Ingenuity Pathway Analysis (IPA): Maximizing the Biological Interpretation of Gene, Transcript & Protein Expression Data with IPA

Joe Salvatore  
Account Manager  
Qiagen Advanced Genomics  
Joe.Salvatore@qiagen.com

Dev Mistry, Ph.D.  
Field Applications Scientist  
Qiagen Advanced Genomics  
Devendra.Mistry@qiagen.com
Introduction
When do you use IPA?

QIAGEN Sample to Insight

Sample Prep → Assay Data → Sequence-Level Statistics → Biology of Interest (Genes, Variants, etc.) → Annotation & Comparative (Statistical) Analysis → Annotation & Biological Interpretation

Upstream Analysis → ‘Primary’ → ‘Secondary’ → ‘Tertiary’

Sample Insight

INGENIUTY VARIANT ANALYSIS

INGENIUTY PATHWAY ANALYSIS

Clinical Insight

HGMD®
Qiagen software allow you to perform a wide range of analyses

- **Gene Expression Analysis**
- **Variant Calling and Causal Variant Identification**
- **De Novo Assembly and Genome Finishing**

**Gene, Transcripts, Proteins, Metabolomics Data Interpretation**

**Epigenomics Analysis**
**ChIP/Histone/Bisulfite Seq**

**Metagenomics Analysis**

**Mol. Bio/Classical Sequencing**
High Impact Disease-Centric Studies

**OncoLand**

**Body Map (GTEx)**
- Microarray and RNA-Seq for Normal Tissue Panel
- 9500+ samples
- 200+ patients
- 50+ distinct tissues

**Immunoland**
- Infectious Diseases
- Alzheimer
- Arthritis
- Asthma
- COPD
- Crohn's Disease

**CVMLand**
- Diabetes
- Obesity
- Heart Failure
- Hypertension...
Sample to Insight: Secondary, Tertiary and Multi-omics analysis

RNA Seq  DNA Seq

Input Files: .fastq, .chp

Data Analysis

Input Files: .xls, .txt, .diff

INGENUITY PATHWAY ANALYSIS

Input Files: .vcf

INGENUITY VARIANT ANALYSIS

Interpretation
Sample to Insight: Secondary, Tertiary and Multi-omics analysis

**Secondary Analysis**
- Fastq to differential expression
- Heatmap clustering
- PCA/PCOA
- Volcano Plot
- Venn Diagram

**Tertiary Interpretation**
- Pathways enrichment
- Diseases and biological processes enrichment
- Activity and enrichment comparison
- Knowledgebase mining

**Datasets Mining and Management**
- Gene expression pattern across wide range of projects and datasets
- Correlation between gene expression/mutation/survival rates etc.
- Transcript variants and gene fusions

**Input Files:** `.fastq`, `.chp`

**Input Files:** `.xls`, `.txt`, `.diff`

---

**Data Analysis**

---

**INGENUITY PATHWAY ANALYSIS**

---

**OmicsSoft Corporation**
What is Ingenuity Pathway Analysis?
What can IPA do?

**Dataset**

- RNA seq, microarray, qPCR, proteomics, metabolomics etc.
- Core analysis, Iso/Bioprofiler etc.

**No Dataset**

- SNAI1? Cell Migration? Glioblastoma?
- Searching and exploring, Iso/Bioprofiler etc.

No dataset

Accessing knowledge base to learn about a gene, phenotype or disease
The Ingenuity Knowledge Base

The Ingenuity Ontology

Ingenuity Content

Ingenuity Findings

Ingenuity® Expert Findings – Manually curated Findings that are reviewed, from the full-text, rich with contextual details, and are derived from top journals.

Ingenuity® ExpertAssist Findings – Automated text Findings that are reviewed, from abstracts, timely, and cover a broad range of publications.

Ingenuity Modeled Knowledge

Ingenuity® Expert Knowledge – Content we model such as pathways, toxicity lists, etc.

Ingenuity® Supported Third Party Information – Content areas include Protein-Protein, miRNA, biomarker, clinical trial information, and others

Species: human, mouse and rat
Data from other species can be mapped to human, mouse and rat orthologues
Species Supported

- **Human, Mouse, Rat in full content**
- **IPA uses HomoloGene to map other identifiers to human/mouse/rat orthologs** (though supporting content for the additional species will be specific to human, mouse, and rat)
  - Arabidopsis thaliana
  - Bos taurus (bovine)
  - Caenorhabditis elegans
  - Gallus gallus (chicken)
  - Pan troglodytes (chimpanzee)
  - Danio rerio (zebrafish)
  - Canis lupus familiaris (canine)
  - Drosophila melanogaster
  - Macaca mulatta (Rhesus Monkey)
  - Saccharomyces cerevisiae
  - Schizosaccharomyces pombe
Ingenuity Curated “Findings”

- **Species**
- **Synonyms**
- **Experimental method**
- **Direction of change**
- **Tissue context**
- **Cell line context**
- **Original source**
- **Post-translational modification**

**Mutant mouse Cystatin c [Cst3] gene (homozygous knockout) in mouse increases growth of pancreatic carcinoma in mouse that involves transgenic SV40 viral SV40 large T-antigen [Large T Antigen] protein in mouse pancreas.**

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Ingenuity Expert Findings</td>
</tr>
</tbody>
</table>

**Enzastaurin, an inhibitor of human AKT3 protein, is in Phase II clinical trial to treat mammary neoplasm.**

ClinicalTrials.gov, Invest New Drugs 2002 08 1;20(3):241-51.

