

## IPA Upload and Analyze Your Data Tutorial

### **Overview:**

Use IPA to upload and analyze your 'omics data.

### Scenario:

You have RNA-seq, microarray, miRNA, proteomic, genomic, SNP, or metabolic data that you want to analyze with IPA. You need to know the rules, restrictions, and best practices for preparing the data for being uploaded into IPA and how to start an analysis

## Tasks:

- Format the data in a way that IPA can upload.
- Set Upload options and assign the ID and Observation columns (measurement values).
- Set analyisis parameters: If the datasets are large, adjust expression cutoff values(s) to restrict their size.
- Run analysis

To understand more about interpreting the results of the analysis, see Analysis Results Tutorial

1) Prepare your data for uploading (as spreadsheet format).

a) If your data is not in an Excel spreadsheet or tab delimited text file, transfer your data into one. IPA can also upload CuffDiff formatted files directly. See <u>Cuffdiff file import</u>.

b) If using Excel, perform basic calculations in the spreadsheet if necessary. Example: compute average ratio for experiment samples vs. control samples and p-values for replicates. Ideally, upload fold changes, log ratios, or log fold changes (which are the same thing as log ratios).

- c) Make sure there is only one header row. (IPA can be set to ignore the first row when doing the import.)
- d) Move the molecule IDs to the first column. If ID's are in the first column, IPA will scan down approximately 100 rows to guess the ID type(s) in the column. See <u>Data Upload definitions</u> for the types of IDs that are recognized.

e) IPA uses text in the header row to guess the column types. For example if it detects the text "fold change" it will assign it to Expr Fold Change.

$\diamond$	A	Observation	1 C	D	E	Header Row
1	Probe Set ID	log2Ratio	p-value	FCAbsolute	[Normal](normalized)	[Tumor](normalized)
2	AFFX-BioB-M_at	-1.01846921	3.11E-04	2.025768	1.0998152	0.08134599
3	AFFX-BioB-3_at	-1.06497304	2.92E-04	2.092131	1.2452649	0.18029186
4	AFFY-RioDn-5 at	-1 01386279	9 86F-04	2 019311	1 0488484	0.034985613
	Molecule ID	Exp1	Exp2		Calculations	3



f) IPA allows a maximum of 20 observations per imported file. An observation is one "comparison" between an experiment and control. For example if you have a dataset with three time points ratio'ed to a 0 time point, and each with a set of fold changes and a set of p-values for each time point, that would be three observations. See <u>Data Upload definitions</u> for more on this topic. If your data has more than 20 observations, then reduce the data to 20 observations in the spreadsheet or pick and choose up to 20 to import during the upload process.

g) IPA allows up to 8 "measurement values" per observations. For example: fold change, p-values, and the average intensity of expression would equate to three measurements. Any more than 8 will be excluded during the upload process.

#### 2) Launch IPA

3) Upload your data into IPA.

- a) Select File>Upload Dataset... from the menu.
- b) Select the dataset that you modified in Step 1.
- c) Click Open.

Your data appears in the "Dataset Upload - ..." window.

Flexit	ole Format	More Info			
: 💿 Ye	5 O No				
Affym	netrix	Specify the id			
experiments: Not s	specified/applicable	Select relevar			
s to specify the colu	e column names that contain identifiers and observation				
at Summany (0)					
et summary (0) \					
INFER OBSERVATIO	More Info				
Ignore 👻	Ignore 👻				
logFC	P.Value				
	:      Ye: Affyrr experiments: Not s s to specify the colur et Summary (0) \ INFER OBSERVATIO Ignore  Infer OBSERVATIO Ignore  Infer OBSERVATIO	Affymetrix       experiments:     Not specified/applicable       s to specify the column names that conta       et Summary (0) \       INFER OBSERVATIONS       More Info       Ignore       Ignore	: • • Yes O No Affymetrix • experiments: Not specified/applicable • s to specify the column names that contain identifiers and obse et Summary (0) \ INFER OBSERVATIONS More Info Ignore • Ignore • Ignore • Ignore •		

d) Select "Flexible Format" if it's not already selected (The other options are for legacy support).

e) You can try clicking the Infer Observations button to see if IPA can guess the columns in your dataset. If that works well,

skip to the Save step below. Otherwise, click the button again and continue with the next step.

f) If you have one header row in your dataset, make sure the Contains Column Header is set to Yes.

g) Select the appropriate Identifier Type(s). Note: IPA will guess the appropriate ID types from the IDs in the first column, but

it is always a good idea to double check. Do not select all identifier types at once. This can lead to mis-mapping.

g) If the data came from a commonly known array platform, select it from the pull down menu, otherwise leave the value "Non

Specified/applicable". Setting the array platform properly here will automatically select the same platform as "reference" in



the create core analysis page and will improve the accuracy of the p-value calculations used in analysis.

h) Select the appropriate ID and Observation labels from the dropdown menus for each of the columns.

i. Set the column with your molecule identities to **ID**. Check the Dataset Summary tab to view the breakdown of mapped vs. unmapped molecules from your uploaded data.

