Lecture 7
Mutation and genetic variation
Thymidine dimer
Repair endonuclease

Damage in one strand of DNA

1. Excision of damaged DNA

2. DNA polymerization to fill gap

3. Remaining nick sealed by DNA ligase
2. Purifying selection

- a form of selection acting to eliminate harmful (deleterious) alleles from natural populations.
- example: human recessive diseases like Tay-Sachs or porphyria.
<table>
<thead>
<tr>
<th>First Letter</th>
<th>Second Letter</th>
<th>Third Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>Phe</td>
<td>T</td>
</tr>
<tr>
<td>TTT</td>
<td>TTC</td>
<td>TTA</td>
</tr>
<tr>
<td>TCC</td>
<td>TCA</td>
<td>TCG</td>
</tr>
<tr>
<td>Leu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Ser</td>
<td>T</td>
</tr>
<tr>
<td>CTT</td>
<td>CCT</td>
<td>CAT</td>
</tr>
<tr>
<td>CTC</td>
<td>CCC</td>
<td>CAC</td>
</tr>
<tr>
<td>CTA</td>
<td>CCA</td>
<td>CAA</td>
</tr>
<tr>
<td>CTG</td>
<td>CCG</td>
<td>CAG</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>A</td>
</tr>
<tr>
<td>ATT</td>
<td>ACT</td>
<td>AAT</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Asn</td>
</tr>
<tr>
<td>ATC</td>
<td>ACC</td>
<td>AAC</td>
</tr>
<tr>
<td>ATA</td>
<td>ACA</td>
<td>AAA</td>
</tr>
<tr>
<td>Met</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>ATG</td>
<td>ACG</td>
<td>AAG</td>
</tr>
<tr>
<td>G</td>
<td>Val</td>
<td>G</td>
</tr>
<tr>
<td>GTT</td>
<td>GCT</td>
<td>GAT</td>
</tr>
<tr>
<td>GTA</td>
<td>GCC</td>
<td>GAG</td>
</tr>
<tr>
<td>GTG</td>
<td>GCA</td>
<td>GAA</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Asp</td>
</tr>
<tr>
<td>G</td>
<td>GGT</td>
<td>GGT</td>
</tr>
<tr>
<td>T</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>C</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

Amino acids and stop codons are shown in the table.
Phylogenetic reconstructions

3. Maximum Likelihood

- Jukes-Cantor Model
- Kimura - 2 parameter Model
Mutation hotspots

Cytochrome b
Molecular clock

Emile Zuckerkandl

and

Linus Pauling, 1965
Synonymous substitutions
v.s.
Non-synonymous substitutions

Ka/Ks is an indicator of selection
General classes of mutations

Point mutations

“Copy-number” mutations

Chromosomal mutations

Genome mutations
Point mutations

There are four categories of point mutations:

1. transitions (e.g., A \rightarrow G, C \rightarrow T)

2. transversions (e.g., T \rightarrow A, C \rightarrow G)
Point mutations

There are four categories of point mutations:

1. transitions (e.g., A → G, C → T)
2. transversions (e.g., T → A, C → G)
3. insertions (e.g., TTTGAC → TTTCCGAC)
4. deletions (e.g., TTTGAC → TTTC)

- in coding regions, point mutations can involve silent (synonymous) or replacement (nonsynonymous) changes.
- in coding regions, insertions/deletions can also cause frameshift mutations.
Indels are insertions and deletions

STOP making sense: effective frameshifts
Loss of function mutations in the cystic fibrosis gene
“Copy-number” mutations
“Copy-number” mutations

• these mutations change the numbers of genetic elements.
“Copy-number” mutations

- these mutations change the numbers of genetic elements.

- *gene duplication* events create new copies of genes.
“Copy-number” mutations

• these mutations change the numbers of genetic elements.

• **gene duplication** events create new copies of genes.

• one important mechanism generating duplications is **unequal crossing over**.
Unequal crossing-over can generate gene duplications
Unequal crossing-over can generate gene duplications.
Unequal crossing-over can generate gene duplications

lethal?

neutral?
“Copy-number” mutations

• these mutations change the numbers of genetic elements.

• **gene duplication** events create new copies of genes.

• one mechanism believed responsible is **unequal crossing over**.

• over time, this process may lead to the development of **multi-gene families**.
α and β-globin gene families

Chromosome 11

Chromosome 16
Timing of expression of globin genes

![Graph showing the timing of expression of globin genes](image)

- **α-like family includes three functional genes:**
  - $\alpha_1$, $\alpha_2$, $\zeta$ (zeta)

- **β-like family includes five functional genes:**
  - $\beta$, $\epsilon$ (epsilon), $\delta$ (delta),
  - $G_\gamma$ (G-gamma), $A_\gamma$ (A-gamma)

Figure 5-7 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.
Retrogenes may also be created

• retrogenes have identical exon structures to their “progenitors” but lack introns!
Retrogenes may also be created

- retrogenes have identical exon structures to their “progenitors” but lack introns!

