

КЛАСТЕРНЫЙ АНАЛИЗ ДАННЫХ ЭКСПЕРИМЕНТА ПО ВЛИЯНИЮ ИОНОВ МЕДИ НА ПЕРВИЧНЫЕ РЕАКЦИИ ФОТОСИНТЕЗА

Киселева Д.Г., Плюснина Т.Ю., Хрущев С.С.

Московский государственный университет им. М.В. Ломоносова, Биологический ф-т,
каф. биофизики, Россия, 119992, Москва, Ленинские горы, МГУ

При анализе данных биологических исследований традиционно применяются статистические методы. Методы кластеризации, в свою очередь, получили распространение относительно недавно, с ростом вычислительной мощности компьютера. Особенности кластерного анализа заключаются в отсутствии первичных допущений и требований к распределению данных и использовании не сопоставимо больших объёмов выборок. В последнее годы объем накопленных экспериментальных биологических данных значительно вырос, и во многих случаях статистический анализ представляет собой довольно трудоемкий процесс, тогда как кластерный анализ при обработке большого количества данных имеет значительные преимущества. Для оценки возможностей кластерного анализа при работе с биологическими данными были использованы данные эксперимента по действию ионов меди на проростки гороха *Pisum sativum* L. Пробы выдерживались в присутствии токсиканта и без него в климатической камере в течение 10 дней. На протяжении всех дней была проведена регистрация индукционных кривых индукции флуоресценции хлорофилла а. Для анализа кинетики индукции флуоресценции мы использовали JIP-тест. Полученные показатели JIP-теста в дальнейшем использовались для кластеризации. На первом этапе была выполнена оценка наличия корреляции между параметрами JIP-теста и факторами, действующими на проростки (Cu^{2+} , день инкубации). Корреляционный анализ выявил, что в рассматриваемом эксперименте *Pisum sativum* L. подвергались стрессу как в результате действия токсиканта Cu^{2+} , так и в результате условий самого эксперимента (удаление корней и усиков). Было показано, что различия между разными днями инкубации являются более существенными, чем различия между контрольными образцами и обработанными токсикантом Cu^{2+} . Далее была проведена предварительная подготовка данных, которая заключалась в уменьшении их неоднородности, возникшей из-за разного количества измерений по каждому растению, нормализации данных и снижении их размерности с помощью метода главных компонент, выбора количества главных компонент. Для дальнейшей кластеризации были выделены две главные компоненты. Кластеризация проводилась тремя методами: иерархическим, k-средних и OPTICS. Для характеристики полученных кластеров был проведен анализ параметров JIP-теста в каждом из них. Было показано, что кластеры были сгруппированы по степени подверженности растения стрессу. Комплексный анализ показал, что в данном эксперименте фактор дня инкубации вызывал больший стресс у растений, чем фактор действия токсиканта Cu^{2+} . Для выявления действия Cu^{2+} необходимо использовать его в больших концентрациях.

Исследование выполнено в рамках научного проекта государственного задания МГУ №121032500060-0 при частичной поддержке гранта РФФИ № 20-64-46018 и гранта РФФИ № 20-04-00465