MGI
High Throughput Sequencing Platform

Innovate Life Technology
MGI Tech Co., Ltd

MGI
华大智造

*MGISEQ-2000



About MGI

MGI Tech Co., Ltd. (MGI) a subsidiary of BGI Group, is committed to enabling effective and affordable healthcare solutions for all. Based on its proprietary technology, MGI produces sequencing devices, equipment, consumables and reagents to support life science research, medicine and healthcare. MGI's multi-omics platforms include genetic sequencing, mass spectrometry and medical imaging. Providing real-time, comprehensive, life-long solutions, its mission is to develop and promote advanced life science tools for future healthcare.

MGI 华大智造

Outline

About MGI 01
MGI Sequencing platform 03
MGI Sequencing technology 04
MGI Genetic Sequencer 05
- MGISEQ-2000 05
- MGISEQ-200 06

Automated sample preparation system 07
- MGISP-100 07

Library Preparation for Sequencing 08
- Whole Genome Sequencing 09
- Whole Exome Sequencing 11
- mRNA/ Total RNA Sequencing 13
- Small RNA Sequencing 16
- Methylation Sequencing 17
- Metagenomics Sequencing 19
- Rad Sequencing 20
- Library Preparation Modules 20

Ordering information 21

MGI Sequencing platform

MGISEQ Series adopt combinatorial Probe-Anchor Synthesis (cPAS) and DNA Nanoballs (DNB) technology to provide high sequencing accuracy and efficiency to our customers. DNB technology enables rolling-circle replication (RCR), which ensures that each copy is amplified from the original template and no replication errors are generated. cPAS chemistry allows accurate imaging and base calling. Currently, MGI has developed a series of sequencers, automated sample preparation systems and libraries preparation kits for various applications.



*MGISEQ-200

MGI sequencing technology

DNB

DNB is the unique technology that allows DNA linear amplification in a single-tube solution. The workflow includes DNA fragmentation, adapter ligation and single-stranded circulation to produce DNB.

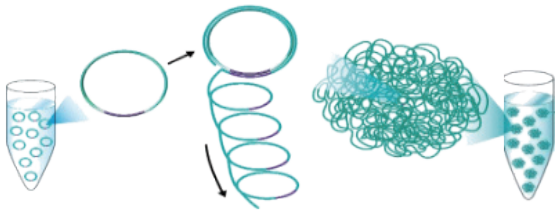
▶ High intensity of fluorescent signals and high sequencing accuracy

RCR increases the quantity of “DNA fragments (to be processed)” and thus enhance the intensity of fluorescent signals and accuracy.

▶ Zero-error accumulation using RCR

RCR enables each copy to be amplified from the original template and generate no replication errors, delivering high sequencing accuracy.

▶ Maximized efficiency of DNA loading on patterned array



Patterned array

Patterned array is the DNB loading technology. The coating surface of the semiconductor chip makes each loading spot positively charged. Therefore, negatively charged DNB can easily attach to the positively charged chip surface through electrostatic adhesion. Patterned array ensures each spot is attached to a single DNB and thus prevents the signal interference. Overall, patterned array enables high sequencing accuracy and high chip utilization.

▶ Increasing sequencing signals using nanoscale patterned chip

The Patterned Array of spots on Flow Cells prevents cross interference between fluorescent signals and enable high intensity of signals.

▶ Precision machined semiconductor chip enhances the attachment of DNB to chip

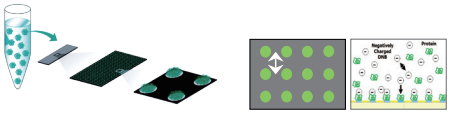
The size of DNB matches the area of effective spot on the chip. Thus, each spot is only attached to a single DNB, avoiding the signal interference.

▶ Precise pattern enables efficient imaging system and DNB loading

Patterned array and DNB technology maximize the utilization of imaging system and chip surface area, and sequencing accuracy.

▶ Low Duplication，No Index Hopping

DNB is amplified in solution and loaded without PCR reagents such as polymerase, primers and dNTP. Hence the duplicate rate is extremely low using DNB technology and patterned array.



cPAS

cPAS is the innovative technology developed by BGI. By improving combinatorial Probe-Anchor Ligation(cPAL), cPAS increases the sequencing accuracy with higher speed and read length of sequencing.

▶ MGI proprietary enzymes react completely with 60 seconds

Our biochemistry team has studied a large quantity of reaction conditions and screened tens of thousands of sequencing enzymes to successfully complete biochemical reactions within just 60 seconds.

▶ Real-time sub-pixel registration, image processing and base calling

Advanced Real-time Image Processing software, Sub-pixel Registration and Multi-thread Parallel Compression Algorithms result in accurate real-time imaging and base calling and high industry-leading data-processing speed. Furthermore, GPU empowers MGISEQ-2000 with high speed to analyze high throughput data in real-time.

MGI Genetic Sequencer

MGISEQ-2000

Versatile and flexible large-scale sequencer

The MGISEQ-2000 provides cost-effective solutions for exome, transcriptome and whole-genome sequencing. Its dual flowcell (FC) system supports two FC and several different read length to meet a wide range of sequencing needs.



Features

- Competitive pricing and low running cost
- Various read lengths, short turnaround time, a wide range of applications
- Multiple throughput modes support different sequencing project needs

Performance parameters

FCS chip					
Lane/chip	2 lanes	Read length	PE150	Run time	≤48 hours
Number of reads	Single chip: ~375 M Dual chip: ~750 M	Output per run	Single chip：~112.5 Gb Dual chip：~225 Gb		
FCL chip					
Lane/chip	4 lanes	Read length	SE50, SE100*, PE50*, PE100, PE150, SE300*	Run time**	≤48 hours
Number of reads	Single chip:~1800 M Dual chip:~3600 M	Output per run	Single chip：75 ~ 540 Gb Dual chip：150 ~ 1080 Gb		
Data quality	Q30≥85% ***				

* SE300 is in development.SE100 and PE50 are customized.
** Run time for PE100 with dual chip at full capacity.
*** The percentage of bases > Q30 is averaged over the entire run. This result is performed under BGI-E8 quality control library.