ii. Set the column with the first observation value to Observation 1

Raw Data (54676) \ Dataset Summary (45520) \ EDIT OBSERVATION NAMES INFER OBSERVATIONS More Info ID Observation 1 Observation 1 Exp Log Ratio Exp p-value Exp Log Ratio . P.Value ID 211372\_s\_at Exp p-value 1.21999999999999998-10 205403\_at Exp False Discovery Rate (q-value) 1.63E-9 204787\_at Exp Intensity/RPKM/FPKM/Counts 1.16E-8 4 5 204900\_x\_at Exp Other 1.48F-8 6 206028\_s\_at Variant Loss/Gain 3.19000000000001E-8 236901\_at 3.640000000000002E-8 Variant ACMG Classification 7 215049\_x\_at 9.2399999999999994E-8 8 Phospho Ratio 9 203645\_s\_at 1.340000000000001E-7 Phospho Fold Change 10 222866\_s\_at 4.0699999999999998E-7 Phospho Log Ratio 11 220496\_at 4.099999999999999999E-7 Phospho p-value 12 226517\_at 4.27E-7 Phospho False Discovery Rate (q-valu - 4.37E-7 13 223660\_at 14 224856 at 2.52835454 5.2E-7 0.95534076000000001 5.5499999999999998E-7 15 202441\_at 2.5498113199999999 16 224840\_at 5.57000000000002E-7

A second pull down menu appears.

iii. Set the new dropdown menu to the measurement type (i.e. Expr Fold Change, Expr p-value, etc.).

iv. Repeat the last two steps for each of the observations if there is more than one.

▲ Note: If there is more than one measurement type per observation, be sure to assign each batch of measurement columns to the **same** observation. For example, if the data set has log ratio and p-value columns like the example shown above, this is NOT two observations-- do NOT assign one column to Observation 1 and the other to Observation 2. This is one observation (Observation 1) with two measurements as shown above.

If there are multiple observations, there are shortcuts available to make assigning the batches easier (as described below).

## Shortcuts:

- Ignore: Right-click a column and select Ignore.
- Repeat Selection: Select a column(s), right-click the selected columns and select Repeat Selection. The



assignments made for the selected columns will be repeated for the columns to the right of the selected ones.

- Header Names -> Observation Name: This selection will take the name found in the column header and use it to label the observation.

- Group In: Select more than one column, right-click the selected columns and select Group In. The selected columns will be grouped as an observation.

ID	
Ignore	
Group in	-
Repeat Selection	
Header Names -> Observation Names	-

v. Leave any extraneous columns set to Ignore.

i) You can customize the names of the column headers.

i. Click the EDIT OBSERVATION NAMES button. This is especially important to do if uploading a

multi-observation dataset like the example below.

ii. Select the name from the drop down menu or type in a new one.

iii. Click OK.



Dataset Upload	- Edit Dataset: Cardiac Differentiation CuffDiff 2.3, 0.01	c	r 🛛 🖂
1. Select File F	Ingenuity File Format A or B		
2. Contains Colu	nn H O Yes 🖲 No		
3. Select Identi	Edit Observation Names	vpe found in the dataset.	
	To label each observation, select an existing name from the pull-down lists,	platform as a reference set for da	ata analy
5. Use the dro	or create a new label by typing directly into the Observation Name field. Then click OK.	ations, select the appropriate ex	, ,
5. Use the dro	Edit Observation Names	ations, select the appropriate ex	pression
Raw Data (4)	Observation Name		
	1. Sample_Cardiac_PrecursorsvsSample_Embryonic_stem_cells		
EDIT OBSERV	2. Sample_CardiomyocytesvsSample_Embryonic_stem_cells		
ID	3. Sample Mesodermal cellsvsSample Embryonic stem cells	Sample_Car 💌 Sample_Ca	ar
		Override 🔻 Exp Log Ra	atio
1 1922		3.2431920	551
2 7037		2.4259510	
3 2354		2.8245352	
4 1821 5 4339		2.8156718 2.5580563	
6 2417		1.9583796	
7 1382		2.7571643	
8 2414		2.1348927	
9 1707		1.8923113	
10 1281		3.5681704	
11 1926	OK CANCEL	2.4556879	429
12 93960	0.011000CT10.00 0.0010	3.9220979	097
13 43596	2.50344891193 5e-05 10.1174	2.5934678	17: 🕶
•			
		REPLACE DATASET CANCEL	HELP

j) Click **Save** at the bottom of the window.

A Save Dataset dialog opens.

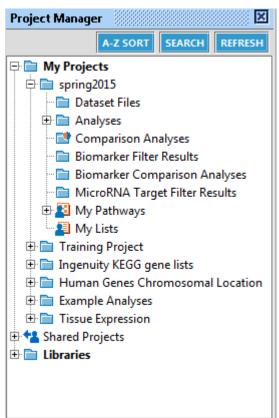
k) Click **New Project** to create a new one.

I) Enter a name for your dataset.

m) Click Save.

IPA saves the data in the Project Manager under the project folder that you selected.





