Pathway Studio[®] Explore 1.1, Affymetrix Edition

Training Manual

Version 2.0

Page **1** of **68**

Preface

This manual is for users of Pathway Studio® Explore, Affymetrix Edition Version 1.1.

Instructions for the installation of Pathway Studio Explore are not included in this manual. Please see the Pathway Studio Explore Installation Guide for information on the installation of Pathway Studio Explore and the ResNet Explore Database. The Installation Guide can be found on Ariadne's Technical Support site: http://www.ariadnegenomics.com/support/pathway-studio-explore/training-material/.

How to Use this Manual

This manual is designed to walk you through some major workflows in Pathway Studio Explore. Each section has an introduction to the tools in Pathway Studio examples of how they can be used, followed by a hands-on work example. Each example builds on information from the previous exercises, and sometimes utilizes files generated from previous examples.

Two example data files accompany this training manual:

- GDS2126.gepr and GDS2126.txt This files contains the information needed to import a microarray experiment
- SCLC genes this is an MS Excel spreadsheet containing genes differentially expressed between normal lung and small cell lung tumor samples.

You will need to download these files to be able to reproduce the hands-on exercises. You can find these files in the same location in the Support section of Ariadne's website as this Training Manual.

ResNet® Explore Database and Pathway Studio® Explore software registered trademarks of Ariadne Genomics, Inc.

MS Excel® is a registered trademark of Microsoft Corporation.

Affymetrix GeneChip® is a trademark of Affymetrix, Inc.

Partek® is a registered trademark of Partek Incorporated.

Other Training and Support Resources

The Ariadne Technical Support site has additional resources available to enable Pathway Studio Explore users found at:

http://www.ariadnegenomics.com/support/pathway-studio-explore

Access to Technical Support

Technical Support for Ariadne's products can be easily accessed from Pathway Studio:

Ε	🛃 Pathway Studio® Explore - [Folders]	
	脂 Home	
	🧃 Database 🔻 📑 Import 👻 🥥 Tools 🔻	
	Search Database	<mark>ب</mark>
	ResNet Explore 1.0 (Mammal) C:\Users\heatwole\Documents\EZPathway1 Data\resnet7explore.gpy Folders	
	Database Release Notes	~
	Index of Database Content	~
	Quick Start	~
	Support & Training	^
	Welcome to Pathway Studio Explore	*
	Search for Answers	-
	Training Manual	-
	Downloads	
	Submit a Question	-
	News & Updates	~
	About Pathway Studio® Explore	~

Links within the Information Pane in Pathway Studio take you directly to the support website. Registration and log-in to this website is required.

In addition, you can access the support site directly at:

http://www.ariadnegenomics.com/support/pathway-studioexplore

A link is provided here to submit questions directly to Technical Support.

In addition this Training Manual, you can access the Quick Start Guide through the Information Pane. This Quick Start Guide will walk you through the steps of some common analysis workflows for Pathway Studio.

Quick Start Guides

🚰 Pathway Studio® Explore - [Folders]	
📑 Home	
间 Database 🔻 🛃 Import 👻 😳 Tools	•
Search Database	<i>P</i> -
ResNetExplore 1.0.0.0 (Mammal) C:\Users\Virginia\Documents\EZPathway1 Data\ResNetExplore.gpc Folders & New Pathway & New Group	,
Database Release Notes	~
Index of Database Content	×
Quick Start	^
Quick Start Hands-on guide to key functionality of Pat Studio® Explore	∧ hway
Quick Start Hands-on guide to key functionality of Pat Studio® Explore Search the ResNet® Explore Database for a gene/protein (or other entity) and build a p	hway boathway
Quick Start Hands-on guide to key functionality of Pat Studio® Explore Search the ResNet® Explore Database for a genc/protein (or other entity) and build a p Examine experimental data to find: ontolog pathways, main regulators associated with	hway bathway gies, results
Quick Start Hands-on guide to key functionality of Pat Studio® Explore Search the ResNet® Explore Database for a gene/protein (or other entity) and build a p Examine experimental data to find: ontolog pathways, main regulators associated with Examine a gene/protein list to find: ontolog pathways, main regulators associated with	hway boathway gies, results gies, list
Quick Start Hands-on guide to key functionality of Pat Studio® Explore Search the ResNet® Explore Database for a gene/protein (or other entity) and build a Examine experimental data to find: ontolog pathways, main regulators associated with Examine a gene/protein list to find: ontolo pathways, main regulators associated with	hway boathway gies, results gies, list
Quick Start Hands-on guide to key functionality of Pat Studio® Explore Search the ResNet® Explore Database for a gene/protein (or other entity) and build a p Examine experimental data to find: ontolog pathways, main regulators associated with Examine a gene/protein list to find: ontolo pathways, main regulators associated with Support & Training News & Undates	A bathway gies, results gies, list
Quick Start Hands-on guide to key functionality of Pat Studio® Explore Search the ResNet® Explore Database for a gene/protein (or other entity) and build a p Examine experimental data to find: ontolog pathways, main regulators associated with Examine a gene/protein list to find: ontolo pathways, main regulators associated with Support & Training News & Updates About Pathway Studio® Explore	bathway bathway gies, results gies, list

DENTIFY ENRICHMENT OF ONTOLOGIES DENTIFY ENRICHMENT OF ONTOLOGIES DENTIFY ENRICHMENT OF ONTOLOGIES Subset of the second of th	Pathway Studio Quick Start Guide	x
Introduction: You can run an enrichment analysis to seel if any known gene ontologies are statilicially enriched in your microarray dataset using analytical to off interpreting gene expression data to gain insights into biological mechanisms. GSEA evaluates data at the level of gene. GSEA is a statistically enriched in your microarray dataset to correlated with the compared sample classes. The entologies available in Pathway Studio & Explore include: Gene Ontology groups and Ariadne Ontology groups. You can add to this field by importing other groups into your database (not shown in this compile) Outime 9 . Open an experiment in Pathway Studio Explore. 1 . Open an experiment from the Folder Viewer. 1 . Open an experiment from the Folder Viewer. 1 . Open an experiment from the Folder Viewer. 1 . Folders 1 . Experiments 1 . 1 .	(IDENTIFY ENRICHMENT OF ONTOLOGIES	
Outline: • Open an experiment in Pathway Studio Explore. • Run the Gene Set Enrichment Analysis algorithm "Gene Set Enrichment Analysis". Procedure: 1. Open an experiment from the Folder Viewer. ↓ Folders ↓ Experiments ↓ ↓ View ↓ @ Import ↓ @ Export ↓ . New ↓ Folders ↓ Depriments ↓ ↓ Folders ↓ Depriments ↓ ↓ Projects Name ↓ Projects Open Preview Gene Set Enrich Gene Set Enrich	Introduction: You can run an enrichment analysis to see if any known gene ontologies are statistically enriched in your micrearray dataset using ontologies are statistically enriched in your micrearray dataset using analysicia tool for interpreting gene expersion data to gain insights into biological mechanismu. OSEA evaluates data at the level of gene. GSEA is a statistical text used to compare two sample groups (i.e. thorow's, normal fissues) and determine which of three gene sets tend to correlated with the compared sample classes. The entologies available in Pathway Studio® Explore include: Gene forology groups and Ariadne Ontology groups. You can add to this list by importing other groups into your database (not shown in this eample).	н
Procedure: 1. Open an experiment from the Folder Viewer. Folders × Folders × F	Outline: • Open an experiment in Pathway Studio Explore. • Run the Gene Set Enrichment Analysis algorithm "Gene Set Enrichment Analysis".	
Folders × Polders × P	Procedure: 1. Open an experiment from the Folder Viewer.	
	Polders ×	
L: View Binport P in Projects Coss126 Open Preview Gene Set Enrich Coss126 Open	🕢 🕟 📓 Folders 🕨 🎉 Experiments 🕨	
Codes Codes Pathways Ontologies Experiments Projects Coss126 Open Preview Gene Set Enrich Codes	🔢 🚉 View 🕶 🚯 Import 👻 😰 Export 👻 🌅 Nev	
Detrusys Ortologies Experiments Directos Projects Project Construction Construction Construction Construction Construction Construction Construction Construction Construction	Folders Name	
Projects GDS2126 Open Preview Gene Set Enrich Sith Matematic E	Ontologies Experiments	
Preview Gene Set Enrich	Projects GDS2126 Open	
Gene Set Enrich	Preview	
	Gene Set Enrich	
	 Run Gene Set Enrichment Analysis" (GSEA). [Experiment Data Viewer: Tools → Gene Set Enrichment Analysis] 	
 Kun "Gene Set Enrichment Analysis" (GSEA). [Experiment Data Viewer: Tools → Gene Set Enrichment Analysis] 	Edit = Select = Tarle =	
Kun Gene set Enrichment Analysis [GSEA]. [Experiment Data Viewer: Tools → Gene Set Enrichment Analysis]	cuit • IIII Select • III 10015 •	
A: Nun Vene set Enrichment Analysis (OS-bA). [Experiment Data Viewer: Tools - Edit - Select - Tools - Beid Dathmark form Colonian	s Ctd OAxe C Save Selection as Group	
	5.080e-01 Find Pathways/Groups Enriched with	

Table of Contents

Preface	2
Table of Contents	4
Section 1: Introduction to Pathway Studio Explore and ResNet Explore Database	6
What can Pathway Studio Explore do for you?	6
MedScan Reader 3.0	7
ResNet Explore Database	8
Curated Pathway Collection	8
Ariadne Gene Ontology/Gene Ontology	9
Section 2: Introduction to the Pathway Studio interface	10
Opening a Database	11
Information Pane	11
Folders View	12
Exercise One: Open the local ResNet Explore Database	13
Section 3: Building Pathways	14
Search ResNet for an Entity	14
Import a list of Entities	14
• Exercise Two: Search ResNet Explore for Entities and Relationships and Import a Pro	otein List17
Build Pathway Tools in Pathway Studio	21
Build Pathway Tool – Quick Menus	21
Build Pathway Tool – Advanced Menu Desktop	22
Graph View/ Entity Table view / Relation Table view	23
Customizing tables, filtering by significance	23
Viewing Details about Entities/Relationships	24
Creating new Entities and Relationships	25
Customization of Pathway Layouts	26
• Exercise Three: Building Pathways	28
Saving Pathway Images	
Exporting Pathways	
Section 4: Importing Experimental Data	
• Exercise Four: Import data	
Section 5: Experimental data analysis tools	41
	_

Fisher's Exact Test	41
• Exercise Five: Experimental Data Analysis – Fisher's Exact Test for Enriched Groups and Pathways	46
Gene Set Enrichment Analysis	50
• Exercise Six: Experimental Data Analysis – Gene Set Enrichment Analysis	52
Sub-Network Enrichment Analysis	56
Defining the Sub-Networks	56
Both enrichment algorithms, Fisher's Exact Test and Gene Set Enrichment Analysis, can be used to identify enriched Sub-Networks	59
• Exercise Seven: Experimental Data Analysis – Sub-Network Enrichment Analysis (with GSEA)	60
Appendix A: Definitions	65
Appendix B: Deleting Entities and Relations from a Local Database (Pathway Studio Explore and ResNet Explore)	67

Section 1: Introduction to Pathway Studio Explore and ResNet Explore Database

Congratulations on your decision to utilize Pathway Studio Explore as an important tool to support your biological research. Pathway Studio Explore helps you to interpret experimental data in the context of pathways, gene regulation networks and protein interaction maps; interpret microarray and proteomics data, classify and prioritize proteins, draw pathway diagrams, and automatically update your pathways with newly published facts using MedScan Technology.

What can Pathway Studio Explore do for you?

• Identify relationships among proteins, small molecules, cell processes and diseases

What is known to interact with my protein? What processes are associated with my protein?



Build and reconstruct pathways from your microarray and other high-throughput data

Analysis Tools to find enriched networks in experimental data



MedScan Reader 3.0

MedScan technology is a Natural Language Processing Technology used to extract relationship information from biomedical literature. Ariadne utilizes MedScan technology to build the ResNet Explore Database, the mammalian (human, mouse and rat) database, of relationship information summarized from all abstracts in PubMed as well as information contained in 61 free full-text journals. ResNet Curator was applied to the extracted data to condense information and remove some technical false positives.

Here is an example of how MedScan works. MedScan reads sentences and identifies entities (proteins, complexes, small molecules, diseases etc.) in sentences, here depicted in red.

Sentence in scanned literature: "Axin binds beta-catenin and inhibits GSK-3beta."

• Identify Proteins in Dictionary (in red): [Entities] "Axin binds beta-catenin and inhibits GSK-3beta."

Next MedScan utilizes pattern rules to identify described relationships between the sentences

- Identify Interaction Type (in blue):
 [Relationships]
- "Axin binds beta-catenin and inhibits GSK-3beta."

The extracted facts are the relationships, which are added to the ResNet database:

Extracted Facts: Axin - beta-catenin relation: Binding Axin -> GSK-3beta relation: Regulation, effect: Negative

The sentences containing identified relationships are available for examination within the Pathway Studio interface as well as the literature reference.

ProtModification Properties	Pathways	
Relation Type: ProtModification Add Remove	v Hide empty properties	Declare New Property Add Remove
Category	Property	Value
Common Properties	MedLine Reference	15308560:10068
Local Properties	Sentence	In vitro phosphorylation of PLSCR1
All References	Journal	Blood
Reference 1	Journal Reference	v104 i12 p3731
Reference 2	Journal Link	http://www.bloodjournal.org/cgi/co
Reference 3	CellType	erythrocyte
Reference 4		
		OK Cancel

Types of Relationships Identified by MedScan

The MedScan dictionary contains lists of entity names. (There are eight types of relationships between entities identified by MedScan. To see a definition of entity types and relationships types, see Appendix A. All relations have directionality except Binding.)



ResNet Explore Database

The ResNet Explore database includes almost 700,000 unique relationships derived from over 19 million PubMed abstracts as well as 61 full-text journals. In addition to information extracted by MedScan, ResNet Explore includes MeSH terms for diseases based on the Medical Subject Headings (MeSH) from the National Library of Medicine (http://www.nlm.nih.gov/mesh/meshhome.html), GO terms from The Gene Ontology Consortium (http://www.geneontology.org/), Ariadne Ontology terms, and a collection of Ariadne curated reference pathways (227 receptor signaling pathways, 21 cellular process pathways and 39 metabolic pathways). Most small molecules in ResNet have identifiers from either PubChem (http://pubchem.ncbi.nlm.nih.gov/) or from the American Chemical Society (http://www.cas.org/).

Curated Pathway Collection

The ResNet Explore Database includes a collection of reference pathways including a large number of receptor signaling pathways, cellular process pathways and metabolism pathways. These can be found in the Pathways folder. Ariadne scientists have built these pathways based on general knowledge in order to provide you with useful building blocks to extend with your own specific expertise.



Receptor pathways curated into 227 signaling pathways

21 Cellular Process pathways

39 Metabolic Pathways

Ariadne Gene Ontology/Gene Ontology

In addition to the well-known "GO" vocabulary from the Gene Ontology Consortium, Ariadne has prepared a robust ontology that has been optimized to provide the best results from analysis with Pathway Studio. Ariadne's Gene Ontology is designed to sort genes into the appropriate category, rather than to assign multiple categories to a gene. The ontology is relatively flat, containing only two groups, cellular process and molecular function. Additionally, the Ariadne Gene Ontology has been designed to minimize redundancy in the classification, and consequently avoids much of the redundancy found in analysis results produced using the Gene Ontology Consortium's vocabulary.



Organizes almost 9000 genes into 505 groups with three-tier biological hierarchy (created by Ariadne scientists)

This more succinct ontology can be used with Gene Set Enrichment Analysis and Fisher's Exact test to obtain meaningful results.

Section 2: Introduction to the Pathway Studio interface

The major panes in the interface include: Information Pane (1), Folders Views (2), List Pane (3), Graph (and Relationship and Entity) View (4), Experiment Pane (5) (see following two figures).

