Choose a product to login

INGENUITY® PATHWAY ANALYSIS
Comprehensive pathway and network analysis of complex 'omics data

INGENUITY® iREPORT
For current customers

INGENUITY® VARIANT ANALYSIS
Rapidly find causal variants using a knowledge-driven approach

New: Streamline login and launch by installing the IPA client on your computer

Need Help with Ingenuity Products?
Live Support 6AM - 5PM Pacific Time, M-F

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

Ted King
Account Manager
Qiagen Advanced Genomics
Ted.King@qiagen.com
(240) 731 -7102

Dev Mistry, Ph.D.
Field Applications Scientist
Qiagen Advanced Genomics
Devendra.Mistry@qiagen.com
Overview

- Introduction to IPA
- Search and Explore
  - Growing a network out of a molecule
  - Bioprofiler (Advanced Analytics)

- Large Dataset Analysis
  - Uploading your dataset(s) and starting a core analysis
  - Core Analysis
    - Canonical Pathways
    - Upstream Regulators
      - Causal Network (Advanced Analytics)
    - Diseases and Functions
    - Regulator Effect
    - Networks
  - Comparison Analysis

- Questions/Answer
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

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

INGENIUNITY Variant Analysis

INGENIUNITY Pathway Analysis

Sample to Insight
What can IPA do?

Large RNA seq dataset in form of a huge pile of papers

Methodical analysis by IPA in form of organized binders on a bookshelf
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
Peer-reviewed publications citing QIAGEN’s Ingenuity products

14,311 publications and growing!
Two different types of analyses by IPA

- Deep pathway understanding of a single gene/protein
- Biological understanding of large data sets
How can IPA help you?

- Deep pathway understanding of a **single** gene/protein
- Drug/therapeutic target discovery
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
Gene/Protein Expression Analysis: IPA Core Analysis

- **Canonical Pathways**
  - Check whether molecules from your dataset belong to pre-defined pathways
  - Predict the pathway activation/inhibition

- **Upstream Analysis**

- **Diseases and Functions**

- **Regulator Effect**

- **Networks**

---

Sample to Insight
Gene/Protein Expression Analysis: 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

Sample to Insight
Gene/Protein Expression Analysis: IPA Core Analysis

- **Canonical Pathways**
- **Upstream Analysis**
- **Diseases and Functions**
  - Predicts affected biology based on gene expression and predicts directional change on that effect
    - "Increase in EMT"
    - "Decrease in proliferation"
- **Regulator Effect**
- **Networks**

Diagram showing a network of genes and their relationships, including:
- ITGB1
- CDH1
- SNAI2
- CLDN4
- TGFβ1
- ADIPOQ
- PTEN

Delineated areas indicate:
- Invasion of tumor cells
Gene/Protein Expression Analysis: IPA Core Analysis

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

Models pathway interactions from predicted upstream regulators through differentially expressed genes, to biological processes.
Gene/Protein Expression Analysis: IPA Core Analysis

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

- Predicts non-directional gene interaction map
1. Scroll down
2. Change memory to >1000mb
3. Save
Don’t worry too much about notes or if you fall behind during the point and click training. We have manuals/videos for everything.
Uploading your dataset
Suggested Format for uploading RNA-seq data

