Manual for LINT

1. Introduction

Being one rising stars of "omics" family, lipidomic analyses generate hundreds of thousands of datasets that waiting for in-depth interpretation and data mining. We introduce an interactive web-based tool, LINT-web, to meet the increasing demand of lipidomic data analysis. The LINT-web built on an interactive interface that provides biological and clinical scientists a toolbox to deal with common lipidomic analysis. The web-based tool offers all the users an integrative lipidomic analysis approach that interpret lipid ontology based on a multi-omic integration associated data mining.

The LINT-web consists of two aspects: (1) Single Omic Dataset Analysis, and (2) Multiomic Dataset Analysis. The *Single Omic Dataset Analysis* mostly processes the common statistical analyses on lipidomic dataset including ***lipid differential analysis***, ***multivariate comparison***, ***dimensionality reduction***, ***lipid class summarization***, and ***lipid ontology analysis*** (based on LION/web, www.lipidontology.com). The *Single Omic Dataset Analysis* also integrate basic statistical analyses for metabolomic, transcriptomic, and proteomic datasets. The *Multiomic Dataset Analysis* also named *Intra-omic Dataset Analysis* mostly connect two omic datasets to ***create correlation network***. The data mining of such intra-omic correlation network would provide users ***a broader ontological view of lipidomic results***.

2. Single Omic Dataset Analysis

The Single Omic Dataset Analysis aspect consist statistical processing functions for lipidomics, metabolomics, transcriptomics and proteomics data analyses. Users can click the "Single Omic Dataset Analysis" tab to access and upload your data files and description files. Users can select the data type of aimed "omics" dataset. In addition, the website provides multiple dataset examples to download and demonstration process on the "Try example data" table (Figure 1).

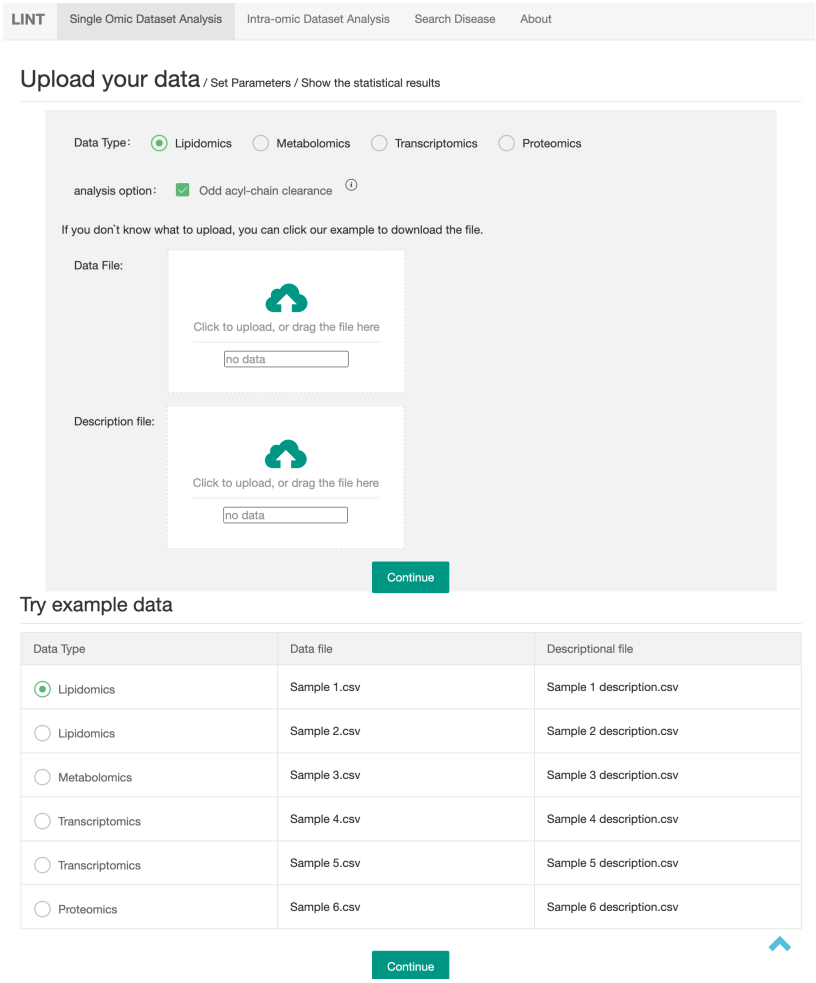


Figure 1 Single omic dataset analysis uploading page

2.1. Lipidomic Analysis

2.1.1 Data Preparation and Upload

LINT-web can deal with main stream lipidomic profiling datasets, the users should firstly convert the lipidomic result table to comma separated value (csv) file; Two files are needed: one contains all the lipid names and their intensities (*data file*), and another contains all the group factors (*description file*).

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	LipidIon	A1	A2	A3	B1	B2	B3	C1	C2	C3	D1	D2	D3
2	Cer(d28:0)+H	9296073	8113669	1.94E+07	1.53E+07	1.11E+07	6949586	1.13E+07	1.39E+07	1.74E+07	2.41E+07	1.09E+07	1.43E+07
3	Cer(d30:0)+H	7809082	6673210	1.39E+07	1.15E+07	8567149	5430360	7827687	1.01E+07	1.23E+07	1.98E+07	7971844	9475264
4	Cer(d30:0)+H	7953969	6837046	1.47E+07	1.29E+07	8513927	6053213	8217586	1.04E+07	1.39E+07	1.96E+07	9094334	1.01E+07
5	Cer(d32:0)+H	6.53E+07	5.47E+07	1.13E+08	8.70E+07	7.42E+07	4.79E+07	5.47E+07	8.34E+07	9.11E+07	1.45E+08	6.08E+07	6.93E+07
6	Cer(d32:0)+H	6526320	6752325	1.21E+07	9967571	9074643	5430770	6261949	8664675	1.10E+07	1.78E+07	6851503	8247776
7	Cer(d18:1_14:0)+H	1.98E+07	1.17E+07	1.41E+07	2.97E+07	6.07E+07	2.47E+07	2.78E+07	2.16E+07	1.14E+07	5.60E+07	2.03E+07	4.32E+07
8	Cer(d18:1_15:0)+H	2.37E+07	1.45E+07	2.04E+07	2.82E+07	3.98E+07	2.32E+07	2.87E+07	2.48E+07	1.39E+07	3.89E+07	2.01E+07	4.10E+07
9	Cer(d34:0)+H	3.07E+07	2.00E+07	4.69E+07	3.35E+07	2.68E+07	1.83E+07	2.12E+07	3.25E+07	3.30E+07	4.93E+07	2.12E+07	2.83E+07
10	Cer(d34:0)+H	6.02E+07	4.69E+07	9.29E+07	7.51E+07	6.42E+07	4.24E+07	4.86E+07	7.29E+07	6.90E+07	1.14E+08	4.69E+07	6.32E+07
11	Cer(d18:0_16:0)+H	191761.4	462353.8	574402.5	1053798	2648486	641049.3	828052.2	537196.8	213321.9	1600548	1642541	3980791
12	Cer(d18:0_16:0)+H	3.29E+07	2.24E+07	3.14E+07	6.23E+07	1.23E+08	5.55E+07	3.62E+07	3.32E+07	2.19E+07	1.08E+08	4.50E+07	8.05E+07
13	Cer(d18:0_16:0)+H	7517346	5448476	6273227	5940047	1.15E+07	6435525	1.00E+07	8907793	5208973	1.22E+07	4422760	9079358
14	Cer(d34:0+O)+H	1833378	1800564	7807036	2850524	2883052	1587968	2303757	3227068	4840955	3857929	2072199	2426981
15	Cer(d18:1_16:0)+H	1969127	1313222	1995496	7581895	1.35E+07	4223372	2211646	1013123	1186796	1.73E+07	4251870	7326317
16	Cer(d34:1)+H-H2O	3954676	1883825	4135501	1.05E+07	2.35E+07	7254127	2312166	4039792	1076484	2.58E+07	5665711	1.38E+07
17	Cer(d18:1_16:0)+H	1174324	2955753	2009533	4502662	5985865	4229977	2649954	2377781	1719654	4042612	3299428	4509372
18	Cer(d18:1_16:0)+H	78390.14	223527.7	189592.7	718920	4299596	1670348	332174.6	2082490	581143.6	6719620	1228487	5877969
19	Cer(d18:1_16:0)+H	2.34E+08	1.48E+08	1.88E+08	2.75E+08	5.71E+08	2.27E+08	3.00E+08	2.61E+08	1.39E+08	5.30E+08	1.84E+08	4.05E+08
20	Cer(d34:1)+H-H2O	7.49E+07	4.54E+07	6.27E+07	8.51E+07	1.79E+08	7.12E+07	1.00E+08	8.12E+07	4.40E+07	1.70E+08	5.89E+07	1.31E+08
21	Cer(d18:2_16:0)+H	1.62E+07	1.08E+07	1.09E+07	2.42E+07	4.49E+07	2.13E+07	2.08E+07	1.73E+07	1.10E+07	4.27E+07	1.70E+07	3.21E+07
22	Cer(d34:2)+H-H2O	4615470	3241210	2870787	6748928	1.40E+07	6236339	5580200	5069591	2902954	1.36E+07	5271212	8977381
23	Cer(d18:1_17:0)+H	8169488	5658972	7244613	8280000	1.92E+07	7292411	1.23E+07	1.00E+07	5471612	1.72E+07	6835477	1.41E+07

LipidSearch Dataset

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Metabolite name	Aml 12_1	Aml 12_2	Aml 12_3	Aml 12_1i	Aml 12_1i	Aml 12_1i	Aml 12_1-	Aml 12_1-	Aml 12_1-	Aml 12_2i	Aml 12_2i	Aml 12_2i-3
2	BMP 36:2|BMP 18:1_18:1	1179000	1371000	1187000	1256000	1136000	1205000	1414000	1336000	1593000	868300	1248000	1226000
3	BMP 38:5|BMP 18:1_20:4	31920	40470	40980	55430	42750	54670	79810	79850	86030	30260	39370	34910
4	BMP 40:7|BMP 18:1_22:6	139000	70710	66380	103600	90680	109500	176200	153400	142000	68030	75990	77780
5	BMP 42:10|BMP 20:4_22:6	13610	7198	3082	13030	13580	10340	34230	30030	22520	6432	7590	8278
6	BMP 44:11|BMP 22:5_22:6	30560	46900	43520	46020	37330	55730	31560	55810	55980	68960	95120	47740
7	BMP 44:12|BMP 22:6_22:6	28940	26450	19810	37590	27700	38180	71600	68340	51880	11150	18870	23490
8	CAR 18:1	621000	104000	95830	84410	114400	105700	235100	177300	86970	108400	118300	107900
9	CE 18:3	61570	63360	60940	73670	54220	61300	51630	52570	53660	44760	54510	41830
10	CE 20:3	139900	77180	70870	81750	77650	66850	38750	57490	35390	42250	60670	52170
11	CE 20:4	223000	280700	247900	233800	259400	287600	112000	149600	129800	107100	216100	191800
12	CE 20:5	80750	82270	90200	95000	101800	105300	47000	56910	43050	40460	70770	60600
13	CE 22:4	37740	41060	42220	35290	40120	39470	19560	24800	20170	11740	31350	24670
14	CE 22:5	117800	143300	154300	144100	169300	167900	60110	93270	82230	70210	125200	110500
15	CE 22:6	195300	193100	198700	223300	211500	236900	110700	125500	116500	98980	168700	134000
16	Cer 34:1;2O|Cer 18:1;2O/16:0	1636000	1605000	1319000	1989000	1442000	1493000	2190000	1989000	2406000	1637000	1506000	1704000
17	Cer 36:3;2O|Cer 18:2;2O/18:1	23020	19540	21100	35050	23720	18350	31240	35640	28630	15470	11520	15390
18	Cer 36:3;2O|Cer 18:2;2O/18:1	172000	13630	11230	4976	25250	N/A	82890	N/A	8834	26800	27010	10040
19	Cer 41:2;2O|Cer 18:2;2O/23:0	110400	106300	80010	214200	183600	94750	138300	228200	126000	92070	154600	87280
20	Cer 42:0;4O|Cer 18:0;3O/24:0;(2(	3342000	258400	106000	222900	125800	224800	392400	443600	147600	92260	128400	95290
21	Cer 42:1;2O|Cer 18:1;2O/24:0	491300	535600	432700	620000	314100	521000	523100	483000	500700	292300	379000	406500
22	Cer 42:1;4O|Cer 18:1;3O/24:0;(2(	856600	94590	38780	84830	57430	71170	174100	122900	42420	38530	56350	38690
23	Cer 42:2;2O|Cer 18:1;2O/24:1	1193000	1233000	937900	1081000	687100	765800	707100	670000	703400	849000	982100	944300

MS-DIAL Dataset

Figure 2. Example of two LINT acceptable lipidomic data files. LINT-web only takes "*LipidIon*" or "*Metabolite name*" as legal name for lipidomic data file, and it just targets to the lipidome identified by LipidSearch (Thermo Fisher Scientific) and MS-DIAL (http://prime.psc.riken.jp/compms/msdial/main.html) for now. More lipid identification software will be supported soon in the future.