Application

Research

WGS, WES, targeted sequencing, whole transcriptome sequencing, RNA-Seq, non-coding RNA sequencing, ChIP-Seq, WGBS, ddRAD sequencing, single-cell sequencing (DNA, RNA)

Clinical Application

NIPT, PGS/PGD, Chromosome CNV detection, complex disease detection, monogenic disorder detection, tumor gene mutation detection and pathogenic microorganism detection

MGISEQ-200

Dedicated high-efficiency desktop sequencer

MGISEQ-200 is a dedicated benchtop sequencer which meets different requirements of read length, time and output. Its key applications are targeted sequencing, small WGS and pathogen screening.



Features

- Various read length options and short turnaround time
- Build-in automated system saves the costs of reagents
- Integrated LIMS (Laboratory Information Management System) enables the fully-automated sample tracking

Performance parameters

FCS					
Lane/flow cell	1 lanes	Read length	SE50, SE100*, PE50*, PE100, SE200*	Run time**	≤48 hours
Number of reads	~300 M	Output	15 ~ 60 Gb		
Data quality	Q30≥75% ***				

* SE200 is in development. SE100 and PE50 are customized.
** Run time for PE100 at full capacity
*** The percentage of bases > Q30 is averaged over the entire run. This result is performed under BGI-E8 quality control library

Application

Research

small WGS, targeted sequencing, WES, transcriptome sequencing, RNA-seq, small RNA sequencing, ChIP-Seq and single-cell sequencing (DNA/RNA)

Clinical Application

NIPT, PGS/PGD, Chromosome CNV detection, pathogenic microorganism detection, tumor gene mutation detection, complex disease detection and monogenic disorder detection

Solutions for a wide range of applications

Sequencer	MGISEQ-200	MGISEQ-2000	
Highlights	Rapid and efficient solution for targeted sequencing and small genome sequencing	Cost-effective solution for genomic research	
Key application	Targeted sequencing, microbial genome sequencing	Exome, transcriptome, genome sequencing etc.	
Flow cell type	FCS	FCS	FCL
Lane/chip	1 lane	2 lanes	4 lanes
Throughput	Mid throughput	Mid throughput	High throughput
Chip/run	1	2	2
Data output	~ 60 Gb	~ 225 Gb	~ 1080 Gb
Run time*	≤48 hours	≤48hours	≤48hours
Number of reads/ chip	~300 M	~375 M	~1800 M
Maximum read length	PE100	PE150	PE150

* Run time for PE100 at full capacity

Automated sample preparation system

MGISP-100

Automated sample preparation system MGISP-100 is a dedicated library preparation platform for NGS. Its automated workflow enables batch operation and eliminates complicated manual operation, delivering affordable reproducible results on NGS library preparation.



Function

- Automated nucleic acid extraction, PCR, library preparation etc.
- Supports a wide range of NGS applications, such as NIPT, PGS, human whole genome sequencing, pathogen identification *, single bacteria identification*

* kits will be released in the future, please see www.mgitech.cn for updated information

Features

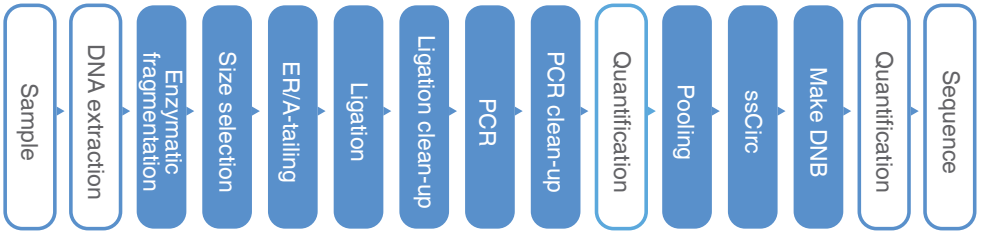
- One stop solution integrating PCR, temperature control and many functional modules
- Simplified library preparation on a single instrument. Minimal labor requirement and human errors
- Cost-effective, short turnaround time. Batch operation is ideal for small and medium laboratory
- Simple-to-follow interface and instruction
- Real-time sample tracking
- LIMS (Laboratory Information Management System) supports automated sample tracking throughout the sequencing workflow
- Full contamination control. Build-in ISO5 standard cleanroom fan filter unit
- Open platform and customizable system setting

Performance parameters

Sample/run	16 samples		
Turnaround time*	Manual operation 1hour + run time 6hours		
Sample type	plasma, saliva, FFPE, gDNA, WGA product etc.		
Volume of liquid handler	2 - 200 μl		
Accuracy of liquid handler accuracy	liquid	2 μl	200 μl
	CV	< 5%	< 1%
	accuracy	< ±10%	< ±1%

