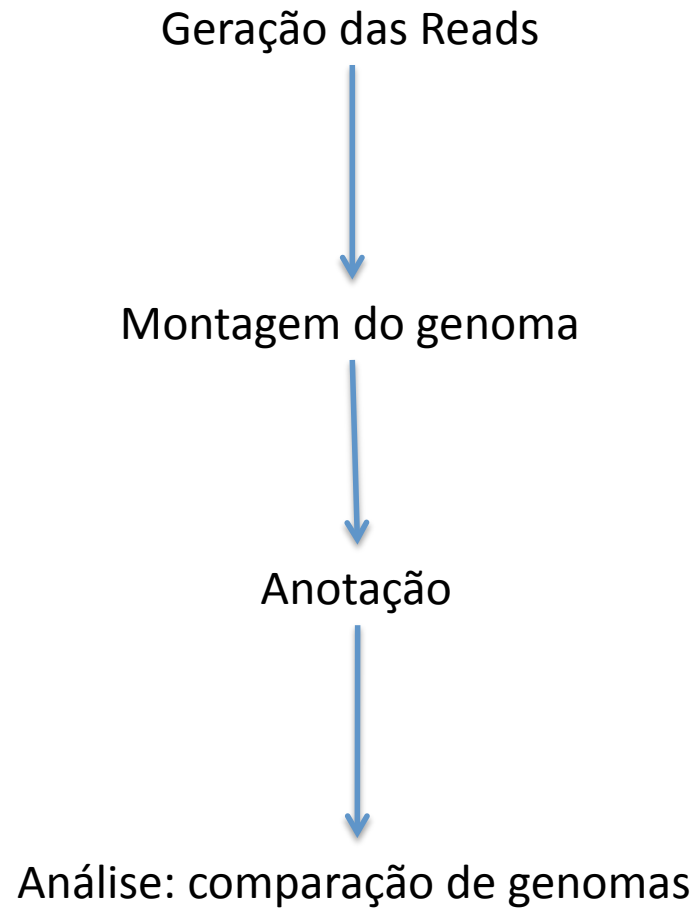


Introdução a Anotação e Comparação de Genomas

Tiago Mendes

Sequenciamento de genomas:



Sequenciamento de genomas:

Geração das Reads



Montagem do genoma

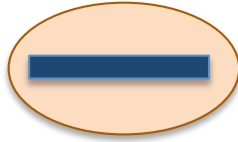


Anotação

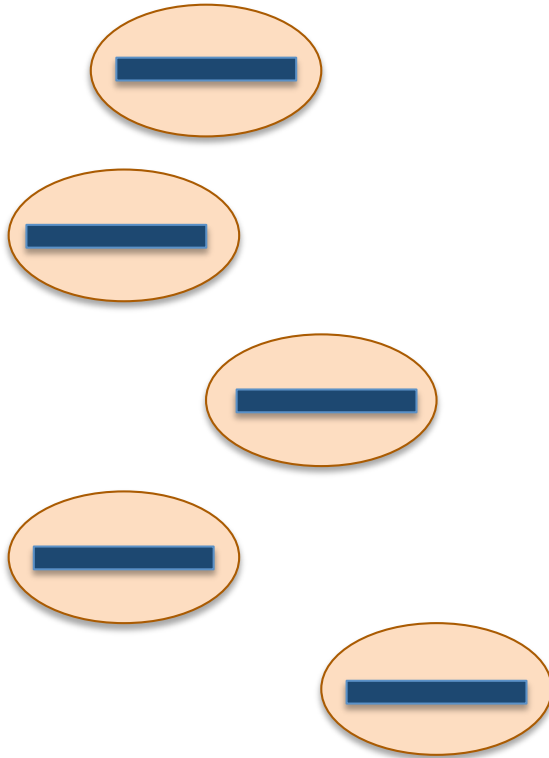


Análise: comparação de genomas

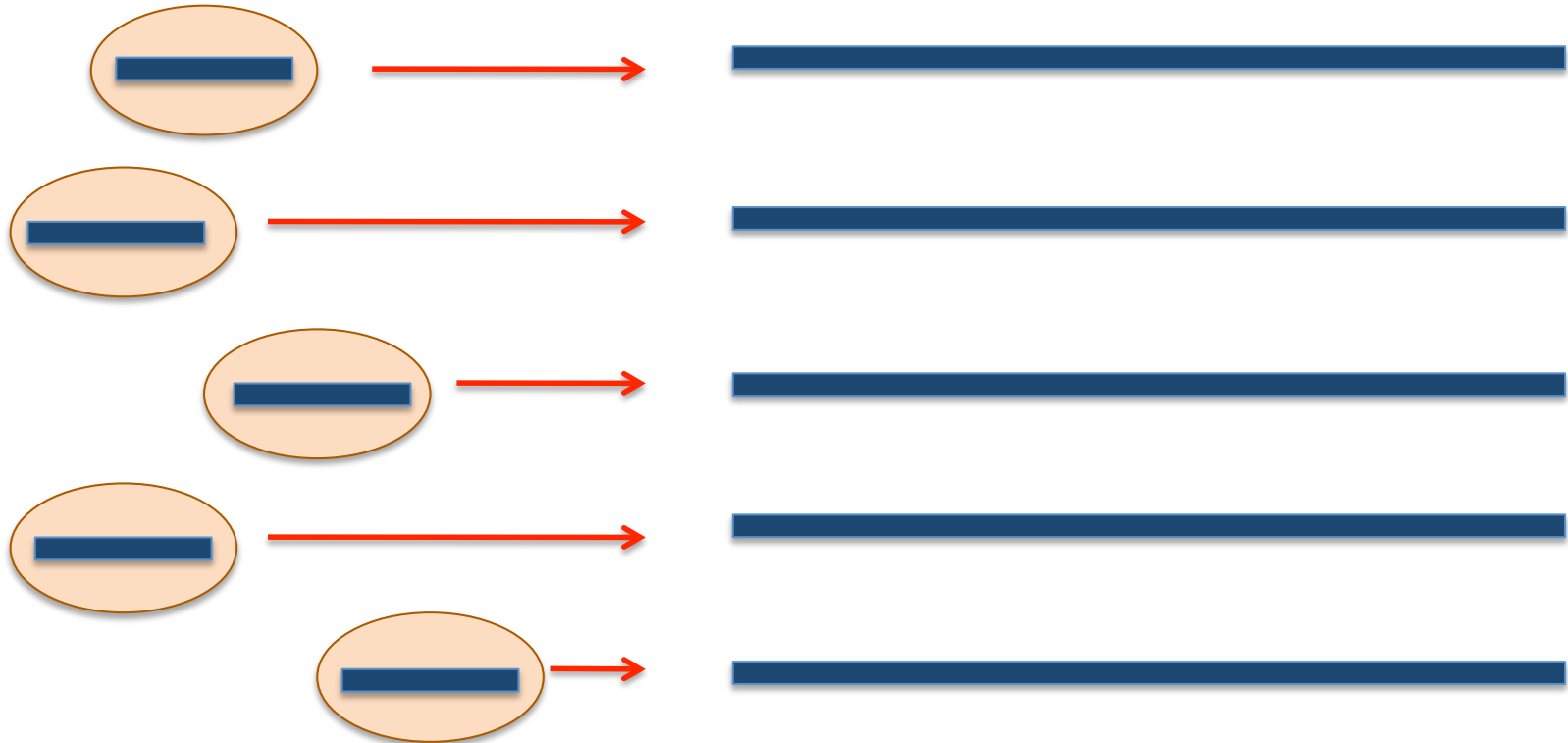
Geração de reads:



Geração de reads:



Geração de reads:



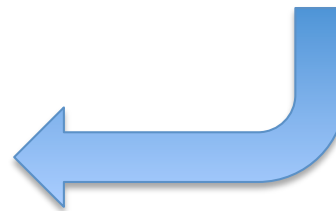
Geração de reads:



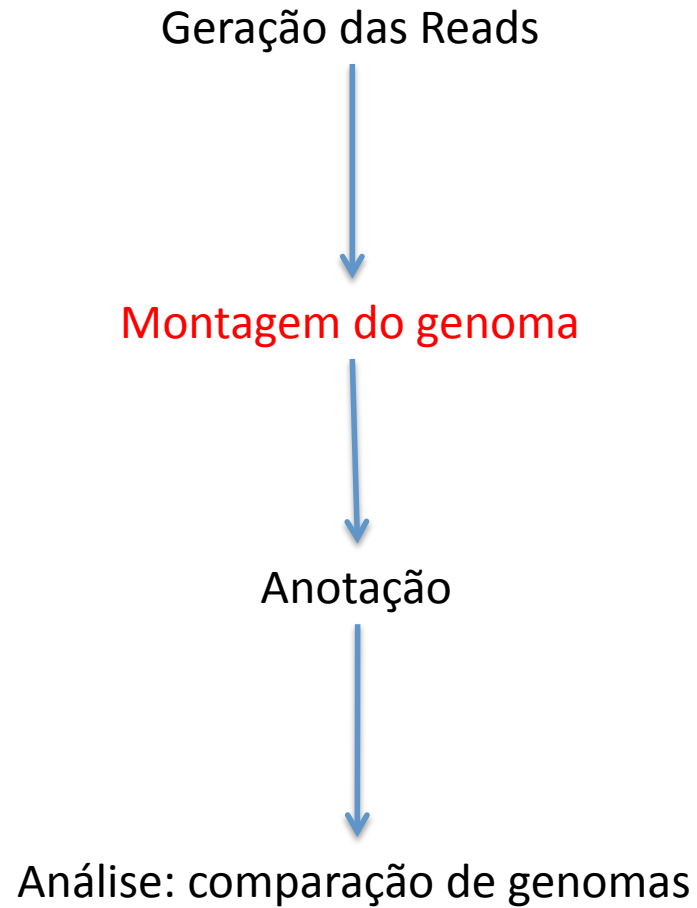
Geração de reads:



SEQUENCIAMENTO



Sequenciamento de genomas:



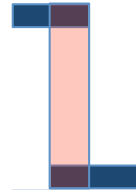
Montagem do genoma:



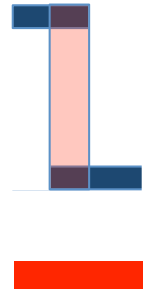
Montagem do genoma:



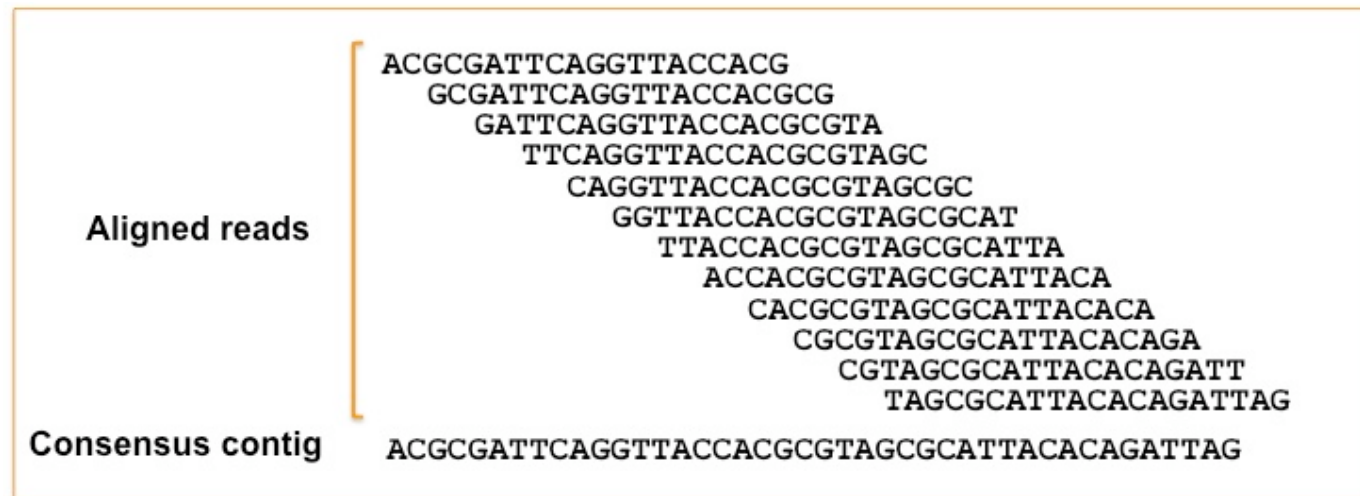
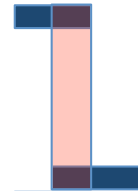
Montagem do genoma:



Montagem do genoma:



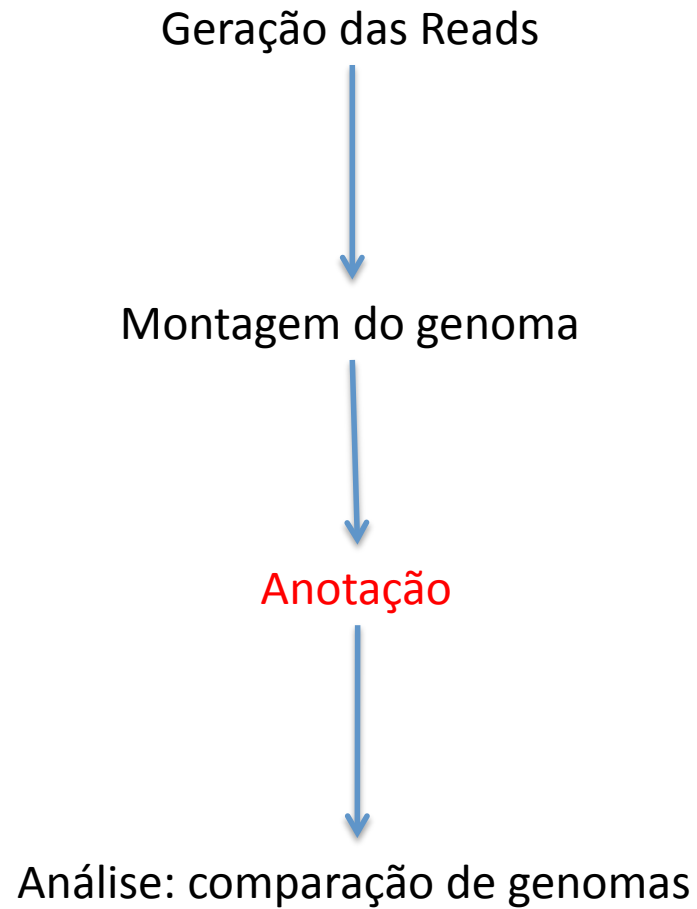
Montagem do genoma:



Montagem do genoma:



Sequenciamento de genomas:



Anotação:

TCTACCGCAAACAACACCGT**CGAGGTGACTTACGAGGATACTCTCCCAAGTTGTACCGACTAACGCAGAG**TGAAGTGGTCGTGAGGCGGGAGAGAGAGGGGGGGGTGAGG
AGTTGGTATGCTTCCCTACCTGTGCGTCTACCATTGGCAGTGTAGTCTCTGAGATGCCTCATTGGCGGTA CTCTGACCTCGCTTTTCTCTCTCCCGTTGCTGAGCCGTGACTTCTG
CTCTGGTTGTGGCGGTAGGCTAGGGGAGGGAGGGGGAGGGGGGACGGCGGTGTTCTGTCCTCCCC**CTTTGCTTGTGGAGATGCAGGCATGTTTTG**CATGTGCTGAGGACAAC
GATGTACTGTTGGCTGCAACCTGGAAGTGTGGGGCAGAGAGAGAGGGCTCGAGGAATGGAGCATGGAGGCTCATGGAGAGGTCTTGTGGTGCCTTTGATGTGGTGGCATCACACA
CCTCACTCGAAACTACTCGCCAGCGAGCCTCTCCTTTCCCGCATGCGTGTGGCTCTCTCCTCTCTCTTCCGGTCTCGGCTCTTCTTCAAGGCAGCGCCAACAGCCGCAAAA
ACAAAGCGAGGCAC**AGAGGAGGAGTACTCAGACACGA**GTAATGCCGAAGCAGGTCAATTCATGCAGCAAACATGCCCGCAGGAGAAACGCCATGGCGCTGTCCGCCGACTG
CAGGTGTGCCTATGGCGAGCGCCCCCTCCCCCTCACTGAGCGCGTGCCTGCCGGCACGCCTGCGCACCATCGCCACGCGTCTTTTTTGGTTTGCAGGCGATTGGGCTCTTTCGTC
CTCTCAACCTTACGACATTTGCGCTCCGTCTCCTGCCTCGCACACTCCCTCACCTCC**TCCTCCCTCTTTCTCTTCCCTTCCTTCCTTCGACGCCGGCGCC**ACGCGCACACAGG
CACAGGGACAGACACACCTATGGACAGCGTGTCTGTGCAACATGTGGCAGACTCTGCGGCAGCGATATCGTCTTCAAGCCACATCCCTCGCAGCCTTCTGGAGTCCAACCTGC
ACGGTGTGCAAGTCTCCGACGGCGGGCGCTACACCGCAACTCGAGCCTCGTGTGCGCGAAACCAACAGCCTCTGCATCTTCAAAAACCTCGTGCAGAGGAATCACTCGCTCGTGG
TCTTCCCTCGACAAGACGCAGCGGCTGCTCCTTCTACGTGAACCGCCCCTGCGTGTGGTGGGCCGCAACCAGAACCTTCCAGGAGGTGGCGCTGCGGCGGGCGGCTGCC
GACGGCGTTTTCCGTCGCTCGCCGTGCCTCTGGCGGTGGCGCCGTCTTTCATGACGAGGGGAACCTCTGCCT**TGCTTCATCACGCACCGCACGCGTATG**CACCGGAGAAGACAAT
CCAGCTGATCCGCTTGGGCTCTGTGTGAATTATGCGATCGACCCTGCACGGCTGACCACGACAAGGCGGCACGACCTTCTTGGACGGAAAAAAGATCACCGGATCTGCAATGC
GGGTGCAGCGCGAAATTGCGTACCACACTGCACGCTGCTGCTGATAACCCACTGGCGTGTGCTCGGTGCTACCTTACCCCGAAGGGGAGTACGTGGCGTTCAAGACGTCGTCG
GTGGGCTCGGTGCGAAG**CCCTGTACCCACTCGCGGAGTCGGTCCACATTGCGAGTGGC**AGGGCGCCATGGCCTCACTCAAGAGGAACATGGCGGAGTTTTTTCTAACTGAG
GGTGATCGAGTGTGGAAGCAGCGGCACCGTGGGAGCTCGACGTGAGGGAGCTGCGGCAGTCCTTCGCCACCGCACGTCCTGCTGCGCGGACACGCCTCTTTTTCCCTTGATG
