

## II. Population Genetics

# Hardy Weinberg Equilibrium

---

### Lecture 9

Instructor: Su-In Lee

## Contents

---

- Properties of alleles in a population
  - Allele frequencies
  - Genotypes frequencies
  - Hardy-Weinberg equilibrium (HWE)

## Alleles

---

- Alternative forms of a particular sequence
- Each allele has a frequency, which is the proportion of chromosomes of that type in the population

## Alleles Frequency Notation

---

- For two alleles
  - Usually labeled  $p$  and  $q = 1 - p$
- For more than 2 alleles
  - Usually labeled  $p_A, p_B, p_C \dots$
  - ... subscripts  $A, B$  and  $C$  indicate allele name

## Genotype

- The pair of alleles carried by an individual
  - If there are  $n$  alternative alleles ...
  - ... there will be  $n(n+1)/2$  possible genotypes
- **Homozygotes**
  - The two alleles are in the same state
- **Heterozygotes**
  - The two alleles are different

## Genotype Frequencies

- Since alleles occur in pairs, these are a useful descriptor of genetic data ...
- However, in any non-trivial study we might have a lot of frequencies to estimate ...
- $p_{AA}, p_{AB}, p_{AC}, \dots, p_{BB}, p_{BC}, \dots, p_{CC} \dots$

## The simple part

- Genotype frequencies lead to allele frequencies...
- For example, for two alleles:
  - $p_A = p_{AA} + \frac{1}{2} p_{AB}$
  - $p_B = p_{BB} + \frac{1}{2} p_{AB}$
- However, the reverse is also possible!

## Hardy-Weinberg Equilibrium

- Random union of gametes
- Relationship described in 1908
  - Hardy, British mathematician
  - Weinberg, German physician
- Shows **n** allele frequencies determine  **$n(n+1)/2$**  genotype frequencies
  - Large populations

## Required Assumptions

- Diploid, sexual organism
  - Non-overlapping generations
- Autosomal locus
- Large population
- Random mating
- Equal genotype frequencies among sexes
- Absence of natural selection

## Random Mating: Mating Type Frequencies

| <u>Mating</u>     | <u>Frequency</u> |
|-------------------|------------------|
| $A_1A_1 * A_1A_1$ |                  |
| $A_1A_1 * A_1A_2$ |                  |
| $A_1A_1 * A_2A_2$ |                  |
| $A_1A_2 * A_1A_2$ |                  |
| $A_1A_2 * A_2A_2$ |                  |
| $A_2A_2 * A_2A_2$ |                  |
| <u>Total</u>      | <u>1.0</u>       |

## Mendelian Segregation: Offspring Genotype Frequencies

| Mating            | Frequency | Offspring |          |          |
|-------------------|-----------|-----------|----------|----------|
|                   |           | $A_1A_1$  | $A_1A_2$ | $A_2A_2$ |
| $A_1A_1 * A_1A_1$ |           |           |          |          |
| $A_1A_1 * A_1A_2$ |           |           |          |          |
| $A_1A_1 * A_2A_2$ |           |           |          |          |
| $A_1A_2 * A_1A_2$ |           |           |          |          |
| $A_1A_2 * A_2A_2$ |           |           |          |          |
| $A_2A_2 * A_2A_2$ |           |           |          |          |

## Next Generation ...

## Conclusion

- Genotype frequencies are function of allele frequencies
  - Equilibrium reached in one generation
  - Independent of initial genotype frequencies
  - Random mating, etc. required
- Conform to binomial expansion
  - $(p_1 + p_2)^2 = p_1^2 + 2p_1p_2 + p_2^2$

## Simple HWE Example

- If the defective alleles of the cystic fibrosis (CFTR) gene have a cumulative frequency of 1/50 what is:
  - The proportion of carriers in the population?
  - The proportion of affected children at birth?

## Checking Hardy-Weinberg Equilibrium

- A common first step in *any* genetic study is to verify that the data conforms to Hardy-Weinberg equilibrium
- Deviations can occur due to:
  - Systematic errors in genotyping,
  - Unexpected population structure,
  - Presence of homologous regions in the genome,
  - Association with trait in case-control studies.
- Which of these causes would you expect to increase the proportion of heterozygotes?

## Testing Hardy Weinberg Equilibrium

- Consider a sample of  $2N$  alleles
- $n_A$  alleles of type A
- $n_B$  alleles of type B
- $n_{AA}$  genotypes of type AA
- $n_{AB}$  genotypes of type AB
- $n_{BB}$  genotypes of type BB

## Simple Approach

---

- Calculate allele frequencies and expected counts
- Construct chi-squared test statistic
- Convenient, but can be inaccurate, especially when one allele is rare

## Recommended Reading

---

- Wigginton, Cutler, Abecasis (2005), A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* **76**:887-93
- This paper describes an efficient method for testing Hardy-Weinberg equilibrium and includes many important historical references

## Acknowledgement

---

- This set of slides is based on the slides by:
  - Goncalo Abecasis (Univ of Michigan)