Molecular Phylogenetics

Basis of Phylogenetics
Molecular Phylogenetics

Basis of Molecular Phylogenies

Overview

- Phylogenetics Definitions
- Genetic Variation and Evolution
- Source of Information for Phylogenetic Analyses
- Molecules As Documents of Evolutionary History
- Morphology vs. Molecules
- Molecular Tools
  - Molecular Data
  - Which Gene to Use?
- Applications of Molecular Phylogenetics
Molecular Phylogenetics

Basis of Phylogenetics

Phylogenetics Definitions

- field of biology which does deal with identifying and understanding the evolutionary relationships among the many different kinds of life on earth, both living (extant) and dead (extinct)
- procedures for constructing evolutionary trees
  - a phylogeny describe the historical relationships among lineages or organisms or their parts (genes)
- using trees to test hypotheses
Molecular Phylogenetics

Basis of Phylogenetics

The Evolution of Species or Genes Can be Modeled as a Bifurcating Process

- two populations become reproductively isolated and diverge due to random mutational processes

- over time, this process may repeat itself, so that at any time, each population can be said to be most closely-related to some other population with which it shares a direct common ancestor

If genomes evolve by accumulation of mutations

- the amount of nucleotide sequence difference between a pair of genomes should indicate how recently those two genomes shared a common ancestor
Basis of Phylogenetics

Phylogenetic Divergence

- consists of changes in characters, such as amino acids in a protein, or nucleotides in DNA

  the longer two populations remain reproductively isolated, the more divergence will occur

Given the existence of homologous characters across a set of populations, it should be possible to work backwards in time, ascending the tree, until a common ancestor of all populations in the set is reached.
How does genetic variation occur?

- Every organism possesses a **genome** that contains all the biological information needed to construct and maintain a living example of that organism.

- The biological information contained in a genome is encoded in the nucleotide sequence of its DNA or RNA molecules and is divided into discrete units (**genes**).

- The information stored in a gene is read by proteins, which attach to the genome and initiate a series of reactions (**gene expression**).
DNA segments (large or small) can be altered, duplicated, or deleted.

**Genetic Variation Exists and Evolution Depends on It**

**How does genetic variation occur?**

DNA segments (large or small) can be:
- altered
- duplicated
- deleted

**Point mutations** or other small changes (ex. $A \rightarrow G$) generate a new version of a gene (i.e., a new allele).

**New loci** are generated by gene duplication events.
Molecular Phylogenetics

Basis of Phylogenetics

Genetic Variation Exists and Evolution Depends on It

How does genetic variation occur?

DNA replication
- every time a cell divides, it must make a complete copy of its genome
- must be extremely accurate in order to avoid introducing mutations, or changes in the nucleotide sequence of a short region of the genome

Inevitably, some mutations occur
- errors in DNA replication
- effects of chemical agents or radiation
Molecular Phylogenetics

Basis of Phylogenetics

Genetic Variation Exists and Evolution Depends on It

How does genetic variation occur?

Silent mutations
- result in a change that has no effect on the functioning of the genome
- include virtually all changes that happen in the noncoding components of genomes

Mutations in somatic cells
- important only for the organism in which it occurs and has no potential evolutionary impact because a somatic cell does not pass on copies of its genome to the next generation

Mutations in germ cells
- can be transmitted to the next generation and will then be present in all the cells of an individual who inherits that mutation
How does genetic variation occur?

New alleles appear in a population because of mutations that occur in the reproductive cells of an organism.

- many genes are polymorphic
- two or more alleles for that gene are present in a population
- each of these alleles has its own allele or gene frequency

Allele frequencies vary over time by

- natural selection
- random drift
Molecular Phylogenetics

Basis of Phylogenetics

What is phylogenetic analysis and Why should we perform it?
Inferring a phylogeny is an attempt to produce a best estimate of an evolutionary history based upon incomplete information. We do not have direct information about the past; we only have access to contemporary species and molecules.
Phylogenetic analysis has two major components:

1. **Phylogeny Inference or “Tree Building”**
   - The inference of the branching orders, and ultimately the evolutionary relationships between “taxa”
   - (entities such as genes, species, populations, etc.)

2. **Character and Rate Analysis**
   - Using phylogenies as analytical frameworks for rigorous understanding of the evolution of various traits or conditions of interest

---

**Molecular Phylogenetics**

**Basis of Phylogenetics**

What is phylogenetic analysis and Why should we perform it?

