

# Опыт использования технологии мономолекулярного нанопорового секвенирования при изучении штаммов бактерий, депонированных в государственной коллекции патогенных микроорганизмов и клеточных культур «ГКПМ-Оболенск»

В.И.Соломенцев, А.А.Сизова, Ю.П.Скрябин, А.А.Кисличкина, В.Б.Фролов,  
Н.В.Майская, С.А.Иванов, С.В.Дентовская, А.П.Анисимов, А.Г.Богун

ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии»,  
Оболенск, Российская Федерация

Одним из наиболее информативных методов изучения бактерий является полногеномное секвенирование. Мономолекулярное секвенирование является относительно новым и одним из наиболее перспективных способов исследования геномов микроорганизмов. Данная технология позволяет получать прочтения большой длины, что облегчает реконструкцию полных геномов и позволяет изучать любые их особенности.

Принцип работы секвенатора MiniON производства Oxford Nanopore Technologies основан на измерении силы тока, протекающего через нанопору при прохождении через нее молекулы ДНК. Данный секвенатор обладает компактными размерами, стоимость одного раунда секвенирования начинается от 500 долларов США, само оборудование поставляется бесплатно в составе стартового набора. За один раунд секвенирования возможен анализ до 12 бактериальных геномов. Оборудование позволяет проводить прямой анализ РНК без этапа обратной транскрипции. Все это создает большие перспективы для использования секвенатора MiniON в микробиологических лабораториях.

В данной работе мы описываем опыт эксплуатации данного оборудования в отделе коллекционных культур ФБУН ГНЦ ПМБ, а также полученные нами результаты.

*Ключевые слова:* методы, нанопоровое секвенирование

**Для цитирования:** Соломенцев В.И., Сизова А.А., Скрябин Ю.П., Кисличкина А.А., Фролов В.Б., Майская Н.В., Иванов С.А., Дентовская С.В., Анисимов А.П., Богун А.Г. Опыт использования технологии мономолекулярного нанопорового секвенирования при изучении штаммов бактерий, депонированных в государственной коллекции патогенных микроорганизмов и клеточных культур «ГКПМ-Оболенск». Бактериология. 2018; 3(4): 60–68. DOI: 10.20953/2500-1027-2018-4-60-68

## Experience of using the technology of monomolecular nanopore sequence for the study of bacterium strains deponated in the state collection of pathogenic microorganisms and cell cultures “SCPM-Obolensk”

V.I.Solomentsev, A.A.Sizova, Yu.P.Scriabin, A.A.Kislichkina, V.B.Frolov,  
N.V.Mayskaya, S.A.Ivanov, S.V.Dentovskaya, A.P.Anisimov, A.G.Bogun

State Research Center for Applied Microbiology and Biotechnology, Obolensk, Russian Federation

One of the most informative methods of bacterial research is complete genome sequencing. Monomolecular sequencing is a relatively new and most promising way to elucidate the genome of a microorganism. The technology provides genomic reading of the longer length, facilitating the reconstruction of complete genomes to study any of their features. A sequencer MiniON (Oxford Nanopore Technologies) measures current flowing through a nanopore with a DNA molecule passing through it. The sequencer is compact, with the cost of one sequence cycle being \$ 500 minimum. The equipment itself is supplied free of charge as a part of the starter kit. Twelve bacterial genomes are possible to analyze in one sequencing round. The direct

### Для корреспонденции:

Соломенцев Виктор Иванович, младший научный сотрудник отдела коллекционных культур ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора

Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ

Телефон: (4967) 36-0000

E-mail: solomentsev@obolensk.org

Статья поступила 07.10.2018 г., принята к печати 25.12.2018 г.

### For correspondence:

Viktor I. Solomentsev, junior research scientist of the collection cultures department, State Research Center for Applied Microbiology and Biotechnology

Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation

Phone: (4967) 36-0000

E-mail: solomentsev@obolensk.org

The article was received 07.10.2018, accepted for publication 25.12.2018

analysis of the RNA, avoiding the reverse transcription step, is possible. All these make the sequencer MiniON promising for use by microbiological laboratories. The paper deals with the operating experience of the equipment by the Department of collected cultures at the State Research Center for Applied Microbiology and Biotechnology, as well as with results obtained.

*Keywords: methods, nanopore sequencing*

**For citation:** Solomentsev V.I., Sizova A.A., Scriabin Yu.P., Kislichkina A.A., Frolov V.B., Mayskaya N.V., Ivanov S.A., Dentovskaya S.V., Anisimov A.P., Bogun A.G.. Experience of using the technology of monomolecular nanopore sequence for the study of bacterium strains deposited in the state collection of pathogenic microorganisms and cell cultures "SCPM-Obolensk". *Bacteriology*. 2018; 3(4): 60–68. (In Russian). DOI: 10.20953/2500-1027-2018-4-60-68

**Д**ля исследования геномов микроорганизмов широко используют такие методы, как полимеразная цепная реакция, анализ варибельности числа tandemных повторов, мультилокусное сиквенс-типирование и другие. Наиболее информативным методом изучения геномов бактерий является полногеномное секвенирование. На сегодняшний день широкое распространение получило массивированное параллельное секвенирование (полногеномное секвенирование). Данный подход основан на изучении структуры всего генома микроорганизма. Современные системы позволяют проводить одновременное определение геномов нескольких десятков, а в некоторых случаях сотен бактерий. Существенным недостатком большинства систем для секвенирования является высокая стоимость оборудования, составляющая десятки миллионов рублей.

Существует несколько платформ массивированного параллельного секвенирования. В основе их работы лежат общие принципы – на первом этапе создается библиотека последовательностей ДНК, к которым затем прикрепляются адаптерные последовательности, на втором этапе с помощью ПЦР происходит клональная амплификация каждой молекулы, в результате чего формируются кластеры одинаковых молекул. Полученные кластеры используются как образцы при определении нуклеотидной последовательности с применением оптических или химических методов детекции [1].