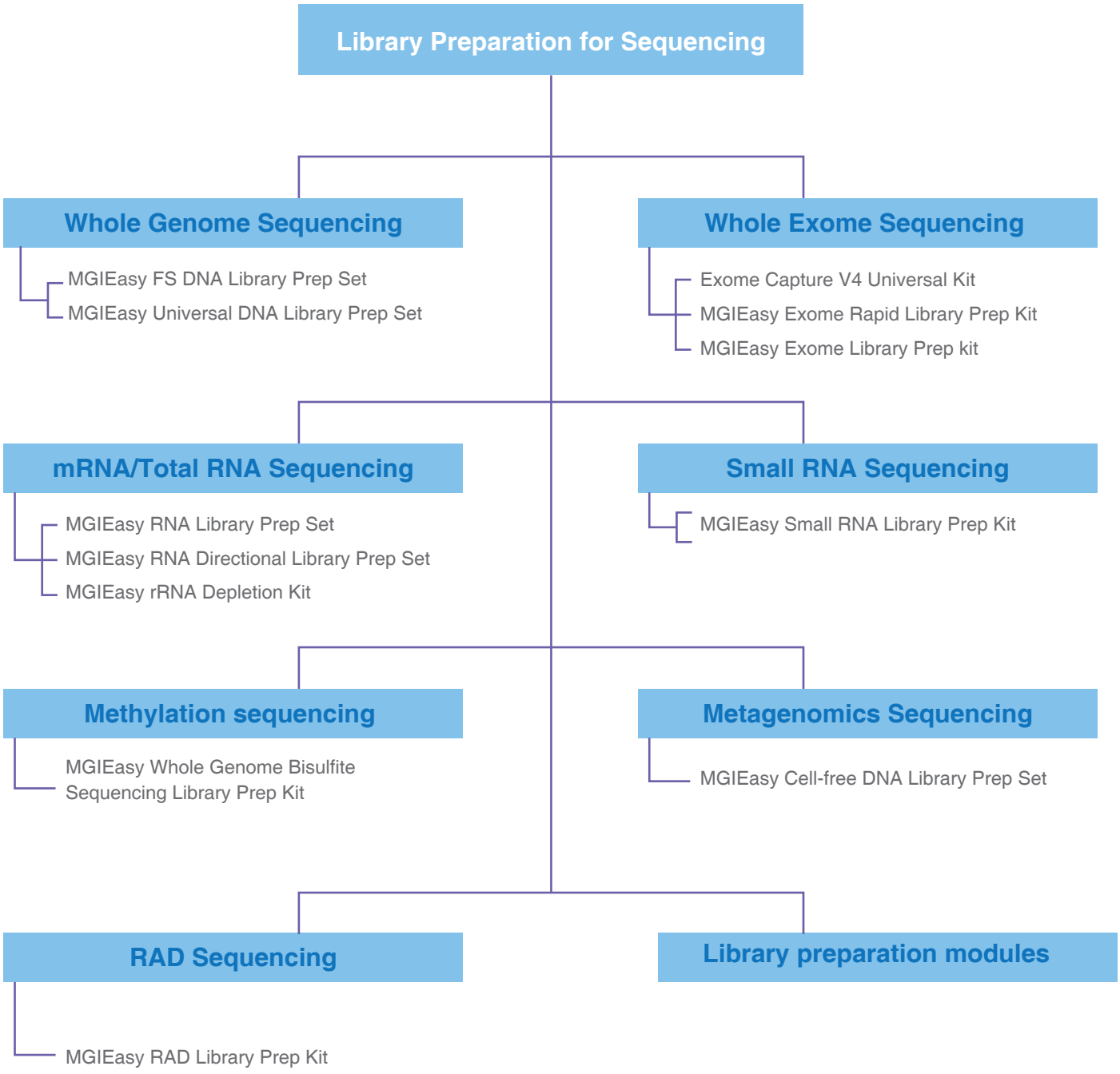
* Turnaround time of library preparation for human whole genome sequencing. Turnaround time might be different between library preparation for different applications.

The automated library preparation workflow



Note: white shows manual operation and blue shows automated steps

Library Preparation for Sequencing



Whole Genome Sequencing

Whole-Genome Sequencing (WGS) is the process by which an organism's entire genome, both nuclear and cytoplasmic, is unraveled and read. It is useful in elucidating the underlying genomic details in studies at either an individual or population scale. WGS has applications in selective breeding, population evolution, disease research, clinical diagnosis and drug developments.
The MGISEQ-based WGS solutions can accurately and rapidly identify the distribution of SNPs across the genome and annotate the variants based on the genome database.

Recommended MGI Process

Library preparation	Sequencing	Bioinformatics Analysis
MGIEasy FS DNA Library Prep Set MGIEasy Universal DNA Library Prep Set Automated sample preparation system MGISP-100	MGISEQ-2000 MGISEQ-200	MegaBOLT Bioformics analysis accelerator (WGS Analysis System)

WGS library preparation

MGI has developed a series of DNA library preparation kits for WGS to meet the needs of various sample types and applications.

▶▶ MGIEasy FS DNA Library Prep Set

The MGIEasy FS DNA Library Prep Set can quickly prepare 0.5-300 ng genomic DNA into a library using enzyme fragmentation method.

▶▶ MGIEasy Universal DNA Library Prep Set

The MGIEasy Universal DNA Library Prep Set provides an efficient and universal library construction solution for various DNA sample types such as fragmented gDNA, FFPE DNA, Meta DNA, Amplicons, cfDNA, ChIP DNA and capture DNA.

Product specification

Products	MGIEasy FS DNA Library Prep Set	MGIEasy Universal DNA Library Prep Set
Assay Time	<5 h	<4 h
Input Quantity	5- 300 ng gDNA	0.5-50 ng fragmented DNA (200-500 bp)
Sample types	gDNA, FFPE DNA, Meta DNA, cfDNA	gDNA, FFPE DNA, Meta DNA, Amplicons, cfDNA, ChIP DNA, capture DNA, etc.
Species Compatibility	Human, animals, plants, fungi and bacteria, such as mouse, rice, Arabidopsis, yeast and E.coli.	
Fragmentation method	Enzyme	Physical
Automation option	Automation-friendly: MGISP-100	Not available
Applications	Human genome sequencing, Plant and Animal genome sequencing, Microbial genome sequencing, Metagenomics sequencing	
Platform Compatibility	MGISEQ-2000, MGISEQ-200	
Recommended Read Length	SE50, PE100, PE150 and etc.	

Features

- Low DNA input
- Compatible with intact and degraded specimens, including FFPE, plasma, and prokaryotic samples with different GC content from human, animal, plant and fungi samples
- Easy to use and suitable for automatic library preparation instrument
- Outstanding quality sequencing data with excellent coverage of genome and variation detection capacity enabling excellent usability of data

Performance

▶ Low duplicate rate and high genome coverage

With low duplicate and high genome coverage rates, the MGI library sequencing data yields a very high proportion of useful reads.

