# metabox: a toolbox for metabolomic data analysis, interpretation and integrative exploration

Kwanjeera Wanichthanarak, Sili Fan, Dmitry Grapov, Dinesh Barupal and Oliver Fiehn

9 January 2017

#### TUTORIALS

A collection of tutorials for using metabox is provided here to illustrate important features of metabox. These tutorials are corresponding to Results and Discussion section of the main manuscript.

#### 0. Example data sets

The example data sets are based on the published studies about lung adenocarcinoma (transcriptomic data (1) and metabolomic data (2)).

The metabolomic data set containing 39 malignant and adjacent non-malignant lung tissue samples was measured by gas chromatography time-of-flight mass spectrometry (GC-TOF-MS) (2) and pre-processed by the BinBase database (3). 462 compounds were measured, 171 of which have associated PubChem CID. This metabolomic data set was prepared in the input format used in metabox, which is in *metabolomics.xlsx*.

Metabox was used for the statistical analysis comparing tumor and control tissues. For users' convenience, we filtered the list of significantly different compounds comparing tumor and control tissues from the analysis result and provided in *pubchem\_stats.txt*.

The transcriptomic data, GEO accession: GSE32863 (1), contain gene expression profiles of 58 lung adenocarcinoma and adjacent non-malignant lung tissues using Illumina HumanWG-6 v3.0 expression BeadChip platform. Differential gene expression analysis comparing lung adenocarcinoma and adjacent non-malignant lung tissues was performed with GEO2R (4). P-values were adjusted with Benjamini and Hochberg false discovery rate at 5% (pFDR < 0.05). Differential gene expression analysis comparing tumor and control tissues reported 171 out of 21,204 genes in which pFDR < 0.05 and |log2FC| > 2. For users' convenience, we filtered the list of significantly different genes comparing tumor and control tissues from the analysis result and provided in *ensembl\_stats.txt*.

File	Description	Input of section
metabolomics.xlsx	Metabolic profiles of 39 lung malignant and adjacent	Section 2
	non-malignant tissues pre-processed by the BinBase	
	database.	
pubchem_stats.txt	Multi-column table of significant compounds and	Section 3
	associated statistical values computed by metabox.	
ensembl_stats.txt	Multi-column table of filtered genes and associated	Section 3
	statistical values computed by GEO2R.	

#### Table 1. List of example data sets

#### 1. Load required libraries and run the metabox using R-terminal

## Load library> library(metabox)> library(opencpu)

## Run metabox on a web browser > opencpu\$browse('library/metabox/www')

#### 2. Using metabox for in-depth analysis of metabolomic data

This tutorial shows that metabox is used for deep analysis (data processing, statistical analysis, metabolic network construction and functional interpretation) of metabolomic data. The example metabolomic data set is in *metabolomics.xlsx*. It is uploaded to metabox for log transformation before statistical analysis using paired t-test (Fig 1). The output from the data normalization and the statistical analysis is transferred to the network construction part to calculate the chemical structure similarity network of the PubChem compounds. The default threshold of correlation coefficient > 0.7 was used. The resulting network is enhanced by further mapping with annotation information. In this case, we applied Functional class scoring option to estimate significantly enriched pathways of network nodes by taking metabolic profiles in to account (see Fig 2 for all the steps).

Alternatively, the output from statistical analysis can be passed to one of the functional analysis method to aid interpretation of measured compounds. In this case, we applied WordCloud option to quickly summarize KEGG pathways of the compounds (see Fig 3 for all the steps). The top ten categories of measured compounds include Central carbon metabolism in cancer, Protein digestion and absorption, Aminoacyl-tRNA biosynthesis, Purine metabolism, Pyrimidine metabolism, Glyoxylate and dicarboxylate metabolism, Arginine and proline metabolism, Pentose and glucuronate interconversions, Galactose metabolism and Fatty acid biosynthesis sorted by the number of compounds in each category. This example shows that after statistical analyses, the result can be interpreted with pathway information.

### **3.** Integrative exploration of significant genes and compounds in biological network context

Metabox is used for joint exploration of the lists of significant genes and compounds from comparisons between tumor and non-tumor tissues in the context of biological networks. The *ensembl\_stats.txt* and *pubchem\_stats.txt* are used to construct the biological network outlining relationships between genes and compounds as *(:Protein)-[:CONTROL]->(from:Gene)-[:CONVERSION]->(:Protein)-[:CATALYSIS]->(to:Compound)*. The resulting network including all attributes can be downloaded for advance exploration using visualization software such as Cytoscape (5). Fig 4 shows all the steps to construct the network with metabox.

