

SAW

1. Introduction

1.1. Author or Institution

1.1.1. BGI: Beijing Genomics Institution

1.1.2. STOmics: Stereo Omics

1.2. Description

1.3. Resources

1.3.1. Github source code HTML

1.3.2. Website Documents

1.4. Paper abstract

1.4.1. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned a...

2. Documents

2.1. Install

2.1.1. Requirements

2.1.1.1. Hardware

2.1.1.2. Software

2.1.2. Install

2.2. Usage

2.2.1. Spatial Transcriptomics Workflow

2.2.2. Spatial Proteomics_Transcriptomics Workflow

2.2.3. Spatial Transcriptomics Workflow Parts

2.2.4. Spatial Proteomics_Transcriptomics Workflow Parts

2.3. Cases

2.3.1. Build index for reference genome

2.3.2. Get Stereo-seq Chip T mask file

2.3.3. Get the protein panel

2.3.4. Spatial Transcriptomics Workflow

2.3.5. Spatial Proteomics & Transcriptomics Workflow

3. Workflow

3.1. Stereo RNA

3.2. Stereo RNA-Protein

4. References

1. Introduction

1.1. Author or Institution

1.1.1. BGI: Beijing Genomics Institution

<https://www.genomics.cn/>

1.1.2. STomics: Stereo Omics

<http://stomics.tech/>

1.2. Description

SAW: Stereo-seq Analysis Workflow Workflow for analyzing Stereo-seq transcriptomic data. Stereo-seq Analysis Workflow (SAW) software suite is a set of pipelines bundled to map sequenced reads to their spatial location on the tissue section, quantify the corresponding gene expression levels and visually present spatial gene expression distribution.

1.3. Resources

1.3.1. Github source code [HTML](#)

<https://github.com/STomics/SAW>

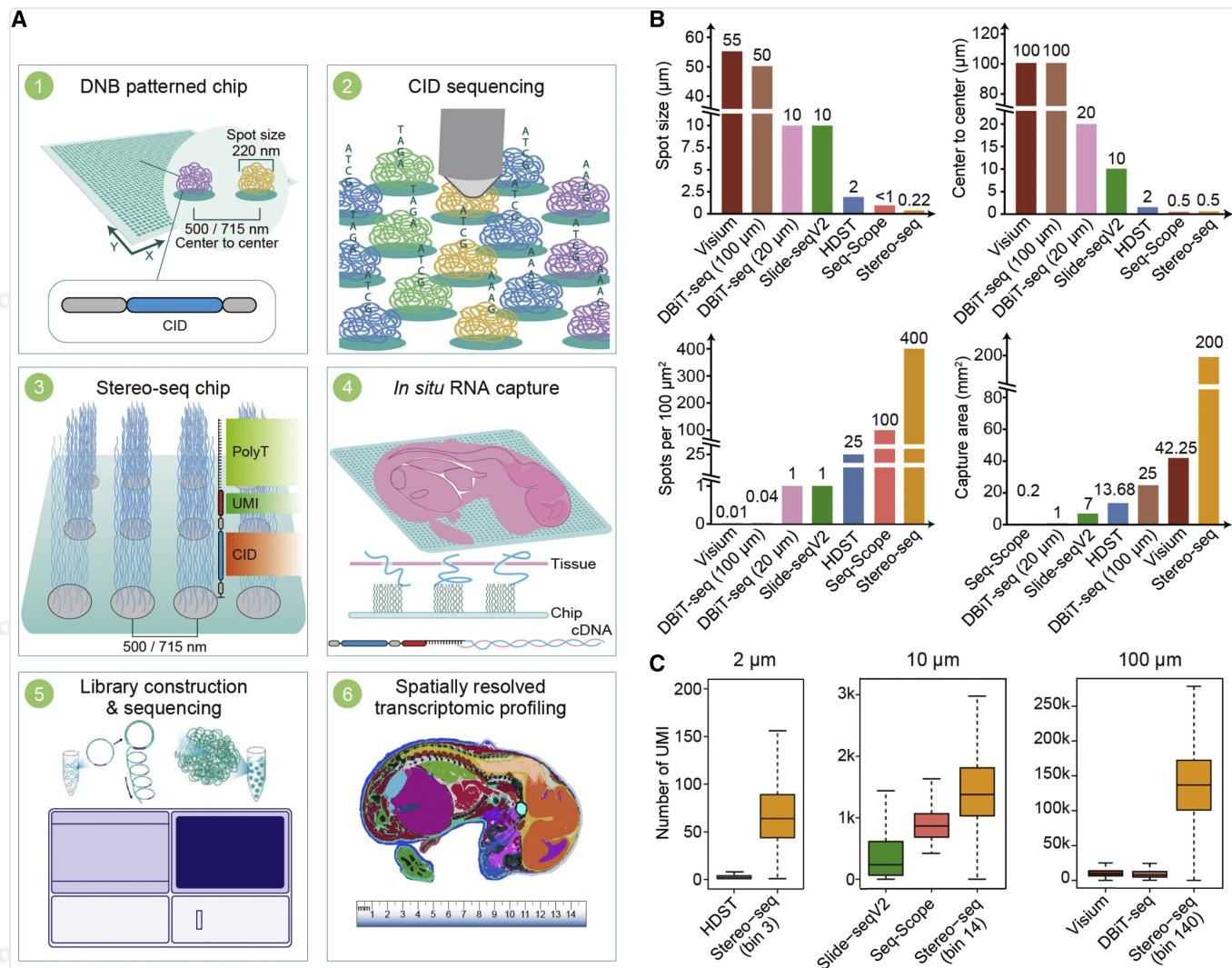
1.3.2. Website Documents

<https://www.stomics.tech/products/BioinfoTools/OfflineSoftware/SAWOperationManuals>

