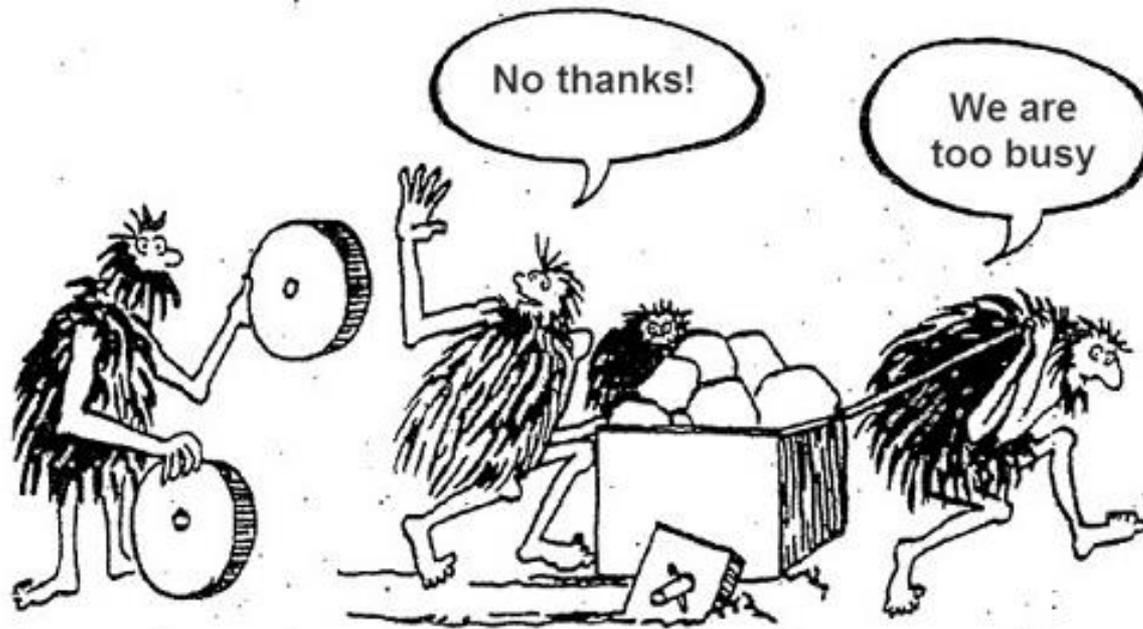


Протеомная Золушка:
оптимизированный анализ больших массивов масс-
спектрометрических данных за одну ночь

