~ Supporting Information ~

DEWNA: Dynamic Entropy Weight Network Analysis and Its Application to the DNA-Binding Proteome in A549 Cells with Cisplatin-Induced Damage

Table of Contents

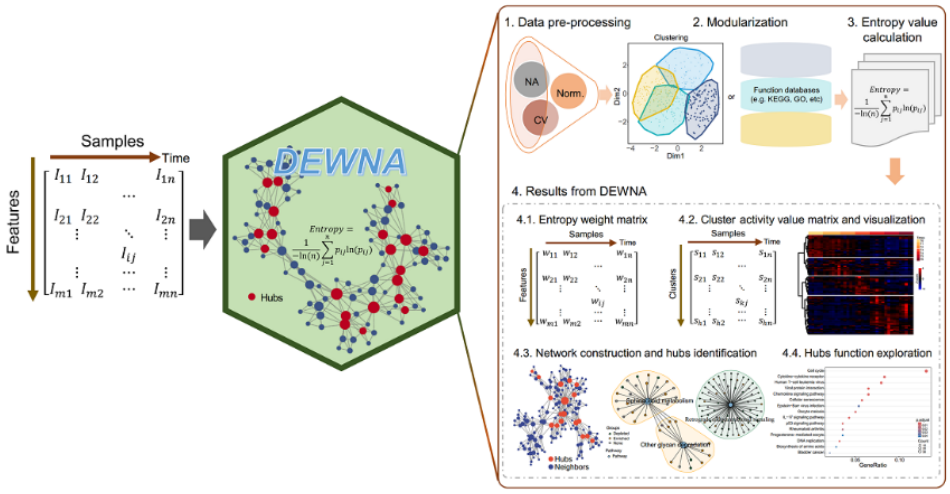
I. Supplementary notes

DEWNA (Dynamic Entropy Weight Network Analysis) is a sophisticated and user-friendly stand-alone software designed to facilitate the dynamic analysis of proteins and their associated functions in response to various treatments. This tool integrates the entropy weight method, providing a comprehensive framework for processing time-course proteome expression data[1]. The entropy weight method assigns weights to indices based on value dispersion, evaluating the perturbation of clusters or pathways by examining protein expression profiles. Higher variability in protein expression indicates a significant response to treatment, leading to greater weight assignment and providing nuanced insights into how treatments affect proteins and pathways over time. Additionally, *DEWNA* integrates co-expression network analysis for each cluster to reveal meaningful multi-scale organizations of co-expressed protein networks and to identify novel therapeutic targets[2]. For pathway analysis, *DEWNA* employs generalized reporter score-based analysis to assess the enriched pathways based on dynamic weighted protein expression matrices[3]. *DEWNA* integrates these methodologies to enhance dynamic proteomic data analysis by allowing users to input time-course proteome expression data, apply entropy weight calculations, and construct co-expression networks, resulting in detailed analysis and visualization of entropy profiles, network diagrams, and temporal activity trends. The software's user-friendly design includes both a web-based platform and a stand-alone version, ensuring broad accessibility and usability.

Users can visit this site: https://www.omicsolution.com/wukong/DEWNA or https://www.omicsolution.org/wukong/DEWNA. **Please Note:** If the online version does not work, which means you cannot open the links, it is probably because our server is down and we will fix it very soon. We also recommend users to install this tool locally (Please check "6. How to run this tool locally?" part below). Then the website homepage can be shown like this:

Welcome to DEWNA

DEWNA (Dynamic Entropy Weight Network Analysis) is a sophisticated and user-friendly stand-alone software designed to facilitate the dynamic analysis of proteins and their associated functions in response to various treatments. This tool implements the entropy weight method, which is a weighting model that assigns weights to indices based on the degree of value dispersion, evaluating the perturbation of clusters or pathways by examining the expression profiles of proteins within these groups. The greater the variability in protein expression upon treatment, the higher the contribution to cluster or pathway perturbation, and therefore, the greater the weight assigned to those proteins. Conversely, proteins with less variability receive lower weights. Additionally, for each cluster, DEWNA integrates the co-expression network analysis to reveal meaningful multi-scale organizations of co-expressed protein network and identifies novel therapeutic targets. On the other hand, for pathway analysis, DEWNA employs the generalized reporter score-based analysis to assess the enriched pathways based on the dynamic weighted protein expression matrices. Therefore, by creatively merging these approaches and extending the functionalities, DEWNA allows for the identification of protein hubs and the elucidation of cluster/pathway entropy weighted profiles during disease progression, providing a comprehensive view of the dynamic changes occurring in response to treatment. The source codes and detailed manual can be accessed at our GitHub. (**Please Note:** If this online version does not work, which means you cannot open the software link, it is probably because our server is down and we will fix it very soon. Or, please try to install this tool and run it locally.)



DEWNA is developed by R shiny (Version 1.3.2), and is free and open to all users with no login requirement. It can be readily accessed by all popular web browsers including Google Chrome, Mozilla Firefox, Safari and Internet Explorer 10 (or later), and so on. We would highly appreciate that if you could send your feedback about any bug or feature request to Shisheng Wang at wsslearning@omicsolution.com.

~~ Enjoy yourself in DEWNA ~~

1. Data Preparation

The uploaded data file formats could be .csv or .txt. Before analysis, users should prepare the proteomics expression data and sample information. The proteomics expression data required here could be readily generated based on results of several popular tools such as Spectronaut[4], MaxQuant[5]. Users then can upload the data and type in right sample information into StatsPro with right formats respectively and start subsequent analysis.

1.1 Expression data

Users can prepare their proteomics expression data as below from any software (e.g. Spectronaut, MaxQuant, etc.) and upload into this tool. In the situation, protein ids/names are sequentially provided in the first two columns of input file. The protein ids/names in the first column could be UniProt ids or protein names. From the second column, proteins expression intensity or signal abundance in every sample should be listed. The data structure is shown as below (rows are proteins and columns are samples):

PG.ProteinGroups	DPC_0h_1	DPC_0h_2	DPC_0h_3	DPC_0h_4	DPC_3h_1	DPC_3h_2	DPC_3h_3	DPC_3h_4	DPC_6h_1	DPC_6h_2
A0A024RB	25189.73	13440.21	19565.14	17490.47	27999.1	11120.79	8287.486	10820.16	21856.66	11681.11
A0A096LP	43690.06	40345.23	37148.71	37204.88	39428.96	39318.33	36062.08	34250.29	34548.79	30284.49
A0A0B4J2	1060724	836391.3	1103099	1063019	973396.6	821988.6	936122.9	1124465	917335	867221.2
A0A0B4J2I	57807.32	60119.08	63978.09	69136.92	68128.56	31073.24	72903.39	71065.42	45719.14	48018.71
A0AV96	14720.61	16155.04	15082.32	14969.27	14932.68	13164.33	14025.51	12534.12	14209.58	13106.73
A0AVK6	16522.56	18600.74	17865.75	18601.68	18643.31	5125.668	14807.14	16103.46	20163.18	17906.59
A0AVT1	17222.84	24763.66	30250.37	28884.49	16928.29	28179.63	22732.48	16664.99	25190.92	23202.59
A0FGR8	20893.18	17165.7	18176.4	20293.75	18935.83	12087.23	14606.6	11881.54	25152.06	17099.2
A0JLT2	37593.66	28032.06	28467.56	21959.01	34682.7	23316.45	35804.23	32016.78	44664.74	40806.98
A0MZ66	29056.75	32836.46	30119.31	30416.28	35155.96	32258.38	33450.51	30839	31779.2	37992.16
A1L0T0	35199.13	35886.66	36962.29	38643.55	29708.71	31485.84	25030.82	31670.05	36484.25	31963.77
A1L390	12021.93	12513.96	11311.04	10823.97	11329.15	9397.313	12337.43	9379.057	12087.48	10786.13
A2RRD8	46519.11	58789.84	54326.02	62714.16	57500.93	64203.45	65464.58	70372.02	79865.82	86481.73
A3KN83	26430.21	25346.23	26637.75	25969.68	26256.35	25992.66	25837	25177.89	24772.02	24262.8
A5YKK6	27706.62	27926.76	26817.8	27684.27	24960.62	26118.65	25467.98	24301.89	31291.05	30275.13
A6NCC3	35944.17	20451.28	27322.03	25434.8	41856.32	7629.363	2581.511	10482.42	26577.31	25864.84
A6NDG6	36172.88	32129.89	29492.28	36041.41	32832.6	28935.81	26838.32	24399.04	38261.67	38333.74
A6NDU8	48340.96	50244.54	40859.61	44103.27	37341.36	42296.37	38651.26	37392.52	47460.55	39691.67
A6NED2	2379.226	9749.278	11220.31	10147.28	8857.214	6130.552	6130.417	3188.316	11724.42	11549.22
A6NFI3	26165.06	32865.58	25063.64	30182.51	29647.58	31624.83	32113.08	31675.63	26228.95	32020.46
A6NHG4	10506.64	6973.583	7115.409	8546.165	3228.381	4923.56	2169.056	6023.871	7572.204	4797.863
A6NHL2	75163.5	66692.17	69335.76	77987.73	69796.09	56109.2	56733.93	61412.86	74370.55	63234.49
A6NHR9	46648.08	49067.9	47149.73	47286.44	49437.65	52611.63	50571.77	49839.58	41032.63	43034.47

Samples

Protein IDs

Intensities

1.2 Samples information data

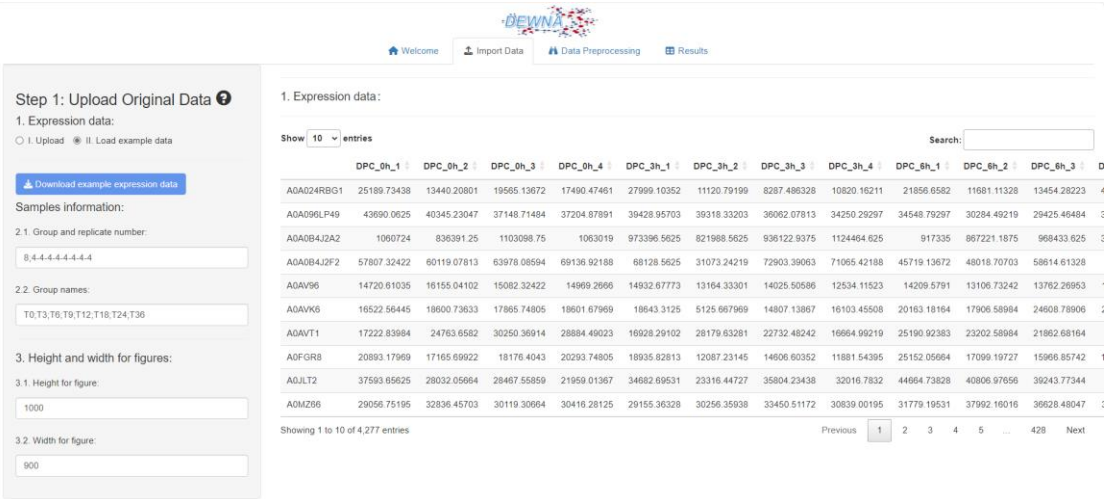
Sample information here means that users should provide group number, replicate number, and group names, then type them in *DEWNA* (see below). This tool will use these information to calculate corresponding results and enable, for example, filtration strategy for different group respectively in a later step:

2.1 Group and replicate number: Type in the group/timepoint number and replicate number here. Please note, the group/timepoint number and replicate number are linked with ";", and the replicate number of each group is linked with "-". For example, if you have two groups/timepoints, each group has three replicates, then you should type in "2;3-3" here. Similarly, if you have 3 groups/timepoints with 5 replicates in every groups, you should type in "3;5-5-5". The example data has 8 timepoints and 4 replicates in each group, so here is “8;4-4-4-4-4-4-4-4-4”.

