

# Análise de Sequências Biológicas

**Ocorrência:** 4º semestre

**Carga horária:** TP 60h; OT 7.5h

**ECTS:** 5,0

**Área disciplinar:** Biotecnologia, Informática

**Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):**

O objetivo desta UC é introduzir os princípios fundamentais da análise de sequências biológicas, partindo do conceito de sequência (DNA, RNA e proteínas), associada à estrutura das biomoléculas, para promover a compreensão das aplicações dos vários algoritmos e para proporcionar a capacidade de escolher os algoritmos mais adequados e interpretar os resultados obtidos. Serão estudados os algoritmos mais frequentes, devendo os alunos ser capazes de os perceber e os transcrever para código, e serão analisadas as várias ferramentas disponíveis de análise e tratamento de sequências de modo a que os alunos as dominem adequadamente.

**Conteúdos programáticos:**

1. Sequências biológicas – DNA, RNA e proteínas. Contribuição da sequência para as estruturas secundária e terciária das biomoléculas.
2. Alinhamento de pares de sequências – matriz de pontos e programação dinâmica; alinhamento com intervalos.
3. Alinhamento de múltiplas sequências - programação dinâmica, métodos progressivos e sequência consensual. Cadeias de Markov, introdução e aplicações. O caso ClustalW.
4. Pesquisa de Motivos Estruturais – Entropia, matrizes de posição e amostragem de Gibbs.
5. Alinhamento de sequência e pesquisa em bases de dados

5.1. Algoritmo BLAST e as várias aplicações

5.2. Filogenia, árvores filogenéticas e cladogramas.

6. Exploração de software de Bioinformática e interação com bases de dados

7. Aplicações – previsão de genes, análise de dados de expressão genética e previsão de estrutura

**Bibliografia principal:**

Baxevanis, A.D., Ouellette, B.F.F., Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Wiley Interscience, ISBN 978-0-471-38391-8

Borodovsky, M., Ekisheva, S., Problems and Solutions in Biological Sequence Analysis, Cambridge University Press, ISBN 978-0-521-61230-2

Mount, D., Bioinformatics: Sequence and Genome Analysis, Cold Spring Harbor Laboratory Press, ISBN 978-0-879-69712-9

Durbin, R., Eddy, S.R., Krogh, A., Mitchison, G., Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press, ISBN 978-0-521-62971-3