MGI

Microorganism RNA Sequencing Package

FAQ

Product Overview

The MGI Microorganism RNA Sequencing Package is based on the reagents, the automated sample preparation system, the high-throughput sequencing platform, and the data processing system, all are developed by MGI independently. It can cover the whole process from nucleic acid extraction to report output. It is committed to performing rapid, accurate, and comprehensive RNA sequencing and data processing for unknown microorganisms that cause infections in humans or animals.

This instruction summarizes some common problems when using this package so that users can complete RNA sequencing of microorganisms smoothly.

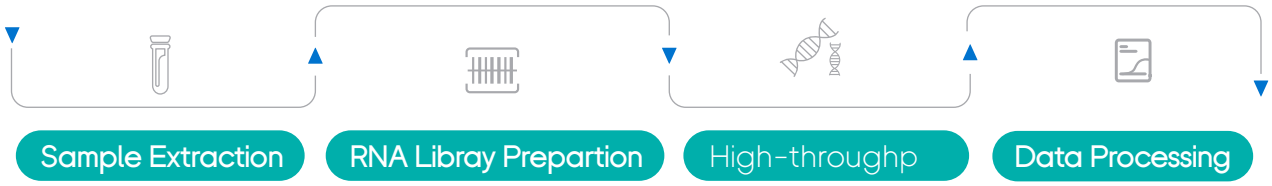

Figure 1 The Workflow of RNA Sequencing of Microorganisms

FAQ

Sample Extraction

Which are the suitable sample types for the MGI Microorganism RNA Sequencing Package

This package uses the MGIEasy Nucleic Acid Extraction Kit for samples' nucleic acid extraction. It is suitable for respiratory tract samples such as throat swabs, nasal swabs, and alveolar lavage fluid.

How does the MGIEasy Nucleic Acid Extraction Kit work?

The kit contains superparamagnetic beads which can effectively bind nucleic acids. The MLB buffer can lyse viruses and release viral DNA/RNA. After adding ethanol, superparamagnetic beads can adsorb nucleic acids through hydrogen bonds and electrostatic interactions without adsorbing proteins and impurities. The magnetic beads bound to nucleic acids are washed with MW1 to remove non-specifically adsorbed proteins or proteinase K and then washed with MW2 to remove the salts adsorbed on the magnetic beads. Finally, it is eluted with nuclease-free water to obtain a high-purity nucleic acid solution.

How to avoid aerosol contamination?

PCR amplification will produce many amplified products, which can increase the risk of aerosol contamination. So strict operation area division requires during the experiment. The operation areas should divide into a pre-PCR area and a post-PCR area at least.

Table 1 The Operation Partition

Area	Pre-PCR Area	Post-PCR Area
Experimental operation	Nucleic Acid Extraction;	RNA Library Preparation 2 (Library Preparation) PCR Products;
	rRNA Depletion;	DNB Preparation;
	RNA Library Preparation 1 (Double-stranded cDNA Synthesis);	Sequencing Reagent Cartridge Preparation
	RNA Library Preparation 2 (Library Preparation) Reagents Preparation	

How to inactive potentially virus-containing samples?

Samples can incubate at 56℃ for 30 minutes.

Pre-extraction treatments for various sample types.

- Throat swabs, nasal swabs: if swabs store in preservation solution, please aspirate the supernatant for nucleic acid extraction; if there is no preservation solution, please immerse swab tips in 500 µL 1×PBS buffer (pH 7.4), mix and aspirate the supernatant for nucleic acid extraction.
- Alveolar lavage fluid: mix and aspirate the supernatant for nucleic acid extraction.

Are components of nucleic acid extraction products DNA or RNA?

The MGIEasy Nucleic Acid Extraction Kit mainly uses for viral nucleic acid extraction, including DNA and RNA. But it is not applicable for nucleic acid extraction from other microorganisms. Extraction products are mixed solutions of DNA and RNA. Since extraction products are used for high-throughput sequencing, adding enhancer buffer is unnecessary during the extraction process. When using MGISP-100RS for automatic extraction, the preparation method of lysate & binding buffer is as follows: 160 µL buffer MLB, 200 µL absolute ethanol, 15 µL proteinase K solution, and 15 µL magnetic beads M.