### FIGURES

					U	Jpload Da
Upload Data Example Data	MetaData	Feature	xpression			
eet Name:	download				Search:	
e first sheet will be read as default. Click me to ange.	feature_index	It Kno	wnorUnknown	ret_index 10	quant_mz 0	PubChem
Upload A File metabolomics.xlsx	1	TRUE		566749	217	6912
ou want to upload a new file, please refresh the page first.	2	TRUE		924754	325	64959
Upload an Excel File.	3	TRUE		702391	353	1188
need, enter a sheet name in the		TRUE		309905	144	6287
herwise the first sheet will be us		IIt. TRUE	I.	979035	352	6030
	6	TRUE		856953	258	6029
	7	TRUE		730534 331223	441	1175
	9	TRUE		385903	99	1176
	10	TRUE		586393	217	17473
	11	TRUE		623732	226	445675
	12	TRUE		670802	218	6057
	13	TRUE		779834	202	6305
	Showing 1 to 462 of	of 462 entries				
Principal C 1. Sample 2. Select data proc	omponenet Analysis cessina met	thods ar	nd Click Apply		Information of 5	Selected Samples
Normalization Use PCA score pla						
None     Sample-specific				N		
mTIC	15		/:	т		-
Loess Normalization + Batch     Median Correction	10		1			
Batch Median Correction		/	(			
2. Data Transformation	5					
None     None		/•/•				
log     power transformation	0			-		
U D						
3. Feature · · · · · · · · · · · · · · · · · · ·	° (/•·	. /				
None     Auto scaling	0301		elect to select		Remove Folio	wing Samples: A
Pareto scaling			ested e.g. out ot will display	liers		
Range scaling		mation c			51	
Apply						
		nationic	n alem.			
Lowmood Normalized Data				G' Settings	Condition	
Normalized data can be dow	nloaded he		Use dropo	lown to cho	oose inform	ation to disp
Normalized data can be dow	Result Table	re	Use dropo Fil	lown to cho	oose inform id(s) that w	
Normalized data can be dow Study Design Experiment Factor:	Result Table To use other meth	re	Use dropo Fil	lown to cho	oose inform id(s) that w	ation to disp vill be remov
Normalized data can be down	Result Table To use other meth Parametric	l <b>'O</b> rods, Cilck met	Use dropc Fil. nod name	lown to cho I in sample	nose inform id(s) that w	ation to disp vill be remov and Click Ap and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected • Repeat Measure Factor: Condition •	Result Table To use other meth Parametric Use paired 1 jest on nonpara-Parar	rods, Click met factor Condition	Use dropc Fill hod name on, with a Benjamini & Ho To choose	lown to cho I in sample chberg (1995) adjus e other met	noise inform id(s) that w ment, to compare N thod, Click	ation to disp vill be remov and Click Ap and T. method nam
Normalized data can be dow Stury Design Nothing selected • Repeat Measure Factor: Constion • Samele Sam <sup>3</sup> . Select factor to be com	Result Table To use other meth Parametric Use paired 1 test on nonpara-Parar Parametric Signe	rods, Click met factor Condition	Use dropc Fill hod name on, with a Benjamini & Ho To choose	lown to cho I in sample chberg (1995) adjus e other met	noose inform id(s) that w ment, to compare M thod, Click	ation to disp vill be remov and Click Ap and T. method nam
Normalized data can be dow Stury Design Deperment Factor Nothing selected Sample Size 3. Select factor to be com Statistical methods	Result Table To use other meth Parametric Use pared () set on nonpara-Parar Par/Paran download	rods, Click met factor Condition	Use dropc Fill hod name on, with a Benjamini & Ho To choose	lown to cho I in sample chberg (1995) adjus e other met	noise inform id(s) that w ment, to compare N thod, Click	ation to disp vill be remov and Click Ap and T. method nam
Normalized data can be dow Stury Design Nothing selected • Repeat Messure Factor: Condition • Sample Stag 3. Select factor to be com	Result Table To use other meth Parametric Use pared () set on nonpara-Parar Par/Paran download	reds, Cilck met factor Conditi metric d-Rank Test or	Use dropc Fill hod name on, with a Benjamini & Ho To choose	lown to cho I in sample chberg (1995) adjus e other met	noose inform id(s) that w ment, to compare M thod, Click	ation to disp vill be remov and Click Ap and T. method nam
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample Sec. 3. Select factor to be com Statistical methods * uill automatically be chos	Result Table To use other meth Parametric Use paired I test on nonpara-Parar Par/ed. en. PubChem	re tods, Click med factor Conditis metric ed-Rank Test of outlier	Use dropc Fil. hod name To choose In factor Condition, with a p_value_N_vs_T	fown to cho i in sample chberg (1995) adjus e other mei Benjamirt & Hochbr	noise inform id(s) that w ment, to compare M thod, Click	ation to disp vill be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample Study 3. Select factor to be com Statistical methods 1 will automatically be chos Click Submit! to compute	Result Table Provide the methic Parametric Use patred 1 jest on nonpara-Parar Parametric download en. PubChem	re toda, Click med factor Conditis metric ad-Rank Test or outlier exist?	Use dropp Fil. bostnere To choose in factor Condition, with a p_vealue_N_ve_T	fown to cho i in sample chberg (1995) adjus e other mei Benjamini & Hochber tdr_scijusted_p 0.0016	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp vill be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample Soc. 3. Select factor to be com Statistical methods s tatistical methods s tatistical methods s tatistical with automatically be chos Click Submit! to compute	Result Table     Result Table     Offo use other meth     Parametric     Use paired 1 jest on     nonpara-Param     Parametric     download     One     O	re tactor Condition factor Condition metric outlier fatse true	Use dropc Fill holiners on, with a Bergemont A Ho To choose p_velae_N_vel_1	lown to che I in sample cheers (1995) adjus e other mei Benamin & Hochber fdr_adjusted_s 0.0016 0.8266	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp vill be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample Soc, 3. Select factor to be com Statistical methods 19 39 Click Submit! to compute Excluded Semples Normatization Applied Sample Romalizator: Nore	Result Table 075 use other metry Parametric Use paired 1 Set on nonpara-Parar Parametric downtoad En. PubChem	re tactor Condition factor Condition metric outlier tatse true true	Use dropcd Fill bod.neme To choose n factor Condition, with a 0,0001 0,7192 0	lown to che l in sample cherg (1995) aquis e other mei Benerin & Hoche Benerin & Hoche 0.0016 0.0005	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp vill be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Stury Design Experiment Factor: Nothing selected Sample Size 3. Select factor to be com Statistical methods * 1 will automatically be chos Click Submit! to compute Excluded Samples Normalization Applied	Result Table     Result Table     Offo use other meth     Parametric     Use paired 1 jest on     nonpara-Param     Parametric     download     One     O	re tactor Condition factor Condition metric outlier fatse true	Use dropc Fill holiners on, with a Bergemont A Ho To choose p_velae_N_vel_1	lown to che I in sample cheers (1995) adjus e other mei Benamin & Hochber fdr_adjusted_s 0.0016 0.8266	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp vill be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample Soc, 3. Select factor to be com Statistical methods 19 39 Click Submit! to compute Excluded Semples Normatization Applied Sample Romalizator: Nore	nloaded he:           Result Table           9 To use other meth           Data and the second se	re ods, Click med factor Conditis metric outlier exist? false true talse true false true	Use dropc Fil. bol.name To choose 1 fador Cenetition, with a 9, value, N, ve. 7 0 0,0001 0,7192 0 0,0564 0,4803	lown to chin l in sample cherg (1999) adus e other met Bersenri & Hoote 0.0016 0.0006 0.0006 0.0006 0.0004 0.0047	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp ill be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample Soc, 3. Select factor to be com Statistical methods 19 39 Click Submit! to compute Excluded Semples Normatization Applied Sample Romalizator: Nore	Pesult Table           Pesult Table           O'D use other meth           Parametric           Use parted 1 pick of the pick of the pick of pick of the pick of th	re noda, Click med factor Conditis metric d.Rank Test or outlier exist? false true false true true true true	Use dropc Fill holiners on, with a Bergemont A Ho To choose p_vetue_N_vet_1 0.0001 0.7192 0 0.0004 0.0004 0.0004 0.0004 0.0004 0.0004	lown to chel I in sample cherg (1996) adjus o Other me: Bessenet & Hochel 0.0256 0.0266 0.0266 0.0565 0.0545 0.0532	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow Study Design Experiment Factor: Nothing selected Sample See, 3. Select factor to be com Statistical methods Statistical methods T will automatically be chos Click Submit! to compute Excluded Samples Normalization: None Data Transformatication: None Padare Normalization: None	nloaded he:           Result Table           9 To use other meth           Data and the second se	reads, Circk met factor Conditis metric exist? false twe false twe false twe twe false twe twe false	Use dropcd Fill bod.netre To choose In factor Condition, with a pvalueN_veT 0 0.0001 0.7162 0 0.0564	lown to chel in sample cherg (1999) adjur e other men browner & hoster for_adjusted_1 0.055 0.065 0.065 0.055 0.052 0.055 0.055 0.055	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and T. non, parametric 0.0002 (0.0016) 0.4155 0.0001 0.0015
Normalized data can be dow           Stury Design           Depriment Factor:           Nothing selected           Sample Bog.3. Select factor to be com           Statistical methods           ************************************	Provide description           Result Table           OTo use other method           Data and the second	re noda, Click med factor Conditis metric d.Rank Test or outlier exist? false true false true true true true	Use dropc Fill holiners on, with a Bergemont A Ho To choose p_vetue_N_vet_1 0.0001 0.7192 0 0.0004 0.0004 0.0004 0.0004 0.0004 0.0004	lown to chel I in sample cherg (1996) adjus o Other me: Bessenet & Hochel 0.0256 0.0266 0.0266 0.0565 0.0545 0.0532	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap and T. method nam to compare N and T.
