Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека

**Федеральное казённое учреждение здравоохранения «Ростовский-на-Дону ордена Трудового Красного Знамени научно-исследовательский противочумный институт»** **Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека**

(ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора)

УТВЕРЖДАЮ

Врио директора ФКУЗ

Ростовский-на-Дону противочумный институт

Роспотребнадзора,

\_\_\_\_\_\_\_\_\_Н.Е. Гаевская

«\_11\_» \_\_\_06\_ 2024 г.

**Методические рекомендации**

**«Определение серогруппы у штаммов *Vibrio cholerae* на основе данных полногеномного секвенирования с помощью программы «*VibrioTyper*»**

Ростов-на-Дону

2024 г.

**Учреждение - разработчик:**

ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора

**Авторы:** Водопьянов А.С., Писанов Р.В., Водопьянов С.О.

**Уровень внедрения -** учрежденческий

**Область применения:** методические рекомендации предназначены для специалистов, занимающихся молекулярно-генетическим анализом штаммов холерного вибриона.

**Рассмотрены** на заседании Ученого совета ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора протокол № 6 от 11.06.2024 г.

**Аннотация**

Программа «*VibrioTyper*» предназначена для анализа данных полногеномного или фрагментарного секвенирования штаммов *Vibrio cholerae* с целью определения их серогруппы.

Определение серогруппы «*по фрагментам*» проводится на основе ранее выявленных и заложенных в программу коротких 40 нуклеотидных фрагментов, характерных для каждой из серогрупп. Расчет процента соответствия проводится по формуле 100 \* <количество выявленных фрагментов> / < количество фрагментов, характерных для данного серогруппы>.

Определение серогруппы «*по генам*» проводится на основе ранее составленного перечня всех возможных генов, входящих в О-кластер и «матрицы соответствия», содержащей данные о наличии различных генов у разных серогрупп.

На первом этапе в анализируемом геноме сначала проводится поиск генов с помощью алгоритма Смита-Ватермана с не менее чем 80% соответствием. На втором этапе выявляется серогруппа, имеющая аналогичный или максимально похожий набор генов.

**Установка и запуск программы**

«*VibrioTyper*» доступен для скачивания на сайте Ростовского противочумного института по адресу https://antiplague.ru/scientific-activity/publication/vibrio-typer/

Для запуска программы достаточно распаковать скачанный архив с программой и запустить файл *VibrioTyper.*bat

**Входные данные**

В качестве исходных данных используется файл в fasta-формате, содержащий либо набор контигов (протяженных нуклеотидных последовательностей – результат сборки *de novo* коротких ридов, полученных в результате полногеномного секвенирования), либо фрагмент генома, содержищий гены, кодирующие О-кластер. Предпочтительное расширение файла «.fasta» или «.fa».

Возможна пакетная обработка множества fasta-файлов (каждый файл содержит геном одного штамма)

**Работа с программой**

Внешний вид программы представлен на рисунке 1. Необходимо заполнить поле «*Название работы*» - эти данные будут использованы в качестве имени файла с результатом.

По умолчанию анализ проводится в режиме «*по генам*». При необходимости поиска и «*по фрагментам*» необходимо установить флажок «*Искать по фрагментам*».

При нажатии на кнопку «*Выбрать геномы и начать анализ*» открывается стандартное диалоговое окно выбора исходного файла (файлов) для анализа. После выбора анализ начинается автоматически. В процессе работы внизу окна программы индикатор в виде тонкой бегущей полосы отображает прогресс анализа текущего файла. Если анализируется несколько геномов, индикатор в виде широкой бегущей полосы отображает общий прогресс анализа. Ориентировочное время работы программы от 5 до 120 минут для каждого штамма (fasta-файла).

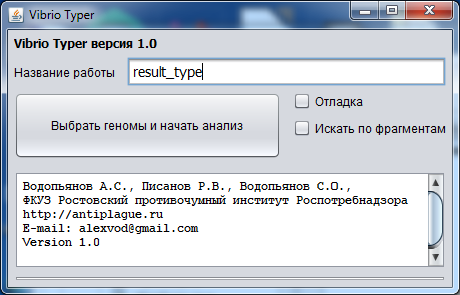


Рисунок 1 – внешний вид программы

В процессе работы получаемые данные отображаются в окне результата. После завершения работы программы результат сохраняется в текстовом файле, расположенном в папке с исходными fasta-файлами при этом имя файла совпадает с данными, введенными в поле «*Названием работы*». При копировании этого файла в текстовый редактор (или MS Excel, LibreOffice Calc) данные отображаются в табличной форме (Таблица 1).

Таблица 1 – Пример отображения результата

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Файл** | **Серогруппа по фрагментам** | **Процент** | **Близкие типы** | **Серогруппа по генам** | **Процент** |
| 13827 | не определено | 11% | O47(11%) O186(11%) O128(3%) | O47 | 98% |
| 13902 | не определено | 14% | O186(14%) O47(11%) O175(2%) | O47 | 98% |
| 13910 | O59 | 100% | O59(100%) O13(7%) O141(2%) | O59 | 99% |
| 13941 | O23 | 100% | O23(100%) O191(1%) O179(0%) | O23 | 99% |
| 14148 | O44 | 88% | O44(88%) O1 Inaba(3%) O57(1%) | O44 | 99% |

Результаты анализа каждого генома отображается в отдельной строке с указанием идентифицированной серогруппы «*по фрагментам*» (с указанием процента соответствия и трех наиболее близких типов) и серогруппы «*по генам*» (с указанием процента). Совпадение выявленной серогруппы «*по фрагментам*» и «*по генам*» свидетельствует о высокой достоверности результата.

**Обратная связь**

Все пожелания, замечания, выявленные ошибки можно направлять по электронной почте ([alexvod@gmail.com](mailto:alexvod@gmail.com)), телефону 8(863)240-22-66 или на почтовый адрес: 344002, г. Ростов-на-Дону, ул. М. Горького, 117/40, Водопьянову Алексею Сергеевичу.