



IPA 2022 Summer Release

Gene Chen
Senior Manager, IPA certificated analyst

What's New in the IPA Summer Release (July 2022)

- Visualize your data in new ways with Canonical Pathway bubble charts
- Visualize OmicSoft single cell data in Land Explorer
- Improve readability of your networks with new node label placement
- Expand your research with support for new species for dataset upload
- Take advantage of inferred disease & phenotype networks with search
- Speed your work by opening Canonical Pathways directly from links on Gene Views
- Content updates
- Version or date of third-party databases
- Identifier Mapping
 - Identifier Source Versions
- System Requirements
 - Minimum Software Requirements
 - Minimum Hardware Requirements

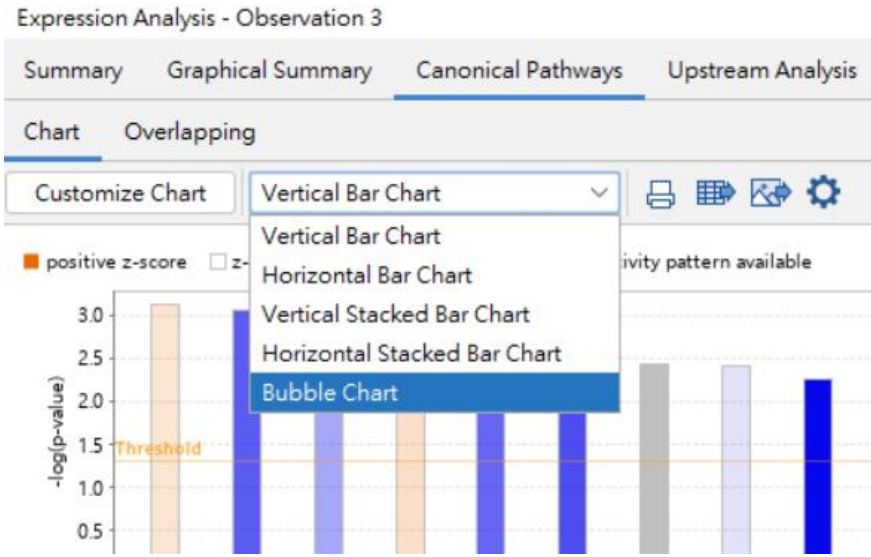
