#50

10.1101/2022.07.12.499681
in bioRxiv

Single-nucleus ATAC-seq elucidates major modules of gene regulation in the development of non-alcoholic fatty liver disease

Fumihiko Takeuchi
fumihiko@takeuchi.name
National Center for Global Health and Medicine, Japan

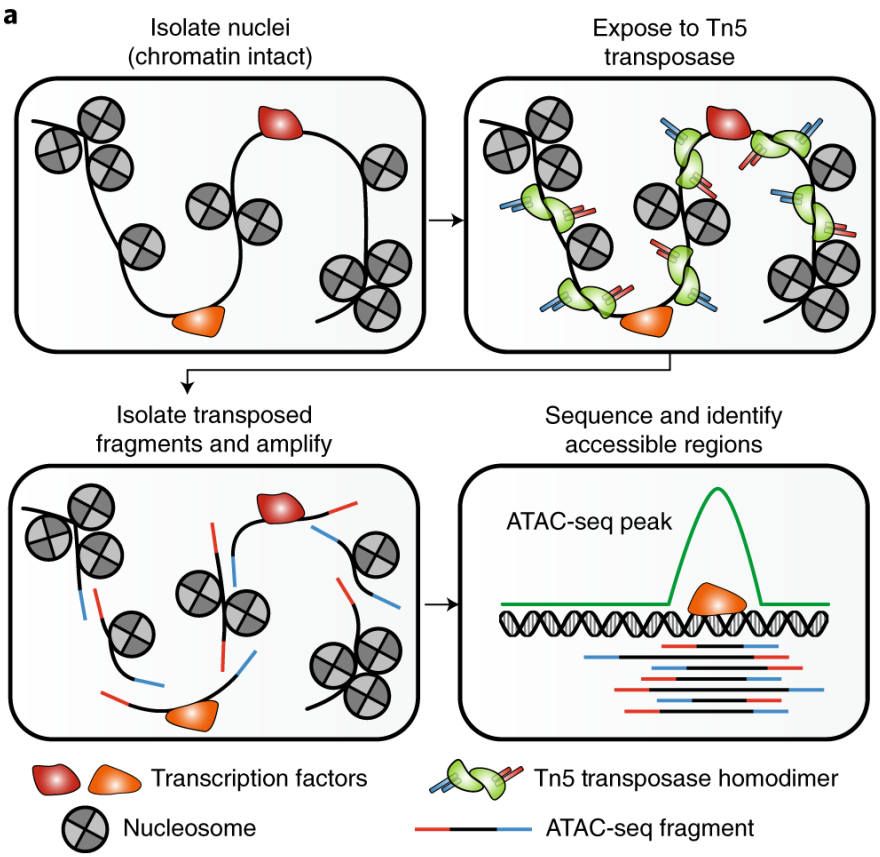
14 Sept 2022　OzSingleCell 2022

Methods. ATAC-seq

ATAC-seq

- Detect open chromatin
- If there is nearby
  - Gene ➔ the gene is expressed
  - Transcription factor binding motif ➔ the transcription factor is binding (indirect evidence)
  - Measure both genome-wide

