EVOLUTIONARY GENOMICS: The Ups and Downs of Evolution
Dennis Normile

ATAMI, JAPAN--Some 200 geneticists came together last month in this hot springs resort in the foothills of Mount Fuji to celebrate the 70th birthday of renowned evolutionary geneticist Masatoshi Nei. Born and educated in Japan, Nei has spent more than 30 years at U.S. universities, most recently Pennsylvania State University, University Park, and has trained many of the scientists making presentations here. In addition to conveying their appreciation, participants discussed cancer genes, speciation, and the impact of replication timing on genetic fidelity.
Focus of our discussions

- heritable change in molecular structure and function
- in some, but not all cases, change that becomes fixed in populations or lineages

Consider molecular evolutionary changes at two levels

- Changes in DNA:
  - Point mutation; mutations of single genes; small alterations in sequence or number of nucleotides
  - Chromosomal mutations; alterations that are more extensive than point mutations; four types – deletions, duplications, inversions, translocations
  - Scope extends from point mutations in introns or exons, to changes in the size and composition of genomes

- Changes in gene products:
  - RNA
  - Proteins… polypeptide chains… amino acid sequences

Important research themes in molecular evolution include:

- Evolutionary changes in structure and function of molecules
- Reconstructing evolutionary histories of genes and organisms; molecular phylogenetics

Mutations – A Review

Example of a point mutation – a substitution in this case

Overview of four classes of chromosomal mutations
Molecular structure strongly affects function. Endorphins are brain signal molecules that contribute to good and even euphoric feelings in humans. Boxed portion of endorphin molecule is the shape that is recognized by receptor molecules on appropriate target cells in the brain. Boxed portion of morphine molecule, an opiate drug, is a close match; morphine affects emotional state by mimicking endorphin. (This discovery was a major milestone in neurochemistry)

Protein function is an emergent property that rises from the structure of the molecule

Tetrameric Human hemoglobin
Motoo Kimura advanced the Neutral Theory of Molecular Evolution in 1968. Two observations underlie the theory:

1. Most natural populations harbor high levels of genetic variation higher than would be expected if natural selection were the evolutionary force primarily responsible for influencing the level of genetic variation in populations.

2. Many mutations in sequences of genes do not alter the proteins encoded by those genes:
   - virtually always true for synonymous substitutions
   - sometimes true for non-synonymous substitutions
  1111
   - If protein function is not altered by a mutation, the allelic variant that results from that mutation is unlikely to be influenced by natural selection...

The Neutral Theory of Molecular Evolution

The Neutral Theory holds that, because most mutations are selectively neutral at the molecular level...

- the majority of evolutionary change that macromolecules undergo results from random genetic drift
- much of the variation within species results from random genetic drift

- Kimura developed a mathematical model showing that the rates at which neutral substitutions accumulate is a function of the mutation rate (not to selection forces)
- by this theory, levels of molecular variation in genomes are strongly influenced by a balance between mutation, which generates variations, and genetic drift, which can eliminate it.
**Functional importance** of an amino acid *varies* with, among other things, location in the molecule.

Some portions of the molecules phenotype are susceptible to neutral evolution, other areas to adaptive evolution.

---

**Molecular Evolution of Lysozyme**

- Evolution of Novel Gene Function
- Convergent Evolution
Lysozyme and the evolution of novel function in three lineages of animals

Evolution of Function in Lysozyme

• Presumed original (ancestral) function: defense against Pathogens
  - occurs in almost all animals
  - kills bacteria by digesting polysaccharide in cell wall

• Evolved (derived) function: digestion/nutrition
  - functions in this capacity in select lineages, including
  - ruminant hooved mammals
  - langurs (primates)
  - Hoatzin (bird)

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>LANGUR</th>
<th>BABOON</th>
<th>HUMAN</th>
<th>RAT</th>
<th>COW</th>
<th>HORSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langur*</td>
<td></td>
<td>14</td>
<td>18</td>
<td>38</td>
<td>32</td>
<td>65</td>
</tr>
<tr>
<td>Baboon</td>
<td>0</td>
<td></td>
<td>14</td>
<td>33</td>
<td>39</td>
<td>65</td>
</tr>
<tr>
<td>Human</td>
<td>0</td>
<td>1</td>
<td></td>
<td>37</td>
<td>41</td>
<td>64</td>
</tr>
<tr>
<td>Rat</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>55</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Cow*</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Horse</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

* species with foregut fermentation
above diagonal: number of differences
below diagonal: number of amino acids uniquely shared
## Lysozyme function and foregut fermentation

Comparisons of Lysozyme Amino Acid Sequences of Different Species.

