

А. А. Акуневич

ОПТИМИЗАЦИЯ БЕЛОК-КОДИРУЮЩИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С УЧЁТОМ НУКЛЕОТИДНОГО СОСТАВА FOWLPOX ВИРУСА

Научный руководитель: канд. биол. наук, доц. В. В. Хрусталёв

Кафедра общей химии,

Белорусский государственный медицинский университет, г. Минск

A. A. Akunevich

OPTIMIZATION OF PROTEIN-CODING SEQUENCES IN ACCORDING TO THE FOWLPOX VIRUS NUCLEOTIDE CONTENT

Tutor: PhD, associate professor V. V. Khrustalev

Department of General Chemistry,

Belarusian State Medical University, Minsk

Резюме. Проанализирован нуклеотидный состав штаммов Fowlpox вируса и выявлен характер мутационного давления. С учётом этого создан компьютерный алгоритм для оптимизации белок-кодирующих последовательностей, встраиваемых в Fowlpox вирусный вектор.

Ключевые слова: Fowlpox вирус, нуклеотидный состав, вектор.

Resume. We analyzed the nucleotide content of the Fowlpox virus strains and revealed the mutational pressure type. Based on this information, we created a computer algorithm to optimize the protein-coding sequences inserted into the Fowlpox viral vector.

Keywords: Fowlpox virus, nucleotide content, vector.

Актуальность. Fowlpox вирусы (FPV) рассматриваются в качестве векторов в терапии опухолевых заболеваний [1]. Мутационная стабильность генов, переносимых в составе подобного вектора, является одной из важнейших характеристик биологического лекарственного средства. Изучение нуклеотидного состава FPV позволяет разработать подходы к эффективной терапии, основанной на FPV-векторах.

Цель: проанализировать нуклеотидный состав штаммов FPV и создать компьютерный алгоритм для оптимизации белок-кодирующих последовательностей, встраиваемых в FPV-вектор.

Материал и методы. Анализировались полногеномные последовательности FPV девяти штаммов (GenBank ID NC_002188.1, MH734528.1, MH719203.1, MH709125.1, MH709124.1, MF766432.1, MF766431.1, MF766430.1, AJ581527.1). Подсчёт показателей нуклеотидного состава осуществлялся с использованием алгоритма VVK Protective buffer (<https://chemres.bsmu.by>). Статистическая обработка и написание алгоритма для оптимизации нуклеотидных последовательностей проводились в программе Microsoft Office Excel. Достоверность различий устанавливалась с помощью t-теста для относительных величин. Оптимизация нуклеотидной последовательности осуществлялась на примере HLA-аллелей 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 (IPD-IMGT/HLA). Отбор наименее часто встречающихся у человека кодонов осуществлялся по верхнему пределу частоты со значением $10 \cdot 10^{-4}$: TTA ($7,7 \cdot 10^{-4}$), CTA ($7,2 \cdot 10^{-4}$), ATA ($7,5 \cdot 10^{-4}$), 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 ($7,4 \cdot 10^{-4}$).

Результаты и их обсуждение. Распределение плотности генов FPV составляет $43,76 \pm 0,51\%$ и $64,89 \pm 0,69\%$ на 5'- и 3'-конце «плюс» цепи соответственно, $56,24 \pm 0,51\%$ и $35,11 \pm 0,69\%$ на 3'- и 5'-конце комплементарной цепи соответственно. Подобное распределение отражает ассиметричный сдвиг в плотности генов: она является высокой на 3'-концах обеих цепей.

Анализ нуклеотидного состава кодирующих участков штамма FPV SD15-670.2 показал, что частоты использования аденина и тимина в четырёхкратно вырожденных сайтах ($35,47 \pm 0,87\%$ и $40,18 \pm 0,93\%$ соответственно) достоверно выше ($P < 0,0001$, для 257 генов), чем гуанина и цитозина ($13,05 \pm 0,66\%$ и $11,29 \pm 0,63\%$, соответственно). Частоты использования нуклеотидов в двукратно вырожденных сайтах ($A2f3p = 34,26 \pm 0,69\%$; $T2f3p = 39,47 \pm 0,72\%$; $G2f3p = 10,19 \pm 0,43\%$; $C2f3p = 16,08 \pm 0,59\%$) повторяют характерное для четырёхкратно вырожденных сайтов распределение, что является общим для всех изученных штаммов вируса и свидетельствует о наличии в FPV мутационного АТ-давления.

В ранних генах количество вариабельных позиций, содержащих цитозин или тимин в двукратно вырожденных сайтах, достоверно выше, чем количество аналогичных позиций с гуанином или аденином ($66,67\%$ к $33,33\%$, $P = 0,0152$). В поздних генах достоверных отличий обнаружено не было ($56,25\%$ к $43,75\%$, $P = 0,4870$).

