

Мини-курс по биоинформатике
ОИПИ НАН Беларуси

Юра Корноушенко, Аня Хадарович, Ваня Кашин, Рома Сергеев, Юра Грушецкий

15 февраля 2017 г.

Биоинформатические сервисы

- **Введение**
- **Обучающие ресурсы**
 - Rosalind
 - Coursera (edX, MIT OpenCourseWare, Stanford Online, Udacity)
- **Базы данных**
 - Protein Data Bank, формат PDB
 - SCOP, CATH, Pfam (Uniprot)
 - GeneOntology, классификация, GO-термины

Биоинформатические алгоритмы

- **Задача: Краткое биологическое введение**
 - цепочка ДНК, аминокислотная последовательность (белок)
 - мутации (замены, удаления/вставки)
- **Формулировка: Выравнивание**
 - простой случай (расстояние Хемминга, расстояние редактирования)
 - сложный случай, целевая функция, матрица замен
 - виды выравнивания
- **Решение: Введение в алгоритмы**
 - оптимизация на ориентированном графе без циклов
 - динамическое программирование
 - блок-схема алгоритма
- **Реализация: Введение в Python**
 - синтаксис языка, операторы
 - необходимые функции, чтение из файла и запись в файл
 - формат fasta
- **Другие задачи**
 - Скрытые Марковские модели (НММ)
 - Нейронные сети (Deep Learning)

Молекулярная эволюция и филогенетический анализ

- **Общие принципы молекулярной эволюции**
 - Теория нейтральной молекулярной эволюции и естественный отбор
 - Концепция молекулярных часов
 - Генетический дрейф
 - Закрепление мутаций в популяции. Эффект бутылочного горлышка (bottleneck effect), эффект основателя (founder effect)
- **Филогенетический анализ.**
 - Филогенетические деревья
 - * Оценка числа деревьев. Дихотомии, политомии.
 - * Корневые, некорневые деревья. Определение положения корня.
 - * Homology, homoplasy.
 - Методы построения филогенетических деревьев:
 - * Метод максимальной экономии (maximum parsimony).
Консенсусное дерево (consensus trees).
 - * Дистанционные методы:
 - Оценивание длин ветвей дерева, используя метод наименьших квадратов.
 - Superimposed substitutions. Эволюционные модели для расчета истинных генетических дистанций между последовательностями. Уточнение моделей, используя Гамма-дистанции.
 - Метод объединения ближайших соседей.
 - Метод максимального правдоподобия:
 - * Вычисление вероятностей, используя матрицы частот аминокислотных замен.
 - * Пример функции правдоподобия для некорневого дерева.
 - Бутстрэп анализ.

Алгоритмы сборки для данных NGS

- Что такое высокопроизводительное секвенирование
- Как собрать геном из набора прочтений, особенности сборки de novo
- В каких задачах используются данные NGS

Структурная биоинформатика

- **Введение**

- Основные задачи структурной биоинформатики и методы их решения
- Биоинформатические ресурсы для работы со структурными данными
- Алгоритмы предсказания структур белков и белковых комплексов

- **Докинг**

- Введение
- Подготовка входных данных в Chimera
- Докинг в программе AutoDock Vina

- **Виртуальный скрининг**

Координаты

*

Где: ОИПИ НАН Беларуси, г. Минск, ул. Сурганова, 6, комн. 100.

Как записаться: отправить письмо на электронный адрес bioinformatics@tut.by, в теме письма указать "мини-курс".

Контактное лицо: Грушецкий Юрий Евгеньевич,
grushetsky@newman.bas-net.by,

(8-029)2526815 (МТС)

(8-029)1322300 (Велком)

(8-025)7445681 (Лайф)