Species: human, mouse and rat
Data from other species can be mapped to human, mouse and rat orthologues
Peer-reviewed publications citing QIAGEN’s Ingenuity products

16,047 publications and growing!

As of 8/26/2016

Publication Count

As of 8/26/2016

<table>
<thead>
<tr>
<th>Year</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>1</td>
</tr>
<tr>
<td>2004</td>
<td>18</td>
</tr>
<tr>
<td>2005</td>
<td>104</td>
</tr>
<tr>
<td>2006</td>
<td>297</td>
</tr>
<tr>
<td>2007</td>
<td>686</td>
</tr>
<tr>
<td>2008</td>
<td>1306</td>
</tr>
<tr>
<td>2009</td>
<td>2306</td>
</tr>
<tr>
<td>2010</td>
<td>3786</td>
</tr>
<tr>
<td>2011</td>
<td>5608</td>
</tr>
<tr>
<td>2012</td>
<td>7704</td>
</tr>
<tr>
<td>2013</td>
<td>9972</td>
</tr>
<tr>
<td>2014</td>
<td>12513</td>
</tr>
<tr>
<td>2015</td>
<td>14861</td>
</tr>
<tr>
<td>2016</td>
<td>16,047</td>
</tr>
</tbody>
</table>
Autophagy Suppresses Tumorigener through Elimination of p62

Robin Mathew,1,5,6 Cristina M. Karp,3,5,6 Brian Beaudoin,2,5,6 Nhan Vuong,3 Guanghua Chen,3,4 Anupama Reddy,6 Gyan Bhanot,3,5,6 Celine Gelinas,1,2 Robert S. DiPaola,4,5 Vassiliki Karantali,4,5 and Eileen White1,2,3,5,6

1University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, NJ 08854, USA  
2Center for Advanced Biotechnology and Medicine, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854, USA  
3Division of Molecular Biology and Biochemistry, Rutgers University, 604 Allison Road, Piscataway, NJ 08854, USA  
4Division of Medical Oncology, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School,  
675 Hoes Lane, Piscataway, NJ 08854, USA  
5The Cancer Institute of New Jersey, 195 Little Albany Street, New Brunswick, NJ 08903, USA

As of 8/26/2016
How can IPA help you?

Biological understanding of large data sets

- Differential gene expression, array and RNA-seq (transcriptomics)
- Differential protein expression (proteomics)
- Metabolomics
- miRNA expression
- Gene List
  - Chip-seq
  - siRNA screening
- Methylation
- Protein phosphorylation
What are the upstream molecular events?

Can we form new hypothesis/conclusions?

Can we do multi-omics analysis?

How are these events different among the samples?
Uploading your dataset
Suggested format for uploading RNA-seq data

Accepted Files Formats
- .txt (tab delimited text files)
- .xls (excel files)
- .diff (cuffdiff files)

Max RPKM in Excel

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>RPKM (Experimental)</th>
<th>RPKM (Control)</th>
<th>RPKM (MAX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYC</td>
<td>500</td>
<td>1</td>
<td>1 =MAX(B2:C2)</td>
</tr>
<tr>
<td>SNAI1</td>
<td>1000</td>
<td>300</td>
<td>1000</td>
</tr>
<tr>
<td>CDH1</td>
<td>14</td>
<td>2000</td>
<td>2000</td>
</tr>
</tbody>
</table>
Dataset with Multiple Observations

Observation = experimental group (time point, treatment, cell line etc.)

