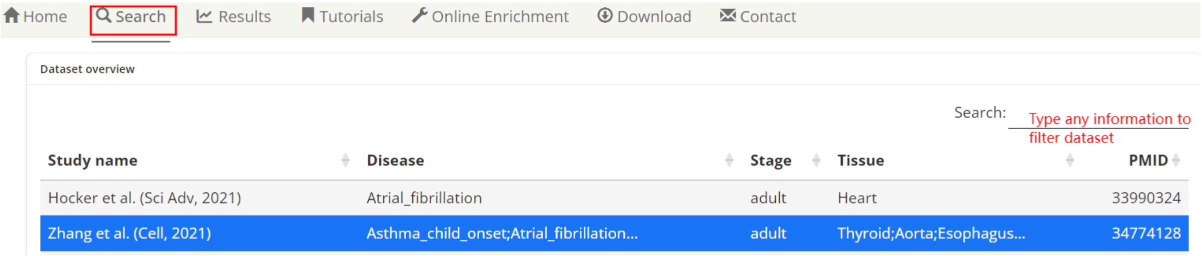
DRCTDB tutorials

Introduction

DRCTdb contains well-processed single-cell multiomics data from 16 studies, encompassing transcriptome and epigenetic information from over 4 million cells across 28 tissues. Through DRCTdb, users can easily browse relationships and regulatory mechanisms between SNPs of 42 genetic diseases and cell types in different human tissues based on GWAS and single-cell multiomics data. Moreover, DRCTdb also provides data downloads and online tools, allowing users to download well-processed single-cell multiomics data and analysis results from DRCTdb.

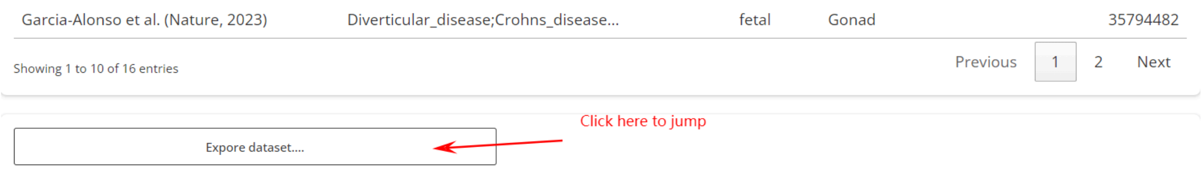
Select dataset

The first step is to select a dataset. Users can click the "Search" button in the navigation bar to go to the dataset selection page. Then, click on the corresponding dataset. Example below:



Study name	Disease	Stage	Tissue	PMID
Hocker et al. (Sci Adv, 2021)	Atrial_fibrillation	adult	Heart	33990324
Zhang et al. (Cell, 2021)	Asthma_child_onset;Atrial_fibrillation...	adult	Thyroid;Aorta;Esophagus...	34774128

Finally click the "Explore dataset" button at the bottom to view the data within each dataset.



Garcia-Alonso et al. (Nature, 2023)	Diverticular_disease;Crohns_disease...	fetal	Gonad	35794482

