

PROCEDURE

Here we respectively introduce the functions and parameter configuration of the nine criteria that belong to classic computational criteria and functional criteria.

In Classic Computational Criteria

Click the 'CLASSIC' button.

Precision

Function of the criteria

We assume that the normal tissue near the tumor can approximately represent the normal state of the tumor tissue. Therefore, we use the paired tumor tissue and the normal tissue near the tumor as the gold standard to evaluate the performance of the DEAs. Proteome abundance datasets containing paired matched tumor tissues and adjacent normal tissues were used to evaluate the DEAs (RankComp v1/v2, PenDA, Peng method, Quantile, T-test and Wilcoxon signed-rank test).

The meaning of the parameters

There are seven methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.

How to get sample results

- 1) Go to the 'Precision' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding five files.
- 4) Choose the methods you want.
- 5) Click the 'CONFIRM' button.
- 6) The chart after running will appear on the right.

Example

- 1) Go to the 'Precision' tab.
- 2) Go to the 'parameters' tab. Choose the default parameters.
- 3) UPLOAD DATA: <https://github.com/xmuyulab/IDEPA->

[XMBD/tree/main/data/test_run_data/methodsComp/data.csv](https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/methodsComp/data.csv)

- 4) UPLOAD NORMAL COHORT: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/methodsComp/normal.txt
- 5) UPLOAD TUMOR COHORT: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/methodsComp/tumor.txt
- 6) UPLOAD PAIRED DATA: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/methodsComp/paired_data.csv
- 7) UPLOAD PAIRED SAMPLES: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/methodsComp/paired_samples.txt
- 7) Click the 'CONFIRM' button, the chart after running will appear on the right.



Type One Error

Function of the criteria

Due to the volatility of protein abundance data, it is highly that a DEA produces a significant number of false-positive DEPs. To study magnitude of this problem, we compared the number and characteristics of the false-positive (FP) detections from DEAs with null data constructed from normal tissue samples.

The meaning of the parameters

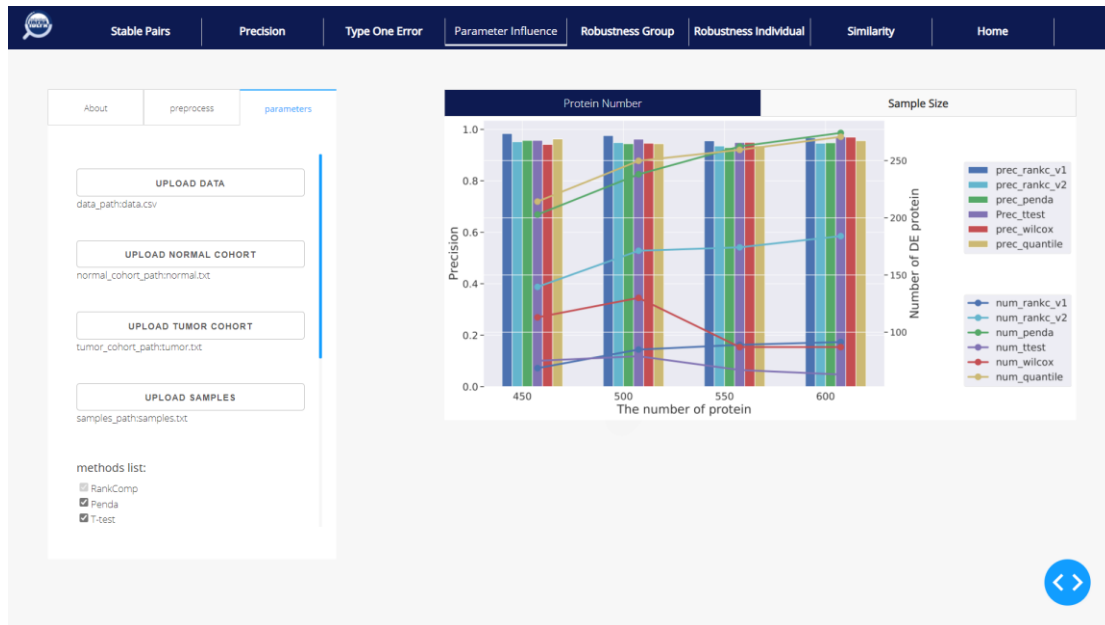
- 1) There are five methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.
- 2) The nd_sample_size means null data with sample size.
- 3) The log_label_list, na_label_list and data_label_list are lists of the corresponding labels. It should be noted that the number of list elements entered in each parameter must be the same as the number of files entered, otherwise it will cause an error.

How to get sample results

- 1) Go to the 'Type One Error' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding three files.
- 4) Choose the methods you want.
- 5) Modify the four parameters according to your needs.
- 6) Click the 'CONFIRM' button.
- 7) The chart after running will appear on the right.

Example

- 1) Go to the 'Type One Error' tab.
- 2) Go to the 'parameters' tab. Choose the default parameters.
- 3) UPLOAD DATA: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/data.csv
- 4) UPLOAD NORMAL COHORT: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/normal.txt
- 5) UPLOAD TUMOR COHORT: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/tumor.txt
- 6) UPLOAD SAMPLES: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/samples.txt
- 7) Click the 'CONFIRM' button, the chart after running will appear on the right.



Stable pairs

Function of the criteria

A pair of proteins are considered as of statistically stable ordering if their REOs are the same in most of the samples. We used REOA package (<https://github.com/pathint/reoa>) to identify protein pairs of stable REOs.

The meaning of the parameters

- 1) If `label_specific_protein` is true, the output consistent pair and reversed pair must contain a specific protein and this specific protein is in the `specific_protein_path` file. If `label_specific_protein` is false, it is not required.
- 2) The `sp` threshold refers to the threshold of the number of stable pairs, reversal pairs, and consistent pairs.
- 3) The `n_visual` refers to how many consistent pairs and reversal pairs will be provided for the user to choose.