Pathway Studio® Explore - [Folders]						- • •
Home Database Import To Toc Database CAUSers/beastwoie/Documents/EZPathway1 DatabaseRtTeoplore.gpy Folders Folders New Group Database Release Notes Index of Database Content Quick Start Support & Training News & Updates About Pathway Studio® Explore	Folders × Fol	Pathways > Ariadne Sign. ort ← Export ← Name Pathways Atlas of Signaling	aling Pathways > Tools = Description Cell Process Regulation	Info Receptor Signaling	Find in this folder	2.
3 1 matches for 'mybl2' Ø 104 matches for B Edit ▼ Select ▼ O Tools ▼	or'abc' ×	-			Find in this table	3
Name Type ABCB6 Protein ABCA1 Protein HEATR6 Protein ABCA2 Protein ABCA3 Protein	Description ATP-binding cassette, sub-fam ATP-binding cassette, sub-fam HEAT repeat containing 6 ATP-binding cassette, sub-fam ATP-binding cassette, sub-fam	iily B (MDR/TAP), member 6 iily A (ABC1), member 1 iily A (ABC1), member 2 iily A (ABC1), member 3			#	



Opening a Database

To before working with Pathway Studio, you must first select to open the Explore Database (if it has not opened by default). The database hosted locally on your computer's hard drive.



To open a local database go to the Information Pane and select: Database > Open local.

Information Pane

The Information Pane contains subheadings that provide useful links and information.



The header in the Information Pane indicates the database that is currently active in the Pathway Studio application.

Opening the "Index of the Database Content" tab reveals a summary of the information contained in the current database, such as totals for entity and relationship types, numbers of groups, pathways and experiments as well as ontologies.

Folders View



The Folders View allows you to browse through the complete folder tree to find specific files. Pathways and Ontologies are provided in their respective folders. In the Experiments Projects folder you can create folders and subfolders as needed, and you now have the flexibility to save related items (group, pathway and experiment) in the same folder. This allows you easily organize your work.

Note: When working with this database you have complete freedom to manipulate the folders. This includes the ability to (accidently) delete reference pathways and ontology groups. It is advisable to keep all your work in the Experiments and Projects folders.

• Exercise One: Open the local ResNet Explore Database

Begin Exercise:

Objective: To learn how open the Explore Database.

Let's begin using Pathway Studio by opening the local ResNet Explore Database. (To do this, the database must first be installed on your local computer.) Note: the first time you use Pathway Studio Explore, the Explore Database should automatically open.



- 1. In the Information Pane, go to Database > Open Local.
- 2. Select resnet7explore (gyp) to open the ResNet Explore database.
- 3. In the top of the Information Pane see "ResNet Explore 1.0 (Mammal)"
- 4. Select "Index of Database Content" to see a summary of the content in the local database: types and numbers of collections, experiments, entities and relationships.

End Exercise: Open the local ResNet Explore Database

Section 3: Building Pathways

Pathway Studio Explore provides the tools to build pathways (networks) from the entities and relationships found in the ResNet Explore database. There are many ways to start to build pathways and only some examples will be shown here. You can start by searching the database to find an entity of interest and then query for relationships to that entity. Alternatively you can import a list of entities and query the relationships they have with each other and what other relationships in the database connect to them.

Search ResNet for an Entity

The Quick Search Box in the Information Pane allows for a keyword search of the entire database. Simply type the keyword in the box and select the magnifying glass to the right of the box. To define more specific searches, utilize the options in the drop-down menu to the right of the magnifying glass. Note: you can search for relationships as well as entities.



The results of the search will appear in the bottom list pane.

Import a list of Entities

You can import a list of entities by selecting the Import > Gene List. The Import Wizard appears

Entity type:	Protein		-	Lookup in the Database
	Protein			
Identified by:	Small Molecule			
	Cell Process			
	Complex		Not	e: ambiguous matches are
	Disease		high	lighted; please use the Delete
	Functional Class		Sele	cted button to resolve ambiguities
	Treatment			
	Input#	Entrez GeneID?	Match	Name in DB
Load from File				
Loud Hommen				
Paste from Clinhoard				
r uste nom enpoond				
Delete Selected				
ase select entity type, look u	up the entities in the d	latabase and press Fini	h to show matche	ed entities as a new group.
			< Pack Nr	avt > Cancel He

-

Lookup in the Database

Note: ambiguous matches are highlighted; please use the Delete Selected button to resolve ambiguities.

Name in DB

<Back Next > Cancel Help

Import Wizard

Load from File.

Paste from Clipboard

Delete Selected

Entity type: Protein

Identified by: Entrez GeneID

Entrez GeneID Name Name+Alias GenBank ID Microarray ID Use Mapfile IPI ID Unigene ID Swiss-Prot ID KEGG ID Homologene I

Homologene ID RGD ID m^{iPC}

Please select entity type, look up the entities in the database and press Finish to show matched entities as a new group.

miRBase ID MGI ID Hugo ID Note: Select this option for importing any entity list of IDs including: proteins/genes, small molecules, cell objects, cell processes, complexes, diseases, functional classes and treatments.

Once the type of entity to import is selected, a second menu is available to identify the specific ID types.

You can simply copy a list of IDs from an Excel spreadsheet or txt file to your clipboard, and then select Paste from Clipboard. The list will appear in the window. Select Lookup in the Database to find the entities.

Entity type	e: Protein		•	Lookup in the Databa	ise
Identified by	: Entrez GeneID		•		
			Not high Sele	e: ambiguous matches ar lighted; please use the Do cted button to resolve an	e elete nbiguities
	Input#	Entrez GeneID?	Match	Name in DB	-
	1	6192	OK	RPS4Y1	
Load from File	2	4232	OK	MEST	
	3	4057	OK	LTF	=
Paste from Clipboard	4	25805	OK	BAMBI	
	5	5947	OK	RBP1	
Delete Celested	6	4311	OK	MME	
Delete Selected	7	8836	OK	GGH	
	8	7447	OK	VSNL1	
	9	2938	OK	GSTA1	
	10	11130	OK	ZWINT	
	11	1164	OK	CKS2	
	12	4321	OK	MMP12	
	13	22943	OK	DKK1	
	14	11197	OK	WIF1	
	15	11272	OK	PRR4	
ana alat atitut a laal	· · · · · · · · · · · · · · · · · · ·		ish to show weatch a		
ease select entity type, loop	cup the entities in th	ie uatabase and press Fil	iish to show matche	o enucies as a new group	•

Note: if the imported ID list produces some ambiguous mapping, you have the opportunity to manually select the desired mapping. Use "Delete Selected" to remove any unwanted mapping. Choose to save your imported list as a group.

Folders Save	New Pathway	New Group ×	
Name		Description	Info
😢 LTF		lactotransferrin	591 neighbors
SERPINI1		similar to neuroserpin	92 neighbors
😣 DKK1	Save Group As		7 neighbors
🛞 SPP1	-		90 neighbors
MAGEA4	Save in: 🎍	Projects	neighbors
MAGEA6			neighbors
PTPRO	Name: Ne	w Group	neighbors
🛞 MMP12			2 neighbors
🛞 MME		Save Cancel	5 neighbors
MAGEA3			neighbors
🔡 DLK1		deita-like 1 nomolog (Urosophila)	138 neighbors
CDKN3		cyclin-dependent kinase inhibitor 3 (CDK2-associated d	95 neighbors
SERPINA3		serpin peptidase inhibitor, clade A (alpha-1 antiproteina	263 neighbors
COT A4			

Once the entities that are the starting point in pathway building have been either identified by searching ResNet Explore or imported and saved as a Group, you are ready to start to build a pathway.

• Exercise Two: Search ResNet Explore for Entities and Relationships and Import a Protein List

Begin Exercise:

Objective: To learn how to find entities and relationships contained in the ResNet Explore database either by searching the database or by importing a list of entities.

Let's begin by searching the ResNet Explore database for a specific protein:

1. Select the Search Database by Keyword option in the Information Pane and type in the protein name: MYBL2

🔀 Pathway Studio® Explore - [New Group]
📄 File 🔻 💷 Window 👻 🕜 Help 👻
iii Home
🧃 Database 🔻 💰 Import 👻 🥥 Tools 👻
mybl2
ResNet Explore 1.0 (Mammal)

The results of your search are found in the List Pane at the bottom of the screen.

🕒 Edit 🔻	Select 🔻 🕥 Tools 👻			
Name	Description	Entrez GeneID	Connectivity	#
MYBL2	similar to Myb-related protein B (B-Myb)	4605, 510420, 17865,	218	1

2. Examine the more specific search options found in the drop down menu just to the right of the Keyword search box:



- a. Search Entities by Keyword can specify specific entity types for your search
- b. Search Entities by Attribute use this option to search specific annotation fields for a selected entity type (left figure below)
- c. Search Relations by Attribute use this option to search specific annotation fields for a selected relationship type (right figure below)

Search Entities by Attributes	8	Search Relations by Attributes	23
Search For:		Search For:	
Оbj Туре	Total #	Ођј Туре	Total #
✓ Protein	106144	✓ Regulation	449638
Cell Process	2179	ChemicalReaction	8630 =
Functional Class	3512 =	Expression	108889
Disease	3936	DirectRegulation	19837
Complex	314	MolSynthesis	10084
Small Molecule	1220 +	Binding	50559 +
Add Condition	alue Logic and	Add Co Attribute Operation	Nalue Logic and
Alias		# of References	
Cell Localization		Authors	
Connectivity		CellLineName =	
Danio rerio Chromosom		Connectivity	
Description		Effect	
EC Number		ISSN	
Entrez GeneID		Issue	
FunctionalClass		Journal	
GenBank ID	OK Cancel	Journal Link	OK Cancel
		Lournal Reference	

Leave the tab in the List Pane containing the search results for MYBL2 open for now.

Now let's import of list of entity IDs. Use the MS Excel file provided with this training manual.

3. Open the MS Excel file that accompanies this manual (SCLC genes) and copy all the column contents with the header "name" to the clipboard.

	А	В	С
1	Entrez GeneID	Probesets	Name
2	4232	202016_at	MEST
3	5947	203423_at	RBP1
4	4311	203434_s_at	MME
5	8836	203560_at	GGH
6	7447	203797_at	VSNL1
7	2938	203924_at	GSTA1
8	4321	204580_at	MMP12
9	22943	204602_at	DKK1
10	11197	204712_at	WIF1
11	11272	204919_at	PRR4
12	5274	205352_at	SERPINI1
13	1469	206224_at	CST1
14	12	206262 at	

4. Go to Import > Gene List. The Import Wizard is displayed.

Note: Use Import > Gene List to import any list of entities: proteins (genes), small molecules, cell objects, cell processes, complexes, diseases, functional classes and treatments. The "Identified by" ID list will reflect the selected entity type.

Note: Recall that in Pathway Studio the gene and the product of the gene (protein) are merged into one entity.

- 5. Select "Paste from Clipboard" to copy the list of protein names into the Wizard.
- 6. For "Entity type" select: Protein. For "Identified by" select: "Name+alias".
- 7. Select "Lookup in the Database" to identify the imported entities in the ResNet database.
- 8. "Not found" in the Match column indicates the entity was not found in (note that "Name" from the column header in the MS Excel sheet is indicated as "not found").

Identified by	Name+Alias		•	1. I I. ² I. I	
				vote: ambiguous matches a nighlighted; please use the D Selected button to resolve ar	re elete nbiguities
	Input#	Name?	Match	Name in DB	
1 10 50	1	Name	not found		
Load from File	2	MEST	OK	MEST	
hata farm Clintanad	3	RBP1	OK	Pdxp	
aste from Clipboard	4	MME	ambigous	MME	1
	4	MME	ambigous	MMP12	
Delete Celested	5	GGH	OK	GGH	
Delete selected	6	VSNL1	OK	VSNL1	
	7	GSTA1	OK	GSTA1	
	8	MMP12	OK	MMP12	
	9	DKK1	OK	DKK1	
	10	WIF1	OK	WIF1	
	11	PRR4	ambigous	PRR4	
	11	PRR4	ambigous	PVRL4	
	12	SERPINI1	OK	SERPINI1	
	13	CST1	OK	CST1	

- 9. Items with ambiguous mapping are highlighted (Note: lower case and capital letters are recognized differently).
- 10. Remove undesired mapping by highlighting the rows and click "Delete Selected" button to remove.
- 11. When all undesired mapping is removed, select "Finish" to import the list. The imported list will appear in a New Group window.

Image: Save with the second	👩 Folders 🗄	New Group ×			
View V Select V Name Description Info SERPINIL similar to neuroserpin 92 neighbors DKXI similar to Dickkopf-1 (hdkk-1) 237 neighbors MAGEA4 similar to MAGE-84 39 neighbors MAGEA5 hypothetical protein (LOC782979) 14 neighbors MMMP12 matrix metallopeptidase 12 (macrophage elastase) 262 neighbors MMRE similar to Tortein 70 neighbors MMME similar to Tortein (CO781934 70 neighbors MAGEA5 hypothetical protein (LO781934 70 neighbors OCKN3 cyclin-dependent kinase inhibitor 3 (CDK2-associated d 95 neighbors SSRPINA3 serpin peptidase inhibitor 3 (CDK2-associated d 95 neighbors GGFA similar to Ribonucleoside-diphosphate reductase M2 c 41 neighbors GGH similar to Ribonucleoside-diphosphate reductase M2 c 43 neighbors SKRDINL visinin-like 1 59 neighbors SVN11 visinin-like 1 59 neighbors BAG2 BCL2-associated athanogene 2 25 neighbors BAG2 BCL2-associated athanogene 2 25 neighbors	🗷 🔡 🛃	🔻 👔 Edit 👻 🕥 Tools 👻	Find in this fo	lder 🔎	•
Name Description Info © SERPINII © DKK1 similar to neuroserpin 92 neighbors © MAGEAA similar to Dickkopf-1 (hdkk-1) 237 neighbors © MAGEAA similar to MAGE-B4 39 neighbors © MAGEAA similar to MAGE-B4 39 neighbors © MAGEAA similar to protein toC782979 14 neighbors © MMPL2 matrix metallopeptidase12 (macrophage elastase) 262 neighbors © MMPL similar to Mme protein 655 neighbors © MMPL similar to Mme protein 655 neighbors © MME similar to Mme protein 655 neighbors © CKN3 serpin peptidase inhibitor 3 (CKC2-associated du 95 neighbors © GSTA1 glutathione S-transferase A1 111 neighbors © GGH similar to human gamma-glutamyl hydrolase 43 neighbors © VSNL1 visinn-like 1 59 neighbors © MAGEA3 bL2-associated duh 111 neighbors © GGH similar to human gamma-glutamyl hydrolase 43 neighbors © WSNL1 visinn-like 1 59 neighbors © MAGEA2	View 🕶	Select 🔻			
SERPINII similar to neuroserpin 92 neighbors DKCl similar to Dickkopf-1 (hdkk-1) 237 neighbors MAGEA4 similar to MAGE-84 39 neighbors MAGEA5 hypothetical protein LOC782979 14 neighbors MMP12 similar to protein tyrosine phosphatase, receptor type, U 42 neighbors MMREA5 hypothetical protein LOC782979 262 neighbors MMRE similar to me protein 262 neighbors MMRE similar to Tortein LOC781934 70 neighbors MAGEA5 hypothetical protein LOC781934 70 neighbors MAGEA5 hypothetical protein LOC781934 70 neighbors MAGEA5 serpin peptidase inhibitor 3 (CDK2-associated d 95 neighbors SERPINA3 serpin peptidase inhibitor, clade A (alpha-1 antiproteina	Name	Description		Info	-
@ GGH similar to human gamma-glutamyl hydrolase 43 neighbors @ VSNL visinin-like 1 59 neighbors @ BAG2 BCL2-associated athanogene 2 25 neighbors @ MEST mesoderm specific transcript homolog (mouse) 46 neighbors	SERPINII DKKI MAGEA4 MAGEA4 MAGEA6 MPTRU MMF12 MME MAGEA3 CDKN3 SERPINA3 SGTA1 GRM2	similar to neuroserpin similar to DickkopF-1 (hdkk-1) similar to MAGE-B4 hypothetical protein LOC782979 similar to protein tyrosine phosphata matrix metallopeptidase 12 (macroph similar to Mme protein hypothetical protein LOC781934 cyclin-dependent kinase inhibitor 3 (serpin peptidase inhibitor, clade A (a glutathione S-transferase A1 similar to Kihonurlencide- chichospha	ise, receptor type, U hage elastase) CDK2-associated d Ipha-1 antiproteina te reductase M2 c	92 neighbors 237 neighbors 39 neighbors 14 neighbors 42 neighbors 262 neighbors 655 neighbors 263 neighbors 263 neighbors 263 neighbors 111 neighbors	н
Image: System of the	GGH	similar to human gamma-glutamy	nvdrolase	43 neighbors	
BAG2 BCL2-associated athanogene 2 25 neighbors MEST mesoderm specific transcript homolog (mouse) 46 neighbors	VSNL1	visinin-like 1	.,	59 neighbors	
MEST mesoderm specific transcript homolog (mouse) 46 neighbors	😣 BAG2	BCL2-associated athanogene 2		25 neighbors	
	B MEST	mesoderm specific transcript homole	og (mouse)	46 neighbors	-

12. Choose Save to save the group (provide a destination folder and name for the group – create a new folder if desired).

