2024-03-15

EFFECTS OF ENZYMATIC AND CHEMICAL BISULFITE CONVERSION ON CELL-FREE DNA FRAGMENTATION

Caroline G. Ploeger
Brigham Young University - Provo

Honors Thesis

EFFECTS OF ENZYMATIC AND CHEMICAL BISULFITE CONVERSION ON CELL-FREE DNA FRAGMENTATION

Caroline G. Ploeger (Catterton)

Submitted to Brigham Young University in partial fulfillment of graduation requirements for University Honors

Department of Cell Biology & Physiology
Brigham Young University
April 2024

Advisor: Dr. Timothy Jenkins

Honors Coordinator: Dr. Roy Silcox

Faculty Reader: Dr. Johnathon Hill, PhD

ABSTRACT

EFFECTS OF ENZYMATIC AND CHEMICAL BISULFITE CONVERSION ON CELL-FREE DNA FRAGMENTATION

Caroline G. Ploeger (Catterton)

Department of Cell Biology & Physiology

Bachelor of Science

This project aims to investigate the DNA fragmentation produced by enzymatic bisulfite conversion compared to DNA fragmentation produced by traditional chemical bisulfite conversion techniques for the analysis of cytosine methylation in seminal cell-free DNA (cfDNA). Traditional chemical bisulfite (HSO_3^-) conversion methods are known to introduce biases and damage DNA, potentially compromising methylation analysis accuracy. In contrast, enzymatic conversion, utilizing biological molecules, is hypothesized to offer more precise visualization with reduced fragmentation (Lambert et al., 2019). While previous studies have analyzed methylation conversion efficiency in genomic DNA, this project specifically examines how these conversion techniques impact cfDNA.

20 samples of seminal cfDNA were bisulfite-converted, 20 samples of cfDNA were enzymatically converted, and 20 samples of cfDNA were not converted and used as a control. Agilent Femto Pulse DNA analysis was used to compute DQN (DNA quality number) scores, length of sequences, and fragmentation values within each sample. Using

these data, bioinformatic analyses were ran on DNA fragmentation using the R programming language. Results indicate that bisulfite DNA conversion contributes to significantly more fragmented sequences than enzymatic conversion in cfDNA. ANOVA and Tukey's test analyses were conducted on DQN scores, revealing a statistically significant difference in DQN values between bisulfite-converted cfDNA and enzymatically converted cfDNA samples ($p = 0.048$). Analyses indicate that the average base pair length found in enzymatic-converted cfDNA were significantly longer ($p=0.0044$) than samples that were bisulfite converted. These outcomes suggest that enzymatic conversion decreases bias when converting cfDNA samples compared to using alternative methylation conversion techniques.

ACKNOWLEDGEMENTS

I extend my sincere appreciation to Dr. Timothy Jenkins and his lab at Brigham Young University for their instrumental role in initiating my journey into undergraduate research and providing unwavering support throughout the completion of my Honors Thesis. Dr. Jenkins's consistent willingness to assist me in achieving my academic ambitions has been invaluable, and I am truly grateful for his guidance. Additionally, I am thankful for the advice and encouragement offered by Isaac Stirland and Chad Pollard, Ph.D. candidates in the Jenkins lab, whose insights significantly shaped the direction of my research endeavors.

Furthermore, I wish to express my gratitude to Dr. Jonathon Hill and Dr. Roy Silcox for their assistance and kindness throughout the thesis process as members of my committee.

I would also like to acknowledge the Honors Program at Brigham Young University for providing me with the opportunity to publish my thesis, actively engage in the program, and experience spiritual enrichment throughout the journey.

Lastly, I owe a debt of gratitude to my husband, Collin Ploeger, whose steadfast academic and emotional support have been a constant source of strength. His assistance with bisulfite conversions and unwavering belief in my abilities have been instrumental in seeing this thesis through to completion.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

While polymerase chain reaction (PCR) and DNA sequencing are common fundamental techniques for amplifying DNA sequences, challenges often arise with the inherent DNA fragmentation that occurs during the process, particularly when utilizing DNA that has been bisulfite-converted (Lambert et al., 2019). In contemporary molecular biology research, understanding DNA methylation patterns is integral to comprehending various biological processes, such as gene expression regulation, cellular proliferation, and chromosomal stability (Li and Tollefsbol, 2011). Abnormal methylation patterns serve as crucial indicators of diseases, including cancer, prompting researchers to explore genomic regions where methylation commonly occurs (Li and Tollefsbol, 2011).

Bisulfite conversion is a widely employed method to distinguish 5-methylcytosine and 5-hydroxymethylcytosine (5mC and 5hmC) from non-methylated cytosines, converting non-methylated cytosines to uracils through deamination (Vaisvila et al., 2021). An overview of the DNA conversion associated with chemical bisulfite conversion is shown in Figure 1. The steps associated with bisulfite conversion include: 1) DNA denaturation; 2) incubation with bisulfite at elevated temperature; 3) removal of bisulfite by desalting; 4) desulfonation of sulfonyl uracil adducts at alkaline pH; and 5) removal of the desulfonation solution (Darst et al., 2010). Despite its prevalence, bisulfite conversion introduces biases and can damage DNA, leading to fragmentation and complicating the analysis of full-length DNA molecules (Vaisvila et al., 2021). Chemical bisulfite conversion is known for utilizing temperatures up to 95°C and pH 5 to convert non-methylated cytosines to uracils most efficiently (Grunau et al., 2001). These harsh chemical reaction conditions cause DNA fragmentation primarily due to depyrimidation

followed by alkali treatment, inducing the formation of abasic sites that subsequently result in cleavage of the DNA phosphodiester bonds (Kint et al., 2018). This thereby leads to DNA degradation. The difficulties in obtaining reliable sequencing data for bisulfite-converted DNA prompted the exploration of alternative techniques.

