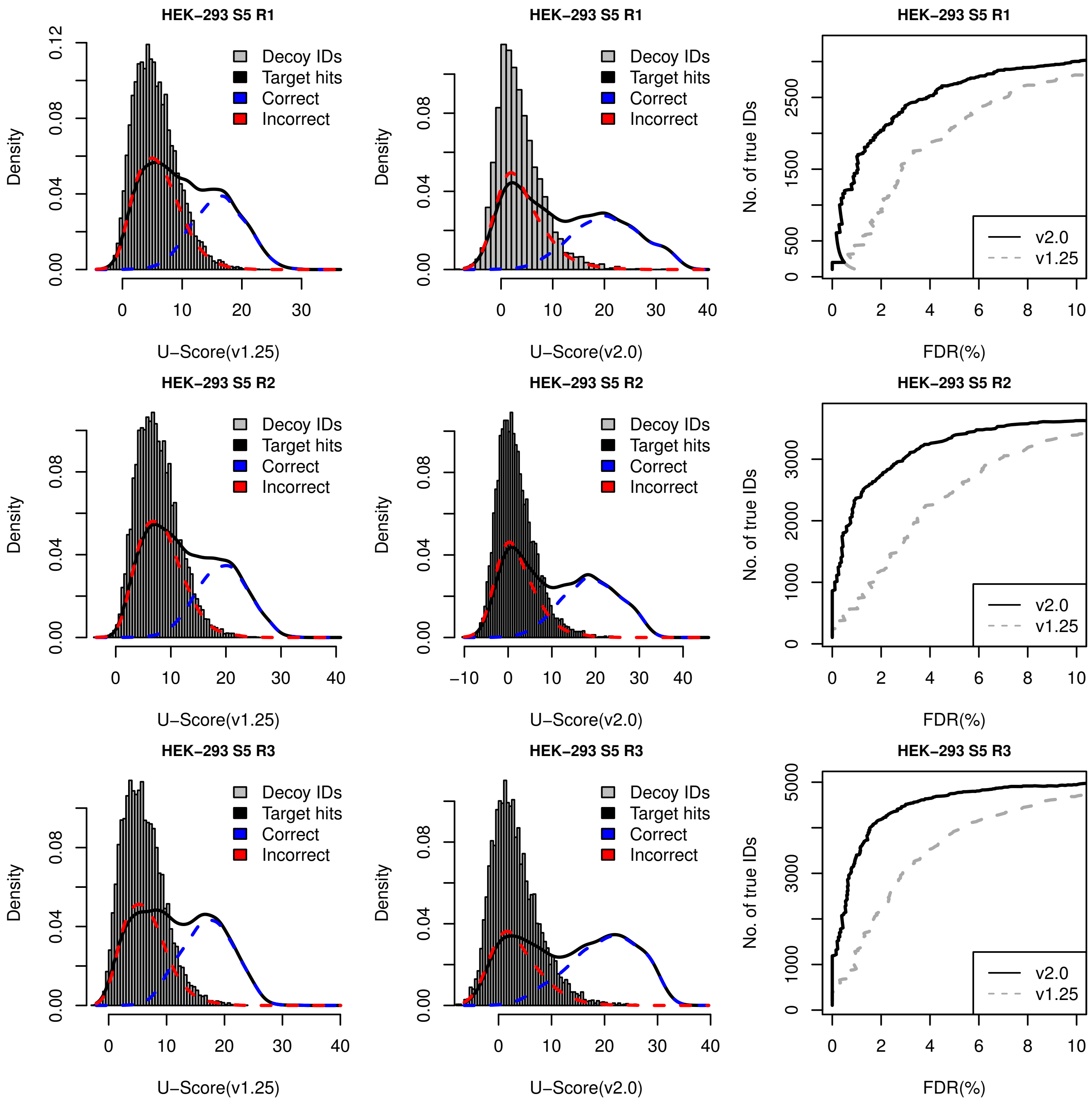

Fusion 25Da R1
Density
0.00
0.02
0.04
0.06
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 25Da R1
Density
0.00
0.02
0.04
0.06
0.08
−30
−10
0
10
20
30
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 25Da R1
No. of true IDs
200
600
1000
1400
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
Fusion 25Da R2
Density
0.00
0.04
0.08
0
10
20
30
40
50
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 25Da R2
Density
0.00
0.04
0.08
0.12
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 25Da R2
No. of true IDs
200
600
1000
1400
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
Fusion 25Da R3
Density
0.00
0.04
0.08
10
20
30
40
50
60
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 25Da R3
Density
0.00
0.02
0.04
0.06
0.08
−10
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
Fusion 25Da R3
No. of true IDs
200
600
1000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25

