

Построение и анализ
эволюционного развития В-клеток
на основе данных иммуносеквенирования

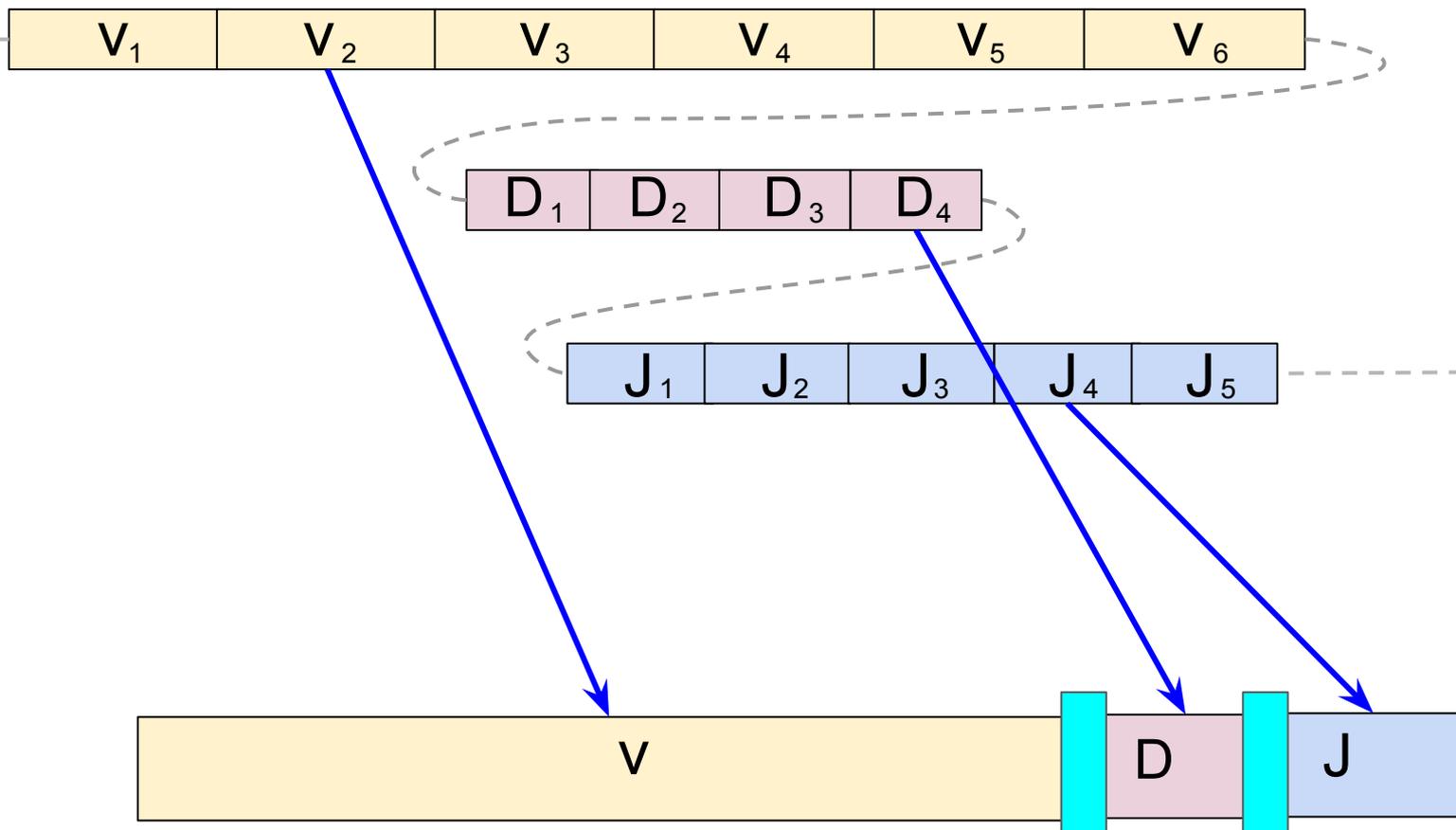
Андрей Слабодкин

научный руководитель: к.ф-м.н. Я. Ю. Сафонова

СПБАУ

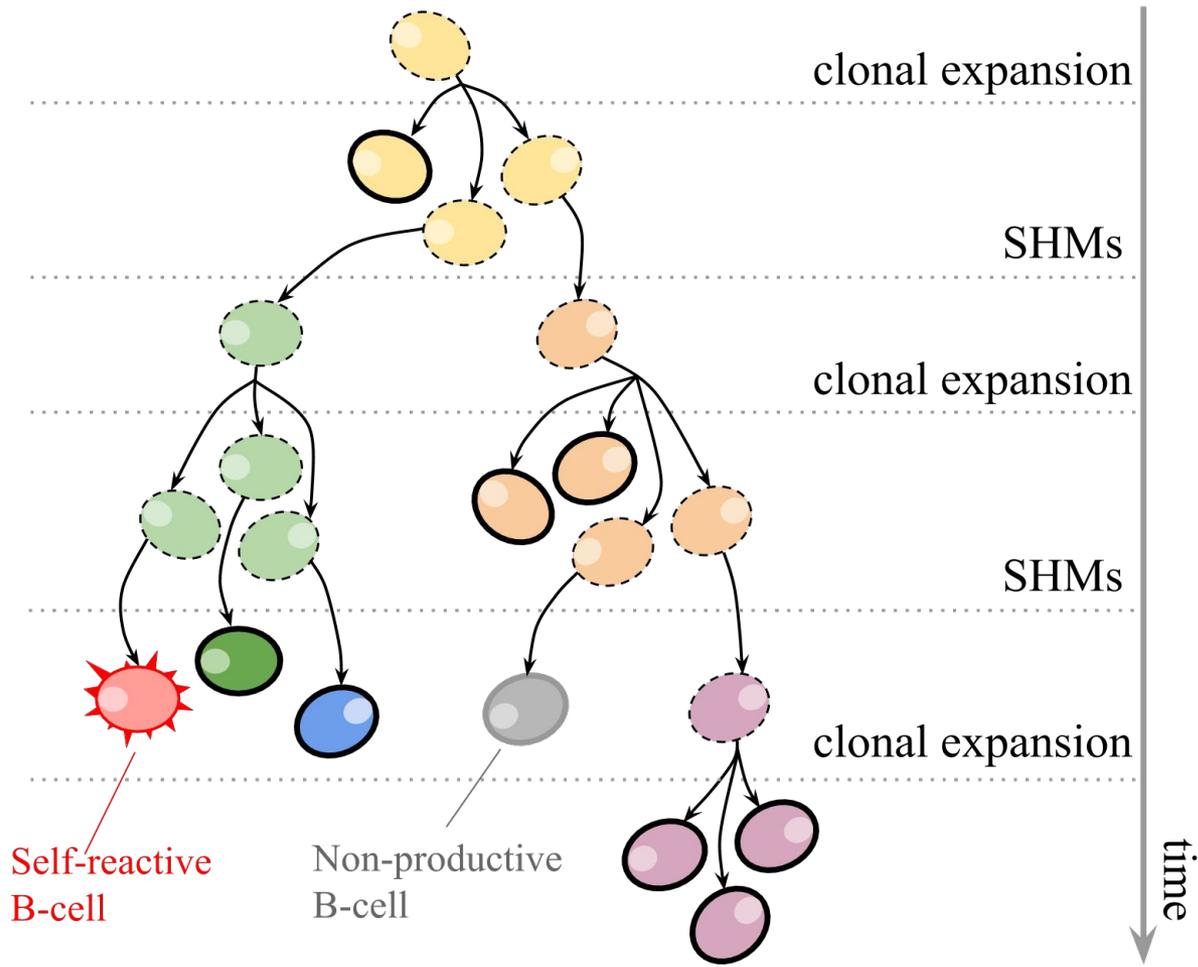
15 июня 2017 г.