Save Group	As
Save in:	Projects/PS6 Training
Name:	Imported Protein List
	Save

End Exercise: Search ResNet for Entities and Relationships and Import a Protein List

Build Pathway Tools in Pathway Studio

Pathway Studio provides both quick short cut menus for pathway building as well as an advanced menu that allows you to define more specific network options.

Build Pathway Tool – Quick Menus

With the entities in a new Pathway View, select the Add tool to see the short cut menu for building pathways.



Each of these short cut menu options has easy-to-interpret submenus. Select the appropriate option for the desired pathway.

Note: If an option is grayed out, it is not an appropriate option for the number of entities selected.

Type(s) of Entities

The Relationship types and Entity types included in the Shortcut menu options:

Neighbors from DB Expression Targets Expression, PromoterBinding Proteins, Complexes, Functional Classes Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes Protein Modification Targets Protein Modification Targets Proteins, Complexes, Functional Classes Protein Modification Targets Protein Modification Proteins, Complexes, Functional Classes Protein Modification Targets Protein Modification Proteins, Complexes, Functional Classes Protein Modification Proteins, Complexes, Functional Classes All All All Direct Interactions Physical Interactions Direct Regulation, Binding
Expression Regulators Expression, PromoterBinding Proteins, Complexes, Functional Classes Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes Protein Modification Targets Prot Modification Proteins, Complexes, Functional Classes Protein Modification Targets Prot Modification Proteins, Complexes, Functional Classes Protein Modification Proteins, Complexes, Functional Classes Protein Modification Proteins, Complexes, Functional Classes Enzymes All All Direct Interactions Physical Interactions Direct Regulation, Binding
Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes Protein Modification Targets Prot Modification Proteins, Complexes, Functional Classes Protein Modification Prot Modification Enzymes All All Direct Interactions Physical Interactions Direct Regulation, Binding
Protein Modification Targets Protein Modification Targets Protein Modification Enzymes All All All All Direct Interactions Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Protein Modification Prot Modification Proteins, Complexes, Functional Classes Enzymes All All All Direct Interactions Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Enzymes All All All All Direct Interactions Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
All All All All All Direct Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Direct Interactions Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Expression Regulation Expression, PromoterBinding Proteins, Complexes, Functional Classes
Protein Modification Prot Modification Proteins, Complexes, Functional Classes
Shortest Path Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Expression Regulation Expression, PromoterBinding Proteins, Complexes, Functional Classes
Protein Modification Prot Modification Proteins, Complexes, Functional Classes
Common Targets Expression Targets Expression, Promoter Binding Proteins, Complexes, Functional Classes
Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Protein Modification Targets Prot Modification Proteins, Complexes, Functional Classes
Common Regulators Expression Regulators Expression. PromoterBinding Proteins. Complexes. Functional Classes
Physical Interactions Direct Regulation, Binding Proteins, Complexes, Functional Classes
Protein Modification Prot Modification Proteins. Complexes. Functional Classes
Enzymes

Type(s) of Relationships

Build Pathway Tool – Advanced Menu Desktop

The Advanced Menu option (Add > Advanced) opens a wizard that provides you to options for more specifically defining the pathway building filter options. After selecting the entities to including in your build pathway algorithm, the four wizard windows allow you to specify: a.) algorithm type for pathway building, b.) select the directionality of the resultant relationships, c.) the entities to be included in the pathway and d.) the relationships to be included in the new pathway.

Graph View/ Entity Table view / Relation Table view

There are three options for viewing resultant pathways: Graph View, Entity Table View and Relation Table View.



Customizing tables, filtering by significance

The Entity Table view and Relation Table view provide customizable tables of summary information of the pathway. Any information deleted from either of these tables is also deleted from the Graph View. Customize the columns of information by selecting Tools > Customize Columns.

🐻 Folders 📰	Imported Prote	ein List 🔄 New Pathway 🗙								
🔄 🔡 Save	🔻 🔒 Print	🔻 🏢 Edit 🔻 🖍 Undo 👻 🥥 Tools	,						Find in this pathway	-
🔠 View 👻 🧾 :	Select 👻 🏭	Add 🔻								
Relation	Туре	Sentence	MedLine Reference	Connec	# of Refere	nces	Owner	Journal Link	Journal Reference	^
→ MYBL2>	Regulation	The BMYB gene is strongly induced at th	11264176:10173.15548681:10116	2		30	Public	http://www.bloodjournal.org/c	v97 i7 p2091, v64 i22 p8167, v23	. =
→ MYBL2> a	Regulation	K Choose the Columns to be Displayed			×	21	Public	http://jcs.biologists.org/cgi/co	v119 i8 p1483, v19 i6 p719, v280	
→ MYBL2> c	Regulation	N				66	Public	http://cancerres.aacrjournals.or	v65 i2 p439, v65 i21 p9751, v7 i4	
→ MYBL2>	Regulation	A Available Columns:	Selected Colum	ns:		2	Public	http://www.bloodjournal.org/c	v105 i10 p3855, v280 i16 p15628	
→ MYBL2> s	Regulation	Authors A	Relation			29	Public	http://www.current-biology.co	v7 i4 p26, v274 i51 p36741, v105	
→ MYBL2> c	Regulation	E CellLineName	I ype Septence			57	Public	http://jcs.biologists.org/cgi/co	v119 i8 p1483, v119 i2 p6, v18 i2	
→ MYBL2> c	Regulation	E Correlation	Add >> MedLine Refer	ence		30	Public	http://www.cancercell.org/cont	v1 i4 p7, v59 i14 p3365, v62 i8 p2	ŝ.
→ MYBL2> I	Regulation	B Effect	Connectivity			1	Public	http://www.bloodjournal.org/c	v95 i12 p3900	
→ MYBL2> v	Regulation	Found In Pathways	# of References	5		5	Public	http://www.cancercell.org/cont	v1 i4 p7, v64 i7 p2561, v65 i7 p28	j. –
→ neuroblasto	Regulation	I Issue	Journal Link			8	Public	http://cancerres.aacrjournals.or	v59 i14 p3365, v275 i28 p21055,	
→ Myeloid Leuk	Regulation	I Journal	Journal Referen	ice		1	Public	http://www.bloodjournal.org/c	v96 i3 p1013	
→ lung cancer	Regulation	F Mechanism				5	Public	http://www.jbc.org/cgi/content	v275 i14 p10692, v97 i7 p2091, v	
→ MYBL2> g	Regulation	A MedlineTA	op			6	Public	http://jcs.biologists.org/cgi/co	v119 i8 p1483, v18 i23 p2837, v2	
→ breast cancer	Regulation	A Organ	Down			2	Public	http://cancerres.aacrjournals.or	v60 i16 p4519	
→ MYBL2+>	Regulation	I				13	Public	http://cancerres.aacrjournals.or	v59 i14 p3365, v119 i8 p1483, v2	
→ MYBL2> c	Regulation	F	Or	Conc		36	Public	http://www.genesdev.org/cgi/c	v19 i6 p719, v278 i11 p9655, v11	
→ cancer>	Regulation	1	ÖK	Canc	.ci	8	Public	http://cancerres.aacrjournals.or	v62 i15 p4499, v275 i28 p21055,	
→ prostate canc	Regulation	1				2	Public	http://cancerres.aacrjournals.or	v62 i15 p4499, v62 i23 p6803	
→ MYBL2> n	Regulation	Raschella Expression of Insulin-like Grow	9408744:12121, 11134182:10669,	2		10	Public	http://edrv.endojournals.org/cg	v18 i6 p801, v107 i1 p73, v275 i2	
→ MYBL2> t	Regulation	Several genes novel to testicular tumorig	11956097:10013	2		1	Public	http://cancerres.aacrjournals.or	v62 i8 p2359	
→ MYBL2>	Regulation	The defects caused by reduced B-Myb le	16551698:10038	2		1	Public	http://jcs.biologists.org/cgi/co	v119 i8 p1483	*
4									•	

Note: In the Relationship View table you can see the number of references in the database that support an individual relationship. High numbers of references can be used as a measure of confidence for a relationship. Lower numbers of references can indicate newly identified relationships or potential false positives from MedScan. The reference sentence is available for you to examine in order for you to determine the accuracy of the interpretation by MedScan. Undesired (false positive) relationships can be manually deleted from the database. See **Appendix B**: Deleting Entities and Relations from a Local Database.

Viewing Details about Entities/Relationships

Detailed information about entities and relationships displayed in the Graph View can be seen by a.) mouse over the object or b.) double click to open the properties dialog.

Bave - Marine - O			Tura ur uns patriway
Protein Properties			
General Notes Found In Pathy	vays Found In Groups		SAVAUS
Name: MYBL2	Type: Protein	Lookup in DB	e cancer de casidess et experiences de casidess et exysters et exysters et exysters et exysters et exysters et exysters et exysters et experiences et exysters et experiences et exysters et experiences et experiences
Description: similar to Myb-rel	lated protein B (B-Myb)		MORFALS VISIANS IN MELK 194 DWEINLING EST SCI LUNG RE3
Properties:	Declare New Proper	ty Add Remove	AFALT
Category ^	Property	Value relid Leuxema	Tiple peroxidation
All Properties	Bos taurus Chromosome position	13	TOP2X Symphome
General Info	Cell Localization	Nucleus 🛊 🗉 📴	
Local Properties	Connectivity	218	CONOT NOT NINE TO NINE TO
Alias	Entrez GeneID	296344	MYBLS WARPI COLTAS" homeost
Ariadne Ontology	Entrez GeneID	17865	DUFELS POPE PRIVE degenera
GenBank ID	Entrez GeneID	4605	SKP2 BRBS: PD00158 met
GO Biological Process	Entrez GeneID	<u>510420</u>	HEPATA BECOME IN DAA
GO Cellular Component	Entrez GeneID	445361 BFIEATE	HEDITES OF POLYMENSE
GO ID	FunctionalClass	DNA binding	CO. AT CONT OF BUS A
GO Molecular Function	Homologene ID	1847 DNA degradation / //	FALO
Microarray ID *	Hugo ID	7548 cell activation pro	dege
		Brein tumor	exphase florolast prolifer

Here is an example of attribute information for a selected entity, viewed in the Properties dialog:

The properties dialog for the relationship provides reference information that supports the relationship. You can view the sentences identified by MedScan that support the relationship.

	Folders Imported Protein List	🚳 New Pathway 🗙			
2	🔚 Save 🔻 🌐 Print 🔻 🛅 B	Edit 👻 🖍 Undo 👻	💭 Tools 🔻		Find in this pathway 🖉 🔻
38	View 👻 😹 Layout 👻 🧾 Select	👻 👬 Add 👻 🎐 Hig	hlight 👻 🍐 Style	e 🔻 🐠	· · · · · · · · · · · · · · · · · · ·
	PromoterBinding Properties			- • •	ACTL SCICHA
	General Linked Entities Found In F	Pathways			acter were coment
	Relation Type: PromoterBinding Add Remove	I Hide empty propert	- Declare	e New Property Add Remove	
	Category	Property	Value		Tipld perceitation
	Common Properties	Journal Link	http://www.ger	nesdev.org/cgi/content/full/16/8/933	NX
	Local Properties	Tissue	serum	I\$	COLEAS BLAN PORTS
	All References	MedLine Reference	11959842:10203		Cros avecence
	Reference 1	Sentence	Robust binding	of E2F4, p130, mSin3B, and HDAC1 to the	COLTAS TOTICS BS
	Reference 2	Journal	Genes Dev	Robust binding of E2F4, p130, mSin3B, and HDAC1 to the er	ndogenous B-myb promoter was
		Journal Reference	v16 i8 p933	observed in both cell lines (Fig. 6 B), in agreement with our lines (Figs. 2 B. 3 . 4 . and 6 A).	experiments using other mouse cell
				m OK Cancel	

Creating new Entities and Relationships

You can create a new entity or a new relationship between entities by using the Add > Entity or Add > New Relation between selected entities options in the New Pathway window.

To add a new Entity, provide appropriate information in the Add New Protein dialog:

Name:	My New Protein		Type:	Protein	 Lookup in DB
Description:				Cell Object Cell Process Complex Disease	
Properties:			De	Functional Class	dd Remove
Category		Property		Small Molecule	
All Properties				Treatment	
General Info					
Alias					
Ariadne Onte	ology				
GenBank ID					
GO Biologica	l Process				
GO Cellular (Component				
GO ID					
GO Molecula	r Function				
Microarray II)				

Note: When you attempt to add a new entity to ResNet, Pathway Studio will first check to see if an entity with the given name already exists in the database. It is advisable not to create a new entity with the same name or identifier as an existing entity.

Use the Add New Protein dialog for all entities you choose to add to ResNet, not just protein entities.

No ent	ities with a name or alias starting with	"My New Protein" has been f	ound.
Name	Alias	Description	
O Use entity set	lected in the list		
 Use entity sel Create new e 	lected in the list ntity (Ignore entities in the list)		
◯ Use entity sel	lected in the list ntity (Ignore entities in the list) Name	Value: My New Prot	ein

If your identifier is unique, complete the dialog by providing the type of primary identifier (in the example shown "name") and select OK. This will create this new entity in ResNet.

Newly created entity:



You can permanently delete an entity or relationship from ResNet by selecting it in a Graph View, then when the mouse is in the white space of the graph, right-click and select "Delete Selected Entities/Relations from the Database." Note: if you do not find this option available in the menu it can be enabled by going to the Information Pane, selecting Tools > Program Options > Menu > Enable Advanced Menu for Pathways > Yes.

Customization of Pathway Layouts

There are multiple layout options for displaying a pathway. (Please see the User's Manual for a description of how each layout is calculated.)

The Layout by Localization options utilizes Gene Ontology cellular localization assignments when placing the entity with respect to cellular objects. Be aware that entities can have more than one legitimate localization assignment within Gene Ontology (an example would be a protein that shuttles between the nucleus and cytoplasm). In this case the layout displays only one of the localization assignments. You can change the assigned localization of an entity by selecting the entity, right-click, choose "Localization in Pathway" and choose the desired cellular localization.

You can change the default layout view by choosing Layout > Set Default Layout.



• Exercise Three: Building Pathways

Begin Exercise:

Objective: To build pathways using both the Quick Menu and the Advanced Menu option and to become familiar with basic functionality used for viewing and modifying pathways.

Build Pathway Tool – Quick Menu

Let's build a simple pathway displaying all connections in the ResNet Explore database to our protein MYBL2.

