Application Note

Genomics

Agilent

Trusted Answers

Comparison of Different Methods to Isolate HMW DNA from Bacteria for Nanopore Sequencing

Authors

Christina Engesser,
NGS Competence Center Tübingen

Whitney Pike,
Agilent Technologies

Abstract

The first step in many sequencing experiments is to isolate nucleic acids from a specimen. As long-read sequencing technologies have advanced, many commercial companies now offer kits specific to the extraction of high molecular weight (HMW) genomic DNA (gDNA). However, the size and quality of the gDNA can vary greatly depending upon the extraction method used. For sequencing facilities, this can affect the type of sequencing that is performed, making knowledge of sample integrity crucial to a successful experiment. In this application note, researchers at the NGS Competence Center Tübingen (NCCT) used the Agilent Femto Pulse system to analyze HMW gDNA from five different commercially available extraction kits. The size of the input material was then compared to the sequenced read lengths to demonstrate the impact that the isolation method can have on sample size and sequencing results.

Introduction

Oftentimes, the goals of a project dictate the methods used for sample preparation. For example, when determining which gDNA extraction kit to use, a researcher must consider if the input material is eukaryotic or prokaryotic, and must also determine the final sample type (DNA, RNA, amplicons, or plasmids). Additionally, the expected fragment length should be considered, as the length of the fragments composing a library will influence the total sequencing output, and different lengths can be achieved with different extraction methods (column versus precipitation). Furthermore, the sequencing type depends on the project's requirements. When trying to identify variants, the goal may be to produce a higher sequencing depth, and short read sequencing would result in the larger output needed for this. On the other hand, if the goal of the project is to close a genome, the longer fragments generated by long read sequencing may be ideal.

Nanopore sequencing enables real-time sequencing of both DNA and RNA, and allows for the analysis of any fragment length, from short to ultralong. Thus, Nanopore sequencing technologies are used for a variety of applications, including whole genome, targeted, whole transcriptome, and metagenomic sequencing. Advantages of sequencing with Oxford Nanopore Technologies systems include a portable or desktop sequencer, simple and rapid library preparation, and a high yield, enabling sequencing of whole genomes. Successful long-read sequencing depends on having high-quality samples of a size sufficient for the desired sequencing outcomes. Several isolation kits are commercially available for obtaining HMW gDNA, but each can produce samples of different sizes or integrity. Quality control (QC) of the initial gDNA prior to sequencing is essential to obtaining the sample size needed for successful sequencing. The Agilent Femto Pulse system is an automated capillary electrophoresis instrument designed for nucleic acid QC. The Agilent gDNA 165 kb kit for the Femto Pulse uses an optimized pulsed-field separation method that enables accurate sizing of gDNA through 165 kb[1], making it ideal for analysis of HMW gDNA prior to long read sequencing.

Researchers at the NGS Competence Center Tübingen (NCCT), a part of the DFG (German Research Foundation) compared the sample quality, size, and sequencing results of HMW gDNA from five commercially available kits. The core facility is a collaboration between several institutes, each specializing in different sample types and project requirements. Common applications of the projects performed by the facilities include bacterial whole genome sequencing, RNA sequencing, and metagenomics, among others. In this application note, the NCCT used the Agilent Femto Pulse and the Oxford Nanopore Technologies MinION Flongle to analyze bacterial gDNA isolated from different extraction kits to demonstrate the importance of sample QC prior to sequencing.

Experimental

HMW gDNA was isolated from bacteria (E. coli 1x10⁹ cells per reaction Lysis with 15 mg/mL Lysozyme (Sigma-Aldrich, p/n L6876) and 100 µL/mL Proteinase K (Thermo Fisher Scientific, p/n EO0492)) using five different extraction kits (Table 1). Purity, yield, size, and sequence results were compared. Purity was measured by Nanodrop, as recommended by Nanopore sequencing protocol. Concentration was measured by both Nanodrop and Qubit, with the values from the Qubit used to calculate the total yield. The size of the gDNA was determined using the Femto Pulse with the gDNA 165 kb kit (p/n FP-1002) prior to library preparation, and compared to the size of the sequenced DNA fragments. Samples were prepared for analysis with the Femto Pulse according to the gDNA 165 kb kit Quick Guide[2], with the following modifications: samples and reagents were left at room temperature for three hours prior to use; the Agilent FP 165 kb Ladder (p/n FP-7002-U035) was diluted 1:20 in Agilent FP gDNA Diluent Marker (p/n FP-8001-0003), and the samples were diluted to a concentration of 400 pg/µl or less. Sequencing was performed using a MinION Flongle with 34 open pores, and the fragment length was reported after 24 hours.

