

# Formulas used for the Analysis Script of CROPMETAPOP

## Sensibility Analysis

Baptiste ROUGER

November 24, 2017

The script has been to work with the GenotypeMonoLocus.csv file produced by CROPMETAPOP. It works with 2 alleles for each gene so the formulas presented here are for 2 alleles only. We computed the indices of the marker under selection and under no selection separately but with the same formulas.

## 1 Locus level

### 1.1 Hs

#### 1.1.1 For each locus : $Hs_l$

$$Hs_l = 1 - (p_l^2 + q_l^2)$$

With the following formulas for  $p_l$  and  $q_l$  :

$$p_l = \frac{2N_{0/0} + N_{0/1}}{2N}$$

$$q_l = \frac{N_{0/1} + 2N_{1/1}}{2N}$$

$N_{0/0}$  represents the number of homozygotes individuals 0/0,  $N_{1/1}$  the number of homozygotes 1/1,  $N_{0/1}$  the number of heterozygotes, and  $N$  the total number of individuals.

#### 1.1.2 Mean of the loci : $Hs$

$$Hs = \frac{1}{L} \sum_{l=1}^L Hs_l$$

With  $L$  the number of locus (under selection or under no selection)

### 1.2 Hobs

#### 1.2.1 For each locus : $Hobs_l$

$$Hobs_l = \frac{N_{0/1}}{N}$$

$N_{0/1}$  the number of heterozygotes and  $N$  the total number of individuals.

#### 1.2.2 Mean of the loci : $Hobs$

$$Hobs = \frac{1}{L} \sum_{l=1}^L LHobs_l$$

With  $L$  the number of locus (under selection or under no selection)

### 1.3 Fis

#### 1.3.1 For each locus : $Fis_l$

$$Fis_l = 1 - \frac{Hobs_l}{2p_lq_l}$$

#### 1.3.2 Mean of the loci : $Fis$

$$Fis = \frac{1}{L} \sum_{l=1}^L Fis_l$$

With  $L$  the number of locus (under selection or under no selection).

## 2 Population level

### 2.1 $\overline{Hs}$

$$\overline{Hs} = \frac{1}{S} \sum_{s=1}^S Hs$$

With  $S$  the number of populations.

We also compute the variance of  $Hs$  with the `statistics.py` library.

### 2.2 $\overline{Hobs}$

$$\overline{Hobs} = \frac{1}{S} \sum_{s=1}^S Hobs$$

With  $S$  the number of populations.

We also compute the variance of  $Hobs$  with the `statistics.py` library.

### 2.3 $\overline{Fis}$

$$\overline{Fis} = \frac{1}{S} \sum_{s=1}^S Fis$$

With  $S$  the number of populations.

We also compute the variance of  $Fis$  with the `statistics.py` library.

## 3 Metapopulation level

### 3.1 Ht

#### 3.1.1 Ht for each locus

$$Ht_l = 1 - (\overline{p}_l^2 + \overline{q}_l^2)$$

With the following formulas for  $\overline{p}_l$  and  $\overline{q}_l$  :

$$\overline{p}_l = \frac{1}{S} \sum_{s=1}^S p_l$$

$$\overline{q}_l = \frac{1}{S} \sum_{s=1}^S q_l$$

With  $S$  the number of populations.

### 3.1.2 Mean of the locus for $Ht$

$$Ht = \frac{1}{L} \sum_{l=1}^L Ht_l$$

With  $L$  the number of loci (under selection or under no selection)

### 3.2 $Gst$

$$Gst = \frac{Ht - \overline{Hs}}{Ht}$$