Технологии NGS (Next Generation Sequencing) имеют ряд недостатков, к которым относятся малая длина единичных прочтений, ошибки в гомополимерных участках, сложности разрешения числа повторов, зависимость точности прочтения от ГЦ-состава фрагмента ДНК и некоторые другие. В качестве альтернативы все большее распространение получают системы мономолекулярного анализа нуклеиновых кислот, которые часто называют технологиями секвенирования третьего поколения. В настоящее время коммерчески доступными являются системы одномолекулярного секвенирования в реальном времени производства компании PacBio и нанопоровое секвенирование на платформе Oxford Nanopore Technologies (ONT).

Принцип секвенирования PacBio основан на изучении скорости включения в состав синтезируемой молекулы ДНК флуоресцентно меченных нуклеотидов с использованием высокочувствительной оптической системы [2]. Принцип действия секвенаторов ONT основан на измерении электрического тока при прохождении молекулы нуклеиновой кислоты через нанопору. Детекция осуществляется в камере с разделенной мембраной, содержащей нанопору. К камере подается напряжение, вызывающее движение ДНК или РНК через пору. При прохождении молекулы сечение поры уменьшается, в результате чего сила тока снижается. Таким

образом, измеряя силу тока, можно определять тип нуклеотида, проходящего через пору в данный отрезок времени.

Оба подхода позволяют получать единичные прочтения очень большой длины [3]. Длина единичных прочтений на платформе PacBio составляет более 80 тысяч оснований. Для ONT типичная длина единичного прочтения составляет более 10 тысяч оснований. На данный момент удалось экспериментально получить непрерывные прочтения длиной более 1 000 000 оснований. Оборудование ONT значительно дешевле (стоимость секвенатора MiniON с набором реактивов начинается от 1000\$) в сравнении с приборами Sequell (350 000\$) и PacBio RSII (700 000\$) производства Pacific Biosciences. Все приборы генерируют объемы данных, достаточные для анализа нескольких бактериальных геномов. Максимальная производительность MiniON составляет 20 000 000 000 оснований, PromethION – до 6 000 000 000 000 оснований при использовании 48 проточных ячеек. Производительность системы для PacBio Sequell составляет 10 000 000 000 оснований. Системы мономолекулярного нанопорового секвенирования несколько уступают по суммарному времени анализа (более 48 ч для ONT против 4 ч для PacBio). Однако первые данные нанопорового секвенирования становятся доступными уже спустя несколько минут после начала анализа. Системы имеют близкую стоимость проведения одного раунда секвенирования – запуск MiniON оценивается в 475–900 долларов США, для PromethION стоимость составляет 600–2000 долларов. Стоимость секвенирования на системе Sequell составляет 400 долларов США и 850 долларов США за запуск для PacBio RSII. Однако оборудование производства компании Pacific Biosciences в настоящее время недоступно на территории нашей страны.

В 2017 г. в России стало коммерчески доступным оборудование компании Oxford Nanopore Technologies – производителя систем мономолекулярного нанопорового секвенирования. Стоимость стартового набора для проведения исследований, в который входит секвенатор MiniON, а также набор реактивов, на территории Российской Федерации составляет немногим более 100 000 рублей. В связи с этим данное оборудование может быть доступным для оснащения широкого круга микробиологических, в том числе и бактериологических, лабораторий. В данной статье нами описывается опыт эксплуатации оборудования MiniON в отделе коллекционных культур ФБУН ГНЦ ПМБ.

## Материалы и методы

**Оборудование для секвенирования и обработки данных**  
Работа выполнялась на секвенаторе MiniON (Oxford Nanopore Technologies). Для управления секвенатором ис-

пользовали портативный компьютер HP Probook 450 G5, с процессором core i7-8550U и HDD жестким диском объемом 1 Тб. На данную модель ноутбука дополнительно установили 8 Гб оперативной памяти и жесткий диск объемом 1 Тб типа SSD с высокой скоростью записи.

Обработка данных осуществлялась на сервере Z-800 производства Hewlett-Packard с двумя 12-ядерными процессорами Intel® Xeon(R) CPU X5675 @ 3.07GHz, поддерживающими технологию Hyper-threading (всего 24 потока), объемом оперативной памяти 96 Гб типа DDR3.

#### **Программное обеспечение**

При проведении работы мы использовали следующие программы и вычислительные алгоритмы.

Albacore – программа, предназначенная для бейсколлинга, процесса перевода данных из формата fast5 в fastq. Помимо этого, Albacore способна осуществлять разбивку общего массива данных на отдельные образцы.

Для сборки геномов использовали программы Canu, Flye, Unicycler. Подробное описание работы программ приведено в разделе «Результаты и обсуждения».

Для визуализации метрических показателей полученных ридов использовали программу Paurve (<https://github.com/conchoecia/pauvre/blob/master/pauvre/bamparse.py>).

Для определения качества сборок использовали программу Quast (<http://bioinf.spbau.ru/quast>). Визуализацию графов, генерируемых сборщиками, осуществляли с использованием программы Bandage (<https://rrwick.github.io/Bandage/>).

#### **Расходные материалы для секвенирования**

В работе использовали проточные ячейки FLO-MIN106, наборы для приготовления библиотек SQK-LSK108 Ligation Sequencing Kit 1D и Rapid Barcoding Sequencing (SQK-RBK004). При проведении тестового запуска применили контрольные образцы Lambda control DNA.

#### **Бактериальные штаммы**

Штаммы *Yersinia pestis* C-678, *Klebsiella pneumoniae* B-7849 и *Staphylococcus aureus* 2015-C45-1803 были секвенированы с использованием набора SQK-LSK108 для приготовления библиотек посредством фрагментации ДНК и лигирования адаптеров.

