# Application of MGIEasy FFPE DNA Extraction Prepacked Kit and MGISP-NE32 Automatic Nucleic Acid Extractor for the Whole Exome Sequencing

## Highlight

- **High extracting efficiency**

  Providing sufficient DNA for the library preparation

- **Extracting high quality DNA**

  Excellent formula, providing relatively high quality DNA for the library preparation and sequencing, help to discover the cancer therapy target.

- **Widely tissue types**

  Compatible with multiple tissue sample, such as thyroid, cervical, colon, breast, help to multiple cancer research.

## Introduction

FFPE samples are one of the most common ways to preserve clinical samples, and FFPE tumor samples serve as valuable study materials for clinical and translational medicine research. Formalin-fix and paraffin-embed can make the tissue sample to be stored for long time, but will damage to the nucleic acid in the sample, for example the formalin make the nucleic acid cross linked with protein, decrease the extracting efficiency, and the paraffin broken the phosphodiester bond leads to nucleic acid degradation result in highly fragmented nucleic acid. So the DNA quality of the FFPE sample always bad, which is a challenge for sequencing, may lead to such as: low yield of library, low SNP calling rate, high duplication rate. To ensure successful and reliable sequencing of FFPE-derived DNA, it is necessary to provide sufficient and relatively good quality DNA for the library preparation and sequencing. In this application note, the MGIEasy FFPE DNA Extraction Prepacked Kit in combination with MGISP-NE32 Automatic Nucleic Acid Extractor were used to extract DNA from five FFPE samples, followed by DNA hybridization, library preparation and sequencing. high-quality sequencing data were achieved.

## Experiment

### Materials

FFPE tumor sections (0.5 cm*0.5 cm, 10 μm thickness) include cervical cancer S1(4 sections), colon cancer S2 (4 sections), thyroid cancer S3 (4 sections), breast cancer S4 (4 sections), and reference sample S5 (1 section), the FFPE reference sample were obtained from GeneWell (PN: GWC0301).

## DNA Extraction

DNA was extracted using MGIEasy FFPE DNA Extraction Prepacked Kit (PN: 940-000113-00) and MGISP-NE32RS automatic nucleic acid extractor (PN:950-000020-00) in accordance with the standard protocol. The FFPE DNA extraction workflow include deparaffinization, washing, lysis, denature and purification. Among them, the deparaffinization step is most important, too long time for deparaffinization may lead to DNA degradation, too short time may lead to low DNA yield.


Figure 1. Workflow of FFPE gDNA extraction

## Library Prep

Using 400ng extracted DNA for fragmentation on Covaris ME220 system, followed by the pre-PCR library preparation using MGIEasy Universal DNA Library Prep Set (PN:1000006985), mix the S1-S4's pre-PCR library in proportion, then transfer 1000 ng the S1-S4 mixed library and the S5 pre-PCR library for hybridization, followed by the post-PCR library preparation via elution and PCR amplification.
The Agilent SureSelect Human All Exon V8 kit and SureSelect XT HS2 kit were used for hybridization. The MGIEasy Exome Capture Accessory Kit (PN:1000007743) was used for blocking. The pre-PCR library was prepared by MGISP-100RS DNA Sequencing Library Preparation System (PN: 900-000206-00).

## Sequencing

The library preparation was followed by making DNB, then sequenced by DNBSEQ-G400RS* Genetic Sequencer (PN:900-000170-00) using DNBSEQ-G400RS High-throughput Rapid Sequencing Set (FCS PE100)* (PN:1000016980).