Product: QIAGEN IPA
Release Number: Summer Release 2022
Release Date: July 30, 2022

What's New in the IPA Summer Release (July 2022)

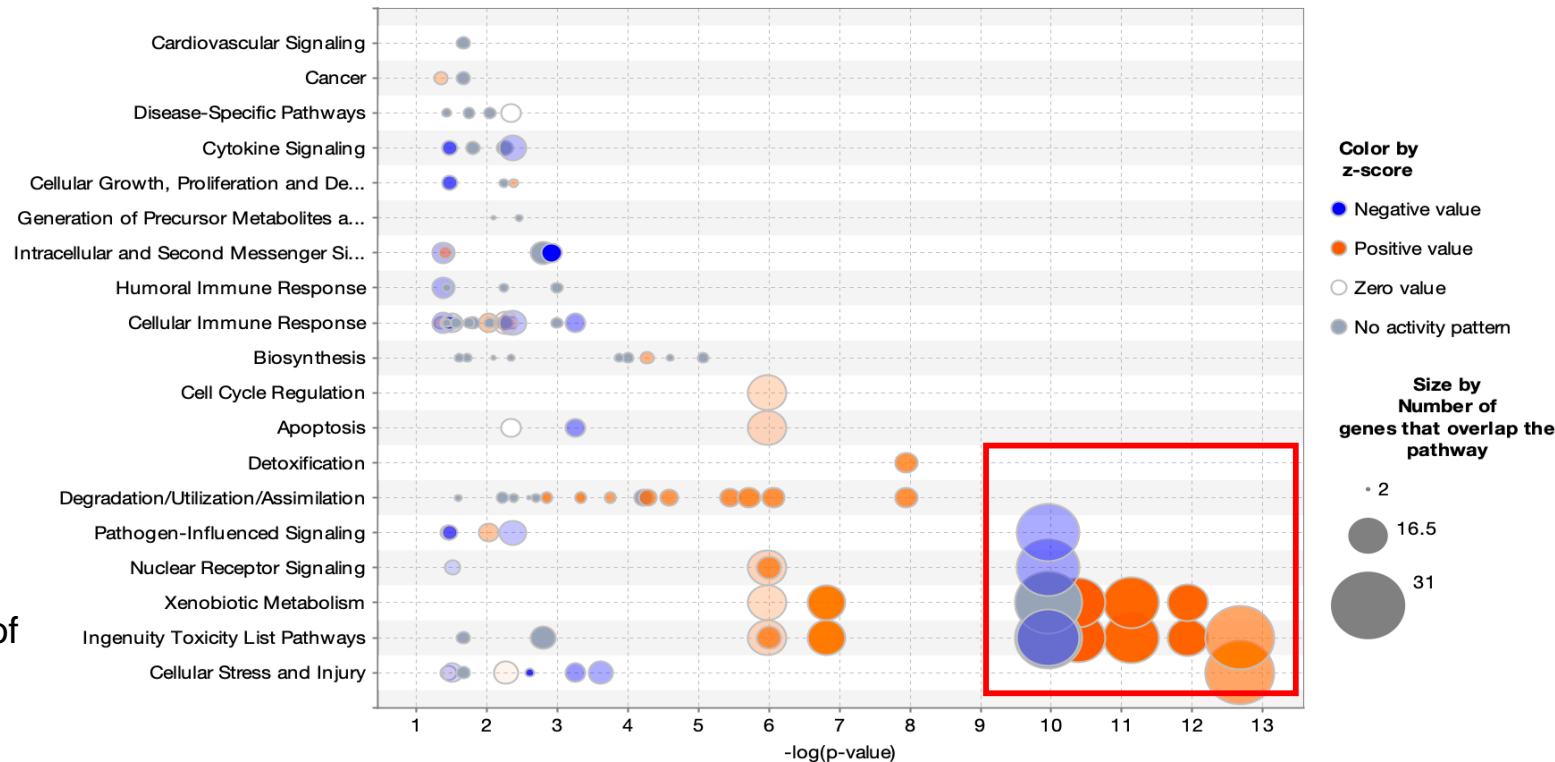
- Visualize your data in new ways with Canonical Pathway bubble charts
- Visualize OmicSoft single cell data in Land Explorer
- Improve readability of your networks with new node label placement
- Expand your research with support for new species for dataset upload
- Take advantage of inferred disease & phenotype networks with search
- Speed your work by opening Canonical Pathways directly from links on Gene Views
- Content updates
- Version or date of third-party databases
- Identifier Mapping
 - Identifier Source Versions
- System Requirements
 - Minimum Software Requirements
 - Minimum Hardware Requirements

