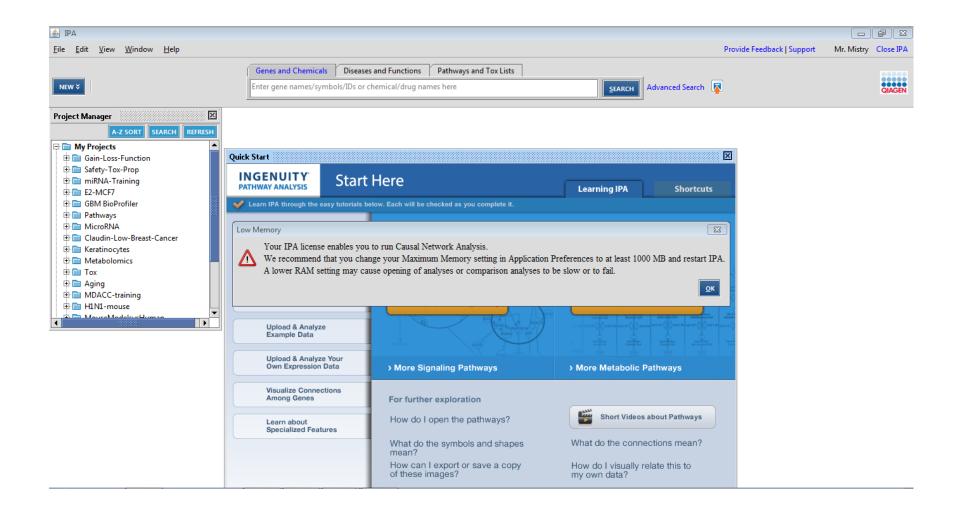
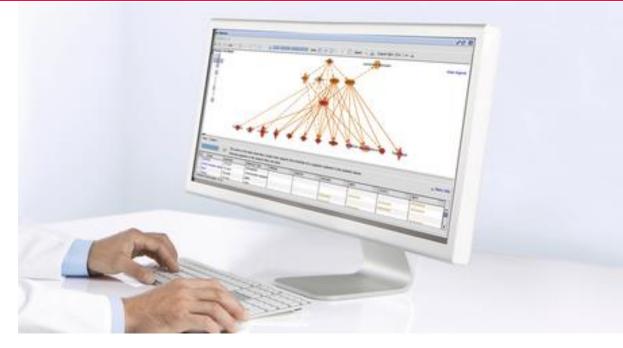


Sample to Insight









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

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

Dev Mistry, Ph.D.

Field Applications Scientist Qiagen Advanced Genomics Devendra.Mistry@giagen.com



- 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

Sample to Insight



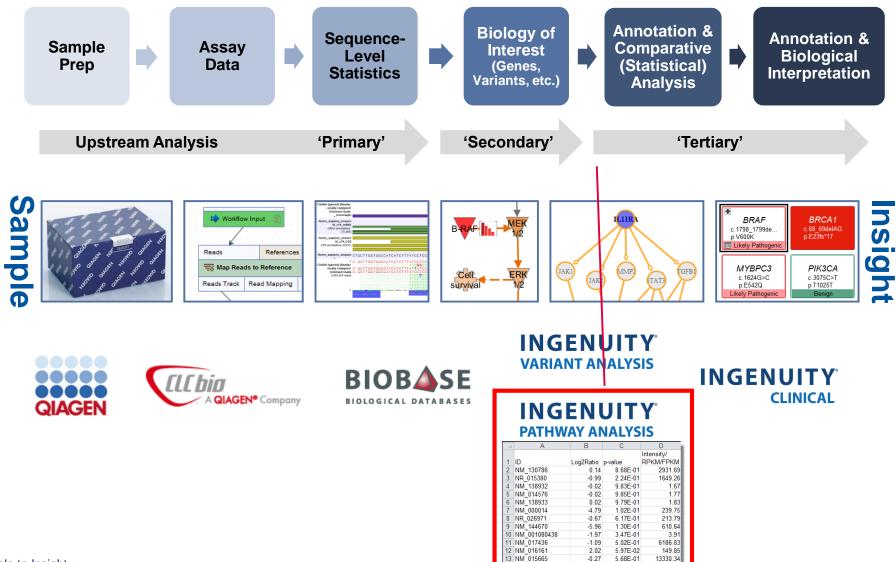
Introduction

- Sample to Insight



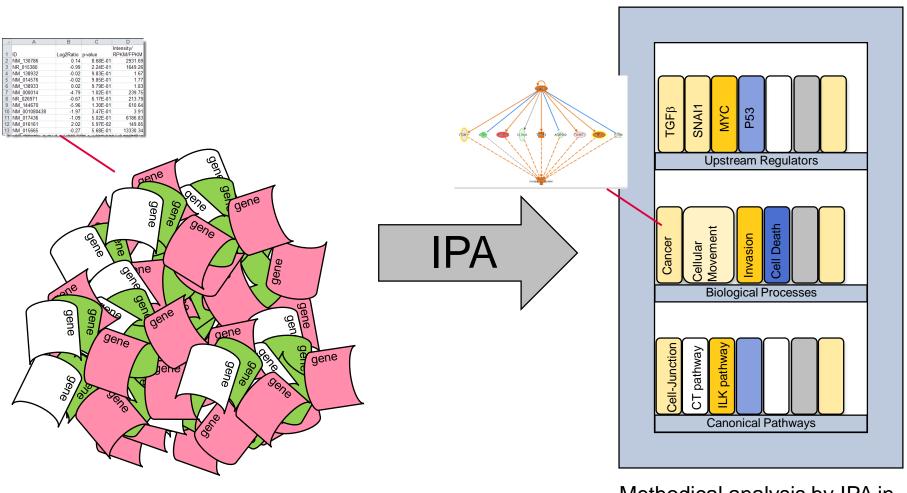
When do you use IPA?

QIAGEN Sample to Insight





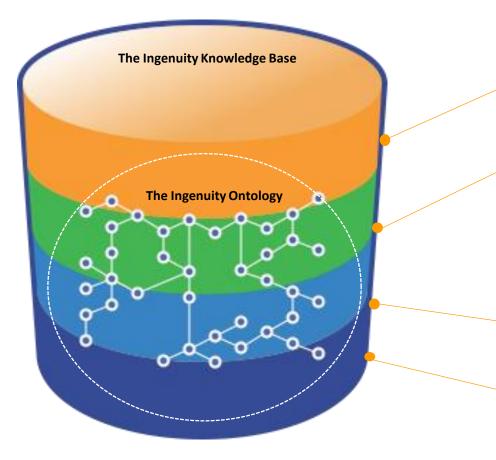
What can IPA do?