Ольга Киселева, Екатерина Поверенная
ИБМХ им. В.Н. Ореховича



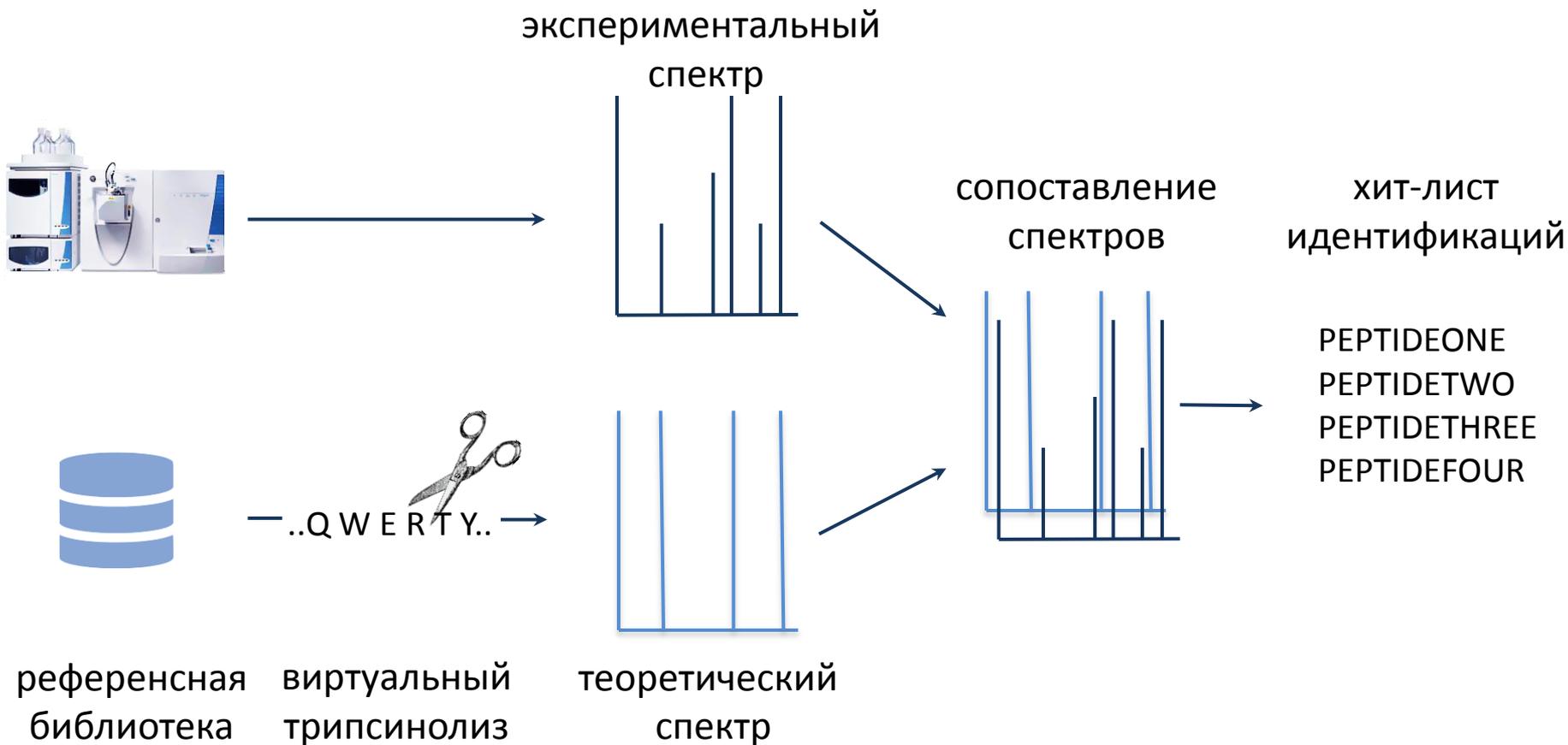
Анализ экспериментальных данных – «Ахиллесова пята» протеомики



-Не надо, спасибо!

-Мы и так слишком заняты!

