A New Approach to DNA analysis from FFPE-derived DNA and cfDNA: NGS library preparation methods for Single-strand and Double-strand DNA

○Shohei Torii, Yuto Kitajima, Ayuka Higuchi, Yoshitaka Kimura, Hiroyuki Matsumoto / Takara Bio Inc.

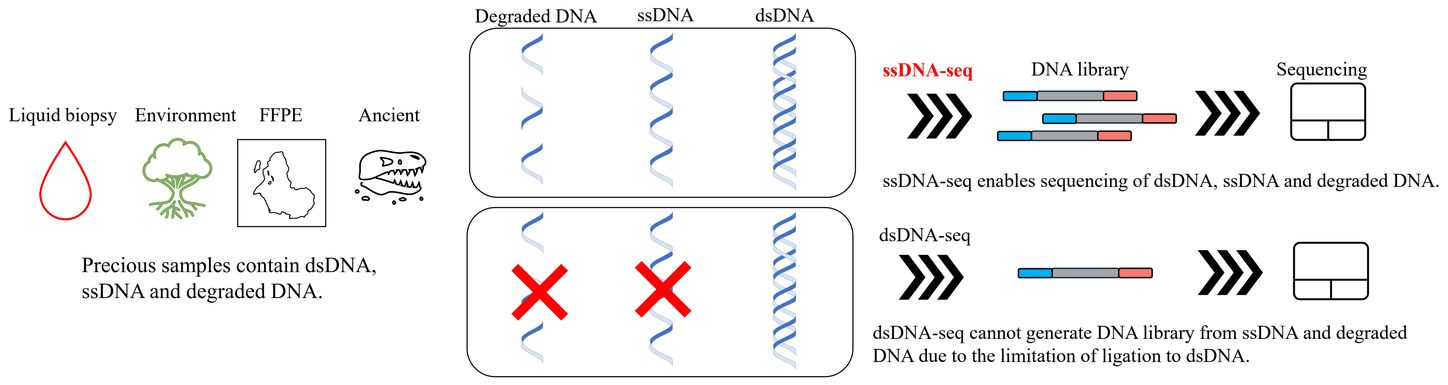
COI Disclosure Information
All presenters are employees of Takara Bio Inc.

Abstract

Background: DNA analysis from formalin-fixed paraffin-embedded (FFPE) samples and cell-free DNA (cfDNA) is essential in various fields such as cancer gene panel analysis, clinical sequencing, infectious disease research, genetic disease research, and forensic science. FFPE-derived DNA and cfDNA are often highly fragmented and consist of a mixture of double-strand DNA (dsDNA) with overhangs, nicked dsDNA, and single strand DNA (ssDNA), which conventional dsDNA-oriented library preparation methods may not fully capture. Therefore, a single strand DNA NGS library preparation method using a single strand DNA ligase (SDL) was developed and optimized for FFPE-derived DNA and cfDNA.
Results: The SDL-based approach yielded sufficient libraries for sequencing from high-degradation FFPE DNA and cfDNA. It improved key metrics compared with commercial kits, including higher mapping rates and lower PCR duplicate rates. With 50 base synthetic ssDNA, the mapping rate increased from ~30% using a conventional ssDNA-seq method to >90% with the SDL-based approach, demonstrating effective recovery of short ssDNA.
Conclusion: These findings suggest that the single strand DNA NGS library preparation method is highly effective at recovering a broader diversity of DNA fragments, including ssDNA and damaged molecules from degraded samples such as FFPE and cfDNA, underscoring its utility for sensitive DNA analysis.

Introduction

①ssDNA-seq resolves challenges that are difficult to overcome with dsDNA-seq in precious samples

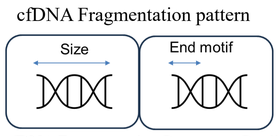


Damaged DNA (for example, nicked DNA and DNA with overhangs) and ssDNA are abundant in cfDNA, FFPE, ancient and environmental samples.
However, dsDNA-seq shows poor recovery of damaged DNA and ssDNA, leading to biased representation of biologically informative fragments.
ssDNA-seq overcomes key limitations inherent to dsDNA-seq by enabling efficient recovery of damaged DNA and ssDNA molecules.

