

# Bioinformatics

## UPGMA

MSc. Vicente Machaca Arceda

Universidad Nacional de San Agustín de Arequipa

June 5, 2021

# Table of Contents

- 1 Introduction
  - Objectives
  - Methods
  
- 2 UPGMA
  - Method

# Table of Contents

- 1 Introduction
  - Objectives
  - Methods
- 2 UPGMA
  - Method

# Introduction

## Objectives

- Understand and implement UPGMA.

# Table of Contents

- 1 Introduction
  - Objectives
  - **Methods**
- 2 UPGMA
  - Method

# Phylogenetics

## Methods

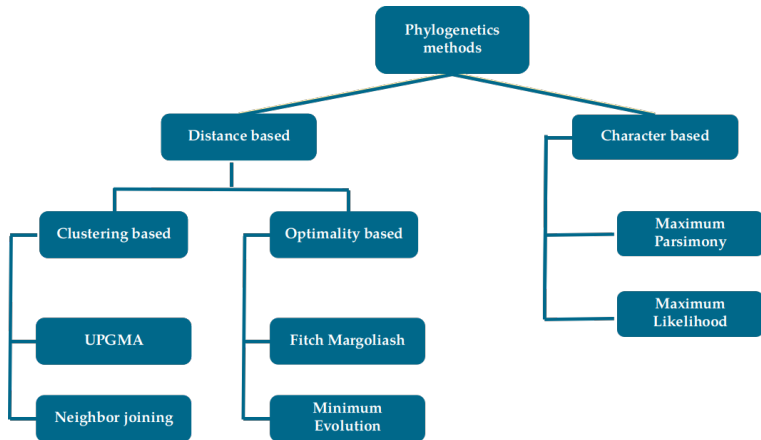


Figure: The most used methods to build phylogenetic trees.

# Table of Contents

- 1 Introduction
  - Objectives
  - Methods
- 2 UPGMA
  - Method

# UPGMA

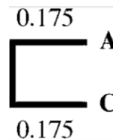
## Unweighted Pair Group Method Using Arithmetic Average

	A	B	C
B	0.40		
C	0.35	0.45	
D	0.60	0.70	0.55

Figure: Distances between sequences (score).

Choose the smallest distance (A, C). Then compute the cluster.

$$\frac{AC}{2} = \frac{0.35}{2} = 0.175$$





# UPGMA

## Unweighted Pair Group Method Using Arithmetic Average

Join C, A columns.

	A	B	C
B	0.40		
C	0.35	0.45	
D	0.60	0.70	0.55

Figure: Distances between sequences (score).

	A-C	B
B	$\frac{0.4 + 0.45}{2} = 0.425$	
D	$\frac{0.55 + 0.6}{2} = 0.575$	0.70

Figure: New matrix

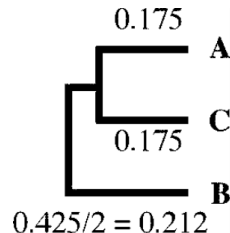
# UPGMA

## Unweighted Pair Group Method Using Arithmetic Average

	A-C	B
B	$\frac{0.4 + 0.45}{2} = 0.425$	
D	$\frac{0.55 + 0.6}{2} = 0.575$	0.70

Figure: Distances between sequences (score).

Choose the smallest distance ( $B, A-C$ ). Then compute the cluster.



# UPGMA

## Unweighted Pair Group Method Using Arithmetic Average

Join *AC*, *B* columns

	A-C	B
B	$\frac{0.4 + 0.45}{2} = 0.425$	
D	$\frac{0.55 + 0.6}{2} = 0.575$	0.70

Figure: Distances between sequences (score).

	B-A-C
D	$\frac{0.7 + 0.6 + 0.55}{3} = 0.617$

Figure: New matrix

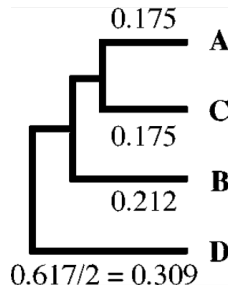
# UPGMA

## Unweighted Pair Group Method Using Arithmetic Average

	<b>B-A-C</b>
<b>D</b>	$\frac{0.7 + 0.6 + 0.55}{3} = 0.617$

Figure: Distances between sequences (score).

Choose the smallest distance  
(B-A-C). Then compute the  
cluster.



# UPGMA

## Unweighted Pair Group Method Using Arithmetic Average

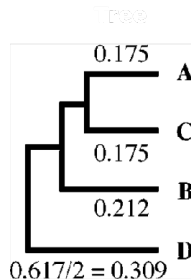
The estimated distances don't match the actual evolutionary distances shown, which illustrates the failure of UPGMA. However, owing to its fast speed of calculation, it has found extensive usage.

	A	B	C
B	0.42		
C	0.35	0.42	
D	0.62	0.62	0.62

$$0.212 + 0.212$$

$$0.309 + 0.309$$

$$0.175 + 0.175$$



# Questions?

