



UNIVERSIDADE DO MINHO

MESTRADO INTEGRADO EM ENGENHARIA INFORMÁTICA

BIOINFORMÁTICA - 4º ANO

LABORATÓRIOS DE BIOINFORMÁTICA

Treponema pallidum subsp. pallidum
str. Nichols

Cláudia Ribeiro - A64460

Duarte Moças - A64319

Jorge Ferreira - A64343

20 de Janeiro de 2016

Conteúdo

Conteúdo	1
1 Introdução	2
2 Desenvolvimento	2
2.1 NCBI	2
2.2 UniProt	3
2.3 GeneOntology	3
3 Conclusão	3

1 Introdução

Treponema pallidum subsp. pallidum Strain Nichols é uma bactéria causadora da doença sexualmente transmissível sífilis, a qual todos os anos provoca novos portadores da mesma. Com o aparecimento e distribuição em larga escala de penicilina, na década de 1940, a transmissão da doença foi bastante mitigada, até à década de 1980-1990. Com o virar do milénio, verificou-se um crescimento das taxas de infeções, em particular nos Estados Unidos da América, Europa, Reino Unido, Canadá e Austrália, sendo que maior parte destes são homossexuais, sendo que a transmissão da doença deve-se, maioritariamente, à pratica de sexo inseguro.

Neste trabalho pretende-se expor as funcionalidades e características dos genes localizados entre a posição 1015201 e 1138011 do genoma do organismo. Para tal, faremos uso das diferentes bases de dados biológicas estudadas, pesquisando por funções, características, localização sub-celular, entre outros atributos.

2 Desenvolvimento

De forma a uniformizar a pesquisa de informação, iniciamos o desenvolvimento do trabalho definindo, primeiramente, os campos (colunas da tabela) referentes a cada gene/-proteína obtido.

Para tal, verificamos quais as informações que cada base de dados biológica oferece, de modo a definir uma estrutura de pesquisa. As bases de dados biológicas usadas foram o NCBI¹, UniProt² e GeneOntology³. A lista seguinte apresenta os campos associados a cada uma das bases de dados:

2.1 NCBI

- Accession
- Start (localização inicial)
- Stop (localização final)
- Strand
- GeneID
- N° EC
- Locus

¹<http://www.ncbi.nlm.nih.gov/>

²<http://www.uniprot.org/>

³<http://geneontology.org/>

- Locus Tag
- Protein Product
- N° de Aminoácidos

2.2 UniProt

- UniProtID
- Revision
- Subcellular Location
- Classification of Protein Function (Função da Proteína)

2.3 GeneOntology

- Protein Function Information (Informação sobre a Função da proteína)
- GeneOntology Terms
- GeneOntology Identifiers

Adicionalmente, existem dois campos, **Description** e **Commentary**, sendo ambas para informações extra ou de síntese.

De modo a realizar o objetivo do trabalho, a obtenção de informação detalhada sobre parte do genome da bactéria em questão, a qual pode ser obtida através do uso de ferramentas como o BioPython⁴, cuja implementação não foi cumprida, foi feita de forma manual.

3 Conclusão

Apesar da conclusão da pesquisa de informação, esta não foi obtida através das ferramentas pretendidas. Com a análise manual, entendemos melhor como deveríamos implementar estas ferramentas, ficando como passo seguinte a implementação das mesmas. Tivemos dificuldade iniciais no contexto do problema devido ao pouco conhecimento que temos sobre a matéria.

⁴http://biopython.org/wiki/Main_Page