- 1. Select "New Pathway" from the Information Pane. A new pathway view will open to the right.
- 2. Click and drag the MYBL2 protein icon from the List Pane below and drag it to the new pathway window.
- 3. With the MYBL2 protein entity selected (highlighted in blue) go to the Add menu and select > Neighbors from DB > All. This will identify all entities connected to MYBL2 in the database.

Pathway Studio	Explore - [New Pathw	/ay]				. •	
] File ▼ 💷 Wi	ndow 🔻 😢 Help 🔻						
🎒 Home 📃	Palette 🔝 Images		🤌 Folders 🔄 New Pathway 🗙				
间 Database 🔻	🚯 Import 🔻 🥥 To	ols 👻 🛛	🖻 🔡 Save 🔻 🖶 Print 💌 🛅 Ed	dit 👻 🖍 Undo 👻 🥥 Tools 👻	Find in this pathway	1	ρ.
mybl2	8	•	🛚 View 👻 📑 Layout 👻 📃 Select 🖲	🖌 👬 Add 👻 🌮 Highlight 👻 🔗 Style 👻 📰	• •		•
ResNet Explo C:\Users\heatwo Database Index of Database Index of Database Quick Start Support & Traini News & Updates About Pathway S	re 1.0 (Mammal) le\Documents\EZPathw ore.gpy rs	ay v v v v v v v	MYBL2	Entity New Relation between Selected Entities Relations between Selected and Unselected Neighbors from DB Direct.Interactions Shortest Path Common Targets Common Regulators Advanced	Expression Targets Expression Regulators Physical Interactions Protein Modification Targets Protein Modification Enzymes All		
🖇 2 matches fo	r 'mybl2' ×			-			
🗎 Edit 🔻 🛄 :	Select 🔻 🥥 Tools 🔻				Find in this table	1	P
lame	Туре		Description			#	
			and the second second	1 1 (1) (1)			

4. The Results Preview window displays a summary of information about the resultant pathway. You can manually review the results here and manually delete any undesired entities or relationships at this time.

Name		Туре		<mark>م</mark> `	I.	Description	Connectivity	1
✓ EP300		Protein		New		E IA binding protein p	2	
myeloid bi	ood cell	Cell Process		New			1	
Cell prolite	ration	Cell Process		New			1	
Cell contac	τ	Cell Process		New			1	-
<+	MYBL2	lame	Protein	i ype		Expression	positive	
	MIDL2		Protein	1		Proteincation	unknown	
Done. Total ti	me: 00:00:0	00 Total obj	ects che	ecked: 3	376	lu sustadada 0		

- 5. To visualize the resultant pathway, select "Finish".
- 6. Select the "Fit All Entities to Window" button (found in upper right) to view the entire pathway in the window.



7. The Advanced Visualization Tool bar provides many functions to optimize your pathway view. Select the icon to display the bottom tool bar row.



8. Select: Resize > Size All Entities to Labels to better visualize the names of the entities.



Build Pathway Tool – Advanced Menu

Now let's use a different Build Pathway algorithm to build a pathway from the list of imported proteins.

1. Right-click the group icon from the saved list of imported proteins. Select "Show contents in Bottom Pane." This will open the group list in the List Pane.

🔁 Pathway Studio® Explore - [Folders]					-	
📄 File 💌 💷 Window 💌 😧 Help 💌						
📑 Home	🍺 Folders 🛪					
🧃 Database 🕶 💰 Import 👻 🥥 Tools 🕶	🕢 🕨 📑 Folders 🔸 🕌 Projects	Find in this folder	<i>P</i> -			
mybl2 🔎 👻	🏢 🛋 View 🔻 👪 Import 👻 👔) Export 👻 📑 N	ew 🔻 🥥 Tools 💌			
ResNet Explore 1.0 (Mammal) CAUser/Nestwole/Document/EZPathway1 Data/vesnet?explore.gpy	Inders Inderse Inders Inders Inders Inder	Name mybl2 pro	Description Description Open Preview Sonw Contents in Bottom Pane Send To Export as RNEF Find Similar Pathways/Groups Cut Copy Copy Contents Delete Rename Properties) ,		
B Edit ▼ Select ▼ O Tools ▼					Find in this table	
Name Description			Entrez GeneID	Conne	ctivity #	
COL1A1 Collagen type L alpha 1			12842 1277 29393	Conne	187	1
TGFB1 transforming growth facto	r, beta 1		59086, 7040, 21803		2523	2
BGF epidermal growth factor (b	eta-uroqastrone)		25313, 1950, 13645		1923	3
3 CCND1 cyclin D1			595, 12443, 58919		1044	4
TP53 tumor protein p53			22059, 24842, 7157		2414	5 👻

2. Select "New Pathway" the Information Pane and select all the entities in the List Pane. Click and drag the entire list to the new pathway window. Select "Fit all Entities to Window" in the upper right to visualize all the entity icons.

🖉 🔚 Save 🔻 🌐 Print 🔻 🛅 Edit 🔻 🖒 Undo 👻 🥥 Tools 👻 🛛 Find in this pathway	-
🐹 View 🔻 😹 Layout 👻 🗌 Select 👻 🎊 Add 🔻 🎐 Highlight 👻 🖓 Style 👻 🐠 🛞 💮 👘 🛞 🛞	Ð
•	
• • •	
•	
• • •	
• •	
• • •	
• •	
• •	

3. Choose Select > All to select all entities. Then choose Advanced from the Add menu. The Advanced Build Pathway wizard will appear.

Let's find common transcriptional regulators shared by at least two members of our group.

- 4. In the Select Algorithm window, select "Add common regulators". This will identify upstream regulators that shared by two or more members of the group. Recall that some relations have directionality, which allows us to define direction of influence.
- 5. Next select the type of entity or entities for the regulators. In this example, select proteins.

Select Algorithm Type	s	et Filter Parameters			
Add neichbors		Entity Type	Use Filter	Operation	Value
the four partice		✓ Protein	ALL	\$	
# of expansion		Cell Process			
Add direct interactions		Functional Class			
0		Disease			
		Complex			
Add shortest path		Small Molecule			
		Treatment			
Add common regulators					
Add neighbors from group					
Group:				Ch	eck all Clear all
<back next=""> Cancel Help</back>			< <u>B</u> ack	Next > C	ancel Help

6. Finally, select the types of relation the transcriptional regulators will have with the targets. In this example, we will select Expression and PromoterBinding. Although regulation could also be utilized, the category of Regulation contains by far the largest number of relationships and is less specific, so we will leave it out for now.

Relation Type	Use Filter		Operation	Value
Regulation				
ChemicalReaction				
 Expression 	ALL	\$		
DirectRegulation				
MolSynthesis				
Binding				
PromoterBinding	ALL	ŧ		
ProtModification				
MolTransport				
			Che	ck all Clear all
			Che	ck all Clear all

7. When the algorithm is finished, select Finish to view the pathway. Note: at the bottom of the Results Preview window there is a summary of the resultant pathway. In the example below, 16 shared potential transcriptional regulators were identified and 42 relations between those regulators and the starting group of entities will be added to the pathway view.

Name	Type	ັρι	Description		Connectivity	1
CSF2	Protein	New	colony stimulating fa	actor 2	2	5
⊻ IL4	Protein	New	interleukin 4		2	
HGF	Protein	New	hepatocyte growth	factor	2	
✓ IFNG	Protein	New	interferon, gamma o	or immune type	6	
CTN	Protein	New	catanin (cadharin ac	enristed protein) hets 1	2	1
Direction	En	ity Name	Entity Type	Relation Type	Effect	_
	MM	P12	Protein	Expression	positive	
	MM	E	Protein	Expression	positive	
Done. Tota	al time: 00:1	00:04 Total	objects checked: 58	1		

- 8. Use the "Fit All Entities to Window" button to maximize the pathway view.
- 9. Select "Resize > Fit All Entities to Labels" in the Advanced Visualization tool bar.
- 10. As this pathway depicts upstream transcriptional regulators, select "Hierarchical Layout" from the layout menu. This will display the transcriptional regulators above the targets they regulate.
- 11. If needed, adjust the x and y-axis scale of the image by selecting Advanced Scaling from the menu in the far upper right. (Uncheck "Keep Aspect Ratio" before changing horizontal and vertical proportions.)

	Advanced Graph Scaling
Find in this pathway	Scale entity positions Scale entity sizes Scale entity labels Horizontal: Vertical: 100 %
Show Magnifier	✓ Keep Aspect Ratio
Advanced Scaling	Apply Close

12. The new pathway contains the original group list and new transcriptional regulators. Let's select the members of the original group. From the List Pane, choose Select > All, then Select > Mirror Selection to active pathway. The original members will be selected (blue line around selected entities). Alternatively you can Select > All, Edit > copy the list, then in the pathway view Select > Entities on Clipboard.



Now let's find shared transcriptional targets for our original list of proteins.

- 13. With the original group members selected, go to the Add menu and select Advanced.
- 14. Select algorithm type > Add common targets. This will find targets shared by two or more from our selected list.
- 15. Select Entity type Proteins and Relation type Expression and PromoterBinding.
- 16. Select Finish to view the pathway. Adjust the visualization by first selecting Layout > Hierarchical Layout, fit the pathway to the window.
- 17. Select the members of the original group by Select > All and then Select > Mirror Selection to Active Pathway

Let's add highlighting (a color halo) to the upstream regulators and the downstream targets.

- 18. With the original list selected, go to Select > Invert Selection. Now the added upstream regulators and downstream targets are selected. (Note: relations will also be selected which is ok as highlight won't apply to relationships.)
- 19. Choose Highlight > blue to highlight the upstream regulators and downstream targets. None of the members of your original list will be highlighted.
- 20. Save the pathway with the name "Upstream Regulators and Downstream Targets."



End Exercise: Building Pathways

Saving Pathway Images

Right-click on the pathway to access the four options for saving an image:

a.) Copy Picture to clipboard (available to paste into many programs),

b.) Save Picture As (saves the pathway image as a *.gif - default or *.jpeg, *.png, *.tif, and *.bmp),

c.) Save Picture with Legend (save an image of a pathway with a legend included. There are also options to scale the image size -height/width and a choice of image resolutions from screen resolution up to 1200 dpi).

d.) Save Pathway as HTML (export and save the pathway to an HTML file which enables the visualization of a pathway outside of Pathway Studio such as Internet Explorer. The HTML figure has hyperlinks for entities.)

Options for saving a pathway image are also found in the Advanced Visualization Tool bar menu when selecting the camera icon.



Exporting Pathways

From the folder view select the pathway file for export and select Export. A menu box appears with options for destinations to export the file.



The third option, "Selected Items as RNEF" opens an export wizard that allows for more flexibility in selecting data about the pathway for export.

	Туре	Source	Owner
MYBL2 network	Pathway		Admin
xport options for reference	ed items:		
	the referenced items		
Do not export content of		200	
Do not export content of Save complete informa	tion for entities and relation	/113	
Do not export content of Save complete informa Save only major proper	tion for entities and relatio ties for entities and relatio	ns	
Do not export content of Save complete informa Save only major proper Use custom filter to exp	tion for entities and relation ties for entities and relation ort specific data	ns Filter	

Section 4: Importing Experimental Data

Experimental data will be directly imported into Pathway Studio from Partek® after an initial analysis of the differentially expressed genes has been performed. See instructions provided by Partek for explanations of the data generated and exported into Pathway Studio Explore.

For the purpose of this Training Manual, two files of microarray data, GDS2126.gepr and GDS2126.txt have been provided. The following exercise demonstrates how to import this file so that the experiment data is available for use in the remaining training exercises. This experiment contains gene expression data from synovial tissues from patients with rheumatoid arthritis and osteoarthritis and well as normal controls.

• Exercise Four: Import data

Note: This example is not your standard workflow for moving experimental data into Pathway Studio Explore. This example is solely to generate an experiment data file for further training examples to be used here. When you use Pathway Studio Explore for your research needs, your data will be imported directly from Partek.

Begin Exercise: Import data (non standard workflow)

Objective: import a file of microarray data into Pathway Studio Explore.

1. Place the files, GDS2126.gepr and GDS2126.txt on your computer desktop.

2. With Pathway Studio Explore open, double click on the GDS2126.gepr file. The following dialog will appear:

Import and Analyze Experiment			X
Destination folder:	Projects		
Experiment Name:	gds2126		
p-value cutoff:	(no cutoff)	(filters out row only if p-value exceed specified cutoff)	s in all samples
Post-Import Steps:			
 GSEA Ariadne Pathways an GSEA Ariadne Pathways an GSEA Gene Ontology - OA GSEA Gene Ontology - RA GSEA Gene Ontology - RA SNEA - OA vs. Ctrl SNEA - RA vs Ctrl 	ıd Ontology - OA vs. (ıd Ontology - RA vs C .vs. Ctrl vs Ctrl	.trl trl	
Open the experiment to colo	or pathways in the res	ults Start	Cancel

3. Select the desired destination folder where the experiment will be imported.

4. Some experimental analysis options have been selected by default (Gene Set Enrichment Analysis and Sub-Network Enrichment Analysis). For the purpose of this training session, uncheck all the analysis. We will introduce and explain these tools below prior to running these analyses. When the import has completed, select "Finish"

Import and Analyze Experiment			×
Destination folder:		Projects/Training	
Experiment Name:	gd	s2126	
p-value cutoff:		(filters out row only if p-values in all sampl exceed specified cutoff)	les
Experiment import complete			
Experiment Import started Mapping probes to entities Mapping completed. 11950 pr No tools selected	robes	were mapped to entities using property LocusLink ID	*
		Start	h

The Experiment Table will open automatically.

📑 gds2126 ×								
Link 🝸 🔻 🗈 Edit 👻 🔄 Select 👻 🥥 Tools 👻								
III View ▼ 🤫 Colors ▼ Find probe								
Name	OA vs. Ctrl	OA vs. Ctrl : p	RA vs Ctrl	RA vs Ctrl : pv				
МАРКЗ	0.2180	5.080e-01	-0.1990	5.290e-01				
TIE1	-0.2968	4.940e-01	-0.4092	6.780e-01				
CYP2C19	-1.0393	4.050e-01	-0.0494	9.820e-01				
CXCR5	0.7845	5.740e-01	1.0190	5.220e-01				
CXCR5	-0.1713	9.370e-01	1.2458	3.220e-01				
DUSP1	-0.2578	7.140e-01	-0.4002	5.080e-01				
MMP10	-0.4918	8.010e-01	-0.9154	6.080e-01				
DDR1	0.0956	8.620e-01	-0.1134	7.950e-01				
EIF2AK2	0.8167	3.820e-01	0.8275	3.410e-01				
HINT1	-0.1156	8.740e-01	-0.3861	3.480e-01				
RABGGTA	0.1931	7.510e-01	0.4647	1.380e-01				
MAPK11	-0.5302	6.360e-01	-0.5809	6.550e-01				
YWHAE	-0.8455	5.440e-01	-0.0109	9.900e-01				
PCAF	0.3495	8.400e-01	0.5316	6.780e-01				
SMAD5	0.2637	7.810e-01	-0.0301	9.820e-01				
POLG	-0.3515	3.660e-01	-0.2671	3.880e-01				
LTMK1	0.1080	9.350e-01	0.6239	3.840e-01				

End Exercise: Import data

Section 5: Experimental data analysis tools

The experimental data analysis tools in Pathway Studio provide two different enrichment analysis algorithms, the Fisher's Exact Test and Gene Set Enrichment Analysis. Identification of enrichment of defined gene sets (pathways and groups) as well as user-defined sub-network identified in the ResNet database are both possible.

gene	Known Gene Sets	Sub-Networks
sets	(ontologies, curated pathways)	(user defined from ResNet)
algorithms		
Fisher's Exact Test (experimental values <u>not</u> utilized)	Find Pathways/Groups Enriched with Selected Entities	Find Sub-Networks Enriched with Selected Entities
Gene Set Enrichment Analysis (GSEA) (experimental values <u>are</u> considered in the analysis)	Gene Set Enrichment Analysis	Sub-Network Enrichment Analysis

Names of the tools as they appear in Pathway Studio

Fisher's Exact Test

Fisher's Exact test is a statistical test used to determine if there are nonrandom associations between two categorical variables. You can use the Fisher's Exact test to see if there are gene groups (such as ontology groups) or pathways that are statistically enriched in your list of genes. In addition you can identify if subnetworks are enriched.