- 4) Start a Core Analysis on your dataset.
  - a) Select File>New>Core Analysis... from the application menu.
    - A Create Core Analysis dialog appears.
  - b) Select the dataset that you saved in step 3 and click Next.
- 5) Set Filters and General Settings for Analysis parameters.

a) Set the **Reference Set** parameter to the molecule set (reference set) that should be viewed as the complete universe when calculating the statistical significance. You may have already set this when uploading the data.

b) You may either leave the Network Analysis parameters, Optional Analysis, Data Source, Confidence, Species, and Tissue & Cell Lines optionson their default settings or provide appropriate filters of your choice.



## 6) Set Cutoffs

It is important to analyze the most significant molecules in your dataset. For example, if you uploaded the data for an entire microarray, you need to set a cutoff so that IPA analyzes only the significantly differentially expressed genes.

- a) Enter values for the **Cutoffs**.
- b) Click the Recalculate button to recalculate the "analysis ready" molecules (i.e. that passed your cutoffs)

Note: We recommend that the number of "analysis ready" molecules should ideally be less than ~3000 in order to focus on the most relevant molecules in your dataset. This will minimize noise in the results. Adjust the Cutoff values more stringently to restrict how many are analyzed. Also, the maximum number of molecules that can be analyzed in one analysis is 8000. You may use the expression value cutoff and also the analysis filter options to keep your "analysis ready" molecule size within the allowable limits.

Set Cutoffs								
Expression Value Type	Cutoff Range F	ocus On						
Exp Fold Change	1.5 -8.3234 to 3.9795 E	Both Up/Downregulated 💌 🖪	ECALCULATE 673 analysis-rea	dy molecules across observations				
eview Dataset Exampl	e Prostate Data Observation: C	Observation 1 (672) 🔻						
Analysis-Ready (672)	Mapped IDs (726) Unmapped ID	s (249) All IDs (975)						
· · ·	Mapped IDs (726) \Unmapped ID						Symbol A1CF - CD69 (p1 of 7)	• «
· · ·			🕅 📐 Symbol	T 🕅 Entrez Gene Name	X Location	Type(s)	Symbol A1CF - CD69 (p1 of 7)	• «
ADD TO MY PATHWAY	ADD TO MY LIST		X Symbol	T X Entrez Gene Name APOBEC1 complemen				
ADD TO MY PATHWAY	ADD TO MY LIST CREATE DATASET			APOBEC1 complement		Type(s)		
ADD TO MY PATHWAY 7 xp Fold Change -1.075 -1.128	ADD TO MY LIST CREATE DATASET		A1CF	APOBEC1 complemen ATP-binding cassette,	tation factor Nucleus	Type(s) other		
ADD TO MY PATHWAY	ADD TO MY LIST CREATE DATASET		A1CF ABCA4	APOBEC1 complemen ATP-binding cassette, ATP-binding cassette,	tation factor Nucleus sub-family Plasma Membrane	Type(s) other transporter		

7) Click the **Run Analysis** button at the bottom right of the window.

#### The Start Analysis dialog appears

- a) Select the same project that contains your dataset.
- b) Set the analysis name. It is often helpful to include the cutoff values used in the analysis in the name.
- c) Click OK.

An Analysis Running dialog appears to tell you that the analysis is in progress. When the analysis is complete an email is sent to



you and the name of the analysis appears in the **Project Manager** window in bold text.

8) Open the analysis and explore the results.

a) When the analysis is finished, open it by double clicking on it in the **Project Manager**. If you analyzed a multi-observation dataset, the analysis will appear in a folder.

## **Core Analysis Results Tutorial**

Note: This article pertains to a Core Analysis based on expression data. For details on the other types, see Core Analysis Types.

## Overview

A **Core Analysis**has multiple ways of helping you find biological insights by automatically using the curated information from the QIAGEN Knowledge Base to put molecules in your 'omics dataset into context. For this tutorial, we will walk through the interpretation of a transcriptomics dataset.

The tabs in the analysis are as follows:

- The Summary tab displays the top results for all the analyses. See <u>Analysis Summary Page</u>.
- Canonical Pathways tab displays your molecules of interest within well-established signaling or metabolic pathways. See <u>Canonical Pathways Overview</u>.
- Upstream Analysis tab predicts which upstream regulators (any molecule that can influence the transcription or expression of another molecule) might be activated or inhibited to explain the expression changes in your dataset. See <u>Upstream Regulator Analysis</u>. It also connects upstream regulators into signaling cascades called <u>Mechanistic Networks</u>. If you have Advanced Analytics in IPA, this tab can also display <u>Causal Networks</u>, or hierarchical networks or regulators controlled by a master regulator.
- Diseases & Functions relates molecules in your dataset to known disease states and biological functions. See <u>Downstream Effects Analysis Tutorial</u>.
- **Regulator Effects** displays hypotheses for how a phenotype, function or disease is regulated in your dataset by activated or inhibited upstream regulators. See <u>Regulator Effects</u>.
- Networks displays non-directional interaction networks of molecules based on known relationships in the QIAGEN Knowledge Base to your molecules of interest. See <u>What are</u> <u>Networks?</u>
- Molecules displays all of the molecules from the analyzed dataset in a table format.
   "Analysis-ready" molecules will have their identifier displayed in bold font, and are all those that passed the filters and cutoffs you may have applied and were evaluated for overlap with



Canonical Pathways, as targets of Upstream Regulators etc in the Core Analysis.