Штаммы *Y. pestis* C-771, *Y. pestis* C-776, *Y. pestis* И-1996, *Y. pestis* И-3113, *Y. pestis* C-590, *Y. pestis* C-359, *Y. pestis* C-197, *Y. pestis* C-235 были секвенированы с использованием набора SQK-RBK004 для быстрого приготовления библиотек.

#### **Выделение ДНК**

Для выделения ДНК использовали рутинный метод фенол-хлороформной экстракции. Методические приемы, направленные на получение молекул с высокой молекулярной массой, не применяли. Измерение концентрации ДНК осуществляли на флуориметре Qubit 2.0 с помощью набора реактивов Qubit® dsDNA HS Assay Kit.

#### **Подготовка библиотек для секвенирования**

Приготовление библиотек для секвенирования осуществляли с использованием двух наборов. Набор Ligation Sequencing kit (SQK-LSK108) применяли для приготовления библиотек посредством фрагментации ДНК и лигирования адаптеров. Данный набор входил в стартовый комплект; его применили для секвенирования контрольной фаговой ДНК и штаммов *Y. pestis* C-678, *K. pneumoniae* B-7849 и *S. aureus* 2015-C45-1803.

Набор Rapid Barcoding Kit (SQK-RBK004) закупили дополнительно и использовали для секвенирования штаммов чумного микроба. Данный набор позволяет быстро приготовить набор из нескольких образцов ДНК и секвенировать их на одной проточной ячейке.

#### **Общелабораторное оборудование**

Для работ, помимо наборов для подготовки библиотек и секвенирования, необходимы следующие реактивы и оборудование: пробирки Eppendorf DNA LoBind tubes объемом 1,5 мл, пробирки для ПЦР объемом 0,2 мл, свободная от нуклеаз вода, свежеприготовленный 70% этанол, водяная баня со льдом, микроцентрифуга, таймер, амплификатор, автоматические регулируемые пипетки с максимальным объемами 2 мкл, 20 мкл, 100 мкл, 200 мкл и 1000 мкл, а также наконечники к ним.

#### **Секвенирование на приборе MinION**

Секвенирование осуществлялось на трех ячейках FLO-MIN106. На первом этапе использовали набор для приготовления библиотек Ligation Sequencing kit (SQK-LSK108). На одной ячейке были проанализированы четыре образца. Секвенирование каждого из образцов осуществлялось в течение 2–4 ч. После секвенирования образцов ячейка промывалась согласно рекомендациям производителя.

Образцы для следующих двух раундов секвенирования были подготовлены с использованием набора Rapid Barcoding Kit (SQK-RBK004). Каждый образец для секвенирования состоял из нескольких библиотек ДНК. Секвенирование каждого образца осуществилось в течение 20–48 ч. Ячейки повторно не использовались.

### **Результаты и обсуждение**

#### **Подготовка оборудования к работе**

Для работы Oxford Nanopore MiniON требуется компьютер с высокими системными требованиями. Необходимы операционная система Windows, MacOS или Ubuntu, процессор серии i7 или Xeon с не менее чем четырьмя ядрами, не менее 16 Гб оперативной памяти, высокоскоростной накопитель типа SSD объемом не менее 1 Тб и с USB-портом версии 3.0. и выше [4].

Поскольку секвенатор MiniON является мобильным устройством, мы решили использовать для работы с ним переносной персональный компьютер (ноутбук), чтобы получить переносной комплект оборудования. На момент поставки секвенатора MiniON в лабораторию (конец 2017 г.) выбор моделей ноутбуков с необходимыми параметрами был ограничен, а стоимость существующих превышала 200 000 рублей. В связи с этим мы выбрали модель, позволяющую установить дополнительные компоненты, – HP Probook 450 G5 с процессором core i7-8550U и установленным жестким диском на 1 терабайт типа HDD и одним модулем оперативной памяти объемом 8 Гб. На данную модель были дополнительно установлены 8 Гб оперативной памяти и накопитель объемом 1 Тб типа SSD. Суммарная стоимость компьютера после установки дополнительных компонентов оказалась существенно ниже стоимости имеющихся в продаже моделей на период покупки оборудования. Внешний вид оборудования представлен на рисунке 1.





Рис. 1. Секвенатор MiniON и компьютер для обработки данных размещены в боксе биологической безопасности.

В процессе секвенирования происходит генерирование данных в специализированном формате fast5. Данный формат отражает особенности прохождения электрического тока через нанопору. Однако данные в этом формате не могут обрабатываться большинством программ для биоинформационного анализа. В связи с этим после анализа необходима конверсия данных в более распространенные форматы, например FASTQ.

Подобный перевод (конверсия) возможен параллельно с генерированием данных или может осуществляться с использованием специализированного вычислительного оборудования (сервер или рабочая станция) после окончания эксперимента.

### Результаты секвенирования образцов, подготовленных с использованием набора SQK-LSK108

Все протоколы для проведения анализа MiniON доступны на сайте сообщества (<https://community.nanoporetech.com>), доступ к которому пользователь получает при приобретении оборудования.

Информация о количестве работающих пор отображается на протяжении всего запуска. С течением времени количество работоспособных пор уменьшается. На рисунке 2 представлены диаграммы состояния проточной ячейки после 4 мин и 20 ч секвенирования соответственно. Интересно отметить, что, несмотря на значительно возросшее количество инактивированных ячеек, количество пор, в которых осуществлялся сиквенс в момент составления диаграммы, через 20 ч секвенирования оказалось большим (182 поры), чем через 4 мин после начала секвенирования (165 пор). Это означает что при оптимальных условиях секвенирование может происходить в течение длительного времени.

Тестовый запуск осуществили в присутствии представителей поставщика оборудования с использованием ДНК фага λ и соответствующего протокола для проверки оборудования. На рисунке 3 представлен отчет по количеству и длине полученных ридов. В таблице 1 представлены метрические показатели запуска, на рисунке 3 – распределение индивидуальных прочтений.

