The MGI Automation System Integrated with DNBSEQ Platform Enables the Detection and Tracing of Pathogenic Microorganisms

The ATOPlex Technology Compatible with MGISP-Smart 8 and DNBSEQ-G99 Facilitates SARS-CoV-2 Research

In this study, MGISP-Smart 8 & MGISP-100 automated systems were used to complete library preparation of the target pathogen (SARS-CoV-2) based on ATOPlex technology. Afterwards, DNBSEQ-G99 and MGI metarget-COVID were utilized for SARS-CoV-2 sequencing, detection and tracing, respectively. This workflow can also be applied in other pathogens (eg. Monkeypoxvirus) and MGI provides a complete set of combinational products for infectious pathogen detection.

Recommended applications: Pathogenic microorganisms - specific pathogen detection

Recommended models: MGISP-Smart 8RS, MGISP-100RS and MGISP-960RS (Automation system)

DNBSEQ-G99ARS, DNBSEQ-T7RSand DNBSEQ-G400RS (Sequencing platform)

- ### ATOPlex technology enables pathogen detection and tracing

  The MGI unique ATOPlex multiplex PCR technology fullfill the detection of different pathogenic microorganisms, with high detection sensitivity (LOD of SARS-CoV-2: 10 copies/mL) and high genome coverage, enabling detection and tracing of various samples.

- ### The automation system is highly compatible with nucleic acid extraction and library preparation

  The MGISP-Smart 8 & MGISP-100 automated systems from MGI can accurately, stably and efficiently complete the whole process including nucleic acid extraction, normalization, library preparation, pooling and DNB making.

- ### Efficient and high-quality sequencing data output

  DNBSEQ sequencing technology exhibits many excellent features such as high accuracy, low duplication rate and low index hopping rate. Among them, DNBSEQ-G99 has fast sequencing speed, a built-in computing module integrating sequencing and bioinformatics and only takes 4.5 hours to complete SE100+10+10 sequencing. Its data output is efficient and high-quality.

- ### A complete set of combinational products for pathogen detection and tracing

  The combinational products, including nucleic acid extraction kits & library preparation kits, automation systems, sequencing platforms, and bioinformatics analysis software, fully enabling pathogen research.

Background

The rapid development of massively parallel sequencing (MPS) has facilitated its wide application in scientific research and clinical diagnosis. This technology not only improves sequencing accuracy and speed, but also significantly reduces cost[1]. Specifically, targeted sequencing (TS) can focuses on specific genes, making it more cost-effective. Additionally, TS is widely used in pathogen detection, clinical diagnosis, genetics disease screening with higher sequencing depth, coverage, accurac*Rudi*-sensitivity[2]. The ATOPlex customized targeted sequencing platform based on MGI's self-developed ultra-high plex PCR technology has the characteristics of high sensitivity, high accuracy, easy to operate and low cost. It has been widely recognized by researchers and applied in many scenarios such as individual identification and kinship analysis[3], noninvasive prenatal testing (NIPT)[4], environmental monitoring[5], and the tracking of SARS-CoV-2[6].

The process of MPS library preparation is complicated with high manual operation cost, high error rate and pollution risk, especially when dealing with large number of samples[7]. The automated equipment can effectively circumvent these problems and ensure the accuracy and reliability on the premise of saving labor cost. Therefore, more and more hospitals and laboratories choose automated systems to free the hands of doctors and researchers[8]. The automated system independently developed by MGI covers multiple series. Among which, MGISP-Smart 8, a newly launched automated sample preparation system, has an independent and controllable eight-channel pipette that can not only complete flexiable library preparation from 1 to 48 samples/run, but also automatically cover the whole process from sample transfer, nucleic acid extraction, library preparation, pooling to DNB making. Additionally, MGI has also developed MGISP-100 automated system focusing on the MPS field. It is a small and compact low throughput sample preparation system, which can quickly, stably and efficiently complete a series of experimental operations such as nucleic acid extraction and library preparation. Therefore, the automated solution composed of MGISP-Smart 8 and MGISP-100 was used for library preparation in this study. This solution can ensure the uniformity of library preparation, thereby improving the accuracy and consistency of sequencing results, reducing the running cost of MPS, and making it truly unattended.

