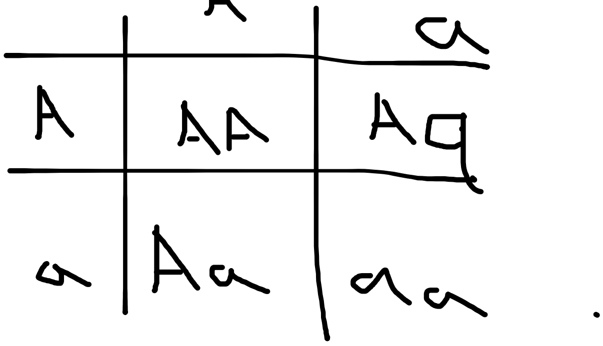
Chap 1: POPULATION GENETICS

1.1 <u>Segregation Principle</u>

- Gamete randomly receive one of adult's 2 alleles (Aa) for each autosomal gene with equal probability.
- Gamete randomly receive one of the adult's two sex chromosomes (XY) with equal probability.
- Occurs **independently** for genes on different chromosomes.
- Example:
    - Aa x Aa



        - By **segregation principle**, E_1=offspring get A from father, E_2=offspring get A from mother -> $P(E_1) = P(E_2) = 0.5$
        - Events are **independent**.
            - P(AA)-> 25%
                - P(AA)= $P(E_1 \cap E_2) = P(E_1) \times P(E_2)$
                    - multiplication law for independent events as E1 and E2 are independent
                - 0.5*0.5=**0.25**
            - P(Aa)-> 50%
                - $P[(E_1 \cap E'_2) \cup (E'_1 \cap E_2)]$
                - $P(E_1 \cap E'_2) + P(E'_1 \cap E_2)$ as mutually exclusive
                - $[P(E_1) \times P(E'_2)] + [P(E'_1) \times P(E_2)]$ as E1 and E2 are independent
                - $(0.5 * 0.5) + (0.5 * 0.5)$=**0.5**
            - P(aa) -> 25%

1.2 <u>Law of probability</u>

- Event -> something that may or may not occur.
- Complement Law:
    - P(E') = 1-P(E)
- Conditional probability:
    - $P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$
- Independent events: (multiplication law)
    - $P(E_1 \cap E_2) = P(E_1) \times P(E_2)$
- Mutually exclusive: (addition law)
    - $P(E_1 \cap E_2) = 0$

- $P(E_1 \cup E_2) = P(E_1) + P(E_2)$

1.3 Gene Disorders

- 4000 (rare) diseases, caused by single gene disorders/Mendelian disease
  - Depend on whether gene is:
    - Autosomal / sex-linked
    - Dominant / recessive
- Example:
  - Autosomal dominant
    - Huntington's disease (frequency 1 in 40,000)
  - Autosomal recessive
    - Sickle cell anemia (frequency 1 in 625)
  - X-linked recessive
    - Red-green color blindness (frequency 1 in 125)
- **Prevalence VS Incidence**
  - Prevalence (frequency) -> total number/proportion **of population** affected by disease **at a given time**.
  - Incidence -> total number/proportion **of new births** that are affected by disease.
    - If incidence is constant, incidence = prevalence

1.4 Genotype Number and Frequency

- N – total population size
- N^{AA} – number of AA individuals in population
- f^{AA} – frequency of AA in population $f^{AA} = \frac{N^{AA}}{N}$
- $N^{AA} + N^{Aa} + N^{aa} = N$ and $f^{AA} + f^{Aa} + f^{aa} = 1$

1.5 Allele Numbers and Frequency

- N^{A} – number of A alleles in population
- f^{A} – frequency of A allele in **allele pool** (allele pool (single allele) from everyone in population)

1.6 Calculating allele frequencies from genotype frequencies

- $N^A = 2N^{AA} + N^{Aa}$ $N^a = 2N^{aa} + N^{Aa}$
- $f^A = \frac{N^A}{2N}$ $f^a = \frac{N^a}{2N}$
- $f^A = \frac{N^A}{2N} = \frac{2N^{AA}+N^{Aa}}{2N} = \frac{N^{AA}}{N} + \frac{1}{2}\frac{N^{Aa}}{N}$
- $f^A = f^{AA} + \frac{1}{2}f^{Aa}$ $f^a = f^{aa} + \frac{1}{2}f^{Aa}$ (note: f^A+f^a=1)

1.7 Modelling assumptions (assumption to derive a math model) in population genetics

- 1.7.1. Random mating
  - Adults choose mating partner random (mate selection independent of genotypes)
  - Same as random sampling (**with replacement**) from the allele pool
  - Thus:
    - **Single random mating**

- P(offspring is AA) = $(f^A)^2$
- P(offspring is Aa) = $2f^A f^a = 2f^A(1-f^A)$
- P(offspring is aa) = $(f^a)^2 = (1-f^A)^2$

- **N random mating**
  - $N^{AA} \sim Bi(N, (f^A)^2)$
  - $N^{Aa} \sim Bi\big(N, 2f^A(1-f^A)\big)$
  - $N^{aa} \sim Bi(N, (1-f^A)^2)$
    - $X \sim Bi(n, p)$ n=no. trials, p=probability of success
      - $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$
      - $E(X) = np$
        - $E(N^{AA}) = N(f^A)^2$
        - $E(f^{AA}) = \frac{1}{N} E(N^{AA}) = (f^A)^2$
      - $Var(X) = np(1-p)$
        - $var(N^{AA}) = N(f^A)^2(1-(f^A)^2)$
        - $var(f^{AA}) = \frac{1}{N^2} var(N^{AA})$
        - Mean genotype frequencies are independent of N, but S.D. shrink as N increase.

- 1.7.2 Large population ($\boldsymbol{N \cong 10^4}$) assumption
  - As N increases, genotype frequency become concentrated on mean value, ignore variability in f^{AA}
  - $f^{AA} = E(f^{AA})$