

Разработка базы данных и подбор параметров программы Mascot для оптимизации диагностики онкозаболеваний органов дыхания

Очередыко Елена Сергеевна¹, Медведева Мария Витальевна²

1 - Московский государственный университет имени М.В.Ломоносова, Москва, Россия; 2

- Московский государственный университет имени М.В.Ломоносова, Факультет

биоинженерии и биоинформатики, Москва, Россия

E-mail: MsCorleone96@gmail.com

Рак лёгкого чаще, чем другие типы рака, приводит к летальному исходу. В настоящее время среди разрабатываемых методов ранней диагностики рака лёгкого одним из наиболее перспективных считается анализ белкового состава конденсата выдыхаемого воздуха (КВВ). Для идентификации белков в КВВ чаще всего используют тандемную масс-спектрометрию и анализ масс-листов с помощью программы Mascot, однако, особенности пробы затрудняют однозначную интерпретацию белков и стандартизацию анализа. Ещё одна проблема – отсутствие эффективных инструментов для хранения и обработки большого объёма информации из масс-спектров без потерь промежуточных данных о пробоподготовке и особенностях доноров, поскольку, как показали предшествующие эксперименты, анамнез пациента оказывает прямое влияние на состав пробы.

Целью работы было создание базы данных (БД) для решения проблемы хранения информации и калибровка существующего интерфейса для интерпретации масс-спектров КВВ.

Образцы КВВ были собраны у 103 человек, из них 43 пациента с диагностированным раком легкого, находящихся в отделении торакальной хирургии МНИОИ им. П.А. Герцена, 28 здоровых некурящих донора, 17 больных хронической обструктивной болезнью легких и 15 - внебольничной пневмонией, находящихся на лечении в пульмонологическом отделении ГКБ № 57 г.Москвы. Анализ проб КВВ проводился с использованием ВЭЖХ-МС/МС. Полученные спектры обрабатывались с использованием программы Mascot (Matrix Science, version 2.0.04). Для создания БД использовали MySQL Workbench.

В результате проведенной работы, была создана БД, позволяющая загружать всю необходимую информацию о доноре и пробе (включая анамнез и масс-хроматограммы) и осуществлять первичный анализ полученных результатов, что должно упростить и автоматизировать процесс обработки данных, а также сделать возможным сравнение данных по комплексу параметров для формирования наиболее полной картины состояния организма. Информация о пробах КВВ была загружена в базу и использована для калибровки Mascot: были отобраны параметры, которые оказывают влияние на результат работы программы. Первичный анализ результатов работы программы Mascot показал, что каждому однозначно определённой пептиду могут соответствовать несколько идентифицированных белков, поэтому дальнейшую диагностику рака лёгкого рациональнее строить на основании найденных пептидов.

Слова благодарности

Авторы выражают признательность Филимонову И.С., Федорченко К.Ю., Митрофанову С.И., Пономарёву О.А. за неоценимую помощь в написании работы.