<table>
<thead>
<tr>
<th>ID</th>
<th>Log2Ratio</th>
<th>p-value</th>
<th>Intensity/(\text{RPKM}/\text{FPKM})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_130786</td>
<td>-0.14</td>
<td>8.68E-01</td>
<td>2931.69</td>
</tr>
<tr>
<td>NR_015380</td>
<td>-0.99</td>
<td>2.24E-01</td>
<td>1649.26</td>
</tr>
<tr>
<td>NM_138932</td>
<td>-0.02</td>
<td>9.83E-01</td>
<td>1.67</td>
</tr>
<tr>
<td>NM_014576</td>
<td>-0.02</td>
<td>9.85E-01</td>
<td>1.77</td>
</tr>
<tr>
<td>NM_138933</td>
<td>0.02</td>
<td>9.79E-01</td>
<td>1.83</td>
</tr>
<tr>
<td>NM_000014</td>
<td>-4.79</td>
<td>1.02E-01</td>
<td>239.75</td>
</tr>
<tr>
<td>NR_026971</td>
<td>-0.67</td>
<td>6.17E-01</td>
<td>213.79</td>
</tr>
<tr>
<td>NM_144670</td>
<td>-5.96</td>
<td>1.30E-01</td>
<td>610.64</td>
</tr>
<tr>
<td>NM_001080438</td>
<td>-1.97</td>
<td>3.47E-01</td>
<td>3.91</td>
</tr>
<tr>
<td>NM_017436</td>
<td>-1.09</td>
<td>5.02E-01</td>
<td>6186.83</td>
</tr>
<tr>
<td>NM_016161</td>
<td>2.02</td>
<td>5.97E-02</td>
<td>149.85</td>
</tr>
<tr>
<td>NM_015665</td>
<td>-0.27</td>
<td>5.68E-01</td>
<td>13330.34</td>
</tr>
</tbody>
</table>

Max RPKM recommended for RNA seq
Verify the differential expression calculation

- Recommend Log$_2$(ratio) differential expression

$$\text{Log}_2\left(\frac{\text{Experimental Condition Exp.}}{\text{Control Exp}}\right)$$

- Ratio differential expression

$$\left(\frac{\text{Experimental Condition Exp.}}{\text{Control Exp}}\right)$$

- Fold Change
  - If increased differential expression

$$\left(\frac{\text{Experimental Condition Exp.}}{\text{Control Exp}}\right)$$

  - If decreased differential expression

$$-1 \left(\frac{\text{Control Exp.}}{\text{Experimental Condition Exp.}}\right)$$

Fold change will never have values between 1 and -1
Case Study
RNA Seq: Claudin Low vs Luminal Breast cancer cell lines
Epithelial to Mesenchymal Transition


Mesenchymal / stem cell-like breast cancer

Ratio Claudin-low to Luminal
5 vs 5 cell lines, RNA-Seq data

Luminal Breast cancer

Luminal cell lines

Claudin-low cell lines

Breast development

Luminal cells in ducts

Luminal progenitors

MaSC (stem cell)

E-Cadherin
Claudins
Occludins
ZO-1
Desmoplakin
Cytokeratins

N-Cadherin
Fibronectin
Collagen I/III
Snail
αSMA
Vimentin

Tight-junction dissociation
Loss of microvilli

Adherent-junction and desmosome dissociation
Loss of apical-basal polarity

αSMA expression
Cytoskeleton reorganization
Front-back polarity
Migration

MMPs up-regulation
Basement membrane degradation
Invasion

Sample to Insight

Proprietary and Confidential
Verify the biology
- Can IPA identify cancer and EMT related pathways and biological functions in this dataset?
- What are some of the relevant pathways?
- What are some of the relevant biological functions?

Identification of transcriptional regulators
- What are the transcriptional regulators that are causing the gene expression changes in this dataset?
- Are they activated or inhibited?

Hypothesis generation
- Are the predicted upstream regulators increasing or decreasing downstream biological functions?
P value and Z Score
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 insignificant (<0.05) for random datasets
- **Gene expression direction is not taken into account** for this calculation
Z-score: Activation Prediction

- - + + + + + + +

Gene expression from Knowledge Base (literature)

↓ ↓ ↑ ↑ ↑ ↑ ↑ ↓

Gene expression in your dataset

1 1 1 1 1 1 1 -1

+1 score for the consistent and -1 for the inconsistent relationships

\[
z = \frac{x}{\sigma_x} = \frac{\sum_i x_i}{\sqrt{N}} = \frac{N_+ - N_-}{\sqrt{N}} = \frac{(7-1)}{\sqrt{8}} = 2.12 \text{ (predicted activation)}\]

- **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
Upstream Regulators, Mechanistic Networks and Causal Networks

Upstream Regulators

Mechanistic Networks

Advanced Analytics: Causal Network Analysis

Master Regulators

Upstream Regulators

Your Dataset Genes/Molecules
Sample to Insight
Upstream Analysis

Sample to Insight
Diseases and Functions

Sample to Insight
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.