Example: jingwei in Drosophila yakuba
Retrogenes may also be created

• retrogenes have identical exon structures to their “progenitors” but lack introns!

Example: *jingwei* in *Drosophila yakuba*
Retrogenes may also be created

- retrogenes have identical exon structures to their “progenitors” but lack introns!

**Example:** *jingwei* in *Drosophila yakuba*

```
Alcohol dehydrogenase (Adh) → mRNA
```

Chromosome 2  Chromosome 3
Retrogenes may also be created

• retrogenes have identical exon structures to their “progenitors” but lack introns!

**Example**: *jingwei* in *Drosophila yakuba*
Retrogenes may also be created

- retrogenes have identical exon structures to their "progenitors" but lack introns!

Example: *jingwei* in *Drosophila yakuba*

Alcohol dehydrogenase (Adh) → Chromosome 2 → mRNA ↓ cDNA → "jingwei" → Chromosome 3
Whole-genome data yields data on gene families

Table 5.1  Sizes of gene families

These data are from analyses of protein-coding regions in whole-genome sequences. In most cases, genes were designated as members of the same family if at least 30% of the amino acids in the resulting protein are identical. Modified from Gu et al. 2002.

<table>
<thead>
<tr>
<th>Size of gene family</th>
<th>Yeast</th>
<th>Fruit flies</th>
<th>Roundworms</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>415</td>
<td>404</td>
<td>665</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>113</td>
<td>188</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>46</td>
<td>93</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>21</td>
<td>71</td>
</tr>
<tr>
<td>6-10</td>
<td>19</td>
<td>52</td>
<td>104</td>
</tr>
<tr>
<td>&gt;10</td>
<td>8</td>
<td>38</td>
<td>98</td>
</tr>
</tbody>
</table>
“Copy-number” mutations

- transposable elements (TEs) are common.
“Copy-number” mutations

• transposable elements (TEs) are common.

• three major classes of TEs are recognized:
“Copy-number” mutations

- transposable elements (TEs) are common.

- three major classes of TEs are recognized:

1. insertion sequences (700 – 2600 bp)
“Copy-number” mutations

- **transposable elements** (TEs) are common.

- three major classes of TEs are recognized:
  1. insertion sequences (700 – 2600 bp)
  2. transposons (2500 – 7000 bp)
“Copy-number” mutations

- transposable elements (TEs) are common.

- three major classes of TEs are recognized:
  1. insertion sequences (700 – 2600 bp)
  2. transposons (2500 – 7000 bp)
  3. retroelements
Chromosomal inversions lock blocks of genes together
Changes in chromosome number are common
Changes in chromosome number are common

- Robertsonian fusions and fissions are common and can have major effects on speciation
Changes in chromosome number are common

• Robertsonian fusions and fissions are common and can have major effects on speciation

• in mammals, chromosome numbers range from $N = 3$ to $N = 42$. 
Changes in chromosome number are common

- Robertsonian fusions and fissions are common and can have major effects on speciation

- in mammals, chromosome numbers range from $N = 3$ to $N = 42$.

- in insects, the range is from $N = 1$ (some ants) to $N = 220$ (a butterfly)
Changes in chromosome number are common

- Robertsonian fusions and fissions are common and can have major effects on speciation

- in mammals, chromosome numbers range from $N = 3$ to $N = 42$.

- in insects, the range is from $N = 1$ (some ants) to $N = 220$ (a butterfly)

- karyotypes can evolve rapidly!
Muntiacus reevesi

Muntiacus muntjac
Muntiacus reevesi; $N = 23$

Muntiacus muntjac; $N = 4$
Genome mutations
Genome mutations

- **polyploidization events** cause the entire genome to be duplicated.
Genome mutations

• **polyploidization events** cause the entire genome to be duplicated.

• polyploidy has played a major role in the evolution of plants.
Genome mutations

• **polyploidization events** cause the entire genome to be duplicated.

• polyploidy has played a major role in the evolution of plants.

• ancient polyploidization events have also occurred in most animal lineages.
Generation of a tetraploid

Parent

Gametes

2n + 2n

Mutation causes production of diploid gametes; self-fertilization occurs

First generation offspring

Self-fertilizes, mates with 4n sibling, or "backcrosses" to parent

Gametes

2n + 2n

Second generation offspring

4n 4n 4n

Figure 5-11 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.
Where do new genes come from?
Where do new genes come from?

An example: the antifreeze glycoprotein (AFGP) gene in the Antarctic fish, *Dissostichus mawsoni*

Convergent evolution of an AFGP gene in the arctic cod, *Boreogadus saida*