From data to biological insight using QIAGEN OmicSoft and QIAGEN IPA: Single-cell sequencing of normal human liver

Discovery Team, QIAGEN Digital Insights

Legal disclaimer

QIAGEN products shown here are intended for molecular biology applications. These products are not intended for the diagnosis, prevention or treatment of a disease.

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual. QIAGEN kit handbooks and user manuals are available at www.qiagen.com or can be requested from QIAGEN Technical Services or your local distributor.

Objective: Explore the cellular landscape of the human liver with single-cell sequencing

Which cell types comprise a normal human liver?

- How many distinct cell types are in a representative liver sample?

- What gene expression markers best distinguish cell types within a liver?

- Do published cell type signatures correspond to liver cells clustered by expression similarity?

What is the biology shared by cell clusters in scRNA data?

- Which molecular pathways and biological processes are highlighted in the clusters?

- What transcriptional programs underpin each cell type?

What additional biological information can we get by comparing each cluster analysis to >60,000 curated analyses?

Single-cell technology detects a relatively limited set of expressed genes – we can automatically analyze the sets of genes in the clusters to infer the overall biological signatures of groups of individual cells

Agenda

QIAGEN Sample to Insight

To the QIAGEN OmicSoft Lands: Single-cell dataset curation

From the QIAGEN OmicSoft Lands to Array Studio: The single-cell sequencing human liver dataset is processed

From Array Studio to QIAGEN IPA: Highlight the biology of the cell types in the liver

Discover hidden biology

Conclusions

Single-cell sequencing of human liver analyzed using QIAGEN Digital Insights

Background

Data source: MacParland, S. A. et. al (2018). Single-cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nature communications, 9(1), 4383.

- scRNA-seq of whole human liver cellular landscape from liver grafts of five healthy neurologically deceased donors

- Human liver tissue dissociation using collagenase procedure then collection of dissociated single cells

- Single-cell sequencing protocol (10x Genomics Single Cell 3′ v2): Reverse transcription, amplification of cDNA and cDNA library preparation

- Sequencing on HiSeq 2500 followed by alignment to human GRCh38 and CellRanger (10X Genomics) analysis pipeline (UMI counts per gene per cell) and clustering using R

Our processing: Alignment to Human B37 using OSA, quantification using RSEM and dimension reduction using t-SNE

Top gene markers from single-cell inference reports based on new cellular clusters
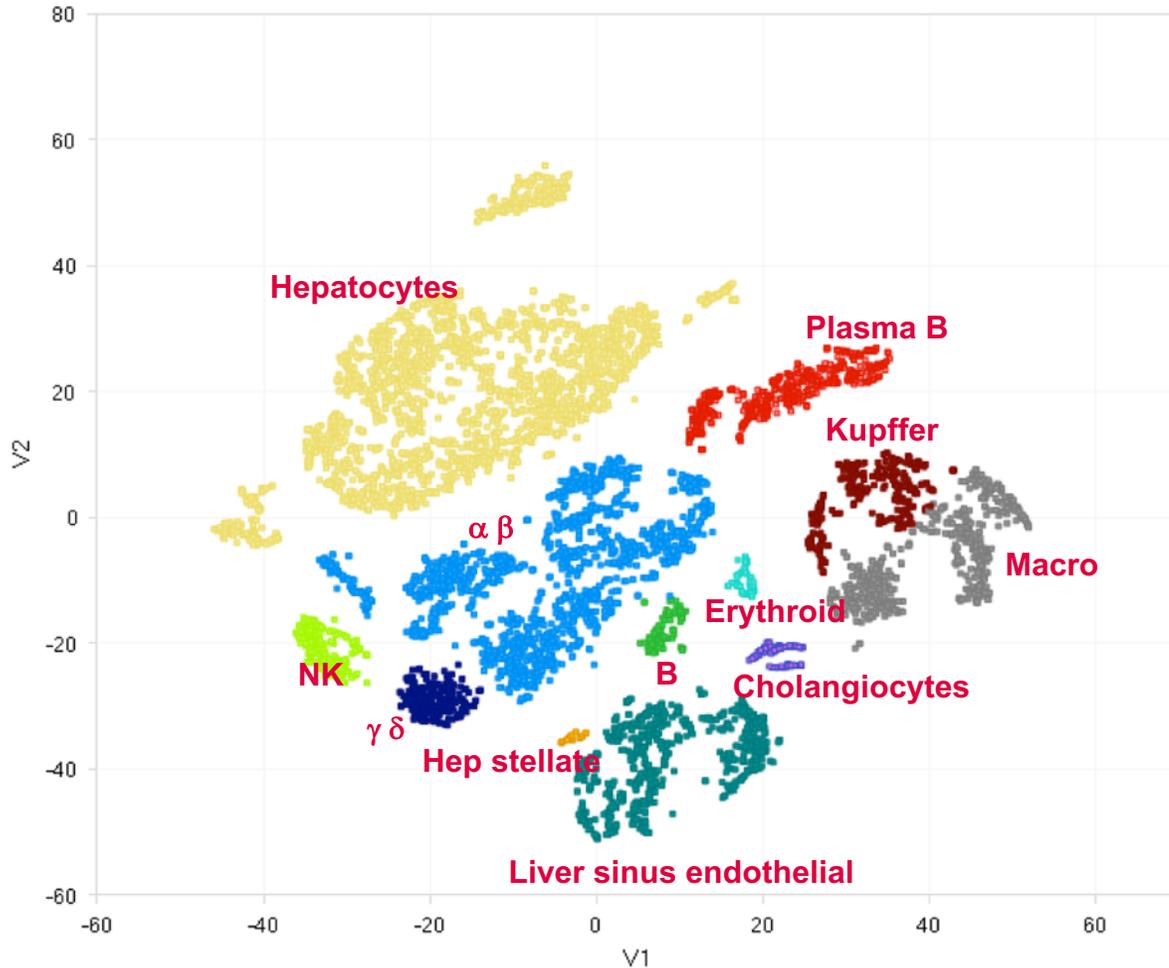

Differentially expressed genes

Hepatocytes vs Others.SeuratMarkerTest | SCInferenceReport

100%   1626 Rows, 15 Column

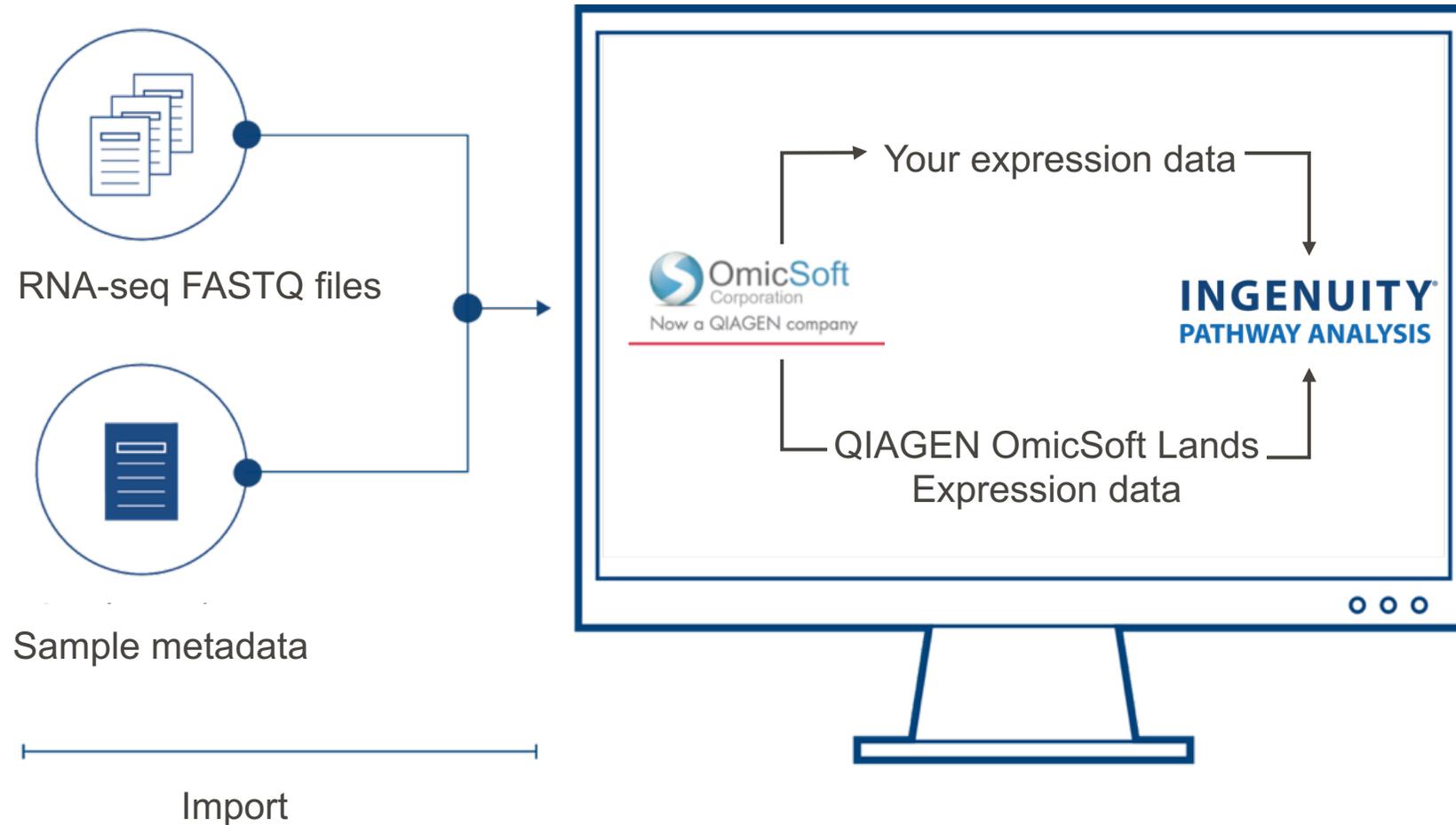| GeneID	Cell List Item => Hepatocyte vs other cells Estimate	Cell List Item => Hepatocyte vs other cells Fold Change	Cell List Item => Hepatocyte vs other cells Raw Pvalue	Cell List Item => Hepatocyte vs other cells Bonferroni	Hepatocyte	other cells	Gene Name
TTR	3.8197	14.1199	0.0000e+000	0.0000e+000	7.2005	3.3808	TTR
ANG	3.1857	9.0987	0.0000e+000	0.0000e+000	4.8304	1.6447	ANG
RBP4	3.0908	8.5197	0.0000e+000	0.0000e+000	6.0040	2.9132	RBP4
APOH	2.9393	7.6706	0.0000e+000	0.0000e+000	5.2150	2.2756	APOH
APOC3	2.9330	7.6368	0.0000e+000	0.0000e+000	8.9273	5.9943	APOC3
APOA1	2.8627	7.2740	0.0000e+000	0.0000e+000	7.1251	4.2623	APOA1
ORM1	2.8424	7.1721	0.0000e+000	0.0000e+000	8.2257	5.3833	ORM1
APOC1	2.8058	6.9923	0.0000e+000	0.0000e+000	7.4948	4.6891	APOC1
APOC4-APO...	2.6798	6.4077	0.0000e+000	0.0000e+000	5.6073	2.9275	APOC4-APO...
APOA2	2.6715	6.3707	0.0000e+000	0.0000e+000	8.1465	5.4750	APOA2

Color by CellListItem => Hepatocyte vs other cells.FoldChange
-22.09   6.01
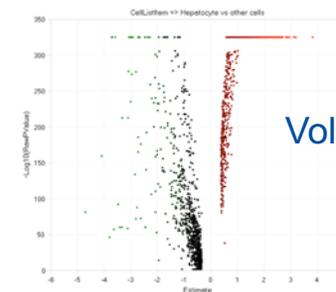
Volcano plot