Product: QIAGEN IPA
Release Number: Summer Release 2022
Release Date: July 30, 2022

Bubble charts present multi-variate data in an easily understood graphical form that can help explain your data to your audience. Now in IPA, you can generate bubble charts for your Core Analysis Canonical Pathway scores.



NRF2 activator turns on multiple pathways related to xenobiotic metabolism, toxicity, and cellular stress (see bottom right area of Figure 1), and with a higher number of overlapping genes relative to other pathways in the chart



Alternatively, by sizing the pathway bubbles in proportion to the negative log of the right-tailed Fisher's Exact test p-value and coloring them by pathway type (signaling or metabolic)

Expression Analysis - Observation 3

Summary Graphical S Customize Chart

Chart Overlapping

Customize Chart

Options fo

Filter by Gene(s)

Include:

Exclude:

P-Value Type for Use in Bubble Chart

Fisher's Exact Test p-value

Score Cutoff

Display only entities that have a $-\log(p\text{-value})$ greater than: (between 0 and)

an absolute value z-score greater than: (between 0 and)

Y-Axis and X-Axis

Y-axis (horizontal axis)

$-\log(p\text{-value})$

z-score

Pathway category

X-axis (horizontal axis)

$-\log(p\text{-value})$

z-score

Pathway category

Size of Bubbles

Number of genes that overlap the pathway

Ratio of number of overlapping genes to total number of genes in the pathway

Absolute value of the z-score

$-\log(p\text{-value})$

Color of Bubbles

Pathway category Signaling Metabolic

$-\log(p\text{-value})$

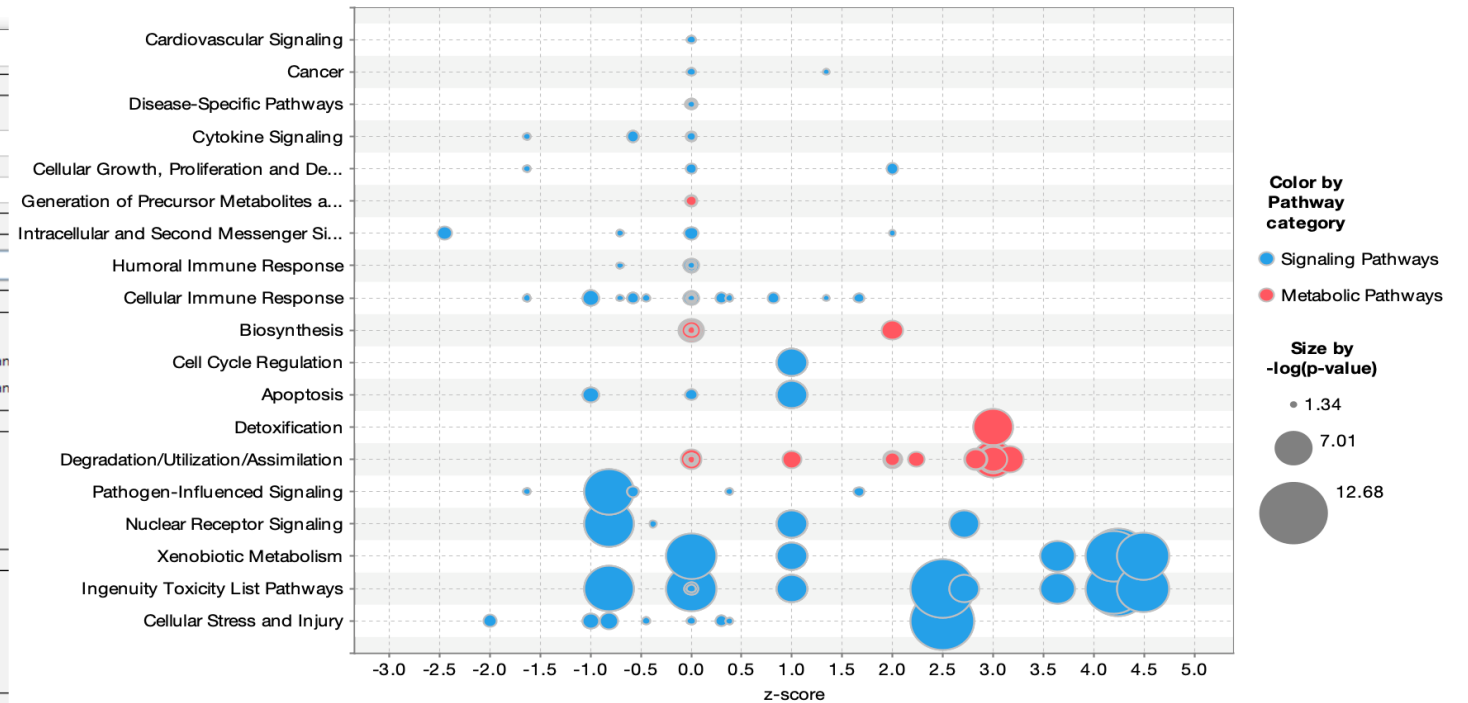
z-score

Number of genes that overlap the pathway

Ratio of number of overlapping genes to total number of genes in the pathway

Select Font Sizes

Entity Names: Scores: Legend:



You can see that the detoxification pathway shown in previous figure is a metabolic pathway, and is highly statistically significant, even though the number of overlapping genes for that detoxification pathway is rather small