For more information about the Fisher's Exact test see: <u>http://mathworld.wolfram.com/FishersExactTest.html</u>

In gene expression analysis, the Fisher's Exact test is typically run on the list of genes that have been determined (ex. by fold change / p-value) to be statistically significantly differently expressed between experimental conditions. The experimentally derived values are not utilized in the Fisher's Exact test when calculating enrichment in the gene list.

You can launch the Fisher's Exact test from the Group view by selecting the proteins in the group, then rightclick and select "Find Pathways/Groups Enriched with Selected Entities".

Polders SCLC genes ×											
🐼 🔚 Save 🔻 🗎 Edit	😰 📊 Save 🔻 🏢 Edit 🔻 🥥 Tools 👻										
View - Select -											
Name	Description	Info									
3 SERPINII	similar to neuroserpin	101 neighbors									
BKK1	dickkopf homolog 1 (Xenopus	s la 297 neighbors									
😣 MAGEA4	Open										
😢 MAGEA6	Preview										
😂 PTPRO											
B MMP12	Find Sub-Networks Enriched with Selected Entities										
See MME	Find Pathways/Groups Enriched with Selected Entities										
😂 MAGEA3	Build Pathway from Selection										
, 🍪 CDKN3											
SERPINA3	Сору										
😂 GSTA1	Delete										
😂 RRM2											
SGH GGH	Properties										
😣 RBP1	retinol binding protein 1, cellu	lar 168 neighbors									
SNL1	visinin-like 1	67 neighbors									

You can also launch the Fisher's Exact Test from the Experiment Pane. First filter the view of your results table by fold change / p-value to identify the list of statistically significant genes.

Filter a table by selecting the filter icon and choosing: "Filter Probes by Value"



You can define two filters, in this example fold change range and p-value range.

ter Probes by Value			E
elect samples of interest	Se	elect All	Unselect All
Sample	Т	уре	
OA/N	S	ample	
RA/N	S	ample	
Eiltering conditions (cossify at least one on	in may or n y	alua cut	toff)
Filtering conditions (specify at least one - n	in, max or p-v	alue cui	(on)
	min		max
Hide probes within range range	-1.2000	to	1.2000
Hide probes with p-values exceeding	(no cutoff)		
		OK	Cancel

Note: Applying this filter does not remove the rows that are filtered out. The rows containing values that do not meet the filter criteria are displayed in gray.

To run an analysis on only the probes that did meet the filter criteria, you must first choose "Select Unfiltered Probes."

You can apply this filter to one or more than one								
dataset.	In addition, you can choose Highlight							
Probes by	y value to highlight the probes in the list							
that meet	the filter criteria.							

Name	OA/N	OA/N : pvalue
CD19	0.1164	7.711e-01
CCR7	0.1429	7.293e-01
THPO	-0.5120	6.201e-01
GSTT2	0.6160	2.957e-01
RABEPK	-0.1451	4.270e-01
IRAK1	-0.3257	1.490e-01
APBB1	-0.1878	7.976e-02
NR3C1	0.1583	6.077e-01
ABO	-0.9554	4.373e-02
BIRC3	-2.0447	3.001e-05
GPR33	0.9192	7.073e-03
CCR9	0.8968	2.277e-03
ISG15	0.5487	6.431e-02
EPHA1	0.4140	1.652e-01
00.054	0.0004	0.000 04

In addition, you can determine this list of genes in an analysis outside of Pathway Studio, and import the resultant list into Pathway Studio for Fisher's Exact test analysis.

To run the Fisher's Exact test from the Experiment Pane, select the genes to be included in the analysis by choosing "Select Unfiltered Probes", then right-click on the gene name column and select "Find Pathways/Groups Enriched with Selected Entities."

☐ GDS2126 ×									GDS2126	i ×			
	Link 🖓 🕶 Edit 👻 🛄 Select 👻 🥥 Tools 👻									📄 Link 📑	7.	🗎 Edit 👻 📗	Select 👻 🔘 T
	III View	🤫	Colors 🔻		Find probe		ρ.	-		III View 🕶	- 3	Colors 🔻	
	Name		OA/N	p-values for	RA/N	p-values fo	or R	^		Name		OA/N	p-values for
	disease st	tate]				-		disease state			
	DDIT4		-3.2274	1.108e-07	-2.1787	6.397e-04				DDIT4		-3.2274	1.108e-07
	H1FX		-2.2321	3.932e-07	-2.1971	7.210e-04				H1FX		-2.2321	3.932e-07
	H1FX		-1.1628	7.150e-07	-1.1660	1.276e-03				H1FX		-1.1628	7.150e-07
	MMP3		3.8320	3.986e-06	4.3299	5.974e-05				MMP3		3.8320	3.986e-06
	RAB8A		0.6196	5.130e-06	0.5693	9.755e-03				RAB8A		0.6196	5.130e-06
	ІТРКС		-3.1132	6.559e-06	-2.0506	5.059e-06				ІТРКС		-3.1132	6.559e-06
	TSC22D3		-1.4353	7.789e-06	-1.1027	1.186e-03				TSC22D3		1 //252	7 790 0 06
	MAQA		-2.1788	8.079e-06	-3.3359	6.010e-06				MAOA	Bi	uild Pathway fr	om Selection
T	FAS	Build Pa	thway from Sele	ction		-05				Save Selection as Group			
	HNI	Save Se	ection as Group	•		-03				HNRNPA1	Fi	nd Pathways/G	Groups Enriched wit
	VEG	Find Pat	thways/Groups E	nriched with Seleo	ted Entities	e-01				VEGFA	Fi	nd Sub-Netwo	rks Enriched with S
	GAE	Find Sul	b-Networks Enric	hed with Selected	Entities	e-03				GADD45A	C	opv	
	GNE	~				e-05				GNE	-		
	тст	Сору				<u>e-01</u>				TCTA	Se	elect Unfiltered	Probes
	SLC	Select U	nfiltered Probes			e-05				SLC36A1	Br	ring Selected Pi	robes Together
	MA	Bring Se	elected Probes To	gether		e-03				MAFF	Pr	roperties	
	TNF	Properti	es			e-02				TNFRSF11A		2.0172	2.275e-05
	РРАРZВ	- ·	-1.1413	2./50e-05	-1./81/	4.347e-05				PPAP2B		-1.1413	2.750e-05
	TRIM14		1.9815	2.888e-05	1.7270	5.693e-04				TRIM14		1.9815	2.888e-05
	HSPA1A		-2.0447	3.001e-05	-2.1024	1.782e-04				HSPA1A		-2.0447	3.001e-05
	FADS1		-1.8955	3.342e-05	-1.7986	3.151e-05				FADS1		-1.8955	3.342e-05
	RGS19		1.0190	3.780e-05	1.0598	5.787e-05		-		RGS19		1.0190	3.780e-05
	•	Þ	۲ III				Þ			4	•	• III	

🔢 View)e	<i>.</i> •					
Name		OA/N	p-values for	RA/N	p-values fo	r R	
disease stat	(e						
DDIT4		-3.2274	1.108e-07	-2.1787	6.397e-04		
H1FX		-2.2321	3.932e-07	-2.1971	7.210e-04		
H1FX		-1.1628	7.150e-07	-1.1660	1.276e-03		
MMP3		3.8320	3.986e-06	4.3299	5.974e-05		
RAB8A		0.6196	5.130e-06	0.5693	9.755e-03		
ТРКС		-3.1132	6.559e-06	-2.0506	5.059e-06		
TSC22D3		1 //050	7 790 - 06	1 1027	1 196- 02	1	
MAOA	B	uild Pathway	from Selection				
FASN	S	ave Selection	as Group				
HNRNPA1	F	ind Pathways	s/Groups Enriched wit	h Selected Er	ntities		
VEGFA	F	nd Sub-Networks Enriched with Selected Entities					
GADD45A	c	Copy					
Chir							
GNE	S	elect Unfilter	ed Probes				
GNE TCTA			Dealage Teasther				
GNE TCTA SLC36A1	В	Iring Selected	Probes rogether				
GNE TCTA SLC36A1 MAFF	B	ring Selected	refores rogerner				
GNE TCTA SLC36A1 MAFF TNFRSF11Å	B	roperties 2.0172	2.275e-05	1.4092	2.796e-02	J	
GNE TCTA SLC36A1 MAFF TNFRSF11Å PPAP2B	P	ring Selected Properties 2.0172 -1.1413	2.275e-05 2.750e-05	1.4092 -1.7817	2.796e-02 4.347e-05]	
GNE TCTA SLC36A1 MAFF TNFRSF11Å PPAP2B TRIM14	P	ring Selected 2.0172 -1.1413 1.9815	2.275e-05 2.750e-05 2.888e-05	1.4092 -1.7817 1.7270	2.796e-02 4.347e-05 5.693e-04	J	
GNE TCTA SLC36A1 MAFF TNFRSF11Å PPAP2B TRIM14 HSPA1A	P	Properties 2.0172 -1.1413 1.9815 -2.0447	2.275e-05 2.750e-05 2.888e-05 3.001e-05	1.4092 -1.7817 1.7270 -2.1024	2.796e-02 4.347e-05 5.693e-04 1.782e-04	J	
GNE TCTA SLC36A1 MAFF TNFRSF11Å PPAP2B TRIM14 HSPA1A FADS1	P	ring Selected 2.0172 -1.1413 1.9815 -2.0447 -1.8955	2.275e-05 2.750e-05 2.888e-05 3.001e-05 3.342e-05	1.4092 -1.7817 1.7270 -2.1024 -1.7986	2.796e-02 4.347e-05 5.693e-04 1.782e-04 3.151e-05	J	

The Find Pathways/Groups Enriched with Entities dialog opens. Here you can select by check box the groups and pathways to include in the analysis. In Pathway Studio you have available the classic Gene Ontology as well as Ariadne Ontology and the Ariadne reference pathway collection to use in the analysis. In addition you can include any pathways and groups you have created in the analysis.

Find Pathways/Groups Enriched with Entities	×
Find Pathways/Groups enriched with 308 selected entities	
(1 non-entities in selection will not be used)	
Look in:	
Pathways Ariadne Metabolic Pathways V Ariadne Signaling Pathways User's pathways V Ariadne Ontology GO C cellular_component Molecular_function	
biological_process	
User's groups	
Expand the content of functional classes, cell processes and complexes in target gene se	ts
OK Cancel	

Note: Some types of entities, such as functional classes in the Cellular Process pathways contain proteins that have been mapped to them. To include these proteins in the analysis make sure the check box labeled, "Expand the content of functional classes and complexes in target gene sets" is selected.

The results of the enrichment analysis are displayed in a table in the List Pane, ranked by p-value.

Enriched Pathways/Group	Enriched Pathways/Groups ×										
🔚 Save 👔 Edit 🕶 🛄 Se	🔚 Save 👔 Edit 🔻 🛄 Select 🔻 🥥 Tools 👻 Find in this table 🔎										
Name	Туре	Total Entities	Expanded # of Entities	Overlap	Percent Overlap	Overlapping Entities	^ 🔎 p-v	Data Source	#	-	
🛅 Estrogen-like	Group	9	9	2	22	PGR,ESRRB	0.00120421	Ariadne Ontology	1		
🚳 Melanogenesis	Pathway	51	682	13	1	ITPKC, ADM, TGFA, VEGFA, CXCL12, PRKCB, CSF1R, ED	0.00286401	Ariadne Signaling Pathways	2		
ECM degradation	Group	14	14	2	14	MMP3,MMP14	0.00298662	Ariadne Ontology	3		
🚳 Adipocytokine Signaling	Pathway	52	780	14	1	ADIPOQ,IL6R,FASN,ADFP,TNFRSF11A,FABP4,ACSL1	0.00334572	Ariadne Signaling Pathways	4		
Concogenes Concogenes	Group	273	273	6	2	RABL3, CSF1R, ETS2, RET, MYBL1, MAFF	0.00531976	Ariadne Ontology	5		
Gonadotrope Cell Activat	Pathway	71	698	12	1	ITPKC, ADM, TGFA, VEGFA, MMP3, PRKCB, MMP14, CS	0.00976601	Ariadne Signaling Pathways	6		
🛅 Adiponectin	Group	2	2	1	50	ADIPOQ	0.0117851	Ariadne Ontology	7		
CADD45	Group	3	3	1	33	GADD45A	0.0176263	Ariadne Ontology	8	-	

The "# of Entities" indicates the number of entities shown in the group or pathway. The "Expanded # of Entities" includes the total number of entities shown on the pathway and contained in complexes and functional classes. The "Overlap" is the number of proteins shared in common between your input group and the resultant group or pathways. The Percent Overlap expresses the overlap of the input list with the entities in the identified object in percent value. The overlapping entities are listed in the Overlapping Entities column. Data Source indicates the source of the group or pathway.

You can save the results of the analysis by exporting to MS Excel: Select > All, Tools > Send Data to Excel. In addition, you can save the results in Pathway Studio by selecting "Save" from the tool bar just above the table.

• Exercise Five: Experimental Data Analysis – Fisher's Exact Test for Enriched Groups and Pathways

Begin Exercise:

Objective: To run a Fisher's Exact test analysis on a filtered list from gene expression data to identify ontology groups and pathways enriched in the significantly differentially expressed gene set.

Let's use the osteoarthritis and rheumatoid arthritis synovial tissue data set to run the Fisher's Exact test.

Fisher's Exact Test

1. Filter the table by selecting the filter icon and selecting "Filter Probe by Value".

i gds2126 ×										
🗃 Link 🕎 🗣 Edit 🔻 🛄 Select 👻 🥥 Tools 👻										
III View		Filter Probes by	Value				Find probe			
Name		Highlight Probe	es by Value		Ctrl	RAV	s Ctrl : pv			
МАРКЗ	•	Don't Filter by V	alue		0	5.29)e-01			
TIE1		Filter Probes by	Active Pathway		2	6.78	0e-01			
CYP2C19		Hide Unmanner	d Prohes		4	9.820	0e-01			
CXCR5	_	ooro	J	5.220	0e-01					
CXCR5		-0.1713	9.370e-01	1.245	58 3.220e-01		0e-01			
DUSP1		-0.2578	7.140e-01	-0.400)2	5.080	0e-01			
1 10 10 10										

- 2. Set the filter for the Osteoarthritis data set (check that box) to include only genes with a log change range greater than 1.5 (i.e. set the range to be outside of -1.5 and 1.5), and set the p-value cut off at 0.05.
- 3. Select "Set Filter." All genes that don't meet the filter criteria and displayed in gray.
- 4. Right click and choose "Select Unfiltered Probes." Then right click and select "Find pathways/groups Enriched with Selected Entities". This runs the Fisher's Exact test.

