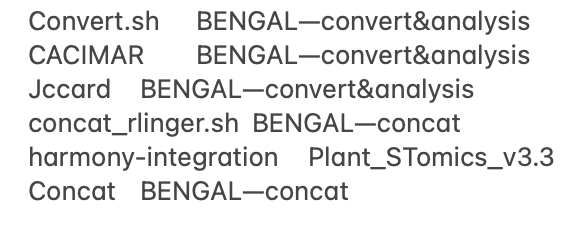
<https://www.stomics.tech/helpcenter/zh/usermanual/E.Project/H.Image.html>

#基于平台预设基础镜像构建

建立环境！



相应的运行环境



参考文章

2023(developmental cell)The maturation and aging trajectory of Marchantia polymorpha at single-cell resolution

将该物种的根的数据拆分出来，加入到root的目录中Mar\_root.rds √

数据读取，根据root、leaf、stem标签拆分root内容，将root内容重命名后保存到root\_data目录里面。 ventral scales和root √

~~Library(Seurat)~~

~~data <- readRDS(“ALL\_CELL.rds”)~~

~~data\_root <- subset(x = data, idents = c("Naive CD4 T", "CD8 T"), invert = FALSE)~~

~~saveRDS(file = “ROOT\_CELL.rds”, data\_root)~~

work.sh流程

s0\_concat.sh

pre-processe.py 输出 adata.write\_h5ad(filename=sample.iat[i,0]+'.h5ad',compression="gzip")

concat\_by\_homology\_multiple\_species\_by\_gene\_id\_s1.R one2one方法输出h5ad

concat\_by\_homology\_multiple\_species\_by\_gene\_id\_s2.R many\_higher\_expr方法输出h5ad

concat\_by\_homology\_multiple\_species\_by\_gene\_id\_s3.R many\_higher\_homology\_conf方法输出h5ad

s0\_rliger\_concat.sh （输出结果到Rliger）

concat\_by\_homology\_rligerUINMF\_multiple\_species\_s1.R 生成似乎跟上一步一样的.text文件，输出在Rliger，有一个metadata的.tsv文件

convert\_format.R 遍历Rliger里面的h5ad文件，转变为rds文件。问题：这个文件夹应该没h5ad文件吧

s1\_convert.sh (if判断，不同的情况使用不同的脚本)

convert\_format.R （stype==anndata\_to\_seurat）跳出Rliger，遍历当前目录里面的h5ad文件，转变为rds文件。

Seurat2h5ad.R （stype==seurat\_to\_anndata）跳出Rliger，遍历当前目录里面的rds文件，转变为h5ad文件。

s1\_integration.sh （输出结果到integration）

LIGER\_integration.R (optparse rliger SeuratWrappers)

fastMNN\_integration.R

seurat\_CCA\_integration.R

seurat\_RPCA\_integration.R

harmony\_integration.py (scanpy, ValueError: numpy.ufunc size changed, may indicate binary incompatibility. Expected 232 from C header, got 216 from PyObject)

scANVI\_integration.py缺少安装包pymde (scvi, AttributeError: module 'jax' has no attribute 'Device')