1.4. Paper abstract

1.4.1. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball–patterned arrays

Ao Chen, Sha Liao, Mengnan Cheng, Longqi Liu, Xun Xu, Jian Wang. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 2022, doi: <https://doi.org/10.1016/j.cell.2022.04.003>



2. Documents

<https://github.com/STomics/SAW>

<https://www.stomics.tech/products/BioinfoTools/OfflineSoftware/SAWOperationManuals>

2.1. Install

<https://github.com/STomics/SAW>

2.1.1. Requirements

2.1.1.1. Hardware

Stereo-seq Analysis Workflow (SAW) should be run on a Linux system that meets the following requirements:

1. 8-core Intel or AMD processor (>24 cores recommended)
2. 128GB RAM (>256GB recommended)
3. 1TB free disk space or higher
4. 64-bit CentOS/RedHat 7.8 or Ubuntu 20.04

2.1.1.2. Software

1. **Docker** : a container platform, >=20.10.8
2. **Singularity** : a container platform, >=3.8
3. **SAW** in the **Singularity Image File (SIF)** format
4. **ImageStudio** >= v3.0
5. **StereoMap** >= v3.1

2.1.2. Install

```
1 # 1. Quick installation of Singularity
2
3 ## On Red Hat Enterprise Linux or CentOS install the following dependencies:
4 $ sudo yum update -y && \
5     sudo yum groupinstall -y 'Development Tools' && \
6     sudo yum install -y \
7     openssl-devel \
8     libuuid-devel \
9     libseccomp-devel \
10    wget \
11    squashfs-tools \
12    cryptsetup
13
14 ## On Ubuntu or Debian install the following dependencies:
15 $ sudo apt-get update && sudo apt-get install -y \
16     build-essential \
17     uuid-dev \
18     libgpgme-dev \
19     squashfs-tools \
20     libseccomp-dev \
21     wget \
22     pkg-config \
23     git \
24     cryptsetup-bin
25
26 ## Install Go
27 $ export VERSION=1.14.12 OS=linux ARCH=amd64 && \
28     wget https://dl.google.com/go/go$VERSION.$OS-$ARCH.tar.gz && \
29     sudo tar -C /usr/local -xzvf go$VERSION.$OS-$ARCH.tar.gz && \
30     rm go$VERSION.$OS-$ARCH.tar.gz
31
32 $ echo 'export GOPATH=${HOME}/go' >> ~/.bashrc && \
33     echo 'export PATH=/usr/local/go/bin:${PATH}: ${GOPATH}/bin' >> ~/.bashrc && \
34     source ~/.bashrc
35
36 ## Install singularity on CentOS without compile
37 $ yum install -y singularity
```

```
1 # 2. Quick download SAW from DockerHub
2
3 # Currently, the latest version of SAW in DockerHub is 07.1.0. You can download SAW by running the following command:
4 singularity build SAW_<version>.sif docker://stomics/saw:<version>
5
6 # singularity build SAW_7.1.sif docker://stomics/saw:07.1.1
7 # singularity build SAW_7.0.sif docker://stomics/saw:07.0.1
8 # singularity build SAW_6.1.sif docker://stomics/saw:06.1.4
9 # singularity build SAW_6.0.sif docker://stomics/saw:06.0.2
10 # singularity build SAW_5.5.sif docker://stomics/saw:05.5.4
11 # singularity build SAW_5.4.sif docker://stomics/saw:05.4.0
12 # singularity build SAW_5.1.sif docker://stomics/saw:05.1.3
13 # singularity build SAW_4.1.sif docker://stomics/saw:04.1.0
14 # singularity build SAW_4.0.sif docker://stomics/saw:04.0.0
15 # singularity build SAW_2.1.sif docker://stomics/saw:02.1.0
16 # singularity build SAW_2.0.sif docker://stomics/saw:02.0.0
17 # singularity build SAW_1.0.sif docker://stomics/saw:01.0.0
```

2.2. Usage

[https://cdn-](https://cdn-newfile.stomics.tech/%E5%88%86%E6%9E%90%E6%B5%81%E7%A8%8B%E8%BD%AF%E4%BB%B6%E5%8C%85%E4%BD%BF%E7%94%A8%E6%89%8B%E5%86%8Cv7.1.pdf)

