

Curso GB-500: Análise de Sequências Biológicas

Professor: Catherine Matias (<http://stat.genopole.cnrs.fr/~cmatias>)

A Dra. Catherine Matias é pesquisadora sênior do CNRS no Laboratoire de Probabilités et Modèles Aléatoires (LPMA), Universités Pierre et Marie Curie et Paris Diderot, em Paris (França) e tem experiência nas áreas de Estatística e Genômica. Dentre as linhas de pesquisa destacam-se: modelos de evolução de sequências e alinhamento, coevolução de sistemas e reconstrução filogenética (sistema parasita-hospedeiro) e métodos estatísticos aplicados a inferência, análise e evolução de redes metabólicas.

Período: 20 a 24 de Outubro de 2014

Local: Laboratório Nacional de Computação Científica (LNCC/MCTI) – Petrópolis – RJ

Público-Alvo: Pós-graduandos, pós-doutorandos ou pesquisadores que atuem na área Bioinformática ou áreas afins.

Pré-Requisitos: Noções elementares de probabilidade e estatística: probabilidade de eventos, probabilidade condicional; estimativa de parâmetros, Método de máxima verossimilhança.

Carga Horária: 36 horas (dia todo, início 9:00)

Número de Vagas: 20

Objetivos:

- Entender os modelos e conceitos probabilísticos e estatísticos que baseiam os métodos clássicos de análise de sequências;
- Ser capaz de utilizar corretamente as ferramentas disponíveis;
- Entender os algoritmos e ser capaz de desenvolver esses algoritmos em casos específicos.

Metodologia de Trabalho:

- **Teoria:** Apresentação e estudo de conceitos teóricos relacionados a modelos probabilísticos.
- **Prática:** Implementação de algoritmos e/ou scripts simples (em R) e utilização de ferramentas disponíveis na internet.

Ementa:

- Modelos probabilísticos para a análise de sequências: cadeias de Markov (tempo discreto), Modelos escondidos de Markov (HMM); Aplicação na detecção de motivos;

- Evolução de sequências: cadeias de Markov (tempo contínuo); Modelos clássicos de evolução: de Jukes-Cantor até general time reversible (gtr) e mais... Estimativa de velocidades de evolução e distâncias evolutivas entre sequências;
- Comparação de sequências: alinhamento de sequências (matriz de pontuação);
- Introdução à filogenia.

Avaliação: O conceito final será baseado no desempenho dos alunos em uma prova escrita e exercícios práticos.

Referências:

1. R. Durbin, S. Eddy, A. Krogh, and G. Mitchison. Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge University Press, Cambridge, UK, 1998.
2. Z. Yang. Computational Molecular Evolution. Oxford Series in Ecology and Evolution. Oxford University Press, 2006.
3. O. Gascuel and M. A. Steel, editors. Reconstructing evolution: new mathematical and computational advances. Oxford university press, Oxford, 2007.
4. J. Felsenstein. Inferring phylogenies. Sinauer Associates, 2004.

Informações importantes:

Data de Inscrição: 15 a 30 de setembro de 2014.

Obs.: Alunos regulares do Curso de Pós-graduação do LNCC deverão efetuar sua matrícula no período de 15 a 19 de setembro de 2014.

Resultado da Seleção: 03 de outubro de 2014.

Idioma: As aulas serão ministradas em inglês e as perguntas poderão ser feitas e esclarecidas em português.

Os participantes selecionados **NÃO** terão auxílio para acomodação e transporte para Petrópolis.

Os interessados poderão enviar CV resumido e carta de motivação para o e-mail ananeri@lncc.br, identificando no Assunto: Curso de análise de sequências biológicas, ou entrar em contato pelo telefone (24) 2233-6024.