Importing Affymetrix CEL Files with TAC for BRB-ArrayTools

Version 1.0

Вy

Dr. Qian Xie

BRB-ArrayTools Development Team

arraytools@emmes.com

Jan. 2021

Contents

1.	Intro	oduction	1
2.	Dow	vnload and install TAC Software	2
3.	Con	vert CEL files to TXT file with TAC	4
4.	Imp	ort the TXT file in BRB-ArrayTools	7
5.	Exar	nples of other array types	11
5	5.1.	Affymetrix Mouse Gene 1.0 ST Array	12
5	5.2.	Affymetrix HT HG-U133+ PM Array	13

1. Introduction

<u>Affymetrix Expression Console Software is now part of the Transcriptome Analysis Console (TAC)</u> <u>Software</u>. In this instruction, you use the TAC software to convert Affymetrix CEL files to a probe summarization TXT file and then import the TXT file in BRB-ArrayTools to collate a BRB-ArrayTools project.

2. Download and install TAC Software

You can download the TAC software from the website <u>https://www.thermofisher.com/us/en/home/life-science/microarray-analysis/microarray-analysis-instruments-software-services/microarray-analysis-software/affymetrix-transcriptome-analysis-console-software.html</u>. Current version is TAC 4.0.2.15 (Nov. 2020). After installing the TAC software, you need to set up the software by clicking the tab "Preferences":

• Set Library Folders: C:/Users/%username%/Documents/TACX

applied biosystems	TAC 4.0	?	í	-		×
New Analysis Open Existing Result Preferences						
Download Library Files Library Folder: C:\Users\xieq\Documents\TACX					Brow	se
▲ NetAffx Settings						
NetAffx Username:						
NetAffx Password:						
Internet Proxy Settings						
Installed Array Types						
Array Type Version Description						

- Download the Library files for your array type:
 - Click the "Download Library Files" button, a NetAffx Account Information window pops up. If you have an account, you may enter your account information here and click "OK". If you do not have an account, you need to click "Register Now" to create an account in your default browser and then enter here.

applied biosystems	TAC 4.0	?	í	-		×
New Analysis Open Existing Result Preferences						
Download Library Files Library Folder: C:\Users\xieq	\Documents\TACX				Brows	se
 NetAffx Settings 						
NetAffx Username:						
NetAffx Password:						
Automatic Update Library Files						
 Internet Proxy Settings 						
Installed Array Types	Not Affer A consult information					
Array Type Version Description	Enter your NetAffy amail address and password					
	Enter your NetArix email address and password.					
	Email: dian_xie@sra.com					
	Register Now					
	<u>regists rivi</u>					
	OK Cancel					

 The NetAffx Account Information will automatically fill in the NetAffx Settings field and a window pops up for selecting array types. Select miRNA-4_0 and click the "Download" button. The library files for the array type will be downloaded and the relevant information will keep under the "Installed Array Types" field.

applied biosystems	TAC	4.0	(? () _ □	×
New Analysis Open Existing Result Preferences	NetAffx Library Files				
Download Library Files Library Folder: C:\Users\xieq	Array Type	Information		Brow	/se
▲ NetAffx Settings	MedTA-1 0		^		
NetAffx Username: qian_xie@sra.com					
Automatic Update Library Files	☐ MG_U74Bv2				
▼ Internet Proxy Settings	□ MG_U74Cv2			-	
Installed Array Types	miRNA-2_0				
Array Type Version Description	miRNA-3_0				
· · · · ·	miRNA-3_1				
	miRNA-4_0				
	miRNA-4_1				
	MOE430A				
	MOE430B				
	MoEx-1_0-st-v1				
	MoGene-1_0-st-v1				
	MoGene-1_1-st-v1				
	☐ MoGene-2_0-st				
	MoGene-2_1-st		~		
		Select array type(s) to download.			
			Download Cancel	1	

3. Convert CEL files to TXT file with TAC

You keep all CEL files in a folder. Here uses the NCBI GEO GSE123377 dataset as an example. You may download the dataset <u>https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE123377&format=file</u>.

You click the tab "New Analysis" to start a new study.

 Import Files > Import CEL Files: a file browser window pops up > Select CEL files (NCBI GEO GSE123377) > Open.

applied biosyst	tems	TAC 4.0	? (ì _ □ ×
New Analysis	Open Existing Result Prefe	rences	
Import Files 💌	Remove Selected File(s)	d New Column Sample Attributes	
File Name	Market CEL Files		×
	← → ~ ↑ 📕 «	Datasets > Users > GSE123377miRNA4 v 💍 $>$ Search GSE123	377miRNA4
	Organize - New fo	der 📲 🔹	· 🔳 🕜
	HotLinks	Date modified Type	Size ^
	OneDrive	GSM3502404_4877_6_miRNA-4_0_CEL 8/28/2014 7:31 AM GIMP 2.10.22	2,8
	This PC	GSM3502405_4877_7_miRNA-4_0_CEL 8/28/2014 7:42 AM GIMP 2.10.22	2,8
	3D Objects	GSM3502406_4877_8_miRNA-4_0_CEL 8/29/2014 7:32 AM GIMP 2.10.22	2,8
	Deskton	GIMP 2.10.22	2,8
	Documents	GIMP 2.10.22	2,8
	Downloads	GSM3502409_4877_35_1_miRNA-4_0 8/29/2014 10:40 AM GIMP 2.10.22	2,8
	Downloads	GIMP 2.10.22	2,8
	Music	GIMP 2.10.22	2,8
	Pictures	GIMP 2.10.22	2,8
	Videos	GIMP 2.10.22	2,8
	🐛 Windows (C:)	GSM3502414_4877_34_miRNA-4_0_CEL 8/29/2014 10:52 AM GIMP 2.10.22	2,8 🗸
		/ <	>
	File	ame: "GSM3502414_4877_34_miRNA-4_0_CEL" "GSM3502388_4877_1. >	~
	_	Open	Cancel
Result Name: A	nal,	er Festeseiskund te neuristustustustustustus	Browse
			Run Analysis

• In the column "Condition", change Type: Comparison and assign the column values. Half of the files are "1" and another half of the files are "2". These assignments are random. It is for the program to process. Click the lower-left button "Run Analysis".

