TME2 - PHYG

Witold PODLEJSKI & Jérémie PERRIN

October 14, 2019

1 Exercise 1 : Parsimony
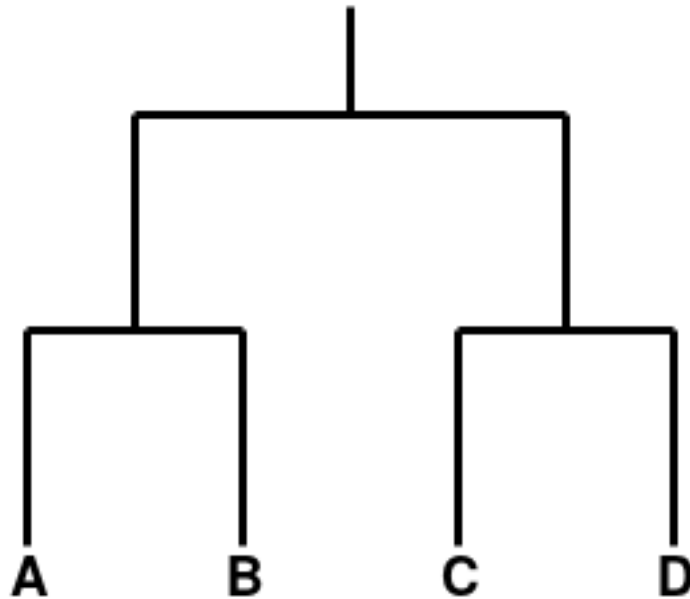
1.1 Question 1

1.2 Question 2

1.3 Question 3

1.4 Question 4

We assume that the topogy of the graph is the following :

Lets compute the parsimony scores of that tree with the Sankoff and fitch algorithms.

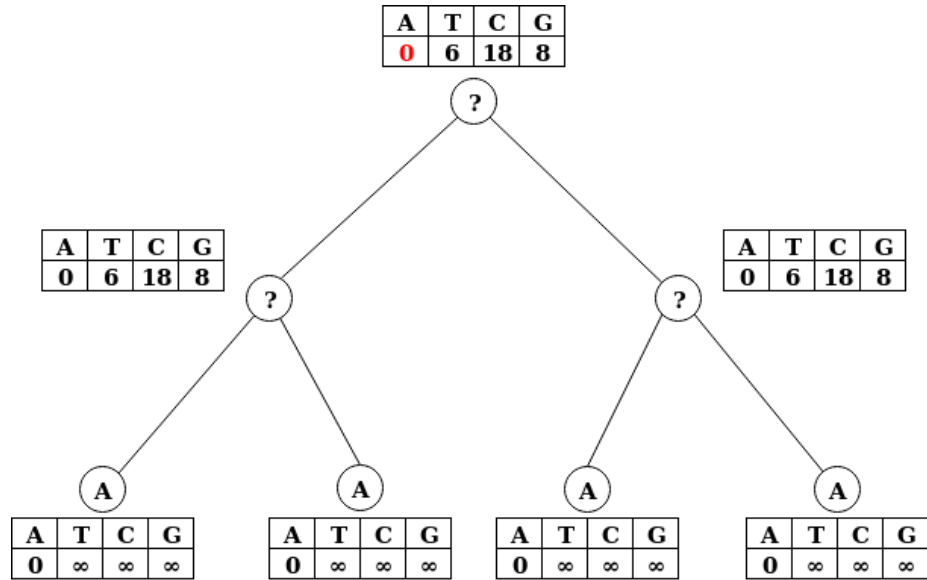## • Sankoff

We compute score tree nucleotide by nucleotide.

A	T	C	G
0	6	18	8

(?)

A	T	C	G
0	6	18	8

(?)

A	T	C	G
0	6	18	8

(?)

(A)

A	T	C	G
0	∞	∞	∞

(A)

A	T	C	G
0	∞	∞	∞

(A)

A	T	C	G
0	∞	∞	∞

(A)

A	T	C	G
0	∞	∞	∞

Figure 1: Sankoff score tree for the first nucleotide

A	T	C	G
11	6	8	6

(?)

