

QC-pipeline

Руководитель: А. Коробейников

Студент: К. Федоров

Задача

Удаление артефактов из ридов:

- 1) Adapters (Illumina, etc)
- 2) Cloning Vectors (pUC19, etc)
- 3) Contaminations (H.Sapiens 😊, etc)

Существующие решения

- Поиск векторов: `cross_match`
- Поиск адаптеров: `cutadapt`, `fastX clipper`

Основные подходы:

- Выравнивание с помощью `blast`
- Некоторые утилиты ищут только точные совпадения

Цель

- Хотим научиться искать искать адаптеры и вектора за приемлемое время
- Желательно комбинировать различные подходы: точный поиск, выравнивание и т.д. для ускорения работы
- В качестве источников данных о контаминациях используется база UniVec

Спасибо за внимание