Marine Ecological Genetics

02. Hardy-Weinberg equilibrium | computer practical

- Calculate allele and genotype frequencies
- Test for Hardy-Weinberg equilibrium
- Learn to interpret HW test results
- Become familiar with microsatellite data

Martin Helmkampf

Access files for practical

Download files from GitHub:

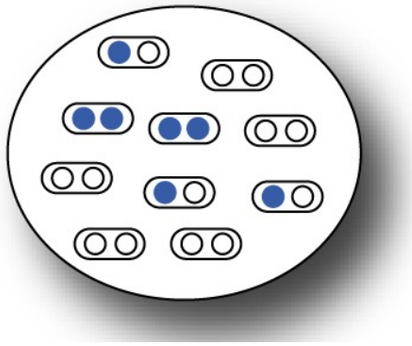
https://github.com/mhelmkampf/meg24.git (Code | Download ZIP)

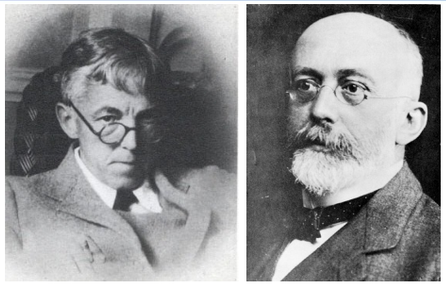
Alternatively, run git from terminal:

```
git clone https://github.com/mhelmkampf/meg24.git
```

Open R script called **Mar_Ecol_Gen_week2.R** in RStudio

HARDY-WEINBERG (1908)





Godfrey H. Hardy (1877-1947)

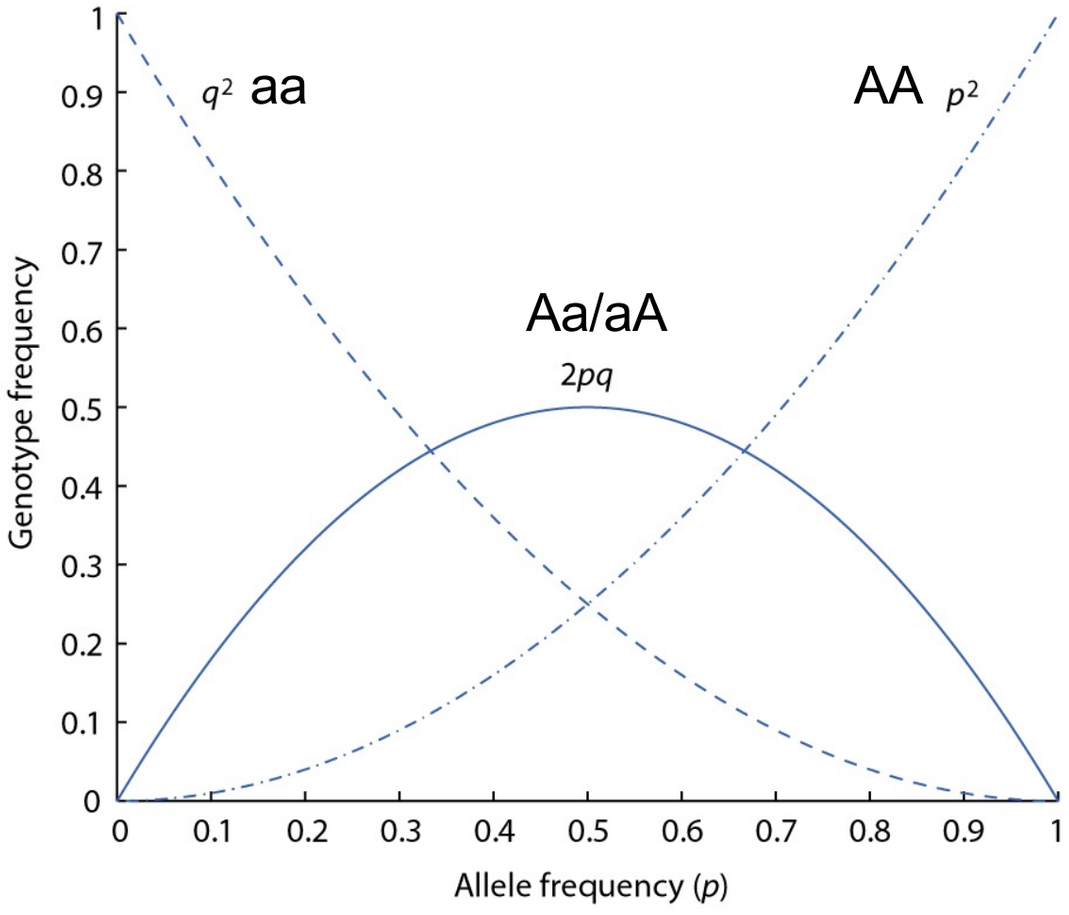
Wilhelm Weinberg (1862-1937)

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

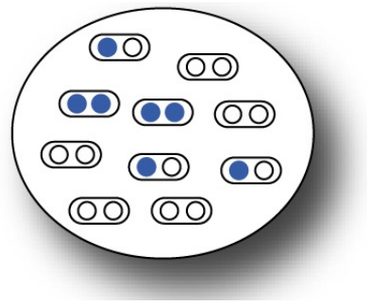
$$p^2 + 2pq + q^2 = 1$$

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*



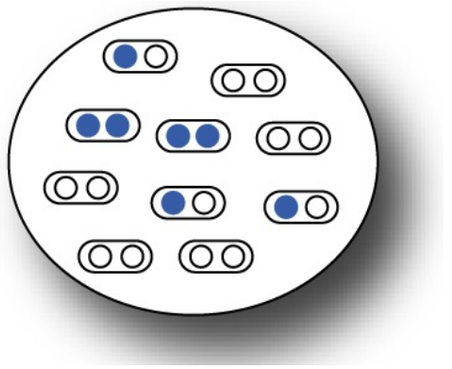
Assuming an “ideal” population, i.e. :

- 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
- ...

HETEROZYGOSITY



In one population

H_o = proportion of heterozygote individuals, observed heterozygosity

$H_e = 2pq = 1 - p^2 - q^2$, expected heterozygosity (assuming HW equilibrium)

$$F = \frac{H_e - H_o}{H_e}$$

Fixation index: *proportion by which heterozygosity is reduced or increased relative to the heterozygosity of a population at HW equilibrium with the same allele frequencies*.

Divided by H_e −> *proportion* (of expected heterozygosity)

Varies between -1 and 1

$F < 0$: heterozygote excess

$F > 0$ heterozygote deficit (homozygote excess)

May be averaged over several loci −> reduces bias

May be extended to *k* alleles

Is this population in Hardy-Weinberg equilibrium?

Exercise 1

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

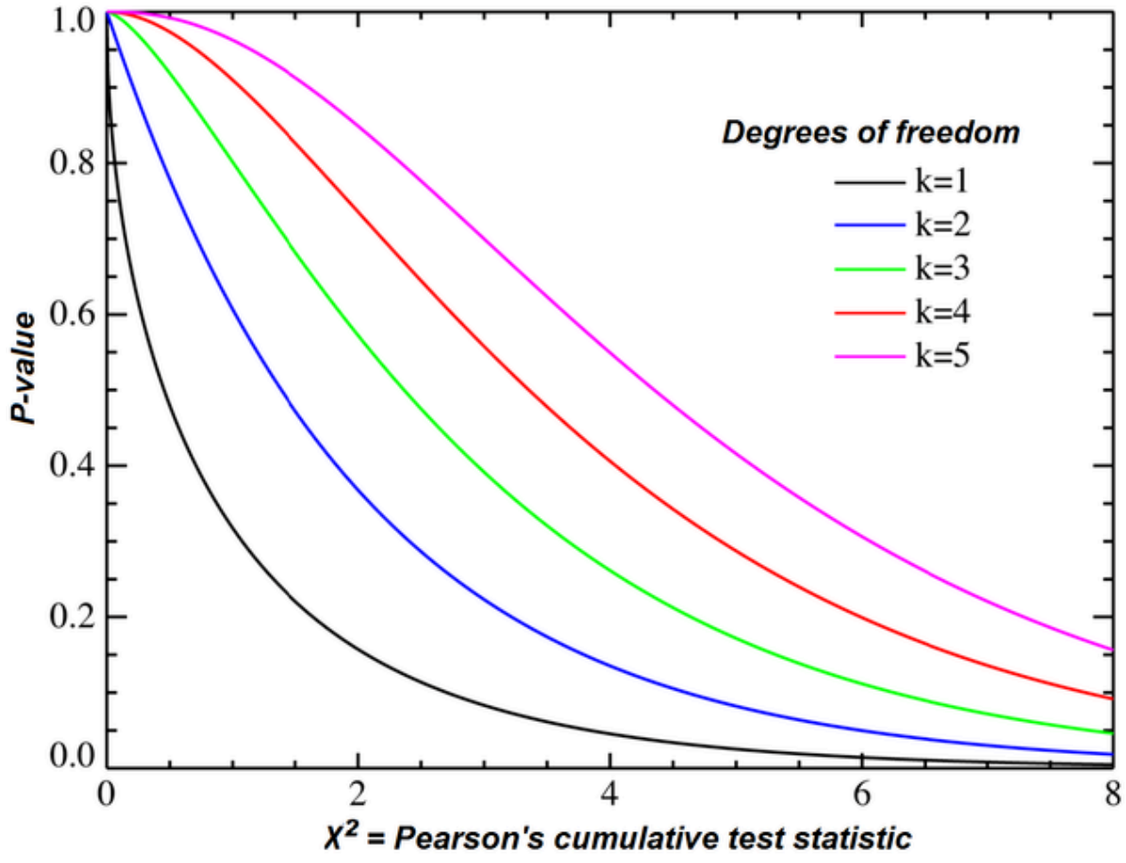
Pearson’s chi-squared test

Exercise 1

Chi-square statistic:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Chi-square distribution:



Genepop format

yellowblue.txt

```
Example data
Color
POP
ind01, yy
ind02, yy
ind03, yy
ind04, yy
ind05, yy
ind06, yy
ind07, yb
ind08, bb
ind09, bb
ind10, bb
```

puella_barbados.txt

```
Microsatellite genotypes of Hypoplectrus puella from Barbados
g2
gag010
h24
hyp001
hyp015
hyp018
e2
hyp008
hyp016
pam013
POP
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
```

Free text

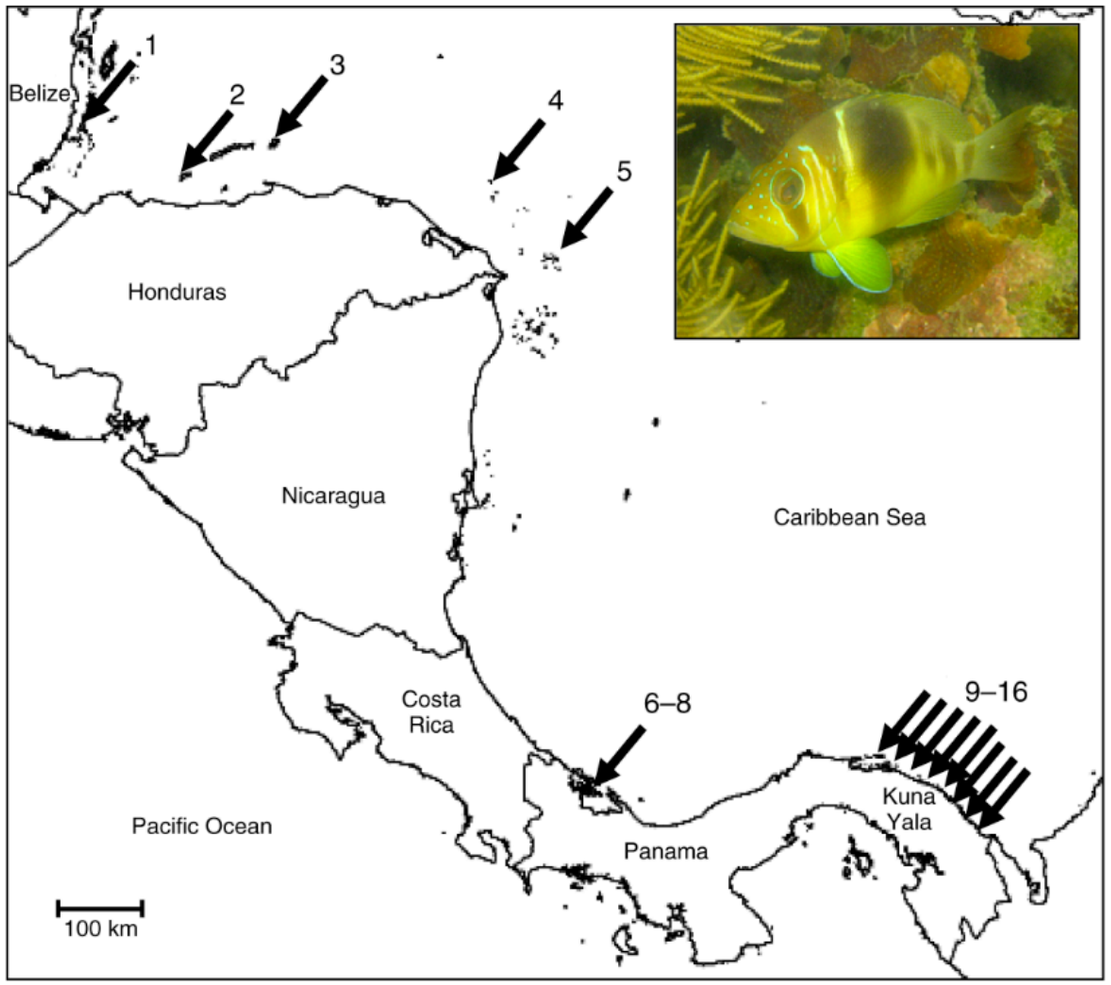
Locus names

Sample / population

Alleles (000 = missing data)

Barred hamlet (*H. puella*) microsatellite dataset

Exercise 3



Ecology, 90(11), 2009, pp. 3087–3098

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

Oscar Puebla,[1,2,3] Eldredge Bermingham,[1,2] and Frédéric Guichard[2]

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

Take-home messages

- Testing for HWE can tell us whether a population is behaving like an ideal population or not

- Deviation from HWE indicates that the population is under the influence of genetic, evolutionary or demographic processes

- Tests should be conducted over multiple loci, because not all loci may show the same pattern with respect to HWE (requires correction for multiple tests)

- Microsatellites are a widely used type of marker in population genetics (as well as forensics and other applications)