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDs</td>
<td>12hr Fold Ch</td>
<td>24hr Fold Ch</td>
<td>48hr Fold Ch</td>
<td>12hr p-val</td>
<td>24hr p-val</td>
<td>48hr p-val</td>
<td>48hr p-val</td>
</tr>
<tr>
<td>ENST0000031</td>
<td>-4.2142056</td>
<td>-3.8649127</td>
<td>-2.0739169</td>
<td>-1.368</td>
<td>4.87E-02</td>
<td>1.147</td>
<td>4.52E-02</td>
</tr>
<tr>
<td>ENST0000026</td>
<td>1.2606281</td>
<td>1.4772771</td>
<td>1.20865587</td>
<td>-1.309</td>
<td>3.07E-02</td>
<td>1.55</td>
<td>1.05E-02</td>
</tr>
<tr>
<td>ENST0000023</td>
<td>-1.9114326</td>
<td>1.14430339</td>
<td>-1.1000889</td>
<td>1.324</td>
<td>2.66E-02</td>
<td>1.401</td>
<td>1.96E-03</td>
</tr>
<tr>
<td>ENST0000033</td>
<td>-3.1518905</td>
<td>-2.0669915</td>
<td>5.91865976</td>
<td>-1.056</td>
<td>5.01E-03</td>
<td>1.201</td>
<td>1.57E-02</td>
</tr>
<tr>
<td>ENST0000042</td>
<td>19.9882489</td>
<td>10.5381272</td>
<td>18.594255</td>
<td>1.18</td>
<td>8.43E-03</td>
<td>1.201</td>
<td>1.57E-02</td>
</tr>
<tr>
<td>ENST0000043</td>
<td>3.40116856</td>
<td>6.25461217</td>
<td>1.29044621</td>
<td>1.18</td>
<td>8.43E-03</td>
<td>1.201</td>
<td>1.57E-02</td>
</tr>
<tr>
<td>ENST0000058</td>
<td>1.353577</td>
<td>1.34114811</td>
<td>1.46894852</td>
<td>1.204</td>
<td>4.09E-04</td>
<td>1.362</td>
<td>4.50E-02</td>
</tr>
<tr>
<td>ENST0000051</td>
<td>6.25461217</td>
<td>1.29044621</td>
<td>1.29044621</td>
<td>1.376</td>
<td>2.10E-02</td>
<td>1.392</td>
<td>1.32E-02</td>
</tr>
<tr>
<td>ENST0000057</td>
<td>5.91865976</td>
<td>5.91865976</td>
<td>5.91865976</td>
<td>1.224</td>
<td>2.14E-02</td>
<td>1.224</td>
<td>2.14E-02</td>
</tr>
<tr>
<td>ENST0000052</td>
<td>1.7919726</td>
<td>1.7919726</td>
<td>1.7919726</td>
<td>1.224</td>
<td>2.14E-02</td>
<td>1.224</td>
<td>2.14E-02</td>
</tr>
<tr>
<td>ENST0000036</td>
<td>-1.0449733</td>
<td>-2.3454026</td>
<td>-1.1711674</td>
<td>-1.211</td>
<td>3.19E-02</td>
<td>1.142</td>
<td>2.54E-02</td>
</tr>
<tr>
<td>ENST0000046</td>
<td>-1.221893</td>
<td>1.05473552</td>
<td>1.09623163</td>
<td>1.369</td>
<td>1.14E-02</td>
<td>1.142</td>
<td>2.54E-02</td>
</tr>
<tr>
<td>ENST0000021</td>
<td>1.37791758</td>
<td>1.55958078</td>
<td>-1.0968653</td>
<td>1.296</td>
<td>6.16E-03</td>
<td>1.637</td>
<td>1.65E-03</td>
</tr>
<tr>
<td>ENST0000029</td>
<td>3.05193778</td>
<td>2.62791144</td>
<td>1.46286398</td>
<td>1.317</td>
<td>1.81E-02</td>
<td>1.121</td>
<td>4.10E-02</td>
</tr>
<tr>
<td>ENST0000034</td>
<td>4.24606155</td>
<td>-3.7234355</td>
<td>-2.5639044</td>
<td>1.317</td>
<td>1.81E-02</td>
<td>1.121</td>
<td>4.10E-02</td>
</tr>
<tr>
<td>ENST0000049</td>
<td>1.77642563</td>
<td>2.13692829</td>
<td>1.30861714</td>
<td>-2.187</td>
<td>3.00E-02</td>
<td>-3.93</td>
<td>4.6E-02</td>
</tr>
</tbody>
</table>
Dataset Upload - Claudin Low.xls

1. Select File Format: Flexible Format
2. Contains Column Header: Yes
3. Select Identifier Type: RefSeq
4. Array platform used for experiments: Not specified/applicable
5. Use the dropdown menus to specify the columns that contain identifiers and observations. For observations, select the appropriate expression value type.