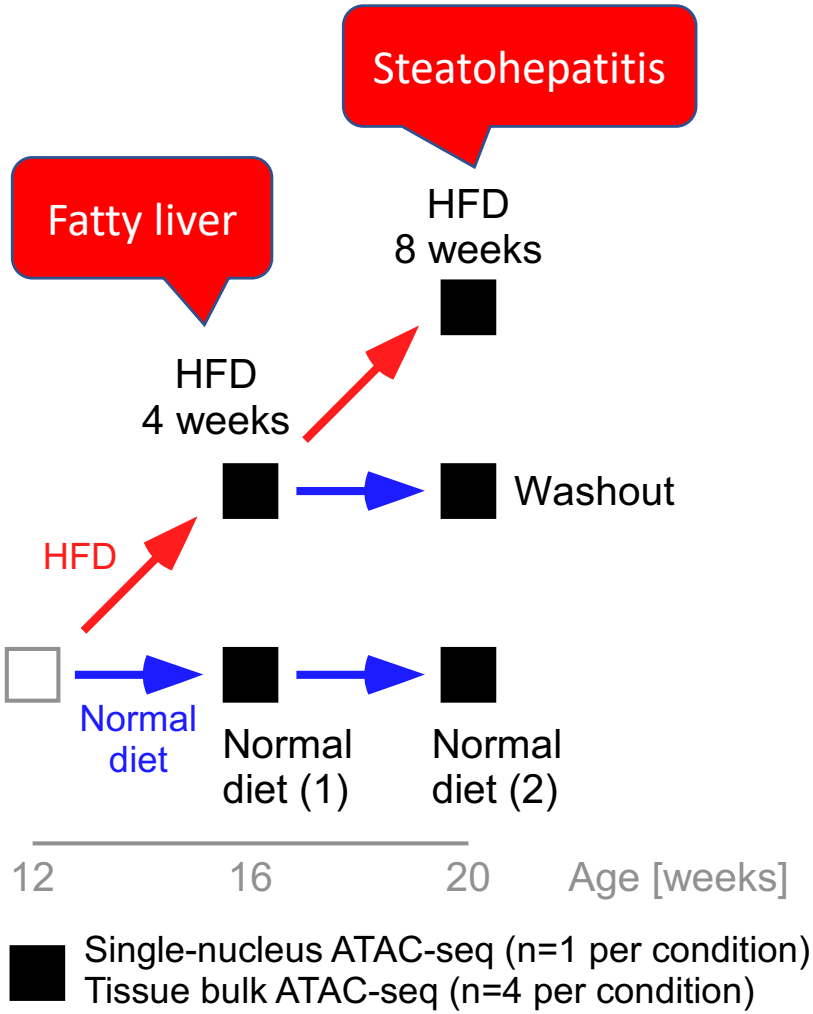
	ChIP-seq	Single-cell ATAC-seq
Transcription factor	1	All
Cell	Bulk aggregate	Distinguish each



10.1038/s41596-022-00692-9

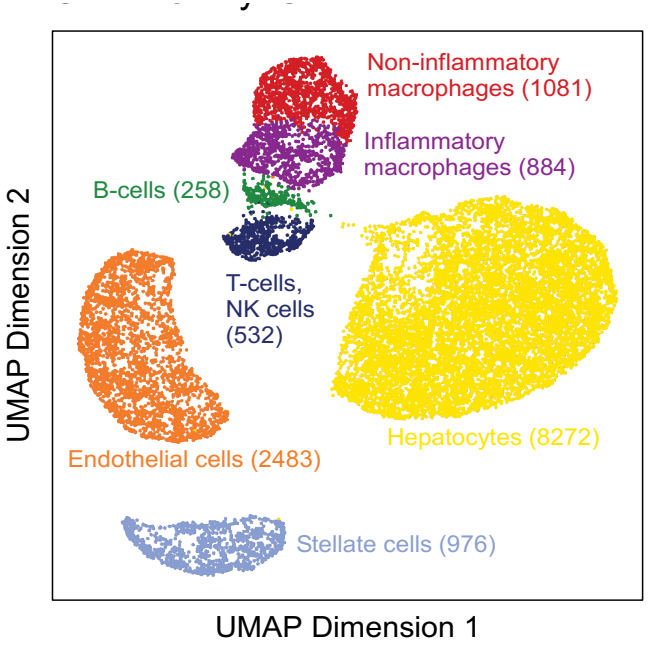
Methods. High-fat diet model for non-alcoholic fatty liver disease

- Male Spontaneously Hypertensive Rats (SHR/Izm)
- High-fat atherogenic diet (HFD)
- ATAC-seq of livers
  - Single-nucleus ATAC-seq (10x Genomics)
  - Bulk ATAC-seq

