

Побойнев В. В.

РЕЗУЛЬТАТЫ АПРОБАЦИИ АЛГОРИТМА ДЛЯ ПРЕДСКАЗАНИЯ ВТОРИЧНОЙ СТРУКТУРЫ БЕЛКОВ И ПЕПТИДОВ

Научный руководитель: канд. биол. наук, доц. Хрусталёв В. В.

Кафедра общей химии

Белорусский государственный медицинский университет, г. Минск

Актуальность. Для большинства аминокислотных остатков не существует резко выраженных предпочтений для формирования того или иного элемента вторичной структуры. Вероятность существования каждого элемента вторичной структуры возрастает, если соответствующие аминокислоты образуют кластеры. Влияние на фолдинг данного кластера оказывают фланкирующие элементы вторичной структуры, что учитывается в алгоритме PentaFOLD 2.0 для предсказания вторичной структуры белка.

Цель: протестировать алгоритм для предсказания вторичной структуры белков и пептидов PentaFOLD 2.0.

Материалы и методы. В данной работе использовались 3D структуры белков человека, относящиеся к четырём структурным классам: альфа-спиральные, бета-структурные, альфа + бета и альфа/бета. Каждая выборка включает по 100 аминокислотных последовательностей, полученных из международной базы данных PDB. Максимальный процент сходства аминокислотных последовательностей белков в каждой выборке друг с другом не превышал 25% по алгоритму Decrease Redundancy.

Результаты и их обсуждение. Чувствительность к альфа-спиралям у альфа-спирального (66,90%; 65,36%; 65,04% и 67,15%) и бета-структурного паттерна (31,48%; 32,02%; 31,27%; 30,12%) алгоритма PentaFOLD не зависит от структурного класса белка (структурные классы белков расположены в порядке – альфа + бета, альфа/бета, бета, альфа). Специфичность к альфа-спиралям демонстрирует ярко выраженную зависимость от структурного класса. По альфа-спиральному паттерну наибольшего значения (74,50%) специфичность достигла в выборке альфа-спиральных белков, наименьшего (17,06%) – в выборке бета-структурных белков, специфичность в выборке белков альфа/бета класса (52,89%) превысила таковую в выборке белков альфа + бета (48,43%). Похожая ситуация наблюдается и с предсказанием бета-тяжей. Чувствительность к ним как по бета-паттерну, так и по альфа-паттерну находится примерно на одном и том же уровне для белков альфа + бета, бета и альфа классов (61,33% и 29,54%; 59,68% и 27,91%; 55,89 и 26,18%). Несколько более высокие значения характерны для чувствительности к бета-тяжам среди альфа/бета белков: 66,31% и 35,94%. Что касается специфичности к бета-тяжам, то самые высокие её значения достигаются в выборке бета-структурных белков (60,89% и 69,21%), наименьшие (5,48% и 9,25%) – в выборке альфа-спиральных белков. Для белков класса альфа + бета и альфа/бета специфичность к бета-тяжам практически идентична (33,80% и 45,95%; 32,31% и 47,51%). Что касается неструктурированных участков белков, то результаты предсказаний алгоритма PentaFOLD 2.0 по альфа- и бета-паттернам отличаются незначительно. При обработке четырёх выборок чувствительность и специфичность для неструктурированных участков как по альфа-, так и по бета-паттерну находится в пределах 60-70%.

Выводы. С помощью четырёх выборок был протестирован оригинальный алгоритм PentaFOLD 2.0, определены чувствительность и специфичность данного алгоритма для различных элементов вторичной структуры белков по двум паттернам. Наилучшие результаты алгоритм PentaFOLD достигает при использовании альфа-спирального паттерна для альфа-спиральных белков и бета-структурного паттерна – для бета-структурных. С помощью экспериментальных данных (анализа спектра кругового дихроизма, комбинационного рассеяния или инфракрасного спектра) можно отнести белок к преимущественно альфа-спиральному, преимущественно бета-структурному, или к белку, содержащему как бета-тяжи, так и альфа-спирали. При наличии такой информации эффективность работы PentaFOLD для альфа-спирального или бета-структурного белка достигнет максимума.