Raw Data (29074) \ Dataset Summary (29073)

ID | Observation 1 | Observation 1 | Observation 1 |
---|---------------|---------------|---------------|
1  | NM_130786     | 1.102         | 8.68E-01      |
2  | NR_015380     | -1.992        | 2.24E-01      |
3  | NM_138932     | -1.012        | 9.83E-01      |
4  | NM_014576     | -1.011        | 9.85E-01      |
5  | NM_138933     | 1.016         | 9.79E-01      |
6  | NM_000814     | -27.663       | 1.02E-01      |
7  | NR_026971     | -1.587        | 1.30E-01      |
8  | NM_144670     | -62.184       | 3.47E-01      |
9  | NM_001080438  | -3.912        | 3.47E-01      |
10 |               |               |               |

More Info

SAVE & CREATE ANALYSIS  CANCEL  HELP
Large dataset analysis using IPA
Check whether molecules from your dataset belong to pre-defined pathways
Predict the pathway activation/inhibition

Canonical Pathways

Upstream Analysis

Diseases and Functions

Regulator Effect

Networks
IPA Core Analysis

- **Canonical Pathways**
- **Upstream Analysis**
  - Predict what regulators caused changes in gene expression
  - Predicts directional state of regulator
- **Diseases and Functions**
- **Regulator Effect**
- **Networks**
Upstream Regulators, Mechanistic Networks and Causal Networks

**Upstream Regulators**

**Mechanistic Networks**

**Advanced Analytics: Causal Network Analysis**

Master Regulators

Upstream Regulators

Your Dataset Genes/Molecules
IPA Core Analysis

- Canonical Pathways
- Upstream Analysis
- Diseases and Functions
- Regulator Effect
- Networks

Diseases & Functions

- Predicts effected biology based on gene expression and predicts directional change on that effect
  - “Increase in EMT”
  - “Decrease in proliferation”

Sample to Insight

Diseases & Functions

- ITGB1
- CDH1
- SNAI2
- CLDN4
- TGFβ1
- ADIPOQ
- PTEN

Invasion of tumor cells
Models pathway interactions from predicted upstream regulators, through differentially expressed genes, to biological processes.
IPA Core Analysis

- Canonical Pathways
- Upstream Analysis
- Diseases and Functions
- Regulator Effect

Networks: Predicts non-directional gene interaction map
P value and Z Score
Genes in the reference universe

Overlapping P-value

Genes from previous literature that belong to
• A canonical pathway OR
• Downstream of an upstream regulator OR
• Upstream of a disease or function

- Different from the “Expression P-value” uploaded with your dataset
- Calculated using Fisher’s exact test
- The statistical test looks for an unexpectedly large overlap given the number of molecules in each category
- p-values should be significant (<0.05) for random datasets
- Gene expression direction is not taken into account for this calculation
Overlapping P-value: Only comparing names NOT the direction of gene expression

Genes in the reference universe

Overlapping Molecules

Genes in your dataset

B
C
D
E
F
G
K
L
M
N
O
P

Genes in a Canonical Pathway from Knowledge-base (literature)

A
B
C
E
H
I
J

*Alternatively these can be genes in an upstream network or disease/bio function
Overlapping P-value: Only comparing names NOT the direction of gene expression.

Genes in the reference universe

Overlapping Molecules

Genes in your dataset
- B (↓)
- C (↑)
- D
- E (↑)
- F
- G
- K
- L
- M

Genes in a Canonical Pathway from Knowledge-base (literature)
- A
- B (↓)
- C (↓)
- E (↑)
- H
- I
- J

*Direction of gene expression for Gene C is different between your dataset and literature.*
Z-score: Activation Prediction

Knowledgebase (literature)

Activated

↑
↓
↑
↑
↑
↑

Your dataset

Activated (?)

↑
↓
↑
↑
↑
↑

↑
↓
↑
↑
↑
↑

↓
↓
↓
↓
↓
↓
Z-score: Activation Prediction

- - + + + + + + + 

\[ z = \frac{x}{\sigma_x} = \frac{\sum_i x_i}{\sqrt{N}} = \frac{N_+ - N_-}{\sqrt{N}} \]

= \frac{(7-1)}{\sqrt{8}} = 2.12 (= predicted activation)

Gene expression from Knowledge Base (literature)

\[ \begin{align*}
\downarrow & \quad \downarrow & \quad \uparrow & \quad \uparrow & \quad \uparrow & \quad \uparrow & \quad \uparrow & \quad \downarrow \\
1 & \quad 1 & \quad 1 & \quad 1 & \quad 1 & \quad 1 & \quad 1 & \quad -1 
\end{align*} \]

Gene expression in your dataset

\[ \begin{align*}
\downarrow & \quad \downarrow & \quad \uparrow & \quad \uparrow & \quad \uparrow & \quad \uparrow & \quad \uparrow & \quad \downarrow \\
1 & \quad 1 & \quad 1 & \quad 1 & \quad 1 & \quad 1 & \quad 1 & \quad -1 
\end{align*} \]

\[ +1 \text{ score for the consistent and } -1 \text{ for the inconsistent relationships} \]

• z-score is a statistical measure of the match between expected relationship direction and observed gene expression
• z-score > 2 or < -2 is considered significant
• Note that the actual z-score is weighted by the underlying findings, the relationship bias, and dataset bias
Comparison Analysis: Comparing Multiple Experimental Groups