Results and discussion

The integrity and size of the input sample is crucial to the success of sequencing experiments and the ability to achieve long reads. Sample fragmentation and degradation can cause issues with library preparation, leading to suboptimal sequencing results. With nanopore sequencing, the entire length of a fragment can be sequenced. The desired length will depend upon the goals of the project – for more sequencing output, smaller fragments are more desirable, as they will be preferentially sequenced over larger fragments. If ultrahigh molecular weight (UHMW) samples greater than 200 kb are necessary for a project, it is important to use a nucleic acid extraction kit that can help achieve these long lengths, while also eliminating smaller fragments. Thus, QC of the input sample is crucial to the design of sequencing projects.

The NCCT is a core facility which processes samples for a wide variety of projects and has significant experience with several commercially available extraction kits and sequencing methods. When performing long-read sequencing with the Oxford Nanopore Technologies sequencing systems, the NCCT's workflow includes determining sample purity, concentration, and size prior to sequencing. To demonstrate how different extraction kits can impact sequencing results, bacterial gDNA was isolated using five HMW gDNA extraction kits and sequenced with the MinION Flongle. QC of the extracted gDNA was performed prior to library preparation. The purity and yield of the samples were measured by both Nanodrop and Qubit (Table 2). The size of the input DNA for each extraction method was determined using the Femto Pulse. The resulting libraries were sequenced using the MinION Flongle.

Table 1. Kit comparison.

Kit	Kit Type	Expected Fragment Size	Hands-on Time	Pros	Cons
A	Glass beads	From 50 to ≥ 500 kb and into the Mb range	~ 45 minutes	Fast workflow; accurate defined fragment length	Not high throughput due to manual intervention steps; DNA can tear off the beads
B	Gravity-flow, anion-exchange tips	Up to 150 kb with an average size of 50–100 kb	~ 2-3 hours	Easy workflow; flexible input requirements	Not high throughput; higher input can lead to blocks, reagents need to be pushed through the column which takes time (gravity based); needs a lot of bench space for more samples; requires extra kit (gDNA buffer set)
C	Magnetic	200 to > 500 kb	~ 1.5 hours	Clean workflow; higher throughput	Magnetic stand limits the throughput
D	Magnetic disks	HMW DNA (50 kb to 300 + kb) and UHMW DNA (50 kb to 1 Mb +)	~ 1.5 hours	Clean workflow; can generate ultrahigh molecular weight fragments	Magnetic stand limits the throughput; DNA can tear off the disk
E	Column	> 50 kb	~ 30 minutes	Easy workflow; flexible input requirements; very fast; high yield	No HMW DNA; columns do not fit well into collection tubes

Table 2. gDNA yield and purity.

Kit	Nanodrop Purity* 260/280	Nanodrop Purity* 260/230	Qubit Concentration [ng/µl]	Total Amount [µg]
A	1.84	0.8	23.4	1.872
B	1.69	0.93	27.2	2.72
C	1.91	1.83	97	9.7
D	2.14	1.76	4.89	0.489
E	1.88	2.24	155	5.425

**Optimal purity OD 260/280 ~ 1.8; OD 260/230 2.0-2.2.*

Extraction kit A (glass beads)

Extraction kit A resulted in gDNA that appeared as a single large fragment with a peak size greater than 140 kb on the Femto Pulse (Figure 1A). The sequencing results displayed a fairly even distribution of fragments sized through 100 kb (Figure 1B). Further analysis of the sequencing reads exhibited the presence of ultrahigh molecular weight fragments from 100 kb through 400 kb (Figure 1C).

