

## АЛГОРИТМЫ КЛАСТЕРИЗАЦИИ ДЛЯ ОЦЕНКИ СПОСОБНОСТИ НАБОРА СТРУКТУРНЫХ ДЕСКРИПТОРОВ ПРОГНОЗИРОВАТЬ АЛЛОСТЕРИЧЕСКИЕ ВЗАИМОДЕЙСТВИЯ РЕЦЕПТОРА ВИТАМИНА D

Пац К. М. (Университет ИТМО)

Научный руководитель – к.т.н., доц. Сергушичев А. А.

(Университет ИТМО), PhD, проф. Молнар Ф. (Назарбаев Университет)

Работа посвящена исследованию способности набора структурных дескрипторов описывать аллостерические взаимодействия рецептора витамина D. Исследование проводилось с использованием алгоритмов кластеризации и методов понижения размерности. Полученные результаты могут быть использованы при разработке модели, предсказывающей взаимодействия ядерных рецепторов с различными малыми молекулами и коактиваторами/корепрессорами белковой природы.

**Введение.** Структурные дескрипторы широко применяются для оценки подобия двух структур, а также позволяют делать выводы о принадлежности белковых молекул к тому или иному семейству. Существуют белковые молекулы, использующие аллостерическую регуляцию в механизмах своей работы. Примером может служить рецептор витамина D (VDR). Известно, что поведение различных сайтов связывания данного рецептора зависит от связанных с рецептором молекул – лигандов, коактиваторов или корепрессоров, молекул ДНК. В предыдущих исследованиях было показано, что помимо задач классификации белковых молекул, структурные дескрипторы могут быть использованы для задач кластеризации. При анализе лигандов рецептора витамина D при помощи таких дескрипторов, как % заполняемости сайта связывания, объем лиганда, а также длина спиралей 6 и 7, удалось выявить кластеры, объединяющие слабые агонисты, сильные агонисты, а также лиганды, содержащие заместители в положении 2 $\alpha$ . Таким образом, в данной работе производится более масштабное исследование предсказательной способности структурных дескрипторов, включающее 15 дескрипторов и 126 структур рецептора витамина D из базы данных PDB.

**Основная часть.** В работе используется ранее разработанный список структурных дескрипторов, включающий такие параметры, как расстояние между вторичными структурами, их относительная ориентация, длина, тип, поверхность доступная растворителю. Далее из базы данных PDB ([www.rcsb.org](http://www.rcsb.org)) загружаются трехмерные структуры рецептора витамина D человека. Данные структуры проходят предобработку на предмет восстановления недостающих боковых цепей, оптимизации водородных связей и энергетической минимизации. Затем с помощью ранее разработанного пайплайна для выбранных рецепторов рассчитывается набор структурных дескрипторов. Таким образом для каждого из рецепторов формируется их структурный профиль, при этом в результате образуется два набора расчетов – для структур, прошедших предобработку, и для структур, использованных в исходном виде. Далее при помощи различных алгоритмов кластеризации в комбинации с методами понижения размерности (PCA, t-SNE) проводится исследование полученных результатов расчета для каждого из двух наборов. Таким образом, становится возможным сделать выводы о дескрипторах, вносящих наибольший вклад в формирование кластеров, а также выяснить, является ли предобработка структур из PDB необходимым шагом для выявления зависимостей взаимодействия рецептора витамина D от его структурных особенностей.

**Выводы.** Использование структурных дескрипторов для анализа аллостерического поведения рецептора витамина D является новым и перспективным направлением исследований. Использование модели, основанной на структурных дескрипторах, может помочь избежать

трудоемкого анализа результатов молекулярной динамики, которая является стандартом при моделировании взаимодействий «белок-лиганд» с течением времени.

В ходе работы проведено сравнение различных алгоритмов кластеризации применительно к набору структурных дескрипторов рецептора витамина D, проведено сравнение кластеров, образуемых при использовании обработанных и необработанных трехмерных структур, а также сделаны выводы о дескрипторах, вносящих наибольший вклад в формирование данных кластеров.

Результаты работы могут быть использованы при дальнейшей разработке модели, основанной на структурных дескрипторах, для предсказания аллостерических взаимодействий белков семейства ядерных рецепторов. Такая модель может стать удобным инструментом анализа результатов экспериментов по молекулярной динамике, позволяющим сократить время на поиск интересующих закономерностей и причинно-следственных связей в образовании аллостерических взаимодействий ядерных рецепторов.