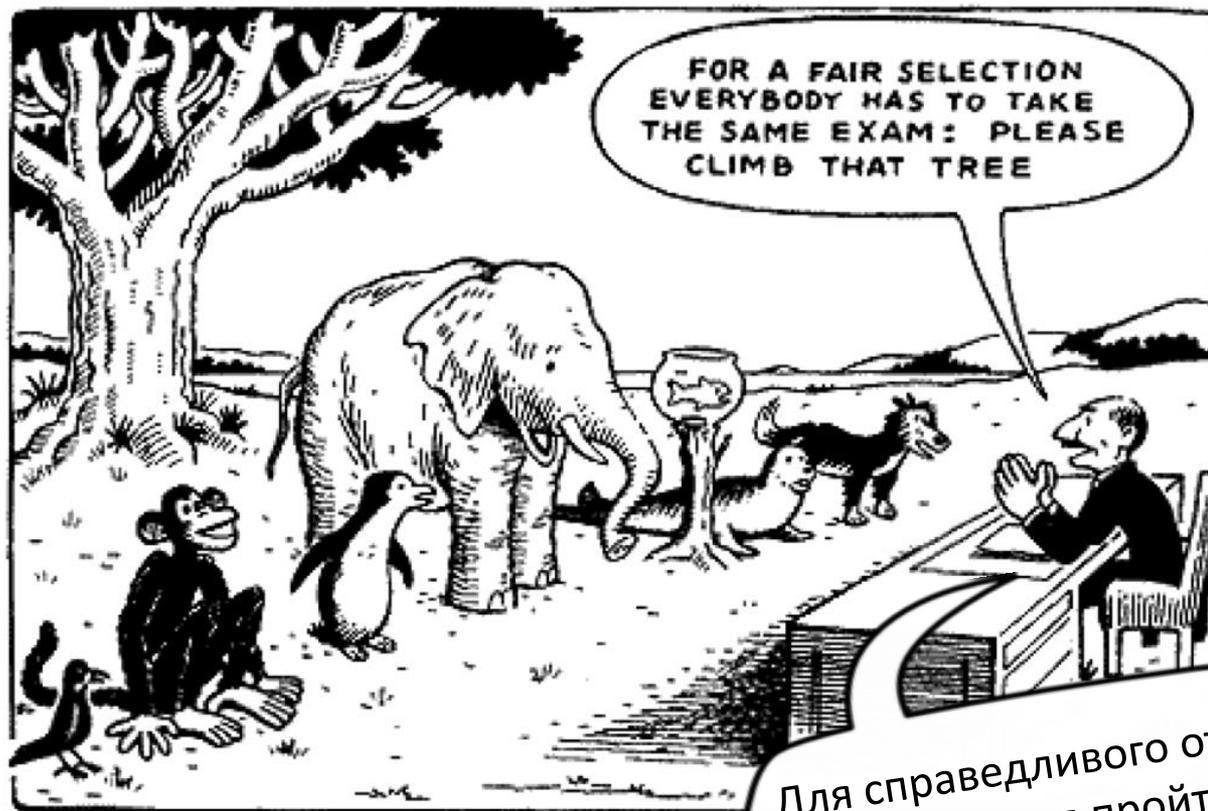
Стратегия интерпретации экспериментальных данных по референсным библиотекам



Оптimum условий интерпретации экспериментальных данных

- ~ аппаратное обеспечение и ОС
- ~ размер референсной библиотеки
- ~ размер анализируемого спектра
- ~ поисковый алгоритм или комбинация алгоритмов