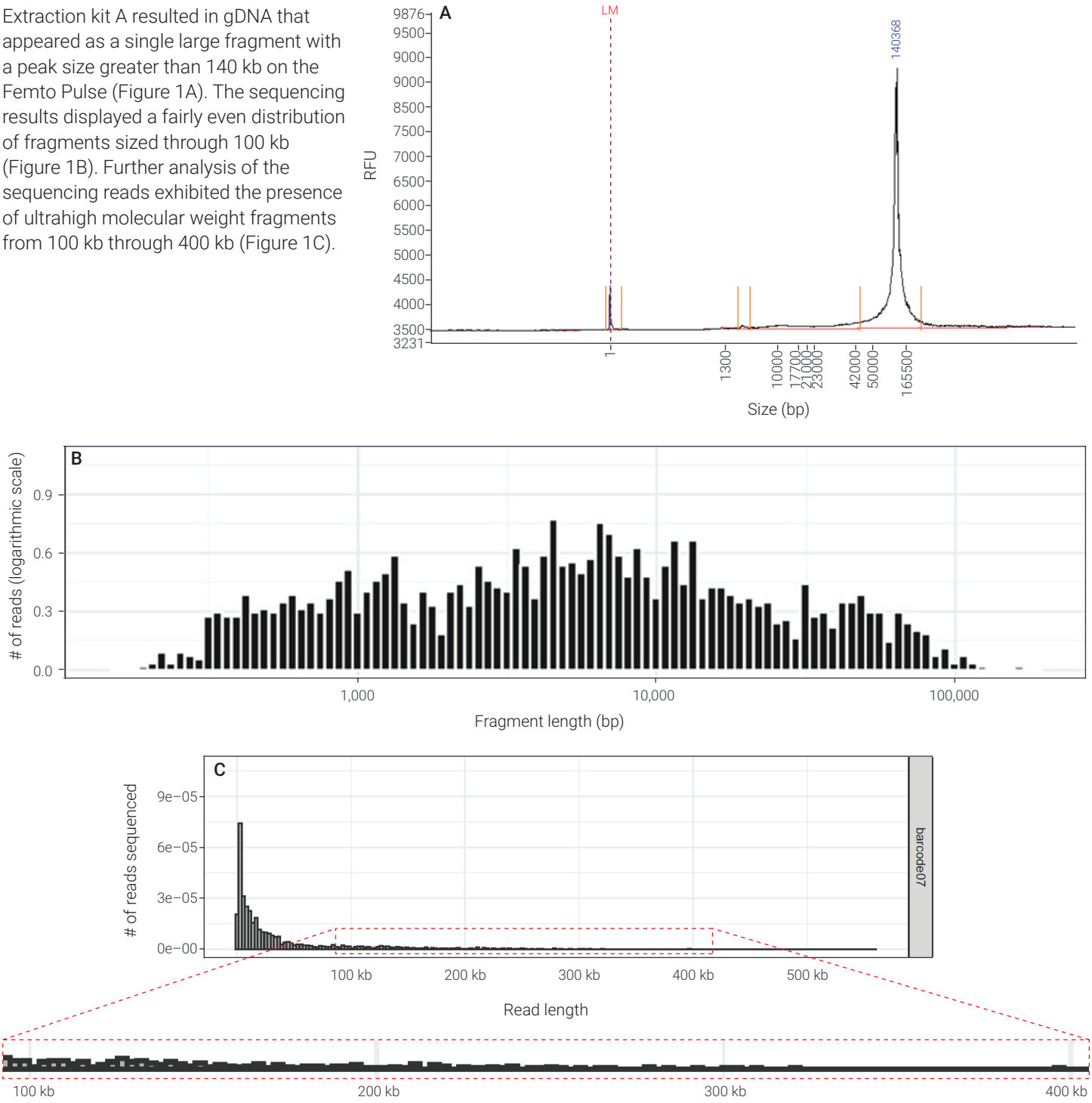

Figure 1. Bacterial gDNA isolated using extraction kit A was analyzed with the Agilent Femto Pulse system and sequenced with the MinION Flongle A) Femto Pulse electropherogram and digital gel image of the input gDNA B) Summary of sequencing read lengths in logarithmic scale C) Further analysis of the number of reads sequenced by read length indicates the presence of sequenced fragments between the sizes of 100 and 400 kb.

