

# Whether the addition of nanodroplets increases the efficiency of DNA fragmentation in different sonication equipments

Haofan Zheng

Advisor: Dr. Haaland, Dr. Marron

Client: Dr. Pattenden, Dr. Mehrab-Mohseni.

March 26, 2021

## **Abstract**

Nanodroplets are an cavitation agent for use in DNA fragmentation, and our purpose is to prove that the addition of decafluorobutane(DFB) nanodroplets when sonicating can enhances efficiency and consistency of fragmentation. After applying the analysis of variance in models from three experiments, DNA fragmentation is not effective in samples without any treatment during sonication. Besides, DNA fragmentation efficiency will increase when processing samples with the addition of nanodroplets, compared with samples without the addition. Futhermore, the efficiency and consistency of fragmentation in samples with the addition can meet the industry standard(samples processed by microtube) in one experiment roughly. Hence this treatment may be applied in industry as an alternative economical way for extracting DNA fragments in concentration.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Experimental Design</b>	<b>3</b>
<b>3</b>	<b>Data collection</b>	<b>4</b>
<b>4</b>	<b>Method</b>	<b>5</b>
<b>5</b>	<b>Results</b>	<b>5</b>
5.1	Excutive Summary . . . . .	5
5.2	Experiment 1(LE220) . . . . .	6
5.3	Experiment 2(LE110) . . . . .	9
5.4	Experiment 3(Qsonica) . . . . .	12
<b>6</b>	<b>Discussion</b>	<b>16</b>
	<b>References</b>	<b>17</b>
<b>7</b>	<b>Appendices: Method</b>	<b>18</b>
7.1	One-way Anova . . . . .	18
7.2	Two-way Anova . . . . .	18
7.3	Logistic regression + Two-way Anova . . . . .	19
7.4	Interaction Term . . . . .	19
<b>8</b>	<b>Appendices: Additional plots and explanation</b>	<b>20</b>
8.1	Confidence Intervals Adjustment in Experiment 1(LE220), 2(LE110) . . . . .	20
8.2	Outliers Analysis in Experiment 3(Q sonica) . . . . .	20
8.3	Further Analysis Result in Experiment 3 . . . . .	22
<b>9</b>	<b>Appendices: R output</b>	<b>25</b>
9.1	Experiment 1(LE220) . . . . .	25

9.2	Experiment 2(LE110) . . . . .	26
9.3	Experiment 3(Q sonica) . . . . .	26

# 1. Introduction

DNA fragmentation is essential in study on cellular processes, and more concentrated fragments in certain range from 150 base pairs(bp) to 300 bp can reduce bias in result[1]. There are many methods for DNA fragmentation including enzymatic digestion, microwave irradiation, and sonication[2, 3, 4]. In spite of sonication as an important step during DNA fragmentation, it's not efficient and consistent to produce fragments in range[5]. Based on this problem, one research shows that a perfluorocarbon nanodroplet formulation is an effective cavitation enhancement agent, enabling rapid and consistent fragmentation of genomic DNA in a standard ultrasonic water bath[6]. According to this, our clients invent a new treatment as the addition of the decafluorobutane nanodroplet, so this experimental data analysis aims to prove that the addition of DFB nanodroplets is an effective treatment during sonication.

Based on the results from three experiments, it's shows that the efficiency and consistency can be improved by the addition of DFB nanodroplets, and this treatment during sonication can be an alternative approach to produce more fragments in range, compared to samples without the addition as no effectiveness in fragmentation. However, DNA fragmentation in samples with the addition of nanodroplets meet the industry standard in one experiment(Experiment 1) but not in another experiment(Experiment 2), so more samples are required for futher data analysis to obtain a more detailed and comprehensive conclusion.

# 2. Experimental Design

This study involoes three different experiments, whose design is indicated in Table 2.1.

Experiment	Sonicator	Groups		
Experiment 1	LE220	DFB plus	DFB minus	Covaris microtube
Experiment 2	LE110	DFB plus	DFB minus	Covaris microtube
	Qsonica	DFB plus		DFB minus
Experiment 3	Qsonica	DFB plus		DFB minus
	Qsonica	DFB plus		DFB minus

Table 2.1: Experimental design for different sonicators.

*DFB plus group: with the addition of nanodroplets, DFB minus group: without nanodroplets, Covaris microtube group : processed by microtube as the industry standard.*

Shown in the Table 2.1, there is no replication of experiment 1 and 2, but the researchers conducted experiment 3 three times. According to this experimental design, we can compare the outcome among

different groups in three experiments and do further analysis.

### 3. Data collection

The researchers recorded the sample sonication time points and the belonging group when measuring the its fragment size during experiment 1 and 2. In experiment 3, the row and column labels of each sample in the plate and its belonging group are recorded, and the fragment size is measured at the same time after the sonication in two groups. Here are variables and responses in three experiments:

- Experiment 1(LE220):

Full Predictors(variable)

- Time (second)
- Group

Response(outcome)

- Fragment size (bp)

- Experiment 2(LE110):

Predictors(variable)

- Time (second)
- Group

Response(outcome)

- Fragment size (bp)

- Experiment 3(Qsonica):

Predictors(variable)

- Row label
- Column label
- Group

Response(outcome)

- Fragment size (bp)

Shown in the list above, the primary response is the DNA fragment size(bp) in three experiments. For further logistic regression application, the **new** outcome is defined as an indicator(0/1), pointing out whether the fragment size is in our target region. The new defined response is **1** if the fragment size is in range(150bp-300bp), and **0** if its size is out of the range.

## 4. Method

See details in Appendices section, please.

## 5. Results

### 5.1 Excutive Summary

Here is the brief summary of the results from three experiments:

- Experiment 1: Fragmentation in DFB minus group is not effective. It's hard to prove that the DNA fragmentation in which group is better, the DFB plus group or Covaris microtube group, but easy if we do the comparison at a specific sonication time. The fragamentation seems to meet the industry standard(Covaris microtube group) at the time 240 sec, see in experiment 1 result.
- Experiment 2: Fragmentation in DFB minus group is not effective. It's hard to prove that the DNA fragmentation in which group is better, the DFB plus group or Covaris microtube group, but easy if we do the comparison at a specific sonication time, see in experiment 2 result.
- Experiment 3: Fragmentation in DFB plus group has significantly higher efficiency tham DFB minus group, convincingly this time. In addition, DNA fragmentation efficiency in DFB plus group is guaranteed to be full, see in experiment 3 result.

## 5.2 Experiment 1(LE220)

There are three groups – DFB minus(N=72)<sup>1</sup>, DFB plus(N=54) and Covaris microtubes(N=54) groups in experiment 1. The sample size at different time point is equal in the same group. The sonication time points in each group is displayed in Table 5.1.

Experiment	Group	Time point
Experiment 1	DFB minus	720 sec, 840 sec, 1008 sec, 1209 sec
	DFB plus	80 sec, 120 sec, 240 sec
	Covaris microtube	80 sec, 120 sec, 240 sec

Table 5.1: Sonication time table for three troupes in experiment 1.

We take a look in DFB minus group at first. Figure 5.1 shows the distribution of fragment size at the given time points along the vertical line, each black dot represents fragment size in one well. The area between the two red lines is the target region(150bp-300bp). Similiar as Figure 5.1, Figure 5.2 shows the distribution of fragment size in DFB plus and Covaris microtube group.

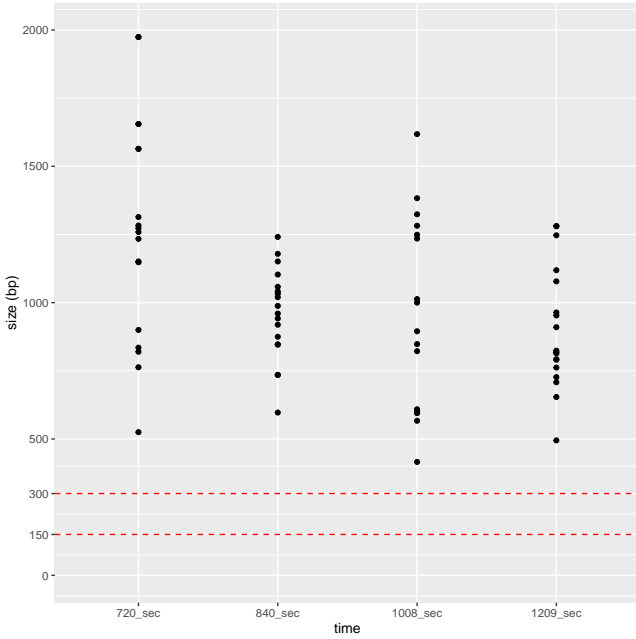


Figure 5.1: Distribution of fragment size over-time in DFB minus group.