[newfile.stomics.tech/%E5%88%86%E6%9E%90%E6%B5%81%E7%A8%8B%E8%BD%AF%E4%BB%B6%E5%8C%85%E4%BD%BF%E7%94%A8%E6%89%8B%E5%86%8Cv7.1.pdf](https://cdn-newfile.stomics.tech/Stereo-seq_Analysis_Workflow_User_Manual_A9.pdf)

https://cdn-newfile.stomics.tech/Stereo-seq_Analysis_Workflow_User_Manual_A9.pdf

2.2.1. Spatial Transcriptomics Workflow

```
1 # SAW v7.1
2 usage: bash stereoPipeline_v7.1.sh -splitCount -maskFile -fq1 -fq2 -refIndex -genomeFile -speciesName -tissueType -annotationFile -outDir -imageRecordFile -imageCompressedFile -doCellBin -rRNARemove -threads -sif
3           -splitCount : count of splitted stereochip mask file, usually 16 for Q
4 fq data and 1 for Q40 fq data
5           -maskFile : stereochip mask file
6           -fq1 : fastq file path of read1, if there are more than one fastq fil
e, please separate them with comma, e.g:lane1_read_1.fq.gz,lane2_read_1.f
q.gz
7           -fq2 : fastq file path of read2, if there are more than one fastq fil
e, please separate them with comma, not requested for Q4 fastq data, e.g:l
ane1_read_2.fq.gz,lane2_read_2.fq.gz
8           -refIndex : reference genome indexed folder, please build index befor
e SAW analysis run
9           -speciesName : specie of the sample
10          -tissueType : tissue type of the sample
11          -annotationFile : annotations file in gff or gtf format, the file mus
t contain gene and exon annotations
12          -outDir : output directory path
13          -imageRecordFile : image file(*.ipr) generated by ImageStudio softwar
e, not requested
14          -imageCompressedFile : image file(*.tar.gz) generated by ImageStudio s
oftware, not requested
15          -doCellBin : [Y/N]
16          -rRNARemove : [Y/N]
17          -threads : the number of threads to be used in running the pipeline
18          -sif : the file format of the visual software
```

2.2.2. Spatial Proteomics_Transcriptomics Workflow

```

1 # SAW v7.1
2 usage: bash stereoPipeline_pt_v7.1.sh -splitCount -maskFile -rnaFq1 -rnaFq
2 -adtFq1 -adtFq2 -proteinList -refIndex -genomeFile -speciesName -tissueT
ype -annotationFile -outDir -imageRecordFile -imageCompressedFile -doCellB
in -rRNAremove -threads -sif
3     -splitCount : count of splited stereochip mask file, usually 16 for Q
4 fq data and 1 for Q40 fq data
5     -maskFile : stereochip mask file
6     -rnaFq1 : RNA fastq file path of read1, if there are more than one fas
tq file, please separate them with comma, e.g:lane1_read_1.fq.gz,lane2_rea
d_1.fq.gz
7     -rnaFq2 : RNA fastq file path of read2, if there are more than one fas
tq file, please separate them with comma, not requested for Q4 fastq dat
a, e.g:lane1_read_2.fq.gz,lane2_read_2.fq.gz
8     -adtFq1 : ADT fastq file path of read1, if there are more than one fas
tq file, please separate them with comma, e.g:lane1_read_1.fq.gz,lane2_rea
d_1.fq.gz
9     -adtFq2 : ADT fastq file path of read2, if there are more than one fas
tq file, please separate them with comma, not requested for Q4 fastq dat
a, e.g:lane1_read_2.fq.gz,lane2_read_2.fq.gz
9     -proteinList : protein list file which contain protein id, sequences a
nd names.
10    -refIndex : reference genome indexed folder, please build index befor
e SAW analysis run
11    -annotationFile : annotations file in gff or gtf format, the file mus
t contain gene and exon annotations
12    -speciesName : specie of the sample
13    -tissueType : tissue type of the sample
14    -outDir : output directory path
15    -imageRecordFile : image file(*.ipr) generated by ImageStudio softwar
e, not requested
16    -imageCompressedFile : image file(*.tar.gz) generated by ImageStudio s
oftware, not requested
17    -doCellBin : [Y/N]
18    -rRNAremove : [Y/N]
19    -threads : the number of threads to be used in running the pipeline
20    -sif : the file format of the visual software
21
22 # 1GiB=1024M=10241024KB=10241024*1024B
23 # SAW version : v7.1