How about the purity and concentration of extraction products? How to deal with it if the concentration is too low?

Some clinical samples contain relatively low levels of viral nucleic acid, such as throat swabs and nasal swabs. Therefore, they are not suitable to measure the purity of viral nucleic acid using OD260/OD280 and their concentration by Qubit. If samples with known viral species have to know the exact viral concentration, qPCR is recommended for quantification.

Can buffer mixtures be prepared in advance and stored for a long time?

Buffer MLB and absolute ethanol in the lysate & binding buffer can be mixed in advance and stored for a long time. Please add other reagents before use, such as proteinase K solution and magnetic beads M. The well-prepared lysate & binding buffer needs to be used within 30 minutes.
MW1 and MW2 can be added with absolute ethanol and stored for a long time.

What is the throughput and time for automated nucleic acid extraction using MGISP-100RS?

Under different extraction throughputs, the running time is different.

Table 2 Throughputs and Time for Automated Nucleic Acid Extraction

Throughput	8 rxn	16 rxn	24 rxn	32 rxn
Running Time	40 min	55 min	70 min	80 min

In the process of nucleic acid extraction, do you need filter tips for manual operation?

We need to use filter tips to reduce the risk of RNA degradation and cross-contamination between samples.

RNA Library Preparation

Do samples need to use DNase I before RNA library preparation?

It depends on the amount of nucleic acid and DNA ratio in extracted samples. DNase I can be used for DNA digestion if extracted products have enormous RNA yield and a high proportion of genomic DNA contamination. However, DNase I digestion will cause a certain amount of RNA loss, so the total RNA input needs to be appropriately increased by 20% to 30% compared to the expected input for library preparation. Typically, the amount of nucleic acid extracted from samples such as throat and nasal swabs is low, so DNase I digestion is not required.

Does the host rRNA need to use for RNA library preparation?

Different strategies can select for types of samples and needs. The MGIEasy rRNA Depletion Kit can remove rRNA from the host's total RNA, including humans, mice, and rats.

If rRNA depletion perform on RNA samples, it benefits to improve the utilization of sequencing data. However, it increases the loss of RNA samples and declines library yield, which is particularly unfavorable for the library preparation of low-input samples. If rRNA depletion do not perform, it can minimize the loss of RNA samples to a large extent. Nevertheless, more invalid data (rRNA data) will be generated during sequencing, reducing data utilization. Therefore, users can comprehensively select a suitable method for library preparation based on their specific sample conditions and experimental needs.

The following table shows that effective reads per sample are recommended under different sequencing strategies when rRNA has been removed from RNA samples.

Table 3 Recommended Sequencing Strategies for the MGI Microorganism Sequencing Package

Application	Recommended Read Length	Recommended Flow Cell	Recommended Sample Throughout/ Run	Effective Reads Per Sample (M)
Microorganisms Fast Identification	SE50	FCS	4	≥ 20
Platform of microorganisms Fast Identification and Assembly Evolution	PE100	FCL	4	≥ 100

In the process of RNA library preparation, do you need filter tips for manual operation?

We need to use filter tips when the nucleic acid in products still exists in the state of RNA to reduce the risk of RNA degradation and cross-contamination between samples, such as depleting rRNA and performing reverse transcription.

Why do adapters need to be diluted in some cases?

Whether adapters need to dilute or how to dilute them depends on the RNA input of library preparation. Adding too many adapters during ligation reaction will result in a large number of adapters or adapters dimers remaining in libraries, which can diminish the amount of available sequencing data.

Which modules are in RNA library preparation? And which modules can use MGISP-100RS?

RNA library preparation includes four main modules: rRNA depletion, RNA library preparation 1 (double-stranded cDNA synthesis), RNA library preparation 2 (library preparation), and DNB preparation.

Among them, MGISP-100RS can prepare 16 libraries simultaneously, including three modules of rRNA depletion, RNA library preparation 1 (double-stranded cDNA synthesis), and RNA library preparation 2 (library preparation). The DNB preparation requires manual operations.