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

Methodical analysis by IPA in form of organized binders on a bookshelf





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

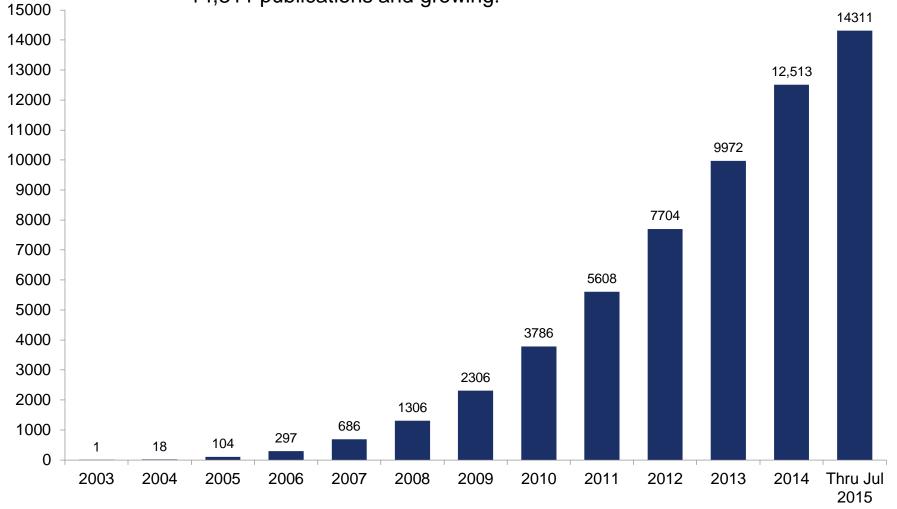


Species Supported

- 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



14,311 publications and growing!





Two different types of analyses by IPA

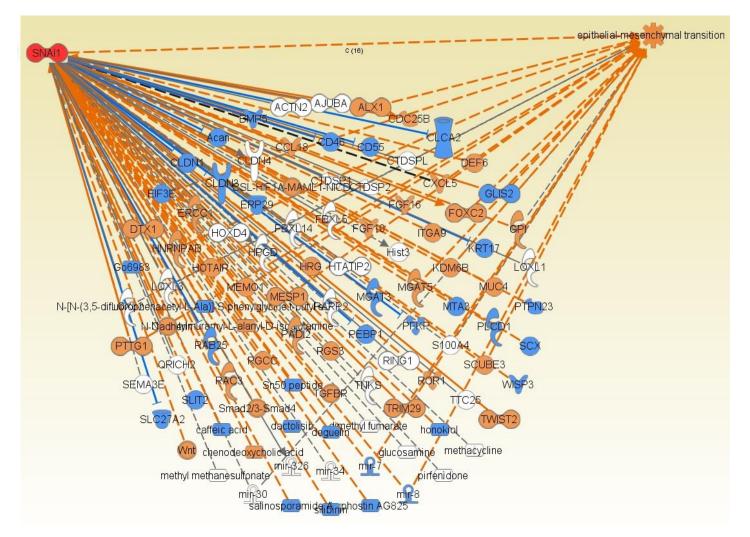
- Deep pathway understanding of <u>a single gene/protein</u>
- Biological understanding of <u>large data sets</u>



How can IPA help you?

Deep pathway understanding of a single gene/protein

□ Drug/therapeutic target discovery



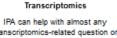


How can IPA help you?

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





transcriptomics-related question or application

Toxicogenomics

Delivers a focused toxicity and

safety assessment of candidate

compounds, and provides a more

complete understanding of

pharmacological response, drug

mechanism of action, and mechanism of toxicity



Biomarker Discovery

Identifies the most promising and relevant biomarker candidates within experimental datasets



Metabolomics

Overcomes the metabolomics data analysis challenge by providing the critical context necessary to gain biological insight into cell physiology and metabolism from metabolite data



microRNA Research

Combines filtering tools and microRNA-mRNA content to provide insight into the biological effects of microRNAs



Drug Repositioning

Expression profiling of approved drugs and comparison to profiles of diseased tissue can lead to discovery of new uses for these already approved entities



Proteomics

Perform a comprehensive analysis of your proteomics for a deep understanding of proteins and related biological processes

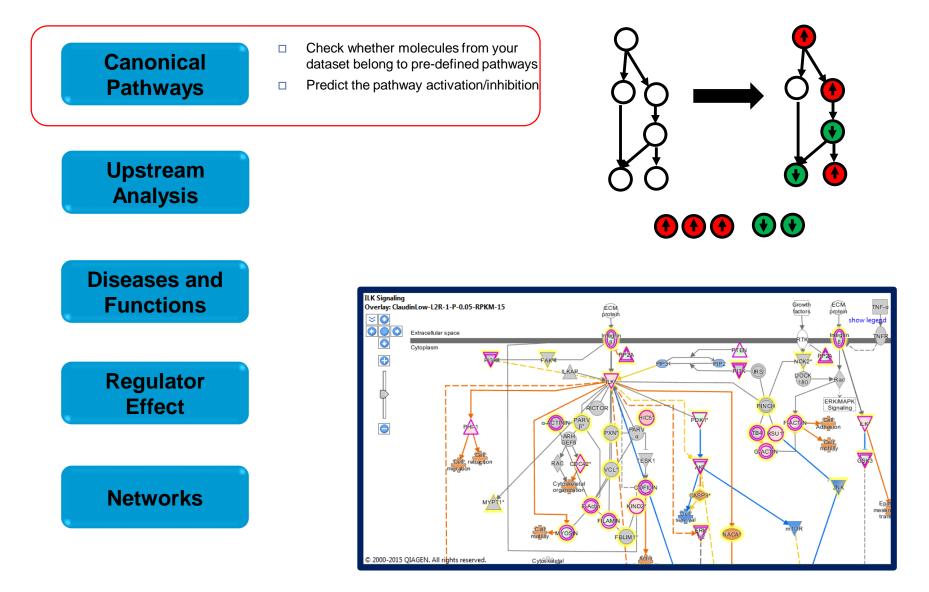


Target Discovery

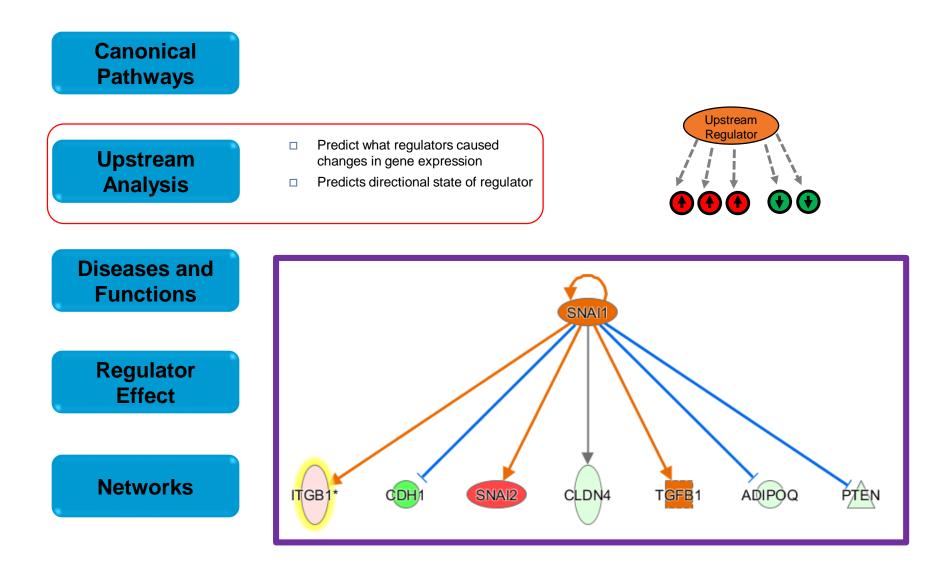
Genes that are shown to be activated in a pathological condition may serve as promising targets for therapeutic development efforts