Extraction kit B (gravity-flow)

Data generated from the Femto Pulse system demonstrated that samples extracted using method B also resulted in a large peak but displayed some smearing to the left of the peak indicating the presence of smaller fragments within the sample (Figure 2A). The distribution of this sample was confirmed with the sequencing results, which showed a large amount of fragments of smaller sizes (Figure 2B), but also some as big as 500 kb (Figure 2C).

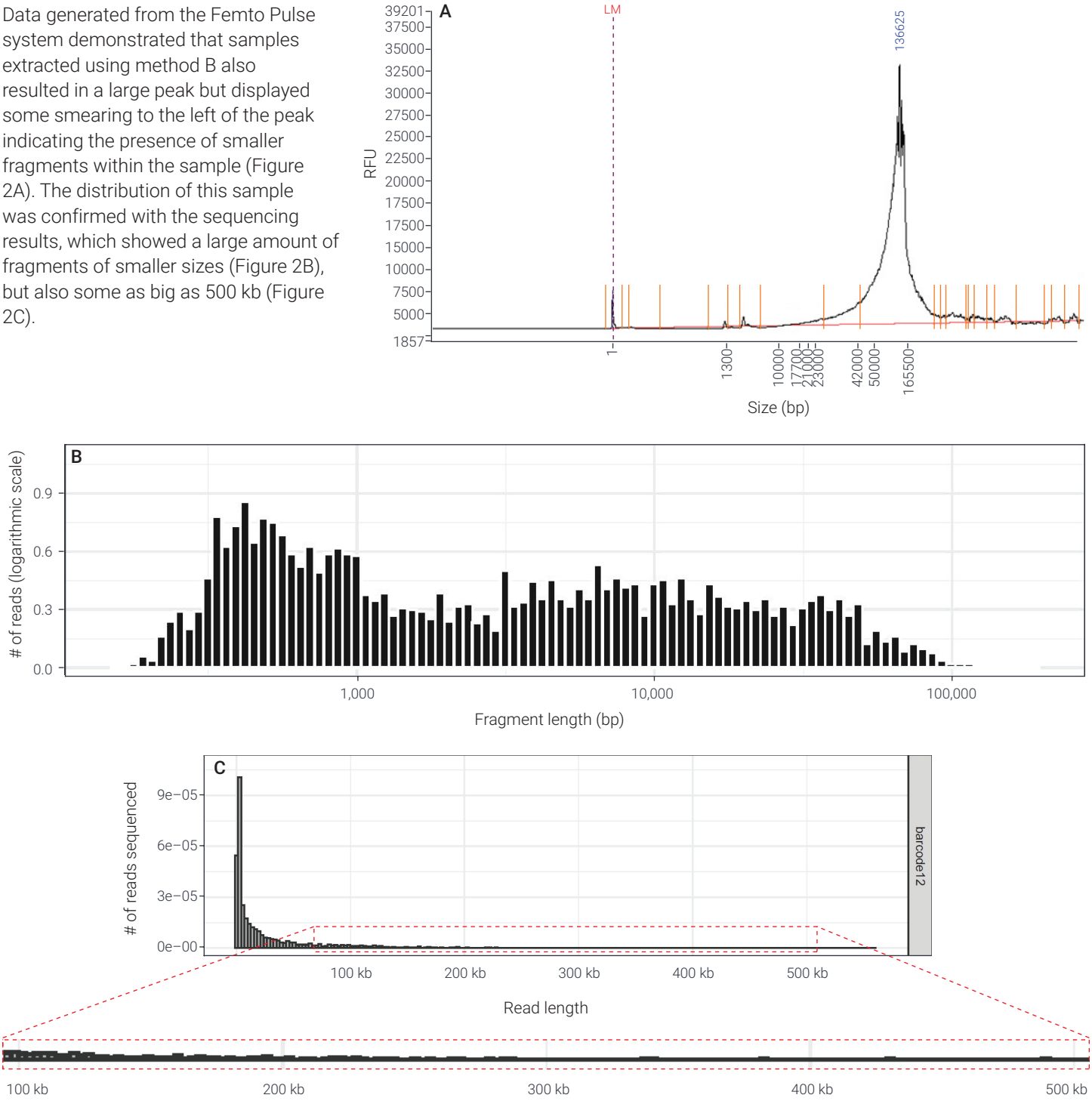

Figure 2. Bacterial gDNA isolated using extraction kit B was analyzed with the Agilent Femto Pulse system and sequenced with the MinION Flongle A) Femto Pulse electropherogram and digital gel image of the input gDNA B) Summary of sequencing read lengths in logarithmic scale C) Further analysis of the number of reads sequenced by read length indicates the presence of sequenced fragments up to 500 kb.