Normalized data can be dow           Stury Design           Depriment Factor:           Nothing selected           Sample Bog.3. Select factor to be com           Statistical methods           ************************************	Provided Her           Result Table           O'D use other methy           O'D use other methy           Parametric           Use patrent Jiset on           Parametric           Bary Gell           Bary Gell           Other State           Other State <tr< td=""><td>reads, Click met factor Condition metric ad-Rank Test of exist? false true true true true true true true tru</td><td>Use dropc Fill hod name To choose factor Condition, with a p_value_N_ve_T_1 0.0001 0.0564 0.4693 0.00056 0.5425</td><td>lown to chef l in sample cherg (1990) edus e other met Benannt &amp; Hoche 0.0016 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0002 0.0002 0.0002 0.0002</td><td>bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:</td><td>ation to disp illi be remov and Click Ap and T. method nam .so compare N and T. .so compare N</td></tr<>	reads, Click met factor Condition metric ad-Rank Test of exist? false true true true true true true true tru	Use dropc Fill hod name To choose factor Condition, with a p_value_N_ve_T_1 0.0001 0.0564 0.4693 0.00056 0.5425	lown to chef l in sample cherg (1990) edus e other met Benannt & Hoche 0.0016 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0006 0.0002 0.0002 0.0002 0.0002	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap and T. method nam .so compare N and T. .so compare N
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	Persuit Table           Persuit Table           O'D use other meets           Parametric           Use patrent 1/set on           Portage           Orbus of the meets           Orbus of the m	re coda, Click med factor Conditis metric outlier 0 exist? false true false true false true false true false true false true	Use dropc Fil. bol.rame To choose 1 fador Condition, with a 9_vvalue,N_vve_T 0_00001 0_7192 0_00004 0_00004 0_00004 0_00005 0_00005 0_00005 0_00005 0_00005 0_0005 0_0005 0_0005 0_0005 0_0005 0_0005 0_0005 0_0005 0_0005 0_0005 0_00000000	Cover to chef in sample cherg (1999) adur o other mei Berennet A Hoche 0.0006 0.0006 0.0007 0.0647 0.0645 0.0002 0.0022 0.0252 0.0252	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap and Click Ap method nam 
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	Provided Her           Result Table           O'D use other meth           Parametric           Use parter (1) factor           D'D use other meth           Bue parter (1) factor           Barger           G012           G012           G012           G030           G029           1176           1174           17473	Control of the second of the s	Use dropc Fill bodyneme on, with a Bergemont A Ho To choose p_vetae_N_veta 1 0.0001 0.7192 0 0.0004 0.0003 0.00050	own to chef         in sample           cherg (1999) adus         o           cherg (1999) adus         o           cherg (1999) adus         o           fdadjusted_1         o           0.0016         0.856           0.0026         0.9647           0.6459         0.0022           0.0021         0.0022           0.0032         0.0032           0.0396         0.0112	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap and T. method nam method nam (non_peramétric 0.0002 0.6046 0 0.8739 0.4185 0.0001 0.9833 0.0002 0.0001
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	nloaded her Result Table Of buse other netty Parametric Use patent Jest on nonpaga-Parar Bearded	re- tactor Conditions into a conditions of fairs for a conditions of fairs for a conditions to conditions to a	Use dropc Fill bod.rears To choose To choose the condition, with a pvalueN_var 0.0001 0.7192 0 0.0001 0.7192 0 0.0004 0.0001 0.0001 0.0001 0.0005 0.0005 0.0005 0.0005 0.0005 0.0002 0.0002 0.0002	Image: Control of the second of the	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap method nam . to compare N and 7. . to compare N and 7. . to compare N and 7. . compare
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	Provided Her           Result Table           O'D use other methy           O'D use other methy           Parametric           Use patrent Justice           Darametric           Othor and the state of th	re- metric stater constitution factor constitution factor	Use dropc Fil. http://www.initianality.com/ To choose factor Condition, with a p_value_N_ve_T 0.0001 0.7102 0.0004 0.0004 0.0004 0.0005 0.0000	town to chelin sample           in sample           chera (1990) adus           other me:           boons           tdr_stiputed_j           0.055	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap and T. method nam 
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	Persuit Table           Persuit Table           O'D use other methy           Parametric           Use patrent 1/set on           Parametric           Othouse days           O	re- tator conditioned tator conditioned extension of the second tator conditioned tator conditioned tat	Use dropc File bod rearse on, with a Bengannet A Ho <i>To choose</i> n factor Condition, with a p-yvalue, N_yva ( 0,0001 0,7192 0 0,000500000000	Town to chef           in sample           cherg (1999) adus           o ther mei           booms           fdr_adjusted_1           0.0016           0.869           0.0022           0.9647           0.8458           0.0022           0.0252           0.0386           0.0112           0.0366           0.0252           0.0366           0.0366	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap method nam non_peramitic 00002 0.6046 0.6739 0.4185 0.0001 0.4185 0.0001 0.6353 0.0002 0.0019 0.00000000
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	Provided Her           Result Table           O'D use other methy           O'D use other methy           Parametric           Use parted type of development           COL           PubChem           6012           6052           6053           6053           6053           6053           6053           6053           6053           6053           6054           1175           1177           1174           1174           1074           6055           7018	re- metric tacks Condition tacks Condition tacks Total tacks tack	Use dropc Fill box reads not with a Bergement A Ho To choose n factor Condition, with a p_vatue_N_ve_1 [] 0.0001 0.7192 0 0.0004 0.0003 0.0005 0.0000	own to chef         in sample           cherg (1999) adus         o chef           o chef (1999) adus         o chef           o cons         o cons	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and T. remothed name to compare N and T. 
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	nioaded her. Result Table Of Dirac ether rest Parametric Use patrent Jisel on nonpaga-Parar References of the second PubChem    0912 0	re- metric tracer Condition tracer Condition tracer Condition tracer Condition tracer Test on tracer tracer Condition tracer tracer Condition tracer tracer Condition tracer tracer Condition tracer tracer tracer Condition tracer traco	Use dropc Fill bod.nears no., with a Bergenetic A Hoto To choose n factor Condition, with a pvalueN_var_ 0.0001 0.7192 0 0.0001 0.7192 0 0.0001 0.7192 0 0.0001 0.7192 0 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001 0.0003 0.0004	Cover to chef in sample others (1999) adus other me: beaure A hosted 0.0056 0.0056 0.0056 0.0052 0.0052 0.0051 0.0052 0.0051 0.0051 0.0052 0.0051 0.0051 0.0052 0.0051 0.0051 0.0052	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap method nam . b compare N and 7. . b compare N and 7. . b compare N and 7. . compare N
Normalized data can be dow           Strip Design           Experiment Factor:           Northing selected           Stript Sezio:           Scale Catalon           Sc	Nioaded her           Result Table           Ofbuilde other methy           Ofbuilde other methy           Parametric           Use patrent Juer           Bard Address           Bard Address           Bard Address           Bard Address           Object Juer on           Bard Address	And and a second	Use dropc Fill bostness on, with a Benjamori A Ho To choose in factor Condition, with a p_value, N_v va. ( p_value, N_v va. ( p	Torm to chef           in sample           cherg (1999) adus           other mei           cherg (1999) adus           other mei           cherg (1999) adus           other mei           cherg (1999) adus           oberg (1	bose inform id(s) that w	ation to disp illi be remov and f. method nam to compare N and f. 0.0002 0 0.0016 0.0011 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0001 0.0015 0.0005 0.0005 0.0015 0.0005 0.00
Normalized data can be dow           Study Design           Experiment Factor:           Nothing selected           Study Design           Study Design           Study Design           Study Design           Study Experiment Factor:           Study Experiment Factor:           Statistical methods           Statistical methods	Persuit Table           Persuit Table           O'D use other methy           D'D use other methy           Parametric           Use patront Jiest on           Pondations Support           0012           04090           0012           04090           01118           0207           00000 <td>re- tractor Conditioned tractor Conditioned trac</td> <td>Use dropc Fil. bod same To choose factor Condition, with a p_value_N_ve_T 0.0001 0.7162 0.0001 0.7162 0.0003 0.0005 0.5425 0.0003 0.0005 0.5425 0.0003 0.0005</td> <td>town to chef in sample others (1990) adus other me: bearing a thore 0.055 0.05</td> <td>bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:</td> <td>ation to disp illi be remov and Click Ap method nam to compare Nam 7. (b compare Nam 7. (b compare Nam 7. (b compare Nam 7. (c compare Nam</td>	re- tractor Conditioned tractor Conditioned trac	Use dropc Fil. bod same To choose factor Condition, with a p_value_N_ve_T 0.0001 0.7162 0.0001 0.7162 0.0003 0.0005 0.5425 0.0003 0.0005 0.5425 0.0003 0.0005	town to chef in sample others (1990) adus other me: bearing a thore 0.055 0.05	bose inform id(s) that w ment, to compare N thod, Click rg (1995) adjustment Search:	ation to disp illi be remov and Click Ap method nam to compare Nam 7. (b compare Nam 7. (b compare Nam 7. (b compare Nam 7. (c compare Nam

