

 <p><b>INSTITUTO FEDERAL</b> Brasília</p>	<p>Instituto Federal de Brasília Campus Taguatinga Superior em Computação</p> <p><b>Ferramentas de Bioinformática 3</b></p> <p><b>Alinhamento Múltiplo</b></p> <p>1/2020</p>
--	--

Nome: Danyelle da Silva Oliveira Angelo.

Estamos estudando um conjunto de sequências de snoRNAs, um tipo de RNA não codificador pequeno responsável por modificações químicas em outros ncRNAs. Queremos descobrir se determinados snoRNAs presentes em diferentes espécies são ortólogos.

Dois genes homólogos são designados ortólogos quando estes foram separados por um evento de especiação, onde cada cópia do gene divergiu para duas espécies distintas. Sendo assim, genes ortólogos partilham um ancestral comum e têm funções iguais.

Para inferir se um conjunto de snoRNAs são ortólogos, foi dada a ideia de realizar um alinhamento múltiplo usando o programa Clustal Omega.

1) A partir das sequências contidas no arquivo [snoRNAsOrtologos.fasta](#).

a) Qual tipo de snoRNA será comparado?

**SNORD102 ou snoRNA U102**

b) Procure sobre este snoRNA no banco de dados RFAM (<https://rfam.xfam.org/>).

i) Qual classe de snoRNA estas sequências são classifica- das?

**C/D box**

ii) Quais sequências de motivos são conservadas nessa família de snoRNA?

**UGAUGA (C Box) e CUGA (D Box)**

c) Liste quais organismos serão comparados em nosso alinhamento múltiplo.

- **Canis familiaris (cachorro)**
- **Dasyus Novemcinctus (tatu-galinha)**

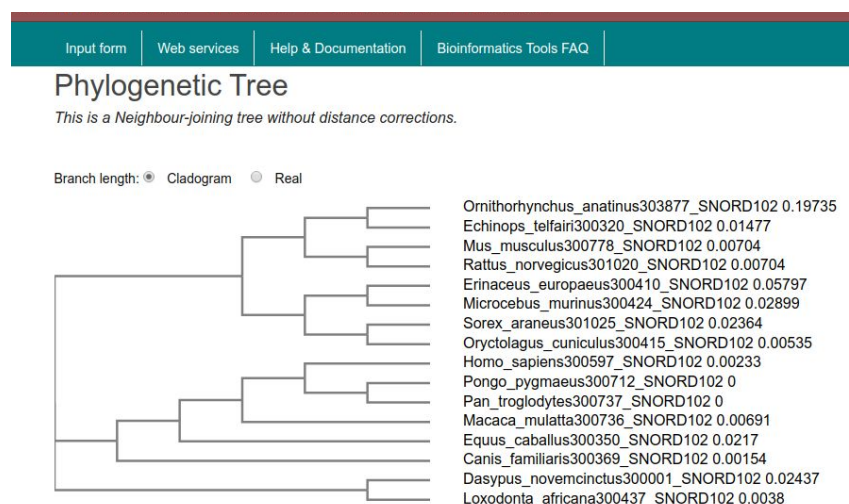
- Echinops Telfairia (Ouriço-cacheiro)
- Equus Caballus (cavalo)
- Erinaceus Europaeus (ouriço-terrestre)
- Homo sapiens (homem)
- Loxodonta Africana (elefante da savana)
- Macaca Mulatta (macaca mulatta)
- Microcebus Murinus (Lêmure-rato-cinza)
- Mus Musculus (camundongo)
- Ornithorhynchus Anatinus (Ornitorrinco)
- Oryctolagus Cuniculus (Coelho-europeu)
- Pan Troglodytes (Chimpanzee)
- Pongo Pygmaeus (orangotango-de-bornéu)
- Rattus Norvegicus (ratazana)
- Sorex araneus (musaranho)

d) Utilize a ferramenta web *Clustal Omega* (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) para alinhar as sequências de snoRNAs. Obs.: Use todos os parâmetros *default*. Note que as sequências do arquivo fasta, embora serem de RNAs, estão no formato DNA. A partir da saída do programa, realize as tarefas a seguir:

- Faça o download do arquivo de alinhamento (anexe o arquivo junto com seu documento de respostas no envio da atividade).

Ok

- Tire um *printscreen* da árvore filogenética/guia gerada.



- Crie uma espécie de *ranking* que mostra quais organismos são mais próximos do *Homo sapiens* quando comparada esta família de snoRNAs.

De acordo com a **árvore guia** os organismos mais próximos do Homo Sapiens são:

1. Dasypus Novemcinctus
2. Erinaceus europaeus
3. Ornithorhynchus Anatinus
4. Equus Caballus
5. Echinops Telfairia
6. Mus Musculus
7. Sorex Araneus
8. Rattus Norvegicus
9. Loxodonta Africana
10. Microcebus Murinus
11. Oryctolagus cuniculus
12. Macaca Mulatta
13. Pongo pygmaeus
14. Canis Familiaris
15. Pan troglodytes

- iv) Busque os motivos conservados dessa família de snoRNA. O alinhamento realizado mostra conservação desses motivos em diferentes organismos? Justifique sua resposta.

