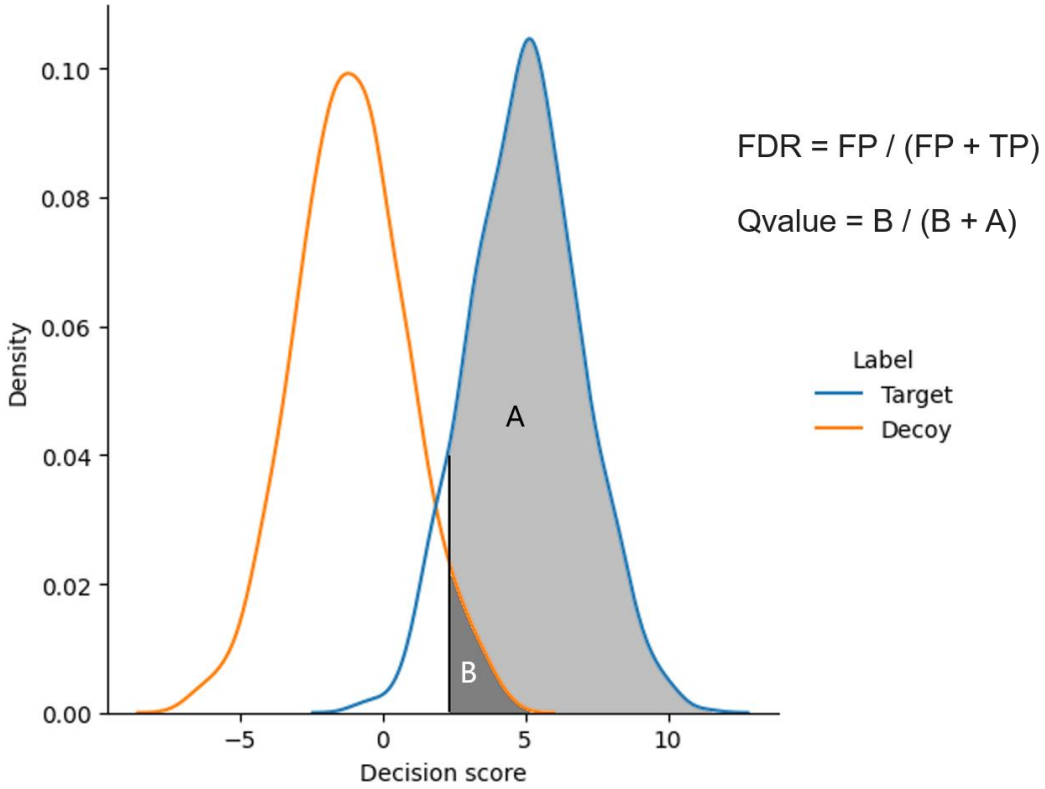
Supplementary Material

Figure 1. FDR and q values 1
Figure 2. The shapley beeswarm plot 3
Figure 3.Quantification result of multi-species proteomics data 6
Figure 4. The decision score distribution of the initial classifier. 7

Table 1. Yeast benchmarking samples preparation 2
Table 2. The number of target and decoy in training dataset 2
Table 3. The precision score of classifiers on test data set 4
Table 4. Average deviation of iRT peptides results. 4
Table 5. The accuracy of the result within ±0.5 log2 fold change tolerance range 4
Table 6.Mouse protein ratio within differentially expressed proteins (%) 5

Figure 1. FDR and q values



FDR is the expectation value of q value. Q value is calculated by decision score generating from a target-decoy binary classifier, which means the possibility of a datapoint with specific decision score being classified as false discovery.

Table 1. Yeast benchmarking samples preparation

	Yeast(μL)	iRT(μL)	Mouse Kidney(μL)	Yeast dil factor
Group 1	1.5625	23.4375	25	8
Group 2	0.390625	24.60938	25	32

