# Nanodigmbio Target Enrichment Solutions Compatible with MGI Automation System and DNBSEQ Platform to Enable Tumor-related Variation Detection and WES Research

## Data Evaluation of Nanodigmbio LungCancer Panel and NEXome Core Panel based on MGISP-Smart 8 and DNBSEQ-G400

In this study, two panels (lung cancer and WES) of Nanodigmbio were chosen to test the compatibility of their target enrichment solutions with MGI's automation system and DNBSEQ sequencing platform. The results showed that the Nanodigmbio target enrichment solutions achieved high-quality library preparation equivalent to manual operations, and high-quality sequencing with a high accuracy in mutation identification. This comprehensive research showed that the automation system and DNBSEQ sequencing platform developed by MGI is perfectly compatible with the Nanodigmbio target enrichment solutions.

Application recommended: Oncology (lung cancer) and whole-exon sequencing (WES)
Model recommended: MGISP-Smart 8RS, MGISP-100RS, MGISP-960RS (Automated system)
DNBSEQ-G99ARS, DNBSEQ-G400RS, DNBSEQ-T7RS (Sequencing platform)

- **Nanodigmbio target enrichment solutions exhibit stable and outstanding performance**

  With an efficient target capture performance, the Nanodigmbio LungCancer Panel and NEXome Core Panel can be applied to lung cancer-related gene mutation detection and WES study, respectively.

- **Automation system is highly compatible with sample pretreatment and library preparation**

  The MGI self-developed MGISP-Smart 8 automation system can flexibly realize library preparation of 1-48 samples/run. It completes the whole process from library preparation, hybrid capture, homogenization, pooling to DNB preparation accurately, stably and efficiently, ensuring the homogeneity and accuracy of library samples.

- **Perfect compatibility with DNBSEQ sequencing platform**

  Nanodigmbio target enrichment solutions can be perfectly adapted to DNBSEQ genetic sequencers of MGI.

- **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.

# Background

Massively parallel sequencing (MPS) can sequence millions or billions of DNA molecules simultaneously and can effectively reduce the sequencing cost, improve the sequencing accuracy and detection speed. Currently, it has been widely applied in disease preliminary diagnosis, cancer monitoring and drug resistance mechanism, etc[1,2]. However, complicated manual operation and high-skilled personnel requirements are the main challenges during MPS library preparation[3], which could be effectively solved by automatic equipment[4].

MGI has self-developed the MGISP automated sample preparation system to simplify customers' work. This system covers three series of products: sample pretreatment product,sample preparation system, integrated testing system.

Among them, MGISP-Smart 8 is a professional laboratory automatic sample preparation system. This system, equipped with 8 independent and controllable pipetting channels, is able to not only flexibly realize the library preparation of 1-48 samples per round, but also cover the automatic operation of the whole process from sample subpackaging, nucleic acid extraction, library building, pooling to DNB preparation. Moreover, this system can also be integrated with microplate reader to realize automatic quantitative pipetting, which greatly reduces manual operation time.

Nanodigmbio Biotechnology is a service provider dedicated to providing professional and high-quality target enrichment products and closed-loop solutions for global customers. Based on its proprietary technology, this company has developed a variety of target enrichment solutions compatible with MGI DNBSEQ sequencing platform. DNBSEQ sequencing platform is more and more favored by global customers with its advantages of high accuracy and sensitivity, low duplication rate and low index hopping rate, etc.

In this study, two target enrichment solutions of Nanodigmbio, LungCancer Panel and NEXome Core Panel, were tested to evaluate the compatibility of Nanodigmbio target enrichment solutions with MGISP system and DNBSEQ sequencing platform of MGI.

LungCancer Panel is specifically designed for lung cancer analysis, with selected lung cancer-related genes as its targeted region, covering genes related to the treatment of non-small cell lung cancer in NCCN Guidelines. This panel covers ~218 kb region of genome, and can detect many mutation types such as base substitution, insertion/deletion, gene rearrangement and amplification. NEXome Core Panel contains ~400,000 independently synthesized and quality controlled single-stranded DNA probes, targeting 34.7 Mb genome region (19,613 genes). As a core full-exon Panel, this product can be combined with different sub-panels to meet different application requirements.

The Nanodigmbio target enrichment solutions combined with MGISP automation system can fulfill high-quality library preparation. Moreover, the subsequent sequencing of the libraries on DNBSEQ platform has the characteristics of high target rate, high enrichment specificity of target genes, high coverage uniformity, low GC preference and accurate identification of related mutations. This study shows that the Nanodigmbio target enrichment solutions can be perfectly adapted to the MGISP automation system and DNBSEQ sequencing platform developed by MGI.

# Materials and methods

## Sample preparation

In this study, the pan-tumor gDNA standard (PN: GW-OGTM800) was purchased from GeneWell and used for the following experiments.