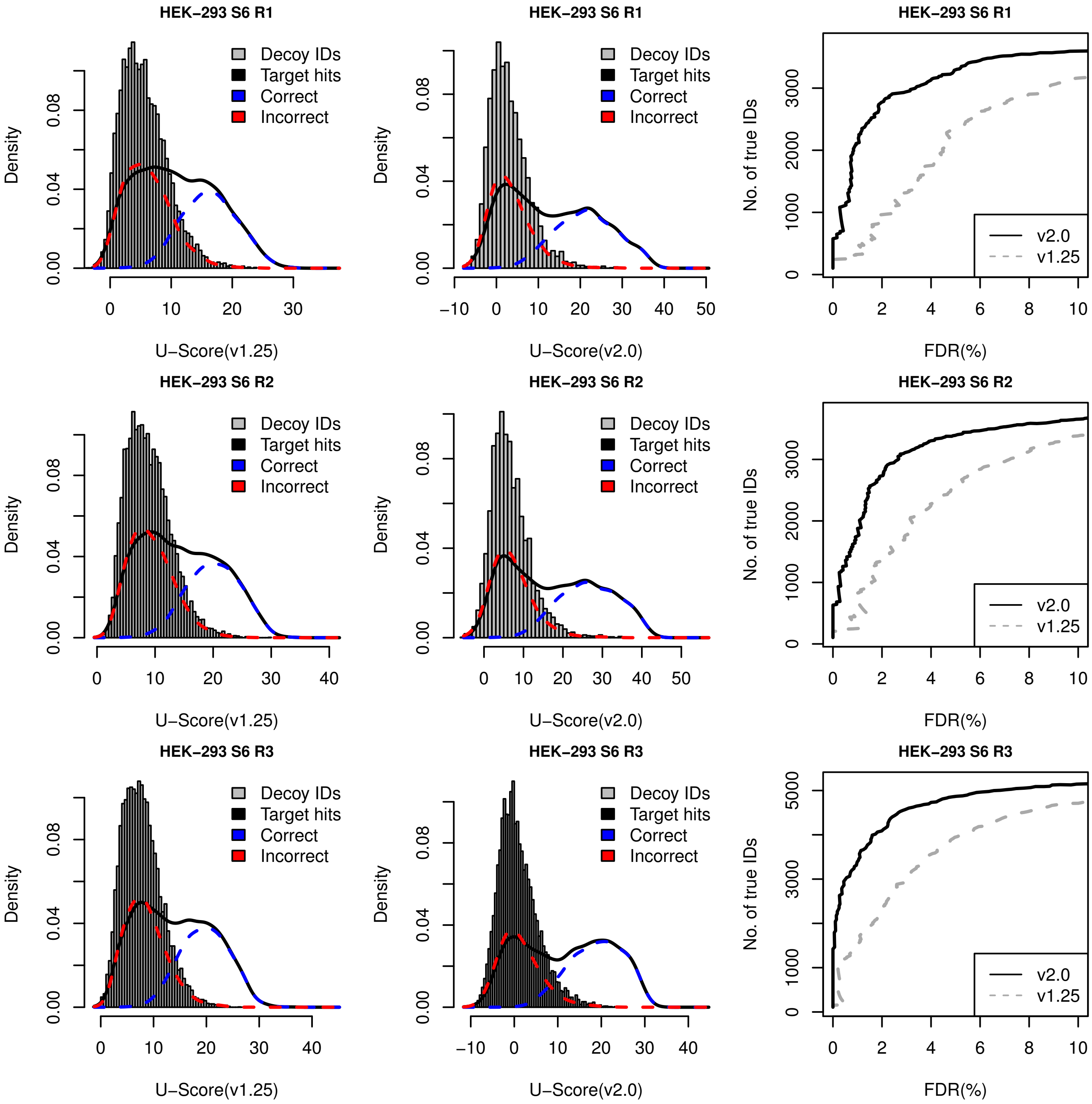

HEK−293 S1 R1
HEK−293 S1 R2
HEK−293 S1 R3
Density
No. of true IDs
U−Score(v1.25)
U−Score(v2.0)
FDR(%)
Decoy IDs
Target hits
Correct
Incorrect
v2.0
v1.25
0.00
0.04
0.08
0.12
0
1000
2000
3000
4000
−10
0
10
20
30
40
50
2
4
6
8

