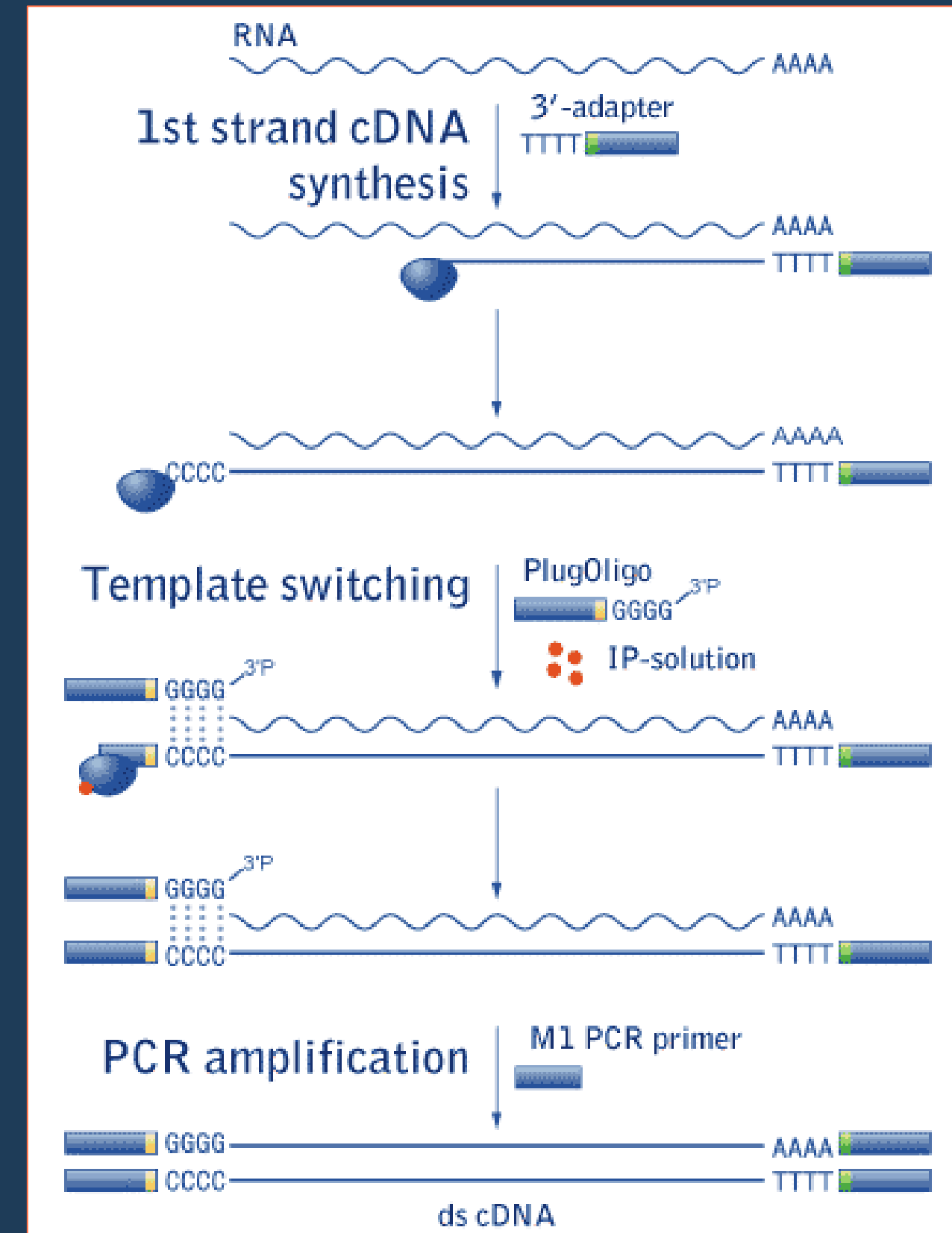
