

Ingenuity Pathway Analysis (IPA) hands-on workshop

(Introduction and data upload)

by

MSc. Julio A. Finalet Ferreiro

IPA use for cancer research: examples

1)

Haematologica. 2015 Jul;100(7):e275-9. doi: 10.3324/haematol.2015.124305. Epub 2015 Mar 20.

Post-transplant molecularly defined Burkitt lymphomas are frequently MYC-negative and characterized by the 11q-gain/loss pattern.

Ferreiro JF¹, Morscio J², Dierickx D³, Marcelis L², Verhoef G³, Vandenberghe P¹, Tousseyn T², Wlodarska I⁴.

⊖ Author information

¹KU Leuven, University of Leuven, Center for Human Genetics, Belgium.

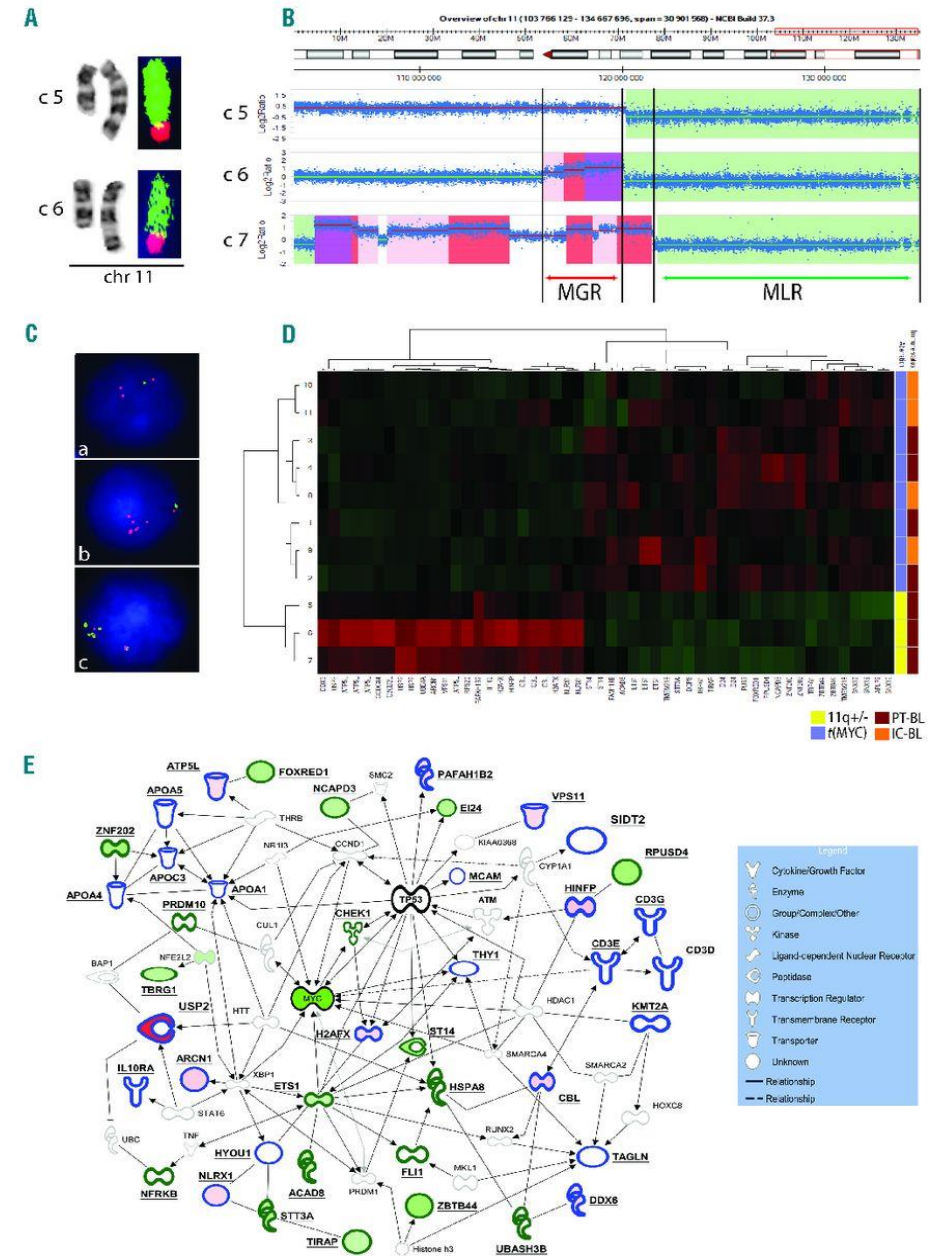
²KU Leuven, University of Leuven, Translational Cell and Tissue Research and KU Leuven, University Hospitals Leuven, Department of Pathology, Belgium.

³KU Leuven, University Hospitals Leuven, Department of Hematology, Belgium.

⁴KU Leuven, University of Leuven, Center for Human Genetics, Belgium iwona.wlodarska@uzleuven.be.

Genomic and transcriptomic data on PT-mBL.