The sequencing platform based on the DNBSEQ technology launched by MGI has the advantages of high accuracy and sensitivity, ultra-low duplication rate and index hopping rate. The main genetic sequencers of MGI include DNBSEQ-G400, DNBSEQ-G99 and DNBSEQ-T7 can meet the research demands of medicine, scientific research, public health, food safety and other related fields.

Viruses are among the most lethal infectious pathogens in the world[9]. The widesspread of infectious diseases such as tuberculosis, polio, smallpox and diphtheria has resulted in high morbidity and mortality. Animal diseases, for example rinderpest, spread along trade routes and accompanying armies, causing great damage to livestock and related populations[10]. In recent years, Ebola, AIDS, Influenza A, Middle East respiratory syndrome (MERS), severe acute respiratory syndrome (SARS) and COVID-19 have caused millions or tens of millions of deaths around the world11. Timely diagnosis, treatment and intervention of infectious pathogens can effectively curb the spread of the virus. Therefore, rapid and accurate detection, diagnosis and research of viruses are undoubtedly essential[9]. MGI has developed a complete set of combinational products for the detection of infectious pathogens: Amplifying the target fragment of pathogens such as monkeypox HIV, influenza and SARS-CoV-2 with the ATOPlex technology on the automation system (MGISP-Smart 8 & MGISP-100), followed by sequencing on the DNBSEQ platform, and analysis on related bioinformatics software to complete detection and tracing.

Study Description

The combinational products were evaluated with the SARS-CoV-2 standard. During the process, the targeted amplification of virus was completed with the ATOPlex RNA Library Prep Set V3.1 on the MGISP-100 and MGISP-Smart 8 automated system. The sequencing process was completed on the DNBSEQ-G99 genetic sequencer, and the SARS-CoV-2's detection and tracing were completed with the MGI metargetCOVID software.

The final results showed that both the library and sequencing results were of high quality, and the pathogen identification and typing results were correct and accurate. In addition, the combinational products can also be used in the study of other pathogens, for instance, monkeypox virus detection and genotyping.

Materials and Methods

Sample preparation and extraction

A commercial standard COVID-19 RNA Control (high copy)-L (GeneWell PN: TEST09, specification: $4*10^8$ copies/RNA dry powder) having the same sequence as SARS-CoV-2 isolate Wuhan-Hu-1 (GenBank: MN908947.3) was used in this study. The lowest detection limit of ATOPlex RNA Library Prep Set V3.1 is 10 copies/mL. In this study, $4*10^4$ and $4*10^5$ copies/mL RNA samples were selected after gradient dilution, named GW 4 and GW 5, respectively. Since standards were selected in this study, no extraction operation was involved. For biological samples, MGIEasy Nucleic Acid Extraction Kit with MGISP-100 is recommended for automated extraction (Figure 1).

Library preparation and sequencing

In this study, 10 μL of each sample was input for library preparation. In order to avoid aerosol contamination caused by PCR, there are two physical isolation areas to complete this process. (1) Anterior area: The reverse transcription reagents and multiplex PCR amplification reagents were prepared manually, and RNA reverse transcription reagent dispensing, reverse transcription reaction, and multiplex PCR amplification reagent dispensing were performed on MGISP-100. (2) Posterior area: Product purification, library reagents dispensing, PCR (35 cycles), PCR product purification, library preparation and DNB making were performed on MGISP-Smart 8.

