

РАЗРАБОТКА И АНАЛИЗ МЕТОДОВ ГЕНЕРАЦИИ CELL-PENETRATING PEPTIDES (CPP) И ПОИСК НОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ В ОТКРЫТЫХ ГЕНОМНЫХ ДАННЫХ

Петровский О. Д.¹, Кучеренко И. А.¹
Научный руководитель – аспирант ФИТиП, Сердюков А. Н.¹
¹Университет ИТМО
fallenwaker@gmail.com

Введение

Борьба с раком является одной из наиболее актуальных тем современной науки. Для лечения опухолевых заболеваний с наибольшим процентом ремиссии применяется комбинированный подход таргетных методов (антитела, CAR-T) и химиотерапии, обладающей системным токсическими побочными эффектами, без которой эффективность лечения заметно снижается [1]. Одним из потенциальных методов решения проблемы токсичности является исследование белков, проникающих через клеточную мембрану (CPP – Cell-Penetrating Peptides) для таргетной доставки терапевтических молекул в клетки опухоли [2]. На данный момент ведутся исследования и даже клинические испытания препаратов на основе конъюгатов таких пептидов и терапевтических молекул или наночастиц для транспорта их внутрь клеток [3], что позволяет доставить в клетку крупные соединения, не проходящие через мембрану.

Одной из ключевых проблем в дизайне таких лекарств является низкая специфичность получающихся белков – они могут проникнуть в любые клетки, а не только в целевые, хотя есть примеры CPP, специфичных к поверхностным белкам меланомы. В интернете есть базы данных CPP, например CPPsite 3.0 [4].

Главная цель проекта – разработка модели для предсказания новых CPP и поиск их последовательностей в геномах. На данном этапе работа посвящена анализу общедоступных источников данных, а также разработке плана дальнейшего исследования.

Основная часть

Для достижения указанной цели были выделены следующие задачи:

1. Анализ существующих публикаций о CPP;
2. Поиск открытых баз данных таких белков;
3. Разработка генеративной модели CPP и последующий анализ их характеристик, метрик качества и поиск их последовательностей в геномах;
4. Разработка классификационной модели и выделение дескрипторов, отвечающих за определение принадлежности к классу CPP белков, поданных на вход данной модели в виде нуклеотидной последовательности.

Активные исследования CPP ведутся уже более 20 лет, однако для анализа литературы было принято решение выбрать публикации не старше 2010 года, ввиду того, что более ранние публикации использовали другие методы и протоколы исследований, несовместимые с более современными.

Главной отличительной особенностью разрабатываемого метода является поиск предсказанных последовательностей в геномных базах данных или геномов на наличие последовательностей, ассоциированных с CPP. В то время как большинство инструментов фокусируются на генерации “идеального” CPP, разрабатываемое в рамках данной работы решение дает возможность найти такие последовательности в геномах.

В качестве основных публично-доступных источников данных были взяты базы данных семейства CPPsite (версии 1.0-3.0). В качестве дополнительных источников были

взяты открытые наборы данных для инструмента CellPPD-MOD [5], однако важным их отличием является отсутствие метаданных, кроме принадлежности классу CPP.

На первом этапе главным инструментом для сравнения с разрабатываемыми моделями был выбран CPPpred-En [6]. Также планируется сравнение с его модификациями.

Выводы

На данном этапе проведен анализ публикаций и поиск доступных баз данных, найден существующий инструментарий для предсказаний характеристик клеточно-проникающих белков. Начато составление обобщенной базы данных на основе открытых источников, а также подготовка инструментария для анализа геномов. Составлен план дальнейшего исследования, планируется воспроизвести уровень производительности инструмента CPPpred-En и с его помощью найти в геномах новые последовательности CPP белков.

Литература

1. Gautam M., Kumar C.D., Ramachandran P., Kumar A., Gopalsamy I.K., Swaminathan G., Cordani M., Dinesh Kumar L. Shifting tides from non-targeted chemotherapy to targeted therapies and small RNA-based bio drugs: A changing paradigm in cancer therapy // *European Journal of Pharmacology*. — 2026. — Vol. 1011. — 178446. — DOI: <https://doi.org/10.1016/j.ejphar.2025.178446>.
2. Gessner I., Neundorf I. Nanoparticles Modified with Cell-Penetrating Peptides: Conjugation Mechanisms, Physicochemical Properties, and Application in Cancer Diagnosis and Therapy // *International Journal of Molecular Sciences*. — 2020. — Vol. 21, No. 7. — 2536. — DOI: <https://doi.org/10.3390/ijms21072536>.
3. Jang Y.E., Kwon M., Kim S.G., George N.P., Hwang J.S., Basith S., Lee G. CPPpred-En: Ensemble framework integrating a protein language model and conventional features for highly accurate cell-penetrating peptide prediction // *Computers in Biology and Medicine*. — 2025. — Vol. 195. — 110590. — DOI: <https://doi.org/10.1016/j.compbimed.2025.110590>.
4. Moreno-Vargas L.M., Prada-Gracia D. Exploring the Chemical Features and Biomedical Relevance of Cell-Penetrating Peptides // *International Journal of Molecular Sciences*. — 2024. — Vol. 26, No. 1. — 59. — DOI: <https://doi.org/10.3390/ijms26010059>.
5. CellPPD-MOD [Электронный ресурс]. — URL: <https://webs.iiitd.edu.in/raghava/cellppdmod/index.html> (дата обращения: 15.02.2026).
6. CPPpred-En [Электронный ресурс]. — URL: <https://github.com/YE-MJ/ CPPpred-En> (дата обращения: 17.02.2026).

Автор _____ Петровский О. Д.

Автор _____ Кучеренко И. А.

Научный руководитель _____ Сердюков А.Н.