```

2.2.3. Spatial Transcriptomics Workflow Parts

```
1 # SAW v7.1 manual
2 # for saw_v7.1_manual spatial_Transcriptomics_workflow
3 ## part1
4 usage: bash stereoPipeline_v7.1_manual_part1.sh -genomeSize -splitCount -maskFile -fq1 -fq2 -speciesName -tissueType -refIndex -annotationFile -imageRecordFile -imageCompressedFile -sif -threads -outDir
5      -splitCount : count of splitted stereochip mask file, usually 16 for 0
6      4 fq data and 1 for Q40 fq data
7      -maskFile : stereochip mask file
8      -fq1 : fastq file path of read1, if there are more than one fastq file, please separate them with comma, e.g:lane1_read_1.fq.gz,lane2_read_1.fq.gz
9      -fq2 : fastq file path of read2, if there are more than one fastq file, please separate them with comma, not requested for Q4 fastq data, e.g:lane1_read_2.fq.gz,lane2_read_2.fq.gz
10     -speciesName : specie of the sample
11     -tissueType : tissue type of the sample
12     -refIndex : reference genome indexed folder, please build IT before SAW analysis run
13     -annotationFile : annotations file in gff or gtf format, the file must contain gene and exon annotations
14     -rRNARemove : [Y/N]
15     -imageRecordFile : image file(*.ipr) generated by ImageStudio software, not requested
16     -imageCompressedFile : image file(*.tar.gz) generated by ImageStudio software, not requested
17     -sif : the file format of the visual software
18     -threads : the number of threads to be used in running the pipeline
19     -outDir : output directory path
20
21 ## part2
22 usage: bash stereoPipeline_v7.1_manual_part2.sh -SN -dataDir -registJson -speciesName -tissueType -outDir -imageRecordFile -imageCompressedFile -cellBin -threads -sif
23     -SN : sample id
24     -dataDir : output directory of gene expression matrix result
25     -registJson : manual registration json file
26     -speciesName : specie of the sample
27     -tissueType : tissue type of the sample
28     -outDir : output directory path
29     -imageRecordFile : QC-success: image file(*.ipr) found in /dataDir/03. register, QC-failed: image file(*.ipr) generated by ImageStudio software
30     -imageCompressedFile : image file(*.tar.gz) generated by ImageStudio software
```

```
31      -doCellBin : [Y/N]
32      -threads : the number of threads to be used in running the pipeline
33      -sif : the file format of the visual software
```

2.2.4. Spatial Proteomics_Transcriptomics Workflow Parts

```

1 # for saw_pt_v7.1_manual spatial_Proteomics_and_Transcriptomics_workflow
2 ## part1
3 usage: bash stereoPipeline_pt_v7.1_manual_part1.sh -genomeSize -splitCount
        -maskFile -rnafq1 -rnafq2 -adtFq1 -adtFq2 -proteinList -refIndex -annotationFile
        -speciesName -tissueType -rRNAremove -imageRecordFile -imageCompressedFile -sif -threads -outDir
4         -splitCount : count of splitted stereochip mask file, usually 16 for 0
4 fq data and 1 for Q40 fq data
5         -maskFile : stereochip mask file
6         -rnafq1 : RNA fastq file path of read1, if there are more than one fas
tq file, please separate them with comma, e.g:lane1_read_1.fq.gz,lane2_rea
d_1.fq.gz
7         -rnafq2 : RNA fastq file path of read2, if there are more than one fas
tq file, please separate them with comma, not requested for Q4 fastq dat
a, e.g:lane1_read_2.fq.gz,lane2_read_2.fq.gz
8         -adtFq1 : ADT fastq file path of read1, if there are more than one fas
tq file, please separate them with comma, e.g:lane1_read_1.fq.gz,lane2_rea
d_1.fq.gz
9         -adtFq2 : ADT fastq file path of read2, if there are more than one fas
tq file, please separate them with comma, not requested for Q4 fastq dat
a, e.g:lane1_read_2.fq.gz,lane2_read_2.fq.gz
10        -proteinList : protein list file which contain protein sequences and n
ames.
11        -refIndex : reference genome indexed folder, please build IT before SAW
analysis run
12        -annotationFile : annotations file in gff or gtf format, the file mus
t contain gene and exon annotations
13        -speciesName : specie of the sample
14        -tissueType : tissue type of the sample
15        -rRNAremove : [Y/N]
16        -imageRecordFile : image file(*.ipr) generated by ImageStudio softwar
e, not requested
17        -imageCompressedFile : image file(*.tar.gz) generated by ImageStudio s
oftware, not requested
18        -sif : the file format of the visual software
19        -threads : the number of threads to be used in running the pipeline
20        -outDir : output directory path
21
22 ## part2
23 usage: sh stereoPipeline_pt_v7.1_manual_part2.sh -SN -dataDir -registJson
        -speciesName -tissueType -outDir -imageRecordFile -imageCompressedFile -do
-CellBin -threads -sif
24        -SN : sample id
25        -dataDir : output directory of gene expression matrix result
26

```

```

27     -proteinList : protein list file which contain protein sequences and n
28     ames
29     -registJson : manual registration json file
30     -speciesName : specie of the sample
31     -tissueType : tissue type of the sample
32     -outDir : output directory path
33     -imageRecordFile : QC-success: image file(*.ipr) found in /dataDir/04.
34     register; QC-failed: image file(*.ipr) generated by ImageStudio software
35     -imageCompressedFile : image file(*.tar.gz) generated by ImageStudio s
36     oftware
37     -doCellBin : [Y/N]
38     -threads : the number of threads to be used in running the pipeline
39     -sif : the file format of the visual software
40
41 # 1GiB=1024M=10241024KB=10241024*1024B
42 # SAW version : v7.1