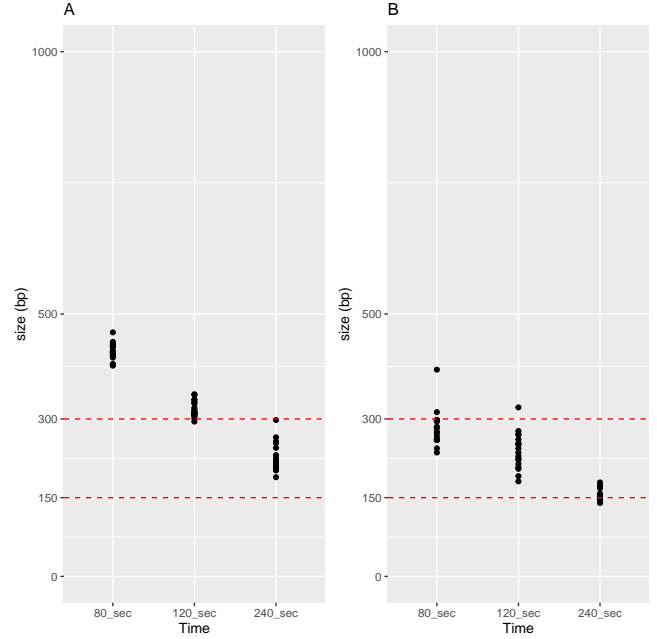


Figure 5.2: Distribution of fragment size over-time in DFB plus group(Panel A) and Covaris microtube group(Panel B).

As shown in Figure 5.1, there are no points in our target area, implying that DNA fragmentation in DFB minus group is not effective. However, from Figure 5.2, we can see that DNA fragments are more concentrated and fall into the target region overtime.

After doing exploratory analysis as previous plots, one-way anova is applied to DFB minus data, Figure 5.3 shows the result after fitting the model. In this Figure, each point is the average of fragment size at the corresponded time point, together with 95% confidence interval.

<sup>1</sup>N: sample size.



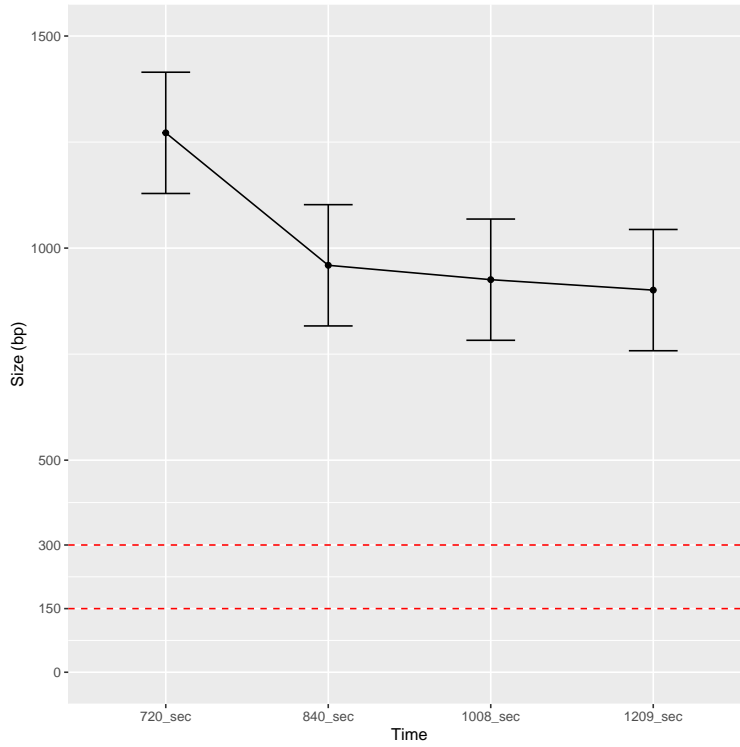


Figure 5.3: DNA fragment average size in DFB minus group over time.

*Shown are averages of fragment size at sonication time 720 sec, 840 sec, 1008 sec and 1208 sec with 95% confident interval. The area between the two horizontal red lines is the target region (size: 150bp - 300bp).*

In Figure 5.3, no point falls into the target region and there is no overlapping area between the confidence intervals and our target region, though the fragment size average is decreasing with time increasing (p value < 0.05). Therefore, the fragmentation of DNA in DFB minus group is not efficient.

Then as for DFB plus group and Covaris microtube group, we applied two-way anova and logistic regression to them for comparison. Figure 5.4 shows the average of fragment size (Panel A) and the probability of fragment size in our target region (Panel B) overtime in two groups. In panel A, each point is the average fragment size at the time point in the belonging group, together with 95% confidence interval<sup>2</sup>. Similar as panel A, panel B shows the the probability of fragment size in our target region with 95% confidence interval at given sonication time in two groups.

From the Figure 5.4 panel A, we can see that fragment size average is always lower in Covaris microtube group than DFB plus group. Based on the R output, the difference is significant between different time points (p value < 0.05), groups (p value < 0.05), and the interaction terms between time and group (p value < 0.05)<sup>3</sup>. In Panel B, with R output, the difference is also significant between different time points (p value < 0.05), groups (p value < 0.05), and the interaction terms between time and group (p value < 0.05). The efficiency and consistency of Covaris microtube group is better, since the average fragment sizes are always in target region and probability of size in range always is always larger than 0.5 at three time points. However, at time 240 sec, the fragment size is in our target region with probability 1 in DFB plus group but less than 0.75 in Covaris microtube group. It seems that efficiency increases and decreases overtime in Covaris microtube group but increases all the time in the DFB plus group, guaranteed adequate efficiency at the end (time = 240 sec) in DFB plus group.

<sup>2</sup>There is an adjustment in calculating the CIs, see in appendices.

<sup>3</sup>Interaction term: see in the appendices, please.

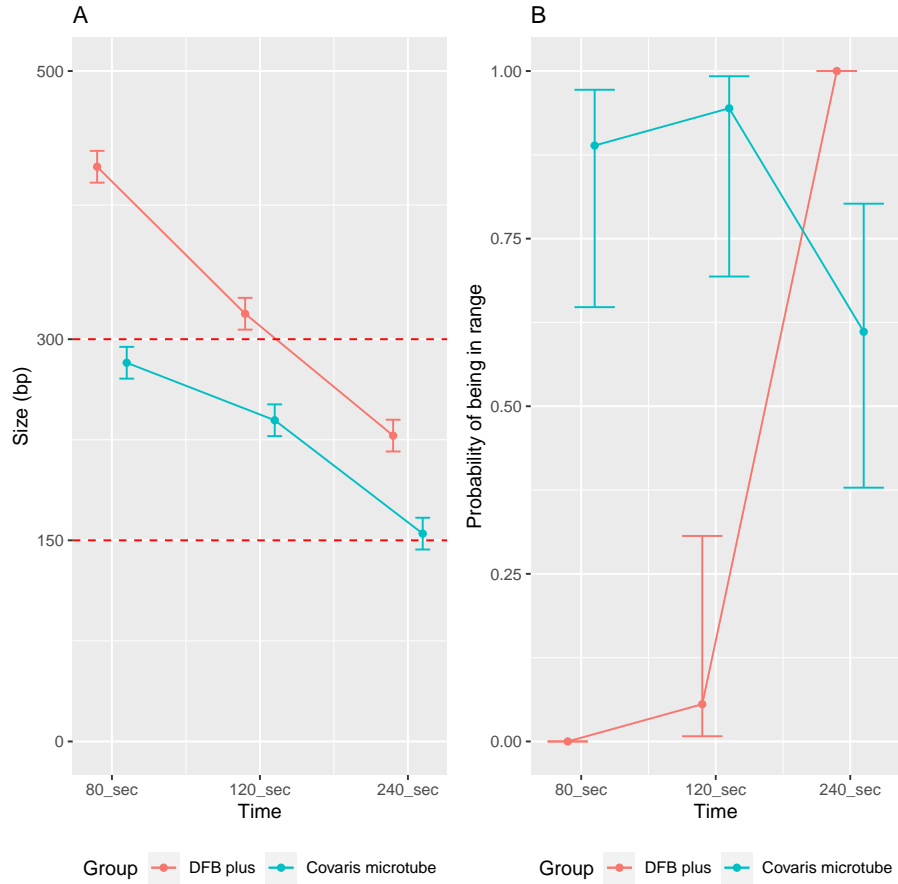


Figure 5.4: Comparison on Fragment average size and probability of being in the target region between DFB plus group and Covaris Microtube group overtime.

*Shown are the averages size with 95% confidence interval (Panel A) and probability of falling in the target region(Panel B) at sonication time 80 sec, 120 sec, 240 sec in DFB plus group(red) and Covaris microtube group(cyan).*

## 5.3 Experiment 2(LE110)

There are three groups as similar as experiment 1(LE110)– DFB minus(N=30), DFB plus(N=30), and Covaris microtubes(N=30) groups in experiment 1. The sample size at different time point is equal in the same group. The sonication time points in each group is displayed in Table 5.2.

