February, 2002
Genetics 453

Evolutionary Genetics

Coalescents

Joe Felsenstein

Department of Genome Sciences
University of Washington, Seattle

email: joe@genetics.washington.edu
Fig. 3  a, Genealogical tree for 134 types of human mtDNA (133 restriction restriction sites) used. The tree accounts for the site differences observed and confirms the inferred human population at an earlier time than this one. The
A random-mating population
A random-mating population
A random−mating population
A random-mating population
A random-mating population
A random-mating population
A random-mating population
A random-mating population
A random-mating population
A random-mating population
A random-mating population
A random-mating population
Genealogy of gene copies, after reordering the copies
Genealogy of a small sample of genes from the population
under the Wright–Fisher model

each gene comes from a random copy in the previous generation

a chance of \( \frac{1}{2N} \) that another one comes from the same copy

hence it takes about \( 2N \) generations for two lineages to coalesce

\[ 
\text{under the Wright–Fisher model}
\text{each gene comes from a random copy in the previous generation}
\text{a chance of } \frac{1}{2N} \text{ that another one comes from the same copy}
\text{hence it takes about } 2N \text{ generations for two lineages to coalesce}
\]
Coalescent trees of gene copies within species (Kingman, 1982)

Random collision of lineages as go back in time (sans recombination)

Collision is faster the smaller the effective population size

\[ \text{Average time for } k \text{ copies to coalesce to } k-1 = \frac{4N_e}{k(k-1)} \]

Average time for two copies to coalesce = \( 2N_e \) generations

Average time for \( n \) copies to coalesce = \( 4N_e \left( 1 - \frac{1}{n} \right) \) generations

In a diploid population of effective population size \( N_e \)
Two levels of variability

(1) Randomness of mutation

\[ \text{affected by the mutation rate } u \]

\[ \text{can reduce variance of number of mutations per site per branch by examining more sites} \]

(2) Randomness of coalescence of lineages

\[ \text{affected by effective population size } N_e \]

\[ \text{coalescence times allow estimation of } N_e \]

\[ \text{can reduce variability by looking at} \]

(i) more gene copies, or

(ii) more loci
Coalescent with population size changing through time

During a population bottleneck there is expected to be a burst of coalescence down near the root of the tree.

Effects of population size become difficult to see.
A recombining coalescent

Different markers have slightly different coalescent trees
"Out of Africa" hypothesis

Europe

Asia

Africa

(Vertical scale is not time or evolutionary change)

Tree of gene copies, compared with the phylogeny of the species for the case in which effective population size is small compared to the number of generations between speciations.

The tree of gene copies has topology consistent with the phylogeny.
Tree of gene copies, compared with the phylogeny of the species for the case in which effective population size is large compared to the number of generations between speciations.

The tree of gene copies has topology inconsistent with the phylogeny.

East African species related to *D. melanogaster*

- East Africa
  - *D. melanogaster*
  - *D. simulans*

- Seychelles: *D. seychellea*

- Mauritius: *D. mauritianus*

- Madagascar
This freeware-friendly presentation prepared with

- Linux (operating system)

- PDFLaTeX (mathematical typesetting and PDF preparation)

- Idraw (drawing program to modify plots and draw figures)

- Adobe Acrobat Reader (to display the PDF in full-screen mode)

(except that we had to use Microsoft Windows to project this as the X server I have in Linux is not too great)