

BIOS 477/877 Bioinformatics and Molecular Evolution

Instructor: Etsuko Moriyama
(School of Biological Sciences)

Spring 2026

Lecture 3

BIOS477/877 L3 - 1

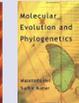
1

Today's topics

➤ Molecular Evolution

[Some textbooks]

- Molecular and Genome Evolution, 1st ed. by Graur, 2016 Sinauer Associates, Inc. [ISBN: 1605354694]
- An Introduction to Molecular Evolution and Phylogenetics, 2nd ed. by Bromham, 2016. Oxford University Press [ISBN: 9780191070693]
- Molecular Evolution and Phylogenetics by Nei & Kumar, 2000 Oxford University Press. [ISBN: 0195135857]



MEGA website
(for Windows, MacOS, Linux)

BIOS477/877 L3 - 2

2

Molecular Evolution & Bioinformatics



- “Bioinformatics and Molecular Evolution” by Higgs and Attwood (2005) Blackwell Publishing [ISBN: 1405106832]
 - Ch 1: how population genetics and molecular evolution are important and related to bioinformatics
 - Ch 3: the basic mechanisms of molecular evolution (PDF available on Canvas)



- “Fundamental Concepts of Bioinformatics” by Krane and Raymer (2003) Benjamin Cummings [ISBN: 0805346333]
 - Ch 3: presents the basics of molecular evolution (PDF available on Canvas)

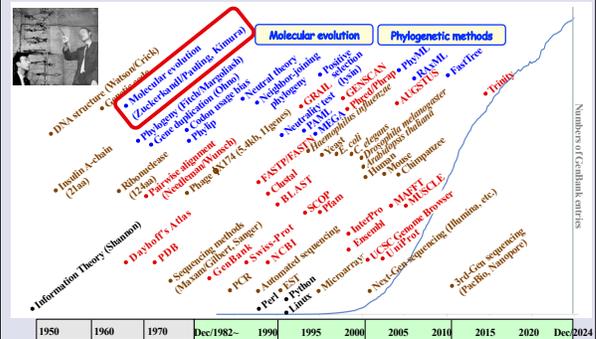


- “Bioinformatics for Beginners” by Choudhuri (2014) Elsevier [ISBN: 9780124104716]
 - Ch 2: presents the foundation of molecular evolution (PDF available on Canvas)
- (all chapters are accessible on Canvas via UNL libraries)

BIOS477/877 L3 - 3

3

History of Genomics, Bioinformatics, and Molecular Evolution



BIOS477/877 L3 - 4

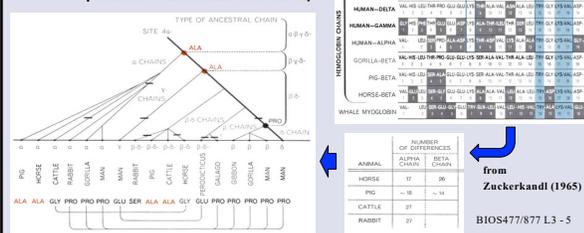
4

Molecular Evolution

➤ DNA molecules:

- ➔ Hereditary material for all living organisms
- ➔ “Documents of evolutionary history”

(Emile Zuckerkandl, one of the pioneers of molecular evolution)



BIOS477/877 L3 - 5

5

Molecular Evolution

What can we do?

- Reconstruct the evolutionary history of genes and genomes (evolution at the molecular level)
- Reconstruct the evolutionary history of populations and species (evolution at the organismal level)
- Can be tracked both in space and in time
- Attempt to build a classification of the living world
- Reconstruct the evolution of adaptation (function)
- Identify the driving forces behind the evolutionary process



BIOS477/877 L3 - 6

6

How to analyze sequences

What's happening in this sequence?

```
>Seq1
ATGGCTCTCACCAACAAGAACATCATCTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
ACCAAGTCGGGAATGGTCAAAACGGGATCTCAAGTTTGTCAACTTTAATATTTGGTT
TTGTTTCCAAAAAATCTACTTTGTTTCCCGCTGGTTAGAACTTG
```