Fig 1. Data normalization and statistical analysis comparing twopaired groups of metabolomic study.

Compute chemical-structure similarity network

Fig 2. Computing chemical similarity network and performing FnClassScoring on network nodes.

npu	-								
	ts								
				similarity network.					
	out summary				analysis is show in il				out overvi
101	v 10 * er	ntries	PubCr	iem column v	vill be used to compu	ite a network	Search:		
þ	quant_m	z 🍦 PubChem	outlier     exist?	p_value_N_vs_T 0	fdr_adjusted_p_value_N_vs_T 0	_non_parametric_p_	value_N_vs_T 0	fdr_adjusted_no	n_param
	217	6912	false	0.0001	0.0016	0.0002		0.0024	
	325	64959	true	0.7192	0.8266	0.6046		0.7125	
	353	1188	true	0	0.0006	0		0.0007	
	144	6287	false	0.9584	0.9647	0.8739		0.9218	
	352	6030	true	0.4893	0.6458	0.4185		0.5687	
	258	6029	true	0.0003	0.0032	0.0001		0.0012	
	441	1175	true	0.0058	0.0252	0.0015		0.0097	
	99	1176	false	0.5425	0.6811	0.9833		0.9943	
	217	11/4	false	0.0103	0.0396	0.0082		0.0329	
		11415	141.040	0.002		0.0010		0.0110	
		of 462 entries	relation	lant		Previous 1	2 3 4	5 41	/ Nex
0.7		oto similarity cor	relation coeffic	aent:					
<i>J.</i> (									
20	mpute 2	2. Use def	ault thre	shold 0.7 and	d Click Compute				
	10 🔻 entri						Searc		
^	gid 🗄	nodename 🍦	nodelabel 👙	nodexref  feature	e_index 🕴 KnownorUnknown 🌵	ret_index	mz exist?	p_value_N_vs	_T 0 1
	222656	L-Malic acid	Compound		C00149  CHEBI:30797  FDB001044		69	TRUE	
	790	Hypoxanthine	Compound	790  HMDB00157  C0	0262  CHEBI:17368  DB04076  FDB00	3949  HYPOXANTHINE	87	TRUE	
)	229 290	229	Compound				151	TRUE	
	6057	L-Tyrosine	Compound	6057IIUMDB00158IIC	00082  CHEBI:17895  DB00135  FDB0		128	TRUE	
3	6140	L.	Compound		00079  CHEBI:17295  DB00120  FDB0		46	TRUE	
,	0140	Phenylalanine	Compound				40	TRUE	
\$	5950	L-Alanine	Compound	5950  HMDB00161  C ALANINE	00041  CHEBI:16977  DB00160  FDB0	00556  L-ALPHA-	155	TRUE	
ŀ	145742	L-Proline	Compound	145742  HMDB00162	C00148  CHEBI:17203  DB00172  FD	B000570  PRO	40	TRUE	
	- and -								
etv	vork								
				To pan – To zoom	panel allows interact click, hold and drag – scroll the mouse w t multiple nodes – pre	background vheel		, in the second s	
				$\Phi$	Z				
	ork legend		<u>t</u>	• <i>t</i>					