Sample	Average Sequencing Depth	Mapping rate	Duplicate rate	Mismatch rate	Coverage (≥1X)	Coverage (≥4X)	Coverage (≥10X)	Coverage (≥ 20X)
NA12878	31.2X	99.56%	2.23%	0.57%	99.09%	98.53%	97.50%	92.56%

Library was prepared from 1ug input Universal Human Reference Genome (NA12878) gDNA using the MGIEasy Universal DNA Library Prep Set and sequenced on MGISEQ-2000 at PE100 read-length.

▶ Excellent accuracy and sensitivity on SNPs and Indels

MGI sequencing platforms demonstrated high accuracy and sensitivity on SNPs and Indel detection. MGI sequencers showed better Indel detection than Illumina platform.

Variation	True-pos	False-pos	False-neg	Precision	Sensitivity
SNPs	3,195,908	12299	14349	99.62%	99.55%
Indel	456,675	36284	24592	92.64%	94.89%

Library was prepared from 1ug input NA12878 gDNA using the MGIEasy Universal DNA Library Prep Set and sequenced on MGISEQ-2000 at PE100 read-length.

Whole Exome Sequencing

Whole Exome Sequencing (WES) is a targeted sequencing technology that focuses on only 2% of an entire region's genome, but covers most of the gene coding regions and >99% of variations in the ClinGen database. Data from WES can be used to detect SNVs, indels, gene arrangements and copy number variations (CNVs). The decreasing cost of DNA sequencing has made WES more accessible, resulting in widespread adoption of the technology for genetic disease scanning, tumor molecular diagnosis and immunotherapy. The MGISEQ-based WES solutions deliver full coverage of coding regions and help identify the exomes with the most significant phenotypic impact.

Recommended MGI Process

Library preparation	Sequencing	Bioinformatics Analysis
Exome Capture V4 Universal Kit MGIEasy Exome Rapid Library Prep Kit MGIEasy Exome Library Prep kit	MGISEQ-2000 MGISEQ-200	MegaBOLT Bioformics analysis accelerator (WES Analysis System)
Automated sample preparation system MGISP-100		

WES library preparation

▶▶ Exome Capture V4 Universal Kit

The Exome Capture V4 Universal Kit is the first exome capture kit designed based on Chinese genetic background. The probe is designed to target only coding regions from RefSeq, CCDS, GENCODE & UCSC Known Genes with a design size of 59 Mb with 22634 exons, and provide excellent coverage of targeted regions with minimal sequencing, allowing more samples to be pooled per flow cell. It is compatible with gDNA, FFPE DNA and plasma sample with an input of 500 ng Genomic DNA PCR amplification. As a universal exome capture kit, it is compatible with MGI high throughput sequencing platforms as well as other NGS platforms.

▶▶ MGIEasy Exome Rapid Library Prep Kit

The MGIEasy Exome Rapid Library Prep Kit can quickly construct a WES library from 300 ng genomic DNA using an enzyme fragmentation method in conjunction with Exome Capture V4 Universal Kit.

▶▶ MGIEasy Exome Library Prep Kit

The MGIEasy Exome Library Prep Kit can prepare 1 μg genomic DNA into a WES library using a non-enzymatic fragmentation method combined with Exome Capture V4 Universal Kit.

Product specification

Products	MGIEasy Exome Rapid Library Prep Kit	MGIEasy Exome Library Prep kit
Assay Time	<6.5 h	
Input Quantity	300 ng gDNA	1 ug gDNA
Sample type	tissues, plasma and FFPE sample	tissues, plasma and FFPE sample
Fragmentation method	enzyme	physical
Compatible with major target-enrichment platforms	MGI Exome Capture, Agilent SureSelect, Roch Nimblegen SeqCap EZ, and IDT xGen Lockdown Probes	
Platform Compatibility	MGISEQ-2000, MGISEQ-200	
Recommended Read Length	PE100, PE150 and etc.	

Performance

▶ Excellent sequencing data quality

With low duplicate and high genome coverage rates, the MGI library sequencing data has a high effective utilization ratio.

Sample	Mapping rate	Duplicate rate	Capture rate	Average Depth	Coverage (≥1X)	Coverage (≥4X)	Coverage (≥10X)	Coverage (≥ 20X)
NA12878	99.71%	5.17%	58.14%	100.36%	99.61%	99.02%	97.36%	92.96%

Library was prepared from 1ug NA12878 gDNA using the Exome Capture V4 Universal Kit and the MGIEasy Exome Library Preparation Kit, and sequenced on MGISEQ-2000 using PE100.

▶ The accuracy and sensitivity of SNP and Indel detection

MGI sequencing platforms demonstrated high accuracy and sensitivity on SNP and Indel detection.

Variation	True-pos	False-pos	False-neg	Precision	Sensitivity
SNPs	40,153	371	748	99.08%	98.17%
Indel	2,802	485	289	85.24%	90.65%

Libraries was prepared from 1ug NA12878 gDNA using the Exome Capture V4 Universal Kit and the MGIEasy Exome Library Preparation Kit, and sequenced on MGISEQ-2000 using PE100.

mRNA/ Total RNA Sequencing

RNA sequencing is a powerful method for comprehensive and rapid analysis of gene expression changes, examination of rare and novel transcripts, and discovery of alternate splicing events, gene fusions, SNPs and allele-specific expression in tissues or cells. It has been widely applied to many research fields, including biological, disease, drug development, agriculture and environmental studies.

Recommended MGI Process

Library preparation ▶ **Sequencing**

MGIEasy RNA Library Prep Set
MGIEasy RNA Directional Library Prep Set
MGIEasy rRNA Depletion Kit

Automated sample preparation system MGISP-100

MGISEQ-2000
MGISEQ-200

mRNA/ Total RNA library preparation

MGI has developed the following library preparation products to support mRNA/ Total RNA sequencing.