Размер генома фага λ составляет приблизительно 45 000 оснований. Однако при секвенировании нами были обнаружены прочтения длиной более 50 000 оснований. В результате анализа установлено, что данное прочтение состоит из нескольких последовательностей, гомологичных разным участкам генома фага λ. Вероятно, подобные артефакты связаны с лигированием нескольких молекул ДНК в процессе приготовления образца для секвенирования.

После завершения тестового эксперимента проточная ячейка был залита буфером для промывки и оставлена на ночь. На следующий день был проанализирован штамм

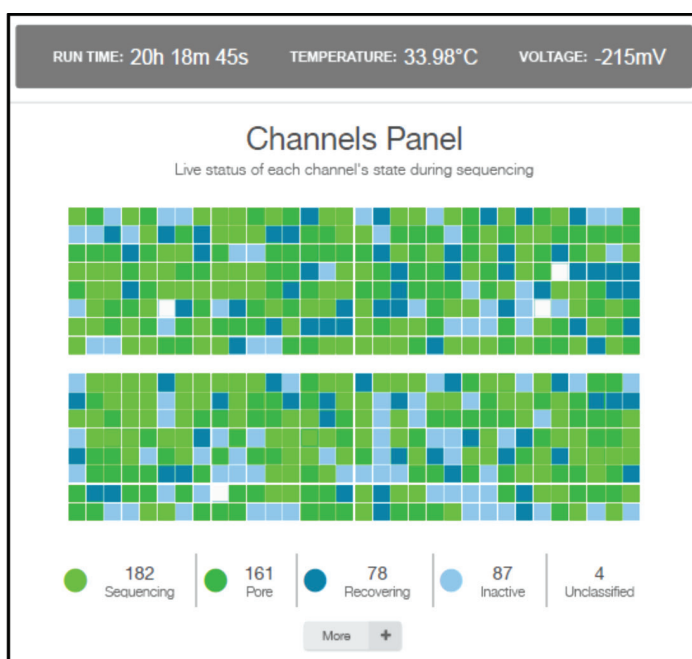
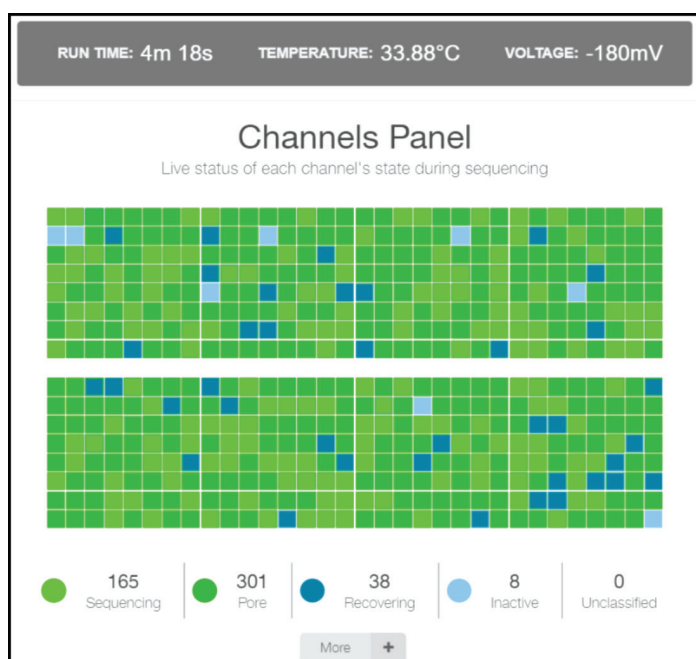


Рис. 2. Диаграмма состояния пор в проточной ячейке через 4 мин и 20 ч после начала секвенирования.

*Y. pestis* C-678. Метрические показатели архива ридов представлены в таблице 2. Распределение по количеству и качеству – на рисунке 4. Стоит отметить снижение качества (показатель Phred Quality – Q) полученных ридов при повторном и последующих запусках на одной ячейке. Данный пока-

затель рассчитывается по формуле  $Q = -10 \log_{10}P$ , где P – вероятность ошибки. Таким образом, при Q = 10 и Q = 20 вероятность ошибочного определения нуклеотида равняется 10% и 1% соответственно. Несмотря на то что средняя длина ридов при повторном использовании проточной ячей-

Таблица 1. Параметры ридов с тесового запуска (ДНК фага λ)

Образец	Риды	Нуклеотиды	Длина			
			Медианная	Сред.	Мин.	Макс.
ДНК фага λ	10 290	70 941 099	6 894.18	6 574	292	51 082

Таблица 2. Параметры ридов при анализе ДНК штамма *Y. pestis* C-678

Штамм	Риды	Нуклеотиды	Длина			
			Медианная	Сред.	Мин.	Макс.
<i>Y. pestis</i> C-678	6527	18 553 788	2842.62	1549	77	58628

Таблица 3. Параметры ридов при анализе ДНК *K. pneumoniae* B-7849

Штамм	Риды	Нуклеотиды	Длина			
			Медианная	Сред.	Мин.	Макс.
<i>K. pneumoniae</i> B-7849	10 664	44 598 141	4 182.12	1 731	5	95 471

Таблица 4. Параметры ридов при анализе ДНК *S. aureus* 2015-C45-1803

Штамм	Риды	Нуклеотиды	Длина			
			Медианная	Сред.	Мин.	Макс.
<i>S. aureus</i> 2015-C45-1803	438	1 146 012	2 616.47	912	5	24 363

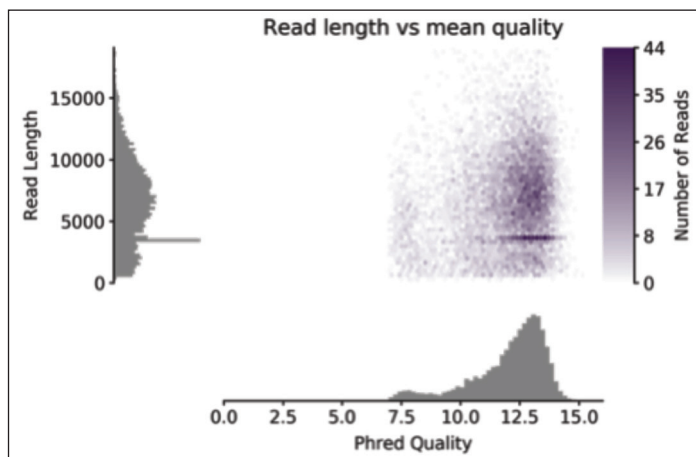


Рис. 3. Распределение ридов при анализе штаммов в пробном запуске с λ фагом.

