# Starling-May18 Projects/Katarina Stuart/KStuart.Starling-Aug18/Sv10\_NZstarlings/Analysis/2024-05-07.StairwayPlot

PDF Version generated by

Katarina Stuart (z5188231@ad.unsw.edu.au)

on

Sep 25, 2024 @12:48 PM AEST

## **Table of Contents**

2024-05-07.StairwayPlot



## **Stairway Plot 2**

https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02196-9

https://academic.oup.com/icb/article/62/6/1849/6698719#389865038

https://github.com/philippinespire/PIRE\_TemporalSims\_HPC/blob/main/stairscripts/stairscript\_n20n50n100.sh

https://evomics.org/wp-content/uploads/2020/01/WoG PopGenTutorial1 DemographicInference1.pdf

https://github.com/AlexLewanski/petrochromis\_project/tree/main/project\_scripts

Manually downloaded Stairway jar files and placed them in the programs folder program needs to be run in this directory

cd /nesi/nobackup/uoa02613/kstuart projects/programs/stairway plot v2.1.1/

## downloading some scripts for SFS

git clone https://github.com/shenglin-liu/vcf2sfs.git

 $\label{lem:cd/nesi/nobackup/uoa02613/kstuart_projects/programs \\ \mbox{git clone } \mbox{\underline{https://github.com/marqueda/SFS-scripts.git} \# \mbox{didn't like this one. Kept erroring out for my pop file : (} \mbox{}$ 

filter VCF

cd /nesi/nobackup/uoa02613/kstuart\_projects/Sv10\_NZstarlings/analysis/stairway

module load VCFtools/0.1.15-GCC-9.2.0-Perl-5.30.1

DIR=/nesi/nobackup/uoa02613/kstuart\_projects/Sv10\_NZstarlings/data/processing\_rawdata/BCFtools/final\_filtering/persite\_missingness\_filter VCF=\${DIR}/starling\_noduprel\_qual\_miss.recode.vcf

 $cat $\{DIR\}/indlist\_PLM.txt $\{DIR\}/indlist\_UHT.txt > indlist\_PLMUHT.txt \\ cat $\{DIR\}/indlist\_MKW.txt $\{DIR\}/indlist\_NWC.txt > indlist\_MKWNWC.txt \\ cat $\{DIR\}/indlist\_CAN.txt $\{DIR\}/indlist\_UHT.txt > indlist\_CAN.txt \\ }$ 

vcftools --vcf \${VCF} --keep indlist\_PLMUHT.txt --max-missing 0.7 --max-alleles 2 --min-alleles 2 --max-meanDP 100 --thin 1000 --recode --recode-INFO-all --out starling\_noduprel\_qual\_miss\_filt\_PLMUHT

vcftools --vcf \${VCF} --keep indlist\_MKWNWC.txt --max-missing 0.7 --max-alleles 2 --min-alleles 2 --max-meanDP 100 --thin 1000 --recode --recode-INFO-all --out starling\_noduprel\_qual\_miss\_filt\_MKWNWC

vcftools --vcf \${VCF} --keep indlist\_MKWNWC.txt --max-missing 0.7 --max-alleles 2 --min-alleles 2 --max-meanDP 100 --recode --recode-INFO-all --out starling\_noduprel\_qual\_miss\_filt\_MKWNWC\_nothin

vcftools --vcf \${VCF} --keep indlist\_CAN.txt --max-missing 0.7 --max-alleles 2 --min-alleles 2 --max-meanDP 100 --thin 1000 --recode --recode-INFO-all --out starling\_noduprel\_qual\_miss\_filt\_CAN

UK: After filtering, kept 25365 out of a possible 68262 Sites

NZ: After filtering, kept 25283 out of a possible 68262 Sites

## Need to grab info for blueprint file

```
module load R/4.1.0-gimkl-2020a R

source("/nesi/nobackup/uoa02613/kstuart_projects/programs/vcf2sfs/vcf2sfs.r")

mygt<-
vcf2gf("/nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway/starling_noduprel_qual_miss_filt_PLMUHT.recode.vcf",
"/nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway/indlist_PLMUHT_POP.txt")

mysfs1<_fold<- fold.sfs(mysfs.1)
write.1D.fsc(mysfs1_fold,"/nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway/PLMUHT.sfs")

#

mygt<-
vcf2gf("/nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway/starling_noduprel_qual_miss_filt_MKWNWC.recode.vcf",
"/nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway/indlist_MKWNWC_POP.txt")

mysfs1<-gt2sfs.raw(mygt, "UK")
mysfs1<-gt2sfs.raw(mygt, "UK")
mysfs1<-fold<-fold.sfs(mysfs1)
write.1D.fsc(mysfs1_fold,"/nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway/MKWNWC.sfs")
```

