diagenode
A Hologic Company

APPLICATION NOTE

Application of the Megaruptor® for shearing of ultra-long DNA fragments in the context of Long Read Sequencing using 3rd Gen Sequencing Technologies

Sylke Winkler (1), Sylvia Clausing (2), Nicola Gscheidel (1), Yannick Duport (1,2), Eugene Myers (1) and Andreas Dahl (2)
1) Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany
2) Deep Sequencing Group, BIOTEC, TU Dresden, Germany

Introduction

The 3rd generation sequencing platforms claim to provide long reads, high throughput, and less biased output by using single molecule sequencing approaches. One pre-requisite for the efficient utilization of the read length delivered by those systems such as the RSII from Pacific Biosciences® is the availability of sequencing libraries with ultra-long inserts. In Figure 1 the standard workflow for the library preparation process is shown:

Figure 1 – ***Schematic of the PacBio® library preparation process***.
Shearing of DNA is the initial step. After shearing, DNA is purified and quantified to start the library preparation process with a defined input.

As any amplification is avoided in the library preparation procedure for PacBio® sequencing the integrity and purity of the DNA is of crucial importance for the final quality of sequencing data. This is achieved by the quality of the provided DNA and the initially performed DNA shearing into large DNA fragments. We have evaluated the Megaruptor® system focussing on shearing performance, reproducibility, usability, and potential effects on the raw data quality.

Shearing performance

Long range sequencing highly depends on the ability of shearing DNA to rather large fragment sizes (20 – 40 kb). To target fragments of a specific size, the shearing quality and precision can be evaluated by the width of the sheared DNA smear in gel electrophoresis, the yield after shearing and purification of the DNA (1x SPRIBeads, Beckman Coulter). These factors allow an estimation of unintentionally created small fragments during the shearing process and the performance of the sample in the library preparation process.

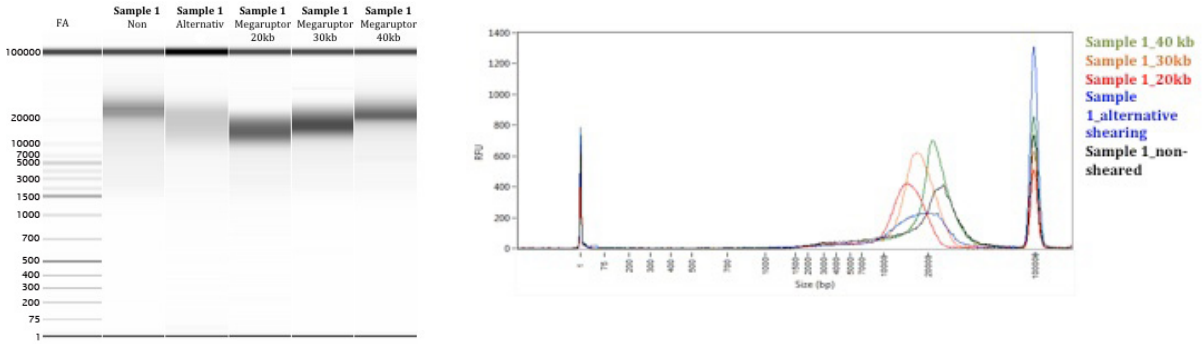


Figure 2
Gelelectrophoretic analysis after shearing of one sample with identical input (10 µg genomic DNA) with one alternative method and three different conditions using the Megaruptor®.

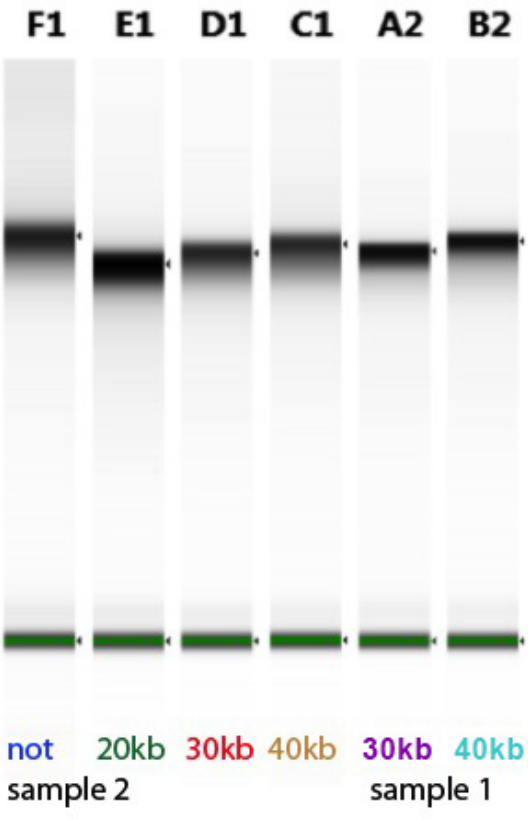
www.diagenode.com Please contact us for more information
Europe Diagenode sa / LIEGE SCIENCE PARK // Rue du Bois Saint-Jean, 3 // 4102 Seraing (Ougrée) // Belgium // Phone: (+32) 4 364 20 50 // E-mail: info@diagenode.com
USA Diagenode Inc. / 400 Morris Avenue, Suite 101 // Denville, NJ 07834 // USA // Phone: (+1) 862 209-4680 // Mail: info.na@diagenode.com