It is not easy to find any pattern from a single sequence

BIOS477/877 L3 - 7

7

Comparing sequences

Comparison gives us a lot more power!

```
>Seq1
ATGGCTCTCACCAACAAGAACATCATCTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
ACCAAGTCGGGAATGGTCAAAACGGGATCTCAAGTTTGTCAACTTTAATATTTGGTT
TTGTTTCCAAAAAATCTACTTTGTTTCCCGCTGGTTAGAACTTG
```

```
>Seq2
ATGGCACTCACCAACAAAAACGTCACTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
ACCAAGTCGGGAATGGCTTAAAGCGGGATCTCAAGTTTGTTCGGAGAAAATCTTTGAAGG
TTCCAAAAGAAATACCTTTGTTTTTTTGTGTTTTTTGTAGAACTTG
```

BIOS477/877 L3 - 8

8

Alignment of sequences

Sequence comparison = Alignment

```
Seq1 ATGGCTCTCACCAACAAGAACATCATCTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
Seq2 ATGGCACTCACCAACAAAAACGTCACTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC

Seq1 ACCAAGTCGGGAATGGTCAAAACGGGATCTCAAGTTTGTCAACTTTAATATTTGGTT
Seq2 ACCAAGTCGGGAATGGCTTAAAGCGGGATCTCAAGTTTGTTCGGAGAAAATCTTTGAAGG

Seq1 TTGTTTTCCAAAAAATCTACTTTGTTTTCCCGCTGGTTAGAACTTG
Seq2 TTCCAAAAGAAATACCTTTGTTTTTTTGTGTTTTTTGTAGAACTTG
```

BIOS477/877 L3 - 9

9

Finding patterns

Can you see any pattern from the alignment below?

```
Seq1 ATGGCTCTCACCAACAAGAACATCATCTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
Seq2 ATGGCACTCACCAACAAAAACGTCACTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC

Seq1 ACCAAGTCGGGAATGGTCAAAACGGGATCTCAAGTTTGTCAACTTTAATATTTGGTT
Seq2 ACCAAGTCGGGAATGGCTTAAAGCGGGATCTCAAGTTTGTTCGGAGAAAATCTTTGAAGG

Seq1 TTGTTTTCCAAAAAATCTACTTTGTTTTCCCGCTGGTTAGAACTTG
Seq2 TTCCAAAAGAAATACCTTTGTTTTTTTGTGTTTTTTGTAGAACTTG
```

What do you notice in the distribution of changes?
Does any region have more (or fewer) changes?

BIOS477/877 L3 - 10

10

Finding patterns: less vs. more

Two regions have different levels of conservation

Fewer changes [more conserved]

```
Seq1 ATGGCTCTCACCAACAAGAACATCATCTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
Seq2 ATGGCACTCACCAACAAAAACGTCACTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC

Seq1 ACCAAGTCGGGAATGGTCAAAACGGGATCTCAAGTTTGTCAACTTTAATATTTGGTT
Seq2 ACCAAGTCGGGAATGGCTTAAAGCGGGATCTCAAGTTTGTTCGGAGAAAATCTTTGAAGG

Seq1 TTGTTTTCCAAAAAATCTACTTTGTTTTCCCGCTGGTTAGAACTTG
Seq2 TTCCAAAAGAAATACCTTTGTTTTTTTGTGTTTTTTGTAGAACTTG
```

Why the difference?

More changes
[less conserved]

BIOS477/877 L3 - 11

11

Difference in sequence conservation

Which region is functionally more important?

Region A [more conserved]

```
Seq1 ATGGCTCTCACCAACAAGAACATCATCTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC
Seq2 ATGGCACTCACCAACAAAAACGTCACTTTGTGGCCGGTCTGGCCGGCATTGGTCTGGAC

Seq1 ACCAAGTCGGGAATGGTCAAAACGGGATCTCAAGTTTGTCAACTTTAATATTTGGTT
Seq2 ACCAAGTCGGGAATGGCTTAAAGCGGGATCTCAAGTTTGTTCGGAGAAAATCTTTGAAGG

Seq1 TTGTTTTCCAAAAAATCTACTTTGTTTTCCCGCTGGTTAGAACTTG
Seq2 TTCCAAAAGAAATACCTTTGTTTTTTTGTGTTTTTTGTAGAACTTG
```