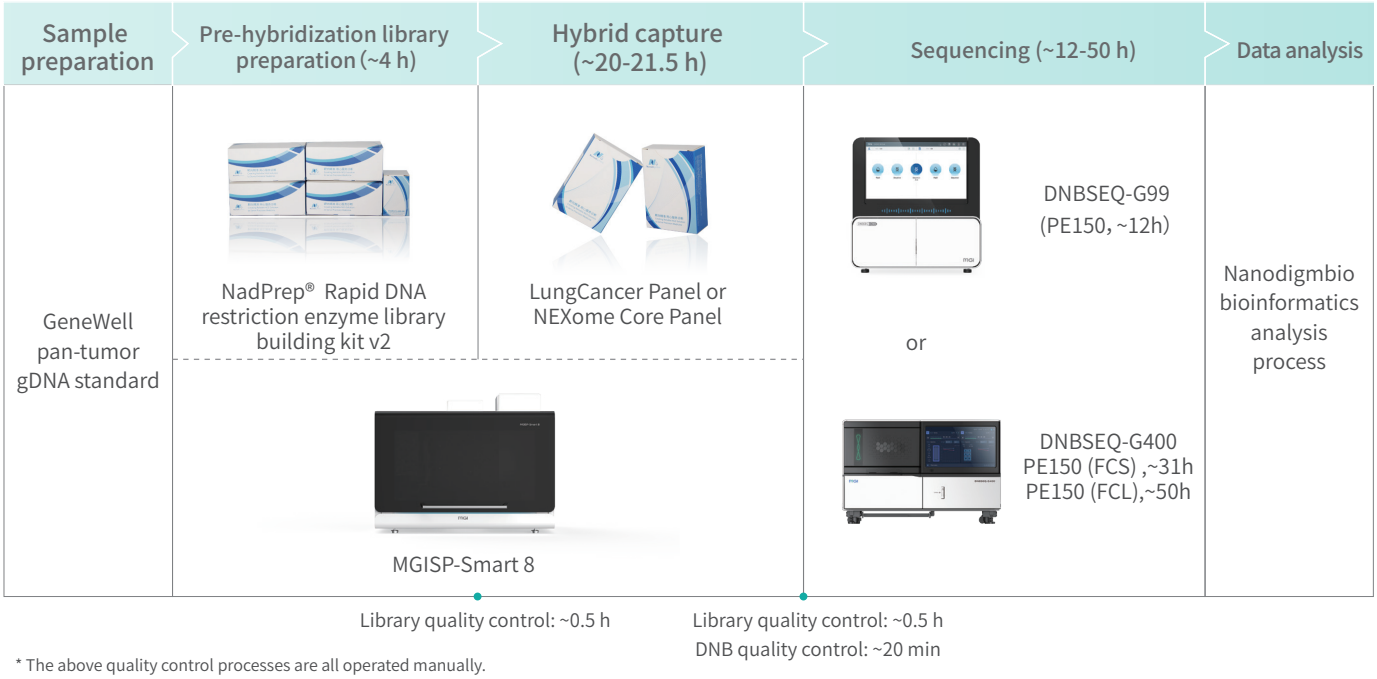
## Library preparation and sequencing

In the stage of library preparation, 50 ng of each sample was used to complete the library preparation process before hybridization by using NadPrep® EZ DNA Library Preparation Module v2 and NadPrep® Universal Adapter (MDI) Module Set B1. The digestion time is 23 min and Pre-PCR cycle is 7. Subsequently, LungCancer Panel or NEXome Core Panel was used for hybridization. For samples used for 1-plex, each sample input was 500 ng, with a total input of 500 ng, Post-PCR cycle number of 11, and hybridization time of 16 h. For samples used for 4-plex, each sample input was 500 ng, with a total input of 2 μg, Post-PCR cycle number of 9, and hybridization time of 16 h. All the above processes were carried out on MGISP-Smart 8. In this study, 1-plex hybridization strategy was adopted for LungCancer Panel and 1-plex or 4-plex hybridization strategy was adopted for NEXome Core Panel. In addition, manual library preparation was also carried out in parallel in the test. Both manual and automatic library preparation procedures could be referred to in relevant instructions.

The obtained double-stranded DNA libraries were prepared into DNB by RCA (Rolling circle amplification) reaction, followed by pair end 150 bp (PE150) sequencing on DNBSEQ-G400 genetic sequencer.

## Bioinformatic analysis

For LungCancer Panel and NEXome Core Panel, 0.2 G and 10 G offline sequencing data was intercepted, respectively. The bioinformatics analysis process was as follows: 1. Use fastp software to evaluate the quality of sequencing reads, eradicate the adaptors and filter the low-quality data. 2. Align the filtered reads to the reference genome by BWA to obtain the comparison bam document; 3. Conduct statistics for bam files, with the statistical content including comparison rate, probe GC preference analysis, target rate analysis, and other basic quality control indexes; 4. Use Vardict to analyze the variation of the captured data, and conduct statistics for the consistency between the measured variation frequency and the theoretical value.