Sample to Insight
Multi-Omics: Variant Data Overlay
## Variant Analysis in IPA

<table>
<thead>
<tr>
<th>Upstream Regulator</th>
<th>Exp Fold Change</th>
<th>Variant Gain/Loss</th>
<th>Molecule Type</th>
<th>p-value of overlap</th>
<th>Target molecules in...</th>
<th>Mechanistic Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1QBPR</td>
<td>+1.034</td>
<td></td>
<td>transcription regulator</td>
<td>5.1E-02</td>
<td>C1QBPR, CTNNB1</td>
<td>all 7</td>
</tr>
<tr>
<td>SIN3B</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>1.88E-03</td>
<td>CDK4, HES1, KL69</td>
<td>all 10</td>
</tr>
<tr>
<td>SPI1</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>5.52E-05</td>
<td>ALDH3A2, RACE2</td>
<td>all 118</td>
</tr>
<tr>
<td>HTT</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>3.43E-05</td>
<td>AGFRN, AZIN1, C...</td>
<td>all 22</td>
</tr>
<tr>
<td>TCF4</td>
<td>+1.458</td>
<td></td>
<td>transcription regulator</td>
<td>4.38E-02</td>
<td>FNL, LF1, NCOR2</td>
<td>all 2</td>
</tr>
<tr>
<td>SMAD5</td>
<td>-1.000</td>
<td></td>
<td>transcription regulator</td>
<td>1.10E-07</td>
<td>COL4A1, COL6A3, F...</td>
<td>all 9</td>
</tr>
<tr>
<td>STAT1</td>
<td>+1.065</td>
<td></td>
<td>transcription regulator</td>
<td>4.09E-02</td>
<td>AKT1, GBP2, PDGFA</td>
<td>all 7</td>
</tr>
<tr>
<td>NOTCH1</td>
<td>+1.046</td>
<td></td>
<td>transcription regulator</td>
<td>1.53E-02</td>
<td>FNL, HES1, PDGFB</td>
<td>all 3</td>
</tr>
<tr>
<td>KDM5A</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>1.33E-02</td>
<td>HMOX1, ITGB1</td>
<td>all 7</td>
</tr>
<tr>
<td>HDAC7</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>3.55E-04</td>
<td>CCNB1, CTNNB1</td>
<td>all 10</td>
</tr>
<tr>
<td>KLF4</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>2.74E-03</td>
<td>C1QBPR, DCTN1, FNL</td>
<td>all 12</td>
</tr>
<tr>
<td>NF2F2L</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>4.01E-02</td>
<td>HMOX1, PSMA7</td>
<td>all 7</td>
</tr>
<tr>
<td>NF2F2L</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>1.44E-02</td>
<td>DSP, FNL, ITGB1</td>
<td>all 3</td>
</tr>
<tr>
<td>TRIM21</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>6.04E-03</td>
<td>AGFRN, GBP2, PSMA6</td>
<td>all 5</td>
</tr>
<tr>
<td>TCF7L2</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>4.55E-02</td>
<td>ALDH3A2, PC4CAM</td>
<td>all 2</td>
</tr>
<tr>
<td>PCGFI</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>8.35E-03</td>
<td>CDK4, HES1, NOTCH1</td>
<td>all 3</td>
</tr>
<tr>
<td>NFIC</td>
<td>+1.000</td>
<td></td>
<td>transcription regulator</td>
<td>3.52E-02</td>
<td>CCNB1, PDGFA</td>
<td>all 2</td>
</tr>
<tr>
<td>NF2F3A</td>
<td>+2.720</td>
<td></td>
<td>transcription regulator</td>
<td>5.09E-08</td>
<td>AKT1, ACC20, CC1</td>
<td>all 21</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>+1.061</td>
<td></td>
<td>transcription regulator</td>
<td>3.29E-08</td>
<td>ACTN4, ACTR3, C...</td>
<td>all 22</td>
</tr>
<tr>
<td>SOX31</td>
<td>+1.347</td>
<td></td>
<td>transcription regulator</td>
<td>3.75E-02</td>
<td>CDC24BP8, CTNNB1</td>
<td>all 2</td>
</tr>
</tbody>
</table>
IVA export to IPA

Sample to Insight

Title, Location, Date
Transcriptomics and Proteomics profiles have similar enrichment and activation.
### Variant Analysis in IPA

#### EEC-Variant-RPKM-20

**Upstream Regulators**
- **C1QBP**: 1.000 (transcription regulator)
- **SD1B**: 1.000 (transcription regulator)
- **SP1**: 0.000 (transcription regulator)
- **HTT**: 0.000 (transcription regulator)
- **TCF4**: 1.000 (transcription regulator)
- **SPDEF**: <1.000 (transcription regulator)
- **STAT3**: >1.000 (transcription regulator)