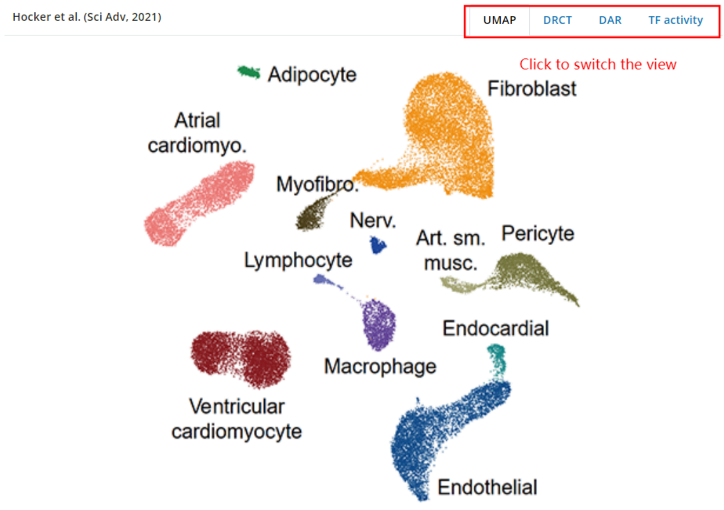
Showing 1 to 10 of 16 entries

Explore dataset

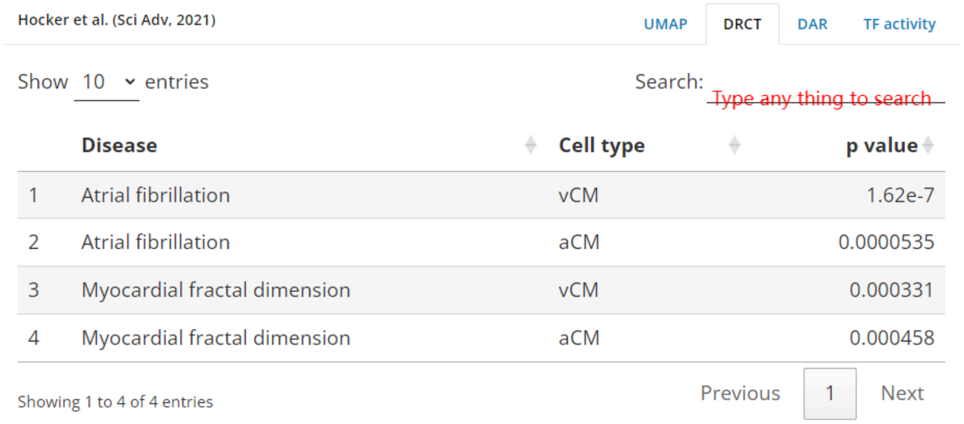
Left panel

Then we arrive at the main page of each dataset. The left section has four buttons:

1. **UMAP:** Displays the UMAP embedding of the dataset.
2. **DRCT:** Shows the Disease Related Cell Type.
3. **DAR:** Displays the Differentially Accessible Region.
4. **TF Activities:** Shows the top 10 transcription factors with the highest activities in each cell type.



Except for UMAP, which is an image, the other three buttons display tables. Similarly, you can search and filter these tables using the search box in the top right corner. An example is shown in the image below.



	Disease	Cell type	p value
1	Atrial fibrillation	vCM	1.62e-7
2	Atrial fibrillation	aCM	0.0000535
3	Myocardial fractal dimension	vCM	0.000331
4	Myocardial fractal dimension	aCM	0.000458