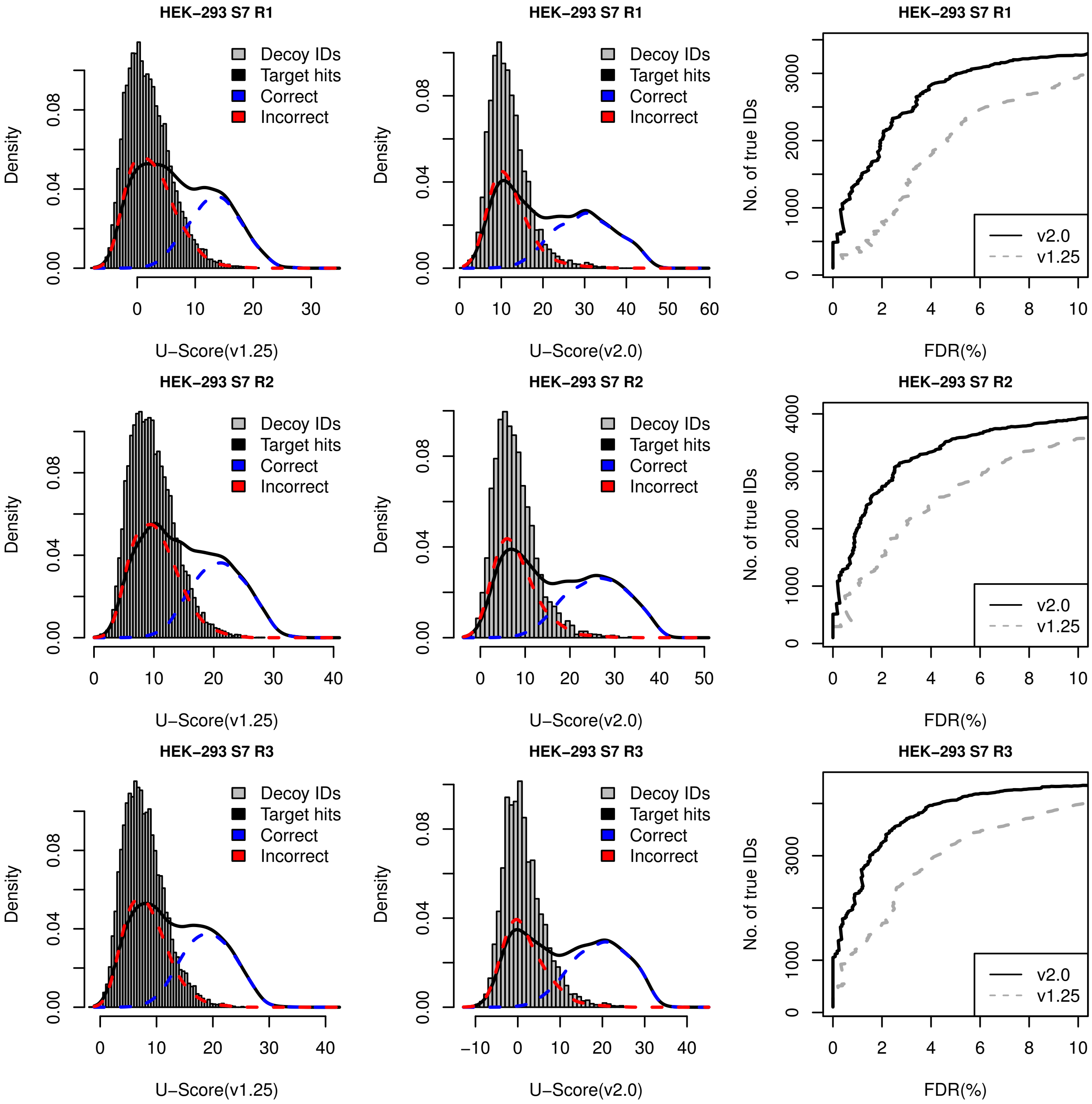

HEK−293 S2 R1
Density
0.00
0.04
0.08
0.12
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S2 R1
Density
0.00
0.04
0.08
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S2 R1
No. of true IDs
0
500
1500
2500
3500
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
HEK−293 S2 R2
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S2 R2
Density
0.00
0.04
0.08
−10
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S2 R2
No. of true IDs
0
1000
2000
3000
4000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
HEK−293 S2 R3
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S2 R3
Density
0.00
0.04
0.08
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S2 R3
No. of true IDs
0
1000
3000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25

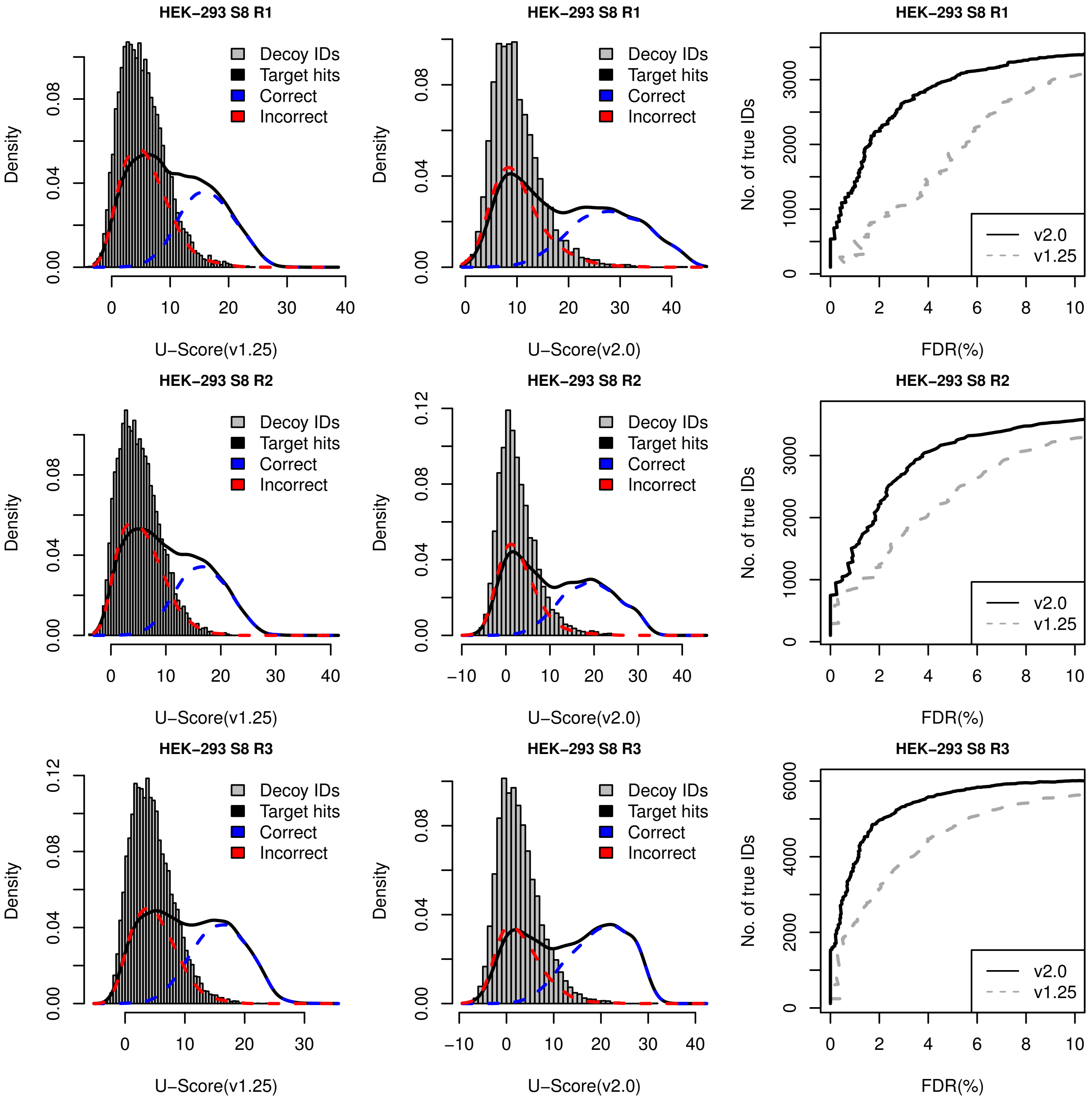

HEK−293 S3 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
10
20
30
40
U−Score(v1.25)
HEK−293 S3 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
−10
0
10
20
30
40
50
U−Score(v2.0)
HEK−293 S3 R1
No. of true IDs
0
1000
2000
3000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
HEK−293 S3 R2
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
HEK−293 S3 R2
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v2.0)
HEK−293 S3 R2
No. of true IDs
0
1000
2000
3000
4000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
HEK−293 S3 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
HEK−293 S3 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
50
U−Score(v2.0)
HEK−293 S3 R3
No. of true IDs
0
1000
3000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)