applied biosystems	TAC 4.0	? (ì _ □ ×
New Analysis Open Existing Result Prefe	ences	
Array Type: miRNA-4_0 Analysis Type: Expre	sion (Gene) V Summarization: RMA+DABG (All Organisms) V	Version: version 1 ~
Import Files Remove Selected File(s) Add	New Column Sample Attributes 💌 Clear Sorting Show Graph 💌	
File Marrie (07)	Condition	
File Name (27)	Type: Comparison ¥	
GSM3502388_4877_1_miRNA-4_0CEL	1	^
GSM3502389_4877_2_miRNA-4_0CEL	1	
GSM3502390_4877_3_miRNA-4_0CEL	1	
GSM3502391_4877_4_miRNA-4_0CEL	1	
GSM3502392_4877_5_miRNA-4_0CEL	1	
GSM3502393_4877_16_miRNA-4_0CEL	1	
GSM3502394_4877_17_miRNA-4_0CEL	1	
GSM3502395_4877_18_miRNA-4_0CEL	1	
GSM3502396_4877_19_miRNA-4_0CEL	1	
GSM3502397_4877_20_miRNA-4_0CEL	1	
GSM3502398_4877_21_miRNA-4_0CEL	1	
GSM3502399_4877_22_miRNA-4_0CEL	1	
GSM3502400_4877_23_miRNA-4_0CEL	1	
GSM3502401_4877_24_miRNA-4_0CEL	2	
GSM3502402_4877_32_miRNA-4_0CEL	2	~
Result Name: Analysis_1 Output Fold	r: [C:\Users\xieq\Documents\TACX\AnalysisResults	Browse
	Algorithm Settings Comparison S	Setup Wizard Run Analysis

• When the analysis finishes, a new window pops up with the results.

applied biosystems				Analysi	s_1.tacx) () _	6 ×
Sample QC View Summary Vi	ew Gene Vi	iew					
Sample Table					PCA Hybridization Controls (3') Hybridization Controls (5') Spike-In Controls Signal Box Plot		
Apply View Filters Show/Hid	e Columns Exp	oort 🔻 🛛 Add Co	olumn Reanalyze	nples Create Line Graph	Color By: Condition V Shape By: <none> V Default View</none>		≡
File Name	Hybridization Controls (3') Threshold	Hybridization Controls (5') Threshold	Spike-In Controls Conc Threshold	ก	PCA Mapping 41.1% (CHP)		Condition
GSM3502388 4877 1 miRNA-4 0	Pass	Pass	Pass 1				<u> </u>
GSM3502389 4877 2 miRNA-4 0	Pass	Pass	Pass 1		PCA2 14.2%		
GSM3502390 4877 3 miRNA-4 0	Pass	Pass	Pass 1				
GSM3502391 4877 4 miRNA-4 0	Pass	Pass	Pass 1				
GSM3502392 4877 5 miRNA-4 0	Pass	Pass	Pass 1				
GSM3502393 4877 16 miRNA-4 0	Pass	Pass	Pass 1				
GSM3502394_4877_17_miRNA-4_0	Pass	Pass	Pass 1				
GSM3502395 4877 18 miRNA-4 0	Pass	Pass	Pass 1				
GSM3502396_4877_19_miRNA-4_0	Pass	Pass	Pass 1				
GSM3502397_4877_20_miRNA-4_0	Pass	Pass	Pass 1				
GSM3502398_4877_21_miRNA-4_0	Pass	Pass	Pass 1				
GSM3502399_4877_22_miRNA-4_0	Pass	Pass	Pass 1				
GSM3502400_4877_23_miRNA-4_0	Pass	Pass	Pass 1				
GSM3502401_4877_24_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502402_4877_32_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502403_4877_33_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502404_4877_6_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502405_4877_7_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502406_4877_8_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502407_4877_9_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502408_4877_10_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502409_4877_35_1_miRNA-4	Pass	Pass	Pass 2				
GSM3502410_4877_11_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502411_4877_12_miRNA-4_0	Pass	Pass	Pass 2			PCA1 17.8%	
GSM3502412_4877_13_miRNA-4_0	Pass	Pass	Pass 2				
GSM3502413_4877_14_miRNA-4_0	Pass	Pass	Pass 2		•		
GSM3502414_4877_34_miRNA-4_0	Pass	Pass	Pass 2				
					100323		
	?7 Selected: 0 🕇	+		Show Filtered Only	Left Mouse Action: Rotate v Right Mouse Action: Lasso v Zoom: Size:		

• Click the "Gene View" in the new window. And click "Customize Annotations" to select information columns for output. In the popped up window, check "Top Assignment" to pick the first subfield in each annotation column and select gene information columns you need. Click "OK".



• Click "Show/Hide Columns", a window will pop up. You can select columns for output. Click "OK".