Experiment	Group	Time point
Experiment 2	DFB minus	60 sec, 90 sec, 120 sec, 210 sec, 240 sec
	DFB plus	15 sec, 30 sec, 90 sec, 120 sec, 210 sec
	Covaris microtube	15 sec, 30 sec, 90 sec, 120 sec, 210 sec

Table 5.2: Sonication time table for three groups in experiment 2.

The analysis is the same as we do in experiment 1, so we take a look in DFB minus group at first. Figure 5.5 shows scatter plot and box plot on the distribution of fragment size at the given time points along the vertical line, each black dot represents fragment size in one well. The area between the two red lines is the target region(150bp-300bp), while we can find some outliers in the box plot.<sup>4</sup> Similar as Figure 5.5, Figure 5.6 shows the distribution of fragment size in DFB plus and Covaris microtube group.

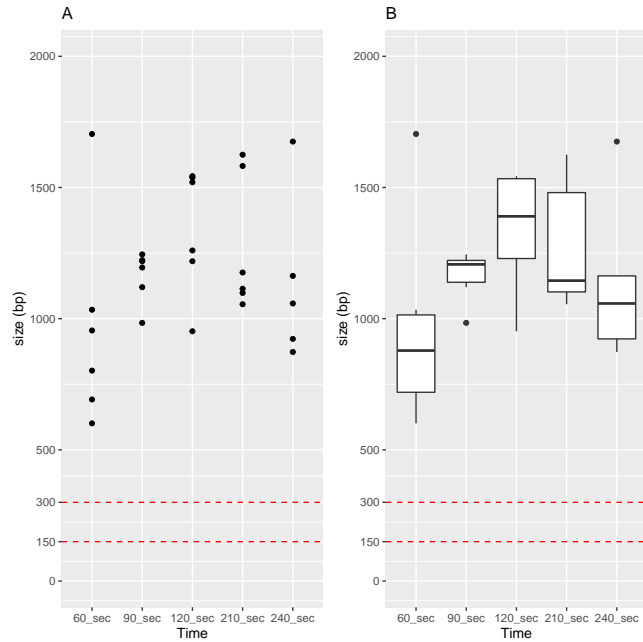


Figure 5.5: Scatter plot(Panel A) and box plot(Panel B) on distribution of fragment size overtime in DFB minus group.

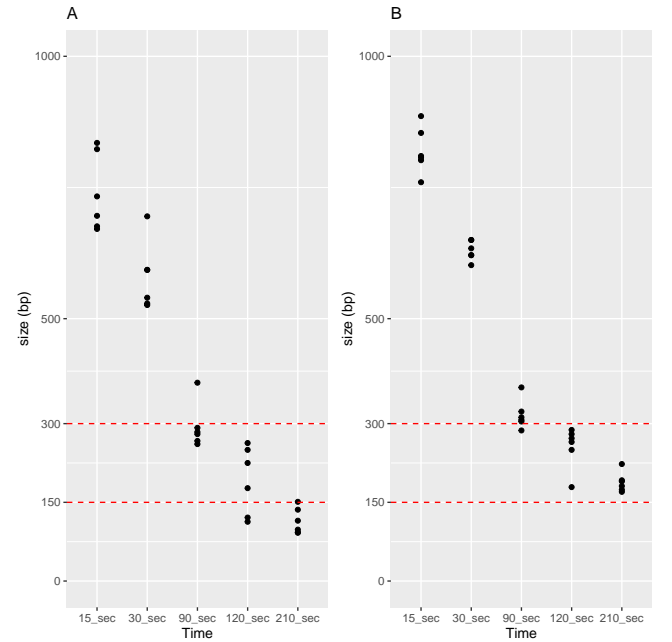


Figure 5.6: Distribution of fragment size overtime in DFB plus group(Panel A) and Covaris microtube group(Panel B).

As shown in Figure 5.5, there are no points in our target area, implying that DNA fragmentation in DFB minus group is not effective. However, from Figure 5.6, we can see that DNA fragments are more concentrated and fall into the target region overtime.

After doing exploratory analysis as previous plots, quantile regression is applied to DFB minus data, since there is an outlier at time Figure 5.3 shows the result after fitting the model. In this Figure, each point is the average of fragment size at the corresponded time point, together with 95% confidence interval.

<sup>4</sup>There is one extreme value in the Figure 5.5 Panel A and B not shown, because the value in fragment size of this sample is 4823, larger than 2000.

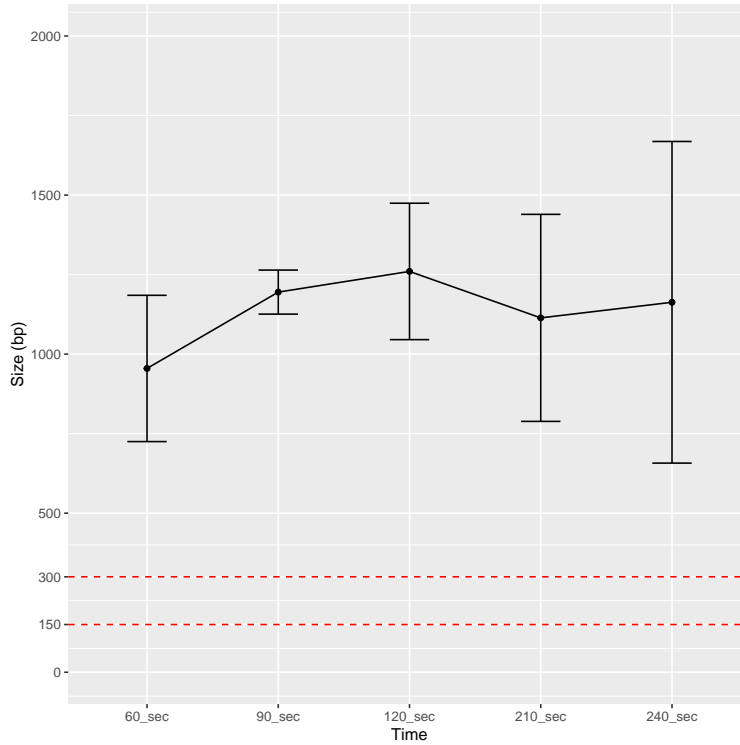


Figure 5.7: DNA fragment median size in DFB minus group over time, including extreme value in the model.

*Shown are median sizes of fragment size at sonication time 60 sec, 90 sec, 120 sec, 210 sec and 240 sec with 95% confident interval. The area between the two horizontal red lines is the target region (size: 150bp - 300bp).*

Since there is an extreme value as the outlier at 240 sec, quantile regression as an outlier resistant model is applied for analysis. In Figure 5.7, no point falls into the target region and there is no overlapping area between the confidence intervals and our target region, and the fragment size average is almost the same with time increasing ( $p$  value = 0.446), different from experiment 1. Therefore, the fragmentation of DNA in DFB minus group is not efficient.

Then as for DFB plus group and Covaris microtube group, we applied two-way anova and logistic regression to them for comparison. Figure 5.8 shows the average of fragment size (Panel A) and the probability of fragment size in our target region (Panel B) overtime in two groups. In panel A, each point is the average fragment size at the time point in the belonging group, together with 95% confidence interval. Similar as panel A, panel B shows the the probability of fragment size in our target region with 95% confidence interval at given sonication time in two groups.