```

2.3. Cases

2.3.1. Build index for reference genome

A genome index has to be constructed before performing data mapping. The index files are used as reference when aligning reads. You can prepare the indexed reference before run SAW as follow:

```

1  referenceDir=/Full/Path/Of/Reference/Folder/Path
2  mkdir $referenceDir/STAR_SJ100
3  export SINGULARITY_BIND=$referenceDir
4
5  singularity exec <SAW_version.sif> mapping \
6      --runMode genomeGenerate \
7      --genomeDir $referenceDir/STAR_SJ100 \
8      --genomeFastaFiles $referenceDir/genome/genome.fa \
9      --sjdbGTFfile $referenceDir/genes/genes.gtf \
10     --sjdbOverhang 99 \
11     --runThreadN 12

```

Because of update of mapping module for computational efficiency, from SAW V6.1, index for reference genome has to be reconstructed with the same commands. For more information, refer to "script/pre_buildIndexedRef"

2.3.2. Get Stereo-seq Chip T mask file

- If you want to access mask file (.h5/.bin) for your own data, please contact BGI-FAS team.
- To access mask file for published paper, please go to [CNGBdb > STOmicsDB > Collections](#).

2.3.3. Get the protein panel

If you need to run SAW Proteomics & Transcriptomics workflow, please confirm you chose the correct species for protein panel. Access the protein panel: SAW/Documents/UserManual/ProteinPanel.

2.3.4. Spatial Transcriptomics Workflow

For SAW_v7.1, please use the [stereoPipeline_v7.1.sh](#) to run the whole workflow.

For SAW_v7.1_manual, please use the [stereoPipeline_v7.1_manual_part1.sh](#) and [stereoPipeline_v7.1_manual_part2.sh](#) to finish the manual processing.

```

1  cd <Your Working Directory>
2
3  ulimit -n 10240
4  ulimit -v 33170449147
5  NUMBA_CACHE_DIR=<Your Working Directory>
6
7  dataDir=<Your Working Directory>/rawData
8  outDir=<Your Working Directory>/result
9
10 export SINGULARITY_BIND=$dataDir,$outDir
11
12 ## Choose from the following scenarios
13
14 ## Scenario 1: input image and run cell bin
15 bash stereoPipeline.sh \
16   -sif $dataDir/SAW/SAW_<version>.sif \
17   -splitCount 1 \ ## 16 or 64 for Q4, 1 for Q40
18   -maskFile $dataDir/mask/SN.h5 \
19   -fq1 $dataDir/reads/lane1_read_1.fq.gz,...,$dataDir/reads/laneN_read_
20   1.fq.gz \
21   -fq2 $dataDir/reads/lane1_read_2.fq.gz,...,$dataDir/reads/laneN_read_
22   2.fq.gz \ # [optional] when the sequenced data is in Q40 format
23   -speciesName <speciesName> \
24   -tissueType <tissueName> \
25   -refIndex $dataDir/reference/STAR_SJ100 \
26   -annotationFile $dataDir/reference/genes.gtf \ ## GFF or GTF
27   -rRNARemove : N \
28   -threads 16 \
29   -outDir $outDir/result \
30   -imageRecordFile $dataDir/image/<SN_date_time_version>.ipr \ # [option
31   al] when image is given and has passed QC
32   -imageCompressedFile $dataDir/image/<SN_date_time_version>tar.gz \ # [
33   optional] when image is given and has passed QC
34   -doCellBin Y # [optional] when you want to do the cell segmentation a
35   nd get cell gene expression data
36
37 ## Scenario 2: input image but no need for cell bin
38 bash stereoPipeline.sh \
39   -sif $dataDir/SAW/SAW_<version>.sif \
40   -splitCount 1 \ ## 16 or 64 for Q4, 1 for Q40
41   -maskFile $dataDir/mask/SN.h5 \
42   -fq1 $dataDir/reads/lane1_read_1.fq.gz,...,$dataDir/reads/laneN_read_
43   1.fq.gz \
44   -fq2 $dataDir/reads/lane1_read_2.fq.gz,...,$dataDir/reads/laneN_read_
45   2.fq.gz \ # [optional] when the sequenced data is in Q40 format