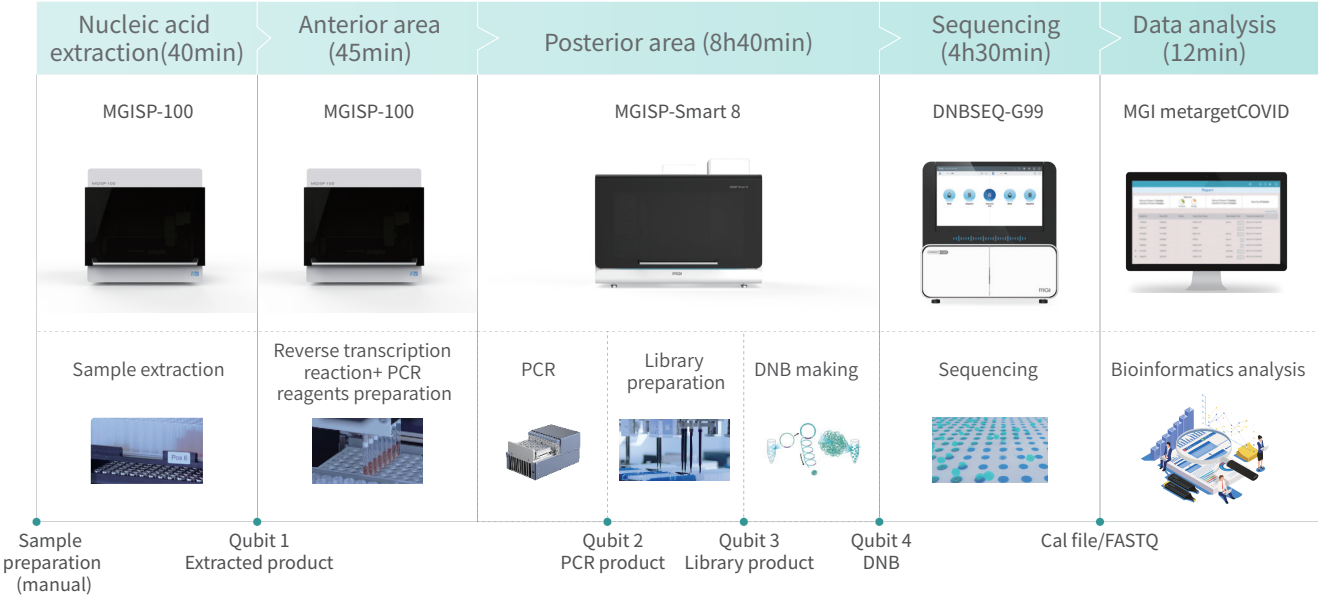
MGISP-Smart 8 can complete not only multiplex PCR, purification, fragmentation, adapter ligation, pooling, DNB making and other reactions, but also the pipetting of eight different volumes at one time, replacing a lot of manual pipetting such as pooling, normalization, reagent dispensing (Figure 1).

In this study, flexible throughput of 8 ~ 40 samples/run can be achieved. The deck layout of MGISP-Smart 8 in the library preparation process is shown in Figure 2. The deck contains 30 SBS standard plate positions and 1 trash can position. In the purification step of the posterior area, the single well reservoir with cap contained in MGISP-Smart 8 could prevent reagents from volatilizing during the automated dispensing of large volumes of ethanol. When adapter ligation is performed, the PCR plate adapter C of MGISP-Smart 8 could support the automated pipetting step of the UDB-Barcode reagent in the sealed membrane state without

manual intervention. Before DNB preparation, MGISP-Smart 8 could also complete these steps of pooling, normalization and quantification. During the library preparation process, the proportion of automated operation time is as high as 90.5%, indicating this process was highly automated (Figure 3). In addition, manual library preparation was carried out in parallel in this test to evaluate the feasibility of automated system, the operation procedures can be referred to the manual (https://www.mgi-tech.com/Home/Products/reagents_info/id/71.html),and then these libraries were sequenced on DNBSEQ-G99 for SE100+10+10 strategy.

Data Analysis

The MGI metargetCOVID software was used for bioinformatics analysis in this study and it can be installed in the DNBSEQ-G99. The whole process includes: FASTQ QC, short read alignment, virus content detection, correct strand bias & cutoff adapter, variant detection & lineage distribution detection, summary report output.



Note:
1.Standards were selected for testing in this study, so sample extraction was not involved. For biological samples, it is recommended to use MGISP-100 for automated extraction
2.This process shows the duration of the experiment for 8 samples.

Figure 1. The whole workflow of the detection and tracing of pathogenic microorganisms including nucleic acid extraction, library preparation, sequencing and bioinformatics analysis

❶ Comb.carrier K: Compatible with 1000 µL, 200µL and 50µL tips
❷ Standard plate position: Compatible with SBS standard plate consumables such as PCR plates and deep well plates
❸ PCR plate adapter C: Achieve pipetting of PCR plates at the time of membrane sealing
❹ Temp.control module: Achieve low-temperature storage or high-temperature heating of reagents or samples
❺ Magnetic module A: For magnetic beads adsorption on deep well plates to achieve purification and extraction of samples
❻ Heater & shaker module: Dual functions of heating and shaking
❼ Single well reservoir adapter A: Achieve dispensing of large volumes of liquid to prevent the reagent volatilization
❽ Magnetic module B: Support magnetic bead adsorption on PCR plates to achieve purification and extraction of samples
❾ PCR module: Achieve enzyme reaction or PCR
❿ Trash can: Garbage recycling