applied biosystem	s				Analysis	_1.tacx								? ()	-	б×
Sample QC View	Summary	View Gene Vie	w													
Gene Table Sum	imary Samp	le Table				Scatter Plot	Volcano Plot	Hierarchical Cl	ustering Sa	nple Signals						
Comparison: 1 vs	2 🔹 🗲 What	t is the differential ex	xpression between 1 a	nd 2?		Color Settings	Show Filte	red Only			_					=
Apply View 💌 Filte	ers Show/H	Hide Columns Expo	Add Column	Customize Annotation	Gene List miRNA Interaction				1.	Avg (log2) v	s 2 Avg (lo	g2)				
ID count: 36222	Target Gene Symbols	Transcript ID(Array Design)	Accession	Probe Set Name		14.4										•
20535182		hsa-mir-521-2	MI0003163	MI0003163_x_st She	w/Hide Columns		×								•	
20525120		bta-miR-2285v	MIMAT0025579	MIMAT0025579_	4.4											
20512215		ppc-miR-2273	MIMAT0011730	MIMAT0011730	1 Avg (log2)											
20514738		aae-miR-1174	MIMAT0014265	MIMAT0014265_	2 Avg (log2)									\$		
20525478	ABLIM1	hsa-miR-6758-3p	MIMAT0027417	MIMAT0027417_	1 Standard Deviation									F		
20518208		tca-miR-iab-8-3p	MIMAT0018699	MIMAT0018699_	2 Standard Deviation								8			
20522571		osa-miR5814	MIMAT0023286	MIMAT0023286_	1 Expressed								1			
20529108		ptc-miR3627b	MIMAT0030391	MIMAT0030391_	2 Expressed							<u></u>				
20535950		mmu-mir-1892	MI0008320	MI0008320_st	Fold Change											
20533134		ENSG00000212618	ENSG0000212618	ENSG00002126	P-val											
20529180		prd-miR-239-3p	MIMAT0030463	MIMAT0030463	F-981											
20514958		tgu-miR-2965	MIMAT0014468	MIMAT0014468_	FDR P-Val							5				
20511655		gma-miR1509b	MIMAT0011201	MIMAT0011201	Expressed in Both Conditions						1.18					
20534298		HBII-55	HBII-55	HBII-55_st	Target Gene Symbols											
20512653		dps-miR-375-5p	MIMAT0012253	MIMAT0012253_	Validated Gene Symbols						6 -0					
20523287		cgr-miR-505-3p	MIMAT0023980	MIMAT0023980_	Transcript ID(Array Design)					100						
20504857		xtr-miR-16c	MIMAT0003562	MIMAT0003562	Accession											
20507154		oan-miR-16c-5p	MIMAT0006847	MIMAT0006847_	Chromosome											
20508315		gga-miR-16c-5p	MIMAT0007739	MIMAT0007739_	Strand											
20515169		tgu-miR-16c	MIMAT0014586	MIMAT0014586_	Suanu											
20521318		aca-miR-16b-5p	MIMAT0021801	MIMAT0021801_	Start			88	22							
20525668	ASB1	hsa-miR-6853-3p	MIMAT0027607	MIMAT0027607_	Stop			. 466								
20506499		mmu-miR-582-5p	MIMAT0005291	MIMAT0005291_	Probe Set Name											
20524626		mmu-miR-6341	MIMAT0025084	MIMAT0025084_												
20537586		hsa-mir-6841	MI0022687	MI0022687_st		OK	Cancel	· *								
20524200		nta-miR6025d	MIMAT0024732	MIMAT0024732_				100°6								
20521034		asu-miR-375-5p	MIMAT0021525	MIMAT0021525_st				5 · ·								
20533255		ENSG0000238377	ENSG0000238377	ENSG0000238377_st	_			•								
20534966		rno-mir-99a	MI0000883	MI0000883_st	_	1.6	ALC: NO									
20534678		mmu-mir-92a-2	MI0000580	MI0000580_st	_		1539									
20534115		ENSG00000253091	ENSG00000253091	ENSG00000253091_st												
20501929		gga-miR-200a-5p	MIMAT0026533	MIMAT0026533_st	~											
P Find in Table	Coun	t: 36222 Selected: 0	t+		Show Filtered Only	0	1.6	3.2 4	.8 6.	1 8 2 Avg	3 (log2)	9.6	11.2	12.8	14.4	16

 Click "Export" > Export..., an Export window pops up. You can select the output contents. Click "OK".

applied biosystem:	s				Analysis	1.tacx							0	i _	ъх
Sample OC View	Summary	View Gene Vie	w												
Gene Table Sum	mary Samp	le Table				Scatter Plot	Volcano Plot	Hierarchical	Clustering Sar	mple Signals					
Comparison: 1 vr.3	2 - + What	t is the differential ex	pression between 1 a	nd 22		Color Setting	s Show Filte	ared Only U	se EDR PaVal						=
comparison. 1 vs z	+ Wild	ris die differendares	pression between 1 a	10 21		Color Setting	3 - 310W THE				-				-
Apply View 🔻 Filte	ers Show/H	Hide Columns Expo	rt 💌 Add Column	Customize Annotations	Gene List miRNA Interaction	5				P-val vs Fold	Change				
ID count: 36222	Target Gene Symbols	Transcript ID(Array Design)	Accession	Probe Set Name		45									
20535182		hsa-mir-521-2	MI0003163	MI0003163_x_st	^					•					
20525120		bta-miR-2285v	MIMAT0025579	MIMAT0025579_st	-										
20512215		ppc-miR-2273	MIMAT0011730	MIMAT0011730_st				_							
20514738		aae-miR-1174	MIMAT0014265	MIMAT0014265_st	Export										
20525478	ABLIM1	hsa-miR-6758-3p	MIMAT0027417	MIMAT0027417_st	E						•				
20518208		tca-miR-iab-8-3p	MIMAT0018699	MIMAT0018699_st	Export Options										
20522571		osa-miR5814	MIMAT0023286	MIMAT0023286_st	O Table Columns Only			-			••				
20529108		ptc-miR3627b	MIMAT0030391	MIMAT0030391_st	Sample Signals Only						•				
20535950		mmu-mir-1892	MI0008320	MI0008320_st	 Table Columns + Sample 	Signals				• •	•				
20533134		ENSG0000212618	ENSG00000212618	ENSG00000212618_st	Include All Decimal Point	5					s .				
20529180		prd-miR-239-3p	MIMAT0030463	MIMAT0030463_st	Row Options										
20514958		tgu-miR-2965	MIMAT0014468	MIMAT0014468_st	Current Rows in Table	Selected Rows	 All Rows 			•	7				
20511655		gma-miR1509b	MIMAT0011201	MIMAT0011201_st	- Column Ontions						-				
20534298		HBII-55	HBII-55	HBII-55_st	Current Columns O All	Columns				· · · ·					
20512653		dps-miR-375-5p	MIMAT0012253	MIMAT0012253_st	Content columns () An	columna				· Jo J					
20523287		cgr-mik-505-3p	MIMA10023980	MIMAT0023980_st	Sample Level Options					- X. 1	9°				
20504857		xtr-mik-16C	MIMAT0003562	MIMATUUU3562_st	Current Samples	Samples			1 1	- 32 4	è .				
2050/154		oan-mik-16c-5p	MIMA10006847	MIMAT0006847_st	Include DABG					- 78 d					
20508315		gga-mik-16c-5p	MIMAT0007739	MIMATUUU/739_st						🛣 🔰	•				
20515169		tgu-mik-16c	MIMAT0014586	MIMATUU14586_st						- 35 4	. .				
20321318	ACD1	aca-mik-100-5p	MIMAT0027607	MIMAT0027607_ct											
20525000	MODI	mmu miP 502 50	MINAT0027607	MIMATOOE201 st		OK	Cancel		• .	100					
20506499		mmu-mik-582-5p	MIMAT0005291	MIMAT0005291_st						·	5.00				
20324020		hea mis 6941	MINAT0023004	MI0022697 ct					• •	- 38 A	6 S (
20537300		nto miR6025d	MINU022007	MIMAT0024722 ct		1				~ 38 P	S	•			•
20524200		asu-miR-375-5n	MIMAT0024732	MIMAT0024732_St						·2					
20521054		ENISG00000238377	ENSG00000238377	ENSG00000238377 et	-				8				•		
20534966		mo-mir-99a	MI0000883	MI0000883 et		0.5				- 55388		•			
20534678		mmu-mir-92a-2	MI0000580	MI0000580 st	-				• •	· 3400	8				
20534115		ENSG0000253091	ENSG0000253091	ENSG0000253091 st							0000				
20501929		gga-miR-200a-5p	MIMAT0026533	MIMAT0026533_st		0				. 200	10				
9 Find in Table	∧ ∨ Coun	t: 36222 Selected: 0	t +		Show Filtered Only	-11.31	-6.96	-4.29	-2.64 -1.62	2 1 Fold Chan	1.62 ige	2.64	4.29	6.96	11.31