scanorama\_integration.py缺少安装包scanorama

scvi\_integration.py

1环境建设不顺利，使用选项界面和使用conda或者pip的效果都不好，而且每次安装还需另外开容器检验，很费时，效率很低；

2打开容器太慢，输出结果也很满

3看到云服务器环境即是python，又是R，可以只使用一个R或python吗

4现在能用的plant\_omicsv3.3的R使用的还可以，python问题很多，建环境问题很多

4环境建好理应使用环境重新测试脚本，那些脚本和输入源是否已经传输到云平台

5要实现work.sh是不是要使用流程，包括流程中的那些嵌套脚本是否已经完全准备到云平台

2024/8/8

WDL（Workflow Description Language）是一种用于描述和执行复杂科学工作流的领域特定语言。WDL的语言逻辑基于以下几个主要概念：

Task（任务）：任务是工作流的最小单位，用于执行特定的操作或计算。每个任务都有输入和输出，并定义了任务的执行命令。

Input（输入）：输入是任务所需的数据或参数。输入可以是文件、字符串、整数等等。输入可以在任务级别和工作流级别进行定义。

Output（输出）：输出是任务的结果或生成的文件。输出可以是文件、字符串、整数等等。输出可以在任务级别和工作流级别进行定义。

Command（命令）：命令是任务中实际执行的操作或计算。命令可以是任何可执行的外部工具或脚本。命令可以使用输入数据和参数，并生成输出数据。

Workflow（工作流）：工作流是任务的组合，按照特定的顺序和依赖关系进行执行。工作流定义了任务之间的数据流和依赖关系。

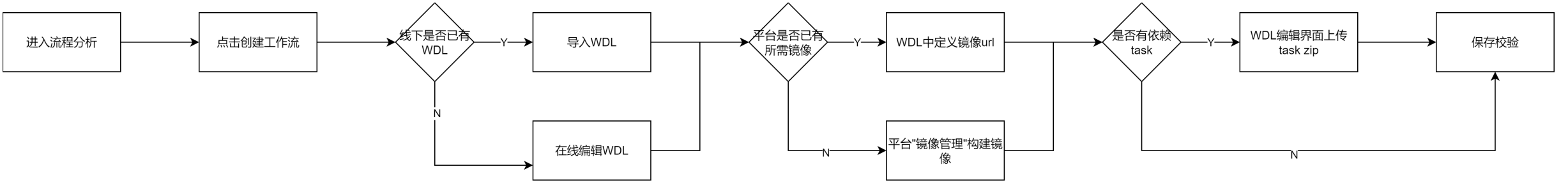
Call（调用）：调用用于在工作流中执行任务。通过调用，可以将输入数据和参数传递给任务，并将任务的输出连接到其他任务。

Conditional（条件语句）：条件语句用于根据特定的条件来决定工作流的执行路径。可以使用if语句和boolean表达式来实现条件逻辑。

Parallelization（并行处理）：WDL支持并行处理，可以同时执行多个任务。可以使用scatter语句将输入分割成多个部分，并使用gather语句将结果合并。

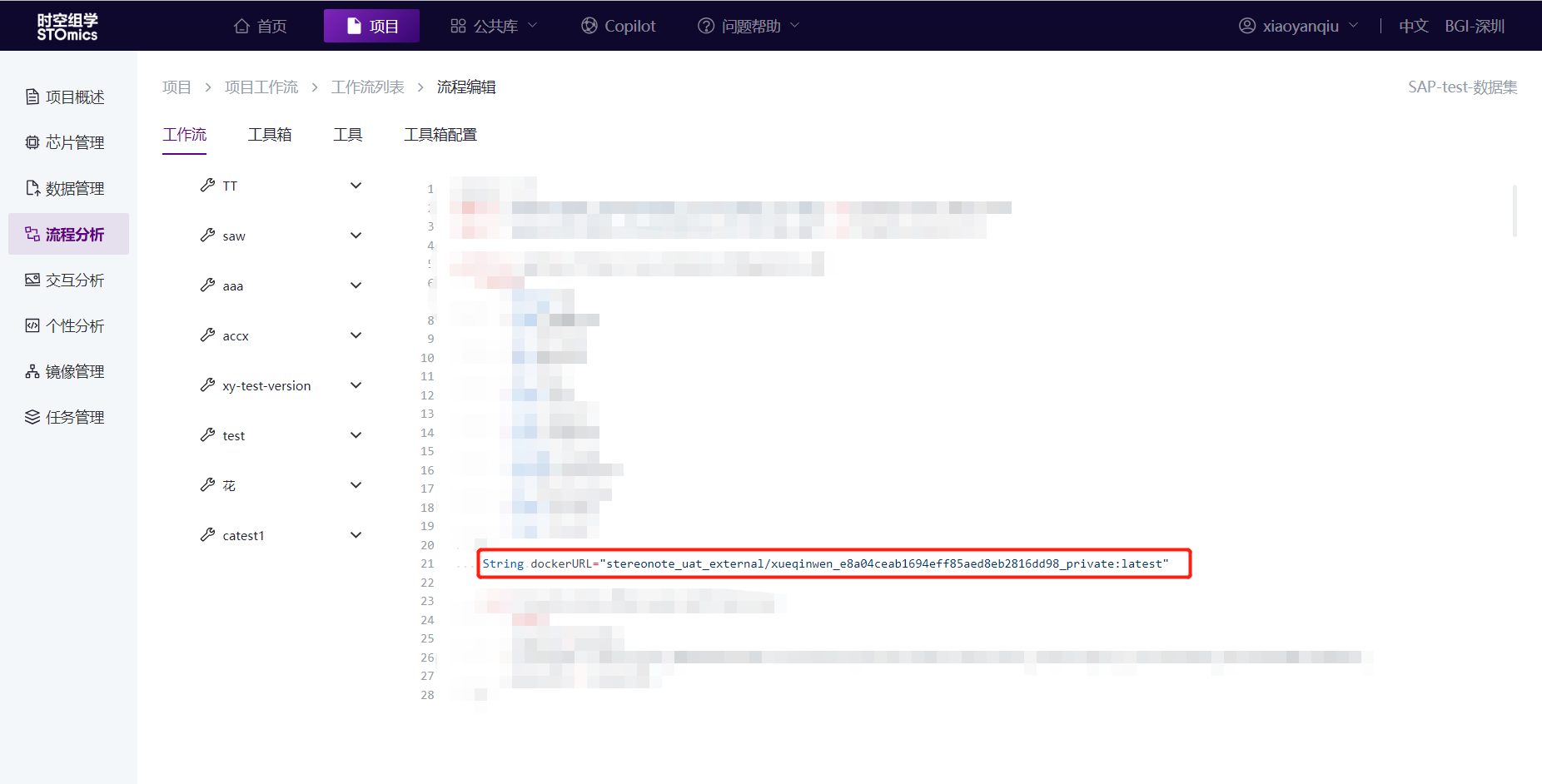
Parameterization（参数化）：WDL支持参数化，可以通过在执行工作流时提供不同的参数值来重用同一个工作流定义。