Figure 2. The MGISP-Smart 8 deck layout display

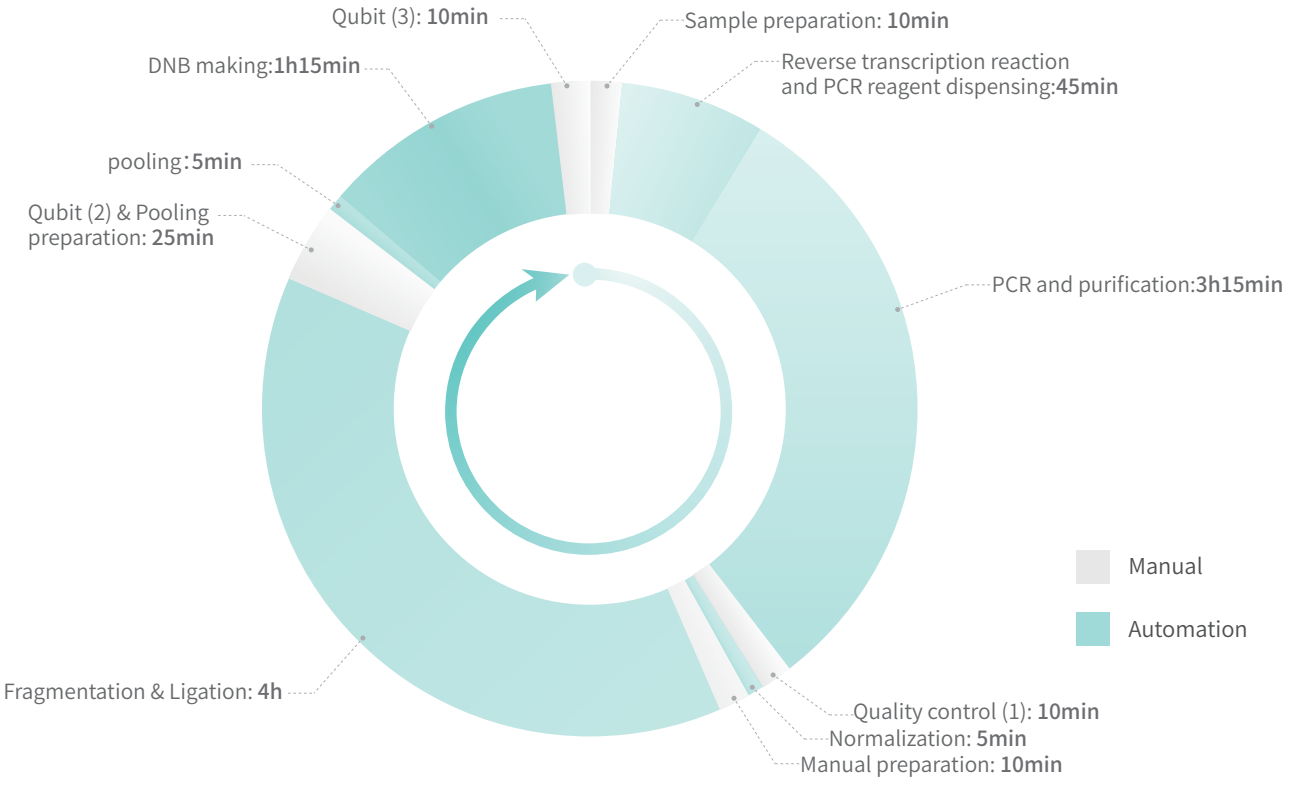


Figure 3. The display of time required for automated and manual steps when preparing 8 libraries based on the automated system

Result

The MGI automation system enables high-quality libraries preparation

In this study, SARS-CoV-2 standards and ATOPlex RNA Library Prep kit were used for library preparation based on the automation system (MGISP-100 & MGISP-Smart 8) (Figure 1). With the same input, the RT-PCR products in automated solution are more than 150 ng, and the library yield also meet the criterion (> 20 ng) (Table 1). Then all the libraries were mixed in equal moles to prepare DNB, the final concentration of automated and manual solution is 21.6 ng/μL and 14.0 ng/μL, respectively, also meeting the QC criterion (> 8 ng/μL). The above results indicate that MGI automated system can prepare high-quality libraries and even slightly better than manual solution.