VDJ-рекомбинация

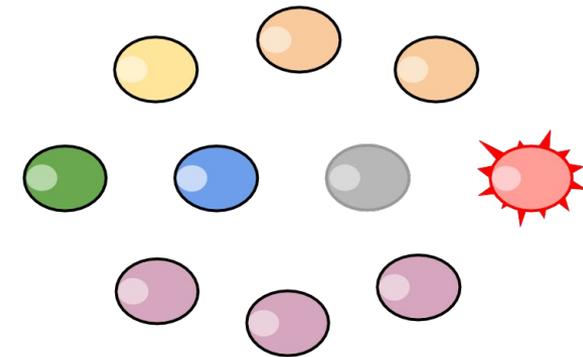


Клональные семейства и эволюционные деревья

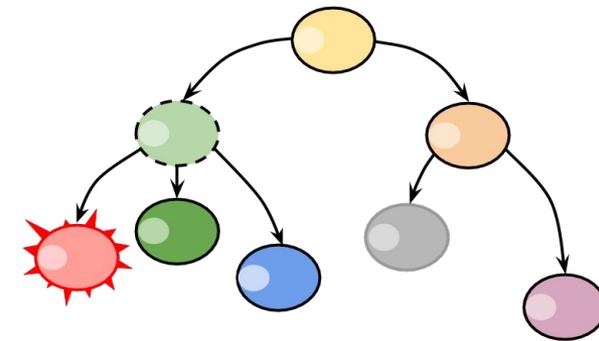
Secondary diversification



Clonal lineage



Clonal tree



- Филогенетика?

- Филогенетика
 - нет общего для всех предка
 - есть внутренние вершины
-
- Основная идея:
 - разбить на клональные семейства
 - применить филогенетический алгоритм внутри каждого семейства
 - решает только первую проблему

Цель:

- Научиться восстанавливать эволюционное развитие В-клеток по Rep-Seq данным.

Задачи:

- Научиться разбивать репертуар на клональные семейства
- Научиться восстанавливать рёбра эволюционного дерева внутри каждого семейства
- Разработать метрики для оценки результатов двух предыдущих пунктов.

