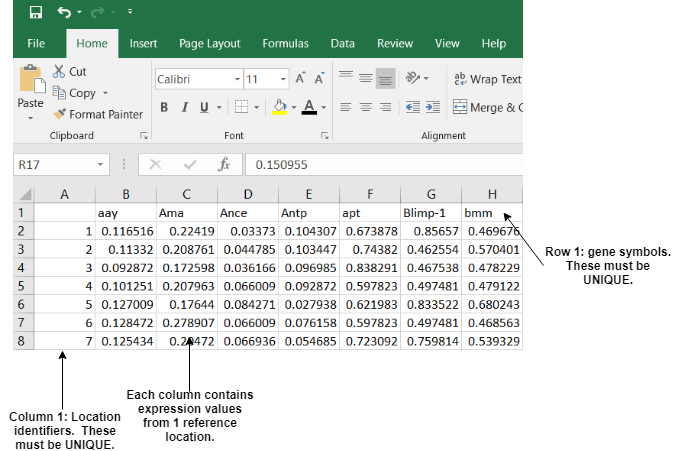
#### Step 1: Please upload data sets

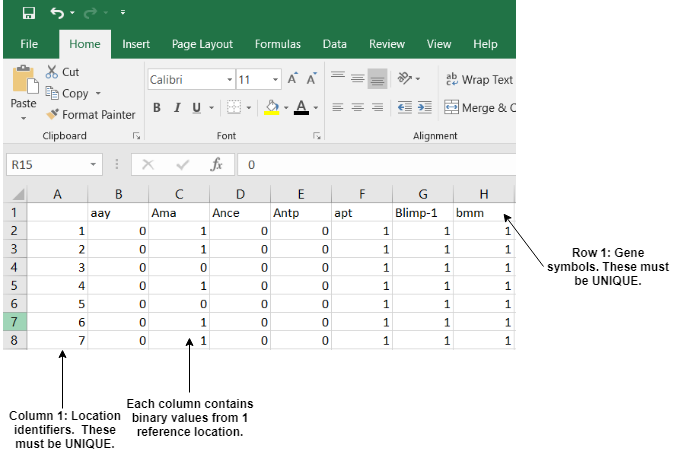
* Please upload a reference scRNA-seq raw file

Description: Upload a .csv file with a reference gene expression matrix for locations in the tissue. The rows are locations and the columns are genes. The first column is a numeric vector which contains the indices of locations. The first row is a character vector which contains the symbols of genes. The file is organized as follows:



* Please upload a reference scRNA-seq binarized file

Description: Upload a .csv file with a binarized reference gene expression matrix for locations in the tissue. The rows are locations and the columns are genes. The first column is a numeric vector which contains the indices of locations. The first row is a character vector which contains the symbol of genes. The value of gene expression is 0 or 1. The file is organized as follows:

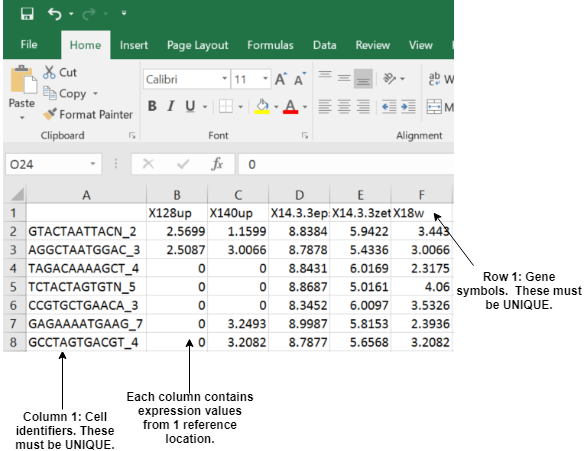


* Please upload a scRNA-seq raw file

Description: Upload a .csv file with a gene expression matrix for cells which are needed to predict their positions. The columns are genes and the rows are cells. The first column is a vector which contains the indices or names of cells. The first row is a character vector which contains the symbol of genes. The file is organized as follows:

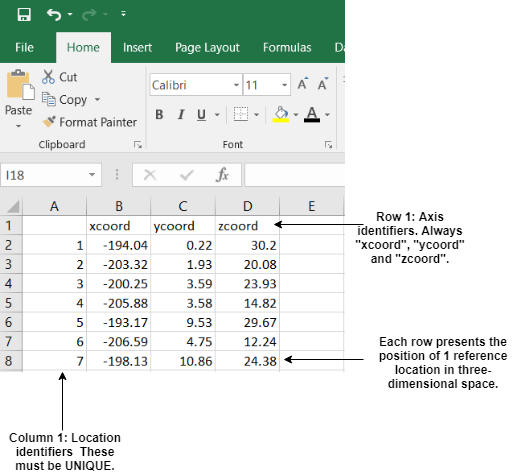
* Please upload a scRNA-seq normalized file

Description: Upload a .csv file with a normalized gene expression matrix for cells which are needed to predict their positions. The normalized matrix is obtained by dividing each entry by the total number of UMIs for that cell, adding a pseudocount and taking the logarithm of that. All entries are finally multiplied by a constant. The columns are genes and the rows are cells. The first column is a vector which contains the indices or names of cells. The first row is a character vector which contains the symbol of genes. The file is organized as follows:



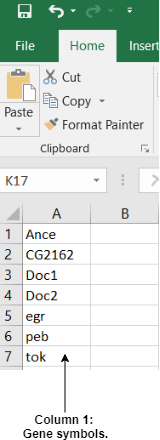
* Please upload a 3D geometry file

Description: Upload a .csv file with three-dimensional coordinate value of (x, y, z). The coordinate (x, y, z) indicates the physical position in the tissue. The first column is a numeric vector which contains the indices of locations. The first row is a character vector which contains the three-dimensional name of axis. The file is organized as follows:



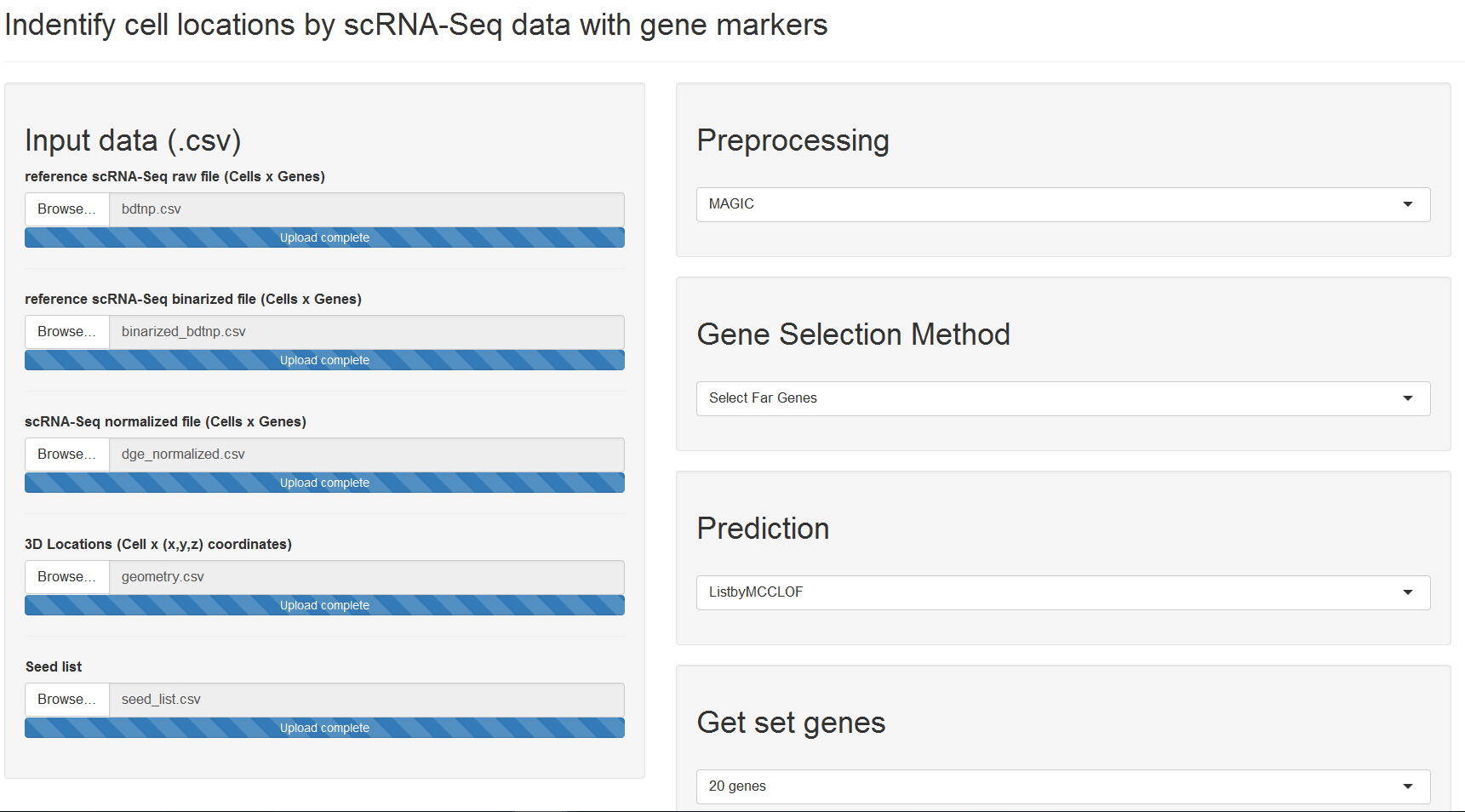
* Please upload a seed list:

Description: Upload a .csv file with a seed gene list in the first column.



#### Step 2: Choose a function for preprocessing

There is a drop-down box providing 2 kinds of functions for pre-processing the raw gene expression data set. The default option is “Use MAGIC”. If you want to use the raw scRNA-seq data, please choose the “Use Raw file” option (showed in Fig. 1).

Figure1. The user interface of identifying cell locations by scRNA-seq data with gene markers

#### Step 3: Choose a function for gene selection

There is a drop-down box providing 11 kinds of functions for the gene selection, which covers all functions in the SCTCwhatateam R package and external python functions. The default option is “Select Genes with highest variations” (showed in Fig. 1).

#### Step 4: Choose a function for prediction of cell locations

There is a drop-down box providing 3 kinds of functions for cell locations prediction, which covers all the functions in the SCTCwhatateam R package. The default option is “by MCC” (showed in Fig. 1). The prediction method will output a cell by location matrix which records the 10 most possible positions for each cell.

#### Appendix

The five input data sets are not compulsory for each method. But the gene name in these data sets should be consistent. Table 1 shows the requirement files for each method. If you use that method, please make sure you upload the corresponding requiring data sets.

Table 1. The requirement files for each method

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Reference scRNA-Seq raw file | Reference scRNA-Seq binarized file | scRNA-Seq normalized file | 3D geometry file | Seed list |
| Use MAGIC | √ |  |  |  |  |
| Use Raw file | √ |  |  |  |  |
| Use Normalized file |  |  | √ |  |  |
| Select Genes with highest variations | √ |  |  |  |  |
| Select Genes with highest variations with seeds | √ |  |  |  | √ |
| Select Genes by high ranking from Google PageRank | √ |  |  |  |  |
| Select Genes by MAD Expression |  | √ |  |  |  |
| Select Genes by MAD Distance | √ |  |  |  |  |
| Select Genes by Influence | √ |  |  |  |  |
| Select Genes by Influence with Seeds | √ |  |  |  | √ |
| Select Genes by high ranking by Influence | √ |  |  |  |  |
| Select Genes by LinRev | √ |  |  |  |  |
| Select Genes by LinFwd | √ |  |  |  |  |
| Select Genes by LinGen | √ |  |  |  |  |
| By MCC |  | √ | √ |  |  |
| By MCC-LOF |  | √ | √ |  |  |
| By correlation | √ |  | √ |  |  |