2.2 Group names: Type in the group names of your samples. Please note, the group names are linked with ";". For example, there are 8 timepoints in the example data, you can type in "T0;T3;T6;T9;T12;T18;T24;T36 ".

1.3 Download example datasets

If users want to download the example datasets to their own computer and check the data format locally, they can download them from here:

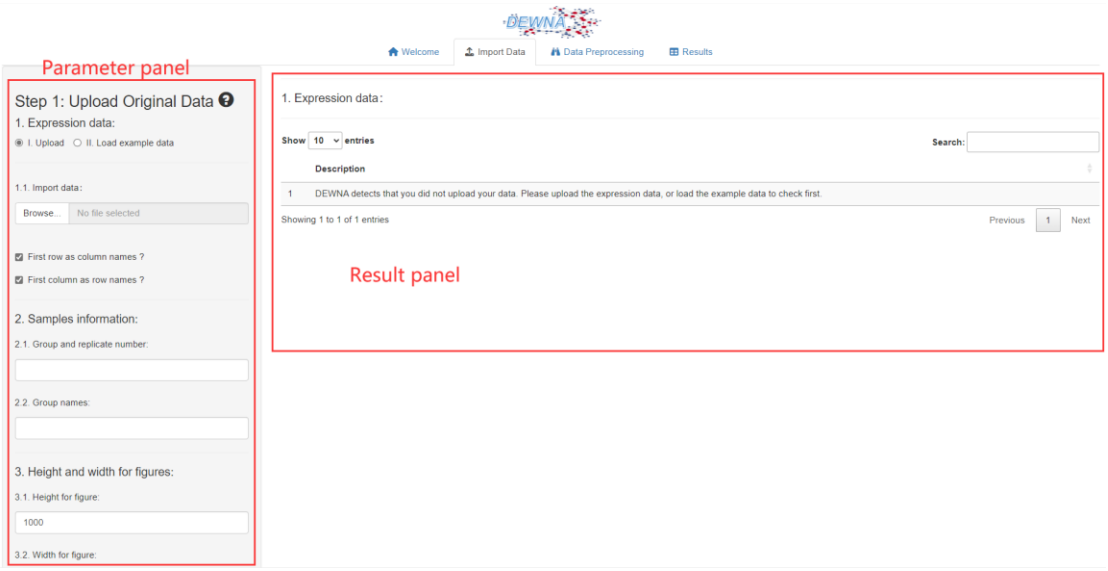


users can download the example data (proteomics expression data) by clicking the “Download example expression data” button. The data are saved as corresponding format and users can open them in other software, such as Excel. Then, users can check the example sample information and understand these parameters better.

2. Import Data

This is the first step, in which users should upload data here or load the example data with the above data formats. By default, we use the example data to show result of every step.

2.1 Uploading data. When users prepare their data (expression and sample information data set), they can upload these data from here:



There are two main panels: first, *parameter panel*, users can adjust parameters here; second, *result panel*, many results after users set the parameters will be shown here and users can also download these results.

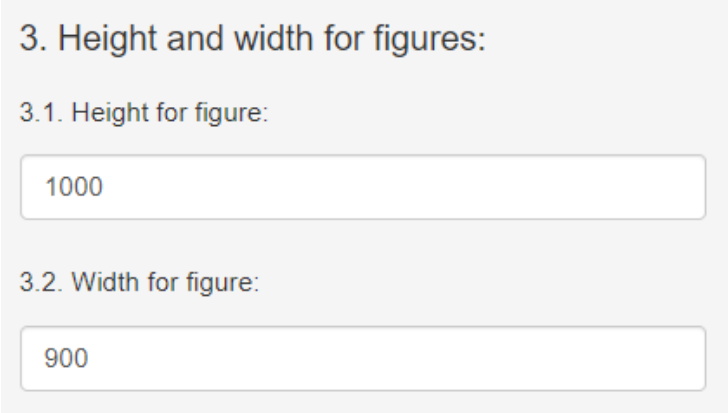
In the *parameter panel* of “Import Data”, there are three parts for users:

a. Upload/Load example data. When users choose this option, they can upload their own data here. Users should select the right format (.csv or .txt) based on their data and then click “Browse” button to import the data; Otherwise, users can select the “Load example data” to check the example data and run this tool using this data.

First row as column names: this means whether the first row is column names. If true, you should choose this parameter.

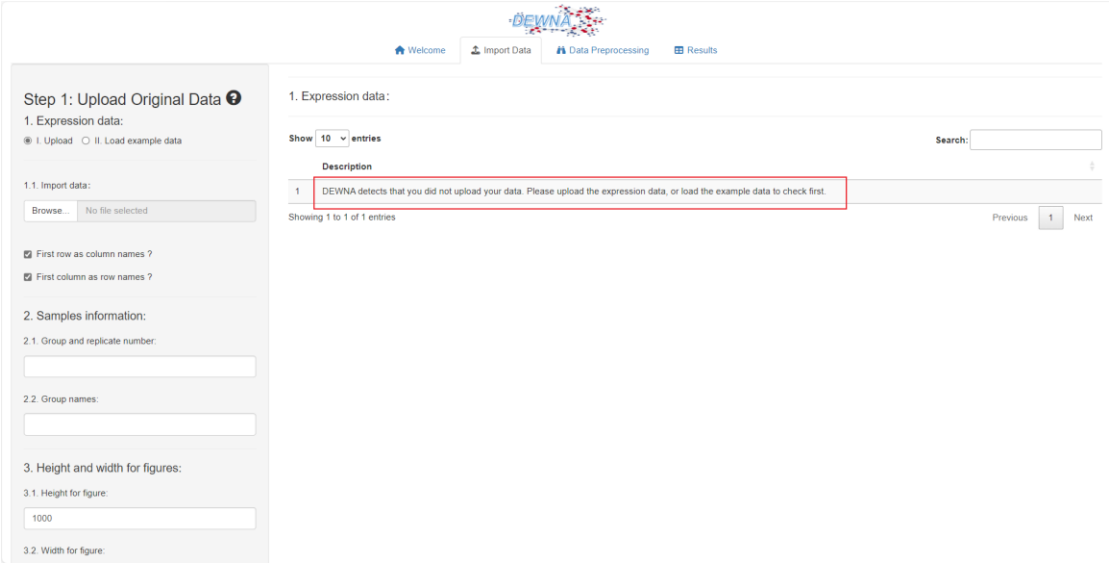
First column as row names: this means whether the first column is row names. If true, you should choose this parameter.

b. Samples information. As described in part 1.2, users can choose this option and download the example data to check them locally.

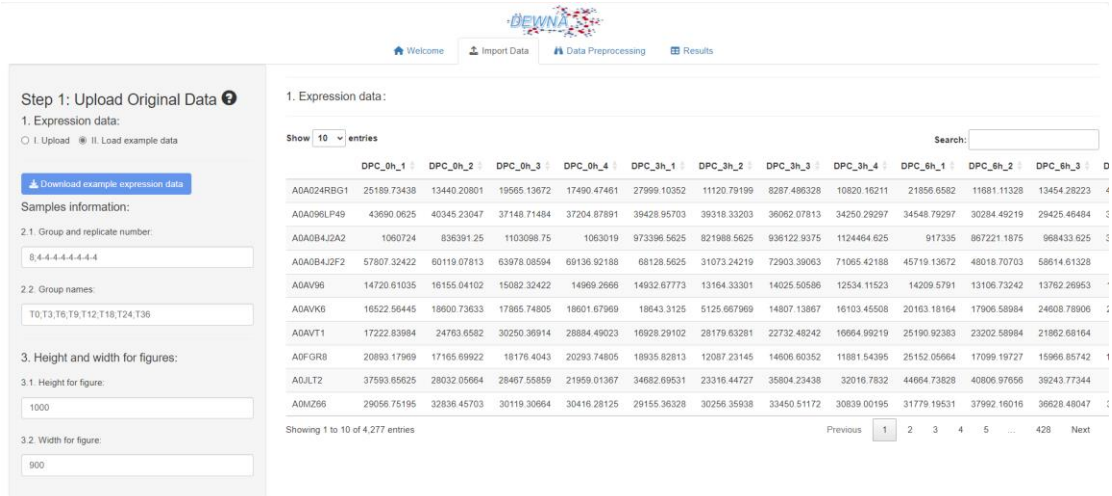


c. Height and width for figures. Users can set up these parameters to control the height and width of figures generated in this tool.

If users do not input anything, this tool will print “DEWNA detects that you did not upload your data. Please upload the expression data, or load the example data to check first.” to warn users:

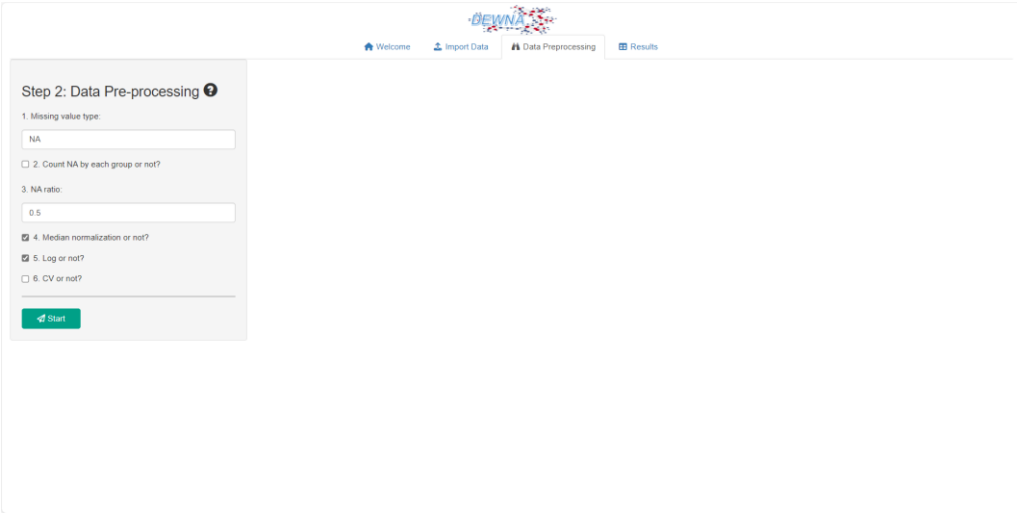


Now, we select the “Load example data” to run this tool as an example:

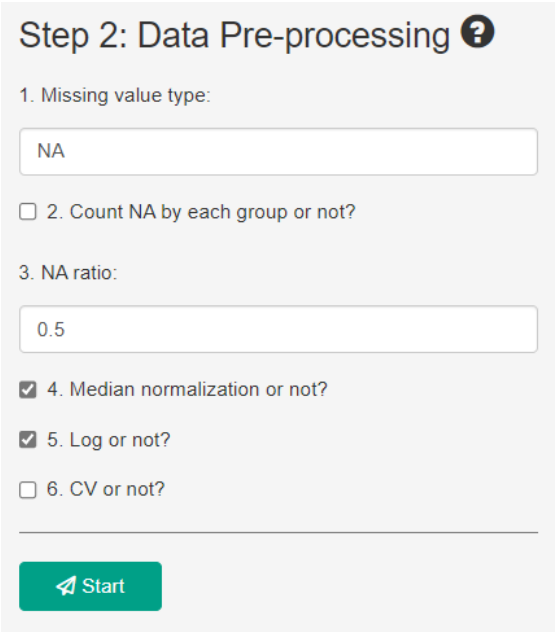


3. Data Pre-processing

Users can pre-process their data in this step, including data filtration (i.e. removing those proteins with high proportion of missing values (NAs) and large coefficient of variation (CVs)), normalization (i.e. normalizing protein intensities using median value of each sample), missing value imputation (i.e. all missing values are derived with the k-Nearest Neighbor method[6]).