• A file browser window pops up. Give an Export File name: GSE123377.txt. Click "Save". The probe summarization TXT file will be saved onto your local drive.

4. Import the TXT file in BRB-ArrayTools

You may modify the title names of the array signals in the TXT file with a text editor before importing it in BRB-ArrayTools. For example, you may shorten the title name "GSM3502388_4877_1_miRNA-4_0_.rma-dabg.chp Signal" to be "GSM3502388".

Start Excel and import the TXT file with the BRB-ArrayTools "General format importer".

E HOME INSERT PAGE	AYOUT I	FORMULAS DATA	REVIEW	/IEW	DEVELOPE	R ADD-IN	S												Qian X
irayTools *																			
Getting started																			
Import data		Data import wizard																	
<u>O</u> raphics		General format impor	ter																
Endrysis wizaru		NCBI GEO Data Sets (GDS)	_									-			_			
<u>Clustering</u>		NCBI GEO Series (GSE	.)	н		J	K	L	M	N	0	Р	Q	R	S	T	U	V	W
Class comparison		Affymetrix Gene ST A	rray Importer																
Sumiyal analysis		RNA-Seq count data I	Importer																
Ouantitative trait analysis																			
Guanatauve and dildiysis																			
Time course analysis																			
Methylation tools																			
New Seq count data analysis	· .																		
Be-niter, normalize and subset the d	11.0																		
Plugins																			
Utilities	-																		
Support	· -																		
Manuais	•																		
License and version information	•																		

Select "Arrays are saved in a horizontally aligned file." And click "Next".

This General Format Importer will allow you to specify the file typ your files. Please note your files must be tab-delimited text files. Data Import Wizard and other importers.	e, data type, and structure of For other file types, please use the
File type:	
$\ensuremath{^{\frown}}$ Arrays are saved in separate files stored in one folder.	Explain More
Arrays are saved in a horizontally aligned file.	

Browse for the TXT file and click "Next".

General I	Format Importer	×
The e	xpression data for all arrays should be stored in one horizontally aligned file.	
File o	C:[Datasets]Users](GSE123377.bd	Browse
	Reset Back Next Cancel	

Select "Single-channel data." And click "Next".

General Format Impor	ter		×
C Dual-channe	els data.	Single-channel data.	Explain More
	Back	Next Cancel	

Select "Affymetrix probeset-summary data." And click "Next".

Non-Affymet	ix single-channel data.
Affymetrix p	obeset-summary data.
ChipType:	Other (other)
🔽 T would lik	a to use my own gene identifiers file rather the one from Bioconductor

Select correct lines and columns for the fields. And click "Next".

ID	Target Gene S	ymbols	Trans	script ID(Arra	ay Design)	Accession	Probe Set	▼ sN
irst d	ata line: ——							
2050	0000	cel-let-	7-5p	MIMATO	000001 MIMA	T0000001_st	11.20606	-
ile fo	rmat: log trans	formed a	Affyme	etrix signal	value —			
Probe	Set ID:	Col 5: Probe Set Name						
Data f begins	or first array in column:	Col 6	GSM3	502388_487	7_1_miRNA	4_0rma-dabg.	chp Signal	•
Data for second array Col 7: GSM3502389_4877_2_miRNA-4_0rma-dabg.chp Signal							•	
- Dat	a columns for t	he FIRS	T arra	y:				_
Log	ged Signal:	Col 6	: GSM3	3502388_48	77_1_miRNA	-4_0rma-dabg	.chp Signal	-
Dete Call(ction optional):						2	-
							Evolain More	1
							Explain More	

Confirm by clicking "Yes".

Microsoft Excel	
Your data set has 27 arrays.	
Is this correct?	
Yes N	0

Give a BRB-ArrayTools project folder name or use the default one and click "Next".

Project location:	C:\Datasets\Users	
Project folder:	GSE123377Proj	
Project name:	Project.xlsx	

Select correct columns for different gene identifiers and click "Next".

lease select your	Gene identifiers file:			Browse line #:
lease select the a	available gene identifiers: —			mie #.
nique ID (Well or oot ID, etc.):	Col 5: Probe Set Name	Ŧ	EntrezId:	•
one ID (IMAGE ATCC ID, etc.):		٠	Gene Name ,Title or Description:	Col 1: ID
niGene Cluster ID:		•	GenBank Accession:	Col 4: Accession
ene Symbol:	Col 2: Target Gene Symbols	•	Map Location:	•
isembl ID:		•	microRNA ID:	Col 3: Transcript ID(Array Desi
Annotate the pr	niect with these gene ids, inst	ead of	fusing or	

Check "I do not have an experiment descriptor file. Please create a template with just array Ids." And click "Next".

ita Impo	t Wizard	
I do	not have an experiment descriptor file. Please create a template with just array ids.	
File:	Iment descriptor file: C:\Datasels\Users\ExpDescFile.xls	Browse

Confirm by clicking "OK".