<table>
<thead>
<tr>
<th>ID</th>
<th>Molecules in Network</th>
<th>Score</th>
<th>Focus Molecule</th>
<th>Top Diseases and Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ANXA9, BFSP1, BIN1, CD164, CNFN, CTFS2, CUTA, EVPL, FAM110A, GRB2, HEI2, KIF13A, KIF26A, KIRRE1, LONRF2, MDS2, NCKAP5, NOD2, NUTM1D2, OTUD5, PEX7, PEX13, PPM1, RIN3, SH2D5, SH2D9, SH3D9, SH3GL1, SH3GL2, SLCE2A1, SNX4, SNX7, SOG2, TIAP1, TULP1, ZNF609</td>
<td>34</td>
<td>35</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
<tr>
<td>2</td>
<td>ARRB1, BPH1, CCDC86, CMAS, CNBP, DDX11, DKC1, DNAH1, DNAH3, FHL, GNL3, GPBAR1, HDR2, IGF2R, INA3, MYBBP1A, NAA15, NAA50, NHP2, NOC3L, NOP56, PABPC1, PPRC1, PPP2R5D, RPL11, RPL7A, SAFB2, SB35B, SRPK3, SXTBPS5, SUN3, TCO1, THUMP1, TXN2, ZNHIT6</td>
<td>34</td>
<td>35</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
<tr>
<td>3</td>
<td>C2orf44, CAB, CAB39L, CCDC186, CEP85L, COBL1, COL4A3BP, CSNK1G1, CSNK1G2, DCAF7, DCK7, FANCB, FANCE, HIBBC1, KIAA1324, KIAA1724, LEL2, MARCH4, MARK2, MARK5, MARK6, MYO10, NRD1, PIK3HA1, POM121/POM121C, PPPI8, SEC13, SEHIL, SH2BPSL1, SOG1A1, SPDL1, STRA13, STRADA, TMC4</td>
<td>34</td>
<td>35</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
<tr>
<td>4</td>
<td>ABCA3, ANO2, BRWD1, C10orf25, C1QTNF6, CRML1, DCK2, EN2, ESR1, FAM102A, FAM103A, KCNK6, MACROD1, METTL7A, ML15, NDC1, PHP5A, PSD4, Rab1l, RAB11A, RAB11FIP3, RAB11FIP4, RERG, RTN2, SEMA4A, SLC18B, SLC4A4, SLC4A5, T2CN, TMRRS3, TCTC, VPS13D, ZNF107, ZNF1141, ZNF703</td>
<td>34</td>
<td>35</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
<tr>
<td>5</td>
<td>ABC17, ABL1M1, AKTIP, BRICD5, C10orf2, CALCOCO2, CNBP1P1, COMTD1, FAM107A, GADD45GDI1, HOOK1, HOOK2, DOK (complex), KATNAL1, KHRDBP3, KLC1, KAT5, KIAA1101, KIAA1104, LPP, NCOR1, NCOA4, RPAF2, RPS6KA6, SAFP1, SH3BP2, SNIP2</td>
<td>34</td>
<td>35</td>
<td>Connective Tissue Disorders, Developmental Disorder, Hereditary Disorder</td>
</tr>
</tbody>
</table>
Comparison Analysis:
Comparing Multiple Observations (Experimental Groups)
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
What can IPA do?
QUESTIONS?

CONTACTS:

- General: support@ingenuity.com
- +1 650 381-5111
  6am-5pm Pacific Time (M-F)

QIAGEN Redwood City
1700 Seaport Blvd., 3rd Floor
Redwood City
CA 94063, USA