Challenges of Conventional dsDNA-seq
- Underrepresentation of ultra-short, nicked, overhang-containing, and ssDNA fragments
- Causing end/size biases and altering native fragmentation patterns
- Lower conversion efficiency for crosslinked, chemically damaged DNA (e.g., FFPE DNA)
- Poor recovery of degraded and modified molecules in ancient or environmental samples

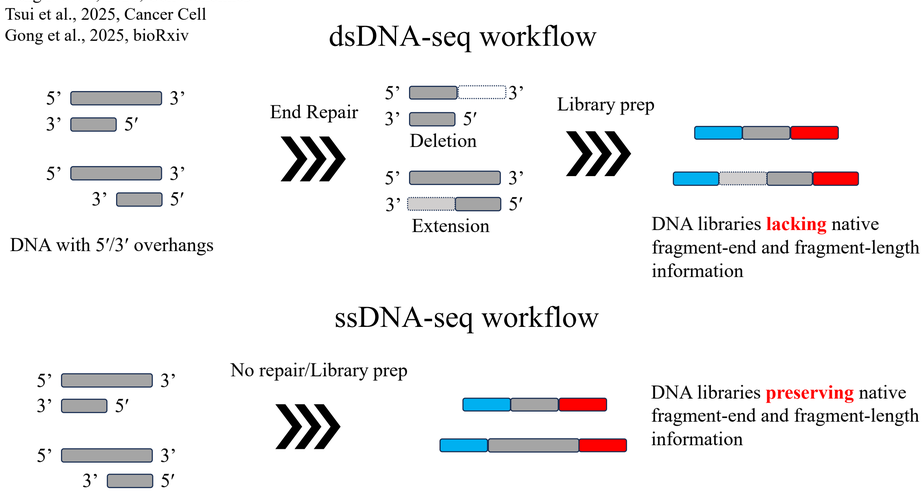
Advantages of ssDNA-seq
- Efficient recovery of nicked DNA, DNA with overhangs, and ssDNA
- High-resolution capture of termini, size profiles, and fragmentation motifs
- Improved library complexity for severely degraded samples
- High compatibility with ultra-short cfDNA (~50 nt), microbial cfDNA, mtDNA
- Suitable for low-input and damage-rich sample types (FFPE, ancient DNA)

②ssDNA-seq provides high-quality data for fragmentomics analyses.



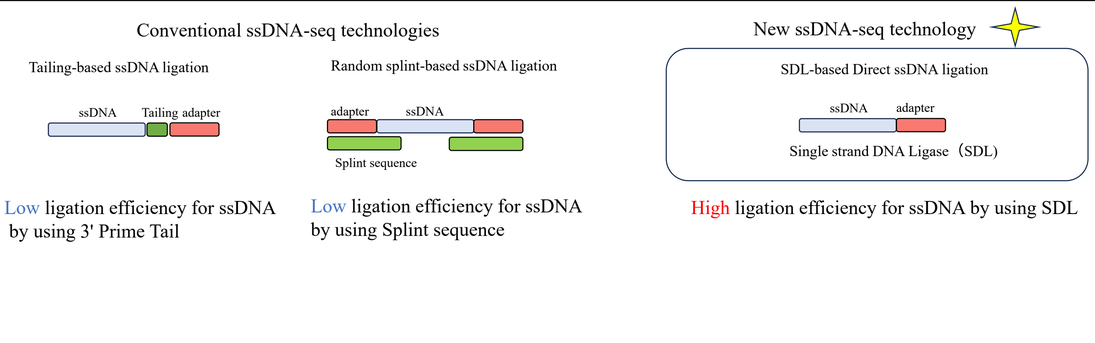
cfDNA Fragmentation Patterns as Predictors of Cancer
Jiang P. et al., 2026, Cell Genomics
Tsui et al., 2025, Cancer Cell
Gong et al., 2025, bioRxiv

These studies indicated distinct fragmentomic patterns, including size and end-motif features, that are strongly linked to cancer detection. Integrating these features into machine-learning models improves cancer detection and tissue-of-origin prediction (Tsui et al., 2025; Gong et al., 2025), and ssDNA-seq improves cancer detection by preserving the native 3′ ends and adding new end-motif features (Jiang et al., 2026).



