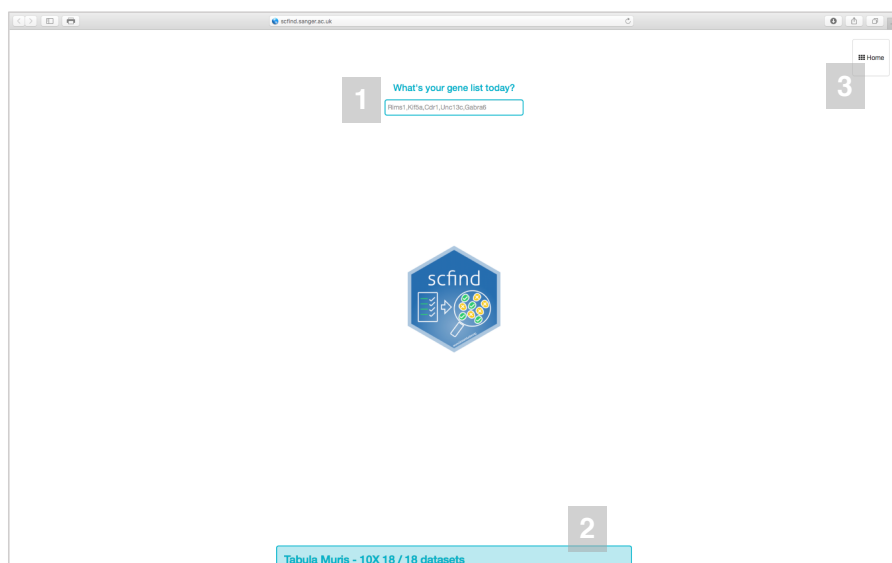




*Search engine for genes in
large single-cell sequencing collections*
<https://scfind.sanger.ac.uk/>
(User Guide)

The User interface of Scfind



- 1 Search box:**
Users can input a list of gene names, separate by either commas or spaces
- 2 Dataset check box:**
Where the user can find all the datasets in the index and limit the search result by dataset / tissue
- 3 Home button:**
To the home page of our index collection

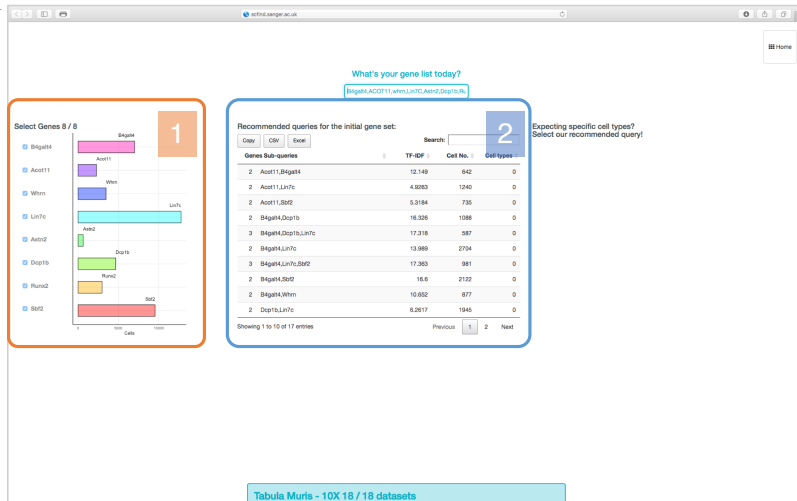
4 Steps : From your favorite gene list to your target cell type

I

What's your gene list today?

B4gat4,ACOT11,whrn,Lin7C,Astrn2,Dcp1b,Ru

II



Step I: Input your favorite gene list

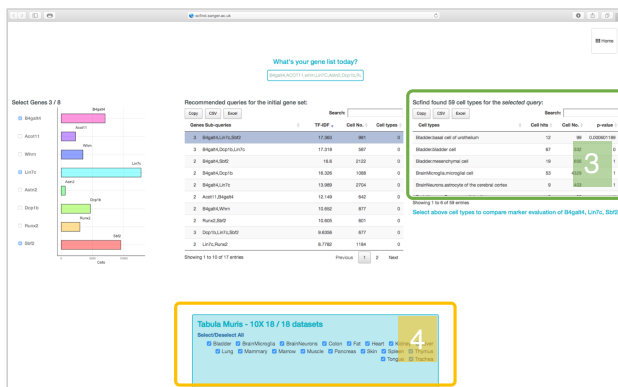
- The bar chart **1** is a summary of the number of cells in the index expressing the gene

- User can always amend the initial gene list by clicking on ☒

Step II: select one of our recommended sub-query on table **2** according to TF-IDF score

- subqueries are ranked by TF-IDF score (*Highest = likely the best*)
- click on **tfidf** to sort TF-IDF score in order

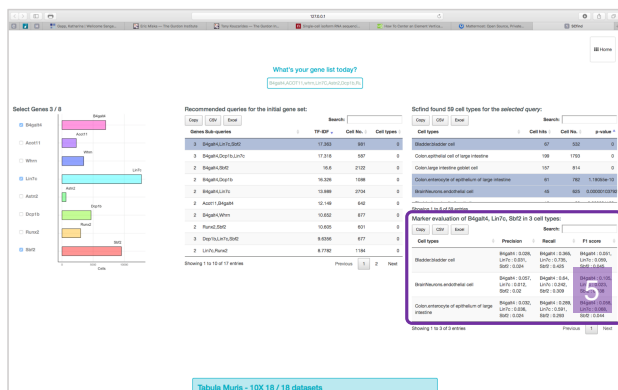
III



Step III: Select the enriched cell types on table **3**

- Cells expressing all genes in the query are grouped by cell type. The user can rank them based on enrichment by clicking on **pval**
- User can limit the dataset by clicking on ☒ in **4**

IV



Step IV: Check the best representing genes

- Click on any cell type you are interested in **3**
- **5** allows you to compare precision / recall / f1 values of each gene corresponding to each cell type
- User can click on **Copy** **CSV** **Excel** to export the tables as raw text, csv file or .xlsx file