Motoo Kimura

Aaron Flynn

Methods and Logic

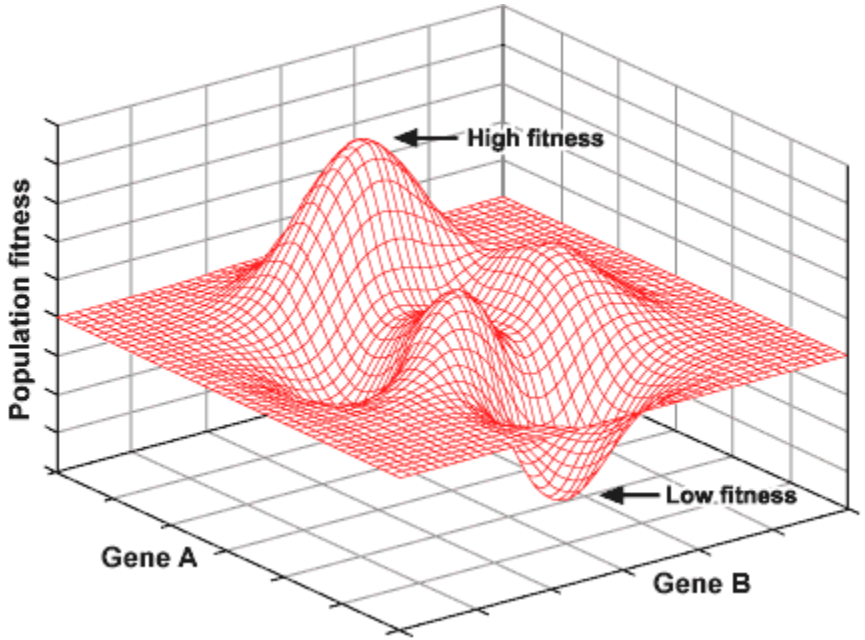
March 3, 2015

Neo-Darwinism

- Darwin proposed evolution occurs primarily by natural selection
- He did write “Variations neither useful nor injurious would not be affected by natural selection, and would be left a fluctuating element”

Theoretical Population Genetics and the Modern Evolutionary Synthesis

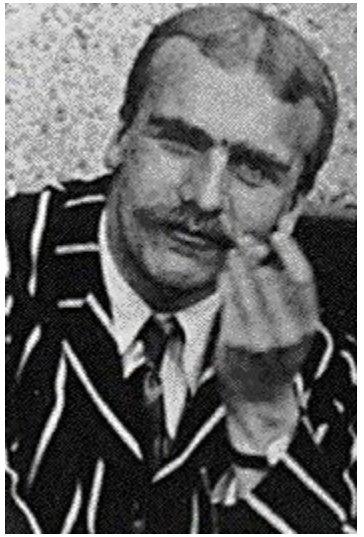
- Combined Darwinian evolution through natural selection with Mendelian heredity
- Important figures: R.A. Fisher, J.B.S. Haldane, Sewall Wright



R.A. Fisher

- Proposed the fundamental theorem of natural selection
- Coined the term “null hypothesis”
- Developed ANOVA, Fisher’s exact test

J.B.S. Haldane



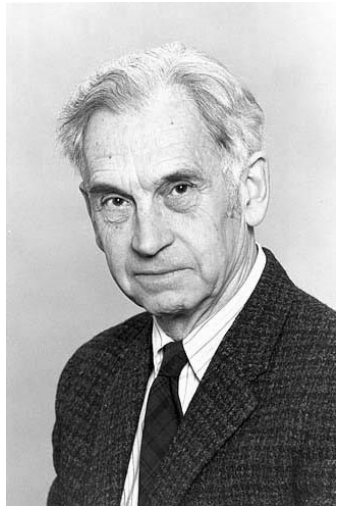
- Known for primordial soup theory and malaria hypothesis, also conceptualized in vitro fertilization and hydrogen economy
- First to demonstrate genetic linkage in mammals, first human genetic mapping
- Investigated blood acidification by drinking HCl, sealing himself in decompression chamber at 7% CO_2
- Coined many biological terms including cis, trans, coupling, repulsion, clone
- Estimated $2x10^{-5}$ mutations per gene per generation for hemophilia

