From Bi203 – non-majors' evolution course

Population Genetics

- The Hardy-Weinberg equilibrium.
  - No selection.
  - No gene flow.
  - Large population.
  - Random mating.
  - Mutation is not detectable.

$$P^2 + 2PQ + Q^2 = 1.0$$

Hardy-Weinberg provides a rigorous test of evolution.

Hardy-Weinberg Equilibrium

Flipping Coins

P ½ = 0.50

Q ½ = 0.50

1.00

Allele Frequencies Add to 1

$P + Q = 1.0$

Flipping Two Coins

P P ½ x ½ = 0.25

P Q ½ x ½ = 0.25

Q P ½ x ½ = 0.25

Q Q ½ x ½ = 0.25

1.00

If we flip two coins 100 times, how many times do you expect to get at least one 'heads'?

Coin 2 \ Coin 1	Q	P
Q	25	25
P	25	25

Hardy-Weinberg Equilibrium

- Expected genotype frequencies:

$$P^2 + 2PQ + Q^2 = 1.0$$

- If genotype frequencies do not match the Hardy-Weinberg expectation then….
  - Selection?
  - Non-random mating?
  - Gene flow from other populations?

Application of HW

- The population genetics of a rare genetic disorder: Cystic Fibrosis (CF).
  - 1/2500 people of Caucasian descent are born with CF.
  - CF is known to be caused by a recessive allele (gene).
    - Individuals with two copies of the CF gene (c/c) will have this disease.
    - Individuals who are heterozygous (C/c), will not have the disease.
      - They are carriers.
  - Why is CF rare in populations?
  - What is the frequency of the CF allele?
  - How many people are carriers?

Population Genetics of a Rare Genetic Disorder

- Why is CF rare in populations?
  - Individuals with this disease have low survival and reproduction.
    - Selection against the CF gene.
- What is the frequency of the CF disease?
  - If 1/2500 people are born with this disease then 1/2500 = 0.0004

Genotype	Number of People	Frequency
C/C		
C/c		
c/c	1	0.0004
Total	2500	1.0

If we could prevent people with CF from reproducing could we eliminate this disease?

1. Yes, but it would take too long.
2. Yes, but it would be unethical.
3. No, this is not possible.
4. Both 1 and 2 are true.

Population Genetics of a Rare Genetic Disorder

- How many people are carriers?
  - The frequency of the c allele = square root (q) = square root (0.0004) = 0.02
  - The frequency of the C allele = 1 – q = 1 – 0.02 = 0.98
  - The frequency of carriers (heterozygous for the CF allele) = 2pq = 2*0.98*0.02 = 0.0392 (that's 3.92% of the population).
  - The number of people who are carriers out of 2500 = 2pq * 2500 = 0.0392 * 2500 = 98

Genotype	Number of People	Frequency
C/C	2401	0.9604 p^2
C/c	98	0.0392 2pq
c/c	1	0.0004 q^2
Total	2500	1.0

Population Genetics of a Rare Genetic Disorder

- Why doesn't selection completely eliminate this disease from human populations?
  - There will always be carriers.
  - The disease will be carried from one generation to the next undetected.
  - Only appears if two people who are carriers have children.
- This principle led to the demise of the Eugenics movement.

Genotype	Number of People	Frequency
C/C	2401	0.9604 p^2
C/c	98	0.0392 2pq
c/c	1	0.0004 q^2
Total	2500	1.0

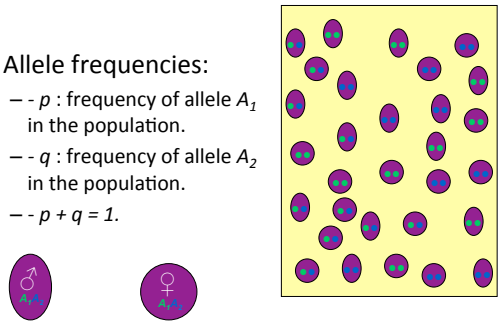
From Bi358 – evolution for bio majors

The HW Principle

- Assumptions.
  - Very large population.
  - Random mating.
  - No selection.
  - No mutation.
- Provides a description of the genetic composition of a diploid population in allelic terms.
  - Predictable/tractable.
  - Provides a clear testable hypothesis.

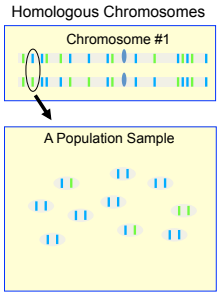
The Principle

- Allele frequencies:
  - *p* : frequency of allele A_1 in the population.
  - *q* : frequency of allele A_2 in the population.
  - $p + q = 1$.