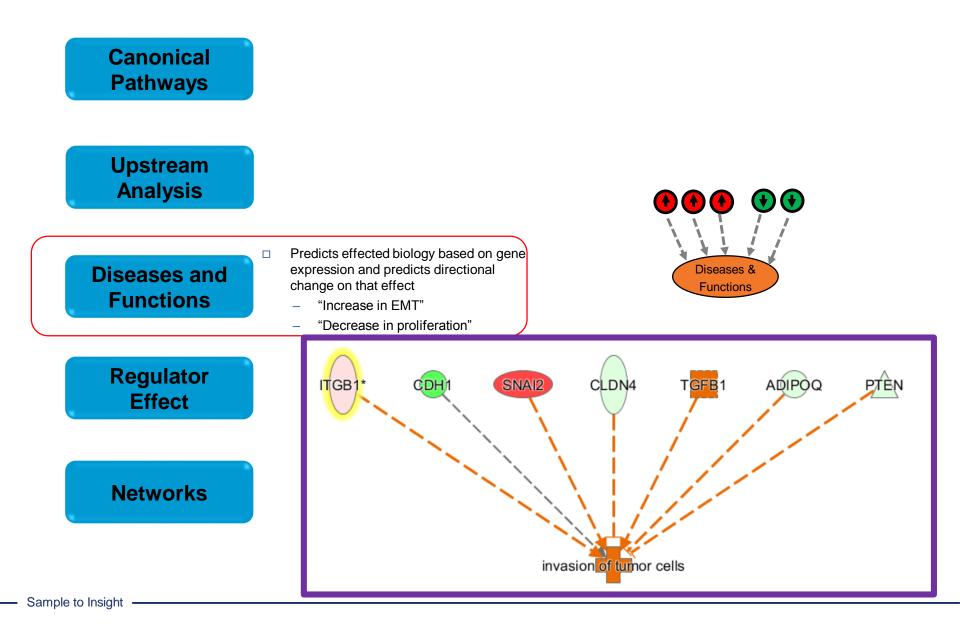
Gene/Protein Expression Analysis: IPA Core Analysis