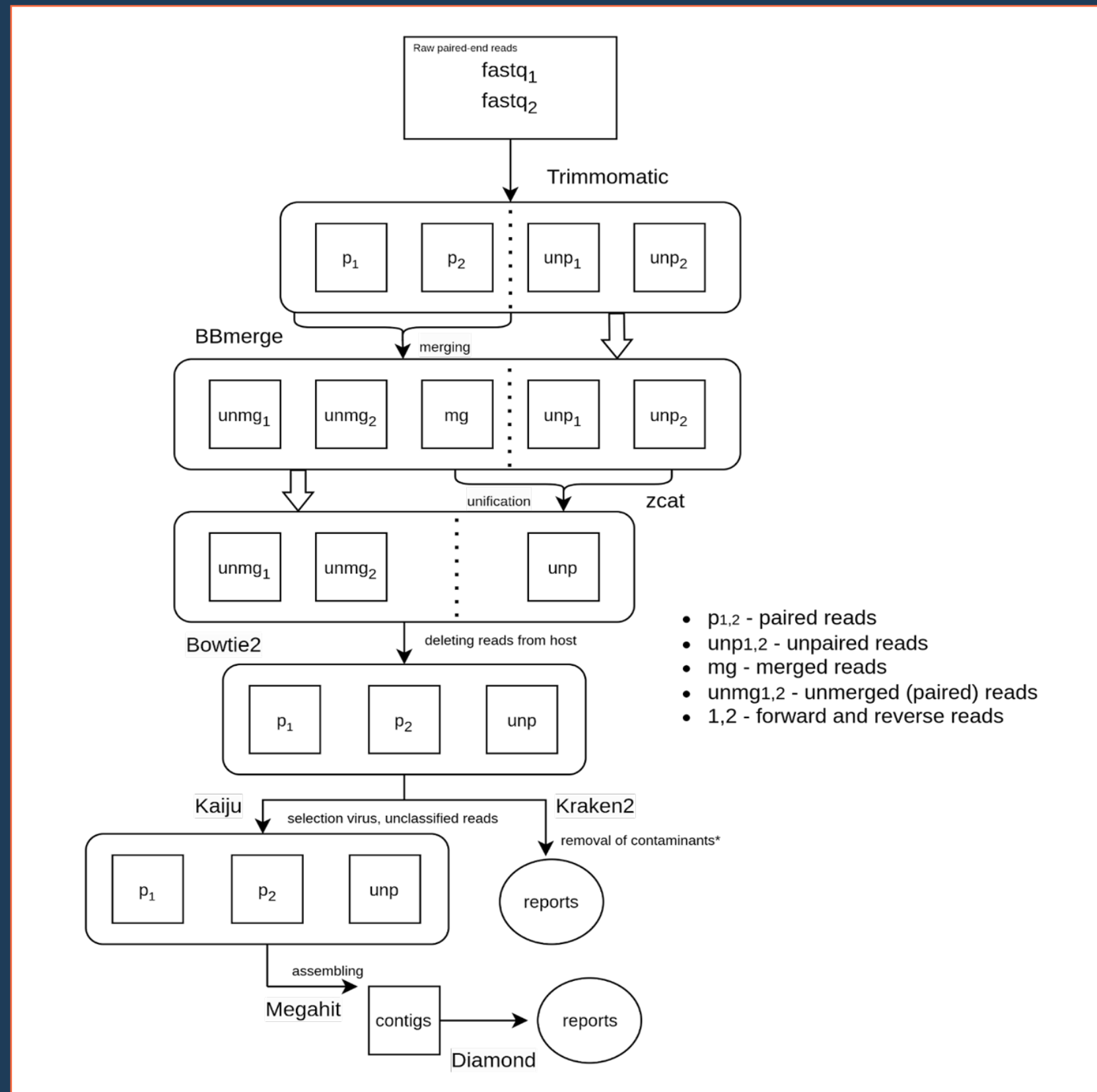