总体而言，WDL的语言逻辑是基于任务、输入、输出、命令和工作流的组合，以及条件、并行处理和参数化等机制来描述和执行复杂的科学工作流。通过定义任务和任务之间的依赖关系，WDL可以实现复杂的数据处理和分析流程。



# 调用环境

String dockerURL=”stereonote\_hpc/yangdong\_91fc6b8c4f3c4e77b46ba5853b0261b2\_private:latest”



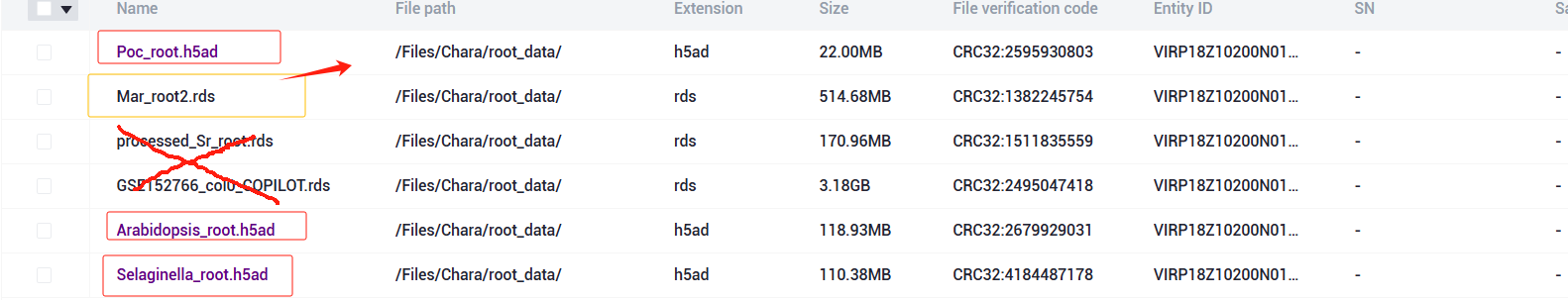
2024/8/12

首先是那些输入文件，pre-processe.py的运行，传参

第一个问题，我们在干嘛？现在有root, stem, leaf的数据，我们现在是分析root吗？

第二个问题，确定分析root之后，我们输入的文件被锁定在root\_dir, 我们发现根有.h5ad和.rds的文件，为什么有这样的差别，为什么要分析这些物种，这些物种的联系在哪里，分析这些物种的意义又在于什么？