Sewall Wright

- Hypothesized genetic drift and fitness landscapes
- As a first grader knew how to extract cube roots
- Disagreement with R.A. Fisher over fitness landscapes

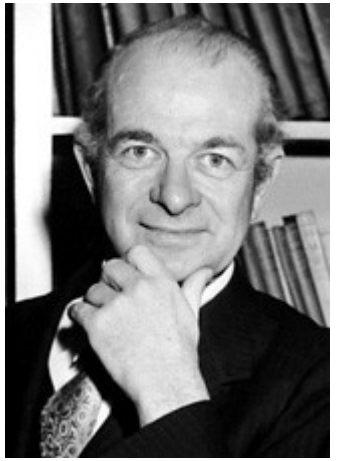
Ernst Mayr

- Proposed biological species concept: a species is defined as members of populations that actually or potentially interbreed
- Developed peripatric speciation theory to explain evolution
- Did not think the gene was a target of selection, and disagreed with mathematical approaches to genes and was a critic of Haldane, Woese



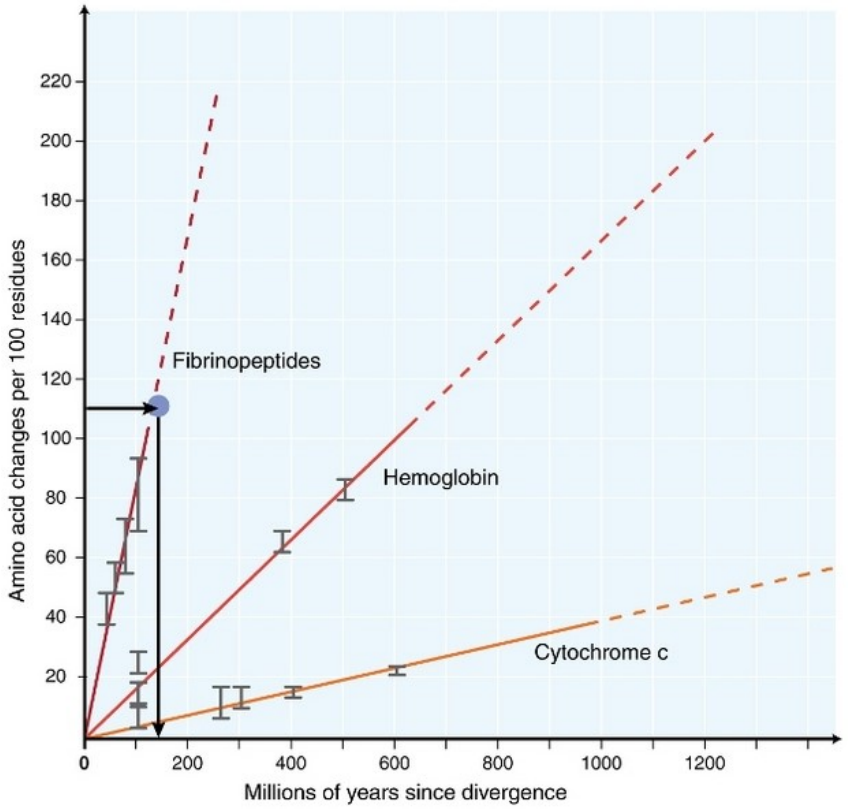
Linus Pauling

- Nobel prize in Chemistry and Nobel Peace prize
- Introduced ideas of electronegativity, orbital hybridization, resonance
- Discovered alpha helix and beta sheet, hemoglobin modification in sickle cell anemia; proposed triple helix structure of DNA



Molecular Clock Hypothesis

- Important figures: Emile Zuckerkandl, Linus Pauling, Emanuel Margoliash
- Amino acid residues in a protein and DNA bases in a genome change through spontaneous mutation at a constant rate
- Can estimate how long ago two species diverged from a common ancestor based on DNA or protein differences