Extraction kit C (magnetic)

Samples from extraction method C when analyzed on the Femto Pulse also showed a large single peak, with ample smearing to the left. Additionally, a small smear ~ 100 bp in length was observed close to the Lower Marker in the electropherogram (Figure 3A). This could indicate the presence of small fragments, or potentially degraded sample. The sequencing results for the DNA from this method showed some reads longer than 500 kb (Figure 3B-C).

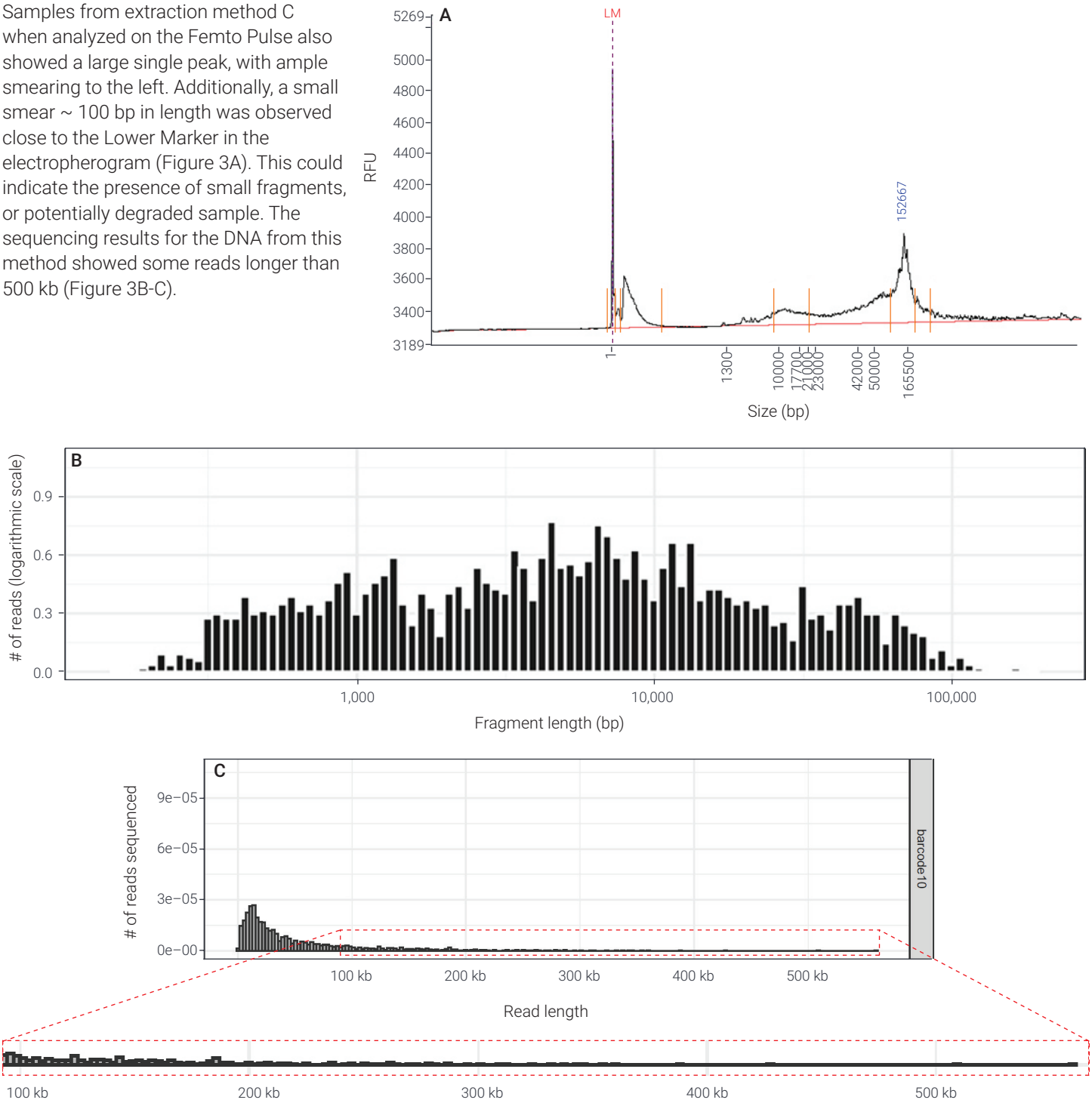

Figure 3. Bacterial gDNA isolated using extraction kit C was analyzed with the Agilent Femto Pulse system and sequenced with the MinION Flongle A) Femto Pulse electropherogram and digital gel image of the input gDNA B) Summary of sequencing read lengths in logarithmic scale C) Further analysis of the number of reads sequenced by read length indicates the presence of sequenced fragments through 500 kb.

Extraction kit D (magnetic disks)

In contrast to the sharp peaks and large sizes of the gDNA from methods A-C on the Femto Pulse, gDNA extracted from methods D-E appear to be more spread out. The Femto Pulse electropherogram of the samples from extraction method D showed a large variety of fragments of varying sizes, as indicated by the presence of several peaks. While some of these peaks are degraded and small fragments, for long read sequencing, the most important fragments are those greater than 100 kb. This extraction method produced gDNA that showed several peaks even larger than 200 kb, making it ideal for isolation of HMW gDNA (Figure 4A). The sequencing data exhibited a large number of fragments of a smaller size, but also displayed fragments > 400 kb (Figure 4B-C).