A	T	C	G
6	0	8	4

(?)

A	T	C	G
13	6	4	4

(?)

(T)

A	T	C	G
∞	0	∞	∞

(T)

A	T	C	G
∞	0	∞	∞

(C)

A	T	C	G
∞	∞	0	∞

(G)

A	T	C	G
∞	∞	∞	0
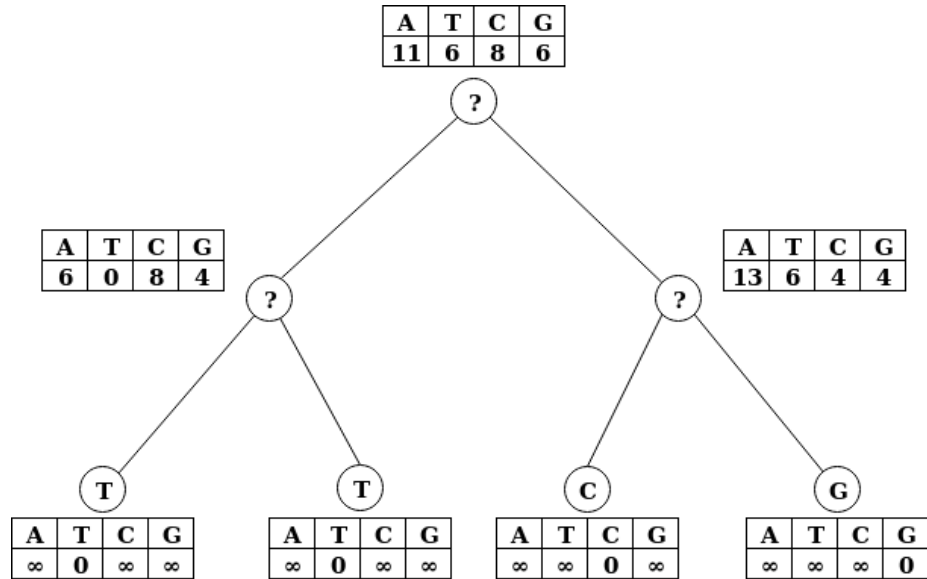
Figure 2: Sankoff score tree for the second nucleotide

A	T	C	G
13	6	4	4

?

A	T	C	G
18	8	0	8

?

A	T	C	G
8	4	8	0

?

C

A	T	C	G
∞	∞	0	∞

C

A	T	C	G
∞	∞	0	∞

G

A	T	C	G
∞	∞	∞	0

G

A	T	C	G
∞	∞	∞	0
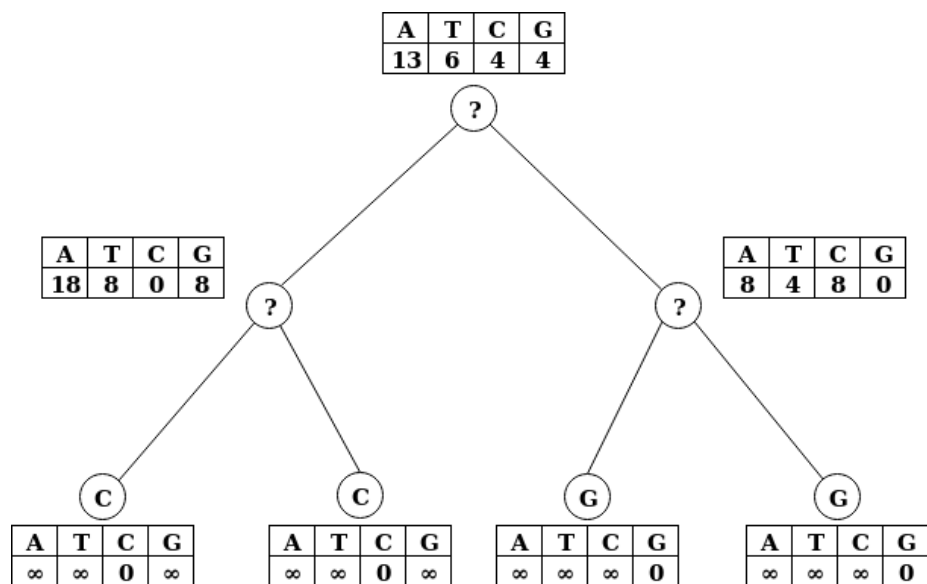
Figure 3: Sankoff score tree for the third and fourth nucleotide

A	T	C	G
14	6	6	6

?

