

УДК 025.04.036:681.2

© Н. Д. Архипов, Д. Б. Архипов, 2020

СЕКВЕНС РНК ВИРУСА COVID-19: КОРОТКИЙ ОБЗОР

Первая статья по массивному параллельному РНК-секвенсу была напечатана лишь в 2018 г., т.е. за два года до начала коронавирусной пандемии. Первая удовлетворительная статья по секвенсу COVID-19 поступила в редакцию Nature 3 февраля 2020 г. и была опубликована 12 марта. В этой статье сообщалось о РНК-секвенсе последовательности из 29 903 нуклеотидов этого вируса. Пионерская статья в журналах Американского химического общества на эту тему (23 марта) посвящена бумажным чипам.

Кл. сл.: аналитическое приборостроение, РНК-секвенс, COVID-19

ИСТОРИЧЕСКАЯ СПРАВКА

Пандемия COVID-19 является беспрецедентным событием в истории человечества, и если в КНР, по сообщениям правительства, ее удалось ликвидировать уже в конце февраля, то в Европе и особенно США число инфицированных этим вирусом перевалило за миллион.

Проект "Геном человека" обошелся не менее чем в \$ 20 млрд, и его главной задачей было создание методики идентификации личности через ДНК-секвенс. В 2005 г. J. Rothberg и его 56 соавторов построили первый ДНК-секвенатор и разработали первую методику по массивному параллельному секвенсу нуклеиновых кислот [1]. Иногда этот прибор называют крупномасштабным параллельным пиросеквенатором. Себестоимость этого прибора составляла \$ 500 тыс.; себестоимость реактивов для секвенирования ДНК из миллиона нуклеотидных пар — \$ 100 тыс. Ротберг назвал свой пиросеквенатор 454 Life Sciences, в 2007 г. фирма Roche приобрела методику за \$ 154.9 млн, а прибор под названием Roche-454 производился до 2016 г. Впоследствии Ротберг активно участвовал в Neanderthal Genome Project. Первоначально была сделана попытка секвенирования ДНК неандертальца из пещеры Виндия в Хорватии, но из-за сильного загрязнения образца бактериями попытка закончилась неудачей. Гораздо более успешным оказался секвенс неандертальцев из Денисовской пещеры в Солонешенском районе Алтайского края.

ПРИБОРНО-АНАЛИТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОБЛЕМЫ

Roche Diagnostics (Коннектикут) занимается в первую очередь медицинской диагностикой.

В марте 2020 г. эта фирма объявила о создании теста на COVID-19, выполняемого за 3.5 ч. Вплоть до февраля 2020 г. разработка РНК-секвенаторов не являлась столбовым направлением в аналитическом приборостроении, хотя известны по меньшей мере две компании, которые выпускали РНК-секвенаторы еще до эпидемии COVID-19. Во-первых, это англо-американская компания Oxford Nanopore Technologies, разработавшая массивный параллельный РНК-секвенс не позднее конца 2017 г. [2], во-вторых, калифорнийская компания Illumina Systems, основанная в 1999 г.

Именно на приборе Illumina Mini Seq авторы работы [3] определили последовательность из 29 903 нуклеотидов COVID-19 и нашли степень сходства (89.1 %) с вирусом SARS, поразившим Южный Китай в 2002–2003 гг. (скончалось 774 человека). В настоящее время Illumina выпускает NovaSeq 6000, который позволяет анализировать макромолекулы как РНК, так и ДНК.

В работе [4] вариант секвенирования не сообщается, но зато сообщается об использовании ПЦР в реальном времени. Авторы этой статьи выдвигают гипотезу о первичном заражении через летучих мышей, которыми, оказывается, торгуют на зоологическом рынке Уханя. Наконец, статья [5], третья и последняя по рассматриваемой тематике за март–апрель 2020 г., посвящена перспективам использования бумажных чипов.

COVID-19 изучают всего три месяца, и неизвестно, насколько легко он мутирует. Метилирование РНК — это задача, которой никто не занимался, хотя эта проблема представляется вполне решаемой. Изучение метилированных ДНК и РНК находится в стадии исследования, и твердо известно лишь то, что это едва ли не главная причина мутаций.