	1	2
1	samples	conditions
2	Day0_1	Day0
3	Day0_2	Day0
4	Day0_3	Day0
5	Day1_1	Day1
6	Day1_2	Day1
7	Day1_3	Day1
8	Day2_1	Day2
9	Day2_2	Day2
10	Day2_3	Day2
11	Day4_1	Day4
12	Day4_2	Day4
13	Day4_3	Day4
14	Day8_1	Day8

Figure 3. Example of description file. The description file indicating group factors of data file should be separately uploaded. It pointed out all the sample names to their experimental conditions.

LINT-web also allows users to omit the odd acyl-chain lipids. Users can check the "Odd acyl-chain clearance" under analysis option.(Figure 1)

2.1.2 Data Processing

Once the data are uploaded, users can click the "Continue" button and access the data processing interface (Figure 4).

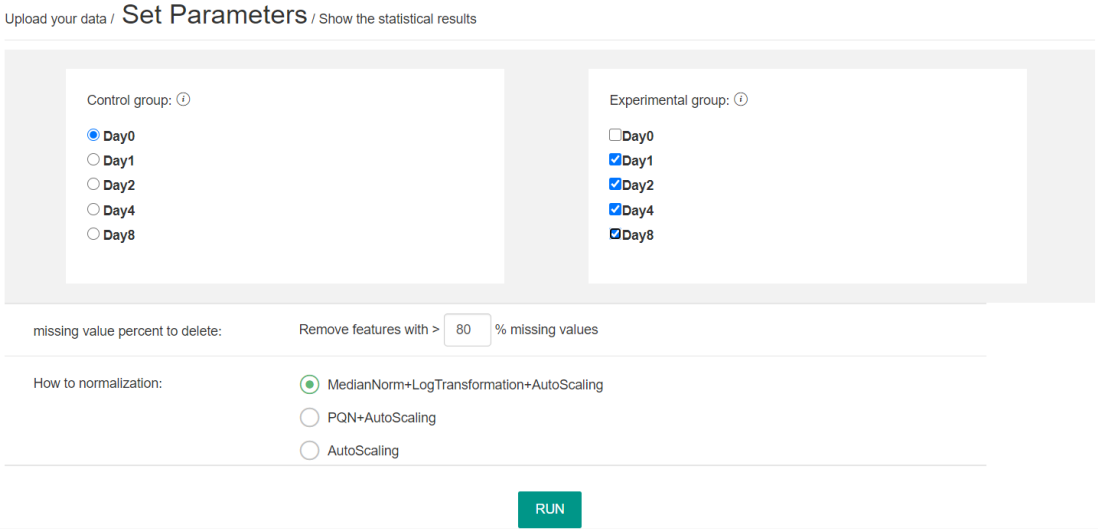


Figure 4. Data processing interface

1) Comparison Groups Setting

Users can set the control group and experimental group(s). Only one control group can be selected, but multiple experimental groups can process. It should be noticed that two group comparison is statistically different to the multiple group comparison.

2) Missing Value Handling

Users can deal with the missing value through lipidomic dataset. Typing a numeric value in the "missing value percent to delete" and lipid hits with missing values percentage lager than the value will be omitted. The rest of the missing values would be refilled with a small value (half of the smallest value through the whole dataset).

3) Normalization Options

LINT-web provides three normalization options commonly used in lipidomic data processing. Option A is sample normalization by median, data log transformation and autoscaling. Option B is sample probabilistic quotient normalization (PQN) and autoscaling. Option C is data autoscaling merely. Users can choose the normalization options according to concrete data situations.

2.1.3 Statistical Analysis and Ontology

LINT-web organizes several modules to perform statistical analyses of lipidomics. Users can click each tab and view the plots and download the results (Figure 5). Some modules offer an interactive interface that users can change the settings and update the plots or results by click the "Update" button.

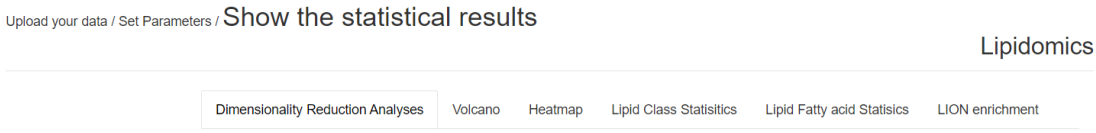


Figure 5. Results of statistical analyses

1) Dimensionality Reduction Analyses

Dimensionality reduction usually helps users to pick out the mostly different features in the lipidomic dataset. Users can access the "Dimensionality Reduction Analyses" interface to process principal component analysis (PCA) or orthogonal partial least squares discriminant analysis (OPLS-DA).

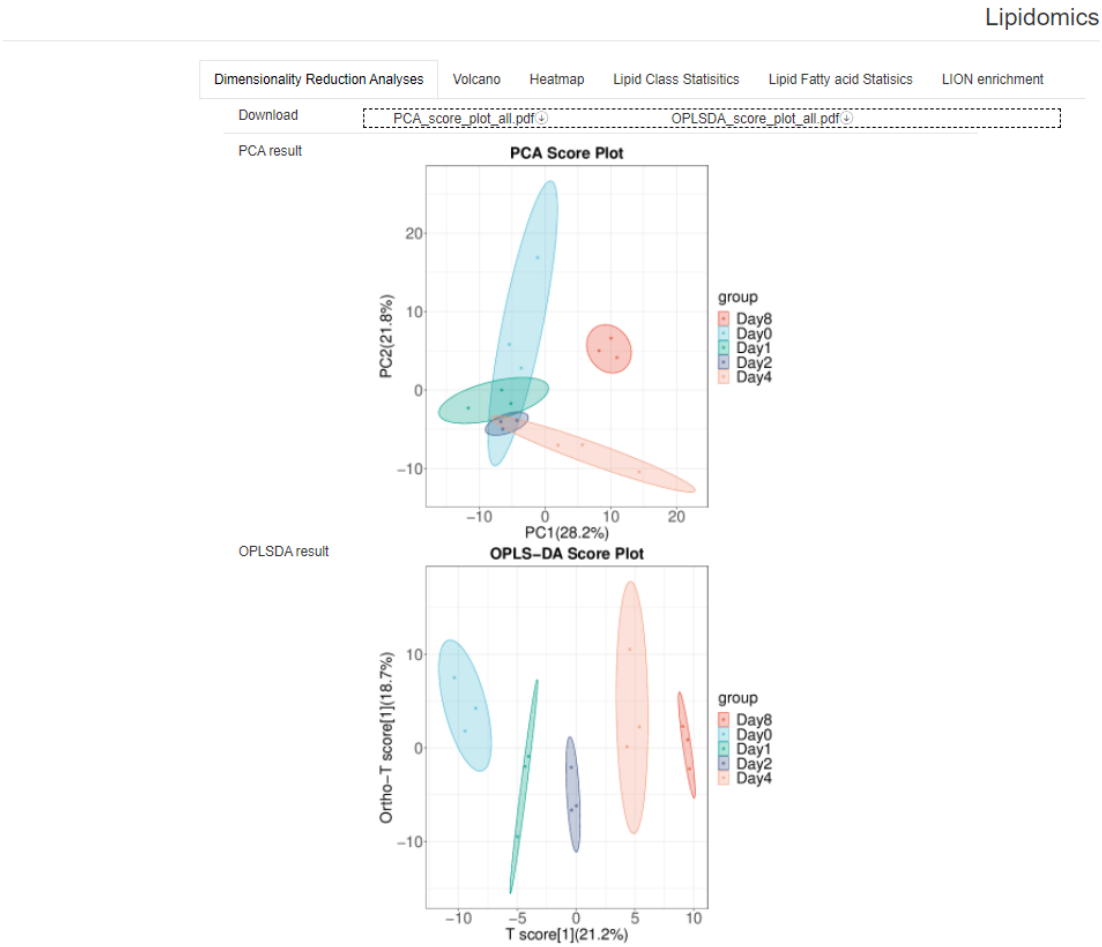


Figure 6. Dimensionality reduction results. The PDF format plots can be downloaded by clicking the "download" button.

2) Volcano Plot

When users only select one control group and one experimental group, the volcano plot is available. Users click the "Volcano" tab can find out the plot and a few parameters can be adjusted (Figure 7). The "Show lipid class" can plot the volcano plot using the summarized lipid classes information. The "Ignore subclass" will overlook the lipid chemical bond linkage information on the volcano plot.

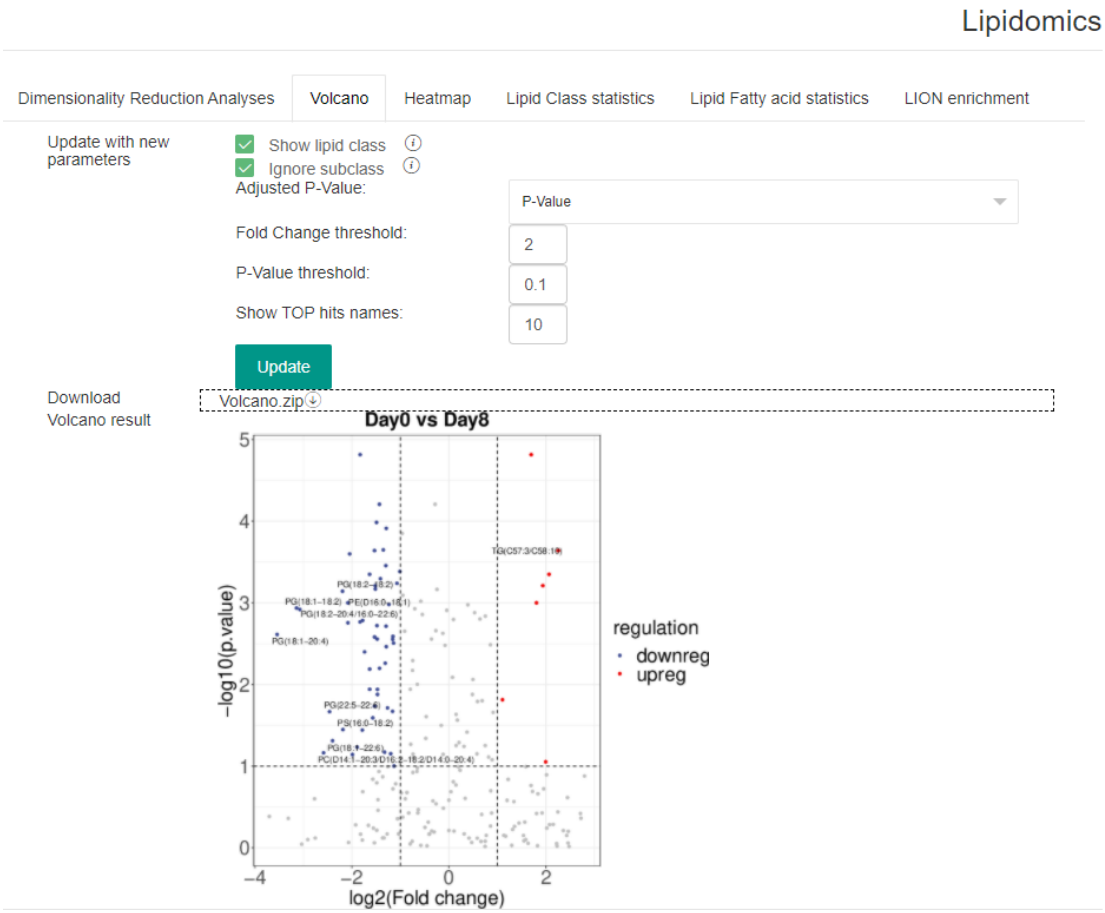


Figure 7. LINT-web let users to set the fold change and p-Value threshold to screen off the non-significantly altered lipids. In addition, the users can set the top lipids to show their names on the plot. All the pdf plots can be downloaded on the same web page.