un eti -	nel	loos	oring								
		lass so	0								
Estimate	e enriche	d functional	classes for the inpu	ıt.							
nputs											
ntity lists	iea P	ubChem or	uninrat or ensemt	bl) will be used for	r enrichment an	alusis					
Input su		Input da					show i	in interactive	a tahlas	Inpu	ut overvi
how 10								ies will be u		arch:	
		outlier		npute enric							
PubCh	em 🗄	exist?	p_value_N_vs_1	ſ ∲ fdr_adjuste	ed_p_value_N_v	vs_T∲n	non_paran	metric_p_value_N_vs	_T 🕴 fdr_adjuste	d_non_parametric_p_valu	ue_N_vs
6912	t	false	0.0001	0.0016		0.00	002		0.0024		
64959	t	true	0.7192	0.8266		0.60	046		0.7125		
1188	t	true	0	0.0006		0			0.0007		
6287	1		0.9584	0.9647		0.87			0.9218		
6030			0.4893	0.6458		0.41			0.5687		
6029			0.0003	0.0032		0.00			0.0012		
1175			0.0058	0.0252		0.00			0.0097		
1176			0.5425	0.6811		0.98			0.9943		
11/4			0.0103	0.0396		0.00			0.0329		
1/4/3				0.0112		0.00	- 10		0.0110		
howing 1	to 10 of	462 entries						Previou	s 1 2 3	4 5 47	Ne
ntity type	e:		s	elect method: 0						Mesh	
compoun	d			median			*	Mesh annotation	n is available for Com	pound only.	
elect ent	tity-level	statistics:	•								
p_value											
	e_N_VS_	Т					*				
					0.4.4.6	<b>-</b>					
Comput	te 4	4. Choo			Select F	Pathway		l Select colu	mn p_value	e_N_vs_T	
Comput	te 4	4. Choo	ose metho ck Compu		Select F	Pathway		l Select colu	mn p_value	e_N_vs_T	
	te 4	4. Choo					y and			e_N_vs_T	
	te 4	4. Choo			Table annota	of analj ation-er	y and lysis r ntity p	esult, node bairs are sho	attributes, e		
twork	to 4	4. Choo and Cli		te	Table annota Top te	of anal ation-er	y and lysis r ntity p s are	result, node pairs are sho colored.	attributes, e own.	edge lists and	Sun
twork	Enric	4. Choo and Clin	ck Compu		Table annota Top te Colors	of analg ation-er en terms s in the	y and lysis r ntity p s are table	esult, node bairs are sho colored. and nodes	attributes, e own. are the san	edge lists and	Sur
work	Enric	4. Choo and Clin	ck Compu	te	Table annota Top te Colors	of analg ation-er on terms s in the legend	y and lysis r ntity p s are table	result, node pairs are sho colored. and nodes the Network	attributes, e own. are the san	edge lists and ne. Search:	Sur
work onsole w 10 *	Enric entries	4. Choc and Clin thment	<u>ck Compu</u> Node Edge	AnnotationPair	Table annota Top te Colors Color	of analg ation-er en terms s in the legend P	lysis n ntity p s are table l is in	result, node pairs are sho colored. and nodes the Network	attributes, e own. are the san c panel.	edge lists and ne. Search:	
work	Enric entries id ¢	4. Choo and Clin chment s gid	Node Edge nodename Arginine and proline metabolism	AnnotationPair	Table annota Top te Colors Color	of analg ation-er en terms s in the legend P	y and lysis n ntity p s are table i s in p_adj \$	result, node pairs are sho colored. and nodes the Network no_ot_entities	attributes, e own. are the san panel. annotation_size	edge lists and ne. Search: member	1407,37
work	Enric entries id ¢	4. Choc and Clin shment gid () hsa00330	Node Edge nodename Arginine and proline metabolism Arginine biosynthesis Pentose and glucuronate	te AnnotationPair AnnotationPair Pathway Pathway Pathway	Table annota Top te Colors Color nodexref	of analy ation-er on terms s in the legend p 0.0114 c	lysis n ntity p s are table ! is in i p_adj ()	result, node pairs are sho colored. and nodes the Network no_ot_entities	attributes, e pwn. are the san ( panel. annotation_size ( Compound (84)	me,           Search:           member           1018,41640,9483,864,4	1407,37 7,38891
work	Enrice entries id 0 434555	4. Chocand Clin	Node Edge nodename Arginine and proline Marginine and proline biosynthesis Pertoarante interconversion Buticosin and neomych	te AnnotationPair AnnotationPair Pathway Pathway Pathway	Table annota Top te Colors Color nodexref hsa00330	of anal; ation-er en terms s in the legend p 0.0114 c 0.0224 c	y and lysis m ntity p s are table i s in p_adj ( 0.3808	result, node pairs are sho colored. and nodes the Network no_ot_entities () Compound (7) Compound (5)	attributes, e own. are the sam panel. annotation_size Compound (84) Compound (26)	edge lists and           Bearch:           member           1019.41640.9483.864.4           41407.3771.36414.4433	1407,37 7,38891
work onsole w 10 •	entriet id 0 434555 434542	4. Chocand Clin	Node Edge nodename Arginine and protilee metabolism Arginine and protilee metabolism Pentose and glucuroate interconversion Buttrosin and pentose and glucuroate interconversion Buttrosin and pentose and pentose and glucuroate interconversion Buttrosin and pentose and pentose and glucuroate interconversion Buttrosin and pentose a	te AnnotationPair ondelabel () Pathway Pathway Pathway Pathway	Table annota Top te Colors Color nodexref () hsa00300 hsa00200 hsa00400	of anal; ation-er n terms s in the legend p 0.0114 c 0.0224 c 0.0228 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	esult, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (7) Compound (3)	attributes, e wm. are the sam; panel, annotation_itime i Compound (26) Compound (27) Compound (37)	me.           Search:	1407,37 7,38891
work onsole w 10 •	entrie: id \$434566	4. Chocand Clin	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table annota Top te Colors Color nodexret () hsa00330 hsa00220	of anal; ation-er n terms s in the legend p 0.0114 c 0.0224 c 0.0228 c	y and lysis n ntity p s are table i s in p_adj () 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_ct_entities () Compound (7) Compound (7)	attributes, e own. are the sam panel. annotation_size Compound (84) Compound (84) Compound (87)	member           1019.41640.9483.864.4         41407.3771.36414.4433           39145.3771.8141.40296	1407,37 7,38891
twork	entriet id 0 434555 434542	4. Chocand Clin	kode Edge     kode     ko	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table annota Top te Colors Color nodexref () hsa00300 hsa00200 hsa00400	of anal; ation-er n terms s in the legend p 0.0114 c 0.0224 c 0.0228 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	esult, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (7) Compound (3)	attributes, e wm. are the sam; panel, annotation_itime i Compound (26) Compound (27) Compound (37)	edge lists and           Bearch:           mmbber           1018,41640,3483,864,4           41407,3771,36414,4433           39145,3771,8141,40294           41407,10648,919	1407,37 7,38891
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the second sec	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table annota Top te Colors Color nodexref () hsa00300 hsa00200 hsa00400	of anal; ation-er n terms s in the legend p 0.0114 c 0.0224 c 0.0228 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	esult, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (7) Compound (3)	attributes, e wm. are the sam r panel. anotation_size [ Compound (84) Compound (87) Compound (87) Compound (27)	member           1018,41640,9483,664,4           39145,3771,8641,40296           41407,3771,8641,40296           41407,3771,554	11407,37 7,38891 9,35781,
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the second sec	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table annota Top te Colors Color nodexref () hsa00300 hsa00200 hsa00400	of anal; ation-er n terms s in the legend p 0.0114 c 0.0224 c 0.0228 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (5) Compound (3) Compound (3)	attributes, e wm. are the sam r panel. anotation_size [ Compound (84) Compound (87) Compound (87) Compound (27) Previous	member         norm           1018,41640,3483,664,4         41407,3771,3844,4433           39145,3771,8141,40296         41407,3771,584           11         2         3	5 1
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the second sec	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table annota Top te Colors Color nodexref () hsa00300 hsa00200 hsa00400	of anal; ation-er n terms s in the legend p 0.0114 c 0.0224 c 0.0228 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (5) Compound (3) Compound (3)	attributes, e wm. are the sam r panel. anotation_size [ Compound (84) Compound (87) Compound (87) Compound (27) Previous	Pegge lists and Peg. Search: 1018,41640,9483,064,4 41407,3771,30444,4433 39145,3771,8141,40296 41407,10648,919 41407,3771,504	5 1
wwork onsole  w 10 •  d  d  d  d  d  d  d  d  d  d  d  d  d	Image: Constraint of the second sec	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table annota Top te Colors Color nodexref () hsa00300 hsa00200 hsa00400	of analy ation-er n terms s in the legend p 0 0.0114 c 0.0224 c 0.0228 c 0.0288 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_at_entities Compound (7) Compound (5) Compound (7) Compound (3) Compound (3)	attributes, e wm. are the sam r panel. anotation_size [ Compound (84) Compound (87) Compound (87) Compound (27) Previous	adge lists and           Bearch:           member           1019.41640.9483.864,4           41407.3771,30414,4433           39145.3771,30414,4433           41407,3771,8141,4029           41407,3771,554           1         2           3         4	11407,37 7,38891 9,35781,
