NOV 09, 2023

DOI:
dx.doi.org/10.17504/protocols.io.kqdg3xeqqg25/v1

Protocol Citation: zehnpjr 2023. Mitochondrial Genome Assembly and Annotation. **protocols.io** https://dx.doi.org/10.17504/protocols.io.kqdg3xeqqg25/v1

Protocol status: Working
We use this protocol and it has worked for us in the past.

Created: Jun 09, 2023

Last Modified: Nov 09, 2023

Mitochondrial Genome Assembly and Annotation

zehnpjr[1]

[1]Central Michigan University

zehnpjr

DISCLAIMER

I have used the protocol, and it has worked for me, however I am not guaranteeing that this protocol will work.

ABSTRACT

This is a step by step protocol for assembling invertebrate mitochondrial genomes, annotating the genomes , publishing the genomes and then step by step protocol for making a phylogenetic tree using the data from the mitochondrial genomes

PROTOCOL integer ID:
83144

Funders Acknowledgement:
NSF
Grant ID: ANT-1043745
NSF
Grant ID: OPP-0132032

Novoplasty v4.3.1

1 After QC, we ran our illumina short reads through a program called Novoplasty v.4.3.1. for mitochondrial genome assembly.

CITATION

Dierckxsens N, Mardulyn P, Smits G (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data..
LINK
https://doi.org/10.1093/nar/gkw955

Next you want to upload a seed and a reference file. The seed is where you would want the assembler to start assembling the program. I used the COI of an already published mitochondrial genome. Next you will want to upload the reference genome. I used the already published mitochondrial genome of the sea spider Achelia bituberculata.
Then you want to create a config.txt file for your assembly. It should loo like this:

Project:

Project name= A_M
Type= mito
Genome Range= 12000-22000
K-mer = 33
Max memory=
Extended log= 0
Save assembled reads= no
Seed Input= /mnt/home/zehnp1jrCMICH/NOVOPlasty/NOVOPlasty/Seed.fasta
Extend seed directly= no
Reference sequence= /mnt/home/zehnp1jrCMICH/NOVOPlasty/NOVOPlasty/reference.fasta
Variance detection=
Chloroplast sequence=

Dataset 1:

Read Length = 151
Insert size = 300
Platform= illumina
Single/Paired = PE
Combined reads=
Forward reads = /mnt/home/zehnp1jrCMICH/NOVOPlasty/NOVOPlasty/M_R1.fastq
Reverse reads = /mnt/home/zehnp1jrCMICH/NOVOPlasty/NOVOPlasty/M_R2.fastq

Heteroplasmy:

MAF =
HP exclude list =
PCR-free=

Optional:

Insert size auto= yes
Use Quality Scores= no

Then you will want to create a submission file (.sh file) it should look like this:

```
#!/bin/sh --login
#SBATCH --time=5:00:00
#SBATCH --nodes=1
#SBATCH --cpus-per-task=20
#SBATCH --mem=250G
#SBATCH --mail-user=zehnp1jr@cmich.edu
#SBATCH --mail-type=FAIL,BEGIN,END
#SBATCH -J A_Mitochondrial



cd ${SLURM_SUBMIT_DIR}

perl NOVOPlasty4.3.1.pl -c config.txt
```

scontrol show job $SLURM_JOB_ID

Then you want to submit the job. If using a cluster its a good idea to submit as a job rather than run in the terminal as you will most likely run out of resources and your task will be terminated in the terminal.

2 Submit consensus file from Artemis to MITOS for Annotation.

CITATION

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013). MITOS: improved de novo metazoan mitochondrial genome annotation..
LINK
https://doi.org/10.1016/j.ympev.2012.08.023

This will be the .fasta file from NOVOPLasty.

3 Next you will want to take annotations from MITOS2 and manually check gene boundaries in Artemis

CITATION

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B (2000). Artemis: sequence visualization and annotation..
LINK
https://doi.org/

4 Ammothea_clausi_example.tbl Mitogenome_notes_aftermitos (1).txt
Next I followed these steps for my submission to NCBI.

5 Phylogenetic tree building with protein coding gene data Phylogenetic_tree_notes.txt