3) Heatmap

The heatmap can usually provide users an impression of lipidomic dataset. Users can access the heatmap on the "Heatmap" tab (Figure 8).

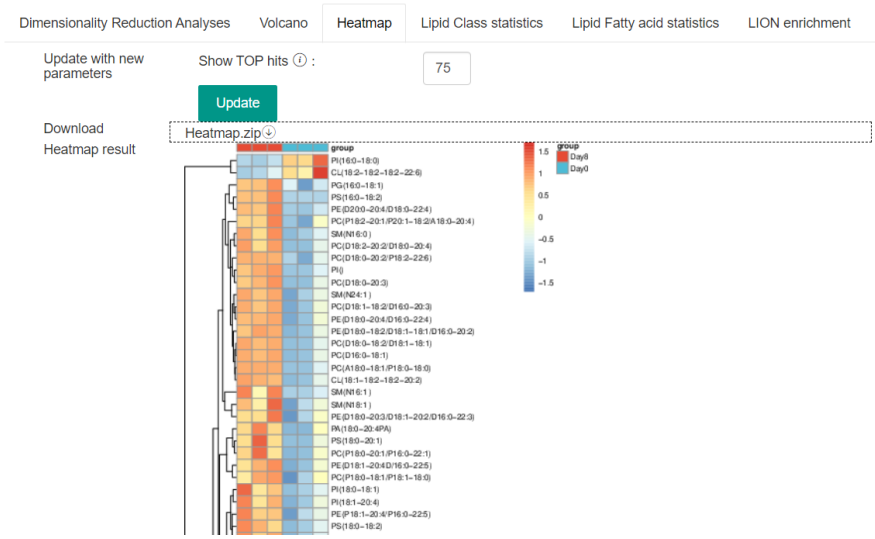


Figure 8. Users can set the number to only show the top hits on the heatmap. All the plots can be downloaded on the same web page.

4) Lipid Class Statistics

Users can access the lipid class summarized information through the "Lipid Class Statistics" tab (Figure 9). The "Ignore subclass" will overlook the lipid chemical bond linkage information while processing the lipid class summary. The summarization of lipid class has three different plots: box-whisker plots (Figure 10), cumulative histograms (Figure 11) and heatmaps (Figure 12).

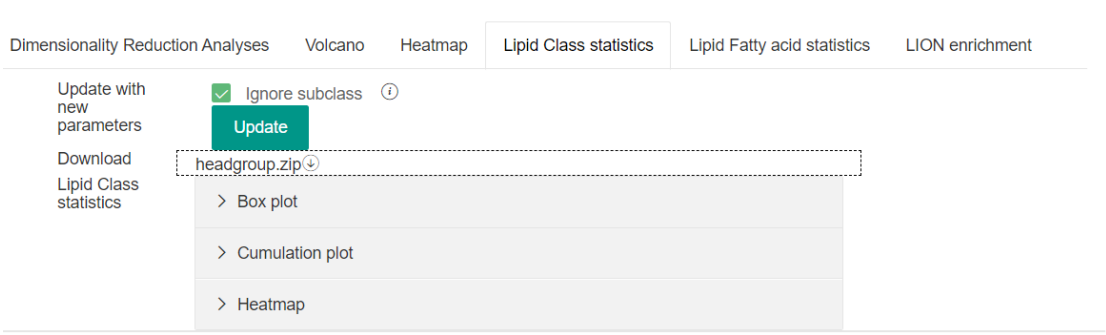


Figure 9 "Lipid Class Statistics" tab overview

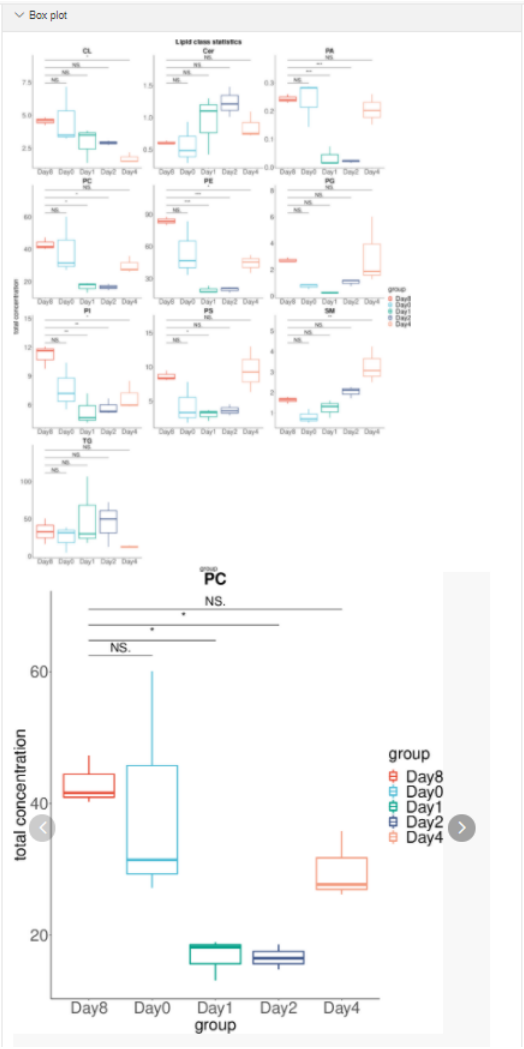


Figure 10. Box-whisker plots

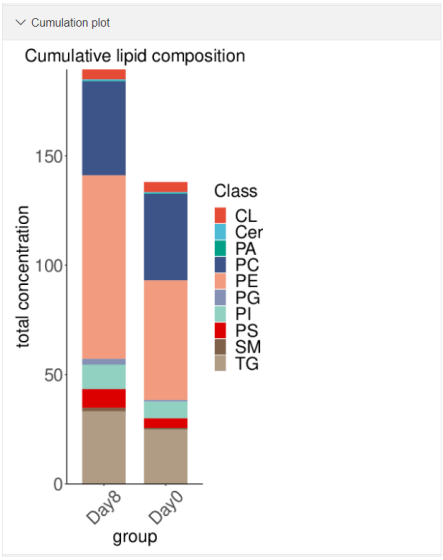


Figure 11. Cumulative histograms

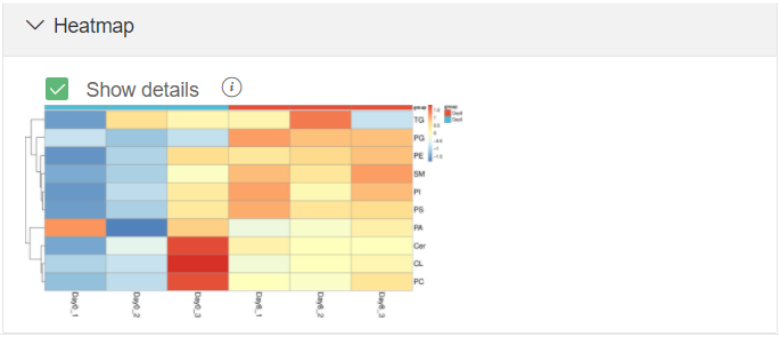


Figure 12. Heatmap

5) Lipid Fatty Acid Statistics

LINT-web provides a way to classify all lipids based on their linked acyl-chain differences (i.e., the attached fatty acids). Users can access three types of plot on the "Lipid Fatty acid statistics" tab (Figure 13); (a) grid plot (Figure 14) which illustrates the fatty acid chain length and numbers of double bond on a grid, (b) heatmap (Figure 15) can present the mostly altered lipid with a certain acyl-chain, and (c) box-whisker plot (Figure 16) indicates the lipid with certain acyl-chain altering in different groups.

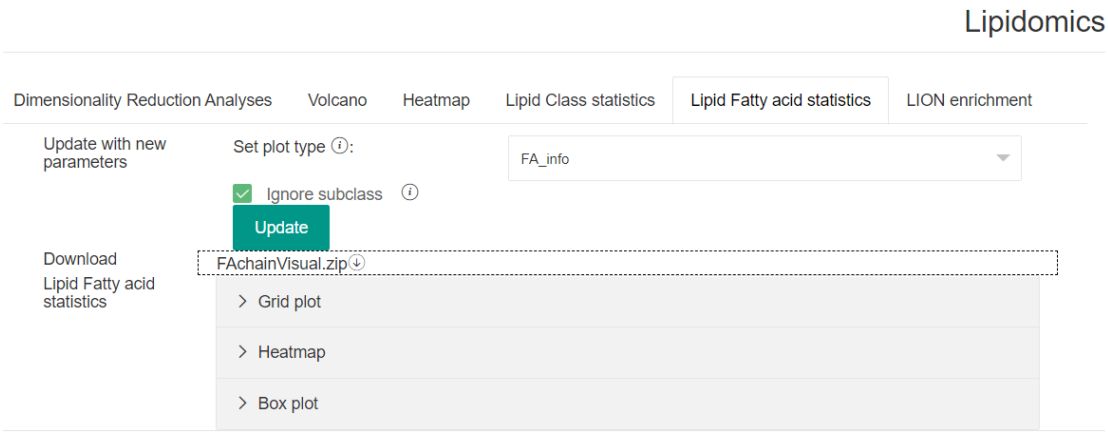


Figure 13. "Lipid Fatty acid statistics" tab overview. Users can set two summarization manner which the "FA_info" only consider the lipids with clear acyl-chain information, but the "all_info" consider all lipids including those with ambiguous acyl-chain information. All the plots can be downloaded on the same web page.

Lipid Fatty acid statistics

Grid plot

Figure 14. Grid plot

Heatmap

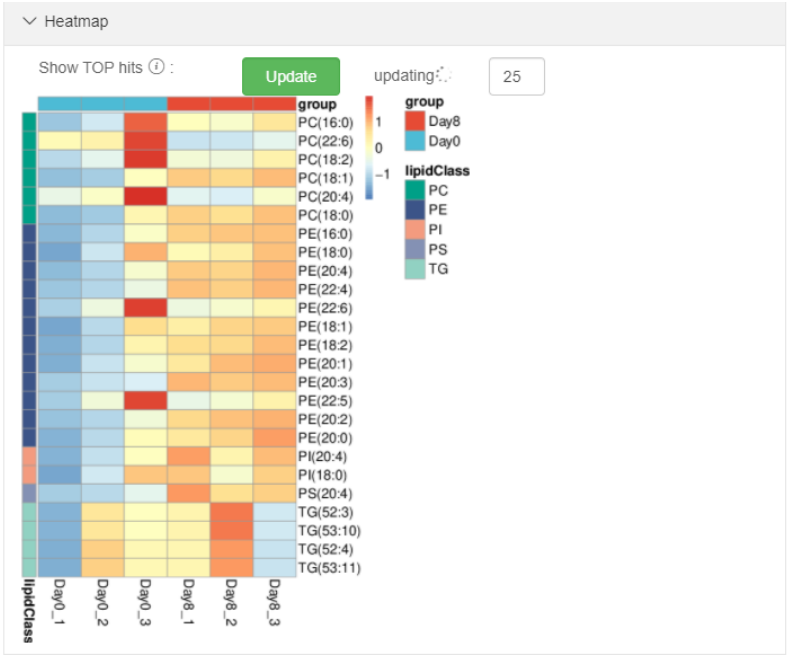


Figure 15. Heatmap

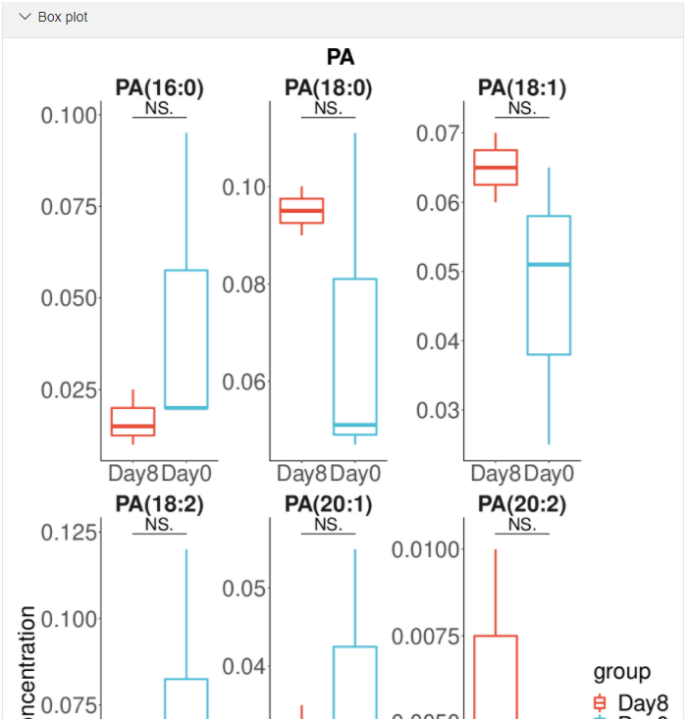


Figure 16. Box-whisker plot

6) LION enrichment

The LINT-web support lipidomic ontology analysis (Figure 17) using a published database, LION/web (www.lipidontology.com). Although we aim to interpret the lipid ontology using an intra-omic approach, the LION/web database can supplement the shortage when analyzing only lipidomic dataset.