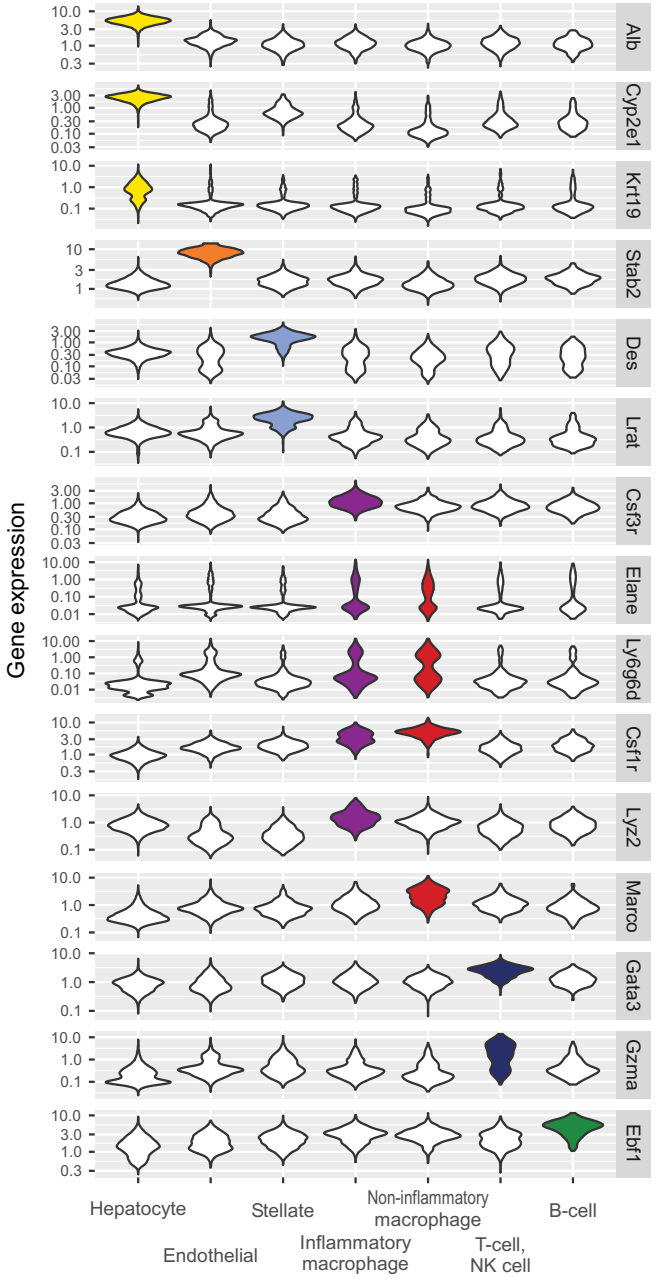


Observed cell types

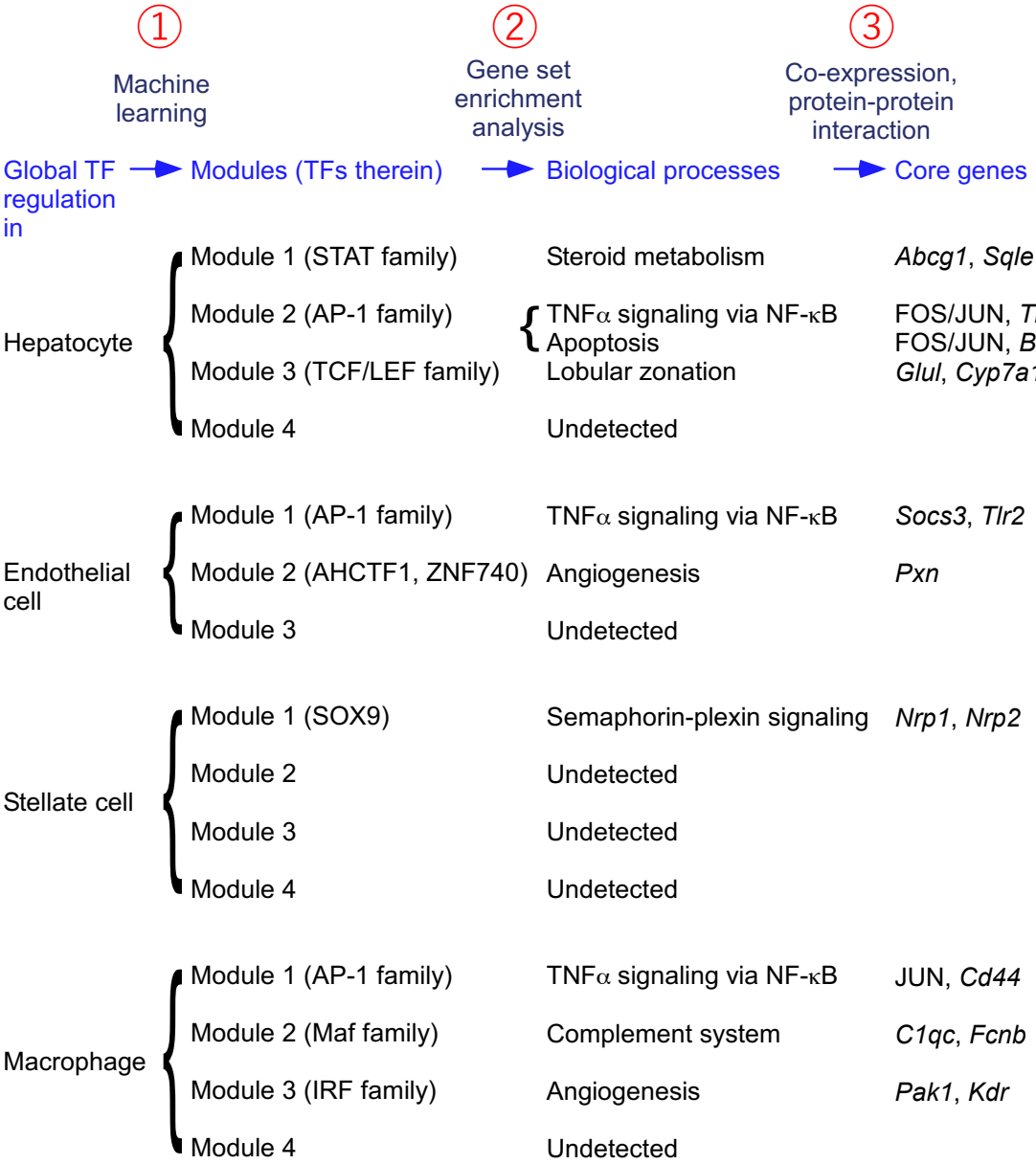
- By similarity of chromatin opening, nuclei were grouped into 16 clusters
- Cell type assigned by marker gene expression
  - Hepatocytes
    - 7 clusters
  - Endothelial cells
    - 3 clusters
  - Stellate cells
    - 2 clusters
  - White blood cells
    - 4 clusters