Julio Finalet Ferreiro et al. Haematologica 2015;100:e275-e279



Genomic and transcriptomic data on PT-mBL. (A) Partial karyotype of case 5 (c5) and case 6 (c6) showing chromosome 11 abnormalities and FISH images of both derivative chromosomes painted with WCP11 (green) and WCP8 (red) (case 5), and WCP18 (red) (case 6). The aberrations were eventually described as der(11)(11pter->11q23.3::11q23.3->11q13::8q22q24.3) and der(11)t(11;18)(q23.3;q12), respectively. Note (inverted) duplication of 11q13q23.3 in case 5 and a normal appearance of this region in case 6. (B) Chromosomal view of chromosome 11q23q24 and imbalances identified by array CGH analysis in cases 5–7. Gained regions are highlighted in red-scale (increased intensity reflects an increased amplification level), while lost regions are marked in green. Note a variable level of 11q23.3 gain, a common loss of 11q24qter and the defined MGR (~4 Mb) and MLR (~13.5 Mb). (C) Examples of interphase FISH analysis performed in case 5 (a) and case 6 (b, c). The applied probes include the 11q-MGR/MLR FISH assay (a) (b), and two probes from the duplicated (RP11-284O21-SpectrumOrange) and amplified (RP11-784K23-SpectrumGreen; *Online Supplementary Figure S2*) area in case 6 (c). Note the duplicated and amplified red/11q23.3 signal in (a) and (c), respectively, and loss of the green/11q24 signal in both cases. In (c), note two red and five green signals in the cluster, illustrating various levels of gain within the MGR. (D) Hierarchical clustering of mBL cases using the dysregulated genes located in the MGR and MLR. (E) Interaction network found by Ingenuity Pathway Analysis involving genes targeted by the 11q-gain/loss aberration (bold edges). Solid and interrupted lines represent direct and indirect interactions, respectively. Notably, most of the interactions in this network are direct protein-protein interactions. The molecules with blue and green edges are encoded by MGR- and MLR-associated genes, respectively. The data obtained by comparison of cases 5–7 (PT-mBL with the 11q-gain/loss pattern) (11q+/-) by cases 1–4 (PT-MYC-translocation-positive mBL) [(t(MYC))] were overlaid in this network. Molecules which are down- and up-regulated in cases with 11q+/- when compared to cases with t(MYC) are filled in green and red, respectively.

2)



Integrative Genomic and Transcriptomic Analysis Identified Candidate Genes Implicated in the Pathogenesis of Hepatosplenic T-Cell Lymphoma

Julio Finalet Ferreiro¹, Leila Rouhigharabaei¹, Helena Urbankova¹, Jo-Anne van der Krogt¹, Lucienne Michaux¹, Shashirekha Shetty², Laszlo Krenacs³, Thomas Tousseyn⁴, Pascale De Paepe⁵, Anne Uyttebroeck⁶, Gregor Verhoef⁷, Tom Taghon⁸, Peter Vandenberghe¹, Jan Cools^{1,9}, Iwona Wlodarska^{1*}

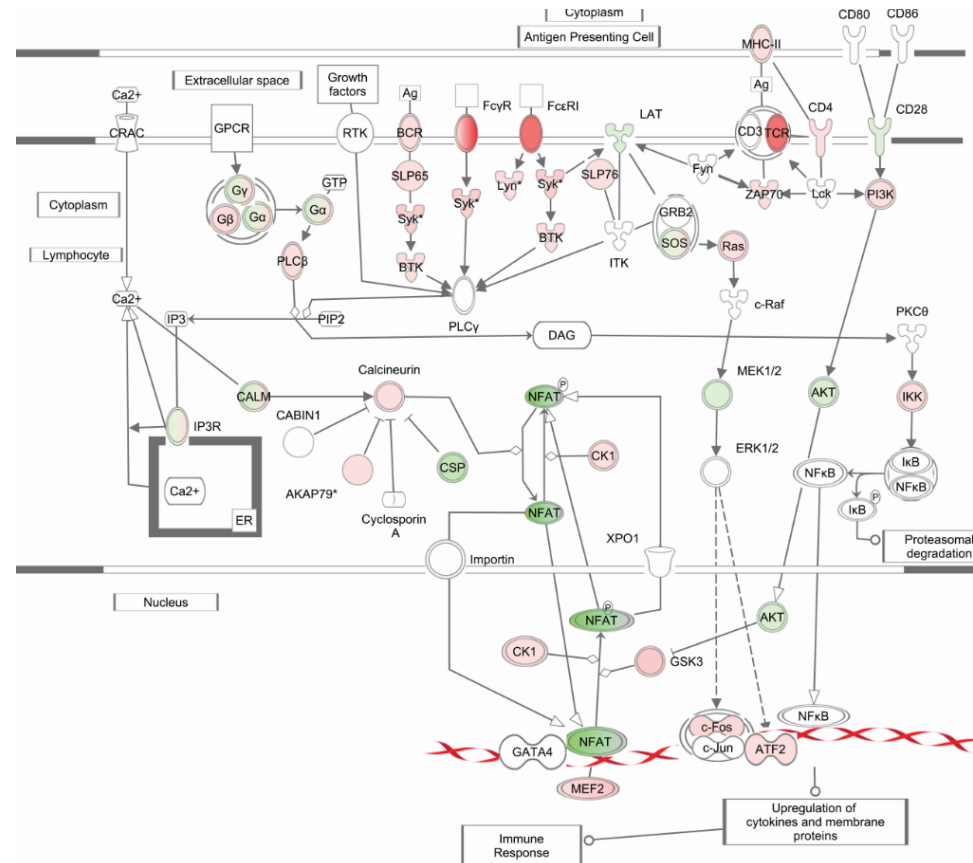


Figure S3

3)

Non-IG Aberrations of *FOXP1* in B-Cell Malignancies Lead to an Aberrant Expression of N-Truncated Isoforms of *FOXP1*

Leila Rouhigharabaei^{1,3}, Julio Finalet Ferreiro^{1,3}, Thomas Tousseyn², Jo-Anne van der Krogt¹, Natalie Put¹, Eugenia Haralambieva³, Carlos Graux⁴, Brigitte Maes⁵, Carmen Vicente^{1,6}, Peter Vandenberghe¹, Jan Cools^{1,6}, Iwona Wlodarska^{1*}