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the second sec	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table           annota           Top te           Colors           Color           nea00320           hsa00220           hsa0040           hsa00524           hsa00430	of analy ation-er n terms s in the legend p 0 0.0114 c 0.0224 c 0.0228 c 0.0288 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	esult, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (3) Compound (3) Compound (3)	attributes, e swn. are the san panel, anotation_size [ Compound (28) Compound (27) Compound (27) Previous resulting in annotation	adge lists and           Bearch:           member           1019.41640.9483.864,4           41407.3771,30414,4433           39145.3771,30414,4433           41407,3771,8141,4029           41407,3771,554           1         2           3         4	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the state of	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge Redename Arginine and metabolism Pentose and plucaronale Interconversion Batirosin and hypothesis Totarine and hypothesis	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table           annota           Top te           Colors           Color           nea00320           hsa00220           hsa0040           hsa00524           hsa00430	of analy ation-er n terms s in the legend p 0 0.0114 c 0.0224 c 0.0228 c 0.0288 c	y and lysis m ntity p s are table l is in p_adj = 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_of_entities Compound (7) Compound (5) Compound (3) Compound (3)	attributes, e swn. are the san panel, anotation_size [ Compound (28) Compound (27) Compound (27) Previous resulting in annotation	edge lists and           nee.           Search:           1019,41640,9483,664,4           1019,41640,9483,664,4           39145,3771,8141,40296           41407,3771,8641,4433           1         2           3         4           41407,3771,564           1         2           1         2           4         4           4	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the state of	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	Node Edge nodename Arginine and proline a	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table           annota           Top te           Colors           Color           nea00320           hsa00220           hsa0040           hsa00524           hsa00430	of analy ation-er n terms s in the legend p 0 0.0114 c 0.0224 c 0.0228 c 0.0288 c	y and (ysis n ntity p s are table is in 1 p_#¶ 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_ct_entities Compound (7) Compound (5) Compound (7) Compound (3) Compound (3)	attributes, e swn. are the sarr panel. anotation_size Compound (84) Compound (87) Compound (87) Compound (87) Compound (87) Previous Previous	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the state of	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge Redename Arginine and metabolism Pentose and plucaronale Interconversion Batirosin and hypothesis Totarine and hypothesis	te AnnotationPair nodelabel Pattway Pattway Pattway Pattway Pattway Pattway Pattway	Table           annota           Top te           Colors           Color           nea00320           hsa00220           hsa0040           hsa00524           hsa00430	o of analy attion-ein in terms in the legend 0.0214 c 0.0228 c 0.0082 c 0.0082 c 0.0080 c	y and (ysis n ntity p s are table is in 1 p_#¶ 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_ct_entities Compound (7) Compound (5) Compound (7) Compound (3) Compound (3)	attributes, e wm. are the sam panel. amotation_size Compound (24) Compound (26) Compound (27) Previous resulting m annotation les that are top ten ann	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the state of	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge nodename Arginine Arginine Interconversion Buttroin and neomydh Interconversion Buttroin and Rentos and Buttroin and Buttroin and Rentos and Rento	te AnnotationPair nodelabel Pathway Pathway Pathway Pathway Pathway	Table           annota           Top te           Colors           Color           nea00320           hsa00220           hsa0040           hsa00524           hsa00430	o of analy attion-ein in terms in the legend 0.0214 c 0.0228 c 0.0082 c 0.0082 c 0.0080 c	y and (ysis n ntity p s are table is in 1 p_#¶ 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_ct_entities Compound (7) Compound (5) Compound (7) Compound (3) Compound (3)	attributes, e wm. are the sam panel. amotation_size Compound (24) Compound (26) Compound (27) Previous resulting m annotation les that are top ten ann	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • 4 4 4 4 4	Image: Constraint of the state of	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge nodename Arginine Arginine Interconversion Buttroin and neomydh Interconversion Buttroin and Rentos and Buttroin and Buttroin and Rentos and Rento	te AnnotationPair nodelabel Pathway Pathway Pathway Pathway Pathway	Table           annota           Top te           Colors           Colors           nedexref           hsa00220           hsa00220           hsa0040           hsa0040           hsa0040           hsa00430	of anal/ ation-er in terms: in the legend 0.0224 c 0.0224 c 0.0224 c 0.0228 c 0.0228 c 0.0209 c	y and hysis n ntity p s are table is in i p_#4 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_ct_entities Compound (7) Compound (5) Compound (7) Compound (3) Compound (3)	attributes, e swn. are the sam; panel, annotation_size [ Compound (28) Compound (28) Compound (27) Previous resulting in annotation les that are top ten ann annotated a	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • 4 4 4 4 4	10         4           Enrice         entries           ed         434566           434565         434565           10         of 48           10         of 48	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge nodename Arginine Arginine Interconversion Buttroin and neomydh Interconversion Buttroin and Rentos and Buttroin and Buttroin and Rentos and Rento	te AnnotationPair nodelabel Pathway Pathway Pathway Pathway Pathway	Table           annota           Top te           Colors           Colors           nedexref           hsa00220           hsa00220           hsa0040           hsa0040           hsa0040           hsa00430	of anal/ ation-er in terms: in the legend 0.0224 c 0.0224 c 0.0224 c 0.0228 c 0.0228 c 0.0209 c	y and hysis n ntity p s are table is in i p_#4 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_of_entities Compound (5) Compound (5) Compound (3) Compound (3) Compound (3) Compound (3)	attributes, e swn. are the sam; panel, annotation_size [ Compound (28) Compound (28) Compound (27) Previous resulting in annotation les that are top ten ann annotated a	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	11407,37 7,38891 9,35781, 9,35781, 5 P
twork onsole w 10 • ank * 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	10         4           Enrice         entries           ed         434566           434565         434565           10         of 48           10         of 48	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge nodename Arginine Arginine Interconversion Buttroin and neomydh Interconversion Buttroin and Rentos and Buttroin and Buttroin and Rentos and Rento	te AnnotationPair nodelabel Pathway Pathway Pathway Pathway Pathway	Table           annota           Top te           Colors           Colors           nedexref           hsa00220           hsa00220           hsa0040           hsa0040           hsa0040           hsa00430	of anal/ ation-er in terms: in the legend 0.0224 c 0.0224 c 0.0224 c 0.0228 c 0.0228 c 0.0209 c	y and hysis n ntity p s are table is in i p_#4 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_of_entities Compound (5) Compound (5) Compound (3) Compound (3) Compound (3) Compound (3)	attributes, e swn. are the sam; panel, annotation_size [ Compound (28) Compound (28) Compound (27) Previous resulting in annotation les that are top ten ann annotated a	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	11407,37 7,38891 9,35781, 9,35781, 5 P
twork	10         4           Enrice         entries           ed         434566           434565         434565           10         of 48           10         of 48	4. Chocc and Cli threat s gid hsa00330 hsa00220 hsa00430	ck Computer Node Edge nodename Arginine Arginine Interconversion Buttroin and neomydh Interconversion Buttroin and Rentos and Buttroin and Buttroin and Rentos and Rento	te AnnotationPair nodelabel Pathway Pathway Pathway Pathway Pathway	Table           annota           Top te           Colors           Colors           nedexref           hsa00220           hsa00220           hsa0040           hsa0040           hsa0040           hsa00430	of anal/ ation-er in terms: in the legend 0.0224 c 0.0224 c 0.0224 c 0.0228 c 0.0228 c 0.0209 c	y and hysis n ntity p s are table is in i p_#4 0 0.3808 0.3808 0.3808 0.3808	result, node pairs are sho colored. and nodes the Network no_of_entities Compound (5) Compound (5) Compound (3) Compound (3) Compound (3) Compound (3)	attributes, e swn. are the sam; panel, annotation_size [ Compound (28) Compound (28) Compound (27) Previous resulting in annotation les that are top ten ann annotated a	adge lists and           Berch:           member           1018.41640.5483.664.4           41407.3771.36414.4433           39145.3771.8141.40286           41407.3771.554           1         2           1         2           1         2           1         2           41407.3771.554	7,38891 9,35781,/ 5 ► mapj