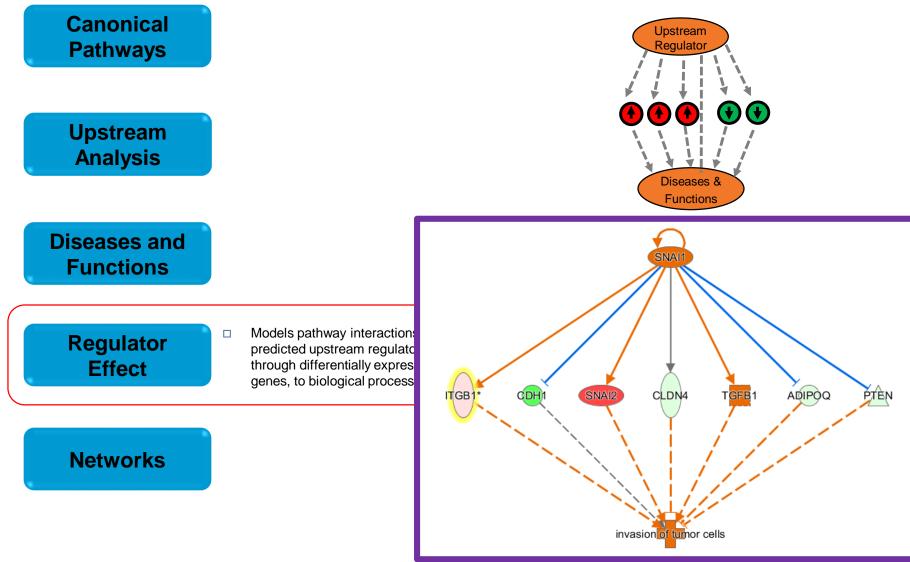




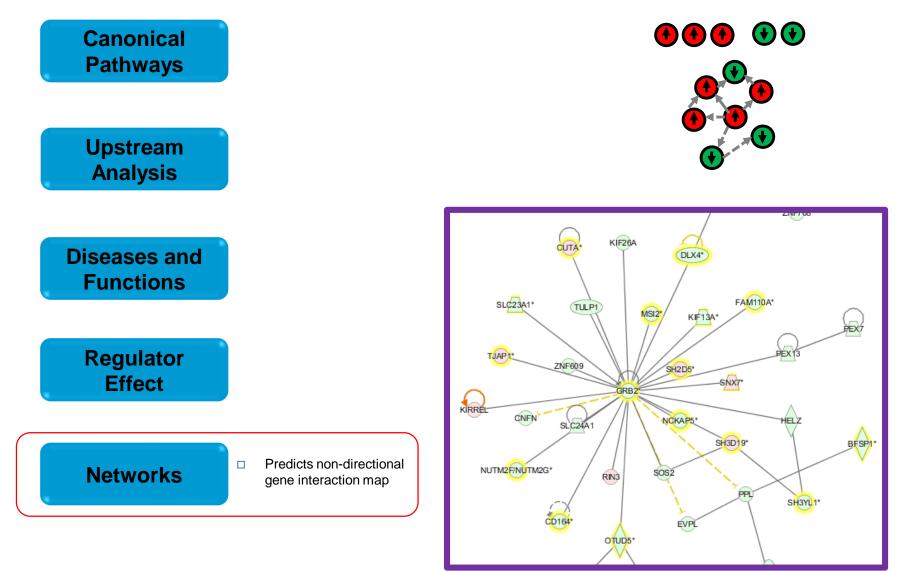












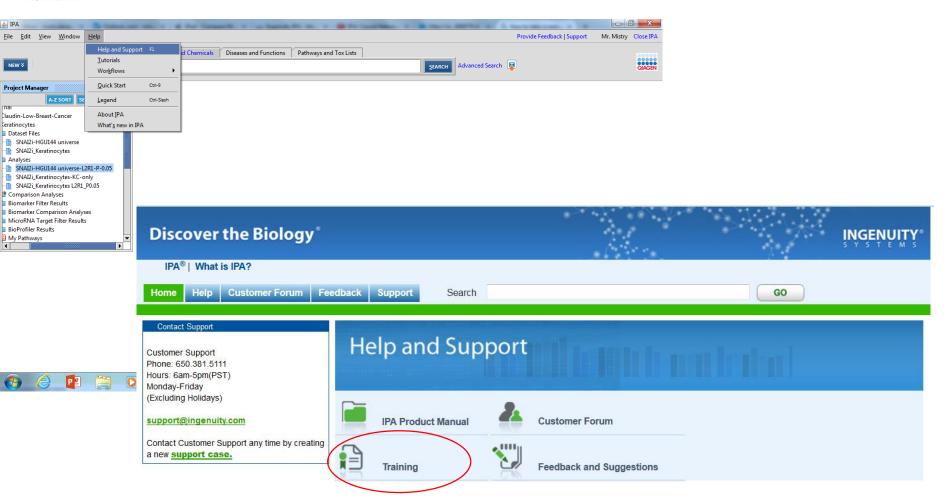


Change Preferences

🖆 IPA								
<u>File Edit View Window H</u>	<u>H</u> elp					Provide Fee	edback Support Mr. Mistry	Close IPA
New	•	Genes and Chemicals	Diseases and Functions	Pathways and Tox Lists]			
<u>O</u> pen		· · · · · · · · · · · · · · · · · · ·	ls/IDs or chemical/drug na			Advanced Search		QIAGEN
Save	Ctrl-S				SEARCH			QIAGEN
S <u>a</u> ve As								
<u>U</u> pload Dataset	Ctrl-U							
Batch Upload Datasets				🖆 Applicatio	n Preferences			X
Search Datasets and Analyses								
Refresh Project Manager	FS			Application	Graph Filter Cache	_		
Vie <u>w</u> References				False Di	scovery Rate (q-value)			-
Export <u>D</u> ata	Ctrl-E			Comparison	Analysis Bar Chart Coloring			
Export <u>I</u> mage	Ctrl+Shift-E				lored square to set the color that	appears for each observation.		
Send <u>T</u> o	• • •				ors: 01 02 03 04 05			017 018 019 020
Sha <u>r</u> e	• •			Charteen				
Prop <u>e</u> rties	• •							
Pre <u>f</u> erences	•	Application Preferences		Destants				
	Ctrl-P	Change Password		Projects	· · Samanla and Tissue Everyosian I	Veniente	1. Scroll	down
<u>P</u> rint		-			y Sample and Tissue Expression I	nojecis	1. 001011	uowii
Close IPA	Ctrl-Q			Notification				
					Analysis Status E-mail			
					Analysis Status E-mail Analysis-Running popup			
					Analysis-Running popup			
				Workflow In	tructions			
					Workflow Instructions			
					worknow instructions			
				Quick Start				
					Quick Start Screen on startup			
					Zuick start sereen on startup			
				Auto hide st	atus har			
							(to) 1000p	nh
👌 🌔 📳		🛂 🕘	PA		^{hide status bar} 2. Char	ige memory	<u>y to >100011</u>	di
				Maximum M	emory			
				2000mb				
				20001110				
							2 (Save
					mum number of Pathways, Netw	orks, or Lists that can be open	ed in a window.	Jave
				3 🔻				
						1000000		
							RESTORE DEFAULTS	SAVE CANCEL

Sample to Insight





Don't worry too much about notes or if you fall behind during the point and click training. We have manuals/videos for everything.

Sample to Insight



Uploading your dataset



R	equired		Recomm	ended	rec	ax RPKM commended RNA seq
	A		В	С	D	
				(Intensity/	
_1(<	Log2Ratio	p-value	RPKM/FPKM	
2	NM_130786		0.14	8.68E-01	2931.69	
3	NR_015380		-0.99	2.24E-01	1649.26	
4	NM_138932		-0.02	9.83E-01	1.67	
5	NM_014576		-0.02	9.85E-01	1.77	
6	NM_138933		0.02	9.79E-01	1.83	
7	NM_000014		-4.79	1.02E-01	239.75	
8	NR_026971		-0.67	6.17E-01	213.79	
9	NM_144670		-5.96	1.30E-01	610.64	
10	NM_001080438		-1.97	3.47E-01	3.91	
11	NM_017436		-1.09	5.02E-01	6186.83	
12	NM_016161		2.02	5.97E-02	149.85	
13	NM_015665		-0.27	5.68E-01	13330.34	
- P			in an a constant of		~~~~~~~~~~~~	

p	Ctl		
300	10	0 =Max(B2:	C2)
1000	10	0 1000)
10		1 10)
		1000 10	1000 100 1000



Verify the differential expression calculation

Recommend Log₂(ratio) differential expression

```
Log_2\left(\frac{Experimental Condition Exp.}{Control Exp}\right)
```

Ratio differential expression

```
(Experimental Condition Exp.)
```

- Fold Change
 - If increased differential expression

(Experimental Condition Exp.) Control Exp

□ If decreased differential expression

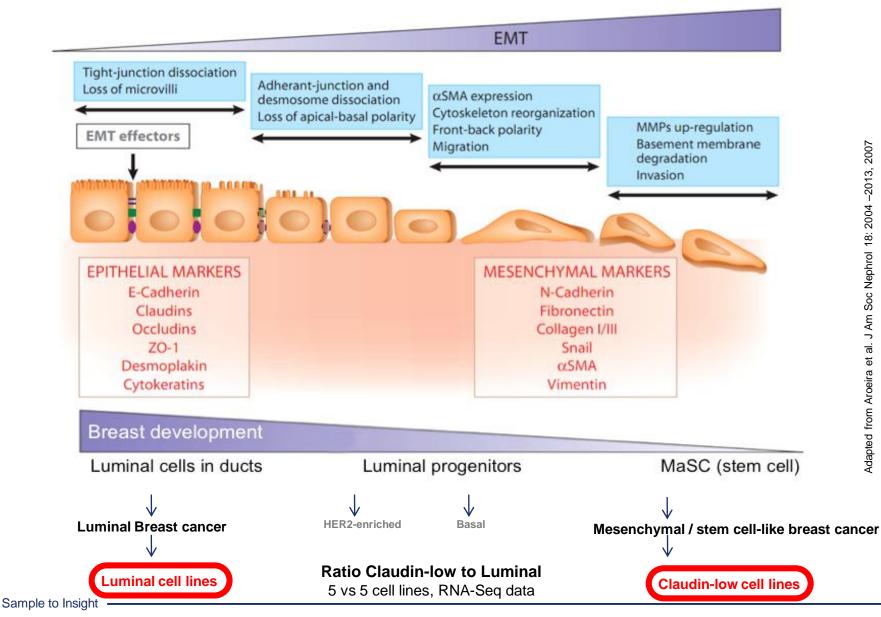
```
-1 \left(\frac{Control Exp.}{Experimental Condition Exp.}\right)
```

Fold change will never have values between 1 and -1



Case Study RNA Seq: Claudin Low vs Luminal Breast cancer cell lines





Proprietary and Confidential



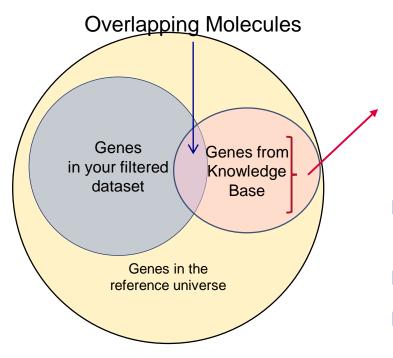
- Verify the biology
 - Can IPA identify cancer and EMT related pathways and biological functions in this dataset?
 - □ What are some of the relevant pathways?
 - □ What are some of the relevant biological functions?
- Identification of transcriptional regulators
 - What are the transcriptional regulators 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