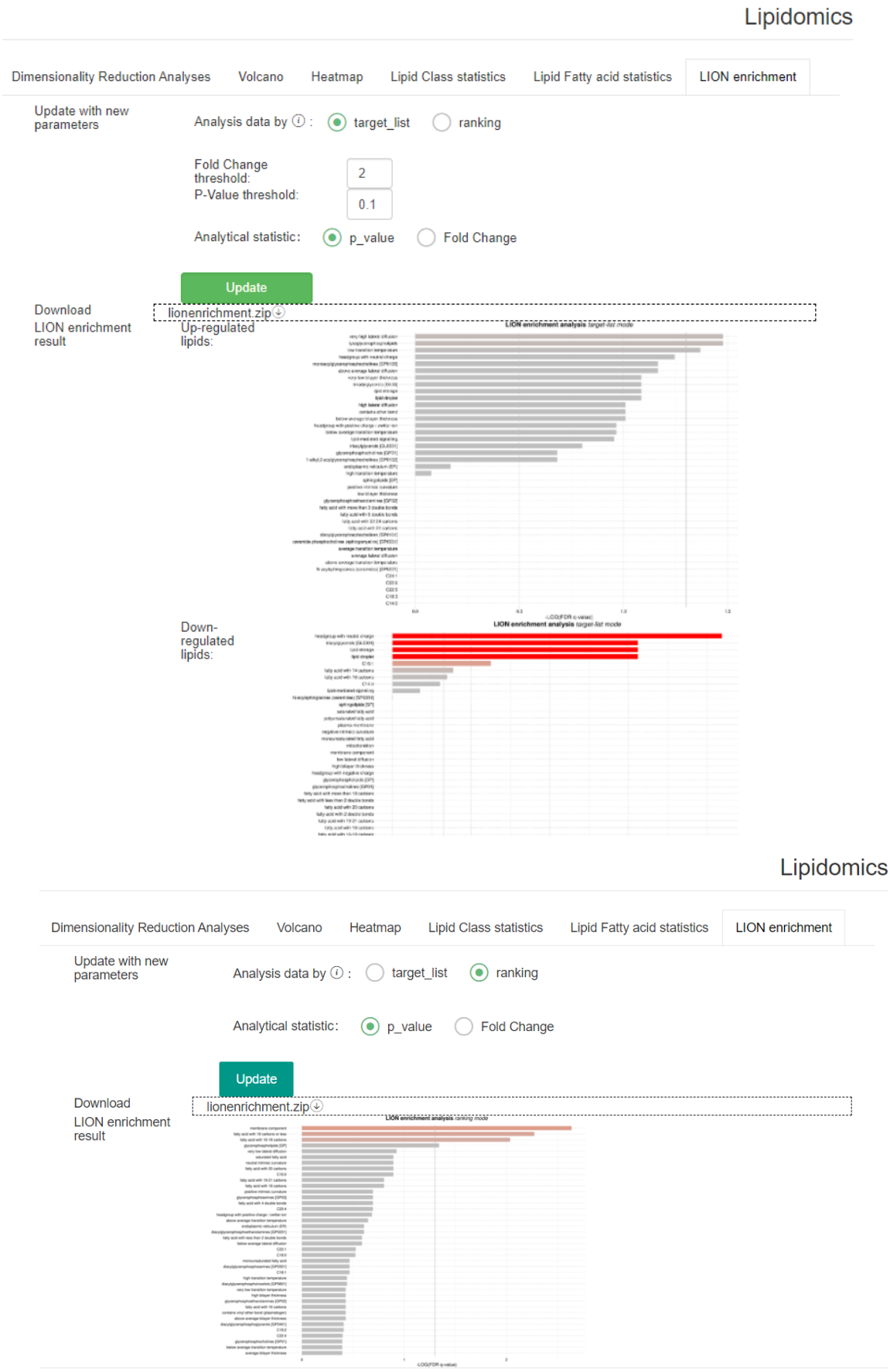


Figure 17. The LINT-web offers “target_list” and “ranking” mode to process such lipid

enrichment analysis, which more detail information can be found on the lipidontology.com website. Plots of lipid enrichment results can be downloaded on the same web page.

2.2 Transcriptomic Analyses

LINT-web provides the modules to analyze two types of transcriptomic datasets that users can either upload RNA-seq results or microarray results as they need to (Figure 18).

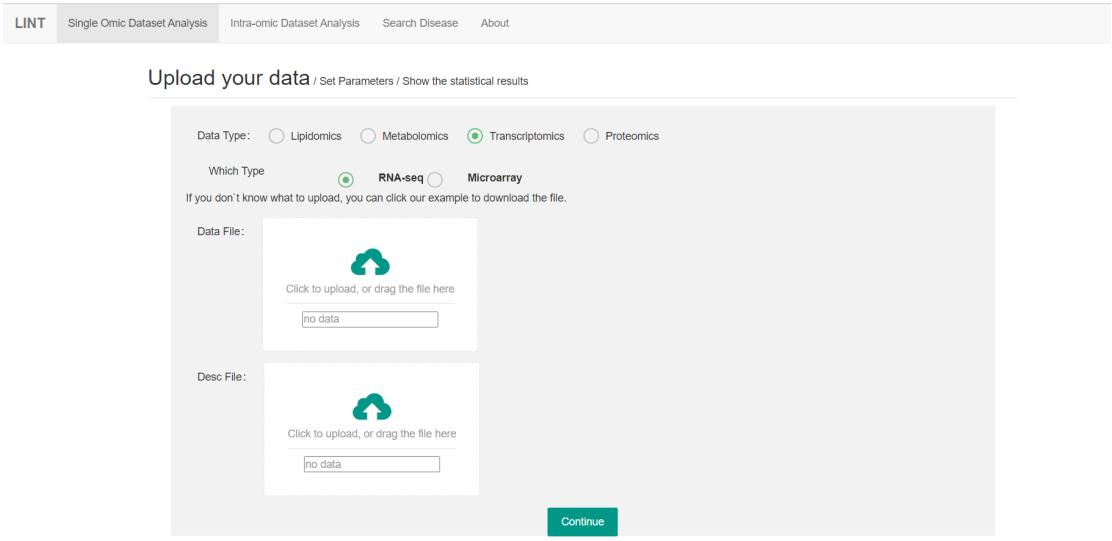
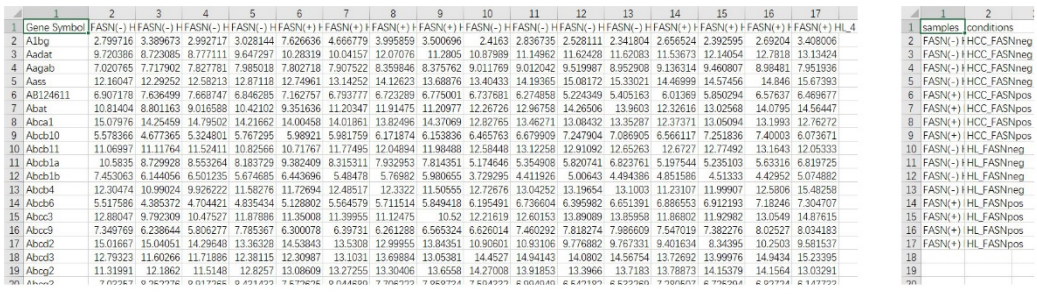


Figure 18. Transcriptomic analyses data uploading interface

2.2.1 Data Preparation and Upload

Users should firstly convert the transcriptomic result table to comma separated value (csv) file or tab separated value (tsv) file; (Figure 19).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	Gene Symbol	FASN(-) H	FASN(-) H	FASN(-) H	FASN(-) H	FASN(+) H	FASN(+) H	FASN(+) H	FASN(+) H	FASN(-) H	FASN(-) H	FASN(-) H	FASN(-) H	FASN(+) H	FASN(+) H	FASN(+) H	FASN(+) HL_4
2	A1bg	2.799716	3.389673	2.992717	3.028144	7.626636	4.666779	3.995859	3.500696	2.4163	2.836735	2.528111	2.341804	2.656524	2.392595	2.69204	3.408006
3	Aadat	9.720386	8.723085	8.777111	9.647297	10.28319	10.04157	12.07076	11.2805	10.87989	11.14962	11.62428	11.62083	11.53673	12.14054	12.7818	13.13424
4	Aagab	7.020765	7.717902	7.827781	7.985018	7.802718	7.907522	8.359846	8.375762	9.011769	9.012042	9.519987	8.952908	9.136314	9.460807	8.98481	7.951936
5	Aass	12.16047	12.29252	12.58213	12.87118	12.74961	13.14252	14.12623	13.68876	13.40433	14.19365	15.08172	15.33021	14.46999	14.57456	14.846	15.67393
6	AB124611	6.907178	7.636499	7.668747	6.846285	7.162757	6.793777	6.723289	6.775001	6.737681	6.274858	5.224349	5.405163	6.01369	5.850294	6.57637	6.469677
7	Abat	10.81404	8.801163	9.016588	10.42102	9.351636	11.20347	11.91475	11.20977	12.26726	12.96758	14.26506	13.9603	12.32616	13.02568	14.0795	14.56447
8	Abca1	15.07976	14.25459	14.79502	14.21662	14.00458	14.01861	13.82496	14.37069	12.82765	13.46271	13.08432	13.35287	12.37371	13.05094	13.1993	12.76272
9	Abcb10	5.578366	4.677365	5.324801	5.767295	5.98921	5.981759	6.171874	6.153836	6.465763	6.679909	7.247904	7.086905	6.566117	7.251836	7.40003	6.073671
10	Abcb11	11.06997	11.11764	11.52411	10.82566	10.71767	11.77495	12.04894	11.98488	12.58448	13.12258	12.91092	12.65263	12.6727	12.77492	13.1643	12.05333
11	Abcb1a	10.5835	8.729928	8.553264	8.183729	9.382409	8.315311	7.932953	7.814351	5.174646	5.354908	5.820741	6.823761	5.197544	5.235103	5.63316	6.819725
12	Abcb1b	7.453063	6.144056	6.501235	5.674685	6.443696	5.48478	5.76982	5.980655	3.729295	4.411926	5.00643	4.494386	4.851586	4.51333	4.42952	5.074882
13	Abcb4	12.30474	10.99024	9.926222	11.58276	11.72694	12.48517	12.3322	11.50555	12.72676	13.04252	13.19654	13.1003	11.23107	11.99907	12.5806	15.48258
14	Abcb6	5.517586	4.385372	4.704421	4.835434	5.128802	5.564579	5.711514	5.849418	6.195491	6.736604	6.395982	6.651391	6.886553	6.912193	7.18246	7.304707
15	Abcc3	12.88047	9.792309	10.47527	11.87886	11.35008	11.39955	11.12475	10.52	12.21619	12.60153	13.89089	13.85958	11.86802	11.92982	13.0549	14.87615
16	Abcc9	7.349769	6.238644	5.806277	7.785367	6.300078	6.39731	6.261288	6.565324	6.626014	7.460292	7.818274	7.986609	7.547019	7.382276	8.02527	8.034183
17	Abcd2	15.01667	15.04051	14.29648	13.36328	14.53843	13.5308	12.99955	13.84351	10.90601	10.93106	9.776882	9.767331	9.401634	8.34395	10.2503	9.581537
18	Abcd3	12.79323	11.60266	11.71886	12.38115	12.30987	13.1031	13.69884	13.05381	14.4527	14.94143	14.0802	14.56754	13.72692	13.99976	14.9434	15.23395
19	Abcg2	11.31991	12.1862	11.5148	12.8257	13.08609	13.27255	13.30406	13.6558	14.27008	13.91853	13.3966	13.7183	13.78873	14.15379	14.1564	13.03291

Data file

	1	2
1	samples	conditions
2	FASN(-) H	HCC_FASNneg
3	FASN(-) H	HCC_FASNneg
4	FASN(-) H	HCC_FASNneg
5	FASN(-) H	HCC_FASNneg
6	FASN(+) H	HCC_FASNpos
7	FASN(+) H	HCC_FASNpos
8	FASN(+) H	HCC_FASNpos
9	FASN(+) H	HCC_FASNpos
10	FASN(-) H	HL_FASNneg
11	FASN(-) H	HL_FASNneg
12	FASN(-) H	HL_FASNneg
13	FASN(-) H	HL_FASNneg
14	FASN(+) H	HL_FASNpos
15	FASN(+) H	HL_FASNpos
16	FASN(+) H	HL_FASNpos
17	FASN(+) H	HL_FASNpos
18		
19		

Description file

Figure 19. Example of LINT acceptable transcriptomic data files. One contains all the gene names and their expressions (*data file*), and another contains all the group factors (*description file*). If

users can provide “batch” information in the *description file*, LINT-web will align the samples and eliminate the batch effect of RNA sequencing automatically.