A	T	C	G
7	2	8	2

?

A	T	C	G
18	8	0	8

?

T

A	T	C	G
∞	0	∞	∞

G

A	T	C	G
∞	∞	∞	0

C

A	T	C	G
∞	∞	0	∞

C

A	T	C	G
∞	∞	0	∞
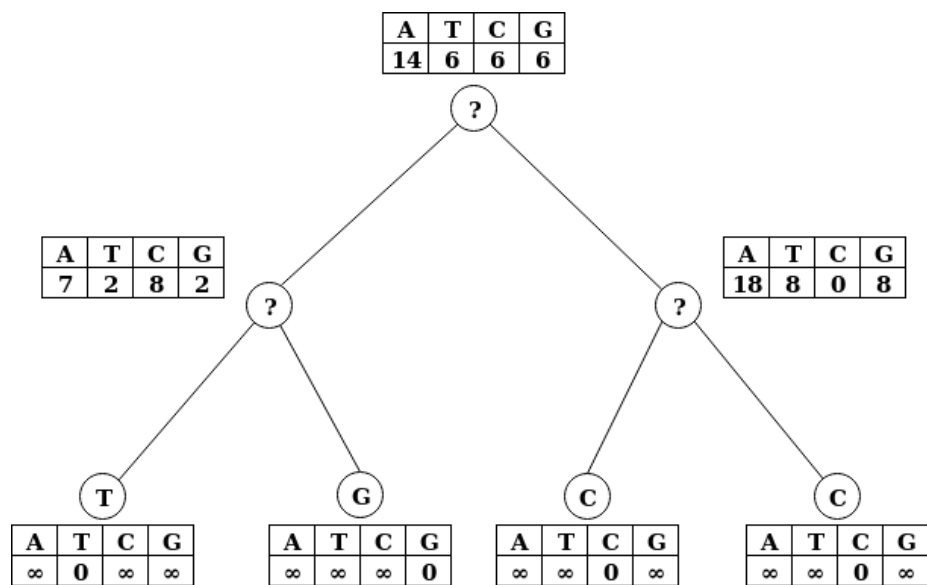
Figure 4: Sankoff score tree for the fifth nucleotide
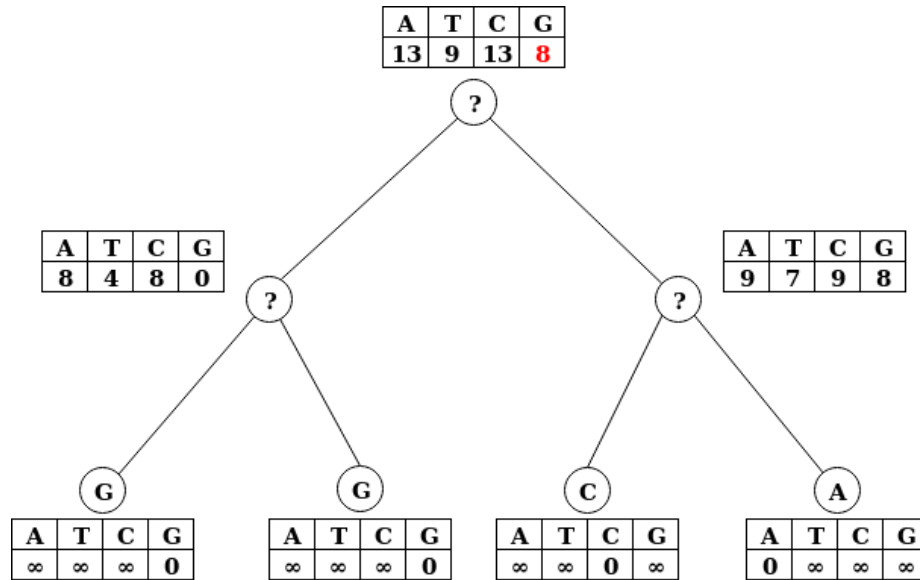
Figure 5: Sankoff score tree for the sixth nucleotide

- **Sankoff**

Again we proceed nucletide by nucleotide.