*Land Explorer for IPA license is required to view the linked pages

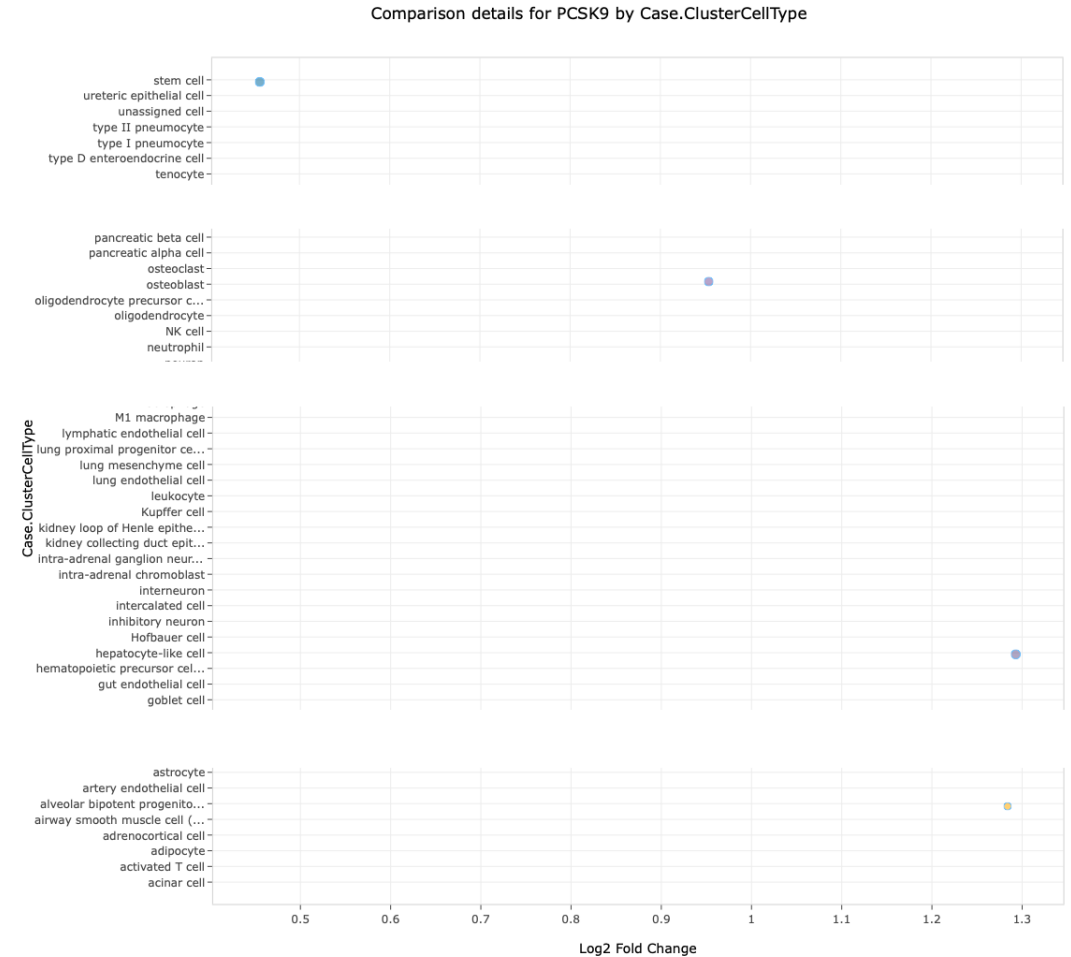
Now you can easily explore single cell expression for any gene among the public data curated by OmicSoft. These single cell views are available via new links on IPA Gene Views

OmicSoft Land Explorer: Sample-level experimental data

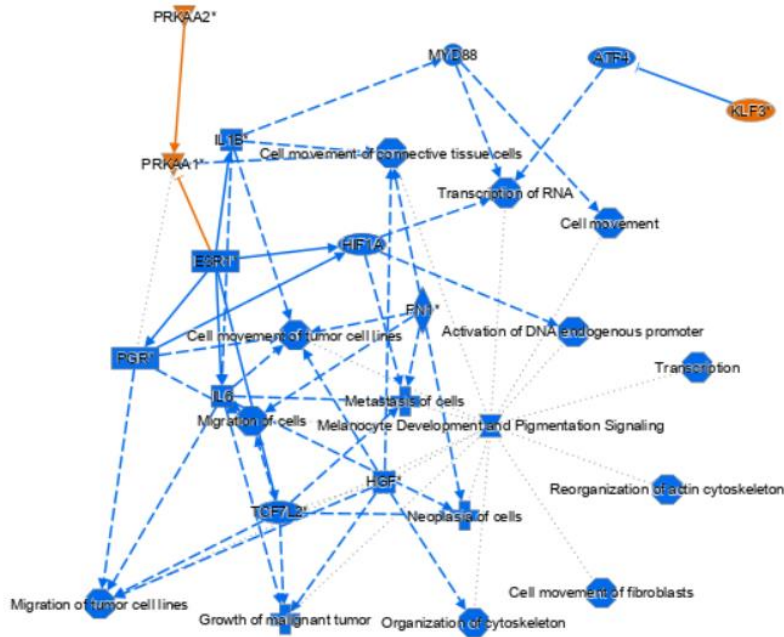
Data Type / Data Source	Normal Tissue	Cell Lines	Oncology Consortia	Oncology Studies
RNA-seq expression:	Solid tissue (GTEx), Solid tissue (HPA), Blueprint	Cancer cell lines (CCLE)	TCGA, TARGET, BeatAML, ICGC, CGCI, CCLE+GTEx+TCGA, ENCODE RNA-associated gene knockdown	General oncology, Hematology, Pediatrics, Mouse studies
Microarray expression:	Solid tissue (GTEx)	Cancer cell lines (CCLE), GSK cell	TARGET, expO, METABRIC,	General oncology, Hematology, Metastasis, Pediatrics, Mouse
Differential regulation:	Solid tissue (GTEx)			
Alteration frequency:				
Survival by expression:				
Single Cell differential regulation:	Human Cell Landscape (HCL), Tabula Sapiens			