6	TRUE	460567	158	5810	true	
7	TRUE	1067178	237	14985	true	
owing 1 to 45	2 of 462 entries					
Download	I Statistical Analysi	s Result				
	,,		1. CI	noose Wor	dCloud	

Fig 3. Computing WordCloud for measured compounds.

Compute Word							
Compute WordU	oud for the Ir	pur.					
nputs							
intity lists (e.g. Pu	ibChem or u	niprot or ensembi)	will be used for	WordCloud g	eneration.		
Input summary						show in inter	active tables. Input overvie
how 10 * entr	ies	PubChe	em colum	n will be	used	to compute	WordCloud. Search:
mz   PubChe	m i outli exist	r 0 p_value_N	vs_T () fdr_d	adjusted_p_va	lue_N_vs_	T  non_parametr	ic_p_value_N_vs_T   fdr_adjusted_non_parametric_p_value
6912	false	0.0001	0.0016	8		0.0002	0.0024
64959	true	0.7192	0.8266			0.6046	0.7125
1188	true	0	0.0008	5		0	0.0007
6287	false	0.9584	0.9643			0.8739	0.9218
6030	true	0.4893	0.6458			0.4185	0.5687
6029	true	0.0003	0.0032			0.0001	0.0012
1176	true	0.5425	0.6811			0.9833	0.9943
1174	false	0.0103	0.0396	5		0.0082	0.0329
17473	false	0.002	0.0112	2		0.0019	0.0115
* howing 1 to 10 of -	462 entries						Previous 1 2 3 4 5 47 Next
intity type:		Sel	ect annotation: ( sh annotation is a				
compound		Me	sn annotation is a	wallable for Co	mpound on	y.	
Compute 2	Choos	se Pathwa	y and Clic	ck Comp	oute		
tputs	Cloud N	Top t Color	en terms	are colo	ored. C		and annotation-entity pairs are shown table and nodes are the same.
w 10 * entries							Search:
ank * id 🕴	gid 0	nodename 0	nodelabel ()	nodexref	freq ()	member	
434820	hsa05230	Central carbon metabolism in cancer	Pathway	hsa05230	20		17,864,446,36946,113,36858,1330,1774,36522,36414,41407,919,404
434764	hsa04074	Protein digestion and absorption	Pathway	hsa04974	18	38680,335,39090,996	,1019,664,446,36946,1996,36858,1330,1774,39522,36414,41407,2
434556	hsa00230	Purine metabolism Aminoacyf-tRNA	Pathway	hsa00230	15		1086,1852,2984,224,40618,40408,38522,36414,41308,24424,2884
434620	hsa00970	biosynthesis	Pathway	hsa00970	15	38680,335,39090,996	064,446,36946,1996,36856,1330,1774,39522,36414,41407,554
434558	hsa00240	Pyrimidine metabolism	Pathway	hsa00240	11	10427,11317,11095,1	1539,8762,37669,5655,58414,37988,33313,38746
494800	keennean	Glycxylate and steadousciets	Detheaser	keennitan	10	8007800044338087	10800 A0880 18444 A1407 17887 1774
etwork							
L-Pro	line			ļ	Alpha-	Tocopherol	Pseudourid
L-Glutan	) nic acid	I			Сарі	V Tylic acid	Dodecanoic
Stigma					Eter	<b>V</b> aric acid	D-Ribose
Sugma	steroi				Stea	aric aciu	D-Ribose
L-Tyro	sine			Aden	osine i	monophosph	ate Adenosin
Xanthe	osine				qui	nic acid	Galactosylgly
cde legend		Nodes the	t are not	the part	s of th	e top ten anr	notation terms. notations or not annotated are in gre e in the WordCloud panel.
R <sub>a</sub> Certui cator residuitor	Netw	ork panel	allows int				ensuring at the second se
	To pa To zo	n – click, l om – scro	hold and Il the mou	drag bao use whe	ckgrou el	ind	ox around the nodes
VordCloud						-	
				Electronic and a second and a s	Protein digestion and absorption		and Barrow The State of State of State of State Barrow State of State of St
					Protei		