* The above quality control processes are all operated manually.
* This solution can process 1~24 samples/run.
* The newly upgraded MGISP-Smart 8 can be integrated with microplate reader to complete the quantitative pipetting steps of PCR library, DNB samples and extraction products automatically. The whole process of library building is automatic.

# Results

## High-quality libraries can be prepared on MGISP automation system

In this study, the library preparation was completed on the MGISP-Smart 8 automation system by using Nanodigmbio target enrichment solutions (LungCancer Panel or NEXome Core Panel). Under the same input of 50 ng, the average yield of Pre-PCR products was >1800 ng, meeting the experimental requirements of >1000 ng (Figure 1A, B). The yield of post-PCR products also met the experimental requirements (Fig. 1A, B). Additionally, the library yield at each quality control point in the automatic group was slightly higher than the manual one. The fragment size of the library was as expected (~400 bp), and no non-specific peaks such as primer dimer were observed (Figure 1C, D).

The above results show that the MGISP-Smart8 automation system developed by MGI can be used by LungCancer Panel and NEXome Core Panel to complete high-quality library preparation. Moreover, NEXome Core Panel can obtain good results regardless of the hybrid strategy of 1-plex or 4-plex (1-12plex are officially recommended).

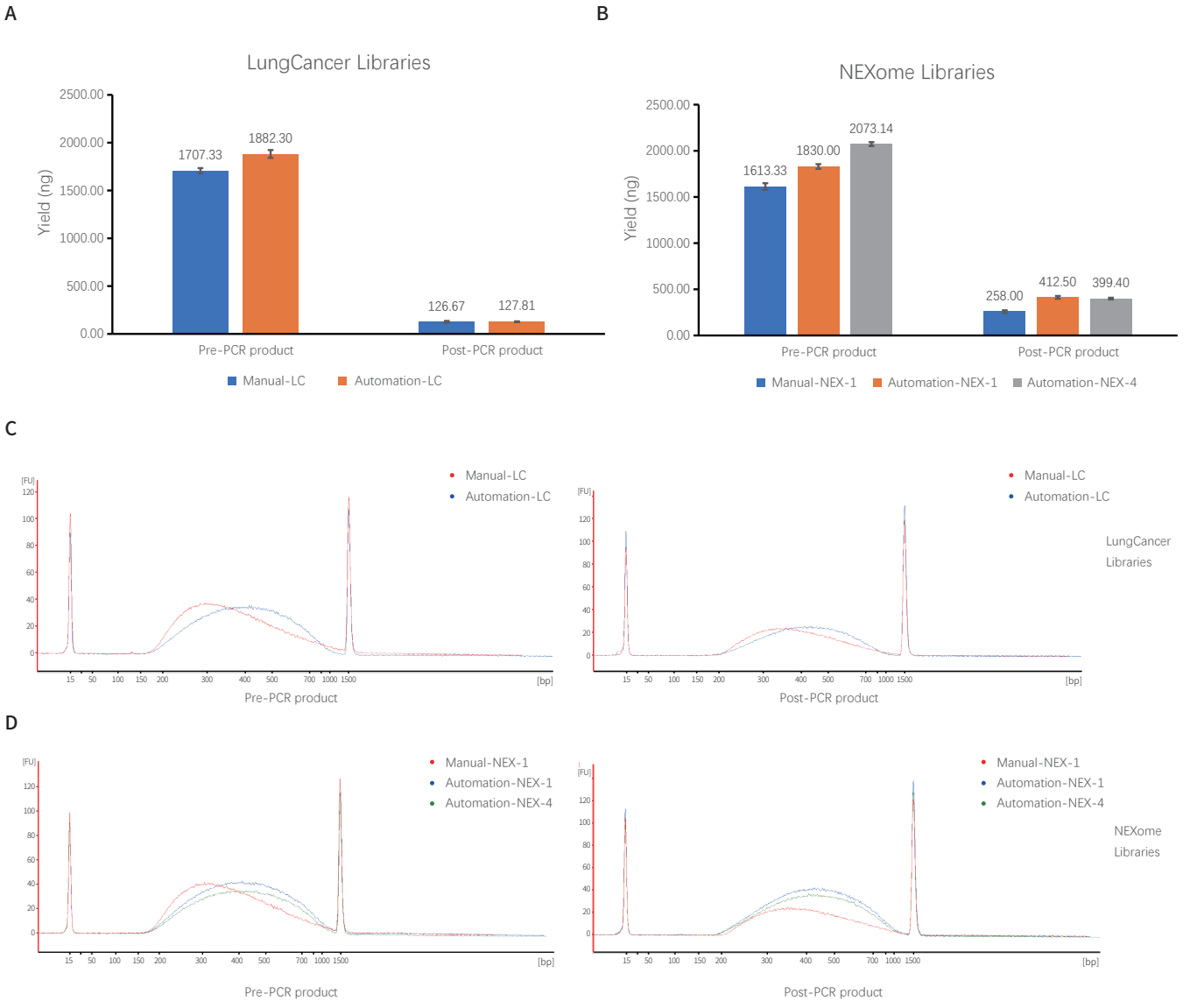



Figure 1. The MGISP-Smart 8 facilitates Nanodigmbio target enrichment solutions to complete high-quality library preparation. (A, B) Analysis of library yield at each quality control point; (C, D) Analysis of fragment size at each quality control point. Manual-LC stands for the library prepared manually with LungCancer Panel; manual-NEX-1/4 stands for the library prepared manually with NEXome Core Panel with 1-plex or 4-plex hybrid strategy, the rest nomination follows similar strategy.

## Nanodigmbio target enrichment solutions are compatible with DNBSEQ sequencing platform

The Q30, "Mapping rate" and "Target rate" of automation-LC group are 94.62%, 99.58% and 90.10%, respectively. This indicates that the LungCancer Panel has a high target rate and high enrichment specificity to target genes. The "% target > 0.2× mean" and "% target > 0.5× mean" are 99.95% and 95.74% respectively. This indicates that this panel has a high target coverage rate (Figure 2A).

The Fold-80 Base Penalty is 1.36, indicating that this panel has an excellent uniformity of target coverage (Figure 2B). When the intercepted data amount is 0.2G, the average sequencing depth of automatic-LC reaches 692.23× (Figure 2C).

The above sequencing results are all equivalent to the manual results, indicating that LungCancer Panel is very suitable for the MGISP automation system and DNBSEQ sequencing platform developed by MGI.