□ GDS2126 ×									
📄 Link 🝸 🗸 🗈 Edit 🔻 🧾 Select 👻 🥥 Tools 👻									
III View ▼ 🧐 Colors ▼ Find probe 🔎 ▼									
Name		OA/N	p-values for \triangle	RA/N	p-val 🔺				
disease state									
H1FX		-1.1628	7.150e-07	-1.1660	1.276				
MMP3		3.8320	3.986e-06	4.3299	5.974				
RAB8A		0.6196	5.130e-06	0.5693	9.755				
ІТРКС		-3.1132	6.559e-06	-2.0506	5.059				
TSC22D3		-1.4353	7.789e-06	-1.1027	1.186				
MAOA		-2.1788	8.079e-06	-3.3359	6.010				
FASN	D.,	a area	Calendian	2 6 6 9 7	2.025				
HNRNPA	DU Car	and Patriway from Selection							
VEGFA	Sal	ave selection as oroup							
GADD454	Fin	nd Pathways/Groups Enriched with Selected Entities							
GNE	Fin	ind Sub-Networks Enriched with Selected Entities							
ТСТА	Co	opy							
SLC36A1									
MAFF	Sel	lect Unfiltered Pro	bes						
TNFRSF1:	Bri	ng Selected Prob	es logether						
PPAP2B	Pro	operties							
TRIM14		1.9815	2.888e-05	1.7270	5.693				
HSPA1A		-2.0447	3.001e-05	-2.1024	1.782				
FADS1		-1.8955	3.342e-05	-1.7986	3.151				
RGS19		1.0190	3.780e-05	1.0598	5.787				
ADH1A		-2.6042	3.812e-05	-4.7348	7.362				
NPR1		-1.2410	4.103e-05	-1.4594	8.106 🚽				
4	Þ	< III			•				

	0032120 ×							
	📄 Link 🕎 🔻	🗎 🗈 Edit 🔻 🛄	Select 👻 🕥 To	lect 🔻 😳 Tools 👻				
	🏢 View 👻 🧐	Colors -	Colors Find probe					
	Name	OA/N	p-values for \triangle	RA/N	p-val 📩			
	disease state							
	H1FX	-1.1628	7.150e-07	-1.1660	1.276			
	MMP3	3.8320	3.986e-06	4.3299	5.974			
	RAB8A	0.6196	5.130e-06	0.5693	9.755			
	ITPKC	-3.1132	6.559e-06	-2.0506	5.059			
	TSC22D3	-1.4353	7.789e-06	-1.1027	1.186			
	MAOA	-2.1788	8.079e-06	-3.3359	6.010			
	FASN	-3.2558	8.461e-06	-3.5697	2.835			
	HNRNPA1	1 2500	1 100 - 05	1.0250	1 41 2			
I	VEGFA	Build Pathway from Selection						
	GADD45A	Save Selection as Group						
	GNE	Find Pathways/Groups Enriched with Selected Entities						
	ТСТА	Find Sub-Networks Enriched with Selected Entities						
	SLC36A1	Сору						
	MAFF	с н. сн н						
	TNFRSF11A	Select Unfiltered I	Probes					
	PPAP2B	Bring Selected Pro	obes logether					
	TRIM14	Properties						
	HSPA1A	-2.0447	3.001e-05	-2.1024	1.782			
	FADS1	-1.8955	3.342e-05	-1.7986	3.151			
	RGS19	1.0190	3.780e-05	1.0598	5.787			
	ADH1A	-2.6042	3.812e-05	-4.7348	7.362			
	NPR1	-1.2410	4.103e-05	-1.4594	8.106 🖵			
	۰ ۲	•			•			

- 5. Select "Ariadne Signaling Pathways" and "Ariadne Ontology" for the analysis and then select "OK" to run the analysis. (Make sure the check box for "Expand the content of functional classes and complexes in target gene sets" is selected.
- 6. See the results in the List Pane.
- 7. Find the top pathway in the list and right-click and select "Open" to see it in a pathway view.

Pathway Studio® Explore - [Adipocytokine Signaling]										×		
🔹 📄 Folders 🖂 Ner	w Pathway	Adipocytok	ine Signaling ×			🗒 gds2126 ×						
📔 🗵 🖬 🖛 📾	- 🗈 -	မာ • 🔘 •	· Find in this pa	thway	o 🗸	Link 🖓 🕶	🗈 Edit 👻 🛄	Select 👻 🥥 To	ools 🔻			
े 🔎 🐹 View 🕶 👧 🔹	- 🔲 - 🕷	i + 🧐 + 🍐	• 📰 😑 –	- •	\odot	📰 View 🔻 🧐	Colors -			Find probe		o 🗸
						Name	OA vs. Ctrl	OA vs. Ctrl : p	RA vs Ctrl R	A vs Ctrl : pv		
E		therity	Cyrtok ine			H1FX	-2.2321	2.480e-03	-1.9489 5.	970e-02		
	ý	<u></u>	Ý			TRIM14	1.9815	1.890e-02	1.6201 8	110e-02		
	3	- <u>~</u>				TNFRSF11A	2.0172	1.690e-02	1.9371 5.	730e-02		
	Λ		¥			TF	-2.4251	2.510e-02	-2.8490 5.	850e-02		Е
		H 💭 🐧				HDC	2.2631	2.720e-02	0.5737 6.	580e-01		
C	L	10 19				ADH1B	-2.6801	3.760e-02	-4.3567 6.	.060e-02		
P			(and Der			Scrg1	2.8258	2.720e-02	0.6012 8	620e-01		
1 1	11					SDC4	-1.7723	3.390e-02	-1.5538 6	370e-02		
						MTHFD2	-1.7694	4.080e-02	-1.3384 6	540e-02		
-		X 1 T				ACSL1	-1.5598	2.510e-02	-1.2185 9	320e-02		
						1/1.50	4.0450	4.500 00	8 6420 8	.680e-02		
		, LA	TI.			Open			6	.880e-02		
	(iii) (iii)	2 🗠 🛛	a			Preview			1.	720e-02		
			<1			Find Pathways/Gro	oups Enriched wi	s 3.	130e-02			
						Find Sub-Network	s Enriched with S	1	720e-02			
						Find Similar Pathw	/ays/Groups	2	500e-02			
Η		2 🥮 🗴	r (e) (e)			Build Pathway from	m Selection	2	2.660e-02			
		- WAT	Hank.	-		Com	440e-02					
erci/iso acias de pren	egred ome stion divisi Is	acid acid e los oxid import	acid acid acid meta	tenso plycero ot 5 de		Copy Come Contractor		5	5.760e-02			
						Copy Contents					_	F.
Repriched Pathways/Grou	ins x					open cocation						
Children and States						Mirror Selection T	o Active Pathway				_	
📋 Edit 🔻 🛄 Select 🔻 🕻	🔉 Tools 🔻					Mirror Selection Fr	rom Active Pathy	vay		Find in this table		
Name	Туре	# of Entities	Expanded # of En	Overlap	Pe	Select Contents or	n Active Pathway)-va	Data Source	#	^
Adipocytokine Signaling	Pathway	52	780	10		Properties			68143	Ariadne Signaling Path		1 🗏
🔁 Aromatic aminoacid me	Group	125	125	3	-	2 111110000	опцодопца		0.00003429	Ariadne Ontology		2
Adiponectin	Group	2	2	1		50 ADIPOQ			0.00561144	Ariadne Ontology		3
GADD45	Group	3	3	1		33 GADD45/	A		0.00840575	Ariadne Ontology		4
Lipid transport	Group	55	55	2		3 ADEP,EA	BP4		0.0103221	Ariadne Ontology		5
His metabolism	Group	4	4	1		20 HDC			0.0111925	Ariadose Ontology	-	7
Melanogenesis	Pathway	51	682	7		1 VEGEA C	XCI 12 ADM MAG	DA ADHIB ADHIA	0.0195043	Ariadne Signaling Path		R
Maf	Group	7	7	1		14 MAFF			0.0195074	Ariadne Ontology		9
Burnero emuno : r									0.0255074			-

8. If the proteins in the pathway are not colored by the experimental values (here red and blue), then select the "Link" button above the Experiment Pane. This will apply the colorized values of the experiment to the nodes on the pathway.

🕅 gds2126 ×									
🔄 Link 🔽 🗣 🗎 Edit 🔻 🛄 Select 🔻 🥥 Tools 🔻									
Eink colorin	Link coloring and selection with active pathway								
Name	OA vs. Ctrl	OA vs. Ctrl : p	RA vs Ctrl						
H1FX	-2.2321	2.480e-03	-1.9489						
TON 11									

9. With the "Link" function active (the pathway nodes colored by the experiment values) select the second dataset, rheumatoid arthritis, by clicking that column header. Do you see any change in the expression levels of any of the proteins in the pathway?

Osteoarthritis

Rheumatoid arthritis



10. Find the members of an ontology group that overlap with your gene list, right click on the ontology group and select "Copy Contents." This will copy the proteins in the ontology to the clipboard. Note: if the ontology contains a child ontology group, the contents of that lower level will not be copied.

About Pathway Studi	io 🗸					
	Open Preview					
	Find Pathways/Groups Enriched with Selected Entities Find Similar Pathways/Groups Build Pathway from Selection					
	Сору					
	Copy Contents					
	Open Location					
 Ø Enriched Path Bedit ▼ Se 	Mirror Selection To Active Pathway Mirror Selection From Active Pathway Select Contents on Active Pathway					
Name	Properties					
ECM degradation	la -	Bathurse				
ALK -> STAT signa	aing	Group				
GADD45		Group				
		Group				

11. In the Experiment Pane choose: Select > Entities on Clipboard. The proteins of the ontology group will be highlighted in the Experiment table.

RA and AO synovial tissue ×								
📄 Link 🝸 🖲	- 🗈 Edit - 📃	Select 🔻 🕥 To	ools 🔻					
💷 View 🔻 💡	Colors 🔻	Find probe	۶	•				
Name	OA vs Normal	p-values for	RA vs Normal	p-values for R	^			
disease state								
JMJD6	-4.0087	1.989e-003	-1.3712	1.553e-001				
MMP1	10.2186	2.008e-003	25.3763	4.960e-003	1			
PDE4A	-3.6573	2.028e-003	-3.2914	1.160e-002				
STXBP5L	-3.3703	2.075e-003	-1.0146	9.752e-001				
STMN2	4.6250	2.323e-003	2.1641	1.698e-001				
MT2A	-3.0628	2.353e-003	-2.4260	5.068e-002				
NID2	3.8366	2.357e-003	3.2449	4.084e-003	Ξ			
3.8-1	4.0255	2.387e-003	2.7571	3.152e-002				
CD27	10.2953	2.406e-003	32.1093	2.663e-005				
MMP7	-4.0085	2.596e-003	-3.8596	1.181e-002				
CNR1	-3.4803	2.792e-003	-4.8477	1.177e-002				
LAMA2	-3.4851	2.812e-003	-10.5789	1.398e-003				
APOBEC3G	3.0230	2.857e-003	6.0628	2.301e-004				
C6orf32	-3.9680	2.861e-003	-1.2433	3.332e-001				

Right-click and select "Bring Selected Probes Together" to see all overlapping probes groups together in the experiment list.

- 12. Results of the analysis can be saved by exporting to MS Excel: Select > All, Tools > Send Data to Excel.
- 13. Leave the experiment tab open for use in the following exercise.

End Exercise: Experimental Data Analysis – Fisher's Exact Test

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) is similar to the Fisher's Exact test in that it identifies statistical enrichment in experimental data of known groups (such as ontologies) and curated pathways. Gene Set Enrichment analysis differs from Fisher's Exact test in that the **rank** (based on the absolute value of the ratios of the experimental data values) of the genes in the experimental dataset is taken into consideration when identifying enrichment and the entire dataset can be used (no statistical threshold needs to be initially defined).

Potential Advantages of Gene Set Enrichment over Fisher's Exact Test:

- Threshold relevance In Fisher's Exact test only the subset of genes determined by a relevance threshold is considered. This list can be variable depending on the user defined chosen threshold (ex. fold change or p-value).
- In Fisher's Exact Test the rank position of the gene in the experimental results is not considered (it is either on the list or its not). In GSEA rank is considered.
- GSEA is able to identify when many members of a pathway are changed even if none are changed above the threshold used in the Fisher's Exact test (correlation).

It is unnecessary and not recommended to filter your data set by fold change/p-value prior to running the Gene Set Enrichment Analysis; however you can filter to remove genes with high p-values if desired.

Note: The Gene Set Enrichment Analysis algorithm was developed at the Broad Institute, (<u>http://www.broad.mit.edu/gsea/</u>). The Broad Institute has curated a large number of gene sets that can be used in this analysis. Gene sets downloaded from the Broad Institute can then be imported into Pathway Studio by selecting Advanced in the Import menu in the Information Pane. Select Gene Sets/Gene Sets in Broad Institute Format to download these gene sets.

📑 gds2126 ×								
📄 Link 🕎 🔻	Eink ▼ In Edit ▼ Select ▼ Tools ▼							
💷 View 👻 🧐	Colors 🔻		Build I	Pathway fro	m Selection			
Name	OA vs. Ctrl	OA vs. Ct	Save S	election as	Group	1		
H1FX	-2.2321	2.480e-03	Find Pathways/Groups Enriched with Selected Entities Find Sub-Networks Enriched with Selected Entities					
TRIM14	1.9815	1.890e-02						
TNFRSF11A	2.0172	1.690e-02	02 Gene Set Enrichment Analysis					
TF	-2.4251	2.510e-02	0e-02 Sub-Network Enrichment Analysis					
HDC	2.2631	2.720e-02	.720e-02 Remap Experiment Data					
ADH1B	-2.6801	3.760e-02	Send 9	elected Ro	ws to Excel			
Scrg1	2.8258	2.720e-02	Send 9	elected Ro	ws to Text Format			
SDC4	-1.7723	3.390e-02	bend s	refected no	ins to reaction dat			
MTHFD2	-1.7694	4.080e-02	Experi	ment Prope	erties			
ACSL1	-1.5598	2.510e-02	-1.21	85	9.320e-02			
KLF9	-1.6458	4.500e-02	-1.64	30	8.680e-02			
ADIPOQ	-2.3233	2.580e-02	-6.57	86	6.880e-02			
IGHM	6.0716	2.510e-02	7.631	1	1.720e-02			

Launch GSEA from the Tools menu in the Experiment Pane.

Select the dataset for the analysis, the gene sets (ontologies/pathways,) and the preferred algorithm. You will get similar results from both the Mann-Whitley U-Test or the Kolmogorov-Smirnov algorithms however the

Mann-Whitley U-Test will run much faster. As with the Fisher's Exact test, check the box "Expand the contents of functional classes and complexes in target gene sets, to include those genes in your analysis.

Gene Set Enrichment Analysis of g	jds2126	
Sample to analyze:	OA vs. Ctrl 🔹	Gene Set Categories
Enrichment algorithm:	Mann-Whitney U-Test (faster) 🔹 🔻	
Enrichment p-value cut-off:	0.05	cellular_component cellular_function biological_process User's groups
complexes in target gene se	ts	
Limit the analysis to the hig (no filter saved with the exp	hlighted/filtered probes eriment)	
		Run Cancel

Select "Run" to compute the results and view the results in the bottom List Pane. Enriched gene sets are ranked by p-value.

☑ Significant Gene Sets for GDS2126: OA/N ×										
🔓 Save 👔 Edit 🔻 🔲 Select 👻 🥥 Tools 💌									P	٣
Name	Туре	Total Entities	Expanded # of En	# of Measured En	Measured Entities	Median change	^ 🔎 p-va	Gene Set Category	#	*
🚳 T-cell receptor -> NFATC sig	Pathway	37	172	49	PIK3C2G, TRBV19, TRAV20, CD86, CD80, CD8	1.48681	2.18275e-06	Ariadne Signaling Path	4	
🚳 Melanogenesis	Pathway	51	682	515	ALDH3B2, AOC2, CALML3, GPR183, GPR18,	1.02095	2.88109e-06	Ariadne Signaling Path	5	
Focal Adhesion Regulation	Pathway	41	308	253	PIK3C2G, ITGBL1, TAOK3, EGF, PLG, HGF, FG	1.08994	1.10461e-05	Ariadne Signaling Path	6	
🚳 Gap Junction Regulation	Pathway	51	639	469	GPR183, GPR18, GUCY1A2, SGSM3, TAOK3,	1.00207	1.23365e-05	Ariadne Signaling Path	7	
Mast Cell Activation	Pathway	64	529	383	CALML3, AP2S1, CLTB, IFNA16, PIK3C2G, ST	1.0827	1.37767e-05	Ariadne Signaling Path	8	
T-cell receptor -> ATF/CREB	Pathway	49	189	66	PIK3C2G, TRBV19, TRAV20, RAC1, CD86, CD	1.43807	1.45816e-05	Ariadne Signaling Path	9	
TGF family	Group	25	25	16	TGFB1, BMP4, BMP7, TGFB2, TGFB3, BMP2, I	-1.20278	9.44121e-05	Ariadne Ontology	10	
T-cell receptor -> CREBBP si	Pathway	36	176	47	CALML3, PIK3C2G, TRBV19, TRAV20, CD86,	1.48681	0.000134352	Ariadne Signaling Path	11	-

• Exercise Six: Experimental Data Analysis – Gene Set Enrichment Analysis

Begin Exercise:

Objective: To run the Gene Set Enrichment Analysis (GSEA) on a gene expression data set to find ontology groups and pathways that are enriched in the dataset and to compare the results of GSEA to the results obtained by the Fisher's Exact test.