Sample name	Automated/manual library preparation	Input (μL)	RT-PCR yield (ng)	Library yield (ng)
GW 4-1	Manual	10	1356.0	30.8
	Automation	10	2226.0	47.5
GW 5-1	Manual	10	3000.0	32.3
	Automation	10	6480.0	44.8

Table 1. The primary comparison of libraries prepared by manual and automated solutions.

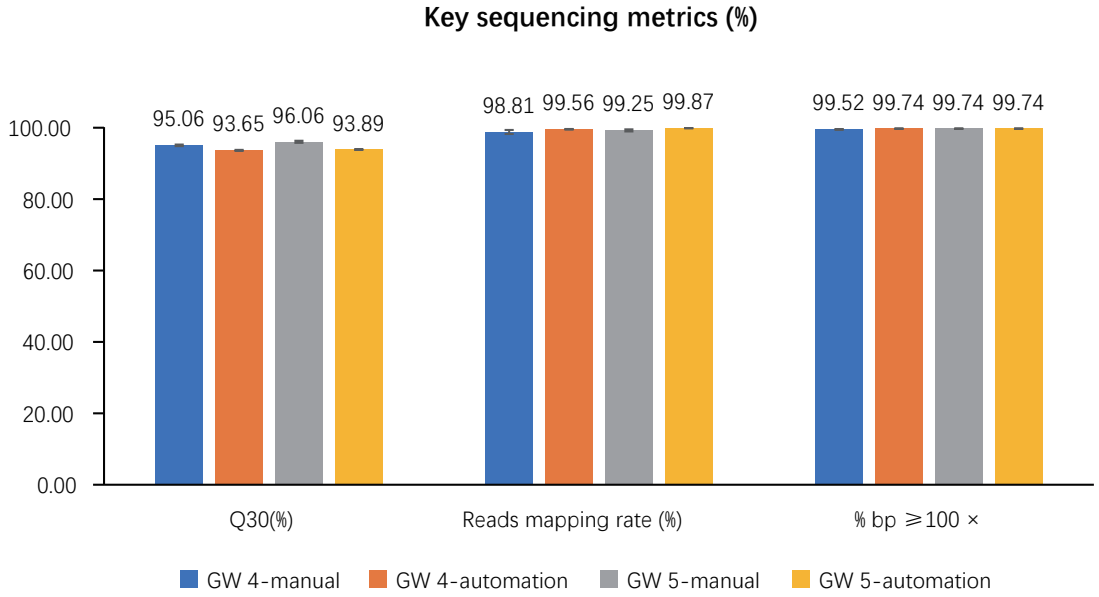


Figure 4. A comparison of key sequencing metrics between automated and manual solutions. GW 4-automation refers to the libraries of GW4 prepared by the automated system. The meanings of other sample names follow this analogy. Primary sequencing metrics show that there is no difference between automated and manual solutions.

The sequencing quality of the libraries prepared by automated system is high

To evaluate the sequencing quality of libraries from MGI automated system, researchers analyzed the sequencing data from GW 4 and GW 5 libraries. The results show that for both GW 4 and GW 5, the Q30 obtained from automated system are higher than 93%, mapping rate are higher than 98%, and target coverage (% bp ≥ 100×) are higher than 99%, which are comparable to manual (Figure 4). In addition, when mapping the GW 4-automation sequencing results to the reference genome, and the log values of read depth are around 4.8, indicating high coverage uniformity (Figure 5).

The above results show that ATOPlex library preparation workflow combined with MGI automated system and DNBSEQ sequencing platform can obtain high quality data with higher sequencing coverage and depth.

MGI provide a complete combinational products for the detection and tracing of pathogenic microorganisms

Further analysis for the automated solution show that the concentration of SARS-CoV-2 detected from GW 4 is 10^4 orders of magnitude, accounting for more than 80%. Similarly, the concentration detected from GW 5 is 10^5 orders of magnitude and account for more than 90%. All identification results are positive, in line with expectations, and basically consistent with the manual solution (Table 2). Further variant calling analysis also reveal that GW 4-automation library contain neither SNPs (single nucleotide polymorphisms) nor INDELs (insertion-deletion mutations) (Table 3).

