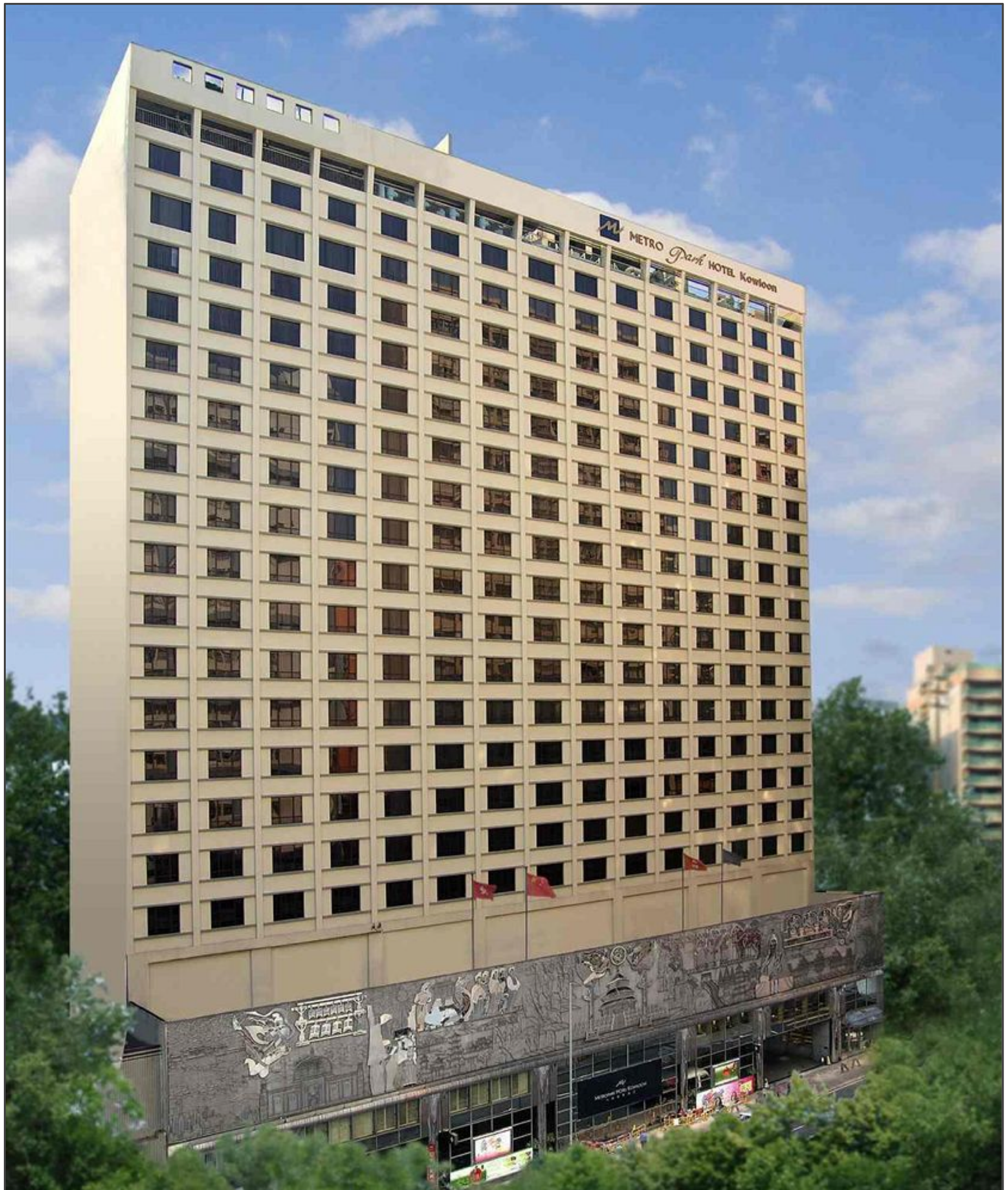
Which Animal Gave Us SARS?

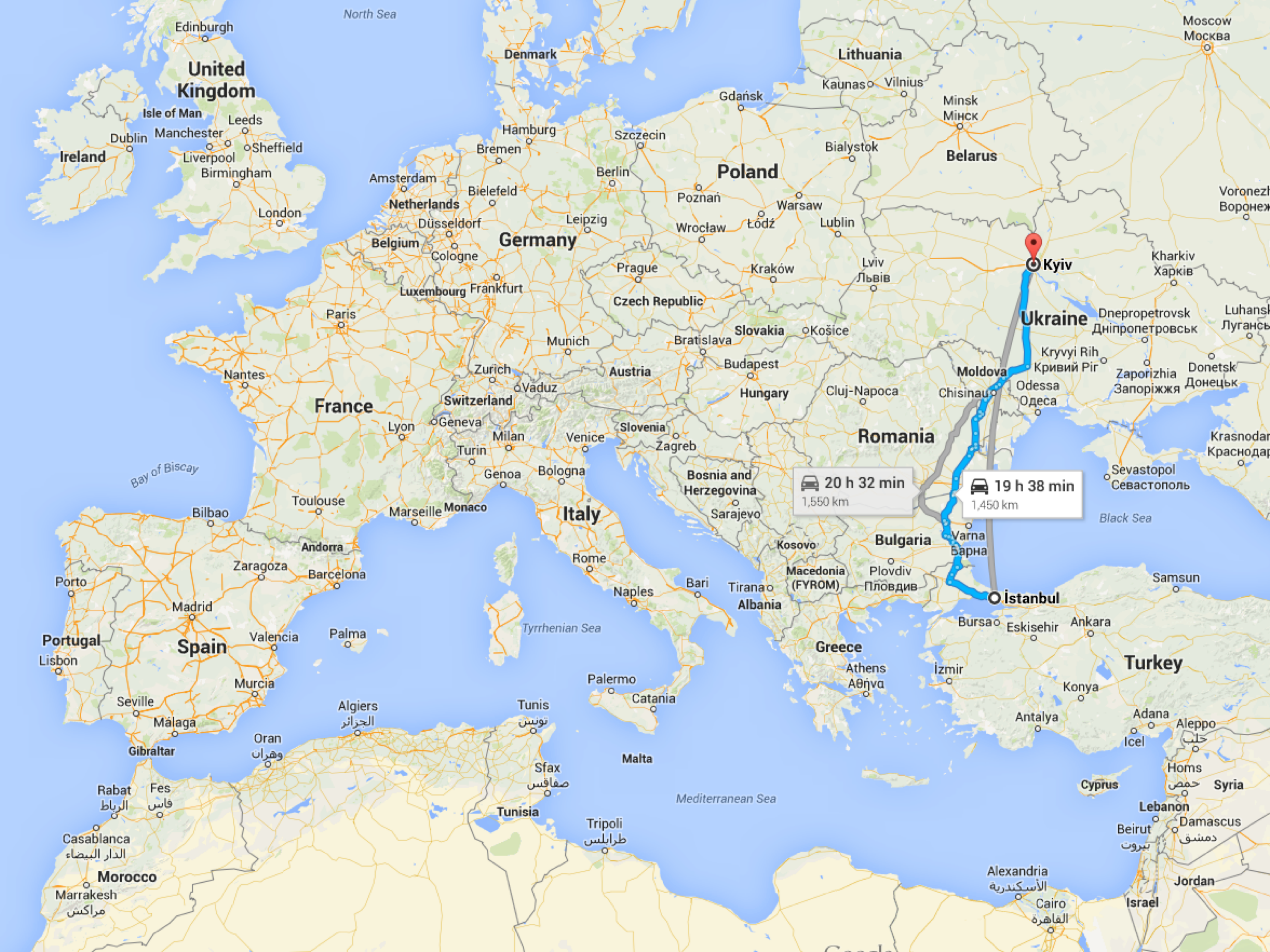
Evolutionary Trees Part 1: The Neighbor-Joining Algorithm

Phillip Compeau and Pavel Pevzner

Bioinformatics Algorithms: An Active Learning Approach

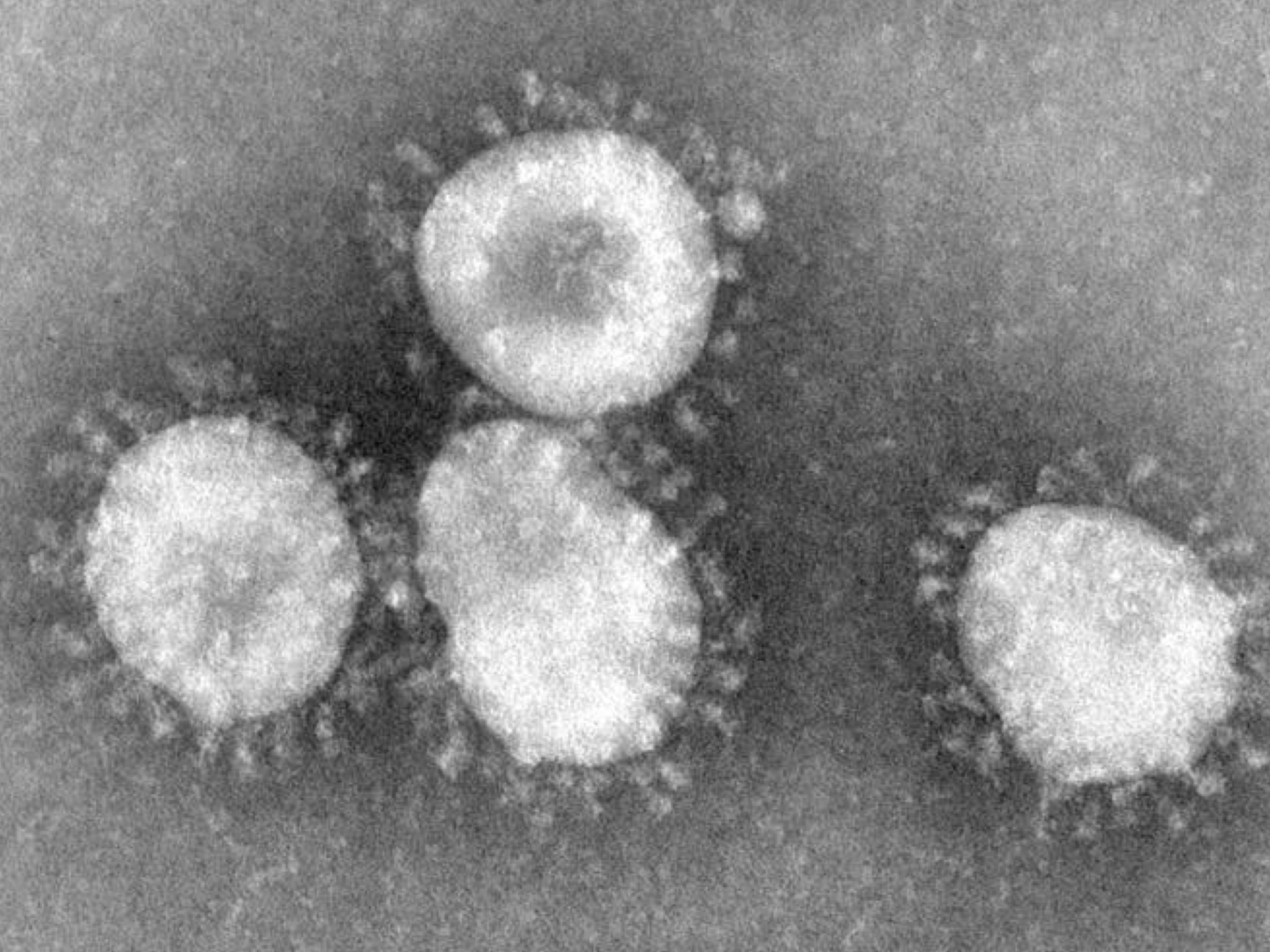

METRO Park HOTEL Kowloon

Modern boundaries are shown for reference.
Cities with repeated plague outbreaks, 14th to 18th century
Spreading of the plague by marine routes
Approximate date of the first outbreak
1347
1348
1349
1350
1351
relatively unaffected
London
Amsterdam
Rotterdam
Antwerp
Aachen
Lübeck
Danzig (Gdansk)
Kiev
Paris
Strassburg (Strasbourg)
Nürnberg
Vienna
approaching from Asia, 1346
Bordeaux
Lyon
Florence
Venice
Trieste
Genoa
Marseille
Lisbon
Barcelona
Ancona
Ragusa (Dubrovnik)
Rome
Valencia
Seville
Naples
Constantinople (Istanbul)
approaching from Asia, 1346


North Sea
Edinburgh
United Kingdom
Isle of Man
Leeds
Manchester
Sheffield
Liverpool
Birmingham
London
Dublin
Ireland
Denmark
Hamburg
Bremen
Szczecin
Gdańsk
Lithuania
Kaunas
Vilnius
Minsk
Мінск
Belarus
Moscow
Москва
Bialystok
Poland
Berlin
Amsterdam
Bielefeld
Netherlands
Düsseldorf
Belgium
Cologne
Germany
Leipzig
Poznań
Warsaw
Łódź
Wrocław
Lublin
Voronezh
Воронеж
Kyiv
Kharkiv
Харків
Lviv
Львів
Prague
Kraków
Czech Republic
Luxembourg
Frankfurt
Paris
Ukraine
Dnepropetrovsk
Дніпропетровськ
Luhansk
Луганськ
Slovakia
Košice
Bratislava
Munich
Budapest
Austria
Zurich
Vaduz
Switzerland
Nantes
France
Hungary
Cluj-Napoca
Moldova
Chisinau
Kryvyi Rih
Кривий Ріг
Zaporizhia
Запоріжжя
Donetsk
Донецьк
Odessa
Одеса
Lyon
Geneva
Milan
Turin
Venice
Slovenia
Zagreb
Romania
Krasnodar
Краснодар
Sevastopol
Севастополь
Bay of Biscay
Genoa
Bologna
Bosnia and Herzegovina
Sarajevo
20 h 32 min
1,550 km
19 h 38 min
1,450 km
Black Sea
Toulouse
Marseille
Monaco
Italy
Bilbao
Andorra
Kosovo
Varna
Варна
Bulgaria
Zaragoza
Barcelona
Rome
Macedonia (FYROM)
Plovdiv
Пловдив
Samsun
Porto
Madrid
Naples
Bari
Tirana
Albania
İstanbul
Bursa
Eskisehir
Ankara
Portugal
Spain
Valencia
Palma
Tyrrhenian Sea
Greece
Lisbon
Athens
Αθήνα
İzmir
Turkey
Konya
Murcia
Palermo
Catania
Seville
Málaga
Algiers
الجزائر
Tunis
تونس
Antalya
Adana
Icel
Aleppo
حلب
Gibraltar
Oran
وهران
Malta
Homs
حمص
Syria
Rabat
الرباط
Fes
فاس
Sfax
صفاقس
Cyprus
Tunisia
Mediterranean Sea
Lebanon
Beirut
بيروت
Damascus
دمشق
Casablanca
الدار البيضاء
Tripoli
طرابلس
Morocco
Marrakesh
مراكش
Alexandria
الأسكندرية
Jordan
Cairo
القاهرة
Israel


Ireland
1 case
China
Canada
29 cases
United States
1 case
Guangdong
Province
Metropole
Hotel
Hong Kong
Special
Administrative
Region
195 cases
Vietnam
58 cases
Singapore
71 cases
Philippine
Sea
South
China
Sea
Bay
of
Bengal
Arabian
Sea
Indian
The Spread
of ?????

The Spread of SARS



Questions about SARS

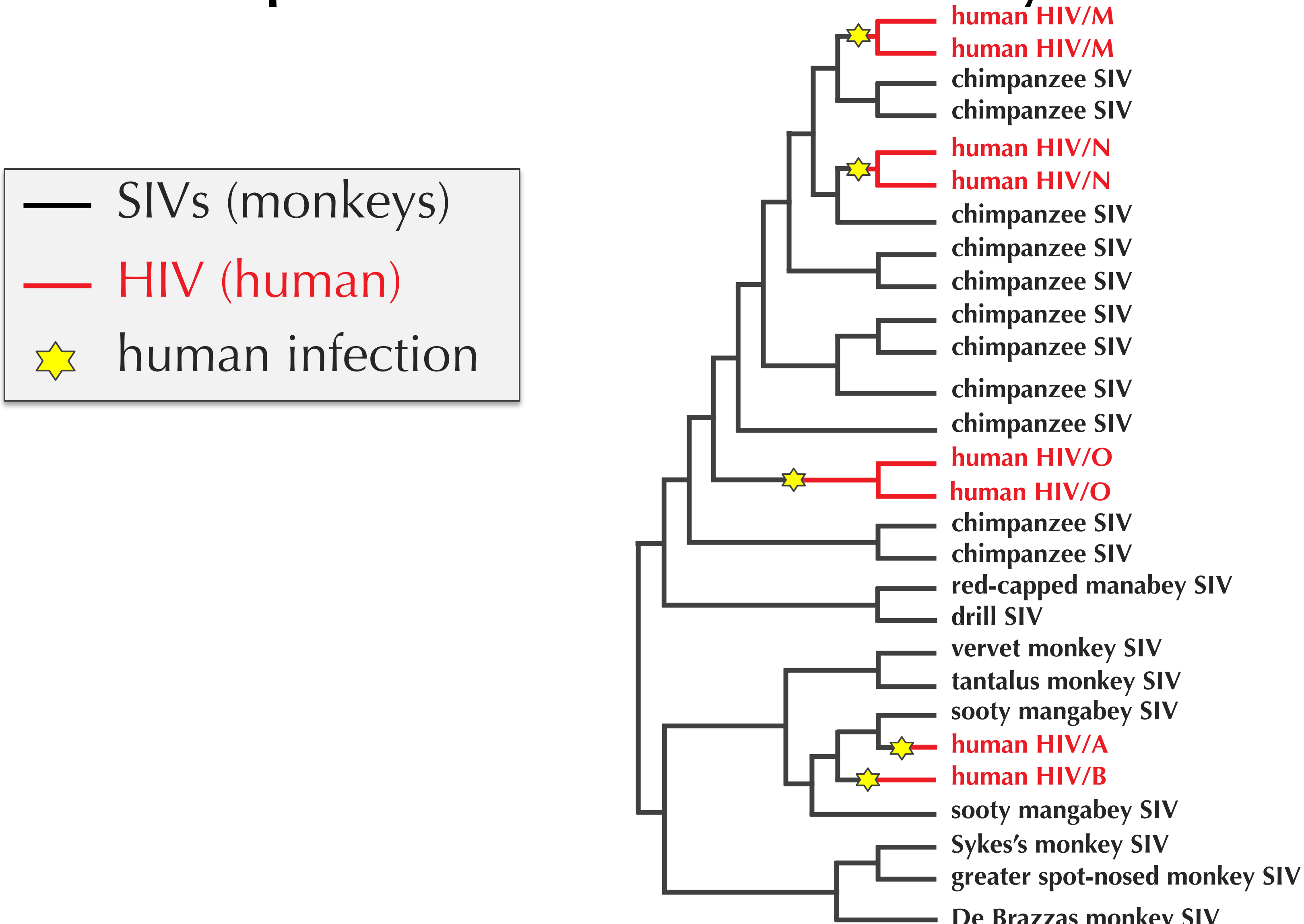
Which animal gave us SARS? How does SARS compare to other viruses and how did it mutate over time?