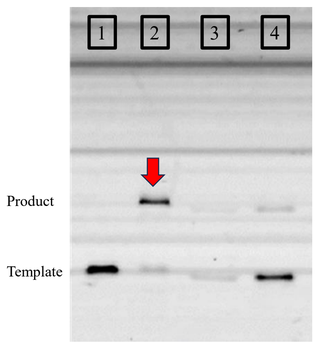
End repair in conventional dsDNA-seq generates libraries that lose native fragment information.
In contrast, ssDNA-seq does not require end repair and preserves native fragment information.

③Single strand DNA Ligase is a novel enzyme that shows high ligation efficiency for ssDNA



We developed Single strand DNA Ligase (Takara Bio Inc) and compared its ligation efficiency with that of conventional enzyme.

Experiment: Ligation between ssDNA and ssDNA

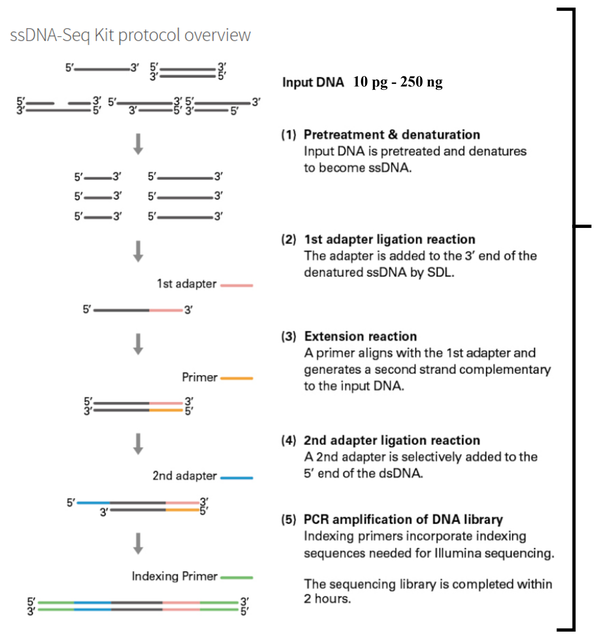


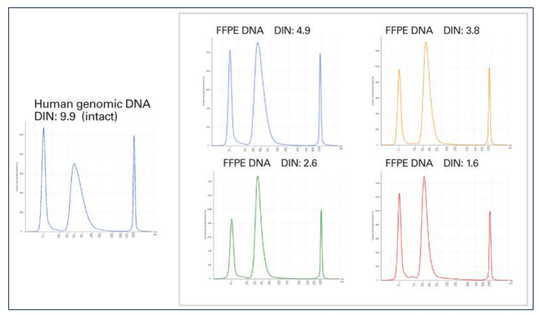
1.No enzyme
2.Single strand DNA Ligase (SDL)
3.Conventional enzyme A
4.Conventional enzyme A + additive

Both enzymes used were commercially available products.

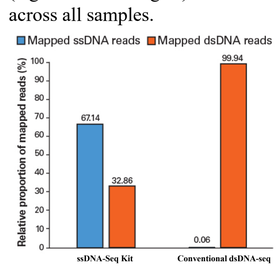
SDL achieves higher ligation efficiency than conventional enzyme.

④The SDL-based NGS library preparation performance: ssDNA-Seq Kit





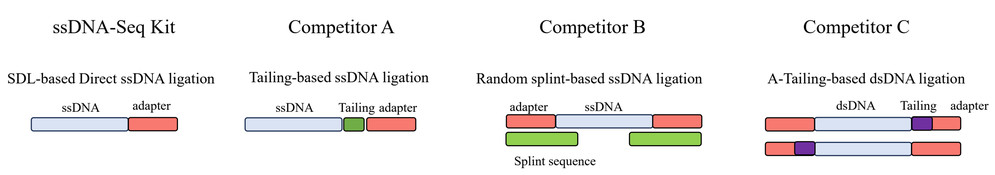
DNA libraries were prepared using the ssDNA-Seq Kit (Takara Bio Inc.) from human genomic DNA and FFPE-derived DNA samples with varying degrees of fragmentation. Input DNA, 10 ng for each sample, was fragmented using an ultrasonic generator (Covaris). Evaluation of the prepared libraries with the TapeStation HD5000 (Agilent Technologies) confirmed expected DNA library traces across all samples.