▶▶ MGIEasy RNA Library Prep Set

The MGIEasy RNA Library Prep Set provides a solution for RNA quantification and transcriptome studies with all kinds of eukaryotic and prokaryotic species.

▶▶ MGIEasy RNA Directional Library Prep Set

The MGIEasy RNA Directional Library Prep Set enables strand specific studies to more accurately detect and quantify transcripts, reveal gene structure, and probe sense and antisense transcripts.

Product specification

Products	MGIEasy RNA Library Prep Set	MGIEasy RNA Directional Library Prep Set
Assay Time	<7 h	
Input Quantity	10 ng- 1 ug total RNA	
Objects	mRNA enriched by oligo(dT) beads, RNA enriched by rRNA depleted with rRNA depletion kit	
Sample types	tissues, plasma and FFPE sample	
Species Compatibility	Human, animals, plants, fungi and bacteria, such as mouse, rice, Arabidopsis, yeast and E.coli.	
Applications	RNA-Seq, Transcriptome Sequencing, total RNA sequencing, lncRNA Sequencing	
Platform Compatibility	MGISEQ-2000, MGISEQ-200	
Recommended Read Length	25 M raw reads（SE50）/8 Gb raw data（PE100/PE150）	
Detection of strand orientation	Not available	Available

Features

- As low as 10 ng total RNA per sample
- Compatible with human, animal, plant, fungi and bacteria samples with high integrity or degraded total RNA such as FFPE and plasma samples
- Easy to use with simple protocol and short operation time in 7 hours
- Provide two options of size selection of 150-300 bp or 200-400 bp
- Accurate and comprehensive mapping of transcripts
- High uniformity enables high uniform from transcripts 5 'to 3' ends and superior 3'end coverage
- Accurate identification transcript strand orientation to enhance transcript annotation and detect antisense transcripts using MGIEasy RNA Directional Library Prep Set

Performance

▶ High concordance of gene expression

The concordance of gene expression was measured using different input quantities of Universal Human Reference RNA (UHRR). Pearson and Spearman correlation tests demonstrated, stable and accurate results, yielding r values >0.995 for all data analyzed (Fig. 1).

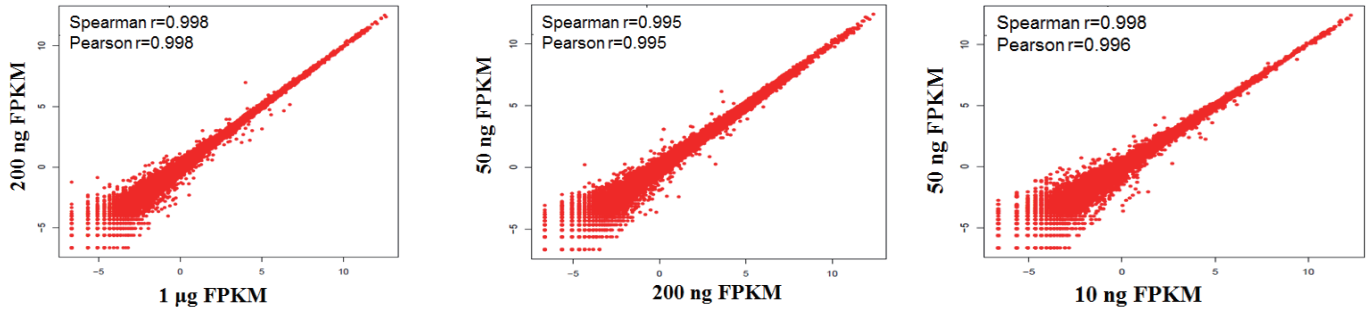


Fig. 1 Gene expression level reproducibility and concordance in different input amounts of total RNA

The libraries were prepared from an input of UHRR ranging from 10 ng -1 μg using the MGIEasy RNA Library Prep Set and sequenced on MGISEQ-2000 at PE100 read-length. After data filtering, approximately 8 Gb of data was collected per library for analysis.

▶ High uniformity

The results in Fig.2 demonstrate that the libraries constructed using MGIEasy RNA Library Prep Set have high uniformity throughout the entire length of a transcript with supreme 3'end coverage.

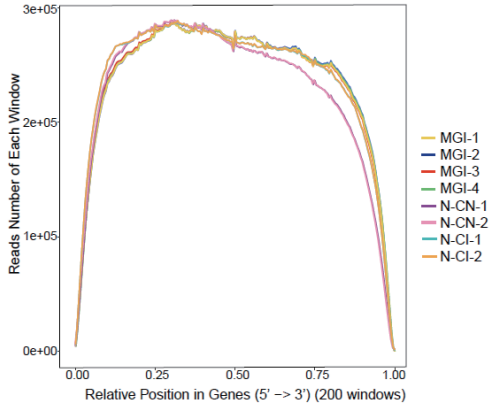


Fig.2 Reads randomness of different platforms and kits

The libraries MGI -1/2/3/4 were generated from UHRR using the MGIEasy RNA Library Prep Set and sequenced on MGISEQ-2000 at PE150 read-length. The libraries N-CN-1/2 and N-CI-1/2 were prepared using Company-N and Company-I kits respectively and sequenced on "N" platform at PE150 read-length. After data filtering, about 10 Gb were collected per library for analysis. The analysis of sequencing data was based on the same instruction.