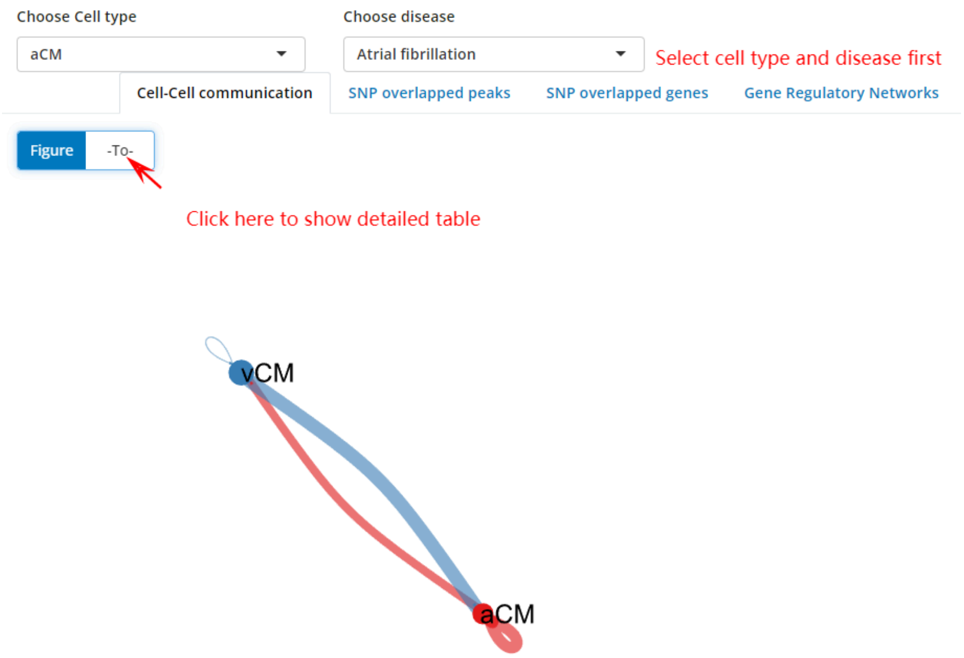
Showing 1 to 4 of 4 entries

Right panel

Then, on the right side of the page, there are four modules:

1. **Cell-Cell Communication**:
2. **SNP Overlapped Peaks**
3. **SNP Overlapped Genes**
4. **Gene Regulatory Networks**

Except for the Cell-Cell Communication module, the content of the other three modules is cell type specific, meaning the content will change based on the selected cell type and disease.



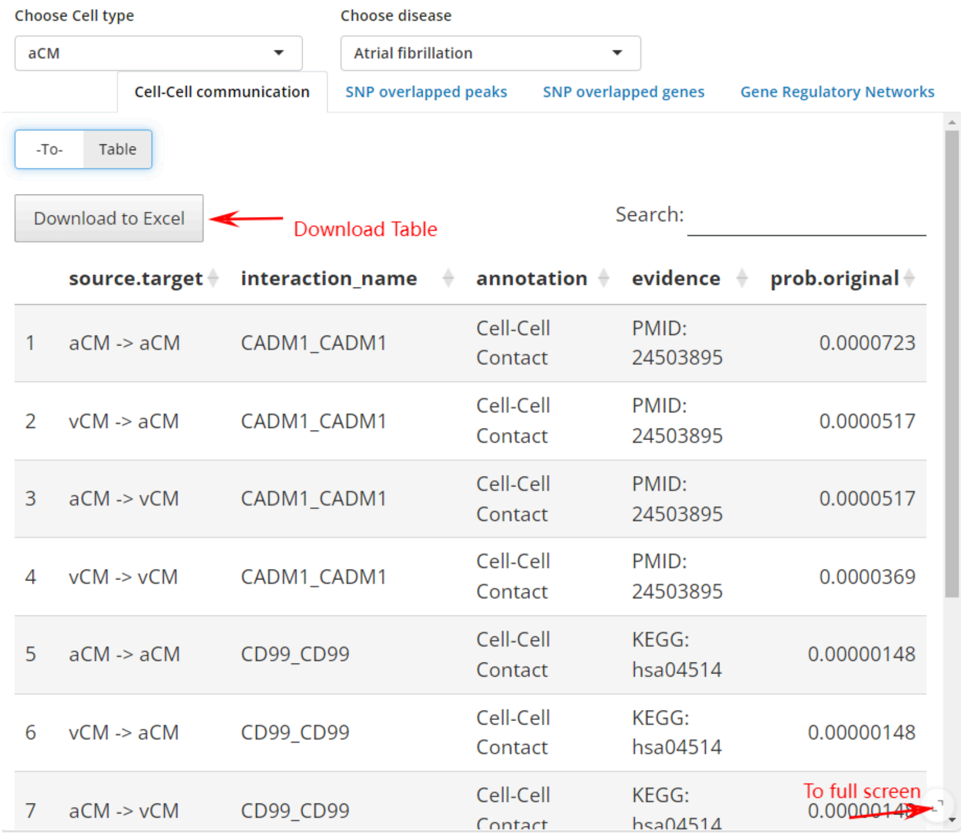
Choose Cell type: aCM

Choose disease: Atrial fibrillation

Select cell type and disease first

Click here to show detailed table

Except for the Gene Regulatory Networks module, users can toggle between image and table views by clicking the "To" button in the top left corner. All tables in this section can be downloaded to an Excel file by clicking the "Download to Excel" button. Additionally, there is a button in the bottom right corner to switch to full-screen mode for displaying the tables.



