

Московский физико-технический институт (государственный университет)

Факультет биологической и медицинской физики

Кафедра кафедры молекулярной и трансляционной медицины

Диссертация допущена к защите

зав. кафедрой

\_\_\_\_\_ Лазарев В.Н.

«\_\_\_\_\_» \_\_\_\_\_ 2017 г.

**Выпускная квалификационная работа  
на соискание степени  
МАГИСТРА**

**Тема: Количественный протеоеномный  
анализ туберкулеза  
и ещё чего-нибудь**

Направление: 010900 – Прикладные математика и физика

Магистерская программа: 010982 – Физико-химическая биология и биотехнология

Выполнил студент гр. 1114 \_\_\_\_\_ Смоляков А.В.

Научный руководитель,

к. б. н. \_\_\_\_\_ Лазарев В.Н.

Работа выполнена в ФГБУ ФНКЦ ФХМ ФМБА России

Москва – 2017

# Оглавление

1.	Список сокращений . . . . .	3
2.	Введение . . . . .	4
3.	Литературный обзор . . . . .	5
3.1.	Mycobacterium tuberculosis . . . . .	5
3.2.	Применение масс-спектрометрии в протеомике . . . . .	5
3.3.	Orbitrap . . . . .	5
3.4.	Количественная протеомика . . . . .	5
3.5.	Безметочная квантификация . . . . .	5
	Относительная квантификация по интенсивности MS-1 сигнала . . . . .	5
3.6.	Количественный анализ по количеству спектров . . . . .	5
4.	Материалы и методы . . . . .	7
4.1.	Получение бактерий . . . . .	7
4.2.	Проведение масс-спектрометрического эксперимента . . . . .	7
4.3.	Контроль качества . . . . .	7
4.4.	Протеогеномика . . . . .	7
	Создание поисковых баз . . . . .	7
	Идентификация новых белков . . . . .	7
	Уточнение N-конца . . . . .	7
	Анализ SAP . . . . .	7
4.5.	Идентификация пептидов и белков . . . . .	7
4.6.	Количественный анализ . . . . .	8
	Анализ в MaxQuant . . . . .	8
	Анализ в Progenesis LC-MS . . . . .	8
	Сравнение результатов программ . . . . .	8
5.	Результаты и обсуждение . . . . .	9
6.	Выводы . . . . .	10
	<b>Список литературы . . . . .</b>	<b>11</b>

## 1. Список сокращений

## 2. Введение

### **3. Литературный обзор**

#### **3.1. *Mycobacterium tuberculosis***

#### **3.2. Применение масс-спектрометрии в протеомике**

#### **3.3. Orbitrap**

#### **3.4. Количественная протеомика**

#### **3.5. Безметочная квантификация**

Вне зависимости от выбранного метода безметочного количественного анализа, эксперимент должен включать в себя следующие шаги: 1. пробоподготовка, включая извлечение белков, их очистку, трипсинолиз и прочие шаги 2. разделение пептидов при помощи различных хроматографических методов с последующий MS/MS анализом 3. анализ полученных результатов: идентификация, количественный анализ и статистический анализ В целом, методы безметочного количественного анализа можно разделить на две группы: на основе интенсивности ионов или за счет spectral counting [1].

#### **Относительная квантификация по интенсивности MS-1 сигнала**

При ионизации электроспреем, интенсивность MS-1 иона коррелирует с его концентрацией [2]. Впервые количественный анализ белков и пептидов за счет интенсивности LC-MS пиков был проведен на миоглобине. Были проанализированы концентрации в диапазоне от 10 фемтамоль до 100 пикомоль. После

#### **3.6. Количественный анализ по количеству спектров**

Квантификация по спектральному показателю основана на сравнении количества идентифицированных tandemных спектров для одного и того же белка из различных проб. Луи и соавторы изучали связь между относительной концентрацией белка и следующими тремя факторами: процентом идентифицированной (покрытой) части белка, количеством наблюдаемых пептидов и спектральным показателем. Было показано, что среди этих факторов только спектральный показатель имеет сильную линейную корреляцию с относительной концентрацией белка (коэффициент детерминации  $r^2 = 0.9997$ ), в динамическом диапазоне, составляющем два порядка.

Таким образом, спектральный показатель можно использовать для определения относительной представленности белка в данном образце. Зайбелов и соавторы провели исследование мембранных белков *S. cerevisiae*. Они провели идентификацию и количественный анализ, используя подход спектрального показателя и мечением  $^{15}\text{N}$ . Была показана высокая корреляция между двумя методами для белков, с высоким отношением “сигнал/шум”. Так же было установлено, что метод спектрального показателя имеет больший динамический диапазон и воспроизводимость.

## 4. Материалы и методы

### 4.1. Получение бактерий

### 4.2. Проведение масс-спектрометрического эксперимента

### 4.3. Контроль качества

### 4.4. Протеогеномика

#### Создание поисковых баз

В работе использовалось 2 типа баз: белковая и геномная. Белковая база - аннотированные последовательности, для данного штамма. Геномная - база, полученная в результате транслирования генома в шести рамках. Белковые базы для *M.tuberculosis* W-148 и *M.tuberculosis* H37Rv были составлены из аннотированных белков штаммов (NCBI Reference Sequence: NZ\_CP012090.1, 4244 аминокислотных последовательностей для W-148 и ). Геномные базы были получены в результате 6 рамочного транслирования от стоп- до стоп-кодона геномов штаммов *M.tuberculosis* W-148 и *M.tuberculosis* H37Rv, используя программу Artemis версия 16.0.0 [3]. При транслировании использовалась 11 трансляционная таблица NCBI. Минимальная длина рамки была установлена в 100 нуклеиновых кислот. К каждой базе были добавлены последовательности 26 контаминантных белков (кератины, альбумины, трипин).

#### Идентификация новых белков

#### Уточнение N-конца

#### Анализ SAP

### 4.5. Идентификация пептидов и белков

Данные полученные в результате LC-MS/MS эксперимента (Raw формат) были сконвертированы в пик-лист (MGF формат), используя ProteoWizard msconvert [4]. Идентификация проходила против двух белковых и двух геномных баз с использованием Mascot Search Engine version 2.5.1 [5]. Параметры поиска были следующими: триптические пептиды, не более двух пропущенных сайтов трипсинолиза, ошибка массы прекурсера 20 ppm, ошибка массы фрагментов 0.5 Да, заряды прекурсера

2+, 3+, 4+. Oxidation(M), Carbamidomethylation(C) and Deamidated(NQ) были установлены как возможные модификации пептидов. Для подсчета FDR и порогового скоринга использовался поиск против decoy-базы, полученной в результате реверса исходной базы. FDR был выбран на уровне 5%. Пептид считался идентифицированным, если его скор выше порогового скоринга и ранг равен единице. Белок считался идентифицированным, если для него нашлось два и более уникальных пептидов.

#### **4.6. Количественный анализ**

##### **Анализ в MaxQuant**

##### **Анализ в Progenesis LC-MS**

##### **Сравнение результатов программ**



## 5. Результаты и обсуждение

## 6. Выводы

## Список литературы

1. Zhu W., Smith J. W., Huang C.-M. Mass spectrometry-based label-free quantitative proteomics // BioMed Research International. 2009. Vol. 2010.
2. Voyksner R. D., Lee H. Investigating the use of an octupole ion guide for ion storage and high-pass mass filtering to improve the quantitative performance of electrospray ion trap mass spectrometry // Rapid Communications in Mass Spectrometry. 1999. Vol. 13, no. 14. P. 1427–1437.
3. Rutherford K., Parkhill J., Crook J. et al. Artemis: sequence visualization and annotation // Bioinformatics. 2000. Vol. 16, no. 10. P. 944–945.
4. Chambers M. C., Maclean B., Burke R. et al. A cross-platform toolkit for mass spectrometry and proteomics // Nature biotechnology. 2012. Vol. 30, no. 10. P. 918–920.
5. Cottrell J. S., London U. Probability-based protein identification by searching sequence databases using mass spectrometry data // electrophoresis. 1999. Vol. 20, no. 18. P. 3551–3567.