▶ High-Quality Stranded Information

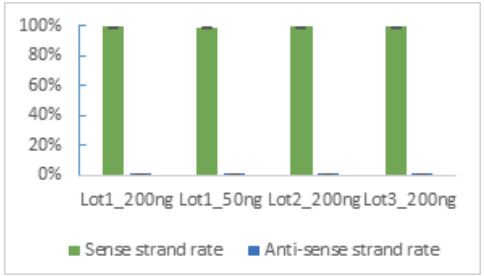


Fig.3a Sense strand rate of MGIEasy RNA Directional Library Prep Set

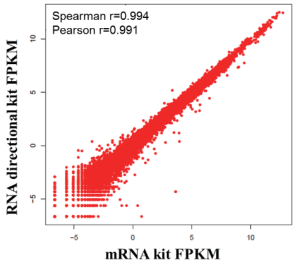


Fig.3b Concordance of MGIEasy RNA Directional Library Prep Set and MGIEasy RNA Library Prep Set RNA

Fig.3 Libraries were prepared with from an enriched poly(A) mRNA of UHRR using the MGIEasy RNA Directional Library Prep Set and sequenced on MGISEQ-2000 at PE100 read-length. After data filtering, approximately 8 Gb were collected per library for analysis.

MGIEasy rRNA Depletion Kit

The MGIEasy rRNA Depletion Kit can efficiently deplete rRNA from 10 ng-1 μg of total RNA from human, mouse and rat samples, including HMW material and even degraded samples (such as FFPE), to enrich messenger and non-coding RNA, increasing the proportion of useful data.

Product specification

Assay Time	~2 hours
Input Quantity	10 ng - 1 μg of total RNA
Sample types	tissues, plasma and FFPE sample
Species Compatibility	Human, mouse, rat
Applications	total RNA sequencing, lncRNA Sequencing
Removed rRNA types	cytoplasmic 5S rRNA, 5.8S rRNA, 18S rRNA, 28S rRNA, mitochodrial ribosomal RNA 12S rRNA and 45S pre-ribosomal RNA, 16S rRNA

Performance

▶ High efficiency on 99% rRNA depletion

Compared with other depletion kits, MGIEasy rRNA Depletion Kit has greater efficiency.
Less than 1% of the initial quantity of rRNA remains after a single depletion.

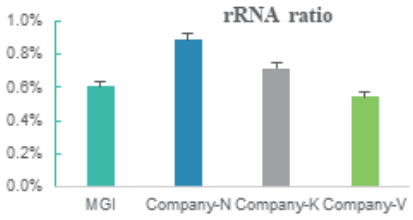


Fig.4a Depletion efficiency of different depletion kits

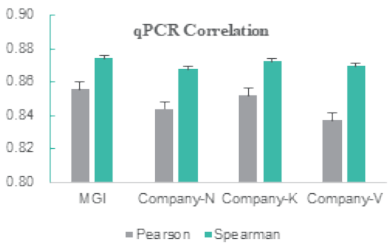


Fig.4b qPCR correlation of different depletion kits

Fig.4 Ribosomal RNA in 200 ng UHRR was depleted with different rRNA depletion kits. The resulting rRNA-depleted RNA were prepared to construct libraries using the MGIEasy RNA Directional Library Prep Set and sequenced on MGISEQ-2000 at PE100 read-length. After data filtering, approximately 8 Gb of data was used per library for analysis.

Small RNA Sequencing

Small RNAs are an endogenous non-coding single-stranded RNA molecules formed of 20-24 nucleotides and include miRNAs, siRNA and piRNA, snRNA and snoRNA. Small RNAs are involved in gene silencing and post-transcriptional regulation of gene expression through RNA-RNA interactions. They are crucial regulators in cell growth and differentiation in development processes. Small RNA Sequencing is a technique used to isolate and sequence small RNA for biological, disease, drug development, plant and animal research.

The MGISEQ-based Small RNA sequencing solution uses supported kits to prepare single-stranded circularized cDNA libraries from total RNA, sequence those libraries, and generates base calling sequencing data of hundreds of small RNA sequences.

Recommended MGI Process

Library preparation	Sequencing
MGIEasy Small RNA Library Prep Kit	MGISEQ-2000 MGISEQ-200

▶▶ MGIEasy Small RNA Library Prep Kit

The MGIEasy Small RNA Library Prep Kit provides an efficient solution for generating libraries suitable for MGI high-throughput sequencing platforms from 10 ng - 1 μg of total RNA

Product specification

Assay Time	~13 hours
Input Quantity	10 ng - 1 μg of total RNA
Sample types	Tissues and plasma
Species Compatibility	Human, animals and plants, such as mouse, rice and Arabidopsis
Applications	Small RNA Sequencing
Platform Compatibility	MGISEQ-2000, MGISEQ-200
Recommended Read Length	SE50
Recommended sequencing data per sample	25 M raw reads

Features

- As low as 10 ng total RNA per sample
- Compatible with various species (human, animal and plant) samples
- Provides two options of size selection based on either bead or gel purification to meet a wide range of library needs

Methylation Sequencing

DNA methylation is an important epigenetic phenomenon. It is essential to know every methylated cytosine in the genome to understand temporal and spatial gene expression and chromatin remodeling. The MGISEQ-based Whole Genome Bisulfite Sequencing (WGBS) solution leverages bisulfite conversion during library preparation and high throughput sequencing technology. The solution enhances epigenetic research with high accuracy, producing a genome-wide DNA methylation map.

Recommended MGI Process

Library preparation	Sequencing
MGIEasy Whole Genome Bisulfite Sequencing Library Prep Kit	MGISEQ-2000

▶▶ MGIEasy Whole Genome Bisulfite Sequencing Library Prep Kit

The MGIEasy Whole Genome Bisulfite Sequencing Library Prep Kit is specifically designed for preparing WGBS libraries for MGI high-throughput sequencing platform from 10-100 ng DNA fragments. This kit is compatible with a variety of sample types, including challenging FFPE samples, and can be widely applied to many applications, such as cell development and differentiation research, human disease research, animal and plant breeding, species and population evolution studies.