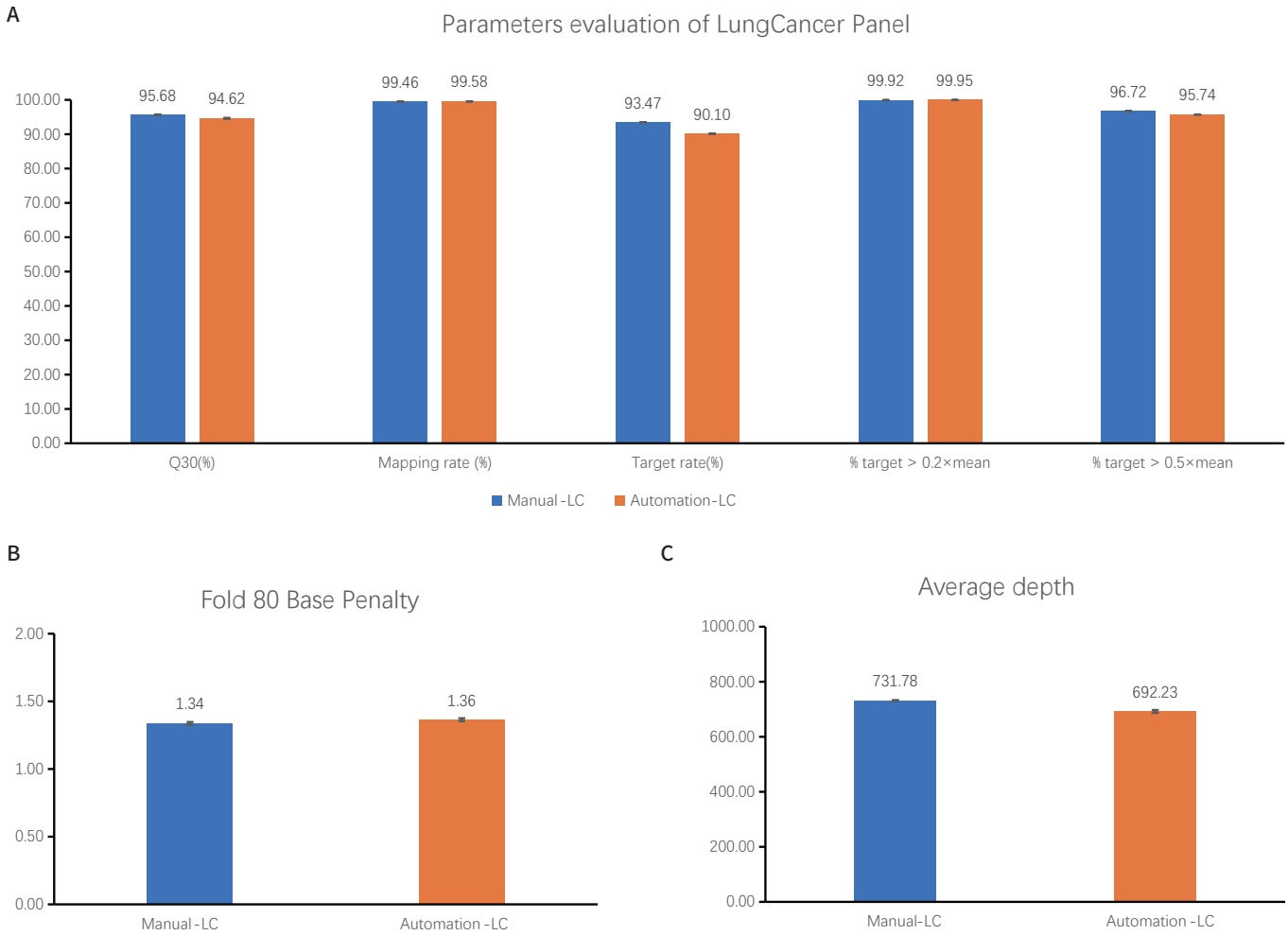



Figure 2. Display of key sequencing parameters of libraries prepared with LungCancer Panel and sequenced on DNBSEQ-G400. (A) Mapping rate: proportion of reads aligned to the reference genome, Target rate: target rate of target gene, %target>0.2×/0.5×mean: target coverage when the average sequencing depth is >0.2×/0.5×. (B) Fold-80 Base Penalty: coverage uniformity. (C) Average depth: average sequencing depth.

For automation-NEX-1 and automation-NEX-4, Q30 is ＞90%, "Duplication rate"is ＜5%, "Mapping rate"is＞99%, "Target rate"is ＞90%, "%target>0.2×mean" is ＞99%, and "%target>0.5×mean"is also ~99%.

The above results indicate that the NEXome Core Panel is also perfectly adapted to MGISP-Smart 8 and DNBSEQ-G400, both 1-plex and 4-plex strategy can obtain ideal sequencing results, and also the above sequencing results are equivalent to manual ones.