 Analysis Match (requires additional licensing). Displays matches against your own analysis and >50,000 other analyses. See <u>Analysis Match</u>.

Use these tools to find insights that are most relevant to your experimental model or question. Note that this tutorial will draw on only a *subset* of these features, due to length considerations.

# Scenario

You have RNA microarray data from the white blood cells drawn from children with childhood exacerbated asthma compared to the convalescent state. For this example, we will use expression data from PBMCs from PMID <u>19620293</u>. You would like to know if (and how) the data supports involvement of immune/inflammatory responses in acute asthma attack. You can also find novel gene-to-disease associations that can be followed up and confirmed with future wet bench experiments.

# Tasks

- Open the Core Analysis for your microarray data.
- Use the **Summary Page**to quickly identify promising directions for exploration.
- Explore the results for areas of importance to your research:
  - View Canonical Pathways that contain significant numbers of genes from your dataset.
  - Use **MAP** (Molecule Activity Predictor) to predict effects on functional endpoints in a canonical pathway.
  - Overlay **Biomarkers** that identify genes in the TREM1 Signaling Pathway that are used as efficacy indicators for asthma treatments.
  - Explore the impact on downstream diseases and functions.
  - View **Upstream Regulators** to find which and immune regulators may have been activated to drive the observed gene expression changes.
  - Use **Regulator Effects** to see a hypotheses of how activation of certain upstream regulators may lead to outcomes like asthma.
  - Evaluate and explore similarity and differences with other analyses with **Analysis Match** (if available on your IPA license).

# Steps

1) Start IPA

2) In the **Project Manager**, double click My Projects > Example Analyses > Analyses > Childhood exacerbated asthma GSE16032 with FDR GEO2R. The analysis opens in a new window displaying the **Summary**tab:





Top Canonical Pathways         Name         Phagosome Formation         Iron homeostasis signaling pathway         IL-12 Signaling and Production in Macrophages         Th1 and Th2 Activation Pathway         Elcosanoid Signaling         Top Upstream Regulators         V Upstream Regulators         V Upstream Regulators         Immunoglobulin         GATA2         Iipopolysaccharide         Iipopolysaccharide         Iigrastim         Bame         fligrastim         Bame         Iipopolysaccharide         Iipopolysaccharide         Iipopolysaccharide         Iigrastim         BTB 09588         RDR 03785         DFP 00129	1 1 1 1 6 7 8 9 >	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	-value 3.74E-07 5.38E-06 2.30E-05 5.23E-05 8.41E-05 Activated Activated Inhibited	0verlap           14.9 %         18/12           13.0 %         17/13           11.6 %         17/14           10.5 %         18/17           15.9 %         10/6
Phagosome Formation         Iron homeostasis signaling pathway         IL-12 Signaling and Production in Macrophages         Th1 and Th2 Activation Pathway         Elcosanoid Signaling         op Upstream Regulators         V Upstream Regulators         dexamethasone         filgrastim         Immunoglobulin         GATA2         lipopolysaccharide         fluicasone propionate         filgrastim         BTB 09588         RDR 03785		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	3.74E-07 5.38E-06 2.30E-05 5.23E-05 8.41E-05 Activated Activated	14.9 %         18/12           13.0 %         17/13           11.6 %         17/14           10.5 %         18/17           15.9 %         10/6
Iron homeostasis signaling pathway IL-12 Signaling and Production in Macrophages Th1 and Th2 Activation Pathway Elicosanoid Signaling OP Upstream Regulators V Upstream Regulators Name dexamethasone filigrastim Immunoglobulin GATA2 lipopolysaccharide V Causal Network Name fluticasone propionate filigrastim BTB 09588 RDR 03785		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	5.38E-06 2.30E-05 5.23E-05 8.41E-05 Activated Activated	13.0 %         17/13           11.6 %         17/14           10.5 %         18/17           15.9 %         10/6
IL-12 Signaling and Production in Macrophages Th1 and Th2 Activation Pathway Eicosanoid Signaling  op Upstream Regulators  V Upstream Regulators  Name dexamethasone filgrastim Immunoglobulin GATA2 lipopolysaccharide  ilipopolysaccharide  ilipopolysaccharide  iliposite filgrastim BTB 09588 RDR 03785		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	2.30E-05 5.23E-05 8.41E-05 Activated Activated	11.6 % 17/14 10.5 % 18/17 15.9 % 10/6
IL-12 Signaling and Production in Macrophages Th1 and Th2 Activation Pathway Eicosanoid Signaling  op Upstream Regulators  V Upstream Regulators  Name dexamethasone filgrastim Immunoglobulin GATA2 lipopolysaccharide  ilipopolysaccharide  ilipopolysaccharide  iliposite filgrastim BTB 09588 RDR 03785		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	2.30E-05 5.23E-05 8.41E-05 Activated Activated	11.6 % 17/14 10.5 % 18/17 15.9 % 10/6
Th1 and Th2 Activation Pathway         Eicosanoid Signaling         op Upstream Regulators         V       Upstream Regulators         dexamethasone		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	5.23E-05 8.41E-05 Activated Activated	<b>10.5 %</b> 18/17 <b>15.9 %</b> 10/6
Elcosanoid Signaling  op Upstream Regulators  Upstream Regulators  Name dexamethasone filgrastin Immunoglobulin GATA2 Ilipopolysaccharide  1233456  Causal Network  Name fluticasone propionate filgrastin BTB 09588 RDR 03785		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	8.41E-05 Activated Activated	<b>15.9 %</b> 10/6
op Upstream Regulators  V Upstream Regulators  Name dexamethasone filgrastim Immunoglobulin GATA2 lipopolysaccharide  V Causal Network  Name fluticasone propionate filgrastim BTB 09588 RDR 03785		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	Activated Activated	
<ul> <li>✓ Upstream Regulators</li> <li>Name         <ul> <li>dexamethasone</li> <li>filgrastim</li> <li>Immunoglobulin</li> <li>GATA2</li> <li>lipopolysaccharide</li> <li>1111456</li> </ul> </li> <li>✓ Causal Network</li> <li>Name         <ul> <li>fluticasone propionate</li> <li>filgrastim</li> <li>BTB 09588</li> <li>RDR 03785</li> </ul> </li> </ul>		alue 9.81E-25 6.62E-24 5.30E-17 4.21E-16	Activated	Predicted Activation
<ul> <li>✓ Upstream Regulators</li> <li>Name         <ul> <li>dexamethasone</li> <li>filgrastim</li> <li>Immunoglobulin</li> <li>GATA2</li> <li>lipopolysaccharide</li> <li>1111456</li> </ul> </li> <li>✓ Causal Network</li> <li>Name         <ul> <li>fluticasone propionate</li> <li>filgrastim</li> <li>BTB 09588</li> <li>RDR 03785</li> </ul> </li> </ul>		9.81E-25 6.62E-24 5.30E-17 4.21E-16	Activated	Predicted Activation
Name dexamethasone filgrastim Immunoglobulin GATA2 lipopolysaccharide  Causal Network  Name fluticasone propionate filgrastim BTB 09588 RDR 03785		9.81E-25 6.62E-24 5.30E-17 4.21E-16	Activated	Predicted Activation
dexamethasone		9.81E-25 6.62E-24 5.30E-17 4.21E-16	Activated	Predicted Activation
filgrastim	1 1 1 1 6 7 8 9 >	6.62E-24 5.30E-17 4.21E-16	Activated	
Immunoglobulin GATA2 lipopolysaccharide ✓ Causal Network Name fluticasone propionate figrastim BTB 09588 RDR 03785	6789 >	5.30E-17 4.21E-16		
Immunoglobulin GATA2 lipopolysaccharide ✓ Causal Network Name fluticasone propionate figrastim BTB 09588 RDR 03785	6789 >	4.21E-16	Inhibited	
GATA2 lipopolysaccharide	6789 >	4.21E-16	minorea	
lipopolysaccharide ✓ Causal Network Name fluticasone propionate filgrastim BTB 09588 RDR 03785	6789 >			
✓ Causal Network       Name       fluticasone propionate       filgrastim       BTB 09588       RDR 03785		8.66E-16		
✓ Causal Network     Name     fluticasone propionate     filgrastim     BTB 09588     RDR 03785			Activated	
Name fluticasone propionate fligrastim BTB 09588 RDR 03785	B			
fluticasone propionate filgrastim BTB 09588 RDR 03785	19 > ==			
fluticasone propionate filgrastim BTB 09588 RDR 03785				
filgrastim BTB 09588 RDR 03785		alue		Predicted Activation
BTB 09588 RDR 03785	•	3.63E-25	Activated	
RDR 03785	•	6.62E-24	Activated	
RDR 03785		4.53E-23	Activated	
		4.53E-23	Activated	
		4.53E-23	Activated	
DFP 00129		4.332-23	Activated	
1 2 3 4 5 6				
op Diseases and Bio Functions $\sim$ Diseases and Disorders				
				<i></i>
Name			llue range	# Molecules
Inflammatory Response		AND ALL DOLL		278
Organismal Injury and Abnormalities			5.00E 05 1.55E 10	486
Connective Tissue Disorders			2.93E-05 - 1.72E-16	154
Inflammatory Disease		46-01-01-0	2.93E-05 - 1.72E-16	213
Skeletal and Muscular Disorders				149
Skeletal and Muscular Disorders		123456789		149
		123456789 >		
imes Molecular and Cellular Functions				
Marra				# Mala auto a
Name			lue range	# Molecules
Cell-To-Cell Signaling and Interaction		12 Tan		194
Cellular Compromise			1.72E-05 - 8.50E-19	98
Cellular Function and Maintenance		1-20 C 4/20 P	2.26E-05 - 3.68E-17	210
Cellular Movement		1000 1010	2.16E-05 - 1.10E-14	197
Cellular Development				242
Celular Development		123456789 >		242
a photo bailed of the photo bailed of the state		123430709 /		
arsigma Physiological System Development and Function				
Name			llue range	# Molecules
Hematological System Development and Function		6n6760-340 -		228
Immune Cell Trafficking		10000	2.08E-05 - 1.32E-15	135
Tissue Morphology		14 4 21 4	2.34E-05 - 3.06E-15	167
Hematopoiesis		August.		115
Lymphoid Tissue Structure and Development		Tracking and		141
Lymphola Tissue Structure and Development		123456789 >	2.772-05 - 1.782-14	141
op Tox Functions $\vee$ Assays: Clinical Chemistry and Hematology				
Name		_	lue range	# Molecules
Name Increased Levels of Creatinine		-	4.03E-01 - 9.02E-03	# Molecules
Increased Levels of AST			2.20E-02 - 2.20E-02	3
		-1		
Increased Levels of Albumin			3.62E-02 - 3.62E-02	
Increased Levels of ALT		-	6.01E-02 - 6.01E-02	
Increased Levels of Red Blood Cells		123456789 >	6.20E-02 - 6.20E-02	7
$\sim$ Cardiotoxicity		123456789 >		
Name			lue range	# Molecules
Cardiac Infarction		and the second second		30
Cardiac Inflammation			2.41E-01 - 3.46E-05	13
Cardiac Arteriopathy			4.25E-01 - 8.30E-05	29
Cardiac Dilation			6.17E-01 - 5.55E-04	24
Cardiac Enlargement			1.00E00 - 5.55E-04	36
-		123456789 >		
imes Hepatotoxicity				
Name		p-va	lue range	# Molecules
Liver Inflammation/Hepatitis		Aug. C.L. C.		41
Liver Damage		100 C 100	3.58E-01 - 1.21E-07	35
Hepatocellular Carcinoma				65
		49-0	5.87E-01 - 1.26E-04	267
Liver Hyperplasia/Hyperproliferation				
Liver Hyperbilirubinemia		123456789 >	7.11E-02 - 6.03E-04	5
$\sim$ Nephrotoxicity				
Name			lue range	# Molecules
Name Renal Necrosis/Cell Death		p-va	lue range 3.08E-01 - 6.77E-06	# Molecules
Kidney Failure		- + + + · · ·	3.62E-01 - 2.45E-05	27
Renal Inflammation		54 - Carl	5.87E-01 - 9.59E-05	27