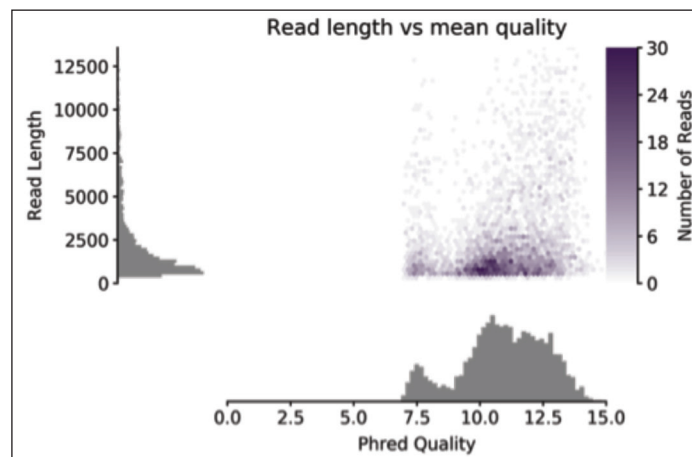


Рис. 4. Распределение ридов при анализе *Y. pestis* C-678.

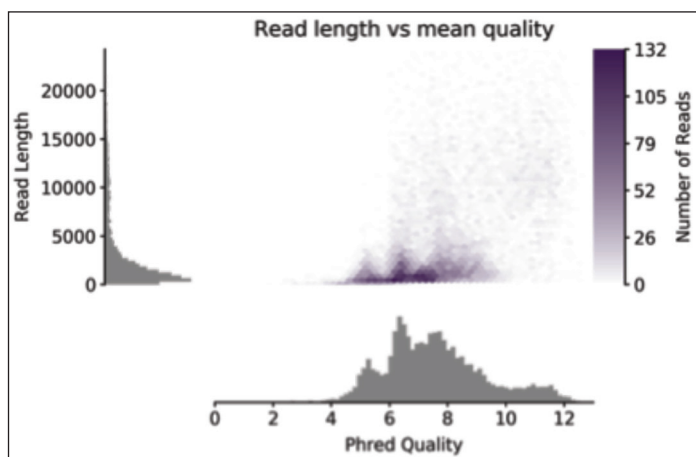


Рис. 5. Распределение ридов при анализе *K. pneumoniae* B-7849.

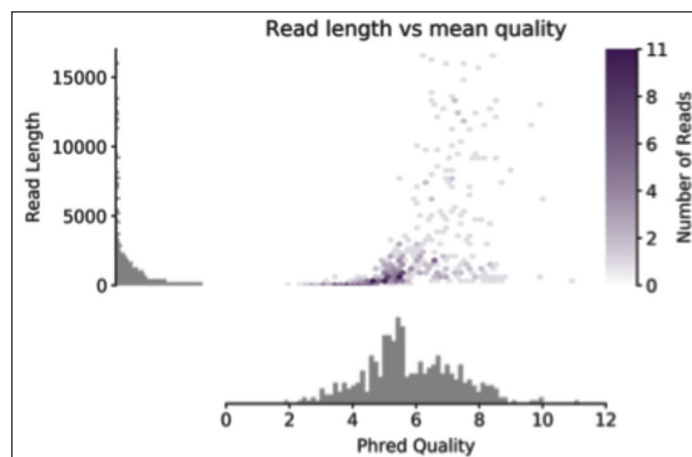


Рис. 6. Распределение ридов при анализе *S. aureus* 2015-C45-1803.

Таблица 5. Анализ геномов штаммов чумного микроба при использовании второй ячейки (протокол SQK-RBK004)

Штамм	Риды	Нуклеотиды	Длина			
			Медианная	Сред.	Мин.	Макс.
<i>Y. pestis</i> C-197	168 733	1 452 447 205	8607,96	4 456	139	145 806
<i>Y. pestis</i> C-235	120 024	1 098 703 801	9154,03	5 331	103	123 411
<i>Y. pestis</i> C-359	107 655	697 016 934	6474,54	3 736	143	100 826
<i>Y. pestis</i> C-739	15 863	78 777 467	4966,11	2 305	137	98 721

Таблица 6. Анализ геномов штаммов чумного микроба при использовании третьей ячейки (протокол SQK-RBK004)

Штамм	Риды	Нуклеотиды	Длина			
			Медианная	Сред.	Мин.	Макс.
<i>Y. pestis</i> C-771	95 035	424 022 177	4 461	2 635	82	87 573
<i>Y. pestis</i> C-776	78 546	374 444 809	4 767	2 563	89	109 376
<i>Y. pestis</i> И-1996	8 975	100 363 313	11 182	5 886	94	132 863
<i>Y. pestis</i> И-3113	53 647	617 169 107	11 504	7 107	150	163 412

Таблица 7. Показатели N25, N50, N75 для данных, полученных на второй ячейке MinION (протокол SQK-RBK004)

Штамм	N25	N 50	N75	Количество ридов		
				N25	N50	N75
<i>Y. pestis</i> C-197	32 284	17 577	8422	8742	26 209	59 899
<i>Y. pestis</i> C-235	30 045	16 951	8699	7214	21 175	46 542
<i>Y. pestis</i> C-359	21 332	11 945	6019	6244	18 727	41 831
<i>Y. pestis</i> C-739	22 398	10 977	4577	643	2068	5159

Таблица 8. Показатели N25, N50, N75 для данных, полученных на третьей ячейке MinION (протокол SQK-RBK004)

Штамм	N25	N 50	N75	Количество ридов		
				N25	N50	N75
<i>Y. pestis</i> C-771	14 676	7641	3962	4651	15 049	34 528
<i>Y. pestis</i> C-776	17 467	8973	4177	3543	11 183	26 530
<i>Y. pestis</i> И-1996	44 759	22 906	10 352	404	1192	2798
<i>Y. pestis</i> И-3113	33 381	21 299	10 728	3165	8794	18 959

ки оказалась значительно меньше (1546 оснований против 6574), максимальная длина оказалась существенно выше (58 628 против 11 236 оснований).

