## Tutorial for Build database

The following is a step-by-step tutorial on building a database, covering the execution logic and key operations of four core scripts

In this case, we used a subset of three types from MISAR to run the analysis. If you want to process all types, please load the file **cell\_type\_gt.csv**.

Step1: Detecting the interaction between peak regions (step1\_detect\_TFs\_for\_peaks.R)

Step2: Fine grained detection based on ArchR (step2\_detect\_ITs\_for\_peaks\_upon\_ArchR.R)

Now, we get TF-peak links

## Step3: Generate TF-TG relationship pairs (step3\_generate\_tf\_tg.py)

python step3 generate tf tg.py Loading data... Data loading done. Extracting TF lists... TF lists extraction done. Building TF peak dictionary... TF peak dictionary built. Saving TF peak union... TF peak union saved. Extracting marker regions... Marker regions extracted and saved. Building TG peak dictionary... 523/523 [00:29<00:00, 17.52it/s] TG peak dictiona ry built. Building TF-TG dictionary... 112/112 [00:04<00:00, 22.67it/s] TF-TG dictionary built and saved.

## Step4: Build the final database (step4\_build\_database.py)

python step4\_build\_database.py

```
Loading data...
Data loading done.
Filtering database...
100%
13217/13217 [00:00<00:00, 80658, 75it/s]
Database filtering done.
Extracting gene index...
Gene index extraction done.
Building inter LR for each TG...
100%
                129/129 [00:00<00:00, 2091.80it/s]
Inter LR building done.
Preprocessing REs befor intra...
809652it [05:58, 2257.96it/s]
100%
790450/790450 [00:36<00:00, 21365.29it/s]
REs preprocessing done.
Building intra TFRE for each TG...
129/129 [00:00<00:00, 3557.53it/s]
Intra TFRE building done.
Filtering ligand interactions...
100%
______
_______
129/129 [00:21<00:00, 6.02it/s]
/home/nas2/biod/zhencaiwei/RegChat-main/Code/Build database/step4 build database.py:227: FutureWarn
ing: DataFrame.applymap has been deprecated. Use DataFrame.map instead.
mask = zscore mat.applymap(lambda x: abs(x) >= zscore threshold)
Filtering receptor interactions...
129/129 [00:21<00:00,
Filtering LR pairs...
100%
129/129 [00:17<00:00, 7.18it/s]
Filtering TF interactions...
100%
```

	129/129 [00:05<00:00, 23.49it/s]
Filtering RE interactions	
100%	
	129/129 [00:05<00:00, 23.72it/s]
Filtering TFRE pairs	
100%	
	129/129 [00:13<00:00, 9.39it/s]
Building final database	
Database construction completed.	
Final database shape: (66769, 4)	

 $Final\ database\ saved\ to:\ /home/nas2/biod/zhencaiwei/RegChat-main/RegChat\_Real\_Datasets/MISAR/outpu$ t/database/intersection\_all\_DB.csv