



БИОИНФОРМАТИКА

Авторы: М. С. Гельфанд

БИОИНФОРМАТИКА, науч. направление на стыке математики, биологии, физики полимеров и биохимии, изучающее структуру, функцию и эволюцию биологич. макромолекул (белков, ДНК и РНК). Возникло в кон. 20 в. Важнейшие задачи Б.: исследование последовательности мономеров и пространственной структуры биологич. макромолекул; идентификация генов и регуляторных сигналов в нуклеотидных последовательностях; предсказание структуры и функции белков; выявление и моделирование метаболич. путей и регуляторных взаимодействий в живой клетке. Б. активно занимается также проблемами молекулярной эволюции. Среди используемых математич. подходов – теория алгоритмов, в особенности методы анализа символьных цепей (напр., нуклеотидных и аминокислотных последовательностей), теория графов, вычислительная геометрия, теория дифференциальных уравнений, теория вероятностей, математич. статистика и распознавание образов. В Б. анализируются также данные об экспрессии генов, концентрации белков в клетке, взаимодействия между белками и между белками и нуклеиновыми кислотами, данные о точечных заменах в белках и нуклеиновых кислотах, частотах аллелей в популяциях и др.

В Б. можно выделить статистический, сравнительный и системный подходы. В основе первого лежит предположение о том, что наблюдаемые явления, которые не сводятся к простым вероятностным моделям, имеют функциональное значение. Напр., при распознавании кодирующих белок областей генома используют характерные периодичности в ДНК, а при предсказании пространственной структуры белков – периодичности в аминокислотных последовательностях, характерные для основных типов их вторичной структуры. Сравнит. подход основан на сопоставлении однородных данных из разных организмов и выделении общих элементов. При этом предполагается, что необходимость сохранения структуры и функции таких

элементов, как, напр., каталитич. центры ферментов и пространственная структура РНК, определяет их устойчивость (неизменяемость) в ходе эволюции. Системный подход состоит в привлечении разнородных данных при анализе одного объекта. Так, при предсказании функции белка в клетке используют сравнение его аминокислотной последовательности с последовательностями аминокислот в уже изученных белках, сведения о пространств. структуре белка и его взаимодействиях с др. белками, об экспрессии гена, кодирующего этот белок и т. п. Применение комплекса биоинформатич. методов даёт возможность составить общее представление о физиологии организма путём компьютерного анализа его генома. В свою очередь это позволяет целенаправленно проводить его дальнейшее эксперим. изучение при существенной экономии времени, труда и материальных затрат.

В более широком понимании к Б. относят исследования математич. и компьютерными методами всех информац. процессов и взаимодействий в живых организмах: моделирование нервной системы, зрения, слуха, мышечной активности, метаболич. процессов и информац. взаимодействий в иммунной системе; математич. исследование экологич. систем; исследование сигнальных систем животных и растений; создание компьютерных систем поддержки биологич. экспериментов, баз данных по биологии и медицине, алгоритмов автоматич. реферирования биологич. текстов.

Литература

Лит.: Математические методы для анализа последовательностей ДНК. М., 1999; Гельфанд *М. С.*, Любецкий *В. А.* Биоинформатика: от эксперимента к компьютерному анализу и снова к эксперименту // Вестник РАН. 2003. Т. 73. № 11; Гасфилд *Д.* Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология. СПб., 2003; Koonin *E. V.*, Galperin *M. Y.* Sequence-evolution-function: computational approaches in comparative genomics. Boston, 2003.