**Target molecules in...**
- C1QBP, C1QBP, CTNNB1
- CDK4, HES1, KLF6
- ALDH1A2, TRACER, ...all 100 (7)
- AIPRN, A2IZ1, C1, ...all 22, 116 (14)
- COL4A1, COL6A3, ...all 7

#### My Pathways

**Overlay**
- **SPDEF**

**Options**
- Display Expression Bar Chart
- Display All Exp Values from Dataset
- Display Isoform information

**Display measurement types**

**Value**
- Variant Gain/Loss
- Variant ACMG Classification
- Exp Fold Change

**Range**
- 0.000

**Current Overlay**
- EEC-Variant-RPKM-20

---

**Sample to Insight**
Gene-view, Isoform-view and Relationship-view
Gene-view and Relationship-view Pages
Search and Explore
(Data not required)
Case Study: SNAI1 knockdown making cells less invasive

Blue = healthy cells

Red = invasive cells

Knockdown SPDEF

Overexpress SPDEF
Bioprofiler and Isoprofiler

(Data not required)
Profile genes, diseases and phenotypes
Find significant isoforms in your RNA-seq with IsoProfiler

Dataset chooser

Genes and their isoforms from the dataset

Isoform Filters

Isoform details on the selected gene
When the identity of the dominantly expressed isoform for a gene differs between the expressed and the control

<table>
<thead>
<tr>
<th>isoform</th>
<th>Max intensity</th>
<th>Tumor intensity</th>
<th>adjacent normal intensity</th>
<th>fold change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1000</td>
<td>100</td>
<td>1000</td>
<td>-10</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>1000</td>
<td>100</td>
<td>10</td>
</tr>
</tbody>
</table>

A is dominant in the adjacent normal, while B is dominant in the tumor = switching
Phospho-proteomics
1) Visualize multiple differentially phosphorylated sites (phospho peptides) on networks and pathways.

![Phospho Fold Change Diagram](image)

<table>
<thead>
<tr>
<th>Protein IPI number</th>
<th>SILAC Ratio H/L Normalized</th>
<th>Average H/L ratio significance</th>
<th>Modified Peptide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPI00421063</td>
<td>11.013</td>
<td>1.86955E-27</td>
<td><em>NINENPVS(ph)QNLK</em></td>
</tr>
<tr>
<td>IPI00421065</td>
<td>10.151</td>
<td>7.21401E-20</td>
<td><em>AS(ph)LEIGFGK</em></td>
</tr>
<tr>
<td>IPI00648068</td>
<td>10.045</td>
<td>1.3945E-25</td>
<td><em>S(ph)LESVLSLGPRPT(ph)GGGSSPPELR</em></td>
</tr>
<tr>
<td>IPI00132481</td>
<td>9.5872</td>
<td>4.97636E-19</td>
<td><em>ac(TAAAPS(ph)QQRPAAR</em></td>
</tr>
<tr>
<td>IPI00857376</td>
<td>9.4597</td>
<td>2.12202E-24</td>
<td><em>QALEQS(ph)AEGLIR</em></td>
</tr>
<tr>
<td>IPI00753701</td>
<td>7.8844</td>
<td>2.4098E-16</td>
<td><em>VLLAADS(ph)EEGDFPS(ph)GR</em></td>
</tr>
<tr>
<td>IPI00226216</td>
<td>6.8911</td>
<td>1.17135E-18</td>
<td><em>ANLS(ph)PSSFR</em></td>
</tr>
<tr>
<td>IPI00130200</td>
<td>6.8667</td>
<td>1.35049E-18</td>
<td><em>ALSEDEPCSSS(ph)AVK</em></td>
</tr>
</tbody>
</table>

**Fig 1. Display multiple phospho sites from an uploaded "phospho" dataset**
Top image: The small badge at the top right of the node indicates how many phospho sites are in the dataset or that passed your cutoffs in an analysis (depending on whether a dataset or analysis is overlaid). In this example, two phospho peptides for Chk1 passed the analysis cutoff for Phospho Fold Change. Clicking the badge shows the differential phosphorylation as a heatmap alongside the phosphorylated peptide sites (if uploaded in the dataset). Bottom image: Example of phosphorylation sites uploaded in the dataset (right column).
Other Useful IPA Features

- miRNA Target Filter
- Biomarker Filter
- Filtered Data
- Collaborative Workspace
- Compare
Tools are interconnected

