Стохастическая модель сплайсинга CD44 с переносом и связыванием ESRP1 и ESRP2

А. А. Пирогов*

Построена дискретная стохастическая модель выбора стыков вариабельной области гена CD44. Модель описывает перенос белков ESRP1 и ESRP2 между ядром и цитоплазмой, их связывание с пре-мРНК и последовательное образование стыков с экзонами v8, v9 и v10. Траектории моделируются алгоритмом Гиллеспи. Параметры подбирались по долям стыков в клетках H358 в контроле и после совместного подавления ESRP1/2; независимой проверки на других данных пока не проводилось. В результате подбора получен набор параметров, согласующийся с данными, при котором доля ESRP1 в ядре после подавления возрастает, что может частично компенсировать снижение его общего количества.

Ключевые слова: альтернативный сплайсинг, CD44, РНК-связывающие белки, стохастическая модель, алгоритм Гиллеспи.

1. Введение

Альтернативный сплайсинг гена CD44 формирует изоформы рецептора с разным набором экзонов; изменения их соотношения описаны при опухолях [1]. Вариант CD44v8–v10 связывают с защитой клеток от окислительного стресса [2]. Включение вариабельных экзонов регулируют ESRP1 и ESRP2 [3]; их действие может зависеть от содержания в ядре. Мы моделируем транспорт и связывание этих белков с пре-мРНК и подбираем параметры по данным H358 в контроле и после подавления ESRP1/2.

2. Модель

Модель рассматривает последовательный выбор стыков конститутивного экзона с v8, затем v8–v9 и v9–v10. В каждый момент обрабатывается одна пре-мРНК. Индекс $g = 1, 2$ нумерует белки ESRP1 и ESRP2, через N_g и C_g обозначены числа их свободных молекул в ядре и цитоплазме. На каждом этапе пре-мРНК находится либо в свободном состоянии ($b = 0$), либо в комплексе с одним из белков ($b = g$).

* *Пирогов А. А.* — НИУ ВШЭ; МГУ имени М.В. Ломоносова, e-mail: aleksei.pirogov@math.msu.ru, ORCID: 0009-0000-5085-4045.

Pirogov A. A. — HSE University; Lomonosov Moscow State University.

Опубликовано на условиях лицензии CC-BY 4.0
Published under CC-BY 4.0

Перенос белков и обратимое связывание задаются переходами

$$\begin{aligned} C_g \to N_g \quad & (k_{\text{in},g} C_g), \quad N_g \to C_g \qquad (k_{\text{out},g} N_g), \\ N_g + U \to B_g \quad & (k_{\text{on},g} N_g), \quad B_g \to N_g + U \qquad (k_{\text{off},g}), \end{aligned}$$

где в скобках указаны интенсивности; U и B_g обозначают свободную и связанную РНК. Переход возможен при наличии исходного состояния. Этап завершается с интенсивностью λ с высвобождением белка. Вероятности выбора стыка для свободной РНК и РНК, связанной с регулятором, равны p_{0j} и p_{bj}, $j = 1, 2, 3$.Вероятность p_{bj} одинакова для ESRP1 и ESRP2. Обе вероятности предполагаются одинаковыми в контроле и после подавления, в то время как при другом исходе обработка прекращается.

Траектории строятся алгоритмом Гиллеспи [4]: время ожидания распределено экспоненциально с параметром $a_0 = \sum_r a_r$, переход r выбирается с вероятностью a_r/a_0. Изначально 80% молекул каждого белка находится в ядре, РНК свободна. Между обработкой молекул распределение белков сохраняется; период установления не задаётся.

3. Данные и подбор параметров

Мы использовали опубликованные данные RNA-seq клеток H358 из работы [3], включающие по три образца в контроле и после совместного подавления ESRP1/2. Для каждого образца вычислены доля первого стыка с v8 среди десяти выбранных стыков того же конститутивного экзона, а также доли стыков v8–v9 и v9–v10 среди доступных стыков, начинающихся соответственно от v8 и v9. Доли стыков приближают доли этапов модели; принадлежность всех стыков одной молекуле не установлена.

В каждом расчёте общее число молекул белка (пула) $R_g = N_g + C_g + \mathbf{1}_{\{b=g\}}$ постоянно: $(R_1, R_2) = (174, 43)$ в контроле и $(28, 24)$ после подавления; его изменение задано по уровням мРНК ESRP1/2. Для обоих белков $k_{\text{on},g} = 0{,}02$, $k_{\text{off},g} = \lambda = 1$; время и числа молекул выражены в условных единицах. Подбирались шесть вероятностей p_{0j}, p_{bj} и восемь констант транспорта для двух условий. Из 6500 случайных наборов 120 лучших пересчитаны по четырём траекториям из 6000 обработанных молекул мРНК CD44 каждая. Контрольные константы импорта и экспорта выбирались логарифмически равномерно из диапазонов $[0{,}008; 0{,}2]$ и $[0{,}002; 0{,}05]$. После подавления каждая константа составляла от 0,25 до 4 её контрольного значения.