Let's use the osteoarthritis and rheumatoid arthritis synovial tissue data set to run the Gene Set Enrichment Analysis and find Ariadne Pathways and Ariadne Ontology groups enriched in the experimental results.

1. If a filter has been applied to the experimental dataset, remove the filter before continuing.



2. Select Tools > Gene Set Enrichment Analysis.

Gene Set Enrichment Analysis of	gds2126	
Sample to analyze: Enrichment algorithm:	OA vs. Ctrl Mann-Whitney U-Test (faster) Mann-Whitney U-Test (faster) Kolmogorov-Smirnov (classic)	Gene Set Categories Pathways Ariadne Metabolic Pathways Variadne Signaling Pathways User's pathways Variadne Ontology
Enrichment p-value cut-off: Expand the content of funct complexes in target gene se	0.05 ional classes, cell processes and ts	GO GO Cellular_component — molecular_function — biological_process — User's groups
Limit the analysis to the hig (no filter saved with the exp	hlighted/filtered probes eriment)	
		Run Cancel

- 3. Select: Sample to analysis: (osteoarthritis vs control), Gene Set Categories: Ariadne Signaling Pathways and Ariadne Ontology, Enrichment algorithm, Mann-Whitney U-Test, p-value cut-off: 0.05, check "Expand the contents of functional classes and complexes in target gene sets." When finished, select "Run" to compute the results. The results appear in the List Pane, sorted by p-value.
- 4. Right-click on the top pathway and select open to view the pathway. See that this pathway contains many functional classes.

Pathway Studio® Explore - [Gonadotrope Cell Activation]											
• • 🔂 🔂 Folde	ers 🔺 Nev	v Pathway 🔝	Gonadotrope Cell A	ctivation ×			gds2126 ×				
	- 🖶	- h - 🗠	- 🔿 - 🛛 F	ind in this pathway	/ P	- 🕞	Link 🝸 🖣	📄 Edit 👻	🔲 Select 👻 🕥 T	ools 🔻	
🗄 🔎 🔉 View	- 🛒 -	- *	🎐 - 👌 - 🔮		• • •		View 🔻 🄫	Colors 🔻	Find probe	_	o - ا
^						Nam	ie	OA vs. Ctrl	OA vs. Ctrl : p	RA vs Ctrl	*
		<u>e</u> -@	· 🐵 💿			MAP	K3	0.2180	5.080e-01	-0.1990	
1		t de				TIE1		-0.2968	4.940e-01	-0.4092	
			to T			CYP2	2C19	-1.0393	4.050e-01	-0.0494	
			a a			CXC	85	0.7845	5.740e-01	1.0190	
		- `@ <i>[</i> @	r 👰 🧟			CXC	85	-0.1713	9.370e-01	1.2458	
4		_ & @	600	—		DUSE	71	-0.2578	7.140e-01	-0.4002	
				I		MM	-	-0.4918	8 010e-01	-0.9154	
		0		9		DDR	1	0.0956	8.620e=01	-0.1134	_
· -		🗢 🗠 🏟				EIE2		0.01/7	2.020 01	0.0075	
		@		d.		Open					
				X		Preview					
				Ψ		Find Path	ways/Groups	Enriched wit	h Selected Entities		
		<u>a</u>				Find Sub-	Networks En	riched with S	elected Entities		_
						Find Simil	ar Dathways	Groups			
						Puild Datk	ai Fatiways/	loction			
		-		•		bullu Pau	iway noni se	lection			_
-		🕘 🙍				Сору					
		Po				Copy Cor	itents				
-						Open Loc	ation				F F
		1 21 25 . 0.1	<u>.</u>		-	Mirror Sel	ection To Ac	tive Pathway			
Significant G	ene Sets for	gds2126: OA vs.	Ct ×			Mirror Sel	ection From	Active Pathw	av		
📄 Edit 🔻 🛄	Select 🔻 🕻	🕽 Tools 🔻				Select Co	atents on Act	ive Dathway	uy		P - 9
Name	Type	# of Entities	Expanded # of	# of Measured	Mea	Select Col	iterits on Aci	ive ratilivay		#	
🔊 Gonadotrop	Pathway	71	698	520	KITL	Properties					1 E
NK Cell Acti	Pathway	59	523	376	EGR1, MA	PK1, VWF,	. :	L.10926 1J	05843e-07 Ariadne Si	gnali	2
🚳 Skeletal Myo	Pathway	70	569	431	KITLG, EPO	, VEGFA, .		L.02328 8.	72919e-07 Ariadne Si	gnali	3
🚳 T-cell recept	Pathway	37	172	49	CD4, CD3E	, SYK, PT	. :	L.48679 2.	20931e-06 Ariadne Si	gnali	4
🚳 Melanogene	Pathway	51	682	515	KITLG, EPO), VEGFA, .		L.02094 3.	01022e-06 Ariadne Si	gnali	5
🚳 Focal Adhesi	. Pathway	41	308	253	KITLG, EPO), VEGFA, .	. :	L.08998 6.	11858e-06 Ariadne Si	gnali	6
🐴 Mast Cell Ac	Pathway	64	529	383	EGR1, MA	PK1, TNFS.		L.08267 8.	70775e-06 Ariadne Si	gnali	7
🗟 Gap Junctio	Pathway	51	639	469	KITLG, EPO), VEGFA, .	. :	1.00208 1.	27845e-05 Ariadne Si	gnali	8
T-cell recept	Pathway	49	189	66	CDC42, PR	KCA, CD		L.43804 1.	48001e-05 Ariadne Si	gnali	9
TGF family	Group	25	25	16	TGFB1, BN	1P2, TGFB		-1.2028 9.	49468e-05 Ariadne Or	ntolo	10
🛛 🐴 T-cell recept	Pathway	36	176	47	CD4, CD3E	, SYK, PT		L.48679 0	.00013587 Ariadne Sie	qnali	11 🔻

5. Select a functional class and right-click to open the properties view. Select the members tab to see the members of the functional class. Select "Open list in bottom pane."

E Function	nal Class Pro	perties			- • •
General	Members	Found In Pathways			
i HRA KRA MRJ RAL RAL RA RA RA RRA RRA	NS S AS NS A S S2				
				Open list in bot	ttom pane
				ОК	Cancel

- 6. With the Functional Class open in the list pane choose Select > All, Edit > Copy. This copies the list of proteins to the clipboard.
- 7. In the Experiment Pane, choose Select > Entities on Clipboard. Next, right-click and choose "Bring Selected Probes Together." This will group all the selected proteins in the table together. If no proteins are selected, there is no overlap between the selected functional class and your dataset.

📑 gds	2126 ×]						
🔄 Link 🝸 👻 🛅 Edit 🔻 🛄 Select 👻 🥥 Tools 👻								
III View ▼ 🧐 Colors ▼ Find probe 🔎 ▼								
Name		OA vs. Ctrl	OA vs. Ctrl : p	RA vs Ctrl				
NRAS		0.3834	8.950e-01	1.4004				
HRAS		-0.2210	4.970e-01	-0.5423				
RRAS		0.0673	9.170e-01	-0.1510				
KRAS		0.0719	9.470e-01	-0.0087				
KRAS		0.4586	4.530e-01	0.2972				
HRAS		0.0530	8.960e-01	-0.2722				
RRAS		0.3117	5.080e-01	-0.0044				
MRAS		-0.0269	9.890e-01	-0.2000				
RALA		-0.1031	9.150e-01	0.5532				
HINT1	В	uild Pathway from	Selection					
RABGGT	S	ave Selection as Gr	oup					
MAPK11	F	ind Pathways/Grou	ups Enriched with	Selected Entities				
YWHAE	F	ind Sub-Networks	Enriched with Sel	ected Entities				
PCAF								
SMAD5	C	Сору						
POLG	В	ring Selected Prob	es Together					
LIMK1	P	roperties						
IL13RA2 ∢			0.5306-01	-0.2243				

- 8. To find the overlap of pathways and groups between the results of the Fisher's Exact test and the Gene Set Enrichment Analysis for a dataset, first choose the tab containing the Fisher's Exact test result (labeled "Enriched Pathways/Groups"). This will allow you to examine what pathways/groups in your results that are unique to one of the algorithms and which are common to both.
- 9. Choose Select > All, Edit > Copy.
- 10. Next, choose the GSEA results tab "Significant Gene Sets". Choose Select > Entities on Clipboard. The groups and pathways that are in both analysis results will be selected.

🕝 Enriched Pathways/Groups 🕖 Significant Gene Sets for gds2126: OA vs. Ct ×									
🗈 Edit 👻 🛄 Select 👻 🥥 Tools 👻									
Name	All	All			# of Measured Entities	Measured entities			
🛃 Gonadot	Entities on Clipboard	Entities on Clipboard			520	KITLG, EPO, VEGFA, MET, EG			
📓 NK Cell A	Minne Colortion To	A stille Distle		523	376	EGR1, MAPK1, VWF, NCR2, T			
📓 Skeletal I	Wirror Selection To	Active Pathy		569	431	KITLG, EPO, VEGFA, MET, IG			
🐴 T-cell rec	Mirror Selection From	m Active Pat	hway	172	49	CD4, CD3E, SYK, PTPRC, CD8			
📓 Melanog	Select Contents on A	Active Pathw	ay	682	515	KITLG, EPO, VEGFA, MET, MA			
🗟 Focal Ad	Invert Selection		Ctrl+I	308	253	KITLG, EPO, VEGFA, MET, MA			
🔏 Mast Celi 🗛	cuvation	Patriway	04	529	383	EGR1, MAPK1, TNFSF10, CDC			
📓 Gap Junctio	on Regulation	Pathway	51	639	469	KITLG, EPO, VEGFA, MET, MA			
🐴 T-cell recep	tor -> ATF/CREB sign	Pathway	49	189	66	CDC42, PRKCA, CD4, PRKCQ			
TGF family		Group	25	25	16	TGFB1, BMP2, TGFB2, INHBA			
🐴 T-cell recep	tor -> CREBBP signali	Pathway	36	176	47	CD4, CD3E, SYK, PTPRC, CD8			
A A A STORE	Indiata a Description	Detlement	E4	520	407	KITLO FRO VECEA MET M			

11. Choose Tools > Send Data to Excel to save the selected groups and pathways to an Excel spreadsheet.

End Exercise: Experimental Data Analysis – Gene Set Enrichment Analysis

Sub-Network Enrichment Analysis

For every entity in the ResNet database, the Sub-Network Enrichment Analysis (SNEA) algorithm uses the relationships in the ResNet database to build "sub-networks" based on user specified criterion. It then uses these sub-networks and the Fisher's Exact test or GSEA algorithm to identify the networks that are significantly enriched.

The user-defined sub-networks consist of a single "regulator" gene and its targets. The significance of the target expression levels in every built network is evaluated. The result is the identification of individual "regulators" which most likely affect the differentially expressed genes, thus providing the one plausible explanation for the observed expression changes in the experiment.

Defining the Sub-Networks

Setting user-defined criteria for building the sub-networks involves first defining the "regulator" (also referred to as the "seed") and then the "neighbors" (the "targets" defined by selecting specific relationship types).

For example, defining a protein seed and promoter binding relationships will give these types of networks:



The network consists of the target proteins with which the seed protein has a promoter binding relationship.

The Sub-Network Enrichment Analysis menu provides some short-cut menu options (see table below) or you can use the "custom" options for more flexibility in sub-network definitions.

Run Enrichment analysis for OA vs. Ctrl in gds2126 against dynamically generated sub-networks of Proteins Sub-networks are generated by connecting entities to their neighbors in the database. The choice of neighbors is: Expression Targets Binding Partners Protein Modification Targets Custom Custom Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. 0.05 enrichment probes (rifter saved with the experiment) Imit the analysis to the highlighted/filtered probes (roffilter saved with the experiment)	Sub-Network Enrichment Ar	nalysis of gds2126	
against dynamically generated sub-networks of Proteins Sub-networks are generated by connecting entities to their neighbors in the database. The choice of neighbors is: Image: Second Se	Run Enrichment analysis fo	or OA vs. Ctrl 🔹 in gds2126	
Sub-networks are generated by connecting entities to their neighbors in the database. The choice of neighbors is: <pre> </pre> <pre> </pre> <pre> </pre> <pre> <pre> </pre> </pre> <pre> <pre> <pre> </pre> </pre> <pre> <pre> <pre> <pre> </pre> </pre> </pre> <pre> <pre> <pre> <pre> <pre> <pre> </pre> </pre> </pre> </pre> </pre> </pre> </pre> </pre> <pre> <pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre></pre>	against dynamically	generated sub-networks of Proteins	
The choice of neighbors is: Expression Targets Binding Partners Protein Modification Targets Custom Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. 0.05 enrichment p-value cut-off. Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel	Sub-networks are generate	- d by connecting entities to their neighbors in the database.	
Expression Targets Binding Partners Protein Modification Targets	The choice of neight	pors is:	
Expression Targets S Protein Modification Targets m Custom Custom Limit the returned results to 100 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Cuint the analysis to the highlighted/filtered probes (no filter saved with the experiment) Imit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Imit Clean up resulting sub-networks by removing neighbors not present in the experiment			
Protein Modification Targets m Custom Custom Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel	Binding Partners		s
Custom Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel	Protein Modificatio	on Targets	m
Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel	Custom		
Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK			
Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel			
Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel			
Limit the returned results to 100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel			
100 sub-networks with best p-values; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK	Limit the returned results t	0	
100 sub-networks with best p-value; use 0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel	100	-	
0.05 enrichment p-value cut-off. Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel	100 SU	ib-networks with best p-values; use	
 Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel 	0.05 er	richment p-value cut-off.	
 □ Limit the analysis to the highlighted/filtered probes (no filter saved with the experiment) ☑ Clean up resulting sub-networks by removing neighbors not present in the experiment OK Cancel 			
Image: Clean up resulting sub-networks by removing neighbors not present in the experiment OK	Limit the analysis to the	highlighted/filtered probes	
Clean up resulting sub-networks by removing neighbors not present in the experiment			
OK Cancel	Clean up resulting sub-	networks by removing neighbors not present in the experiment	
OK Cancel			
		OK	el

See definitions for these short-cut menu options in table below.

Note: Ariadne recommends that you check the box "Clean up resulting sub-networks by removing neighbors not present in the experiment." This way, only measured entities for which experimental data exists (for example proteins/genes in a microarray experiment) are included in the network. (A small molecule or a disease would not have experimental data on a gene expression microarray and would have no expression data associated with it.)

Sub-network preset conditions for short-cut menu:

Sub-network Type	Seed	Neighbors	Relationship	Direction
Expression Targets	Gene	Gene	Promoter Binding,	Outbound from
			Expression	seed
Protein Binding Partners	Protein	Protein	Binding	No direction
Protein Modification	Protein	Protein /	Protein Modification	Outbound from
Targets		Complex		seed

Use the custom menu to select specific seeds and relationships.