Search and Explore
Build, Overlay and other tools

Bioprofiler and Isoprofiler
Profiling genes, isoforms, phenotypes and diseases

Core Analysis
Pathways, Diseases, regulators enrichment
Canonical Pathways

Sample to Insight
Diseases and Functions
### Regulator Effects

**Summary**

<table>
<thead>
<tr>
<th>ID</th>
<th>Context</th>
<th>Node Total</th>
<th>Regulator</th>
<th>Target To</th>
<th>Disease &amp; Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.333</td>
<td>11</td>
<td>↑ SNAI1</td>
<td>↓ ADIPOQ</td>
<td>invasion, all 100% (0/1)</td>
</tr>
<tr>
<td>2</td>
<td>1.890</td>
<td>9</td>
<td>↑ SNAI1</td>
<td>↓ ADIPOQ</td>
<td>invasion, all 100% (0/1)</td>
</tr>
<tr>
<td>3</td>
<td>1.789</td>
<td>7</td>
<td>↑ miR-153</td>
<td>↑ EGFR</td>
<td>invasion, all 100% (0/1)</td>
</tr>
<tr>
<td>4</td>
<td>0.522</td>
<td>35</td>
<td>↑ F2</td>
<td>↑ ALDH</td>
<td>tumor, all 100% (0/1)</td>
</tr>
<tr>
<td>5</td>
<td>0.514</td>
<td>36</td>
<td>↑ F2</td>
<td>↑ ALDH</td>
<td>female, all 100% (0/1)</td>
</tr>
<tr>
<td>6</td>
<td>0.000</td>
<td>25</td>
<td>↑ VEGF</td>
<td>↓ CDH1</td>
<td>endocrine, all 100% (0/1)</td>
</tr>
<tr>
<td>7</td>
<td>0.000</td>
<td>9</td>
<td>↑ WISP2</td>
<td>↓ CDH1</td>
<td>endocrine, all 100% (0/1)</td>
</tr>
<tr>
<td>8</td>
<td>-3.020</td>
<td>60</td>
<td>↑ TNF</td>
<td>↓ ARA</td>
<td>endocrine, all 100% (0/1)</td>
</tr>
<tr>
<td>9</td>
<td>-3.441</td>
<td>21</td>
<td>↑ Estrogen</td>
<td>↑ AXL</td>
<td>endocrine, all 100% (0/1)</td>
</tr>
<tr>
<td>10</td>
<td>-3.667</td>
<td>11</td>
<td>↑ IL2</td>
<td>↑ CD44</td>
<td>invasion, all 100% (0/1)</td>
</tr>
<tr>
<td>11</td>
<td>-4.373</td>
<td>43</td>
<td>↑ Cg</td>
<td>↑ ACPP</td>
<td>hepatoblastoma, all 100% (0/1)</td>
</tr>
<tr>
<td>12</td>
<td>-4.427</td>
<td>42</td>
<td>↑ Cg</td>
<td>↑ ACPP</td>
<td>liver cancer, all 100% (0/1)</td>
</tr>
<tr>
<td>13</td>
<td>-4.860</td>
<td>90</td>
<td>↑ MYC</td>
<td>↑ ABCE1</td>
<td>female, all 100% (0/1)</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Diagram**

[Diagram showing regulatory network with nodes and edges]
Network Analysis

The analysis is composed of 25 networks. To view a network, select the appropriate network(s) and click View Networks. To merge selected networks, click Merge Networks.

Network 1: ANXA9, BFSP1, BIK1, CD164, CNFN, CTSF2, CUTA, EVPL, FAM110A, GRB2, HEI2, KIF13A, KIF26A, KIRREL, LONRF2, MS12, NCKAP5, NOTCH3, NUDT21G, OTUD5S, PEX7, PEX13, PPI, RIN3, SH2D5S, SH2D9S, SHD19, SH3GL1, SH3GL2, SH3GL3, SLC32A1, SLC22A4L, SNX4, SNX7, SOD2, TIAP1, TULIP1, ZNF509

Network 2: ARRB1, BH1L, CCDC86, CMB1, CNBP, DDX21, DKC1, DNAH1, DNAH3, FBL, GNL3, GPR101, HFRD2, IGF2R, INF1, MIA, MYBBP1A, NAA15, NAA50, NHP2L1, NOC3L, NOP56, PABPC1, PHR1, PRR5, RSLI1, RTCA, SAFB2, SIB5, SRRK3, STXBPS, SUN5, TOPO1, THUMPD1, TWN2, ZNHT6

Network 3: CNK1G2, DCAF7, Dock7, FCAN, FCRL1, FIBG1, KIAA1324, KIAA1724, LGL2, MC4R, MARK2, MARK5, MARK6, MKRN1, MYO10, NR1D1, PIK3RAS, POM121/POM121C, PPPM3, SECH1, SH2BPSL, SOGAL1, SPD1, STRA13, STRADA, TMC4

Network 4: ABCA3, AN2, BRWD1, C10orf25, CIQTN5, CRMD1, DCDC2, ENY2, ER1, FAM102A, FAM180A, KCN1G6, MACROD1, METTL7A, MT1S, NDC1, PHEFA, PSD4, Rab11, Rab11FIP3, Rab11FIP4, RENG, RNF2, SEMA4A, SLCO5A4, SLCO4A2, SLCOA5, TGN2, TMPRSS5, TTC2, VPS13D, ZNF107, ZNFI141, ZNF703