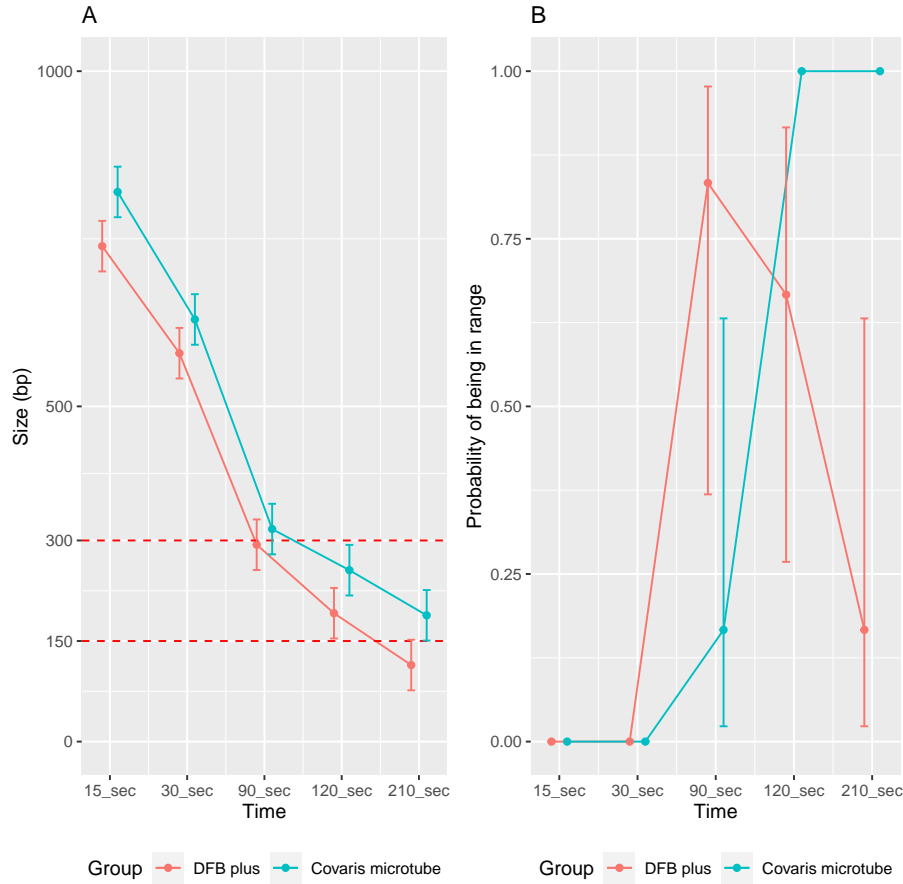


Figure 5.8: Comparison on fragment average size and probability of being in the target region between DFB plus group and Covaris Microtube group overtime.

*Shown are the averages size with 95% confidence interval (Panel A) and probability of falling in the target region with 95% confidence interval (Panel B) at sonication time 15 sec, 30 sec, 90 sec, 120 sec, 210 sec in DFB plus group(red) and Covaris microtube group(cyan).*

From the Figure 5.8 panel A, we can see that fragment size average is always higher in Covaris microtube group than DFB plus group, different from experiment 1. Based on the R output, the difference is significant between different time points ( $p$  value  $< 0.05$ ), groups ( $p$  value  $< 0.05$ ), but not significant between different interaction terms ( $p$  value  $= 0.571$ ). In Panel B, with R output, the difference is also significant between different time points ( $p$  value  $< 0.05$ ), the interaction terms ( $p$  value  $< 0.05$ ), but not significant between groups ( $p$  value  $< 0.168$ ). There are several difference result between experiment 1 and experiment 2. The efficiency and consistency of two groups are similar. However, at time 120 sec and 240 sec, the fragment size is in our target region with probability 1 in Covaris microtube group. It seems that efficiency increases and decreases overtime in DFB plus group but increases all the time in the Covaris microtube group, and guaranteed adequate efficiency at the end(time = 210 sec), different from the experiment 1 result.

## 5.4 Experiment 3(Qsonica)

Different from the previous two experiments, there are only two groups in experiment 3 – DFB minus(N=288) and DFB plus(N=288) groups with three replications, not including sonication time points. In this experiment, there are 192 samples for each replication, 96 in the DFB minus group and 96 in the DFB plus group. The plate with 96 wells is shown below:

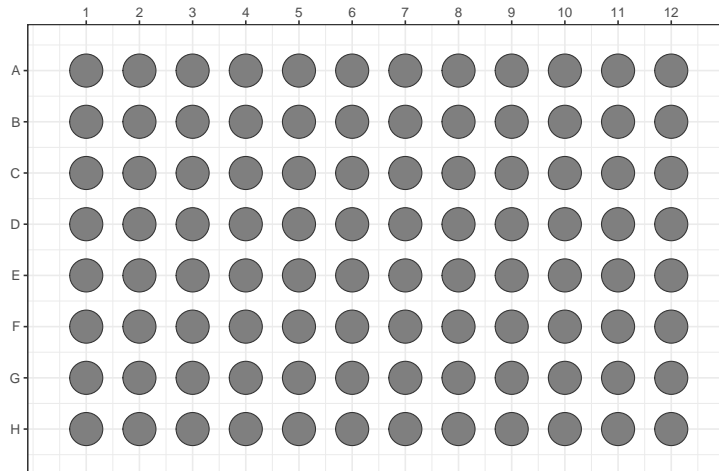


Figure 5.9: 96-well Microplate.

*Shown is the plate in the experiment 3. Row labels: A-H. Column labels: 1-12.*

We take a glare at the data at first. Figure 5.10 shows the distribution of fragment size in DFB minus group and DFB plus group of three replications in experiment 3. For clarification, the three replication are named replication 1, replication 2 and replication 3. There are two outliers in DFB minus group of replication 1 and one outlier in DFB minus group of replication 2, not shown in the figure.<sup>5</sup>

---

<sup>5</sup>In DFB minus group of replication 1, the value of two outlier are 8779 and 9617. In DFB minus group of replication 2, the value of the outlier is 60871.

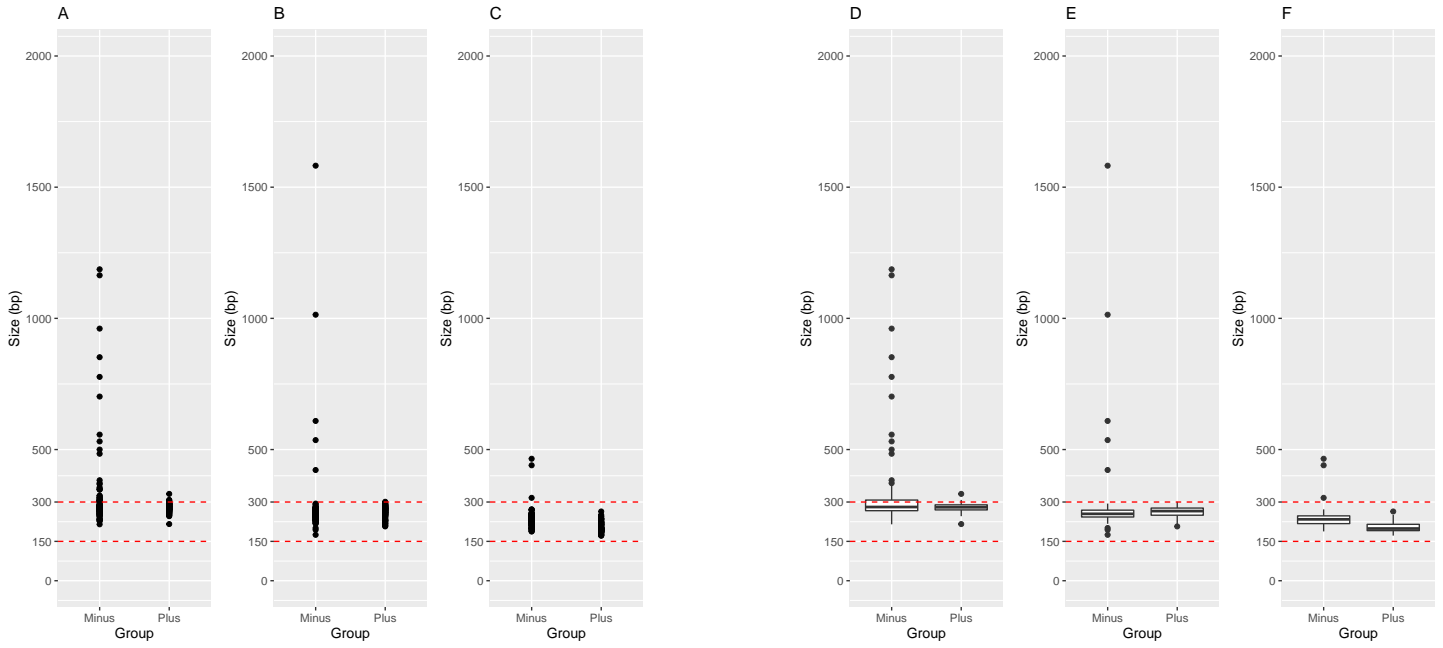


Figure 5.10: Scatter plot(Panel A-C) and box plot(Panel D-F) on distribution of fragment size in DFB minus group and DFB plus group of three replications.

From the Figure 5.10, the fragment sizes in DFB plus group seems lower than in DFB plus group and more concentrated in DFB plus group. In addition, almost all samples in DFB plus group fall into the target region but some samples in DFB minus group fall outside.

After doing the exploratory analysis, the beginning problem is how to deal with the three outliers. Based on the diagnostic plot and analysis<sup>6</sup>, in the model that the response is fragment size, the three outliers will be eliminated from the samples(two in replication 1 and one in replication 2). However, in the model that the response is the new defined one(0/1), the outliers are included, since the response are rescaled and they are not outliers anytime. Then we check if there is difference on fragment size or probability of being in range among row and columns, displayed in Figure 5.11. In Panel A and Panel B, each point is the average of fragment size in one row or column of the plate in all three replications. In Panel C and Panel D, each point is probability of being in target region in one row or column of the plate in all three replications.

