

**Систематическое исследование различных способов выявления и  
характеризации распределения мутационной нагрузки на примере белков  
семейства RAS в экстремально большой выборке образцов колоректального  
рака.**

**Научный руководитель – Серебрянский Илья Генрихович**

***Цалковский Иван Алексеевич***

*Студент (бакалавр)*

Казанский (Приволжский) федеральный университет, Институт фундаментальной  
медицины и биологии, Кафедра генетики, Казань, Россия

*E-mail: piqadolf@gmail.com*

Некоторые мутации в опухолевых клетках повторяются довольно часто, такие мутации называются хотспотами. Идентификация хотспотов отражает потенциально функциональные регионы гена, являясь таким образом важной частью в диагностике и планировании таргетной терапии. Однако сотни хотспотов ни разу не были проверены экспериментально. Для теоретического предсказания были разработаны десятки методов[1].

В этом исследовании взяли 3 наиболее распространенных подхода к идентификации хотспотов (3D, sliding window, single-residue) и сравнили предсказательную способность как каждого из них в отдельности, так и различных комбинаций этих методов.

Подобное сравнение ранее еще не проводилось, потому что данные из открытых источников гетерогенны и не достигают необходимого количества. В качестве источника данных использовали когорту, состоящую из 34,129 образцов, полученную при сотрудничестве с Foundation Medicine. Анализ проводился с помощью скриптов на R и python.

Семейство белков RAS включено в важнейшие для клеточного цикла сигнальные пути, которые часто нарушены в раковых клетках[2]. Однако и эти белки остаются не до конца изученными, также в силу отсутствия достаточного количества данных. Наш анализ показал, что нашего размера выборки вполне достаточно для предсказания практически всех хотспотов KRAS. В качестве промежуточных результатов можно заключить, что подходы sliding window и 3D hotspots в первую очередь позволяют быстрее обнаружить потенциальные single-residue хотспоты, и уже во-вторых, показать новые функциональные участки. Также, с помощью даунсэмплинга, определили оптимальные размеры выборки для каждого варианта.

В дальнейшем сделать сделать кросс-валидацию наших результатов на открытых данных, а также других генах.

### **Источники и литература**

- 1) Emmanuel Martinez-Ledesma, David Flores, Victor Trevino, Computational methods for detecting cancer hotspots, Computational and Structural Biotechnology Journal, Volume 18, 2020, Pages 3567-3576
- 2) Khan I, Rhett JM, O'Bryan JP. Therapeutic targeting of RAS: New hope for drugging the "undruggable". Biochim Biophys Acta Mol Cell Res. 2020;1867(2):118570.