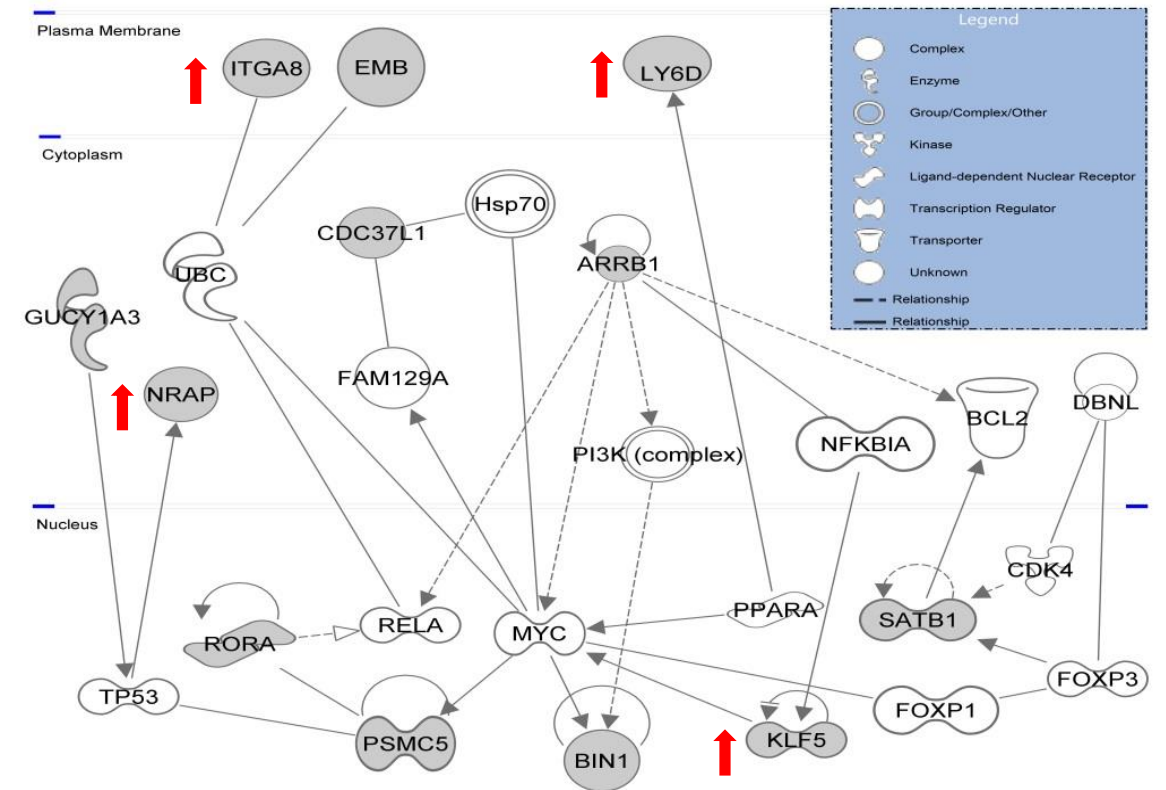
¹ Center for Human Genetics, KU Leuven, Leuven, Belgium, ² Translational Cell and Tissue Research KU Leuven, Department of Pathology UZ Leuven, Leuven, Belgium, ³ Department of Pathology, University of Würzburg, Würzburg, Germany, ⁴ Mont-Godinne University Hospital, Yvoir, Belgium, ⁵ Virga Jesse Hospital, Hasselt, Belgium, ⁶ Center for the Biology of Disease, VIB, Leuven, Belgium

Figure S4.

Interaction network of genes exclusively mutated in t(3;14)-positive case 5 (in grey) with the well know cancer genes specified by IPA. Continuous and discontinues lines indicate direct and indirect interactions, respectively. Red arrows mark genes found to be upregulated in *FOXP1_{FL}* expressing case 5 when compared with cases expressing *FOXP1_{NT}*.

doi:10.1371/journal.pone.0085851.s004

(PPTX)



© 2000-2013 Ingenuity Systems, Inc. All rights reserved.

Summary of the dataset and research objective (1)

Disease and control samples:

Disease: Peripheral T-cell lymphoma :

A rare Lymphoproliferative disorder of T cell origin

Controls: Normal lymph nodes and other T-cell tumor (HSTL)

Research objective:

To gain insight into the biology of the disease

Summary of the dataset and research objective (2)

Research questions:

which:

- molecules (from our dataset) are involved in well characterized biological (canonical) pathways
- canonical pathways are relevant to the disease
- cancer-related genes are relevant for the disease
- genes might define the gene signature of the disease
- genes not present in my dataset are linked, directly or indirectly, to the genes in my dataset
- Publications supports my findings

Dataset after inference analysis with DEseq V2

Tumor_type_A vs Normal Tissue

| ID | Exp Fold Change | Exp False Discovery Rate (q-value) |
|-----------------|-----------------|------------------------------------|
| ENSG00000008311 | -8.838 | 0.011000 |
| ENSG00000108846 | 10.928 | 0.159000 |
| ENSG00000160179 | 10.467 | 0.075000 |
| ENSG00000176244 | -65.239 | 0.002000 |
| ENSG00000103740 | -11.742 | 0.073000 |
| ENSG00000114739 | -10.31 | 0.028000 |
| ENSG00000197381 | -10.105 | 0.004000 |
| ENSG00000152990 | -12.15 | 0.009000 |
| ENSG00000169129 | -7.488 | 0.078000 |
| ENSG00000163568 | 10.918 | 0.084000 |
| ENSG00000129474 | -6.589 | 0.089000 |
| ENSG00000112294 | -6.396 | 0.032000 |
| ENSG00000012779 | 45.744 | 0.023000 |
| ENSG00000139211 | 28.862 | 0.063000 |