- Sample to Insight

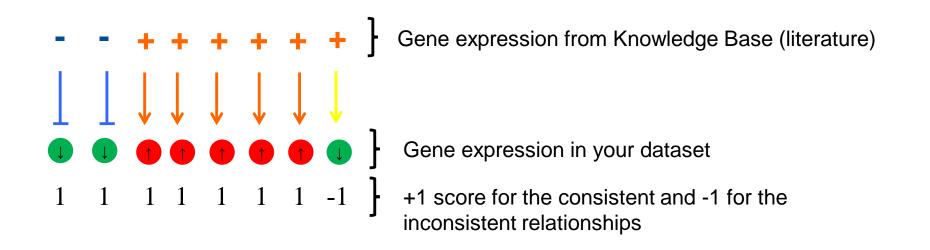




Genes from previous literature that belong to

- 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

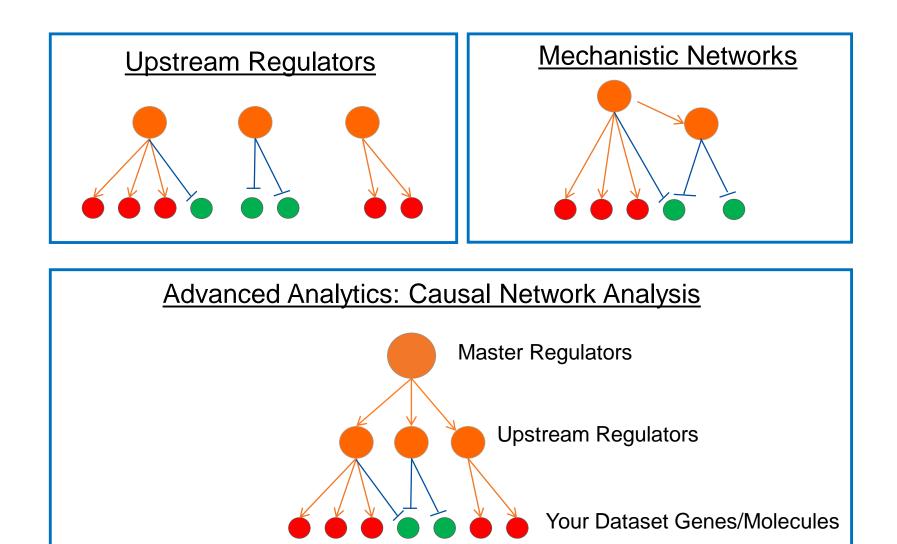




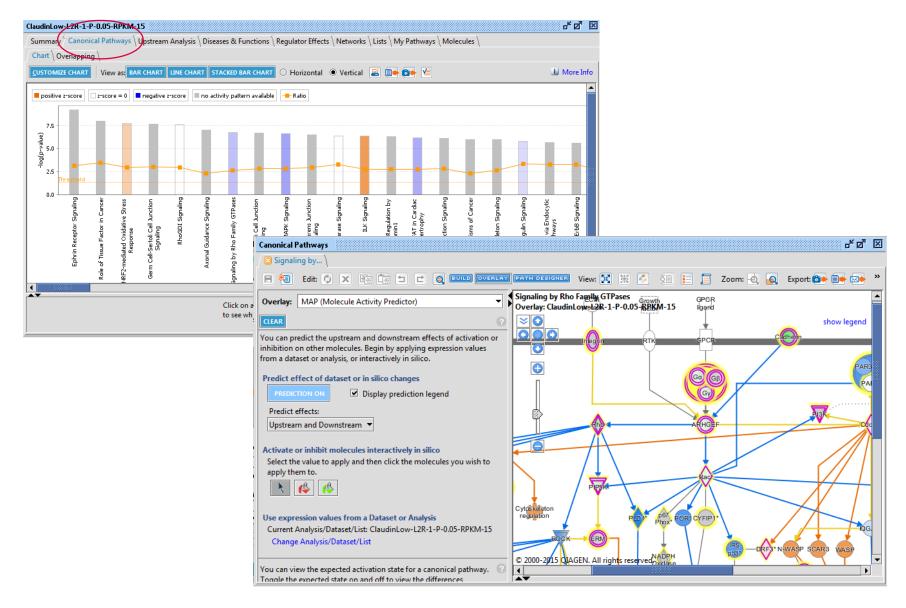
$$z = \frac{x}{\sigma_x} = \frac{\sum_i x_i}{\sqrt{N}} = \frac{N_+ - N_-}{\sqrt{N}} = (7-1)/\sqrt{8} = 2.12 \text{ (= predicted activation)}$$

- z-score is a statistical measure of the match between expected relationship direction and observed gene expression
- <u>z-score > 2 or < -2 is considered significant</u>
- 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