As shown in Figure 2, gDNA sheared with the Megaruptor® shows a discrete and narrow size distribution. The observed differences between the parameter settings for the shearing (20, 30 or 40 kb) and the electrophoretically determined sizes result from the limitation of the capillary electrophoresis system to separate and size large DNA fragments accurately. In our example, the unsheared genomic DNA runs also at a rather small size. However, the obtained fragment sizes correspond nicely with the different parameter settings. Compared to our initially used shearing method, the Megaruptor® shows a more dense size distribution of sheared fragments, and a higher and more reproducible yield (the quantification was done before and after shearing and a subsequent purification by 1x SpriBeads). In summary, the overall yield of the complete PacBio® library preparation process could be increased, which is supposed to be the result of higher content of large DNA fragments. However, based on our experience the initial quality of the provided DNA is the most important parameter for the final yield in the preparation.

Reproducibility

To test the reproducibility of the Megaruptor®, we sheared DNA samples of different origin with variable sizes (Figure 3A) as well as identical size parameters (Figure 3B). We observed a very high reproducibility for all applied parameters that was independent from the input sample.

Sample 1 – Sample 2



File	Indicator	Well ID	Comment
File 2		F1	non-sheared sample 2
File 2		E1	20kb sample 2
File 2		D1	30kb sample 2
File 2		C1	40kb sample 2
File 2		A2	30kb sample 1
File 2		B2	40kb sample 1

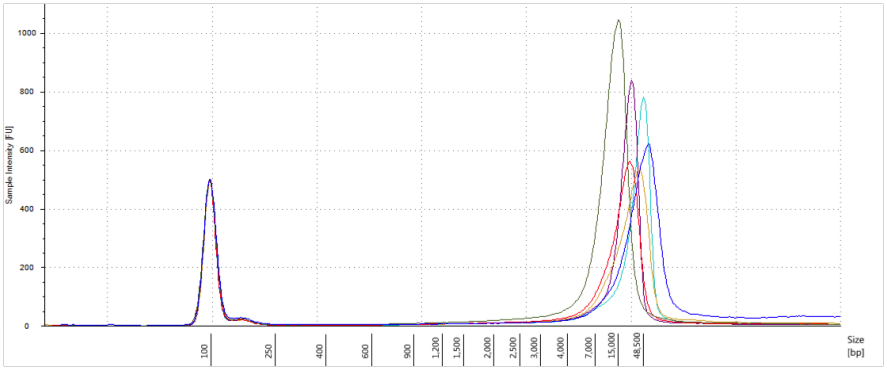
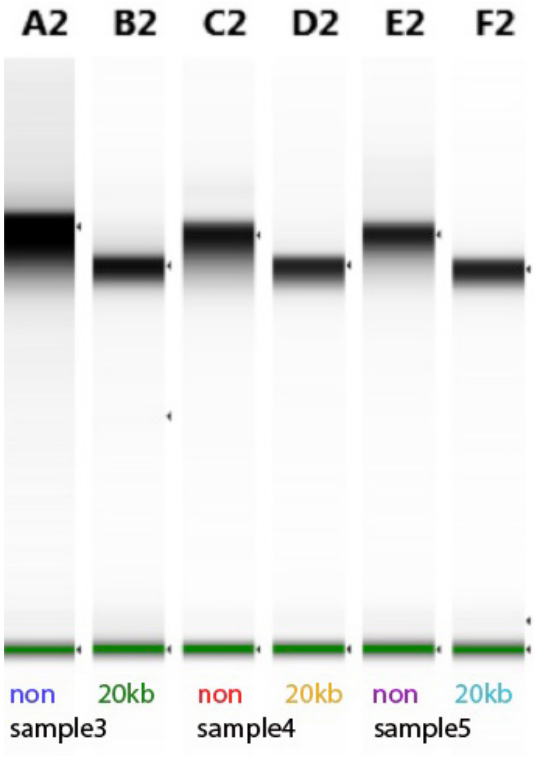


Figure 3A
Electrophoretic Analysis (Tape Station, Agilent) before (F1) and after shearing for 20, 30 and 40 kb for two different DNA samples

Sample 3 – Sample 5



File	Indicator	Well ID	Comment
File 5		A2	Sample 3_non-sheared
File 5		B2	Sample 3_20kb sheared
File 5		C2	Sample 4_non-sheared
File 5		D2	Sample 4_20kb sheared
File 5		E2	Sample 5_non-sheared
File 5		F2	Sample 5_20kb sheared