Intensity Filter:	Detection Call
C EXCLUDE the spot if the intensity i below the minimum.	S EXCLUDE the probeset if the Detection Call contains:
THRESHOLD the intensity at the iminimum value if the intensity is below the minimum.	Any one of the following values (comma-separated): A,M,P,No Ca
Intensity minimum: 10	-

Confirm by clicking "OK".

Microsoft Excel	×
The 'Percent Absent' gene filter will be turned off because y contain the Detection Call data.	our project does not
	ОК

Confirm by clicking "OK".



Confirm by clicking "OK".

BRB-Array	Tools	×
1	There are 27 arrays in this project and the first 20 arra the Filtered log ratio/intensity worksheet automatical	ays are shown in ly.
		ОК

Now the BRB-ArrayTools project is collated successfully.

You may edit the worksheet "Experiment descriptors" to add sample information. You can run all the analysis tools in BRB-ArrayTools.

											7/4/*						
HILE HOME INSERI PAGE LAYOUT FORMULAS DA	IA REVIE	W VIEW	DEVELO	PER ADD	INS												Qian X *
	- 1%	🔐 Wrap Te	xt	General	*			Normal	Ba	ıd	Good			× 🖬	∑ AutoSum	'` ≙ ▼	A
Paste Paste		Marga 8	Contor x	¢ _ 0/ .	e.0 .00	Conditional	Format as	Neutral	Ca	lculation	Check (Cell	Insert	Delete Forma	👽 Fill *	Sort &	Find &
Format Painter		merge o	Center *	3 * 70 3	00.00	Formatting	• Table •							* *	Clear *	Filter -	Select *
Clipboard S Font S	Numb	er 5				Styles					Cells		Editing	^			
A1 • : $\times \checkmark f_x$ Experiment Names																	~
A	В	С	D	E	F	G	н	1	J	К	L	м	N	0	Р	Q	R
1 Experiment Names	Class																
2 GSM3502388_4877_1_miRNA-4_0rma-dabg.chp Signal	СР																
3 GSM3502389_4877_2_miRNA-4_0rma-dabg.chp Signal	CP																
4 GSM3502390_4877_3_miRNA-4_0rma-dabg.chp Signal	CP																
5 GSM3502391_4877_4_miRNA-4_0rma-dabg.chp Signal	CP																
6 GSM3502392_4877_5_miRNA-4_0rma-dabg.chp Signal	CP																
7 GSM3502393_4877_16_miRNA-4_0rma-dabg.chp Signal	Normal																
8 GSM3502394_4877_17_miRNA-4_0rma-dabg.chp Signal	Normal																
9 GSM3502395_4877_18_miRNA-4_0rma-dabg.chp Signal	Normal																
10 GSM3502396_4877_19_miRNA-4_0rma-dabg.chp Signal	Normal																
11 GSM3502397_4877_20_miRNA-4_0rma-dabg.chp Signal	Normal																
12 GSM3502398_4877_21_miRNA-4_0rma-dabg.chp Signal	Normal																
13 GSM3502399_4877_22_miRNA-4_0rma-dabg.chp Signal	Normal																
14 GSM3502400_4877_23_miRNA-4_0rma-dabg.chp Signal	Normal																
15 GSM3502401_4877_24_miRNA-4_0rma-dabg.chp Signal	Normal																
16 GSM3502402_4877_32_miRNA-4_0rma-dabg.chp Signal	Normal																
17 GSM3502403_4877_33_miRNA-4_0rma-dabg.chp Signal	Normal																
18 GSM3502404_4877_6_miRNA-4_0rma-dabg.chp Signal	PDAC																
19 GSM3502405_4877_7_miRNA-4_0rma-dabg.chp Signal	PDAC																
20 GSM3502406_4877_8_miRNA-4_0rma-dabg.chp Signal	PDAC																
21 GSM3502407_4877_9_miRNA-4_0rma-dabg.chp Signal	PDAC																
22 GSM3502408_4877_10_miRNA-4_0rma-dabg.chp Signal	PDAC																
23 GSM3502409_4877_35_1_miRNA-4_0rma-dabg.chp Signal	PDAC																
24 GSM3502410_4877_11_miRNA-4_0rma-dabg.chp Signal	PAT																
25 GSM3502411_4877_12_miRNA-4_0rma-dabg.chp Signal	PAT																
26 GSM3502412_4877_13_miRNA-4_0rma-dabg.chp Signal	PAT																
27 GSM3502413_4877_14_miRNA-4_0rma-dabg.chp Signal	PAT																
28 GSM3502414_4877_34_miRNA-4_0rma-dabg.chp Signal	PAT																
29																	
30																	
Experiment descriptors Gene annotations	Filtered Io	og intensity	Gene	identifiers	Sheet2	Sheet3	(+)							4			Þ
READY 🛗																	+ 100%

5. Examples of other array types

Other types of arrays data can be imported into BRB-ArrayTools by following the same steps given above. Different array types may have different annotation items for output. In the following examples, popular gene information columns are selected. You may select different columns for your needs.

5.1. Affymetrix Mouse Gene 1.0 ST Array

NCBI GEO GSE54773: Affymetrix Mouse Gene 1.0 ST Array. You can follow the above steps to convert CEL files to a probe summarization TXT file and import it in BRB-ArrayTools. You may pay attention to the choices in the following steps:



Click "Customize Annotations" to select information columns for output in TAC.