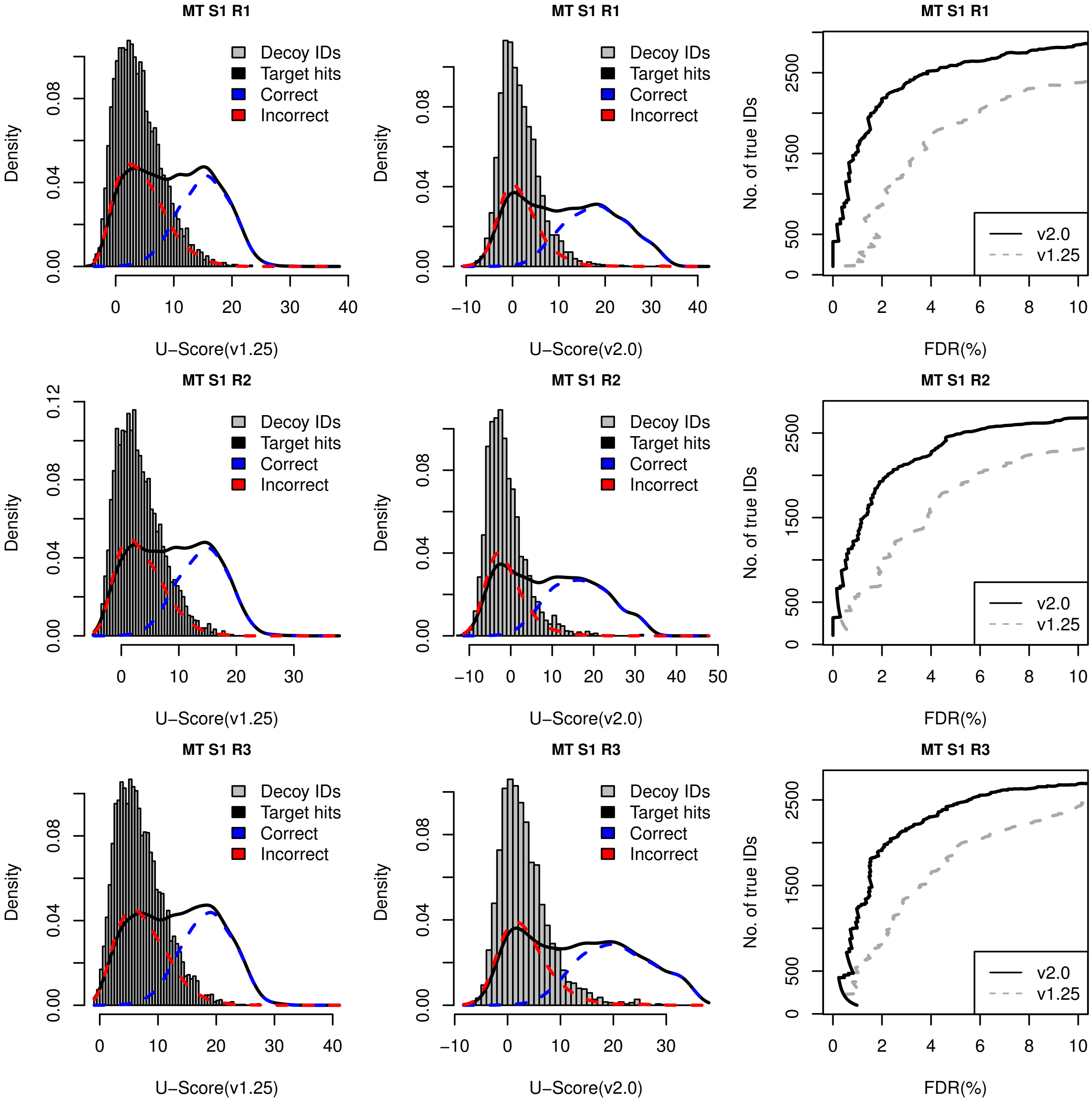

HEK−293 S4 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
U−Score(v1.25)
HEK−293 S4 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
−10
0
10
20
30
40
50
U−Score(v2.0)
HEK−293 S4 R1
No. of true IDs
0
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
HEK−293 S4 R2
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
HEK−293 S4 R2
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v2.0)
HEK−293 S4 R2
No. of true IDs
0
1000
2000
3000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
HEK−293 S4 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0.12
0
10
20
30
40
U−Score(v1.25)
HEK−293 S4 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v2.0)
HEK−293 S4 R3
No. of true IDs
0
1000
3000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)