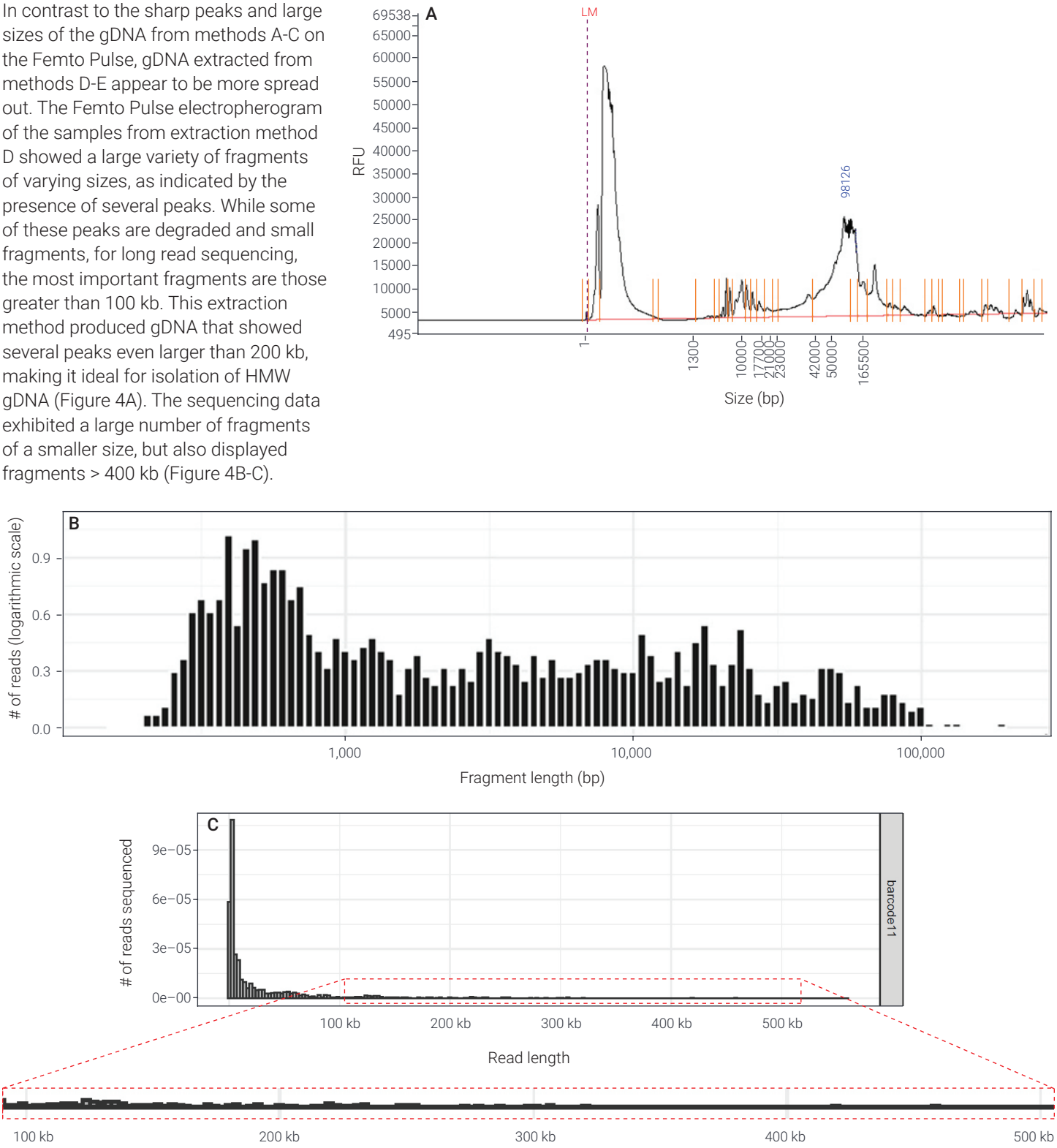

Figure 4. Bacterial gDNA isolated using extraction kit D was analyzed with the Agilent Femto Pulse system and sequenced with the MinION Flongle A) Femto Pulse electropherogram and digital gel image of the input gDNA B) Summary of sequencing read lengths in logarithmic scale C) Further analysis of the number of reads sequenced by read length indicates the presence of sequenced fragments through 500 kb.

Extraction kit E (column)

The gDNA from extraction method E, a column-based kit, appears as a rectangular shaped smear on the electropherogram generated by the Femto Pulse, instead of a single peak, with none of the sample extending beyond 200 kb (Figure 5A). The sequencing results show a large portion of the fragments above 10 kb, but very few of those extend above 200 kb (Figure 5B-C). This column-based extraction method was the only kit that did not produce ultrahigh molecular weight fragments above 200 kb.