Molecular Clock Hypothesis

- 1962: Emile Zuckerkandl and Linus Pauling use electrophoresis and paper chromatography on homologous proteins, notice amino acid differences in hemoglobin between vertebrates changes linearly with time when measured against fossil evidence
- 1963: Emanuel Margoliash proposes genetic equidistance hypothesis, that the number of residue differences between cytochrome c of two species is proportional to the time since the last common ancestor
- 1967: Kimura uses the molecular clock hypothesis to calculate the evolutionary rate of the entire genome of different species

Motoo Kimura:

- Born November 13, 1924
- Optimistic, outspoken, eccentric
- Interested in botany and mathematics from a young age
- Entered botany program after high school to avoid military duty during WWII
- Brought population genetics to Japan
- Diagnosed with ALS in 1993
- Died from a fall on November 13, 1994

Kimura's scientific career

- Inspired to do population genetics by Sewall Wright's 1931 paper 'Evolution in Mendelian populations'
- Japanese biologists did not understand his work, thought it was too mathematical
- He received a PhD from University of Wisconsin in 1956, also sent his dissertation to Kyoto University and it was rejected for being too mathematical in nature

Neutral theory of molecular evolution

- Neo-Darwinism became popular in the 1940s-60s, which said natural selection is more important than mutations
- Neutral theory says most mutations are neutral, fate of mutations determined by random genetic drift
- Theory: Genetic variation accounts for a large fraction of observed genetic diversity
  - Genetic variation that does not result in fitness difference means selection cannot directly affect the frequency of the variation. Genetic variation at those sites will be higher.
- Purifying selection (removal of deleterious mutations) is common
- Positive selection (increased offspring from good mutations) is rare

Evolutionary Rate at the Molecular Level *Nature* (1968)

- Abstract: “Calculating the rate of evolution in terms of nucleotide substitutions seems to give a value so high that many of the mutations involved must be neutral ones.”

Average time for one amino acid replacement in peptide

- Comparing paralogs:
  - Hemoglobin between mammals: one amino acid change per 10,000,000 years per 100 amino acids
  - Mammalian versus avian cytochrome c: one amino acid change per 45,000,000 years per 100 amino acids
  - Triosephosphate dehydrogenase between mammals: one amino acid change per 30,000,000 years per 100 amino acids
  - Average: One amino acid change per 28,000,000 years for a 100 amino acid-length peptide
  - Is 28×10^6 years per amino acid substitution low?

Average time for one base pair replacement in genome

- Assumptions:
  - Size of haploid genome: 4×10^9 bases
  - 100 amino acid-peptide corresponds to 300 nucleotide pairs in a genome
  - 20 percent of nucleotide replacement caused by mutation is synonymous
  - One amino acid replacement corresponds to 1.2 base pair replacements. Average time for one base pair replacement
- $28 \times 10^6\ yr\ \div \left(\frac{4\times10^9}{300}\right) \div 1.2 = 1.8\ yr$

Average time for one base pair replacement in genome

- Assuming mainly nucleotide substitution (mammalian genomes, similar size, similar GC content)
- Accounting for size of codon, synonymous mutations, average time for a base pair replacement in mammalian genome is 1.8 years
- But…Haldane (1957) estimated allele substitution occurred on average once every 300 generations
- No mammalian species should be able to tolerate new alleles once every 2 years

What happens with nearly neutral mutations?

- N_e: effective population size
- s: selective advantage of new allele of pre-existing alleles
- p: frequency of new allele at start
- $|2N_e s| \ll 1$
- Load:
  - $L(p) = 4N_e s\ log_e(1/p)$
  - tells us that for a nearly neutral mutation, the substitutional load can be very low and there will be no limit to the rate of gene substitution in evolution
- Probability of fixation:
  - $u(p) = p + 2N_e sp(1-p)$
  - tells us that the probability of fixation is roughly equal to initial frequency

What happens with nearly neutral mutations?

