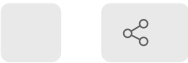




May 23, 2022

# Assembly\_Template

ann.mccartney<sup>1</sup>  
<sup>1</sup>national institute of health



protocol .

ann.mccartney

ERGA\_Assembly\_Strategy

ann.mccartney 2022. Assembly\_Template. **protocols.io**  
<https://protocols.io/view/assembly-template-b9tsr6ne>



protocol ,

May 23, 2022

May 23, 2022

63058

Assembly and Curation 2d

## 1 Data Download

ERGA Nextcloud REP  
Brand  
Barcelona Supercomputer xx63

## 2 Data cleaning

Run Filtlong

**./filtlong -parameters -input -output**

### 3 Assembly of long reads

2d

**canu 1.7.1** [↗](#)

Linux / Darwin

[source](#) by Maryland Bioinformatics Labs

🕒 **48:00:00** Had to stop and restart due to ran out of memory

Best to have a checkpointing system in your PBS script.

Run assembly

**Exact command**

Command used to run canu

macintosh 10

### 4 Scaffolding with HiC

**salsa**

Run Salsa

**./salsa -input -output**

## 5 Polishing with Illumina

**minimap2 2.16** [↗](#)

Linux

[source](#) by Heng Li

**racon 10**

[source](#) by ivan sovic

Run racon

**./racon -parameters -input -output**

## 6 QC

**busco**

Run Busco

**./busco -input -output**

**meryl**

## 7 Curation

**grit**



NCBI data curation protocol  
by Ruth Timme,  
US Food and Drug Administration

PREVIEW

RUN