---

<sup>6</sup>See details in appendices, please

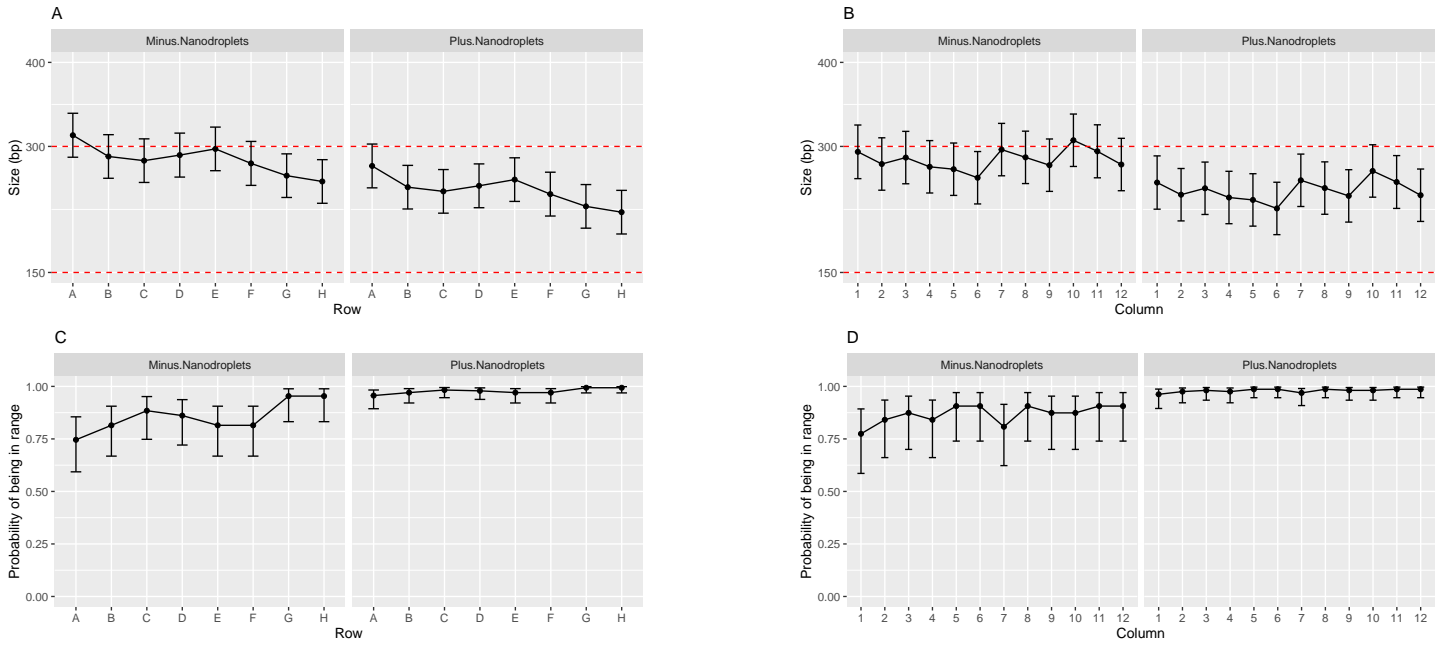


Figure 5.11: Fragment average size and probability of being in target region in each row and column of the plate between DFB plus and DFB minus group.

Shown are average size with 95% confident interval in each row(Panel A) and column(Panel B) of the plate, probability of being in target region with 95% confident interval in each row(Panel C) and column(Panel D) of the plate.

From Figure 5.11 and R output, there is a significant difference on fragment size and probability of being in range between two groups( $p$  value  $< 0.05$ ). However among different rows and columns of the plate, the difference seems not significant(row: $p$  value=0.06, column: $p$  value=0.8). Since in this comparison, the difference of response between two groups is considered, so the model is included samples from both group. However, if one model is built for each group and do further investigation on difference among rows and columns in separate two groups, the result may be slightly different, especially for row label.<sup>7</sup>

Then as for DFB plus group and DFB minus group, we applied two-way anova(outliers excluded) and logistic regression(outliers included) to them for comparison. Figure 5.12 shows the average of fragment size(Panel A)and the probability of fragment size in our target region(Panel B) overtime in two groups. In panel A, each point is the average fragment size in each well of the plate in the belonging group. Similiar as panel A, panel B shows the the probability of fragment size in our target region in each well of the plate in two groups. Thus we have 96 points in DFB minus group and 96 in DFB plus group.

<sup>7</sup>See the appendices, for details, please.



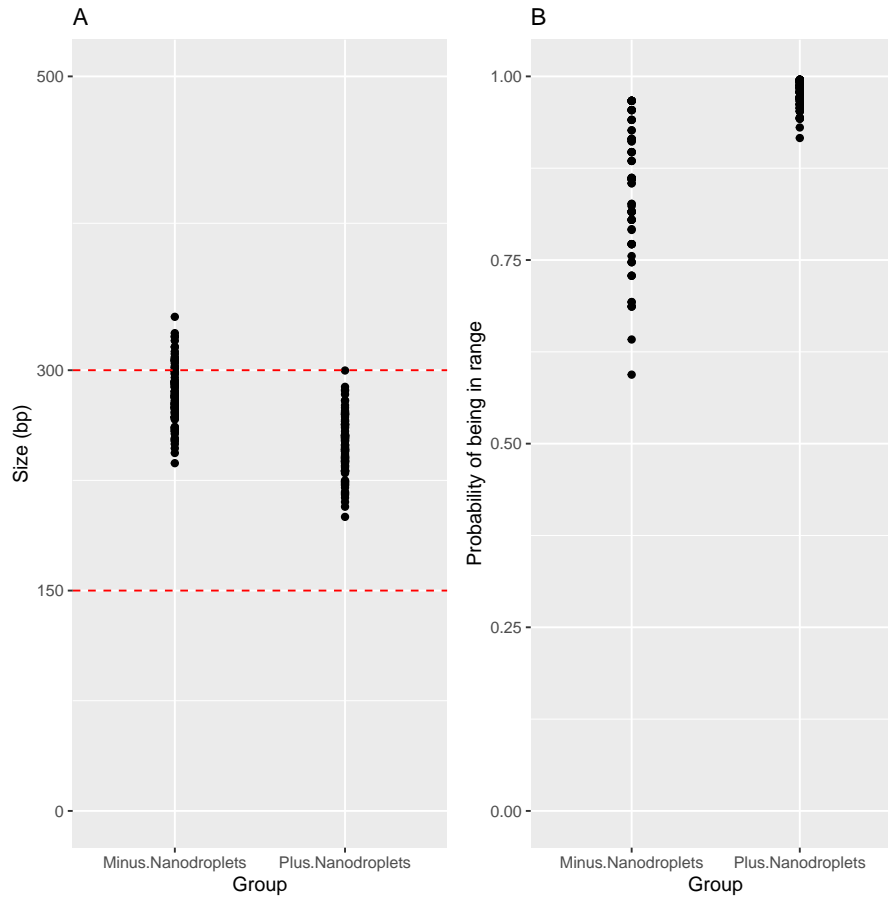


Figure 5.12: Comparison on fragment average size and probability of being in the target region between DFB plus group and DFB minus group.

*Shown are the averages size (Panel A) and probability of falling in the target region (Panel B) in each well of the plate.*

From the Figure 5.12 Panel A, the fragment sizes are lower in DFB plus group than DFB minus group, and all the points are in the range in DFB plus group, while some points in DFB minus group are outside. Based on R output, the difference between two groups show significance( $p$  value  $< 0.05$ ). In Panel B, the probability for all samples in DFB plus group is closed to 1 and concentrated, but the probability for some samples in DFB minus group are below 0.75 and scattered.

According to the model, the prediction of response in each well of the plate can be calculated. Figure 5.13 shows the prediction on whether fragment size is in target region and prediction on probability of being in the target region in each well of the plate.