QIAGEN OmicSoft to QIAGEN IPA

RNA-seq FASTQ files

Sample metadata

Import

Your expression data

QIAGEN OmicSoft Lands Expression data

Summary of a QIAGEN IPA core analysis: Hepatocytes versus others
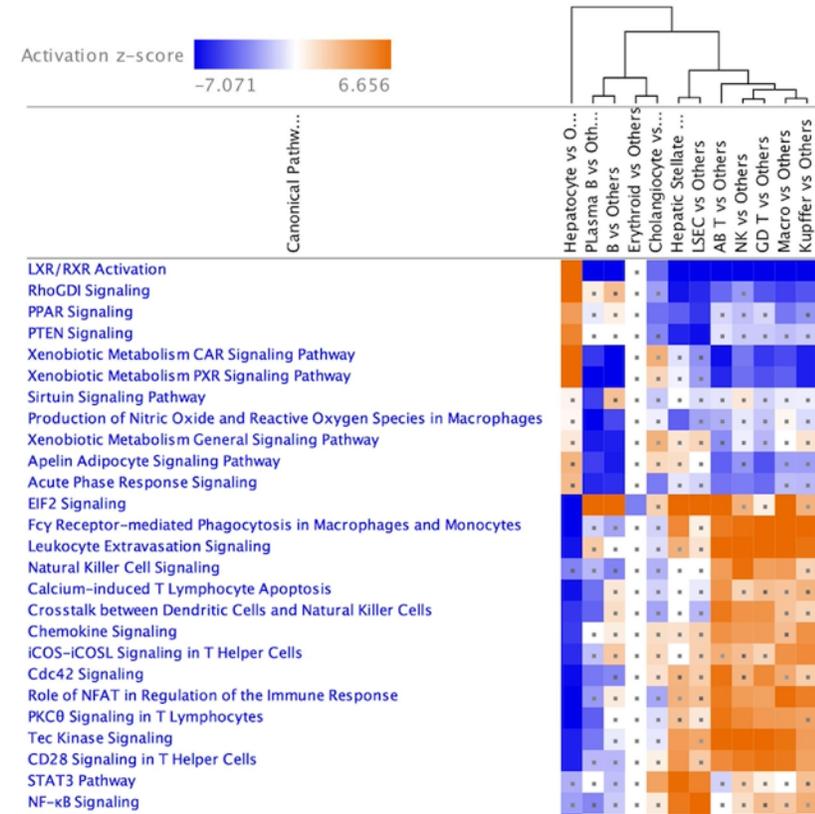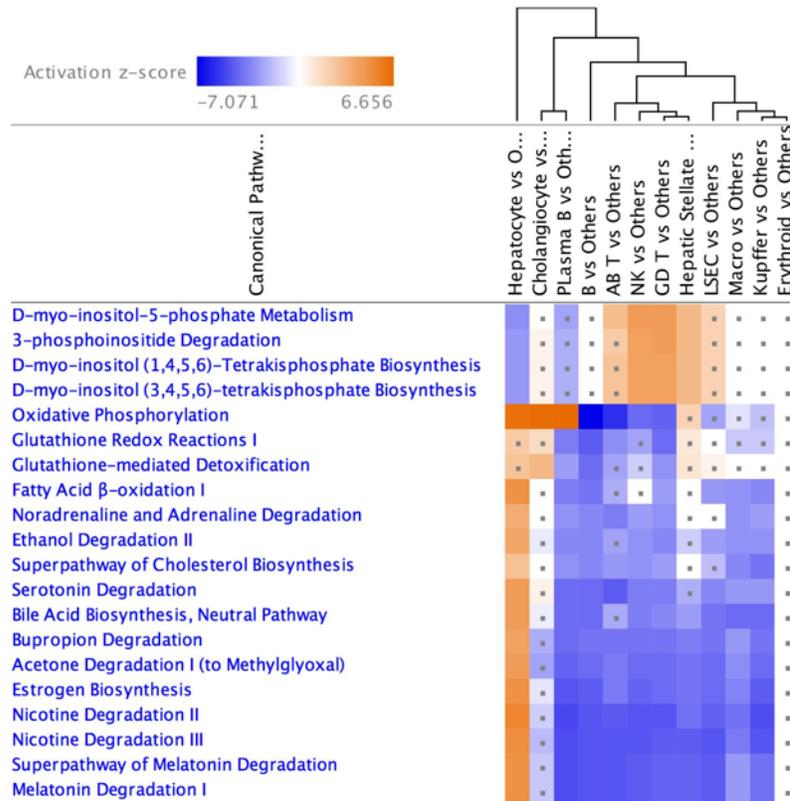

Summary at the gene level:

- Fold change >1.5
- $p < 0.05$

Single-cell RNA-seq of normal human liver reveals cell-specific canonical pathways

Metabolic pathways – liver specificity in metabolic functions

## Signaling pathways reveals the immunobiology of the liver

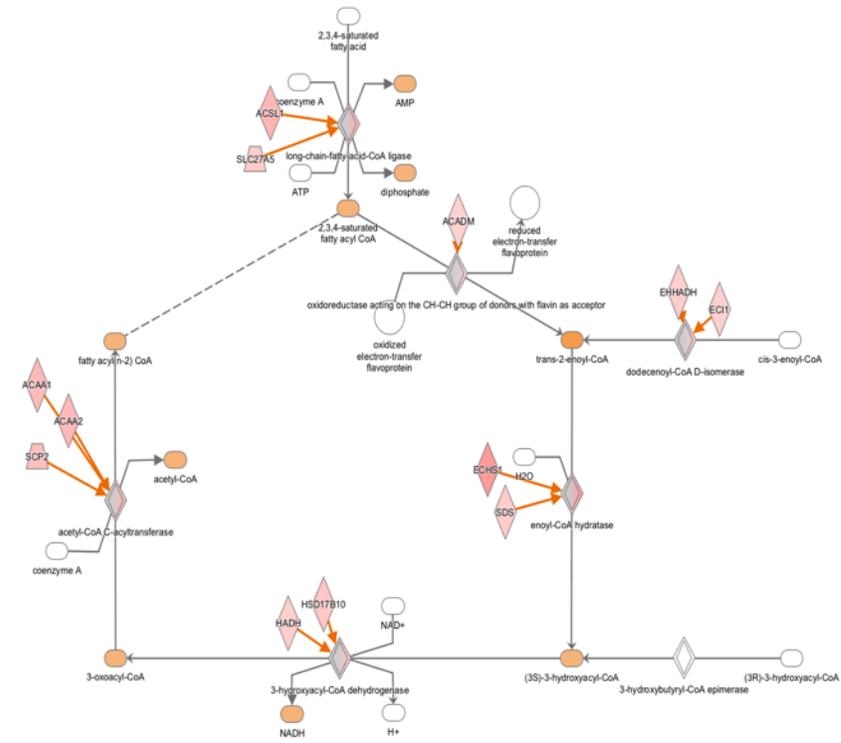● Hepatocytes, other hepato-specific cells and immune cells populate the normal human liver.

Hepatocytes perform important roles in normal liver function (Canonical Pathways)

These metabolic pathways are predicted to be activated in hepatocytes vs others and inhibited in liver immune cells
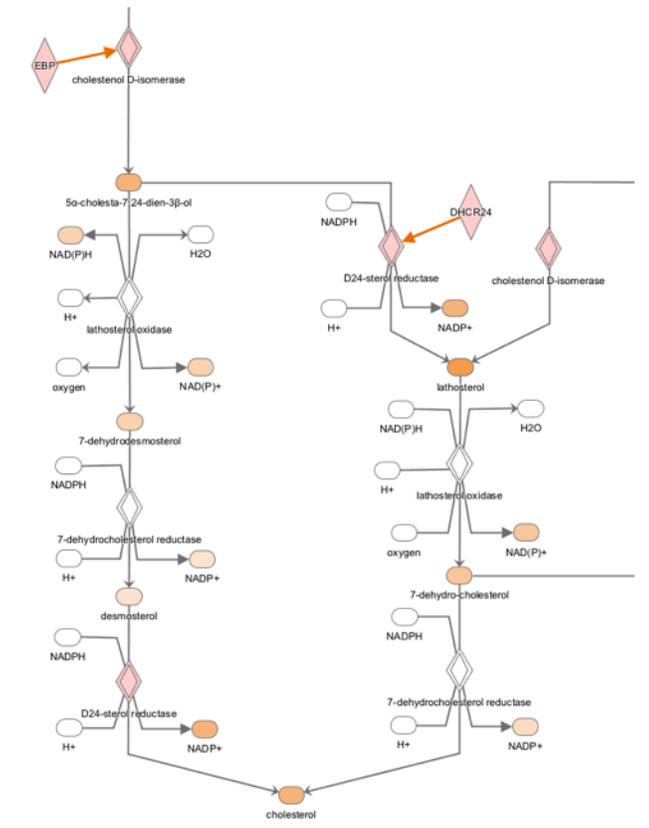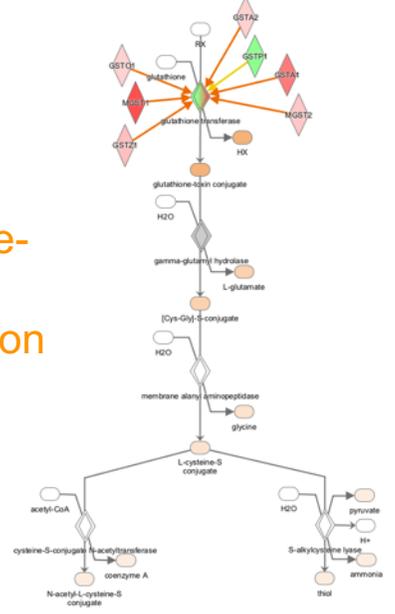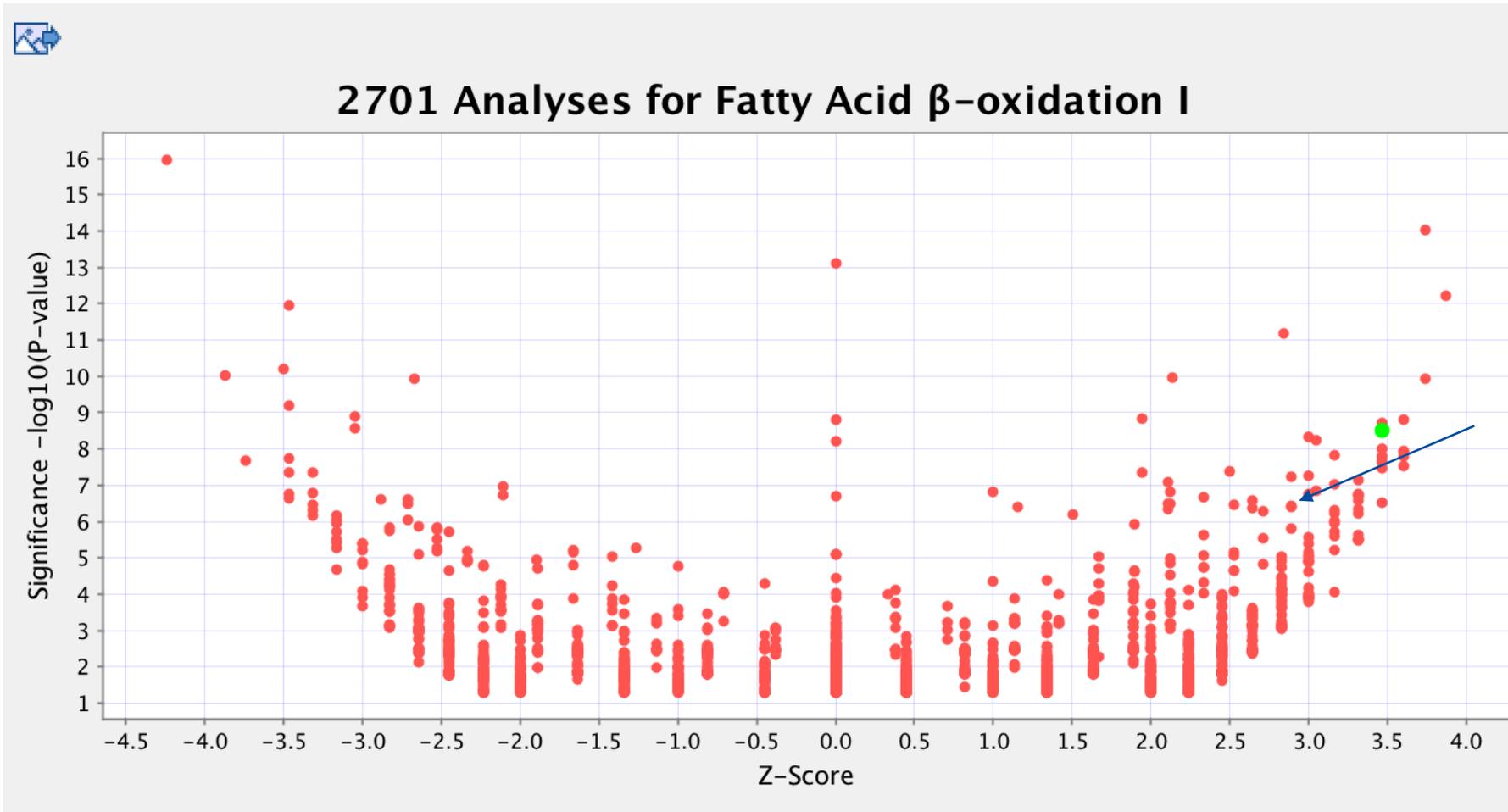

Lipolysis
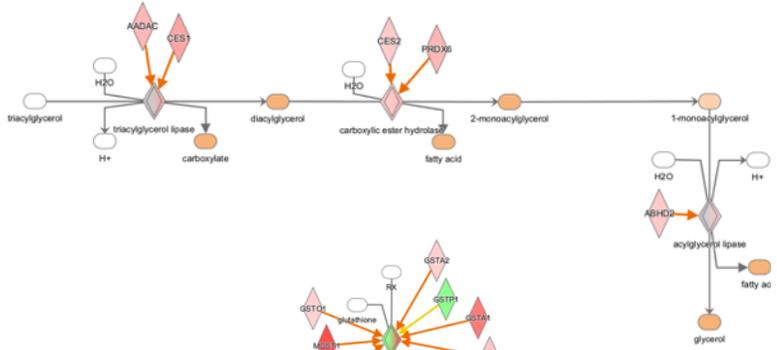
Fatty β-oxidation

Cholesterol biosynthesis

Glutathione-mediated detoxification

Fatty acid β-oxidation (FABO) pathway activity across thousands of datasets

Activity Plot - Fatty Acid β-oxidation I

2701 Analyses for Fatty Acid β-oxidation I

Case/Control Differences

Key	Case	Control
dosage	0.010% w/w	NA
pretreatment	high fat diet	chow diet
subjecttreatment	high fat diet;rosiglitazone	chow diet

Comparison Context

animalstrain	C57BL/6
comparisoncategory	Treatment1 vs. Treatment2
comparisoncontrast	Tissue:ExperimentGroup => subcutaneous adipose tissue -> high fat diet 9 weeks and high fat diet with rosiglitazone 7 weeks vs chow diet 16 weeks
diseasestate	hypercholesterolemia

Activity plot for a canonical pathway

What analysis metadata are significantly associated in FABO pathway-activated cases?