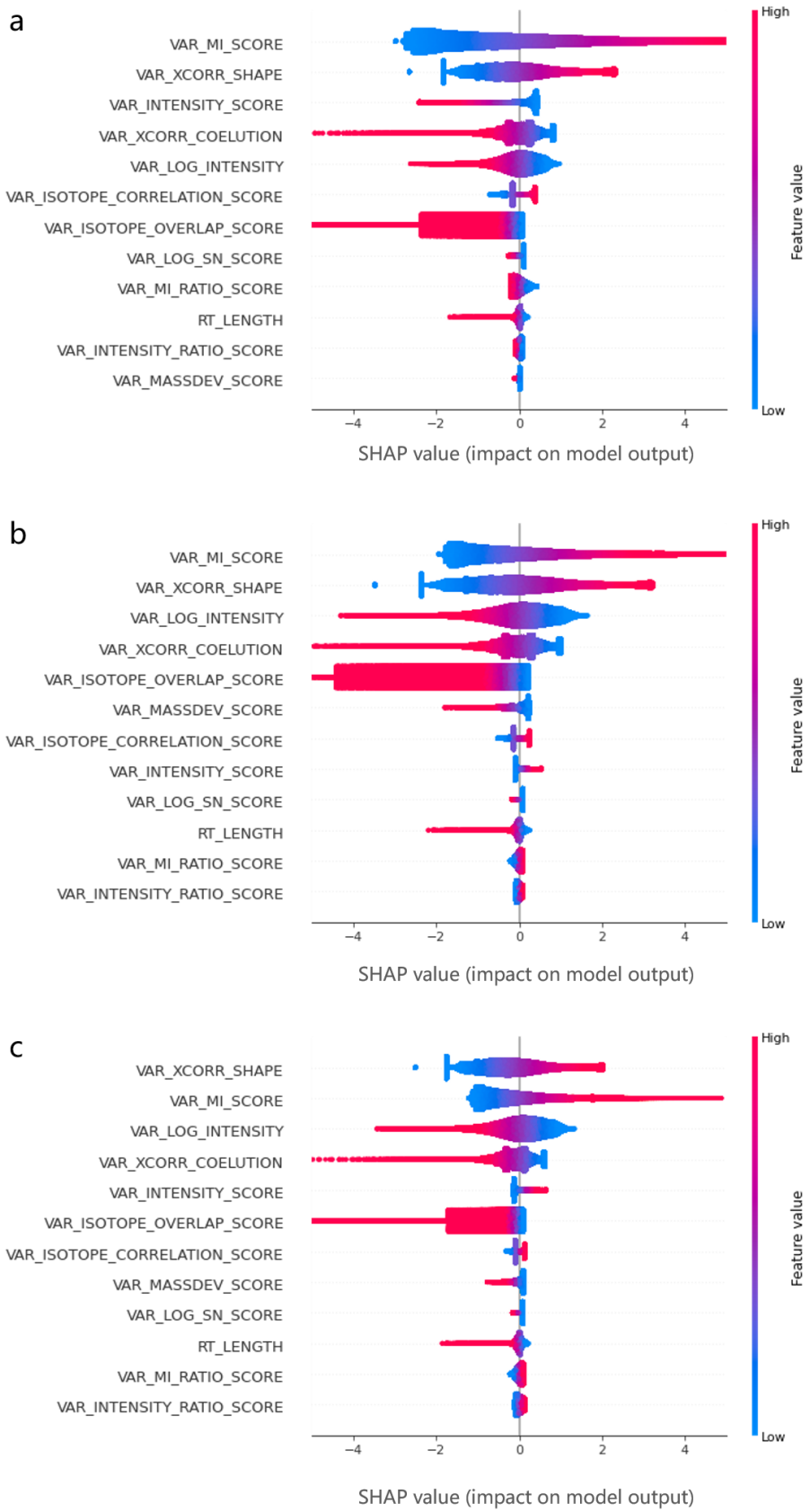
Table notes the concentration of peptides in 2 groups.

Table 2. The number of target and decoy in training dataset

	Indices	Target	Decoy
Total	42531675	27510160	15021515
initial	34025340	22009046	12016294
precision_90	5141123	4627011	514112
qvalue_cutoff	14715566	2699272	12016294
lnl	34025340	22009046	12016294

Table notes the number of target and decoy in every classifier.

Figure 2. The shapley beeswarm plot



The shapley beeswarm plot that displays an information-dense summary of how the top

features in a dataset impact the model's output. Features are sorted according to its importance. Each row(feature) consists of multiple datapoint, red and blue indicates the value of these datapoint on this feature. Positive SHAP value indicates a data be classified as target. Negative a SHAP value indicates a data be classified as target. a) The initial classifier; b) The precision_90 classifier; c) The qvalue_cutoff classifier.