第三个问题，对于pre-processed的认识，这一部的意义是什么，其中的参数似乎不太充足



Poc\_root.h5ad

Mar\_root2.rds

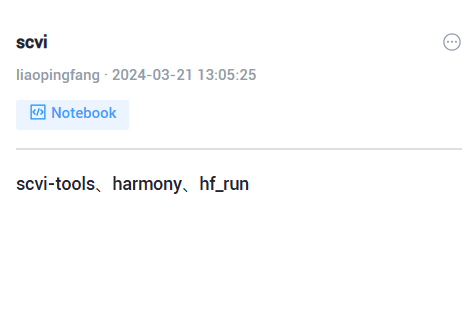
Arabidopsis\_root.h5ad

Selaginella\_root.h5ad

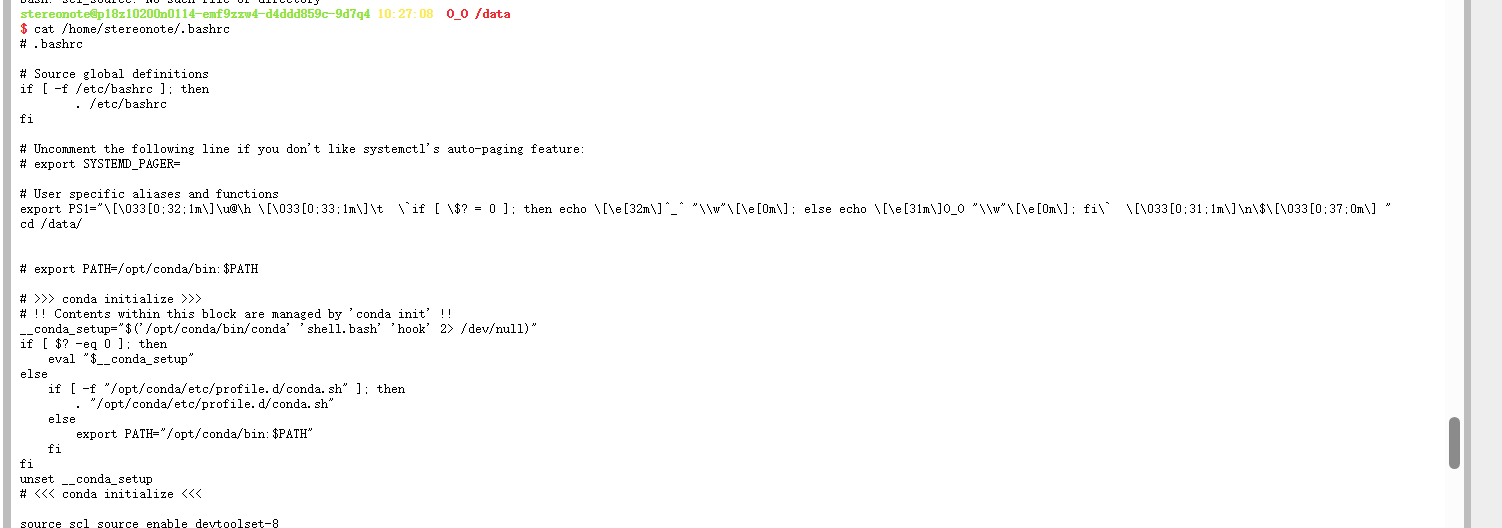
2024/8/15

1完成convert

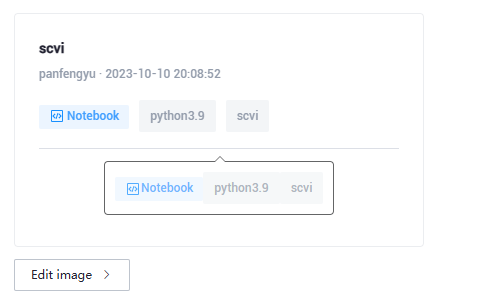
2离线测试scvi等



无skimics



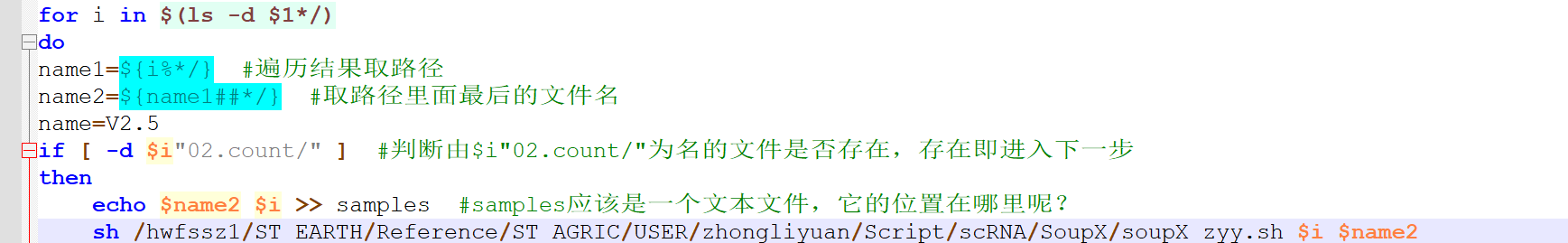
为什么BENGAL--concat可以调用环境，而这个就不可以呢？



Type error subscriptable

ImportError: numpy.core.umath failed to import

Scvi-tools



for i in $(ls -d $1\*/)