Questions about SARS

Which animal gave us SARS? How does SARS compare to other viruses and how did it mutate over time?

To answer these questions, we need to learn how to construct **evolutionary trees** (a.k.a. **phylogenies**).

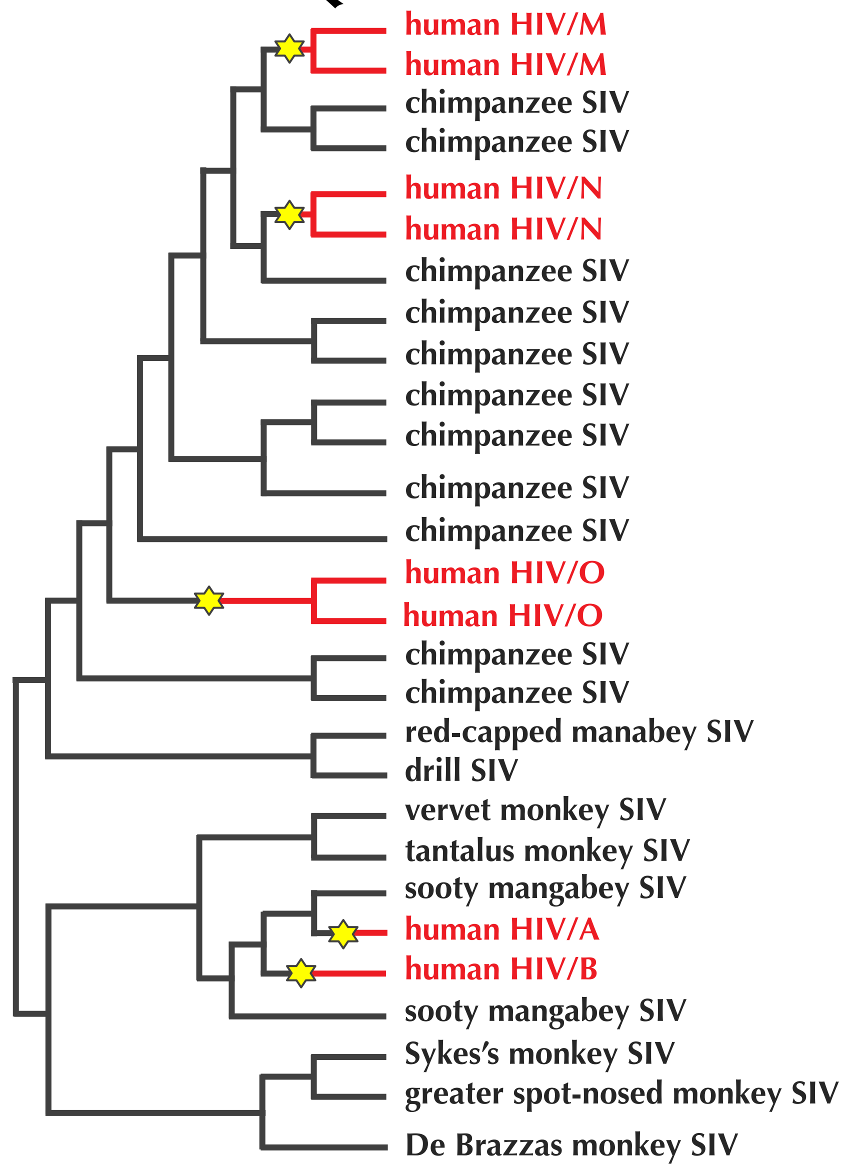
Example: HIV Evolutionary Tree



Two Computational Questions

How do we construct the tree's *structure*?

Can we infer anything about the ancestors?

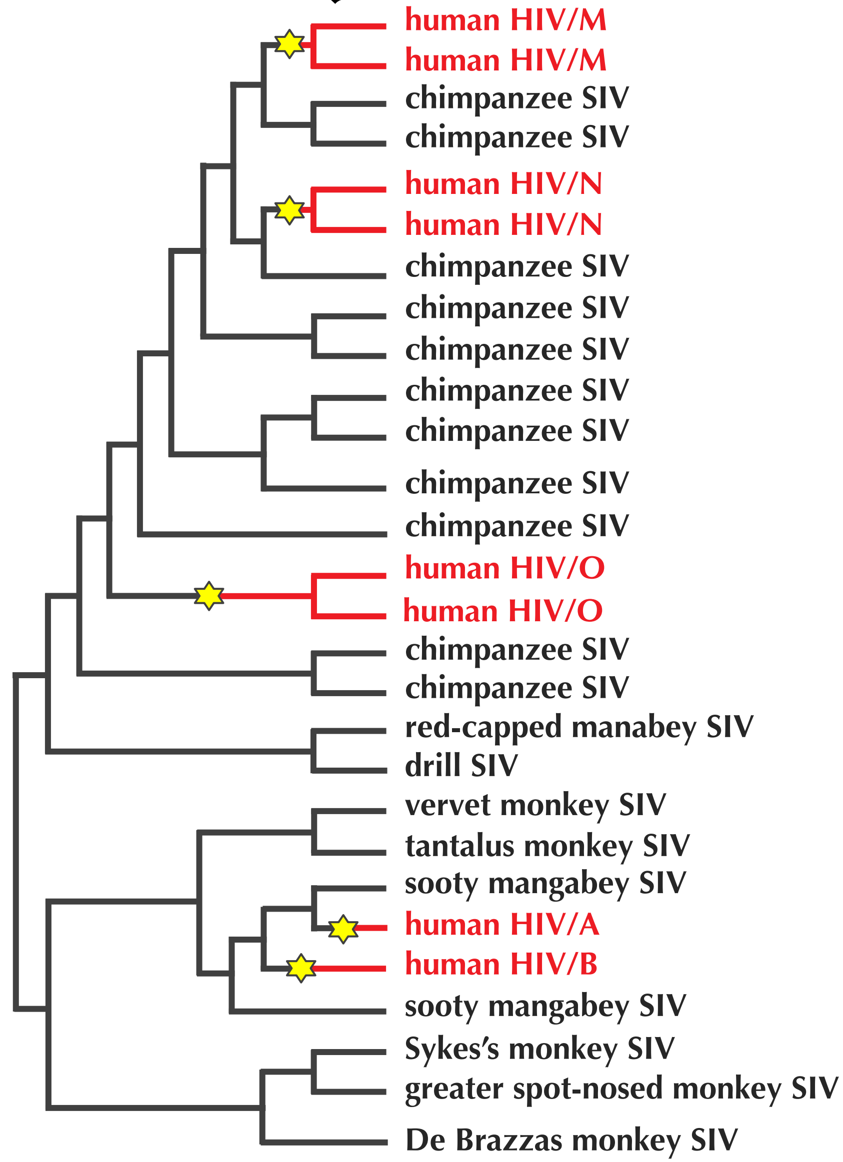


Two Computational Questions

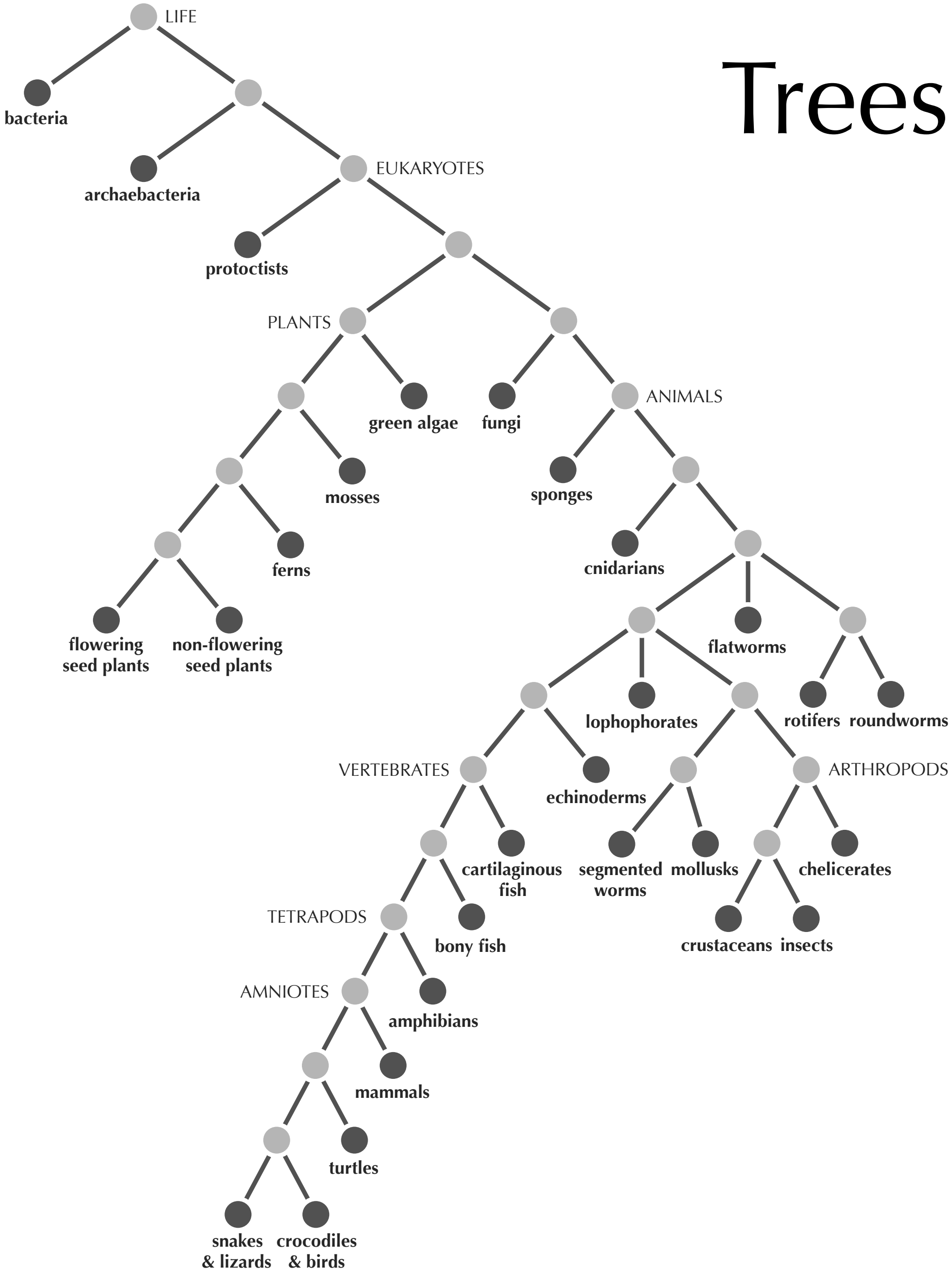
How do we construct the tree's *structure?*

Can we infer anything about the ancestors?

Checkpoint: Any thoughts on how we could answer either question?

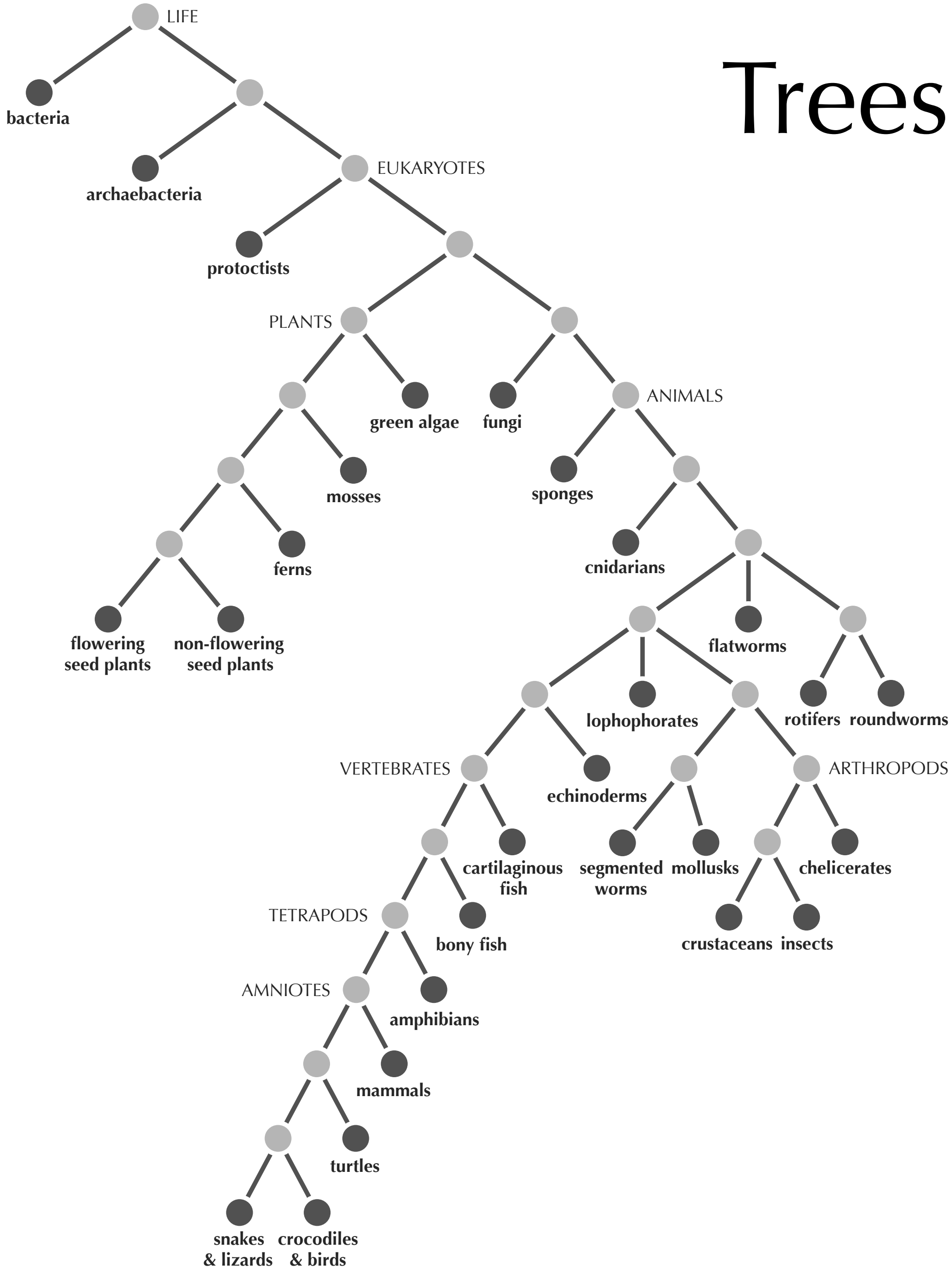


Trees



Tree: Connected graph containing no cycles.