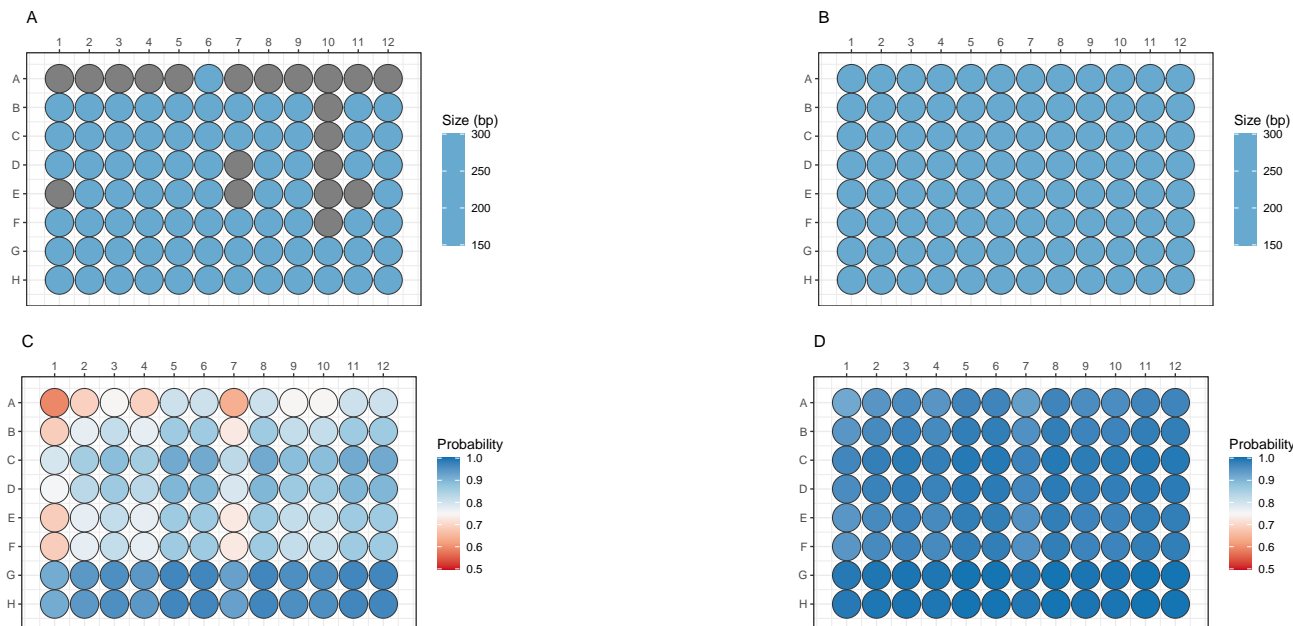


Figure 5.13: Prediction on whether fragment size is in target region and heatmap on probability of being in the target region in each well of the plate.

Shown are size prediction: in target region (blue) or outside (grey) in DFB minus (Panel A) and DFB plus group (Panel B). Probability of being in the target region: high (blue) or low (red) in DFB minus group (Panel C) and DFB plus group (Panel D).

From the Figure 5.13 Panel A and Panel B, there are several grey circles in DFB minus group but all blue in DFB plus group. And in Panel C and Panel D, there are same red or white circles in DFB minus group but almost all dark blue in DFB plus group. Then we can see that if we do the experiment again in future, all fragment size in DFB plus group will be in the target region. However in DFB minus, there will be some samples fall outside the target region. Therefore the DNA fragmentation in DFB plus group shows strongly better efficiency compared with the DFB minus group, guaranteed full efficiency.

## 6. Discussion

Just some thoughts for discussion. The clients are willing to compare the distribution (density) between the two groups and show the concentration on density, but more samples are needed. In addition, if the clients want to know if the DFB plus group meet the industry standard (Covaris microtube group), the specific sonication time and sonicator selection (LE220 or LE110) should be pointed out, because we can't compare them roughly overtime. Moreover, I wonder if we measure the fragment size every 5s or 10s in two sonicators, the specific time for DFB plus group with full fragmentation efficiency may be selected<sup>1</sup>,

<sup>1</sup>For example, in the time period between 90 sec and 120 sec, probability of falling in range may be 1 at 95 sec or 100 sec in DFB plus group, see Figure 5.8 Panel B.

different in two sonicators, but I'm not sure if the sonication time point can be set manually.

# References

- [1] Clifford Meyer and Shirley Liu. Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nature reviews. Genetics*, 15, 09 2014.
- [2] Norio Tanaka, Akihisa Takahara, Taichi Hagio, Rika Nishiko, Junko Kanayama, Osamu Gotoh, and Seiichi Mori. Sequencing artifacts derived from a library preparation method using enzymatic fragmentation. *PLOS ONE*, 15:e0227427, 01 2020.
- [3] Yu Yang and Jun Hang. Fragmentation of genomic dna using microwave irradiation. *Journal of biomolecular techniques : JBT*, 24:98–103, 07 2013.
- [4] Nikolay Pchelintsev, Peter Adams, and David Nelson. Critical parameters for efficient sonication and improved chromatin immunoprecipitation of high molecular weight proteins. *PloS one*, 11:e0148023, 01 2016.
- [5] Leonid Teytelman, Bilge Ozaydin, Oliver Zill, Philippe Lefrancois, Michael Snyder, Jasper Rine, and Michael Eisen. Impact of chromatin structures on dna processing for genomic analyses. *PloS one*, 4:e6700, 02 2009.
- [6] Anna Chiarella, Austin Quimby, Marjan Mehrab-Mohseni, Brian Velasco, Sandeep Kasoji, Ian Davis, Paul Dayton, Nathaniel Hathaway, and Samantha Pattenden. Cavitation enhancement increases the efficiency and consistency of chromatin fragmentation from fixed cells for downstream quantitative applications. *Biochemistry*, 57, 04 2018.

## 7. Appendices: Method

### 7.1 One-way Anova

Model:

$$y_{i,j} = \mu_i + \epsilon_{i,j}$$

$I$ : the number of groups,  $J$ : the number of samples in each group,  $n$ : the number of samples. ( $n = I * J$ )

$i = 1, 2, 3, 4$ , corresponded with 720 seconds, 840 seconds, 1008 seconds, 1209 seconds

$\mu_1, \mu_2, \mu_3, \mu_4$ , corresponded with the mean size of each group

$$\epsilon_{i,j} \sim N(0, \sigma^2)$$

Test:

$$\mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$F = \frac{SST/(I - 1)}{SSE/(n - I)}$$

$SST$ : measures the sum of error between each group

$SSE$ : measures the sum of error within each group

### 7.2 Two-way Anova

Model:

$$y_{i,j,k} = \mu_{i,j} + \epsilon_{i,j,k}$$

$I$ : the number of groups by time,  $J$ : the number of groups by method,  $K$ : the sample in each group,  $n$ : the number of samples. ( $n = I * J * K$ )

$i = 1, 2, 3$ , corresponded with 80 seconds, 120 seconds, 240 seconds

$j = 1, 2$ , corresponded with DFB Plus and Covaris Microtube

$\mu_{11}, \mu_{12}$ , corresponded with the mean size in DFB Plus and Covaris Microtube after 80 seconds

$\mu_{21}, \mu_{22}$ , corresponded with the mean size in DFB Plus and Covaris Microtube after 120 seconds

$\mu_{31}, \mu_{32}$ , corresponded with the mean size in DFB Plus and Covaris Microtube after 240 seconds

$$\epsilon_{i,j,k} \sim N(0, \sigma^2)$$

Test:

$$\mu_{11} = \mu_{12} = \mu_{21} = \mu_{22} = \mu_{31} = \mu_{32}$$

statistics:  $F = \frac{Q/(I-1)(J-1)}{SSE/(n-IJ)}$

*SST*: measures the sum of error between each group

*SSE*: measures the sum of error within each group

## 7.3 Logistic regression + Two-way Anova

Model:

$$\log\left(\frac{p_{i,j}}{1-p_{i,j}}\right) = \mu_{i,j} + \epsilon_{i,j,k}$$

$p$ : probability of the fragmentation size in range(150bp-300bp).

$I$ : the number of groups by time,  $J$ : the number of groups by method,  $K$ : the sample in each group,  $n$ : the number of samples. ( $n = I * J * K$ )

$i = 1, 2, 3$ , corresponded with 80 seconds, 120 seconds, 240 seconds

$j = 1, 2$ , corresponded with DFB Plus and Covaris Microtube

$\mu_{11}, \mu_{12}$ , corresponded with the mean size in DFB Plus and Covaris Microtube after 80 seconds

$\mu_{21}, \mu_{22}$ , corresponded with the mean size in DFB Plus and Covaris Microtube after 120 seconds

$\mu_{31}, \mu_{32}$ , corresponded with the mean size in DFB Plus and Covaris Microtube after 240 seconds

$p_{31}, p_{32}$ , corresponded with the probability of the fragmentation size in range(150bp-300bp) in DFB Plus and Covaris Microtube after 80 seconds

$p_{21}, p_{22}$ , corresponded with the probability of the fragmentation size in range(150bp-300bp) in DFB Plus and Covaris Microtube after 120 seconds

$p_{31}, p_{32}$ , corresponded with the probability of the fragmentation size in range(150bp-300bp) in DFB Plus and Covaris Microtube after 240 seconds

$$\epsilon_{i,j,k} \sim N(0, \sigma^2)$$

## 7.4 Interaction Term

In previous three Methods with models, the interaction terms are not included. For details on models with interaction terms, tap the website link.