Comparison details for PCSK9 by Case.ClusterCellType

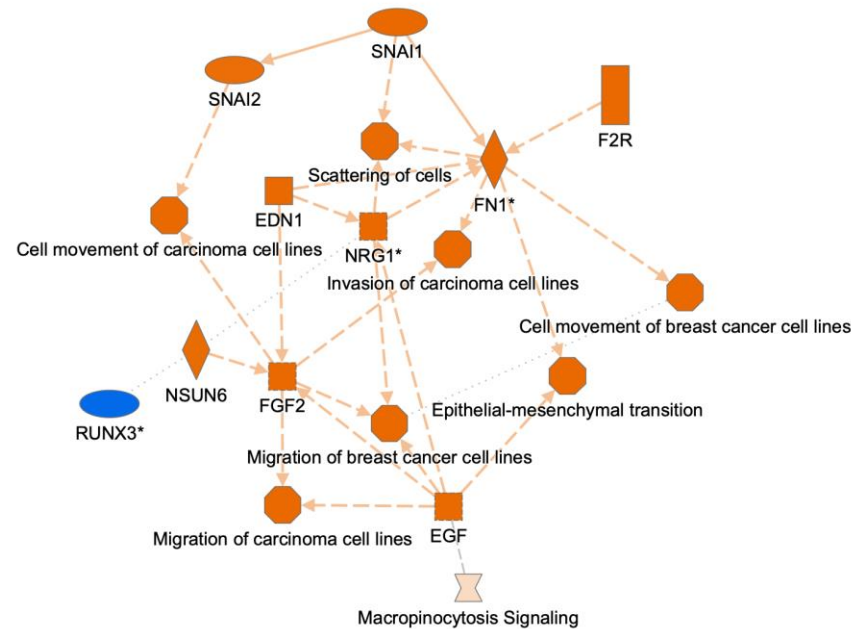
The figure shows how PCSK9 is overexpressed in hepatocyte-like cells, and less so in stem cells and osteoblasts



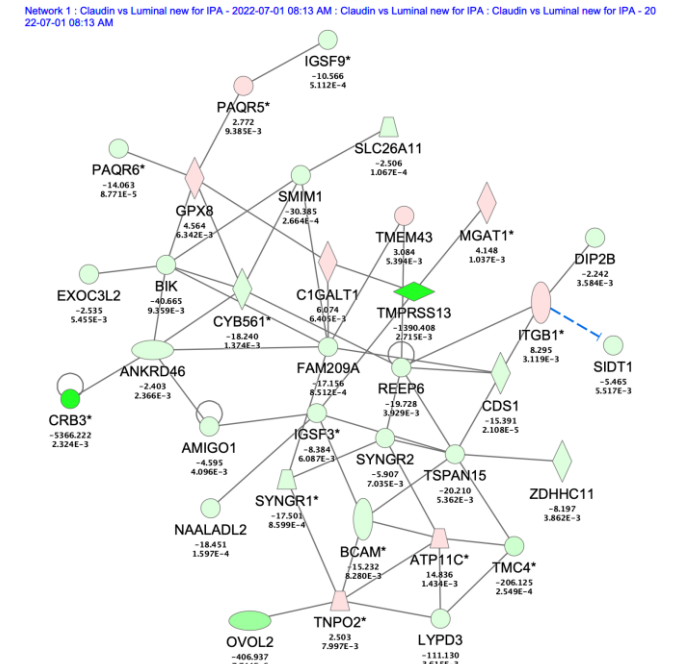
The labels of nodes (e.g., gene names, disease names, etc.) in IPA networks now can be positioned below their node shapes to make them easier to read, especially when they would otherwise be superimposed over dark fill colors. The setting for this feature is found in Preferences > Application Preferences > Graph Appearance, under a section called “Node label position relative to node”



Centered



Below



Below

Now you can analyze datasets from a wider range of species in IPA, including crab-eating macaque, pig, and Chinese Hamster Ovary (CHO) cells. IPA now supports an additional 11 species for a total of 25 supported species

Newly supported species (identifier mapping for dataset upload)