Фирма Illumina Systems сообщает, что стоимость диагностики на новый коронавирус начинается

Число ссылок на статьи по COVID-19

Номер статьи	Число ссылок по Web of Sciences	Число ссылок по Google Scholar
1	5103	8960
2	180	302
3	136	641
4	329	1342
5	2	3

с § 199 и занимает 5 ч. Определение полной РНК-последовательности (total RNA sequence) обходится намного дороже. В заключение приведем в таблице число ссылок на цитируемые статьи (2 мая 2020 г.) по Web of Sciences и Google Scholar.

ВЫВОДЫ

Просмотрены Nature, Science и журналы Американского химического общества за январь–апрель 2020 г. по секвенсу РНК вируса COVID-19. Определенные успехи достигнуты, но в целом проблема не решена из-за возможных мутаций этого коронавируса.

СПИСОК ЛИТЕРАТУРЫ.

1. Margulies M., Egholm M., Altman W.E. et al. Genome sequencing in microfabricated high-density picolitre reactors // Nature. 2005. Vol. 437, no. 7057. P. 376–380. DOI: 10.1038/nature03959
2. Garalde D.R., Snell E.A., Jachimowicz D. et al. Highly parallel direct RNA sequencing on an array of nanopore // Nature Methods. 2018. Vol. 15, no. 3. P. 201–206. DOI: org./1038/nmeth.4577
3. Wu F., Zhao Su, Yu B. et al. A new coronavirus associated with human respiratory disease in China // Nature. 2020. Vol. 579, no. 7798. P. 265–269. DOI: 10.1038/s41586-020-2008-3
4. Zhou P., Yang X.-L., Wang X.-G. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin // Nature. 2020. Vol. 579, no. 7798. P. 270–273. DOI: 10.1038/s41586-020-2012-7
5. Mao K., Zhang H., Yang Z. Can a paper-based device trace COVID-19 sources with wastewater-based epidemiology // Environmental Science and Technology. 2020. Vol. 54, no. 7. P. 3733–3735. DOI: 10.1021/acs.est.0c01174

*Университет информационных технологий,
механики и оптики, Санкт-Петербург*

*Институт аналитического приборостроения РАН,
Санкт-Петербург*

Контакты: *Архипов Николай Дмитриевич,*
nikolos.kage97@mail.ru

Материал поступил в редакцию 03.05.2020

CURRENT STATE FOR RNA SEQUENCE OF COVID-19

N. D. Arkhipov¹, D. B. Arkhipov²

¹University of Information Technologies, Mechanics and Optics, Saint Petersburg, Russia

²Institute for Analytical Instrumentation of RAS, Saint Petersburg, Russia

The epidemic of coronavirus COVID-19 started in Wuhan (China) in December 2019. By March 2020, the Chinese government could suspend the new pandemic but tens of thousands of people outside China already got sick of the new type of a severe acute respiratory syndrome. Pioneer articles on the sequence of COVID-19 were published in the Chinese scientific journals in January 2020. On March 12, 2020 Nature (London) presented two comprehensive and free-of-charge papers on the RNA sequence of this virus.

We report in our short review the current state of the RNA sequence technique and its application to COVID-19. Massive parallel DNA sequence was presented by Rothberg and his group in 2005. The instrument was named 454 Life Sciences. Through a decade, massive parallel sequence in nanopores has been invented. No later than 2017 similar instrument for the RNA sequence was created. It was the massively parallel sequence in nanopores that was used by the Chinese scientists for COVID-19 study. Additionally, polymerase chain reaction in real-time was performed. As a result, within three months the sequence of 29 303 nucleotides of COVID-19 has been detected. Modern sequencers are expensive instruments. For example, the cost of pyrosequencer 454 Life Sciences was \$ 500 000. For this reason, there is still no information about the difference between the Wuhan, Moscow and American coronaviruses. Our results were obtained by study of Nature, Science, and journals of the American Chemical Society.

Keywords: analytical instrumentation, RNA sequence, COVID-19

HISTORICAL REFERENCE

The pandemic of COVID-19 is an unprecedented event in the history of mankind and if in China, according to the government, it was eliminated at the end of February, then in Europe and especially in the United States, the number of people infected with this virus exceeded a million.