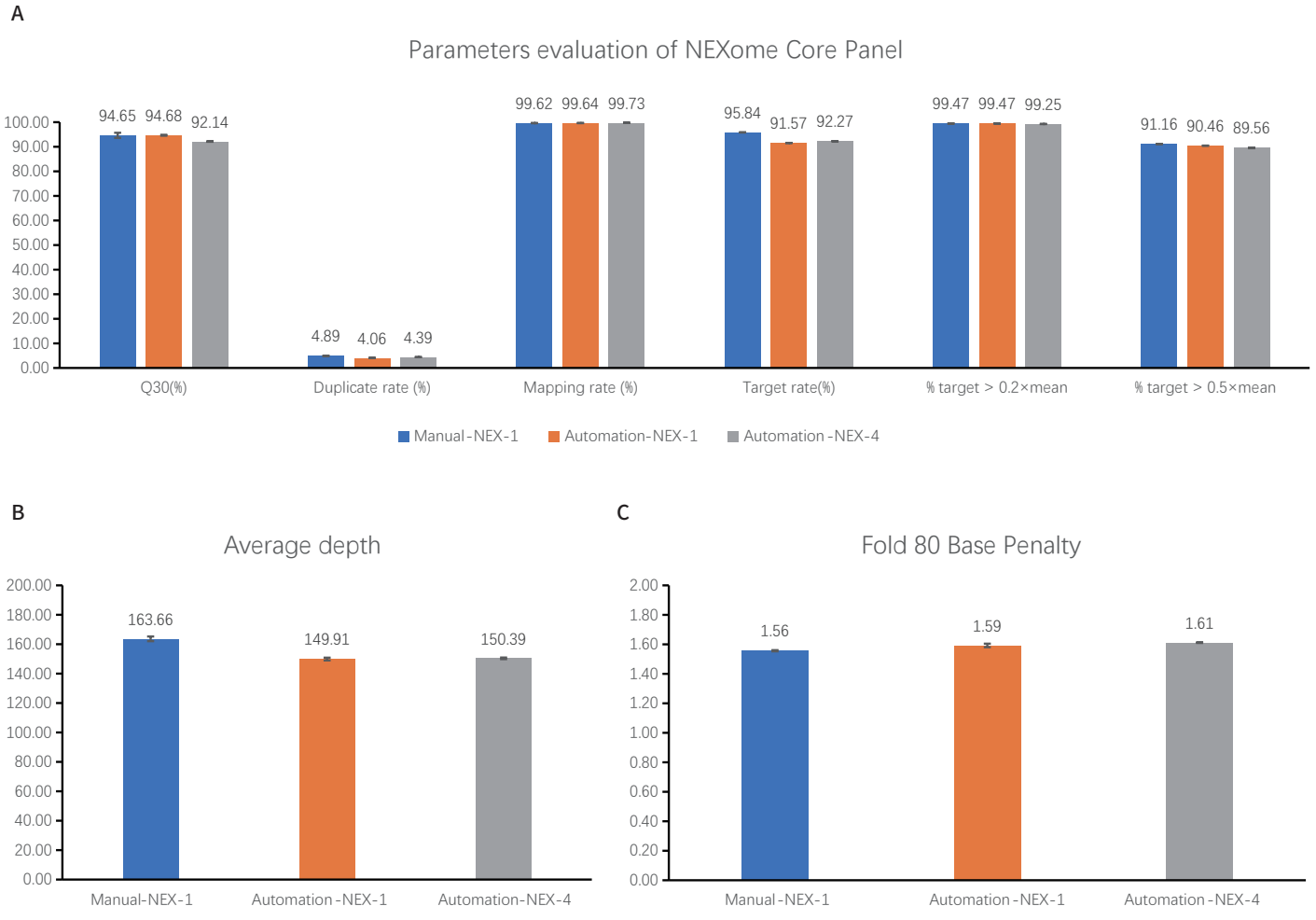



Figure 3. Key sequencing parameters of libraries prepared with NEXome Core Panel and sequenced on DNBSEQ-G400.

## DNBSEQ platform endows the accurate identification of related variations

Further variation detection analysis was carried out for the automation LC group. The results showed that both SNV and InDel variations in tumors could be detected with 100% accuracy, and the measured mutation frequency was highly consistent with the theoretical value (R2 = 0.922) (Figure 5A, B). Similarly, the automation NEX group is also able to detect the related alleles well, and the measured allele frequency is also in good agreement with the theoretical value (R2 is 0.9612 and 0.9620, respectively, Figure 6B) under different hybridization strategies (1-plex and 4-plex). Additionally, the results is no difference from the manual ones (R2 = 0.9625, Figure 6A).

A

| Gene | Mutation site | Theoretical mutation frequency(%) | Detected mutation frequency(%) |
|---|---|---|---|
| *NRAS* | Q61K | 1.00 | 1.37 |
| *KRAS* | A146T | 1.00 | 0.58 |
| *KRAS* | G13D | 4.00 | 3.34 |
| *KRAS* | G12D | 2.00 | 2.08 |
| *PIK3CA* | H1047R | 7.00 | 6.81 |
| *EGFR* | G719S | 4.00 | 4.13 |
| *EGFR* | delE746_A750 | 2.00 | 1.39 |
| *EGFR* | V769_D770_insASV | 3.00 | 1.35 |
| *EGFR* | T790M | 2.00 | 1.44 |
| *EGFR* | L858R | 1.00 | 0.46 |
| *BRAF* | V600E | 7.00 | 5.57 |

B

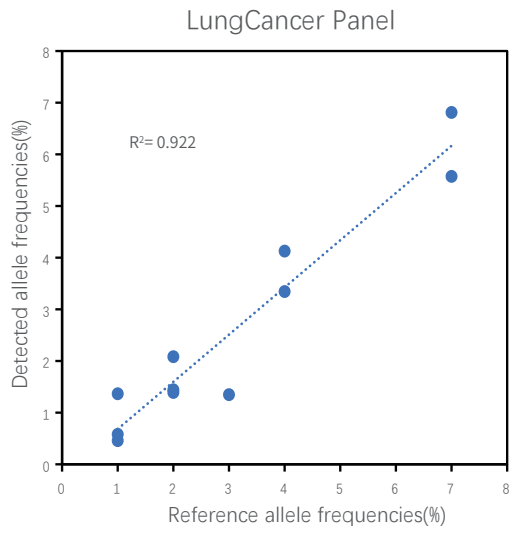
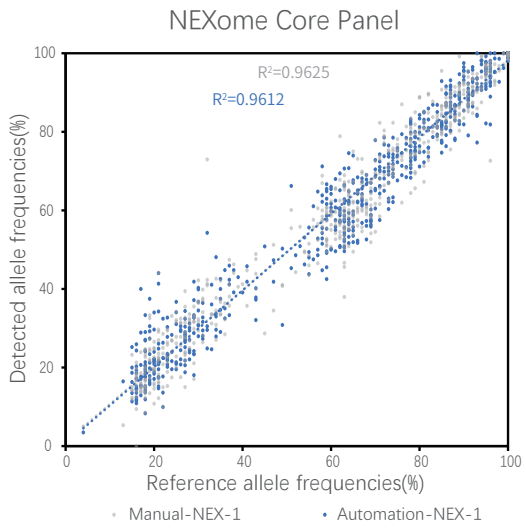



Figure 5. Analysis of variation detection results and mutation consistency of lung cancer libraries.