## 8. Appendices: Additional plots and explanation

### 8.1 Confidence Intervals Adjustment in Experiment 1(LE220), 2(LE110)

In Figure 5.4 Panel B and Figure 5.8 Panel B 95% confidence intervals built on logistic regression is not acceptable, since the data violates the logistic assumption. There is a rule in logistic regression assumption: homogeneity of variance. The variance within each of the populations is equal. At some time points, there are samples new defined response with all 0 or 1<sup>1</sup>, when the variances are 0. So here is the adjustment: if the new defined responses of samples at a time point are all 0, then the confidence interval will be (0,0). If all are 1, it will be (1,1), since they are all 0 or 1 with 100 % to fall in our building confidence interval.

### 8.2 Outliers Analysis in Experiment 3(Q sonica)

Here is the model result with outliers: In Table 8.1, no Predictors shows significance, so we need to do

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Column	11	73930235.12	6720930.47	1.01	0.4346
Row	7	46136251.56	6590893.08	0.99	0.4362
Group	1	13714677.78	13714677.78	2.06	0.1514
Residuals	556	3695023889.29	6645726.42		

Table 8.1: Summary of result from model with outliers in experiment 3

diagnostic plots for the model.

Then we find three outliers in diagnostic plots, shown in Figure 8.1:

---

<sup>1</sup>For example, in Figure5.1 Panel A, the points at time 80 sec are all outside the target region, which means the new defined response of all these samples are 0.

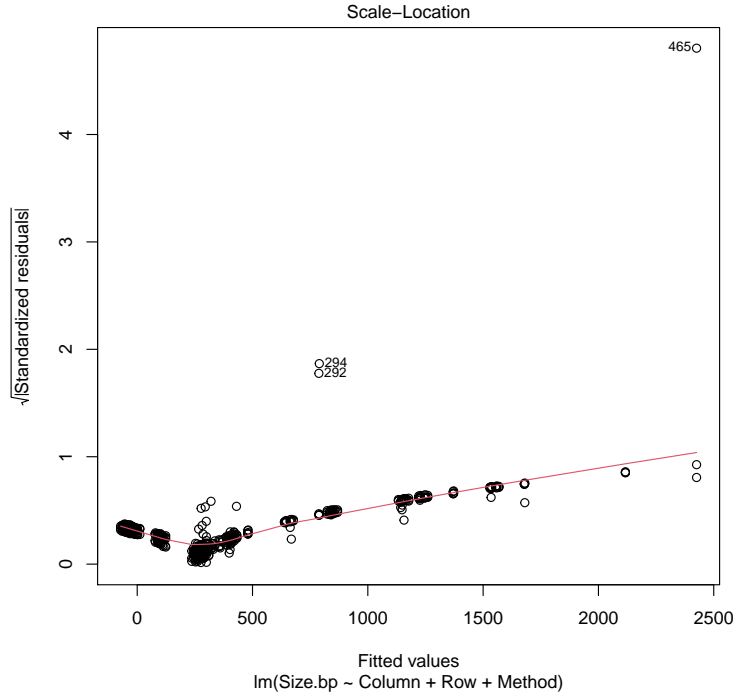


Figure 8.1: Diagnostic plot for outlier detection in model on response(fragment size) without transformation

From Figure 8.1, three outliers are shown, sample 292, 294 and 465. Then we take the log of response to give a try to solve the problem, the result shown in Figure 8.2.

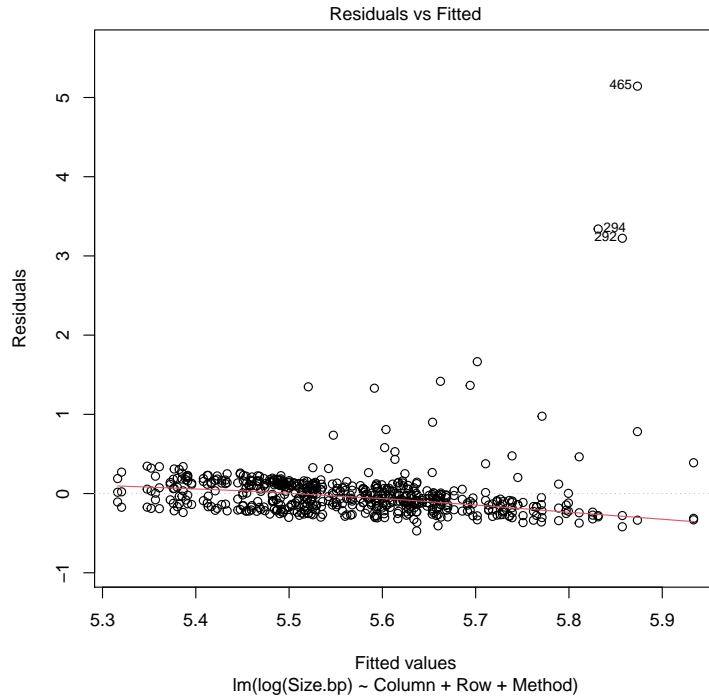


Figure 8.2: Diagnostic plot for outlier detection in model on response(fragment size) with log transformation

The three outliers are still very conspicuous. So to solve the problem, we need to delete the outliers, then

the summary of the model dropped outliers is acceptable with significant difference between two groups, see table 9.7.

### 8.3 Further Analysis Result in Experiment 3

Here we build two models for DFB minus group and DFB plus group for analysis of difference between row and column of the plate:

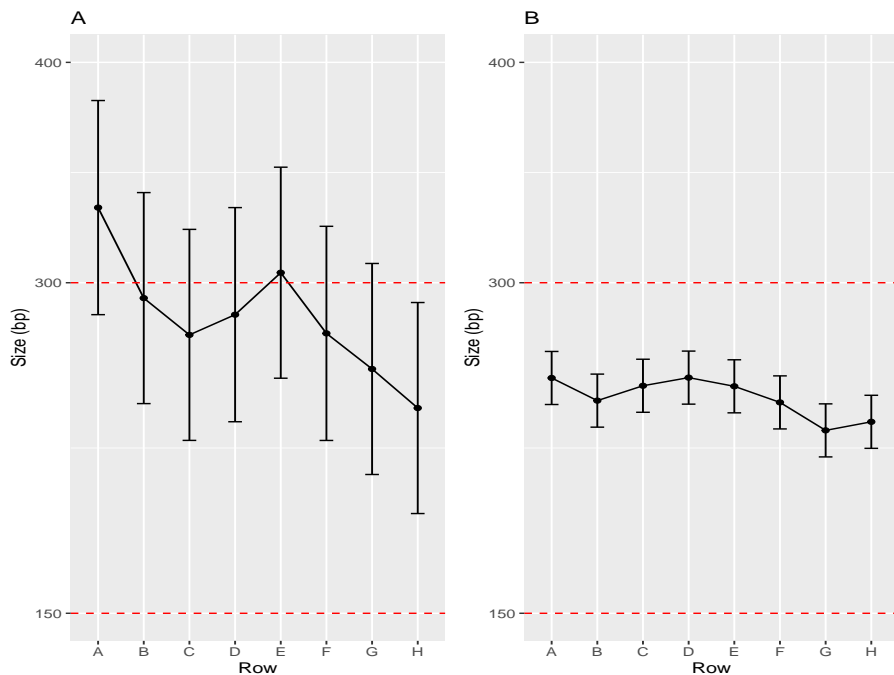


Figure 8.3: Fragment average in each row of the plate with two models in separation of DFB minus group(Panel A) and DFB plus group(Panel B).

In Figure 8.3, we compare Panel A and B, the points in Panel A is more unstable with wide confidence interval. DNA fragmentation efficiency is unsteady among different rows of the plate in DFB minus group than DFB plus group.

Then for column:



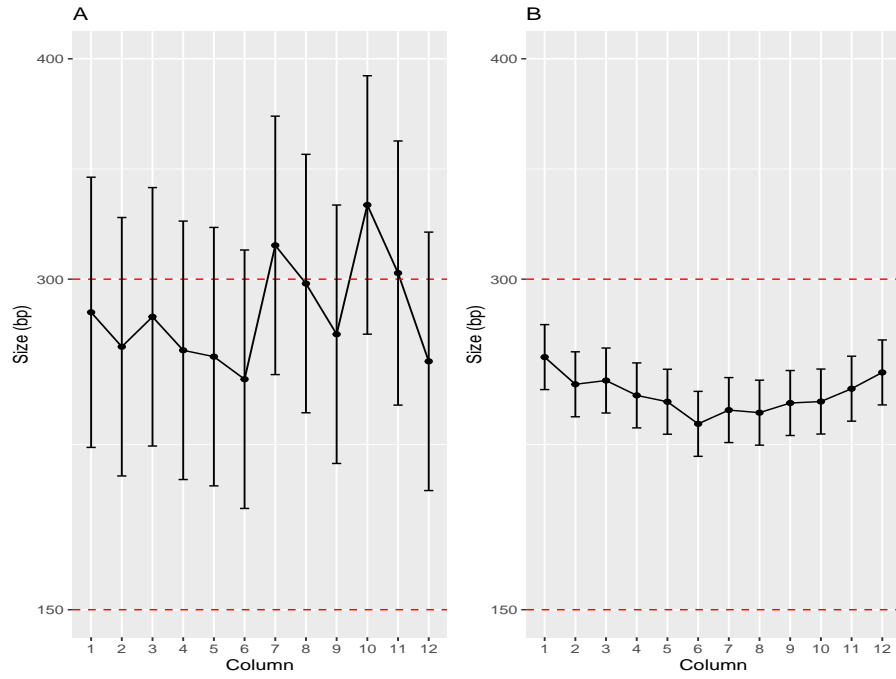


Figure 8.4: Fragment average in each column of the plate with two models in separation of DFB minus group(Panel A) and DFB plus group(Panel B).