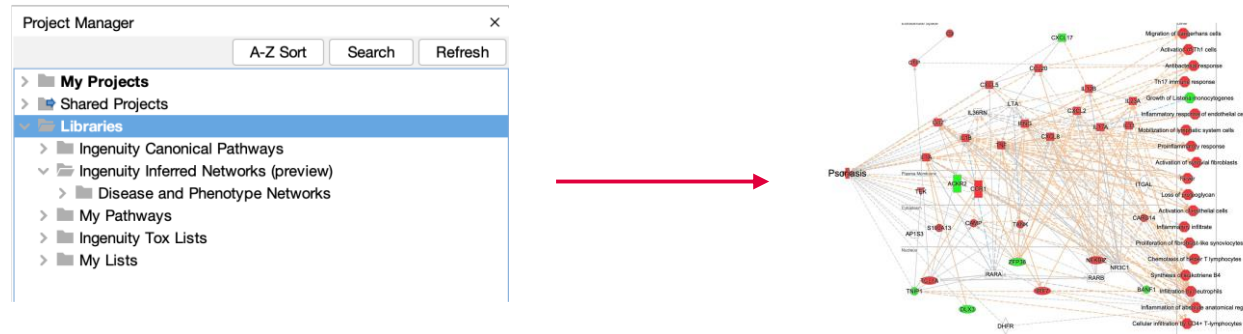
- Bat (Greater horseshoe bat, *Rhinolophus ferrumequinum*)
- Cat (domestic, *Felis catus*)
- Chinese hamster (*Cricetulus griseus*, CHO-K1 genome)
- Crab-eating macaque (*Macaca fascicularis*)
- Frog (Western clawed frog, *Xenopus tropicalis*)
- Golden hamster (*Mesocricetus auratus*)
- Guinea pig, domestic (*Cavia porcellus*)
- Horse (*Equus caballus*)
- Pig (*Sus scrofa*)
- Rabbit (*Oryctolagus cuniculus*)
- Rainbow trout (*Oncorhynchus mykiss*)

The identifiers for these species are mapped according to HomoloGene or NBCI ortholog mapping to the corresponding human, mouse, and rat ortholog information in the QIAGEN Knowledge Base.

The full list of 25 species:

- *Arabidopsis thaliana* (Thale cress)
- Bat (Greater horseshoe bat, *Rhinolophus ferrumequinum*)
- *Caenorhabditis elegans*
- Cat (domestic, *Felis catus*)
- Chicken (*Gallus gallus*)
- Chimpanzee (*Pan troglodytes*)
- Chinese hamster (*Cricetulus griseus*, CHO-K1)
- Cow (*Bos taurus*)
- Crab-eating macaque (*Macaca fascicularis*)
- Dog (*Canis lupus familiaris*)
- Fruit fly (*Drosophila melanogaster*)
- Golden hamster (*Mesocricetus auratus*)
- Guinea pig, domestic (*Cavia porcellus*)
- Horse (*Equus caballus*)
- Human (*Homo sapiens*)
- Mouse (*Mus musculus*)
- Pig (*Sus scrofa*)
- Rabbit (*Oryctolagus cuniculus*)
- Rainbow trout (*Oncorhynchus mykiss*)
- Rat (*Rattus norvegicus*)
- Rhesus Monkey (*Macaca mulatta*)
- *Saccharomyces cerevisiae*
- *Schizosaccharomyces pombe*
- Western clawed frog (*Xenopus tropicalis*)
- Zebrafish (*Danio rerio*)

In 2021 IPA Winter Release, a large library of Disease and Phenotype Networks has been created by leveraging an unsupervised machine learning (ML) model of the literature-derived QIAGEN Knowledge Graph (QKG).



In this release, there are >1,500 of these networks that are now fully discoverable in Search. As shown in Figure, you can search for them in the Pathways and Lists tab using terms such as diseases, phenotypes, or gene names.

Genes and Chemicals Diseases and Functions **Pathways and Lists** Datasets and Analyses

cholestasis Search [Advanced Search](#)

Search Results

Pathways and Lists

[Add To My Pathway](#) [Add To My List](#) [View Report](#) [Activity Plot](#) [Customize Table](#)

The search for cholestasis AND (folder_type:inferred_network) matched 2 pathways and tox lists.

<input type="checkbox"/>	#	Name	Matched Term	Pathway Category	Type	Location
<input type="checkbox"/>	0	Cholestasis	Cholestasis	Ingeniuty Inferred Networks	Ingeniuty Inferred Networks	Libraries > Ingeniuty Inferred Networks > Disease and Phenotype Networks > C > Cholestasis
<input type="checkbox"/>	1	Intrahepatic cholestasis	Intrahepatic cholestasis	Ingeniuty Inferred Networks	Ingeniuty Inferred Networks	Libraries > Ingeniuty Inferred Networks > Disease and Phenotype Networks > I > Intrahepatic cholestasis

Clicking on a Canonical Pathway name in a Gene View will now open the pathway in the IPA client, as shown in Figure 8. The gene of interest will be highlighted in the opened pathway.

IPA Gene View: PDCD1 (Mammalian) > [Interaction Network](#) [Contact Support](#) | [Help Documentation](#)

Review the categorized literature findings and database information for this node.