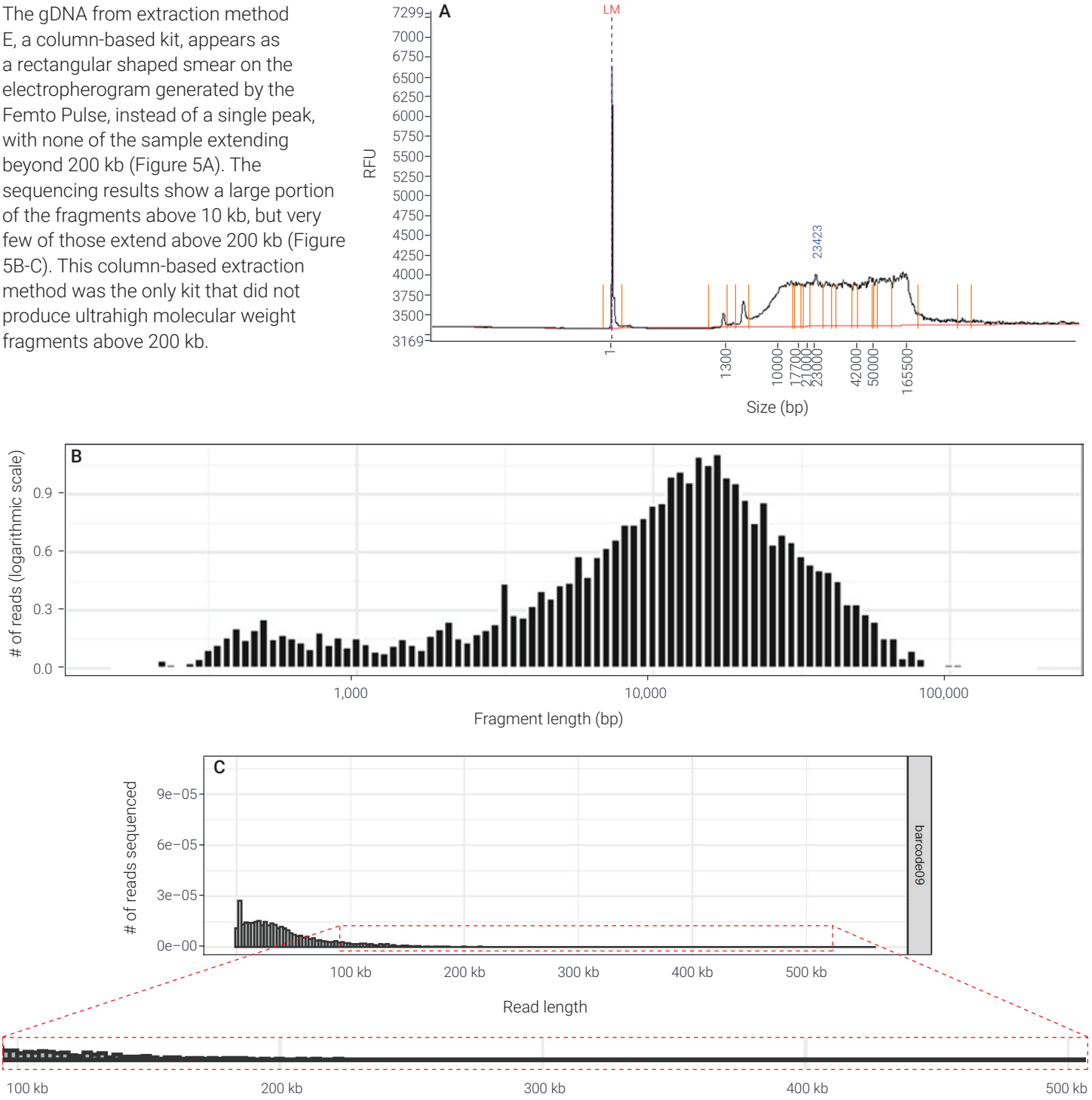

Figure 5. Bacterial gDNA isolated using extraction kit E was analyzed with the Agilent Femto Pulse system and sequenced with the MinION Flongle A) Femto Pulse electropherogram and digital gel image of the input gDNA B) Summary of sequencing read lengths in logarithmic scale C) Further analysis of the number of reads sequenced by read length indicates the presence of sequenced fragments through 200 kb.

Conclusion

In this application note, researchers from the NCCT compared the size and quality of HMW gDNA isolated from five different extraction methods. Multiple factors helped them determine the optimal kit to be used in the long-read sequencing workflow including sample throughput, length of reads needed, and amount of sequencing reads necessary for the planned experiment. For instance, a laboratory analyzing large numbers of samples at once would likely require a high-throughput, fast extraction method, such as the column-based kit E. For experiments requiring read lengths greater than 100 kb, the extraction method chosen must likewise be capable of extracting high or ultrahigh molecular weight DNA, such as with a magnetic-based (D) or gravity-based (A) kit. To help identify the kits that are capable of extracting HMW gDNA fragments greater than 100 kb, the Agilent Femto Pulse system was utilized as a part of the quality control workflow prior to sequencing, to determine the size of the isolated gDNA.

When preparing NGS libraries, the best practice to generate robust results is to routinely perform sample quality assessments prior to, and at certain points throughout the NGS workflow. This can represent a particular challenge when sequencing long nucleic acid fragments. The Femto Pulse system is capable of accurately analyzing fragments through 165 kb, offering the researcher the flexibility to perform sample quality checks recommended for optimal sequencing results.

References

1. Genomic DNA Sizing and Quality Control on the Agilent Femto Pulse System. *Agilent Technologies application note*, publication number 5994-0516EN, **2019**.
2. Agilent Genomic DNA 165 kb Kit Quick Guide for Femto Pulse System. *Agilent Technologies quick guide*, publication number SD-AT000141, **2020**.

www.agilent.com

PR7000-8335

Published in the USA, December 1, 2021
5994-4362EN

Agilent

Trusted Answers