

# Биоинформатика последовательностей

Цель курса – введение в основные концепции анализа генетических последовательностей.

## Программа курса

### **Выравнивания**

1. Задача парного выравнивания. Количество выравниваний. Редакционное расстояние. Алгоритм вычисления редакционного расстояния.
2. Алгоритм выравнивания Миллера-Маерса
3. Локальное выравнивание. Алгоритм Смита-Ватермана
4. Алгоритм выравнивания при общих штрафах за делецию. Оценка времени и памяти.
5. Алгоритм выравнивания для аффинных штрафов за делецию.
6. Статистика выравниваний. Наибольшее общее слово.
7. Статистика выравниваний. Наибольшая общая подпоследовательность.
8. Линейное и логарифмическое поведение веса выравнивания.
9. Матрицы сопоставления аминокислотных остатков. Серия PAM.
10. Матрицы сопоставления аминокислотных остатков. серия BLOSUM.
11. Распределение экстремальных значений. e-value и p-value.
12. Поиск по банку. Хэширование. FASTA, BLAST, BLAST2, быстрое выравнивание.

### **Байесова статистика**

13. Байесова статистика. Априорное и апостериорное распределение. Распределение Дирихле.
14. Байесова статистика. Оценка правдоподобия, матожидания и апостериорная оценка.
15. Байесова статистика. Оценка параметров по результатам наблюдения.

### **Скрытые Марковские модели**

16. НММ. эмиссионные и переходные вероятности. Примеры НММ.
17. Алгоритм Viterbi.
18. Алгоритм Forward-Backward.
19. Оценка параметров НММ при наличии обучающей выборки. Биологические примеры.
20. Оценка параметров при отсутствии обучающей выборки. Алгоритм Баума-Велча.
21. Оценка качества обучения. Селективность и специфичность.
22. Профили. Консенсус, Регулярное выражение, Частотный профиль.
23. Энтропия колонки. НММ профиль. Учет возможности вставок и делеций.
24. Определение параметров НММ - профиля.
25. НММ – формулировка задачи выравнивания. Парные НММ.

### **Коррекции на малые и зависимые выборки**

26. Псевдоотсчеты. Правило Лапласа, Учет фоновых частот, Учет матрицы замен.
27. Взвешивание последовательностей. Метод Герштейна-Сонхаммера-Чотьи.
28. Взвешивание последовательностей. Многогранники Вороного,

### **Множественное выравнивание**

29. Множественное выравнивание. Качество выравнивания. Энтропия, Сумма пар.
30. Множественное выравнивание. Динамическое программирование.

31. Прогрессивное выравнивание. ClustalW. Улучшение выравнивания.
32. Алгоритм множественного выравнивания ProbCons

### ***Локальное множественное выравнивание. Поиск сигналов***

33. Поиск сигналов. Постановка задачи. Распознавание образов.
34. Алгоритм MEME.
35. Гиббс сэмплер.

### ***Вторичные структуры РНК***

36. Вторичная структура РНК. Элементы вторичной структуры.
37. Энергия вторичной структуры РНК.
38. Комбинаторный метод предсказания вторичной структуры.
39. Алгоритм Нуссинофф.
40. Алгоритм Зукера.
41. Субоптимальные структуры.
42. Консенсусные вторичные структуры. Метод ковариаций.
43. Контекстно-свободные грамматики и вторичная структура РНК.