Table 4 Modules of RNA Library Preparation

Module	Operation	Running Time
rRNA Depletion	Automated	2 h 20 min
RNA Library Preparation 1 (Double-Stranded cDNA Synthesis)	Automated	2 h 45 min
RNA Library Preparation 2 (Library Preparation)	Automated	4 h
DNB Preparation	Manual	40 min

In the process of using MGISP-100RS, how about the compatibility of reagent dosage for library preparation?

When using MGISP-100RS, the actual amount of used reagents is generally higher than the theoretical amount.

Therefore, when using MGISP-100RS for rRNA depletion, reagents in the MGIEasy rRNA Removal Kit (32 RXN) (MGI, 1000005953) have to dispense manually before automation. Reagents in this kit can only be used for one automated experiment of 16 RXN.

When using MGISP-100RS for RNA library preparation, if customers use the MGIEasy RNA Library Prep Kit (16RXN) (Cat. No. 1000005274) in the MGIEasy Microorganism Rapid RNA Library Prep Set (16RXN) (Cat. No. 940-000107-00), reagents can be directly used in automated experiments without dispensing.

Suppose customers use the MGIEAsy RNA Library Prep Kit (96 RXN) (Cat. No. 1000005276) in the MGIEasy Microorganism Rapid RNA Library Prep Set (96RXN) (Cat. No. 940-000108-00) for RNA library preparation. In that case, manual dispensing reagents is required (see the table below). The reagents in the kit can perform about four automated experiments with 16 RXN.

Table 5 Dispensing Volume of Each Component of the MGIEAsy RNA Library Prep Kit (96 RXN)-RNA Library Preparation 1 (Synthesis of Double-stranded cDNA)

	Reagent	Dispensing Volume
MGIEAsy RNA Library Prep Kit (96 RXN) (Cat. No. 1000005276)	Fragmentation Buffer	105 µL
	RT Buffer	105 µL
	RT Enzyme Mix	28 µL
	Second Strand Buffer	490 µL
	Second Strand Enzyme Mix	88 µL

Table 6 Dispensing Volume of Each Component of the MGIEAsy RNA Library Prep Kit (96 RXN)-RNA Library Preparation 2 (Library Preparation)

	Reagent	Dispensing Volume
MGIEAsy RNA Library Prep Kit (96 RXN) (Cat. No. 1000005276)	ERAT Buffer	143 µL
	ERAT Enzyme Mix	64 µL
	Ligation Buffer	475 µL
	DNA Ligase	42 µL
	PCR Enzyme Mix	500 µL
	PCR Primer Mix	110 µL

What is the quality control standard of PCR libraries?

Library quality control indexes include fragment distribution size and library concentration. When experimental procedures of identical types of samples are sufficiently stable, it is not necessary to detect the fragment distribution size of each library.

Table 7 Quality Control Standard of PCR Libraries

Quality Control Index	Quality Control Range	Measurement Method
Fragment Distribution Size of PCR Products	150-500 bp	Agilent 2100 Bioanalyzer detection
Concentration of PCR Products	>2.5 ng/μL	Qubit dsDNA quantitative detection

Why is DNB concentration low in some cases?

Please do the following troubleshooting:

- Whether experimental operations are carried out strictly according to the instructions.
- Whether PCR libraries meet the quality control standard.
- When re-preparing DNB, increase the amount of PCR library input appropriately.

What should you do if sequencing reagents have been thawed but cannot be used on time?

- If a kit has been thawed (including dNTPs) and cannot be used on time, it can be frozen and thawed at most once;
- If a kit has been thawed (including dNTPs) and cannot be used on time, it can be temporarily stored at 4°C and used within 24 hours. The reagent cartridge needs to be re-mixed before use;
- If dNTPs and enzymes have been added to a reagent cartridge, the reagent cartridge has been well prepared. If it cannot be used in time, it can be temporarily stored at 4°C and used within 24 hours. The reagent cartridge needs to be re-mixed before use;
- If a reagent cartridge has been well prepared when dNTPs and enzymes have been added. Then sequencer's needle has been inserted into it. If it cannot be used in time, make sure to seal it with tinfoil, store it at 4°C, and use it within 24 hours. Gently mix the reagent cartridge before use. Be careful not to spill the reagent from needle holes when mixing to avoid contamination between the reagents in each hole.