Строение гена иммуноглобулина

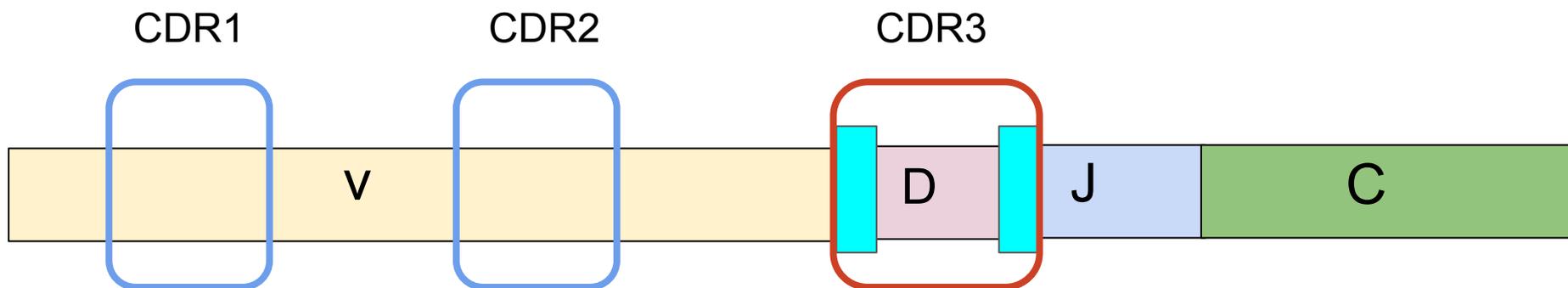


Схема работы алгоритма AntEvolvo

Выравнивание ридов на гермлайн

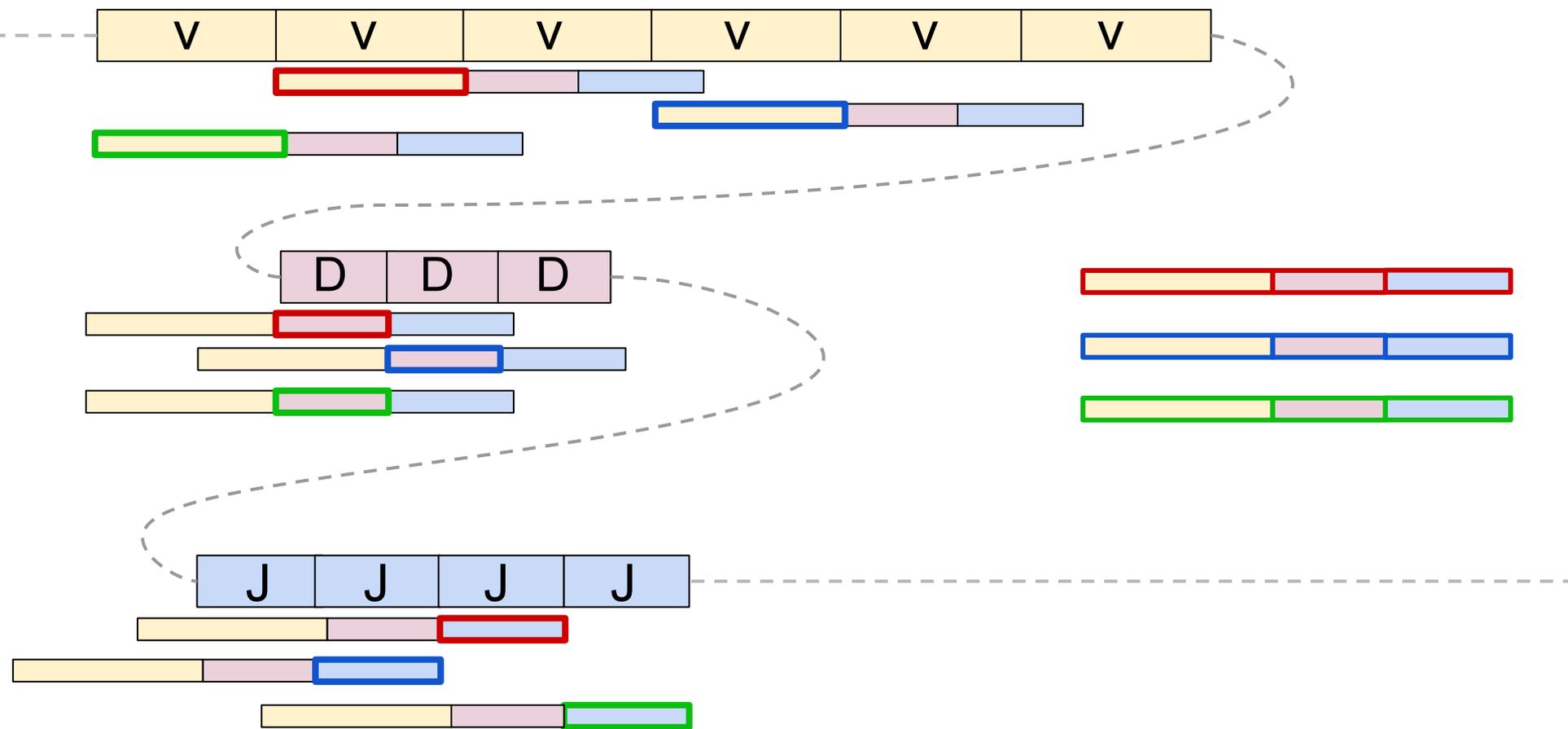
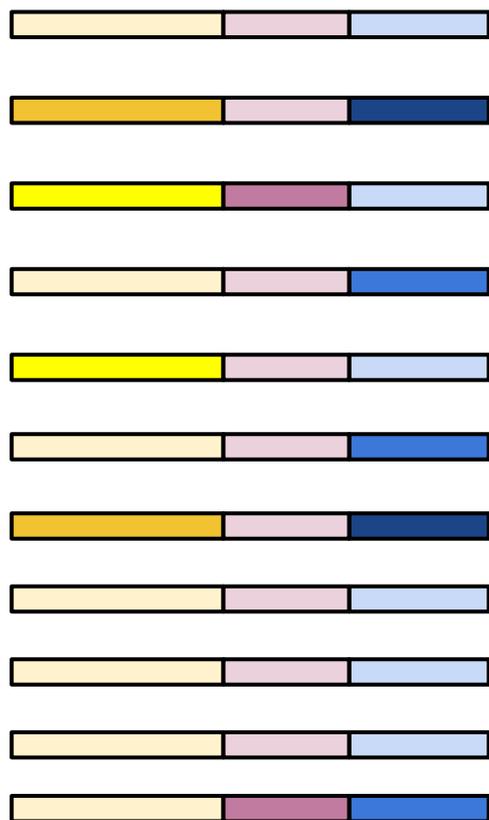


Схема работы алгоритма AntEvolvo

Выровненные на гермлайн
последовательности
вариабельных регионов



Разбиение на VJ-классы

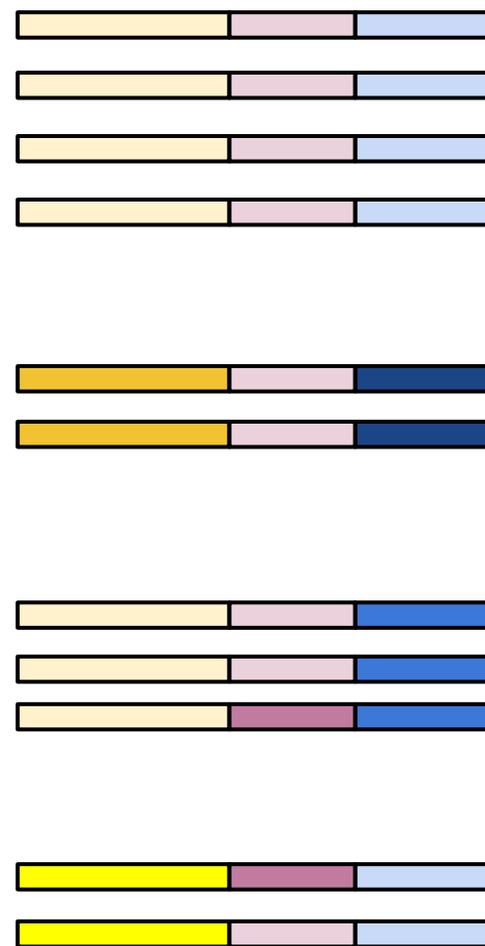
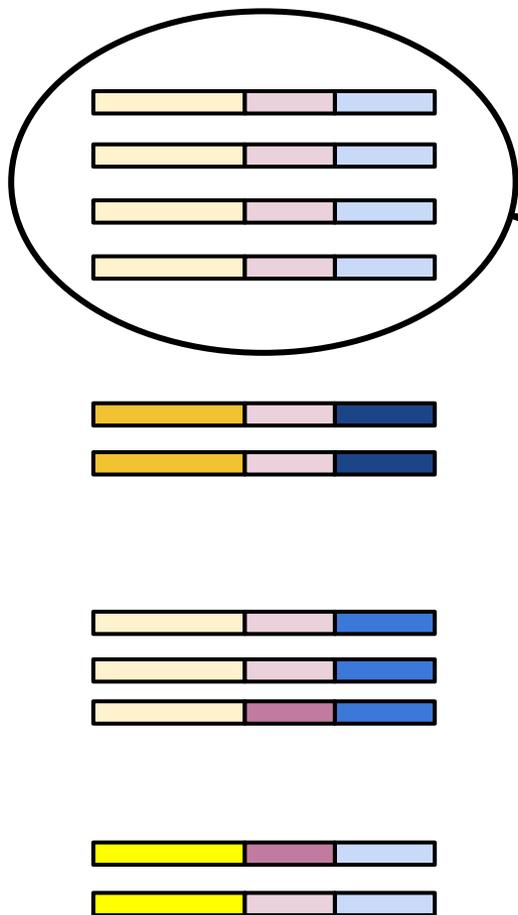