# Summary Tab

The tab (shown above) lists the top results for the different types of analyses that are included in Core Analysis. For this dataset:

- Several immune-related Canonical Pathways are involved, such as the Phagosome Pathway and the IL-12 Signaling and Production in Macrophages Pathway, and the TREM1 Signaling pathway is predicted to be activated.
- Key cytokines are activated as upstream regulators such as TNF and IL1B.
- Biological processes in the categories of Infectious Disease, Respiratory Disease, and Inflammatory Response are involved.
- Regulator Effects networks indicate that several upstream regulators appear to drive homing of leukocytes and adhesion of blood cells.
- Interaction networks are discovered that are involved in processes such as cell cycle and organismal injury.

The p-values associated with the results are indicated on a scale in the Summary with red dots. Many of the results represent categories that contain many sub-categories, so a cluster of red dots is shown with each red dot indicating the p-value for each sub-category.

**Note**: The QIAGEN Knowledge Base is updated on a weekly basis and new pathways and other new assets are released quarterly. Details and information in these screen shots may not appear exactly the same as in the current version of this particular analysis.

# **Canonical Pathways tab**

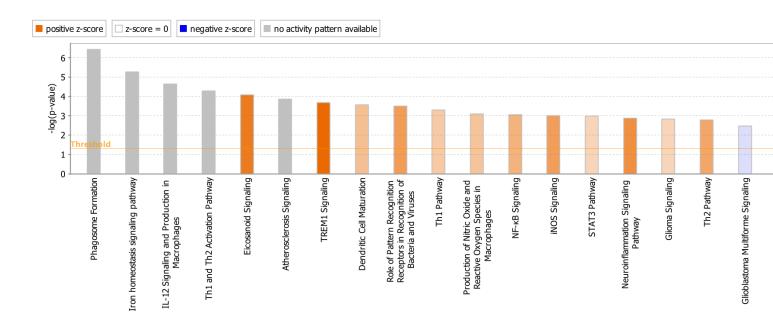
**3)** Click on the **Canonical Pathways** tab. Canonical Pathways provide information about what is known from the literature to occur on the cellular level in signaling and metabolic cascades.

The TREM1 Signaling pathway is predicted to be activated (it has an orange colored bar in the bar chart) based on the expression pattern of the genes in this dataset. TREM1 is an important cell surface signaling molecule involved in the immune response. Click on the orange bar to display results from the TREM1 Signaling pathway.

The **upper pane** (see the graphic below) shows a bar chart where the height of the bars indicates the significance of the overlap of the molecules in your dataset to the pathways in the QIAGEN Knowledge



Base. Significance values calculated based on the Fisher's right-tailed exact test and the -log(p-value) are displayed on the y-axis of the bar chart. The taller the bar, the more significant the overlap of your dataset with the pathway. See <u>Canonical Pathways for a Dataset</u> to learn more about the calculations that are used. The color of the bars indicates whether the pathway is predicted to be activated (orange bars) or inhibited (blue bars), or if the pathway is *ineligible* for such an assessment (gray bars). See <u>Pathway Activity Overlay</u> for more details on the z-score calculation used to color the bars.



