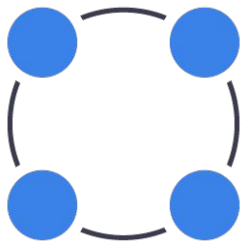


# Аналитические инструменты для базы данных геномных вариаций *Mycobacterium tuberculosis*

Александр Бебяков



Руководитель: Черняева Екатерина  
СПбГУ, Центр геномной биоинформатики  
им. Ф. Г. Добржанского

# GMTV Database



Genome-based Mycobacterium Tuberculosis Variation Database

<http://mtb.dobzhanskycenter.org/>

# Цель проекта

Разработка модуля поиска геномных маркеров резистентности *Mycobacterium tuberculosis* для расширения функционала web-клиента GMTV

# Задачи проекта

Изучить особенности поиска геномных ассоциаций в бактериях на основе данных WGS и используемые методы

Создать приложение поиска маркеров резистентности применительно к текущему состоянию базы GMTV для составления и обновления каталога известных вариаций

Обеспечить сравнение вариантов новых последовательностей с данными в каталоге через web-клиент GMTV

# Исходные данные

- vcf-файлы;
- таблица:
  1. Номер образца
  2. Имя образца
  3. Имя vcf-файла
  4. Препарат: устойчивость(1) или чувствительность(-1)

# Препараты

streptomycin

isoniazid

rifampicin

ethambutol

ethionamide

kanamycin

ofloxacin

para-aminosalicylic

cycloserine

capreomycin

amikacin

pyrazinamide

moxifloxacin

prothionamide

# Скрипт

- Фильтрация vcf
- Анализ
- Вывод по препаратам
  - Вариация
  - Odds Ratio
  - P-value
  - Эффект – аминокислотная замена

# Что сделано

- Скрипт
- 14 препаратов
- 1338 образцов
- Точный тест Фишера
- 2553 мутации ( $p < 10^{-7}$ )
- Кумулятивное пуассоновское распределение [Zhang et al, 2013]
- 668 CDS ( $p < 10^{-5}$ )

$$P(k, \lambda) = e^{-\lambda} \sum_{i=k}^{\infty} \frac{\lambda^i}{i!} = 1 - e^{-\lambda} \sum_{i=0}^k \frac{\lambda^i}{i!}$$



# Литература

- Chen PE, Shapiro BJ: The advent of genome-wide association studies for bacteria. *Curr. Opin. Microbiol.* 2015, 25:17–24.
- Farhat et al: A phylogeny-based sampling strategy and power calculator informs genome-wide association study design for microbial pathogens. *Genome Medicine* 2014, 6:101.
- Zhang et al: Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat Genet* 2013, 45:1255–1260.
- Desjardins et al: Genomic and functional analyses of Mycobacterium tuberculosis strains implicate ald in 198 D -cycloserine resistance. *Nat Genet* 2016, 48:544–551.

**Спасибо за  
внимание**