Выбор оптимального сочетания поисковых алгоритмов



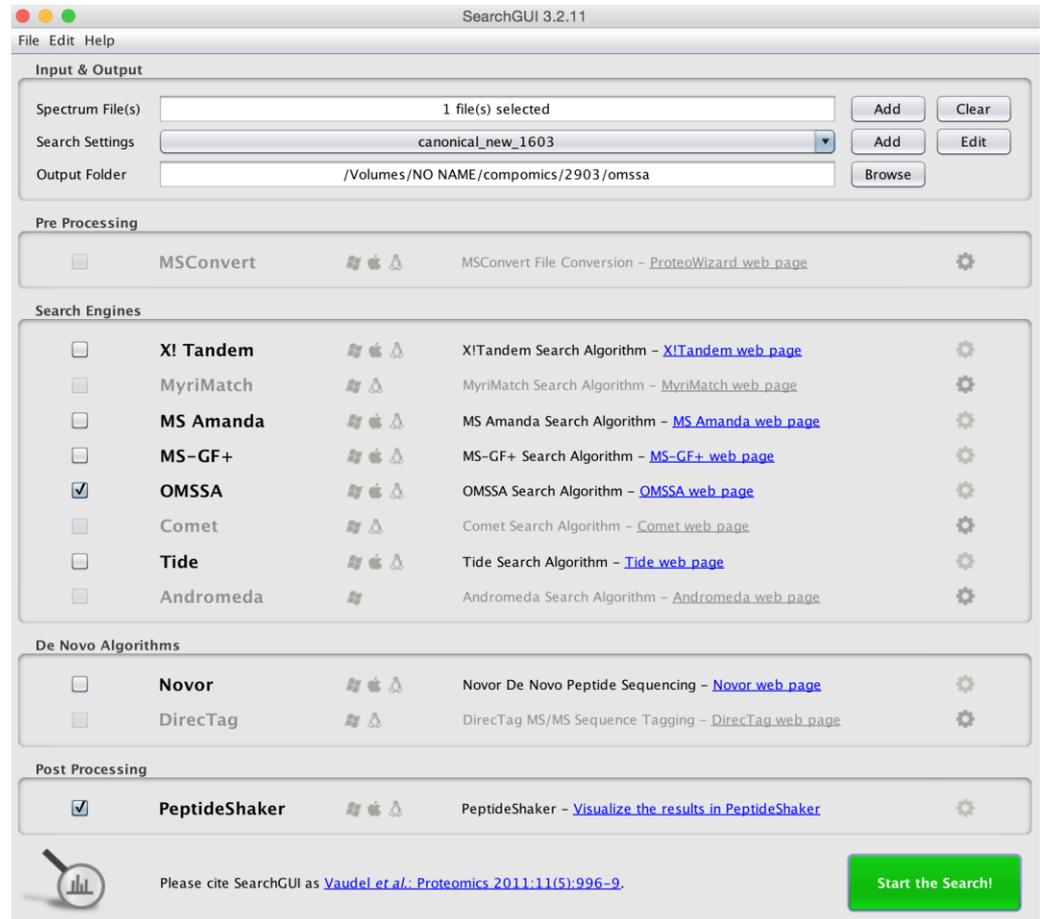
Для справедливого отбора вам всем придется пройти через одинаковое