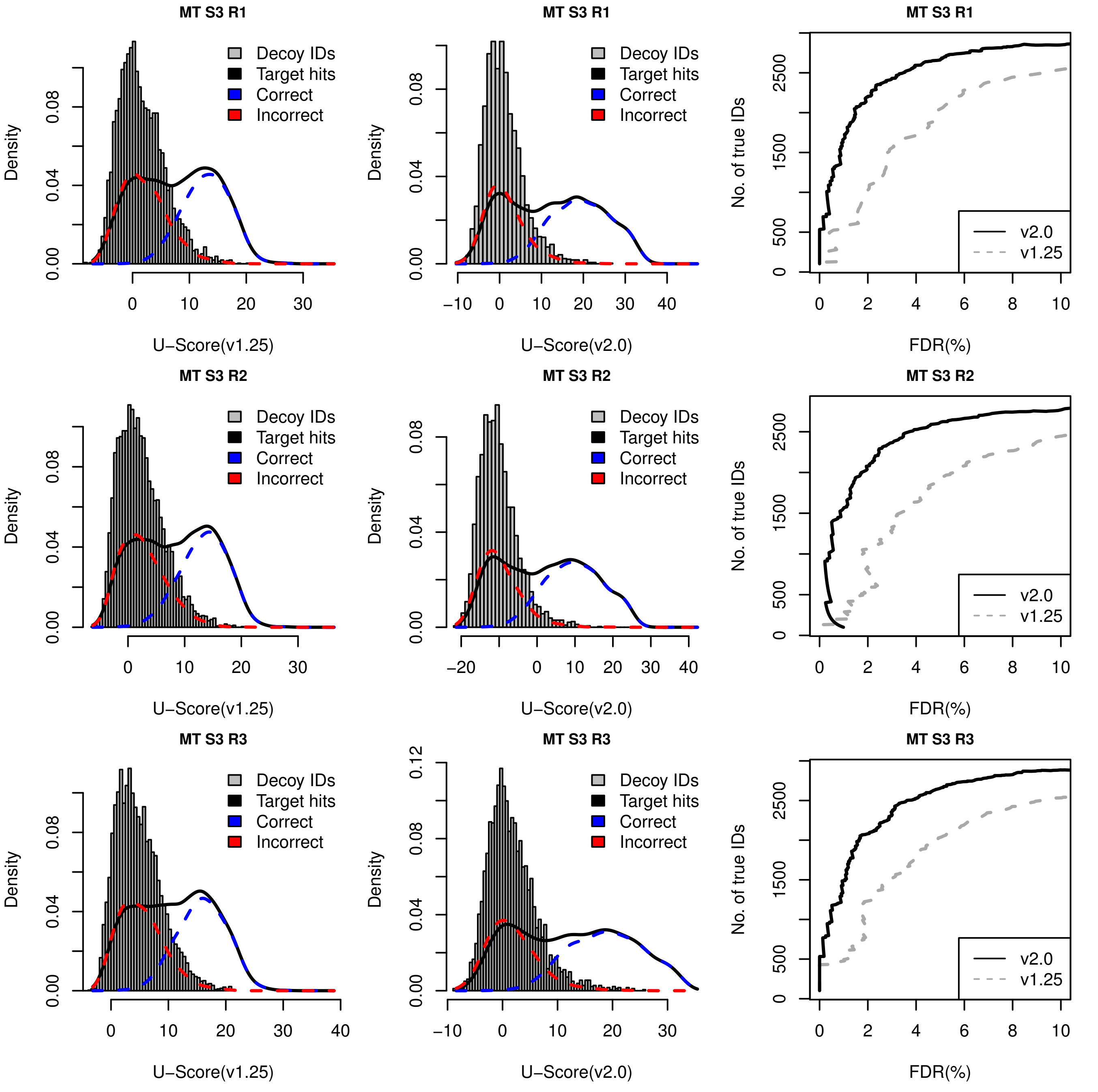

HEK−293 S5 R1
Density
0.00
0.04
0.08
0.12
0
10
20
30
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S5 R1
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S5 R1
No. of true IDs
0
500
1500
2500
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
HEK−293 S5 R2
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S5 R2
Density
0.00
0.04
0.08
−10
0
10
20
30
40
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S5 R2
No. of true IDs
0
1000
2000
3000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
HEK−293 S5 R3
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S5 R3
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S5 R3
No. of true IDs
0
1000
3000
5000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25

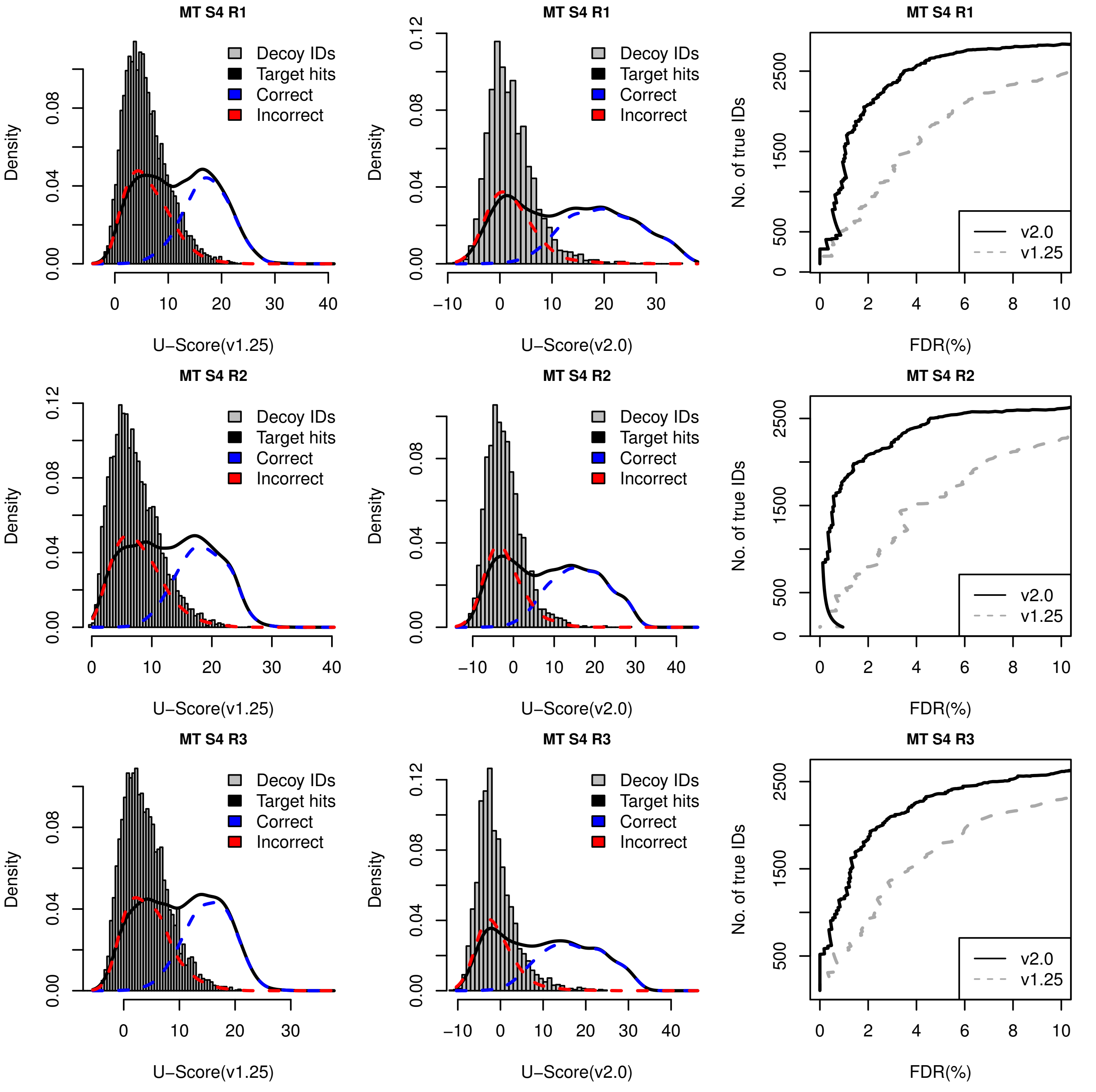

HEK−293 S6 R1
Density
0.00
0.04
0.08
0
10
20
30
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S6 R1
Density
0.00
0.04
0.08
−10
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S6 R1
No. of true IDs
0
1000
2000
3000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
HEK−293 S6 R2
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S6 R2
Density
0.00
0.04
0.08
0
10
20
30
40
50
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S6 R2
No. of true IDs
0
1000
2000
3000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
HEK−293 S6 R3
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S6 R3
Density
0.00
0.04
0.08
−10
0
10
20
30
40
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
HEK−293 S6 R3
No. of true IDs
0
1000
3000
5000
0
2
4
6
8
10
FDR(%)
v2.0
v1.25