3.1 Parameters



1. Missing value type: what the missing values look like in the expression data, for example, Spectronaut [4,7] software usually export “Filtered” as missing values, so users should change this parameter to “Filtered” if their data contain “Filtered”. *DEWNA* will recognize these characters and replace them with NAs. Any other characters indicating a missing value can be similarly defined.

2. Count NA by each group or not: if true, *DEWNA* will count the number of missing values in each group and calculate the NA ratio. Otherwise, it calculates the NA ratio across all groups.

3. NA ratio: the threshold of NA ratio. Those peptides/proteins with NA ratio above this threshold

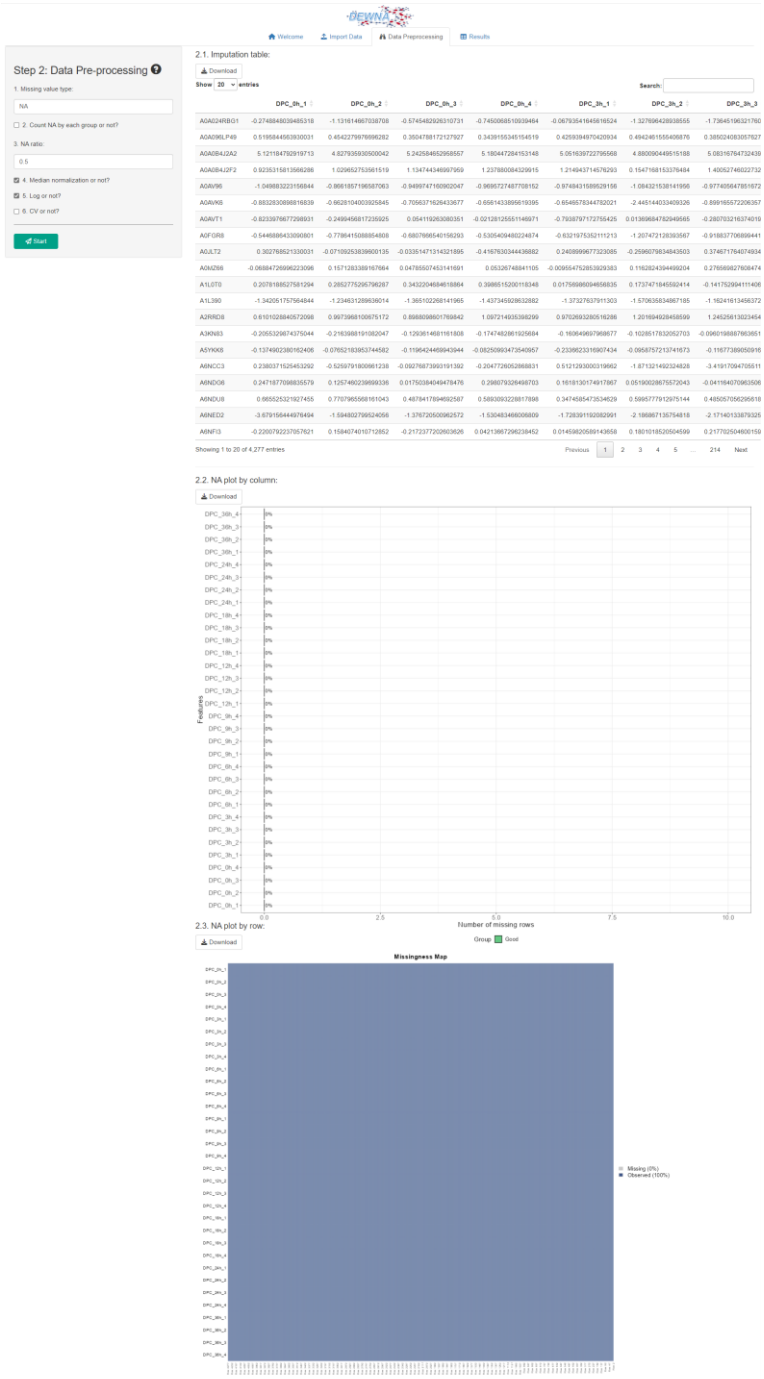
will be removed.

4. Median normalization or not: if true, *DEWNA* will process median normalization for original data. (Note, *DEWNA* was not designed to perform sophisticated normalization analysis. Any normalized datasets with NA can be accepted for analysis).

5. Log or not: if true, the data will be transformed to the logarithmic scale with base 2.

6. CV or not: the threshold of coefficient of variation. If true, those peptides/proteins with CV above this threshold will be removed. “raw scale” here means the CV of each peptide/protein is calculate using the data before logarithm transformation.

If users set these parameters well, then click “calculate” button, the results will appear on the right panel.



3.2 Results of Data Pre-processing

2.1. Imputation table. This part will derive the missing values with the k-Nearest Neighbor method. Users can check how to process missing value problem detailedly in our previous published article [8]. The results are shown as below and users can click “Download” button to save this result in a .csv file:

2.1. Imputation table:

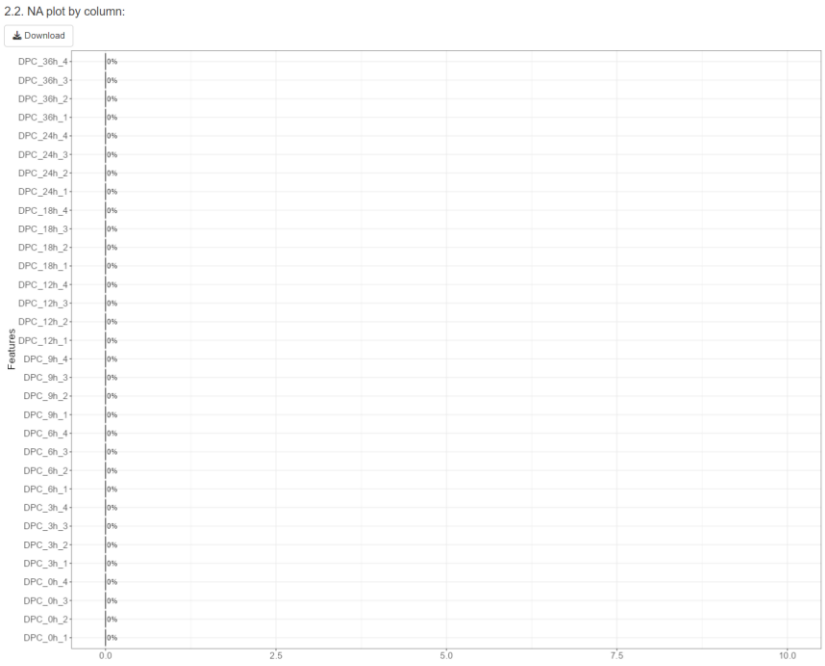
Download

Show 20 entries Search:

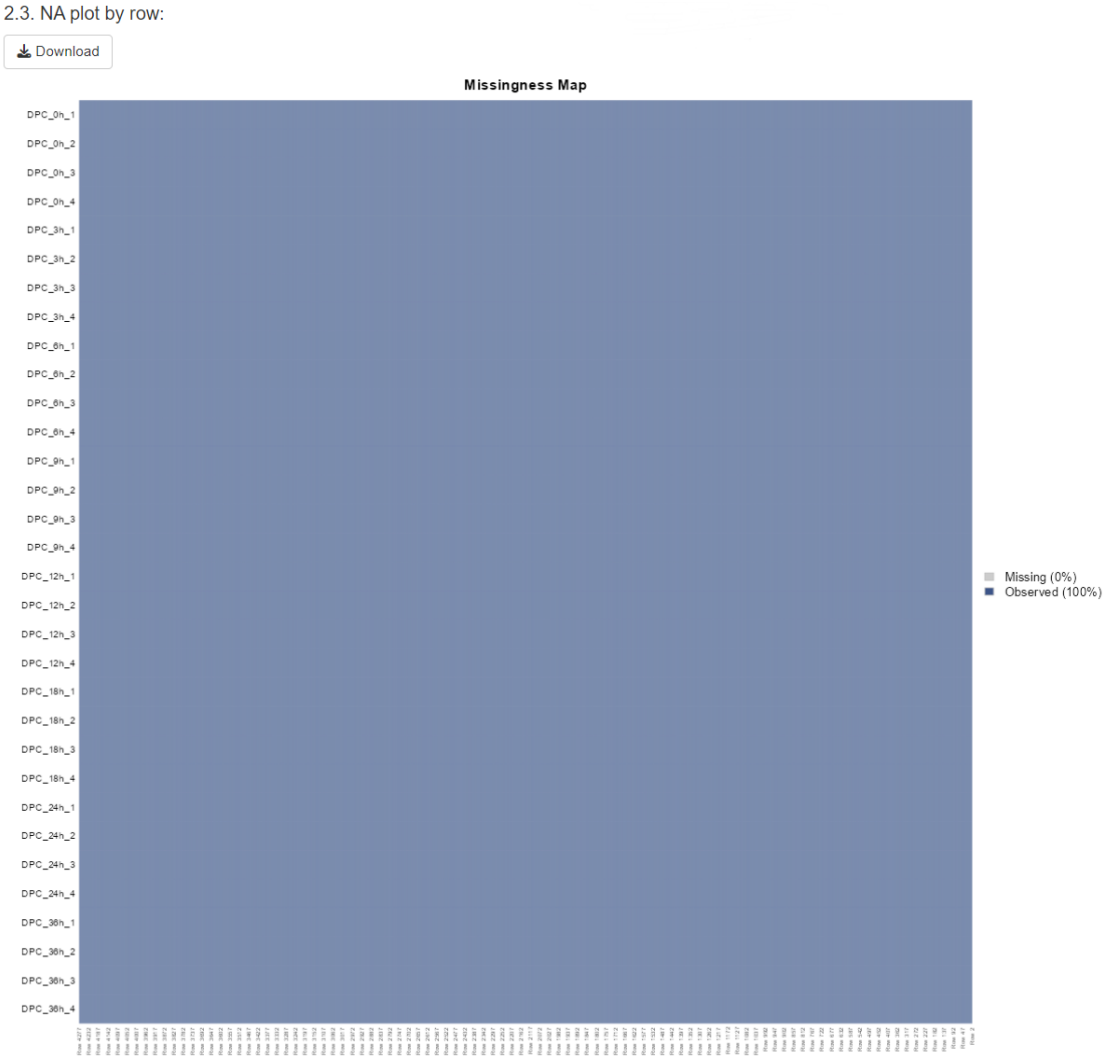
	DPC_0h_1	DPC_0h_2	DPC_0h_3	DPC_0h_4	DPC_3h_1	DPC_3h_2	DPC_3h_3
A0A024RBG1	-0.2748848039485318	-1.131614667038708	-0.5745482926310731	-0.7450068510939464	-0.06793541645616524	-1.327696428938555	-1.736451963217600
A0A096LP49	0.5195844563930031	0.4542279976696282	0.3504788172127927	0.3439155345154519	0.4259394970420934	0.4942461555406876	0.385024083057627
A0A0B4J2A2	5.121184792919713	4.827935930500042	5.242584652958557	5.180447284153148	5.051639722795568	4.880090449515188	5.083167647324390
A0A0B4J2F2	0.9235315813566286	1.029652753561519	1.134744346997959	1.237880084329915	1.214943714576293	0.1547168153376484	1.400527460227320
A0AV96	-1.049883223156844	-0.8661857196587063	-0.9499747160902047	-0.9695727487708152	-0.9748431589529156	-1.084321538141956	-0.977405647851672
A0AVK6	-0.8832830898816839	-0.6628104003925845	-0.7056371626433677	-0.6561433895619395	-0.6546578344782021	-2.445144033409326	-0.899165572206357
A0AVT1	-0.8233976677298931	-0.2499456817235925	0.054119263080351	-0.02128125551146971	-0.7938797172755425	0.01369684782949565	-0.280703216374019
A0FGR8	-0.5446886433090801	-0.7786415088854808	-0.6807666540156293	-0.5305409480224874	-0.6321975352111213	-1.207472128393567	-0.918837706899441
A0JLT2	0.302768521330031	-0.07109253839600135	-0.03351471314321895	-0.4167630344436882	0.2408999677323085	-0.2596079834843503	0.374671764074934
A0MZ66	-0.06884726996223096	0.1571283389167664	0.04785507453141691	0.05326748841105	-0.009554752853929383	0.1162824394499204	0.276569827608474
A1L0T0	0.2078188527581294	0.2852775295796287	0.3432204684618864	0.3986515200118348	0.01756986094656835	0.1737471845592414	-0.141752994111406
A1L390	-1.342051757564844	-1.234631289636014	-1.365102268141965	-1.437345928632882	-1.37327637911303	-1.570635834867185	-1.162416134563720
A2RRD8	0.6101028840572098	0.9973968100675172	0.8988098601769842	1.097214935398299	0.9702693280516286	1.201694928458599	1.245256130234540
A3KN83	-0.2055329874375044	-0.2163988191082047	-0.1293614681161808	-0.1747482861925684	-0.160649697968677	-0.1028517832052703	-0.096019888766365100
A5YKK6	-0.1374902380162406	-0.07652183953744582	-0.1196424469943944	-0.08250993473540957	-0.2336623316907434	-0.0958757213741673	-0.116773890509160
A6NCC3	0.2380371525453292	-0.5259791800661238	-0.09276873993191392	-0.2047726052868831	0.5121293000319662	-1.871321492324828	-3.419170947055110
A6NDG6	0.2471877098835579	0.1257460239699336	0.01750384049478476	0.298079326498703	0.1618130174917867	0.05190028675572043	-0.041164070963506
A6NDU8	0.665525321927455	0.7707965568161043	0.4878417894692587	0.5893093228817898	0.3474585473534629	0.5995777912975144	0.485057056295618
A6NED2	-3.679156444976494	-1.594802799524056	-1.376720500962572	-1.530483466006809	-1.728391192082991	-2.186867135754818	-2.171401338793250
A6NFI3	-0.2200792237057621	0.1584074010712852	-0.2172377202603626	0.04213667296238452	0.01459820589143658	0.1801018520504599	0.217702504600159

