**Секция: Информатика**

**Использование биоинформатических методов для редактирования генов по системе CRISPR/CAS9**
Лебедев Михаил Михайлович, Павленко Иван Александрович
Класс: 11
115522, г. Москва, Южный АО, Район Москворечье-Сабурово, Пролетарский проспект, дом 6, корпус 3, Университетский Лицей № 1511 предуниверситария НИЯУ МИФИ.

lebedev\_mm\_1018@1511.ru, pavlenko\_ia\_1018@1511.ruНаучный руководитель: к.б.н. преподаватель кафедры 94 НИЯУ МИФИ Масловская Е.В.

В настоящее время биоинформатика является одним из самых востребованных и перспективных научных направлений. Для современного редактирования генов по системе CRISPR/CAS используется комплекс белков, именуемых Cas, и в частности белок Cas9. Эта система была найдена у бактерий, где она играет роль иммунитета. После того, как в бактерию попадает ДНК/РНК вируса, эта система активируется, и разрезает чужеродную ДНК/РНК последовательность. Нужно понимать, каким именно образом белки определяют нужную для разрезания последовательность ДНК. Дело в том, что в геноме бактерии хранится своего рода библиотека с кусочками геномов вируса, с которыми до этого встречалась бактерия и ее предки. При построении белков эти кусочки непосредственно встраиваются в белки Cas. После чего Cas система определяет чужеродную ДНК/РНК, и производит разрез в месте совпадения с кусочком ДНК/РНК. Для предотвращения проблемы внесения разреза в собственную ДНК в каждой CRISPR/Cas системе встроен некий "определитель чужеродности" называемый PAM-сайтом, или PAM-последовательностью. Это очень важная последовательность генома, свойственного каждому белку Cas. Таким образом, это защищает бактерию от уничтожения. Однако, эта PAM-последовательность сильно уменьшает вариативность действий ученых по редактированию генома вручную, учитывая то, что ее нельзя поменять. Известно, что каждому роду бактерий свойственна своя PAM-последовательность. Таким образом, можно увеличить вариативность, заимствуя Cas систему у разных родов. Тут встает проблема определения этой последовательности, так как она состоит не в виде генома, но в виде закодированных аминокислот, расположенных в непонятном месте. На данный момент для неточного определения консенсусной PAM последовательности у организма нужно провести от нескольких часов, до нескольких дней в лаборатории, при этом используя довольно дорогое оборудование и реактивы, в свою очередь, целью работы нашей команды являлась реализация метод поиска этой последовательности программно, и добились успешного результата.

На вход программе подаётся CRISPR–кассета. Программа разделяет ее на спейсеры и повторы, при этом отображая все спейсеры и повторы в графической панели. Алгоритм трансформации спейсеров преобразует спейсер в четыре различных состояния: оригинальный, перевернутый, комплементарный и комплементарный перевернутый. Трансформация спейсеров осуществляет быстрый и более тщательный поиск фагов. Поиск фагов осуществляется через BLAST, в котором программа при помощи API находит все бактериофаги и вирусы, содержащие один из состояний спейсера. При помощи запросов к базам данных, программа получает все ДНК нужных вирусов, найденных при помощи BLAST. Далее программа обрабатывает все ДНК-последовательности и вытаскивает из них то, что находится рядом с протоспейсерами, которые комплементарны или равны спейсерам. Алгоритм отсортировывает эти участки и высчитывает вероятность попадания всех рядом прилежащих к спейсеру последовательностей и обрабатывает их при помощи weblogo.

Программа реализована на языке Java, и была протестирована на ряде CRISPR-кассет геномов разных бактерий. Ниже представлены результаты тестирования на Streptococcus Pyogenes, Streptococcus Thermophiles

 . 

Данные логотипы были сгенерированы программой по введенным в нее CRISPR кассетам бактерий, и они полностью одовлетворяют литературным данным.

**Литература.**

1. [A. Bolotin, B. Quinquis, et al. 2005; C. Pourcel, G. Salvignol, et al. 2005; F. J. M. Mojica, C. Díez-Villaseñor, et al. 2005].

2. [Mojica F. J. M., Díez-Villaseñor C., J. et al. 2005, Leenay R.T., Beisel C.L., et al. 2017]

3. [Jiang F, Doudna JA.2017]

4. [Giulia Palermo, Clarisse G. Ricci, 2019]

5. [Mülle M., Lee C.M. 2016, Chatterjee P., Jakimo N. 2018]