Схема работы алгоритма AntEvolvo

Разбиение на VJ-
классы



Внутри каждого класса
строится граф Хэмминга на
последовательностях CDR3

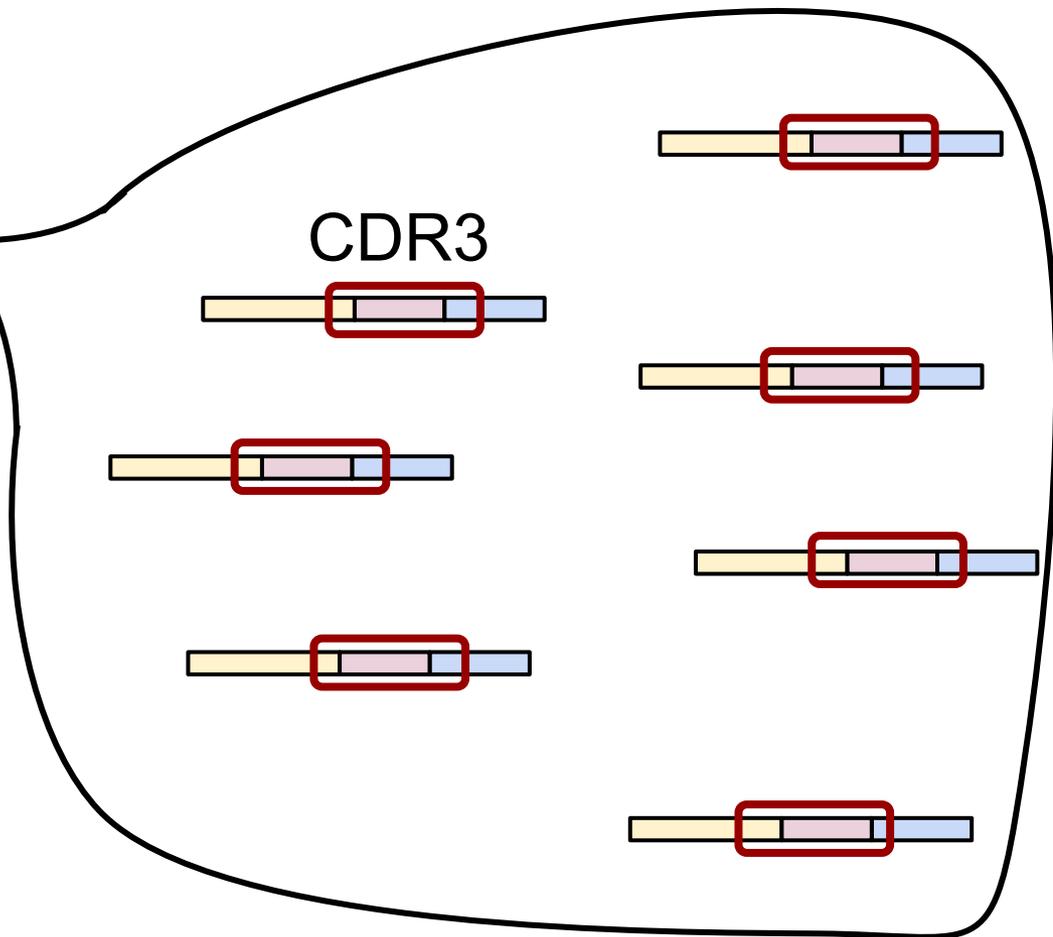
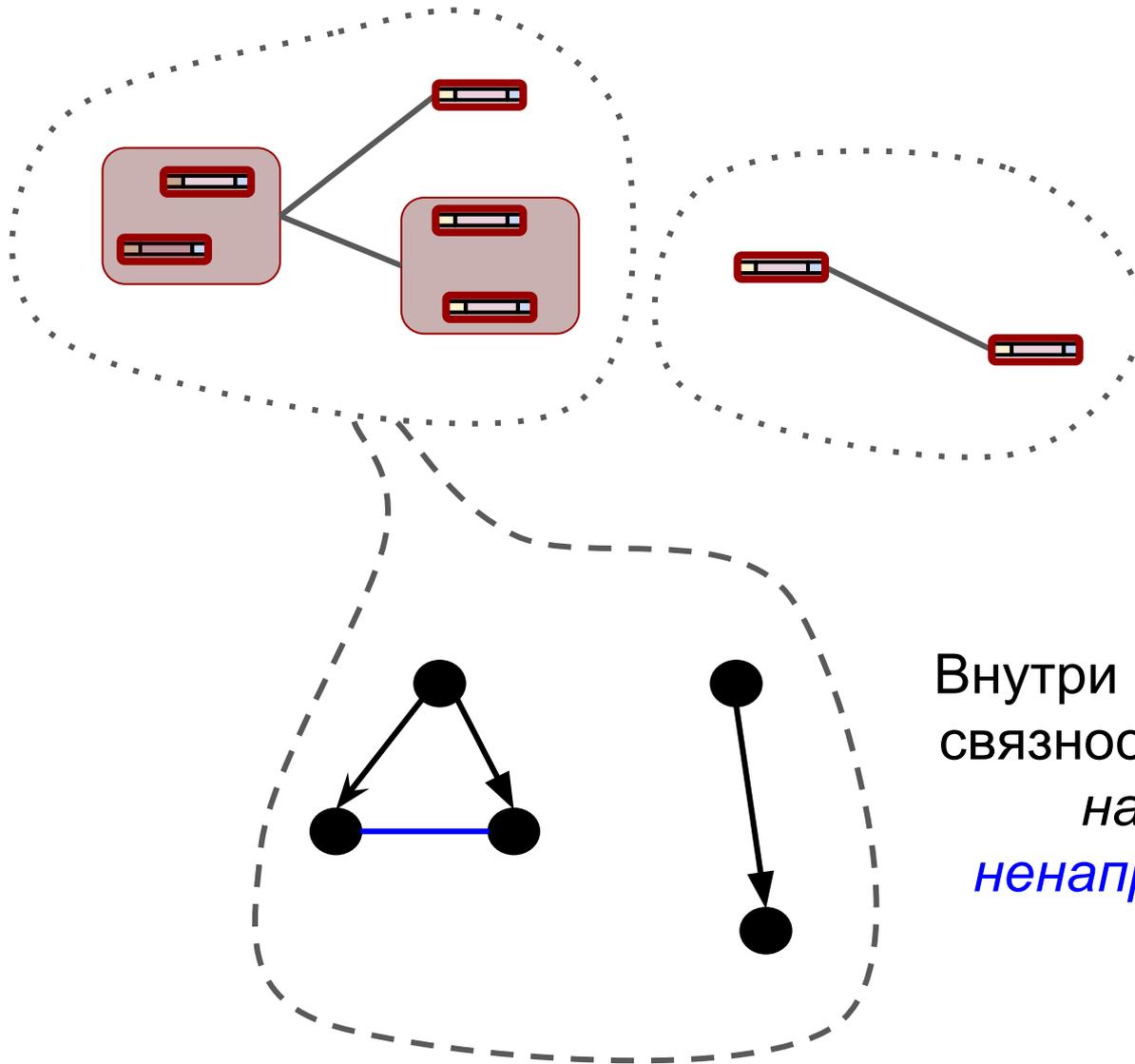


Схема работы алгоритма AntEvolvo

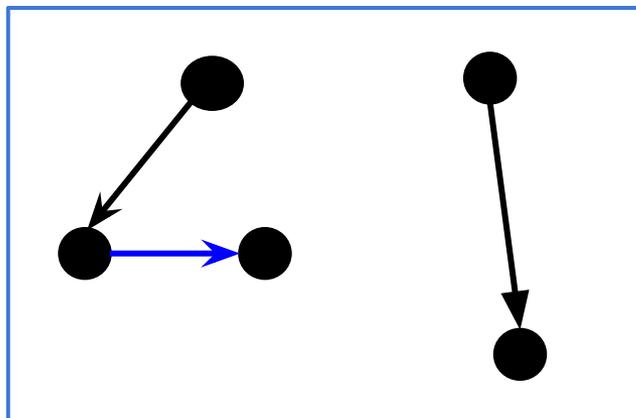
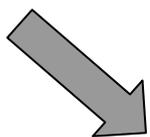
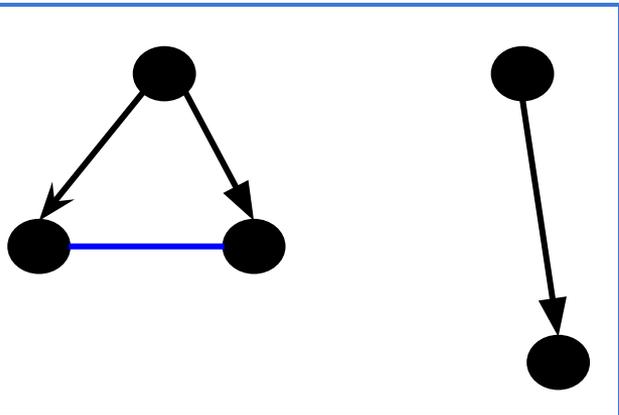


Граф Хэмминга
разбивается на
компоненты
связности

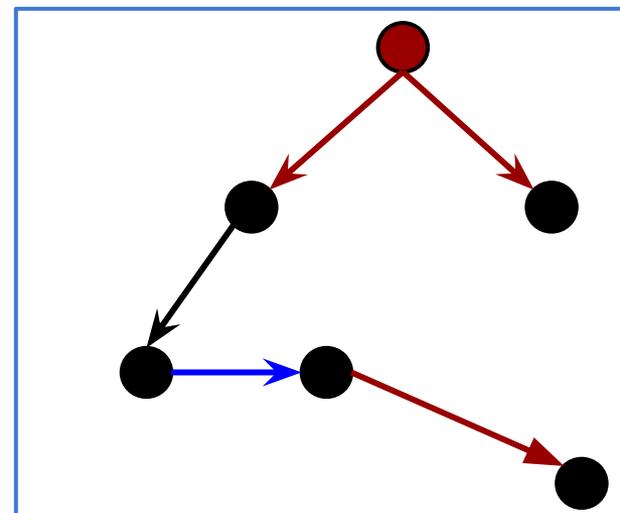
Внутри каждой компоненты
связности берется граф на
направленных и
ненаправленных рёбрах

Схема работы алгоритма AntEvolvo

Внутри каждой компоненты связности берется граф на *направленных* и *ненаправленных* рёбрах