2.2.2 Data Processing

LINT-web can only compare two transcriptomic group at one that users can only select only one control group and one experimental group at the same time (Figure 20).

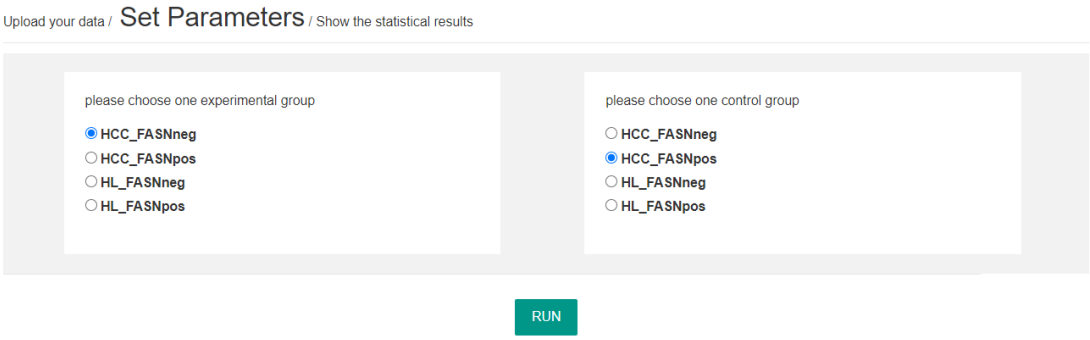


Figure 20. Data processing interface

2.2.3 Gene Differential Expression and Enrichment Analysis

The transcriptomic data analyses are similar to the former lipidomic data processing (Figure 21).

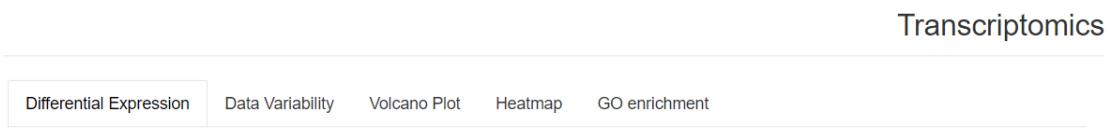
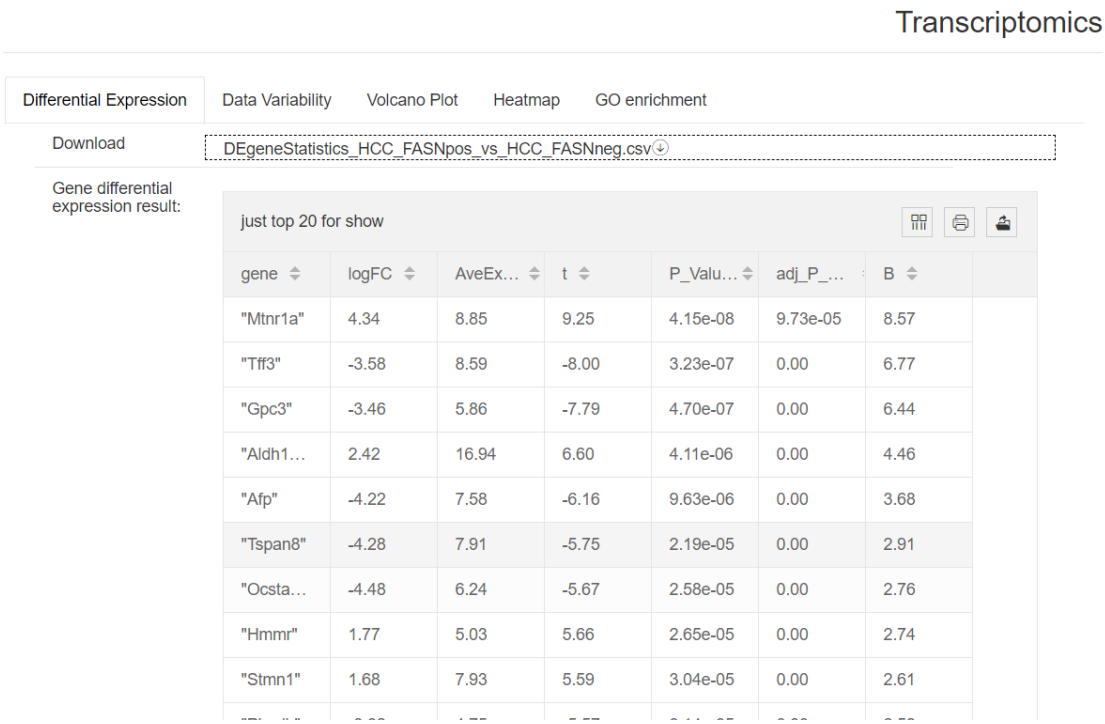


Figure 21. Results of statistical analyses

1) Gene differential expression

The gene differential expression table can be accessed at “Differential Expression” tab (Figure 22), and users can download the csv format table on the same web page.



gene	logFC	AveEx…	t	P_Valu…	adj_P_…	B
"Mtnr1a"	4.34	8.85	9.25	4.15e-08	9.73e-05	8.57
"Tff3"	-3.58	8.59	-8.00	3.23e-07	0.00	6.77
"Gpc3"	-3.46	5.86	-7.79	4.70e-07	0.00	6.44
"Aldh1…	2.42	16.94	6.60	4.11e-06	0.00	4.46
"Afp"	-4.22	7.58	-6.16	9.63e-06	0.00	3.68
"Tspan8"	-4.28	7.91	-5.75	2.19e-05	0.00	2.91
"Ocsta…	-4.48	6.24	-5.67	2.58e-05	0.00	2.76
"Hmmr"	1.77	5.03	5.66	2.65e-05	0.00	2.74
"Stmn1"	1.68	7.93	5.59	3.04e-05	0.00	2.61

Figure 22. Gene differential expression results

2) Data Variability

The data variability of transcriptomic results is evaluated by PCA score plot and heatmap (Figure 23). All the plots can be downloaded on the same web page.

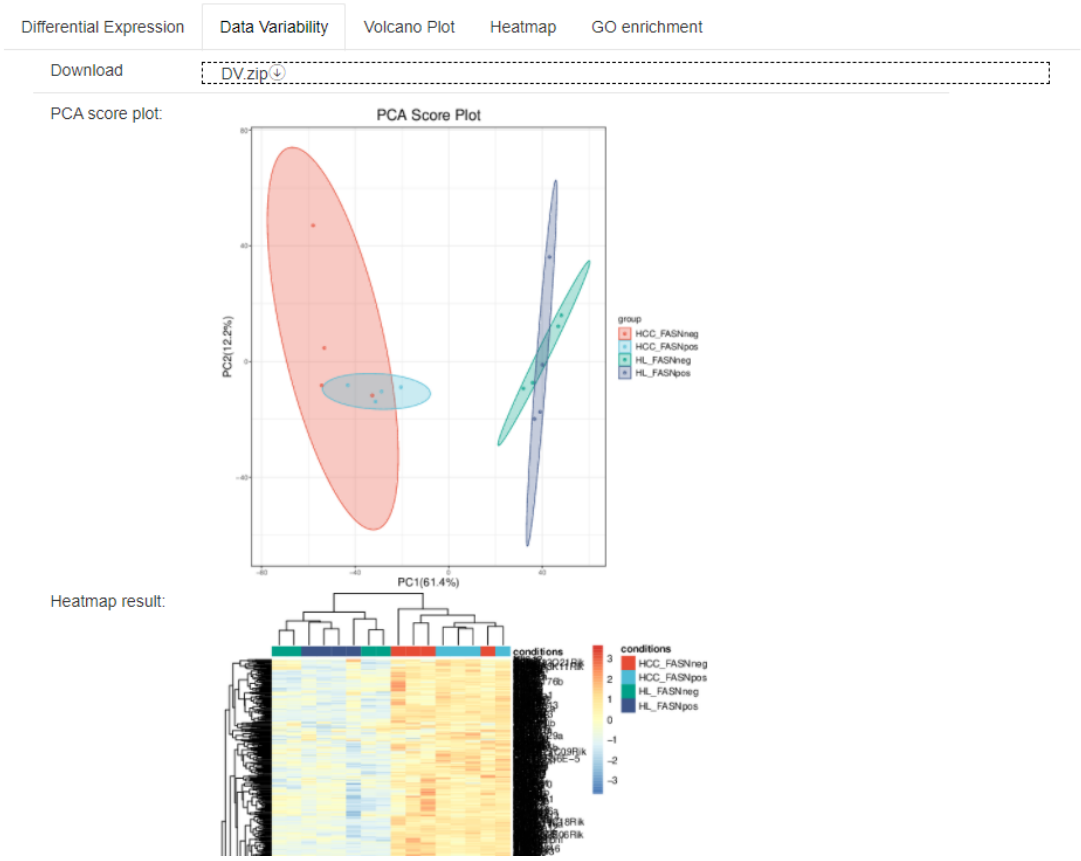


Figure 23. Data variability results

3) Volcano Plot

Using a volcano plot, users can readily figure out the most altered gene expression. Accessing the "Volcano Plot" tab(Figure 24).

Figure 24. Users can set the fold change threshold and p-value threshold to change the screen off hits on the volcano plot. Also, users can justify the numbers of top hits shown on the volcano plots. All the plots can be downloaded on the same web page.

4) Heatmap

A heatmap that only plot the top altered gene hits can be found at "Heatmap" tab (Figure 25).

Figure 25. Users can also customize the number of top hits shown on the heatmap. All the plots can be downloaded on the same web page as PDF format.

5) GO Enrichment analysis

LINT-web integrates GO enrichment analysis (http://geneontology.org/docs/go-enrichment-analysis/) to perform gene ontology analysis. Users can analyze the gene expression enrichment at "GO enrichment" tab (Figure 26).