3000 nuclei per sample

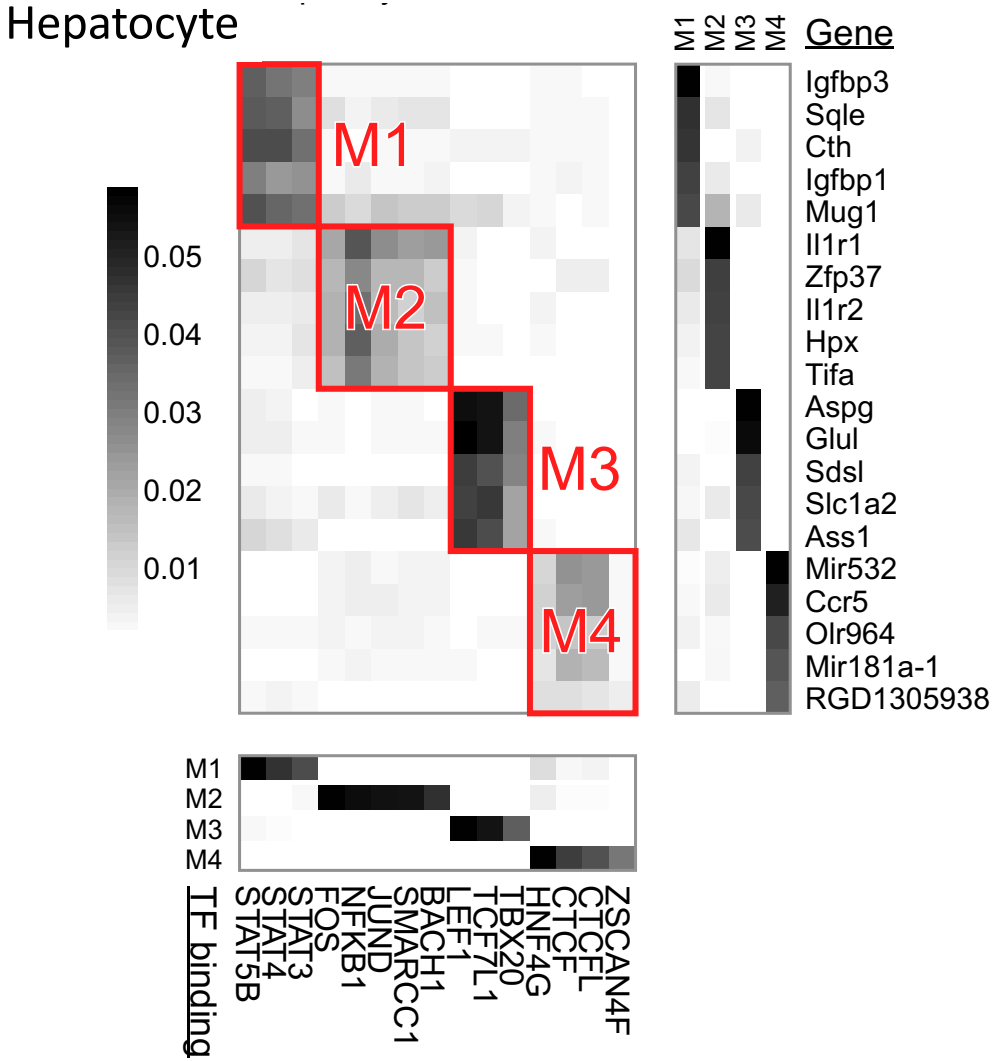


Data-driven discovery of global gene regulation



Conclusion.
Using novel statistical methods, we elucidated a global picture of *in vivo* transcription factor (TF) regulation in each cell type as a set of modules, and discovered core genes.

① Extract major modules of TF regulation



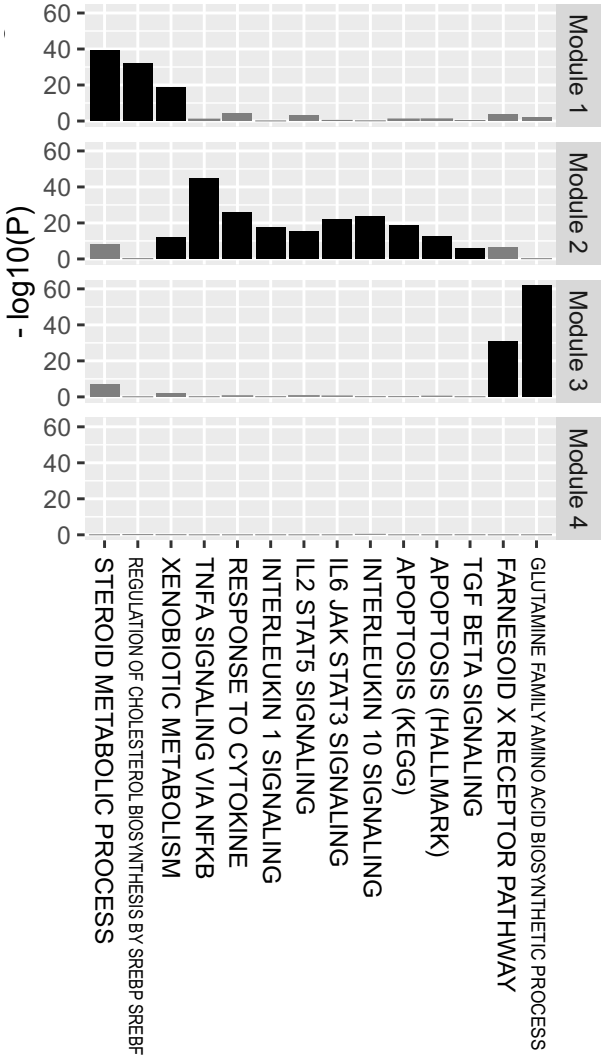
A) Quantify “TF binding” and “gene expression” in single nuclei
- Bindings are totaled genome-wide for each TF

B) Learn which TF regulates which gene

C) Extract modules
- A module is a subset of TFs and genes
- TFs in a module regulate the genes in the same module

② Search biological processes characteristic to a module

Hepatocyte



- Essentially, a module is a list of TFs and genes
- From the list of genes, find characteristic biological processes
  - Search Gene Ontology and pathway databases, using Gene Set Enrichment analysis