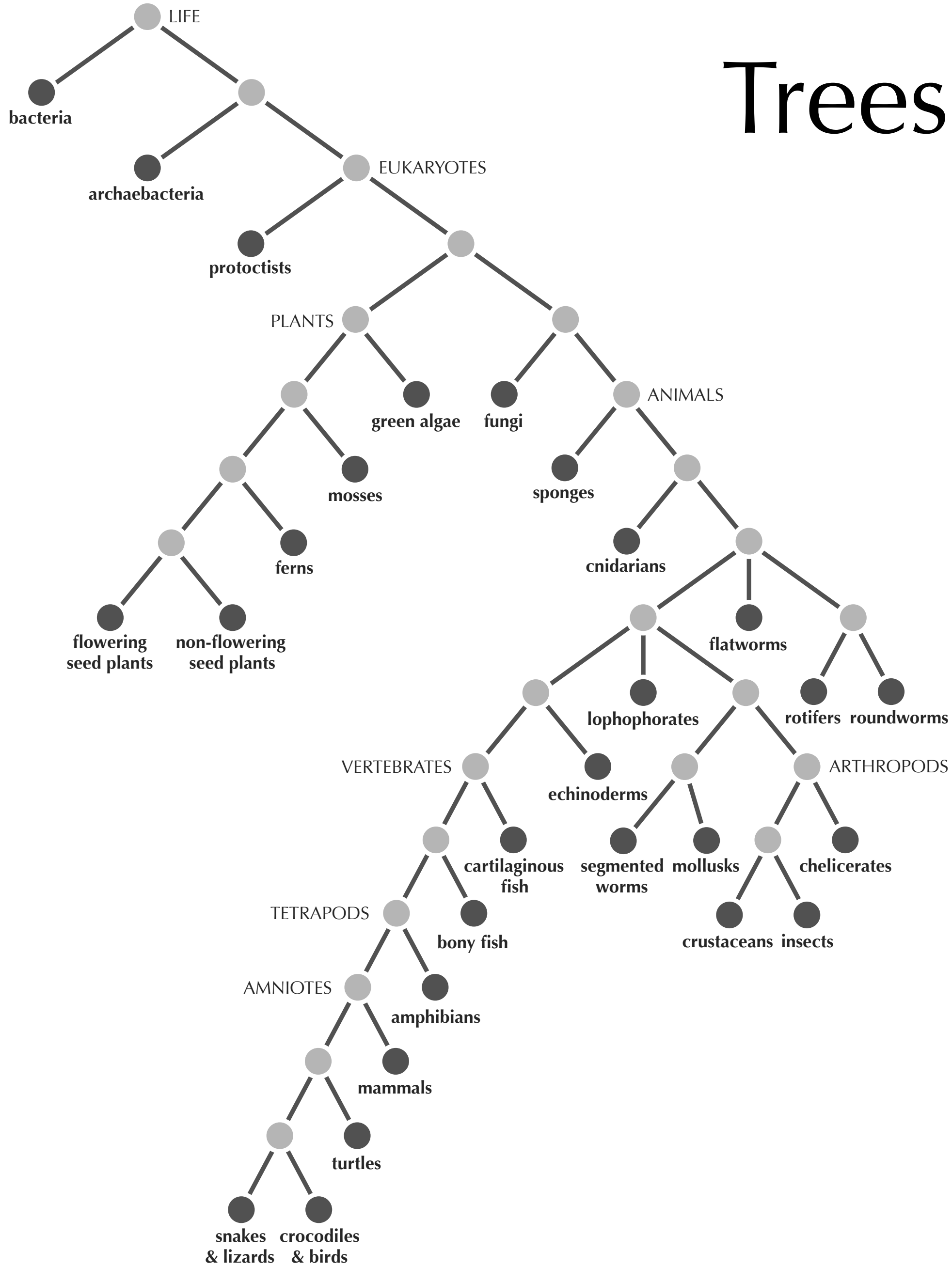
Trees



Tree: Connected graph containing no cycles.

Leaves (degree = 1): present-day species

Trees

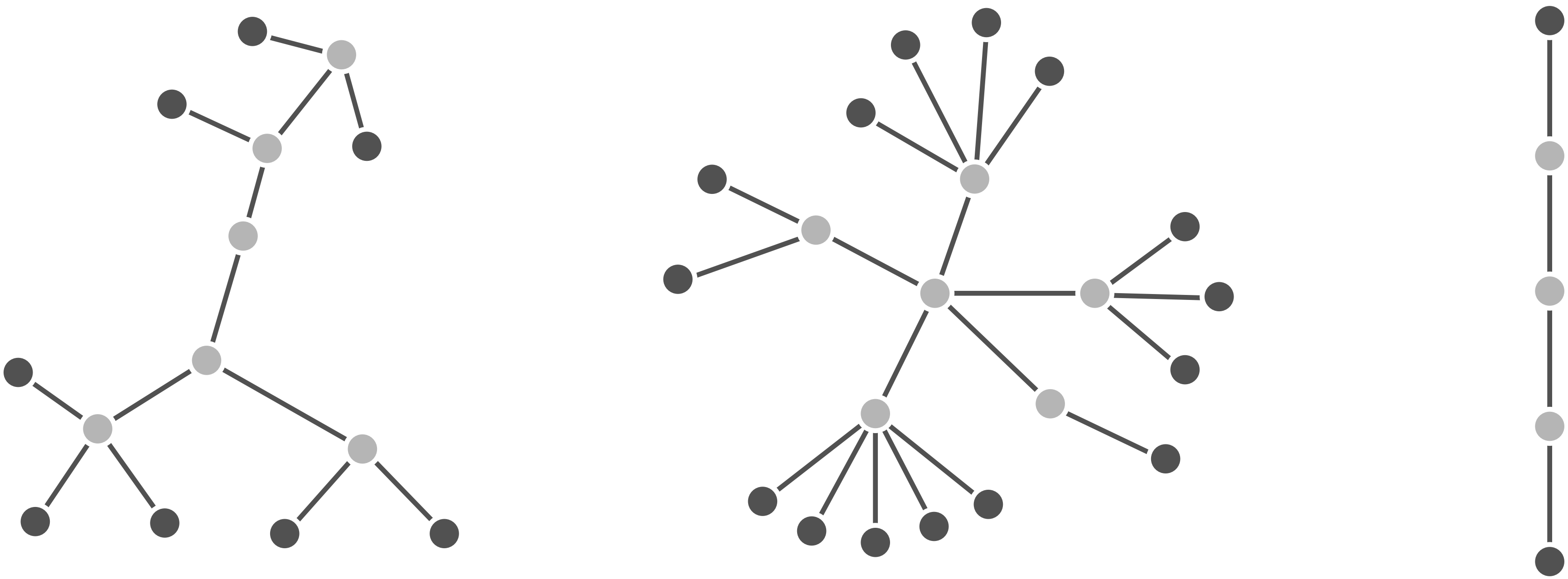


Tree: Connected graph containing no cycles.

Leaves (degree = 1): present-day species

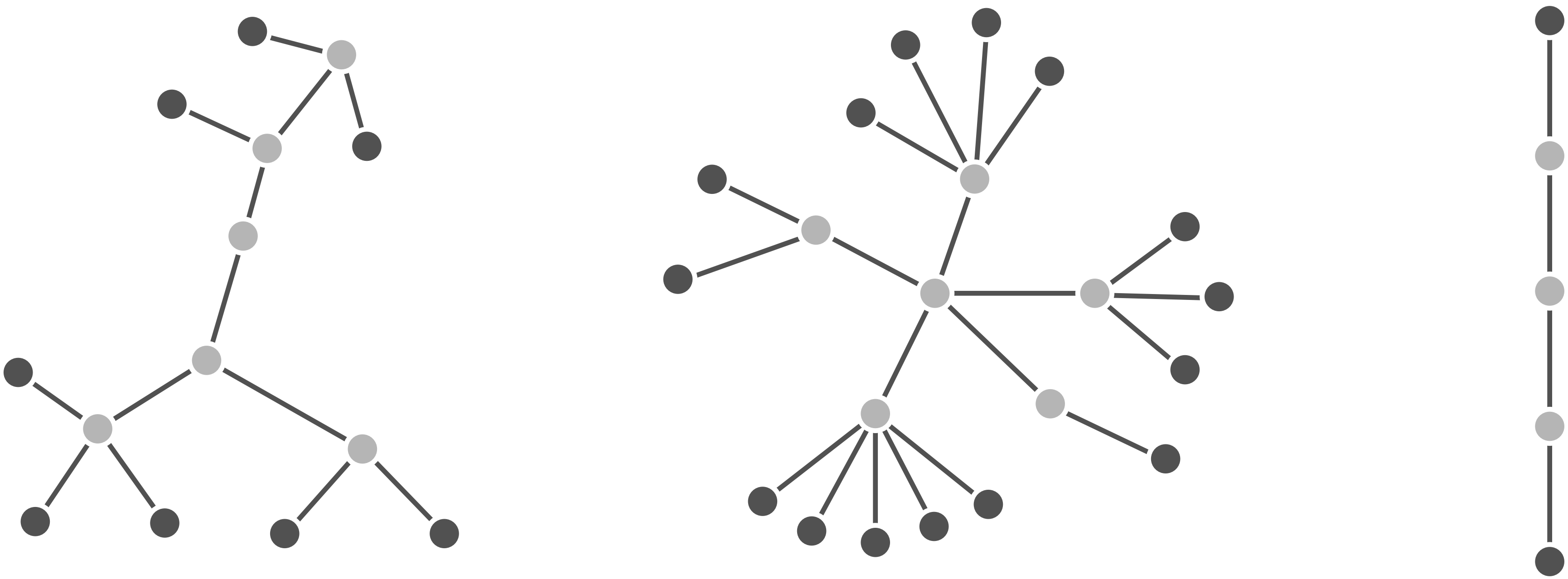
Internal nodes (degree ≥ 2): ancestral species

Trees



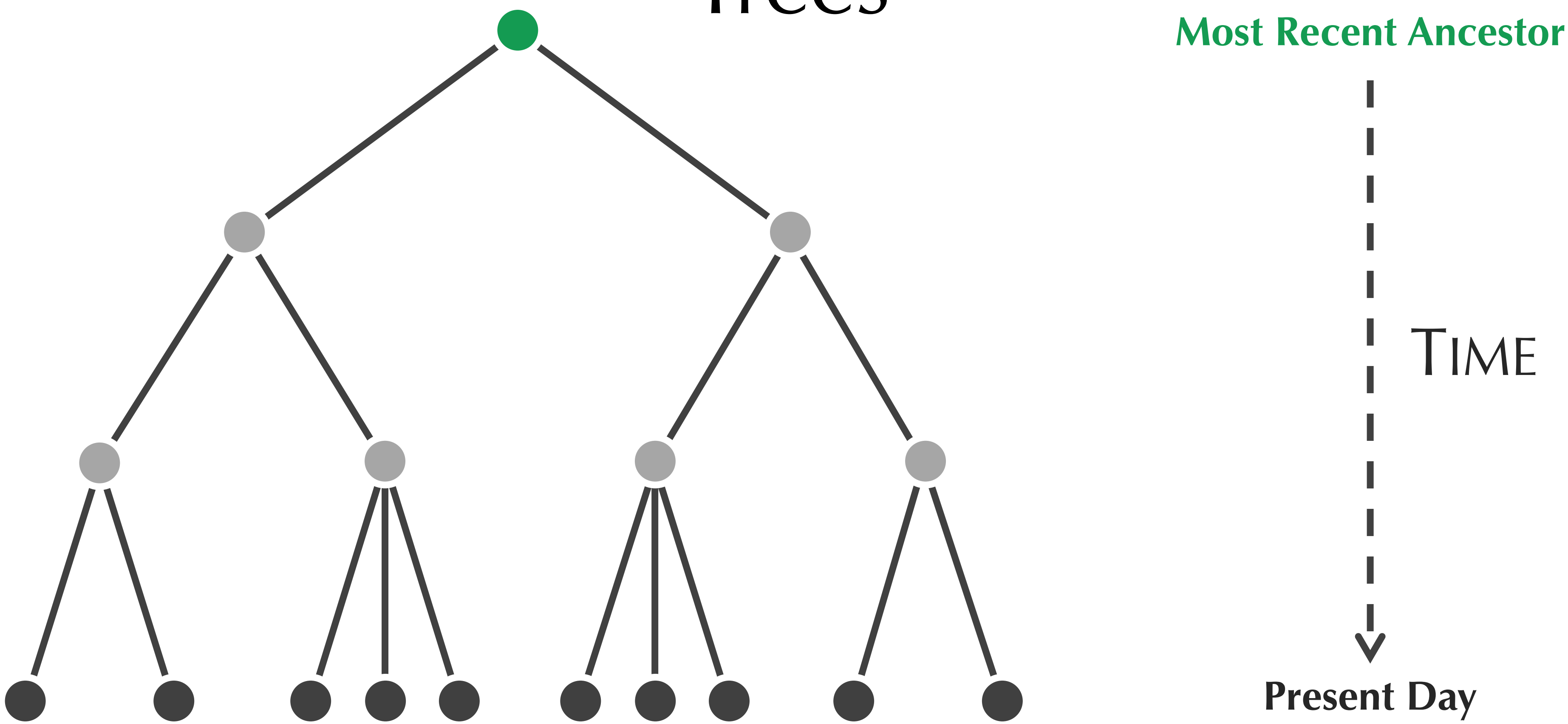
Note: We proved in a previous lecture that every tree with n nodes has exactly $n - 1$ edges.

Trees



Exercise: Prove that there is a unique path connecting any two nodes in a tree.

Trees

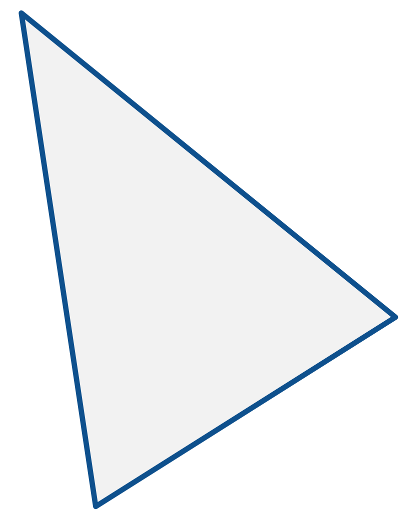


Rooted tree: one node is designated as the **root** (most recent common ancestor).

Definition of a Distance Matrix

Distance matrix: A matrix D representing distances between pairs of n organisms that satisfies three properties:

1. **Symmetry:** $D_{i,j} = D_{i,j}$ for all pairs i, j
2. **Non-negativity:** $D_{i,j} >= 0$ for all pairs i, j
3. **Triangle inequality:** For all i, j, and k, $D_{i,j} + D_{j,k} >= D_{i,k}$.



A Multiple Alignment Defines a Simple Distance Matrix

Species	Alignment
Chimp	ACGTAGGCCT
Human	ATGTAAGACT
Seal	TCGAGAGCAC
Whale	TCGAAAGCAT

A Multiple Alignment Defines a Simple Distance Matrix

$D_{i,j}$ = number of differing symbols between *i*-th and *j*-th rows of a multiple alignment.

SPECIES	ALIGNMENT	DISTANCE MATRIX Chimp	Human	Seal	Whale
Chimp	`ACGTAGGCCT`	0	3	6	4
Human	`ATGTAAGACT`	3	0	7	5
Seal	`TCGAGAGCAC`	6	7	0	2
Whale	`TCGAAAGCAT`	4	5	2	0

A Multiple Alignment Defines a Simple Distance Matrix

$D_{i,j}$ = number of differing symbols between *i*-th and *j*-th rows of a multiple alignment.

Species	Alignment	Chimp	Human	Seal	Whale
Chimp	ACGTAGGCCT	0	**3**	6	4
Human	ATGTAAGACT	**3**	0	7	5
Seal	TCGAGAGCAC	6	7	0	2
Whale	TCGAAAGCAT	4	5	2	0

A Multiple Alignment Defines a Simple Distance Matrix

Exercise: Prove that for any multiple sequence alignment, this way of defining *D* produces a distance matrix.

SPECIES	ALIGNMENT	DISTANCE MATRIX			
		Chimp	**Human**	**Seal**	**Whale**
Chimp	ACGTAGGCCT	0	3	6	4
Human	ATGTAAGACT	3	0	7	5
Seal	TCGAGAGCAC	6	7	0	2
Whale	TCGAAAGCAT	4	5	2	0

Distance-Based Phylogeny

Distance-Based Phylogeny Problem.

- **Input:** A distance matrix.
- **Output:** The unrooted tree "fitting" this distance matrix.

Distance-Based Phylogeny

Distance-Based Phylogeny Problem.

- **Input:** A distance matrix.
- **Output:** The unrooted tree "*fitting*" this distance matrix.

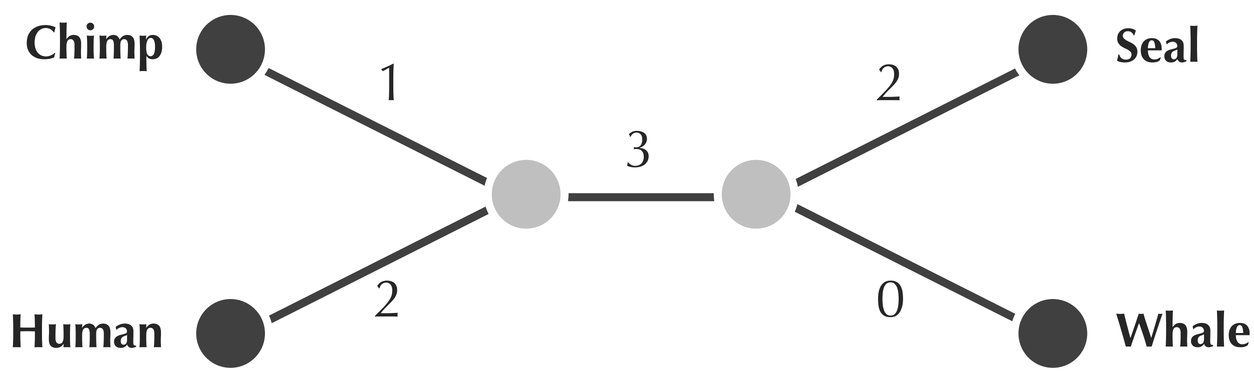
Of course, we are getting a bit ahead of ourselves – we should define what we mean by "fitting"!