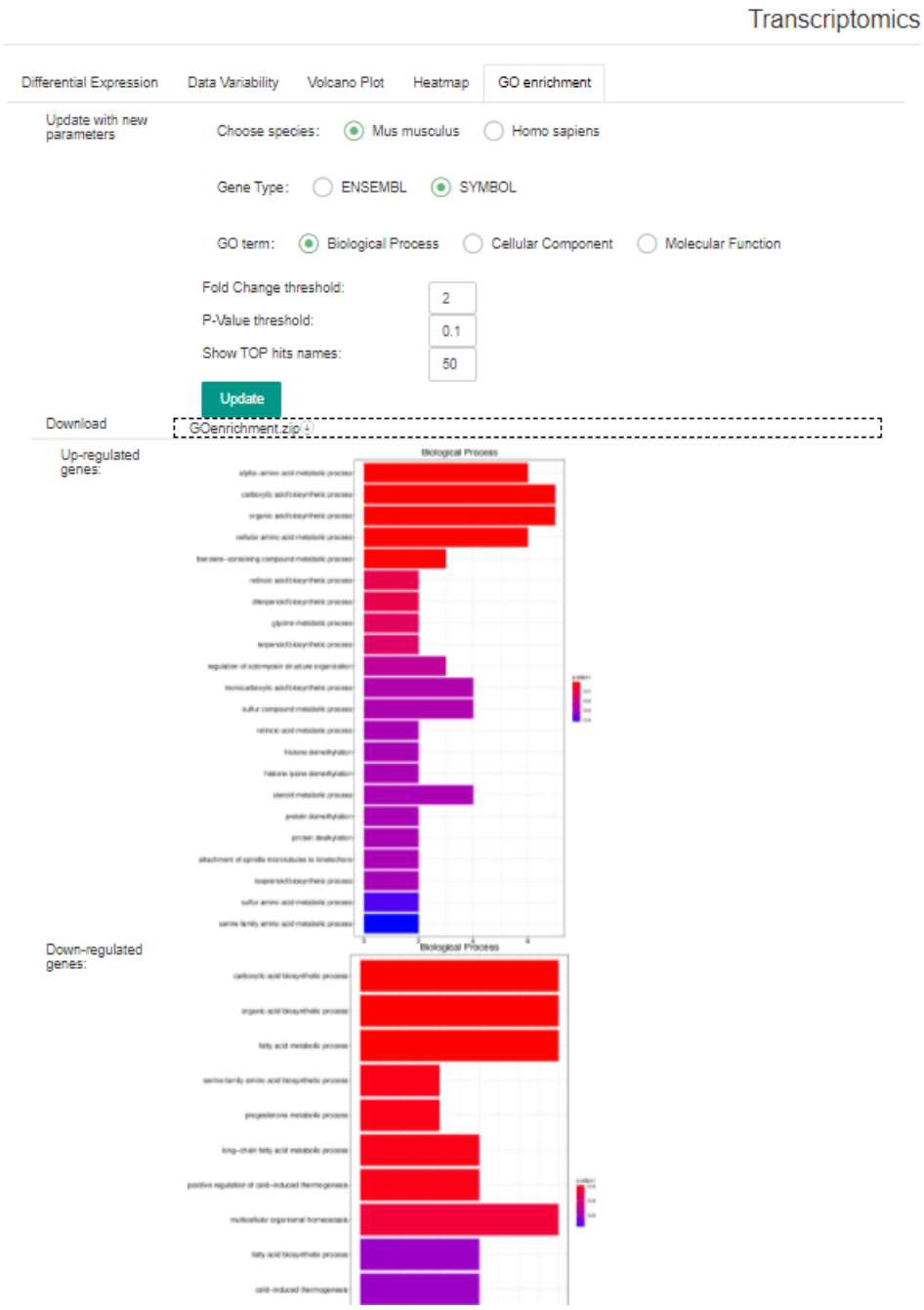


Figure 26. LINT-web can support *Mus musculus* and *Homo sapiens* gene analysis, more species are coming soon. The website supports two types of gene names "ENSEMBL" and "SYMBOL" to process the enrichment, and screen off the non-significantly changed genes using fold change or p-Value. All the plots including upregulated and downregulated gene expression enrichment results can be downloaded on the same web page as PDF format.

2.3 Metabolomic Analyses

LINT-web allows users to analyze only metabolomic dataset at "Metabolomics" tab.

2.3.1 Data Preparation and Upload

Similar to the lipidomic analyses (1.1.1 section), users can upload their metabolomic datasets in "*Data file*" and "*Description file*". Examples of metabolomic data can be downloaded on the same web page.

2.3.2 Data Processing

This section is similar to the lipidomic analyses (1.1.2 section).

2.3.3 Statistical Analysis and Ontology

Similar to the lipidomic analyses, the metabolomic statistical analysis and ontology also consist of a few modules (Figure 27).

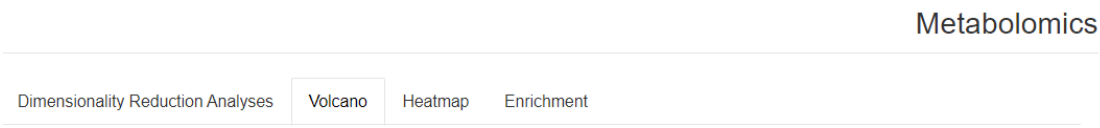


Figure 27. Results of statistical analyses

1) Dimensionality Reduction Analyses

Users can refer to the section 1.1.3.

2) Volcano Plot

Users can refer to the section 1.1.3. While metabolites are not classified as lipid classes, users no need to consider the metabolite classes (Figure 28).

Figure 28.Volcano plot results

3) Heatmap

Users can refer to the section 1.1.3.

4) Metabolite Enrichment Analysis

LINT-web provides users to enrich metabolite pathway using MSEA (doi: 10.1093/nar/gkq329) algorithms. Users can access the interface through "Enrichment" tab (Figure 29).

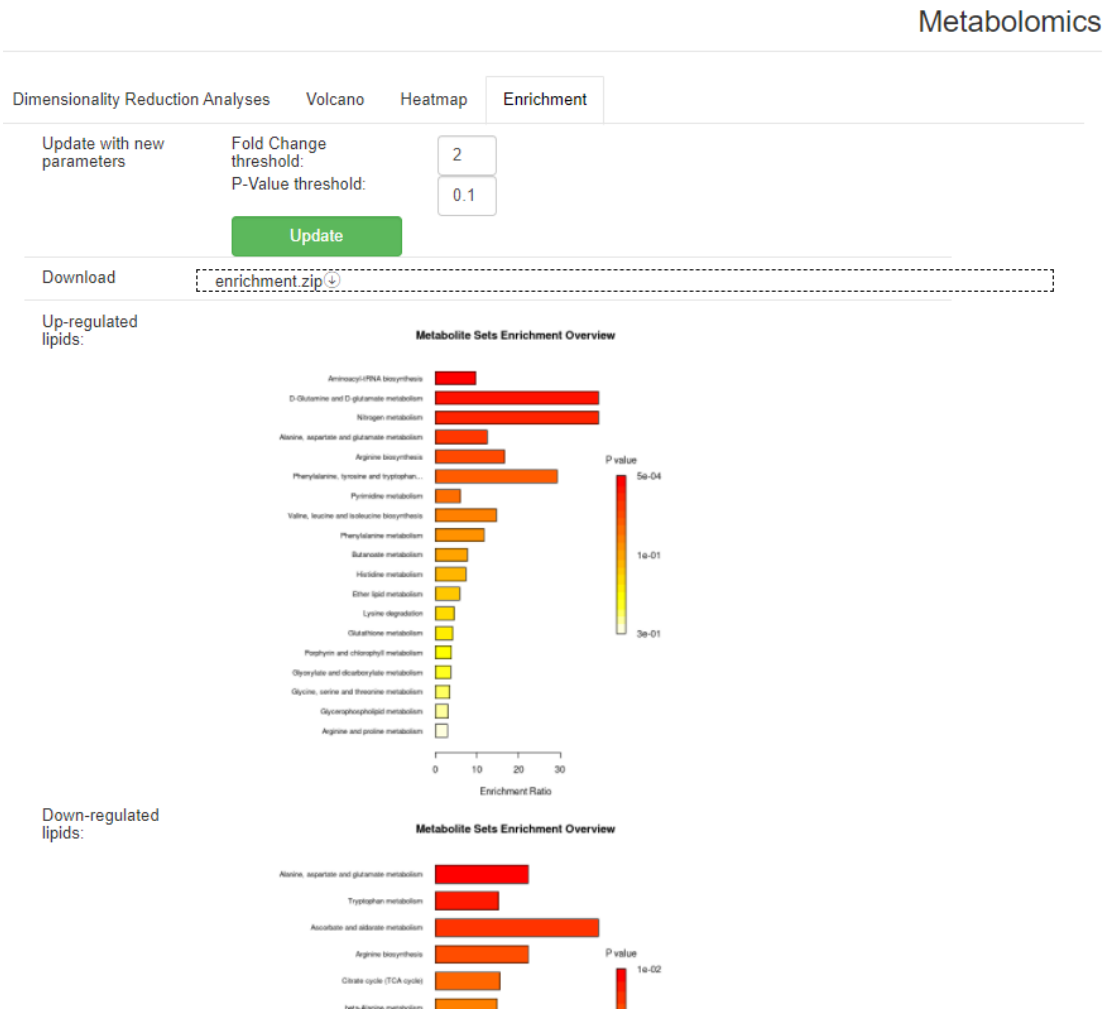


Figure 29. Metabolites with certain fold change and p-Value can be selected for the enrichment processing. All the plots including upregulated and downregulated metabolite enrichment results can be downloaded on the same web page as PDF format.

2.4 Proteomic Analyses

LINT-web allows users to analyze only proteomic dataset at "Proteomic" tab.

2.4.1 Data Preparation and Upload

Similar to the lipidomic analyses (1.1.1 section), users can upload their proteomic datasets in "*Data file*" and "*Description file*". Examples of proteomic data can be downloaded on the same web page.

2.4.2 Data Processing

This section is similar to the lipidomic analyses (1.1.2 section).

2.4.3 Statistical Analysis and Ontology

Similar to the lipidomic analyses, the proteomic statistical analysis and ontology also consist of a few modules (Figure 30).

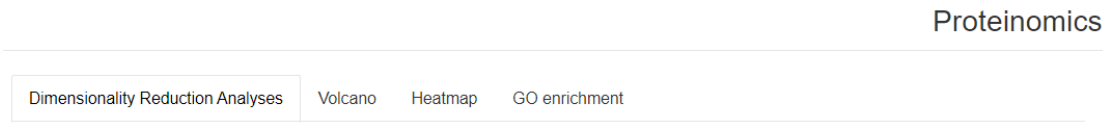


Figure 30. Results of statistical analyses

1) Dimensionality Reduction Analyses
Users can refer to the section 1.1.3.

2) Volcano Plot
Users can refer to the section 1.1.3. While proteins are not classified as lipid classes, users no need to consider their classes (Figure 31).

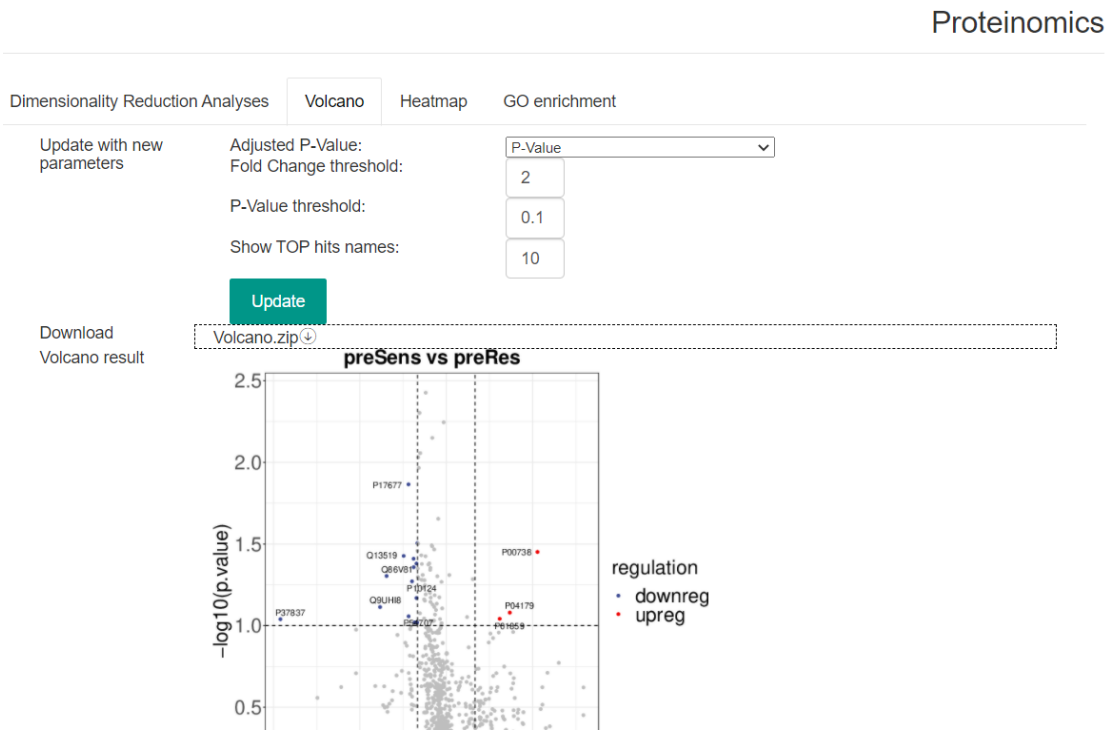


Figure 31. Volcano plot results

3) Heatmap

Users can refer to the section 1.1.3.