Phylogenetic analysis has two major components:

- **Phylogeny Inference or “Tree Building”**
  the inference of the branching orders, and ultimately the evolutionary relationships between “taxa”
  (entities such as genes, species, populations, etc.)

- **Character and Rate Analysis**
  using phylogenies as analytical frameworks for rigorous understanding of the evolution of various traits or conditions of interest
Basis of Phylogenetics

Premises to Phylogenetic Inferences

Tree-like model of evolution (bifurcating process)

- inheritance of ancestral characters
- existence of an evolutionary history defined by changes in these characters

Phylogenetic analysis requires careful thought

- phylogenetic analysis is frequently treated as a black box into which data are fed (often gathered at considerable cost) and out of which The Tree springs
**Molecular Phylogenetics**

**Basis of Phylogenetics**

**Source of Information for Phylogenetic Analyses**

**Pre-Molecular Tools**

- Paleontological record
Molecular Phylogenetics

Basis of Phylogenetics

Source of Information for Phylogenetic Analyses

Pre-Molecular Tools
- Paleontological record
- Morphology
Molecular Phylogenetics

Basis of Phylogenetics

Source of Information for Phylogenetic Analyses

Pre-Molecular Tools

- Paleontological record
- Morphology
- Behaviour
Molecular Phylogenetics

Basis of Phylogenetics

Source of Information for Phylogenetic Analyses

Pre-Molecular Tools
- Paleontological record
- Morphology
- Behaviour
- Genetics
  - chromosomes
  - number
  - structure
  - gene frequencies
Molecular Phylogenetics

Source of Information for Phylogenetic Analyses

Pre-Molecular Tools
- Paleontological record
- Morphology
- Behaviour
- Genetics
  - chromosomes
    - number
    - structure
  - gene frequencies

Molecular Tools
Molecular Phylogenetics

Basis of Phylogenetics

Morphological trait comparisons work for finches and other macroorganisms...
Basis of Phylogenetics

...but does not work for bacteria

B. anthracis

B. japonicum
Molecular Phylogenetics

Basis of Phylogenetics

...but does not work for bacteria

Previous view

bacteria and other microbes are “living fossils” in the course of evolution

“Primitive” microbes as “evolutionary dead-ends”
Molecular Phylogenetics

Basis of Phylogenetics

...but does not work for bacteria

Today’s dominant view

all organisms have evolved, and continue to evolve through time along divergent lines
Zuckerkandl & Pauling (1965)

“We may ask the question where in the now living systems the greatest amount of information of their past history has survived and how it can be extracted”

“Best fit are the different types of macromolecules (sequences) which carry the genetic information”
Molecular Phylogenetics

Basis of Phylogenetics

Molecules As Documents of Evolutionary History

Molecular Phylogenetics

- uses individual nucleotides or amino acids as characters to reconstruct evolutionary history

Zuckerkandl & Pauling (1965)

- comparison of hemoglobin amino acid sequences from a number of species
- molecular clock hypothesis

Fitch & Margoliash (1967)

- comparison of cytochrome c amino acid sequences from a number of species and phylogenetic reconstruction

Wilson & Sarich (1967)

- calibrate genetic distance derived from molecular similarity with time span of fossil record and dated human/chimp divergence at 5 million years
Molecular Phylogenetics

Basis of Phylogenetics

Molecules As Documents of Evolutionary History

The rapid development and power of molecular techniques has produced a **euphoria in evolutionary biology**

- because so many new problems can be addressed, it is a commonly held **misconception** that **all evolutionary problems are solvable with molecular data**

**Technological advances** have made possible for evolutionary biologists to easily obtain DNA sequences for phylogenetic purposes

- **PCR (Polymerase Chain Reaction) technology**
- **DNA sequencing** has become a common laboratory procedure
- **Computer hardware and software** development

**Rapid accumulation of phylogenetic data** encompassing a broad range of topics and applications
Basis of Phylogenetics

Molecules As Documents of Evolutionary History

New journals emphasizing the role of molecular studies in phylogenetic inference
Molecular Phylogenetics

Basis of Phylogenetics

Molecules As Documents of Evolutionary History

Books and Manuals
Molecular Phylogenetics

Morphology vs. Molecules
Molecular Phylogenetics

Morphology vs. Molecules

When we study DNA, we are looking directly at the genetic material. We know that what we are looking at is inherited.