DNA libraries were prepared with the ssDNA-Seq Kit or Conventional dsDNA-seq using a mixture of synthetic ssDNA and dsDNA. The results demonstrate that the dsDNA-seq can capture dsDNA, while the ssDNA-Seq Kit can capture both ssDNA and dsDNA.

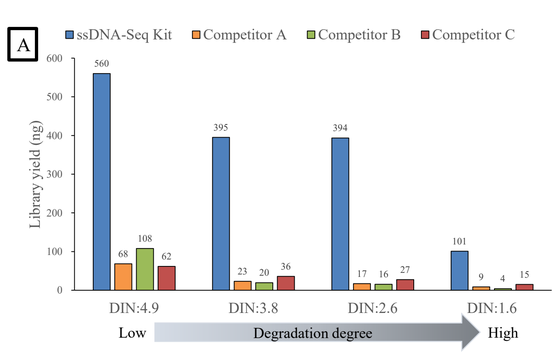
Performance Comparison of ssDNA-seq Kit and Competitor Kits

A. Overview of commercially available ligation technologies



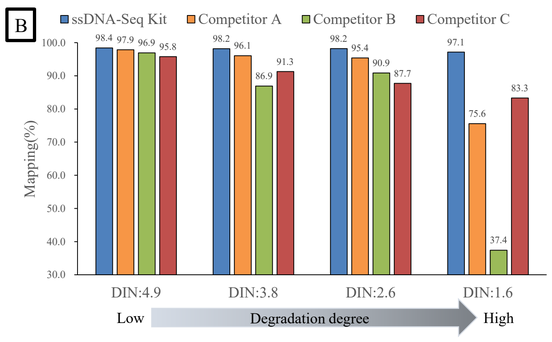
All of the kits used were commercially available products.

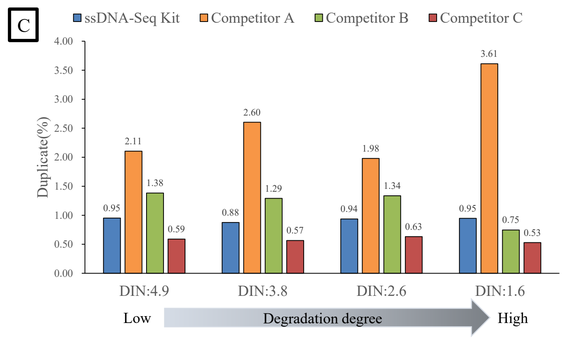
B. FFPE DNA



Highest library yield from FFPE samples as compared to competitor kits.
Panel A. DNA derived from FFPE samples (BioIVT) with different degradation degrees was validated using the TapeStation Genomic DNA (Agilent Technologies). Libraries were prepared according to the protocol of each kit using 10 ng of DNA extracted from FFPE samples
The ssDNA-Seq Kit had the highest library yield among the competitor kits for all FFPE-derived DNA. DIN = DNA integrity number.

Highest mapping rate and low duplicate rate as compared to competitor kits.
Panel B. The mapping rate of Competitor A, B, and C kits decreased significantly as DNA degradation progressed. On the other hand, ssDNA-Seq Kit showed a high mapping rate of over 97% and a low duplicate rate (Panel C) for all FFPE sample-derived DNA, regardless of the degradation degree. Overall, the ssDNA-Seq kit demonstrates superior performance compared with competitor kits.