4) GO Enrichment analysis

Users can refer to the section 2.2.3 “GO Enrichment analysis” part.

3. Multiomics Integrative Analysis

The philosophy of using multiomic integration approach to data mining the lipidomic dataset and reveal lipid ontology is discussed in our manuscript. Such multiomic integrative analysis requires users to upload two omic datasets that come from a same sample, i.e., the intra-omic integrative analysis. Users can access the “Intra-omic Dataset Analysis” through the tab and upload two types of omic datasets (Figure 32). Examples of each omic datasets can be found on the same web page (“Try example data”). Users can also refer to the former sections to find out the “*data file*” and “*description file*” format requirements of each omic dataset.

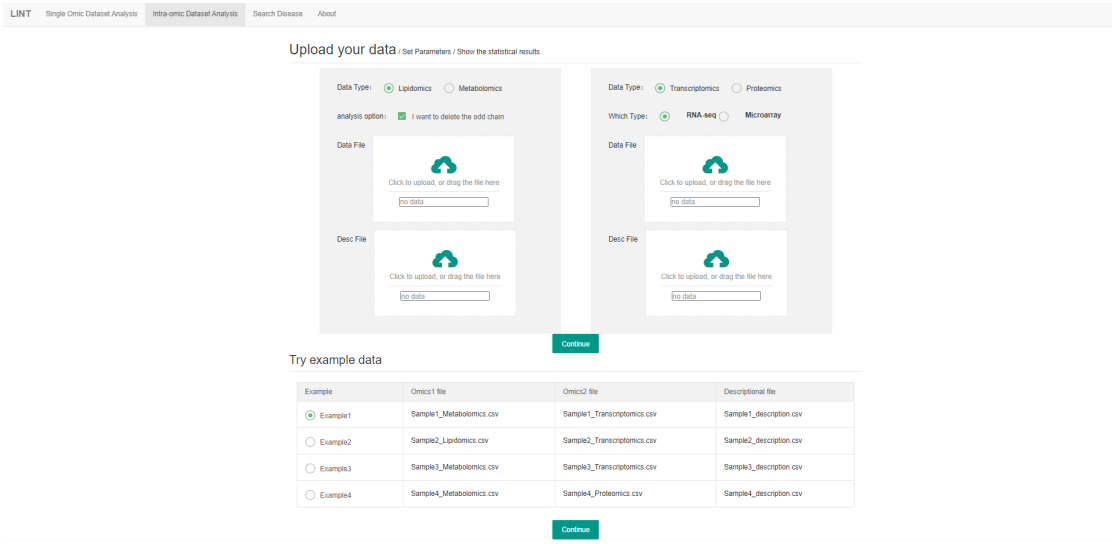


Figure 32. Multiomics integrative analysis uploading page

3.1 Data Preparation and Upload

Similar to the “Single Omic Dataset Analysis” (Section 1), the users can click “Continue” button once the “*data file*” and “*description file*” are well prepared as instructions.

1) Missing value refilling and normalization

LINT-web will refill the missing values of lipidomics/metabolomics/proteomics data and processing the imputation. Half value of the smallest number of the whole file will be refilled if the missing value numbers is less than a certain percentage (set by the users). Normalization options are

described in section 2.1.2.

2) Clustering Options

LINT-web provides four types of clustering algorithms to cluster the interacted feature sets. (Figure 33-35)

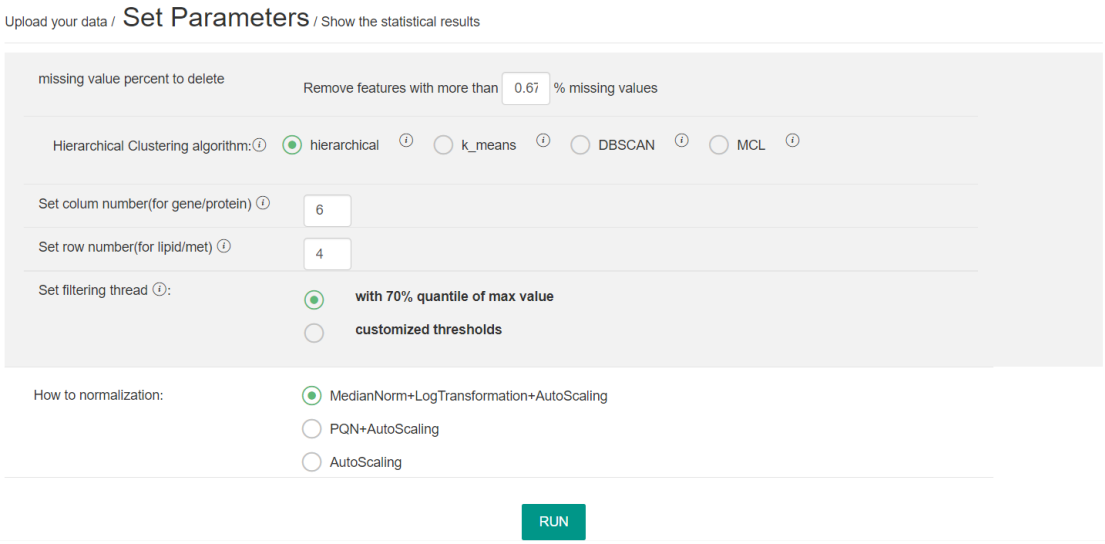


Figure 33. The basic "hierarchical" algorithm works for most situations; The "k-means" and "hierarchical" clustering algorithms can split the correlation matrix as users required (x column numbers and y row numbers) so that a certain x × y submatrices will be presented.

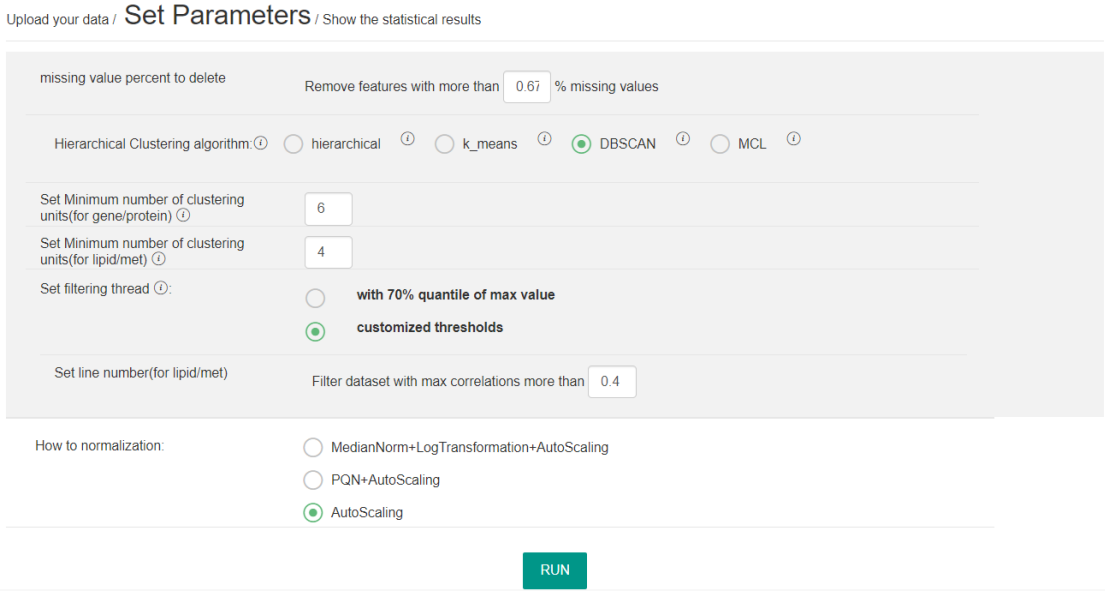


Figure 34. The "DBSCAN" algorithm only split the matrix based on the minimum correlated modules, so the users should enter the minimum number of columns and rows while using such clustering algorithm (Figure 31).

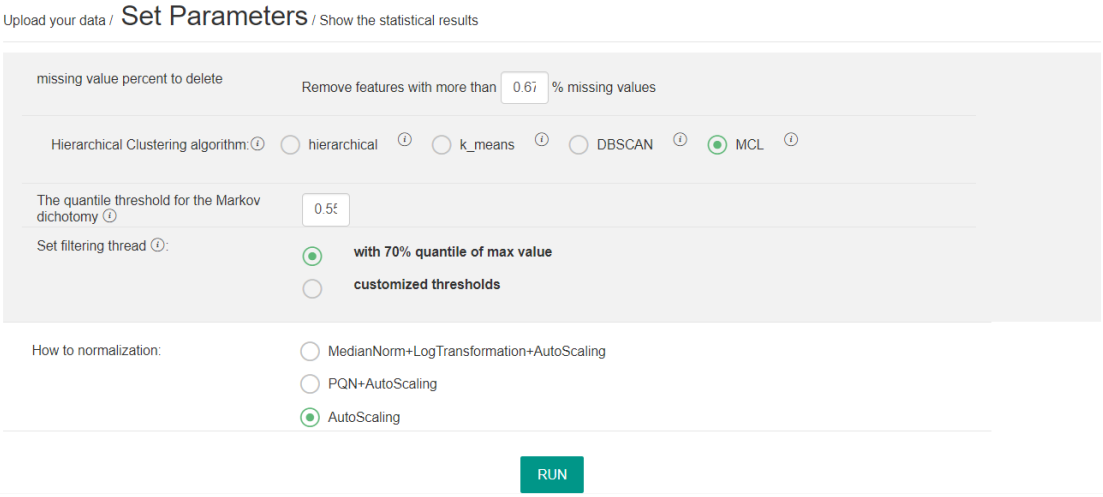


Figure 35. The Markov clustering algorithm ("MCL") can be selected if users wish to split the correlation matrix based on Markov process, and users should set the quantile threshold to process the Markov dichotomy (Figure 32).

LINT-web also provide a filtering threshold to screen off the weak correlated features, and 70% quantile of max value is suggested as default. Users can also change their thread if necessary (Figure 30-31).

3.2 Statistical Analysis and Ontology

1) Pattern Selection

The correlation matrix will be calculated for a while and users will view a split heatmap. The correlation matrix heatmap is split into x × y submatrices based on the parameters setting on the last web page (Figure 36).