Tumor_type_A vs Tumor_type_B

| ID | Exp Fold Change | Exp False Discovery Rate (q-value) |
|-----------------|-----------------|------------------------------------|
| ENSG00000175985 | -2.832 | 0.058 |
| ENSG00000204262 | 2.529 | 0.058 |
| ENSG00000138080 | 2.04 | 0.057 |
| ENSG00000158715 | -2.253 | 0.057 |
| ENSG00000245680 | -2.123 | 0.056 |
| ENSG00000174292 | -2.364 | 0.056 |
| ENSG00000115457 | 2.795 | 0.056 |
| ENSG00000154920 | 2.603 | 0.056 |
| ENSG00000182054 | 2.103 | 0.055 |
| ENSG00000171246 | -2.86 | 0.055 |
| ENSG00000079215 | 2.132 | 0.055 |
| ENSG00000141576 | -2.378 | 0.055 |
| ENSG00000164086 | -2.072 | 0.055 |
| ENSG00000105664 | 2.793 | 0.055 |
| ENSG00000008394 | 2.872 | 0.055 |

Creation of a new project

The image shows the IPA (Ingenuity Pathway Analysis) software interface. The main window has a menu bar (File, Edit, View, Window, Help) and a toolbar with a 'NEW' button. Below the toolbar is a 'Project Manager' sidebar with a tree view showing 'My Projects', 'Shared Projects', and 'Libraries'. A red arrow points to 'My Projects' with the text 'Right-click on "My Projects"'. A context menu is open over 'My Projects', showing 'New Project' and 'VIB11-VIB'. A red arrow points from the 'New Project' option to a 'Create Project' dialog box. The dialog box has fields for 'Workspace' (My Projects), 'Name' (Rare_T-cell-Lymphoma), 'Notes' (3 samples per group), 'Array', and 'Platform' (RNA seq data). At the bottom are 'SAVE' and 'CANCEL' buttons.

IPA IPA

File Edit View Window Help

Genes and Chemicals Diseases and Functions Pathways and Tox Lists

NEW

Project Manager

A-Z SORT SEARCH REFRESH

My Projects

Shared Projects

Libraries

Right-click on "My Projects"

Libraries

VIB11-VIB

New Project

IPA Create Project

Create Project

Workspace: My Projects

Name: Rare_T-cell-Lymphoma

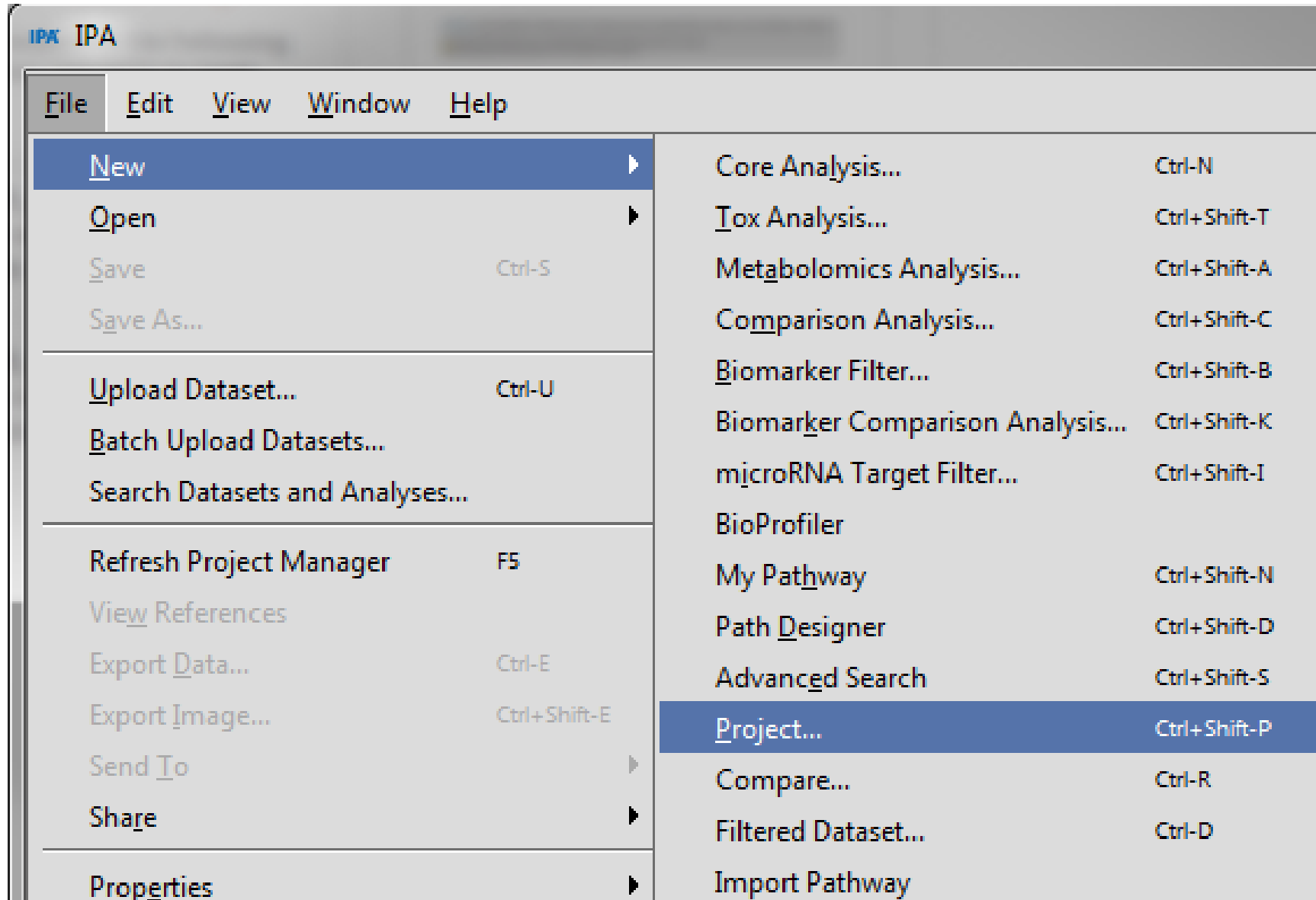
Notes: 3 samples per group

Array:

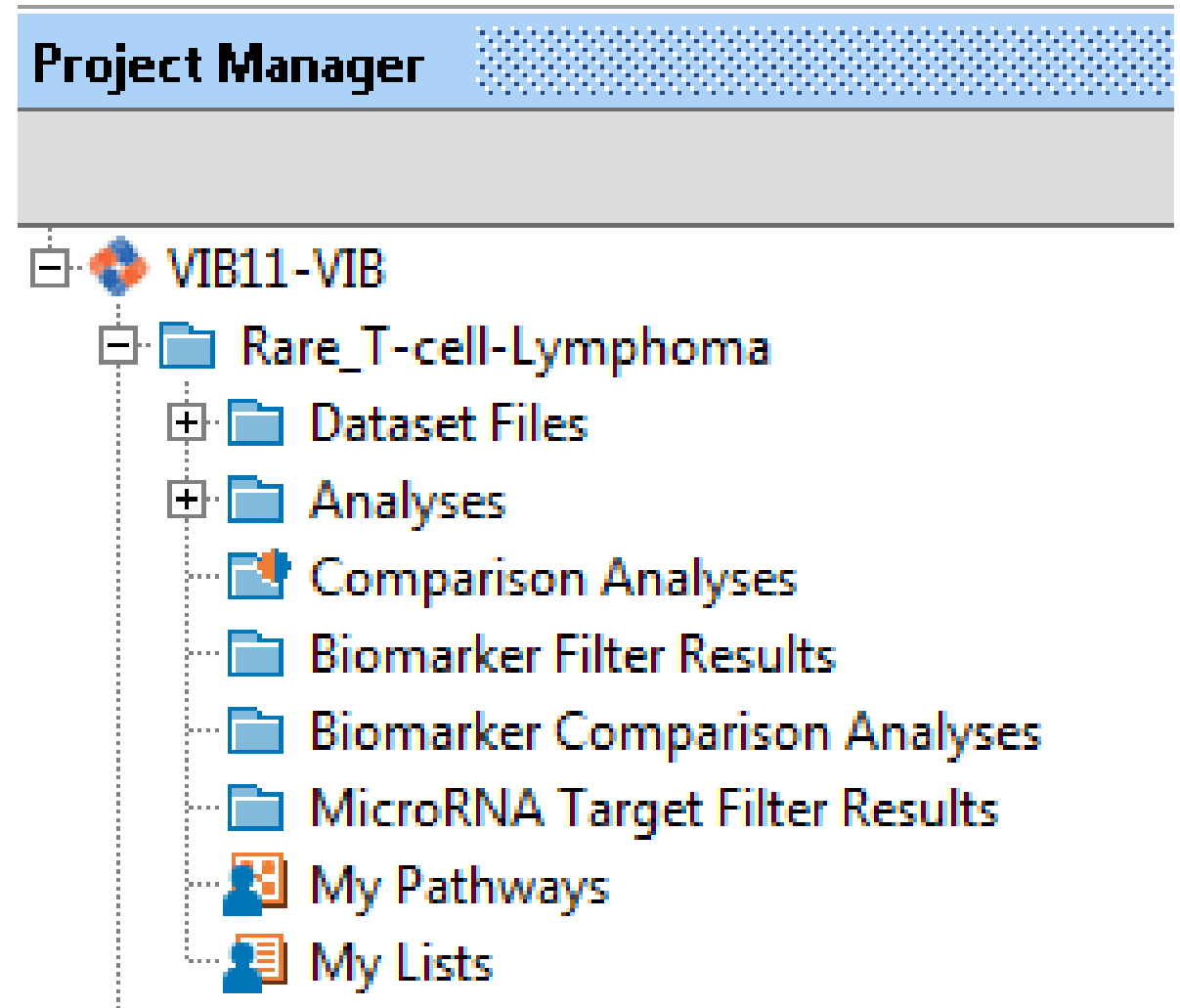
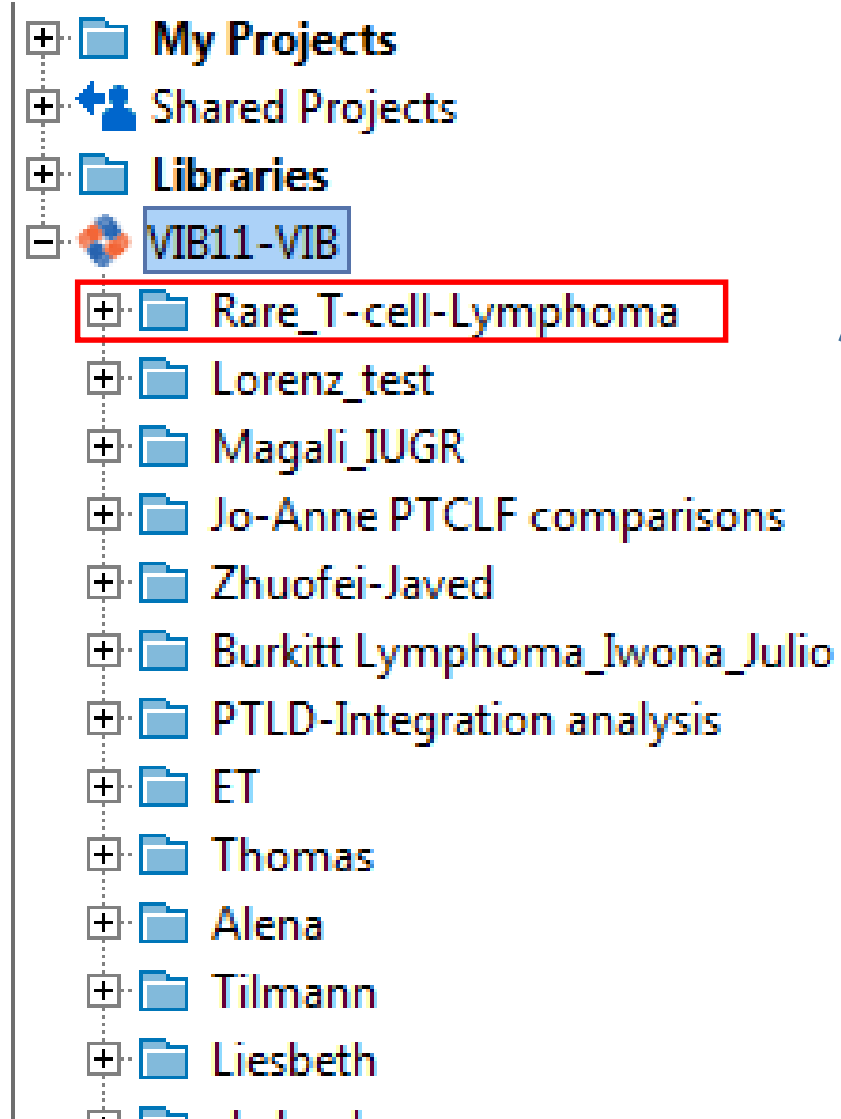
Platform: RNA seq data

