

## ЭКСПЕРТНАЯ СИСТЕМА МОДЕЛИРОВАНИЯ ГЕНОМА ЧЕЛОВЕКА

**Шлихт А.Г., Краморенко Н.В.**

Дальневосточный федеральный университет, Россия, 690091, г.Владивосток,  
Океанский пр-т, д.19а, Тел.:(914)974-15-01, E-mail: [schliht@mail.ru](mailto:schliht@mail.ru)

В созданной авторами экспертной системе (ЭС) моделирования генома человека данные представлены не в традиционном горизонтальном текстовом формате, а в вертикальном (реляционном формате) в виде индексированных записей в таблицах баз данных вплоть до нумерации конкретного нуклеотида в гене и аминокислоты в белке [1]. ЭС интегрирована через систему гиперссылок с всемирными порталами (NCBI, EBI, KEGG, OMIM и др.). ЭС функционирует интерактивном и автоматическом режимах. Наряду с интерактивным анализом и моделированием последовательностей генома, транскриптома, протеома ЭС обеспечивает интеллектуальный подход к анализу данных и позволяет в автоматическом режиме осуществлять поиск мотивов по всему геному и протеому человека, находить консервативные области по заданной структуре мотива протеина. Это, в свою очередь, позволяет решать задачи ДНК-протеин, протеин-протеин взаимодействий. В результате вторичного автоматического поиска и анализа консервативных областей получены ассоциации между промоторной частью генов и мотивами протеинов.

ЭС позволяет построить цепочку по найденному (в результате масс-спектрометрического анализа) пептиду к гену, далее к структуре нуклеотидных и аминокислотных последовательностей. По результатам эксперимента и последующего моделирования ЭС выдает сводный отчет по множеству протеинов, генов, содержащий сводную и статистическую информацию: аминокислотный состав ДНК и первичных транскриптов, пронумерованная нуклеотидная последовательность кодированной и некодированной частей экзонов, информация по результатам сплайсинга экзонов (пронумерованная информационная РНК), результаты преобразования в протеин (пронумерованная аминокислотная последовательность, статистика аминокислот в протеине, экспериментальная и рассчитанная масса протеина, рассчитанные заряды), список найденных консервативных областей с точным указанием координат.

Другой задачей, решаемой ЭС, является осуществление автоматического анализа влияния мутаций в геноме на структуру протеома, с последующим определением мутантной аминокислотной последовательности, заряда, массы, возникновения или утраты консервативных областей в протеине. Данные исследования нашли применение при анализе онкологических заболеваний, вирусологических исследованиях.

### **Литература.**

1. Шлихт А.Г., Краморенко Н.В. Автоматизированная информационная система моделирование генома человека. / XXII Международная конференция «Математика. Компьютер. Образование». – Пущино, 26 – 31 января 2015.