Product specification

Assay Time	~8 hours
Input Quantity	10- 100 ng fragmented DNA (200-500 bp)
Sample types	tissues, plasma and FFPE sample
Species Compatibility	Human, animals, plants and fungi with reference genome sequence, such as Arabidopsis and yeast
Fragmentation method	Physical (Fragmentation is not needed for plasma free DNA)
Applications	Methylation sequencing
Platform Compatibility	MGISEQ-2000
Recommended Read Length	PE100, PE150
Recommended sequencing data per sample	30X clean base (e.g. 90G clean bases for human)

Features

- As low as 10 ng fragmented DNA per sample
- Compatible with human, animal, plant and fungi samples with high integrity or degraded total RNA such as FFPE and plasma samples
- Low duplicate rate and high genome mapping rate enabling high usability of sequencing data
- Excellent CpG coverage and accuracy of methylation detection

Performance

▶ High usability of sequencing data

MGI_NA library was prepared from 100 ng NA12878 gDNA using the MGIEasy Whole Genome Bisulfite Sequencing Library Prep Kit and sequenced at MGISEQ-2000. X_NA library was constructed using the corresponding products for X NGS platform and sequenced on X platform. The result shows that MGI_NA has higher mapping rate and lower duplication rate, resulting in a greater sequencing depth at the same clean data amount, enabling high usability of sequencing data with the MGIEasy Whole Genome Bisulfite Sequencing Library Prep Kit (Table 1).

Table 1 Mapping and coverage rate

Sample ID	X_NA	MGI_NA	Sample ID	X_NA	MGI_NA
input DNA amounts (ng)	1000	100	Duplication rate (%)	28.49	16.29
Clean data amount (Gb)	111	111	Mismatch Rate (%)	0.71	0.72
Read length	PE150	PE100	BS Conversion Rate (%)	99.45	99.55
Clean rate (%)	99.47	96.26	Average sequencing depth	23X	28X
Mapping Rate (%)	87.04	93.52	Coverage at least 4X (%)	97.24	97.08
Unique mapping rate (%)	84.15	89.29	Coverage at least 10X (%)	93.29	94.25
			Coverage at least 20X (%)	61.73	78.15

▶ High CpG coverage

Data produced from NA12878 DNA using Infinium® MethylationEPIC BeadChip was used to compare the CpG coverage between MGI_NA and X_NA. The result shows that MGI_NA data has higher concordance of CpG coverage with microarray compared with X_NA and detects more CpG sites with sequencing depth at 10X and 20X.

Table 2 Analysis data on Concordance of CpG coverage

Table 2 Analysis data on Concordance of CpG coverage

Sample	Sequencing depth	Genome CG number	Concordance CG rate* (%)	Unique CG rate**(%)
MGI_NA	4X	56,434,896	83	8
X_NA	4X			9
MGI_NA	10X		48	35
X_NA	10X			18
MGI_NA	20X		15	70
X_NA	20X			15

*The concordance CG rate of the two platforms in all CG sites detected

**The percentage of CG sites detected by only one of the NGS platforms in all CG sites detected

▶ Excellent accuracy of methylation detection

The genome methylation rate of CpG pattern of MGI_NA is close to that of X_NA. The methylation rate in the CH pattern (CHG and CHH) of NA12878 (b-lymphocyte) should be less than 1%. The chromosomal methylation rates of MGI_NA datasets are at the expected level, while those of X_NA dataset are considerably higher (mostly around 3%), indicating high accuracy regarding methylation levels (Fig.5).

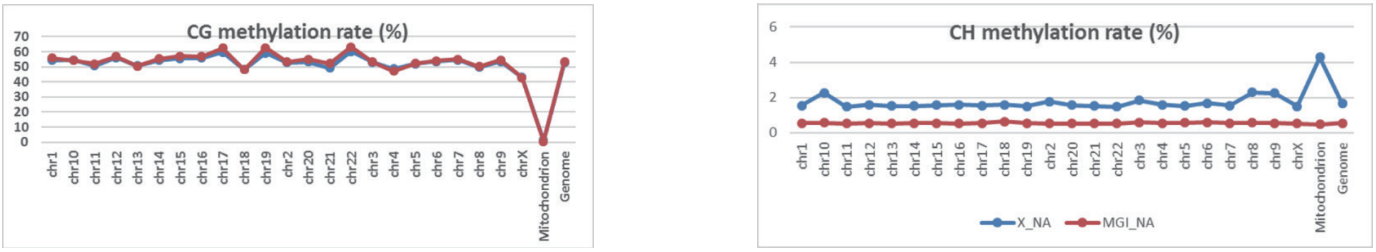


Fig.5 Methylation rate of CG pattern and CH pattern

Metagenomics Sequencing

Metagenomics sequencing is a method for directly sequencing separate microbial populations in pooled samples without isolating and culturing each individual population. It can be used to study the community structure, species classification, phylogenetic evolution, gene function and metabolic network of environmental microorganisms, etc. The method is an effective tool for environmental microbe studies and has been widely used in applications, such as the identification of pathogenic infections, where traditional methods are not useful. MGI high throughput sequencing delivers an accurate and rapid pathogen identification solution for diagnosing infectious disease using microbial genetic information.

Recommended MGI Process

Library preparation	Sequencing	Bioinformatics Analysis
MGIEasy FS DNA Library Prep Set MGIEasy Universal DNA Library Prep Set MGIEasy Cell-free DNA Library Prep Set	MGISEQ-2000 MGISEQ-200	Fast identification of pathogenic infections
Automated sample preparation system MGISP-100		

Metagenomics Sequencing library preparation

MGI has developed a series library preparation kits for Metagenomics Sequencing.