SAVE CANCEL

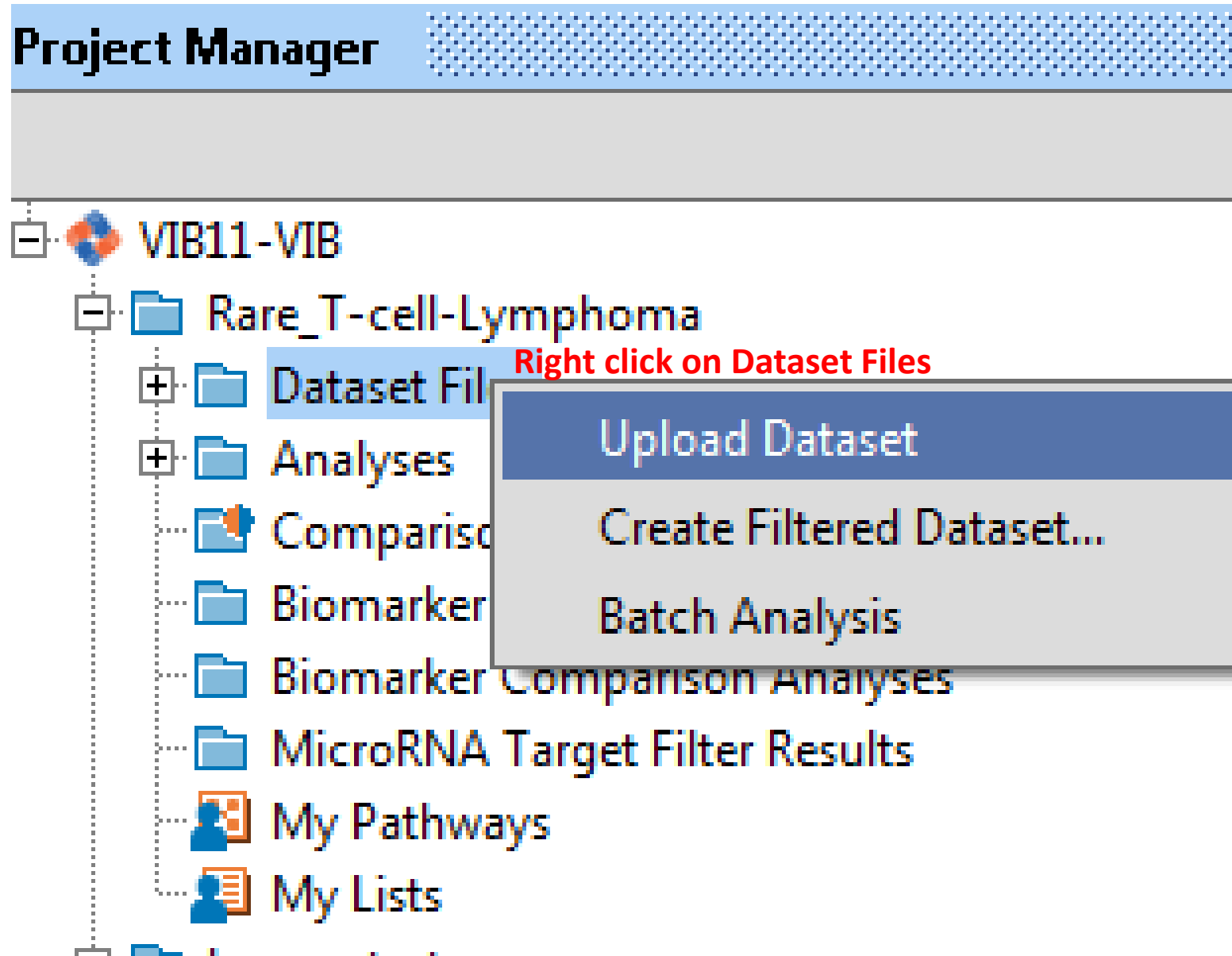
Creation of a new project, alternative



Items in a project



Data upload



Setting up a Core Analysis: General Settings

The screenshot displays the IPA software interface for setting up a core analysis. The 'General Settings' tab is selected, showing various configuration options. The 'Population of genes to consider for p-value calculations' is set to 'Ingenuity Knowledge Base (Genes Only)'. The 'Relationships to consider' are set to 'Direct and Indirect Relationships'. The 'Optional Analyses' section shows a tree structure with 'My Project' and 'My Lists' selected. The 'Analysis Filter Summary' on the right provides a detailed list of included and excluded cell lines and tissues. The 'Set Cutoffs' section shows the 'Exp Fold Change' cutoff set to -1.50972E4 to 1134.0299, with a focus on 'Both Up/Downregulated' molecules. The 'Preview Dataset' section shows a table of 437 analysis-ready molecules, with columns for Exp Fold Change, Exp p-value, Exp False Discovery Rate (q-value), ID, Notes, Symbol, Entrez Gene Name, Location, Type(s), and Drug(s).

