# Project: Statistical Analysis Genomics

**Introduction:**

*The Goal of the analysis*: For ten different loci**, find if the genotype of a given locus is linked to**

**the observed phenotype.**

*Available data*: The phenotype for fly is , a male sine song frequency measured in .

At a particular locus, we also have the genotype of the fly.

*Hypotheses:*

**Null QTL** (Quantitative Trait Loci) Hypothesis: the genotype does not explain the phenotype:

**QTL Hypothesis**: the genotype does explain the phenotype:

*Decision Rule:*

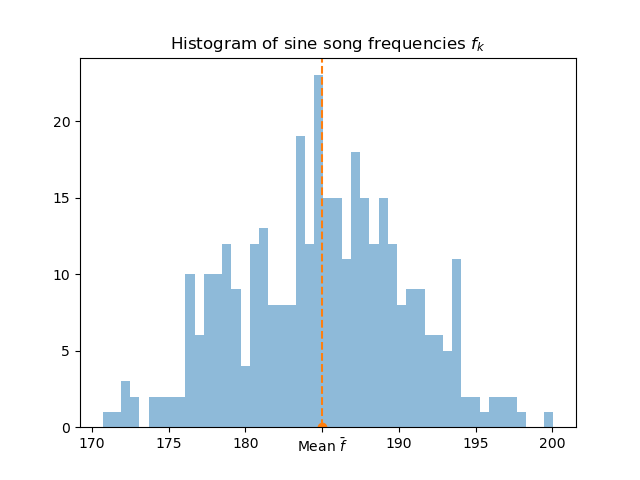
If Ratio > > 1, Then QTL Hypothesis is very Plausible

If Ratio < < 1, Then NQTL Hypothesis is very Plausible.

We recall the expression Ratio derived in lectures:

*(Please visit the Python Notebook associated with this file to see detailed calculations and code of this project.)*

**Task: 1: histogram of sine song frequencies for all backcross males.**

****

**Task: 2: For each locus, compare the two hypotheses and decide whether it is linked to the**

**phenotype or not.**

1. Results our QTL analysis:

The ratio statistics for the 10 loci are:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Locus1 | Locus2 | Locus3 | Locus4 | Locus5 | Locus6 | Locus7 | Locus8 | Locus9 | Locus10 |
| 4.0752 | 3.7519 | 18.5871 | **8.373e+12** | 3.6959 | 5.1475 | 2025.6224 | 3.5706 | 4.5207e+11 | 3.4732 |

**Conclusion**

From these results we can conclude that **Locus4** and **Locus9, are definitely associated the sine song frequencies** of malesDrosophila simulans and Drosophila mauritiana phenotypes*.*

As for the other Loci we do not have enough evidence from data, to assert that any of them is linked with the observed phenotype.