испытание: залезьте вон на то дерево!

SearchGUI и PeptideShaker

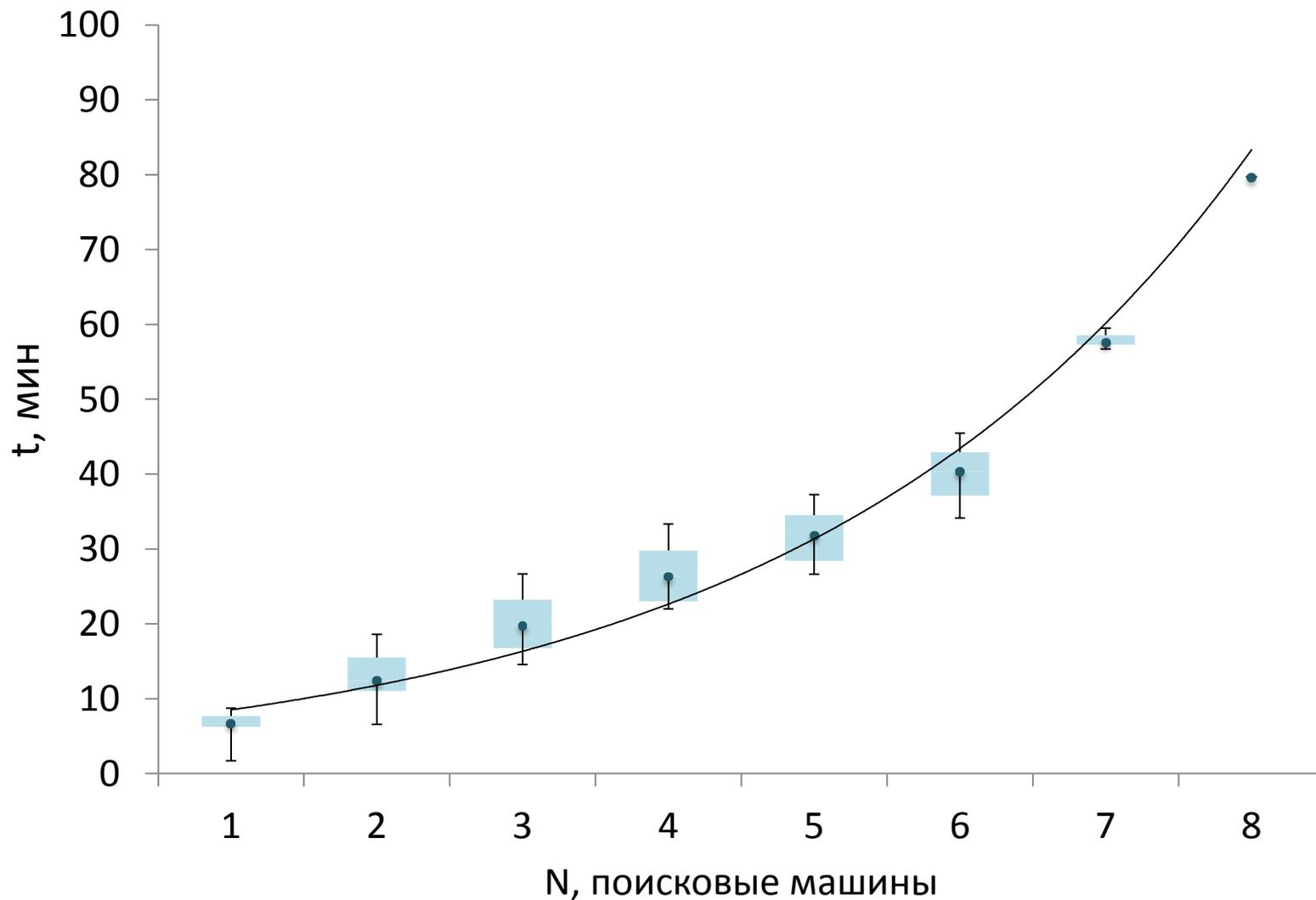
Выбор анализируемого спектра и параметров поиска

Выбор поисковых алгоритмов

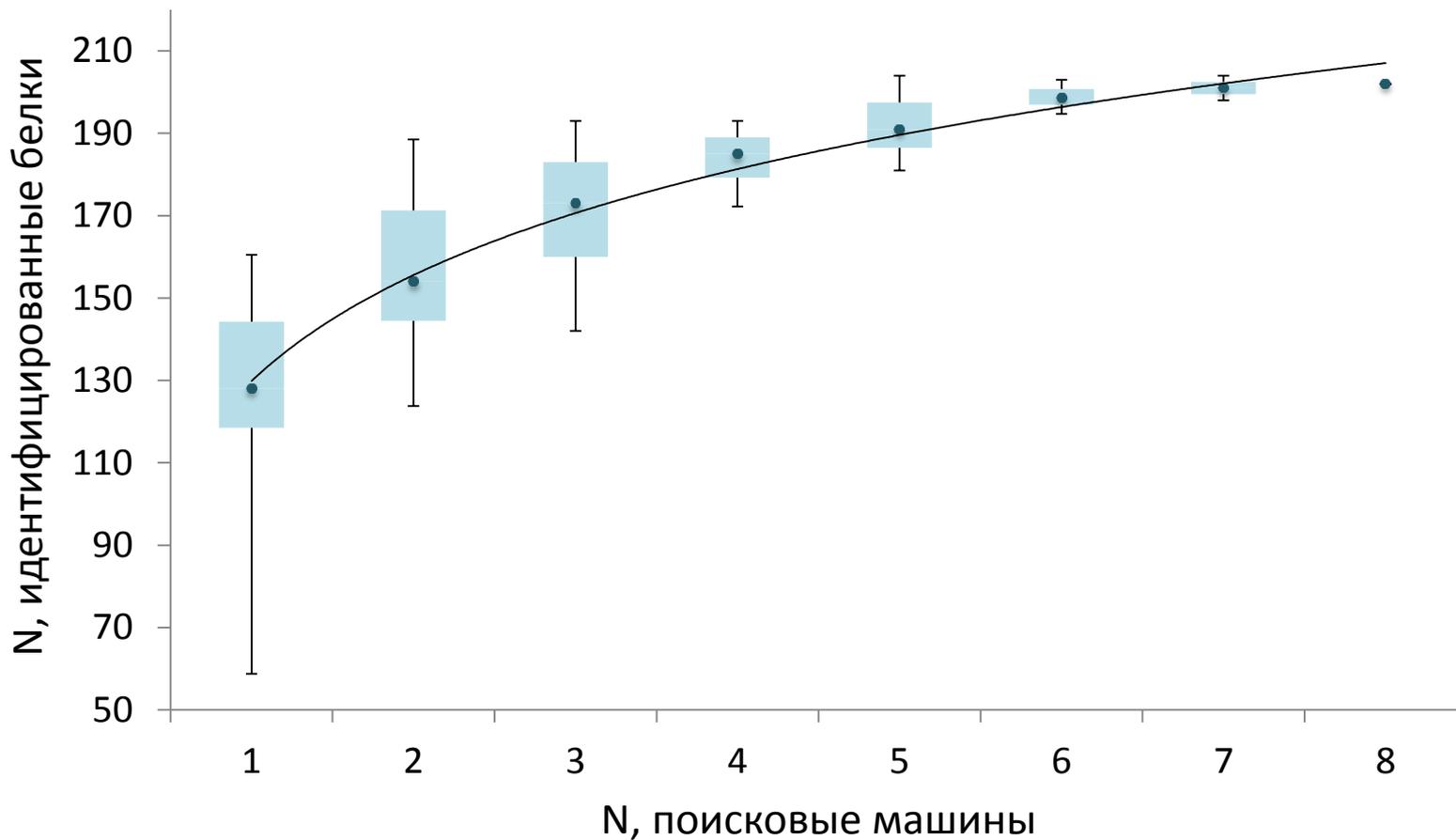
Пост-процессинг результатов интерпретации и визуализация



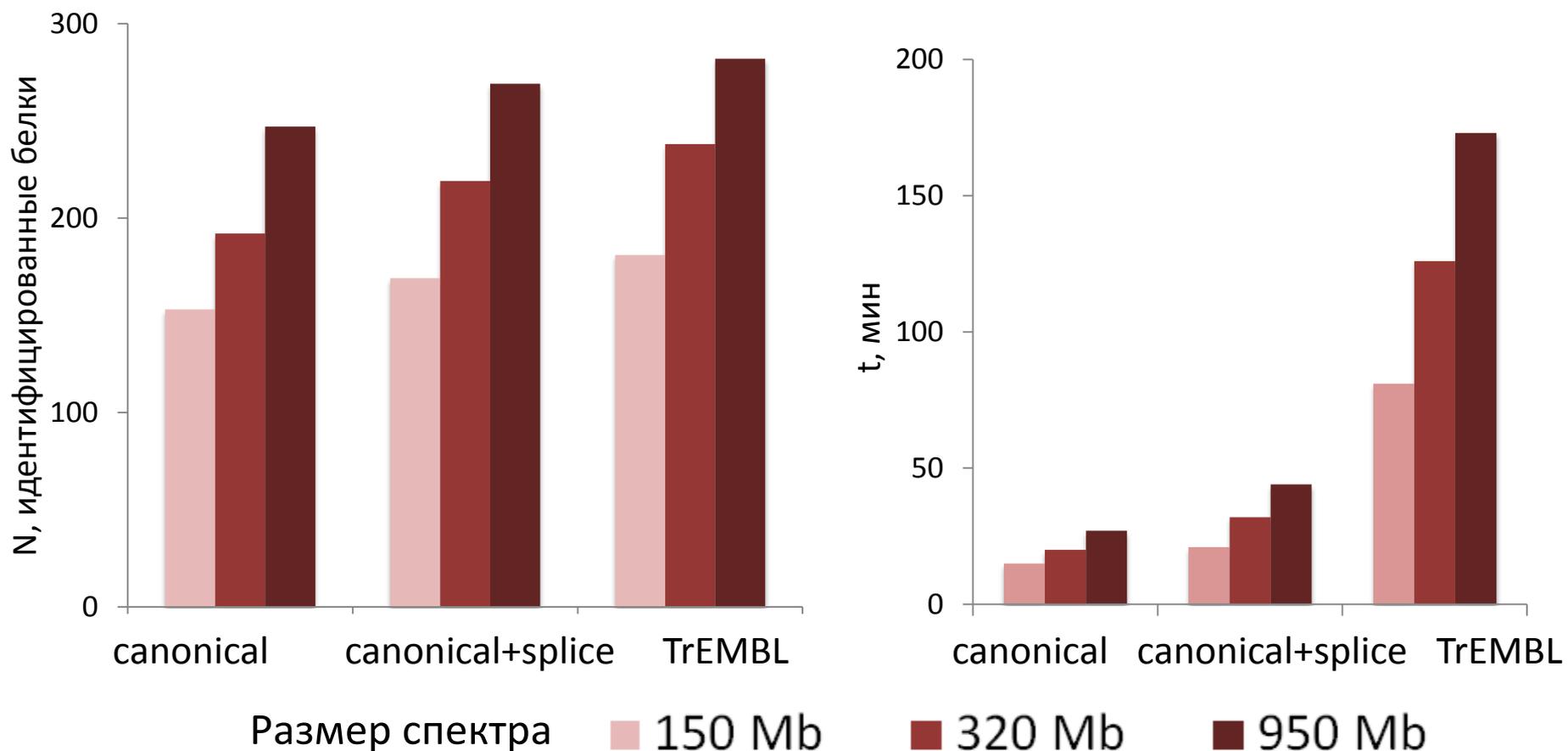
Время интерпретации масс-спектра зависит от комбинации различных наборов поисковых алгоритмов **экспоненциально**



Число идентификаций зависит от комбинации различных наборов поисковых алгоритмов **логарифмически**



