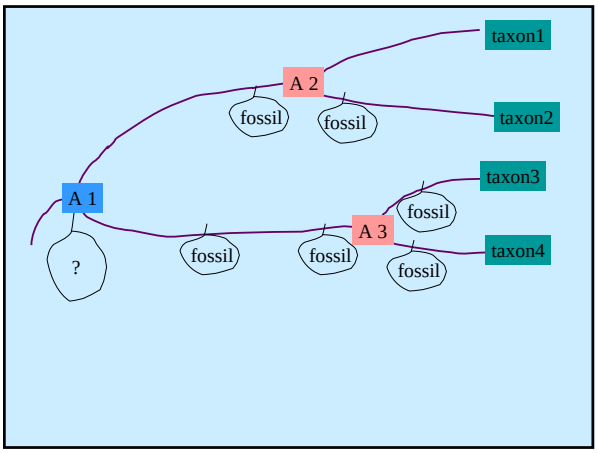
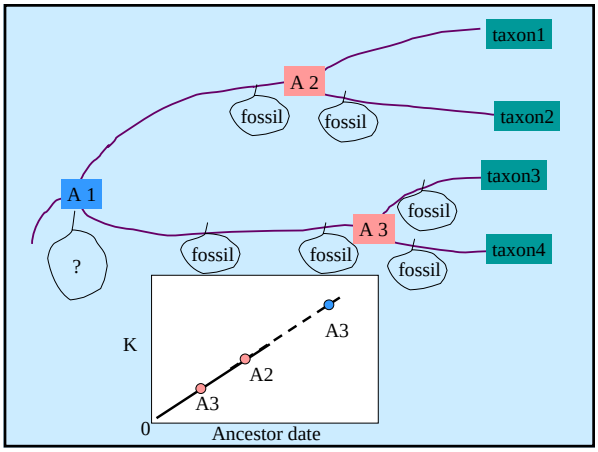
Molecular clocks

Constant rate of evolutionary change
at the molecular level (DNA and proteins)

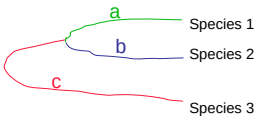
Allows use of modern sequence data to infer
ancestral dates





Relative rate tests

- Does outgroup show equivalent divergence from all members of an ingroup clade?
- Avoids the problems of calibration with respect to absolute time….
- This logic was recognized very early -- Margoliash 1963
- Walter Fitch 1976 proposed test



a+c = b+c

Various technical issues

(even if rates are completely constant)

- Estimating divergence
- Stochastic nature of changes
- Fossil dates
- Fossil record- gaps
- Placing fossil on the tree
- Phylogenetic knowledge

Estimating sequence divergence (K):

How many changes have occurred to give rise to observed differences between two sequences?

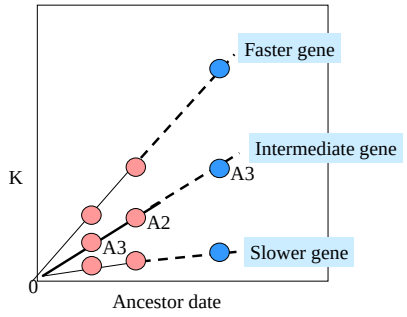
ATGCGCTAGAGGTCTAGCTAGCATGATCGACGCGATGCAAT
↓ ?
ATGCATTAGAGATCTAGCTAGCAAGATCGAGGCGATGCGAT

Need to adjust for multiple "hits" and reversions
unless divergence is very low (when #diff ~ # events)

Estimate of K is based on model of nucleotide substitution
--can be complex or simple.
For low values of divergence, model makes little difference: every change is apparent.

For high divergence values, model can make a big difference.

Using multiple genes--with different rates of evolution--can avoid the problem of rate idiosyncrasies in individuals genes



Some proteins faster than others due to different extent of purifying selection, -- numbers of constrained sites. These patterns more evident as more data were collected.

Examples of rates based on human-rodent (from Li)
Units are substitutions per 10^9 years

Protein	Ka	Ks
Histone 3	0.00	6.4
Actin beta	0.03	3.1
Insulin	0.13	4.0
Alpha globin	0.55	5.1
Fibrinogen	0.55	5.8
Albumin	0.91	6.6

What first led to the idea of the molecular clock?

a. The Neutral Theory of Molecular Evolution Kimura

b. Discovery of the genetic code

c. Protein sequence data for mammals

Zuckerkandl and Pauling 1962, globin sequences in mammals

Early kinds of molecular data

- Protein sequences (1955-->)
- DNA-DNA hybridization
- Ribosomal RNA profiles
- Restriction map profiles (mitochondrial DNA)
- Allozymes

- All used in early molecular clock studies

The Neutral Theory of Molecular Evolution

Motoo Kimura (1924-1994)

1. Theoretical framework for understanding how mutation and selection act on DNA sequences during evolution.
The behavior of neutral alleles provides a null expectation allowing us to recognize the action of selection.

2. Proposal that MOST variation at the molecular level is neutral or nearly neutral.

How does the molecular clock relate to Neutral Theory?-1

- Idea of molecular clock was empirical, initially without explanation, and came first (1962-5)
- This was before we had direct DNA or RNA sequence comparisons
- Evidence for clock-like behavior was a major source of evidence leading Kimura to propose the Neutral Theory of Molecular Evolution (1968)
- If the large majority of molecular changes occuring during evolution are neutral, why might we expect constant rates of evolution?