Введение

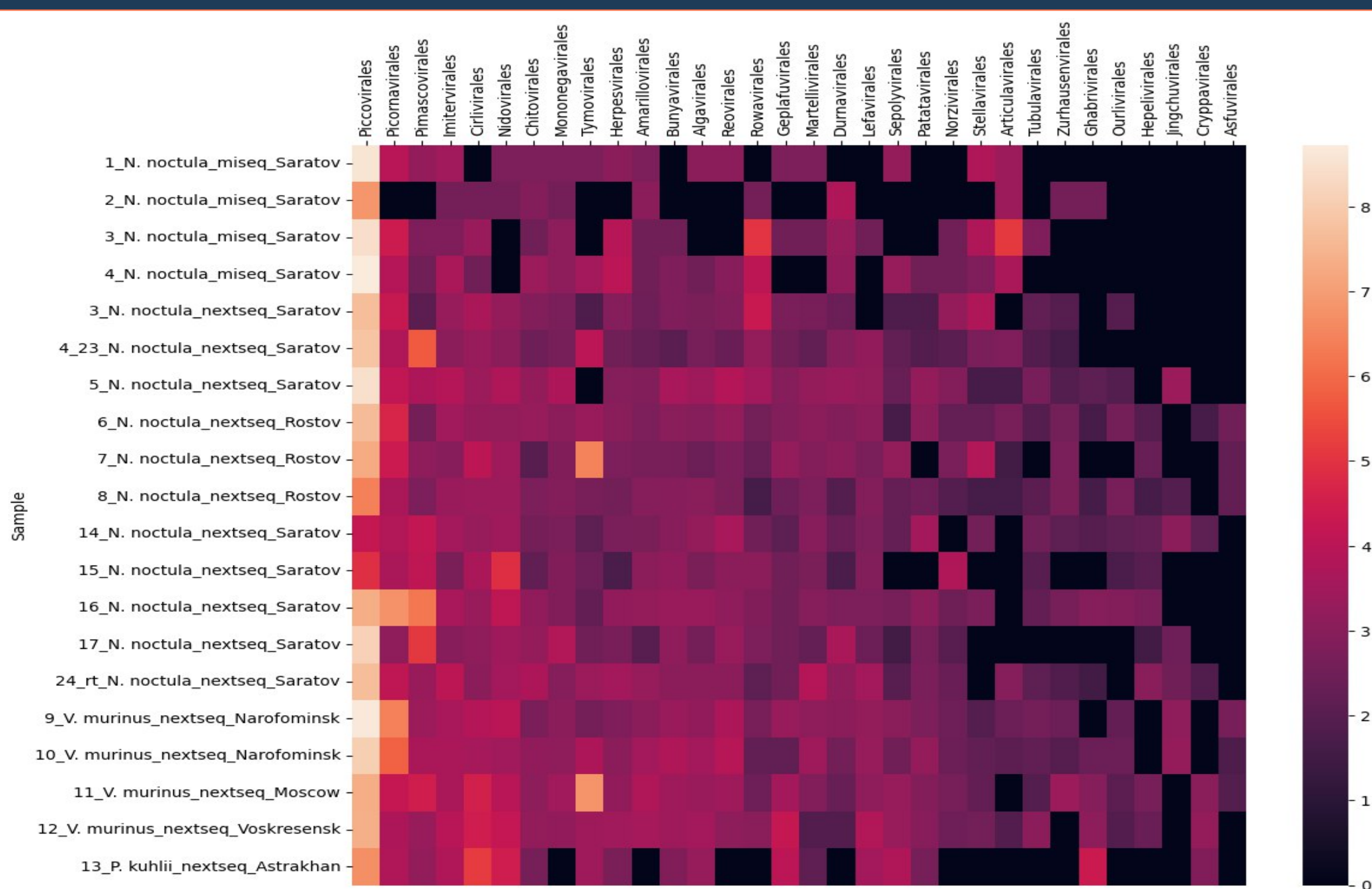
Нами было предложено использовать систему на основе подхода SMART (switching mechanism at the 5' end of the RNA transcript) для идентификации не изученных на сегодняшний день вирусов. При помощи биоинформатического анализа сначала потенциально возможно определить последовательности в данных NGS комплексных образцов, с высокой долей вероятности относящиеся к неизученным на сегодняшний день вирусам. Далее, с использованием единичных праймеров, специфичных к данным последовательностям, и подхода SMART, возможно секвенировать уже более протяженные фрагменты геномов таких вирусов.