# The lower pane:

Displays the molecules from your dataset that are members of the TREM1 pathway (because you clicked on that bar in the bar chart in the previous step). Note that some genes from the dataset may be members of groups or complexes that are on the pathway.

ADD TO MY PATH	WAY ADD TO MY LIST CRE	ATE DATASET CUSTOM	IIZE TABLE 📑 🔳 🛽	4						
/ Symbol	Entrez Gene N 🗵	ldentifier 🛨	Measurement		+ Add/Remove col	umn(s) Expected	×	Location 🗵	Type(s)	X
		Affymetrix 🗶	Expr False Disc 🗵	Expr p-value	🗵 Expr Log Rati	o 🗶				
AKT2*	AKT serine/threoning	236664_at*	3.93E-02	4.91E-04	+-0.466	🕈 Up		Cytoplasm	kinase	
FCGR2B	Fc fragment of IgG re	210889_s_at	3.67E-02	4.19E-04	<b>†</b> 1.052	🕈 Up		Plasma Membrane	transmembrane re	
IL10	interleukin 10	207433_at	2.44E-02	1.81E-04	<b>†</b> 0.858	🕈 Up		Extracellular Space	cytokine	
NLRC4*	NLR family CARD do	1552553_a_at*	1.14E-02	4.16E-05	<b>†</b> 1.810	🕈 Up		Cytoplasm	other	
PLCG2*	phospholipase C gar	204613_at*	2.02E-02	1.23E-04	<b>†</b> 0.614	↑ Up		Cytoplasm	enzyme	
TLR2	toll like receptor 2	204924_at	1.18E-02	4.45E-05	<b>†</b> 1.522	🕈 Up		Plasma Membrane	transmembrane re	
TLR4*	toll like receptor 4	232068_s_at*	1.03E-02	3.44E-05	<b>†</b> 1.828	🕈 Up		Plasma Membrane	transmembrane re	
TLR6*	toll like receptor 6	239021_at*	4.97E-02	8.24E-04	<b>†</b> 0.614	🕈 Up		Plasma Membrane	transmembrane re	
TLR8*	toll like receptor 8	229560_at*	6.41E-03	8.66E-06	<b>†</b> 1.903	↑ Up		Plasma Membrane	transmembrane re	
TREM1	triggering receptor e	219434 at	1.60E-02	7.74E-05	<b>†</b> 1.012	🕈 Up		Plasma Membrane	transmembrane re	

Note that the overall z-score for the TREM1 pathway is 2.53, indicating that it is likely activated (because positive values >2 are considered activating).



The table conveys the following information that can be useful in interpreting the results of your experiment:

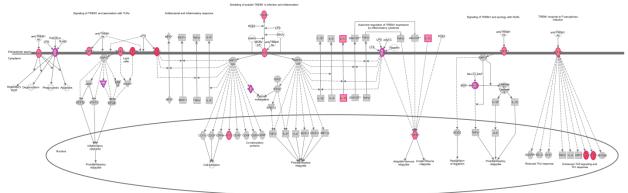
- The Gene Symbol and Entrez Gene name, which identifies the molecules in your dataset that overlap (e.g. participate in) that particular pathway.
- The observed expression changes in the dataset, in this example, Expr False Discovery Rate, Expr p-value, Expr Log Ratio.
- The identifier uploaded from the dataset (Affymetrix IDs in this example).
- The expected "direction" (up or down regulated) for the molecule in the pathway (from the point of view of the activated state of the pathway). This column is available only for pathways that are eligible for Pathway Activity Analysis (see the Eligible Pathways List).
- The predominant cellular location of the molecule, and the molecule type.
- If the molecule has been identified as a biomarker it will indicate the type of biomarker (diagnosis, efficacy, etc.), which is helpful in determining if the molecule has been studied in a particular disease state that is relevant to the one your are studying.
- If there are any drugs that target a given molecule, information on how and in what biological context the molecule has been targeted.

If the pane is too small, you can drag the vertical partition bar up to adjust it.

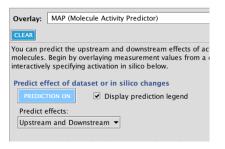
4) Click the **Open Pathway**button at the top right of the lower pane. A new window displays the canonical pathway diagram.

- Canonical pathways are usually directional, following the biological information flow in the cell. Arrows and top to bottom flow indicate upstream and downstream location, respectively.
- Dataset molecules that meet the filters and data value cutoff criteria for up- and down-regulation are shaded red and green, respectively.
- Dataset molecules that do not meet the cutoff or filter criteria are shaded grey in the pathway.
- Pathway molecules that are not in your dataset are white.
- The shapes and positions of the molecules in the pathway define gene type and cellular location, respectively. See <u>Pathway Legend</u>.
- Double outlined shapes represent groups of molecules (protein families, or protein complexes).
   Groups can be multicolored to indicate they include both up- and down-regulated molecules as their members. Right click and choose "Show Members/Membership" to expand groups to see their individual members.
- The pathway shows the biological picture of the activation of the TREM1 pathway, indicating that not only are several of the key receptors up-regulated in asthma, so are a number of downstream effectors.



