

**ПРОГНОЗИРОВАНИЕ ЭФФЕКТИВНОСТИ МИРНК ДЛЯ НОКДАУНА ГЕНОВ****Головкин И.А. (ИТМО), Шатковский Д.П. (ИТМО)****Научный руководитель – кандидат химический наук, Серов Н.С. (ИТМО)**

**Введение.** Использование миРНК является перспективным методом контроля экспрессии генов [1]. Однако эффективность нокдауна зависит от множества факторов, таких как нуклеотидный состав, химические модификации [2] и экспериментальные параметры. Исследования демонстрируют [3], что выбор оптимальных параметров молекул, таких как GC-состав, длина участка и наличие мисматчей, играет существенную роль в достижении высокого уровня подавления целевого гена. Проблема оптимизации параметров миРНК требует комплексного подхода, объединяющего экспериментальные данные и вычислительное моделирование.

**Основная часть.** Предлагаемое решение основывается на комплексном анализе экспериментальных данных и разработке моделей машинного обучения для предсказания эффективности миРНК. В рамках исследования был собран обширный датасет, включающий информацию о последовательностях siRNA, их химических модификациях, а также экспериментальных параметрах, таких как концентрация, метод трансфекции, тип клеток, целевой ген и время инкубации. После очистки и стандартизации датасет составил 3435 записей. Данные прошли корреляционный анализ, и были отобраны информативные дескрипторы для последующего моделирования. Для описания химических свойств нуклеотидов в последовательностях был разработан автоэнкодер, а для представления последовательностей целевых генов использованы эмбединги, полученные с помощью языковой модели Mistral.

Для прогнозирования эффективности молекул были разработаны два классификационных подхода и регрессионная модель. Бинарные классификаторы позволили выделить группы миРНК с высокой и низкой эффективностью. Регрессионная модель, построенная на основе дескрипторов, рассчитанных с помощью библиотеки RDKit, продемонстрировала высокую предсказательную способность ( $R^2 \approx 0.87$ ,  $RMSE \approx 10.89\%$ ), что позволяет достаточно точно оценивать эффективность подавления целевого гена по заданным параметрам. Полученные результаты превосходят показатели всех опубликованных на данный момент моделей.

**Выводы.** Проведенное исследование демонстрирует возможность прогнозирования эффективности миРНК для нокдауна генов с использованием методов машинного обучения. Разработанные модели позволяют предварительно оценивать потенциал молекул для эффективного подавления целевого гена. Практическое применение результатов может способствовать оптимизации дизайна миРНК для экспериментальных и терапевтических целей, что позволит сократить затраты времени и ресурсов при разработке новых методов контроля экспрессии генов.

#### **Список использованных источников:**

1. Arkin M.R., Tang Y., Wells J.A. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality // *Chem. Biol.* – 2014. – 21. – P. 1102.
2. Sioud M., Furset G., Cekaite L. Suppression of immunostimulatory siRNA-driven innate immune activation by 2'-modified RNAs // *Biochem. Biophys. Res. Commun.* – 2007. – 361. – P. 122.
3. Fluiter K., Mook O.R.F., Baas F. The therapeutic potential of LNA-modified siRNAs: reduction of off-target effects by chemical modification of the siRNA sequence // *Methods Mol. Biol.* Clifton NJ. – 2009. – 487. – P. 189.