“Fitting” a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0

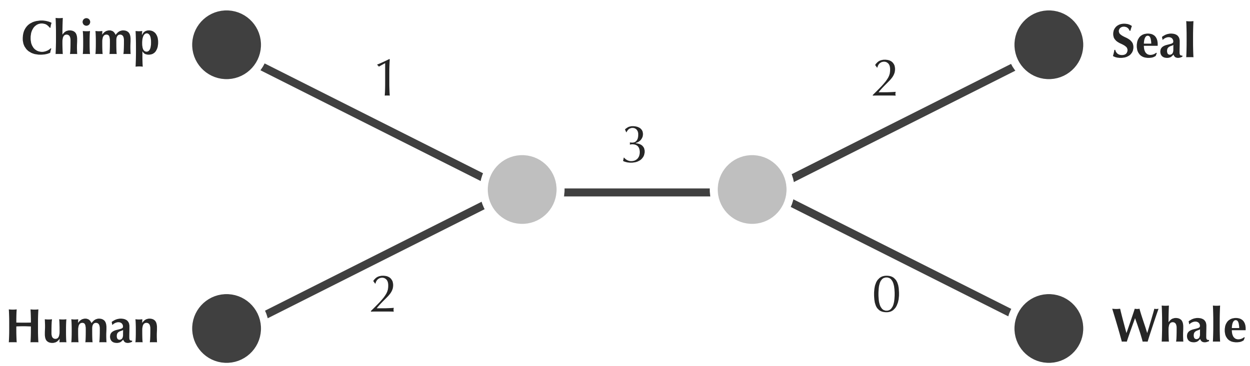
"Fitting" a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



“Fitting” a Tree to a Matrix

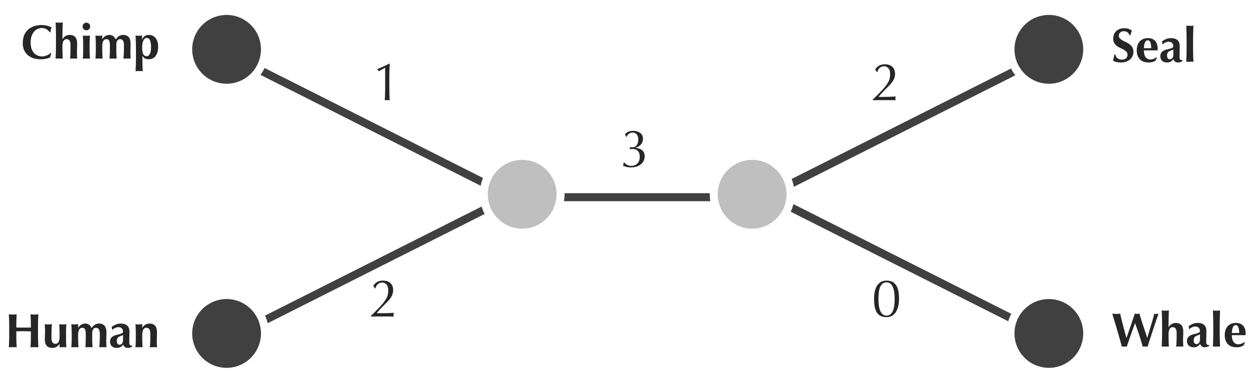
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$d_{i,j}(T)$ = distance between nodes i and j in tree T, computed by summing edge weights from i to j.

"Fitting" a Tree to a Matrix

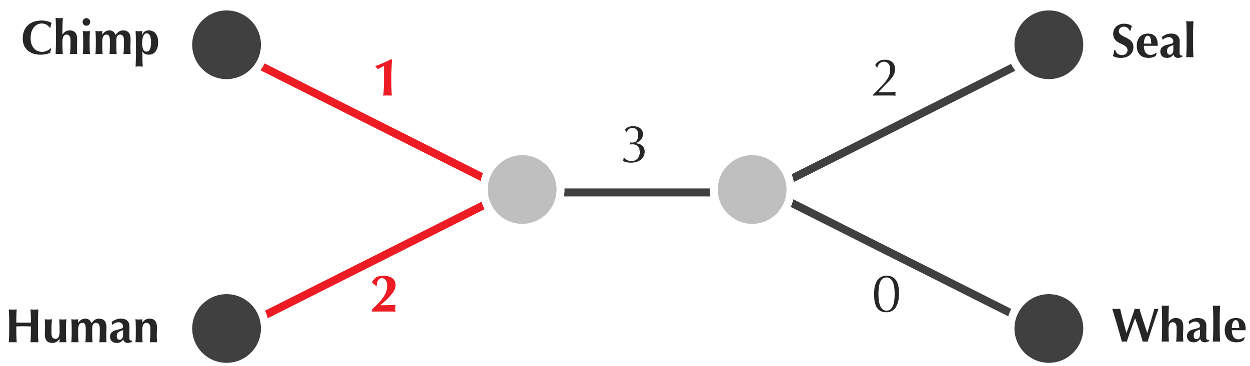
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



We say that T **fits** matrix D if for every pair i and j, $d_{i,j}(T) = D_{i,j}$.

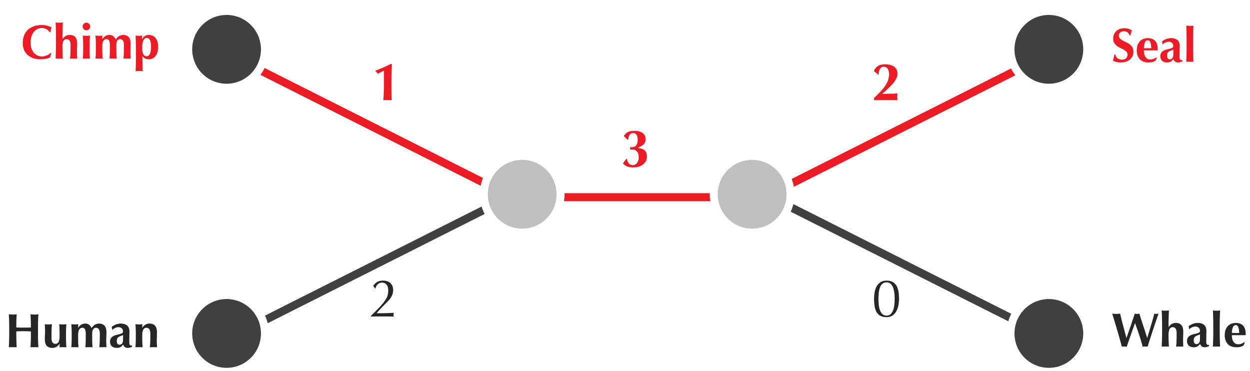
“Fitting” a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



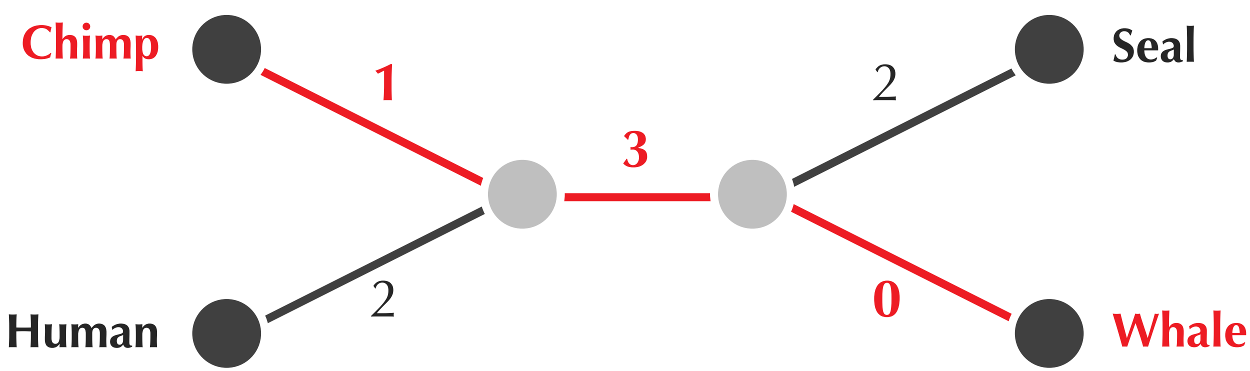
"Fitting" a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



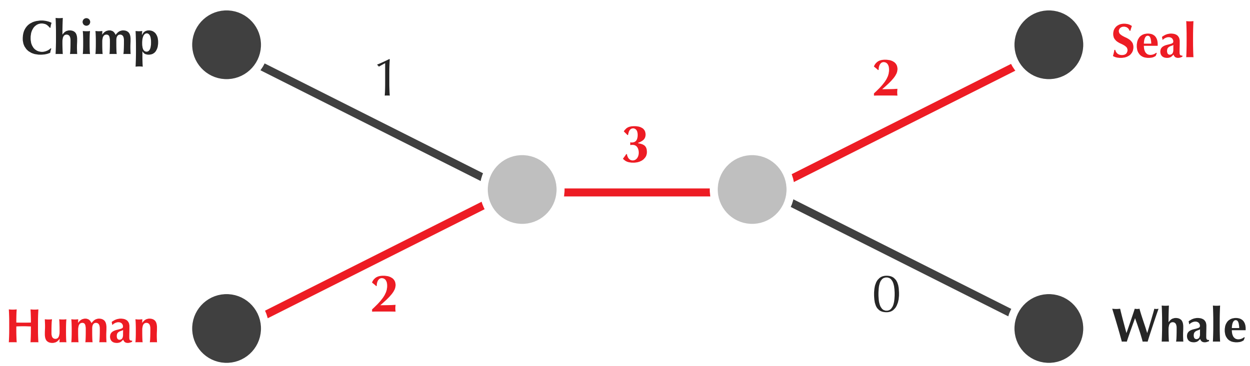
"Fitting" a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	**4**
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



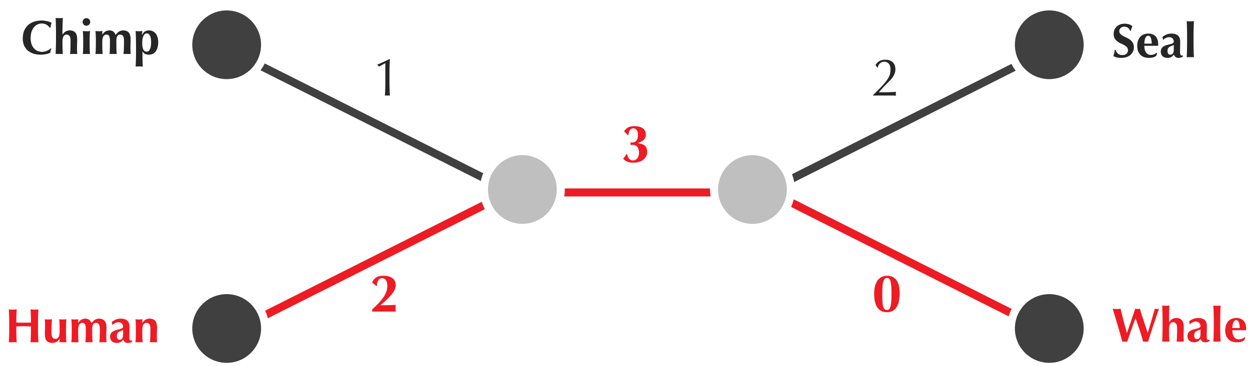
"Fitting" a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



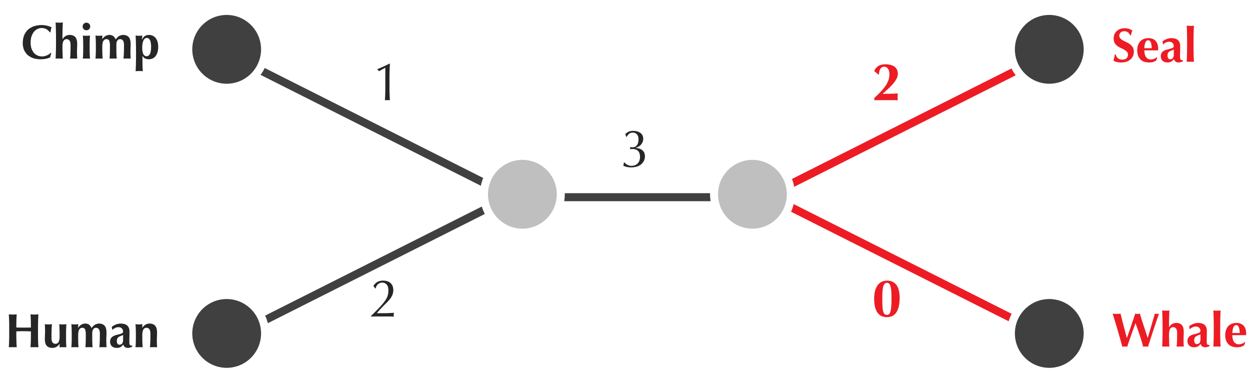
"Fitting" a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	**5**
Seal	6	7	0	2
Whale	4	5	2	0



"Fitting" a Tree to a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	**2**
Whale	4	5	2	0



Return to Distance-Based Phylogeny

Exercise: Find a tree fitting the following matrix.

	v_1	v_2	v_3	v_4
v_1	0	3	4	3
v_2	3	0	4	5
v_3	4	4	0	2
v_4	3	5	2	0

Sometimes, **No** Tree Fits a Matrix

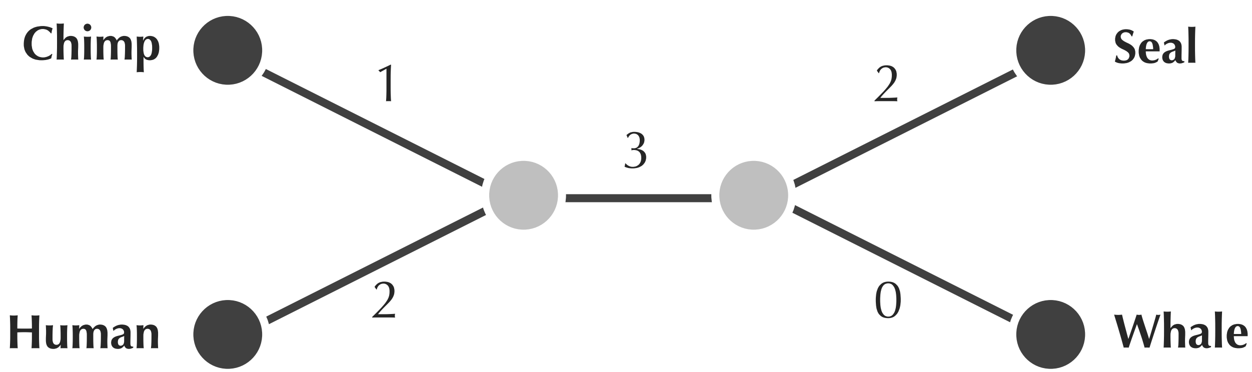
Exercise: Find a tree fitting the following matrix.

	v_1	v_2	v_3	v_4
v_1	0	3	4	3
v_2	3	0	4	5
v_3	4	4	0	2
v_4	3	5	2	0

Additive matrix: distance matrix such that there exists an unrooted tree fitting it.