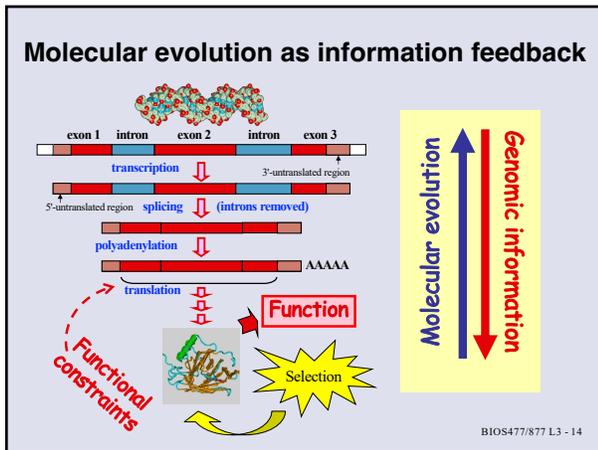
A or B



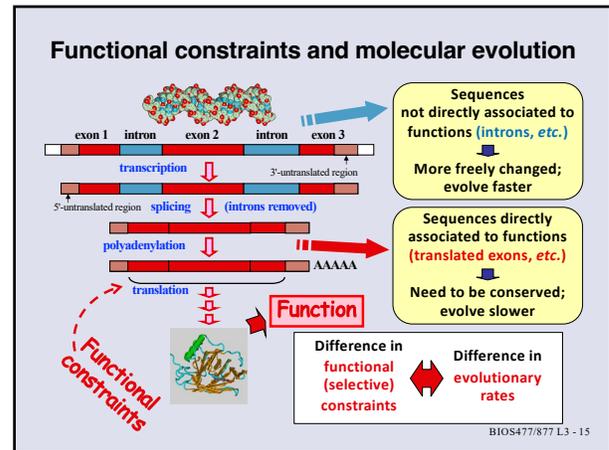
Region B
[less conserved]