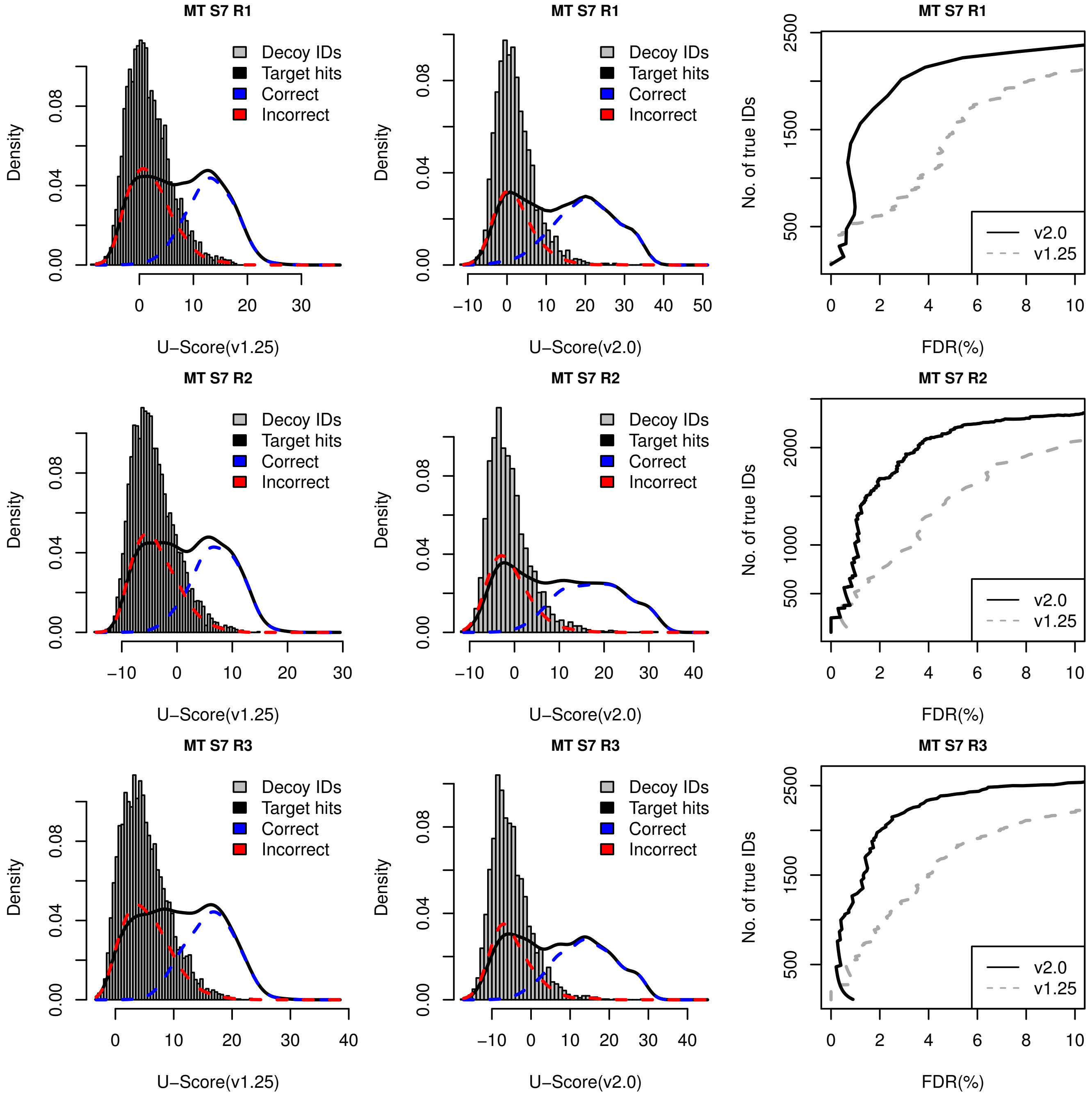

HEK−293 S7 R1
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
U−Score(v1.25)
HEK−293 S7 R1
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
50
60
U−Score(v2.0)
HEK−293 S7 R1
No. of true IDs
0
1000
2000
3000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
HEK−293 S7 R2
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v1.25)
HEK−293 S7 R2
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
50
U−Score(v2.0)
HEK−293 S7 R2
No. of true IDs
0
1000
2000
3000
4000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
HEK−293 S7 R3
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v1.25)
HEK−293 S7 R3
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
−10
0
10
20
30
40
U−Score(v2.0)
HEK−293 S7 R3
No. of true IDs
0
1000
3000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)