Choose Cell type: aCM

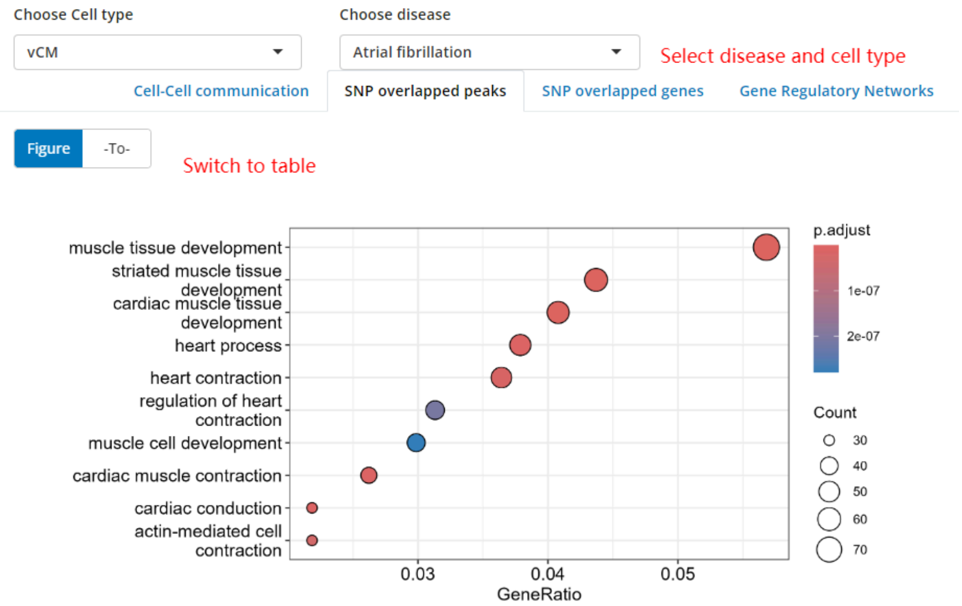
Choose disease: Atrial fibrillation

Download Table

	source.target	interaction_name	annotation	evidence	prob.original
1	aCM -> aCM	CADM1_CADM1	Cell-Cell Contact	PMID: 24503895	0.0000723
2	vCM -> aCM	CADM1_CADM1	Cell-Cell Contact	PMID: 24503895	0.0000517
3	aCM -> vCM	CADM1_CADM1	Cell-Cell Contact	PMID: 24503895	0.0000517
4	vCM -> vCM	CADM1_CADM1	Cell-Cell Contact	PMID: 24503895	0.0000369
5	aCM -> aCM	CD99_CD99	Cell-Cell Contact	KEGG: hsa04514	0.00000148
6	vCM -> aCM	CD99_CD99	Cell-Cell Contact	KEGG: hsa04514	0.00000148
7	aCM -> vCM	CD99_CD99	Cell-Cell Contact	KEGG: hsa04514	0.00000148

To full screen

We also provide enrichment analysis results and graphs for SNP-enriched genes. Similarly, users can click buttons to toggle between views and download the data.