Filter z-score >2 in DiseaseLand and OncoLand

Analysis Name	Project	P-value	Z-sc...	ca...	ca...	ca...	co...	co...	co...	w...
2- normal control [white adipose tissue] NA	MouseDisease	5.88E-13	3.873	normal co...	white adip...	NA	Treatment...	Tissue:Su...	GSE112582.	https://www
4- nonalcoholic fatty liver disease (NAFLD)	MouseDisease	9.58E-15	3.742	nonalcohol...	liver	NA	Disease vs...	AnimalStr...	GSE43106.C	http://www.
5- normal control [liver] NA 8235	MouseDisease	1.13E-10	3.742	normal co...	liver	NA	Treatment ...	Tissue:Co...	GSE67796.C	http://www.
3- obesity [perigonadal adipose tissue] NA	MouseDisease	2.99E-08	3.606	obesity	perigonad...	NA	Treatment ...	SubjectTr...	GSE8831.GF	https://www
1- normal control [white adipose tissue] ros	MouseDisease	1.61E-08	3.606	normal co...	white adip...	rosiglitazo...	Treatment ...	Transfecti...	GSE35011.G	https://www
5- obesity [perigonadal adipose tissue] NA	MouseDisease	1.13E-08	3.606	obesity	perigonad...	NA	Treatment ...	SubjectTr...	GSE8831.GF	https://www
2- normal control [white adipose tissue] ros	MouseDisease	1.23E-08	3.606	normal co...	white adip...	rosiglitazo...	Treatment ...	Transfecti...	GSE35011.G	https://www
32- normal control [bone marrow] triphenyl	MouseDisease	1.61E-09	3.606	normal co...	bone marr...	triphenyl ...	Treatment...	Treatment...	GSE119541.(	https://www
21176- hepatocellular carcinoma (LIHC) [liv	LINCS	2.97E-07	3.464	hepatocell...	liver	PD184352	Treatment ...	Treatment...	GSE70138.G	https://www

Evaluate Metadata   Customize Table   Z-score 3.873 - 2.449   (p1 of 3)   More Info

Selected 0 / 513

FABO activity significantly enriched with high-fat diet and PPAR agonists (anti-diabetics)

PPAR nuclear receptors are involved in the control of metabolism and inflammation in metabolic disease and immunity

Significant metadata in 200 selected analyses (Hepatocyte vs Others)

200 repository analyses selected

Metadata field	Significant term	p-value	Selected analy...	Total analyses ...	Selected an...	Total analy...
case.subjecttreatment	high fat diet;rosiglitazone	1.58E-21	12	19	78	9636
control.treatment	DMSO;differentiation medium	9.74E-16	7	11	200	62813
case.subjecttreatment	high fat diet;pioglitazone	6.15E-15	8	12	78	9636
case.treatment	rosiglitazone;differentiation medium	2.69E-14	6	8	200	62813
case.treatment	rosiglitazone	2.71E-10	6	27	200	62813
control.subjecttreatment	chow diet	2.65E-09	16	311	78	9636
case.subjecttreatment	WY-14643	1.37E-07	4	7	78	9636
case.treatmentgroup	MAPK/ERK signaling inhibitor	1.93E-07	19	2839	48	26436
case.pretreatment	high fat diet	2.28E-07	14	304	28	2804
control.subjecttreatment	high fat diet	2.42E-07	14	322	78	9636
case.subjecttreatment	pioglitazone	4.85E-07	4	9	78	9636
case.treatment	MEHP;differentiation medium	6.32E-07	3	6	200	62813
control.treatment	MEHP;differentiation medium	1.10E-06	3	7	200	62813
case.treatment	tributyltin;differentiation medium	3.75E-06	3	10	200	62813
case.treatment	PD184352	3.03E-05	5	111	200	62813
case.subjecttreatment	fasting	1.00E-04	5	58	78	9636
control.treatment	plating medium	1.50E-04	2	6	200	62813
case.subjecttreatment	5 degrees Celsius	1.93E-04	2	3	78	9636

Selected 0 / 85

Hepatocyte top gene markers are involved in multiple roles and functions (Seurat)
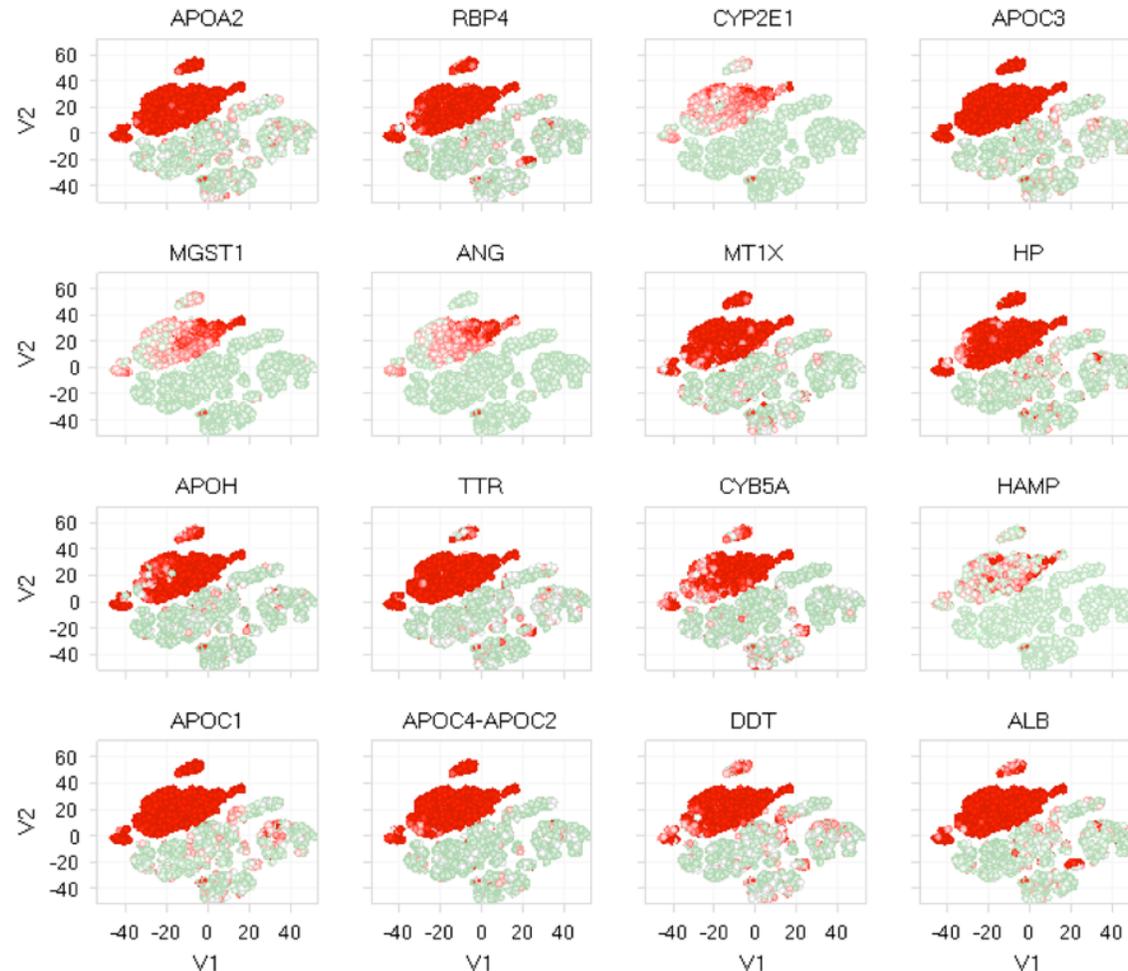
Metabolism:
- Cholesterol
- Triglyceride
- Lipoprotein
- Fatty acids
- Drug
- Mitochondrial

Transport:
- Cholesterol
- Retinol (vit A)
- Thyroid hormones

Coagulation

Blood vessel formation

- Fibrosis
- Inflammatory processes
- Hematopoiesis
- Macrophage polarization (M2)
- Tumor suppressor
- Oxidative stress
- Binding plasma globin
- Iron homeostasis

Single-cell RNA-seq highlights the status of cholangiocytes (Canonical Pathways)

Cholangiocytes line the intra- and extra-hepatic bile ducts composing the biliary epithelium, and are normally quiescent in the liver, but they respond to injury or stress (such as altered hepatic metabolism) by enhanced proliferation

EPCAM+

Cholangiocytes
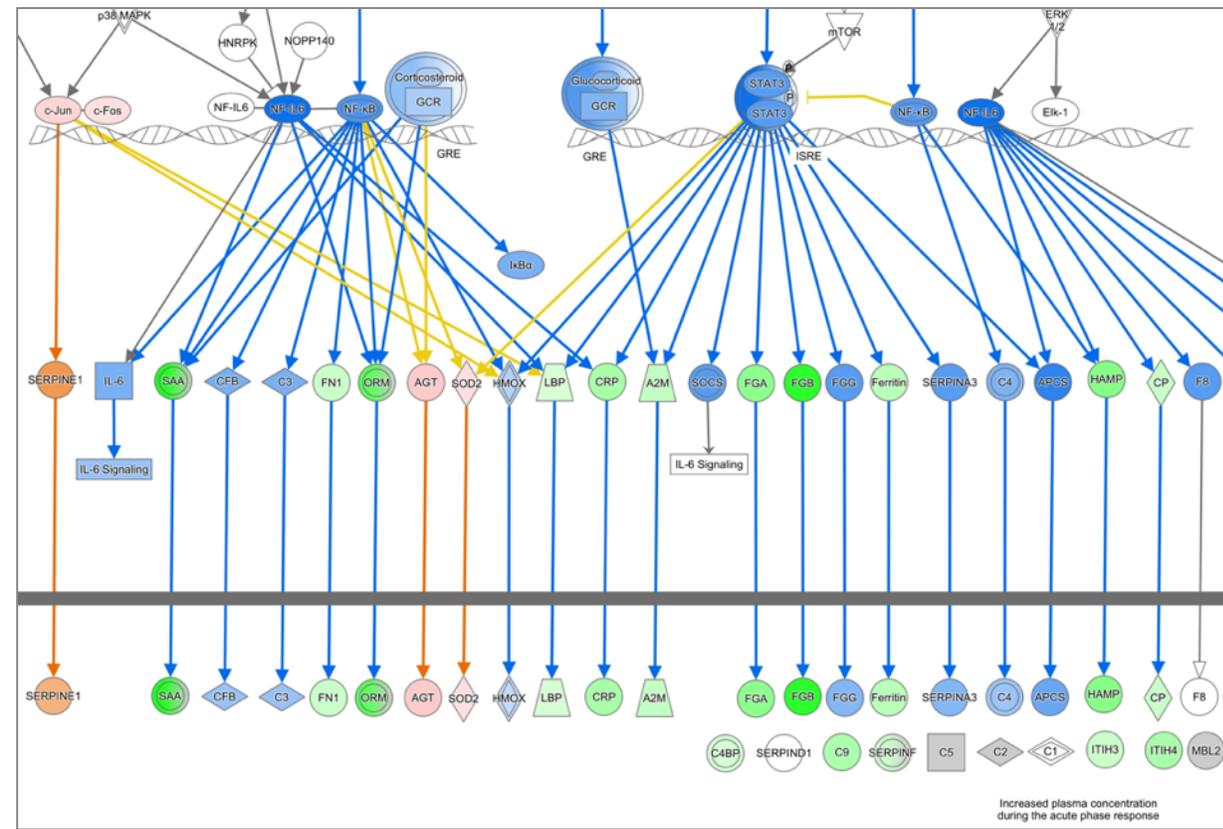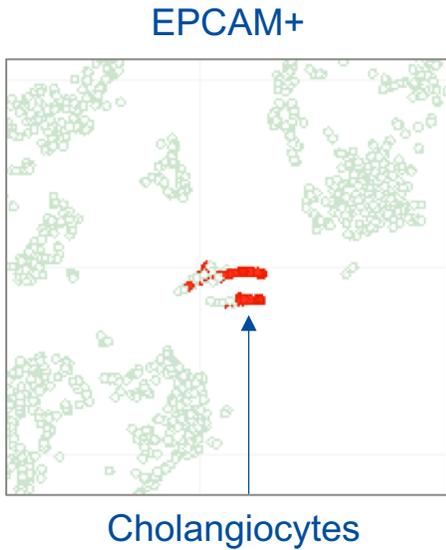
- Acute phase response signaling is predicted to be inhibited
- Indicating absence of stress to the normal liver such as in this analysis
- p-value: 9.89 E-20
- Z-score:-2.065

VEGFA is predicted to be activated in liver sinus endothelial cells (LSECs)

Upstream Analysis

Z-score 4.407, p-value: 3.96 E-20
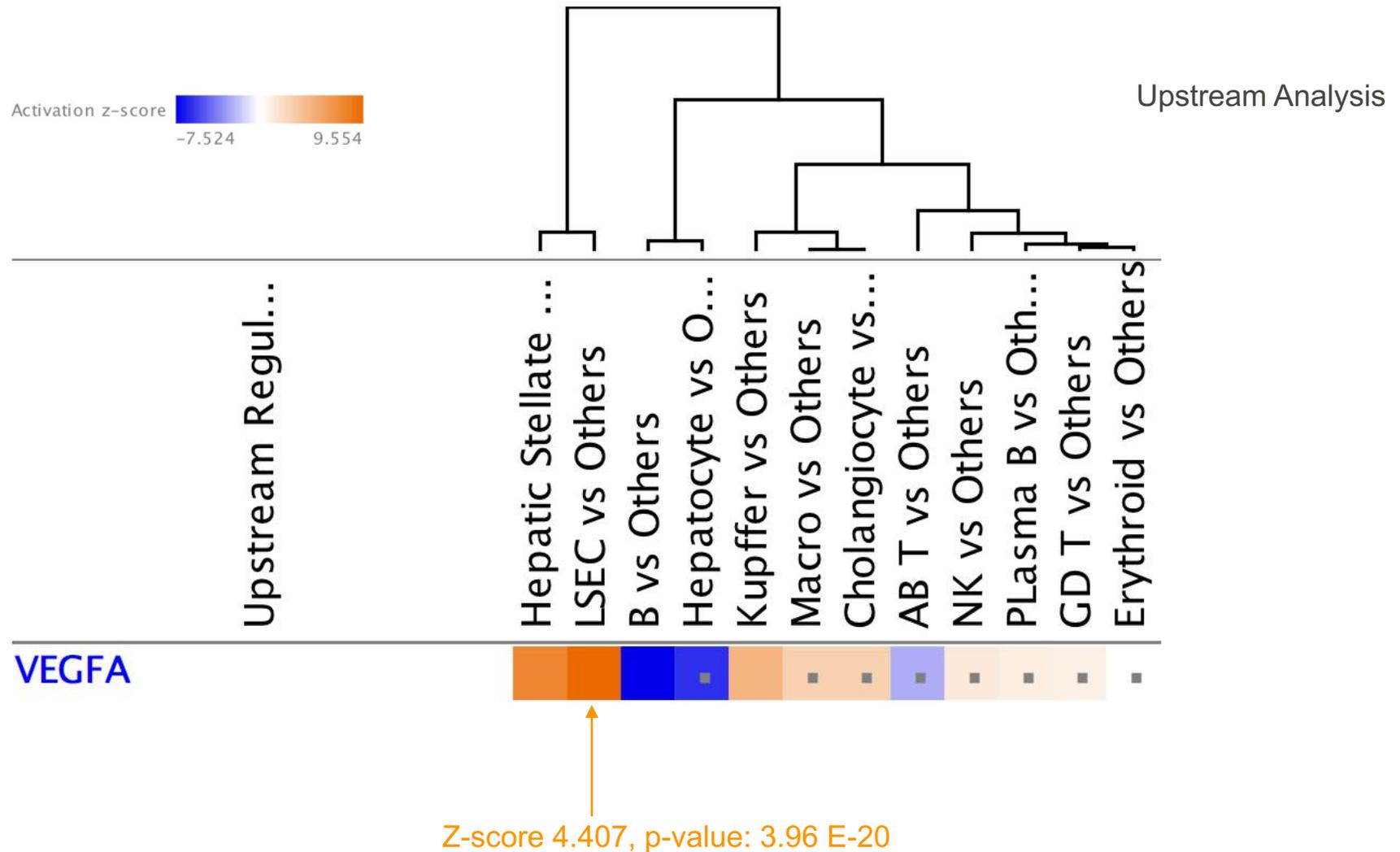
LSECs control tolerance in normal liver; implicated in angiogenesis in normal human liver