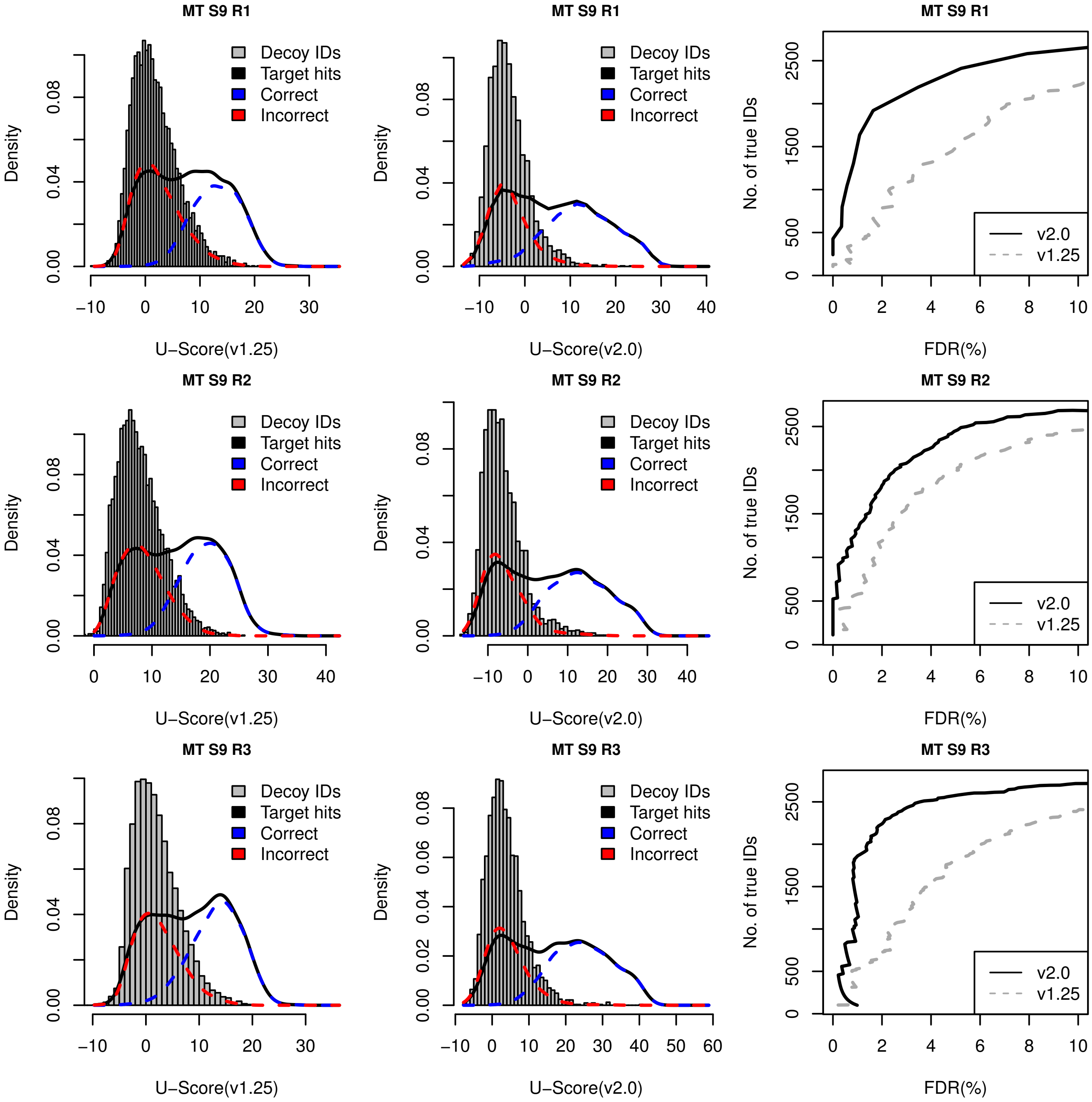

HEK−293 S8 R1
Density
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
0.00
0.04
0.08
0
10
20
30
40
HEK−293 S8 R1
Density
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
0.00
0.04
0.08
0
10
20
30
40
HEK−293 S8 R1
No. of true IDs
FDR(%)
v2.0
v1.25
0
1000
2000
3000
0
2
4
6
8
10
HEK−293 S8 R2
Density
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
0.00
0.04
0.08
0
10
20
30
40
HEK−293 S8 R2
Density
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
0.00
0.04
0.08
0.12
−10
0
10
20
30
40
HEK−293 S8 R2
No. of true IDs
FDR(%)
v2.0
v1.25
0
1000
2000
3000
0
2
4
6
8
10
HEK−293 S8 R3
Density
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
0.00
0.04
0.08
0.12
0
10
20
30
HEK−293 S8 R3
Density
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
0.00
0.04
0.08
−10
0
10
20
30
40
HEK−293 S8 R3
No. of true IDs
FDR(%)
v2.0
v1.25
0
2000
4000
6000
0
2
4
6
8
10

