

Государственное образовательное учреждение высшего профессионального образования

«Московский государственный технический университет имени
Н.Э. Баумана»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКИ И СИСТЕМ УПРАВЛЕНИЯ
КАФЕДРА ТЕОРЕТИЧЕСКОЙ ИНФОРМАТИКИ И КОМПЬЮТЕРНЫХ
ТЕХНОЛОГИЙ

Пояснительная записка
к дипломному проекту на тему:

ИССЛЕДОВАНИЕ МЕТОДОВ МНОЖЕСТВЕННОГО
ВЫРАВНИВАНИЯ НУКЛЕОТИДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
В СООТВЕТСТВИИ С РАМКОЙ СЧИТЫВАНИЯ И
СТОП-КОДОНАМИ

Студент–дипломник _____ Батусов П. В.

Научный руководитель _____ Страшнов П. А.

Москва 2015

Аннотация

Содержание

ВВЕДЕНИЕ

3

ВВЕДЕНИЕ

Современная биоинформатика — это молодая, бурно развивающаяся наука, возникшая в 1976-1978 годах и окончательно оформившаяся в 1980 году со специальным выпуском журнала «Nucleic Acid Research» (NAR) [1]. По сути, это собрание различных математических моделей и методов в помощь биологам для решения биологических задач, таких как: предсказание пространственной структуры белков, расшифровка структуры ДНК, хранение, поиск и аннотация биологической информации.

Основу биоинформатики составляют сравнения. Одна из ключевых задач — поиск сходства последовательностей. Пусть имеются две аминокислотные последовательности, и для одной из них известны ее свойства, тогда, если эти последовательности «похожи», можно сделать предположение, что они выполняют сходные функции. Таким образом, новую отсеквенированную последовательность первым делом ищут в базах данных известных (аннотированных) последовательностей, чтобы после сравнения судить о том, какие функции она выполняет.

Для того чтобы определить на сколько две последовательности «похожи», используют алгоритмы выравнивания. Они основаны на размещении исходных последовательностей мономеров ДНК, РНК или белков друг под другом таким образом, чтобы легко увидеть их сходные участки [2]. Качество выравнивания оценивают, назначая штрафы за несовпадение букв и за наличие пробелов (когда приходится раздвигать одну последовательность для того, чтобы получить наибольшее число совпадающих позиций). При сравнении ищется такой вариант выравнивания, чтобы итоговый счет был максимален.

Алгоритмы множественного выравнивания, аналогично алгоритмам парного выравнивания, представляют собой инструмент для установления функциональных, структурных или эволюционных взаимосвязей между биологическими последовательностями. Для этой задачи существует «золотой стандарт» — это выравнивание, которое бы получилось, если выровнять друг под другом последовательности, которые имеют одинаковую пространственную структуру. Это — биологически обоснованное выравнивание. Несмотря на то, что задача множественного выравнивания была сформулирована более 20 лет назад [3], она до сих пор не теряет своей актуальности.

Список литературы

- [1] Миронов Андрей Александрович. Лекция "Введение в биоинформатику". Режим доступа — URL: http://mipt.ru/dbmp/student/files/bioinformatics/public_lecture/.
- [2] Выравнивание последовательностей [электронная публикация]. — URL: https://ru.wikipedia.org/wiki/Выравнивание_последовательностей.
- [3] Humberto Carrillo, David Lipman "The Multiple Sequence Alignment Problem in Biology" on Applied Mathematics Vol. 48, No. 5. (Oct., 1988).