VEGFA is predicted to be activated in LSEC

VEGFA is expressed only in hepatocytes

CLEC4G+
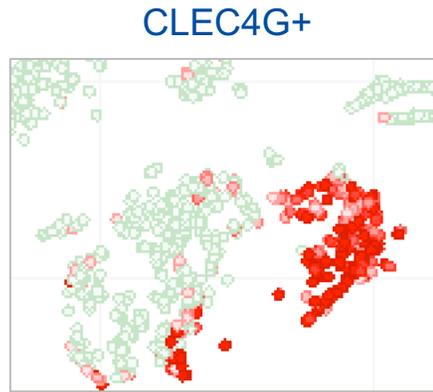
Liver sinus endothelial cells
(LSEC) (multiple sub-clusters)

VEGFA

Angiogenesis
p-value:6.8 E-41

VEGFA

Hepatocytes
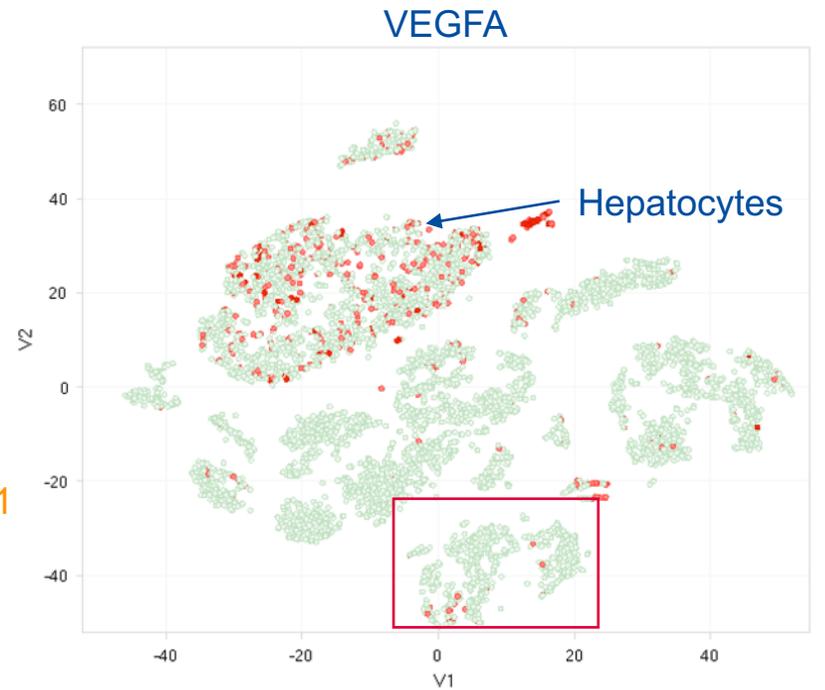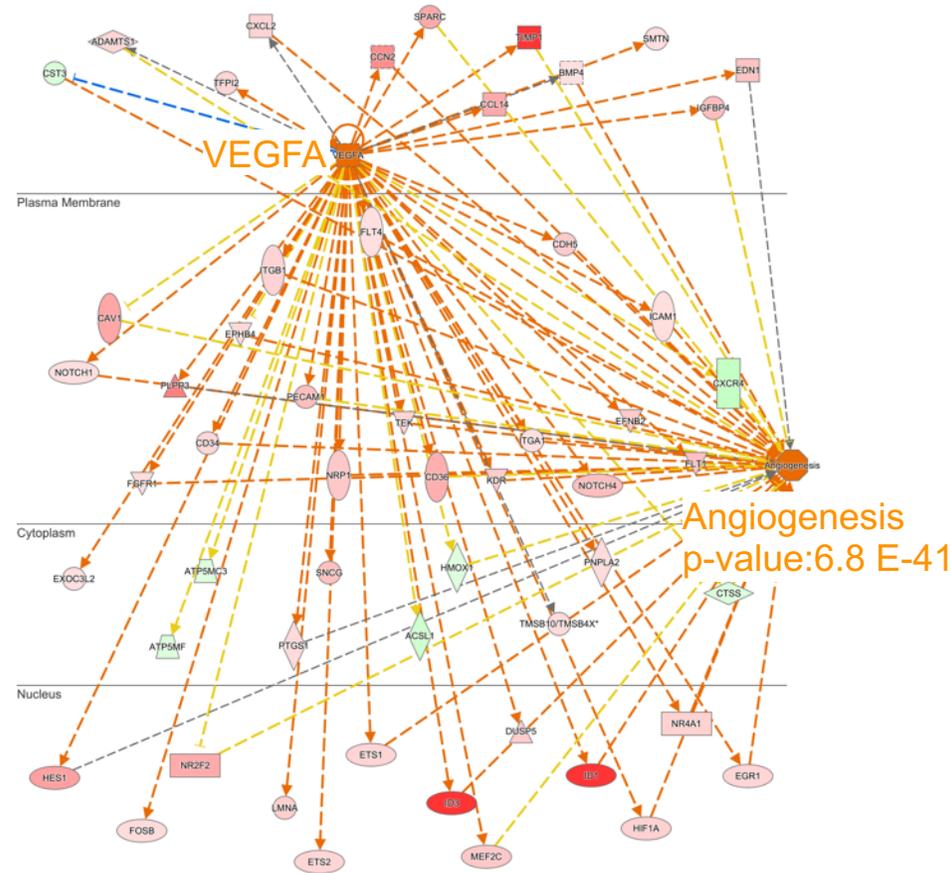
Possible signaling from one cell type to another

Kupffer cells drive host defense, iron metabolism, phagocytosis and liver homeostasis

SPI1 (PU.1, an ETS transcription factor) is predicted to be activated across many cell types, but is detectable mainly in Kupffer cells

Kupffer cell markers

CD5L — Innate immune system (inhibitor of chronic liver injury)

MARCO — Innate antimicrobial immune system

VSIG4 — Potential negative regulator of T cell response
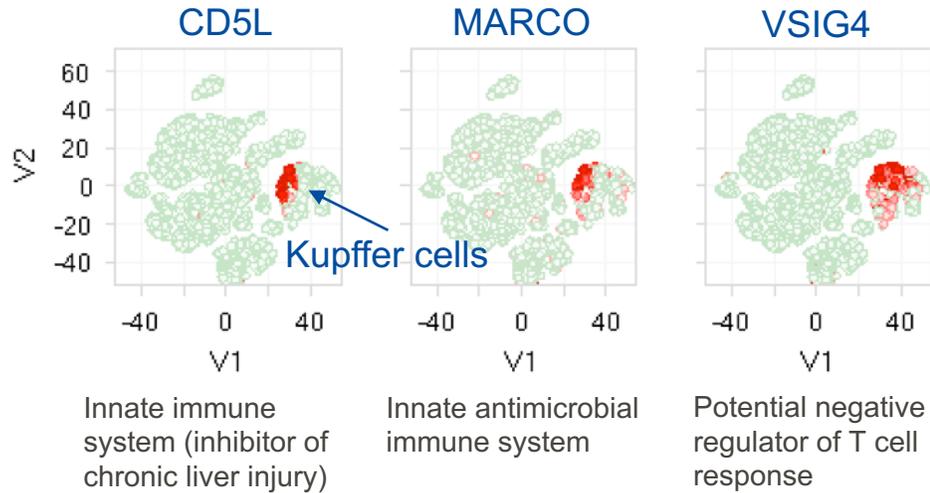
Kupffer cells

SPI1 is predicted to be activated most highly in Kupffer cells

Z-score 2.938, p-value: 6.01 E-17
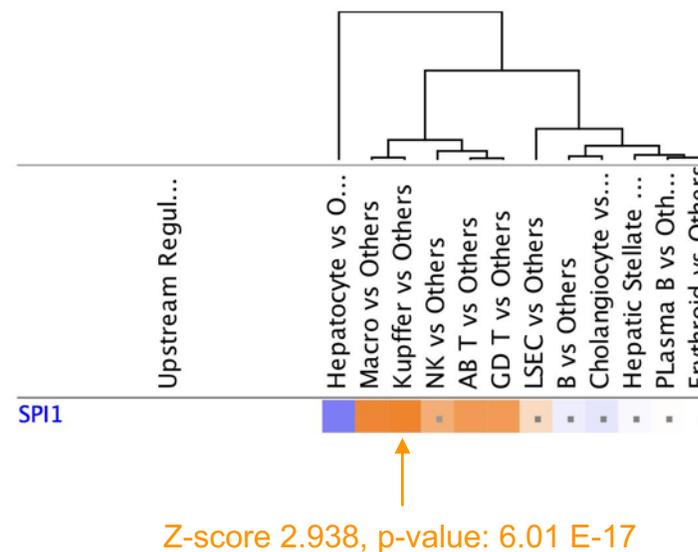
SPI1 is expressed in Kupffer cells
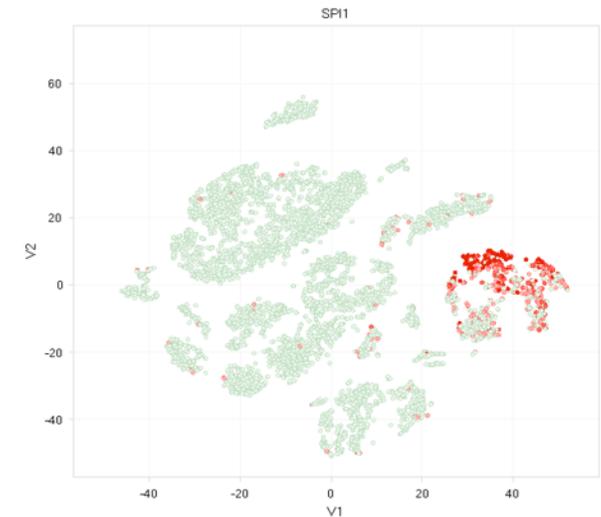
Kupffer cells are involved in immune responses via activation of SPI1

Immune response of leukocytes

p-value: 9.37 E-15
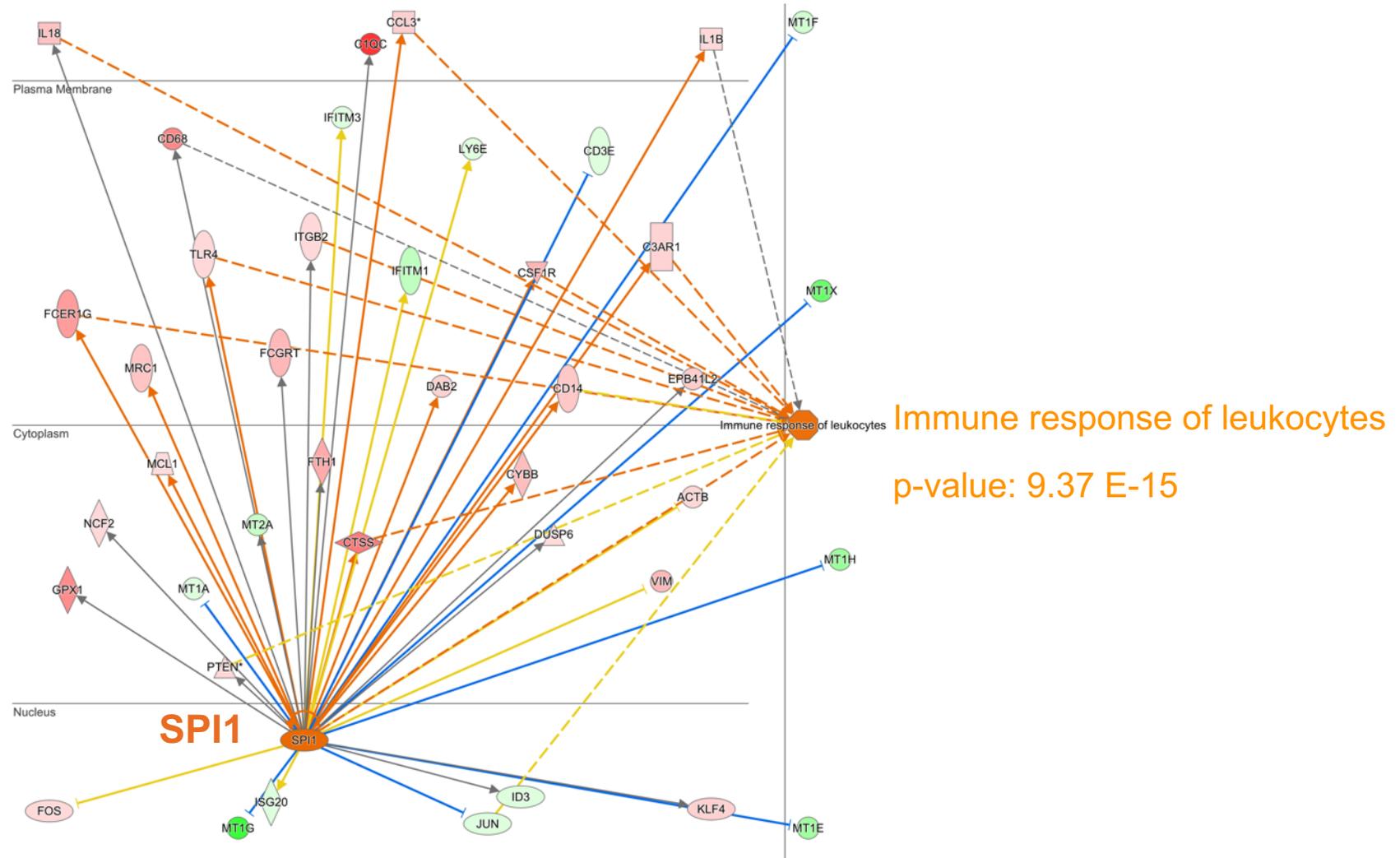
PF4 predicted to be activated in Kupffer cells and drive hepatic fibrosis & innate immunity