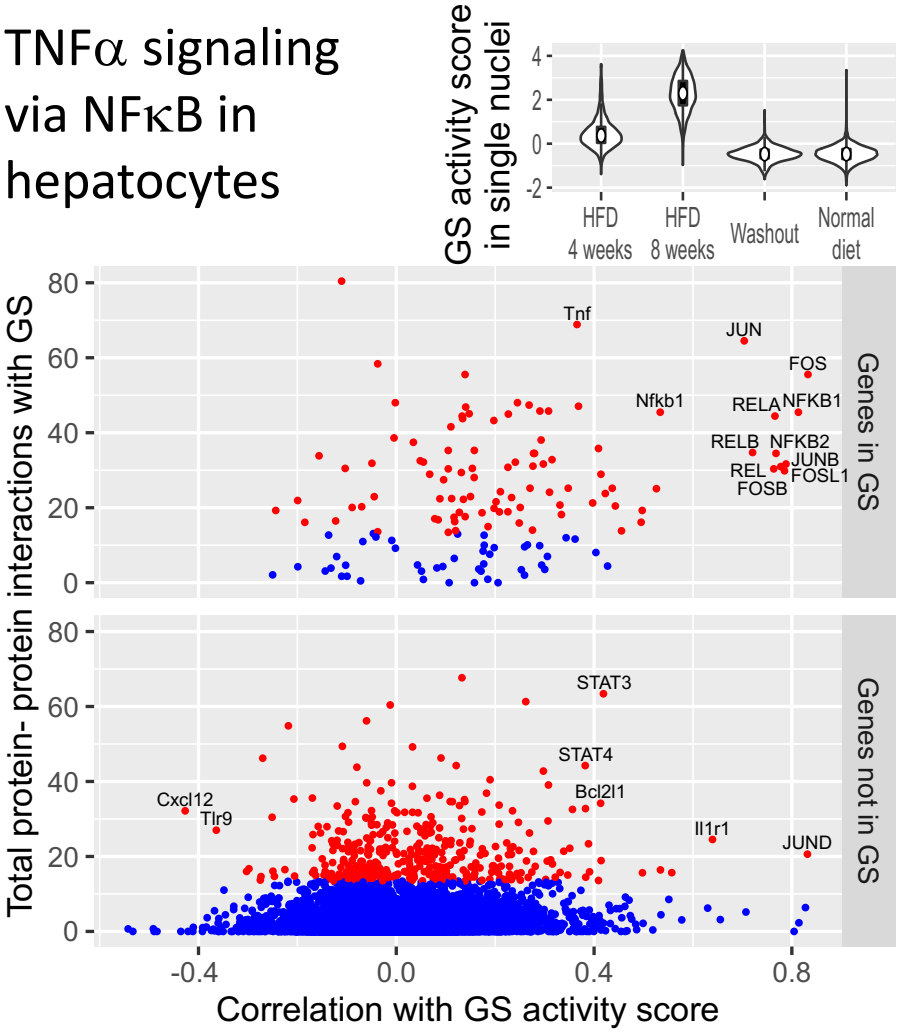
TF modules found in this study

Cell type	TF	Biological process	Previous reports
Hepatocyte	STAT family	steroid metabolism	
Hepatocyte, endothelial, macrophage	AP-1 family	TNFα signaling via NF-κB	AP-1 TFs respond to cytokine stimuli (Hess et al., 2004)
Hepatocyte	TCF/LEF family	zonation in liver lobule	LEF1 TF binds to β-catenin protein and activates Wnt signaling pathway (Sun and Weis, 2011)
Endothelial	AHCTF1, ZNF740	angiogenesis	ZNF740 activates angiogenesis in pulmonary artery endothelial cells of rats (Yu et al., 2018)
Stellate	SOX9	semaphorin-plexin signaling	
Macrophage	Maf family	complement system	In *Mafb*-deficient macrophages of mice, C1q production decreased (Tran et al., 2017)
Macrophage	IRF family	angiogenesis	IRF1 contributes to the commitment of pro-inflammatory M1 macrophages, which produce angiogenic stimulators (Chistiakov et al., 2018).

- The linkage between TF and biological process has been reported in 5 out of 7
- Good indication!

③ Discover core genes of a biological process

TNFα signaling via NFκB in hepatocytes



- Focus on a biological process in a cell type
  - Here, a biological process is defined as a gene set (GS)
- Core genes of a GS
  - Central in co-expression ※
  - Central in protein-protein interaction (STRING database)
- "GS activity" of each nucleus
  - The "average" of TF binding and gene expression for genes in GS
  - ※ Strong positive or negative correlation with GS activity
- Overlapped well with genes know for the disease