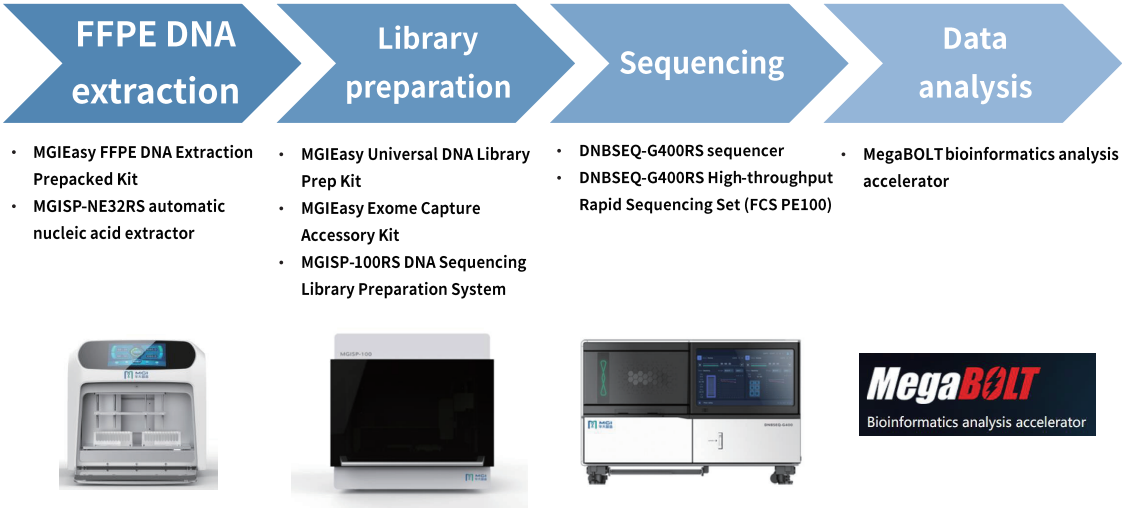


Figure 2. Instrument and reagents involved in the process

# Results

## Quality of genomic DNA derived from FFPE samples

The extracted FFPE DNA concentration was measured by Qubit, the DIN and main peak size were analyzed by Agilent TapeStation. Serious degradation was observed in some of the samples due to long-term storage, which is common for FFPE samples. the DNA yield totally meet the requirement of the library preparation.

Table 1. Measurement of extracted DNA

| Sample ID | Type of Tissue | DIN | Yield (ng) | Main peak (bp) | OD260\230 | OD260\280 |
|---|---|---|---|---|---|---|
| S1 | Cervical | 2.9 | 1001 | 1,919 | 1.71 | 1.82 |
| S2 | Colorectal | 4.5 | 2968 | 2241 | 1.94 | 1.58 |
| S3 | Thyroid | 2.6 | 2332 | 731 | 1.86 | 1.78 |
| S4 | Breast | 2.2 | 612 | 574 | 1.69 | 1.74 |
| S5 | Standards | 7.9 | 1944 | 24,312 | 1.82 | 1.9 |

*DIN lower than 6 means degradation, the lower the DIN, the more serious the degradation

## Quality of the sequencing data

Each sample data was normalized to 7.5 Gb for analysis. the Q30 higher than 85% (the qualified criterion) no matter before or after filtration. The mapping rate, coverage(1X) and SNP calling rate are higher than 99% which means excellent sequencing result.

