

# Exercises in Marine Ecological Genetics

## 02. Hardy-Weinberg equilibrium

- Calculate allele and genotype frequencies
- Test for Hardy-Weinberg equilibrium
- Work with microsatellite data / Genepop format

Martin Helmkamp

<https://github.com/mhelmkampf/meg25>



# Course outline

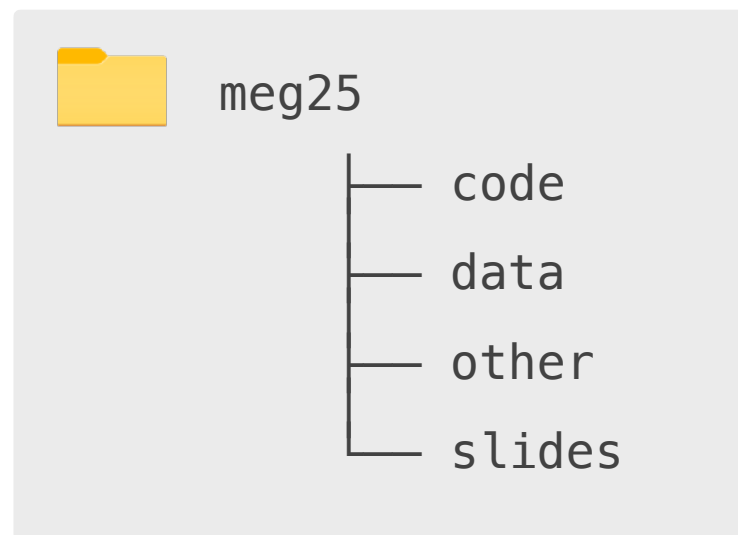
Preliminary, may be subject to change

Class	Date	Topics	Script
01	Apr 11	Introduction, setup	01_intro.R
–	Apr 18	Good Friday	
02	Apr 25	Hardy-Weinberg equilibrium / $N_e$	02_hwe.R
03	May 02	Population structure and gene flow	
04	May 09	Genome assembly and metrics	
05	May 16	SNPs and population genomics	
06	May 23	Recombination and linkage disequilibrium	
–	May 30	Himmelfahrt break	
07	Jun 06	Measures of genetic diversity	
–	Jun 13	Selection and Mutation	
08	Jun 20	Student presentations – no exercises	
09	Jun 27	DNA barcoding	
10	Jul 04	Metabarcoding / eDNA	
11	Jul 11	Metabarcoding / eDNA II	

# Accessing the course materials

All code and data for the course will be provided through a Git repository:

<https://github.com/mhelmkamp/meg25>



# Working with the command line / bash

Recap

## Windows

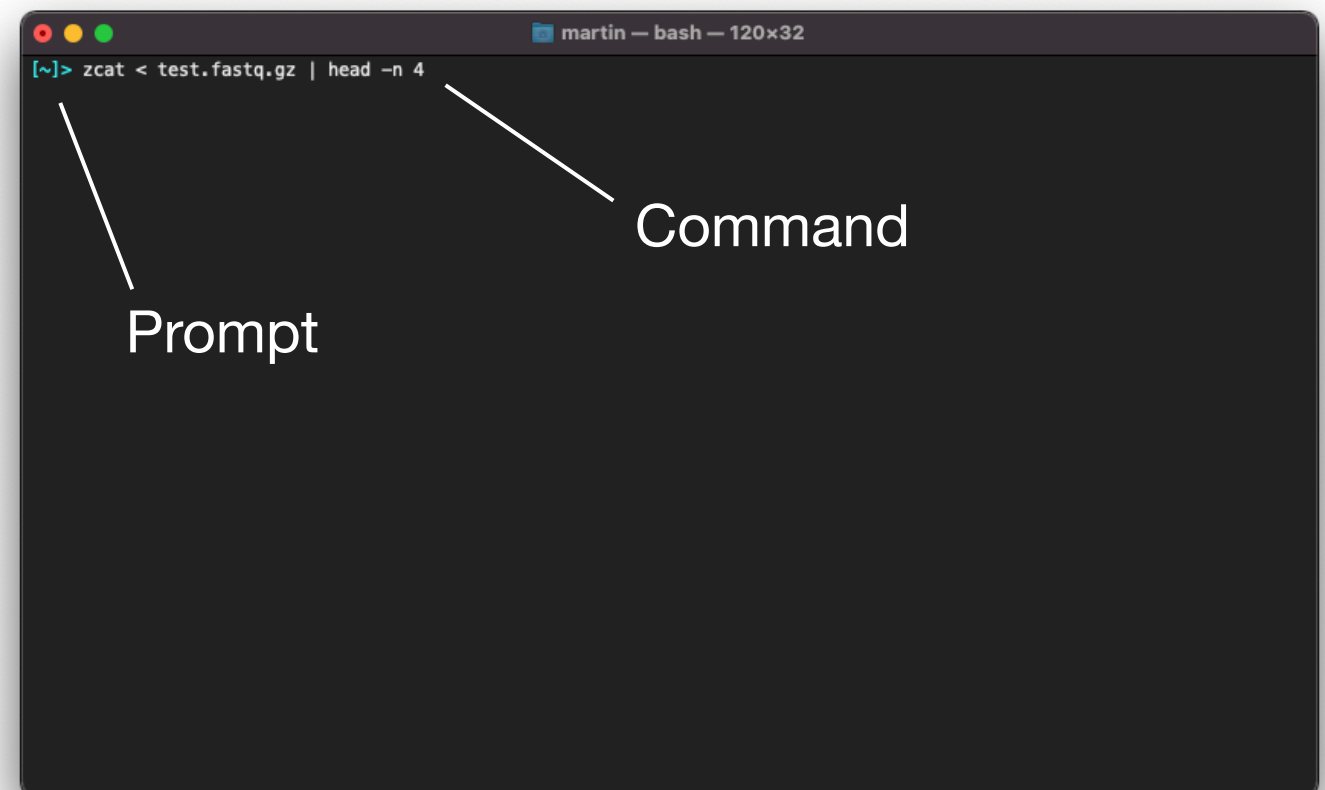
- Launch git bash from Start menu
- Alternatively, install Windows Subsystem for Linux (WSL) on Windows 10 or above (see <https://learn.microsoft.com/en-us/windows/wsl/install>)

## macOS

- Open Terminal app in /Applications/Utilities

## Typical usage

`command [-options] [file]`



# Basic bash commands

## Recap

```
ls                # List directory contents
ls -l            # List directory contents in long format

cd                # Change into home directory
cd <directory_name> # Change into directory
cd ..            # Change into directory above / back

pwd              # Print current (working) directory
```

```
tab key          # Autocomplete file or directory name
up arrow key     # Cycle through previous commands
```

# Basic R commands

## Recap

```
y <- ((6 * 2) + 1) / 20      # Assign value to variable
y                             # Recall
[1] 0.65

c(...)                       # Combine elements into vector
data.frame(...)              # Create data frame (table)

# Example
dat <- data.frame(row.names = c("yy", "yb", "bb"),
                  "observed" = c(6, 1, 3),
                  "expected" = c(yy_e, yb_e, bb_e)
)

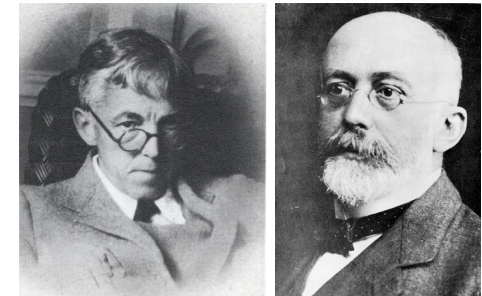
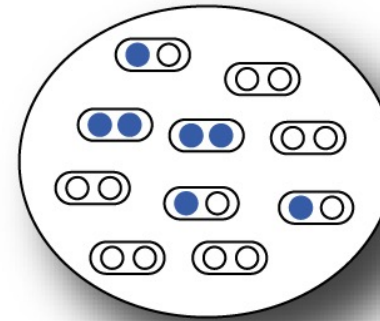
help(function_name)          # Show help page for function
```



## HARDY-WEINBERG (1908)

Godfrey H. Hardy (1877-1947)

Wilhelm Weinberg (1862-1937)

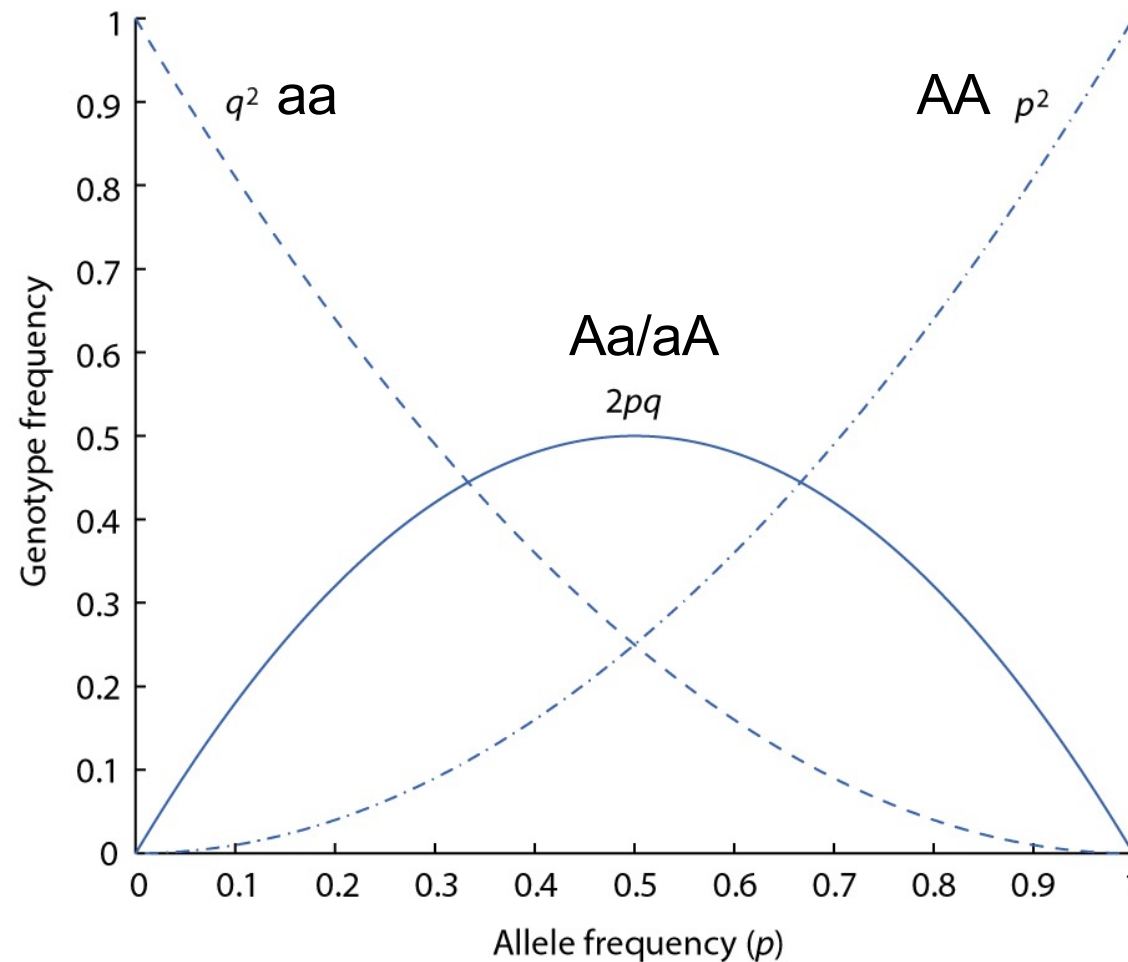


*Establish the relationship between allele frequencies and genotype frequencies in a population*

$$p^2 + 2pq + q^2 = 1$$

AA    Aa/aA    aa

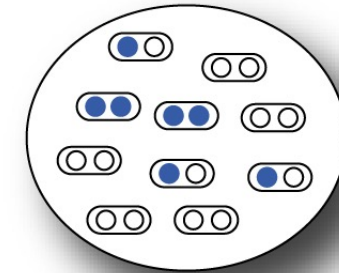
$p$  and  $q$ : allele frequencies for a locus with two alleles (A and a)  
( $p + q = 1$ )



*A single generation of reproduction will result in a population that meets the expected Hardy-Weinberg frequencies, i.e. is at Hardy-Weinberg (HW) equilibrium*

Assuming an “ideal” population, i.e. :

- Diploid organisms
- Sexual reproduction (as opposed to clonal)
- Random mating (as opposed to e.g. assortative) with respect to genotype
- Random union of gametes
- Discrete, non-overlapping generations
- Very large (infinite) population
- No migration
- No population structure
- No natural selection
- Two alleles
- Identical allele frequencies in both sexes



-> Departures from HW equilibrium may indicate:

- Inbreeding
- Assortative mating
- Self-fertilization
- Natural selection
- Population structure
- ...



# Is this population in Hardy-Weinberg equilibrium?

## Exercise 1

Diploid, sexual reproduction, 1 locus, 2 co-dominant alleles (yellow, blue)



# Genepop format

## Exercises 1/2

Example data

Color

POP

ind01, yy

ind02, yy

ind03, yy

ind04, yy

ind05, yy

ind06, yy

ind07, yb

ind08, bb

ind09, bb

ind10, bb

yellowblue.txt

puella\_barbados.txt

Microsatellite genotypes of Hypoplectrus puella from Barbados

g2

gag010

h24

hyp001

hyp015

hyp018

e2

hyp008

hyp016

pam013

POP

Sample / population	Locus names	Alleles (000 = missing data)
barbados ind#735	203235	121127 210248 225231 126132 190192 157157 236244 194220 114128
barbados ind#736	205217	119119 204216 225229 126126 190192 157165 236236 216222 106144
barbados ind#737	203203	135147 224226 233235 126136 190198 157163 238250 202226 126160
barbados ind#738	211217	119121 228230 223231 126130 190190 157157 240244 220226 138160
barbados ind#739	205225	121125 208216 227231 132132 192192 158159 244256 192196 146146
barbados ind#740	217233	119121 216228 227233 126126 192198 157157 242256 194216 000000
barbados ind#741	203209	000000 216222 223229 126130 192198 157157 234236 188226 118130
barbados ind#742	203215	119119 222234 233233 130130 190190 157157 240246 182188 108126
barbados ind#743	203243	111121 216224 231231 126126 190192 157159 234249 188198 118126
barbados ind#744	203225	121135 206232 231231 126130 192198 157159 240249 182222 122156
barbados ind#745	203211	121123 224236 227231 126132 192192 157163 238238 182204 138148
barbados ind#746	227233	119123 216234 227233 126130 190190 157159 236245 186188 124142
barbados ind#747	203223	133141 236240 231231 126126 192198 157157 234236 194204 118126
barbados ind#748	203217	141145 222226 227235 126126 190192 145157 238240 182196 108108
barbados ind#749	205221	125147 210238 223231 126132 192198 157157 234234 188200 106148
barbados ind#750	203235	123123 230234 223231 126132 192198 157157 236256 188214 122136
barbados ind#751	203211	121143 212248 223229 126132 190196 157159 234242 214216 118128
barbados ind#752	213217	123169 204234 223231 126134 190198 157163 236248 193224 142148

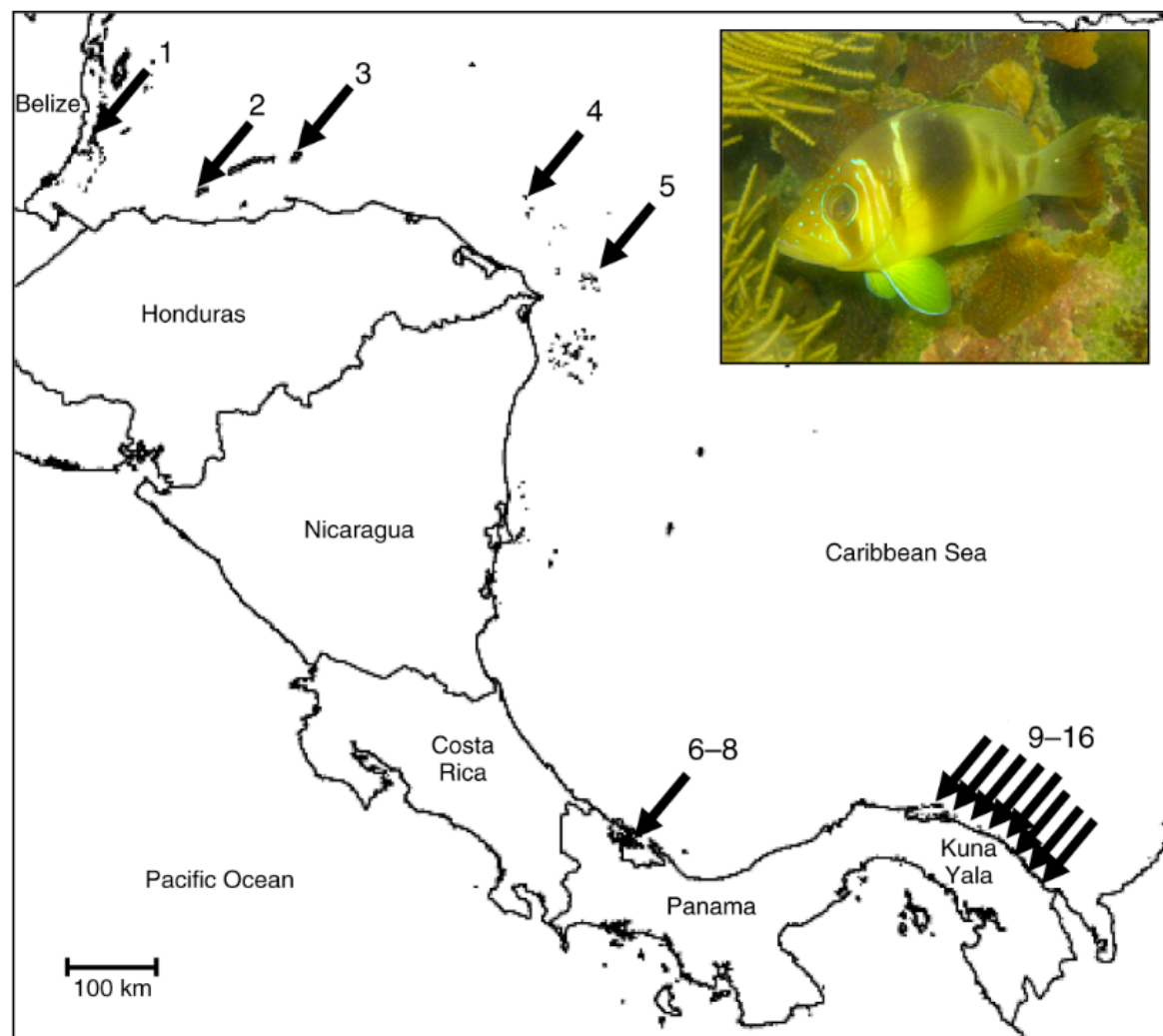
Summer 2025

Exercises in Marine Ecological Genetics  
02. Hardy-Weinberg equilibrium

Carl von Ossietzky  
Universität  
Oldenburg

# Barred hamlet (*H. puella*) microsatellite dataset

## Exercise 2



*Ecology*, 90(11), 2009, pp. 3087–3098  
© 2009 by the Ecological Society of America

Estimating dispersal from genetic isolation by distance  
in a coral reef fish (*Hypoplectrus puella*)

OSCAR PUEBLA,<sup>1,2,3</sup> ELDREDGE BERMINGHAM,<sup>1,2</sup> AND FRÉDÉRIC GUICHARD<sup>2</sup>

- 10 highly variable microsatellite loci
- 854 individuals
- 15 Caribbean locations



# Microsatellites

## Exercise 2

ACAGTTACAGCAGGAAGTTC [...] TATATATATATATATATA [...] CGCTAATG ACAGCACGCTAAC

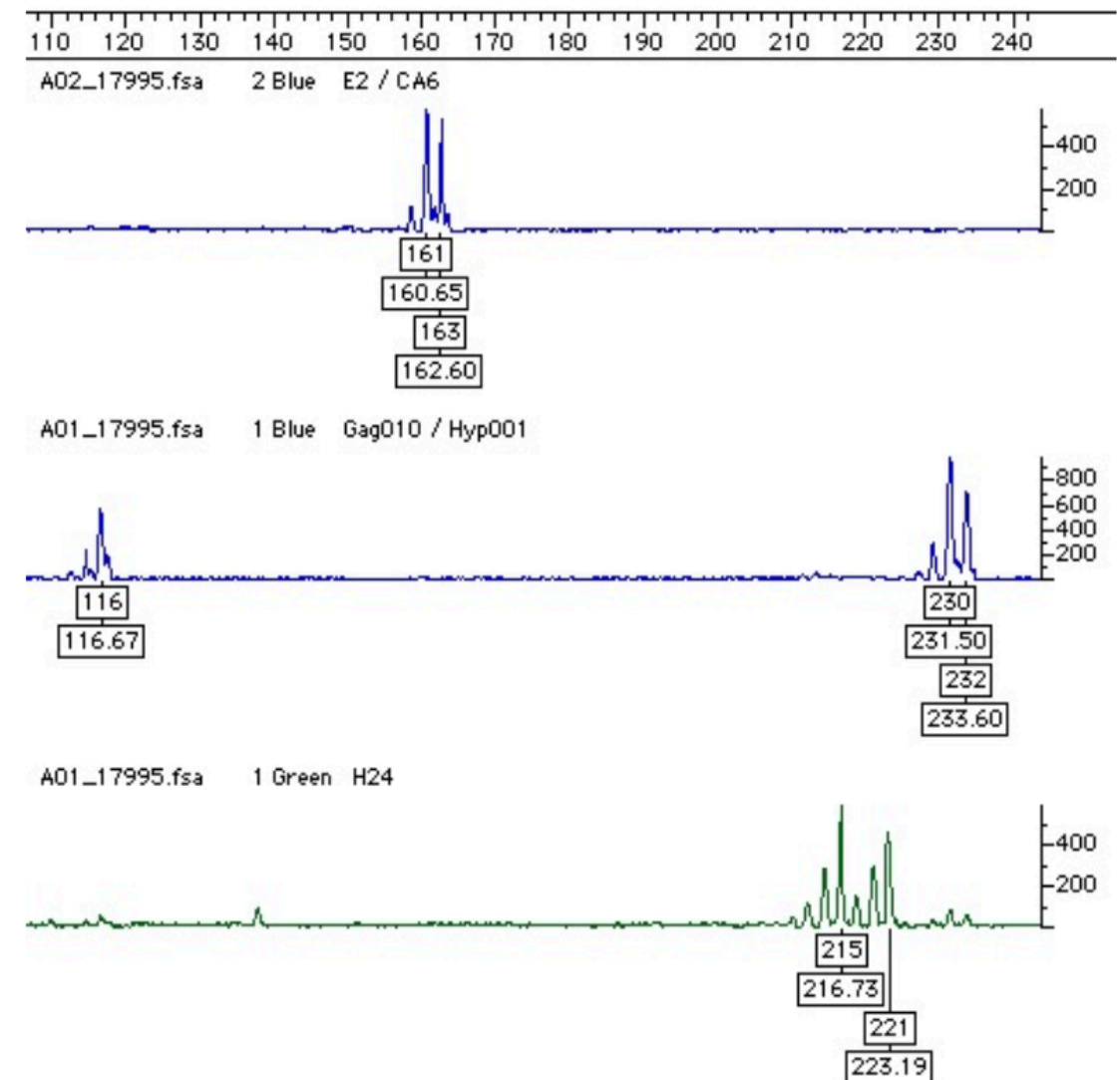
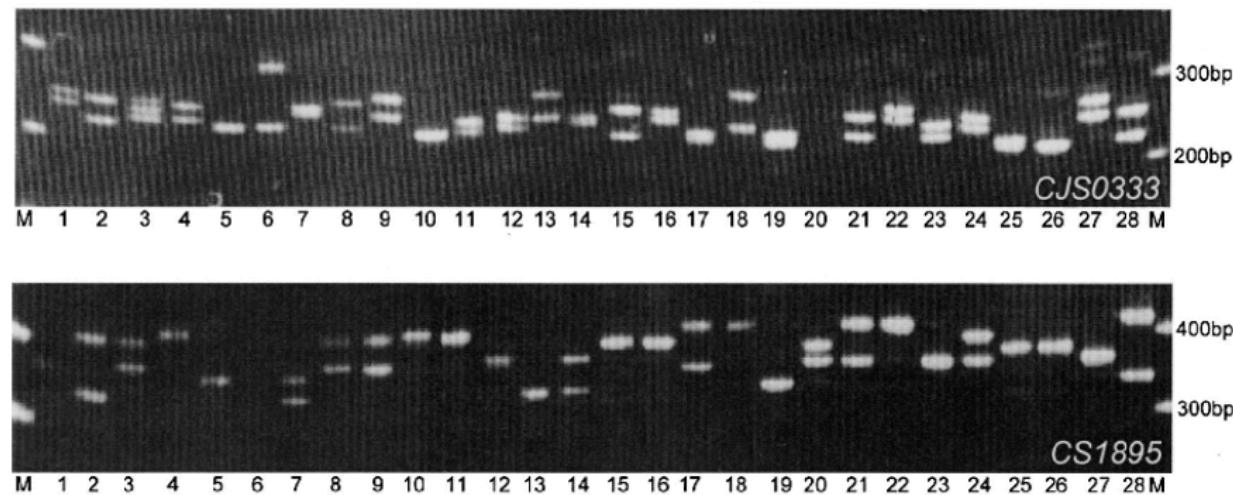
Co-dominant

Many alleles

High mutation rates

High heterozygosities

$(TA)_n$



# Barred hamlet (*H. puella*) microsatellite dataset

## Exercise 3

Hamlets are simultaneous hermaphrodites with external fertilization.



Photograph by Luiz Rocha

Test whether self-fertilization occurs regularly using population genetic data