TTGTCGAGCCGTGCGGGCGGACATGTCTTTCATCGAGGGCGAGGGCAGGCGGGCGGCCAGTGGTGACCTCGCCACCCTTGGCGAGGCGGTTACAAGGCTGCGTCGAAGGAC
TGGCCCTACGCGATGCCAGCCTTACATCCACGGTGTCTCCTCAGCAGCGGTGAGCTCCAGCGGCCTACAAGCG**TTTTCCGTTTGGCCGGATGTAGTCCGGCTA**CTCTCTCGCG
GAGGAGCAGTTGCTGGCGGCACTGCAGCAATGTGTCTTTCGCGACTTGGTGGCTGCGGGGAAGTCGTGGCAGAGACCGAAGTGGGGTTGCACCTGCTTACCACAGTCGAGCAT
CGCCTCGTCACATCCGTTACTGTGAGACGCGCGCCGTGACAGATGCCGCTACTCGTGTGTGGCGCCCTTGGGCGGGTACTCGGGAGGATGGGTGGAGAGGTACCTTTCGGCCC
TGCTGTTGGCCACGCTTGTGATGCCGCTGTGCAAGGCCTGGAGAGCCTTGGCGACATCACAGCAGTCGTGCAGGGCCTCCTGCACGAGACGGGGTCACTCTGCCCGAACTGCC
CGACGTAGCCCGACGCGGCGTTGCTAATGGTAGCGGAGT**GCTTCTGAATGTGTGGCGTGATAAGAATGTCTTTGATCTC**TCGCGCTAATCCGTCAAGCCTTACGCCACCGGGA
GGTATGGGTGTGCTAATAAAGCTGTGGCCCGCAGGTACGGAAAGCAGCGAGCGCATCAGGATGGAGGCCTTTCGCTTCGCCGAGACTCACGGACGTGGCCCTGAGTCTCGGCG
ACCGAGAATCCCGAAGGGGTGACGGGTGCTGTAAGCCCTTGTGCTGCTCTCTCTCCCTCCACCCTCTTGGCCATCGTTGTCCACGTCGATCATTCCAGCGCACCGTCTCT
GTGTCAGTACTGCTGTTTGTATCTTCCCGCTTCTCTCGCCCGCACCCCTCGTTTTCCCTCTCTCGCACGCGACCGTCAGTGTGGCAGCCGTGAGCTTGGATGCTCGCTGTGC
GTGTGT**GTGCGCTTGGCGGTATAATTAATTAATTAATTTAAATT**TATACATTACCAGAATGTCTTGGTGTGTAACCTTTTCTCCAACAGTTGAATCACTAAGACTGACAAAAGAGAAGA
GATAACAATAACCCAGTTGACAGCAGCAGTGACACTAGAGAAGGAACAATGCTGAGTTTACCTCGATAATAAAGATAACAATAATAACAGGTGTGCAAAGGGACGCCAACCGTT
GGTGTACGTCTCTTTTACTATCGCGGCGGTAGAAACAGAATTTTTCCATTTTTTGGGCGACCCCTCCGTTTACTTTTACCTCGATAATAAAGATAACAATAATAACAGGTGTGCAA
GGGACGCCAACCGTTGGTGTACGTCTCTTTTACTATCGCGGCGGTAGAAACAGAATTTTTCCATTTTTTGGGCGACCCCTCCGTTTACTTTTACCTCGATAATAAAGATAACAATA
ATAACAGGTGTGCAAAGGGACGCCAACCGTTGGTGTACGTCTCTTTTACTATCGCGGCGGTAGAAACAGAATTTTTCCATTTTTTGGGCGACCCCTCCGTTTACTTTTACCTCGA
TAATAAAGATAACAATAATAACAGGTGTGCAAAGGGACGCCAACCGTTGGTGTACGTCTCTTTTACTATCGCGGCGGTAGAAACAGAATTTTTCCATTTTTTGGGCGACCCCTCC
GTTTACTTTTACCTCGATAATAAAGATAACAATAATAACAGGTGTGCAAAGGGACGCCAACCGTTGGTGTACGTCTCTTTTACTATCGCGGCGGGAATTTTATCGGCGATATATAA