Showing 1 to 20 of 4,277 entries Previous 1 2 3 4 5 … 214 Next

2.2. NA plot by column. Here shows the result of the NA distribution of every sample. 0% means there is no NA value in the data.

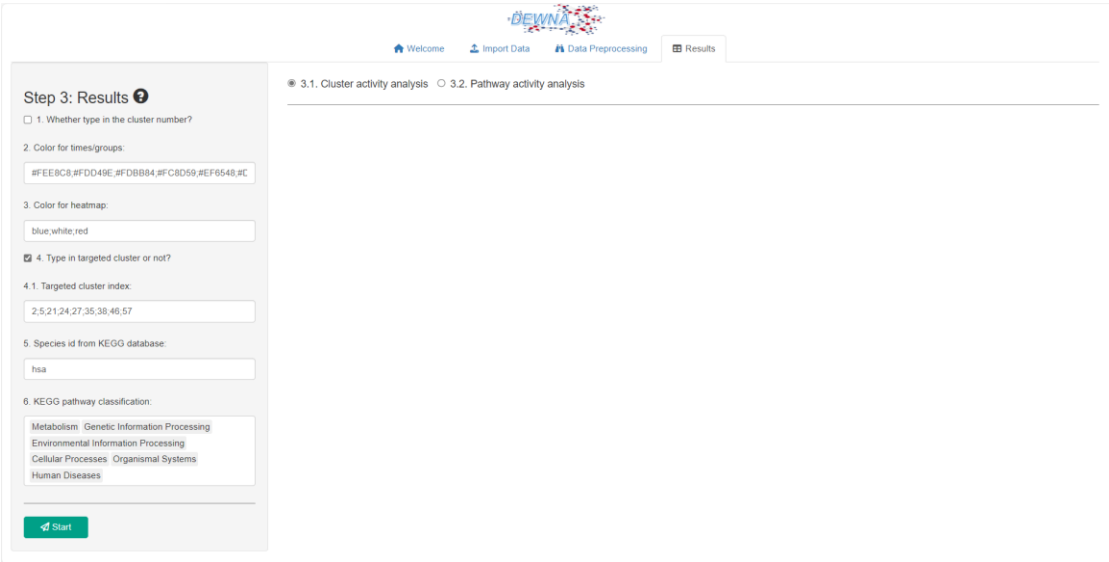


2.3. NA plot by row. Here shows the result of the NA distribution of every protein.

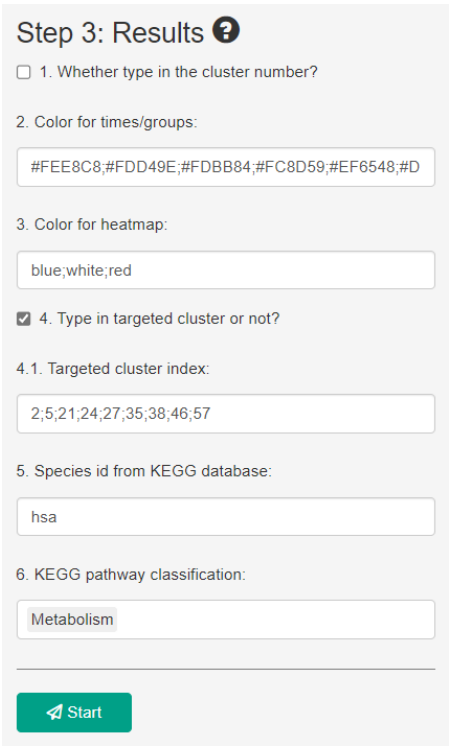


4. Resluts

This step shows all the results generated in *DEWNA*. The pre-processed data were modularized based on one clustering method or one function database (e.g. every pathway in KEGG [9]). Therefore, there are two kinds of results here (shown as below): 1. The results from “3.1. Cluster activity analysis” are those obtained based on the clustering data. 2. The results from “3.2. Pathway activity analysis” are those obtained based on the KEGG database.



4.1 Parameters



1. Whether type in the cluster number? If true, users could type in the number of clusters they want. By default, the number of clusters is the ceiling of the square root of the protein number, for example, if there are 8000 proteins, the cluster number is 90.

2. Color for times/groups: Colors for each timepoint or group. The number of colors should be

equal to that of timepoints or groups.

3. Color for heatmap: Colors for those heatmap plots. By default, there are three color names (blue;white;red) here for the lowest (blue), middle (white), highest (red) values repectively.

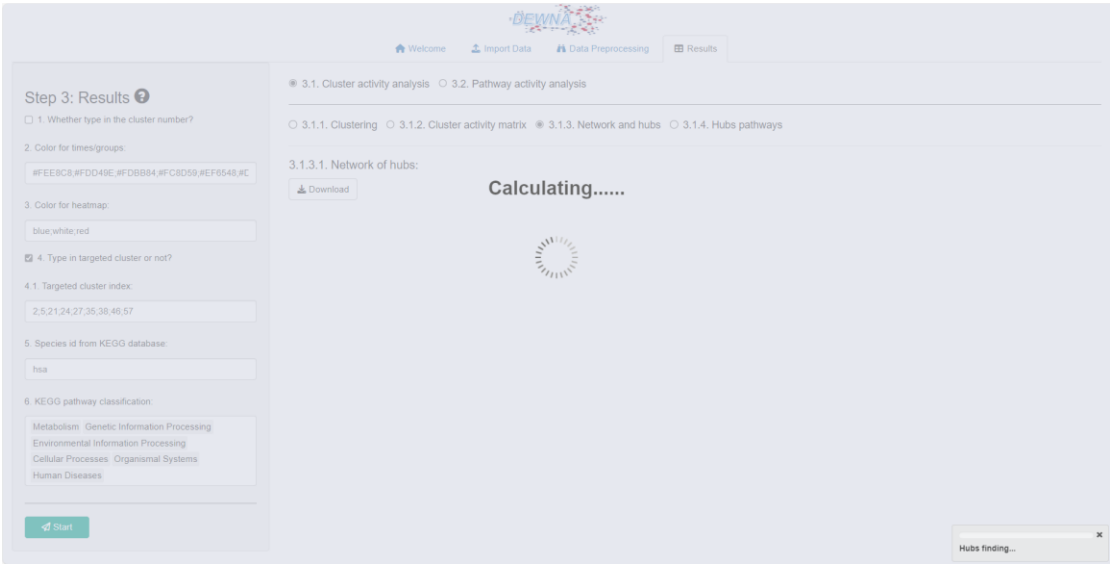
4. Type in targeted cluster or not? If true, users could type in the cluster index of their interest in the “4.1. Targeted cluster index” parameter, which can be obtained from the clustering results in the “3.1. Clustering” part, and this tool would find hubs from these clusters. Otherwise, this tool will analyze all clusters, which will take more time.

5. Species id from KEGG database: The organism id named by KEGG database, for example, Homo sapiens (human) is hsa.

6. KEGG pathway classification: Users can select one or more KEGG pathway classification(s) which are divided into: Metabolism, Genetic Information Processing, Environmental Information Processing, Cellular Processes, Organismal Systems, Human Diseases.