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>LANGUR</th>
<th>BABOON</th>
<th>HUMAN</th>
<th>RAT</th>
<th>COW</th>
<th>HORSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langur*</td>
<td>14</td>
<td>18</td>
<td>38</td>
<td>32</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Baboon</td>
<td>0</td>
<td>14</td>
<td>33</td>
<td>39</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>0</td>
<td>1</td>
<td>37</td>
<td>41</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>55</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Cow*</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Horse</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

* species with foregut fermentation

above diagonal: number of differences
below diagonal: number of amino acids uniquely shared

- **Rapid evolution of lysozyme in langurs**

- **Convergent evolution of lysozyme in langurs and cows**

* species with foregut fermentation

above diagonal: number of differences
below diagonal: number of amino acids uniquely shared
**Transposons** 1000’s of copies scattered around genome

**Tandem Clusters** Clusters containing hundreds of nearly identical copies of a gene

**Multigene Families** Clusters of a few to several hundred copies of related but distinctly different genes

**Satellite DNA** Short sequences present in millions of copies per genome

**Dispersed Pseudogenes** Inactive members of a multigene family separated from other members of the family

**Single-copy Genes** Genes that exist in only one copy in the genome

---

**Genome Evolution**
Genomes continually evolve

Multiple copies of genes have evolved, some then diverging in sequence to become different genes, which in turn have duplicated and diverged (applies to other DNA sequences as well.

**Gene family** two or more genes in a genome, identical or highly similar in nucleotide sequence descended from the same ancestral gene

**Origin of gene families** Repeated gene duplication from errors during DNA replication and recombination

**Hemoglobin Evolution** is an excellent case study

- Gene duplication
  - Multigene Families
- Evolution of molecular function

**Evolution of the Globin Gene** Genes encoding proteins have undergone continual evolution, accumulating increasing numbers of changes over time. Length of lines corresponds to number of nucleotide substitutions in the gene
Globin gene families are well-studied across taxa for sequence, structure and function.

- Hemoglobin multigene families in humans
  - Alpha globin family (on chr. 16)
  - Beta globin family (on chr. 13)
  - Hemoglobin families probably descended from a myoglobin-like ancestral gene

The family of Globin genes. One ancestral form eventually diverged into the 11 forms found in the human genome.
Comparison of Cytochrome c Among Lineages

• Illustrate concepts and principles of molecular evolution
• Illustrate application of our understanding of molecular evolution

Structure of Cytochrome c in *rice*  
Structure of Cytochrome c in *tuna*
Molecules with constant rates of change

• In certain molecules, many of the changes that occur over time involve nucleotide or amino acid substitutions that do not affect the functioning of the molecule and therefore do not affect the fitness of the organism

• Such neutral changes are not influenced by natural selection and therefore accumulate at a rate roughly equal to the mutation rate

• If adaptive changes are few compared to neutral changes, differences between taxa can be used to date lineage separations; rates of divergence at the nucleotide or amino acid level among lineages may serve as “molecular clocks”
Cytochrome C

- Amino Acid sequence known for ~100 species
- Some regions accumulate changes relatively quickly
- Some residues are invariant, those that interact with heme-group, essential to enzyme function
- Alterations may have
  - may have neutral effect
  - may represent minor adaptive species-specific modifications
- Some may indicate major changes in the enzyme (e.g., polar \(\rightarrow\) nonpolar substitution)
- Overall, evolutionarily conservative enzyme; suggests only minor alterations are tolerable

Cytochrome c amino acid sequences, across taxa
Molecules evolve at different rates, some, at constant rates.

**Evolution of Cytochrome c.** Some molecules, including this one, evolve at a constant rate — may be useful as a "molecular clock" in that it provides a tool for reconstructing phylogenies.