Методы



В данном исследовании нами произведено метагеномное секвенирование образцов фекалий летучих мышей из различных регионов РФ. Затем был выполнен биоинформатический анализ, состоящий из нескольких этапов обработки данных. Для подбора праймера для применения методики SMART из полученных данных метагеномного секвенирования с помощью инструмента BLASTX был выделен контиг с низкой степенью гомологии с уже известными вирусами. Этот контиг был классифицирован как *Picornaviridae* sp. с E-value=0.0001 и идентичностью по аминокислотам 29,4 %. После этого продукт SMART был получен и отсеквенирован.

Результаты

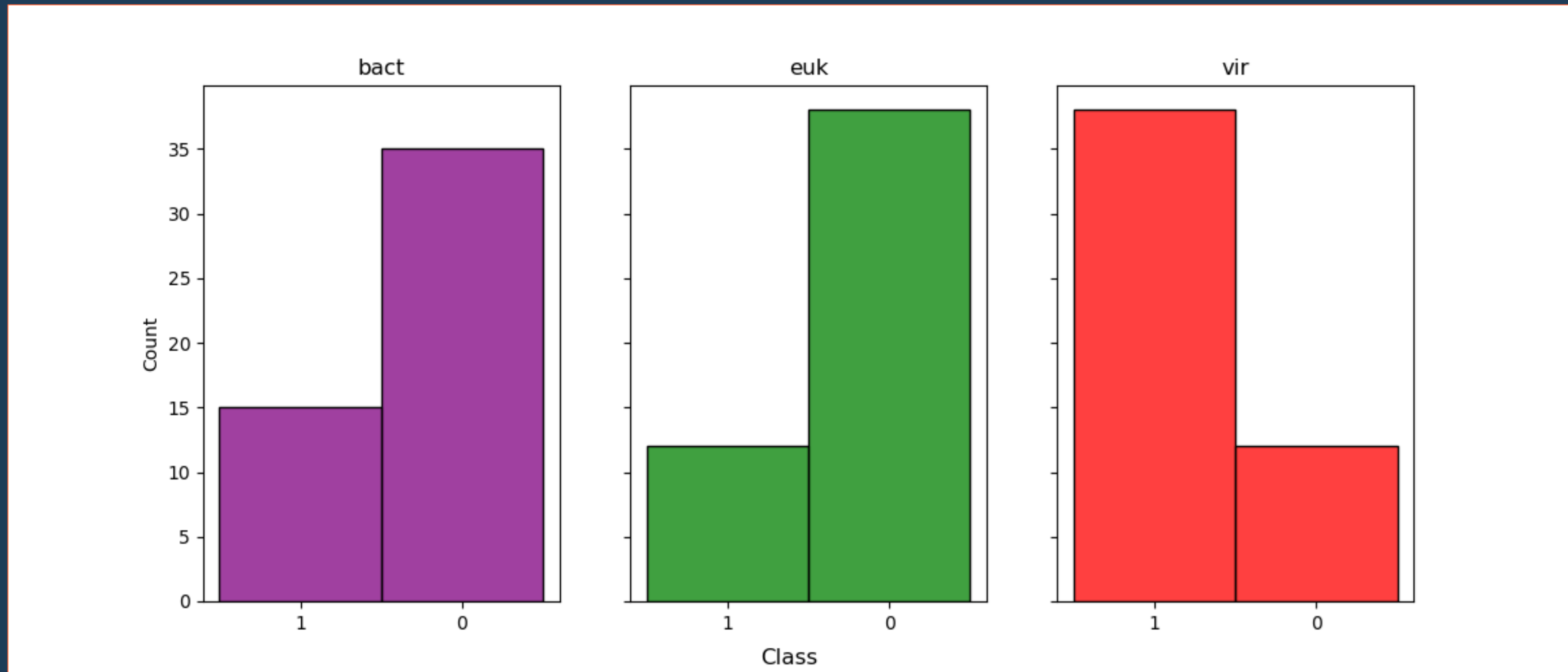
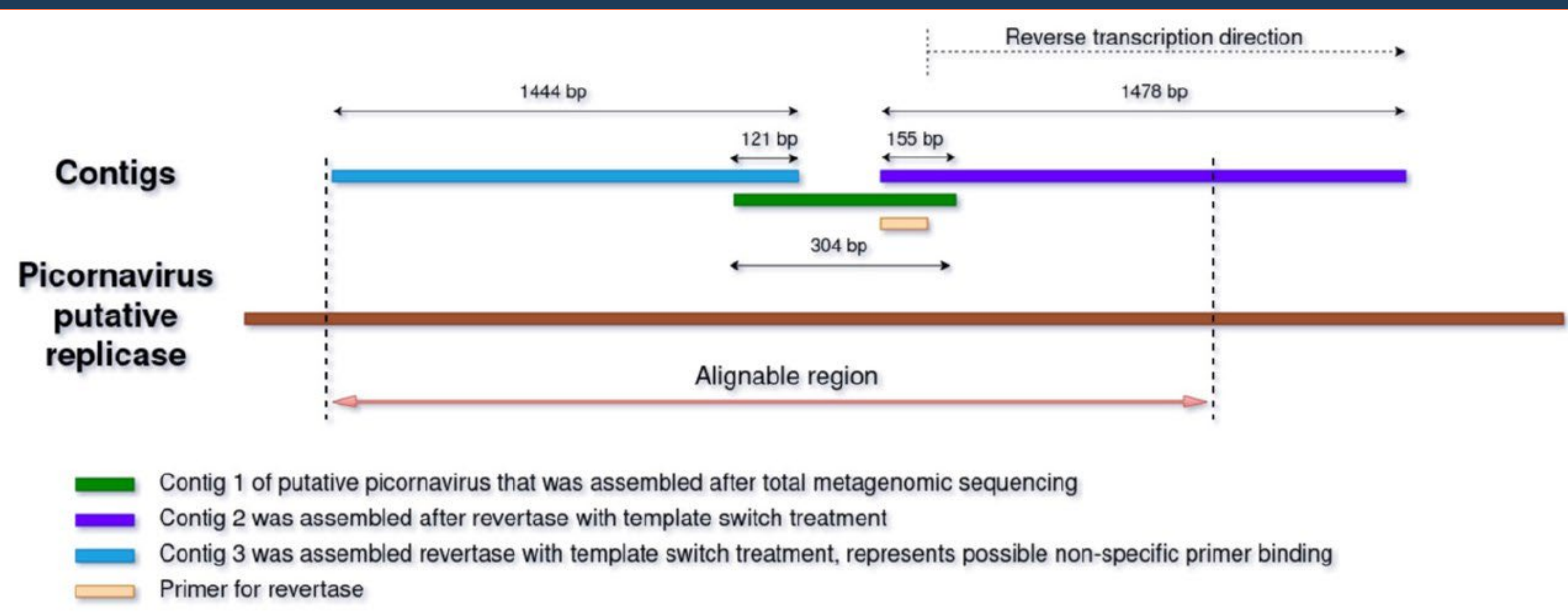
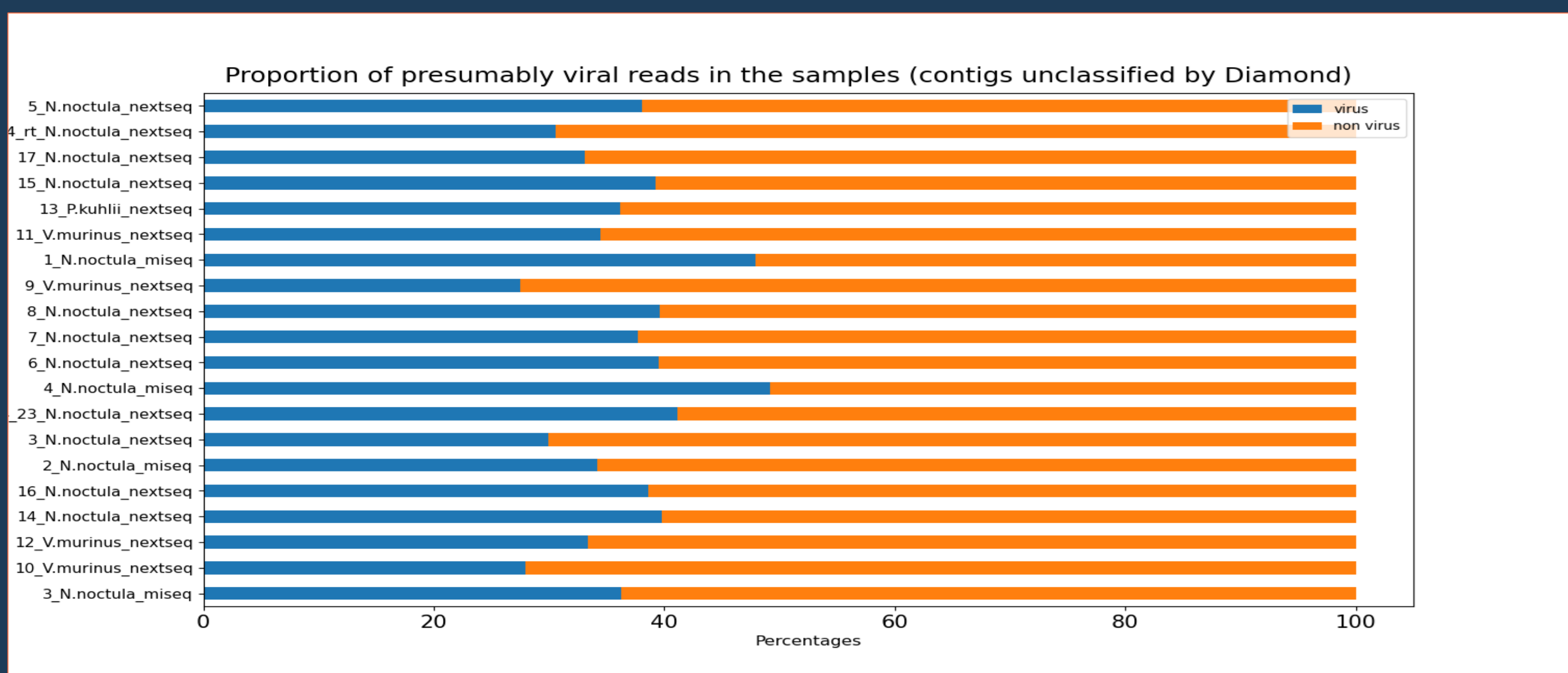