В котором ищется optimum branching



После чего применяются эвристики, связанные с *родственными* рёбрами

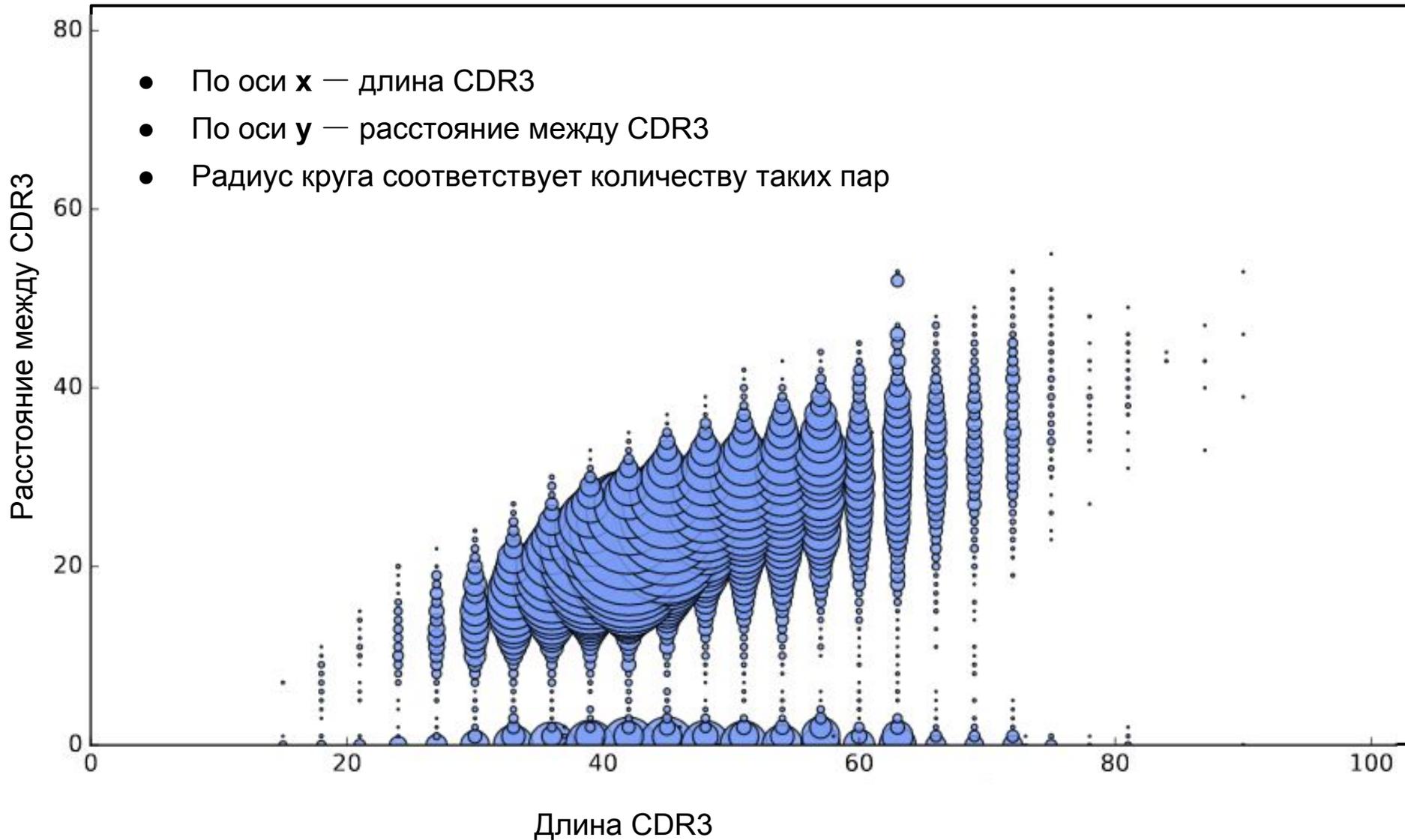
Рассмотрим все пары клонов внутри одного VJ-класса, такие что

- множества их проаннотированных мутаций совпадают либо вложены одно в другое
- их последовательности CDR3 имеют равную длину

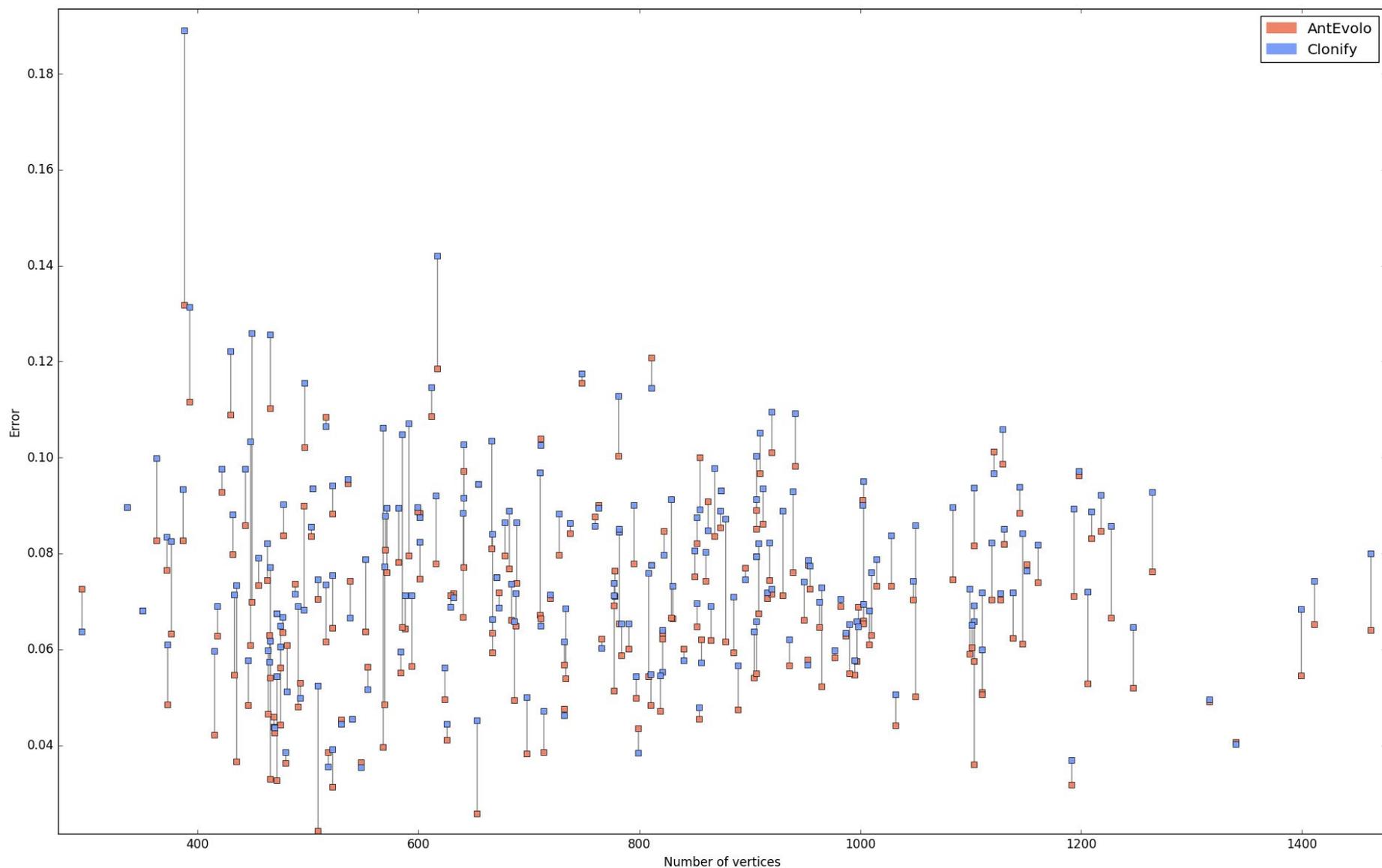
Изобразим эти пары на графике следующим образом:

- По оси x — длина CDR3
- По оси y — расстояние между CDR3
- Радиус круга соответствует количеству таких пар

Результаты

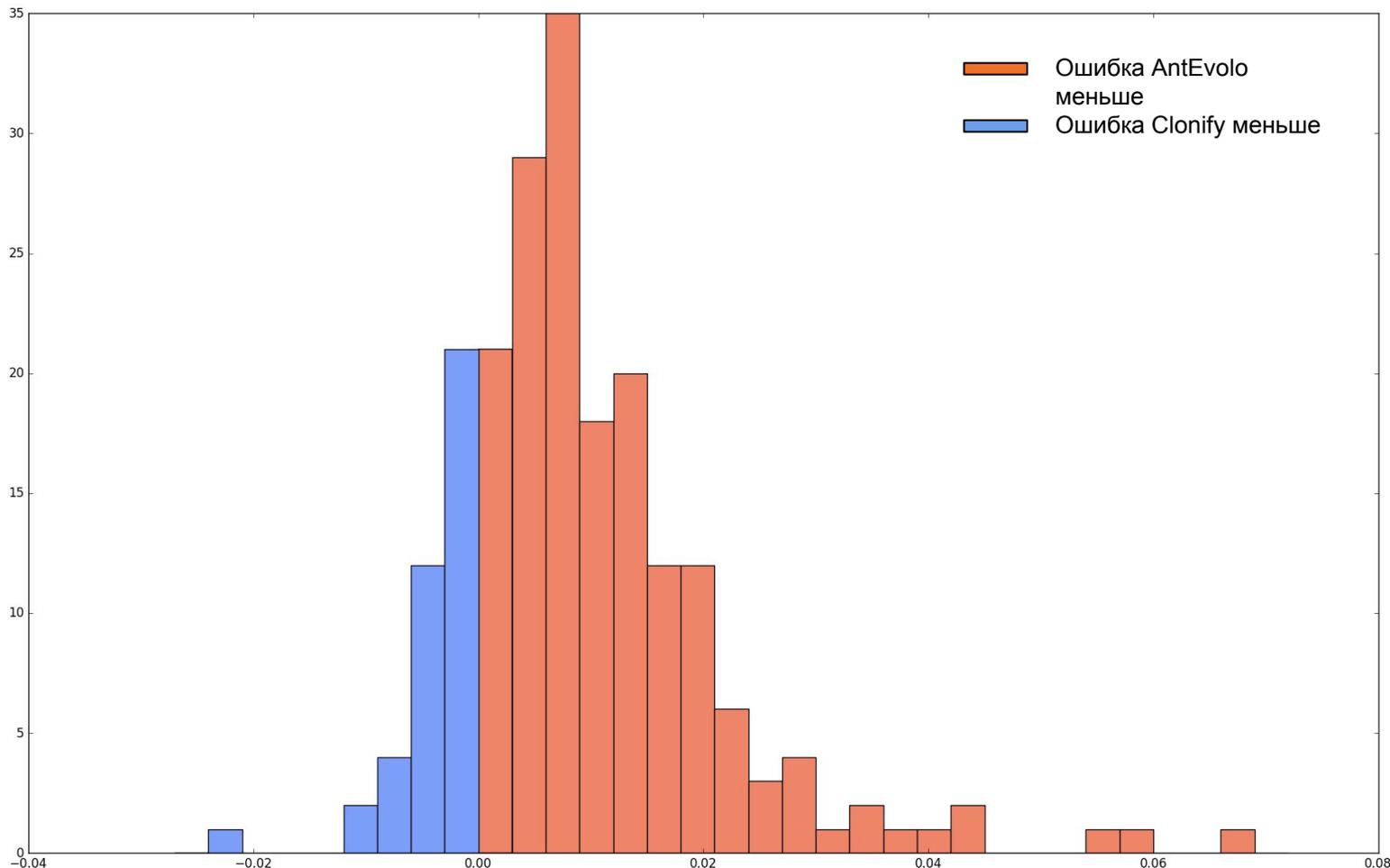


Результаты: сравнение на симулированных данных



Результаты: сравнение на симулированных данных

Гистограмма разности ошибок для предыдущего графика

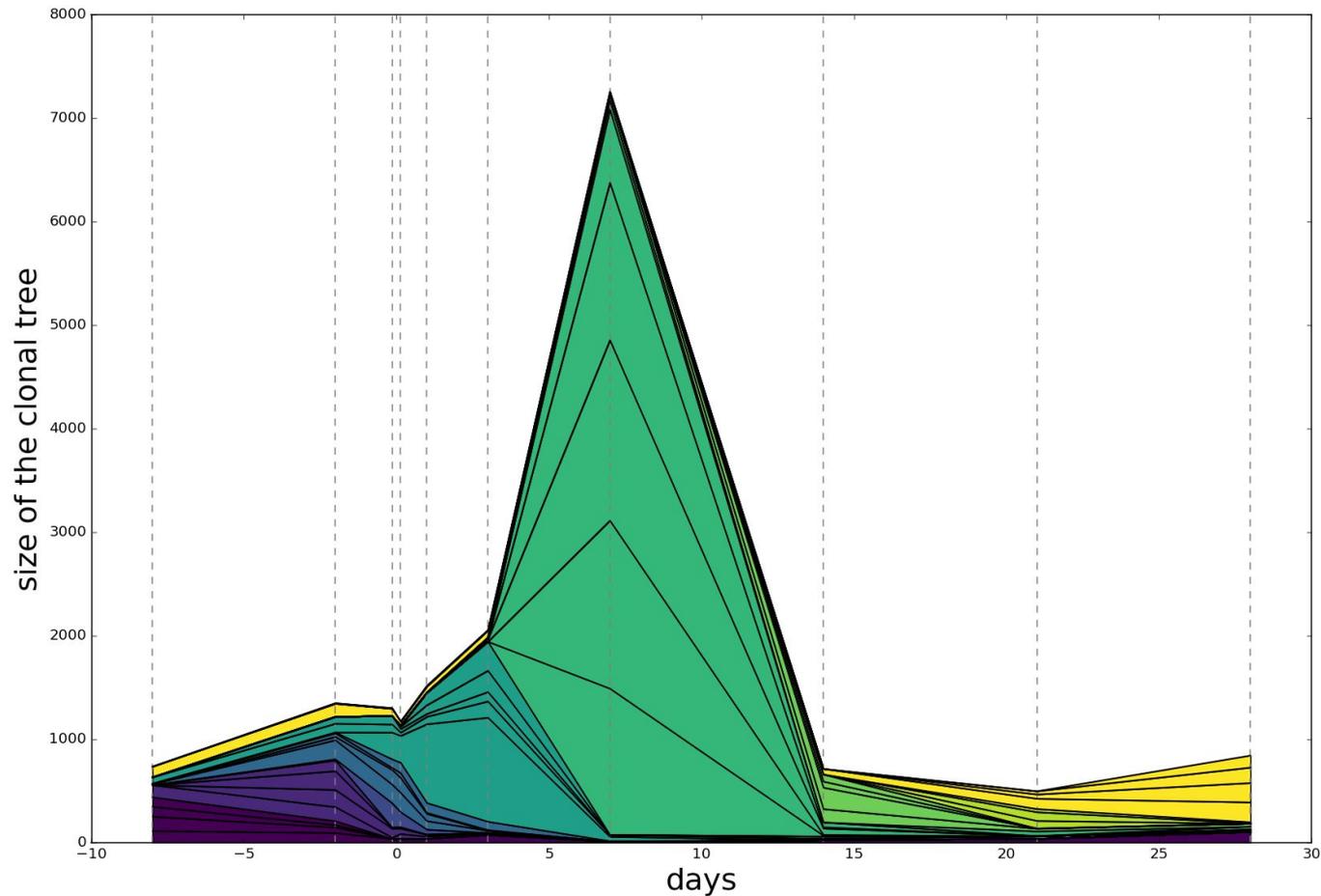


- Разработан алгоритм разбиения на клональные семейства, более эффективный, чем существующие аналоги.