BIOS477/877 L3 - 12

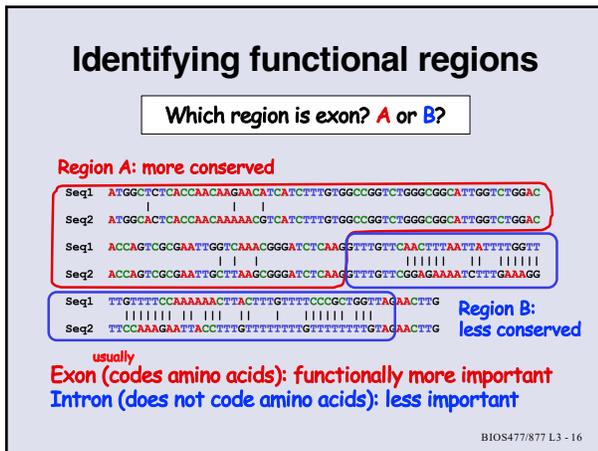
12



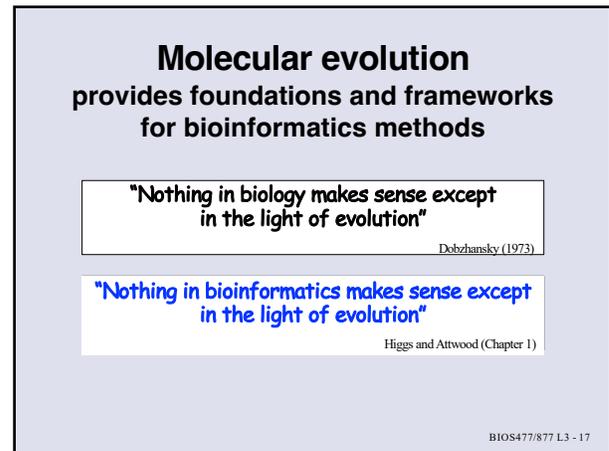
14



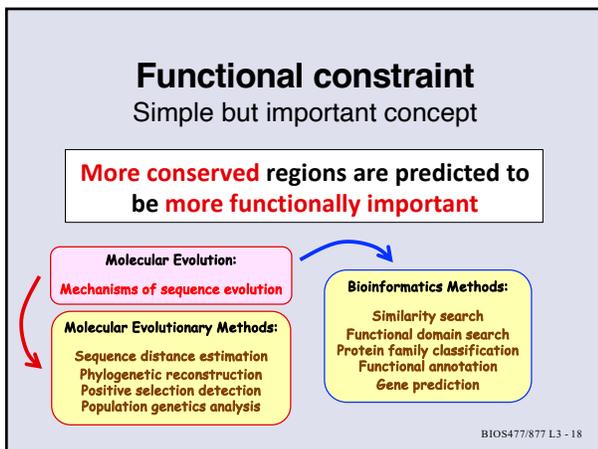
15



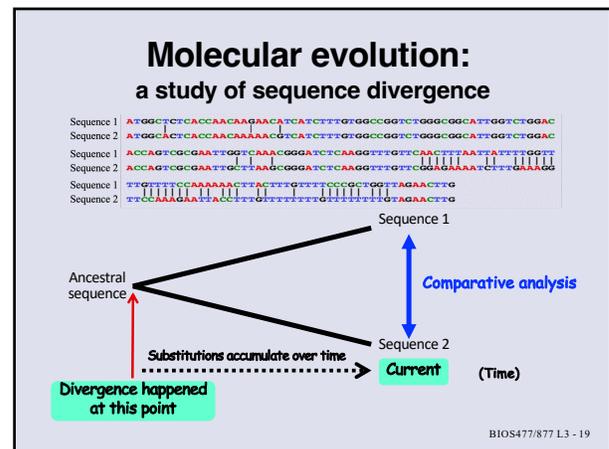
16



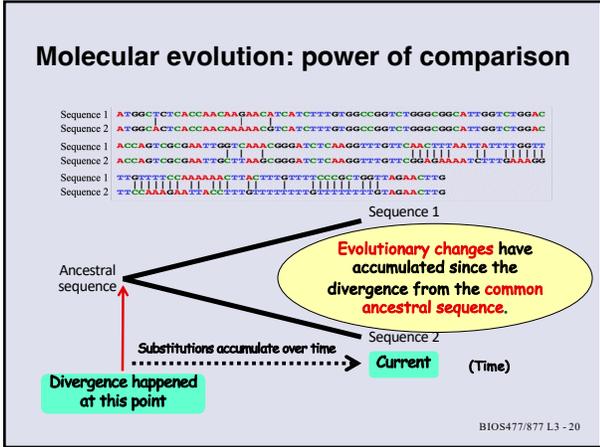
17



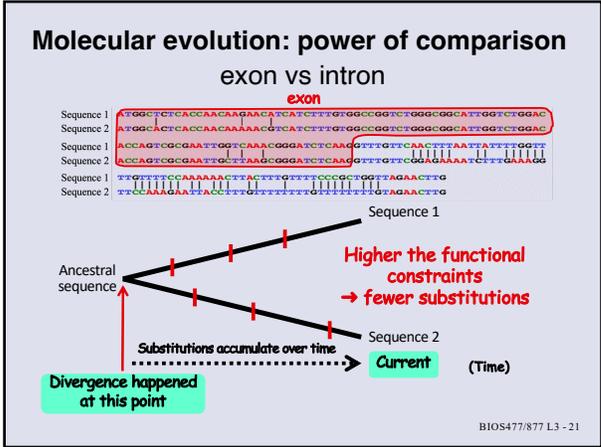
18



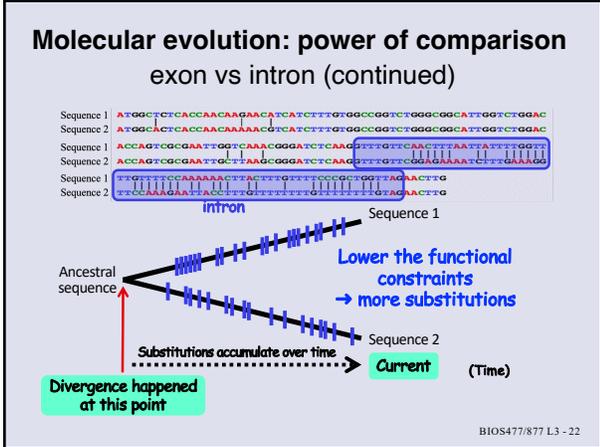
19



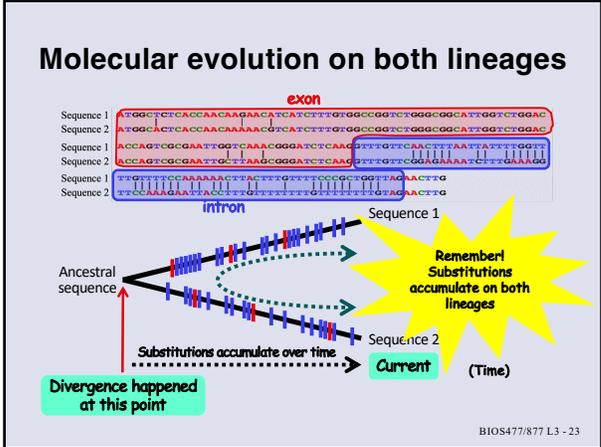
20



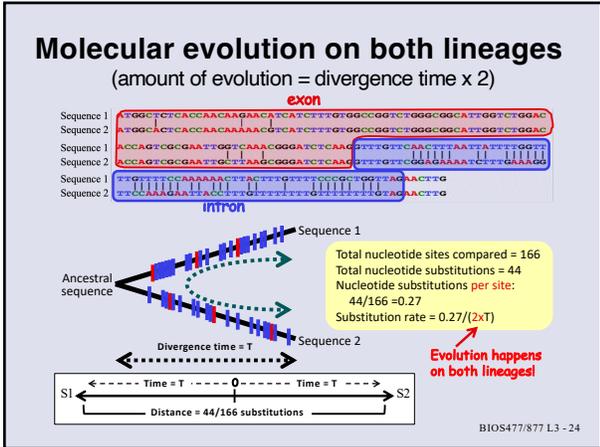
21



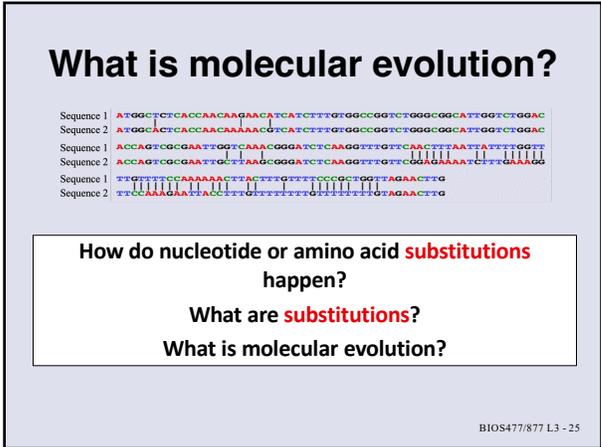
22



23



24



25

Molecular evolution at the population level

Each sequence is a single sample obtained from a large **population**.

Evolution happens **at the population level**.

BIOS477/877 L3 - 26

26

Definition of evolution

Evolution happens **at the population level**.

[Definition]
Evolution: **any change in the frequency of alleles** within a gene pool from one generation to the next. **(population)**

BIOS477/877 L3 - 27

27

What causes changes in allele frequencies?