```

```

39      -speciesName <speciesName> \
40      -tissueType <tissueName> \
41      -refIndex $dataDir/reference/STAR_SJ100 \
42      -annotationFile $dataDir/reference/genes.gtf \ ## GFF or GTF
43      -rRNARemove : N \
44      -threads 16 \
45      -outDir $outDir/result \
46      -imageRecordFile $dataDir/image/<SN_date_time_version>.ipr \ # [option
47      al] when image is given and has passed QC
48      -imageCompressedFile $dataDir/image/<SN_date_time_version>tar.gz \ #
[optional] when image is given and has passed QC
49      -doCellBin N # [optional] when you want to do the cell segmentation a
nd get cell gene expression data
50
51      ## Scenario 3: no image
52      bash stereoPipeline.sh \
53          -sif $dataDir/SAW/SAW_<version>.sif \
54          -splitCount 1 \ ## 16 or 64 for Q4, 1 for Q40
55          -maskFile $dataDir/mask/SN.h5 \
56          -fq1 $dataDir/reads/lane1_read_1.fq.gz,...,$dataDir/reads/laneN_read_
57          1.fq.gz \
58          -fq2 $dataDir/reads/lane1_read_2.fq.gz,...,$dataDir/reads/laneN_read_
59          2.fq.gz \ # [optional] when the sequenced data is in Q40 format
60          -speciesName <speciesName> \
61          -tissueType <tissueName> \
62          -refIndex $dataDir/reference/STAR_SJ100 \
63          -annotationFile $dataDir/reference/genes.gtf \ ## GFF or GTF
64          -rRNARemove : N \
65          -threads 16 \
66          -outDir $outDir/result

```

2.3.5. Spatial Proteomics & Transcriptomics Workflow

For SAW_v7.1, please use the [stereoPipeline_pt_v7.1.sh](#) to run the whole workflow.

For SAW_v7.1_manual, please use the [stereoPipeline_pt_v7.1_manual_part1.sh](#) and [stereoPipeline_pt_v7.1_manual_part2.sh](#) to finish the manual processing.

```
1  cd <Your Working Directory>
2
3  ulimit -n 10240
4  ulimit -v 33170449147
5  NUMBA_CACHE_DIR=<Your Working Directory>
6
7  dataDir=<Your Working Directory>/rawData
8  outDir=<Your Working Directory>/result
9
10 export SINGULARITY_BIND=$dataDir,$outDir
11
12 ## Choose from the following scenarios
13
14 ## Scenario 1: input image and run cell bin
15 bash stereoPipeline_pt.sh \
16     -sif $dataDir/SAW/SAW_<version>.sif \
17     -splitCount 1 \ ## 16 or 64 for Q4, 1 for Q40
18     -maskFile $dataDir/mask/SN.h5 \
19     -adtFq1 $dataDir/ST0mics-ADT/lane1_read_1.fq.gz,...,$dataDir/ST0mics-A
DT/laneN_read_1.fq.gz \
20     -adtFq2 $dataDir/ST0mics-ADT/lane1_read_2.fq.gz,...,$dataDir/ST0mics-A
DT/laneN_read_2.fq.gz \ # [optional] when the sequenced data is in Q40 for
mat
21     -rnaFq1 $dataDir/ST0mics-RNA/lane1_read_1.fq.gz,...,$dataDir/ST0mics-R
NA/laneN_read_1.fq.gz \
22     -rnaFq2 $dataDir/ST0mics-RNA/lane1_read_2.fq.gz,...,$dataDir/ST0mics-R
NA/laneN_read_2.fq.gz \ # [optional] when the sequenced data is in Q40 for
mat
23     -speciesName <speciesName> \
24     -tissueType <tissueName> \
25     -proteinList $dataDir/protein-reference/ProteinPanel.list \
26     -refIndex $dataDir/reference/STAR_SJ100 \
27     -annotationFile $dataDir/reference/genes.gtf \ ## GFF or GTF
28     -rRNAremove N \
29     -threads 16 \
30     -outDir $outDir/result \
31     -imageRecordFile $dataDir/image/<SN_date_time_version>.ipr \ # [option
al] when image is given and has passed QC
32     -imageCompressedFile $dataDir/image/<SN_date_time_version>tar.gz \ #
[optional] when image is given and has passed QC
33     -doCellBin Y # [optional] when you want to do the cell segmentation a
nd get cell gene expression data
34
35 ## Scenario 2: input image but no need for cell bin
36 bash stereoPipeline_pt.sh \
```

```

37      -sif $dataDir/SAW/SAW_<version>.sif \
38      -splitCount 1 \ ## 16 or 64 for Q4, 1 for Q40
39      -maskFile $dataDir/mask/SN.h5 \
40      -adtFq1 $dataDir/ST0mics-ADT/lane1_read_1.fq.gz,...,$dataDir/ST0mics-A
41      DT/laneN_read_1.fq.gz \
42          -adtFq2 $dataDir/ST0mics-ADT/lane1_read_2.fq.gz,...,$dataDir/ST0mics-A
43      DT/laneN_read_2.fq.gz \ # [optional] when the sequenced data is in Q40 for
44      mat
45          -rnaFq1 $dataDir/ST0mics-RNA/lane1_read_1.fq.gz,...,$dataDir/ST0mics-R
46      NA/laneN_read_1.fq.gz \
47          -rnaFq2 $dataDir/ST0mics-RNA/lane1_read_2.fq.gz,...,$dataDir/ST0mics-R
48      NA/laneN_read_2.fq.gz \ # [optional] when the sequenced data is in Q40 for
49      mat
50          -speciesName <speciesName> \
51          -tissueType <tissueName> \
52          -proteinList $dataDir/protein-reference/ProteinPanel.list \
53          -refIndex $dataDir/reference/STAR_SJ100 \
54          -annotationFile $dataDir/reference/genes.gtf \ ## GFF or GTF
55          -rRNAremove N \
56          -threads 16 \
57          -outDir $outDir/result \
58          -imageRecordFile $dataDir/image/<SN_date_time_version>.ipr \ # [option
59      al] when image is given and has passed QC
60          -imageCompressedFile $dataDir/image/<SN_date_time_version>tar.gz \ #
61      [optional] when image is given and has passed QC
62          -doCellBin N # [optional] when you want to do the cell segmentation a
63      nd get cell gene expression data

64      ## Scenario 3: no image
65      bash stereoPipeline_pt.sh \
66          -sif $dataDir/SAW/SAW_<version>.sif \
67          -splitCount 1 \ ## 16 or 64 for Q4, 1 for Q40
68          -maskFile $dataDir/mask/SN.h5 \
69          -adtFq1 $dataDir/ST0mics-ADT/lane1_read_1.fq.gz,...,$dataDir/ST0mics-A
70          DT/laneN_read_1.fq.gz \
71              -adtFq2 $dataDir/ST0mics-ADT/lane1_read_2.fq.gz,...,$dataDir/ST0mics-A
72          DT/laneN_read_2.fq.gz \ # [optional] when the sequenced data is in Q40 for
73          mat
74              -rnaFq1 $dataDir/ST0mics-RNA/lane1_read_1.fq.gz,...,$dataDir/ST0mics-R
75          NA/laneN_read_1.fq.gz \
76              -rnaFq2 $dataDir/ST0mics-RNA/lane1_read_2.fq.gz,...,$dataDir/ST0mics-R
77          NA/laneN_read_2.fq.gz \ # [optional] when the sequenced data is in Q40 for
78          mat
79              -proteinList $dataDir/protein-reference/ProteinPanel.list \
80              -speciesName <speciesName> \
81              -tissueType <tissueName> \
82              -refIndex $dataDir/reference/STAR_SJ100 \
83              -annotationFile $dataDir/reference/genes.gtf \ ## GFF or GTF