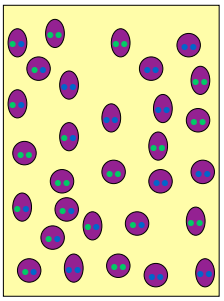


Models in Population Genetics

- Estimating allele frequencies (diploid organism).
  - *P* is the number of green alleles.
  - *Q* is the number of blue alleles.
  - *p* is the frequency of green alleles:
    - $= 4/(4 + 16) = 4/20 = 0.20$
  - The total frequencies of all alleles (2 in this case) must sum to one ….
    - $p + q = 1$
    - $q = 1 – p = 0.80$

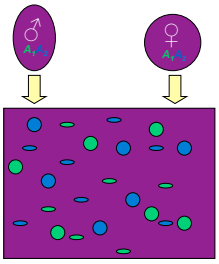


- In this population:
  - *N* (population size) = 100
  - *P* (number of allele A_1 in the population) = 50
  - What is *p* (frequency of allele A_1 in the population)?
  - Answer:
    - $p = P / 2N$
    - $p = 50 / 200 = 0.25$



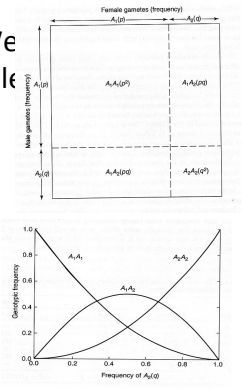
The Principle

- Random union of gametes:
  - The probability (frequency) of two independent events (union of two gametes) is the product of their individual probabilities of occurrence (frequencies).
  - Genotype frequencies:
    - $P_{11}(t+1) = p^2$
    - $P_{22}(t+1) = q^2$
    - $P_{12}(t+1) = pq + qp = 2pq$.



The Hardy We Principle

- Genotype frequencies depend on allele frequencies.
  - Heterozygotes are most abundant if
    $p = q = 0.5$
  - If *p* or *q* are close to 1.0.
    - Nearly all of the less common alleles are present only in heterozygous genotypes.
    - Homozygous genotypes of a rare allele are extremely infrequent.



The Principle

P_{11} = number of A_1 homozygotes.
p_1 = frequency of A_1 allele.

- And so…..
  - $p^2 + 2pq + q^2 = 1$.
  - Genotype frequencies are at a 'neutral' equilibrium.
    - It will not change unless allele frequencies change.
    - Populations will not necessarily return to the original equilibrium after a perturbation.
    - Different populations may differ for *p* and *q*.
  - This will hold true as long as there is no mutation, selection, migration, or drift.

$$p_{t+1} = P_{11(t)} + \frac{1}{2}P_{12(t)}$$

$$p_{t+1} = p_{(t)}^2 + \frac{1}{2}\left(2p_{(t)}q_{(t)}\right)$$

$$p_{t+1} = p_{(t)}^2 + p_{(t)}q_{(t)}$$

$$p_{t+1} = p_{(t)}\left(p_{(t)} + q_{(t)}\right)$$

$$p_{t+1} = p_{(t)} = p$$

Fun with HW

Start with (at time *t*):
$P_{11} = 10$
$P_{12} = 60$
$P_{22} = 30$

- First step:
  - What is *p* ?

$$p = \frac{P_{11} + \frac{1}{2} P_{12}}{N} = (10 + 30)/100 = 0.40$$

$$p = \frac{2P_{11} + P_{12}}{2N} = (20 + 60)/200 = 0.40$$

and ….

$q = 1 - p = 0.60$

Fun with HW

Start with (at time *t*):
$P_{11} = 10$
$P_{12} = 60$
$P_{22} = 30$
$N = 100$

- Next step:
  - What are the expected genotype frequencies ?
  - We know that:

Assuming HW, we expect:
$P_{11} = p^2N = 16$
$P_{12} = 2pqN = 48$
$P_{22} = q^2N = 36$

$p = 0.40$ $q = 0.60$

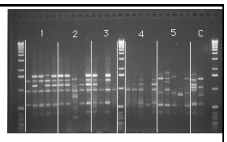
$p^2 = 0.16$ $q^2 = 0.36$

$pq = 0.24$

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

Why is the expected number of heterozygotes less than the observed number?

More Fun with HW



- We know 87% of the moths were melanic…..
- Can we calculate the allele frequency, *p* (for the M allele)?

1. Yes, $P_{MM} = p^2$ so take the square root to get *p* (assuming HW equilibrium). 48%
2. No, we cannot determine the number of black moths that are heterozygous (*2pq*). 18%

Yes, calculate q^2 from P_{mm} and take the square root to get *q* and *p* (assuming HW equilibrium) 34%

More Fun with HW

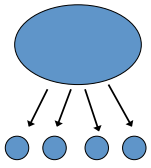
– Assuming HW, the recessive allele would be…

- $P_{mm} = 0.13 = (1\text{-}0.87)$.
- We expect $p_m^2 = 0.13$, *and* $p_m = 0.13^{1/2} = 0.36$.
- So, $p_M = (1 - p_m) = 0.64$.

Gene flow: Continent – Island Model

- Large population (continent)
  - Island is a large population (no drift)
  - Gene flow is in one direction (to the island)
  - Assume *m* individuals are migrants (non migrants = $1 - m$) and q_m is the frequency of A_2 on the continent….
    - $q_1 = (1 - m)q_0 + mq_m$

Continent-Island Model



Change in allele frequency on the island

- Assume that q_m does not change
  - Allele frequency on the island will converge to q_m
  - Change in *q* is a linear function of *m*
  - Over time, *q* will become equal to q_m (asymptotically)
- Remember
  - Continent allele frequency (q_m) does not change.
  - The island population is large enough that drift is not a factor