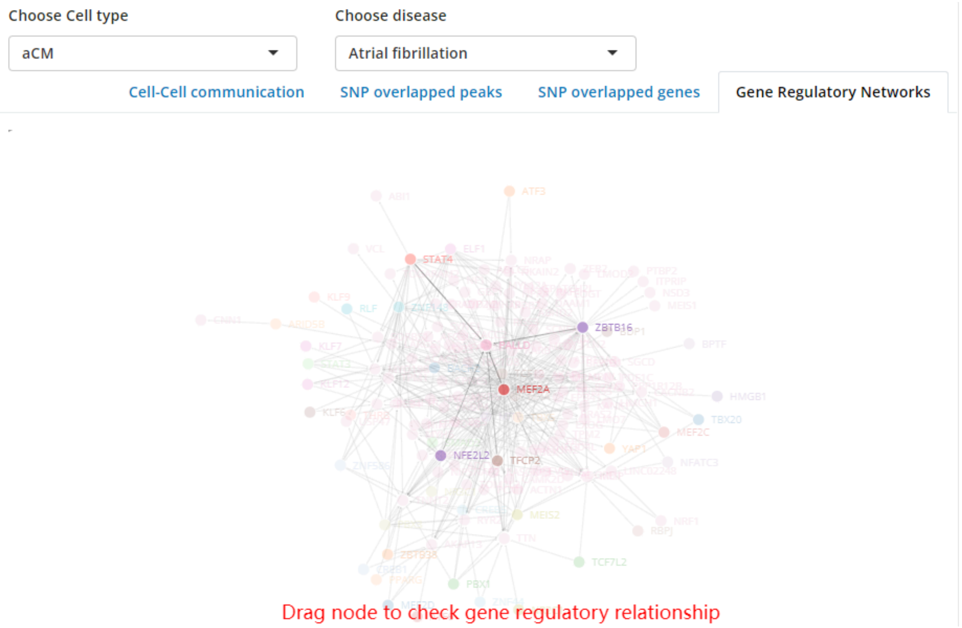

Choose Cell type: vCM

Choose disease: Atrial fibrillation

Select disease and cell type

Switch to table

Finally, in the Gene Regulatory Networks module, users can view the regulatory relationships for specific diseases and cell types. Users can drag nodes to closely examine the relationships between specific genes.

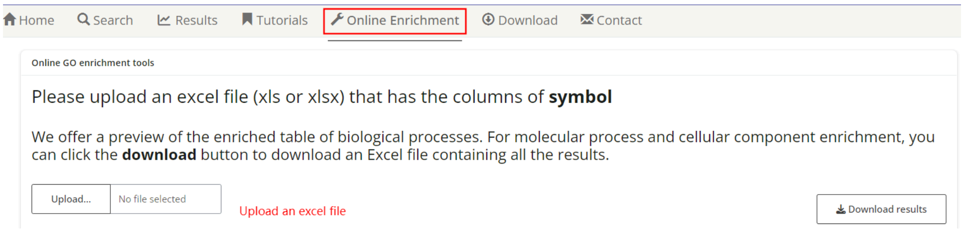

Choose Cell type: aCM

Choose disease: Atrial fibrillation

Drag node to check gene regulatory relationship

Online Enrichment

We also provide an Online Enrichment feature, where users can upload their own gene sets for enrichment analysis.

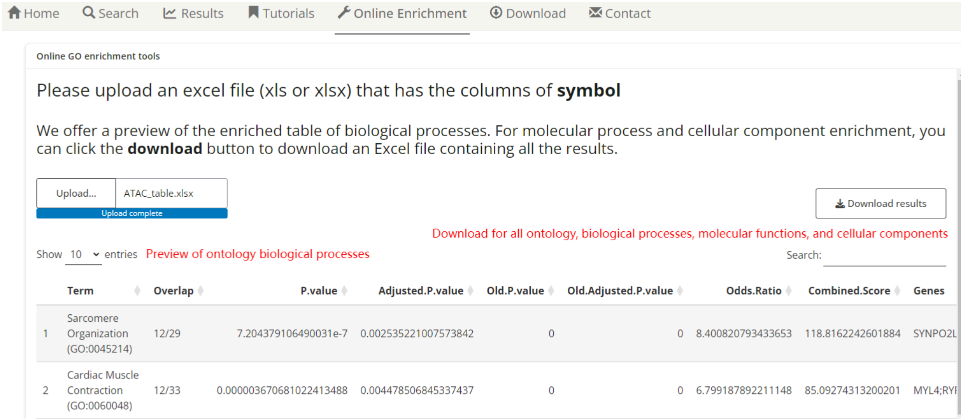


Please upload an excel file (xls or xlsx) that has the columns of **symbol**

We offer a preview of the enriched table of biological processes. For molecular process and cellular component enrichment, you can click the **download** button to download an Excel file containing all the results.