DNA in modern organisms has been passed down ancestral organisms.

DNA should reflect ancestry and be reliable for studying phylogenetic relationships.
**Molecular Phylogenetics**

**Morphology vs. Molecules**

*When we look at a structure of an organism*

We don't know exactly how it is inherited.

- Some aspects of the structure may depend on direct environmental influences and not be inherited at all.

- Since we want to study inherited characteristics to reconstruct phylogenetic relationships, this could make morphology less reliable for phylogenetic studies.
Morphology vs. Molecules

Existence of Variation

One of the problems with studies of morphology is that we may be looking at a group of species that are extremely morphologically similar. Amphibians are highly conserved in terms of skeletal structure, muscles, organs, etc.

Difficult to find a large enough number of characters to do a reliable study of systematics (we need a large number of characters because some may show convergent evolution)
Molecular Phylogenetics

Morphology vs. Molecules

Existence of Variation

DNA sequences have turned out to be variable among species.

It is generally possible to find a large number of characters (base pair positions) that have phylogenetically informative variation.
Morphology vs. Molecules

Cost

It is much less expensive to study morphology than DNA for the same cost we may be able to study more aspects of body structure than we can genes.

DNA technology becomes less expensive and faster all the time, so this disadvantage to morphology is becoming less important.
Molecular Phylogenetics

Morphology vs. Molecules

The Fossil Record

We can include fossil species in studies based on morphology, but, with very rare exceptions, we cannot get DNA from fossils. Studies based on morphology:

- May allow us to find outgroups that are more closely related to the ingroup and therefore more reliable, and to have a better representation of the ingroup.
- Can be superior to those based on molecules, if we are looking at a group for which a good fossil record exists.

Neandertal, Saccopastore (Italy)
Morphology vs. Molecules

Determination of Homology

Morphological Characters

Similar looking structures may be considered homologous, when they are really coded by different genes. This makes the phylogeny unreliable.
Molecular Phylogenetics

Morphology vs. Molecules

Determination of Homology

Similar problems can arise with DNA

Gene Duplication

duplicated copies of the gene both evolve
we might compare the original form of the gene in one species with the duplicate form of the gene in another species
paralogous genes

Alignment Problem

there are often insertions or deletions of several base pairs into a gene so that the same gene in different species has different length
Molecular Phylogenetics

Morphology vs. Molecules

Convergent evolution

Convergent evolution suggests that species may be related, because they have similar characteristics, when in fact those species are NOT related.

When species independently occur in similar environments, they are likely to evolve similar morphological traits through natural selection.
Molecular Phylogenetics

Morphology vs. Molecules

Convergent evolution

The Neutral Theory of Molecular Evolution (Motoo Kimura, 1968)

Discovery of ubiquitous genic polymorphisms in natural populations brought about a great deal of debate over the significance of the variation.

Kimura’s theory

the great majority of the differences in macromolecular structures observed between individuals in a population are of no adaptive significance and have no impact on the reproductive success of the individual in which they arise.
Molecular Phylogenetics

Morphology vs. Molecules

Convergent evolution

The Neutral Theory of Molecular Evolution agrees with

- the mutational molecular substitution rate observed
- the fact that the rate of the substitutions for the less biologically important part of macromolecules is greater than for the active macromolecule domains

but...

DNA data can show convergent evolution just by chance

mutations subject to convergence because there are only four possible bases (A, T, C, and G)

it is not unlikely that by chance, different species independently evolve the same base pair at the same site
Molecular Phylogenetics

Morphology vs. Molecules

Genes vs. Species

Relationships calculated from sequence data represent the relationships between genes, this is not necessarily the same as relationships between species.

Your sequence data may not have the same phylogenetic history as the species from which they were isolated.

Different genes evolve at different speeds, and there is always the possibility of horizontal gene transfer (hybridization, vector mediated DNA movement, or direct uptake of DNA).
Morphology vs. Molecules

There are advantages and disadvantages to both morphological and molecular studies of phylogeny.

Many people currently argue that to really know the phylogenetic tree of life, we need to study both kinds of characteristic.
...EVOLUTION, LIKE GRAVITY, IS JUST A THEORY.