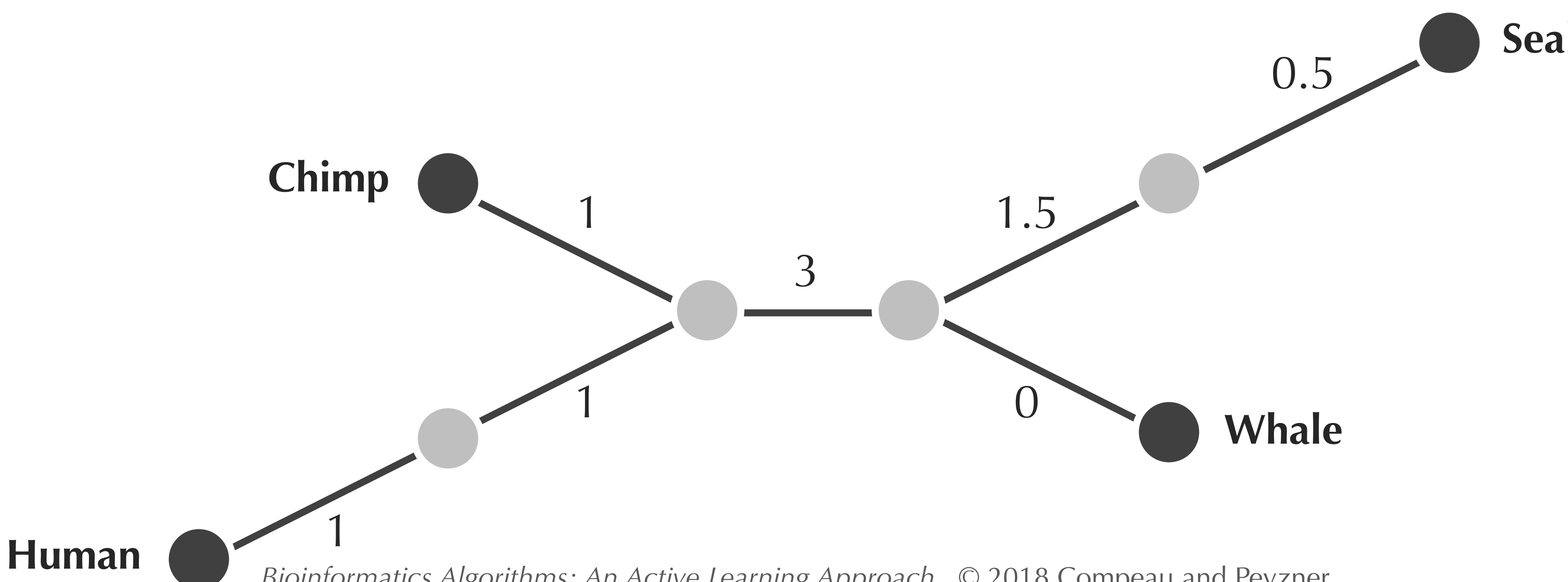
Sometimes, **More Than One** Tree Fits a Matrix

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0

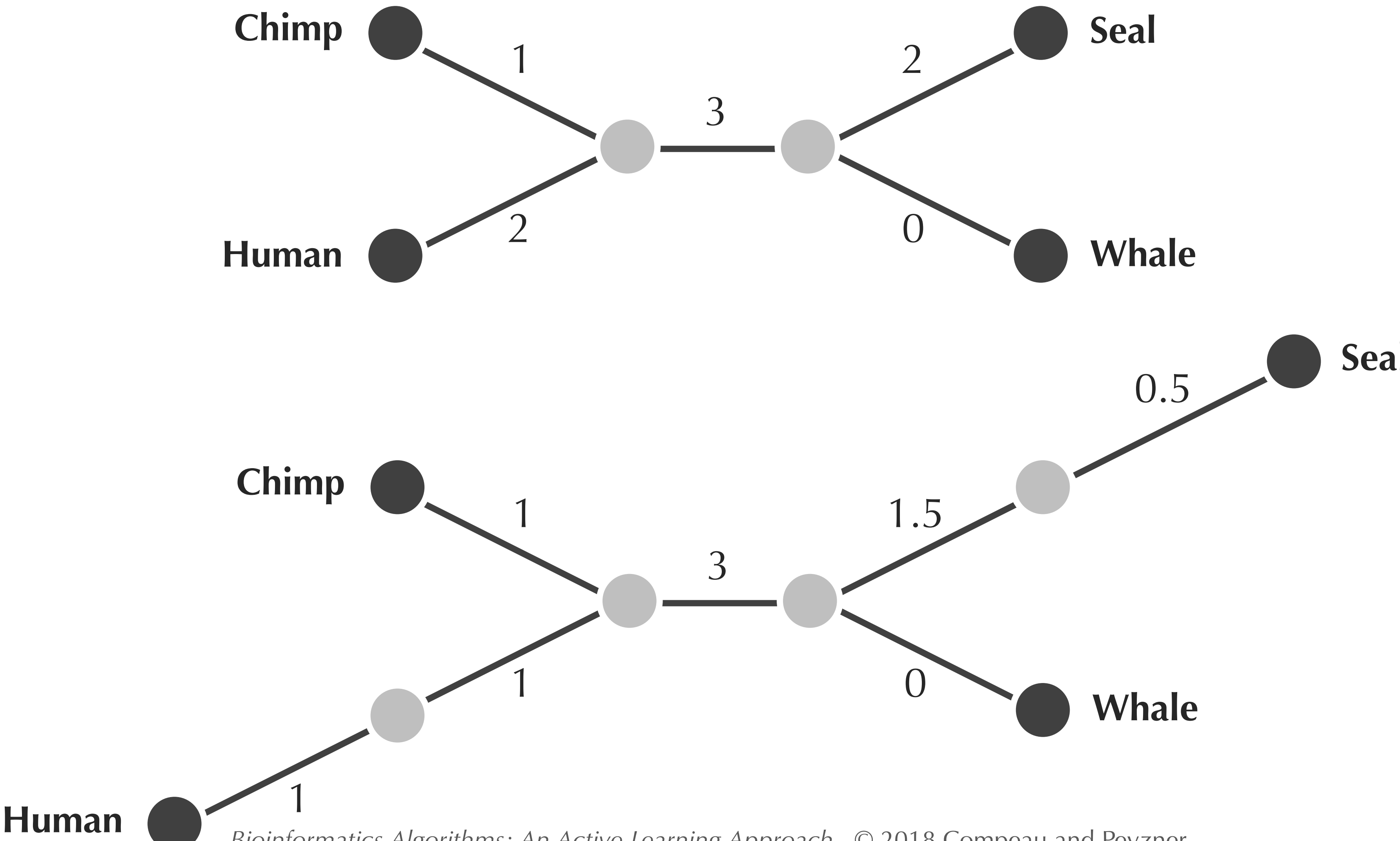


Sometimes, **More Than One** Tree Fits a Matrix

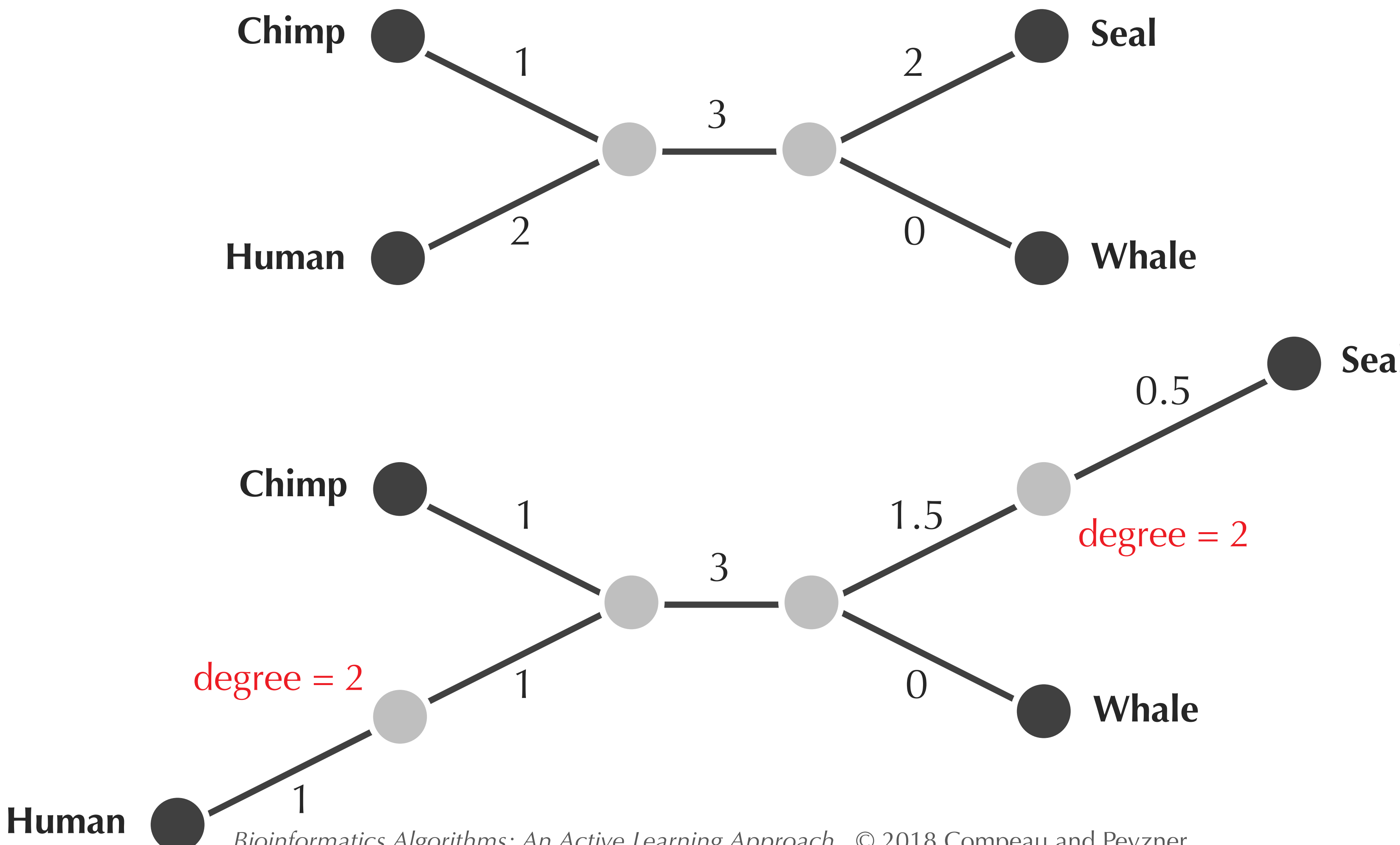
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



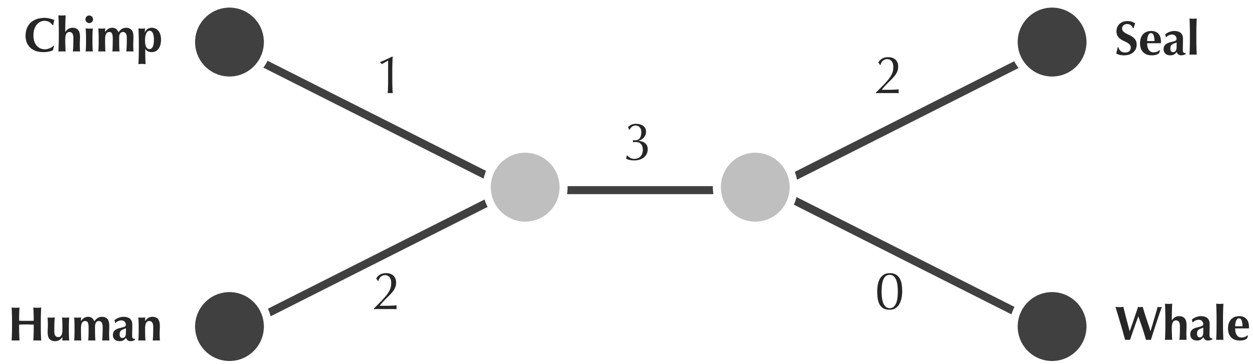
Which Tree is Better?



Which Tree is Better?

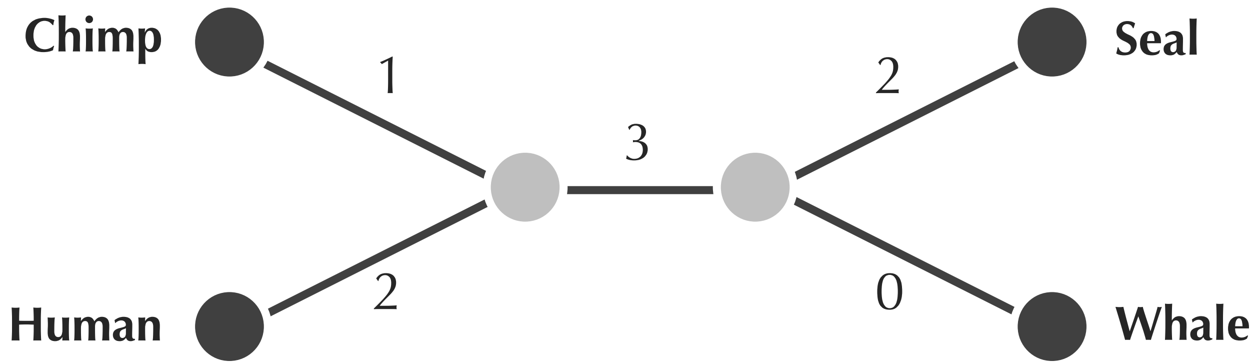


Which Tree is Better?



Simple tree: tree with no nodes of degree 2.

Which Tree is Better?



Simple tree: tree with no nodes of degree 2.

Theorem: There is a unique *simple* tree fitting an *additive* matrix.

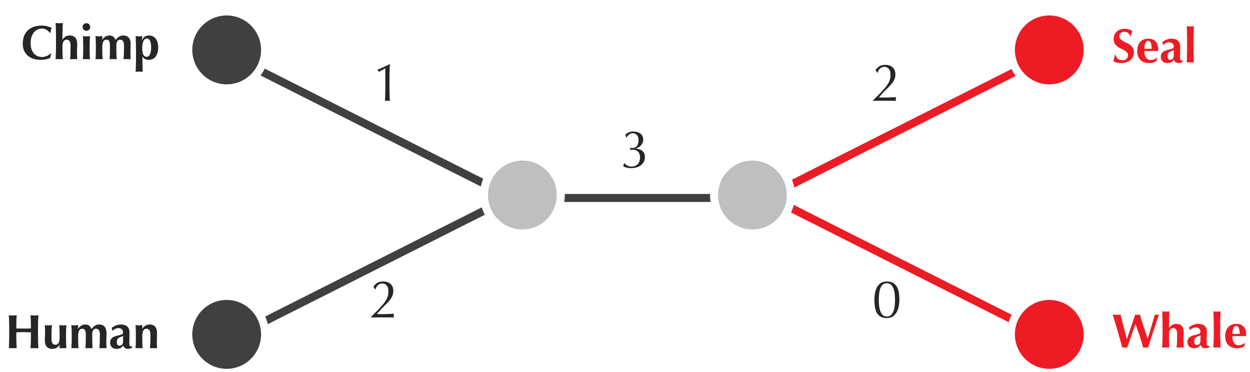
Reformulating Distance-Based Phylogeny

Distance-Based Phylogeny Problem: *Construct an evolutionary tree from a distance matrix.*

- **Input:** A distance matrix.
- **Output:** The simple tree fitting this distance matrix (if this matrix is additive).

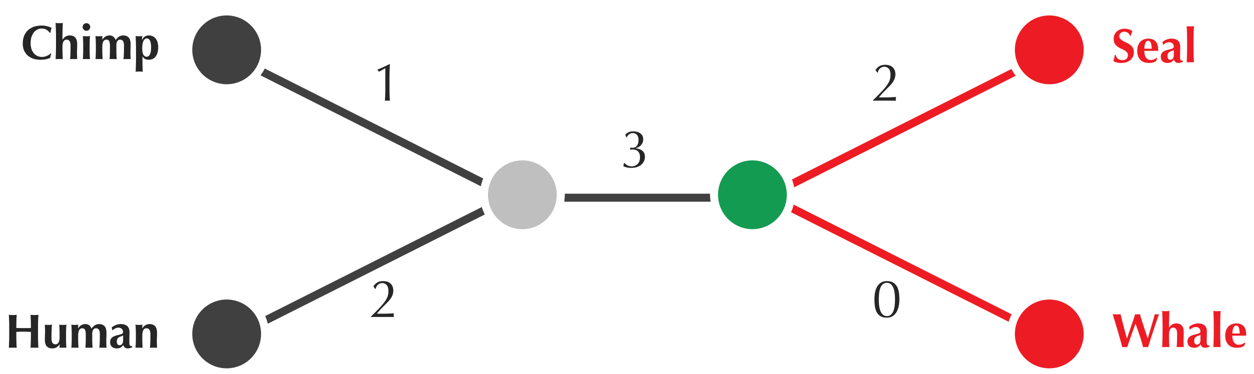
An Idea for Distance-Based Phylogeny

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	**2**
Whale	4	5	2	0



An Idea for Distance-Based Phylogeny

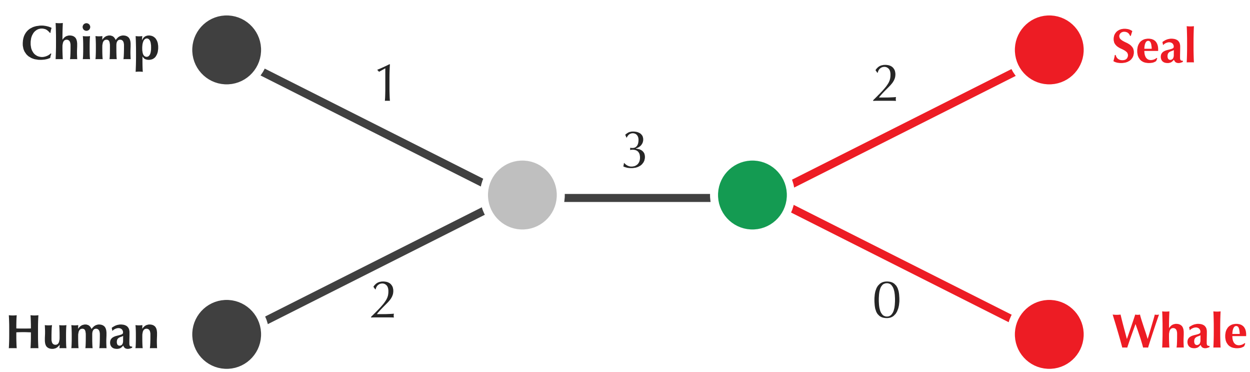
Seal and whale are **neighbors** (meaning they share the same **parent**).



An Idea for Distance-Based Phylogeny

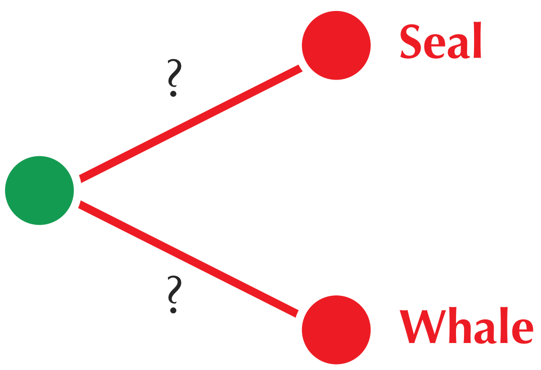
Seal and whale are **neighbors** (meaning they share the same **parent**).

Theorem: Every simple tree with at least three leaves has at least one pair of neighboring leaves.