The interpretation of Figure 8.4 is the same as Figure 8.3, so DNA fragmentation efficiency is unsteady among different columns of the plate in DFB minus group than DFB plus group.

Then Figures on logistic regression on separation of two groups are the following:

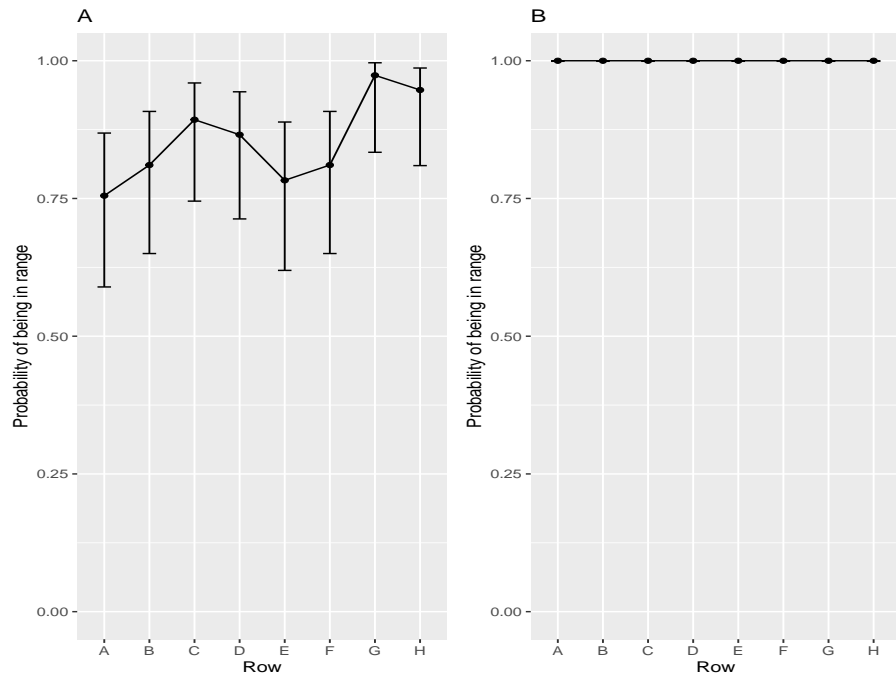


Figure 8.5: Fragment average in each row of the plate with two models in separation of DFB minus group(Panel A) and DFB plus group(Panel B).

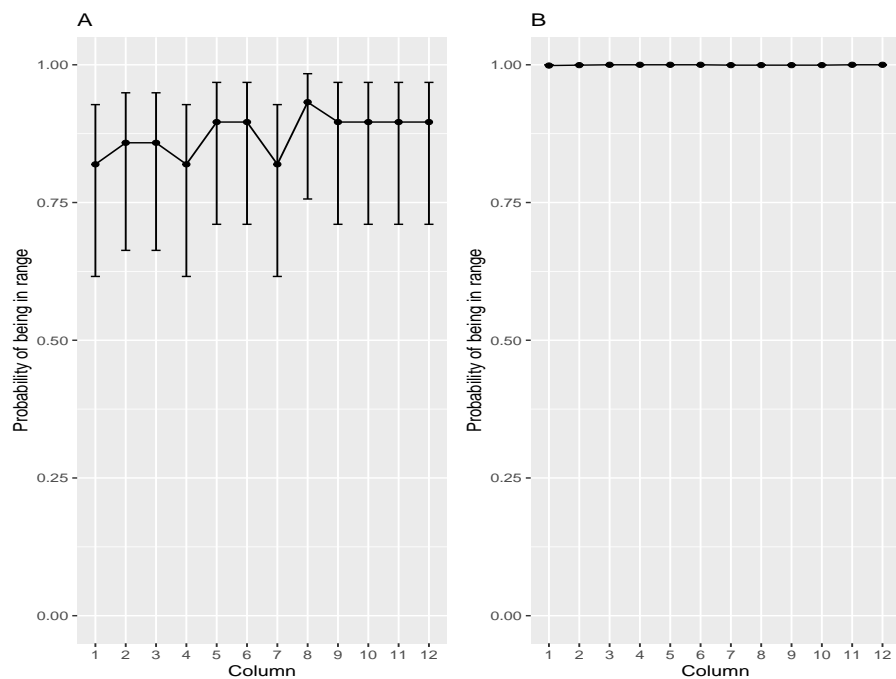


Figure 8.6: Fragment average in each row of the plate with two models in separation of DFB minus group(Panel A) and DFB plus group(Panel B).

Therefore, DNA fragmentation efficiency is unsteady among different rows and columns of the plate in DFB minus group than DFB plus group<sup>2</sup> Panel B.

<sup>2</sup>There is also an adjustment on CIs of logistic regression in Figure 8.5 Panel B and 8.6

## 9. Appendices: R output

### 9.1 Experiment 1(LE220)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	3	1620006.06	540002.02	5.85	0.0013
Residuals	68	6280803.44	92364.76		

Table 9.1: Summary of result from final model(one-way anova) in DFB minus group.  
*The difference at different sonication time points is significant.*

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	2	484753.69	242376.84	377.32	0.0000
Group	1	267704.90	267704.90	416.75	0.0000
Time:Group	2	29487.24	14743.62	22.95	0.0000
Residuals	102	65520.94	642.36		

Table 9.2: Summary of result from final model(two-way anova) between DFB plus and DFB minus group.  
*The difference at different sonication time points, different groups, and different interaction terms between time and group is significant.*

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	2	2.72	1.36	17.48	0.0000
Group	1	5.79	5.79	74.30	0.0000
Time:Group	2	9.80	4.90	62.89	0.0000
Residuals	102	7.94	0.08		

Table 9.3: Summary of result from final model(logistic regression) between DFB plus and DFB minus group.  
*The difference at different sonication time points, different groups, and different interaction terms between time and group is significant.*

## 9.2 Experiment 2(LE110)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	3	1620006.06	540002.02	5.85	0.0013
Residuals	68	6280803.44	92364.76		

Table 9.4: Summary of result from final model(quantile regression) in DFB minus group  
*The difference at different sonication time points is **not** strongly significant.*

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	3443005.23	860751.31	406.98	0.0000
Group	1	51450.82	51450.82	24.33	0.0000
Time:Group	4	6239.10	1559.78	0.74	0.5708
Residuals	50	105749.83	2115.00		

Table 9.5: Summary of result from final model(two-way anova) between DFB plus and DFB minus group.  
*The difference at different sonication time points, different groups is strongly significant, but difference on different interaction terms between time and group is **not** significant.*

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Time	4	6.60	1.65	21.52	0.0000
Group	1	0.15	0.15	1.96	0.1681
Time:Group	4	3.60	0.90	11.74	0.0000
Residuals	50	3.83	0.08		

Table 9.6: Summary of result from final model(logistic regression) between DFB plus and DFB minus group.

*The difference at different sonication time points, different interaction terms is strongly significant, but difference on different group is **not** significant.*

## 9.3 Experiment 3(Q sonica)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Column	11	77061.40	7005.58	0.63	0.8057
Row	7	149897.45	21413.92	1.92	0.0645
Group	1	191241.76	191241.76	17.14	0.0000
Residuals	553	6171931.97	11160.82		

Table 9.7: Summary of result from final model(three-way anova) without outliers between DFB plus and DFB minus groups.

*The difference between DFB plus and DFB minus groups is significant.*

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Method	1	2.25	2.25	29.76	0.0000
Column	11	0.41	0.04	0.49	0.9082
Row	7	0.97	0.14	1.82	0.0803
Residuals	556	42.03	0.08		

Table 9.8: Summary of result from final model(logisitic regression) without outliers between DFB plus and DFB minus groups.

*The difference between DFB plus and DFB minus groups is significant.*