Table 3. The precision score of classifiers on test data set

Classifiers	Precision score
initial	0.662
precision_90	0.977
qvalue_cutoff	0.980
lnl	0.696

Table notes the precision score of every classifier on test dataset.

Table 4. Average deviation of iRT peptides results.

baseline=0.74±1.008	Mean	Median	Top3	Sum
initial	0.54±0.603	0.55±0.304	0.23±0.111	0.49±0.469
precision_90	0.09±0.08	0.45±0.468	0.34±0.215	0.08±0.12
qvalue_cutoff	0.07±0.071	0.35±0.355	0.31±0.443	0.14±0.203
lnl	0.11±0.108	0.42±0.441	0.47±0.539	0.11±0.148

Table notes the average of the deviation between the actual and theoretical log2 fold change of 8 iRT peptides.

Table 5. The accuracy of the result within ±0.5 log2 fold change tolerance range

baseline=75.09	Mean	Median	Top3	Sum
initial	93.95	94.26	96.38	94.22
precision_90	90.83	92.18	96.78	90.43
qvalue_cutoff	92.79	93.73	96.85	91.84
lnl	82.73	87.69	93.88	82.77

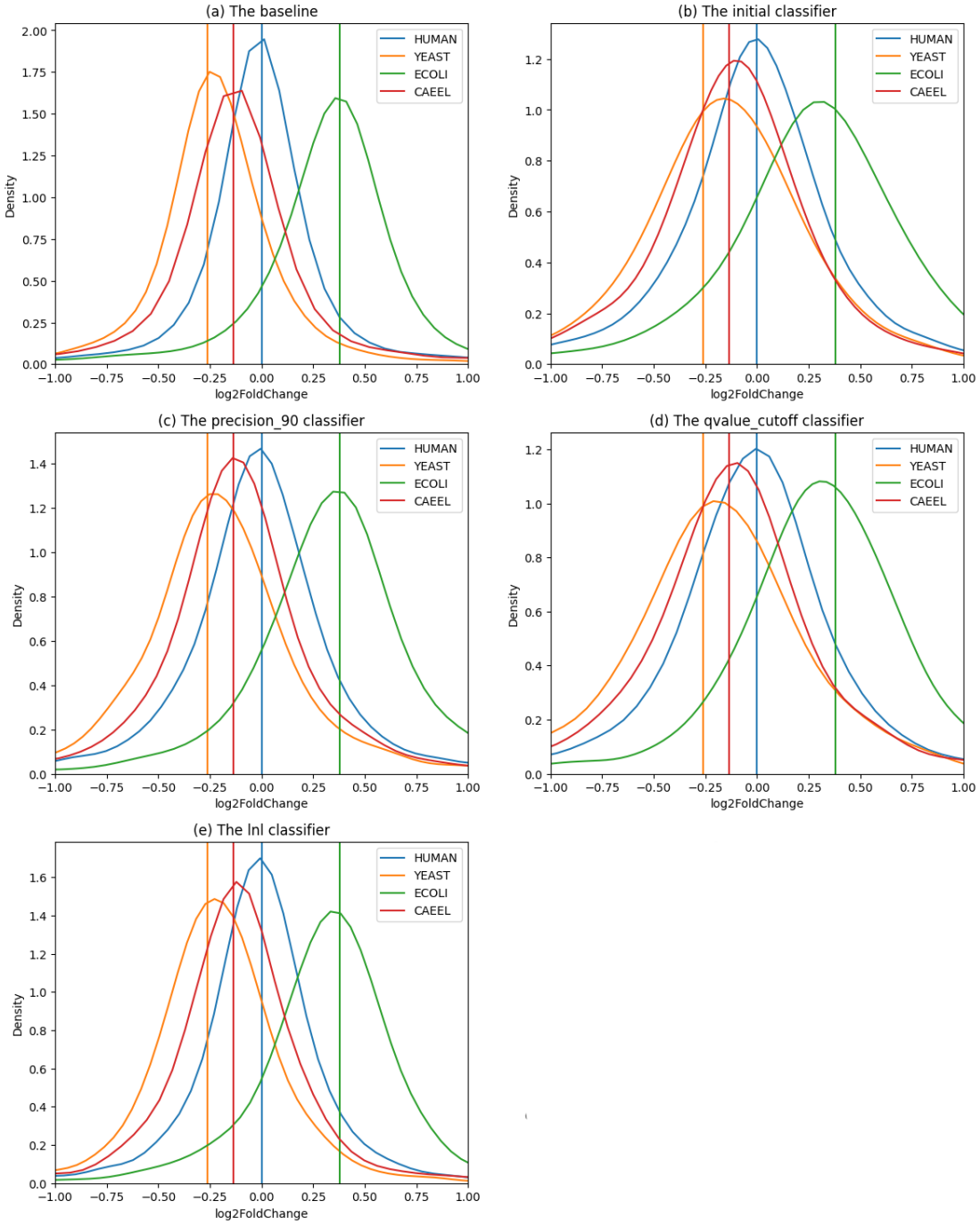
Table notes the accuracy of the result in every method when giving a ±0.5 log2 fold change tolerance range for mouse and yeast proteins.