An Idea for Distance-Based Phylogeny

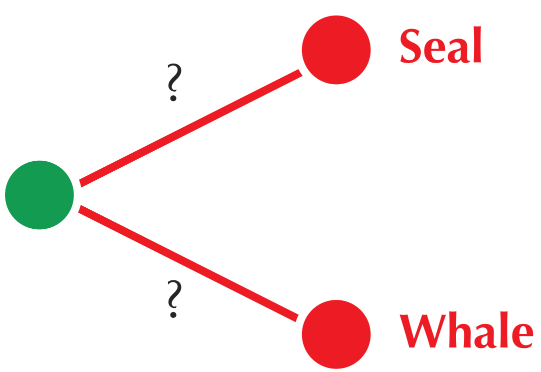
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	**2**
Whale	4	5	2	0



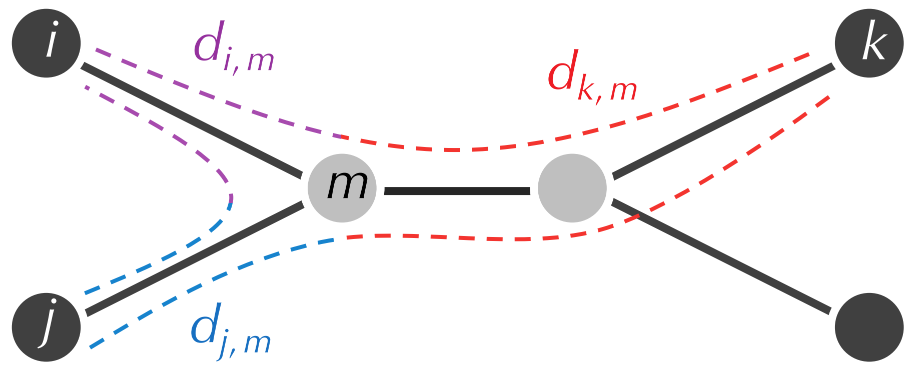
An Idea for Distance-Based Phylogeny

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	**2**
Whale	4	5	2	0

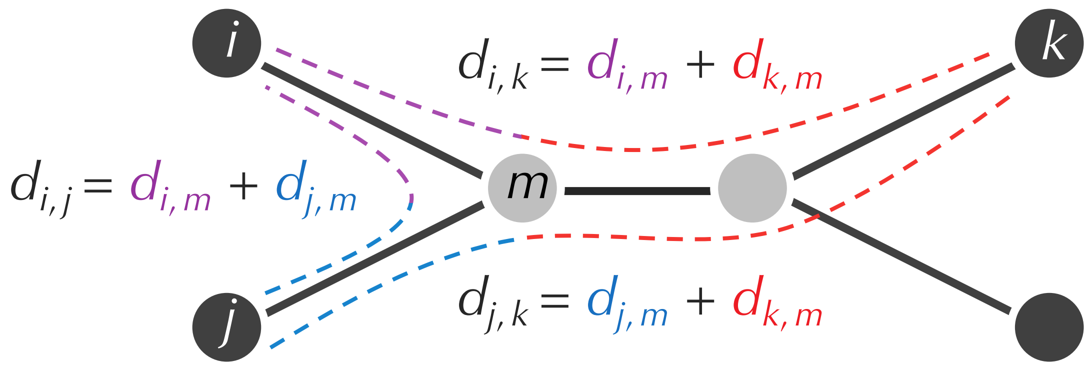
Key Point: How do we compute the unknown distances?



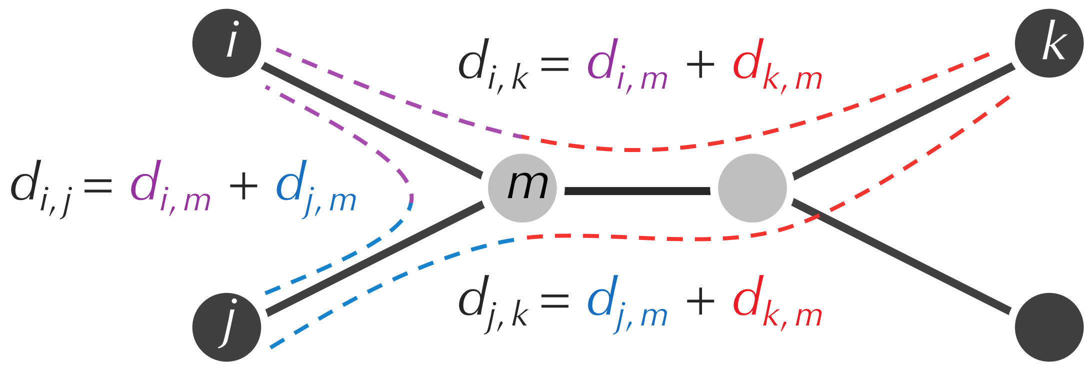
Toward a Recursive Algorithm



Toward a Recursive Algorithm

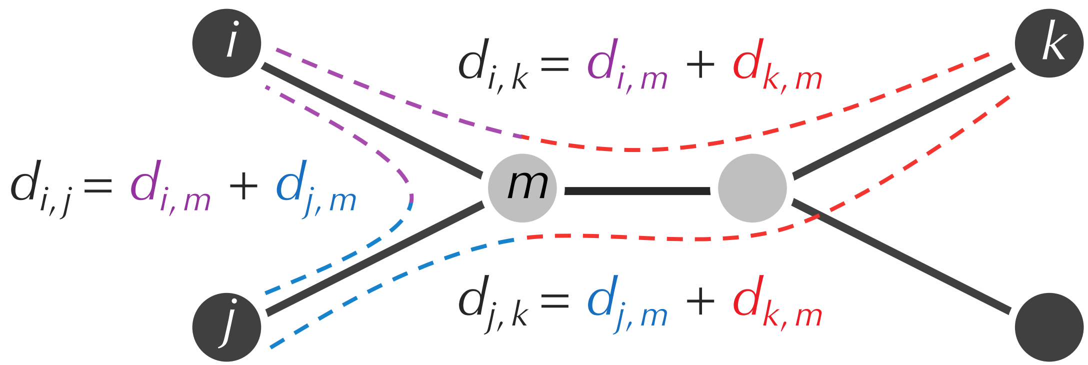


Toward a Recursive Algorithm



$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})]\ /\ 2$$

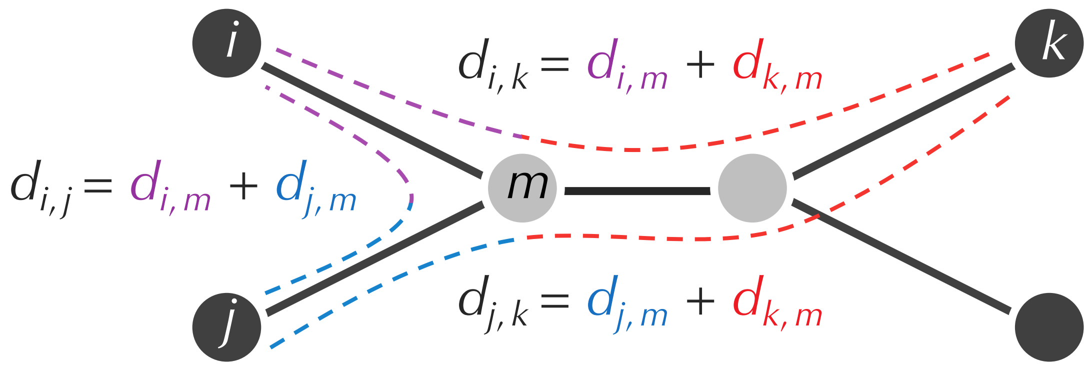
Toward a Recursive Algorithm



$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})] / 2$$

$$d_{k,m} = (d_{i,k} + d_{j,k} - d_{i,j}) / 2$$

Toward a Recursive Algorithm

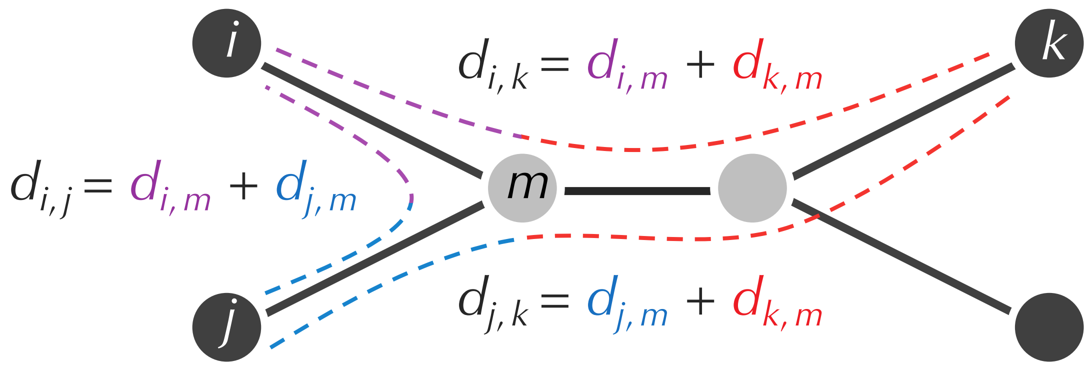


$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})] / 2$$

$$d_{k,m} = (d_{i,k} + d_{j,k} - d_{i,j}) / 2$$

$$d_{k,m} = (D_{i,k} + D_{j,k} - D_{i,j}) / 2$$

Toward a Recursive Algorithm



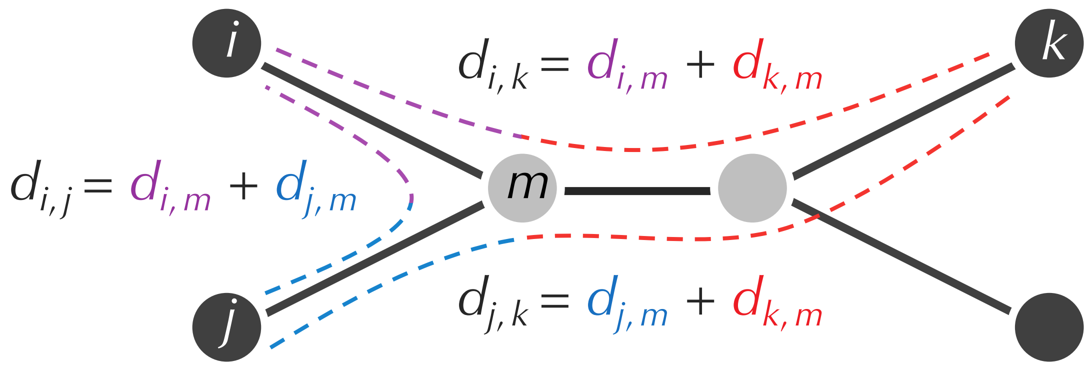
$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})] \,/\, 2$$

$$d_{k,m} = (d_{i,k} + d_{j,k} - d_{i,j}) \,/\, 2$$

$$d_{k,m} = (D_{i,k} + D_{j,k} - D_{i,j}) \,/\, 2$$

$$\therefore \quad d_{i,m} = D_{i,k} - (D_{i,k} + D_{j,k} - D_{i,j}) \,/\, 2$$

Toward a Recursive Algorithm



$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})] \,/\, 2$$

$$d_{k,m} = (d_{i,k} + d_{j,k} - d_{i,j}) \,/\, 2$$

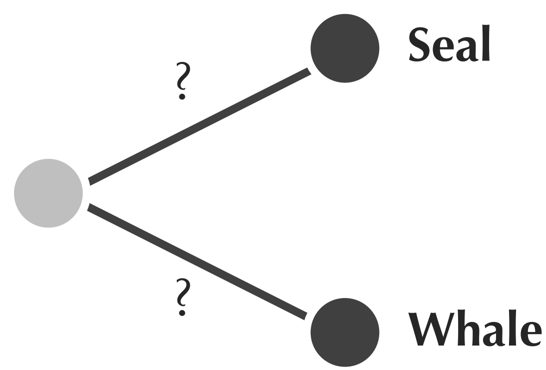
$$d_{k,m} = (D_{i,k} + D_{j,k} - D_{i,j}) \,/\, 2$$

$$\therefore \quad d_{i,m} = D_{i,k} - (D_{i,k} + D_{j,k} - D_{i,j}) \,/\, 2$$

$$d_{i,m} = (D_{i,k} + D_{i,j} - D_{j,k}) \,/\, 2$$

An Idea for Distance-Based Phylogeny

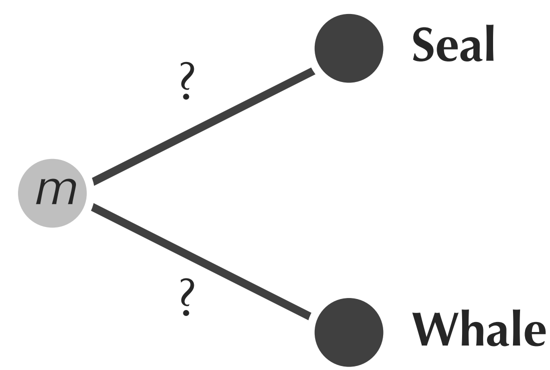
	Chimp	**Human**	**Seal**	**Whale**
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{i,m} = (D_{i,k} + D_{i,j} - D_{j,k}) / 2$$

An Idea for Distance-Based Phylogeny

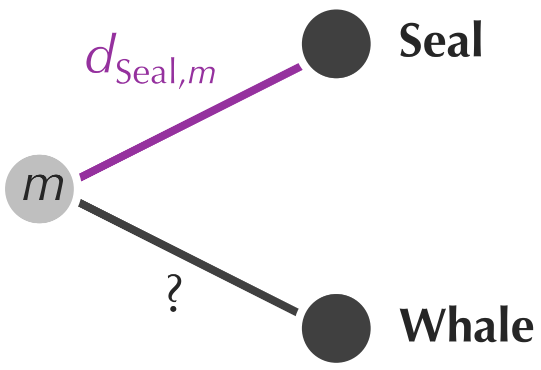
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{i,m} = (D_{i,k} + D_{i,j} - D_{j,k}) / 2$$

An Idea for Distance-Based Phylogeny

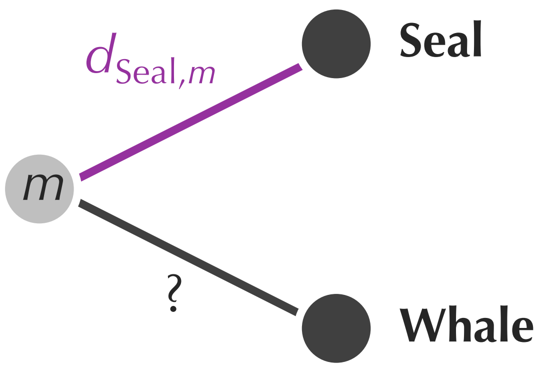
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{\text{Seal},m} = (D_{\text{Seal},k} + D_{\text{Seal},j} - D_{j,k}) \;/\; 2$$

An Idea for Distance-Based Phylogeny

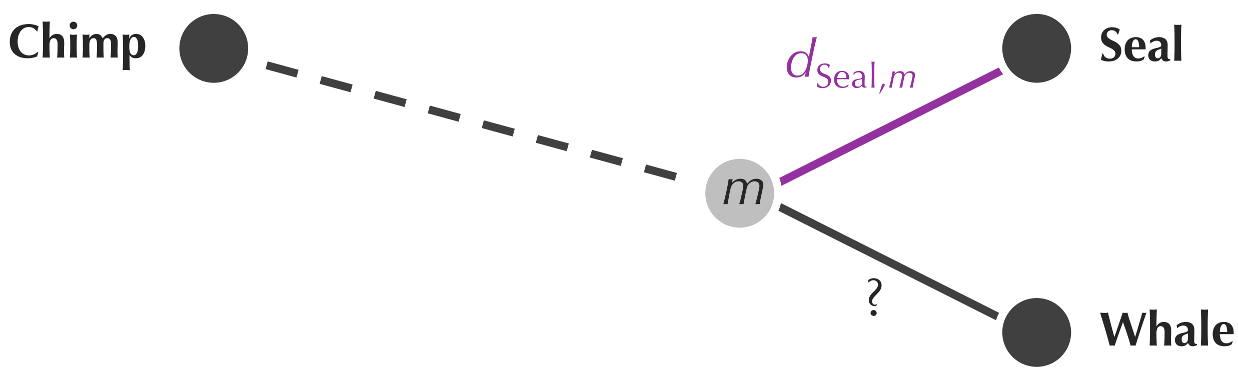
	Chimp	**Human**	**Seal**	**Whale**
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{\text{Seal},m} = (D_{\text{Seal},k} + D_{\text{Seal},\text{Whale}} - D_{\text{Whale},k}) \,/\, 2$$

An Idea for Distance-Based Phylogeny

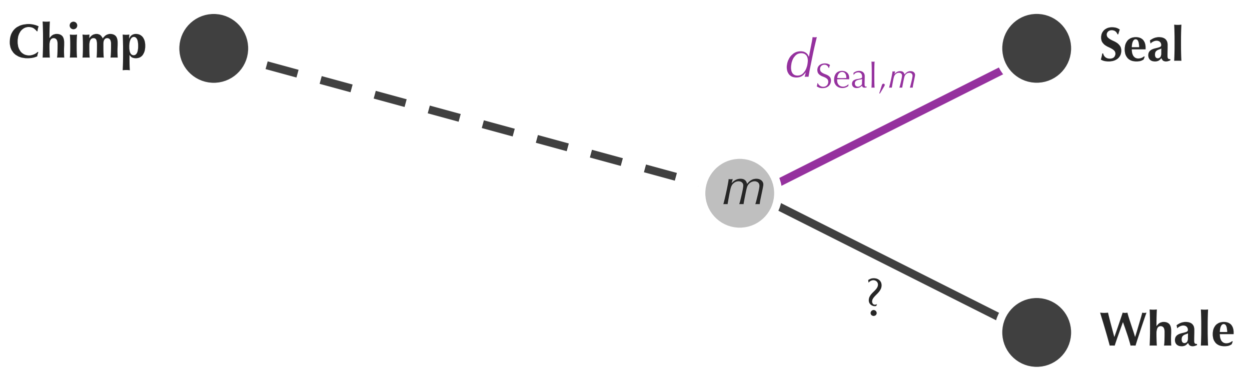
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{\text{Seal},m} = (D_{\text{Seal,Chimp}} + D_{\text{Seal,Whale}} - D_{\text{Whale,Chimp}}) \,/\, 2$$