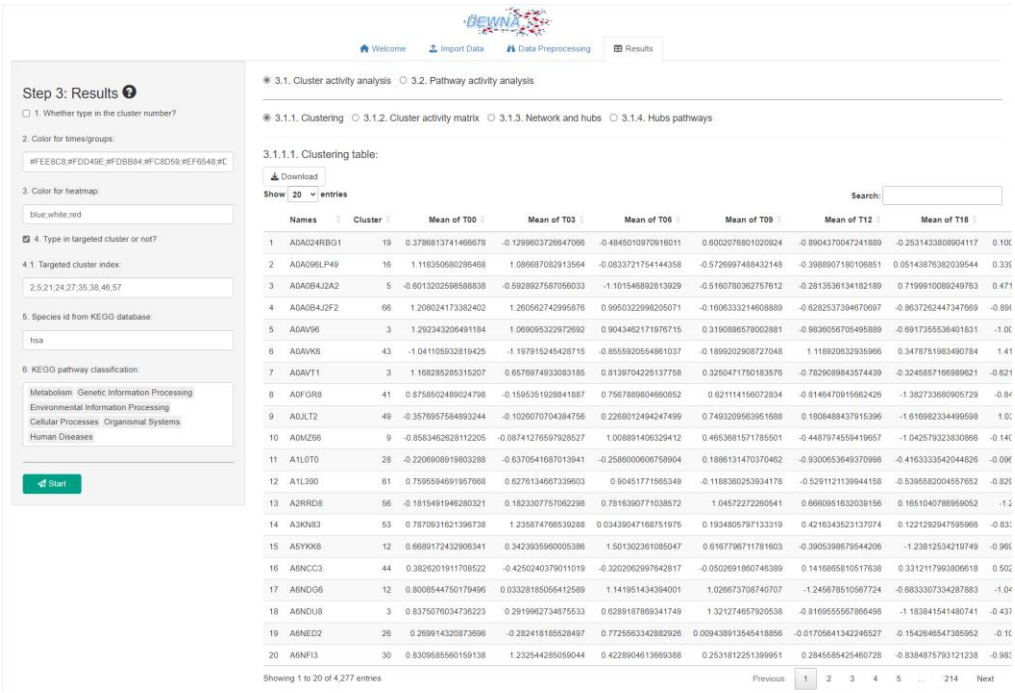
4.2 Results

If after setting these parameters well, users can click “Start” button, *DEWNA* will start to calculate, a process bar will appear in the bottom right corner to tell users where it goes:



Step 3.1. Cluster activity analysis

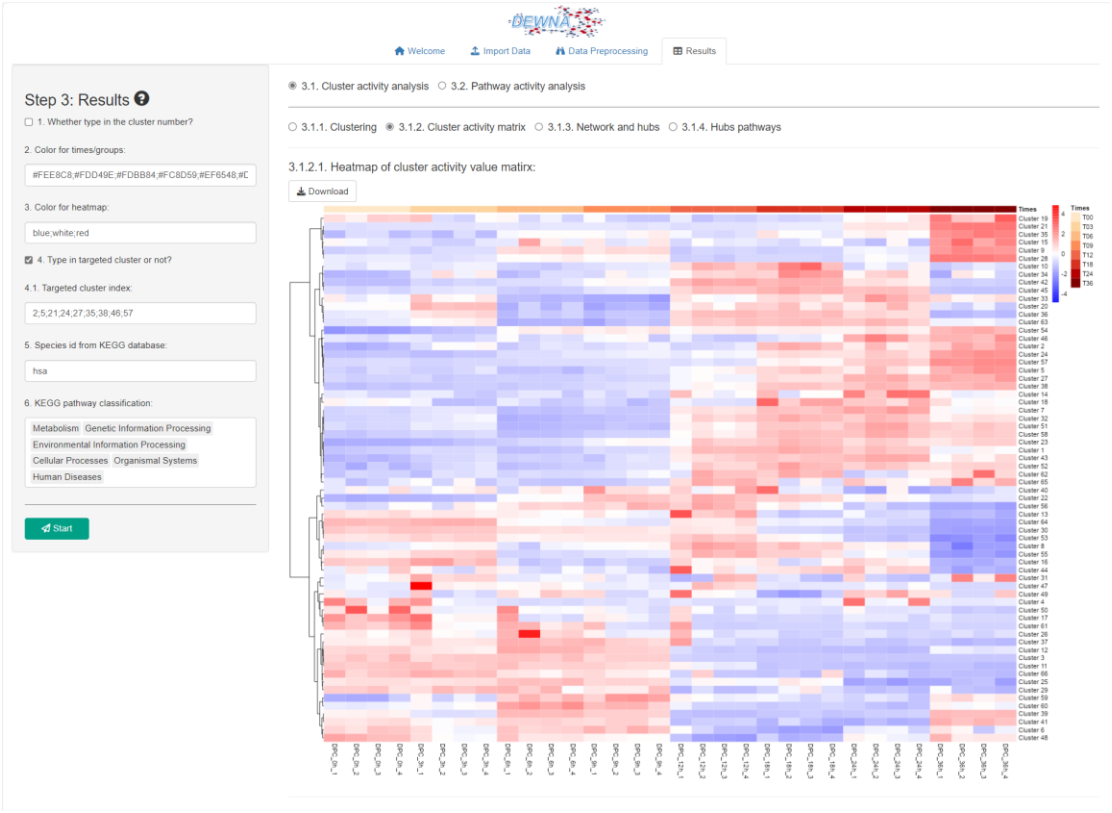
Step 3.1.1. Clustering



Here shows the clustering results, which are derived using the kmeans method by default. In the “3.1.1.1. Clustering table” part, users can obtain the clustering table, in which the first column is the protein ids, the second column is the cluster index (by default, there are 4277 proteins in the example data, thus the cluster number is 66, the cluster index is 1, 2, 3, …, 66), the following columns are the mean values of each timepoint or group.

Step 3.1.2. Cluster activity matrix

This part shows the cluster entropy results. In the “3.1.2.1. Heatmap of cluster activity value matirx” part, users can obtain the heatmap of the cluster activity matrix. From this results, users can review the activity changes of each cluster and pick out those clusters of interest for sequential analyses. For example, by default, we choose clusters “2;5;21;24;27;35;38;46;57” in the “4.1. Targeted cluster index” parameter, which exhit a gradual increase trend. Then, the “3.3. Network and hubs” part will show the results derived from these selected clusters.



In the “3.1.2.2. Cluster activity value table” part, users can obtain the table. The heatmap above is used to show this table.

3.1.2.2. Cluster activity value table:

Download

Show 10 entries

Search:

	DPC_0h_1	DPC_0h_2	DPC_0h_3	DPC_0h_4	DPC_3h_1	DPC_3h_2	DPC_3h_3	DPC_3h_4
Cluster 1	1.470512733179199	1.508929697736735	1.545448250472744	1.558160359358971	1.663287618575284	1.629493027529846	1.645139237396795	1.6755087410181
Cluster 2	1.076151720784559	1.047119973687483	1.088351224737539	1.141120902372676	1.269459573256344	1.432903000633716	1.417231020756922	1.4483667106290
Cluster 3	1.928947096956239	2.129123041492929	1.871800461196986	2.27846018889869	1.791307877089857	1.982327264810091	2.105378890829926	2.07197439068
Cluster 4	1.35636448538628	1.057574509141221	0.684683514374854	1.134893228174811	1.274166841377271	0.7485138080910925	0.6657995803899884	0.64167940696619
Cluster 5	2.583989061600839	2.482681397273237	2.691905571375774	2.675249152778199	2.540701258848886	2.536571635428064	2.637883351184553	2.8595220581600
Cluster 6	0.9673707265676963	0.9338739563111871	1.011920594466023	1.000094492820419	0.9447754859610258	0.8462841883576084	0.8307457443587525	0.82744573309479
Cluster 7	1.01957419431284	1.036064169313094	1.02836825560572	1.017276378629085	1.079637278469469	1.090791468465307	1.091541260996304	1.0757265759499
Cluster 8	1.263492079846213	1.282995634114027	1.305552208512487	1.331306157867538	1.395673402128737	1.446601534257096	1.461560685704808	1.4542255918961
Cluster 9	1.16278817352067	1.189778083579165	1.213415459260657	1.215235249904235	1.173025990701918	1.22110863794491	1.225095724018386	1.2290364204867
Cluster 10	1.17334422043084	1.120680504615815	1.183948275770413	1.196131200035389	1.140164622709815	1.154775076128473	1.188578405020907	1.1968486231011

Showing 1 to 10 of 66 entries

Previous 1 2 3 4 5 6 7 Next

In the “3.1.2.3. Entropy weighted expression table” part, this tool exports the entropy weighted expression matrix for users.

3.1.2.3. Entropy weighted expression table:

Download

Show 10 entries

Search:

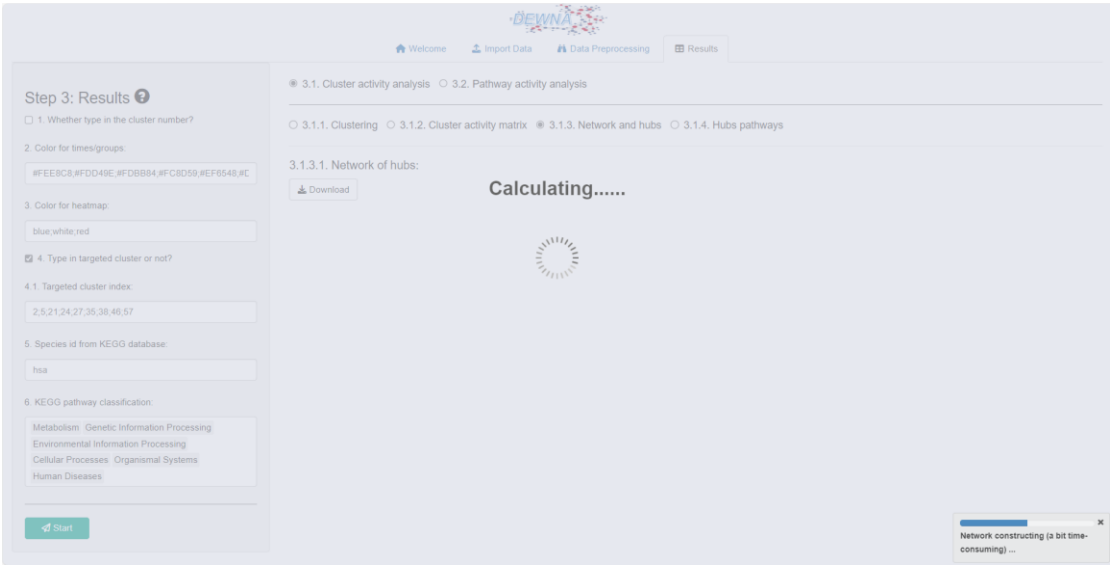
	DPC_0h_1	DPC_0h_2	DPC_0h_3	DPC_0h_4	DPC_3h_1	DPC_3h_2	DPC_3h_3	DPC_3h_4
A0A024RBG1	14.92165044985109	8.239777321928525	12.12297371202288	10.77199205213718	17.2232392846974	7.192650919327646	5.418028765926714	7.190667826030
A0A096LP49	18.73682648774272	17.90695657944164	16.66441855815517	16.58877895544191	17.55925530341173	18.41062136869612	17.06826301047857	16.478582287394
A0A0B4J2A2	496.5288550542998	405.1984619287368	540.1190452507303	517.3497927413785	473.1614345133615	420.1146853430397	483.6154767536706	590.513837975
A0A0B4J2F2	111.8628250784476	120.401372928475	129.4992225389255	139.0958304257047	136.9019317802821	65.65222927359694	155.6954247132014	154.277800906
A0AV96	13.65467518889044	15.50885846114713	14.63379043084909	14.43634444389056	14.38370223807088	13.33258643811034	14.35817684128918	13.0434183781
A0AVK6	11.68331280282028	13.61240610693162	13.21425724682252	13.67545760221351	13.68954658718467	3.95730033078799	11.555397799197	12.774673269578
A0AVT1	8.399336581497105	12.49888389127618	15.43134932680847	14.64556345769918	8.572959525023599	15.00498422625655	12.23523437964055	9.1177466723000
A0FGR8	14.30366562390266	12.16243462774263	13.01618720238766	14.44462361058596	13.46184382108977	9.035030842838575	11.03614144778521	9.125525306004
A0JLT2	25.63374203353975	19.78192797426	20.30395624474272	15.56723309492767	24.55770166085873	17.35880452480403	26.9436880907728	24.491609365915
A0MZ66	17.78147158007467	20.79663820456558	19.27962605958932	19.35209133890288	18.52748738116175	20.2160962026197	22.59168413626705	21.172060860643