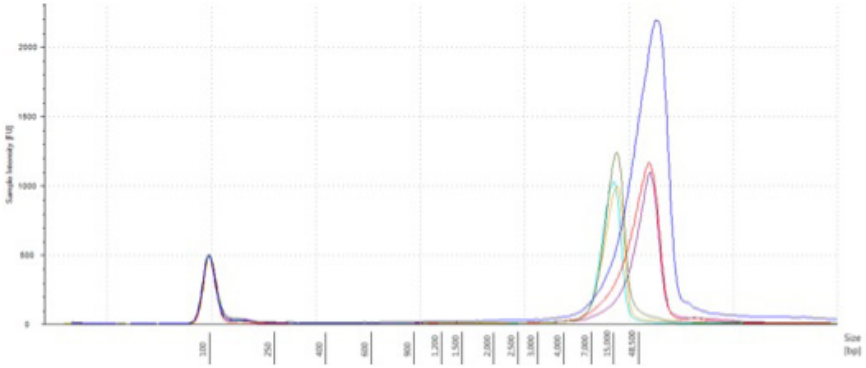


Figure 3B

Electrophoretic Analysis (Tape Station, Agilent) before (A2, C2, E2) and after shearing for 20 kb for three different DNA samples.

Usability

In our hands, the usage of the Megaruptor® is easy and straight-forward. The system provides a walk away solution for shearing of one or two samples. The software is intuitive and due to the small foot print, the instrument can be easily integrated in the lab.

Sequencing performance and Data quality

Apart from shearing robustness and yield after library preparation, the most relevant criteria of the shearing process is performance of given library made from ultra-long DNA fragments in actual single molecule sequencing on the PacBio® RS II. The example presented in Figure 4 is based on a given genomic DNA sample that has been sheared to 40 kb DNA fragments using the Megaruptor®. We managed to achieve an average read length of inserts of 8.14 kb (33 % P1 loading) with an average accuracy of 0.85. The longest insert reads in this sequencing runs are larger than 30 kb in length.

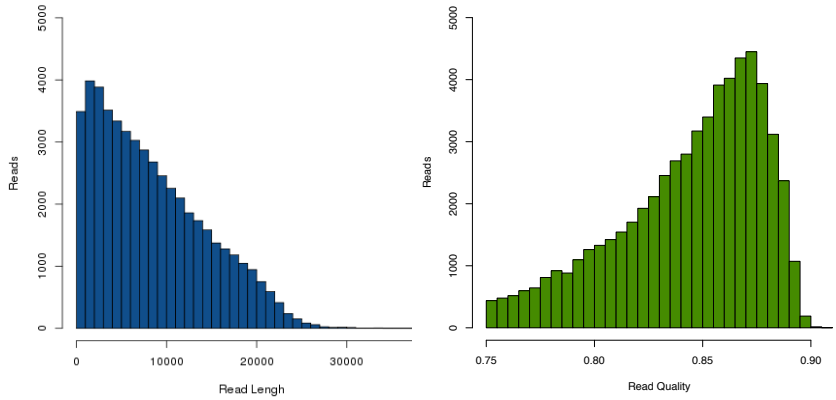


Figure 4
Summary of insert read length in bp and read quality for a library sheared to ultra-long DNA fragments (40 kb) on the Megaruptor® making use of the P4/C2 PacBio® sequencing chemistry, 180 min run).

In respect of data quality on the RS II platform, the Megaruptor® processed samples behaved in the same way as samples that were sheared with alternative technologies. Observed variation in the quality distribution of the obtained sequencing data (see Figure 5) is rather due to variation between different sequencing experiments and conditions than due to the shearing method applied.

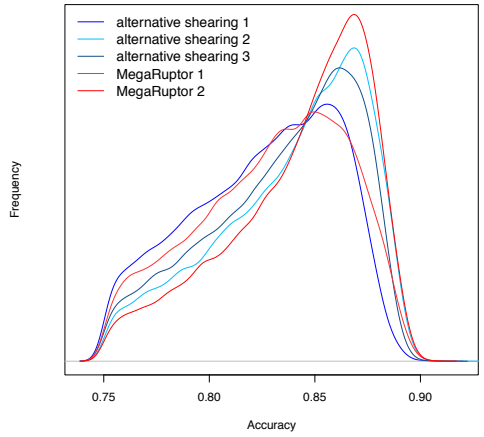


Figure 5
Quality distribution from different PacBio® RSII sequencing runs of different libraries prepared with either Megaruptor® based shearing or an alternative shearing method.

Summary

The Megaruptor® provides a simple, robust, and straight forward method for reproducible DNA shearing to obtain ultra-large DNA fragments. The system has been easily integrated into our workflow for long range sequencing of single molecules using the RSII platform from Pacific Biosciences®. The obtained dense distribution of fragment sizes after shearing and the low content of non-sheared as well as short fragments was beneficial for the performance of the whole library preparation process.