A




B

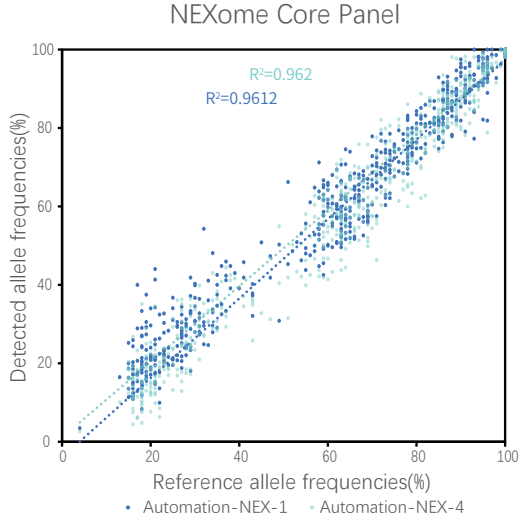
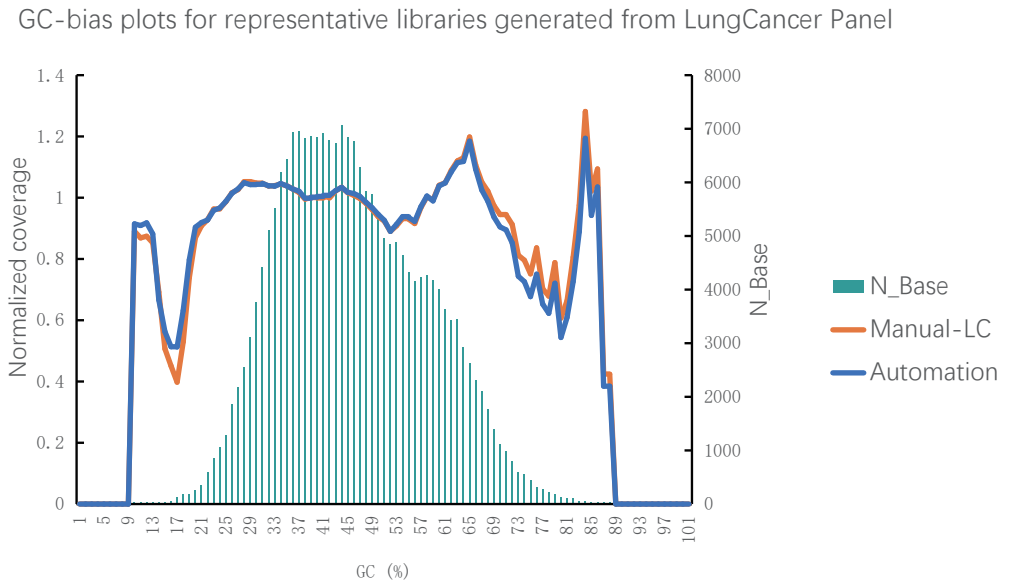



Figure 6. Analysis on consistency of allele frequency in data obtained from the NEXome Core Panel under different hybridization strategies with the allele frequency of standard.

## GC preference analysis of Nanodigmbio targeted enrichment solutions

The GC preference analysis of both solutions shows that the LungCancer Panel can also capture high GC content areas, the NEXome Core Panel has a good coverage under different GC contents, with no difference between automatic and manual strategies (Figure 4A, B).