[Summary](#) [Human](#) [Mouse](#) [Rat](#)

Member Of: [phosphatase](#)

Entrez Gene Name: programmed cell death 1

Synonym(s): B7H1, CD279, hPD-1, hPD-I, hSLE1, LOC100911478, Ly101, PD, PD-1, Pd-1 receptor, Pdc1, PhLP, programmed cell death 1, SLEB2

NCBI CDD Domains (Superfamilies / Multi-Domains): [Immunoglobulin](#), [Immunoglobulin I-set domain](#), [Immunoglobulin like](#), [Immunoglobulin V-set domain](#), [Immunoglobulin V-Type](#)

Protein Functions / Functional Domains: cytosolic tail domain, extracellular domain, immunoglobulin domain, immunoreceptor tyrosine-based inhibition motif, immunoreceptor tyrosine switch motif, protein binding, transmembrane receptor

Subcellular Location: cell surface, cellular membrane, Plasma Membrane, plasma membrane extracellular face

Canonical Pathway: [3-phosphoinositide Biosynthesis](#); [3-phosphoinositide Degradation](#); [D-myo-inositol \(1,4,5,6\)-Tetrakisphosphate Biosynthesis](#); [D-myo-inositol \(3,4,5,6\)-tetrakisphosphate Biosynthesis](#); [D-myo-inositol-5-phosphate Metabolism](#); [PD-1, PD-L1 cancer immunotherapy pathway](#); [Superpathway of Inositol Phosphate Compounds](#); [Systemic Lupus Erythematosus In B Cell Signaling Pathway](#); [T Cell Exhaustion Signaling Pathway](#)

Canonical Pathways

PD-1, PD-L1 cancer immunotherapy pathway

© 2000-2022 QIAGEN. All rights reserved.

What's New in the IPA Summer Release (July 2022)

- Visualize your data in new ways with Canonical Pathway bubble charts
- Visualize OmicSoft single cell data in Land Explorer
- Improve readability of your networks with new node label placement
- Expand your research with support for new species for dataset upload
- Take advantage of inferred disease & phenotype networks with search
- Speed your work by opening Canonical Pathways directly from links on Gene Views
- **Content updates**
- **Version or date of third-party databases**
- **Identifier Mapping**
 - Identifier Source Versions
- **System Requirements**
 - Minimum Software Requirements
 - Minimum Hardware Requirements

Product: QIAGEN IPA
Release Number: Summer Release 2022
Release Date: July 30, 2022

Explore new areas with 10 new Canonical Pathways

- CDX Gastrointestinal Cancer Signaling Pathway
- Immunogenic Cell Death Signaling Pathway
- Macrophage Classical Activation Signaling Pathway
- MicroRNA Biogenesis Signaling Pathway
- Multiple Sclerosis Signaling Pathway
- Pathogen Induced Cytokine Storm Signaling Pathway
- Ribonucleotide Reductase Signaling Pathway
- Role of Chondrocytes in Rheumatoid Arthritis Signaling Pathway
- Role of Osteoblasts in Rheumatoid Arthritis Signaling Pathway
- Role of Osteoclasts in Rheumatoid Arthritis Signaling Pathway

- ~45,000 Expert findings
- ~400,000 cancer mutation findings from ClinVar
- ~18,000 protein-protein interaction findings from BioGrid
- ~2,100 target-to-disease findings from ClinicalTrials.gov
- ~1,900 drug-to-disease findings from ClinicalTrials.gov
- ~800 Gene Ontology findings
- ~300 protein-protein interaction findings from IntAct
- ~180 gene to disease or phenotype associations from the Mouse Genome Database (MGD or "Jax")
- ~40 chemical to cancer findings from the Chemical Carcinogenesis Research Information System (CCRIS)
- ~175 newly mappable chemicals

Land	Repository	Datasets Q1 2022	Datasets Q2 2022	Increase
DiseaseLand	HumanDisease	28,574	29,767	1,193
	MouseDisease	21,912	22,694	782
	RatDisease	7,900	7,900	
	LINCS	28,234	28,234	
OncoLand	OncoHuman (Formerly OncoGEO)	11,692	12,894	1,202
	OncoMouse	1,054	1,054	
	TCGA	4,438	4,438	
	MetastaticCancer	81	81	
	Hematology	4,267	4,267	
	Pediatrics	444	444	
	ENCODE RNA Binding*		486	486
Single Cell Land	SingleCellHuman	207	194	-13
	SingleCellHumanUmi	278	2056	1,778
	SingleCellHumanUmiLite		603	603
	SingleCellHumanHCL	1,476	1,476	
	SingleCellMouse	75	81	6
	SingleCellMouseUmi	751	1,457	706
	SingleCellMouseUmiLite		115	115
Normal Cell and Tissues	Human Tissues (GTEx)	52	52	

This release offers a new source of data called ENCODE RNA binding, which contains RNA-seq experiments of 1,122 samples for two popular cell lines (K562 and HEPG2) after shRNA knockdown targeting various proteins

(<https://www.nature.com/articles/s41586-020-2077-3>):

- RNA-binding proteins
- Transcription factors
- Cofactors
- DNA repair proteins
- Chromatin remodeler proteins
- RNA-polymerase complex
- DNA replication proteins