## Changes made to blueprint file:

```
cd /nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway

#manipulate output of sfs file for input into the edited blueprint file
sfs=$(cat MKWNWC_nozero.sfs)

#move to stairway folder
cd /nesi/nobackup/uoa02613/kstuart_projects/programs/stairway_plot_v2.1.1/

#UK

#edit the blueprint file and save a new copy
sed -e 's/popid: two-epoch_fold/popid: UK/g'\
-e 's/nseq: 30/nseq: 44/g'\
-e 's/nseq: 30/nseq: 44/g'\
-e 's/L: 10000000/L: 25531/g'\
-e 's/estimation: 15/estimation: 22/g'\
-e 's/mu: 1.2e-8/mu: 4.6e-9/g'\
-e 's/generation: 24/generation: 2/g'\
-e 's/groject_dir: two-epoch_fold/project_dir: Sv10_UK/g'\
-e 's/project_dir: two-epoch_fold/project_dir: Sv10_UK/g'\
```

```
-e "7s/.*/SFS:${sfs}/g" \
two-epoch_fold.blueprint > Sv10_UK.blueprint
#nrand: 10
               21
                     31
                            42
#run make bash
java -cp stairway plot es Stairbuilder Sv10 UK.blueprint
#UK
#edit the blueprint file and save a new copy
sed -e 's/popid: two-epoch_fold/popid: UK/g' \
-e 's/nseq: 30/nseq: 44/g' \
-e 's/L: 10000000/L: 25531/g' \
-e 's/estimation: 15/estimation: 22/g' \
-e 's/mu: 1.2e-8/mu: 4.6e-9/g' \
-e 's/generation: 24/generation: 2/g' \
-e 's/xrange: 0.1,10000/xrange: 0.1,1000/g' \
-e 's/project_dir: two-epoch_fold/project_dir: Sv10_UK2/g' \
-e "7s/.*/SFS:${sfs}/g" \
two-epoch fold.blueprint > Sv10 UK2.blueprint
#run make bash
java -cp stairway_plot_es Stairbuilder Sv10_UK2.blueprint
cd /nesi/nobackup/uoa02613/kstuart_projects/Sv10_NZstarlings/analysis/stairway
#manipulate output of sfs file for input into the edited blueprint file
sfs=$(cat PLMUHT_nozero.sfs)
#move to stairway folder
cd /nesi/nobackup/uoa02613/kstuart projects/programs/stairway plot v2.1.1/
#NZ
#edit the blueprint file and save a new copy
sed -e 's/popid: two-epoch_fold/popid: NZ/g' \
-e 's/nseq: 30/nseq: 44/g' \
-e 's/L: 10000000/L: 25283/g' \
-e 's/estimation: 15/estimation: 22/g' \
-e 's/mu: 1.2e-8/mu: 4.6e-9/g' \
-e 's/generation: 24/generation: 2/g' \
-e 's/xrange: 0.1,10000/xrange: 0.1,1000/g' \
-e 's/project_dir: two-epoch_fold/project_dir: Sv10_NZ/g' \
-e "7s/.*/SFS:${sfs}/g" \
two-epoch_fold.blueprint > Sv10_NZ.blueprint
#run make bash
java -cp stairway_plot_es Stairbuilder Sv10_NZ.blueprint
```

## Now run the generated bash script

#!/bin/bash -e

```
#SBATCH --job-name=2023_05_07.stairway_run.sl
#SBATCH --account=uoa02613
#SBATCH --time=00-12:00:00
#SBATCH --mem=25GB
#SBATCH --output=%x_%j.errout
#SBATCH --mail-user=katarina.stuart@auckland.ac.nz
#SBATCH --mail-type=ALL
#SBATCH --mail-type=ALL
#SBATCH --nodes=1
#SBATCH --notes=1
#SBATCH --ropide task

cd /nesi/nobackup/uoa02613/kstuart_projects/programs/stairway_plot_v2.1.1/

bash Sv10_UK.blueprint.sh
bash Sv10_UK.blueprint.sh
bash Sv10_UK2.blueprint.sh
```

## run plotting script

 $\verb|cd/nesi/nobackup/uoa02613/kstuart_projects/programs/stairway_plot\_v2.1.1/|$ 

bash Sv10\_UK.blueprint.plot.sh bash Sv10\_UK2.blueprint.plot.sh

bash Sv10\_NZ.blueprint.plot.sh