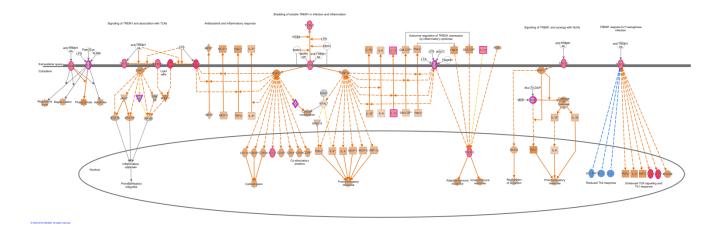


- © 2018 2318 GANGER, AN rights reserved.
- 5) To get a causal picture of the biology represented in the pathway, go to Overlay > MAP (Molecule Activity Predictor) and click the Start Prediction button.



•

This predicts the activity of nodes on the pathway that are not part of your dataset-- i.e. it colors as many gray or white nodes in the pathway orange (activated) or blue (inhibited) as possible based on the expected influence of the molecules in the dataset that are up or down regulated (i.e. the red and green molecules on the pathway). For example, if a gene is upregulated in your dataset, and the literature states that it activates a downstream gene that it is connected to in the pathway diagram, then that downstream gene is predicted to be activated. Orange nodes indicate predictions of activation, and blue nodes indicate prediction of inhibition.





Zooming in on the bottom of the pathway, you can see that several biological functions related to immunity are predicted to be increased in the pathway. For example, both adaptive and immune responses as well as pro-inflammatory response are predicted to be increased.

6) Double click on any relationship on the pathway to view the curated findings and the literature support for it. Clicking on the line between TNFa and the biological function "Proinflammatory response" brings up a relationship summary, which shows the most recently added findings.

	Relationship Summary	
View relationships	between:TNFα Proinflammatory response	ATIONSHIP
Click Add Relation	ship to create a custom relationship.	<b></b>
		-
Latest Ingenuity Fin	dings (Click link above to see all)	
Ingenuity Relations	ships	-
causation [4]		
	<b>B1 protein</b> and <b>TNFA [TNF] protein increases</b>	
proinfiammator	y response of cells.	
		-
P		ок
		, <u> </u>

Click on the blue hyperlink at the top of the relationship summary to bring up a full listing of



## the relationships:

IPA Relationships: TNFα Proinflammatory response					
Review the information that supports the gene-to-function relationship. Click the plus icon to view the reference information.					
Ingenuity Relationships					
causation [4]					
In extracellular space, TN	F-a [TNF] protein increases Proinflammatory response.				
15634892	Gerosa F, Gobbi A, Zorzi P, Burg S, Briere F, Carra G, Trinchieri G. The reciprocal interaction of NK cells with plasmacytoid or myeloid dendritic cells profoundly affects innate resistance functions. J Immunol. 2005 Jan 15;174(2):727-34.				
Source: Ingenuity Expert	Source: Ingenuity Expert Findings				
In nuclei from myeloid ce	In nuclei from myeloid cells, TNFe [TNF] protein increases Proinflammatory response.				
15385460	Lyke KE, Burges R, Cissoko Y, Sangare L, Dao M, Diarra I, Kone A, Harley R, Plowe CV, Doumbo OK, Sztein MB. Serum levels of the proinflammatory cytokines interleukin-1 beta (IL-1beta), IL-6, IL-8, IL-10, tumor necrosis factor alpha, and IL-12(p70) in Malian children with severe Plasmodium falciparum malaria and matched uncomplicated malaria or healthy controls. Infect Immun. 2004 Oct;72(10):5530-7.				
Source: Ingenuity Expert	Findings				
Binding of HMGB1 prot	ein and TNFA [TNF] protein increases proinflammatory response of cells.				
18431461	Klune JR, Dhupar R, Cardinal J, Billiar TR, Tsung A. HMGB1: endogenous danger signaling. Mol Med. 2008 Jul-Aug;14(7-8):476-84.				
Source: Ingenuity Expert Findings					
In nuclei from myeloid cells, TNFa [TNF] protein increases Proinflammatory response.					
16940328	Netea MG, Azam T, Ferwerda G, Girardin SE, Kim SH, Dinarello CA. Triggering receptor expressed on myeloid cells-1 (TREM-1) amplifies the signals induced by the NACHT-LRR (NLR) pattern recognition receptors. J Leukoc Biol. 2006 Dec;80(6):1454-61. Epub 2006 Aug 29.				
Source: Ingenuity Expert	Findings				

If desired, click the blue hyperlinks at the left side of each finding to go the NCBI record for the paper.

7) Click back to the main Core Analysis window then click VIEW REPORT bottom below the Canonical Pathway bar chart. This will open a web page with a detailed description of the TREM1 pathway.

The pathway report discusses the role of the TREM1 pathway in immune function, and shows that several of the proteins on the pathway are targets in drug trials for asthma: