Populational genetics
Hardy-Weinberg equilibrium, departure from HWE

24.10.2005
GE02: day 1 part 2

Yurii Auchenko
Erasmus MC Rotterdam

What is a "population"?

- Two individuals A and B belog to the same genetic population if
  - the probability that they would have an offspring in commomn is greater then zero and
  - this probability is much higher than the probability of A and B having an offspring in common with some individual C, which is said to be belonging to other population

Genetic processes in populations

- **Selection** is a process of differential reproduction

- **Mutation is the process in which one allele is changed to other**

- **Random processes, e.g. drift**

Genetic processes in large populaions

- Assumptions:
  - infinitely large population
  - Generation -> Gametic pool -> Generation abstract
  - Random, independent segregation and aggregation of alleles

Hardy-Weinberg equilibrium

- Consider two alleles, N and D, are segregating in a population. Frequency of D, $P(D) = 0.1$

- If aggregation of alleles is independent and random, what are the expected genotypic proportions?

Solution

- Homozygotes
  - P(N and N) = P(N) x P(N) = 0.9 x 0.9 = 0.81
  - P(D and D) = P(D) x P(D) = 0.1 x 0.1 = 0.01

- Heterozygote
  - P(N and D) = P(N) x P(D) = 0.9 x 0.1 = 0.09
  - P(D and N) = P(D) x P(N) = 0.1 x 0.9 = 0.09

  - Total, P(ND or DN) = 0.18

Hardy-Weinberg equilibrium

- If frequency of allele D is q and the frequency of N is p = (1 − q) then

  - P(DD) = q^2

  - P(ND) = 2 p q

  - P(DD) = p^2

- These proportions are known as HWE

- P(ND) is known as heterozygosity, a measure of marker polymorphism

Task

- Consider three alleles, A_1, A_2 and A_3, segregating in a population
  - $P(A_1) = 0.1$ and $P(A_2) = 0.2$
- Aggregation of alleles is independent and random
- What is
  - Frequency of A_3?
  - How many genotypes can be observed?
  - What are equilibrium proportions?

Solution

- Frequency of A_3?

  $$P(A_3) = 1 - P(A_1) - P(A_2) = 0.7$$

- How many genotypes can be observed?

  Six: A_1A_1, A_1A_2, $\mathbf{A_1}A_3$, $A_2\mathbf{A_2}$, A_2A_3, and A_3A_3

- Generally: if there are n alleles, there could be
  - $n(n+1)/2$ genotypes

Solution

- What are equilibrium proportions?
  - $P(A_1 A_1) = P(A_1) \, P(A_1) = 0.01$
  - $P(A_1 A_2) = 2 \, P(A_1) \, P(A_2) = 0.04$
  - $P(A_1 A_3) = 2 \, P(A_1) \, P(A_3) = 0.14$
  - $P(A_2 A_2) = P(A_2) \, P(A_2) = 0.04$
  - $P(A_2 A_3) = 2 \, P(A_2) \, P(A_3) = 0.28$
  - $P(A_3 A_3) = P(A_3) \, P(A_3) = 0.49$

HWE for multiple alleles

- $P(A_i A_i) = P(A_i)^2$

- $P(A_i A_j) = 2\, P(A_i)\, P(A_j)$

- Heterozygosity is defined as

$$\Sigma_{i > j}\, 2\, P(A_i)\, P(A_j)$$

When HWE is reached?

- If the frequency of genotypes are
  - P(DD) = 0.1, P(ND) = 0.2 and P(NN) = 0.7
- Then
  - What is the frequency of D, P(D)?
  - What genotypic frequencies are expected under HWE?
  - What will be genotypic frequencies after a generation of random mating?

Follow the model…

- What is the frequency of D, P(D)?
  - P(D) = P(DD) + P(ND)/2 = 0.1 + 0.2/2 = 0.2

- Now the gametes start randomly aggregate and we get HWE in one generation:
  - P(DD) = 0.04, P(ND) = 0.32, P(NN) = 0.64

Conclusion

- If genotypes are not in HWE, it is reached in one generation of random mating and segregation/aggregation!

- Normally these assumptions holds

Why HWE may be violated?

- **Genotyping or calling error**
- Sample is not from the same population
- Non-random mating
- Selection

Genotyping or calling error

- SNPs
  - Heterozygotes more difficult to call than homozygotes

- Microsatellites
  - Call errors
  - Missed alleles

Problem

- Three alleles (A_1, A_2 and A_3) are present in a population with frequencies 0.05, 0.15 and 0.8

- Allele A_2 is not called

- What will be genotypic distribution for such locus?

- What is the one expected under HWE?

Answer

- What will be genotypic distribution for such locus?

  – Genotype   11   12   13 22   23   33

  – True distr. (%)   0.25   1.5   8   2.25   24   64

  – Calling   11   11   13 miss   33   33

  – Obs. genot.   11   13   33

  – Observed%   1.75   8   88   / (1 − 0.0225)

                       1.8   8.2   90

Compare with HWE

	11	13	33
– Observed%	1.8	8.2	90

- What is the one expected under HWE?

  – P(1) = 1.8 + 8.2/2 = 5.9%

	11	13	33
– Expected%	0.3	11	89

- Excess of homozygotes!

Why HWE may be violated?

- **Genotyping or calling error**
- Sample is not from the same population
- Non-random mating
- Selection

HWE violation because of incorrect sampling

- Bad sample:
  - 50% are from population where allele frequency is 0.2 (say, European) and
  - other 50% are from population where frequency is 0.9 (Say, Japanese)
- Both populations are in HWE

Resulting population

	AA	Aa	aa
Pop1	0.64	0.32	0.04
Pop2	0.01	0.18	0.81
Mixture	0.325	0.25	0.425

- In mixed population, $P(a) = 0.425 + 0.25/2 = 0.55$

- Expected under HWE:

Expec	0.20	0.50	0.30

- Mixture usually leads to heterozygote deficiency!