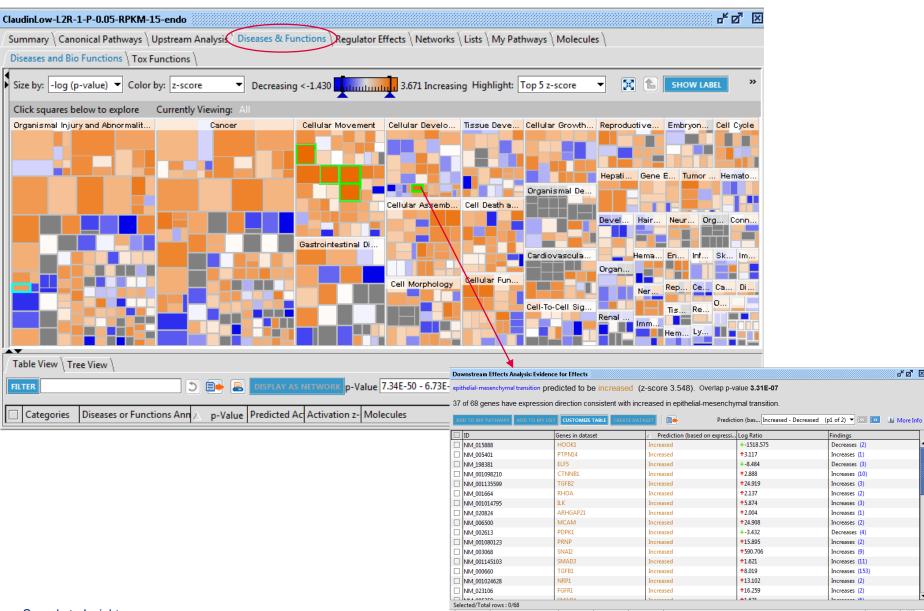






Upstream Regulators	Causal Networks							
ADD TO MY PATHWAY	ADD TO MY LIST DIS	SPLAY AS NETWORK CI	JSTOMIZE TABLE MECH		2 📑 🛃			»
Upstream Reg 💌	Log Ratio 🛛 🕱	Molecule 🛐 💌	Predicted Acti 💌	Activation z-s 💌	🛆 p-value o 💌	Target m 🝸 💌	Mechanis 🝸 🕱	
Sos		group			1.15E-06	↑ABCC1, +all 58	679 (10)	
IDH2	+-7.249	enzyme			1.40E-06	+ADIPOQ, +all 10		
miR-155-5p (miRNAs		mature microrna	Inhibited	-2.464	1.48E-06	↑ABHD16A, 🔩all 47		
FOS	+-70.043	transcription regula		0.713	1.62E-06	↑ABCC1, ↓all 123	1000 (18)	
ID3	↑ 6.045	transcription regula		0.611	2.29E-06	+ADGRG1, +all 52		
SYVN1	+-2.290	transporter	Activated	3.507	7.51E-06	↑ABCC4, ↑all 41	612 (10)	
miR-1-3p (and other		mature microrna		0.248	1.04E-05	↓ABHD11, ↑all 55		
MAPK1	† 1.387	kinase		1.885	1.04E-05	↑ADAM12, ▶all 81	991 (18)	
MTOR	† 1.161	kinase		0.509	1.25E-05	↑ACADL, ↓all 66	948 (17)	
PML	† 4.199	transcription regula		0.178	1.37E-05	↑ACADL, ↓all 37	957 (19)	
miR-27a-3p (and oth		mature microrna		-0.262	1.72E-05	◆BBC3, ◆Call 17		
SDCBP	↑ 5.305	enzyme		1.149	1.97E-05	+CDH1, ↑CTall 9	389 (11)	
AGT	† 1.637	growth factor	Activated	2.660	2.82E-05	↑ACAT2, ↓all 99	1163 (19)	
mir-8		microrna	Inhibited	-2.104	3.24E-05	↑ABL2, ↓AKT1all 19		
SNAI1	† 1.041	transcription regula	Activated	3.473	3.56E-05	↓ADIPOQ, ↑all 22	807 (14)	
Vegf		group	Activated	4.387	3.59E-05	♦ADAM15, ▶all 94	692 (15)	
MGEA5	+-1.305	enzyme		0.000	3.62E-05	↓ABLIM1, ↑ PKM-15		
TP63	† 23.730	transcription regula		0.879	4.24E-05	↑AHR, ↓ARAF		SNAII
7ER1	↑ 370 2/18	transcription regula	Activated	2 003	5 085-05	CCNG2		
Selected/Total mole	cules : 1 / 747						× 1	
0							PIT RELAS	EP300 CTNNB3' ID2 JUN E2F1 SMAD4 ESR1*
PEBP1 ITGB CXCL	3 RAB25 ENAI2 TV	VISTI ADIPOQ ZEBT	PERP CON7 CONS	BARI PIEN OCU		LDN4 TGFB1 ZEB2	FOSL1 KRT18* CLUN3	





- 6 X

-



Regulator Effects

- QIAGEN										
ClaudinLow-L2R-	-1-P-0.05-RPKM-	15-endo								ra X
Summary \ Can	onical Pathways \	Upstream Analysi	s 🛛 Diseases & Fun	ictions Reg	gulator	Effects Network	s 🛛 Lists 🗍 My Path	ways 🛛 Molecules)	
GENERATE NETWO	ORKS ADD TO MY	PATHWAY ADD TO		MIZE TABLE						🔟 More Info
·	\leq									
ID	🗸 Const 🗶	Node Total 💌	Regulator 💌	Regul 🚺		Target To 💌	Targe 🝸 💌			Known Re 💌
1	2.333	ſĨ	1	◆SNAI1	all 1	-	♦ADIPOQall 9		invasionall 1	
2	1.890	9	1	◆SNAI1	all 1		♦ADIPOQall 7		invasionall 1	
3	1.789	7	1	miR-155			↑EGFR, ↑all 5		invasionall 1	
	0.522	35	1	↑F2	all 1		↑ALDHall 33		tumorigeall 1	
· · · · ·	0.514	36	1	↑F2	all 1		↑ALDHall 34		female gall 1	
6	0.000	25	1	Vegf	a#1		♦CDH1, ·all 23		endocrinall 1	
7	0.000	9	1	₩ISP2	all 1	7	♦CDH1, Nall 7		endocrinall 1	555
8	-3.020	60	1	↓TNF	all 1		↓AR, ↑all 58		endocrinall 1	0% (0/1)
9	-3.441	21	1	estrogen	all 1	19	↑AXL,all 19		endocrinall 1	0% (0/1)
10	-3.667	11	1	↓ IL2	all 1	9	+CD44, +all 9		invasionall 1	
11	4.373	43	1	Cg	all 1	41	♦ACPP, *all 41	1	hepatobil <mark>all</mark> 1	0% (0/1)
12	-4 427	42	1	Cg	all 1	40	♦ACPP, *all 40	1	liver cancerall 1	0% (0/1)
13	-4.690	90	1	↑MYC	all 1	88	↑ABCE1,all 88	1	female gall 1	
• •	4.00	lar.	4	L COL	-0.4		1 A 17 1 - 11 - 14	4		100% (1/1)
							ITGB 1.	CDH1 SNA		
	пыци								invas	or of tumor



