

## Устойчивость консенсусного метода независимых компонент для классификации пациентов с немелкоклеточным раком легких

М.К. Чепелева<sup>1</sup>, Н.Н. Яцков<sup>1</sup>, П.В. Назаров<sup>2</sup>

<sup>1</sup> Белорусский государственный университет, Минск, Беларусь

<sup>2</sup> Luxembourg Institute of Health, Strassen, Luxembourg

E-mail: maryna.chepeleva@gmail.com

Метод независимых компонент (МНК) может применяться для декомпозиции неоднородных транскриптомных данных и выделения транскрипционных сигналов, соответствующих биологическим функциям или техническим эффектам [1]. Для использования независимых компонент в качестве признаков транскрипционных сигналов требуется высокая воспроизводимость разложения исходных данных. В работе исследована стабильность МНК, проверена воспроизводимость результатов анализа в зависимости от количества параллельных запусков, усредняемых для получения консенсусного МНК разложения, а также числа образцов.

Разработанный консенсусный МНК [2] применялся к данным экспрессии генов двух типов немелкоклеточного рака легких: плоскоклеточному (LUSC) и аденокарциноме (LUAD) из базы данных TCGA [2]. Стабильность разложения исследована в зависимости от количества параллельных запусков МНК, компонента, объема выборки пациентов. Для классификации пациентов на основании выделенных компонент рассмотрены предсказательные модели из R-пакетов *random forest* и *xgboost*. Для оценки устойчивости разложения рассчитывались индексы Жаккара, коэффициенты детерминации и косинусное сходство между полученными метагенами и списками обогащенных биологических процессов.

Определено оптимальное количество параллельных раундов консенсусного МНК, обеспечивающее воспроизводимое разложение для каждого типа рака. Проведена оценка влияния количества параллельных запусков и объема выборки на качество классификации типов рака легких. Для выделенных компонент определены биологически значимые сигналы, включая клеточный цикл, активность инфильтрирующих иммунных клеток и кератинизация.

1. *Nazarov P.V., Wienecke-Baldacchino A.K., Zinovyev A., et al. // BMC Med Genomics. 2019. V. 12, No. 1. P. 132–149.*
2. *Sompairac N., Nazarov P.V., Czerwinska U., et al. // Int. J. Mol. Sci. 2019. V. 20, No. 18. P. 4414–4441.*
3. *Tomczak K, Czerwińska P, Wiznerowicz M. // Contemp Oncol (Pozn). 2015. V. 19, No. A1. P.68–77.*