No primeiro bloco da imagem abaixo vemos a conservação do motivo C (UGAUGA) já no segundo bloco vemos a conservação do motivo D (CUGA) - destacados em amarelo e laranja. Entretanto esse alinhamento não mostra a conservação dos motivos citados para o organismo Pan Troglodytes, o Chimpanzee (ver linha 50 e 73 no primeiro e segundo bloco respectivamente).

Input form	Web services	Help & Documentation	Bioinformatics Tools FAQ	
Ornithorhynchus anatinus303877_SNORD102			TGCTCGATGATGATTGTCTCGTGTGATTGTTCTTTTGCA--AGAGCAGAAATGAAC-ATA	56
Homo sapiens300597_SNORD102			AGCTTAAATGATGACTGTTTTT-TTT-----GATTGCT--TGAAGCAATGTGAAA-AAC	49
Dasyus novemcinctus300001_SNORD102			AGTTTGATGATGATTGTCTT--TTT-----GATTGCTT-TAAAGCAATGTGAAA-AAC	49
Erinaceus europaeus300410_SNORD102			AGCTTAAATGATGAGTATCTT--GTT-----GAT-GCT--TAAAGCAATGTGAAAAAAG	48
Equus caballus300350_SNORD102			AGCTTAAATGATGATTGCCTT--TTT-----GATTGCT--TGAAGCAACGTGAAA-AAC	48
Echinops telfairi300320_SNORD102			AGCTTGATGATGATTG-----TTT-----GATTGCTT-TGAAGCAACATGTGAAA-AAT	45
Mus musculus300778_SNORD102			AACTTAAATGATGATTGTCTG--TTT-----GATTGCT--TGATGCAATGTGAAA-ACC	48
Sorex araneus301025_SNORD102			AGTTTAAATGATGATTATCTT--TTT-----GATTGCTTTTGAAGCAATATGAAA-AAC	50
Rattus norvegicus301020_SNORD102			AACTTAAATGATGATTGTCTA--TTT-----GATTGCT--TGATGCAATGTGAAA-ACC	48
Loxodonta africana300437_SNORD102			AGCTTGATGATGATTGTCTT--TTT-----GATTGCT--TTGAAGCAATGTGAAA-AAC	49
Microcebus murinus300424_SNORD102			AGCTTAAATGATGATTATCT--TTTT-----GATTGCT--CGAAGCAAAATGTGAAA-AAC	48
Oryctolagus cuniculus300415_SNORD102			AGCTTAAATGATGATTGTC--TTTT-----GATTGCT--TGAAGCAATATGAAA-AAC	48
Macaca mulatta300736_SNORD102			AGCTTAAATGATGATTGTTTTTTT-----GATTGCT--TGAAGCAATGTGAAA-AAC	50
Pongo pygmaeus300712_SNORD102			AGCTTAAATGATGACTGGTTTTTTTTT-----GATTGCT--TGAAGCAATGTGAAA-AAC	50
Canis familiaris300369_SNORD102			AGCTTAAATGATGATTGTC--TTTT-----GATTGCT--TGAAGCAATGTGAAA-AAC	48
Pan troglodytes300737_SNORD102			AGCTTAAATGATGACTGGTTTTTTTTT-----GATTGCT--TGAAGCAATGTGAAA-AAC	50
			. * . * * * * * . * . * * * . . * * * . * * * . *	
Ornithorhynchus anatinus303877_SNORD102			CCATTTTCAACGGCTCTGAGAGCT	79
Homo sapiens300597_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	72
Dasyus novemcinctus300001_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	72
Erinaceus europaeus300410_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	71
Equus caballus300350_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	71
Echinops telfairi300320_SNORD102			CCATTTTCAACGGCTCTGAGAGCT	68
Mus musculus300778_SNORD102			ACATTTTCAACGGCTCTGAGAGTT	71
Sorex araneus301025_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	73
Rattus norvegicus301020_SNORD102			ACATTTTCAACGGCTCTGAGAGTT	71
Loxodonta africana300437_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	72
Microcebus murinus300424_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	71
Oryctolagus cuniculus300415_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	71
Macaca mulatta300736_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	73
Pongo pygmaeus300712_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	73
Canis familiaris300369_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	71
Pan troglodytes300737_SNORD102			ACATTTTCAACGGCTCTGAAAGCT	73
			. * * * * * * * * * * . * * *	

e) A partir do alinhamento realizado, podemos realmente inferir ortologia para essas sequências? Justifique sua resposta.

Dizemos que dois organismos são ortólogos, se e somente se, os mesmos possuem um gene ancestral que deu origem a eles. O gene de dois organismos ortólogos serão parecidos por causa desse ancestral comum, entretanto os mesmos terão algumas diferenças devido ao processo de especiação.

Através do alinhamento feito podemos notar a similaridade destes organismos através dos motivos C/D box (destacados na questão anterior), então sim, através do alinhamento e da definição de ortologia, podemos inferir que esses organismos são ortólogos.