В результате анализа выявлено несколько генов, характеризующихся наибольшим количеством мутаций (таблица 1). Максимальное их число обнаруживается в гене белка сборки вириона, а также в копиях гена, продуктом которого является B22R — гликопротеин поверхности вириона. Теоретически, из вирусного вектора можно удалять перечисленные участки для повышения его стабильности, однако следует оценить, каким образом это повлияет на его сборку.

Табл. 1. Гены FPV, имеющие наибольшее количество мутаций

Ген	Длина (п.о.)	Продукт	Цепь	Число мутаций	Вид гена
FPV080	1092	TGF-β	комплементарная	10	—
FPV030	2454	щелочная фосфодиэстераза	комплементарная	16	—
FPV097	5739	variola B22R	комплементарная	19	ранний
FPV098	5409	variola B22R	комплементарная	21	ранний
FPV107	5334	variola B22R	комплементарная	14	ранний
FPV122	5613	variola B22R	комплементарная	20	ранний
FPV123	5301	variola B22R	комплементарная	18	ранний
FPV197	906	белок сборки вириона	комплементарная	21	поздний

Помимо этого, в геноме FPV присутствуют последовательности, продукты которых могут повлиять на эффективную экспрессию трансгенов и работу вектора в целом. К ним относятся ответственные за иммунное уклонение пептиды (C-type lectin family, IL) и поксвирусные факторы роста (TGF-β, β-NGF, EGF-like protein), стимулирующие пролиферацию опухолевых клеток.

Алгоритм для оптимизации нуклеотидных последовательностей переносимых в составе FPV генов вносит следующие синонимичные замены при условии, что это не ведёт к образованию редких кодонов: (1) Т на С и А на G в третьих положениях кодонов, (2) G на С в третьих положениях редких для человека кодонов, (3) А на С

в кодоне изолейцина АТА, (4) G на A в мотивах «GGGG» и (5) C на G в мотивах «CCCC» в третьем положении кодонов, (6) C на G в CpG, расположенных на стыке кодонов, а также (7) заменяет все кодоны квартета аргинина (CGX) на кодон AGG.

После обработки нуклеотидных последовательностей в алгоритме значения показателей, отражающих насыщенность гуанином и цитозином, увеличиваются (выделены серым цветом в таблице 2). Значение общей частоты использования кодонов квартета аргинина после обработки в алгоритме становится равным нулю. Помимо этого, как правило, алгоритм уменьшает вероятность несинонимичных и нонсенс-мутаций либо не изменяет эти значения.

Табл. 2. Показатели, характеризующие нуклеотидный состав кодирующих участков β_2 -микроглобулина и HLA-A*01:01:01:01

Параметр	Значения для β_2 -микроглобулина		Значения для HLA-A*01:01:01:01	
	до обработки в алгоритме	после обработки в алгоритме	до обработки в алгоритме	после обработки в алгоритме
G+C	0,456	0,592	0,633	0,669
3GC	0,508	0,958	0,787	0,940
3G	0,267	0,417	0,402	0,399
3C	0,242	0,542	0,385	0,541
3A	0,192	0,008	0,090	0,041
3T	0,3	0,033	0,123	0,019
G4f	0,204	0,286	0,359	0,263
C4f	0,222	0,653	0,390	0,659
A4f	0,222	0,020	0,108	0,056
T4f	0,352	0,041	0,144	0,022
G2f3p	0,298	0,516	0,414	0,513
C2f3p	0,281	0,452	0,4	0,442
A2f3p	0,175	0	0,071	0,026
T2f3p	0,245	0,032	0,114	0,019
Arg4	41,667	0	43,716	0
QGRS	0	3	13	16

Заключение. Для FPV характерно симметричное мутационное АТ-давление, снижающее стабильность транскрибируемых в его составе генов с низкой GC-насыщенностью. С учётом этого предложен метод оптимизации нуклеотидных последовательностей для повышения их мутационной устойчивости в составе FPV-вектора. Алгоритм повышает их GC-состав, избегая при этом образования горячей точки для мутаций (мотива метилирования цитозина CpG) и использования редких кодонов.

Информация о внедрении результатов исследования. По результатам настоящего исследования опубликовано 2 статьи в сборниках материалов, 2 тезиса докладов, получено 2 акта внедрения в образовательный процесс кафедры фармацевтической технологии и кафедры общей химии Белорусского государственного медицинского университета.

Литература

1. Phase I trial of a modified Vaccinia Ankara priming vaccine followed by a Fowlpox virus boosting vaccine modified to express brachyury and costimulatory molecules in advanced solid tumors / J. M. Collins, R. N. Donahue, Y.T. Tsai [et al.] // The Oncologist. – 2019. – № 26. – P. 1-6.