```

```

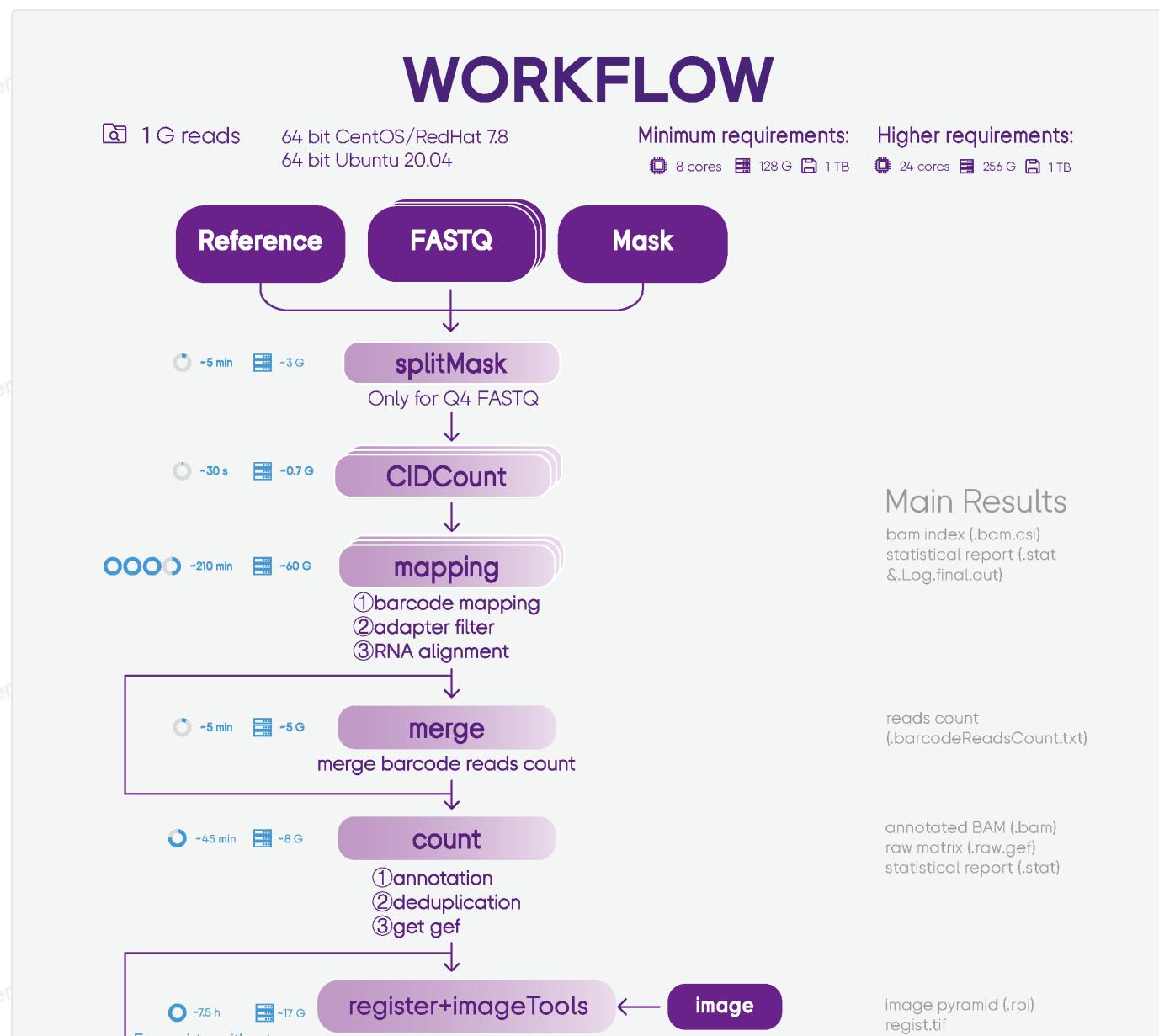
70      -rRNAremove N \
71      -threads 16 \
72      -outDir $outDir/result

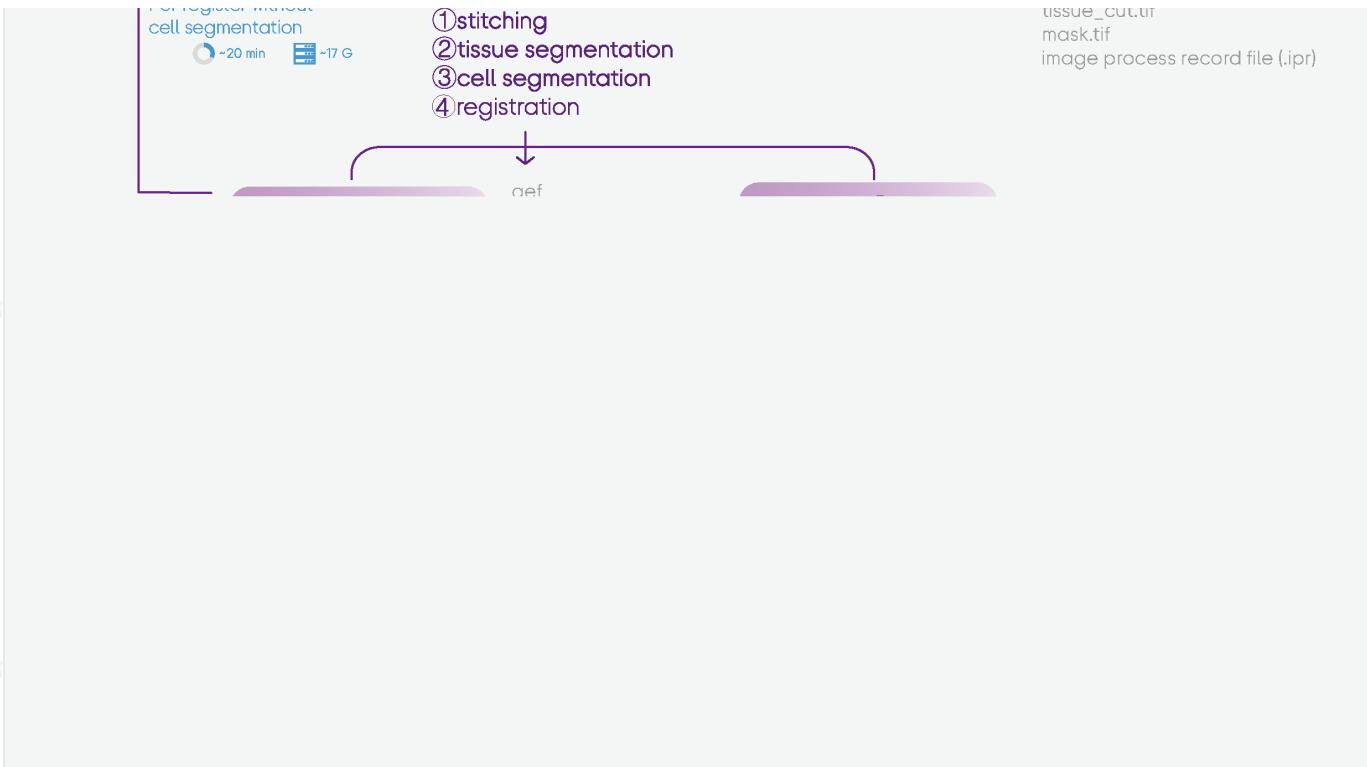
```

3. Workflow

3.1. Stereo RNA

SAW processes the sequencing data of **Stereo-seq** to generate spatial gene expression matrices, and the users could take these files as the starting point to perform downstream analysis. SAW includes thirteen essential and suggest pipelines and auxiliary tools for supporting other handy functions.





3.2. Stereo RNA-Protein

SAW Proteomics & Transcriptomics workflow process sequencing data from one-chip to generate spatial gene and protein expression matrices. You can use these two omics information to start higher-dimensional research. This workflow includes 23 essential and suggest pipelines.

ber

ber

ber

ber

ber

4. References

benben.miao