- Conclusion: New alleles may be produced at the same rate per individual as they are substituted in the population
- Neutral (or nearly neutral) mutations are occurring at the rate of roughly 0.5 per year per gamete
- Assuming average mammalian generation is 4 years, the mutation rate per generation for neutral mutations is 2 per gamete, 4 per zygote or $\frac{2}{4\times10^9} = 5 \times 10^{-10}$ per nucleotide pair per generation

Drosophila

- Esterase-5 MW 10^5 kDa, approximately 1000 amino acids or 3000 nucleotide pairs. Mutation rate would be:
  - $u = (3 \times 10^3) \times (5 \times 10^{-10}) = 1.5 \times 10^{-6}$ per generation
- Drosophila may have 10 times higher mutation rates, which would indicate 1.5×10^{-5} mutations per generation
- Assuming 1 neutral mutation per genome per generation, the mutation rate per nucleotide pair per generation in *Drosophila* is $\frac{1}{2 \times 10^8} = 5 \times 10^{-9}$ which corresponds to 1.5×10^{-5} mutations per generation for 3000 nucleotide pairs

Phage

- According to Watson (1965), phage T_4 *rIIA* gene has DNA replication error rate of $10^{-8}{\sim}10^{-9}$ per base
- Humans have approximately 50 cell divisions from fertilized egg to gamete and 4x10^9 bases per genome
- $(10^{-8}{\sim}10^{-9}) \times 50 \times (4 \times 10^9) = 200{\sim}2000$ mutations
- Conclusion: “This is 100-1000 times larger than the estimate of 2 per generation and suggests that the mutation rate per nucleotide pair is reduced during evolution by natural selection”

Concluding paragraph

- There are many more mutations occurring each generation than previously believed
- These mutations have almost no influence on fitness
- Random genetic drift alters the gene pool of populations
- Random sampling of gametes due to finite population number alters ultimate fate of gene pool
- Mutation is the driving force of evolution at both the genic and phenotypic levels

Concluding paragraph

- “To emphasize the founder principle but deny the importance of random genetic drift due to finite population number is, in my opinion, rather similar to assuming a great flood to explain the formation of deep valleys but rejecting a gradual but long lasting process of erosion by water as insufficient to produce such a result.”

Controversy

- Neo-Darwinism was dominant in biology
- Kimura paper published in Nature, same conclusion published by King and Jukes in Science in 1969 later with the title “non-Darwinian evolution”
- Kimura did not realize his theory would challenge ‘selectionism’
- Evolutionary biologists like Ernst Mayr did not think the gene was a target of selection, and disagreed with mathematical approaches to genes

Conclusions

- Evolutionary changes of proteins are due to neutral mutations and genetic drift rather than natural selection
- More evidence for molecular evolution

Sources

- Steen, TY. Always an eccentric?: a brief biography of Motoo Kimura. *J. Genet.*, Vol. 75, Number 1, April 1996, pp. 19-25.
- Yates, F., & Mather, K. (1963). Ronald Aylmer Fisher 1890-1962. *Biographical memoirs of fellows of the Royal Society of London*, *9*, 91-120.
- Pirie, N. W. (1966). John Burdon Sanderson Haldane 1892-1964. *Biographical memoirs of fellows of the Royal Society of London, 12, 218-249.*
- Crow, JF. Mayr, mathematics and the study of evolution. *J. Biol.*, Vol. 8, Number 13, 23 February 2009.
- Nei M. Selectionism and Neutralism in Molecular Evolution. *Mol. Biol. Evol.*, Vol. 22, Issue 12, December 2005, pp. 2318-2342.
- King, J. L., & Jukes, T. H. (1969). Non-darwinian evolution. *Science*, *164*(3881), 788-798.
- Crow, J. F. (1988). Sewall Wright (1889-1988). *Genetics*, *119*(1), 1.
- Kumar, S. (2005). Molecular clocks: four decades of evolution. *Nature Reviews Genetics*, *6*(8), 654-662.