Showing 1 to 10 of 4,277 entries

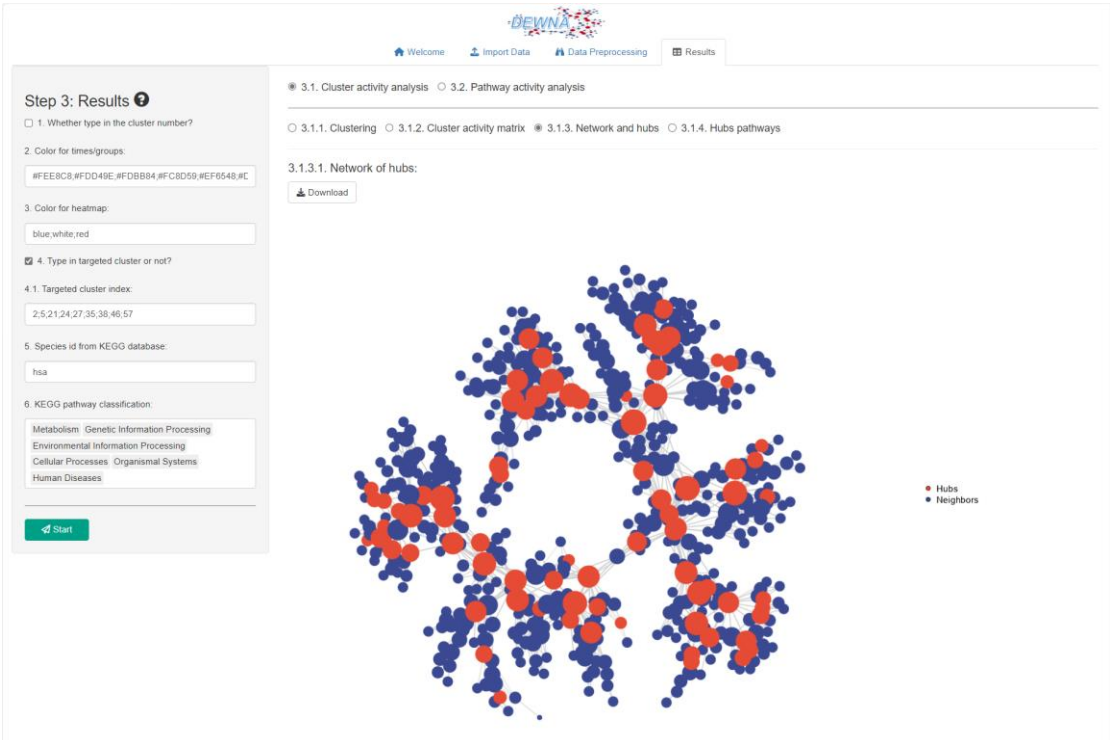
Previous 1 2 3 4 5 … 428 Next

Step 3.1.3. Network and hubs

This part helps users to explore the protein hubs for each cluster using the multiscale embedded gene co-expression network analysis method and plot the network of all protein hubs.



In the “3.1.3.1. Network of hubs” part, this tool shows the network of all protein hubs, in which the red points are hubs and the clue points are the neighbors of those hubs.



In the “3.1.3.2. Entropy weighted expression table of Hubs” part, it shows the protein hubs and their entropy weighted expression, for example, this tool finds 80 protein hubs from the targeted clusters “2;5;21;24;27;35;38;46;57”.

3.1.3.2. Entropy weighted expression table of Hubs:

Download

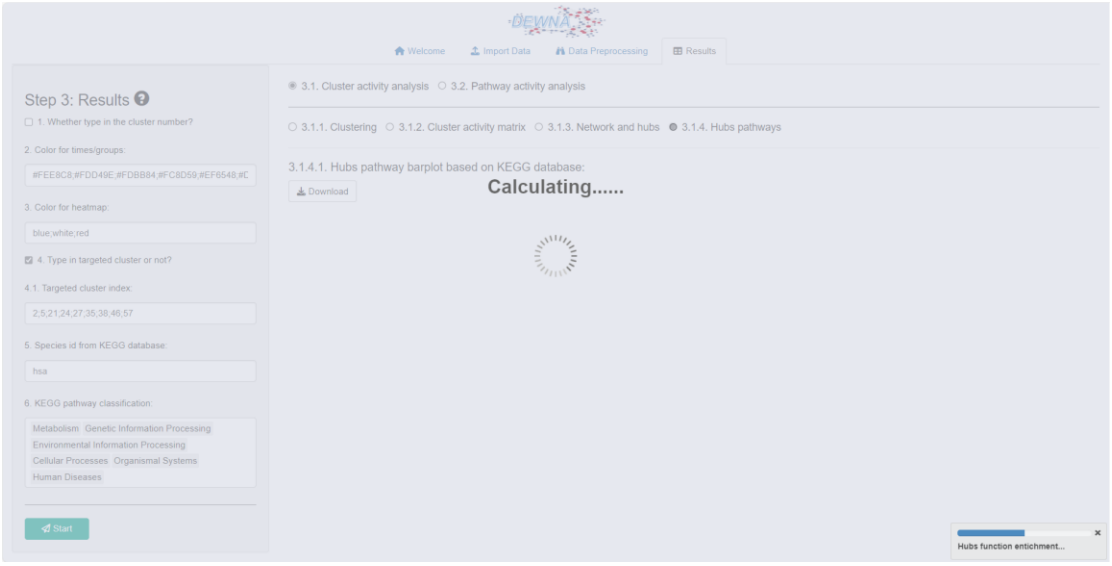
Show 10 entries Search:

	DPC_0h_1	DPC_0h_2	DPC_0h_3	DPC_0h_4	DPC_3h_1	DPC_3h_2	DPC_3h_3	DPC_3h_4
O43242	25.8882566558577	26.85613860621246	28.2399125957546	28.83541279183919	23.31594929806473	24.01443565989878	24.72158623978358	27.55454318547495
O43390	115.9431562031687	122.5415766636071	124.4552446893832	122.116613978806	113.0201279087389	113.7502343595037	118.7943373228184	124.2830342524804
O43592	13.95551061958094	13.50949167368482	13.61717348256535	13.82981814464272	12.88668651854726	13.82507323140456	14.20046291708422	13.53862855570251
O60506	43.95321180295741	44.71504369053867	46.66818988945828	47.43999985507331	46.31972709991335	41.49626452388193	46.43037324388094	47.90016393048749
O60825	5.689563813899634	5.968763265804264	5.78887801029575	5.836964268882309	5.012565751614701	5.972944897234779	5.729587219898522	5.539861789946238
O75369	131.7465456191605	138.0150744474086	138.7253024758305	139.6278566359886	136.3591289386701	145.198713805694	146.5438044716035	149.9987606126825
P04183	17.13258379888165	17.45697431469535	17.01258767845485	15.93445277272646	15.91714670946541	17.22739816605157	17.93229682823374	18.54751052528224
P05387	22.70624667346206	17.23512853787599	20.73577311929658	24.08319699233405	20.05578169184279	23.9751591709893	19.62506222182328	28.36506458210728
P05412	8.575070536083608	17.58887539974177	17.87283207829508	14.51949676589967	21.03043681643489	12.57347251938615	18.00933656069909	14.53135493664421
P05413	89.77873293536045	88.38934350934903	94.77419608162295	95.8320363677423	82.69866494834174	84.71994979704337	91.94061976023907	95.91116841469491

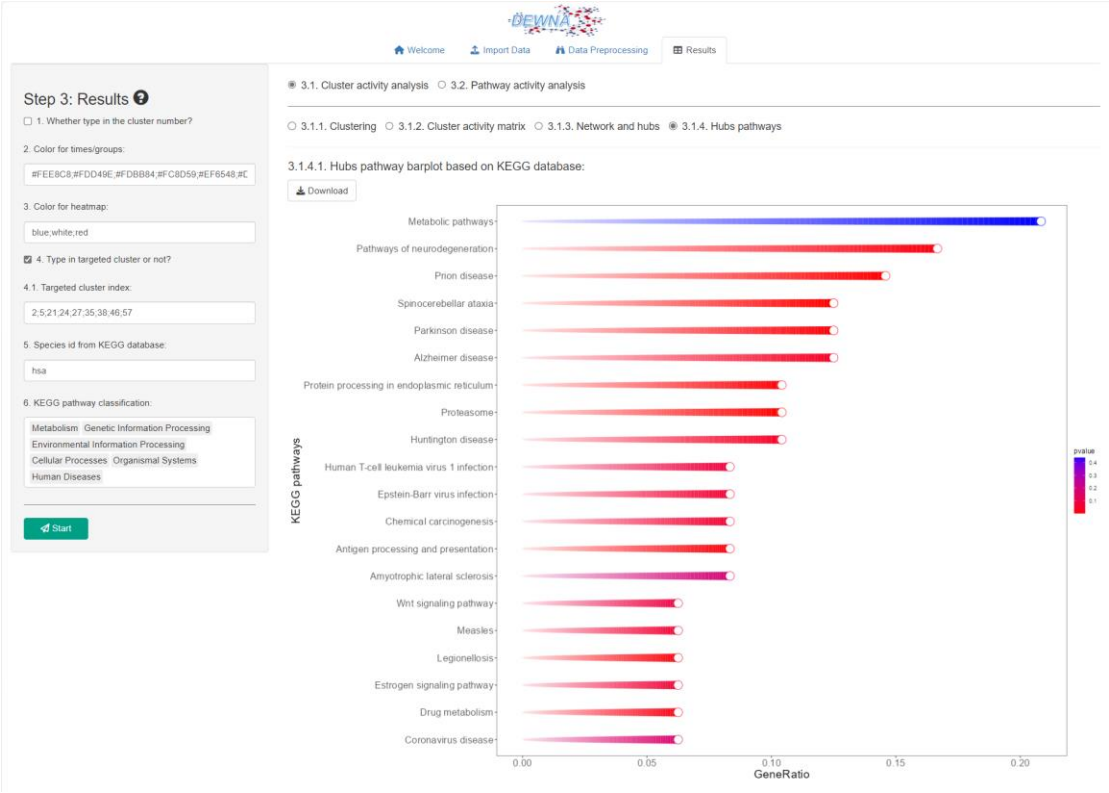
Showing 1 to 10 of 80 entries Previous 1 2 3 4 5 … 8 Next

Step 3.1.4. Hubs pathways

Here shows the KEGG pathway enrichment results of those protein hubs, which were implemented using the clusterProfiler package.



In the “3.1.4.1. Hubs pathway barplot based on KEGG database” part, the plot shows top 20 enriched pathways by default as below.



In the “3.1.4.2. Hubs pathway table based on KEGG database” part, the table shows all the enrichment results.