PF4 is platelet factor 4, a CXC chemokine

PF4 is not expressed in normal human liver
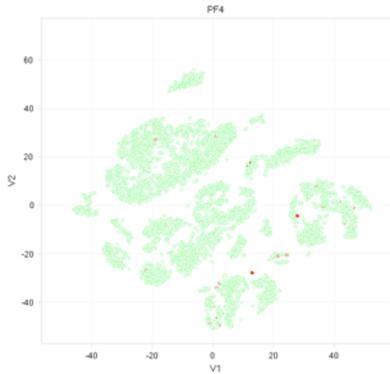
PF4 is predicted to be activated in Kupffer cells
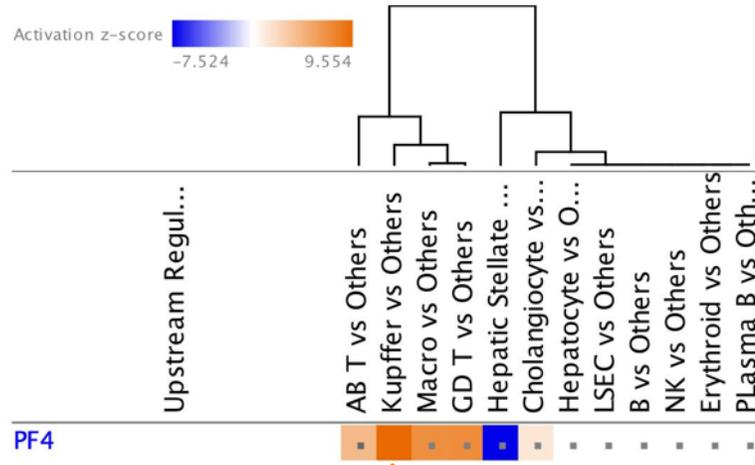
Z-score 2.575, p-value: 1.66 E-04
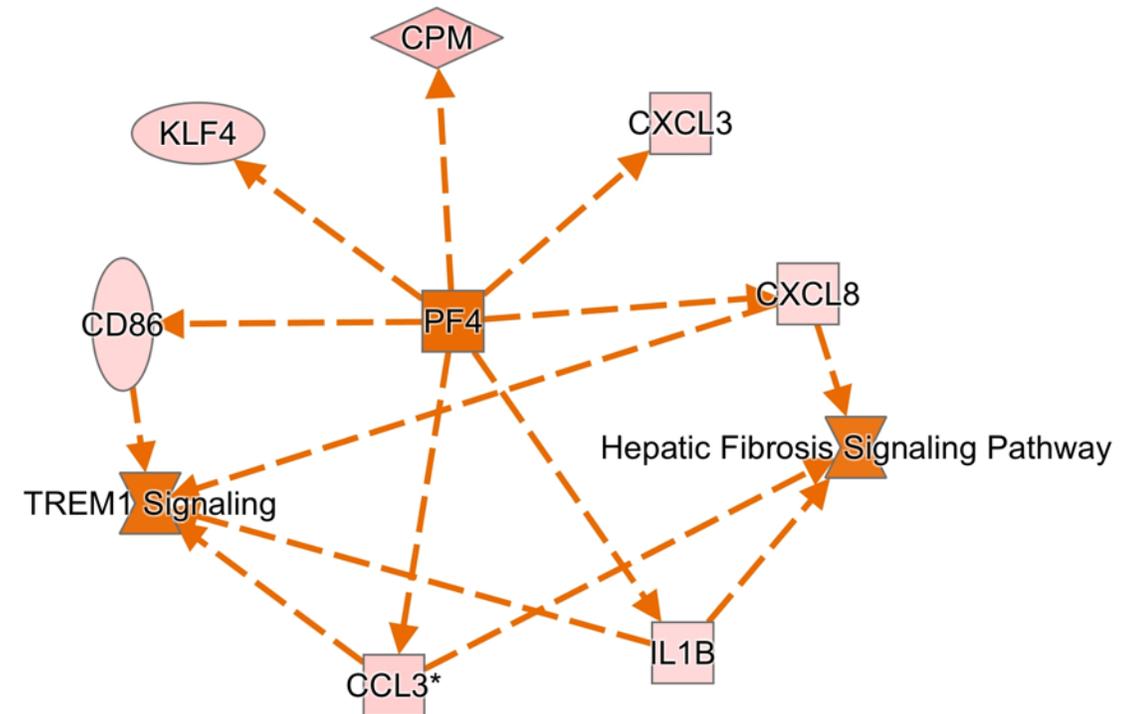
PF4 protein is likely supplied by the blood, activating targets in the Kupffer cells.

Predicted activity (z-score) of PF4 and SPI1 overlaid on t-SNE of normal human liver

PF4 (platelet factor 4)
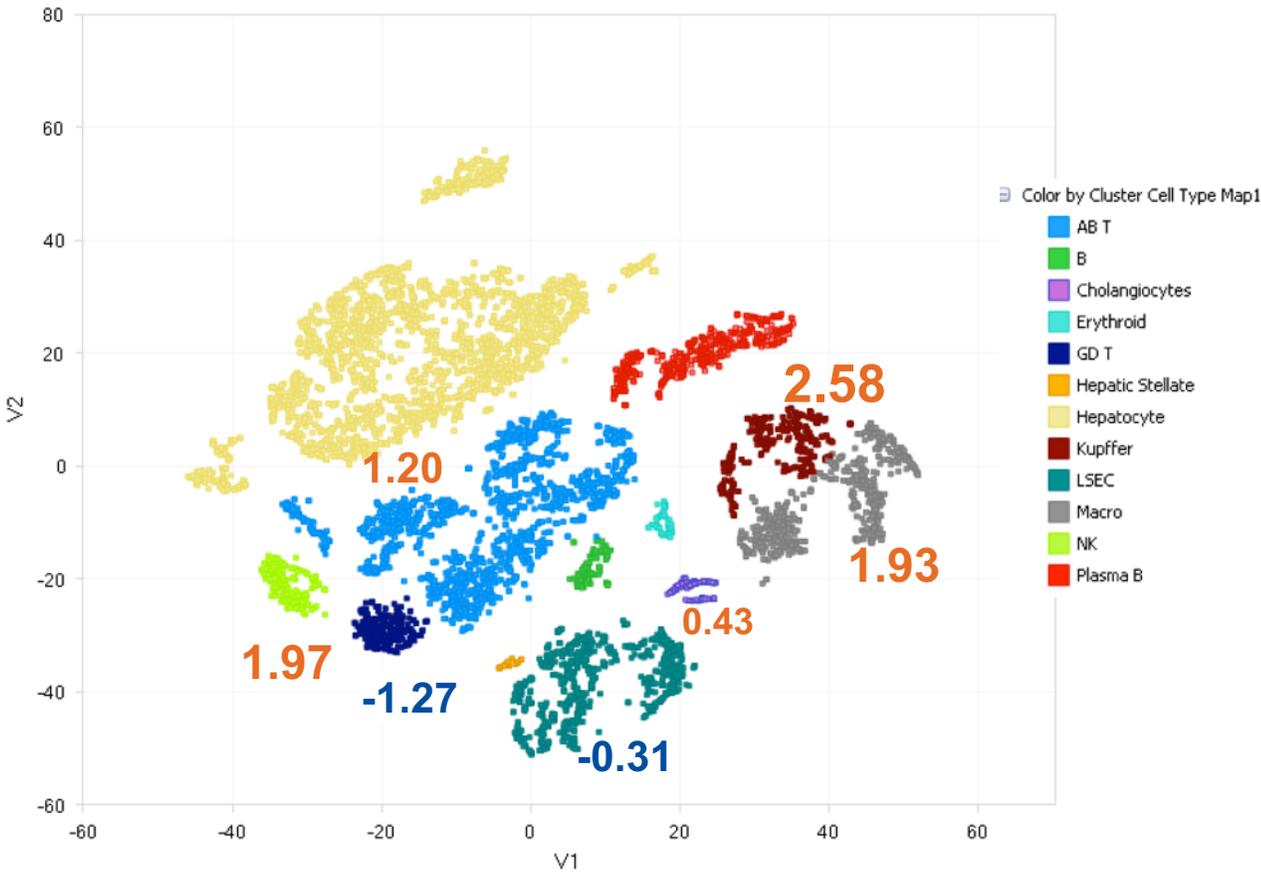
SPI1 (also known as PU.1)
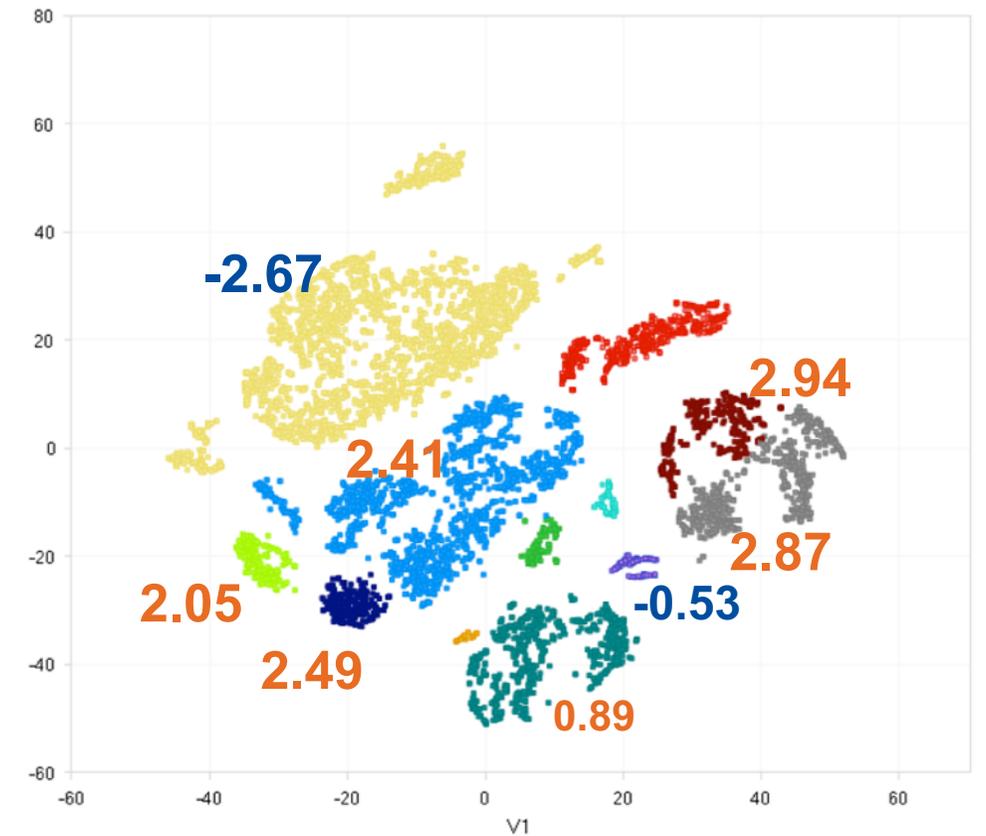
What can we discover about PF4's expression? (QIAGEN OmicSoft Land Explorer)

GTEx B38: (51 normal human tissues)

Blueprint B38 (normal immune cells)

PAX5 is predicted to be activated in resident B cells present in normal human liver