Upload an excel file

After the upload is complete, the system will automatically perform the enrichment analysis. Once the analysis is finished, a preview of the biological processes will be displayed. Users can click the download button to obtain the full set of results, including biological processes, molecular functions, and cellular components.



Online GO enrichment tools

Please upload an excel file (xls or xlsx) that has the columns of **symbol**

We offer a preview of the enriched table of biological processes. For molecular process and cellular component enrichment, you can click the **download** button to download an Excel file containing all the results.

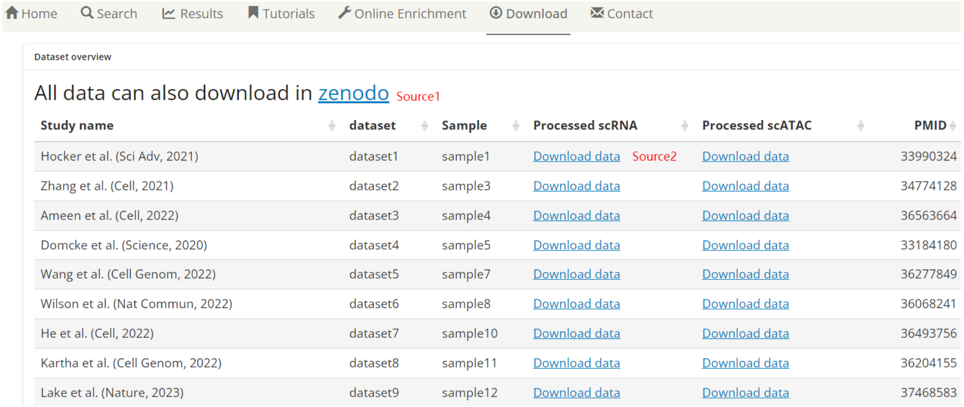
Download for all ontology, biological processes, molecular functions, and cellular components

Preview of ontology biological processes

	Term	Overlap	P.value	Adjusted.P.value	Old.P.value	Old.Adjusted.P.value	Odds.Ratio	Combined.Score	Genes
1	Sarcomere Organization (GO:0045214)	12/29	7.204379106490031e-7	0.002535221007573842	0	0	8.400820793433653	118.8162242601884	SYNPO2L
2	Cardiac Muscle Contraction (GO:0060048)	12/33	0.000003670681022413488	0.004478506845337437	0	0	6.799187892211148	85.09274313200201	MYL4;RY

Data available

Finally, regarding the download feature, all of our processed data can be downloaded from two sources: Zenodo and CowTransfer (recommended for users in mainland China).



Dataset overview

All data can also download in zenodo Source1

Study name	dataset	Sample	Processed scRNA	Processed scATAC	PMID
Hocker et al. (Sci Adv, 2021)	dataset1	sample1	Download data Source2	Download data	33990324
Zhang et al. (Cell, 2021)	dataset2	sample3	Download data	Download data	34774128
Ameen et al. (Cell, 2022)	dataset3	sample4	Download data	Download data	36563664
Domcke et al. (Science, 2020)	dataset4	sample5	Download data	Download data	33184180
Wang et al. (Cell Genom, 2022)	dataset5	sample7	Download data	Download data	36277849
Wilson et al. (Nat Commun, 2022)	dataset6	sample8	Download data	Download data	36068241
He et al. (Cell, 2022)	dataset7	sample10	Download data	Download data	36493756
Kartha et al. (Cell Genom, 2022)	dataset8	sample11	Download data	Download data	36204155
Lake et al. (Nature, 2023)	dataset9	sample12	Download data	Download data	37468583

Code available

All the code for this database is publicly accessible, including the Shiny app and data preprocessing scripts. For more details, please visit github.

Contact

If your have any question, please don't hesitate to contact us

Yunhui: kongyunhui1@gmail.com

Junyao: jyjiang@link.cuhk.edu.hk