After the obtained SARS-CoV-2 consensus sequence was assigned to the SARS-CoV-2 phylogenetic tree by the algorithm, the clade ID information of the consensus sequence was assigned, and then the SARS-CoV-2 lineage was calculated by the Pangolin tool. It was found that the typing information of the GW 4-automation library was completely consistent with SARS-CoV-2 isolate Wuhan-Hu-1 (Figure 6). Table 3 and Figure 6 show the results of GW 4-1-automation.

All these results prove that the ATOPlex technology is perfectly compatible with MGI automated system and DNBSEQ platform, combined with the corresponding software, can detect pathogenic microorganisms quickly and accurately.

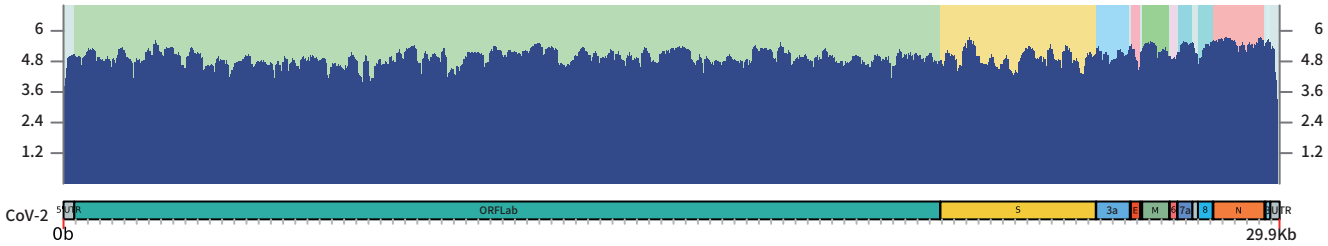


Figure 5. Histograms of coverage range about GW 4-1-automation library. The "histograms of coverage range" means the range and uniformity of sequencing coverage for an entire data set. The X axis is base position in reference chromsome and Y axis is the log value of read depth.

Library preparation	Sample ID	SARS-CoV-2 (copies/mL)	SARS-CoV-2 reads占比(%)	Identification
Manual	GW 4-1	3977.92	83.9	Positive
	GW 4-2	3750.55	83.25	Positive
	GW 4-3	3850.27	83.54	Positive
	GW 5-1	13609.1	93.41	Positive
	GW 5-2	14692	93.78	Positive
	GW 5-3	12122.78	92.8	Positive
Automation	H_2O-1	0.00	0.02	Negative
	H_2O-2	0.00	0.03	Negative
	GW 4-1	5514.85	87.18	Positive
	GW 4-2	4491.04	85.19	Positive
	GW 4-3	4289.13	84.71	Positive
	GW 5-1	17626.61	94.59	Positive
	GW 5-2	19764.91	95.05	Positive
	GW 5-3	16341.06	94.27	Positive

Table 2. The comparison of SARS-CoV-2 detection results between automated and manual solutions

Assemble Size(bp)	Num Ns	Num SNPs	Num INSs	Num DELs	Clade ID	Lineage ID
29903	75(0.25%)	0	0	0	19A	B