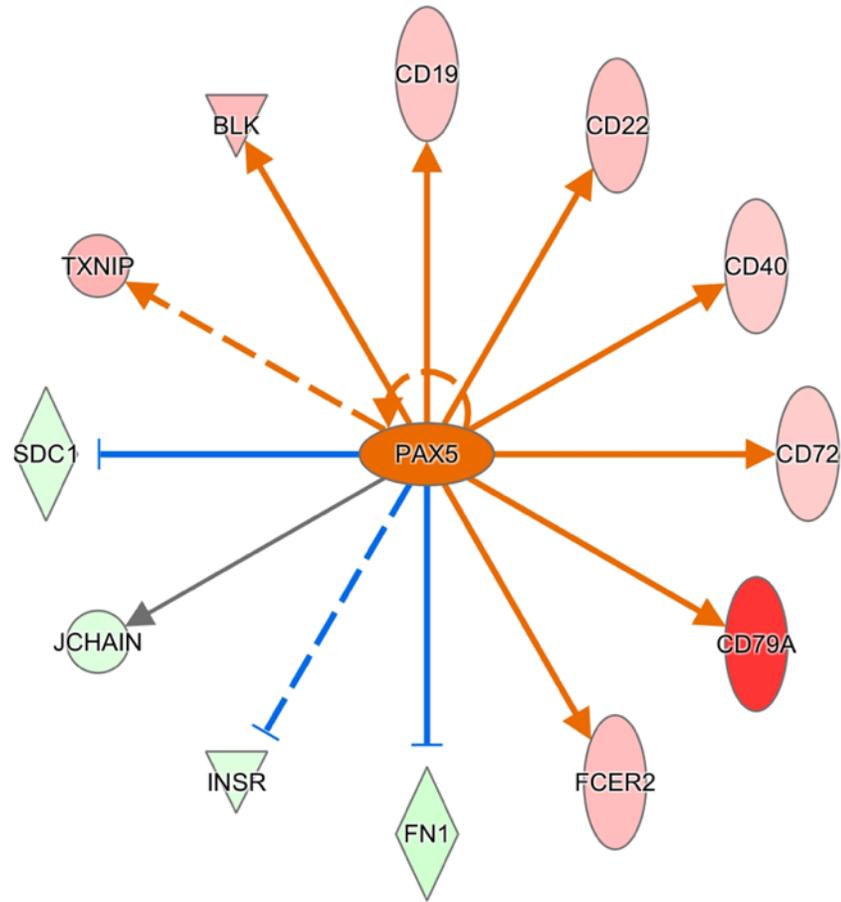## PAX5 is a member of the paired box (PAX) family of transcription factors

PAX5 is almost not detectable in this normal human liver

PAX5 encodes a B-cell lineage specific activator protein that is expressed at early, but not late stages of B-cell differentiation
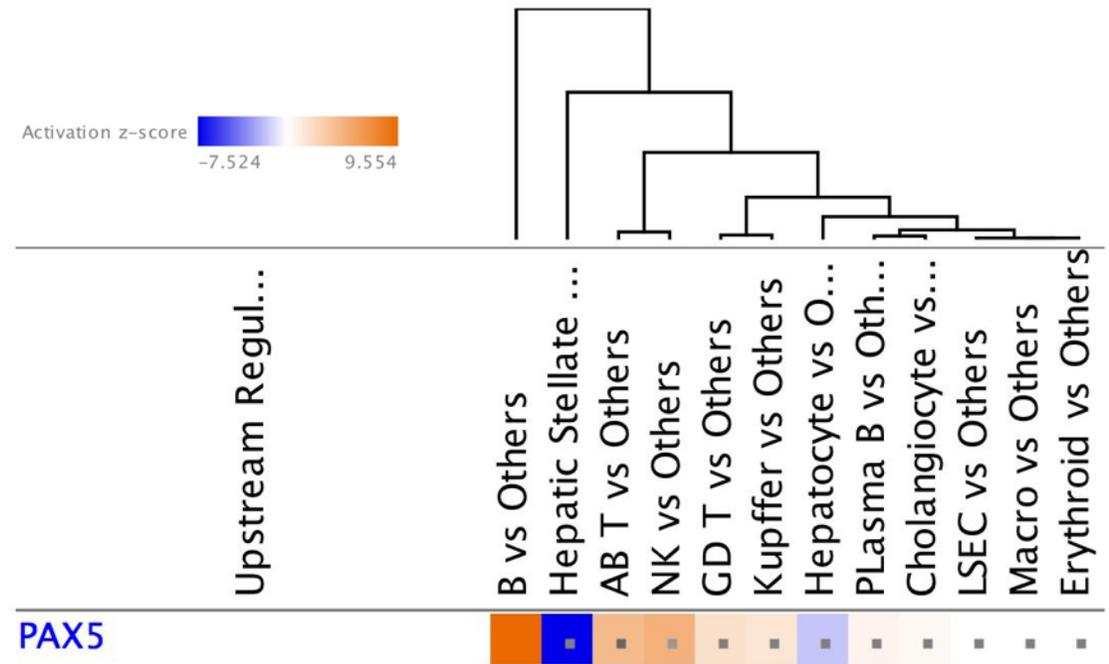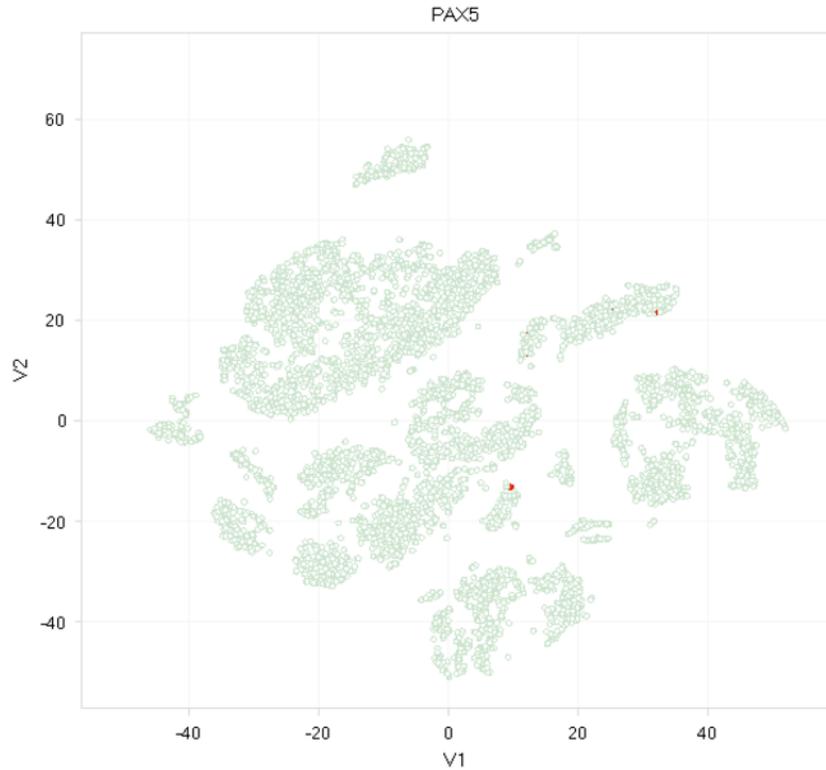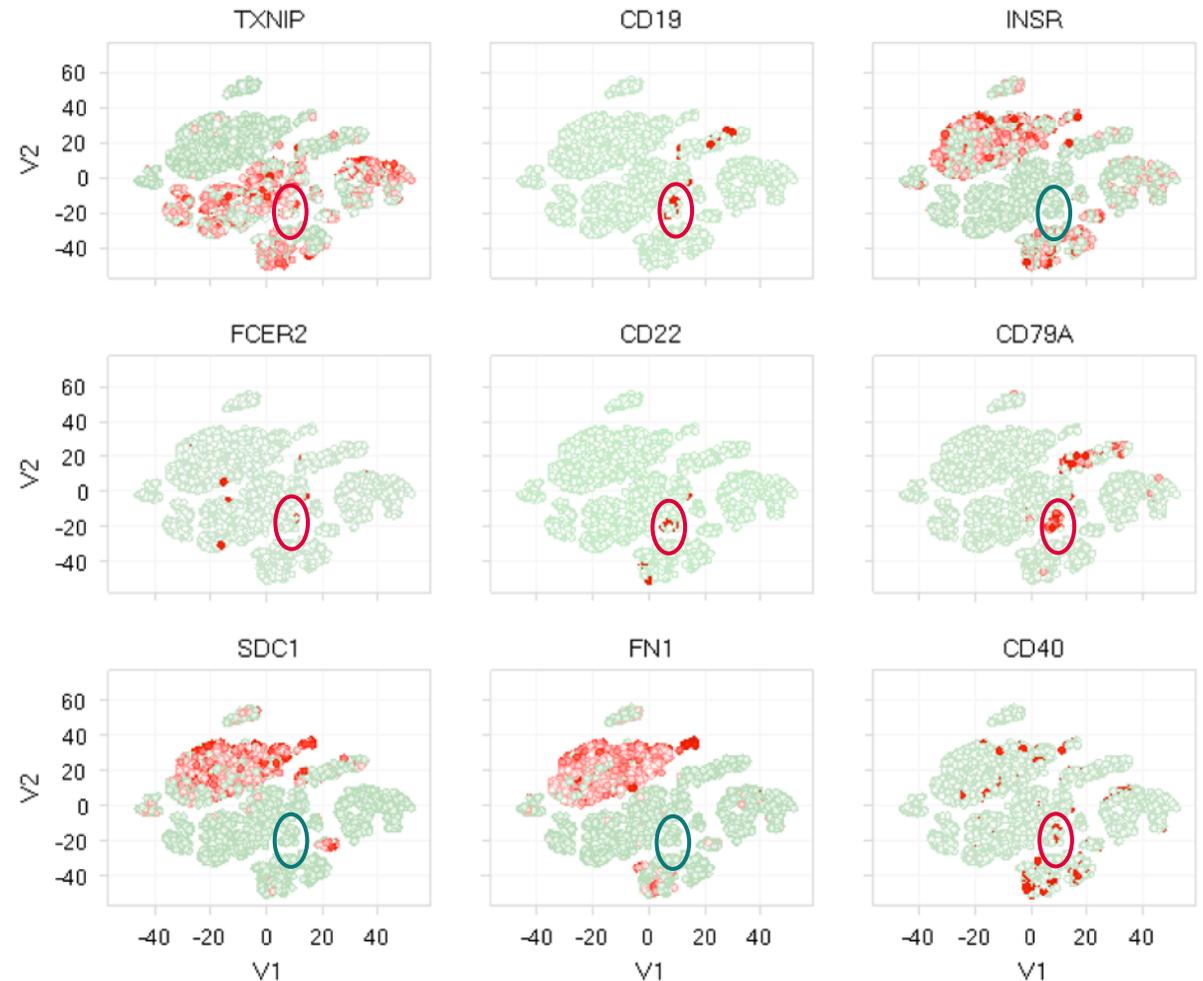

Z-score 3.213, p-value: 6.74 E-07

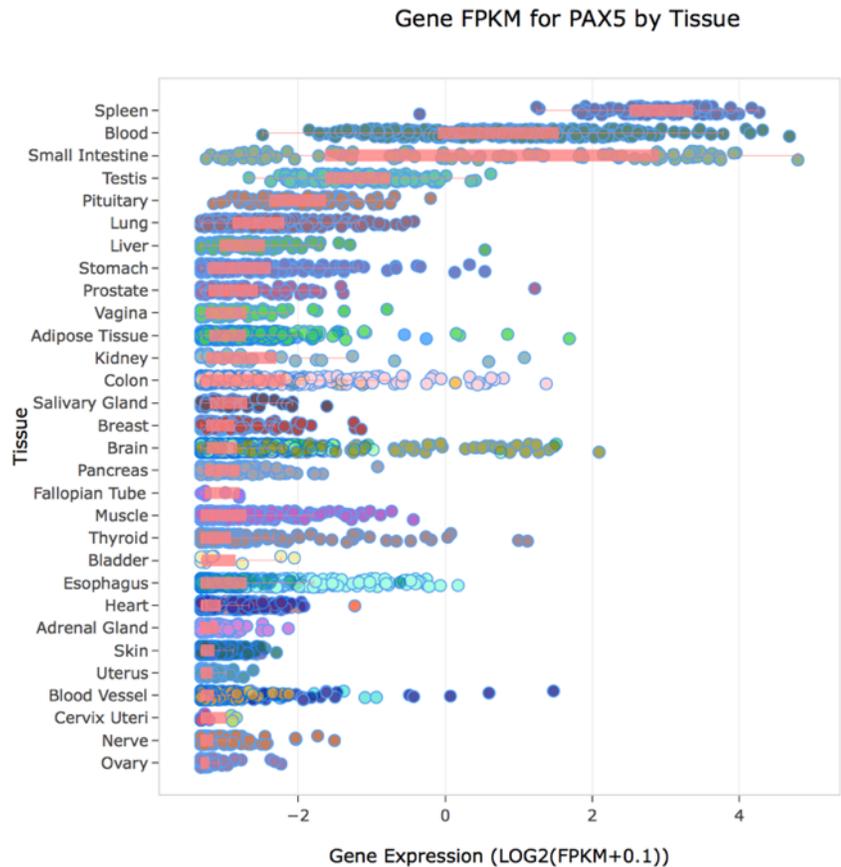# PAX5 targets are expressed in the B cells and others in the liver
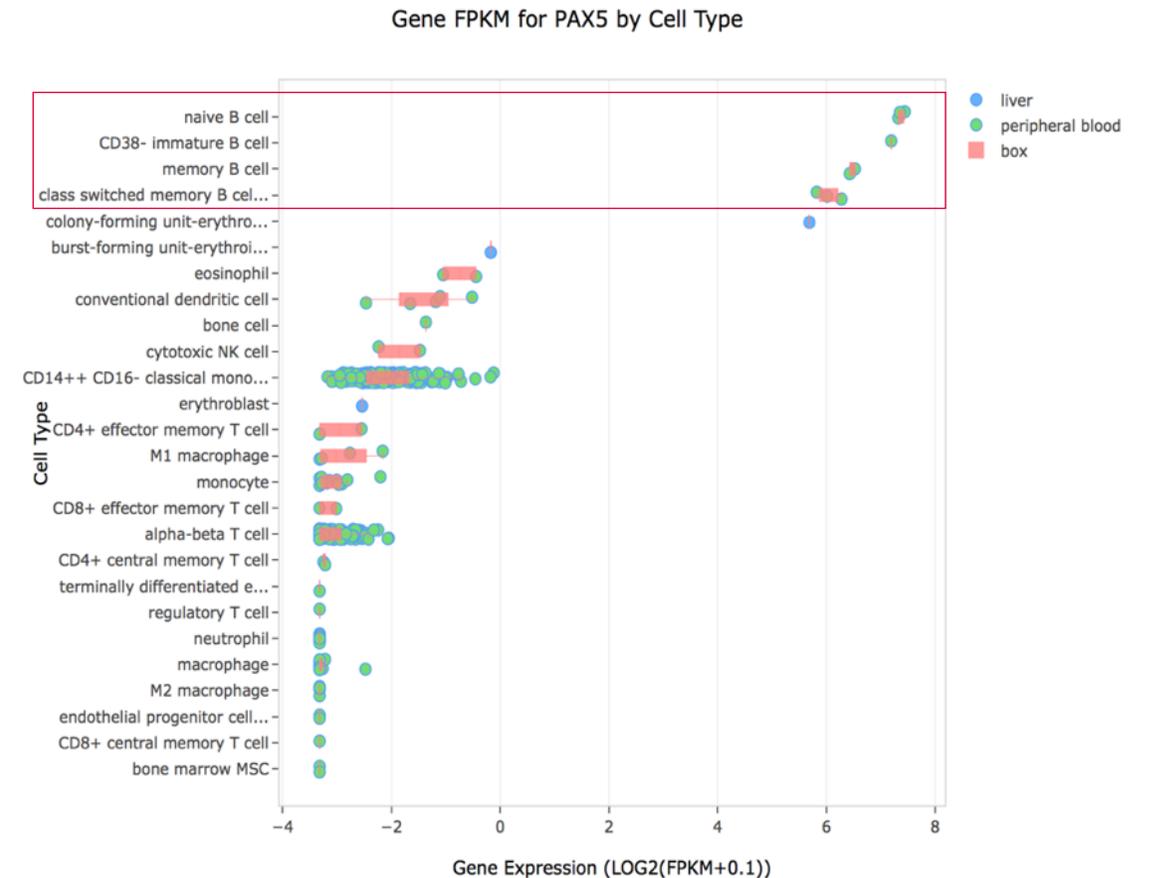

