MUTATION

INTRODUCTION TO MUTATION

READING: Nielsen & Slatkin pp. 27–32, 47–48

   – Mutation plays two key roles in evolution:

   (1) It is an evolutionary force that changes gene frequencies;

   (2) It is the ultimate source of all genetic variation

   • Note: any complete theory of evolution must explain processes that create mutations.

   – Most mutations are rare.

   • Alleles are rarely incorrectly replicated.

   • Ballpark figures for spontaneous mutation rates:

   – Most mutations are deleterious.

   • Most common mutations involve loss of function.

   • Deleterious mutations are generally recessive.

   • Mutation is a destructive force.

   • Mutation is a creative force.

MUTATION AS A FORCE OF EVOLUTIONARY CHANGE

   – Simplest case: two alleles *A* and *a*

   • Reality: consider an exon that is 100 bp long.

   • Approximation: alleles refer to 2 *classes* of alleles, e.g., T/not T (=G, C, or A)

   – Will consider *recurrent mutation*.

   – How do allele frequencies change given recurrent mutation?

   • Rate of mutation = probability a gene copy will mutate per generation

– Let u = rate of mutation from A to a ("forward" mutation rate).
           v = rate of mutation from a to A ("backward" mutation rate).

• Equations of genetic change:

$$p' = p(1-u) + (1-p)v$$

$$\left.\begin{aligned}\Delta p = p' - p &= -up + v(1-p) \\ &= -(u+v)p + v\end{aligned}\right\} \Rightarrow \hat{p} = \frac{v}{u+v}$$

• Observations:

(1) Since u, v are small, Δp due to mutation is typically small.
   – Mutation is a "weak" evolutionary force.

(2) Δp is linear in p.
   – I.e., in contrast to selection no term like $p(1-p)$.
   – rate of creation of new mutations depends on both
      • the frequency of "mutable" copies (e.g., p).
      • the mutation rate per copy (u).

(3) $\hat{p}$ makes sense.
   – proportional to relative mutation rates.
   – at equilibrium, the number of A's converted to a's = number of a's to A's.

(4) Can rewrite above rate equation as $\Delta p = -(u+v)(p - \hat{p})$
   – Shows that the rate of evolution towards equilibrium is proportional to
      • the total mutation rate;
      • the deviation of the current allele frequency from the equilibrium,
         $\hat{p} = v/(u+v)$.

(5) The rate of mutation does not depend on genetic variance.
   – No "variance" factor in the equation for Δp .
   – Mutation as an evolutionary force does not require pre-existing variation
      • in fact, it's <u>most</u> effective as an evolutionary *force* when $p = 0$ or $p = 1$!

• How fast is equilibrium approached?

   – $p_t - \hat{p} = (1 - u - v)^t (p_0 - \hat{p})$

   – Answer: <u>very</u> slowly.
      • e.g., when $u = v = 10^{-6}$, it takes almost 350,000 generations to evolve half
         way to the equilibrium $\hat{p} = \hat{q} = 0.5$.

• <u>Punch Line</u>: Mutation is an extremely weak force in changing allele frequencies.

MUTATION & POPULATION SIZE

• The "randomness" of mutation:

– Have been assuming that population size is infinite.

• Allows us to treat mutation as a *deterministic* force.

• Q: But isn't mutation "random?"
  A: Which allele copies mutate is decided by chance.

• In an infinite-sized population
   – mutations are always occurring if the mutation rate is positive;
   – so mutation is "deterministic" when viewed at the population level.

• In *reality*,
   – Consider a single locus determined by a 500 bp sequence of DNA.
   – Can have $4^{500} \approx 10^{300}$ alleles
   – No real population will carry all these alleles!
   – Reason isn't mutation, rather the constraint of *finite population size*.

• Finite Population Size and Mutation-Selection interactions.

– Some qualitative conclusions, without the details (which we'll consider soon):

• *Deleterious mutations*
   – Genetic diversity not affected much by population size, as long as it's not too small.

• *Favorable mutations*
   – Adaptive evolution must wait for advantageous mutations to arise.

   – Larger populations provide more "opportunities" for advantageous mutations to appear each generation.

   – Is evolution faster in larger populations?