После завершения эксперимента по секвенированию образца ДНК *Y. pestis* C-678 ячейка также была промыта буфером.

Следующий раунд секвенирования проведен с ДНК штамма *K. pneumoniae* В-7849. Качество полученных данных оказалось значительно ниже, чем в предыдущих запусках (табл. 3). Распределение по количеству и качеству показано на рисунке 5. Однако в данном случае удалось сгенерировать больше информации, чем в случае анализа генома чумного микроба.

Последний запуск на данной ячейке был проведен с ДНК штамма *S. aureus* 2015-C45-1803. Метрические показатели ридов представлены в таблице 4. Распределение по количеству и качеству – на рисунке 6.

В связи с тем что эффективность анализа образцов после промывки проточной ячейки существенно снижалась, нами было принято решение использовать наборы реактивов, позволяющие анализировать одновременно несколько образцов ДНК. Мы выбрали набор SQK-RBK004, позволяющий осуществлять быструю пробоподготовку и мультиплексирование до 12 образцов. Было проведено два раунда секвенирования с использованием двух проточных ячеек. В таблицах 5 и 6 представлены результаты анализа геномов штаммов чумного микроба, полученные на второй и третьей ячейке соответственно.

Одним из распространенных метрических показателей, используемых для оценки полученных данных при полноге-

номном секвенировании, является N50 – длина последовательности, которая вместе со всеми последовательностями большей длины составляет более 50% от суммарной длины всех последовательностей, полученных в эксперименте. N25 и N75 определяют минимальную длину для 25% и 75% наиболее длинных последовательностей соответственно. Обычно данный показатель используется для последовательностей, полученных при реконструкции геномов. Однако мы провели анализ архивов сырых ридов, полученных в наших экспериментах. В таблицах 7 и 8 представлены значения для штаммов *Y. pestis* первого и второго запуска по протоколу SQK-RBK004.

Длина ридов и их количество являются важными предпосылками для эффективного анализа генома микроорганизма. Наличие незначительного числа ридов большой длины в случае нанопорового мономолекулярного секвенирования может оказаться более важным, чем присутствие большого количества коротких ридов в архиве. Например, N50, равный 17 577 оснований, полученный при анализе штамма *Y. pestis* C-197, означает, что более половины из 1 452 447 205 оснований, полученных при секвенировании, приходится на риды длиной 17 577 оснований и более. В то же время средняя и медианная длина ридов для данного архива составляет 4456 и 8607 оснований соответственно. Средние значения отражают эффективность работы нанопор в процессе секвенирования, однако малоинформативны для изучения биологически значимых признаков. На рисунке 7 представлено распределение прочтений в зависимости от их характеристик для ДНК четырех штаммов чумного микроба.