Table 3. Variant calling and typing information about GW 4-1-automation library

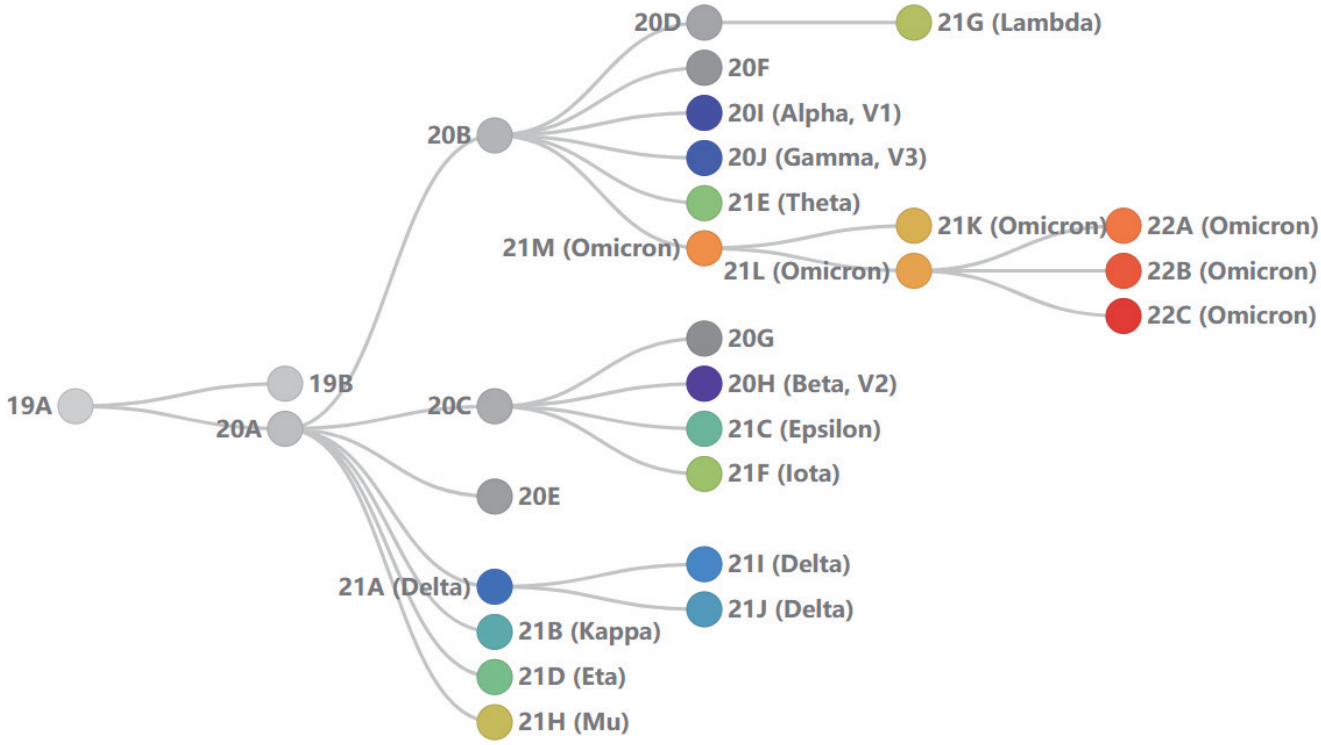


Figure 6. Virus lineage diagram of GW 4-1-automation library

Summary

In this study, the ATOPlex RNA Library Prep Set V3.1 combined with the MGI automation systems (MGISP-100 & MGISP-Smart 8) were utilized for library preparation. Subsequently, DNBSEQ-G99 and MGI metargetCOVID software were used for sequencing and following identification and tracing for SARS-CoV-2, respectively. The results showed that the data generated from this combinational products had high quality, wide coverage and deep sequencing depth, and the identification and typing results were accurate. Notably, the above results were comparable to the counterpart obtained from manual approach.

The ATOPlex technology involved in this study is capable of millionfold amplification and enrichment of extremely low-level viruses. In addition, MGISP-Smart 8 and MGISP-100 automated systems adopted in this study can not only achieve a flexible library preparation completing 8-40 samples/run, but also accomplish pooling, normalization and quantitative reagents preparation, improving the automated efficiency, ensuring the accuracy of results, and saving labor costs. Meanwhile, the DNBSEQ-G99 used for subsequent sequencing, based on the patented DNBSEQ™ technology, can stably obtain high-quality sequencing data. The throughput of a flow cell is 80M reads, and up to two flow cells can be run simultaneously. The whole process of SE100+10+10 sequencing can be completed within 4.5 hours, maximizing sequencing efficiency. The built-in computing module integrates sequencing and bioinformatics, enhancing the sequencing efficiency to the best.

The MGI ATOPlex library preparation solution equipped with the automated system and DNBSEQ sequencing platform, coupled with the self-developed data processing system for analysis, can provide a complete and reliable combinational products for the recognition, identification and tracing of pathogenic microorganisms. This solution can make the monitoring of pathogenic microorganisms faster and easier, and comprehensively contributes to the research of various infectious diseases.