General Settings

Population of genes to consider for p-value calculations:
Reference Set: Ingenuity Knowledge Base (Genes Only)

Relationships to consider:
Affects networks and upstream regulator analysis
☒ Direct and Indirect Relationships
☐ Direct Relationships

Optional Analyses:

- ☒ My Project
 - ☒ My Pathways
 - ☒ My Lists
- ☒ VIB11-VIB
 - ☒ My Pathways
 - ☒ My Lists

Analysis Filter Summary

Consider only molecules and/or relationships where (species = Mouse OR Human) AND (confidence = Experimentally Observed OR High (predicted)) AND (tissues/cell lines = Other Neuroblastoma Cell Lines OR UACC-62 OR CD4+ T-lymphocytes OR BDCA-1+ dendritic cells OR Natural T-regulatory cells OR Other Peripheral blood leukocytes OR Other Teratocarcinoma Cell Lines OR Th1 cells OR Cell line not otherwise specified OR Microglia OR Mammary Gland OR Thyroid Gland OR Other Hepatoma Cell Lines OR MCF7 OR Liver OR HOP-92 OR PC-3 OR MDA-MB-435 OR CAKI-1 OR MOLT-4 OR EK VX OR Trachea OR SK-MEL-5 OR Astrocytes OR K-562 OR BA/F3 OR Cartilage Tissue OR CNS Cell Lines not otherwise specified OR NT2/D1 OR MG-63 OR Ovarian Cancer Cell Lines not otherwise specified OR Cortical neurons OR Mast cells OR Pro-B lymphocytes OR Epidermis OR Small Intestine OR BT-549 OR Vascular smooth muscle cells OR Peripheral blood leukocytes not otherwise specified OR 786-0 OR A375 OR Other Monocytes OR Heart OR Teratocarcinoma Cell Lines not otherwise specified OR Osteoblasts OR U251 OR Smooth Muscle OR Pyramidal neurons OR RBL-2H3 OR SNB-19 OR B lymphocytes not otherwise specified OR Uterus OR Colon Cancer Cell Lines not otherwise specified OR Thymus OR Pancreatic Cancer Cell Lines not otherwise specified OR Myeloid dendritic cells OR PBMCs OR NCI-H522 OR Naive B cells OR Other Macrophages OR SN12C OR Dermis OR CD56bright NK cells OR Activated V α 2 Gamma delta T cells OR Effector T cells OR

Set Cutoffs

| Expression Value Type | Cutoff | Range | Focus On |
|------------------------------------|--------|-------------------------|-----------------------|
| Exp Fold Change | | -1.50972E4 to 1134.0299 | Both Up/Downregulated |
| Exp p-value | | 0.0 to 0.005 | |
| Exp False Discovery Rate (q-value) | | 0.0 to 0.243 | |

Preview Dataset example_for_IPA

Analysis-Ready (437) \ Mapped IDs (439) \ Unmapped IDs (7) \ All IDs (446)

ADD TO MY PATHWAY | ADD TO MY LIST | CREATE DATASET | CUSTOMIZE TABLE

| Exp Fold Change | Exp p-value | Exp False Discovery Rate (q-value) | ID | Notes | Symbol | Entrez Gene Name | Location | Type(s) | Drug(s) |
|-----------------|-------------|------------------------------------|-----------------|-------|----------|------------------------------------|---------------------|-------------|-------------------------------------|
| -8.838 | 0.000 | 0.011 | ENSG00000008311 | | AASS | aminoadipate-semialdehyde sy... | Cytoplasm | enzyme | |
| -43.491 | 0.000 | 0.000 | ENSG00000008563 | | ABCB1 | ATP binding cassette subfamily ... | Plasma Membrane | transporter | dofequidar, tariquidar, OC 144-0... |
| 10.928 | 0.003 | 0.159 | ENSG00000108846 | | ABCC3 | ATP binding cassette subfamily ... | Plasma Membrane | transporter | |
| 10.467 | 0.001 | 0.075 | ENSG00000160179 | | ABCG1 | ATP binding cassette subfamily ... | Plasma Membrane | transporter | |
| -65.239 | 0.000 | 0.002 | ENSG00000176244 | | ACBD7 | acyl-CoA binding domain conta... | Other | other | |
| -11.742 | 0.001 | 0.073 | ENSG00000103740 | | ACSBG1 | acyl-CoA synthetase bubblegu... | Cytoplasm | enzyme | |
| -10.310 | 0.000 | 0.028 | ENSG00000114739 | | ACVR2B | activin A receptor type 2B | Plasma Membrane | kinase | |
| 377.444 | 0.000 | 0.000 | ENSG00000134028 | | ADAMDEC1 | ADAM-like, decysin 1 | Extracellular Space | peptidase | |

0 / 437

Notes:
"Bold" - Focus molecules. Gene/Protein/Chemical identifiers that meet the user-defined cutoff and map to the Global Molecular Network are displayed with bold text.

RUN ANALYSIS **CANCEL**

Setting up a Core Analysis: General Settings > select Genes Only

Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B]

General Settings

Networks Interaction

Node Types

Data Sources An Open Access...

Confidence Experimentally Ob...

Species Human, Mouse

Tissues & Cell Lines 786-O...