В данных метагеномного секвенирования были обнаружены прочтения, относящиеся к множеству вирусных таксонов (рис. слева). На рисунках ниже изображено картирование прочтений на полученный контиг SMART и взаимное расположение полученных контигов, праймеров для ревертазы и репликазы пикорнавируса.

На рисунках ниже представлены результаты тестирования программы DeepVirFinder.

Используя таксономический отчет программы Diamond, мы выбрали 50 случайных контигов каждого домена (эукариоты, бактерии и вирусы) для одного из образцов и подали их на вход нейронной сети.

Как можно заметить, обогащение наблюдается только для вирусного домена, что подтверждает адекватность программы. Затем мы подали ей на вход контиги, которые Diamond никак не классифицировал. Так примерно треть всех считываний для каждого образца может иметь вирусную природу, что, учитывая довольно высокую точность DeepVirFinder, указывает на то, что значительная часть метагеномных данных может содержать неизвестные до сих пор вирусы.



Выводы

В результате метагеномного секвенирования образцов фекалий летучих мышей был обнаружен широкий спектр вирусов, включая вирусы млекопитающих, насекомых и растений, что является вполне ожидаемым результатом, учитывая, что основной пищей летучих мышей являются насекомые. Ярким примером являются многочисленные чтения из семейства *Densovirinae*, большинство из которых, вероятно, принадлежат вирусам насекомых. Однако в данном исследовании мы также стремились подтвердить целесообразность использования методов амплификации нуклеиновых кислот с одиночных праймеров, таких как SMART. Это связано с тем, что значительное число прочтений (от 3 до 34%) не имеет существенной гомологии с уже известными геномами вирусов или других организмов. Для нашего примера мы выбрали прочтение, имеющее очень низкую гомологию с известным пикорнавирусом, и применили метод SMART, в результате чего был получен значительно более длинный фрагмент вирусного генома. Таким образом, мы продемонстрировали возможность метода SMART для амплификации и последующего секвенирования протяженных сегментов (новых) геномов РНК-вирусов. Использование нейросетевого подхода для детекции вирусных патогенов является перспективным направлением, так как многие новые вирусы не имеют заметной гомологии к уже известным.