Table 6. Mouse protein ratio within differentially expressed proteins (%)

baseline=62.56	**Mean**	**Median**	**Top3**	**Sum**
initial	36.66	36.78	23.62	36.17
precision_90	43.61	41.04	19.82	43.82
qvalue_cutoff	39.08	36.57	20.22	41.22
lnl	54.09	45.66	30.76	53.77

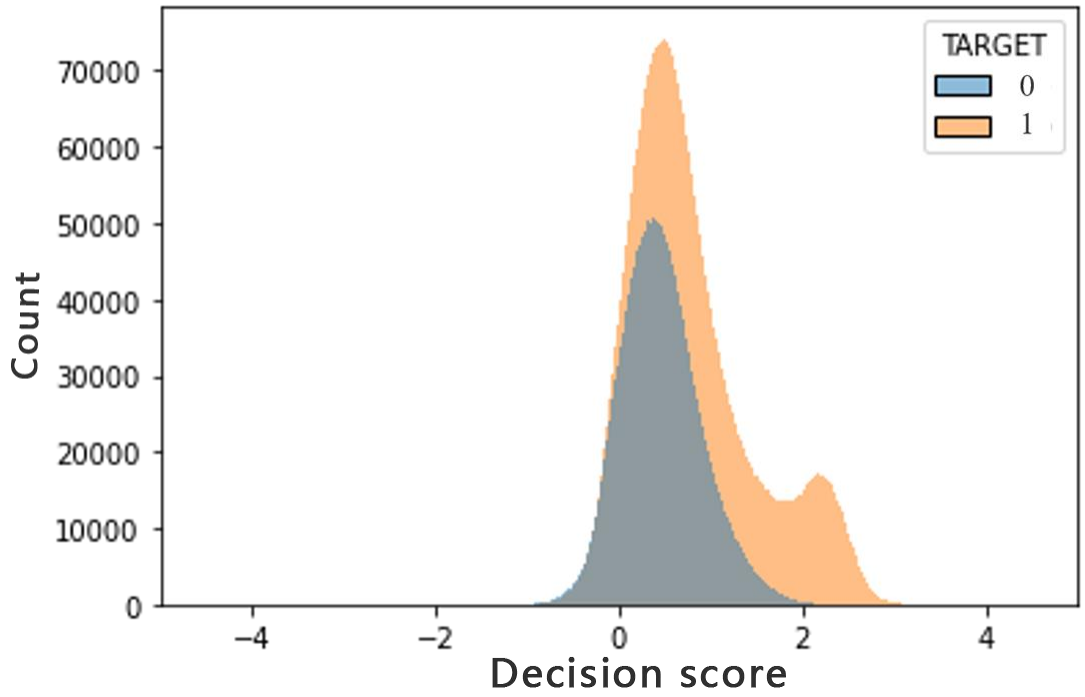
Table notes the ratio of mouse proteins in deferentially expressed proteins within two samples.

Figure 3.Quantification result of multi-species proteomics data



The log2 fold change distribution of multi-species proteomics data. a) The generalized workflow; b) The “initial” classifier + mean quantification method; c) the “precision_90” classifier + mean quantification method; d) the “qvalue_cutoff” classifier + mean quantification method; e) the “lnl” classifier + mean quantification method.

Figure 4. The decision score distribution of the initial classifier.



In labels, 0 is decoys, 1 is targets. There are two peaks in the entire decision score interval on the spike-in dataset, and the first peak almost coincides with the decoy labels, which are considered to be the false targets, i.e. noisy labels.