#### HONOURS 2021 | STATS PIPELINE

# FUNNELING

#### RAW FILES

Each finished job will produce 50,000 files at the individual level and 2 files at the model level.

Values of all individuals/generation Fixed Mutations files SLiM Full Output files

### STATS\_ONE.R

Reads in the raw files and averages all values across the individuals. Smallest unit is now the model seed (replicate ID).

Reads in the mutation + output file and calculates allelic frequencies.

Graphs for one model run (seed level)
BigPopa\_onerun files ( collated data)
Tetris\_onerun files (mean values + µ)

## TANK.PBS TANK.PS1

Stats\_one.R needs to be run for each node folder of each modeltype and optima combination; this script collects the final output and transfers it into a folder called tank for the next script to work on. Tank is situated in Modeltype/Optima.

#### STATS-REPS.R

Reads in the files in the tank folder
(BigPopa and Tetris files) and averages
all values across the seeds. Smallest
unit is now the unique parameter
combination, given by the combination

Graphs for one param combination
(multiple seeds)
PogNoodle files (collated data)
PogNoodled files (mean across seeds)

## AQUARIUM.PBS AQUARIUM.PS1

Stats\_model.R needs to be run for each model type and optima combination, so the PogNoodle and PogNoodled files are moved to a common folder at the root of the optima branch. Move from node level organization to optima level organization.

#### STATS MODEL R

Reads in the files in the aquarium folder (PogNoodle and PogNoodled) and averages all values across combinations. The resulting output contains the mean adaptive walk, mean allelic frequencies and final data on a model (optima x modeltype) level.

Graphs for a model (type x optima)
Shelldon files (collated data)
Richard files (mean across combos)

## OCEAN.PBS OCEAN.PS1

Unlike previous pbs/ps1 scripts, this script copies required files for stats-types.r, as this last step breaks the linear flow and can call in any of the previous data sets

### STATS-TYPES.R

Final piece in the pipeline for tests and more sophisticated analysis utilizing any of the previous data sets. Key function is comparision of ADD and ODE models.

Zephyr files

For more details and the pipeline described here, visit www.github.com/sknief/honours