Source name	Latest version number, download date, or Knowledge Base import date
An Open Access Database of Genome-wide Association Results	Feb-09
APPRIS	human, refseq: rs109_v46 (Feb 22, 2022) human, ensembl: GRCh38 (Feb 22, 2022) mouse, refseq: rs106v29 (Feb 22, 2022) mouse, ensembl: GRCm39 (Feb 19, 2021)
BIND	Dec-08
BioGRID	Version 4.4.208, March 28, 2022
Chemical Carcinogenesis Research Information System (CCRIS)	Oct-06
Clinical Trials	20-Apr-22
ClinVar	14-Apr-22
Cognia	11-Oct-07
Conserved Domain Database	Version 3.19, April 22, 2021
COSMIC	Version 95
DIP	14-Oct-08
DrugBank	27-Apr-16
Entrez Gene	22-Feb-22
Gene Ontology (GO)	22-Mar-22
GVK Biosciences	31-Jul-17
Hazardous Substances Data Bank (HSDB)	Apr-07
HUGO Gene Nomenclature Committee (HGNC)	22-Feb-22
Human Metabolome Database (HMDB)	June 2009 (mapping updated more recently)
Human Phenotype Ontology (HPO)	8-Apr-22
HumanCyc	Version 16.0, February-2012
IntAct	1-Apr-22
Interactome studies	2008
MGI gene IDs	22-Feb-22
MIPS	Apr-08
miRBase	July 22, 2018
miRecords	8-Apr-11
Mouse Genome Database (MGD)	15-Apr-22
NCI Thesaurus	7-Mar-22
Obesity Gene Map Database	Dec-06
OncoTree	oncotree_2019_03_01
Online Mendelian Inheritance in Man (OMIM)	14-Apr-22
Orphanet	7-May-19
PubChem	14-Mar-22
TarBase	18-Jul-11
TargetScan Human	Version 7.2, July 2018
The Human Protein Atlas (THPA)	20.1

Source name	Version
Affymetrix	na36 (7/06/2018)
Applied Bio Systems	various (11/14/2018)
Agilent	various (2/28/2022)
	b151 (chicken, dog, human) b150 (bovine, mouse, zebrafish, Arabidopsis, Macaque, cow) b149 (drosophila, rat, fruit fly) b138 (nematode) b136 (Chimpanzee) (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi)
dbSNP	
EntrezGene	Downloaded on 7/26/2022
Ensembl	Ensembl 106 (Downloaded 07/18/2022) (http://uswest.ensembl.org/info/website/archives/index.html)
Genbank	NCBI-GenBank Flat File Release 204 (https://www.ncbi.nlm.nih.gov/genbank/release/)
HGNC	2/28/2022
	HomoloGene Release 68 (04/09/2014) (http://www.ncbi.nlm.nih.gov/homologene/statistics/) Gene2accession (06/04/2022) Gene_orthologs (06/04/2022)
Homologene	
Illumina	12/8/2010
MirBase	version 22.1 (December 2018)
RefSeq	Release 211 (07/04/2022) (ftp://ftp.ncbi.nlm.nih.gov/refseq/release/release-notes/)
UCSC18hg	5/9/2009 (ftp://hgdownload.cse.ucsc.edu/goldenPath/hg18/database)
UCSC19hg	6/29/2013 (ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/database)
UCSC38hg	11/17/2018 (ftp://hgdownload.cse.ucsc.edu/goldenPath/hg38/database)
Unigene (Retired)	Human(#236), Mouse(#194), Rat(#195), Bovine (#100), Chicken (#46), Dog (#25), Fruitfly (#75), Nematode (#52), Zebrafish(#126), Arabidopsis(#74) (ftp://ftp.ncbi.nih.gov/repository/UniGene/ ,log files)
Uniprot	UniProt release 2021_04, (ftp://ftp.uniprot.org/pub/databases/uniprot/previous_releases/)

Hg38/mm38	
Ensembl Human	Homo_sapiens.GRCh38.106.gtf.gz
Ensembl Mouse	Mus_musculus.GRCm38.106.gtf.gz
RefSeq human	GCF_000001405.40_GRCh38.p14
RefSeq mouse	GCF_000001635.24_GRCm38.p4
Hg19/mm10 from UCSC	
Ensembl Human	April 7th, 2014 (/usr/local/idm/instances/ci/data/download/ucsc/store/19_/ ensGene.txt.gz)
Ensembl Mouse	Feb 6th, 2019
RefSeq Human	Aug 28th, 2019
RefSeq Mouse	Aug 28th, 2019



Better Care with Better Knowledge

Questions?
Genechen@gga.asia