

Оптимизация нуклеотидного состава генов HLA и β_2 -микроглобулина с помощью оригинального компьютерного алгоритма

УО «Белорусский государственный медицинский университет»,
г. Минск, Беларусь

Для биологического лекарственного средства на основе вирусного вектора важным требованием является стабильность переносимых генов во время экспрессии *in vivo*. На основании информации, полученной из расчётов нуклеотидного состава и характера мутационного давления Fowlpox вируса, применён способ защиты от несинонимичных мутаций экспрессируемых в составе вектора генов.

Цель: создать и описать компьютерный алгоритм для оптимизации нуклеотидного состава генов HLA (англ. human leukocyte antigens) и β_2 -микроглобулина для экспрессии в векторе на основе Fowlpox вируса.

Материалы и методы исследования. Написание алгоритма осуществлялось в программе Microsoft Office Excel. Для модификации нуклеотидных последовательностей были выбраны HLA-аллели A*01:01:01:01, A*24:02:01:01, B*08:01:01:01, B*44:03:01:01, C*04:01:01:01, C*12:02:01 (доступны в базе IPD-IMGT/HLA). Для отбора наименее часто встречающихся кодонов у человека был выбран верхний предел частоты со значением $10 \cdot 10^{-4}$. Это кодоны TTA ($7,7 \cdot 10^{-4}$), СТА ($7,2 \cdot 10^{-4}$), АТА ($7,5 \cdot 10^{-4}$), GTA ($7,1 \cdot 10^{-4}$), CGT ($4,5 \cdot 10^{-4}$), CGA ($6,2 \cdot 10^{-4}$), TCG ($4,4 \cdot 10^{-4}$), CCG ($6,9 \cdot 10^{-4}$), ACG ($6,1 \cdot 10^{-4}$), GCG ($7,4 \cdot 10^{-4}$). Подсчёт показателей нуклеотидного состава кодирующих

последовательностей генов HLA и β_2 -микроглобулина осуществлялся с использованием алгоритма VVK Protective buffer.

Результаты. Работа алгоритма включает следующие этапы:

1. замена Т на С и А на G в третьих положениях кодонов, если замены синонимичны;
2. замена G на С в третьих положениях редких для человека кодонов, А на С в кодоне изолейцина АТА;
3. замена всех кодонов квартета аргинина (CGX) на кодон AGG с целью избавления от высокомутабельного мотива CpG;
4. обнаружение CpG на стыке кодонов и замена С на G, если она синонимична и не ведёт к образованию редкого кодона;
5. замена С на G в мотивах "CCCC" в третьем положении кодона, если она синонимична и не приводит к образованию редкого кодона. В противном случае — замена С на Т;
6. замена G на А в мотивах "GGGG".

После обработки нуклеотидных последовательностей алгоритмом значения показателей, отражающих GC-насыщенность (G+C, 3GC, GC4f, GC2f3p), а также значения 3C, G2f3p, C2f3p, C4f увеличиваются во всех последовательностях. Значения показателей A2f3p и T2f3p, A4f, T4f, 3A, 3T снижаются. Значение показателя 3G после обработки увеличивается в кодирующей последовательности β_2 -микроглобулина и несколько снижается в остальных по той причине, что у β_2 -микроглобулина 3GC изначально составляет порядка 50%, а у генов HLA — в среднем около 78%. Так как для генов HLA, в отличие от β_2 -микроглобулина, характерен высокий уровень G4f, после описанных выше замен общий уровень G4f и вероятность синонимичной замены G на Т несколько снижается. Значение общей частоты использования кодонов Arg4 после обработки в алгоритме становится равно нулю во всех кодирующих последовательностях. В то же время алгоритм может дополнительно вносить в их состав от одного до трёх G-квардуплексов.

Выводы. Обработка кодирующих последовательностей генов HLA и β_2 -микроглобулина оригинальным алгоритмом исключает возможность использования редких кодонов, снижает частоту встречаемости CpG-мотивов, разрывает склонные к возникновению мутаций мононуклеотидные фрагменты типа "GGGG" и "CCCC". Наряду с этим алгоритм повышает значение GC-состава генов с изначально низким G+C, учитывая мутационное AT-давление, характерное для Fowlpox вируса. Такая оптимизация нуклеотидного состава обеспечит более эффективную работу генов в составе вирусного вектора *in vivo*.