Advanced Parameters Generate sub-networks as neighbors of: Protein	Connected by:
Functional Class	ChemicalReaction
Complex	Expression
Small Molecule	DirectRegulation
	MolSynthesis
	Binding
	PromoterBinding
	ProtModification
	MolTransport
L	
	OK Cancel

Note: all relationships are outbound from seed to regulators except when relation is binding, which has no directionality.

The results of the SNEA are viewed in a table sorted by p-value. Each network is named by the regulator of the network.

Brriched Sub-networks for gds2126: OA vs. Ct ×							
📄 Edit 🔻 📃 Select 🔻 🥥 Tools	. •					Find in thi	s table
Name	Total # of Neighbors	# of Measured Neighbors	Gene Set Seed	Measured Neighbors	Median change	^́ ₽ p-value	#
Expression Targets of HSD11B1	7	7	HSD11B1	TIMP1, MMP9, TIMP2, MMP7, MMP1, I	1.50483	0.00135834	1
Expression Targets of LCN2	7	7	LCN2	PPARG, HMOX1, CDH1, IRS1, ATF5, VEG	-2.17543	0.00149916	2
Expression Targets of IL18BP	7	7	IL18BP	CCL5, IL1B, TNF, MMP9, IFNG, VCAM1,	-1.42662	0.00223295	3
Expression Targets of DCN	23	21	DCN	RHOA, IL1B, IL6, MMP14, MMP1, NCAN	-1.58195	0.00245946	4
Expression Targets of SPARC	20	20	SPARC	MMP3, MMP1, FN1, MMP2, CDK2, TNF	1.31887	0.00252243	5
Expression Targets of EIF2S1	11	9	EIF2S1	DDIT3, MYC, PHLDA1, ATF3, NFKBIA, C	-2.34404	0.00269447	6
Expression Targets of CCNC	5	5	CCNC	VCAM1, TYMS, CDC2, MYC, CCNH	2.25355	0.00287694	7
Expression Targets of exosome	6	6	exosome	MMP1, IFNG, TNF, SERPINE1, IL6, PRF1	-1.69525	0.00295389	8
Expression Targets of chemokine	62	57	chemokine	AKT1, CSF2, SELE, TNFSF11, CXCL12, CX	1.28138	0.00311814	9
Expression Targets of CCL3	25	24	CCL3	ITGAM, CCL21, IL10, MMP9, IL2, FLT3L	1.31741	0.00332672	10
Expression Targets of SREBF1	105	88	SREBF1	PCSK6, IL8, VEGFA, PDX1, ESR1, PPAT, F	-1.28129	0.00412663	11
Expression Targets of CD8A	120	105	CD8A	ILARA GZMA ICAMI BREI KLRKI CD	1 31605	0.00496164	12

When networks are the same size, those with greater differential gene expression will have better p-values. When differential expression is equal, larger networks will have better p-values than smaller networks.

Better p-value



Both enrichment algorithms, Fisher's Exact Test and Gene Set Enrichment Analysis, can be used to identify enriched Sub-Networks.

	Known Gene Sets (ontologies, curated pathways)	Sub-Networks (user defined from ResNet)
Fisher's Exact Test (experimental values <u>not</u> utilized)	Find Pathways/Groups Enriched with Selected Entities	Find Sub-Networks Enriched with Selected Entities
Gene Set Enrichment Analysis (GSEA) (experimental values <u>are</u> considered in the analysis)	Gene Set Enrichment Analysis	Sub-Network Enrichment Analysis

Names of the tools as they appear in Pathway Studio

• Exercise Seven: Experimental Data Analysis – Sub-Network Enrichment Analysis (with GSEA)

Begin Exercise:

Objective: To run the Sub-Network Enrichment Analysis on a gene expression dataset to identify dynamically defined transcriptional regulatory networks enriched in the dataset and to visualize multiple networks in one view to allow for identification of overlap between networks.

Let's use the osteoarthritis and rheumatoid arthritis synovial tissue data set to run the Sub-Network Enrichment Analysis and find transcriptional regulatory networks enriched in the experimental dataset.

1. If a filter has been applied to the experimental dataset, remove the filter before continuing.



- 2. Select Tools > Sub-Network Enrichment Analysis
- 3. In the Run Enrichment analysis menu, select the osteoarthritis study.
- 4. In the choice of neighbors menu select "Custom"

Sub-Network Enrichment Analy	rsis
Run Enrichment analysis for	OA vs Normal
against dynamically ge	nerated sub-networks of Proteins
Sub-networks are generated b	y connecting entities to their neighbors in the database.
The choice of neighbor	s is:
ſ	Expression Targets Binding Partners Decision Modification Custom
Limit the returned results to	
100 sub-	networks with best p-values; use
0.05 p-va	ue cut-off.
Clean up resulting sub-net	works by removing neighbors not present in the experiment
	OK Cancel

5. Here we will define transcriptionally regulated networks: Select Protein for the (transcriptional) regulator "seed" and select Expression and PromoterBinding to define the transcriptionally regulated "target" network. Note: We will not utilize Regulation in this example, as this is the largest relation category in ResNet.

Advanced Parameters	×
Generate sub-networks as neighbors of:	Connected by:
Protein Functional Class Complex Small Molecule	 Regulation Expression Binding ProtModification MolTransport DirectRegulation PromoterBinding MolSynthesis ChemicalReaction
	OK Cancel

- 6. Select OK
- 7. In the Sub-Network Analysis tool ensure that the box "Clean up resulting sub-networks by removing neighbors not present in the experiment", then select OK to run the analysis.
- 8. View the results in the List Pane. Each network is named by the "seed" (the regulator).

Brriched Sub-network	s for gds2126: OA vs. Ct	x		-				
🗈 Edit 🔻 📃 Select 🔻	😳 Tools 🔻				Find in t	his table	P	Ŧ
Name	Total # of Neighbors	# of Measured Neighbors	Gene Set Seed	Measured Neighbors	Median change	^ 🔎 p-va	#	-
🚳 Neighbors of HSD11B1	7	7	HSD11B1	TIMP1, MMP9, TIMP2, MMP1, MMP7, IL1B, TNF	1.50483	0.00137118	1	
Neighbors of LCN2	7	7	LCN2	PPARG, HMOX1, CDH1, IRS1, ATF5, VEGFA, ADIPOQ	-2.17543	0.00149553	2	
Neighbors of IL18BP	7	7	IL18BP	CCL5, TNF, IL1B, MMP9, IFNG, VCAM1, IL4	-1.42662	0.00224365	3	
Neighbors of DCN	23	21	DCN	RHOA, IL1B, MMP14, IL6, MMP1, NCAN, IFNG, CCL2, IL8	-1.58195	0.00249681	4	
Neighbors of SPARC	20	20	SPARC	MMP3, MMP1, MMP2, FN1, CDK2, MMP14, BMP2, TNFR	1.31887	0.00255866	5	
Neighbors of EIF2S1	11	9	EIF2S1	DDIT3, MYC, PHLDA1, ATF3, CCND1, NFKBIA, NOS2, AT	-2.34404	0.00271077	6	
Neighbors of CCNC	5	5	CCNC	VCAM1, TYMS, CDC2, MYC, CCNH	2.25355	0.00288025	7	
Neighbors of CCL3	25	24	CCL3	ITGAM, CCL21, IL2, MMP9, IL10, FLT3LG, ICOSLG, CXCL1	1.31741	0.00342162	8	
Neighbors of SREBF1	105	88	SREBF1	PCSK6, VEGFA, IL8, ESR1, PDX1, PPAT, FABP6, ACSL1, AD	-1.28129	0.00440584	9	
R Naighborg of CD9A	120	105	CDRA	TINDA GTMAA TOAMAL DEEL VIEWL CORE OVOLO DOOD	1 21605	0.00524220	10	

9. Let's view the top four networks in one pathway. Select the top four networks, right-click and choose "Union Selected Pathways." In the "Combine Pathway (Union)" window, select OK.

1 T	Open Preview				
E	Find Pathways/Groups Enr	iched with Selected Entities			
	Find Sub-Networks Enriche	ed with Selected Entities			
	Find Similar Pathways/Gro	ups			
	Build Pathway from Select	on			
	Сору				
	Copy Contents				
	Open Location				
	Union Selected Pathways				
	Intersect Selected Pathways				
	Subtract Selected Pathways				
- -	Mirror Selection To Active Pathway				
Enriched Sub-n	Mirror Selection From Acti	Mirror Selection From Active Pathway			
📄 Edit 🔻 📃 Sele	Select Contents on Active	Pathway			
Name	Properties		et S		
Neighbors of HSE	Copy Gene Set Seeds		1		
🚳 Neighbors of LCNz	1		LUNZ		
🚳 Neighbors of IL18BP		7	IL18BP		
Neighbors of DCN	23	21	DCN		
Neighbors of SPARC	20	20	SPARC		
Neighbors of EIF2S1	11	9	EIF2S1		
Neighbors of CCNC	5	5	CCNC		
Neighbors of CCL3	25	24	CCL3		
Neighbors of SREBF1	105	88	SREBE1		

Selected Pathways:			
Pathway			
✓ Neighbors of H	SD11B1		
✓ Neighbors of LC	CN2		
Neighbors of IL:	18BP		
✓ Neighbors of D	CN		

The resultant network should look something like this (using the dynamic layout). Note: although your network may look similar, the dynamic layout may generate a slightly different layout than the one see here.



Now let's modify this network to better view the entity names and emphasize the four regulators in this network by applying highlighting.

- 10. Select "Fit all entities to window" to maximize the size of the network in the window.
- 11. Open the Advanced Visualization tool bar and select "Size All Entities to Labels" in the Resize menu.

Page **62** of **68**

12. Next, using the control key, select the four regulators, BP1, CCR7, IV and CD69. Note that selected entities have a blue halo. One of the selected entities will also have small white boxes around it. Click on the corner white box and expand the size of this entity.





- 13. With the other three entities (regulators) still selected, go to the Advanced Visualization tool bar and select "Make Same Size" from the Resize menu. This will expand all the regulators to the same larger size.
- 14. While the four regulators remain selected, go to the highlight menu and select the blue highlight. This will put a blue halo around the regulators. Your network should look similar to this:



Note that two of the four regulators is gray. This is because these regulator were not measured in the gene expression experiment. The Sub-Network Enrichment Analysis is able to utilize all the relationship information in ResNet to identify potentially important regulators for a gene expression experiment, even if that specific regulator was not included on the microarray.

End Exercise: Experimental Data Analysis – Sub-Network Enrichment Analysis (with GSEA)

This Training Manual gave examples of three of the four enrichment analysis options highlighted below: Fisher's Exact Test of known gene sets, Gene Set Enrichment Analysis of known gene sets and Sub-Network Enrichment Analysis (GSEA) of Sub-Networks.

gene sets	Known Gene Sets (ontologies, curated pathways)	Sub-Networks (user defined from ResNet)
algorithms		
Fisher's Exact Test (experimental values <u>not</u> utilized)	Find Pathways/Groups Enriched with Selected Entities	Find Sub-Networks Enriched with Selected Entities
Gene Set Enrichment Analysis (GSEA) (experimental values <u>are</u> considered in the analysis)	Gene Set Enrichment Analysis	Sub-Network Enrichment Analysis

In addition, you can run the Fisher's Exact Test on any list of proteins to find enriched sub-networks. Keep in mind that if you start with a short protein list, most of your enriched sub-networks will also be small. The Find Sub-Networks Enriched with Entities dialog allows you to set a threshold for the size of the sub-network to be returned.

Find Su	ib-Networks Enric	ched with Entities	×
Find S	Sub-Networks en	riched with 41 selected entities	
Sub-r	networks are gene	erated by connecting entities to their neighbors in the database.	
TI	he choice of neig	hbors is:	
	Expression Tar	nets	
	Binding Partne	rs	
	Protein Modifi	cation Targets	
	Custom	-	
			_
Filter	returned results		1
	2	or more selected entities should be present in a sub-network	L 1
Limit	the returned resu	lts to	
	100	sub-networks with best p-values; use	
	0.05	enrichment n-value cut-off	
	0.00	cincinicity voide car on	
Cle	ean up resulting s	ub-networks by removing neighbors not in the original selection	
		OK	el

Appendix A: Definitions

Entity definitions:

Protein: The principal source of proteins and their annotation in ResNet is Entrez Gene. The level of detailzation of proteins in ResNet Mammal is the Gene. This means that if proteins are encoded by the same gene, they will have the same identifier.

Small molecule: Symbolizes metabolites, drug and other chemicals of low molecular weight (< 1 KDa). It also can be used to represent non-biological polymers of larger molecular weight.

Cell object: Symbolizes cellular organelles and other sub-cellular components. A majority of cell process entities coincide with the part of the Gene Ontology cellular component classification that excludes protein complexes.

Cell process: Symbolizes biological processes. A majority of Cell process entities coincide with Gene Ontology biological processes classification.

Disease: Symbolizes diseases and other health conditions and processes.

Treatment: Symbolizes non-chemical treatments and environmental conditions such as cold shock, draught etc.

Complex: This is a container entity that symbolizes several polypeptides that form the complex via physical interaction. The complex is usually well-characterized in the literature, performs well-defined function and is referred in the literature by a specific name. A majority of complex entities coincide with part of Gene Ontology cellular classification that describes protein complexes.

Functional class: This is a container entity that symbolizes functional classes of proteins. Majority of functional class entities coincide with Gene Ontology molecular function classifications.

Relationship definitions:

MolTransport: Indicates that the regulator changes the localization of the target. Describes events of molecular translocation, export, import or release

Regulation: Indicates that the regulator changes the activity of the target. The mechanism of the regulation is either unknown or has not been specified in the sentence describing the relation.

Chemical Reaction: Symbolizes either enzyme catalyzed or spontaneous chemical reaction, i.e. transformation of one set of Small Molecules into another. Usually must have at least one substrate show with the link incoming to control and one product shown with the link outgoing from control. Enzyme is shown as undirected link between Functional Class or Protein entity and control

Binding: Physical interaction between molecules

ProtModification: Indicates that the regulator molecule changes the protein modification of the target molecule. Usually indicates the direct interaction, i.e. the regulator catalyses the chemical modification reaction.

DirectRegulation: Indicates that the regulator influences the target activity by physically interacting with it. Expression: Indicates that the regulator changes the protein level of the target, by means of regulating its gene expression or protein stability.

PromoterBinding: Indicates that the regulator binds the promoter of the target.

MolSynthesis: Indicates that the regulator changes the concentration of the target. Usually Small Molecule is a target in MolSynthesis

Appendix B: Deleting Entities and Relations from a Local Database (Pathway Studio Explore and ResNet Explore)

When you delete a relation or entity from a pathway view, it is only removed from that specific pathway but still remains in the ResNet database.

To permanently delete a relation or entity from the ResNet database, select the relation or entity, move the mouse cursor to an empty area of the pathway, right-click and select "Delete Selected Entities/Relations from Database."



Note: If "Delete Selected Entities/Relations from the Database" does not appear as a menu option, in the Information Pane go to Tools > Program Options. Next, select "Menu" then change "Enable Advanced Menu for Pathways" from "No" to "Yes."

🔁 Pathway Studio® Explore - [Ar	iadne Ontology]
iii Home	
间 Database 🔻 🚯 Import 🔻	😳 Tools 🔻
Search Database	👔 Program Options 🔎 🔻

View	Option	Value
ayout	Auto-hide application menu bar (use Alt to show/hide)	Yes 🌻
Menu General	Enable advanced Database menu	Yes 🌲
	Enable advanced Export menu	No 🌻
	Enable advanced menu for pathways	Yes 🌻
	Shortcut for Build Pathway (F8)	Yes
		No

If you have any questions about Pathway Studio Explore, please contact us at:

Ariadne Support Team

Call Monday – Friday 9:00 am – 5:00 pm Eastern time (GMT -5:00) 866-340-5040 (US and Canada toll-free) or +1-240-453-6301 Email: <u>support@ariadnegenomics.com</u> <u>http://www.ariadnegenomics.com/support/pathway-studio-explore</u>