The Human Genome project cost at least \$20 billion, and its main task was to create a method for identifying a person through a DNA sequence. In 2005, J. Rothberg and 56 co-authors developed the first DNA sequencer and the first technique for a massive parallel nucleic acid sequence [1]. Sometimes this device is called a large-scale parallel pyrosequencer. The cost of this device was \$500 000; cost of reagents for DNA sequencing from million of nucleotide pairs — \$100 000. Rothberg named his pyrosequencer 454 Life Sciences,

Roche acquired the methodology in 2007 for \$154.9 million, and the device under the name Roche-454 was produced until 2016. Subsequently, Rothberg actively participated in the Neanderthal Genome Project. Initially, an attempt was made to sequence the DNA of a Neanderthal from Windia Cave in Croatia, but due to heavy contamination of the sample with bacteria, the attempt ended in failure. The sequence of Neanderthals from the Denisovskaya cave in the Solonshensky district of the Altai Territory turned out to be much more successful.

INSTRUMENTATION AND ANALYTICAL SUPPORT OF THE PROBLEM

Roche Diagnostics (Connecticut) is engaged in primarily medical diagnosis.

In March 2020, this company announced the creation of a test for COVID-19, performed in 3.5 hours. Until February 2020, the development of RNA sequencers was not a main direction in analytical instrument making, although at least two companies are known that produced RNA sequencers even before the COVID-19 epidemic. Firstly, it is the Anglo-American company Oxford Nanopore Technologies, which developed a massive parallel RNA sequence no later than the end

Number of references to articles by COVID-19

Article Number	Number of links by Web of Sciences	Number of links by Google Scholar
1	5103	8960
2	180	302
3	136	641
4	329	1342
5	2	3

of 2017 [2], and secondly, the California company Illumina Systems, founded in 1999.

It was with the use of the Illumina Mini Seq that the authors of [3] determined a sequence of 29 903 nucleotides of COVID-19 and determined the degree of similarity (89.1%) with the SARS virus that affected South China in 2002-2003 (774 people died). Currently, Illumina produces NovaSeq 6000, which allows to analyze macromolecules of both RNA and DNA.

In work [4], the sequencing option is not reported, but real-time PCR is reported. The authors of this article hypothesize about primary infection through bats, which, it turns out, are traded in the Wuhan zoological market. Finally, article [5], the third and last on the topic under consideration for March-April 2020, is devoted to the prospects for the use of paper chips.

COVID-19 has been studied for only three months, and it is unknown how easy it is to mutate. RNA methylation is a task that no one has been involved in, although this problem seems quite solvable. The study of methylated DNA and RNA is under investigation and it is only firmly known that this is almost the main cause of mutations.

Illumina Systems reports that diagnostics cost on a novel coronavirus begins with \$199 and that diagnostics take 5 h. Determination of the total RNA sequence is much more expensive. Finally, the table shows the number of references to the articles cited (May 2, 2020) by Web of Sciences and Google Scholar.

CONCLUSIONS

Nature, Science and the journals of the American Chemical Society issued in Jan-April, 2020 have been looked through for the articles on the RNA sequence

of the COVID-19 virus. Some successes have been achieved, but in general the problem has not been solved due to possible mutations of this coronavirus.

REFERENCES

1. Margulies M., Egholm M., Altman W.E. et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 2005, vol. 437, no. 7057, pp. 376–380. DOI: 10.1038/nature03959
2. Garalde D.R., Snell E.A., Jachimowicz D. et al. Highly parallel direct RNA sequencing on an array of nanopore. *Nature Methods*, 2018, vol. 15, no. 3, pp. 201–206. DOI: org./1038/nmeth.4577
3. Wu F., Zhao Su, Yu B. et al. A new coronavirus associated with human respiratory disease in China. *Nature*, 2020, vol. 579, no. 7798, pp. 265–269. DOI: 10.1038/s41586-020-2008-3
4. Zhou P., Yang X.-L., Wang X.-G. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 2020, vol. 579, no. 7798, pp. 270–273. DOI: 10.1038/s41586-020-2012-7
5. Mao K., Zhang H., Yang Z. Can a paper-based device trace COVID-19 sources with wastewater-based epidemiology. *Environmental Science and Technology*, 2020, vol. 54, no. 7, pp. 3733–3735. DOI: 10.1021/acs.est.0c01174

Contacts: Arkhipov Nikolay Dmitrievich,
nikolos.kage97@mail.ru

Article received by the editorial office on 03.05.2020