MGISP-Smart 8RS Automated Sample Preparation System

DNBSEQ-G99 Genetic Sequencer

References

1. McCombie W R, McPherson J D, Mardis E R. Next-generation sequencing technologies[J]. *Cold Spring Harbor perspectives in medicine*, 2019, 9(11).
2. Suwinski P, Ong C K, Ling M H T, et al. Advancing personalized medicine through the application of whole exome sequencing and big data analytics[J]. *Frontiers in genetics*, 2019, 10: 49.
3. Li R, Shen X, Chen H, et al. Developmental validation of the MGIEasy Signature Identification Library Prep Kit, an all-in-one multiplex system for forensic applications[J]. *International Journal of Legal Medicine*, 2021, 135: 739-753.
4. Yang L, Wu Y, Hu Z, et al. Simultaneous detection of fetal aneuploidy, de novo FGFR3 mutations and paternally derived β-thalassemia by a novel method of noninvasive prenatal testing. *Prenat Diagn*. 2021 Mar;41(4):440-448.
5. Ni G, Lu J, Maulani N, et al. Novel multiplexed amplicon-based sequencing to quantify SARS-CoV-2 RNA from wastewater[J]. *Environmental Science & Technology Letters*, 2021, 8(8): 683-690.
6. Temerozo J R, Fintelman-Rodrigues N, Santos M C, et al. Human endogenous retrovirus K activation in the lower respiratory tract of severe COVID-19 patients associates with early mortality[J]. 2021.
7. Hess J F, Kohl T A, Kotrová M, et al. Library preparation for next generation sequencing: A review of automation strategies[J]. *Biotechnology advances*, 2020, 41: 107537.
8. Holland I, Davies J A. Automation in the life science research laboratory[J]. *Frontiers in Bioengineering and Biotechnology*, 2020, 8: 571777.
9. Kang J, Tahir A, Wang H, et al. Applications of nanotechnology in virus detection, tracking, and infection mechanisms[J]. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 2021, 13(4): e1700.
10. Baker R E, Mahmud A S, Miller I F, et al. Infectious disease in an era of global change[J]. *Nature Reviews Microbiology*, 2022, 20(4): 193-205.
11. Goud K Y, Reddy K K, Khorshed A, et al. Electrochemical diagnostics of infectious viral diseases: Trends and challenges[J]. *Biosensors and Bioelectronics*, 2021, 180: 113112.

Recommended Ordering Information

Category	Product	Cat. NO.
Instruments	Genetic Sequencer DNBSEQ-G99ARS	900-000609-00
	MGISP-100RS Automated Sample Preparation System	900-000206-00
	MGISP-Smart 8RS Automated Sample Preparation System	900-000503-00
Software	MGI metargetCOVID	970-000228-00
Sample extraction	MGIEasy Nucleic Acid Extraction Kit (T-1728)	1000020261
	MGIEasy Nucleic Acid Extraction Kit（T-96）	1000020471
Library Prep	ATOPlex RNA Library Prep Set V3.1（16RXN）	940-000132-00
	DNBSEQ OneStep DNB Make Reagent Kit (OS-DB)	1000026466
Sequencing Reagents	DNBSEQ-G99RS High-throughput Sequencing Set (FCL SE100/PE50)	940-000409-00

* For more ATOPlex customized products, please visit MGI's official website: https://en.mgi-tech.com/products/atoplex/.

MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083

Version: April 2024

+86-4000-688-114
en.mgi-tech.com
MGI-service@mgi-tech.com

Authors: Zhang Hanfei, Han Wanru
Editor-in-Charge: Wang Qiwei
Reviewer: Jiang Yao

1. For StandardMPS and CoolMPS: Unless otherwise informed, StandardMPS and CoolMPS sequencing reagents, and sequencers for use with such reagents are not available in Germany, Spain, UK, Sweden, Belgium, Italy, Finland, Czech Republic, Switzerland, Portugal, Austria and Romania. Unless otherwise informed, StandardMPS sequencing reagents, and sequencers for use with such reagents are not available in Hong Kong. No purchase orders for StandardMPS products will be accepted in the USA until after January 1, 2023.
2. For HotMPS sequencers: This sequencer is only available in selected countries, and its software has been specially configured to be used in conjunction with MGI's HotMPS sequencing reagents exclusively.
3. For HotMPS reagents: This sequencing reagent is only available in selected countries.