An Idea for Distance-Based Phylogeny

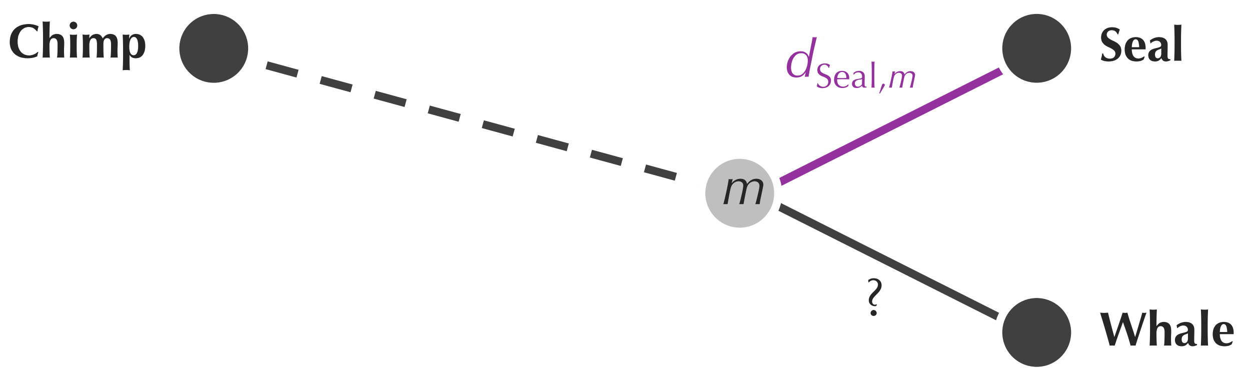
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{\text{Seal},m} = (\quad 6 \quad + D_{\text{Seal,Whale}} - D_{\text{Whale,Chimp}}) \;/\; 2$$

An Idea for Distance-Based Phylogeny

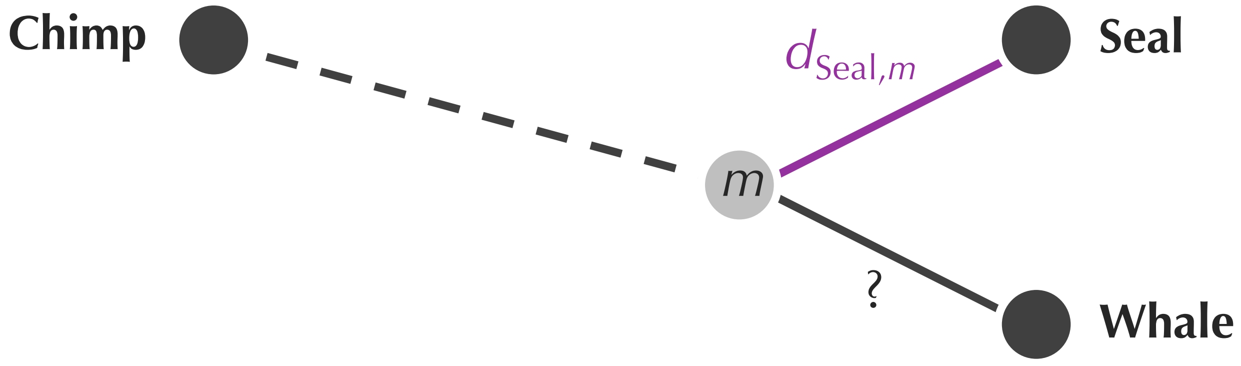
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{\mathrm{Seal},m} = (\quad 6 \quad + \quad 2 \quad - D_{\mathrm{Whale,Chimp}}) / 2$$

An Idea for Distance-Based Phylogeny

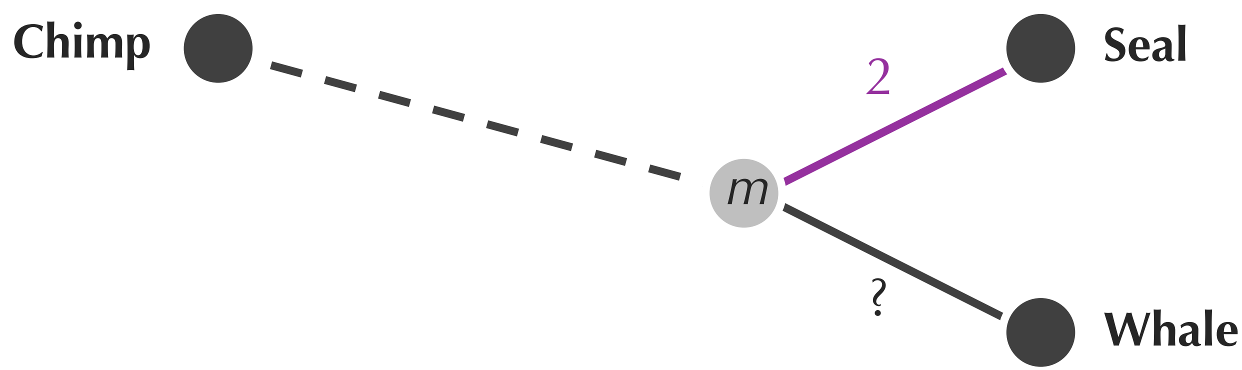
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$$d_{\text{Seal},m} = (\quad 6 \quad + \quad 2 \quad - \quad 4 \quad) / 2$$

An Idea for Distance-Based Phylogeny

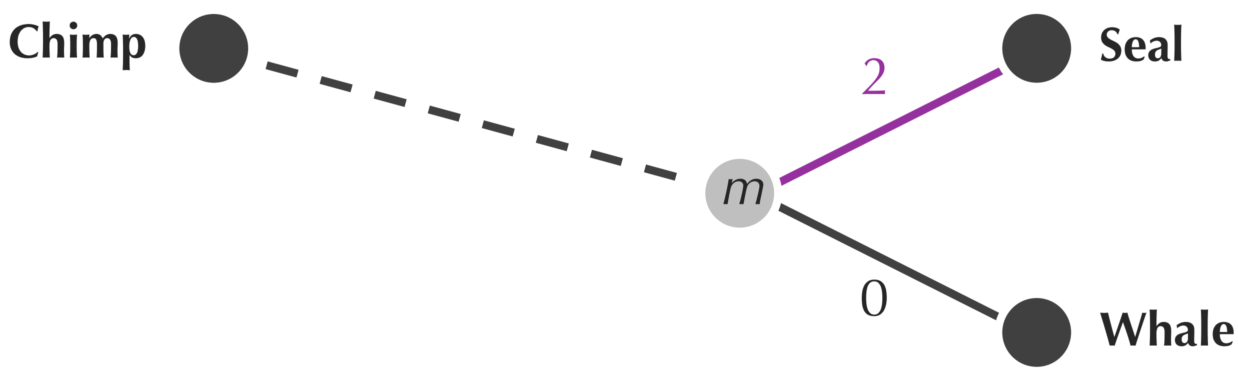
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$d_{\text{Seal},m} = 2$

An Idea for Distance-Based Phylogeny

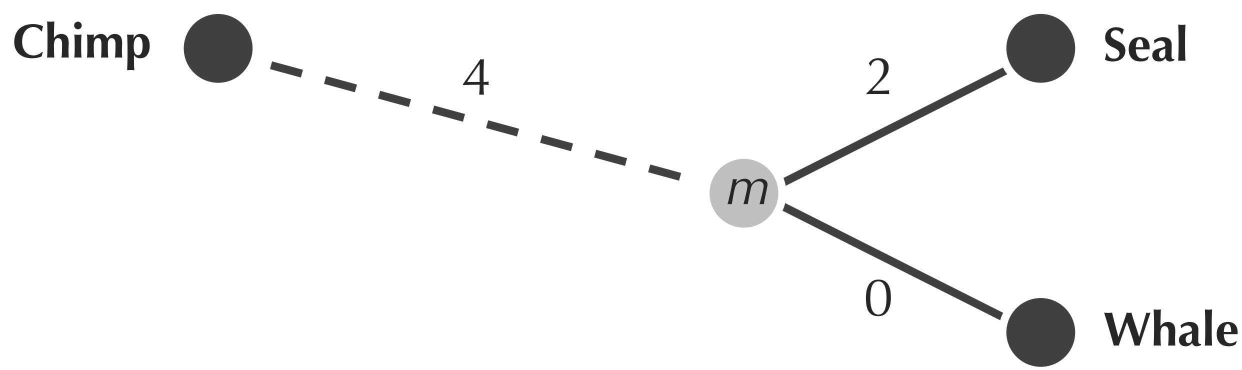
	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



$d_{\text{Seal},m} = 2$

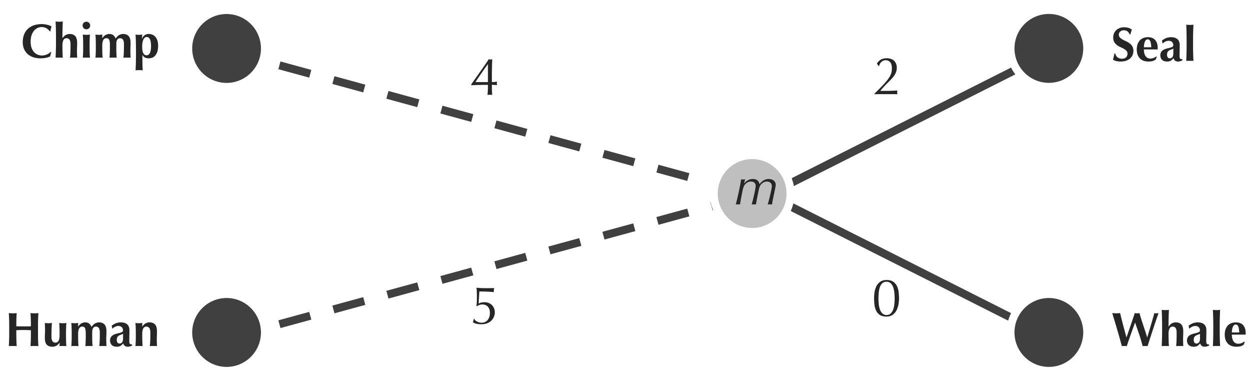
An Idea for Distance-Based Phylogeny

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



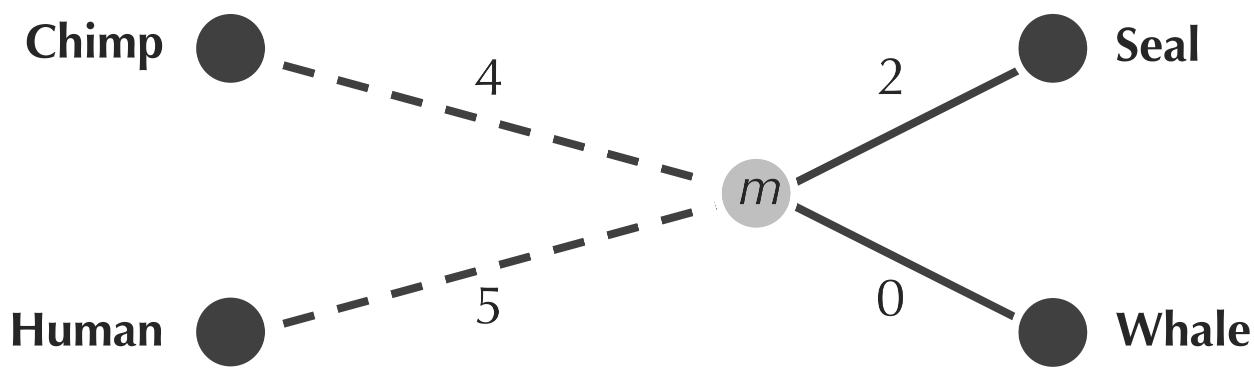
An Idea for Distance-Based Phylogeny

	Chimp	Human	Seal	Whale	*m*
Chimp	0	3	6	4	4
Human	3	0	7	5	5
Seal	6	7	0	2	2
Whale	4	5	2	0	0
m	4	5	2	0	0



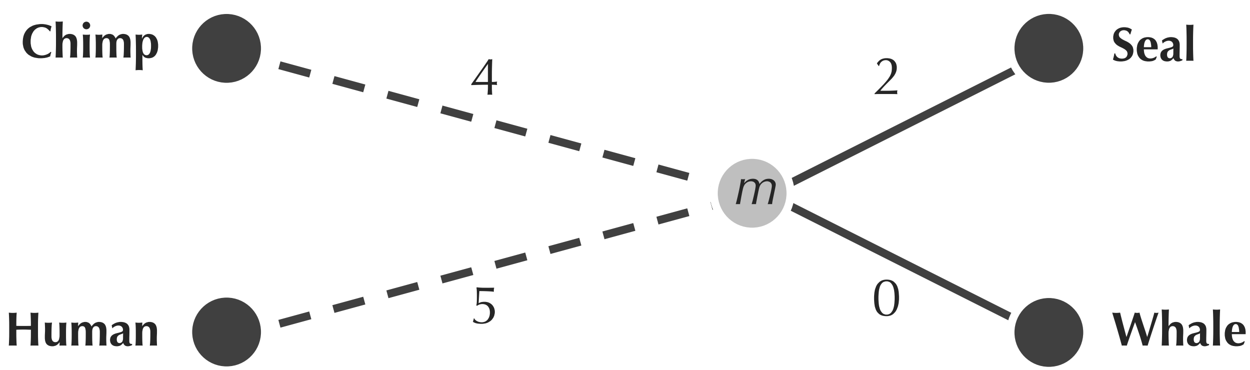
An Idea for Distance-Based Phylogeny

	Chimp	Human	Seal	Whale	*m*
Chimp	0	3	6	4	4
Human	3	0	7	5	5
Seal	6	7	0	2	2
Whale	4	5	2	0	0
m	4	5	2	0	0



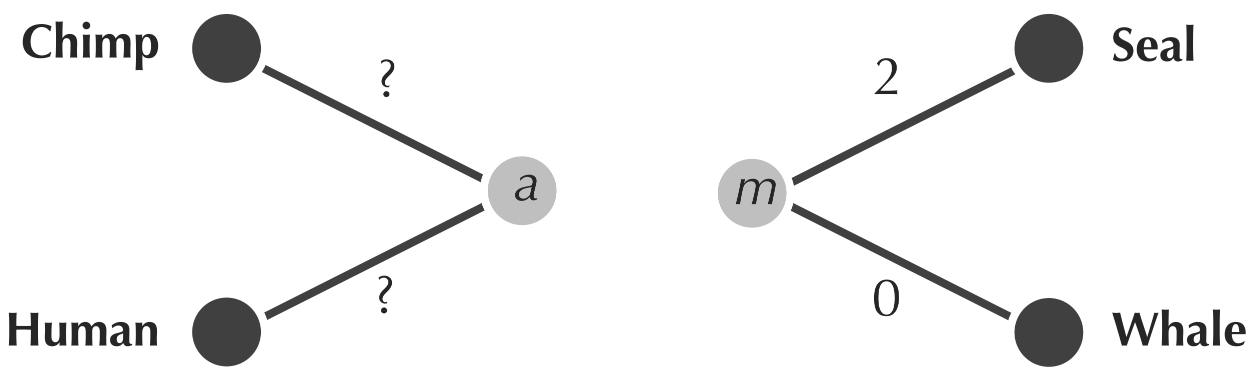
An Idea for Distance-Based Phylogeny

	Chimp	Human	*m*
Chimp	0	3	4
Human	3	0	5
m	4	5	0



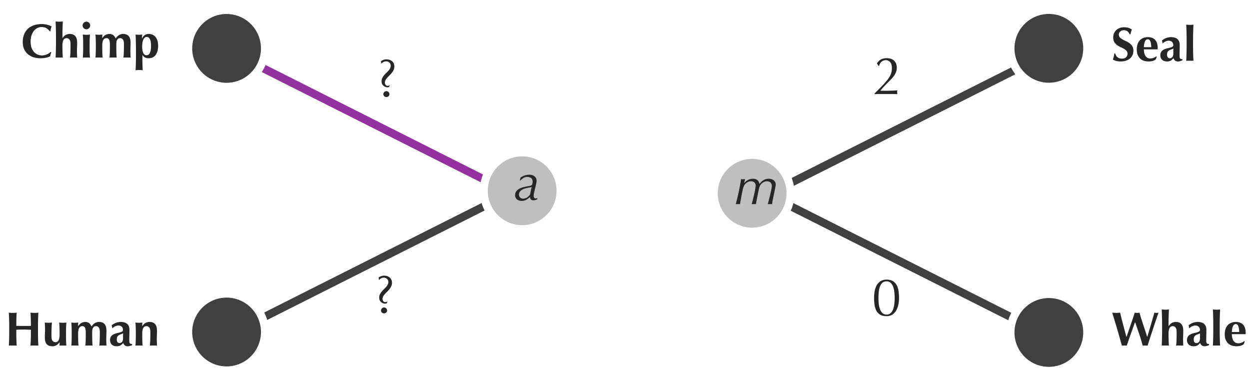
An Idea for Distance-Based Phylogeny

	Chimp	**Human**	***m***
Chimp	0	**3**	4
Human	**3**	0	5
m	4	5	0



An Idea for Distance-Based Phylogeny

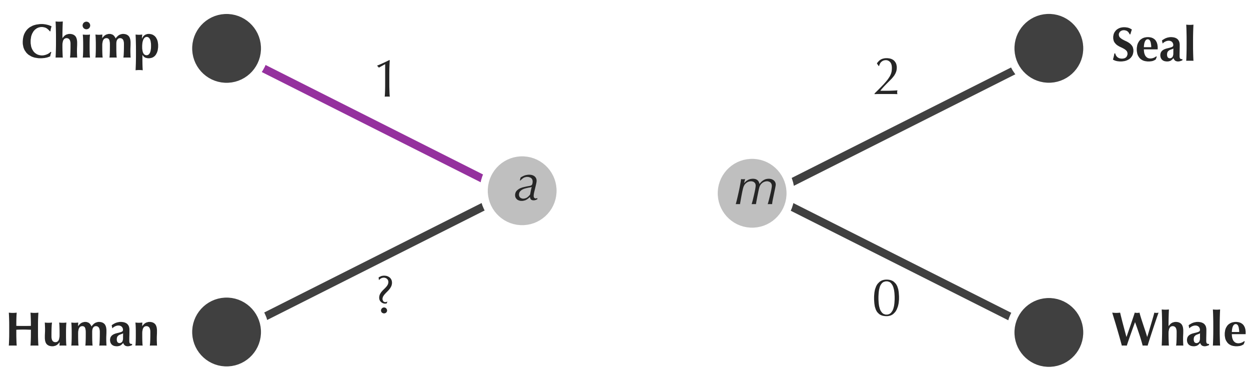
	Chimp	Human	*m*
Chimp	0	3	4
Human	3	0	5
m	4	5	0



$$d_{\text{Chimp},a} = (D_{\text{Chimp},m} + D_{\text{Chimp,Human}} - D_{\text{Human},m}) / 2$$

