(отдельно для последовательностей 1000 п.о. и для последовательностей 15 000 п.о.).

С помощью GLAM2 в последовательностях, расположенных перед точкой старта транскрипции (в области 15000 -7000 п.о.) генов с повышенной экспрессией, найден общий фрагмент, состоящий из большого количества нуклеотидов 'a', «aaaaaaaaaagaaaaaaaaaaaagaaaaaaaaaaaacaa» с небольшими полиморфизмами нуклеотидов 'g' и 'c' перед другими генами. Гомологичный мотив встречался и на участке последовательностях длиной 1000 п.о. При этом он отсутствовал на участках перед генами , которые не увеличивали экспрессию в ответ на окислительный стресс. В последовательностях длиной 1000 после маскирования повторов проявился мотив 'ttttaaaa'. Интересно отметить что этот мотив был найден и в интронах экспрессирующихся генов, в ответ на окислительный стресс.

Данные, полученные в ходе анализа, лягут в основу дальнейших биологических экспериментов, а также будут учтены при разработке программного комплекса для биоинформационного анализа геномов.

Работа выполнена при поддержки грантов ФЦП МОН РФ госконтракт №14.740.11.0006 и №16.552.11.7024.

БИОИНФОРМАЦИОННЫЙ ПОИСК НОВЫХ НАНОСТРУКТУР ДЛЯ РЕГУЛЯЦИИ ГЕНОМА

Пономарева Н.С., Рыжков П.А., Адигеев М.Г., Шкурат Т.П.
Южный федеральный университет, Ростов-на-Дону

Для различных организмов отличается общее количество геномной ДНК, также как, и пропорция кодирующей и некодирующей ДНК в геномах. Более 98% генома человека не кодирует белки, включая большинство последовательностей внутри интронов. В последние годы было показано, что на участках некодирующих белок ДНК у плацентарных животных располагается 2.7 миллиона транскриптомных факторов и около миллиона элементов, которые перекрывают потенциальные промоторы и гены-усилители (Lindblad-Toh, K e.a, 2011), при этом авторы полагают, что только 5% содержимого генома вероятно является выполняют определенные функции , а 3,5 % состоящих из некодирующих элементов, вероятно несут регуляторную роль.

Наше исследование было посвящено биоинформационному поиску новых наноструктур ДНК в некодирующей ДНК, расположенной вокруг паттернов генов значительно изменивших уровень экспрессии в ответ на окислительный стресс, который мы оценивали с помощью ДНК чипов содержащих полногеномный транскриптомный профиль (Shkurat e.a., 2011). Нами была выдвинута гипотеза, что все гены, повысившие экспрессию в ответ на окислительный стресс могут иметь одинаковые мотивы (наноструктуры ДНК) в некодирующей ДНК. Под мотивом распознавания здесь и далее понимается способ описания набора сходных олигонуклеотидов, специфически распознаваемых и связываемых конкретным регуляторным белков. Для поиска мотивов нами была создана интегрированная коллекция баз данных транскрипционных сайтов (ТС) связывания - JASPAR, TRANSFAC, Hocomoco ТФ Homo sapiens, Uniprobe ТФ Mus musculus. В качестве базовой модели мотива распознавания сайтов связывания транскрипционных факторов (ССТФ) в работе использовали классический подход позиционно-весовых матриц (МПВ) (Stormo G.D., 2000). Каждому олигонуклеотиду, «мотиву» заданной длины, МПВ конкретного ТФ сопоставляет численную оценку,