Mutation All

Population of genes to consider for p-value calculations:

Reference Set: Ingenuity Knowledge Base (Genes Only)

Relationships to consider:

☒ Direct and Indirect Relationships
☐ Direct Relationships Only

Ingenuity Knowledge Base (Endogenous Chemicals Only)
Ingenuity Knowledge Base (Genes + Endogenous Chemicals)
Ingenuity Knowledge Base (Genes Only)
User Dataset

Affymetrix
Agilent
CodeLink
Illumina
Life Technologies (Applied Biosystems)

ADVANCED **SAVE AS DEFAULTS**

Set Cutoffs

Setting up a Core Analysis: Networks

Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B]

General Settings ?

Networks Interaction ?

Node Types ?

Data Sources An Open Access... ?

Confidence Experimentally Ob... ?

Species Human,Mouse ?

Tissues & Cell Lines 786-0... ?

Mutation All ?

Generate the following Networks (increases analysis time)

☒ **Interaction networks**

☐ Include endogenous chemicals Molecules per network Networks per analysis

Genes are always included 35 ▼ 25 ▼

☐ **Causal networks**

Score master regulators for relationships to diseases, functions, genes, or chemicals (max 50)

☐ Score using causal paths only

ADVANCED

SAVE AS DEFAULTS

Setting up a Core Analysis: Nodes types, select all

The image shows a software interface for creating a core analysis. The title bar reads "Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B_final]".

Left Panel (Settings):

- General Settings
- Networks Interaction
- Node Types** (highlighted with a red box)
- Data Sources An Open Access...
- Confidence Experimentally Ob...
- Species Human, Mouse
- Tissues & Cell Lines All
- Mutation All

Buttons: ADVANCED, SAVE AS DEFAULTS

Node Types List (Left):

- ☐ Select all
- ☐ biologic drug
- ☐ chemical - endogenous mammalian
- ☐ chemical - endogenous non-mammalian
- ☐ chemical - kinase inhibitor
- ☐ chemical - other
- ☐ chemical - protease inhibitor
- ☐ chemical drug
- ☐ chemical reagent
- ☐ chemical toxicant
- ☐ complex
- ☐ cytokine
- ☐ disease
- ☐ enzyme
- ☐ function
- ☐ G-protein coupled receptor

Node Types List (Right):

- ☒ Select all
- ☒ biologic drug
- ☒ chemical - endogenous mammalian
- ☒ chemical - endogenous non-mammalian
- ☒ chemical - kinase inhibitor
- ☒ chemical - other
- ☒ chemical - protease inhibitor
- ☒ chemical drug
- ☒ chemical reagent
- ☒ chemical toxicant
- ☒ complex
- ☒ cytokine
- ☒ disease
- ☒ enzyme
- ☒ function
- ☒ G-protein coupled receptor

Setting up a Core Analysis: Data sources, select all

Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B]

General Settings ?

Networks Interaction ?

Node Types All ?

Data Sources An Open Access... ?

Confidence Experimentally Ob... ?

Species Human, Mouse ?

Tissues & Cell Lines 786-0... ?

Mutation All ?

- ☐ Select all
- ☒ Ingenuity Expert Information
 - ☒ Ingenuity Expert Findings
 - ☒ Ingenuity ExpertAssist Findings
- ☐ Ingenuity Supported Third Party Information
 - ☐ MicroRNA-mRNA interactions
 - ☒ Protein-protein interactions
 - ☒ Additional sources

Setting up a Core Analysis: Confidence

Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B]

General Settings ?

Networks Interaction ?

Node Types All ?

Data Sources An Open Access... ?

> Confidence Experimentally Ob... ?

Species Human,Mouse ?

Tissues & Cell Lines 786-0... ?

Mutation All ?

ADVANCED **SAVE AS DEFAULTS**

☐ Select all

☒ Experimentally Observed

☒ High (predicted)

☐ Moderate (predicted)

Setting up a Core Analysis: Species, select human and mouse

Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B]

| | |
|---------------------------------|---|
| General Settings | ? |
| Networks Interaction | ? |
| Node Types All | ? |
| Data Sources An Open Access... | ? |
| Confidence Experimentally Ob... | ? |
| > Species Human,Mouse | ? |
| Tissues & Cell Lines 786-0... | ? |
| Mutation All | ? |

ADVANCED SAVE AS DEFAULTS

☐ Select all

- ☐ Mammal
 - ☒ Human
 - ☒ Mouse
 - ☐ Rat
- ☐ Uncategorized

Setting up a Core Analysis: Tissue and cell lines; select all

Create Core Analysis - [analysis : tumor_type_A_vs_tumor_type_B]

General Settings ?

Networks Interaction ?

Node Types All ?

Data Sources An Open Access... ?

Confidence Experimentally Ob... ?

Species Human,Mouse ?

> Tissues & Cell Lines Activated ... ?

Mutation All ?

☐ Select all

☒ Tissues and Primary Cells

- ☒ Tissues and Primary Cells not otherwise specified
- ☒ Cells
- ☒ Nervous System
- ☒ Organ Systems
- ☒ Other Tissues and Primary Cells

☒ Cell Line

ADVANCED

SAVE AS DEFAULTS

Advanced Settings

Select measurement for node coloring: Exp Fold Change

This measurement will be used to calculate directionality (z-scores) in the analysis and will be displayed in color on pathways and networks. If you choose a non-directional measurement (e.g. p-value) then z-scores will not be calculated.

Duplicate Resolution

When IDs map to the same gene, protein, or other molecule:

Apply cutoffs before consolidating IDs: Yes (recommended)

Measurement for resolving duplicates: Exp Fold Change

Consolidate IDs using the measurement value: maximum

For maximum or minimum, the absolute value is used except for Loss/Gain measurements.

OK CANCEL