1.5 Question 5

2 Exercise 2 : Reconstruction using reversal distances

2.1 Question 1

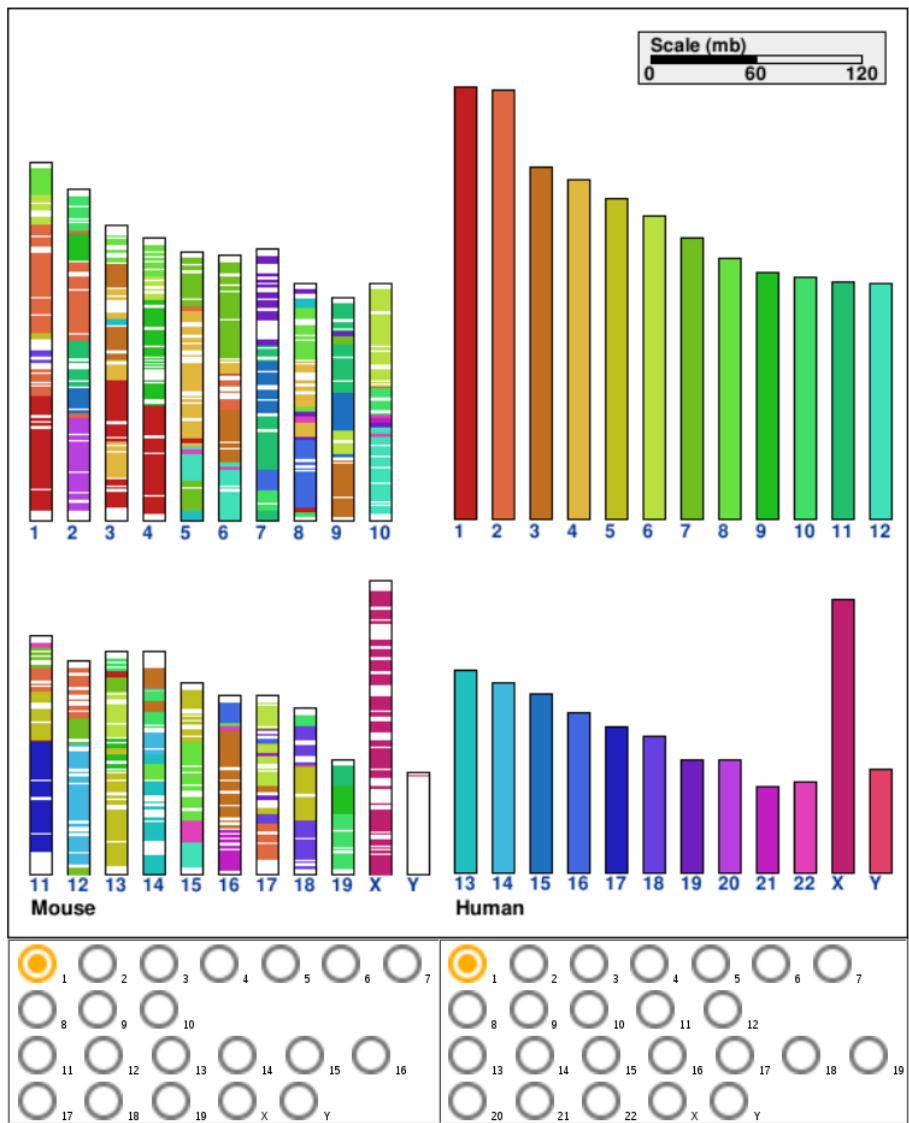So we launch an entire genome comparison between human and mouse with the human as a reference.

