

Сделал fastqc. Проблемы с сырыми ридами:

1. Низкое качество к концу рида
2. Адаптеры не удалены
3. Per base sequence content в начале и конце ридов плохой
4. Есть полиА на конце

С помощью cutadapt:

1. Удалил адаптеры
2. Обрезал полиА
3. Обрезал риды по качеству (<20)
4. Отбросил риды короче 50 пн

Все проблемы решились, per base sequence content в начале ридов не критичен.