A




B

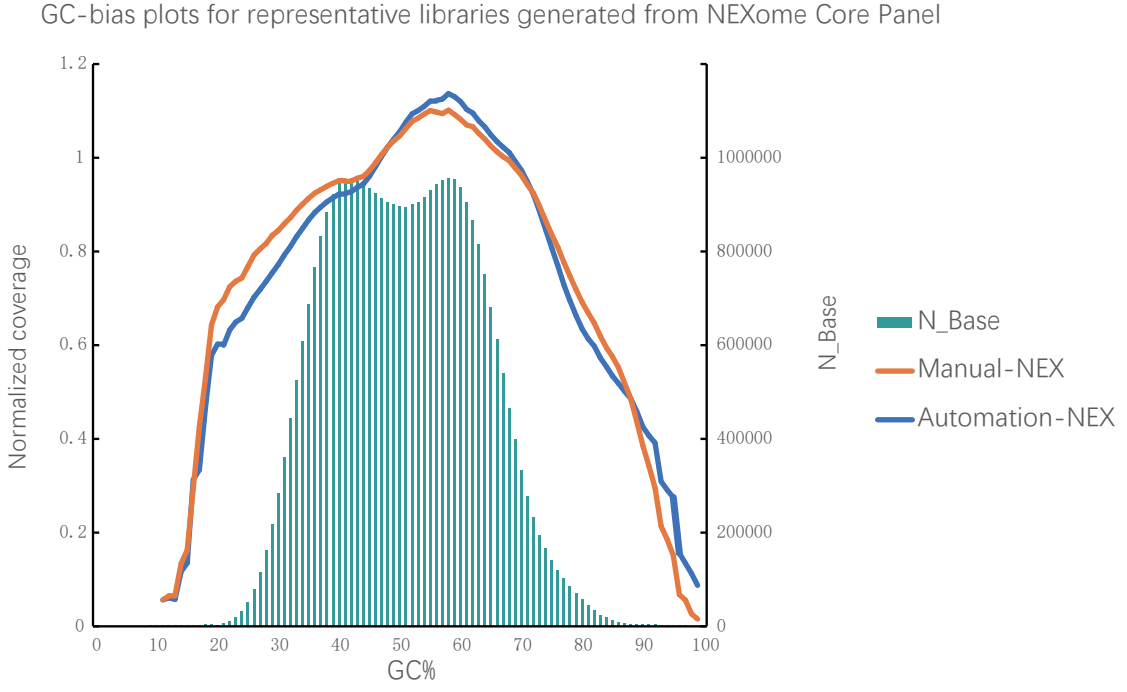



Figure 4. GC preference analysis of the LungCancer Panel and the NEXome Core Panel.

## Conclusion

The comprehensive evaluation of two target enrichment solutions from Nanodigmbio, LungCancer Panel and NEXome Core Panel, shows that those solutions could achieve quick and high-quality library preparation on the MGISP automation system, and followed by sequencing on DNBSEQ platform can output high-quality sequencing data.

MGISP-Smart 8 automation system can not only realize the flexible library preparation of 1~48 samples per run, but also automatically operate pooling, homogenization and quantitative system preparation. This feature could greatly ensure the accuracy of results and saves labor costs while improving automation efficiency. Moreover, DNBSEQ-G400 supports 1-2 slides running per operation and various sequencing reading lengths. PE150-FCL only needs 50 h, making it one of the first choice models for large and medium-sized sequencing laboratories.

DNBSEQ-G99 genetic sequencer

DNBSEQ-G400 genetic sequencer

MGISP-Smart 8 automated sample preparation system

## References


1. Xuchao, Z. et al. Application of next-generation sequencing technology to precision medicine in cancer: joint consensus of the Tumor Biomarker Committee of the Chinese Society of Clinical Oncology. *Cancer Biology & Medicine* 16, doi:10.20892/j.issn.2095-3941.2018.0142 (2019).
2. McInerney-Leo, A. M. & Duncan, E. L. Massively Parallel Sequencing for Rare Genetic Disorders: Potential and Pitfalls. *Frontiers in Endocrinology* 11, doi:10.3389/fendo.2020.628946 (2021).
3. Shi, C. *et al.* Development and clinical applications of an enclosed automated targeted NGS library preparation system. *Clinica Chimica Acta* 540, doi:10.1016/j.cca.2023.117224 (2023).
4. Hess, J. F. *et al.* Library preparation for next generation sequencing: A review of automation strategies. *Biotechnology Advances* 41, doi:10.10 16/j.biotechadv.2020.107537 (2020).

## Recommended Ordering Information

| Category | Product | Cat. NO. |
|---|---|---|
| Instruments | Genetic Sequencer DNBSEQ-G99 | 900-000607-00 |
| | Genetic Sequencer DNBSEQ-G400RS | 900-000170-00 |
| | MGISP-Smart 8RS Automated Sample Preparation System | 900-000503-00 |
| Library Prep | NadPrep® EZ DNA Library Preparation Module v2 (96 rxn) | 1002602* |
| | NadPrep® Universal Adapter (MDI)Module Set B1 (for MGI, 96 rxn) | 1003721* |
| | NEXome Core Panel (16 rxn) | 1001852* |
| | LungCancer Panel v1.0 (16 rxn) | 1001922* |
| | NadPrep® Hybrid Capture Reagents (96 rxn) | 1005101* |
| | NadPrep® NanoBlockers (for MGI, DI, 96 rxn) | 1006208* |
| | NadPrep® M-Amplification Primer Mix (for MGI, DI, 96 rxn) | 1004204* |
| Sequencing Reagents | DNBSEQ-G99RS High-throughput Sequencing Reagent Set（G99 FCL PE150） | 940-001269-00 |
| | DNBSEQ-G400RS High-throughput Sequencing Set (FCL PE150) | 1000016952 |
| | CPAS Barcode Primer 3 Reagent Kit | 1000020834 |

*The relevant products could be searched and ordered at Nanodigmbio homepage http://www.njnad.com/en/.


## MGI Tech Co.,Ltd

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





Version: June 2024



+86-4000-688-114

en.mgi-tech.com

MGI-service@mgi-tech.com



Authors: Hanfei Zhang

Editor-in-Charge: Qiwei Wang

Reviewer: Yao Jiang