An Idea for Distance-Based Phylogeny

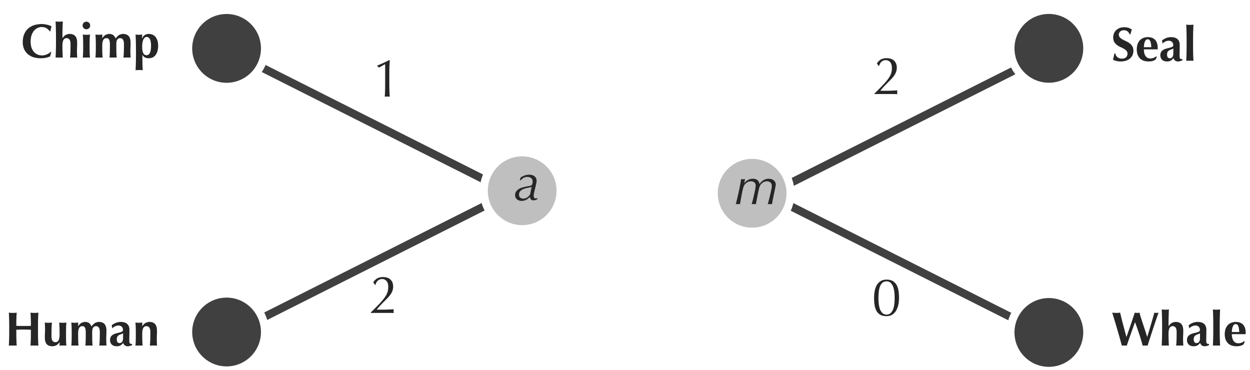
	Chimp	**Human**	***m***
Chimp	0	3	4
Human	3	0	5
m	4	5	0



$d_{\text{Chimp},a} = 1$

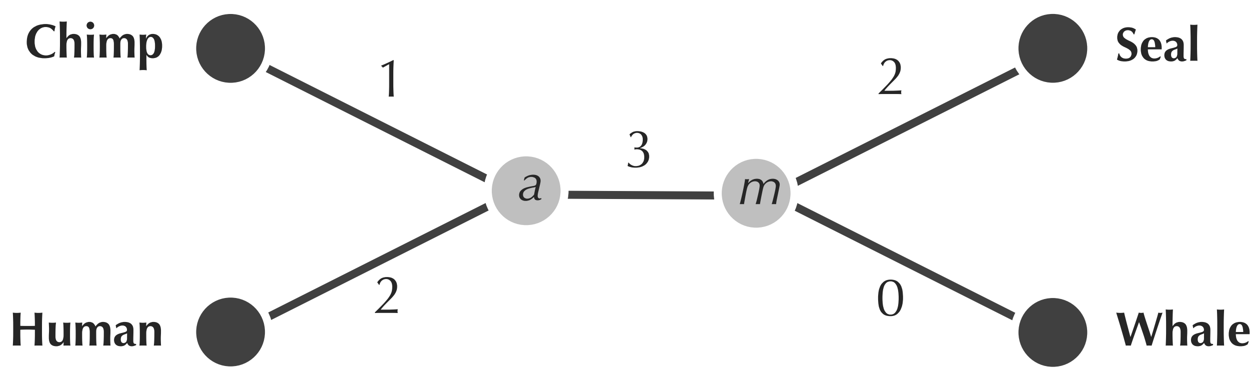
An Idea for Distance-Based Phylogeny

	Chimp	Human	*m*
Chimp	0	3	4
Human	3	0	5
m	4	5	0



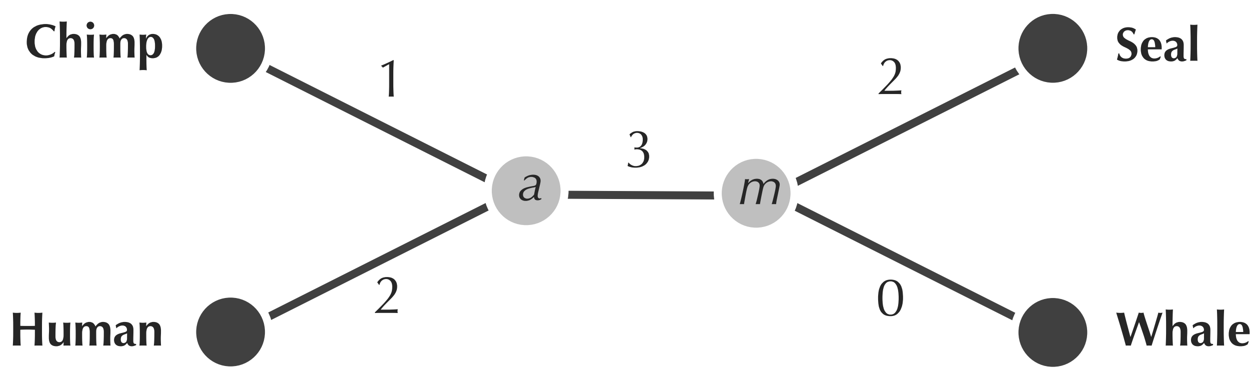
An Idea for Distance-Based Phylogeny

	Chimp	Human	*m*
Chimp	0	3	4
Human	3	0	5
m	4	5	0



An Idea for Distance-Based Phylogeny

	Chimp	Human	Seal	Whale
Chimp	0	3	6	4
Human	3	0	7	5
Seal	6	7	0	2
Whale	4	5	2	0



An Idea for Distance-Based Phylogeny

	0	**1**	**2**	**3**
0	0	13	21	22
1	13	0	12	13
2	21	12	0	13
3	22	13	13	0

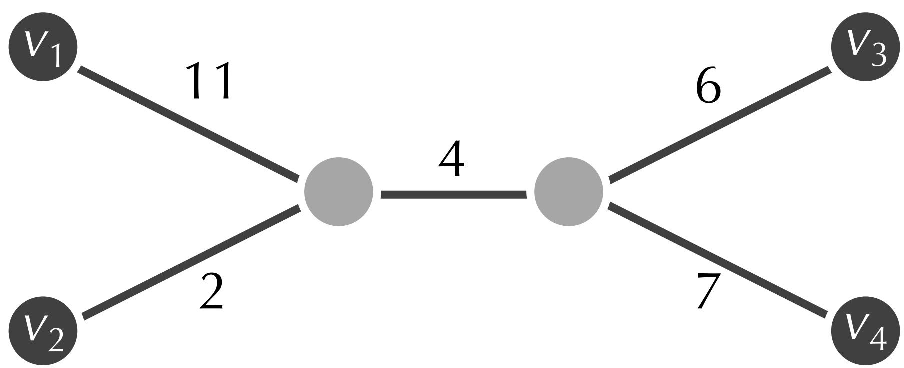
Exercise: Apply this recursive approach to this distance matrix.

What Was Wrong With Our Algorithm?

	v_1	v_2	v_3	v_4
v_1	0	13	21	22
v_2	13	0	12	13
v_3	21	12	0	13
v_4	22	13	13	0

What Was Wrong With Our Algorithm?

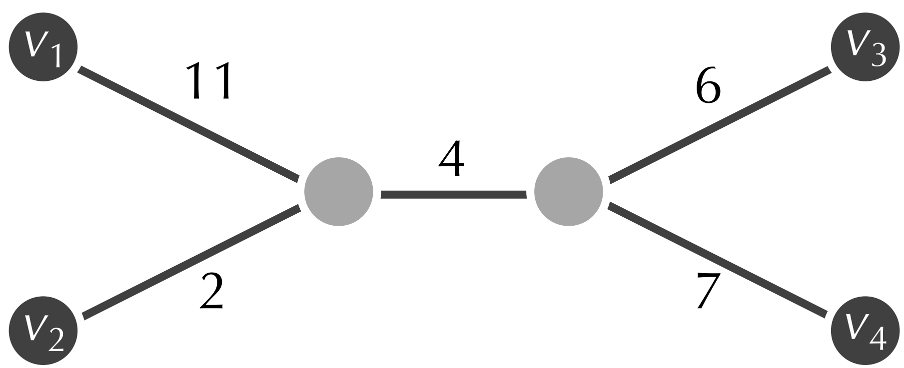
	v_1	v_2	v_3	v_4
v_1	0	13	21	22
v_2	13	0	12	13
v_3	21	12	0	13
v_4	22	13	13	0



What Was Wrong With Our Algorithm?

	v_1	v_2	v_3	v_4
v_1	0	13	21	22
v_2	13	0	**12**	13
v_3	21	12	0	13
v_4	22	13	13	0

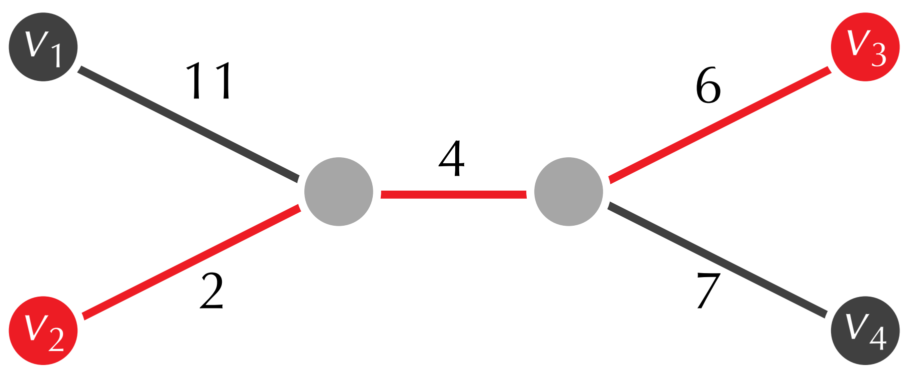
minimum element is $D_{2,3}$



What Was Wrong With Our Algorithm?

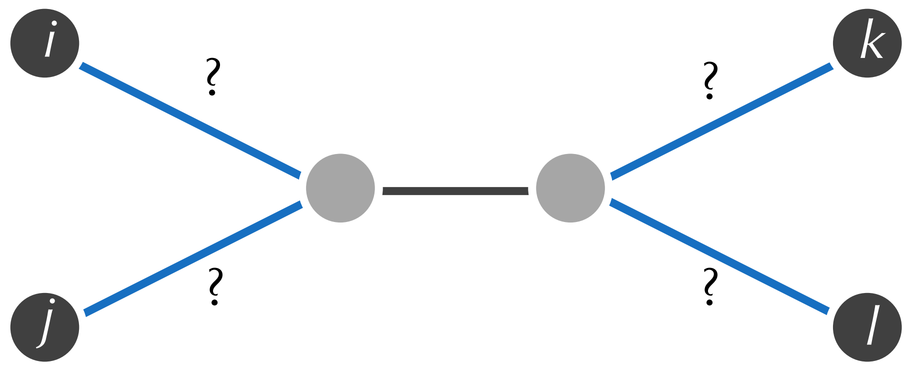
	v_1	v_2	v_3	v_4
v_1	0	13	21	22
v_2	13	0	**12**	13
v_3	21	12	0	13
v_4	22	13	13	0

minimum element is $D_{2,3}$



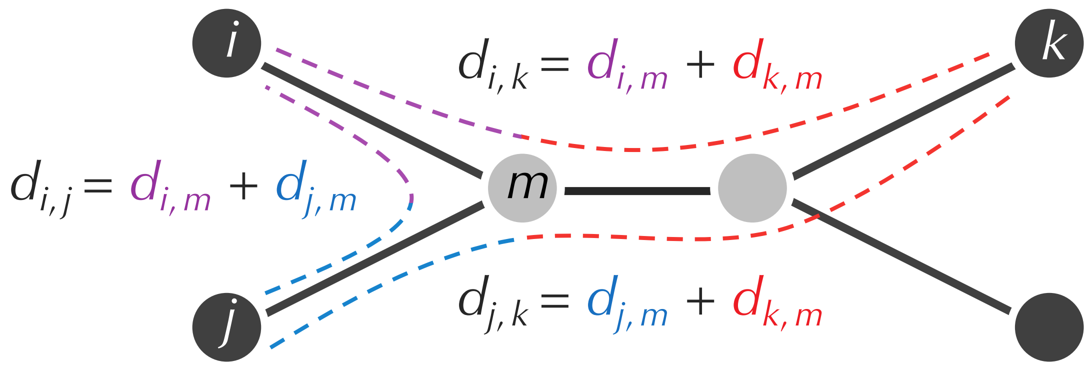
v_2 and v_3 are **not** neighbors!

From Neighbors to Limbs



Rather than trying to infer **neighbors**, let's instead try to compute the length of **limbs**, the edges attached to leaves.

From Neighbors to Limbs



$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})] / 2$$

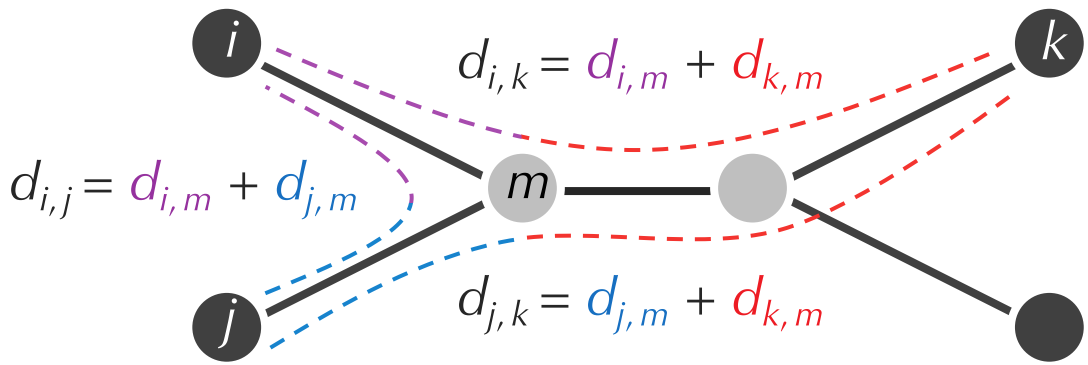
$$d_{k,m} = (d_{i,k} + d_{j,k} - d_{i,j}) / 2$$

$$d_{k,m} = (D_{i,k} + D_{j,k} - D_{i,j}) / 2$$

$$\therefore \quad d_{i,m} = D_{i,k} - (D_{i,k} + D_{j,k} - D_{i,j}) / 2$$

$$d_{i,m} = (D_{i,k} + D_{i,j} - D_{j,k}) / 2$$

From Neighbors to Limbs



$$d_{k,m} = [(d_{i,m} + d_{k,m}) + (d_{j,m} + d_{k,m}) - (d_{i,m} + d_{j,m})] / 2$$

$$d_{k,m} = (d_{i,k} + d_{j,k} - d_{i,j}) / 2$$

$$d_{k,m} = (D_{i,k} + D_{j,k} - D_{i,j}) / 2$$

$$\therefore \quad d_{i,m} = D_{i,k} - (D_{i,k} + D_{j,k} - D_{i,j}) / 2$$

$$d_{i,m} = (D_{i,k} + D_{i,j} - D_{j,k}) / 2$$

Assumes that *i* and *j* are *neighbors*...

Computing Limb Lengths

Limb Length Theorem: *LimbLength*(i) is equal to the minimum value of $(D_{i,k} + D_{i,j} - D_{j,k})/2$ over all leaves j and k.