How does the molecular clock relate to Neutral Theory?-2

- One problem
- What are the time units for the rate (μ) of substitution?
- Mutation -- mistake in replication of DNA -- considered to be dependent on number of replications
  - Roughly proportional to generation time
- So short-lived species should evolve faster than long-lived ones.
- Most studies in mammals, based on proteins, gave roughly constant rates despite major differences in generation times (primates v rodents)

Tomoko Ohta, student of Kimura developed Nearly Neutral Theory

Idea that a large proportion of new Mutations are slightly deleterious

Expectation that the fixation of These will depend on N_e

Why?

Motoo Kimura Neutral Theory	Tomoko Ohta Nearly Neutral Theory
Most evolution at the molecular level is neutral with respect to fitness. Population size does not affect rates.	Many changes are slightly deleterious & fixed by drift. The rate of such changes is higher in small populations.

Negative correlation between population size and generation time, for limited number of species represented by sequences in ~1980.

C.I. Wu, W. H. Li. 1985 Evidence for higher rates of substitution in rodents than in man. PNAS.

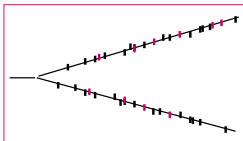
Mammalian rates--

Showed that rodents were
~2x faster for silent sites
~1.3x faster for replacement sites

Hypothesized generation time effect causing greater speedup in the silent sites.

Greater effectiveness of purifying selection reduces rodent rate at replacement sites.

Strength of purifying selection varies among sites.



Change at site under purifying selection (e.g., replacement site)

Change at neutral site (e.g., silent site, pseudogene)

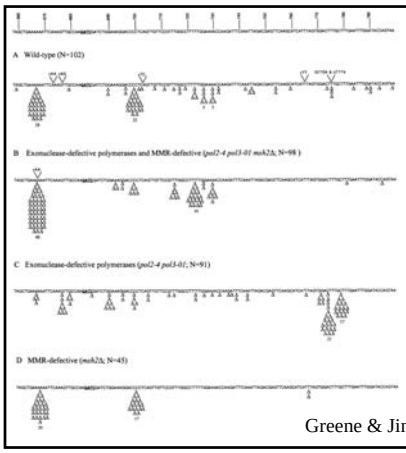
K (neutral sites) > K (selected sites)
Eg: Ks > Ka

K	Substitutions per site
K_a	Substitutions per replacement site (changes amino acid)
K_s	Substitutions per silent site (no change in amino acid)

```
GCA --> GCC          GCA --> GTA
Ala     Ala    Ala   Val
```

What about mutation rates?

- Evidence now shows that they can vary among regions of the genome and among lineages



Counts of mutations in 150 bp region of yeast strains with different repair gene mutations

Greene & Jinks-Robertson 2001

Mutation rates estimated from specific loci in animals
from lab studies or geneaological studies (humans)

Organism	μ(b)
C. elegans	2.3×10^{-10}
Drosophila	3.4×10^{-10}
Mouse	1.8×10^{-10}
Human	5.0×10^{-11}

μ(b) Mutation rate per base pair per replication

From compilation by J Drake et al 1998 *Genetics* 149:1667

The idea of a "Global Clock"

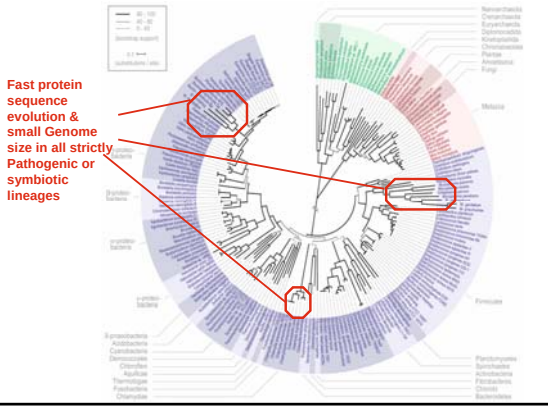
Many examples of unequal rates among lineages.
But many cases in which rates are roughly similar.

R. Doolittle et al. 1960's-1990's: constant rates of protein evolution within large groups

Ochman & Wilson 1987 constant rates in rRNA of bacteria

As more data appeared, more exceptions to rate constancy were noted.

Tree based on Sequenced Genomes Snel & Bork *Science* 2006



uses of the molecular clock

- Explore forces affecting molecular evolution
  - Male driven evolution in animals (Li)
    - Does mutation scale to absolute time or numbers of replications?
    - Does mutation scale to metabolic rate?
    - Does mutation scale to environmental temperature?
      - S. Wright et al. PNAS 2006, tropical v temperate plants
- Date nodes in evolutionary history of organisms
  - Controversies arising from discrepancies between molecular clocks and fossil-based dates

Dating metazoa-1

- Cambrian explosion at 520 Mya, general view was that
  - Origin of Metazoa was just before
  - Metazoan phyla appeared near synchronously
- Earlier Ediacaran fauna considered separate by most paleontologists
- G. Wray et al. used 7 genes to date common ancestor of Metazoa
  - Calibrated using more recent, fossil based dates, checked with relative rate tests, extended to ancestor if relative rate test did not reject equal rates of lineages
  - Arrived at date of about 1 By, almost 2x as old
- Many subsequent studies attacking/supporting Wray et al.

Dating metazoa-2

- Possible reasons for discrepancy
  - Changes in rates over time (faster in the past)
  - Saturation in divergence, so that estimates are not good
  - Poor choice of outgroups for rel rate tests
  - Fossil record is extremely incomplete

For Tuesday Oct 3.

Short questions to be emailed by Friday.

Locate one research paper on the timing of the origin of metazoa or metazoan phyla

Molecular clock or fossil paper

Be ready to present the paper.