Click "Show/Hide Columns" to select output columns for the TXT file in TAC.

applied biosyste	ms					Analysis	2.tacx ⑦ ① _ 5 ×
Sample QC View	w Summary Vi	ew Gen	e View				080
Gene Table St	Immary Sample	Table					Scatter Plot Volcano Plot Hierarchical Clustering WikiPathways Sample Signals Chr Summary
Comparison: 1	/s 2 🔹 🕏 What is	the differen	tial expression between 1 an	d 2?			Color By Strand Up: Vert Color By Strand Up
Apply View 💌 Fi	Iters Show/Hid	le Columns	Export Add Column	Customize Anno	tations Ger	miRNA Interaction	🔹 🛛 er te stander er der te sen geld mit hat give ster som dette stangebilde ber sin er som sin for segnitette var sin for setter som give deter som dat werde 🖄
10		C					2 🛛 dagt "httels von andersetterken mit in skalte "sätttelsendenter spectralisen "in missa varits villged "statistik par varitstratersen anderska skonvert
U	Public Gene IDs	Symbol				UniGene ID	3 with an et al astropacity of a structure of a structure of the struct
count: 28944	10100000000	1111.15	110		6.469.6	Show/Hide Columns	an a
10598389	NM_001290792	Wdr45	WD repeat domain 45	chrX	54636		
10572727	NM_001166645	Zfp882	zinc finger protein 882	chr8	382019	1 Avg (log2)	
10490794	NM_008862	Pkia	protein kinase inhibitor,	chr3	18767	2 Avg (log2)	personal and the second s
10583080	NM_032006	Mmp1a	matrix metallopeptidase	chr9	83995	1 Standard Deviation	Line and a state state state state state and a state and a state and a state and a state of the
10347497	NM_009518	Wnt10a	wingless-type MMTV inte	chr1	22409	2 Standard Deviation	
10577903	NM_029037	Pomk	protein-O-mannose kinase	chr8	74653	Eold Change	
10526410	NM_013560	Hspb1	heat shock protein 1	chr5	15507	- Port Change	a decision of the second statements and the second statement of the second statement of the second statement of
10436849	NM_010508	Ifnar1; A	interferon (alpha and bet	chr16	15975		
10406934	XR_873890	Gm5454	predicted gene 5454	chr13	432800	DR P-Val	
10363901				chr10		Public Gene IDs	A second statement with the second statements of the second statements of the second statements of the second statement of the
10452030	NM_025836	Plin3	perilipin 3	chr17	66905	Gene Symbol	A STREET, A STREET, AND A STREET, A STRE
10575799	NM_172285	Plcg2	phospholipase C, gamm	chr8	234779	 Description 	
10355327	NM_007525	Bard1	BRCA1 associated RING	chr1	12021	✓ Chromosome	Sangadina Algoratin (17 Adarini, 18, 19 - University) Planta, kani keta
10578193	NM_026067	Eri1	exoribonuclease 1	chr8	67276	Strand	address of the sector of the
10543369	NM_001252105	Cadps2	Ca2+-dependent activat	chr6	320405		
10527528	ENSMUST0000	Gm8494	predicted gene 8494 [So	chr5			
10574104	NM_001033207	NIrc5	NLR family, CARD domai	chr8	434341	L Stop	participation of the second
10555303	NM_027629	Pgm2I1	phosphoglucomutase 2-I	chr7	70974	 Entrez ID 	
10466270				chr19		 UniGene ID 	
10381603	NM_020510	Fzd2	frizzled homolog 2 (Dros	chr11	57265	Cytoband	The value of the state of the s
10497335	XM_006530087	1810022	RIKEN cDNA 1810022K0	chr3	69126		2. A fraction
10409021	NM_016785	Tpmt	thiopurine methyltransfer	chr13	22017		
10514658	OTTMUST0000	Ccdc50-ps	Ccdc50 retrotransposed	chr4	277707		OK Cancel
10409190	NM_025495	Cenpp	centromere protein P	chr13	66336		
10608711				chrUn			The second second second set of the second
10496405	XM_006502554	LOC1026	uncharacterized LOC1026	chr3	102634333		a strain haden meine in de heine in heine in service in der bestehen bestehen eine stehen bei der Bestehen der Bestehen in bestehen bestehen bereichen b
10468992	NM_001177843	Frmd4a	FERM domain containing	chr2	209630	Mm.37932	
10386427	NM_001271356	Flcn	folliculin	chr11	216805	Mm.339640	0
10600973	NM_008784	lgbp1	immunoglobulin (CD79A	chrX	18518	Mm.7454	
10545644	NM_001083922	Wbp1	WW domain binding pro	chr6	22377	Mm.1109	ار المولوليون أيش يواريها بطايعا بطبقه فينطيب بالمنام بمنينا فيتعلي مهم الاتف سيموليا ليستبين فطعيا بسبك بلسباست
10464877	NM_133803	Dpp3	dipeptidylpeptidase 3	chr19	75221	Mm.234769	
10411782	NM_001024955	Pik3r1	phosphatidylinositol 3-ki	chr13	18708	Mm.259333	
P Find in Table	∧ ∨ Count: 2	8944 Select	ed: 1 🕇 🖡			Show Filtered Only	유지 유지 유지 유지 유지 유지 유지 유지 유지 유지 유지 유지 유지 유

Assign different gene identifiers in BRB-ArrayTools.

lease select your	Gene identifiers file				
le:	Gene menancis mer		<u>-</u>	Browse Header	Y
lease select the a	available gene identifiers:				
nique ID (Well or oot ID, etc.):	Col 1: ID	Ŧ	EntrezId:	Col 6: Entrez ID	•
one ID (IMAGE ATCC ID, etc.):		•	Gene Name ,Title or Description:	Col 4: Description	•
niGene Cluster ID:	Col 7: UniGene ID	•	GenBank Accession:	Col 2: Public Gene IDs	•
ene Symbol:	Col 3: Gene Symbol	•	Map Location:	Col 5: Chromosome	•
sembl ID:		•	microRNA ID:		•
Annetate the pr	olect with these gone ide	Instead of	fusing		