Результат идентификации зависит от размера подключаемой библиотеки

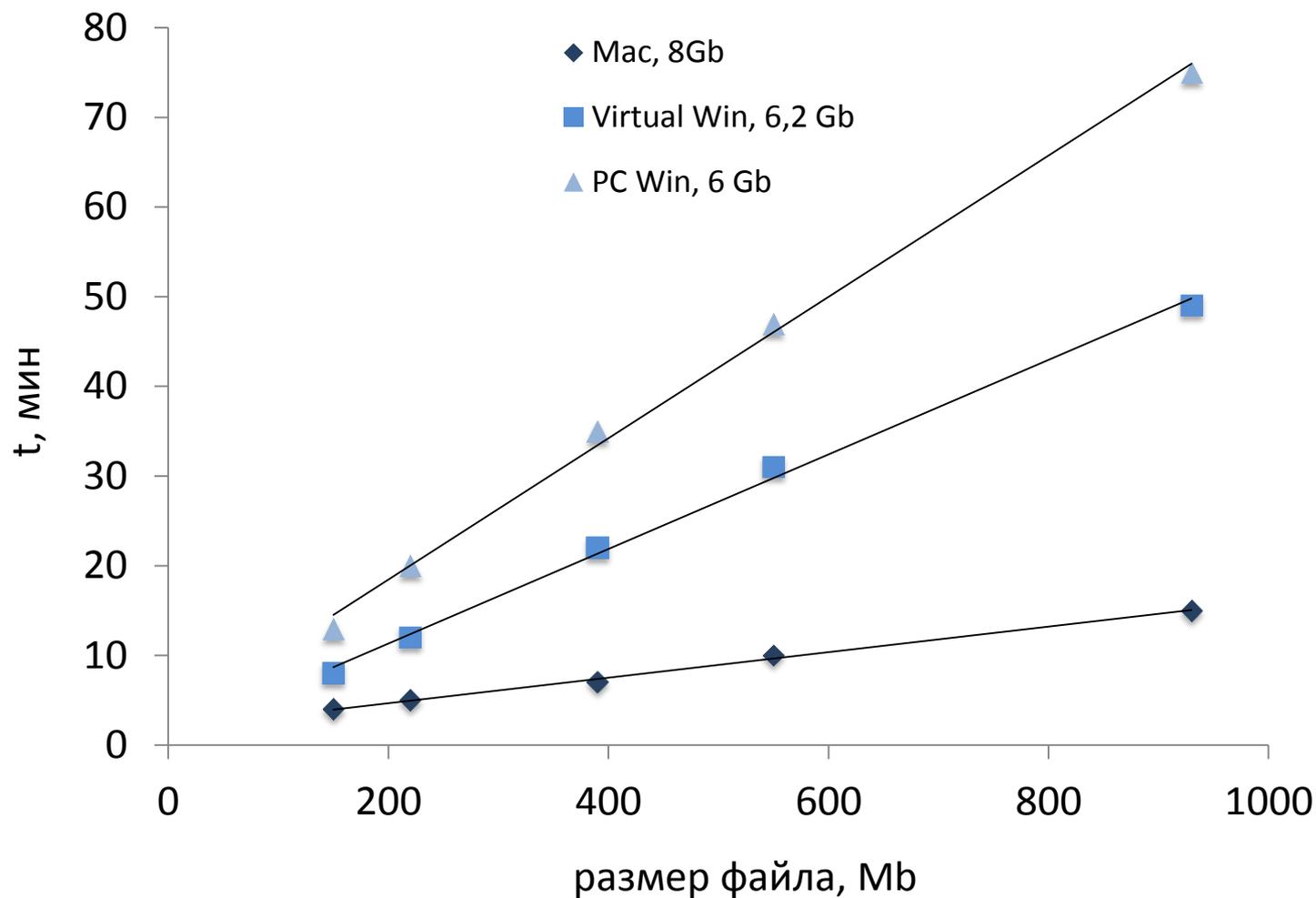


Canonical – 20 208 записей

Splice – 42 153 записей

TrEMBL – 129 478 записей

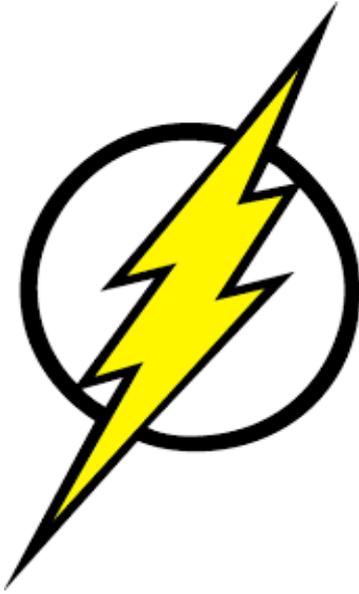
...и размера анализируемого спектра



Наблюдения

- Больше алгоритмов не значит лучше результаты
- Нечётное количество машин позволит избежать конфликт при сопоставлении хит-листов каждого из алгоритмов
- Комбинация разноплановых поисковых алгоритмов даёт лучшие результаты
- Java-приложение позволяет автоматизировать процесс интерпретации. Автоматизация сокращает время интерпретации на 5-10%

Dream team



X!Tandem

почти
непревзойденная
скорость



MS-GF+

быстрый и
эффективный
алгоритм



OMSSA

больше всего
идентификаций



Проанализировано 192 масс-спектра (114 Gb) за 11,5 часов ¹²

Спасибо!

