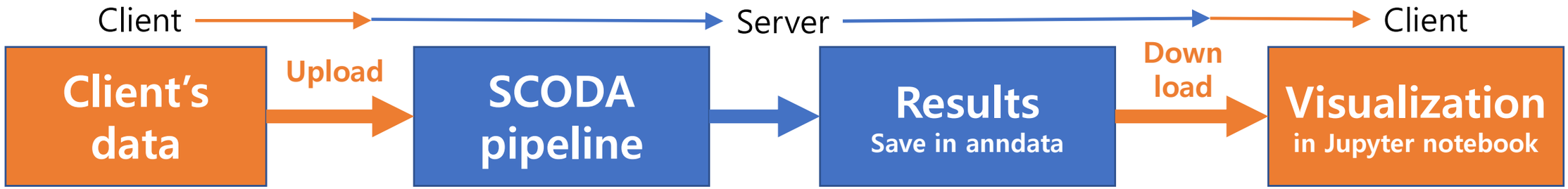
What is SCODA?

- Automated one-stop computing service for **S**ingle-**C**ell (transcript)**O**mics **D**ata **A**nalysis (single-cell RNA-seq)
- Especially useful for
  - In-vivo Tissue/tumor micro environment (TME) study.
  - Immune cell profiling in, for example, autoimmune disease & cancer
- Simply uploading your count data, you get standardized functional analysis results in ten-minutes or so (depending on your data size)
- Results include
  - High precision cell-type annotation (major-type, minor-type and subset)
  - Condition specific DEG and GSA results per-cell-type
  - Possible cell-cell interactions between different cell-types per-condition
  - CNV estimates and tumor-cell identification results

SCODA pipeline overview
(Single-Cell Omics Data Analysis pipeline)



Supported format

1. Compressed 10x mtx files sets + (optional) meta data (csv file)
2. Compressed CSV (cell-by-gene matrix) + (optional) meta data (csv file)
3. Compressed h5ad file

Functions

1. Celltype annotation using HiCAT (v0.6.7)
2. DEG & GSA/GSEA per minor celltype
3. Inference of cell-cell interaction using CellPhoneDB (v2.1.7)
4. (optional) Tumor cell identification using InferCNV (v0.4.2) + icnv addon
5. More will be added soon

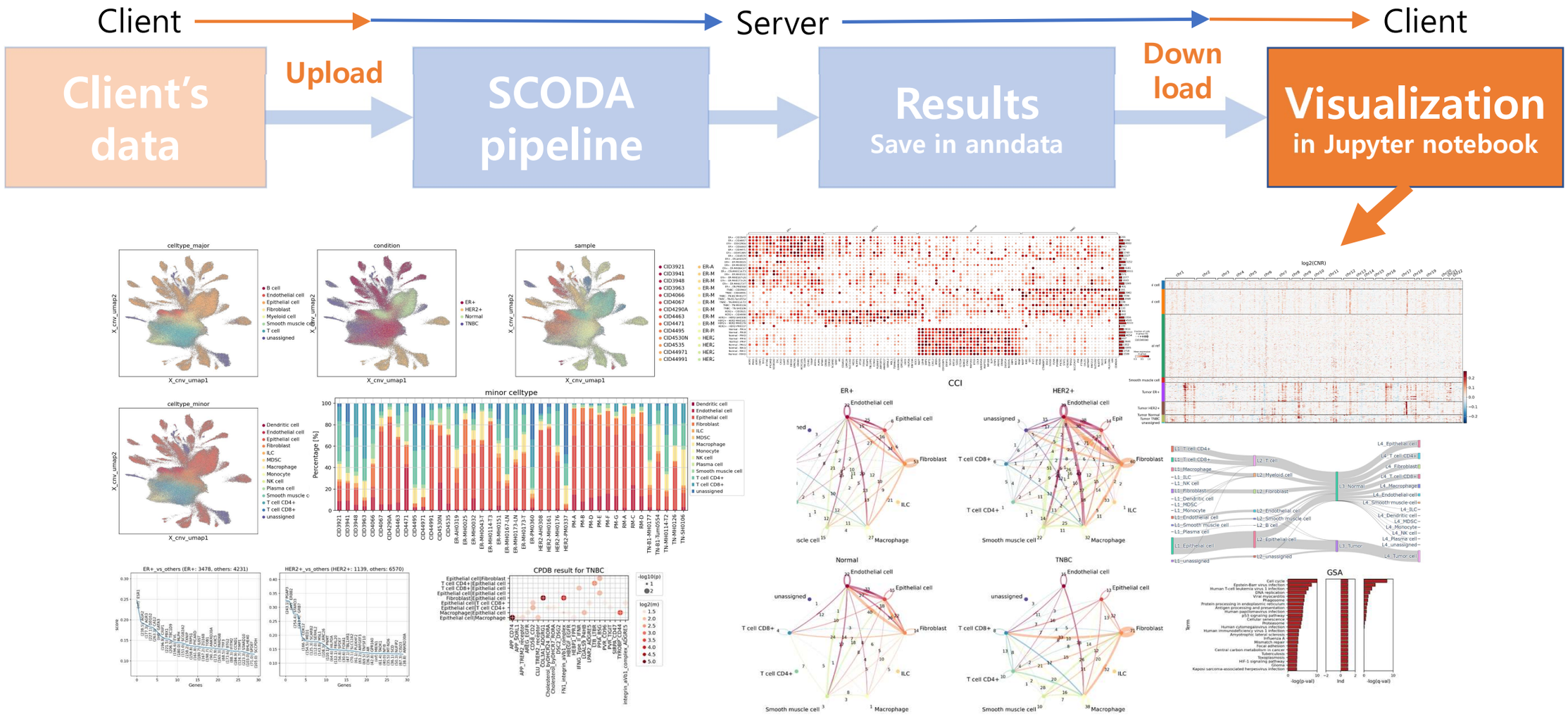
Results saved in anndata format (v0.8.0)