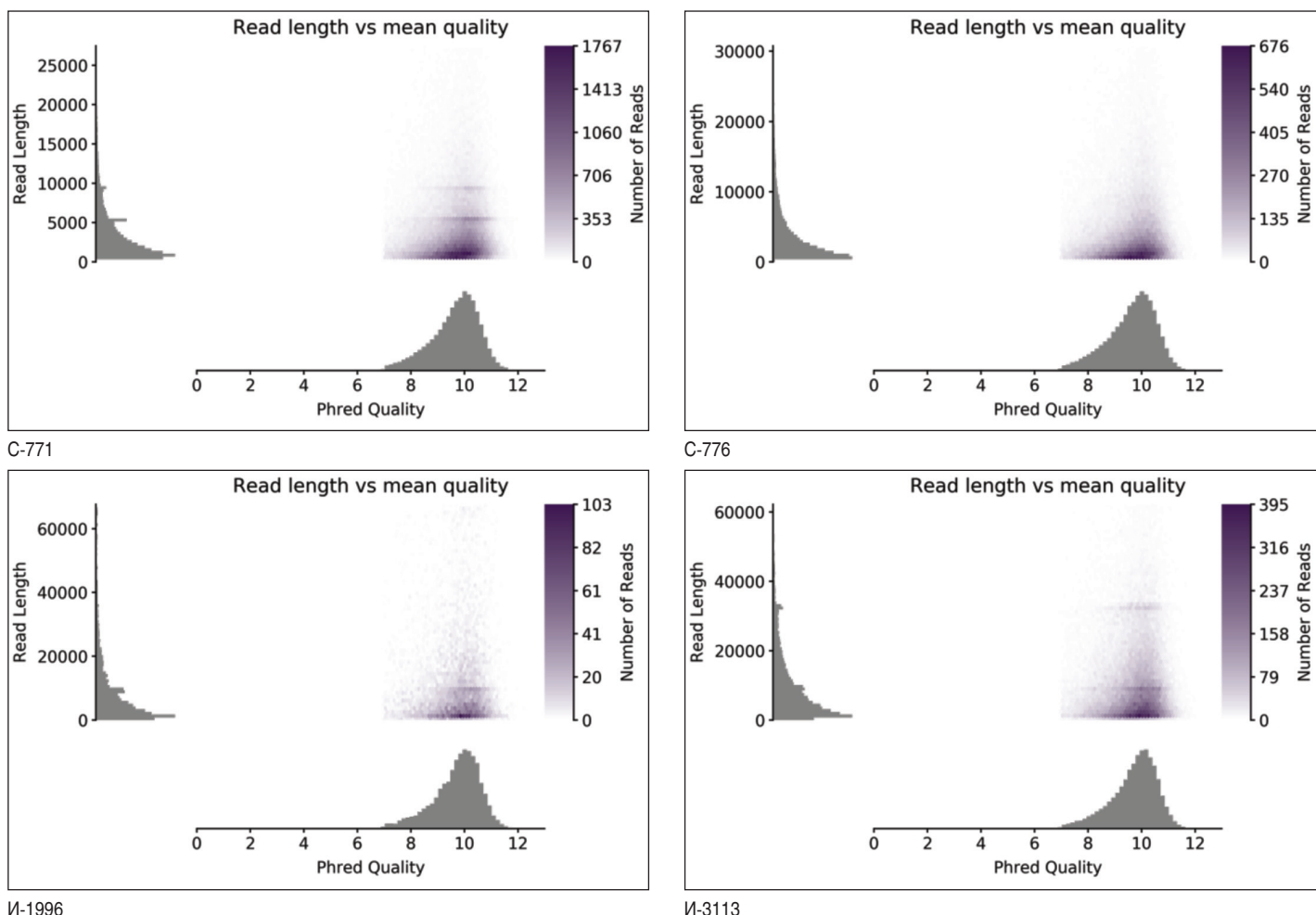


Рис. 7. Распределение ридов при анализе штаммов C-771, C-776, И-1996, И-3113.

Биоинформационный анализ полученных данных секвенирования MiniON несколько отличается от анализа данных массивированного параллельного секвенирования. На первом этапе файлы с данными в формате .fast5 с помощью программы Albacore разбивают на отдельные образцы. Поскольку большинство программ для биоинформационного анализа работают с данными в формате .fast5, необходимо осуществлять их конверсию.

Длинные риды, получаемые в результате мономолекулярного нанопорового секвенирования, облегчают сборку бактериальных геномов. Мы провели сравнение результатов, получаемых при использовании программ Canu, Unicycler и Flye для данных, полученных при использовании набора SQK-RBK004 для быстрого приготовления библиотек. Все программы находятся в свободном доступе и доступны для скачивания на веб-сайте <https://github.com>. Результаты представлены на рисунке 8.

В случае реконструкции кольцевой последовательности бактериальной хромосомы программа Bandage визуализирует полученные данные как кольцевую линию. Если реконструировать кольцевую последовательность не удастся, программа визуализирует полученные последовательности как фрагменты, длина которых изображается пропорционально их размерам.

Нам удалось реконструировать полные кольцевые последовательности бактериальной хромосомы для всех восьми

штаммов *Y. pestis* (рис. 8). Наибольшую эффективность продемонстрировала программа Flye. Однако при анализе штамма *Y. pestis* И-1996 последовательность хромосомы чумного микроба была реконструирована с использованием программы Unicycler, тогда как программа Flye не позволила сделать это.

Следует отметить, что в силу того, что данные мономолекулярного нанопорового секвенирования могут содержать большое количество ошибок, результаты реконструкции бактериальных хромосом нельзя рассматривать как завершённые геномы. Для увеличения точности необходимо использовать дополнительные методы, такие как гибридный анализ с использованием данных массивированного параллельного секвенирования.

### Выводы

Установлено, что одновременный анализ нескольких образцов на одной ячейке является более предпочтительным, чем последовательный анализ образцов на одной ячейке после промывки. Мы смогли получить в результате секвенирования архив ридов, содержащий информацию о 3 326 945 407 нуклеотидах, чем экспериментально подтвердили возможность получения больших объемов данных на оборудовании MiniON. Подобный массив позволяет получить стократное покрытие для десяти бактериальных гено-