3.1.4.2. Hubs pathway table based on KEGG database:

Download

Show 10 entries

Search:

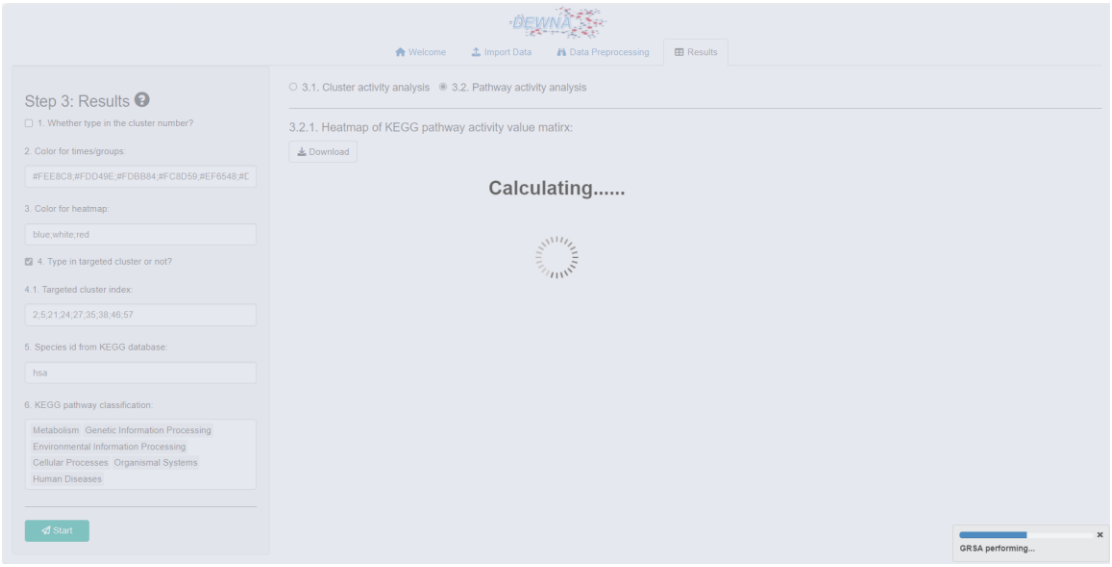
	ID	Description	Count	GeneRatio	BgRatio	pvalue	p.adjust	qvalue	geneID
hsa01100	hsa01100	Metabolic pathways	10	0.20833	2114/11041	0.4399928655373212	0.6997739363665559	0.6701436193142325	O60825/P04183/P082
hsa05022	hsa05022	Pathways of neurodegeneration	8	0.16667	695/11041	0.009502447343194342	0.1579781870806059	0.1512889642798047	O43242/P30405/P458
hsa05020	hsa05020	Prion disease	7	0.14583	417/11041	0.002010001089659721	0.08911004830824763	0.08533688836800922	O43242/P0DMV8/P30
hsa05017	hsa05017	Spinocerebellar ataxia	6	0.125	209/11041	0.0002707575237967991	0.01800537533248714	0.01724297914705931	O43242/P30405/P458
hsa05012	hsa05012	Parkinson disease	6	0.125	397/11041	0.007120567560637054	0.135290783652104	0.1295622067424186	O43242/P30405/P458
hsa05010	hsa05010	Alzheimer disease	6	0.125	555/11041	0.03210406895088198	0.3881673791333912	0.3717313246944229	O43242/P30405/P458
hsa03050	hsa03050	Proteasome	5	0.10417	75/11041	0.0000172317202775962	0.002291818796920295	0.002194777003778043	O43242/P61289/Q132
hsa04141	hsa04141	Protein processing in endoplasmic reticulum	5	0.10417	234/11041	0.003339736988497621	0.1110462548675459	0.1063442567390032	P0DMV8/P27824/Q92
hsa05016	hsa05016	Huntington disease	5	0.10417	431/11041	0.03830284637722179	0.4245232140142082	0.4065477554073541	O43242/P30405/P458
hsa04612	hsa04612	Antigen processing and presentation	4	0.08333	176/11041	0.007033504940966412	0.135290783652104	0.1295622067424186	P0DMV8/P27824/P612

Showing 1 to 10 of 133 entries

Previous 1 2 3 4 5 … 14 Next

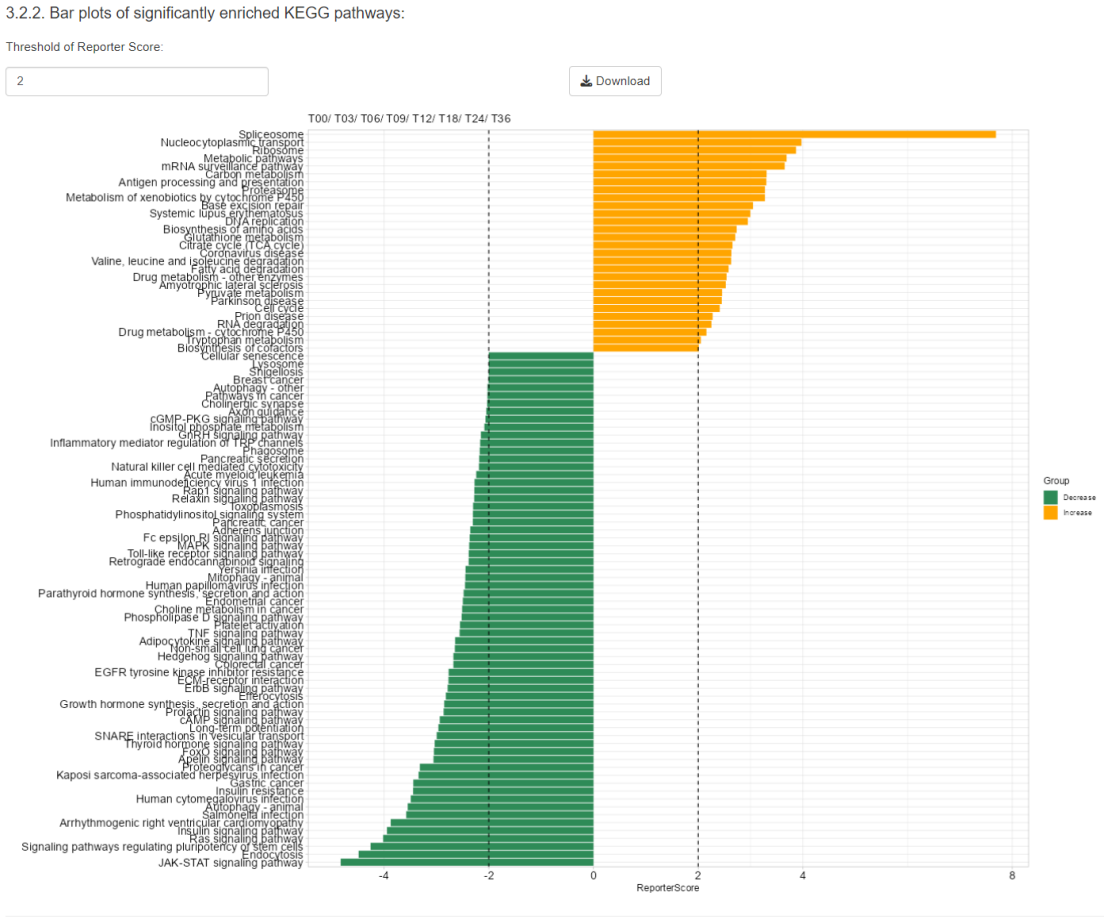
Step 3.2. Pathway activity analysis

This part calculates the pathway activity based on the entropy weighted expression and KEGG database. Users can check the activity change of every pathway.



In the “3.2.1. Heatmap of KEGG pathway activity value matirx” part, if users select all the pathways in the left “6. KEGG pathway classification” parameter, here will show all the pathway activities:

In the “3.2.2. Bar Plots of significantly enriched KEGG pathways” part, users can obtain the enrichment results derived from the generalized reporter score-based analysis. The sign (plus or minus) of the reporter score of each pathway represents the increasing or decreasing trend of the pathway expression.



In the “3.2.3. KEGG pathway activity value and reporter score table” part, users can obtain a pathway activity and reporter score table. The heatmap in the “3.5.1. Heatmap of KEGG pathway activity value matirx” part is visualizing the pathway activity value here.

3.2.3. KEGG pathway activity value and reporter score table:

Download

Show 10 entries Search:

	ID	Description	ReporterScore	DPC_0h_1	DPC_0h_2	DPC_0h_3	DPC_0h_4	DPC_3h_1	DPC_3h
47	hsa03040	Spliceosome	7.68702922495511	2.859265734685003	2.811491717770377	2.926676342174584	2.960647570512229	3.059715608355188	2.9554657356
41	hsa03013	Nucleocytoplasmic transport	3.971406688219677	1.833884396571157	1.856778921984073	1.888828974330788	1.9063510303836	1.870349562032712	1.92542852520
40	hsa03010	Ribosome	3.868946095307709	1.426613236569631	1.369075430851778	1.48307694266788	1.498926659879555	1.903623092607403	1.89824895172
28	hsa01100	Metabolic pathways	3.687641473337201	1.234837446120697	1.186426554046466	1.239139315908255	1.267281813297533	1.271096786245855	1.10882267594
42	hsa03015	mRNA surveillance pathway	3.65150905747925	1.965887780568652	1.989138157812237	2.025781560139102	2.043661071463291	2.037382965761642	2.0538480012
29	hsa01200	Carbon metabolism	3.305852965838621	1.121179654288163	1.101653841586885	1.150768301801851	1.268895223313957	1.061911630248354	1.00305959671
115	hsa04612	Antigen processing and presentation	3.301079017912724	2.448340143493982	2.37708960674393	2.378764556560756	2.45786726114169	2.287060482078171	2.31809356856
48	hsa03050	Proteasome	3.274871030300448	1.028396375374682	1.028022242473303	1.035591717793662	1.039761697805756	0.9925225937139454	0.979922454269
25	hsa00980	Metabolism of xenobiotics by cytochrome P450	3.274671196007341	1.513177318919799	1.487472794941056	1.582036411530894	1.613675587528161	1.506062861419729	1.52440019300
54	hsa03410	Base excision repair	3.047147522625603	1.657221813774888	1.611375314963069	1.657343603610078	1.6402202867482	1.702058718285081	1.67836319338

Showing 1 to 10 of 249 entries Previous 1 2 3 4 5 ... 25 Next

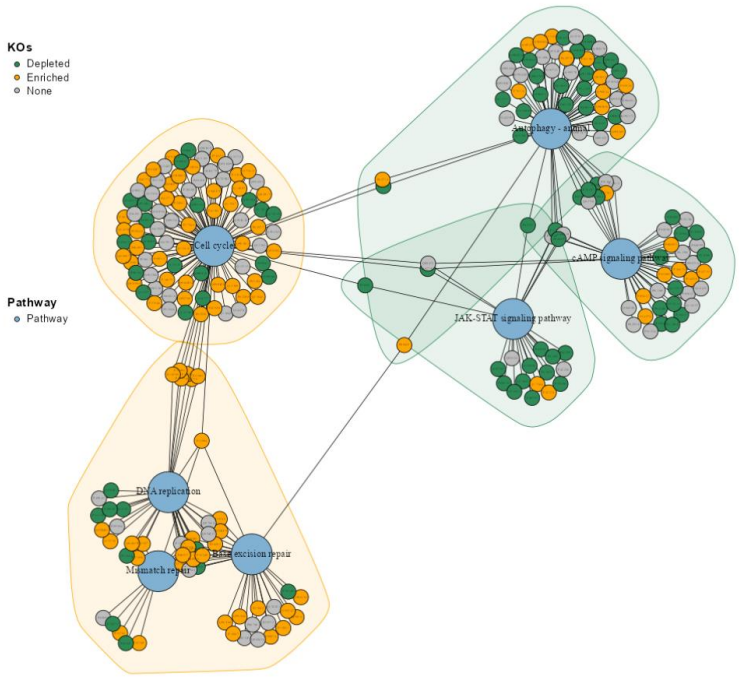
In the “3.2.4. Network of significantly enriched KEGG pathways” part, users can view any pathway-pathway interaction network by setting the parameter “Targeted KEGG pathway ids”, in which all ids are pasted with “;”. The targeted pathway ids can be obtained from the “3.5.3. KEGG pathway activity value and reporter score table” part.