1. Celltype annotation in obs field
2. DEG & GSA/GSEA results for each celltype in uns field
3. CellPhoneDB results in uns field
4. (optional) Tumor cell identification results in obs field

Exploring the results

1. Use SCANpy & SCODA-viz package (open source)
2. Example Jupyter notebook provided
3. With a little bit of programming, users can get deeper insight into the data by comparing cells in different condition

SCODA-viz: visualization lib. (free open source)



Why SCODA ?

Client → 유전체 분석 서비스 → Client

Tissue Sample → Library prep. → Read data (~100GB) → 데이터 분석 ㅜㅜ



Sequence data pre-processing → Downstream (functional) analysis → Biologist Investigation (Data mining)

Client → (단일세포) 유전체 분석 서비스 → **SCODA** → Client **can focus on their original jobs**

Tissue Sample → Library prep. → Read data (~100GB) → 데이터 분석 ^^



Sequence data pre-processing → Downstream (functional) analysis → Biologist Investigation (Data mining)

We offer customized functional analysis service

How to use SCODA?

- With your raw datasets (e.g., CellRanger v3 count data)

1. Go to SCODA home https://mlbi-lab.net to upload and get results.
2. Get Jupyter notebook to explore the results: https://github.com/combio-dku/scoda_explorer
3. Mine the SCODA results in Jupyter notebook on your own computer (or in Google Colab)
4. Contact us if you have larger datasets

- Contact: mlbi.dku@gmail.com for any inquiries or online workshop for using SCODA-viz

Why SCODA?

- Trend in bio/medical research: convergence of data science & bio/medicine.
- Requires complicated communications between biologist and informatician.
- SCODA is an automated one-stop computing service using verified, standardized functional analysis tools for single-cell omics data.
- With SCODA, you don't need to spend your time to do something that are not your field of expertise, e.g.,
  - for complicated computing environment setup
  - develop pipelines to preprocess your datasets
- Can accelerate the research with single-cell RNA-seq data.
- The only thing you need is a little-bit of programming skill, e.g., handling data frame, for your own data mining.
- It provides you with a much higher degree of freedom for your research.

What is SCODA?

- The computing server accepts 10x mtx format (CellRanger v3 output), csv or anndata formatted h5ad file (optionally with meta data)
- SCODA pipeline utilizes verified, suitably licensed (or free) open source SWs (HiCAT v0.6.12, CellPhoneDB v2.1.7, InferCNV v0.4.2, anndata v0.8.0)
- Free visualization library is provided for your own data mining (along with scanpy, searborn, matplotlib packages)
- You only need a little bit of programming skill (in python or R) to explore the SCODA results by yourself to mine biological implications.
- Small group workshop (training) for data mining is available.

- **With SCODA, you can focus more on your original jobs!!! (i.e., biology and/or medical research)**

Try SCODA now

- Are you ready to unravel the mysteries of single-cell gene expression? And gain deeper insights into cellular heterogeneity?
- Look no further!
- Our single-cell RNA-seq data analysis pipeline was designed to empower your research with state-of-the-art computational capabilities.
- The visualization and data mining is fully customizable with a little bit of programming skill.
- You can tailor the analysis pipeline to your specific needs.
- With optional analysis parameter settings, you have control to adapt workflows for your specific requirement.
- Don’t hesitate to contact us (mlbi.dku@gmail.com). We’re here to help you.