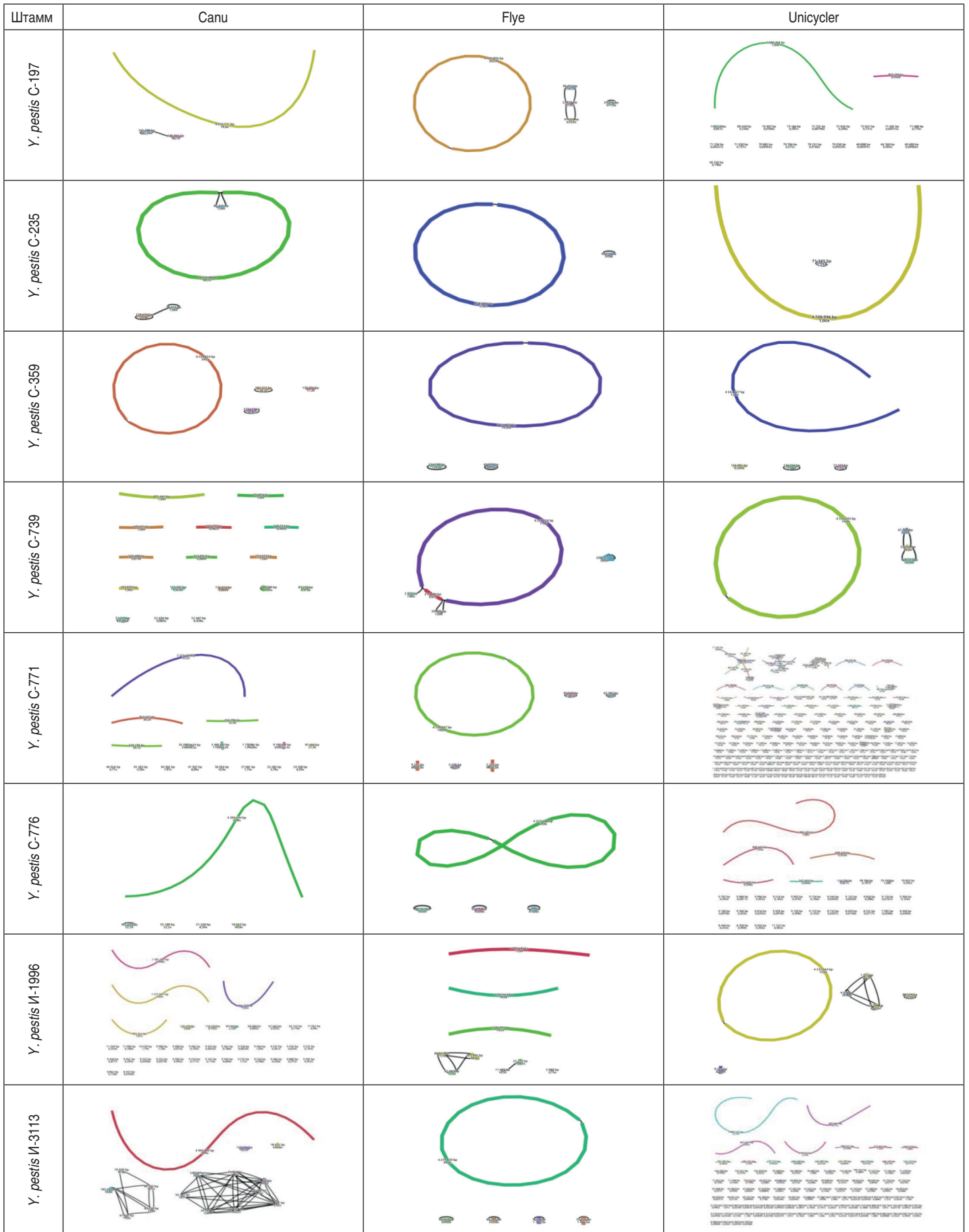


Рис. 8. Результаты реконструкции бактериальных хромосом и плазмид на основании данных мономолекулярного секвенирования.



мов среднего размера. Максимальная длина единичного прочтения составила 163 412 основания.

По нашему мнению, оборудование Oxford Nanopore MiniON в перспективе можно широко использовать в микробиологических лабораториях. Опыт использования MiniON при исследовании вирусов Зика и Эбола показывает высокую его эффективность при работе в экстремальных и полевых условиях [5, 6]. Простота экспериментальных процедур позволяет проводить секвенирование в различных условиях специалистами с минимальной подготовкой. Благодаря низкой цене, возможно комплектование данным оборудованием лабораторий любого уровня. Однако биоинформационный анализ и интерпретация данных являются сложным процессом, требующим мощного вычислительного оборудования и квалифицированного персонала.

### Финансирование

Работа выполнена при финансовой поддержке Российского научного фонда (грант 14-15-00599).

### Литература/References

1. Metzker ML. Sequencing technologies – the next generation. *Nat Rev Genet.* 2010 Jan;11(1):31-46. DOI: 10.1038/nrg2626. Epub 2009 Dec 8.
2. Rhoads A, Au KF. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics.* 2015 Oct;13(5):278-89. DOI: 10.1016/j.gpb.2015.08.002.
3. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. *Trends Genet.* 2018 Sep;34(9):666-681. DOI: 10.1016/j.tig.2018.05.008.
4. Wang JR, Jones CD. Fast alignment filtering of nanopore sequencing reads using locality-sensitive hashing. 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). – IEEE, 2015, pp. 127-130.
5. Hoenen T, Groseth A, Rosenke K, Fischer RJ, Hoenen A, Judson SD, et al. Nanopore sequencing as a rapidly deployable Ebola outbreak tool. *Emerg Infect Dis.* 2016 Feb; 22(2):331-4. DOI: 10.3201/eid2202.151796.
6. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc.* 2017 Jun;12(6):1261-1276. DOI: 10.1038/nprot.2017.066

### Информация об авторах:

Сизова Анжелика Алексеевна, техник-программист отдела коллекционных культур ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Скрябин Юрий Павлович, научный сотрудник лаборатории антимикробных препаратов ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Кисличкина Ангелина Александровна, кандидат биологических наук, старший научный сотрудник отдела коллекционных культур ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Фролов Вадим Борисович, инженер отдела коллекционных культур ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Майская Надежда Васильевна, научный сотрудник отдела коллекционных культур ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Иванов Сергей Андреевич, кандидат биологических наук, ведущий научный сотрудник лаборатории микробиологии чумы ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Дентовская Светлана Владимировна, доктор медицинских наук, главный научный сотрудник лабораторией микробиологии чумы ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Анисимов Андрей Павлович, доктор медицинских наук, заместитель директора по научной работе ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 36-0000

Богун Александр Геннадьевич, кандидат биологических наук, ведущий научный сотрудник отдела коллекционных культур ФБУН «Государственный научный центр прикладной микробиологии и биотехнологии» Роспотребнадзора  
Адрес: 142279, Московская область, Серпуховский р-н, п. Оболенск, ФБУН ГНЦ ПМБ  
Телефон: (4967) 360000

### Information about authors:

Anzhelika A. Sizova, technician-programmer of the collection cultures Department, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Yuriy P. Skryabin, research scientist of laboratory of antimicrobial agents, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Angelina A. Kislichkina, PhD (Biology), senior reseacher of the collection cultures department, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 360000

Vadim B. Frolov, engineer of collection cultures department, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Nadezhda V. Mayskaya, research scientist of the collection cultures department, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Sergey A. Ivanov, PhD (Biology), leading researcher of the laboratory of microbiology of the plague, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Svetlana V. Dentovskaya, MD, PhD, DSc, chief research scientist of laboratory of microbiology of plague, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Andrey P. Anisimov, MD, PhD, DSc, Deputy Director for scientific work, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000

Aleksander G. Bogun, PhD (Biology), leading researcher of the collection cultures department, State Research Center for Applied Microbiology and Biotechnology  
Address: SRCAMB 142279 Obolensk, Serpukhov district, Moscow region, Russian Federation  
Phone: (4967) 36-0000