Genetic drift

24.10.2005
GE02: day 2 part 3

Yurii Aulchenko
Erasmus MC Rotterdam

Simple genetic population

- Model
  - There is a population of n individuals ($2n$ chromosomes)
  - A very large number of copies is generated fom each chromosomes (gamete pool)
  - Next generation is obtained by random sampling of $2n$ chromosomes from this pool

Problem

- Consider a population of 50 people
- One of chromosomes is mutant
- What is the chance that in the next generation the mutation will
  - Disappear?
  - Be still present as single copy?
  - Increase its' frequency?

Solution

- Disappear?
  - P(0 copies M) = 0.99^{100} = 0.366
- Be still present as single copy?
  - P(1 copy M) = 100 0.01 0.99^{99} = 0.37
- Increase its' frequency?
  - P(≥2 copies M) = 1 – P(0 copies) – P(1 copy) =

    1 – 0.366 – 0.37 = 0.264

Drift

- In finite genetic populations allelic frequencies are subject to drift (random changes) because of sampling. Drift may occur because of
  - Small population size
  - Bottleneck effect
    - A large population is reduced very much in size at certain stage
  - Founder effects
    - A small group of founders is sampled from large population to start new one

Bottleneck / Founder effect

- In a population, mutations of some gene are present with frequencies 0.001 (M1), 0.003 (M2) and 0.005 (M3)
- Due to bottleneck or founder effect, the population is reduced to 50 people (100 chromosomes)
- What is the chance that none of these mutations will be present in founders of the new population?
- What is the chance that all 3 mutations will be presents?

Solution

- What is the chance that none of these mutations will be present in founders of the new population?

  $(1 - 0.001 - 0.003 - 0.005)^{100} = 0.405$

- What is the chance that all 3 mutations will be presents?

  Approximate $P(M_1 \geq 1 \ \& \ M_2 \geq 1 \ \& \ M_3 \geq 1)$ by

  $$P(M_1 \geq 1)\ P(M_2 \geq 1)\ P(M_3 \geq 1)$$

  $P(M_1 \geq 1)\ P(M_2 \geq 1)\ P(M_3 \geq 1) = 0.095\ 0.26\ 0.394 = 0.01$

Very small population

- Consider a “population” made of a single self-pollinating plant
- Initially, the plant is heterozygous (genotype AB)

Task

- What is chance that it will be heterozygous in
  - First generation
  - 10^{th} generation
  - *n*-th generation
- After infinite number of generations, what genotypes will be present in the population?

Answer

- What is chance that it will be heterozygous in
  - First generation – $(\frac{1}{2})$
  - 10^{th} generation – $(\frac{1}{2})^{10} = 1/1024$
  - *n*-th generation – $(\frac{1}{2})^n$
- After infinite number of generations, what genotypes will be present in the population?
- When $n \rightarrow \infty$ then $(\frac{1}{2})^n \rightarrow 0$ therefore only AA or BB may be present, with equal chance of ½

Drift

- A population made of $2n$ chromosomes
- k of these are "mutant" (M) and $2n - k$ are "normal" (N). Thus the initial frequency of mutant allele is $p = k/2n$
- After infinite number of generations, probability that
  - Both types are present is 0
  - Only M are present is $k/2n = p$
  - Only N are present is $(2n - k)/2n = 1 - p$
- Expected number of generations before allele is lost is

$$-[2\,k \log_e(p)] / (1 - p) \text{ (if } p = 1/2n \text{ then } 2 \log_e(2n))$$

- Expected number of generations before allele is fixed is

$$-[4\,n\,(1 - p) \log_e(1 - p)] / p \text{ (if } p = 1/2n \text{ then } 4\,n)$$

Drift for 18 chromosomes over 19 generations

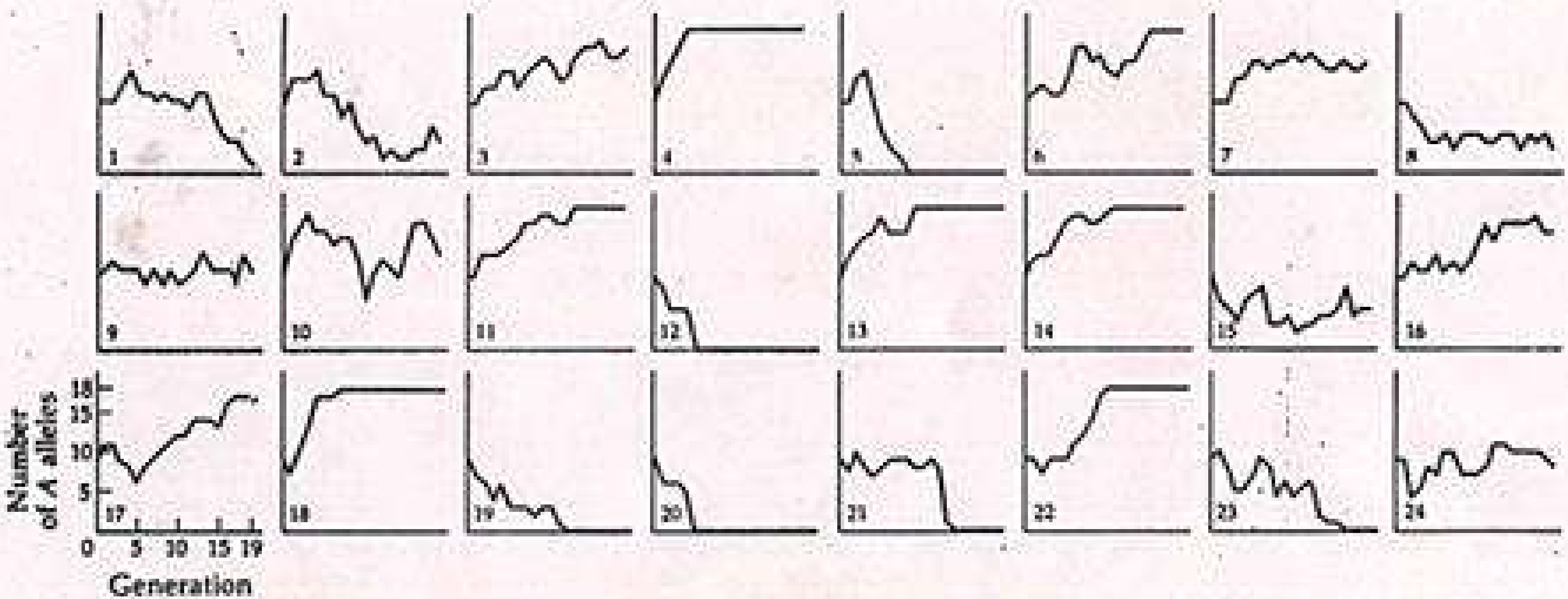

FIGURE 1. Change of allele frequency by random genetic drift over 19 generations in 24 hypothetical populations of size $N = 9$.

Effective number

- The number discussed above does not directly relate to number of people in a population
- n is a so-called “effective” number, it is always smaller then real population size