Network 5: ABC12, ABL1M1, AKT1P, BRCDC5, C10orf82, CALCOCO2, CCR1IP1, COMTD1, FAM107A, GADD45GFD, HOOK1, HOOK2, IKK, KATNAL1, KHORBS2, KIAA, PKC1P, MOPD2, MTR, NCO1, JACOBS, RPCAF, RPS6KA6, SAFIBP1, SIZ1

Sample to Insight

Network Analysis

Networks

<table>
<thead>
<tr>
<th>Molecules in Network</th>
<th>Score</th>
<th>Focus Molecule</th>
<th>Top Diseases and Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANXA9, BFSP1, BIK1,</td>
<td>34</td>
<td>35</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
<tr>
<td>CD164, CNFN, CTSF2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUTA, EVPL, FAM110A,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRB2, HEI2, KIF13A,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF26A, KIRREL, LONRF2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS12, NCKAP5, NOTCH3,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUDT21G, OTUD5S, PEX7,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEX13, PPI, RIN3, SH2D5S,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH2D9S, SHD19, SH3GL1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH3GL2, SH3GL3, SLC32A1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC22A4L, SNX4, SNX7, SOD2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIAP1, TULIP1, ZNF509</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARRB1, BH1L, CCDC86,</td>
<td>32</td>
<td>33</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
<tr>
<td>CMB1, CNBP, DDX21, DKC1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAH1, DNAH3, FBL, GNL3,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPR101, HFRD2, IGF2R, INF1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIA, MYBBP1A, NAA15, NAA50,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NHP2L1, NOC3L, NOP56, PABPC1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHR1, PRR5, RSLI1, RTCA, SAFB2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIB5, SRRK3, STXBPS, SUN5, TOPO1, THUMPD1, TWN2, ZNHT6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNK1G2, DCAF7, Dock7, FCAN, FCRL1, FIBG1, KIAA1324, KIAA1724, LGL2, MC4R, MARK2, MARK5, MARK6, MKRN1, MYO10, NR1D1, PIK3RAS, POM121/POM121C, PPPM3, SECH1, SH2BPSL, SOGAL1, SPD1, STRA13, STRADA, TMC4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCA3, AN2, BRWD1, C10orf25, CIQTN5, CRMD1, DCDC2, ENY2, ER1, FAM102A, FAM180A, KCN1G6, MACROD1, METTL7A, MT1S, NDC1, PHEFA, PSD4, Rab11, Rab11FIP3, Rab11FIP4, RENG, RNF2, SEMA4A, SLCO5A4, SLCO4A2, SLCOA5, TGN2, TMPRSS5, TTC2, VPS13D, ZNF107, ZNFI141, ZNF703</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Comparison Analysis:
Comparing Multiple Observations (Experimental Groups)
Starting a New Comparison Analysis

Select analyses for side-by-side comparison. Click View Comparison to view comparison results.
Comparison Analysis: Gene Expression Comparison on the Pathway Chart
Comparison Analysis: Gene Expression Comparison through Heatmap
Comparison Analysis: Filtering and Sorting the Heatmap
Comparison Analysis: Canonical Pathway, Upstream Analysis and Diseases and Functions

Sample to Insight
Other Useful IPA Functionalities

- Search and Explore
- Bioprofiler
- Isoprofiler
Other Useful IPA Functionalities

- miRNA Target Filter
- Biomarker Filter
- Filtered Data
- Collaborative Workspace
- Compare
Resources: Training Manuals and Videos

“Support on Steroids”

- AdvGx Support: AdvancedGenomicsSupport@qiagen.com
- AdvGx Licensing: BioinformaticsLicense@qiagen.com
IPA strengths

- Powerful Knowledgebase (5 million plus findings)
  - Manual Curation by PhD level staff
  - Links to original sources
- Directionality
  - Causal Networks
- Wide range of tools and functionalities
  - Large dataset analyses tools
  - Pathway building tools
  - Updates
- Easier to use compared to other tools
- Ability to share
- Access to PhD level support staff and field support staff
What can Qiagen bioinformatics products do for you?

- **Biological Processes**
  - TGFβ
  - SNAI1
  - MYC
  - P53

- **Upstream Regulators**
  - Cancer
  - Cellular Movement
  - Invasion
  - Cell Death

- **Canonical Pathways**
  - Cell-Junction
  - ILK pathway

- **Software**
  - WorkBench
  - IPA
  - IVA/HGMD/QCI

- **Journals**
  - Cell
  - Nature
  - Science
  - JBC
Questions?

Thank You!!

Dev Mistry, Ph.D.
Field Applications Scientist
Devendra.Mistry@qiagen.com