▲.bownod functional analysis outputs will download results of functional analysis and the network mapped with annotation terms including corresponding legend.

STATISTICAL ANALYSIS	Query het	erogene	ous network	
<b>di</b> Statistical Analysis	O Query the he	derogeneous ne	twork containing several relationship types from the database.	×
NETWORK CONSTRUCTION	Concernance of			
& BioNetwork +	Inputs			~
» SimpleNetwork	From node: ensembl ac	ijPval log2FC		or Upload a file: Choose File ensembl_stats.txt
HeterogeneousNetwork     Choose Heteroge     Let Correlations -		059066 4 0112751.52E- 01173081.24E-	26 1.55	See Note for input format. 2. Upload ensembl_stats.txt file in From field and pubchem stats.txt file
A Similarity	Separate each To node:	ID by a new line	(try example).	in To field
FUNCTIONAL ANALYSIS	1000000	IJPval log2FC		Choose File publichem_stats.bd
C FriClassScoring	5325915 0. 5312542 0. 656504 0.	0078 -0.685 0098 0.6227 0053 1.1396	239907 94737 06688	See Note for input format.
Overrepresentation		0001 1.2101	1	
A WordCloud	Separate each	ID by a new line	(try example).	
(MISCELLANEOUS)	- For multico	lumn table: 1st (	rate each ID by a new line. column = 10 bists with PubChem or uniprot or ensembl as a header (see below), separate each co ciD for compounds, UniProt entry for proteins, Emembl for genes.	iumn by tab.
↔ ConvertID	1 Conversion			
≢ Tutorial				
	Pattern:			
er, concessure -	3. Con	struct re	om Gene}{:CONVERSION ->(Protein){:CATALYSIS ->(to:Compound) ationship pattern as	
		ick Quer	NTROL]->(from:Gene)-[:CONVERSION]->(:Protein)	)-[:CATALYSIS]->(to:Compound)
FUNCTIONAL ANALYSIS			ise the constructor below to define the relationship pattern (try example).	
FnClassScoring	From node:	0 · Con	pound i Protein i Gene i Pathway	
W Overrepresentation	To node:	0 🕀 Con	pound 💿 Protein 💿 Gene 💿 Pathway	
Sector WordCloud	Other node:	0 © Con	ipound 🔘 Protein 🔘 Gene 🍈 Pathway	
	Relationship		FATION © BIOCHEMICAL_REACTION © CATALYSIS © CONTROL ERSION © GENETIC_ASSOCIATION © MOLECULAR_BINDING	
MISCELLANEOUS				
← ConvertID				Database schema
🗭 Tutorial	Add			
	Query C	llear		
Console Node Edge	Node lists	and edg	e lists are show in interactive tables.	Summary
Show 10 * entries			Search:	
id 0 gid 0	nodename	nodelabel *	nodexref	
4991 985	Palmitic acid	Compound	985((HMDB00220))C00249((CHEBI:15756)(FDB011679))CPD-8475	
9316 91486	Sphinganine	Compound	91486  HMDB00269  C00836  CHEBI:16566  FDB021924	
10548 445675	Uridine diphosphate-N- acetylglucosamine	Compound	445675[[HMDB00280][C00043][CHEBL16264][FDB021930][UDP-N-ACETYL-D-GLUCOSAM	NE
39145 17473	Undine diphosphate glucuronic acid	Compound	17473  HMDB00835  C00187  CHEBI:17250  FDB022325  UDP-GLUCURCNATE	



Network pane of the resulting To pan — click To zoom — click To select mult and drag a bo wherek tegand	g netwo , hold a roll the r tiple noc	ork. and drag mouse des – p	g back wheel ress si	ground			_	
Relationship type ANNOTATION BIOCHEMICAL_REA Node type		CATALYSIS	COP	WERSION RRELATION		ARTIAL_CORRELATION ANIMOTO_SIMILARITY		
Compound DNA	Gene	Pathway	Protein	RNA				
		-	-		Download network	to downloa	d node lists	odao liste
Download network	of Subnetw	work 💿			static image of the		work	
Download network		work 💿					work	

## Fig 4. Constructing biological network from gene lists and compound lists. Relationship pattern between genes and compounds is:

(:Protein)-[:CONTROL]->(from:Gene)-[:CONVERSION]->(:Protein)-[:CATALYSIS]->(to:Compound)

#### REFERENCES

1. Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome research. 2012;22(7):1197-211.

2. Wikoff WR, Grapov D, Fahrmann JF, DeFelice B, Rom WN, Pass HI, et al. Metabolomic markers of altered nucleotide metabolism in early stage adenocarcinoma. Cancer prevention research. 2015;8(5):410-8.

3. Fiehn O, Wohlgemuth G, Scholz M. Setup and annotation of metabolomic experiments by integrating biological and mass spectrometric metadata. Lect Notes Comput Sc. 2005;3615:224-39.

4. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic acids research. 2013;41(Database issue):D991-5.

5. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research. 2003;13(11):2498-504.