      • Two Possibilities:

         (1) *Selection works mainly on existing variation.*

(2) *Evolution is constrained by the appearance of unconditionally advantageous mutations.*

– Which prevails in nature is not known.

• Another possibility: *environmental change.*

– Formerly deleterious mutations become advantageous.

– In this case, evolution is limited by
  • the rate of relevant environmental changes;
  • the *qualities* of deleterious mutations that are maintained.

MUTATION & DRIFT: The infinite alleles model

• Q: How much heterozygosity is maintained if mutation can produce numerous alleles at a locus?

– Can mutation and drift alone explain observed levels of genetic variation in protein and DNA sequence data?

• Consider the infinite alleles model:

– Assume: No mutation is repeated (i.e., every new mutant allele is novel)

• Will focus on f_t.

– Why? Because as *t* becomes large, all homozygotes will have genes descended from a single ancestral mutant allele (i.e., prob.["autozygosity"] —> 0)

– Thus, f_t will thus be a direct measure of homozygosity.

– Intuitively, expect $f \nrightarrow 1$ since recurrent mutation inputs novel alleles; but what balance is reached?

• $f_{t+1} =$ Pr[neither allele mutated] x
  {Pr[descended from one gene in the last generation]
  + Pr[IBD, but *not* descended from the same gene in the last generation]}

  or

$$f_{t+1} = (1-u)^2 \left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_t \right]$$

- As $t \longrightarrow \infty$, $f_t \longrightarrow \hat{f} \approx \dfrac{1}{4Nu+1}$ = eventual probability of homozygosity (or the average frequency of homozygotes).

  – The quantity $4Nu$ is often denoted by θ so that $\hat{f} = \dfrac{1}{1+\theta}$

  – With $k < \infty$ alleles, $\hat{f} = \dfrac{1 + 4Nu/(k-1)}{1 + 4Nu\dfrac{k}{k-1}} = \dfrac{1 + \theta/(k-1)}{1 + \theta\dfrac{k}{k-1}} \rightarrow \dfrac{1}{1+\theta}$ as $k \rightarrow \infty$.

- $\hat{f}$ gives a good idea of homozygosity, but how many different *alleles* are maintained in a population of size $2N$?

  – Minimum number = 1; Maximum, $2N$

  – Define $1/\hat{f}$ = "effective number of alleles" in the population; this is directly related to the amount of heterozygosity.

    • Rationale:

      –Consider a population in H-W proportions with 3 equally frequent alleles.
      - Freq.[homozygotes] = $(1/3)^2 + (1/3)^2 + (1/3)^2 = 1/3$.
      - Then, 1/Freq. of homozygotes = $1/(1/3) = 3$ = # alleles!

      – Consider a pop. whose alleles have <u>unequal</u> freqs., 1/4, 1/4, 1/2
      - Freq.[homozygotes] = $(1/2)^2 + (1/4)^2 + (1/4)^2 = 5/8$ => $8/3 = 2.6$ "equally frequent" of "effective" alleles

    • At equilibrium, effective number of alleles is $1/\hat{f} \approx 4Nu + 1 = \theta + 1$.

  – Punchline: Substantial genetic variability is possible with mutation if $2Nu > 1$

    • i.e., at least one mutant/locus/generation

- Dynamics of Mutation & Drift

  – Q: What is the rate of substitution of selectively neutral mutations?

    1) How frequently does a new neutral mutation arise?
       Ans. $2Nu$, where u is the mutation rate to selectively neutral alleles per generation.

2) How likely is this new mutant to be fixed?
   Ans. Since alleles are "neutral", $1/2N$

3) How frequently will one neutral allele replace another?
   Ans. (how often they arise) x (how often they fix) = $2Nu$ x $(1/2N) = u$ per generation.

– Conclude Rate of neutral substitution = neutral mutation rate.

   • Note: the substitution rate is *independent of N*!

      – Why? While neutral mutations are more likely to fix in a smaller population, the rate at which they arise ($2Nu$) is small. Just the opposite is true for larger pops. The result is that substitution rate is independent of a population's size.

– Property is consistent with a central observation motivating the "molecular clock" hypothesis

– Property is also central to the "neutral theory" of molecular evolution advanced by Kimura.

   • Problem: Some observed substitution rates are constant "per year" rather than "per generation."