do

name1=${i%\*/}

name2=${name1##\*/}

name=V2.5

if [ -d $i"02.count/" ]

then

echo $i >> samples

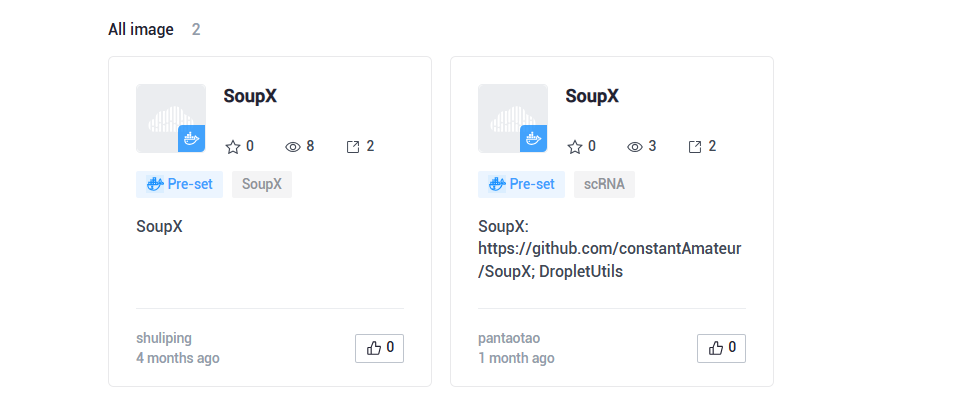
echo $name2 $i >> samples1

sed -i 's/ /\t/g' samples1

echo $name2 $PWD/"$name2"\_soupX\_matrix >> samples2

sed -i 's/ /\t/g' samples2

用python编写



第一个缺library(DropletUtils)和library(getopt)

第二个可以用，包都是齐全的

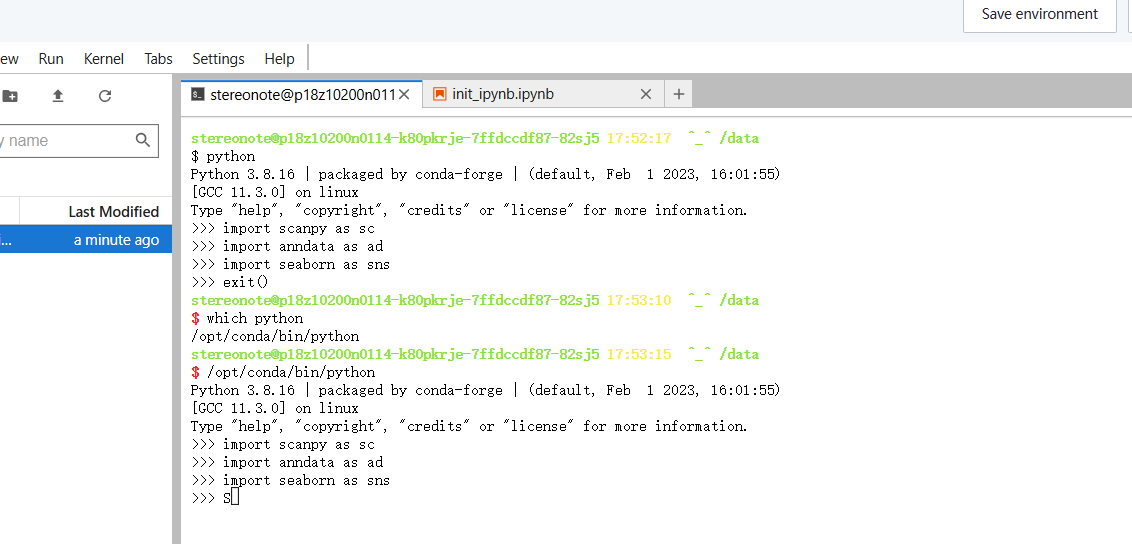
W202408140006977

V2.5R2406140001

import scanpy as sc

import anndata as ad

import seaborn as sns



/opt/conda/bin/python

