Week 7 Population Genomics

Monday, 7 March 2022 8:10 AM

H_0 = NULL (GTR) ← D = data

H_A = ALTERNATE (JC69) ←

$p(H_0 | D)$ = ...

$p(H_A | D)$ = ...

$$-2 \times \left[\frac{\log(p(H_0|D))}{\log(p(H_A|D))} \right] \sim \left[\chi^2 \right]$$

$$\text{eg} \quad \left[df = 10 - 1 = 9 \right]$$

POPULATION GENOMICS

1] LOCUS - physical location on a chromosome

2] GENE - protein coding locus

3] ALLELE - variant at a locus

4] GENOTYPE - combination of alleles at a locus eg. A/A

5] HAPLOTYPE - combination of alleles at neighboring loci [HAPLOID]

... of homologous

6] PLOIDY - # of sets of chromosomes in a cell

7] LINKAGE - alleles at multiple loci on the same chromosome that are inherited together

8] RECOMBINATION - exchange of homologous segments during meiosis (prophase 1)

9) Mendelian principles :

↳ 1 Dominance → simple dominance relationship b/w alleles at a locus

↳ 2 Independent segregation

↳ homologous alleles segregate away from each other

↳ 3 Independent assortment :

↳ alleles at multiple independent loci assort separately

POPULATION : individuals of the same species that interbreed w/ each other in space & time

HARDY - WEINBERG PRINCIPLE

1 population ; at time 't' ; at a single locus ; bi-allelic [A, G]

(diploid, sexually reproducing)

$\# AA = n_{AA}$, $\# AG = n_{AG}$, $\# GG = n_{GG}$

$\Rightarrow n = n_{AA} + n_{AG} + n_{GG}$

$\Rightarrow p(AA) = \frac{n_{AA}}{n}$; $p(AG) = \frac{n_{AG}}{n}$; $p(GG) = \frac{n_{GG}}{n}$

$\Rightarrow \boxed{P(AA) + P(AG) + P(GG) = 1}$ GENOTYPE FREQUENCIES

ALLELE FREQUENCIES

$$p = P(A) = \frac{2n_{AA} + 1 \times n_{AG}}{2n} = P(AA) + \frac{1}{2} P(AG)$$

$$q = P(G) = \frac{2n_{GG} + 1 \times n_{AG}}{2n} = P(GG) + \frac{1}{2} P(AG)$$

$\boxed{P(A) + P(G) = 1}$ $\boxed{p + q = 1}$

$\Rightarrow \quad [\, p(A) + p(G)$

Random mating $p(A) = p$ ♂ $p(G) = q$

		A	G
♀ $p(A) = p$	A	AA p^2	AG pq
$p(G) = q$	G	AG pq	GG q^2

$$\left.\begin{aligned} p^{t+1}(AA) &= p^t(A) \times p^t(A) = p^2 \\ p^{t+1}(GG) &= p^t(G) \times p^t(G) = q^2 \\ p^{t+1}(AG) &= pq + pq = 2pq \end{aligned}\right] \quad \underline{\mathrm{I}}$$

$$p^{t+1}(A) = p^{t+1}(AA) + \frac{1}{2} p^{t+1}(AG)$$

$$= p^2 + \frac{1}{2} \times 2pq$$

$$\left.\begin{aligned} &= p^2 + pq = p(p+q) \\ &= p = p^t(A) \\ \text{|||ly} \quad p^{t+1}(G) &= q = p^t(G) \end{aligned}\right] \quad \underline{\mathrm{II}}$$

allele frequency change

EVOLUTION →

over time

ASSUMPTIONS : (of HWE)

- RANDOM MATING
- NO MUTATION(S)
- NO MIGRATION(S)
- LARGE POPULATION(S)
- NO SELECTION

---

① 'SMALL' POPULATION (FINITE) → n

$p^t(A) = p = \frac{2n_{AA} + n_{AG}}{2n}$

$p^t(G) = q$

t

$p^{t+1}(A) = \frac{x}{2n}$

$p^{t+1}(G) = \frac{2n-x}{2n}$

$t+1$

x A alleles

$2n - x$ G alleles

GENETIC DRIFT

WRIGHT - FISHER PROCESS

P^{t+1} (x A alleles, $2n-x$ G alleles) given p, q

$$= {}^{2n}C_x \, p^x \, q^{2n-x}$$

$$= {}^{2n}C_x \, p^x \, (1-p)^{2n-x}$$

RATE OF DRIFT ~ ① POPULATION SIZE (n)

~ ② p → frequency of

allele in previous generation

eg.

$n = 100$; $p^t(A) = 0.5 = p \Rightarrow q = 0.5$

$p(X = 0 \text{ in } t+1)$

$= {}^{200}C_0 \, (0.5)^0 \, (0.5)^{200}$

$= 6.22 \times 10^{-61}$

$n = 10$; $p^t(A) = 0.5 = p \Rightarrow q = 0.5$

$p(X = 0 \text{ in } t+1)$

$= 0.5^{20} = 9.54 \times 10^{-7}$

$\rightarrow$ smaller pop. size $\Rightarrow \uparrow$ DRIFT