Network Analysis

C	laudinLow	-L2R-1-P-0.05-RPKM-15-endo			- ⁴ 🗠	×
\int	Summary	Canonical Pathways Upstream Analysis Diseases & Functions Regulator Effects Networks Usts My	y Pathway	ys \ Molecules \		
	Networks	Overlapping Networks				
1	FILTER	The second sec	IETWORKS	FUNCTIONS AN	INOTATIONS 📑 🖪 🖉	»
Т	he analysi	s is composed of 25 networks. To view a network, select the appropriate network(s) and click View Networks.	To merce	e selected networks	click Merge Networks	
Г		Molecules in Network	Score		Fop Diseases and Functions	
		↓ANXA9, ↓BFSP1*, ↑BIN1*, ↓CD164*, ↓CNFN, ↓CTPS2*, ↑CUTA*, ↓EVPL, ↓FAM110A*,	34		Connective Tissue Disorders,	
		GRB2*, +HELZ, +KIF13A*, +KIF26A, ↑KIRREL, +LONRF2, +MSI2*, +NCKAP5*,	1.24		Developmental Disorder,	
		+NUTM2F/NUTM2G*, +OTUD5*, +PEX7, +PEX13, +PPL, ↑RIN3, ↑SH2D5*, ↑SH3D19*,		-	Hereditary Disorder	
		↑SH3GL1, +SH3YL1*, +SLC23A1*, +SLC24A1, +SNX4, ↑SNX7*, +SOS2, ↑TJAP1*, +TULP1, +ZNF609				200
ſ	2		Vetwork 2 :	Claudin Low-L2R-2-P-(0.01 : Claudin Low : Claudin Low-L2R	-2-P-0
		↑FBL, ↑GNL3*, +GPBAR1*, ↑IFRD2, ↑IGF2BP2*, ↑MINA*, ↑MYBBP1A*, ↑NAA15, ↑NAA50,			\bigcirc	
		↑NHP2L1*, ↑NOC3L, ↑NOP56*, ↑PABPC4*, ↑PPRC1, ↑RRP9, ↓RSL1D1, ↑RTCA*, ↑SAFB2,			QKI	
	3	↑SF3B5, +SRPK3, ↑STXBP5*, ↑SUN3*, ↑TCOF1*, +THUMPD1, +NWC2, ↑ZNHIT6 ↑C2orf44*, +CA8, +CAB39L*, +CCDC186, ↑CEP85L*, +COBLL1, +COL4A3B*, +CSNK1G1,				
"	L P	+CSNKIG2, +DCAF7, +DOCK7, +FANCB*, +FANCF, +HBEGF, +KIAA1324, +KIAA1324,				~
		+LLGL2*, +LRCH4, +MARK2*, ↑MARK3*, +MARK4, +MKRN1*, ↑MYO10, ↑NRD1*,			ZNH	
		+PLEKHA5*, +POM121/POM121C*, +PPM1B*, ↑SEC13*, ↑SEH1L*, +SH3BP5L, +SOGA1,			(T	
	4	↑SPDL1, ↓STRA13, ↓STRADA*, ↓TMC4*↓ABCA3, ↓ANO2, ↓BRWD1*, ↑C18orf25*, ↓C1QTNF6*, ↑CRIM1, ↓DCDC2, ↑ENY2, ↓ESR1*,			HZAFY*	-
"	4	★ABCAS, ★ANO2, ★BKWD1, +CIGOT25, ★CIQINFO, +CKIM1, ★DCDC2, +ENT2, ★ESK1, ↓FAM102A*, ↓FAM186A, ↓KCNK6, ↓MACROD1, ↓METLTA, ↓MTL5*, ↑NDC1, ↑PHF5A,			$\langle \rangle$	Me
		+PSD4, Rab11, +RAB11A, +RAB11FIP3*, +RAB11FIP4, +RERG, +RTN2*, +SEMA4A,			() -+t	Me
		↑ SLC35B4, ↓ SLC4A2*, ↓ SLC4A5*, ↓ TC2N*, ↓ TMPRSS3*, ↓ TTC9, ↓ VPS13D*, ↓ ZNF107*,		REPIN1*	BMF	
	5	+ZNF141, +ZNF703 ↑ABCF2*, +ABLIM1*, +AKTIP*, +BRICD5, +C10orf82, +CALCOCO2, ↑CCNB1IP1*, +COMTD1			I shi	
"	- '	+FAM107A*, +GADD45GIP1, +HOOK1, +HOOK2*, IKK (complex), +KATNAL1*, +KHDRBS3,		/	ETS12	į
		+KICA* +KRT15 +MPPED2* +NOR1 +PAOR5* +RPAP2 +RPS6KA6 +SAP30RP +SENP2		~		+
_				ACBD4	GATA3	13
					Notch	T
				/	XTEI	1
				\wedge		11
				BLVRB	DAXX*	1
				V	ZEB	1.
						N

POLRIC*

PPP2R3C

TCF4*

ESRP2

MSANTD3

SNAI3

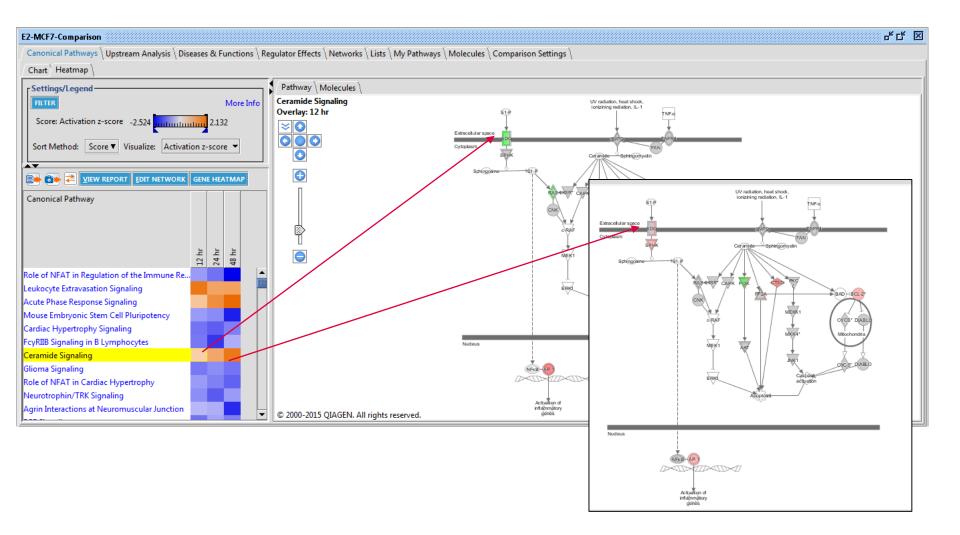
TCF

JUP



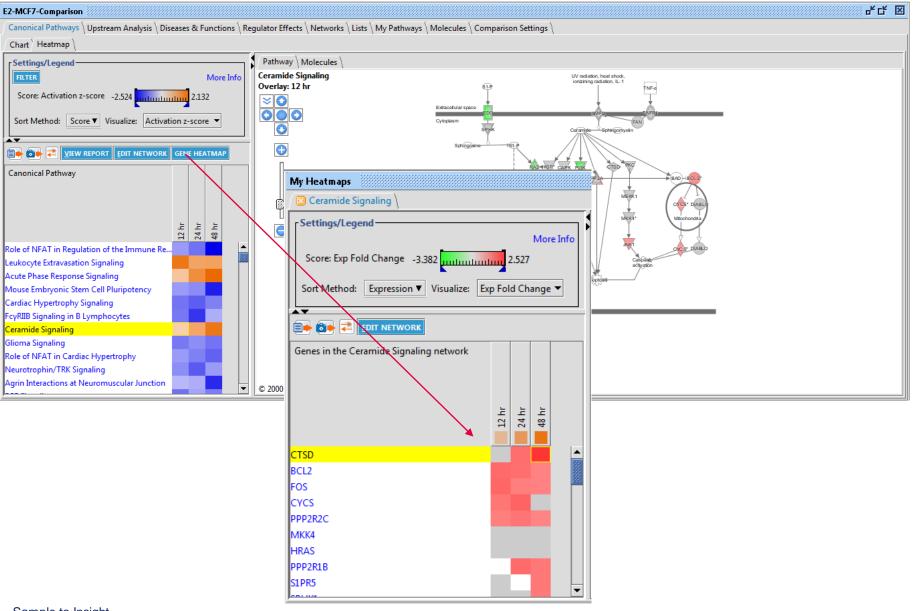
Comparison Analysis: Comparing Multiple Observations (Experimental Groups)