- Разработан пригодный к использованию алгоритм восстановления клональных деревьев

Спасибо за внимание!

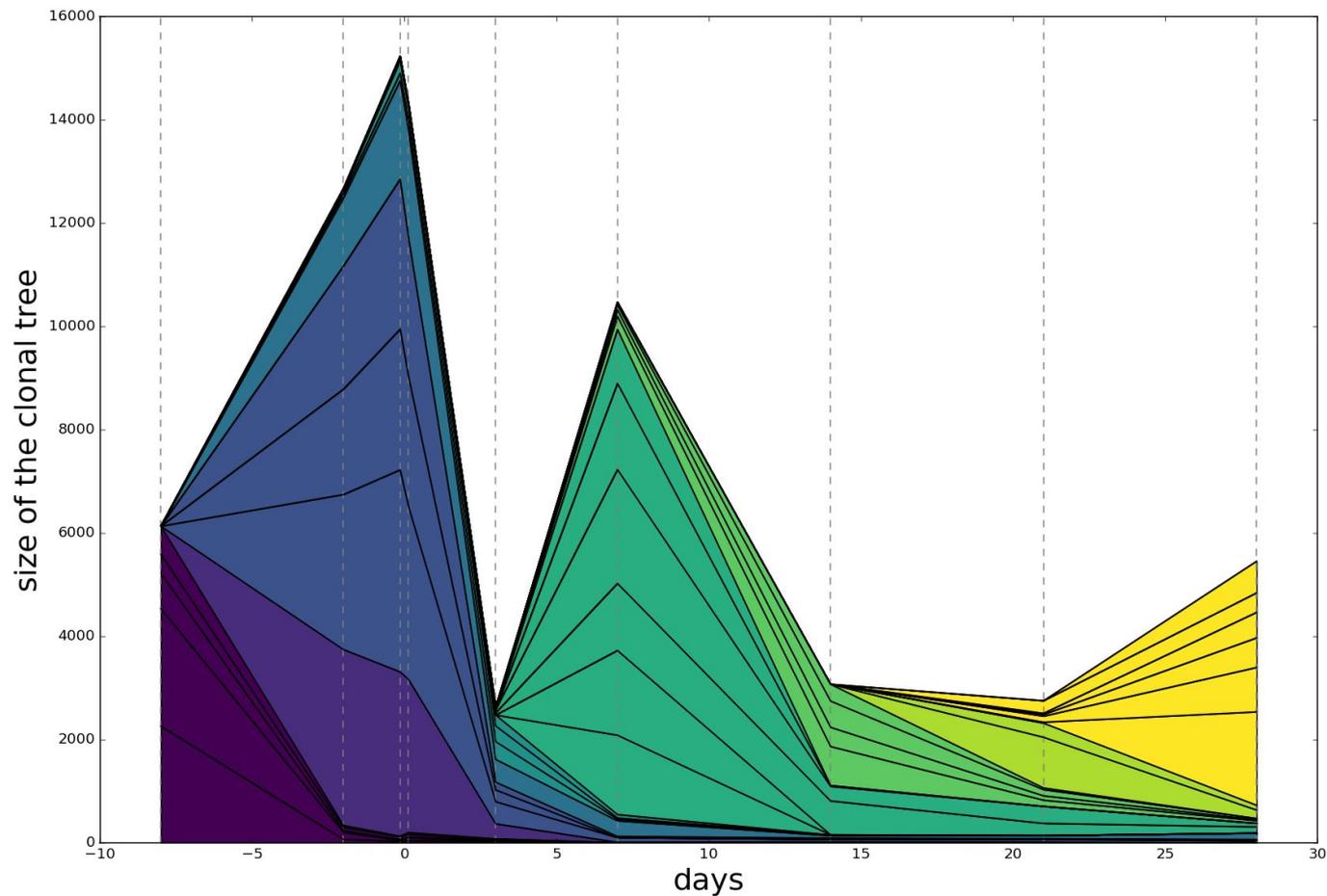
Способ использования инструмента

- Графики изображают изменение размеров пяти крупнейших клональных семейств для каждого из дней



Способ использования инструмента

- Графики изображают изменение размеров пяти крупнейших клональных семейств для каждого из дней



Способ использования инструмента

- Графики изображают изменение размеров пяти крупнейших клональных семейств для каждого из дней