Figure 6: Genome comparison between human (reference) and mouse

As we can see, most of the genome is the same between this two species. However the genes are not in the same chromosomes and they seems mixed by recombination.

2.2   Question 2

Now we look at the homologous genes in the first human chromosome and the fourth mouse chromosome.

Number of synteny blocks: 14
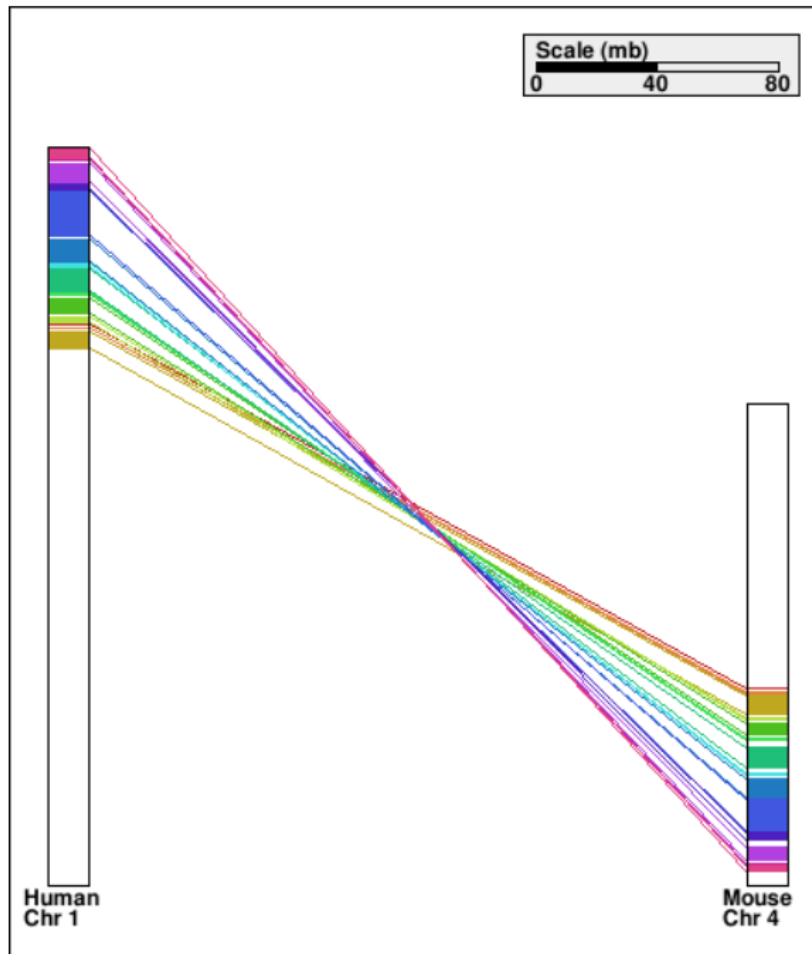Reversal Distance: 1
Breakpoint Reuse: 1.00

Figure 7: Genome comparison of the first human chromosome and fourth mouse chromosome

We can see that there is the same alignment of genes in the two chromosomes, but it is reverse. The reversal distance is about one, that mean that at least one event (recombination) is needed to go to a chromosome to an other.

2.3 Question 3

So by making a whole genome comparison between different mamals (human, mouse, cow and chimp) we obtain the reverse distances between these species. We can write it in the following matrix :

	Human	Mouse	Cow	Chimp
Human	0	302	257	18
Mouse	302	0	360	306
Cow	257	360	0	261
Chimp	18	306	261	0

2.4 Question 4

With this matrix we can launch UPGMA and NJ algorithms, and the results are the following :
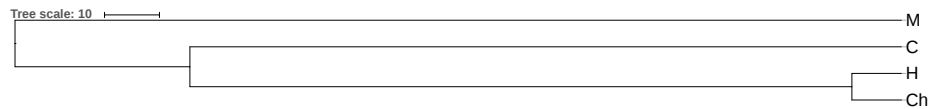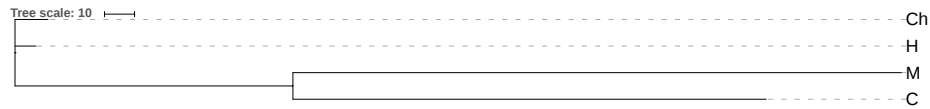
Figure 8: UPGMA tree according to the matrix above

Figure 9: Neighbor-Joinning tree according to the matrix above

These trees are not correct, in fact human and chimp are closer of mouse than cow and the trees show the contrary.