3.2.4. Network of significantly enriched KEGG pathways:

Targeted KEGG pathway ids:

hsa04110;hsa03030;hsa03410;hsa03430;hsa04630;hsa04140;hsa04024

Download



In the “3.2.5. KEGG pathway and protein statistical results table” part, users can obtain the correlation coefficient, p value, Z score for every protein. The network plot above in the “3.5.4. Network of significantly enriched KEGG pathways” part is completed using this table.

3.2.5. KEGG pathway and protein statistical results table:

Download

Show 10 entries Search:

	KOs	id	Description	cor	p.value	p.adjust	Z_score	type
1	P14550	hsa05208	Chemical carcinogenesis - reactive oxygen species	0.6794867741305972	0.00000948920100137706 2	0.00002502177107453126	4.055423533616067	Enriched
2	P14550	hsa01240	Biosynthesis of cofactors	0.6794867741305972	0.00000948920100137706 2	0.00002502177107453126	4.055423533616067	Enriched
3	P14550	hsa01100	Metabolic pathways	0.6794867741305972	0.00000948920100137706 2	0.00002502177107453126	4.055423533616067	Enriched
4	P14550	hsa00620	Pyruvate metabolism	0.6794867741305972	0.00000948920100137706 2	0.00002502177107453126	4.055423533616067	Enriched
5	P14550	hsa00010	Glycolysis / Gluconeogenesis	0.6794867741305972	0.00000948920100137706 2	0.00002502177107453126	4.055423533616067	Enriched
6	P11766	hsa00980	Metabolism of xenobiotics by cytochrome P450	0.5539294354325521	0.00050241958943517 75	0.00094496419701594 28	3.107005567329371	Enriched
7	P11766	hsa00620	Pyruvate metabolism	0.5539294354325521	0.00050241958943517 75	0.00094496419701594 28	3.107005567329371	Enriched
8	P11766	hsa00071	Fatty acid degradation	0.5539294354325521	0.00050241958943517 75	0.00094496419701594 28	3.107005567329371	Enriched
9	P11766	hsa01100	Metabolic pathways	0.5539294354325521	0.00050241958943517 75	0.00094496419701594 28	3.107005567329371	Enriched
10	P11766	hsa00982	Drug metabolism - cytochrome P450	0.5539294354325521	0.00050241958943517 75	0.00094496419701594 28	3.107005567329371	Enriched

Showing 1 to 10 of 10,116 entries Previous 1 2 3 4 5 … 1,012 Next

Every column in the above table means:

KOs: Protein id.

id: KEGG pathway id.

Description: KEGG pathway name.

cor: The correlation coefficient between each KO and the numeric variable. By default, The correlation analysis treats group assignments/timepoints as ordinal (e.g. groups ‘T1’, ‘T2’ and ‘T3’ will be converted to 1, 2 and 3), so the correlation analysis could evaluate if the feature abundance linearly increases or decreases.. The spearman method is used.

p.value: P value of the correlation coefficient test using the cor.test function.

p.adjust: Adjusted p value using the Benjamini-Hochberg method.

Z score: A score is converted from p values in the generalized reporter score-based analysis, a KO with a Z >0 is up-regulated and a KO with a Z<0 is down-regulated.

type: Enriched or Depleted, which is same as up-regulated or down-regulated above.

5. How to run this tool locally?

DEWNA is an open source software for non-commercial use and all codes can be obtained on our GitHub: https://github.com/wangshisheng/DEWNA. If users want to run *DEWNA* on their own computer, they should operate as below:

As this tool was developed with R, you may:

a) Install R. You can download R from here: https://www.r-project.org/.

b) Install RStudio. (Recommendatory but not necessary). You can download RStudio from here: https://www.rstudio.com/.

c) Check packages. After installing R and RStudio, you should check whether you have installed these packages (devtools, shiny, shinyjs, shinyBS, shinyWidgets, readxl, gdata, ggplot2, ggsci, DT, data.table, uwot, pheatmap, RColorBrewer, tidyverse, ggExtra, cowplot, writexl, impute, Amelia, qgraph, MEGENA, clusterProfiler, ReporterScore). You may run the codes below to check them:

```
if(!require(pacman)) install.packages("pacman")
pacman::p_load(devtools, shiny, shinyjs, shinyBS, shinyWidgets, readxl, gdata, ggplot2,
ggsci, DT, data.table, uwot, pheatmap, RColorBrewer, tidyverse, ggExtra, cowplot, writexl,
impute, Amelia, qgraph, MEGENA, clusterProfiler, ReporterScore)
```

Please note, if you find some packages cannot be installed directly using the above command, you can find them in the GitHub source and install them by, for example:

```
library(devtools)
install_github("wangshisheng/DEWNA")
```

d) Run this tool locally

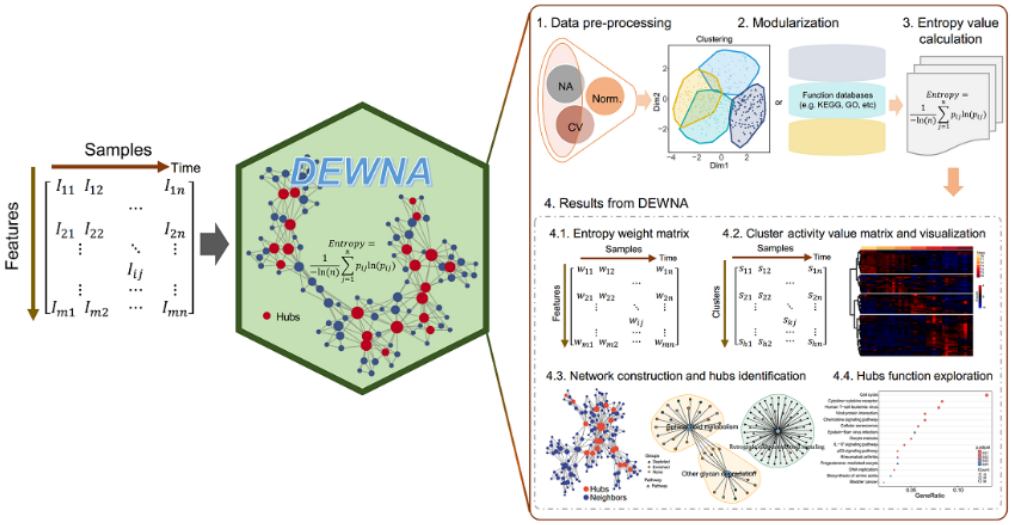
```
if(!require(DEWNA)) devtools::install_github("wangshisheng/DEWNA")
library(DEWNA)
DEWNA_app()
```

Then *DEWNA* will be started as below (same as the online version), and the detailed operation about *DEWNA* can be found in the Supplementary Note parts 1-4 above:

Welcome | Import Data | Data Preprocessing | Results

Welcome to DEWNA

DEWNA (Dynamic Entropy Weight Network Analysis) is a sophisticated and user-friendly stand-alone software designed to facilitate the dynamic analysis of proteins and their associated functions in response to various treatments. This tool implements the entropy weight method, which is a weighting model that assigns weights to indices based on the degree of value dispersion, evaluating the perturbation of clusters or pathways by examining the expression profiles of proteins within these groups. The greater the variability in protein expression upon treatment, the higher the contribution to cluster or pathway perturbation, and therefore, the greater the weight assigned to those proteins. Conversely, proteins with less variability receive lower weights. Additionally, for each cluster, DEWNA integrates the co-expression network analysis to reveal meaningful multi-scale organizations of co-expressed protein network and identifies novel therapeutic targets. On the other hand, for pathway analysis, DEWNA employs the generalized reporter score-based analysis to assess the enriched pathways based on the dynamic weighted protein expression matrices. Therefore, by creatively merging these approaches and extending the functionalities, DEWNA allows for the identification of protein hubs and the elucidation of cluster/pathway entropy weighted profiles during disease progression, providing a comprehensive view of the dynamic changes occurring in response to treatment. The source codes and detailed manual can be accessed at our GitHub. (**Please Note:** If this online version does not work, which means you cannot open the software link, it is probably because our server is down and we will fix it very soon. Or, please try to install this tool and run it locally.)



DEWNA is developed by R shiny (Version 1.3.2), and is free and open to all users with no login requirement. It can be readily accessed by all popular web browsers including Google Chrome, Mozilla Firefox, Safari and Internet Explorer 10 (or later), and so on. We would highly appreciate that if you could send your feedback about any bug or feature request to Shisheng Wang at wsslearning@omicsolution.com.

~~ Enjoy yourself in DEWNA ~~

II. References

(1) Jin, Y.; Shou, Y.; Lei, Q.; Du, C.; Xu, L.; Chen, N.; Ma, W.; Zhu, X.; Zhou, S.; Zheng, Y.; Yu, D. An entropy weight method to integrate big omics and mechanistically evaluate DILI. *Hepatology* **2024**, *79*, 1264-1278.

(2) Song, W. M.; Zhang, B. Multiscale Embedded Gene Co-expression Network Analysis. *PLoS Comput Biol* **2015**, *11*, e1004574.

(3) Peng, C.; Chen, Q.; Tan, S.; Shen, X.; Jiang, C. Generalized reporter score-based enrichment analysis for omics data. *Brief Bioinform* **2024**, *25*.

(4) Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Xuan, Y.; Sondermann, J.; Schmidt, M.; Gomez-Varela, D.; Reiter, L. Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results. *Mol Cell Proteomics* **2017**, *16*, 2296-2309.

(5) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **2016**, *11*, 2301-2319.

(6) Troyanskaya, O.; Cantor, M.; Sherlock, G.; Brown, P.; Hastie, T.; Tibshirani, R.; Botstein, D.; Altman, R. B. Missing value estimation methods for DNA microarrays. *Bioinformatics* **2001**, *17*, 520-525.

(7) Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Miladinovic, S. M.; Cheng, L. Y.; Messner, S.; Ehrenberger, T.; Zanotelli, V.; Butscheid, Y.; Escher, C.; Vitek, O.; Rinner, O.; Reiter, L. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol Cell Proteomics* **2015**, *14*, 1400-1410.

(8) Wang, S.; Li, W.; Hu, L.; Cheng, J.; Yang, H.; Liu, Y. NAguideR: performing and prioritizing missing value imputations for consistent bottom-up proteomic analyses. *Nucleic acids research* **2020**, *48*, e83.

(9) Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **2017**, *45*, D353-D361.