Product specification

Products	MGIEasy FS DNA Library Prep Set	MGIEasy Universal DNA Library Prep Set	MGIEasy Cell-free DNA Library Prep Set
Assay Time	<4 h	<3 h	<3 h
Input Quantity	5- 300 ng gDNA	0.5-50 ng fragmented DNA (200-500 bp)	2-6 ng cfDNA or DNA fragment (150-250 bp)
Sample types	Meta DNA, cfDNA	Meta DNA, cfDNA, Amplicons	cfDNA from plasma, nasopharyngeal swab, sputum, alveolar lavage, cerebrospinal fluid, stool, etc.
Species Compatibility	Human, animals, plants, fungi and bacteria, such as mouse, rice, Arabidopsis, yeast and E.coli.		Meta sample from Human, animals and environment sample
Fragmentation method	Enzyme	Physical	No need
Automation option	Automation-friendly: MGISP-100	Not available	Automation-friendly: MGISP-100
Applications	Human genome sequencing, Plant and Animal genome sequencing, Microbial genome sequencing, Metagenomics sequencing		Pathogen identification
Platform Compatibility	MGISEQ-2000, MGISEQ-200		MGISEQ-2000, MGISEQ-200
Recommended Read Length	SE50, PE100, PE150 and etc.		SE50, PE100

RAD Sequencing

▶▶ MGIEasy RAD Library Prep Kit

The MGIEasy RAD Library Prep Kit supports library preparation for reduced-representation sequencing on the MGI high throughput sequencing platform. The kit provides reagents to prepare a library from 1 μg animal or plant gDNA. This kit offers restriction enzyme digestion, adaptor ligation, samples pooling, size selection using magnetic beads, PCR, purification and single-strand circularization.

Product specification

Assay Time	~7 hours
Input Quantity	1 μg gDNA
Sample types	Tissues and plasma sample
Species Compatibility	Animals and plants, such as pig and rich
Fragmentation method	Enzyme
Applications	Agriculture research
Platform Compatibility	MGISEQ-2000
Recommended Read Length	PE100, PE150

Features

- Compatible with animal and plant samples. Not limited to reference genome. Effectively reduces the genomic complexity of different species.
- Simple and fast operation without ER/A tailing.
- High throughput library preparation solution enables the pooling of 16-64 samples.

Library preparation modules

In addition to high quality Library Preparation Kit, MGI offers flexible library preparation modules and reagents. Library preparation modules are used in a single step or multiple steps. They can be used with other kits to meet the requirements of the workflow. A variety of reagents are available to meet the full range of project needs.

Library preparation module

Type	Products	Configuration	Version	Catalog No.
Adapter	MGIEasy DNA Adapters-96（96-well plate）Kit	96 × 10 μL	V1.0	1000005282
Bead Purification	MGIEasy DNA Purification Magnetic Bead Kit	50 mL	V1.0	1000005279
Circularization	MGIEasy Circularization Kit	16 RXN	V2.0	1000005259

Ordering information

Sequencing platform

Classifications	Product	Catalog No.
Genetic Sequencer	Genetic Sequencer MGISEQ-2000RS	900-000035-00
	Genetic Sequencer MGISEQ-200RS	900-000042-00
DNB Loader	DNB Loader MGIDL-200	900-000037-00
Automated sample preparation system	DNA Sequencing Library Preparation System MGISP-100RS	900-000027-00

Sequencing reagents

Sequencing platform	Product	Configuration	Catalog No.
MGISEQ-2000RS	MGISEQ-2000RS High-throughput Sequencing Set	SE50	1000003703
	MGISEQ-2000RS High-throughput Sequencing Set	PE100	1000002873
	MGISEQ-2000RS High-throughput Sequencing Set	PE150	1000003981
	MGISEQ-2000RS High-throughput Sequencing Set (Small RNA)	SE50	1000006138
MGISEQ-200RS	MGISEQ-200RS High-throughput Sequencing Set	SE50	1000004635
	MGISEQ-200RS High-throughput Sequencing Set	PE50	1000005228
	MGISEQ-200RS High-throughput Sequencing Set	PE100	1000005229

Libraries preparation kits

Applications	Product	Configuration	Version	Catalog No.
WGS	MGIEasy Universal DNA Library Prep Set	16 RXN	V1.0	1000005246
		96 RXN	V1.0	1000005247
	MGIEasy FS DNA Library Prep Set	16 RXN	V1.0	1000005252
		96 RXN	V1.0	1000005253
WES	Exome Capture V4 Universal Kit	16 RXN	V1.0	1000001608
	MGIEasy Exome Library Prep Kit	16 RXN	V1.1	1000003693
	MGIEasy Exome Rapid Library Prep Kit	16 RXN	V1.0	1000005245
mRNA/Total RNA Sequencing	MGIEasy RNA Library Prep Set	16 RXN	V3.0	1000006383
		96 RXN	V3.0	1000006384
	MGIEasy RNA Directional Library Prep Set	16 RXN	V2.0	1000006385
		96 RXN	V2.0	1000006386
	MGIEasy rRNA Depletion Kit	32 RXN	V1.1	1000005953
Small RNA Sequencing	MGIEasy Small RNA Library Prep Kit	24 RXN	V2.0	1000005269
WGBS	MGIEasy Whole Genome Bisulfite Sequencing Library Prep Kit	16 RXN	V1.0	1000005251
cfDNA	MGIEasy Cell-free DNA Library Prep Set	48 RXN	V1.0	1000007037
RAD-seq	MGIEasy RAD Library Prep Kit	64 RXN	V1.0	1000005242

For Research Use Only. Not for use in diagnostic procedures.

Innovate Life Technology
MGI Tech Co., Ltd

High Throughput Sequencing Platform

Innovate Life Technology

MGI Tech Co., Ltd

Contact Us

Address: Building 11, Beishan Industrial Zone,
Yantian District, Shenzhen, China, 518083

Website: en.mgitech.cn
Email: MGI-service@genomics.cn
Telephone: (+86)4000-966-988

Copyright Disclaimer

The copyright of this brochure is solely owed by MGI Tech Co.Ltd., (MGI). The information included is strictly forbidden to be reproduced or transmitted in any form, by any means (e.g electronic, photocopying, recording, translating, or otherwise) without the prior written permission by MGI. All trademarks and icons in the brochure are intellectual property of MGI and its respective producers.

Version: October 2018

Official Wechat

Official Website