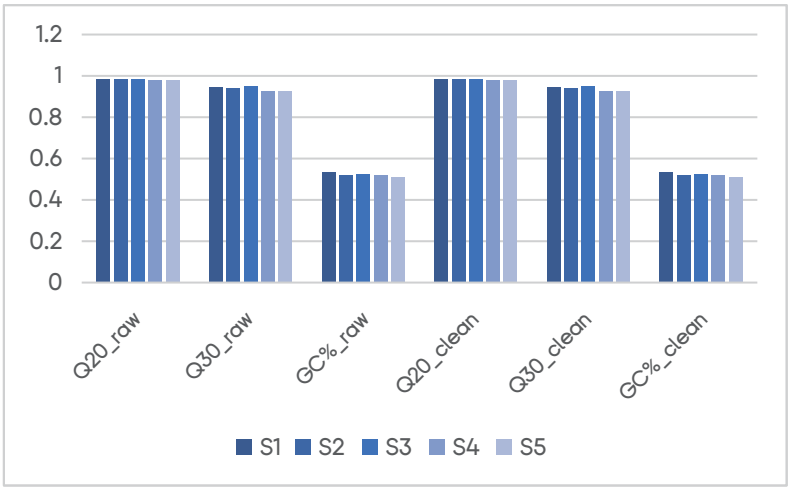


Figure 3. Data quality of the 5 FFPE samples

Table 2. Sequencing data quality

| Sample ID | Mapping_Rate | Cov(1X) | Cov(5X) | Cov(10X) | Cov(30X) | Total SNP | dbSNP calling rate |
|---|---|---|---|---|---|---|---|
| S1 | 99.71% | 99.07% | 98.79% | 97.05% | 93.15% | 22897 | 99.30% |
| S2 | 99.72% | 99.24% | 99.07% | 98.69% | 98.26% | 23633 | 99.27% |
| S3 | 99.67% | 99.06% | 98.47% | 98.14% | 97.24% | 23515 | 99.02% |
| S4 | 99.72% | 99.04% | 98.85% | 98.25% | 97.45% | 23524 | 99.02% |
| S5 | 99.74% | 99.09% | 98.92% | 98.68% | 98.49% | 23633 | 99.42% |

## Bioinformatics

Statistic：Analyze the SNV (frequency higher than 1%) and InDel (frequency higher than 5%) respectively and map the circos chart, as below shown, the Figure 4 is the SNV density chart, the Figure 5 is the InDel density chart, the 6 lines are for different sample, the S5 has 2 replicates, we can see the all the 5 samples SNV and InDel density are consistent.