Anotação:



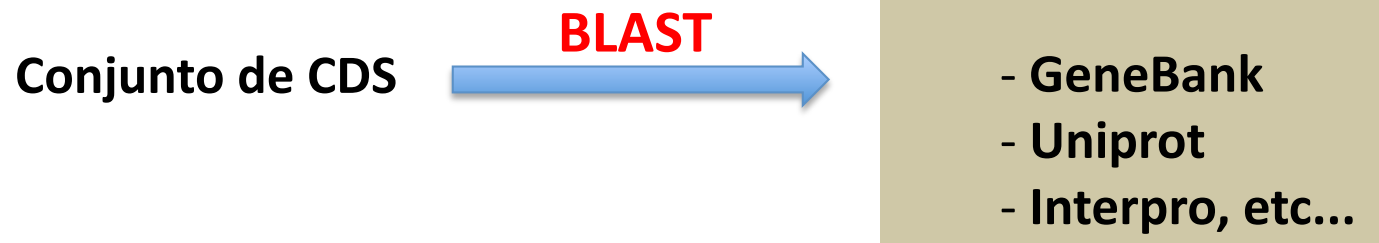
Anotação:

1- Anotação de nucleotídeos:

TCTACCGCAAACAACACCGT **Regiões reguladoras** GAGAGAGAGGGGGGGGTGAGGAGTTGGTAT
GCTTCCCTACCTGTGCGTCTACCATGGCAGTGTAGTCCCTGAGATGCCTCATTGGCGGTA CTGACCTCGC
TTTTCTCTCTCCCGTTCGCTGAGCCGTGACTTCTGCTCTGGTTGTGGCGGTAGGCTAGGGGAGGGAGGGGG
AGGGGGGA **CDS** CTCTTGCTTGTGGAGATGCAGGCATGTTTTGCATGTGCTGAGG
ACAACGATGTACTGGGCTGCAACCTGGAAGTGTGGGGCAGAGAGAGAGGCTCGAGGAATGGAGCATGG
AGGTCATGGAGAGGTCTTGTGGTGCCTTGTGATGTGGTGGCATCACACACCTCACTCGAAA CTACTCGCCAG
CGAGCCTCTCCTTTCCCGCATGCGTGTTGGCTCTCTTCTCTCT **tRNA** GGCA
GCGCCAACCAGCCGCAAAAACAAAGCGAGGCACAGAGGAGGAGTACTCACGACACGAGTAATGCCGAAGC
AGGTCATTT CATGCAGCAAACATGCCCGCAGGAGAAACGCCATGGCGCTGTCCGCCGACTGCAGGTGTGC
CTATGGCGAGCGCCCCCTCCCCCTCACTGAGCGCGTGCGTGCCGGCAGCCTGCGCACCATCGCCACGCGTC
CTTTTTTGTTTTGCAGGCGATTGGGCTCTTTCGTCCTCTCAACC **rRNA** CTCCTGCCT
CGCACACTC **Padrões específicos** CTCTCTTCCCTTCCCTTCCCTTCGACGCCGCGCCACGCGCAC
ACAGGCACAGGGACAGACACACCTATGGACAGCGTGCTCGTGTGCAAACATGTGGCAGACTCTGCGGCAGC
GATATCGTCTTCAGCCACATCCCTCGCAGCCTTCTGGAGTCCAACGCACGGTGTGAAAGTCTCCGACGGC
GGCGCTACACCGAACTCGAGCCTCGTCTCGCCGAAACCAACAGCCTCTGCATCTTCGAAAACCTCGTCGCA
GAGGAATCACTCGCTCGTGGTCTTT **Elementos repetitivos** TGCTCCTCTTCTACGTGAACCGCCCCT
GCGTCGTGGTGGGCCGCAACCAGAACCTCTTCCAGGA

Anotação:

2- Anotação de proteínas

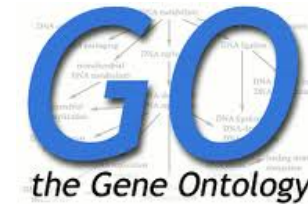


- Proteína nomeadas
- Proteína hipotética conservada
- Proteína hipotética

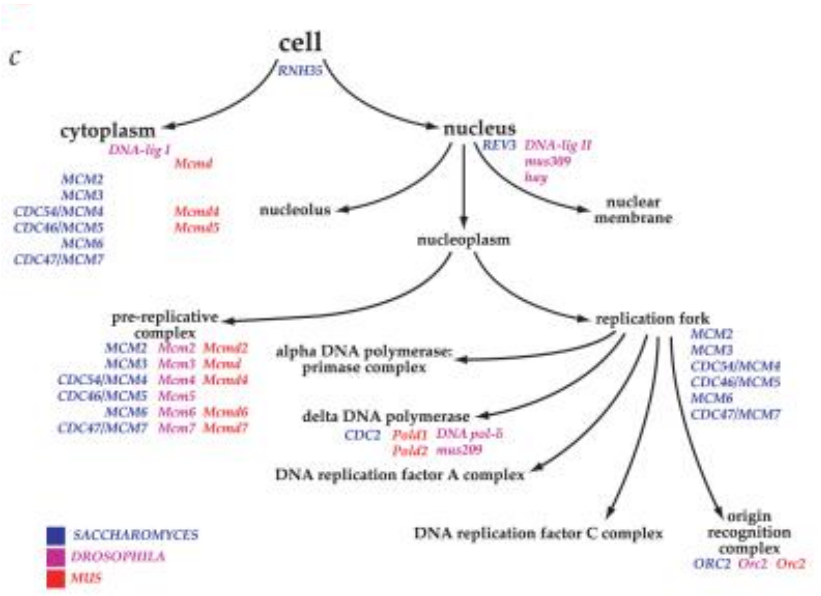
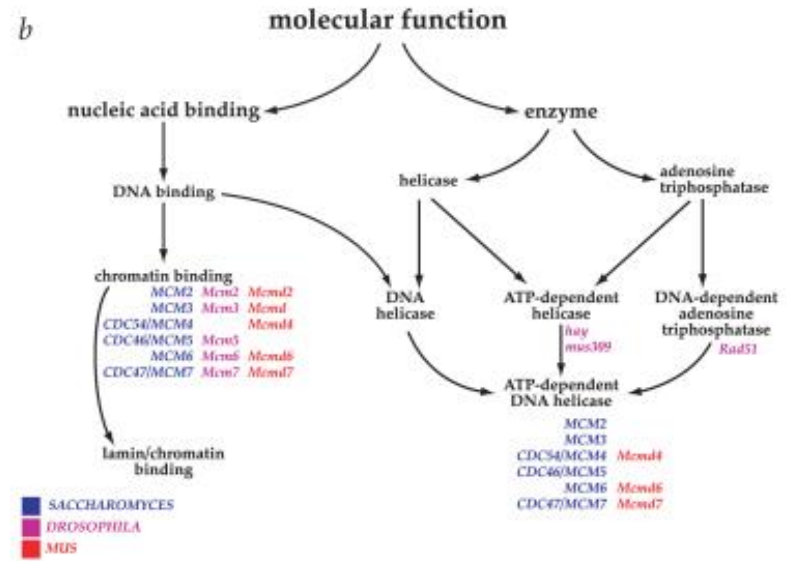
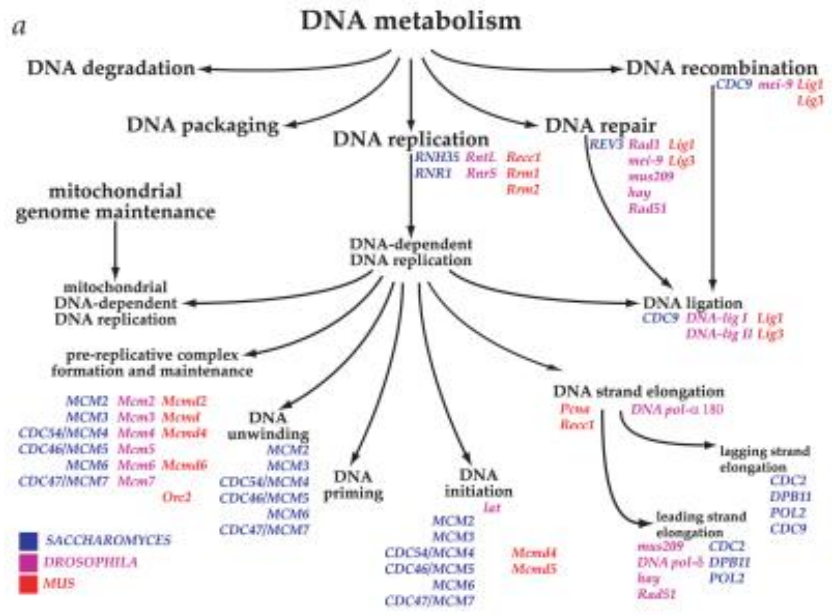
Anotação:

3- Anotação de processos:

Proteínas anotadas



- 1 – Processo biológico
- 2 – Função molecular
- 3 – Componentes celulares



Anotação:

```
graph TD; A[Anotação] --> B[Anotação automática]; B --> C[Anotação manual e curadoria];
```

Anotação automática

Glimmer
GeneScan
RBSfinder
tRNAscan
GeneMark
Repet Masker
BLAST

Anotação manual e curadoria

Sequenciamento de genomas:

Geração das Reads



Montagem do genoma

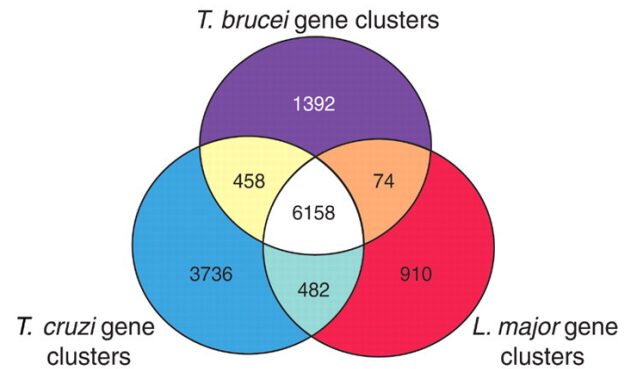


Anotação



Análise: comparação de genomas

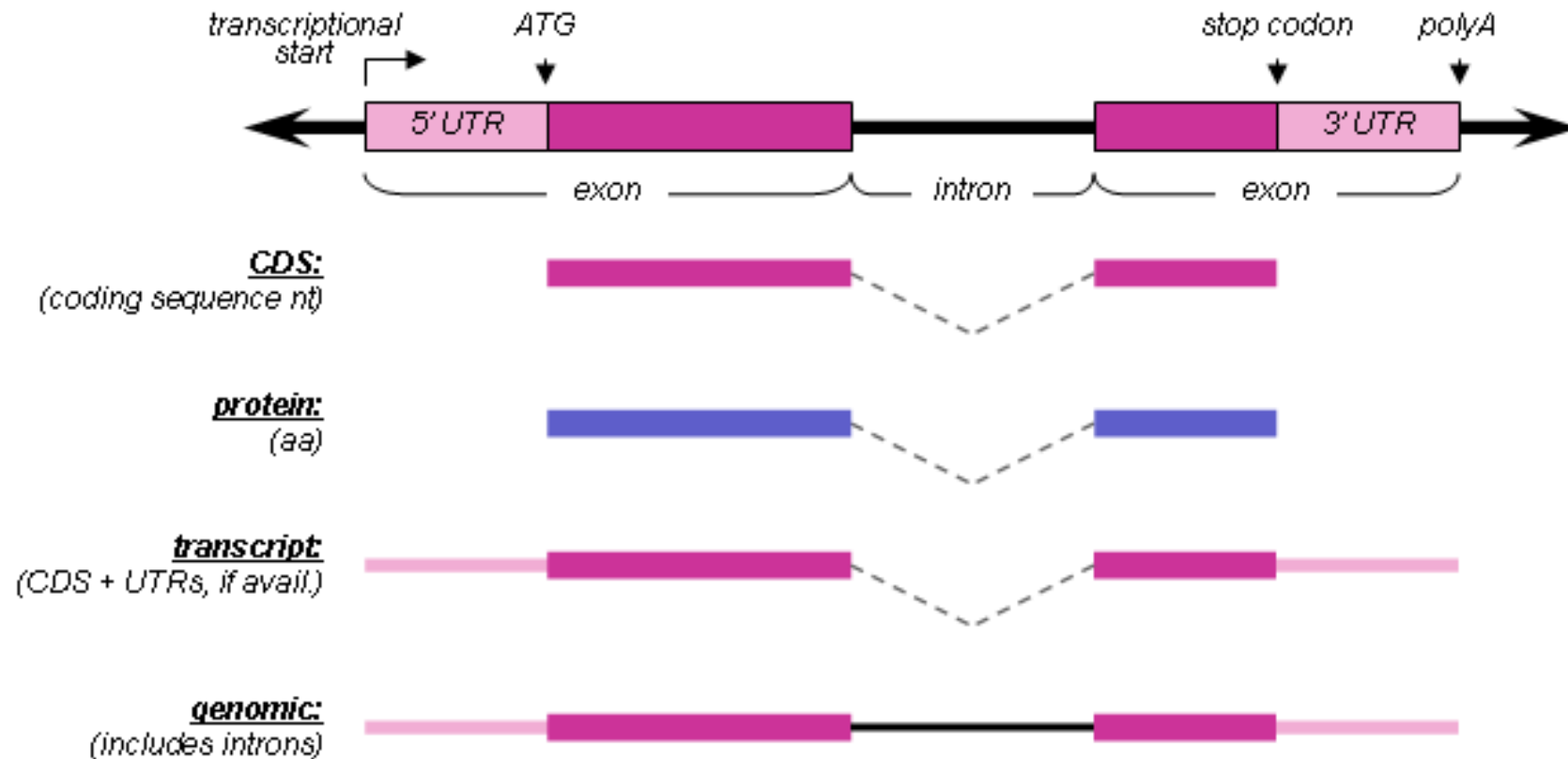
Comparação de genomas:



Tutorial:

- Tutorial: <http://pinguim.fmrp.usp.br/artemis/>
- Anotação manual de proteínas completas;
- Curadoria de um genoma anotado;
- Comparação de três genomas;

Parte I: Anotação manual de proteínas completas



Parte I: Anotação manual de proteínas completas

Artemis Feature Edit: Gene_01

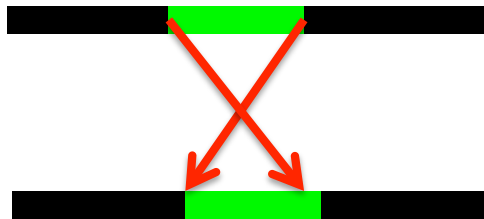
Key:

Location:

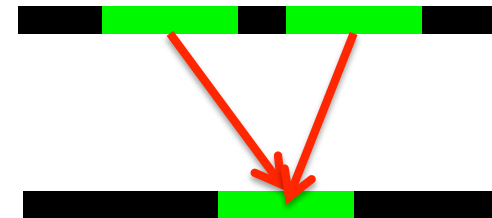
```
/locus_tag="Gene_01"  
/product="putative lypoyltransferase"  
/curation="Tiago Mendes"  
/similarity="Similar to Leishmania braziliensis; lypoyltransferase; 734 aa; 100% id in 705  
aa"
```

Parte III: Comparação de genomas

Inversão



Duplicação/deleção



Alteração da sintenia