How to get sample results

- 1) Go to the 'Stable Pairs' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding four files.
- 4) Modify the four parameters according to your needs.

- 5) Click the 'CONFIRM' button.
- 6) The chart after running will appear on the right.

Example

Parameter Influence

Function of the criteria

In individual-level DEAs, normal tissue samples provide an essential reference for difference analysis of tumor tissue samples. Generally, the stability of the reference group increases as the number of normal tissue samples increases. Also, as the number of proteins increases, the number of differential proteins also increases. We use real protein abundance data to analyze the influence of normal tissue sample size on the performance of the DEAs.

The meaning of the parameters

- 1) There are five methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.
- 2) The sample size label option indicates whether to evaluate the sample size.
- 3) The n protein label indicates whether to evaluate the amount of protein.
- 4) The sample size list is the number gradient of sample size.
- 5) The n protein list means the number of proteins.

How to get sample results

- 1) Go to the 'Parameter Influence' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding four files.
- 4) Choose the methods you want.
- 5) Modify the four parameters according to your needs.
- 6) Click the 'CONFIRM' button.
- 7) The chart after running will appear on the right.

Example

Robustness Individual

Function of the criteria

We evaluated the robustness of the DEAs when it was applied to the same tumor tissue sample but different normal tissue samples as the real difference proteins in tumor, in principle, shouldn't change when different normal tissue sample was used as the reference group.

The meaning of the parameters

- 1) There are four methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.
- 2) The meaning of other parameters is the same as in Robustness Group.

How to get sample results

- 1) Go to the 'Robustness Individual' tab.
- 2) Other steps are the same as in Robustness Group.

Example

Robustness Group

Function of the criteria

To compare the group-level robustness of DEAs under different datasets in the same organization, seven DEAs used to analyze lung dataset and gastric dataset. Then the individualized difference protein results were obtained. Through the binomial distribution test, the difference protein result at the group level can be obtained from the difference protein result at the individual level.

The meaning of the parameters

- 1) There are five methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.
- 2) Robustness is the same algorithm that processes two kinds of data. The first three parameters represent the preprocessing parameters of the two pieces of data.
- 3) The group threshold is the threshold of the binomial distribution test P-value in the

process of obtaining the differential expression results at the individual level to obtain the differential expression results at the group level.

4) The `n_cc` refers to the length of the consistency curve.

How to get sample results

- 1) Go to the 'Robustness Group' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding six files.
- 4) Choose the methods you want.
- 5) Modify the five parameters according to your needs.
- 6) Click the 'CONFIRM' button.
- 7) The chart after running will appear on the right.

Example

Similarity

Function of the criteria

We evaluated the similarity of the results from DEAs. For each sample in each dataset, we calculated the consistency curve of the top N proteins in each pair of algorithms.

The meaning of the parameters

1) There are four methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method. It should be noted that the four methods must be selected, otherwise it will cause errors.

2) The `n_cc` refers to the length of the consistency curve.

How to get sample results

- 1) Go to the 'Similarity' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding three files.
- 4) Modify the parameter according to your needs.
- 5) Click the 'CONFIRM' button.
- 6) The chart after running will appear on the right.

Example

- 1) Go to the 'Similarity' tab.
- 2) Go to the 'parameters' tab. Choose the default parameters.
- 3) UPLOAD DATA: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/data.csv
- 4) UPLOAD NORMAL COHORT: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/normal.txt
- 5) UPLOAD TUMOR COHORT: https://github.com/xmuyulab/IDEPA-XMBD/tree/main/data/test_run_data/tumor.txt
- 6) Click the 'CONFIRM' button, the chart after running will appear on the right.

In Functional Criteria

Click the 'FUNCTIONAL' button.

Pathway Enrichment

Function of the criteria

The DEPs obtained by the DEA at the individual level not only contains the differences between individuals, but also reflects the commonality of the group. We use the individual-level DEAs in IDEPA-XMBD to analyze protein abundance datasets and get DEPs at the individual level.

The meaning of the parameters

- 1) There are three methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.
- 2) The group threshold is the threshold of the binomial distribution test P-value in the process of obtaining the differential expression results at the individual level to obtain the differential expression results at the group level.
- 3) The qvalue thres is the qvalue thres paths from small to large among all the pathways obtained by a certain differential expression algorithm.

How to get sample results

- 1) Go to the 'Pathway Enrichment' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding three files.
- 4) Choose the methods you want.
- 5) Modify the two parameters according to your needs.
- 6) Click the 'CONFIRM' button.
- 7) The chart after running will appear on the right.

Example

Survival Analysis

Function of the criteria

Due to the heterogeneity of cancer, the protein abundance of different individuals on the same type of cancer was very different, and these DEPs between different individuals may be related to the prognosis. Then, we compared the ability of different DEAs to find prognostic proteins.

The meaning of the parameters

- 1) There are three methods in the method list. RankComp represents RankComp v1/v2, and the other options also each represent a method.
- 2) The function of survival threshold is as follows: When the differential expression ratio of a certain protein in all samples is within the range of (survival threshold[0], survival threshold[1]), the protein will be used in subsequent survival analysis.

How to get sample results

- 1) Go to the 'Survival Analysis' tab.
- 2) Go to the 'parameters' tab.
- 3) Follow the prompts to upload the corresponding four files.
- 4) Choose the methods you want.
- 5) Modify the parameter according to your needs.

- 6) Click the 'CONFIRM' button.
- 7) The chart after running will appear on the right.

Example