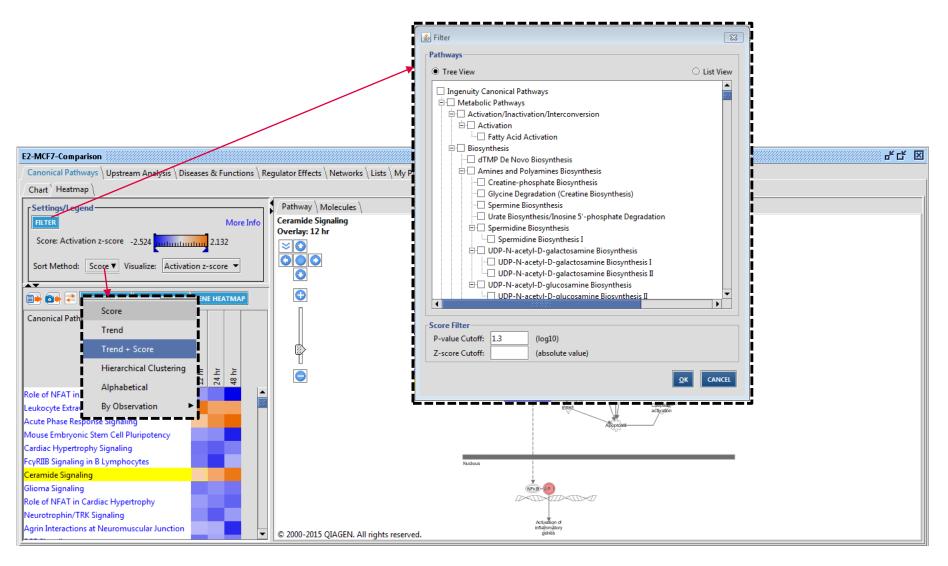




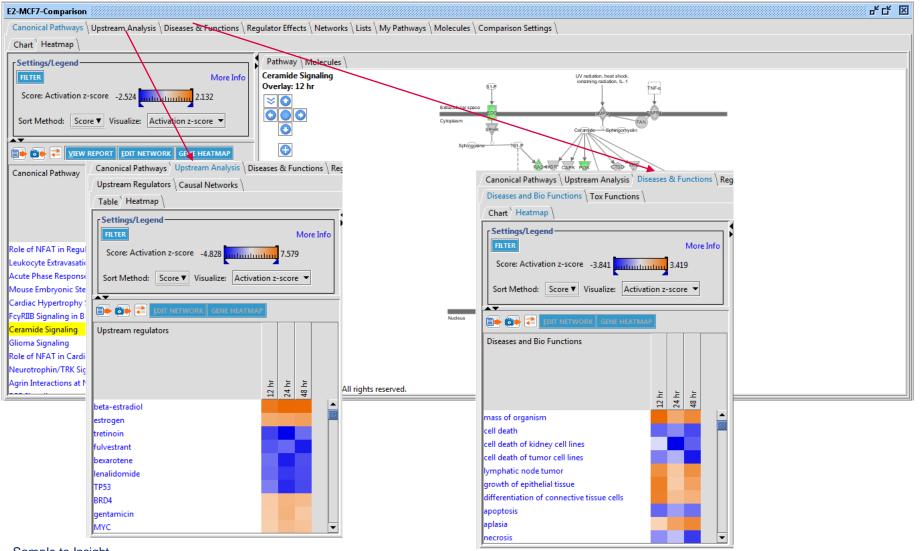
Comparison Analysis: Gene Expression Comparison through Heatmap







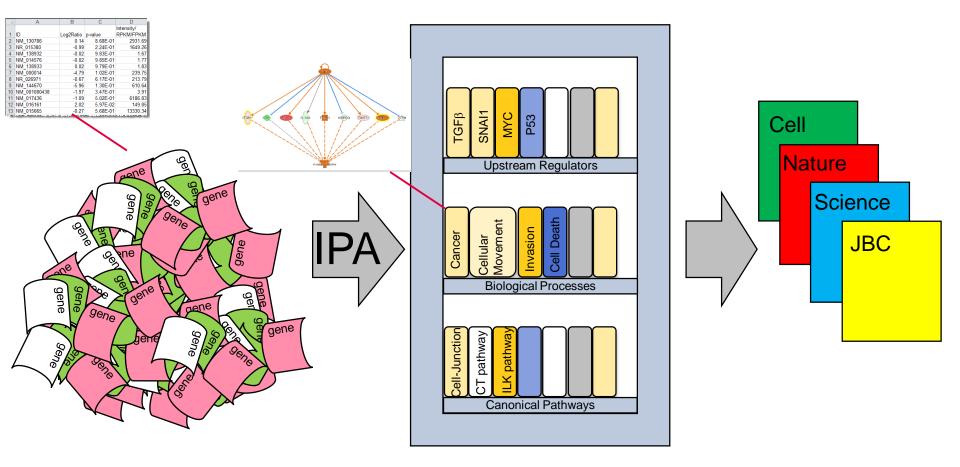
Comparison Analysis: Canonical Pathway, Upstream Analysis and Diseases and Functions



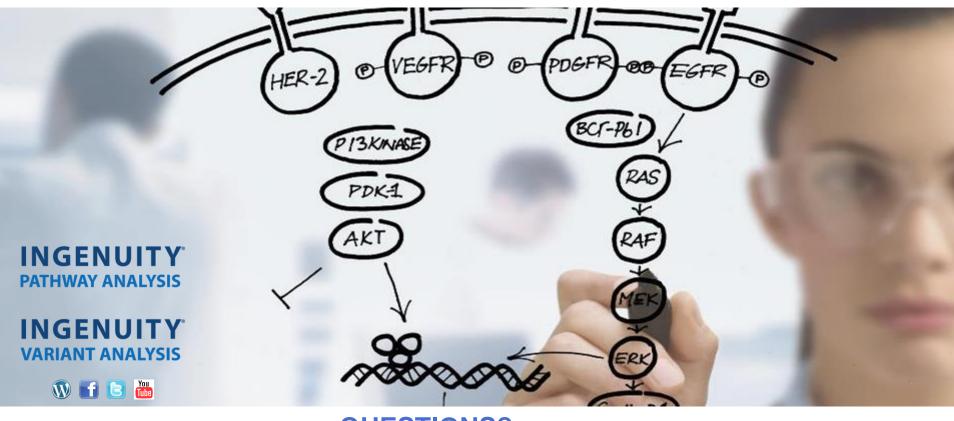
00000



What can IPA do?







QUESTIONS?

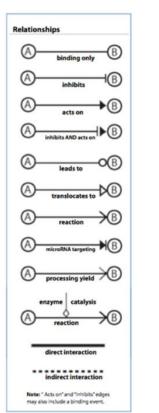
CONTACTS:

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

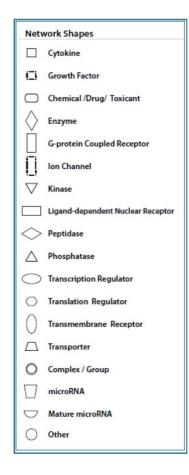
QIAGEN Redwood City

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





	Relationship Labels
A	Activation
в	Binding
С	Causes/Leads to
cc	Chemical-Chemical interaction
CP	Chemical-Protein interaction
E	Expression (includes metabolism/ synthesis
EC	Enzyme Catalysis
1	Inhibition
L	Proteolysis (includes degradation for Chemicals)
LO	Localization
м	Biochemical Modification
miT	microRNA Targeting
MB	Group/complex Membership
nTRR	Non-Targeting RNA-RNA Interaction
P	Phosphorylation/Dephosphorylation
PD	Protein-DNA binding
PP	Protein-Protein binding
PR	Protein-RNA binding
PY	Processing Yields
RB	Regulation of Binding
RE	Reaction
RR	RNA-RNA Binding
т	Transcription
TR	Translocation
UB	Ubiguitination



Introduction to QIAGEN Ingenuity & IPA - www.ingenuity.com