"B cell vs other" PAX5 network

Sample to Insight

PAX5 is highly expressed in the spleen and blood and specifically in B cells

GTEx B38 (51 normal human tissues)

Gene FPKM for PAX5 by Tissue

Blueprint B38 (only normal immune cells)

Gene FPKM for PAX5 by Cell Type
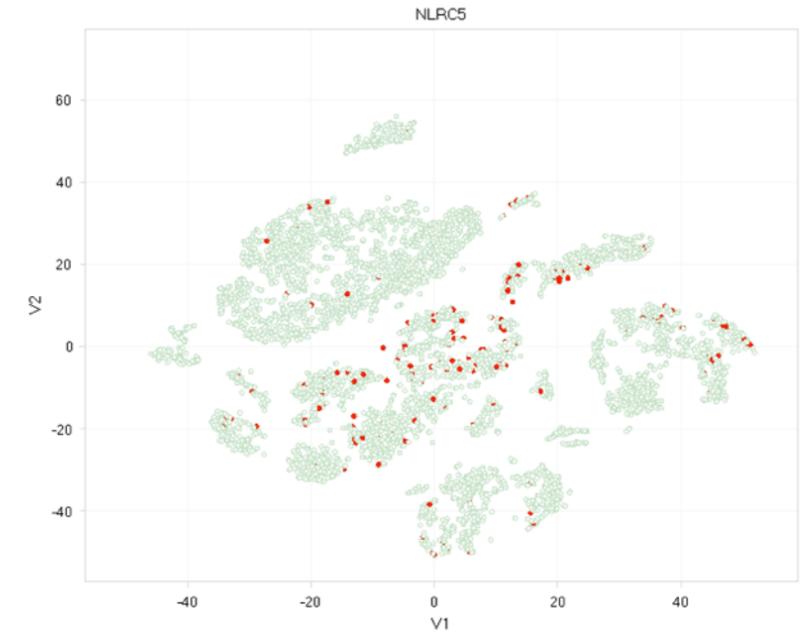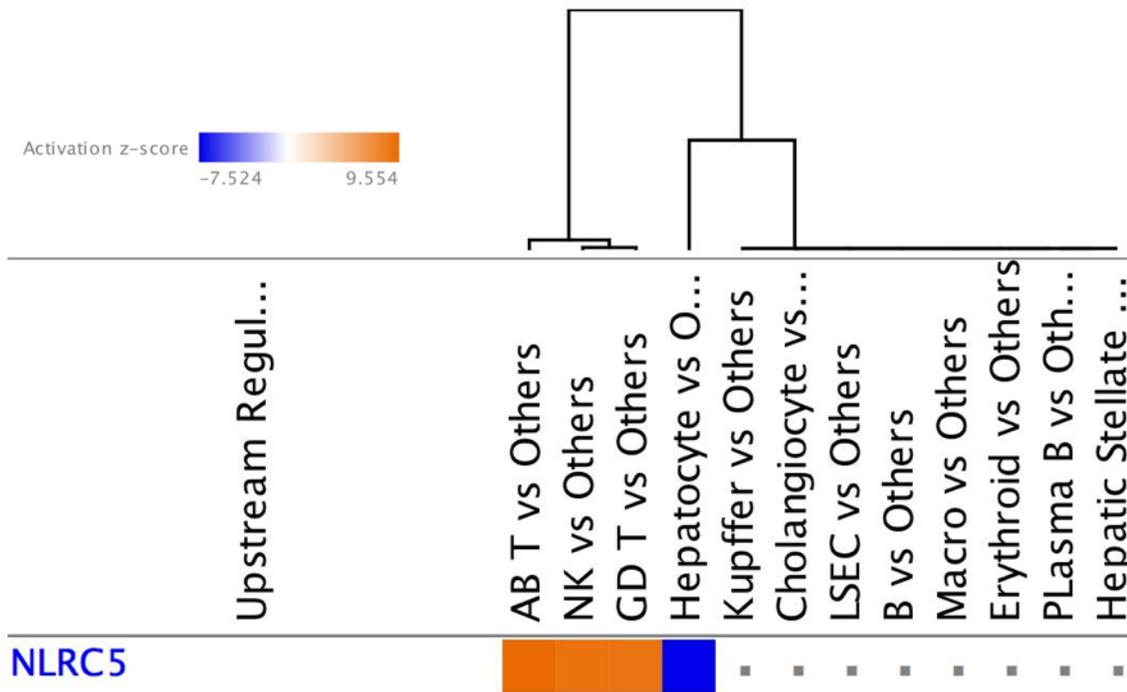
NLRC5 is predicted to be activated in three cell types: $\alpha\beta$ and $\gamma\delta$ T and NK cells

NLRC5 is "NOD-LIKE receptor CARD (caspase recruitment domain) containing 5"

- NLRC5 plays a role in cytokine response and antiviral immunity through its inhibition of NF-kappa-B activation and negative regulation of type I interferon signaling pathways

- NLRC5 is considered a pattern recognition receptor implicated in innate immunity

In what situations is NLRC5 predicted to be activated?

Plot of an entity's activity vs >60,000 analyses
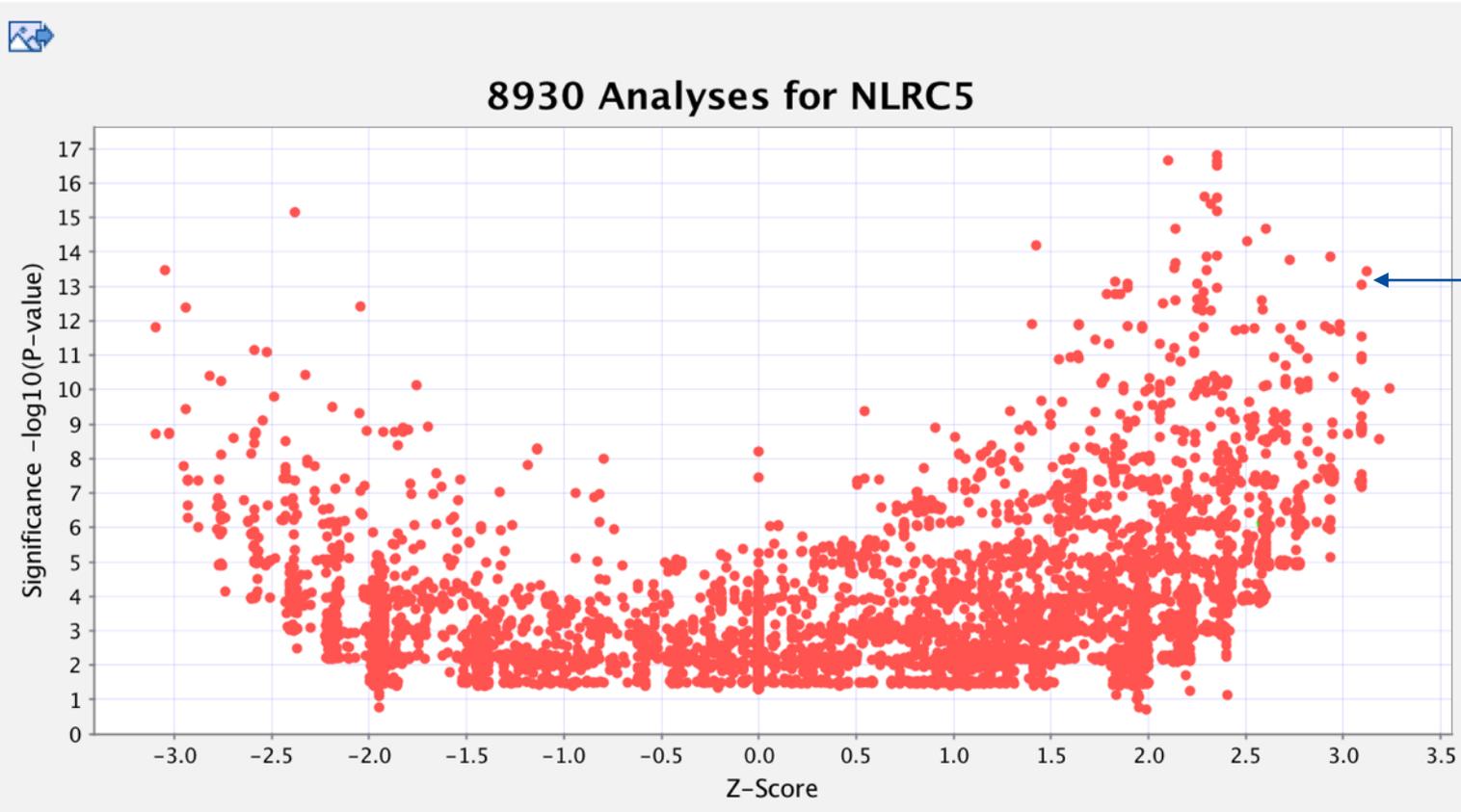
Analyses from the plot with NLRC5 predicted to be activated

With z-score>2

Evaluation of the metadata for analyses where NLRC5 predicted to be activated

NLRC5 is predicted to be activated by treatment with IFN gamma or IFN alpha, response to viral infection, etc.
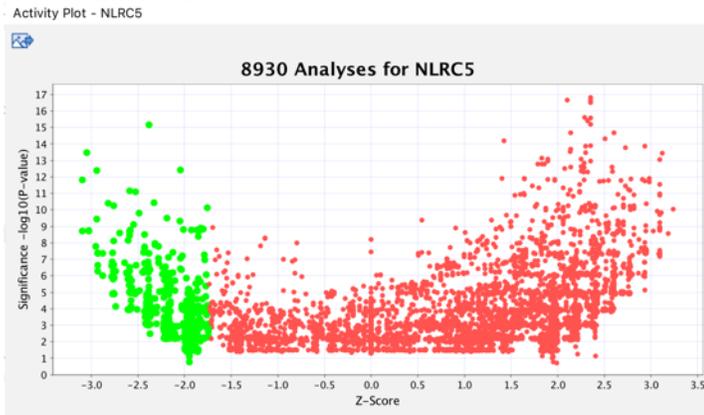
Significant metadata in 200 selected analyses (GD T vs Others)   — ☐ ✕

Customize Table   ⇨ 194 repository analyses selected

Metadata field	Significant term	p-value	Selected analy...	Total analyses ...	Selected a...	Total anal...
case.treatment	IFN alpha	7.15E-16	10	51	194	62813
case.treatment	IFN gamma	1.84E-15	11	81	194	62813
case.treatment	IFN beta 1a	1.48E-11	5	8	194	62813
case.treatment	IFN alpha 2a	2.08E-10	5	12	194	62813
case.subjecttreatment	adenovirus expressing mIFN alpha	6.09E-08	3	6	15	9636
case.pretreatment	Zika virus (ZIKV)	6.09E-08	3	4	8	2804
control.treatment	Infection_none	4.90E-07	6	93	194	62813
case.treatmentgroup	overexpression	1.13E-06	4	37	56	26436
case.treatment	autologous serum;IFN alpha 2b	3.43E-06	3	10	194	62813
case.treatment	Infection_influenza A	4.33E-06	6	135	194	62813
control.treatment	mixed empty vector	5.38E-06	4	37	194	62813
case.treatment	IFN alpha;IFN gamma	9.49E-06	2	2	194	62813
control.treatment	autologous serum	1.03E-05	3	14	194	62813
control.subjecttreatment	radiotherapy	8.09E-05	2	9	15	9636
case.treatment	selumetinib	9.07E-05	5	144	194	62813
case.treatment	Infection_influenza A;rhinovirus	3.37E-04	2	9	194	62813
case.subjecttreatment	radiotherapy	3.41E-04	2	18	15	9636
case.treatment	AZD8330	4.52E-04	4	114	194	62813
case.treatment	IFN alpha 2b	6.14E-04	2	12	194	62813
case.treatmentgroup	MAPK/ERK signaling inhibitor	6.31E-04	15	2839	56	26436