What is the quality control standard of PCR libraries?

Library quality control indexes include fragment distribution size and library concentration. When experimental procedures of identical types of samples are sufficiently stable, it is not necessary to detect the fragment distribution size of each library.

Data Processing

Can PFI be only used for data processing of human samples, whether can it be used for samples from other origins?

The database collected by PFI covers more than 26000 reference genomes of microorganisms, so it is theoretically suitable for the classification requirements involved in this database. However, PFI has only tested sample data from humans and some animals, which has not yet expanded to other environmental samples, such as water bodies, soil, and so forth.
In addition, nucleic acids extracted from environmental samples such as water bodies are usually in a low quantity and poor quality, making it difficult to guarantee the success of metatranscriptomics.

Which host reference sequences are preset in PFI?

PFI presets host sequences for Humans, pigs, goats, sheep, mice, rats, carps, domestic geese, chickens, ducks, cattle, cats, dogs, and rabbits.

What are principles of strains collection for the reference database, and what information is included in the database? How to update and how often?

The reference database is derived from the complete genome sequences of RefSeq, including bacteria (including archaea), viruses, fungi, and parasites. It can be updated remotely, usually once a year.

Why does PFI use NCBI RefSeq as the species identification reference database?

NCBI RefSeq has been strictly verified and entered by NCBI. Compared with other NCBI databases, sequences from NCBI RefSeq are more reliable.

What is the difference between MGI's PFI and BGI's PMseq™ ?

- PFI from MGI is designed for scientific research use. It offers all the identification results, and users can interpret and verify the reports according to their own experience and needs.
- BGI's PMseq™ is mainly used in a clinical scene. Its report generates after interpretation and manual review and only represents the most likely pathogens identified in samples.

Why do proportions vary in different samples for unknown species and numbers of specific reads identified by PFI?

- It is not surprising to obtain different results from various types of samples since samples resources are very complicated and samples are prepared for metatranscriptomic sequencing. For some hosts or environments, their relevant microorganisms included in public databases are relatively few. Correspondingly, there are fewer microorganisms input in our database, so microorganisms from these scenarios are challenging to identify;
- If a kit used for original sample extraction is not specifically designed for microorganisms, it will also cause a part of microbial nucleic acids loss. Moreover, microorganisms in samples account for a limited proportion, resulting in fewer microorganisms identified by PFI;
- If the host sequence is not close and the removal is not performed thoroughly, it will also decrease the proportion of specific reads identified.

What comparison strategy is used in the taxa identification analysis process?

Reads are split into kmers. Then Kmers are assigned by taxonomic labels according to the reference database pre-built. If most kmers of a read are assigned by taxonomy A, then the read will be identified as taxonomy A.

Which scenarios does the MGAP suit?

MGAP aims to assemble interested microbial genomes and trace their phylogenetic relationships. For multi-species sequencing data, PFI must first be used for microbial species identification and reads classification. Next, users can extract targeted sequencing data for MGAP assembly. Based on recommended effective reads per sample (100 Mb per sample for PE100) of the MGI Microorganism Sequencing Package, assembling a viral genome is only recommended. If it is required to assemble the bacterial genome, the increase in effective reads of each sample is needed according to the different conditions of samples.

Contact Us

MGI Tech Co.,Ltd
Address: Building 11,Beishan Industrial Zone, Yantian District,
Shenzhen, CHINA 518083
Email: MGI-service@mgi-tech.com
Website: www.en.mgi-tech.com
Tel: 4000-688-114
Version： November 2022 | MGPB011820205

https://www.linkedin.com/com pany/mgi-bgi

https://twitter.com/MGI_BGI

*All Sequencers and Sequencing Reagents are not available in Germany, the US and Spain.