4. Результаты

Для полученного набора модельные и наблюдаемые средние частоты стыков составили соответственно: первый стык с v8 — 46,6% и 47,0% в контроле, 37,1% и 35,7% после подавления; v8–v9 — 98,8% и 99,0%, 98,7% и 99,5%; v9–v10 — 97,2% и 97,0%, 97,2% и 96,1%. Подбор проведён по обеим группам, поэтому эти значения характеризуют согласование с использованными данными, а не независимый прогноз.

Для выбранного набора вероятности выбора стыков свободной РНК составили $(p_{01}, p_{02}, p_{03}) = (0{,}2759; 0{,}9895; 0{,}9641)$, а связанной РНК — $(p_{b1}, p_{b2}, p_{b3}) = (0{,}5682; 0{,}9855; 0{,}9792)$.

В таблице 1 приведены транспортные константы для этого набора и доли свободных регуляторов в ядре относительно их общего пула.

Параметр	Контроль	Подавление
$k_{\mathrm{in},1}$ (ESRP1)	0,0371	0,0448
$k_{\mathrm{out},1}$ (ESRP1)	0,0085	0,0042
$k_{\mathrm{in},2}$ (ESRP2)	0,0382	0,0632
$k_{\mathrm{out},2}$ (ESRP2)	0,0037	0,0050
Свободный ESRP1 в ядре, % от общего числа белка	81	91
Свободный ESRP2 в ядре, % от общего числа белка	91	92

Таблица 1. Параметры транспорта и доли свободных белков в ядре.

В этом сценарии после подавления константа импорта ESRP1 возрастает, а константа экспорта уменьшается; доля свободного белка в ядре увеличивается с 81% до 91%. Можно предположить, что это частично компенсирует уменьшение общего пула ESRP1 и поддерживает его доступность для связывания с пре-мРНК.

Список литературы

[1] T. Xu, M. Verhagen, R. Joosten et al., “Alternative splicing downstream of EMT enhances phenotypic plasticity and malignant behavior in colon cancer”, *eLife*, **11** (2022), e82006. DOI: 10.7554/eLife.82006.

[2] M. Yanova, E. Stepanova, D. Maltseva, A. Tonevitsky, “CD44 variant exons induce chemoresistance by modulating cell death pathways”, *Frontiers in Cell and Developmental Biology*, **13** (2025), 1508577. DOI: 10.3389/fcell.2025.1508577.

[3] Y. Yang, J. W. Park, T. W. Bebee et al., “Determination of a comprehensive alternative splicing regulatory network and combinatorial regulation by key

factors during the epithelial-to-mesenchymal transition", *Molecular and Cellular Biology*, **36**:11 (2016), 1704–1719. DOI: 10.1128/MCB.00019-16.

[4] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions", *The Journal of Physical Chemistry*, **81**:25 (1977), 2340–2361. DOI: 10.1021/j100540a008 .

Статья поступила 7 ноября 2026 г.

A stochastic model of CD44 splicing with transport and binding of ESRP1 and ESRP2

A. A. Pirogov

We present a discrete stochastic model of alternative splicing of the CD44 variable region. The model describes the transport of ESRP1 and ESRP2 proteins between the nucleus and cytoplasm, their binding to pre-mRNA, and the sequential formation of splice junctions with exons v8, v9 and v10. Trajectories are simulated with the Gillespie algorithm. Parameters were fitted to junction fractions in H358 cells under control conditions and after combined ESRP1/2 knockdown; no independent validation has been performed. Fitting yielded a parameter set consistent with the data, in which the nuclear fraction of ESRP1 increases after knockdown, which may partially offset the decrease in its total amount.

Keywords: alternative splicing, CD44, RNA-binding proteins, stochastic model, Gillespie algorithm.

References

[1] T. Xu, M. Verhagen, R. Joosten et al., "Alternative splicing downstream of EMT enhances phenotypic plasticity and malignant behavior in colon cancer", *eLife*, **11** (2022), e82006. DOI: 10.7554/eLife.82006.

[2] M. Yanova, E. Stepanova, D. Maltseva, A. Tonevitsky, "CD44 variant exons induce chemoresistance by modulating cell death pathways", *Frontiers in Cell and Developmental Biology*, **13** (2025), 1508577. DOI: 10.3389/fcell.2025.1508577.

[3] Y. Yang, J. W. Park, T. W. Bebee et al., "Determination of a comprehensive alternative splicing regulatory network and combinatorial regulation by key factors during the epithelial-to-mesenchymal transition", *Molecular and Cellular Biology*, **36**:11 (2016), 1704–1719. DOI: 10.1128/MCB.00019-16.

[4] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions", *The Journal of Physical Chemistry*, **81**:25 (1977), 2340–2361. DOI: 10.1021/j100540a008.

Received on November 7, 2026