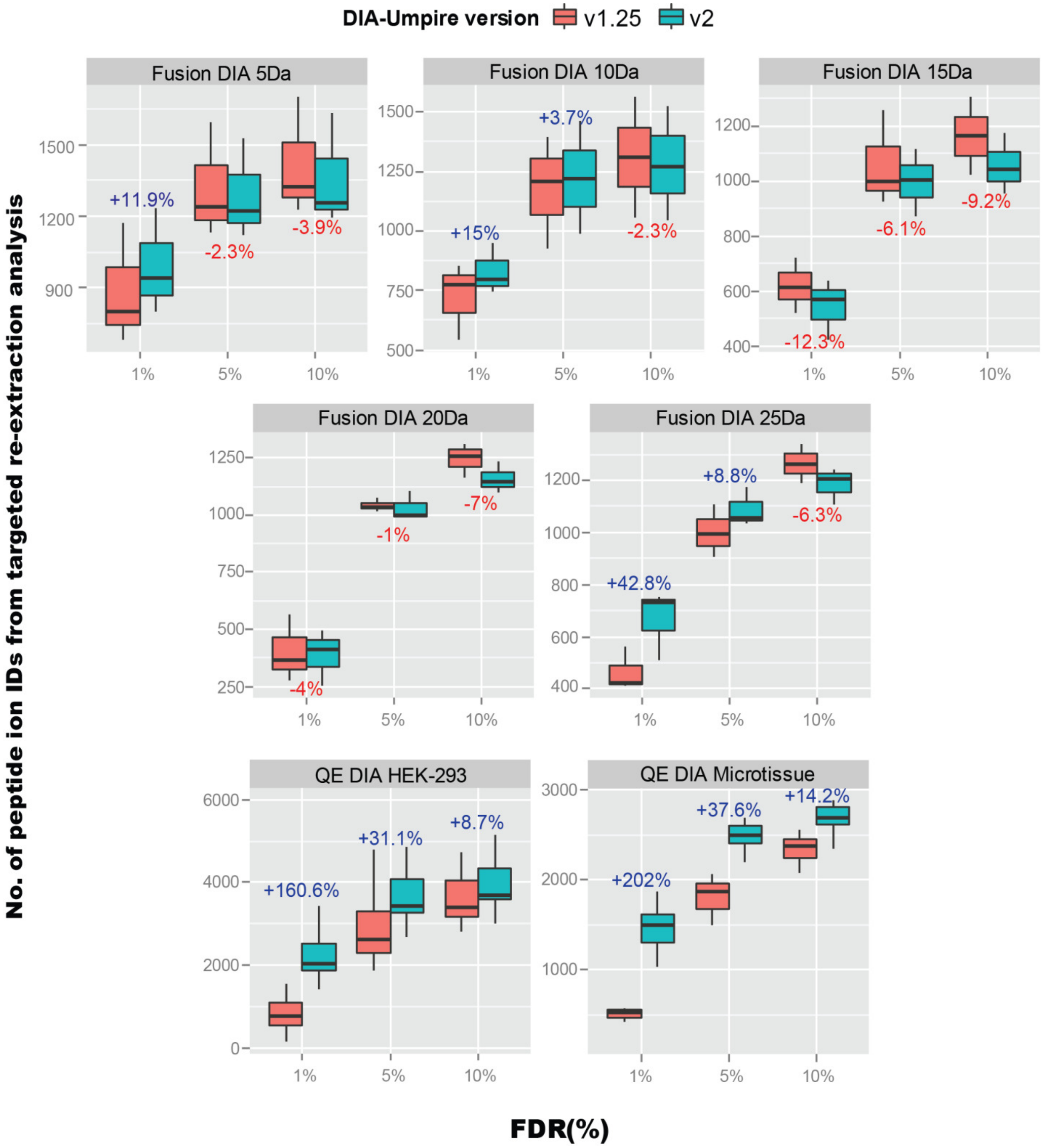

MT S1 R1
MT S1 R1
MT S1 R1
MT S1 R2
MT S1 R2
MT S1 R2
MT S1 R3
MT S1 R3
MT S1 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0.12
U−Score(v1.25)
U−Score(v2.0)
0
10
20
30
40
50
−10
No. of true IDs
0
500
1500
2500
FDR(%)
2
4
6
8
v2.0
v1.25


MT S3 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
U−Score(v1.25)
MT S3 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
−10
0
10
20
30
40
U−Score(v2.0)
MT S3 R1
No. of true IDs
0
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
MT S3 R2
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
U−Score(v1.25)
MT S3 R2
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
−20
0
10
20
30
40
U−Score(v2.0)
MT S3 R2
No. of true IDs
0
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
MT S3 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
MT S3 R3
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0.12
−10
0
10
20
30
U−Score(v2.0)
MT S3 R3
No. of true IDs
0
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)


MT S4 R1
Decoy IDs
Target hits
Correct
Incorrect
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
MT S4 R1
Density
0.00
0.04
0.08
0.12
−10
0
10
20
30
U−Score(v2.0)
MT S4 R1
No. of true IDs
0
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
MT S4 R2
Density
0.00
0.04
0.08
0.12
0
10
20
30
40
U−Score(v1.25)
MT S4 R2
Density
0.00
0.04
0.08
−10
0
10
20
30
40
U−Score(v2.0)
MT S4 R2
No. of true IDs
0
500
1500
2500
v2.0
v1.25
FDR(%)
MT S4 R3
Density
0.00
0.04
0.08
0
10
20
30
U−Score(v1.25)
MT S4 R3
Density
0.00
0.04
0.08
0.12
−10
0
10
20
30
40
U−Score(v2.0)
MT S4 R3
No. of true IDs
500
1500
2500
v2.0
v1.25
FDR(%)


MT S7 R1
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
U−Score(v1.25)
MT S7 R1
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
−10
0
10
20
30
40
50
U−Score(v2.0)
MT S7 R1
No. of true IDs
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
MT S7 R2
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
−10
0
10
20
30
U−Score(v1.25)
MT S7 R2
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
−10
0
10
20
30
40
U−Score(v2.0)
MT S7 R2
No. of true IDs
500
1000
2000
v2.0
v1.25
0
2
4
6
8
10
FDR(%)
MT S7 R3
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
0
10
20
30
40
U−Score(v1.25)
MT S7 R3
Density
0.00
0.04
0.08
Decoy IDs
Target hits
Correct
Incorrect
−10
0
10
20
30
40
U−Score(v2.0)
MT S7 R3
No. of true IDs
500
1500
2500
v2.0
v1.25
0
2
4
6
8
10
FDR(%)


MT S9 R1
Density
0.00
0.04
0.08
−10
0
10
20
30
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
MT S9 R1
Density
0.00
0.04
0.08
−10
0
10
20
30
40
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
MT S9 R1
No. of true IDs
0
500
1500
2500
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
MT S9 R2
Density
0.00
0.04
0.08
0
10
20
30
40
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
MT S9 R2
Density
0.00
0.04
0.08
−10
0
10
20
30
40
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
MT S9 R2
No. of true IDs
0
500
1500
2500
0
2
4
6
8
10
FDR(%)
v2.0
v1.25
MT S9 R3
Density
0.00
0.04
0.08
−10
0
10
20
30
U−Score(v1.25)
Decoy IDs
Target hits
Correct
Incorrect
MT S9 R3
Density
0.00
0.02
0.04
0.06
0.08
−10
0
10
20
30
40
50
60
U−Score(v2.0)
Decoy IDs
Target hits
Correct
Incorrect
MT S9 R3
No. of true IDs
0
500
1500
2500
0
2
4
6
8
10
FDR(%)
v2.0
v1.25



Supplementary Figure 5. Comparison of the numbers of peptide ion identifications from DIA-Umpire v1.25 and v2 targeted re-extraction analysis. The FDRs were estimated by U-score probability calculated by targeted re-extraction step for each DIA dataset. Red and green boxes show the identification numbers from DIA-Umpire v1.25 and v2, respectively. The percentage values shown in the figures are the average improvements from DIA-Umpire v2 compared to the numbers obtained from DIA-Umpire v1.25.