5.2. Affymetrix HT HG-U133+ PM Array

NCBI GEO GSE100833: Affymetrix HT HG-U133+ PM Array Plate. You can follow the above steps to convert CEL files to a probe summarization TXT file and import it in BRB-ArrayTools. You may pay attention to the choices in the following steps:

applied biosyste	ms							Analysis	4.tacx			(? (Ì _	e ×
Sample QC View	v Su	nmary \	/iew G	iene View										
Gene Table Su	immary	Sampl	e Table						Scatter Plot Volcano Plot Hierarchic	al Clustering Wik	iPathways Sample S	ignals Chr Summary		
Comparison: 1 v	s 2 💌	🕻 What	is the diffe	rential exp	ression be	tween 1 ar	nd 2?		Color Settings Show Filtered Only					≡
Apply View 🔻 Fi	Iters 💌	Show/H	ide Colum	ns Export	• Add	Column	Customize Annot	ations Gene List 💌 miRNA Interaction	14	1 A	vg (log2) vs 2 Avg (lo	g2)		
10		~ .			500 D	-		Customize Annotations	×					
10	(log2)	(log2)	Change		val	Symbol	Descri	Annotation File:						6
count: 38			Y					HT_HG-U133_Plus_PM.na36.annot.csv ~		_				
230180_PM_at	5.6	6.76	-2.23	6.16E-15	2.61E-13	DDX17	DEAD (Asp-GI							
232737_PM_s_at	4.34	5.39	-2.06	3.26E-05	0.0002	ENPP3	ectonucleotide	Annotations Assignment:						
240480_PM_at	2.98	4.24	-2.4	0.0003	0.0012	TINAG	tubulointerstit	Iop Assignment All Assignments				1 1		
241547_PM_at	3.22	4.91	-3.23	0.0008	0.0027	A1CF	APOBEC1 com	Select Annotation Column(s) to Add:				•		
207178_PM_s_at	3.94	5.03	-2.13	0.0012	0.0038	FRK	fyn-related Sro	GeneChin Array					· .	
236118_PM_at	5.57	6.61	-2.06	0.0012	0.0038	GATA6-A.	GATA6 antisen	Species Scientific Name				•••		
239065_PM_at	3.15	4.88	-3.31	0.0030	0.0083	DNAJC22	DnaJ (Hsp40)	Annotation Date						
223426_PM_s_at	5.43	6.63	-2.29	0.0031	0.0086	EPB41L4B	erythrocyte me	Sequence Type						
219850_PM_s_at	4.52	5.65	-2.18	0.0032	0.0088	EHF	ets homologo	☐ Sequence Source ☐ Transcrint ID(Array Design) (Already Adde	x1)					
229337_PM_at	5.24	6.45	-2.31	0.0038	0.0101	U5P2	ubiquitin speci	Target Description	(d)					
237530_PM_at	2.44	3.77	-2.51	0.0040	0.0108			Representative Public ID			1.12			
204007_PM_at	11.92	9.89	4.09	0.0065	0.0163	FCGR3B	Fc fragment of	Archival UniGene Cluster			1. A 1.			
210146_PM_x_at	9.96	8.91	2.07	0.0067	0.0167	LILKB2	leukocyte imm	✓ UniGene ID					. *	
2362/9_PM_at	9.4	10.56	-2.23	0.0074	0.0182			Alignments						
208596_PM_s_at	3.37	7.01	-3.11	0.0079	0.0195	UCTIAN:	UDP glucurone	Gene Title				1		
221305_PM_s_at	3.09	5.51	-3.53	0.0079	0.0194	CYD2D6	. UDP glucurone	Gene Symbol (Already Added)						
200735_PM_at	5.01	5.51	-2.05	0.0099	0.0255	CTP2D0	cytochrome Pa	Chromosomal Location		1. A.		· .		
230575_PM_at	0.24	7.65	-2.22	0.0101	0.0239	SUN2 \$100A12	Serum/glucoco	Chigene Cluster Type Fisembl		1.00				
203005_PM_at	9.54	5.26	-2.22	0.0103	0.0246	STUDATZ	stinderin	✓ Entrez Gene						
1332307_PM_a	3.33	9.00	-3.32	0.0122	0.0280	SCIN EGER2	fibroblact grou	SwissProt			A 8			
204575_PM_5_8L	6.05	0.09	-2.07	0.0147	0.0239	LINEAA	hepatocide pu	EC			đ			
215125 DM c at	6.01	7.24	-2.00	0.0152	0.0320	LIGT1A1:	LIDR alugurone	OMIM RefEeg Protein ID						
207781 PM s at	4.06	5.4	-2.53	0.0152	0.0359	7NF711	zinc finger pro	RefSeg Transcript ID						
209949 PM at	10.45	9.28	2.55	0.0103	0.0413	NCE2	neutronhil orte	FlyBase						
207126 PM x at	5.66	6.96	-2.47	0.07312	0.0413	LIGT1A1:	LIDP alucurone	AGI						
205568 PM at	11 23	87	5.76	0.0212	0.0507	AOP9	aquanorin 9	WormBase		1	•			
204532 PM x at	5.25	7.01	-3.4	0.0251	0.0520	LIGT1A1	LIDP glugurone	RGD Name	~					
220435 PM at	2.92	5.13	-4.61	0.0267	0.0548	SIC30A10	solute carrier f	Check/Lincheck All						
204006 PM s at	9.23	8.01	2 33	0.0267	0.0549	ECGR3A	Ec fragment of							
213558 PM at	4.94	6.51	-2.97	0.0287	0.0583	PCLO	piccolo presvn		OK Cancel					
230238 PM at	5.18	6.37	-2.28	0.0293	0.0594	SOWAHA	sosondowah a		OK Cancer					
									1 2.3 3.6	4.9 6.2	7.5	8.8 10.1 11	.4 12.7	14
P Find in Table	^ '	Count	: 38 Selecte	nd: 0 🕇 🖡				Show Filtered Only			2 Avg (log2)			

Click "Customize Annotations" to select information columns for output in TAC.