Selected 0 / 87

NLRC5 is predicted to be inhibited in conditions of respiratory infection

Activity Plot - NLRC5

8930 Analyses for NLRC5

Analysis Name	Project	P-value	Z-s...	c...	c...	c...	c...	c...	c...	c...	w...
114- small intestine carcinoid neuroendoc	OncoGEO	1.55E-12	-3.095	small inte...	small inte...	onalespib	Treatmen...	Treatmen...		GSE96760.	https://ww
58- small intestine carcinoid neuroendoc	OncoGEO	1.97E-09	-3.095	small inte...	small inte...	luminespib	Treatmen...	Treatmen...		GSE96760.	https://ww
2- acute respiratory infection (ARI) [peri	HumanDisease	3.31E-14	-3.049	acute res...	periphera...	NA		Treatmen...	Sampling...	GSE68310.(	http://www
6- acute respiratory infection (ARI) [peri	HumanDisease	1.79E-09	-3.025	acute res...	periphera...	NA		Treatmen...	Sampling...	GSE68310.(	http://www
3- colorectal cancer [colonrectum] NA 22	OncoGEO	1.95E-09	-3.025	colorectal...	colonrect...	NA		Other Co...	GeneticS...	GSE76402.	https://ww
3- normal control [embryo] JAK3 inhibitc	HumanDisease	1.66E-08	-2.950	normal co...	embryo	JAK3 inhi...	Treatmen...	Treatmen...		GSE57896.	http://www
12- focal segmental glomerulosclerosis (	HumanDisease	3.52E-10	-2.943	focal seg...	kidney tu...	NA		Disease1 ...	DiseaseS...	GSE99325.	https://ww
9554- prostate adenocarcinoma (PRAD)	LINCS	4.10E-13	-2.943	prostate ...	HER2;EGFR prostate	neratinib	Treatmen...	Treatmen...		GSE70138.(	https://ww
24- colorectal cancer [colonrectum] NA 2	OncoGEO	4.01E-08	-2.932	colorectal...	colonrect...	NA		Other Co...	GeneticS...	GSE76402.	https://ww
13- colorectal cancer [colonrectum] NA 2	OncoGEO	4.23E-08	-2.932	colorectal	colonrect	NA		Other Co	GeneticS	GSE76402.	https://ww

Evaluate Metadata   Customize Table   Z-score -3.095 - -2.192 (p1 of 2)   More Info

Selected 0 / 309

Significant metadata in 200 selected analyses (GD T vs Others)

Customize Table   192 repository analyses selected

Metadata field	Significant term	p-value	Selected anal...	Total analyses...	Selected an...	Total analys...
case.diseasestate	acute respiratory infection (ARI)	8.12E-31	15	29	192	62813
control.diseasestate	acute respiratory infection (ARI)	8.12E-31	15	29	192	62813

● NLRC5 is an important player in immune responses and participate in the liver in the immunosurveillance

IFNG and NLRC5 are connected and participate in decreasing viral infection in the liver

This contributes to the role of $\gamma\delta$ T cells intra-hepatic immune responses.

IFNG

NLRC5

Viral infection
p-value 8.67 E-40

Can we learn more by comparing a single-cell cluster to other analyses (not SC)?

Comparing $\gamma\delta$ T vs. Others with other datasets in TCGA and OncoGEO (Analysis Match)

Filtering to show only liver or bile duct-based analyses and disease vs. normal as comparison

- The γδ cell clusters derived from normal liver match to bulk liver and bile duct cancer, but with some important differences.

NLRC5, IFNG and TBX21 are involved in human liver and bile duct cancers

All three upstream regulators seems to be activated in normal human liver, but only IFNG is predicted to be activated in liver cancer; however TBX21 and NLRC5 are not predicted to be involved in these selected datasets.
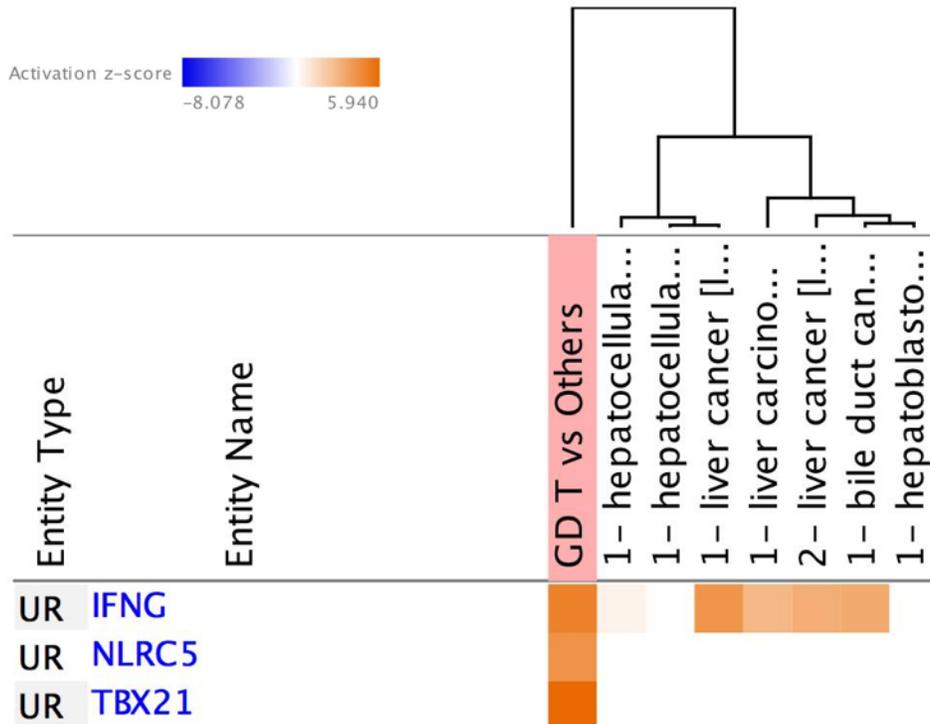
HumanDisease B37
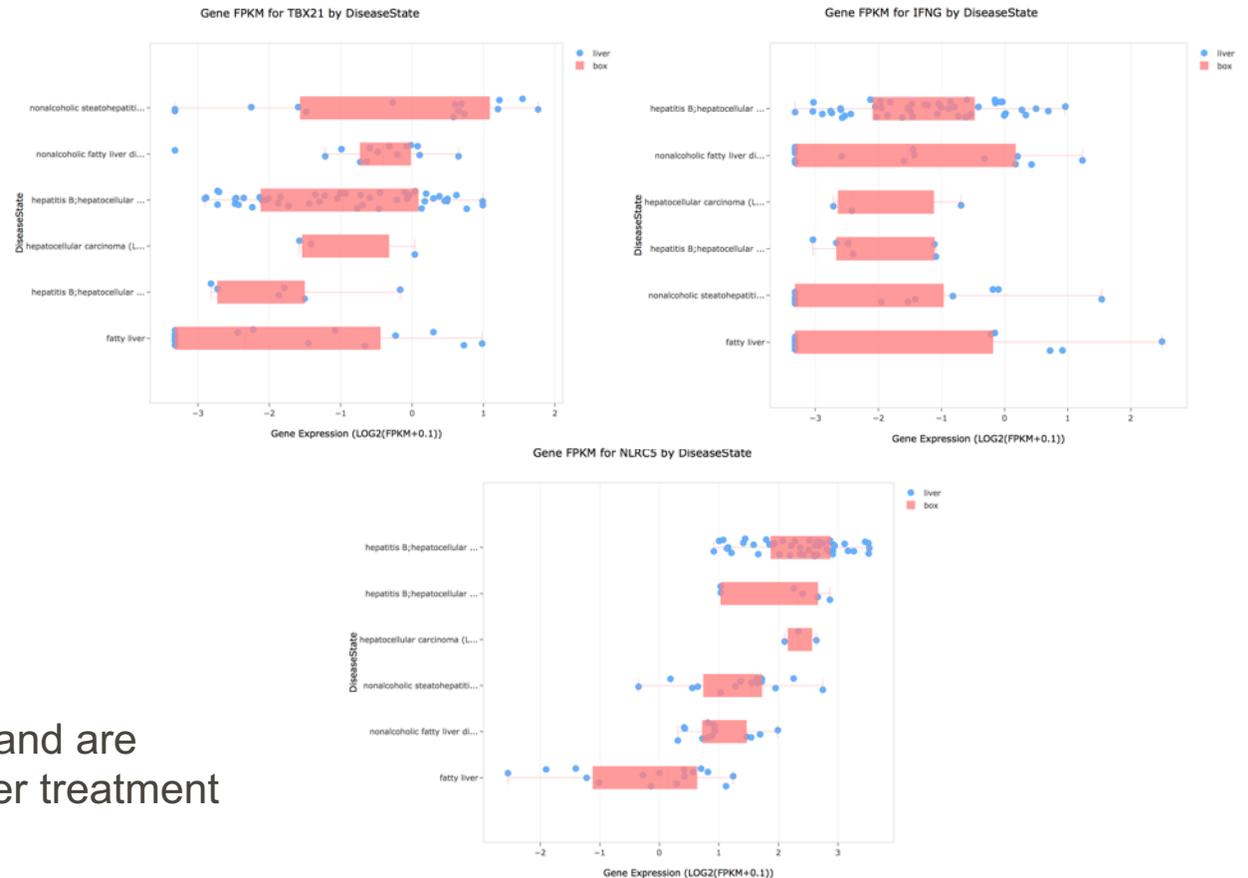
$\gamma\delta$ T cells are known to be involved in anti-tumor responses and are thought to be an important new therapeutic tool in anti-cancer treatment

Conclusion: Single-cell RNA-sequencing analysis of normal human liver

1 — Biological clusters are identified in normal human liver that include parenchymal and non-parenchymal cells

2 — Many upstream regulators and biological processes are involved in a cell specific manner

3 — Identification of the biological clusters could be connected to important expression data in other datasets (cancers and other diseases) to highlight correlation

4 — Detection of hidden biology by understanding the behavior of molecules, pathways and disease & functions across thousands of comparisons
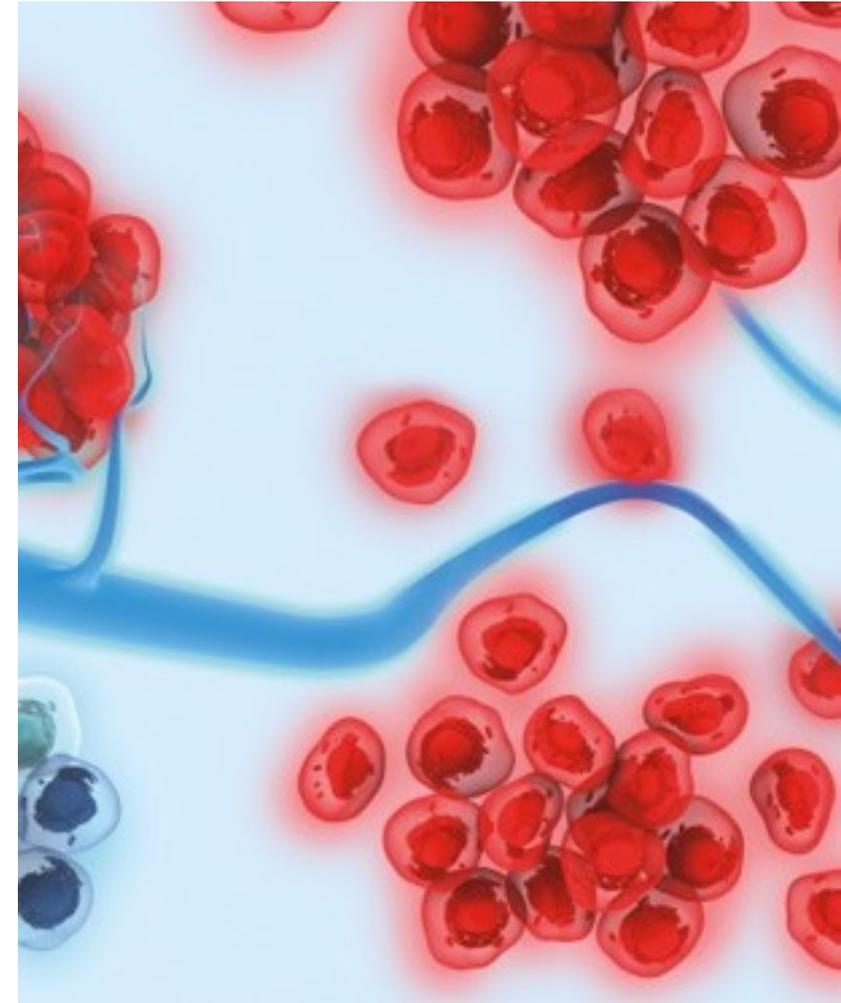
Customer support and additional resources

Contact

Contact us via email or telephone

Reply

A response within ONE business day

Hours

08:00 - 17:00 Pacific
08:00 - 13:00 GMT

Telephone:

- Global: +1 (650) 381-5111
- US toll free: +1 (866) 464-3684
- Denmark toll free: +45 80 82 01 67
- German toll: +49 (0)341 33975301

Email:

ts-bioinformatics@qiagen.com

Websites:

www.qiagenbioinformatics.com

http://tv.qiagenbioinformatics.com

Resources

QIAGEN IPA

- IPA product info: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa

- IPA Analysis Match: https://tv.qiagenbioinformatics.com/video/37242337/exploring-ipas-analysis-match-an

- Land Explorer: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/content-exploration-and-databases/qiagen-omicsoft-land-explorer/

- Coronavirus Network Explorer: https://digitalinsights.qiagen.com/coronavirus-network-explorer/

QIAGEN OmicSoft:

- Product Info: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/qiagen-omicsoft/

QIAGEN CLC Genomics

- Product info: https://digitalinsights.qiagen.com/products-overview/analysis-and-visualization/qiagen-clc-genomics-workbench/

QIAGEN expands integrated coronavirus NGS and software solutions to accelerate COVID-19 research

- QIAseq SARS-CoV-2 Primer Panel converts viral RNA samples into libraries ready for sequencing

- QIAGEN Digital Insights solutions support COVID-19 drug, vaccine and epidemiology research

- For an overview of QIAGEN's coronavirus testing solutions, please visit http://www.qiagen.com/coronavirus.

- To explore QIAGEN's NGS-specific solutions for COVID-19 research, please visit https://go.qiagen.com/CoronavirusNGS

- For details of QIAGEN's SARS-CoV-2 Whole Genome Sequencing Service, please visit https://www.qiagen.com/applications/genomic-services/sars-cov-2-whole-genome-sequencing-services

Trademarks