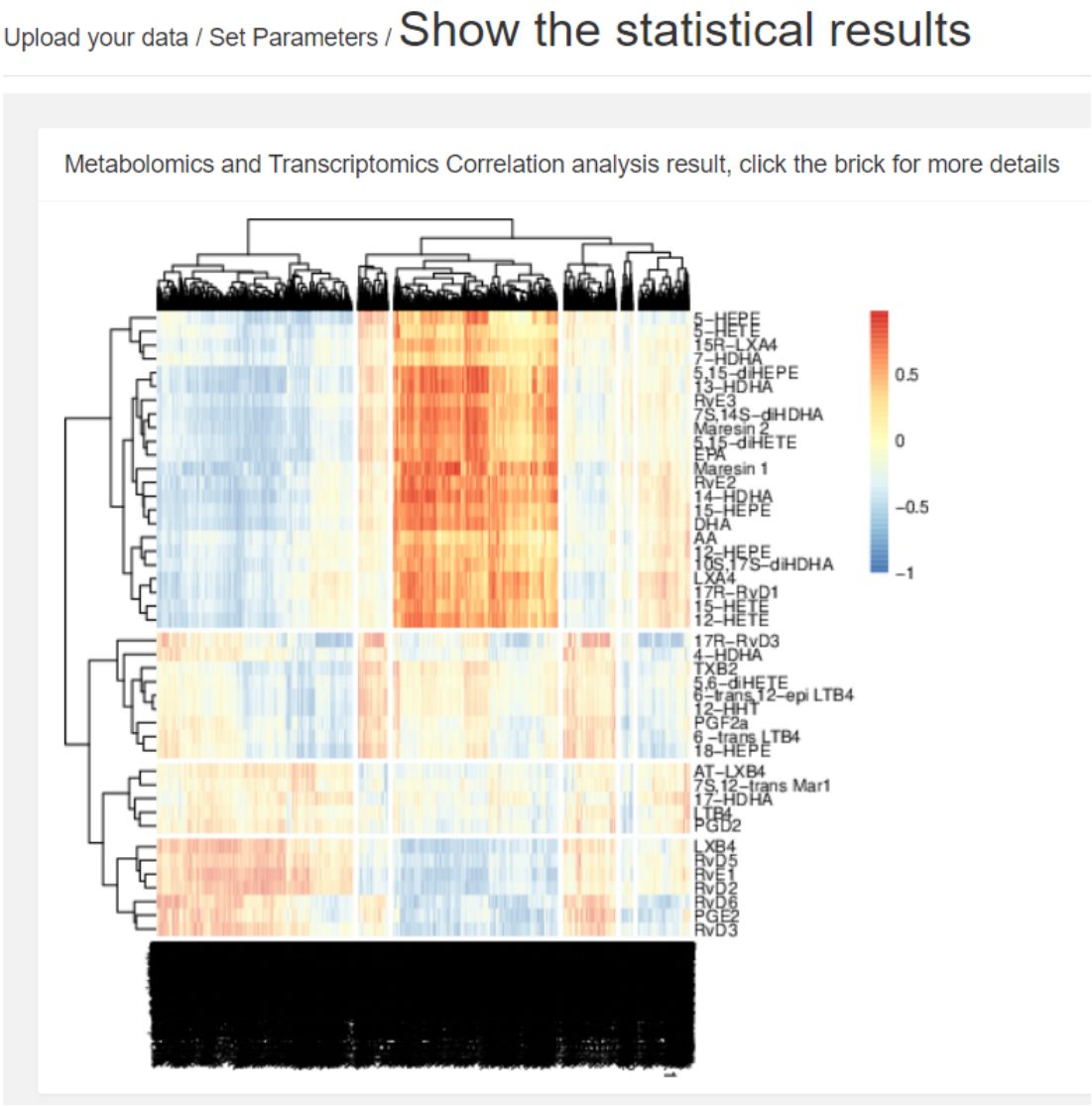


Figure 36. Users can click their interested submatrices to view the detail of strongly correlated lipids and gene (or other data features). The red color indicates a strong positive correlation and bule indicates a strong negative correlation. Users can determine the pattern blocks they interested to gather the information of lipids and genes for the further data mining analyses.

2) Correlated Data Mining

Once the user-interested patterns are selected, the selected data mining procedures will be shown on the next page (Figure 37). For instant, the gene set can perform “GO enrichment analysis” to find out the ontology of user-interested genes are associated to which kinds of biological processes, so that the strongly correlated lipid set may also associated to such biological processes. Lipids in a same pattern block may also share similar biological functions in the targeted sample.

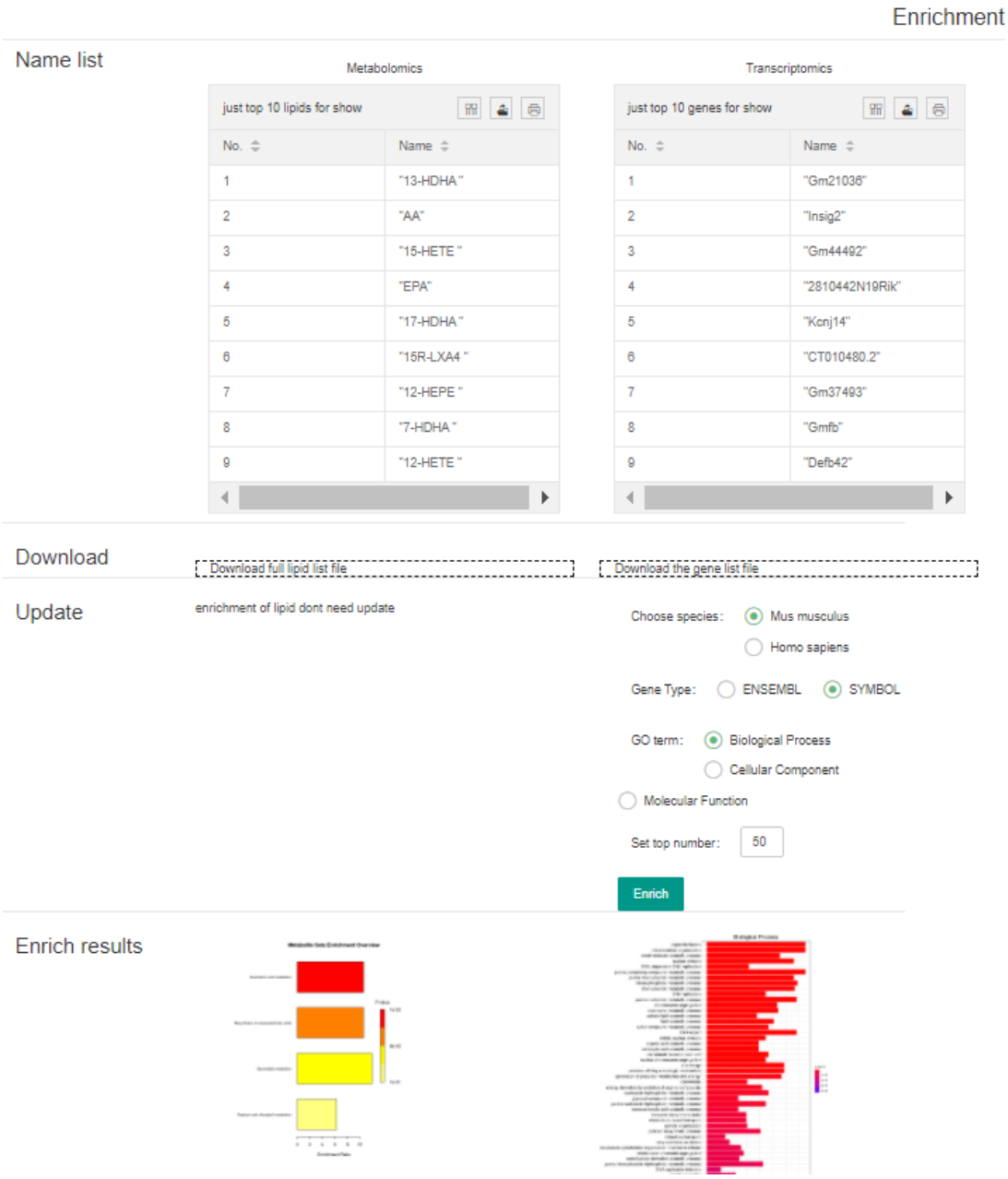


Figure 37. Users can download the selected dataset information. And similar to the “Single Omic Dataset Analysis” (Section 1), the enrichment analysis of each omic dataset can be performed according to the former instruction.

3) Ontology and correlation circos plot

Users can also view a circos correlation plot that present three levels of information of selected pattern block; lipid/metabolite set, gene set and GO terms.

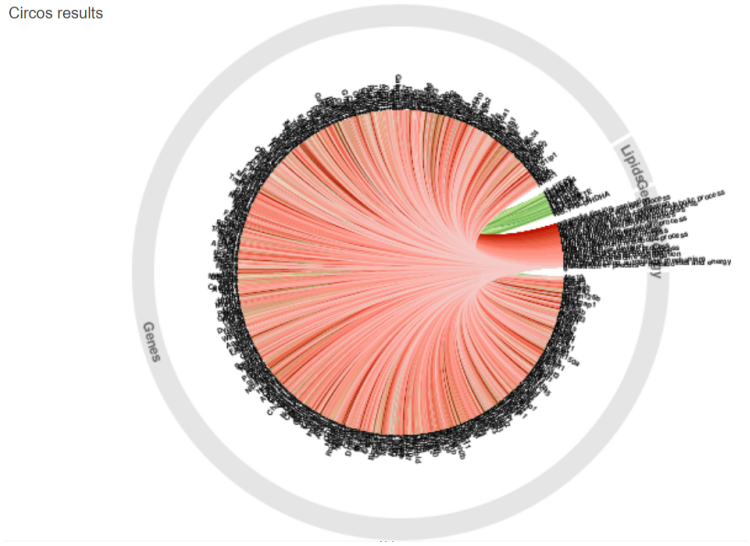


Figure 38. Circos plot of correlation. The red lines indicate the correlation of gene set with GO terms, and the green lines indicate the correlation of gene set with lipids/metabolites. All the plots can be downloaded on the same web page in a PDF format.

4. Searching for Gene Associated Diseases

LINT-web collects most gene associated diseases form KEGG database (https://www.kegg.jp/kegg/disease/) for a handy searching after the data processing and analyses. (Figure 39)

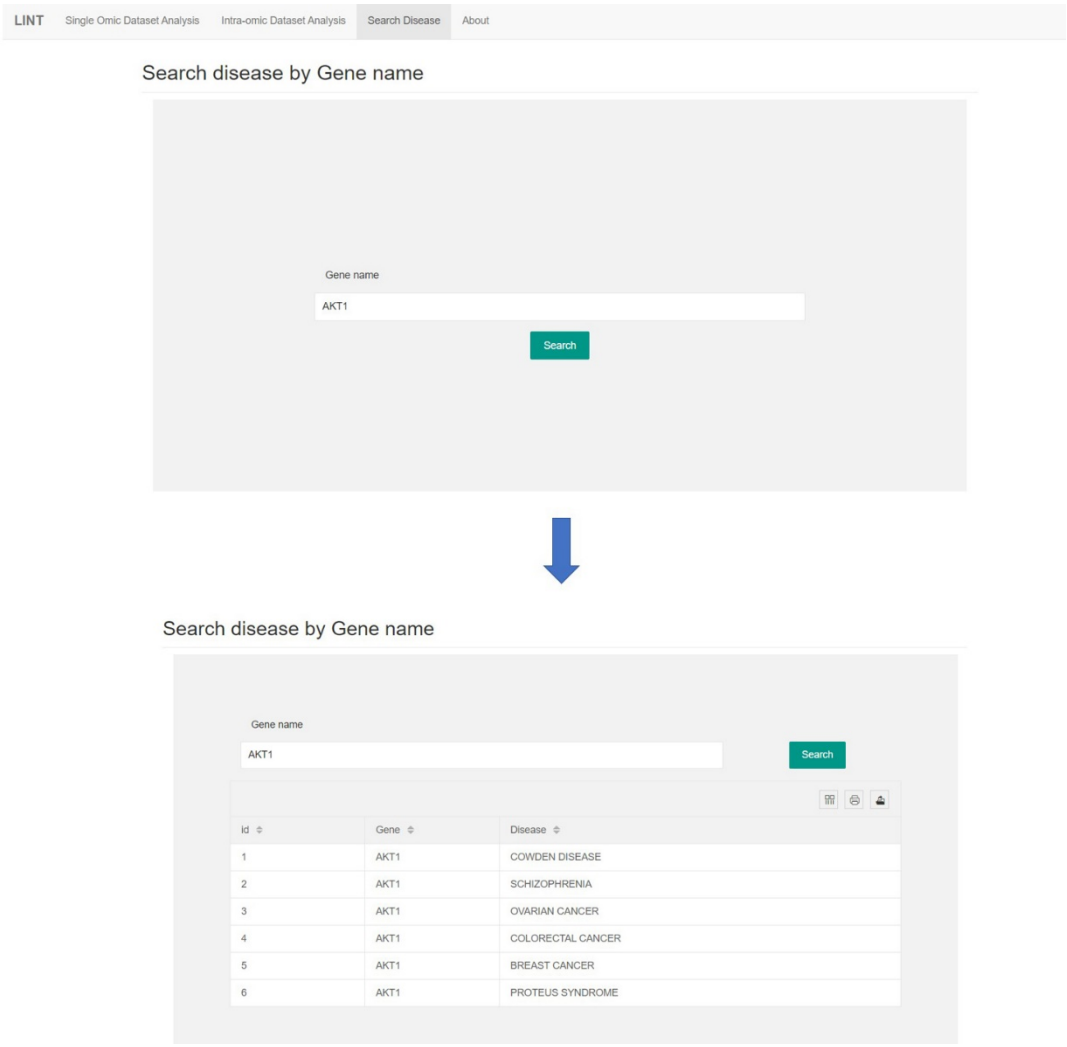


Figure 39. Users can type their interested gene names on "Search Disease" tab and click "Search" button to view the gene associated diseases.