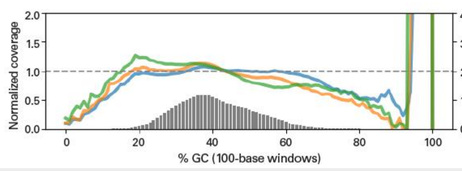
C. cfDNA

D

	Mapping (%)	Duplicate (%)	Library yield (ng)
ssDNA-Seq Kit	**98.8**	**0.92**	**468**
Competitor A	96.2	2.39	40
Competitor B	94.6	1.69	82

Highest library yield, highest mapping rate and low duplicate from cfDNA samples as compared to competitor kits.
Panel D. Libraries were prepared according to the protocol of each kit using 10 ng of cfDNA extracted from human plasma DNA (BioIVT).
The ssDNA-Seq Kit had the highest library yield, highest mapping rate and lowest duplicate rate among the competitor kits.

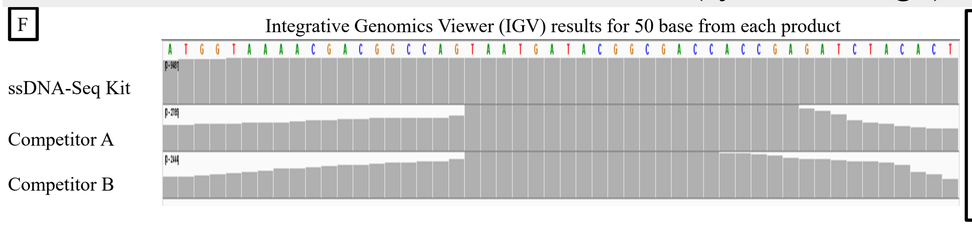
E



cfDNA library coverage uniformity as compared to competitor kits.
Panel E. The vertical gray bars represent the expected GC content distribution using 100 bp windows. The ssDNA-Seq Kit showed the most uniform coverage, independent of GC content.
The ssDNA-Seq Kit, which provides high-quality sequencing results, is optimal for the analysis of cfDNA.

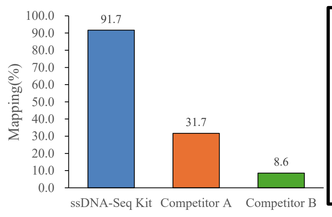
D. Ultra-Short DNA (synthetic oligo)

F



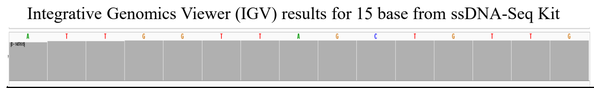
DNA libraries that preserve expected fragment-end and fragment-length information, compared with competitor kits.
Panel F. Libraries were prepared according to the protocol of each kit using 10 ng of 50 base synthetic oligo. The gray bars represent coverage. The ssDNA-Seq Kit achieved the most uniform coverage and retained expected fragment-end/length information, whereas Competitor A and B kits showed reduced end-region coverage and loss of expected fragment features.

G



Highest mapping rate from 50 base synthetic oligos as compared to competitor kits.
Panel G. The ssDNA-Seq Kit showed mapping rates of over 90% for the 50 base synthetic oligo and achieved the highest mapping rate among the competitor kits.

H



Library preparation from 15 base synthetic oligo.
Panel H. The DNA library was prepared from 30 ng of 15 base synthetic oligo. The gray bars represent coverage. Even with a 15 base synthetic oligo, the ssDNA-Seq Kit achieved the uniform coverage and retained the expected fragment-end/length information.

Conclusion

ssDNA-seq improves the recovery efficiency and sequencing quality of short and damaged DNA, supports high-resolution fragmentomics, and expands analytical capability to sample types such as cfDNA and FFPE, for which dsDNA-seq has intrinsic limitations.
Among the ssDNA-seq kits tested in this study, the ssDNA-Seq Kit (Takara Bio Inc.) showed the best performance.

Reference

Jiang P. et al., 2026, Cell Genomics
Zhao Z. et al., 2025, Analyst
Gong et al., 2025, bioRxiv
Tsui et al., 2025, Cancer Cell
Miura F. et al., 2025, NAR
Williams A.D. et al., 2025, Trends Microbiology
Cheng J.C. et al., 2024, Frontiers in Oncology
Steiert T.A. et al., 2023, NAR
Miura F. et al., 2023, Sci Rep
Stiller M. et al., 2016, Oncotarget
Gansauge & Meyer, 2013, PNAS