Click "Show/Hide Columns" to select output columns for the TXT file in TAC.

applied biosyst	ems						Analysis	_4.tacx								0	- 6	ē ×
Sample QC Vie	w Summ	ary View Gene View																ner
Gene Table	ummary S	ample Table						Scatter Plot	Volcano Plot	Hierarchical C	lustering V	VikiPathway:	s Sample S	ignals Chr	Summary			
Comparison: 1	vs 2 🔹 🕏	What is the differential expre	ession betwe	en 1 and	2?			Color Settings Show Filtered Only									≡	
Apply View 💌	ilters 💌 Sh	ow/Hide Columns Export	Add Colu	umn C	ustomize Annota	ions Ger	ne List 💌 miRNA Interaction	14			. 1	1 Avg (log2)	vs 2 Avg (lo	g2)				
ID count: 54653	Gene Symbol	Description	Transcript ID(Array Design)	Chro	Representative Public ID	UniGene ID	Chromosomal Location Ensembl Gene	12.7									ALC OF	,
233570_PM_at			Hs.69429.0	chr9	AJ011377	Hs.6607	Show/Hide Columns			×								
236416_PM_at	ARHGEF7	Rho guanine nucleotide	Hs.13625.0	chr13	AI681617		1 4 4 2)											
220255_PM_at	FANCE	Fanconi anemia comple	g113454	chr6	NM_021922	Hs.3020	1 Avg (log2)											
233318_PM_at			Hs.30697	chr12	BC003386		2 Avg (log2)							•			•	_
230685_PM_at	LINC01184	long intergenic non-prot	Hs.16506		AI799695	Hs.3406	1 Standard Deviation							•	. All	•		
1552669_PM_at	PPP1R3B	protein phosphatase 1, r	Hs2.2832	chr8	NM_024607	Hs.4585	2 Standard Deviation								10 M	•		
214621_PM_at	GYS2	glycogen synthase 2 (liver)	Hs.82614.0	chr12	S70004	Hs.8261	Fold Change								100	•	~	
241468_PM_at	LOC1053	uncharacterized LOC1053	Hs.20160.0	chr17	AA846862		P-val									••••	~	
233785_PM_at	ADAMTS9	ADAM metallopeptidase	Hs.12685	chr3	AB037733	Hs.6560	EDB P-val											
240745_PM_at			Hs.24887	chr10	AW242763		Cons Sumbol						1. 10		e •	۰.	•	
239630_PM_at			Hs.27097	chr12	BF516583		Cene Symbol						:		2.	•		_
1554920_PM_at	SCEL	sciellin	Hs2.1151	chr13	BC020726		Description						1.38	C	1.		•	
230967_PM_s_at	LOC1019	lipopolysaccharide-induc	Hs.12966	chr16	BF433061		 Transcript ID(Array Design) 					· · ·	1990 C		• •	·° ·.		
1558787_PM_a_	TSPAN3	tetraspanin 3	Hs2.3028	chr15	BF962082	Hs.7448	Chromosome								•	ъ°		
213728_PM_at	LAMP1	lysosomal-associated me	Hs.15010	chr13	AI248598	Hs.4944	Strand				• •	. 199	100	ູ້ຈໍ້ວ່				
216846_PM_at	IGLJ3	immunoglobulin lambda	Hs.32678	chr22	AF234254	Hs.7287	Start			•	•••	1000	6 S.Y S	8 .		·		
217554_PM_at			Hs.97109.0	chr1	AV719355		Stop							•				
211076_PM_x_at	ATN1	atrophin 1	g393697	chr12	Z22814	Hs.1437	Representative Public ID			•			380	• •				
207909_PM_x_at	DAZ1	deleted in azoospermia	g1045307	chrY	U21663	Hs.5228					1.22		•					
1555292_PM_at	STRIP2	striatin interacting protei	Hs2.2680	chr7	BC019064	Hs.4899	Conigene ID				100							
234283_PM_at			Hs.30633	chr22	AL110165	Hs.6084	Chromosomal Location			· 3	1 (S			•				
209956_PM_s_at	CAMK2B	calcium/calmodulin-dep	g4096107	chr7	U23460	Hs.3518	Second Ensembl					8.						_
226554_PM_at	ZBTB7A	zinc finger and BTB dom	Hs.10464	chr19	AW445134	Hs.5913	 Entrez Gene 			1000								
1556328_PM_at	PCDH10	protocadherin 10	Hs2.2536	chr4	BC042378	Hs.1928												
242699_PM_at			Hs.12352	chr8	AW384895						· · · · ·							
218318_PM_s_at	NLK	nemo-like kinase	g7706444	chr17	NM_016231	Hs.2087				- 30 000	,	•						
1560818_PM_at	LINC00944	long intergenic non-prot	Hs2.1435	chr12	AI138766	Hs.5346		C	K Cancel	. °°° '	¶•							
235232_PM_at	GMEB1	glucocorticoid modulato	Hs.72031.0	chr1	AI061377		L		Concer		Ť							
223744_PM_s_at	SIAE	sialic acid acetylesterase	g102423	chr11	AF300796	Hs.1005	5 chr11q24 ENSG0 54	2.3	-									-
211264_PM_at	GAD2	glutamate decarboxylase a	g182933	chr10	M81882	Hs.2318.	chr10p11.23 ENSG0 25											
1556797_PM_at	RNF144	RNF144A antisense RNA 1	Hs2.2887	chr2	AK090737	Hs.5590.	chr2p25.2 ENSG0 38											
<	Thirdor	1 F 1 70F		1.40	0000000		· · · · · · · · · · · · · · · · · · ·	1		26	10						10.7	
9 Find in Tabl	p ^ V	Count: 54653 Selected: 0 🕇	ŧ				Show Filtered Only		2.3	3.6	4.9	5.2 2 Av	7.5 g (log2)	8.8	10.1	11.4	12.7	14

Assign different gene identifiers in BRB-ArrayTools.

ease select your	Gene identifiers file:	<u>~</u>	Browse line #:
ease select the a	available gene identifiers:		
ique ID (Well or ot ID, etc.):	Col 1: ID	EntrezId:	Col 10: Entrez Gene
one ID (IMAGE ATCC ID, etc.):	·	Gene Name ,Title or Description:	Col 3: Description
iGene Cluster ID:	Col 7: UniGene ID	GenBank Accession:	Col 6: Representative Public ID
ne Symbol:	Col 2: Gene Symbol 🔹	Map Location:	Col 8: Chromosomal Location 💌
sembl ID:	Col 9: Ensembl	microRNA ID:	Col 4: Transcript ID(Array Desi
Annotate the pr	oject with these gene ids, instead o	f using Organism	s: Human 💌