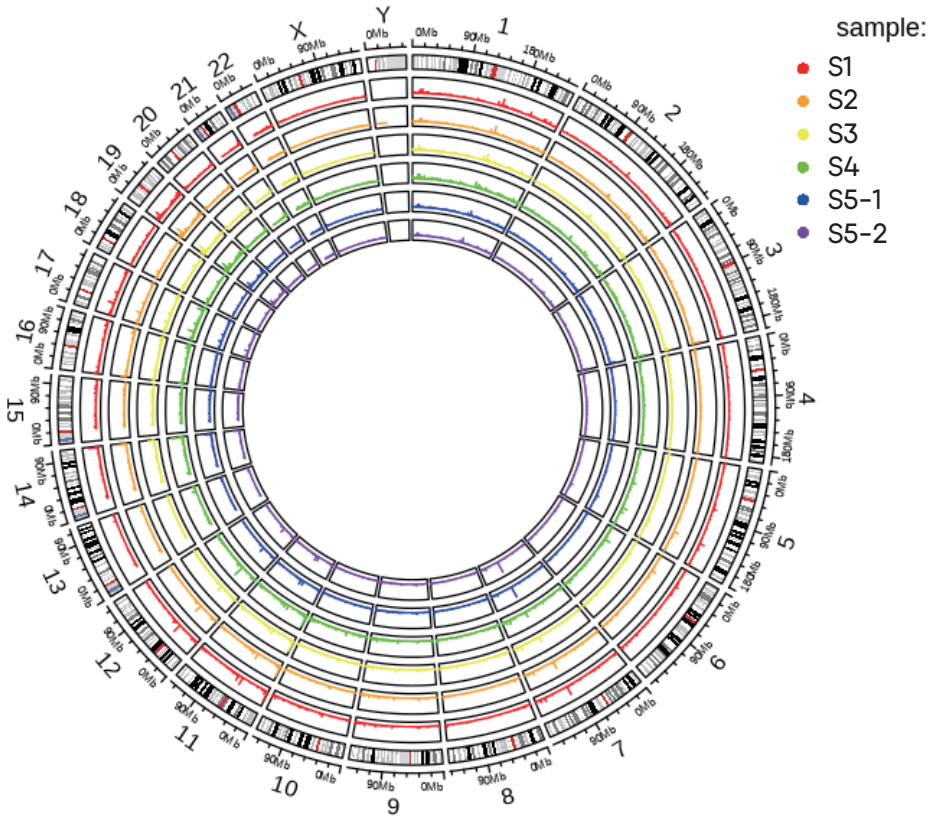


Figure 4. SNV density circos chart

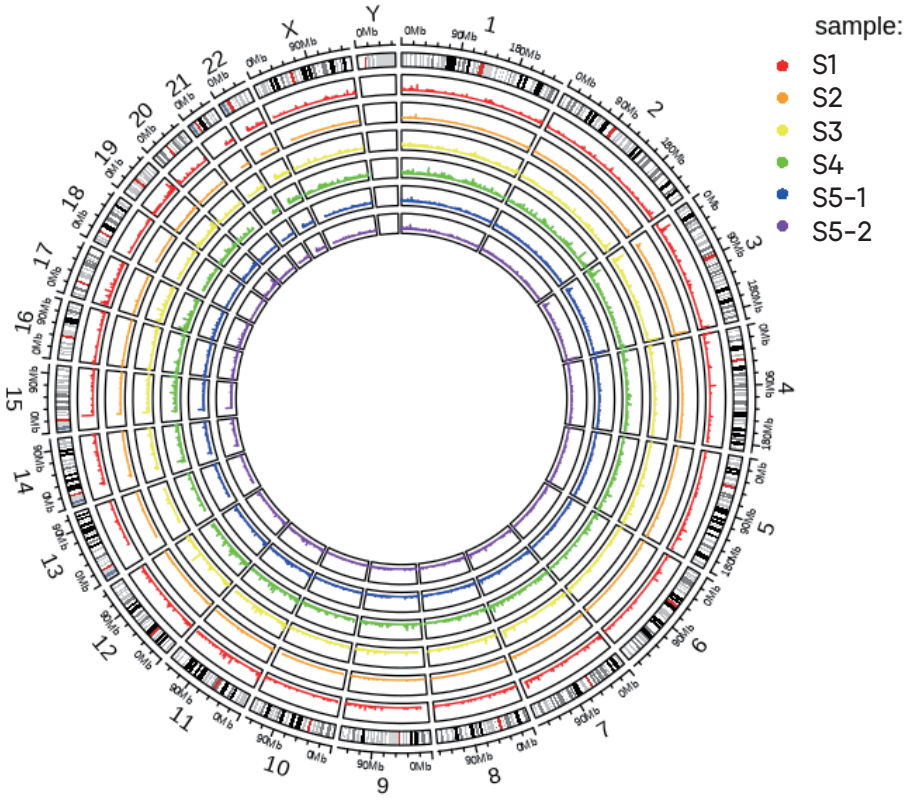


Figure 5. InDel density circos chart

COSMIC is a comprehensive data base which recorded the human cancer related driving gene and published by Sanger institute. It is an authoritative data base for cancer research. So we searched the recorded human cancer related driving genes using the S1-S5 samples' data, the result listed in below table, we can see the all the 4 cancer FFPE samples (S1-S4) can searched out the human cancer related driving genes.

Table 3. COSMIC searching result

| Sample ID | Human cancer related driving genes |
|---|---|
| S1 | 59 |
| S2 | 7211 |
| S3 | 5013 |
| S4 | 2912 |
| S5 | 20381 |

## Conclusion

The DNA extracted by MGIEasy FFPE Genomic DNA Extraction Prepacked Kit in combination with MGISP-NE32 can meet the requirement of downstream DNA hybridization, library preparation and sequencing. the sequencing data quality is excellent, such as the mapping rate, genomics coverage, SNP calling rate, totally meet the requirement of downstream deep bioinformatics analysis, which means MGIEasy FFPE Genomic DNA Extraction Prepacked Kit is a powerful sample treatment tool for the cancer research.

## Ordering Information

| Product name | Specification | Part number |
|---|---|---|
| MGIEasy FFPE DNA Extraction Prepacked Kit | 32 Preps | 940-000113-00 |
| MGIEasy Universal DNA Library Prep Set | 16 Preps | 1000006985 |
| MGIEasy Exome Capture Accessory Kit | 16 RXN | 1000007743 |
| MGISP-NE32RS Automatic Nucleic Acid Extractor | - | 950-000020-00 |
| DNBSEQ-G400RS* Genetic Sequencer | - | 900-000170-00 |
| DNBSEQ-G400RS High-throughput Rapid Sequencing Set (FCS PE100)* | - | 1000016980 |
| MGISP-100RS DNA Sequencing Library Preparation System | - | 900-000206-00 |

## Reference

1. Tang, W., David, F. B., Wilson, M. M., Barwick, B. G., Leyland-Jones, B. R., and Bouzyk, M. M., "DNA Extraction from Formalin-Fixed, Paraffin-Embedded Tissue", Cold Spring Harbor Protocols, doi:10.1101/pdb.prot 5138, 2009.

**MGI Tech Co.,Ltd** | Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083
en.mgi-tech.com | MGI-service@mgi-tech.com