● was changed to ●: **Mutation**
or ● moved into the population: **Migration**

BIOS477/877 L3 - 28

28

Causes of molecular evolution

What causes changes in allele frequencies?

Note: A new green mutant appeared but was lost before the next generation. Not all mutations affect the evolution.

BIOS477/877 L3 - 29

29

Causes of molecular evolution

1. Mutation and migration

- **Mutation**
 - Ultimate source of variation
 - Point mutations (e.g., A → T)
 - Recombinations
 - Insertions and deletions (indels)
- **Migration**
 - can also introduce new variation

BIOS477/877 L3 - 30

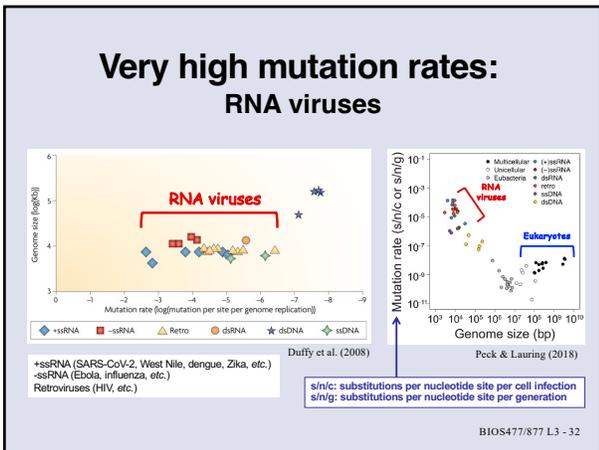
30

Rates of mutation

- Mutation rates are usually very low ($10^{-8} \sim 10^{-10}$ /base/year)
- Higher mutation rates are found in some genomes
 - due to weak (or no) repair systems, etc.
 - e.g., **animal mitochondrial DNA**
some RNA viruses (e.g., retroviruses) ($10^{-3} \sim 10^{-2}$ /base/year!)