2.5 Question 5

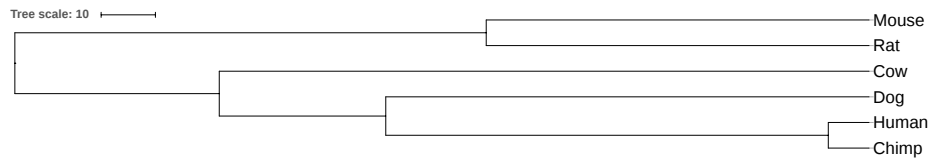With a more complete distance matrix we launch again these algorithms :
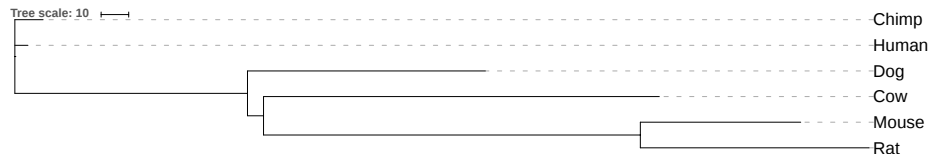
Figure 10: Neighbor-Joinning tree for all mammals

Figure 11: Neighbor-Joinning tree for all mammals

2.6 Question 6

These last trees are still not correct. The used aproach is limited because it is only based on recombination, it also assumes that the good historical sequence of recombinations is the simplest. Maybe we are seeing the result of some convergences between human and cow for example.

3 Exercise 3 : Reconstruction using characters

3.1 Question 1

Divergent evolution creates differences between species and drive population to genomic diversity. A contrario, convergent evolution bring species closer by getting the same characters.
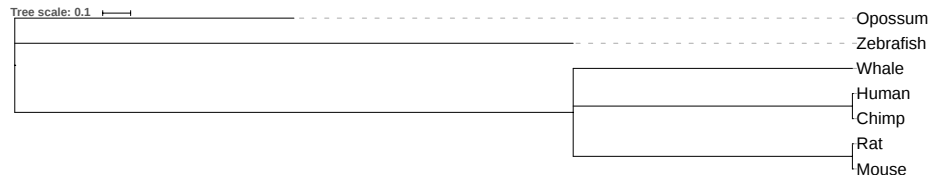
3.2 Question 2

Figure 12: Tree optain with pars method

That tree is correct, the subtree of mamals match with the given correct tree and it seems obvious that the opossum is closer from mamals than zebrafish.
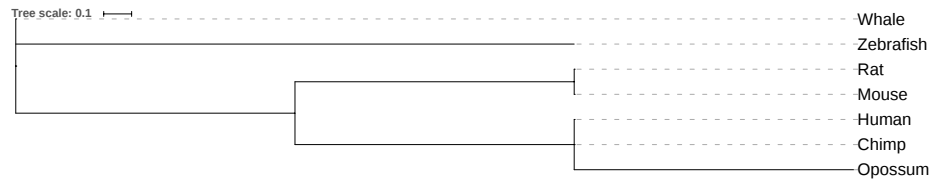
3.3 Question 3

Figure 13: Tree optain with pars method (without enlarged malleolus character)

When we remove the enlarged malleolus character from the computation, it appears that the resulting tree is greatly change. The opossum is now very close from human and chimp, this is due to the opposable thumb witch is a convergent character.

3.4 Question 4

There is two obvious convergence in this table, as it has been said there is the opposable thumb among opossum and primates (human, chimp). There is also the fact that whale and zebrafish live in water witch is a convergence.