BIOS477/877 L3 - 31

31



32

Causes of molecular evolution 2. Natural selection

➤ **Definition:**
Differential reproduction of genetically distinct individuals or genotypes within a population

- **Negative (purifying) selection:**
Deleterious mutations are selected against and removed from the population quickly
- **Positive (advantageous) selection:**
Advantageous mutations will be fixed in the population
- **Neutral mutations:**
free from natural selection

BIOS477/877 L3 - 33

33

Causes of molecular evolution 3. Random genetic drift

- **Random fluctuations** in allele frequencies produced by random sampling of gametes in the process of reproduction
- Direction of the change is random
- Its effect is larger in small populations

8 - 8

12 - 4

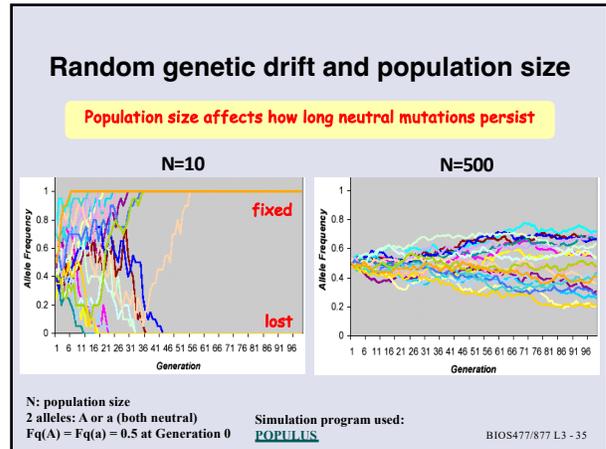
0 - 8

loss (extinction)

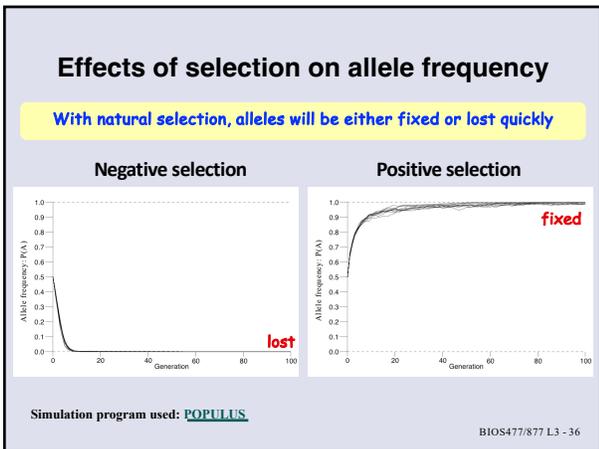
fixation

BIOS477/877 L3 - 34

34



35



36

Mutation vs. Substitution

- **Mutations** are changes in nucleotide sequences that occur due to errors in replication or repair.
- **Substitutions** are mutations that have passed through population process (drift and selection).

Note: this is unrelated to a type of genetic mutations!

Mutations (or migration)

Natural selection

Substitutions

Random genetic drift

Mutations ≠ Substitutions!!

BIOS477/877 L3 - 37

37

Mutation vs. Substitution (continued)

- **Mutations** are changes in nucleotide sequences that occur due to **errors in replication or repair**.
- **Substitutions** are mutations that have passed through population process (**drift and selection**).

Mutations ≠ Substitutions!!

- It is very difficult to directly estimate mutation rates.
- Changes observed by comparing sequences from different species are **substitutions** (after drift and selection). *e.g.*, nucleotide substitutions, amino acid substitutions

Only if no drift nor selection, substitution rate = mutation rate*

*Spontaneous (or raw) mutation rate driven by replication errors, damage, and repair errors

BIOS477/877 L3 - 38

38

Worksheet for calculating the number of nucleotide substitutions

(handout or available also on Canvas)

1. For exon and intron regions
2. For each of three codon positions
 - Count the numbers of:
 - Nucleotide sites compared
 - Nucleotide substitutions
 - Then calculate the number of nucleotide substitutions per site

Fill out the numbers and bring it Thursday!

BIOS477/877 L3 - 39

39