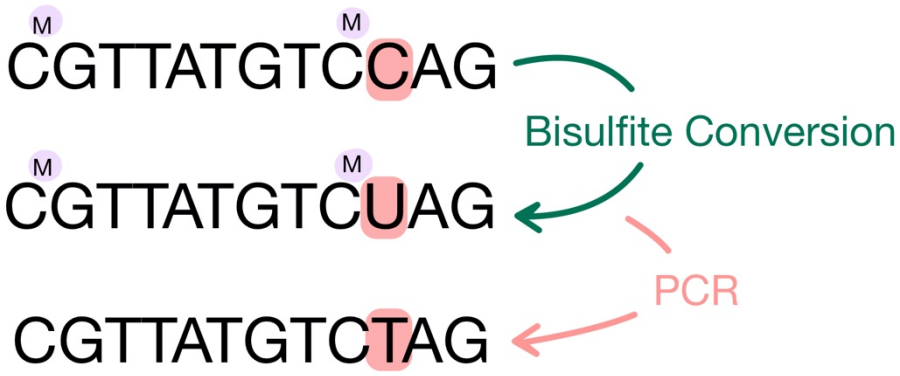


Figure 1: A representation of how bisulfite conversion altars un-methylated cytosines in a DNA sequence to uracils.

Enzymatic Methyl-seq (EM-seq) conversion, as suggested by Wang et al. (2022), presents itself as a less destructive approach, leveraging biological molecules to register cytosine profiling in DNA. This proposed enzymatic conversion offers the potential to mitigate biases and reduce DNA fragmentation errors compared to bisulfite conversion, aligning with prior findings on enzymatic methods being less destructive than chemical modifiers (Wang et al., 2022). In EM-seq, methylation dependent restriction enzymes (MDRE) MspJI and AbaSI are used to detect 5mC or 5hmC throughout the DNA. Two reactions occur. First, enzymes TET2 and T4-BGT act to convert 5mC and 5hmC into products that are resistant to deamination by APOBEC3A. In the second reaction,

APOBEC3A catalyzes the deamination of unmodified cytosines, thereby converting them into uracils (Vaisvila et al., 2021).

This study focuses on seminal cell-free DNA (cfDNA). Cell-free DNA is a unique sample type found in blood plasma and bodily fluids (like semen), known for its high fragmentation and distinctions from genomic DNA. The average fragment size for cfDNA is around 167-bp long. Several laboratories at Brigham Young University (BYU) in Provo, UT, study cfDNA. The Jenkins Lab at BYU utilizes cfDNA in attempt to understand how cfDNA methylation impacts male fertility, and grapples with the challenges posed by DNA fragmentation in bisulfite-converted cfDNA samples. This study seeks to address these challenges by systematically comparing the efficiency of bisulfite conversion and enzymatic conversion in maintaining low DNA fragmentation in seminal cfDNA. The anticipated outcomes include the finding that enzymatic conversion may exhibit lower DNA fragmentation in cfDNA samples compared to bisulfite conversion due to its less-harsh reaction conditions, aligning with the hypothesis that enzymatic conversion provides a more accurate means of visualizing seminal cfDNA (Rubenstein and Solomon, 2023).

In the context of our experiment evaluating the efficiency of enzymatic conversion versus bisulfite conversion in analyzing cytosine methylation within seminal cell-free DNA (cfDNA), the Agilent Femto Pulse system emerges as a crucial tool offering valuable insights. The Femto Pulse system employs pulsed-field capillary electrophoresis to accurately size and measure DNA fragments, serving as a sophisticated analytical platform. This technology facilitates precise evaluation of DNA quality and offers valuable insights into fragmentation patterns, essential for assessing the effects of

conversion techniques on DNA integrity. The system's ability to provide the average base pair length of samples being analyzed provides crucial insights in quantifying the overall size distribution and integrity of DNA fragments in our samples for this study (Pocernich et al., 2019). Relative fluorescent units (RFU) produced by this technology correspond to the amount of genetic material present. This metric will prove vital in assessing the impact of both enzymatic and bisulfite conversion techniques on the size distribution of seminal cfDNA, offering a nuanced understanding of the overall sequencing efficiency of the sample. As common laboratory techniques, like PCR, can have reduction in efficiency due to DNA fragmentation, it is important to determine whether the already fragmented cfDNA is impacted negatively by chemical or enzymatic conversion (Golenberg et al., 1996).

Furthermore, the system's capability to reveal fragmentation patterns is particularly relevant to our study. By identifying the specific size ranges where fragmentation occurs, the nature and extent of cfDNA fragmentation induced by each conversion method can be discerned. This information will help us understand the potential biases and limitations associated with enzymatic and bisulfite conversion, ultimately aiding in the optimization of our experimental procedures.

The DQN Score provided by the Femto Pulse system is also applicable to this investigation. Users of the Femto Pulse system define a size threshold deemed appropriate for their investigation prior to running fragmentation analyses. The DQN score produced by the system is a calculated value from 0-10 representing the concentration of the sample that lies above the selected size threshold. 10 indicates that 100% of the sample lies above the threshold value, and 0 indicates that none of the

sample exceeded the threshold value. As bisulfite-converted and enzymatic-converted cfDNA samples are compared, the DQN score, with the user-defined size threshold of 300 base pairs, allows one to quantitatively assess the proportion of DNA fragments within the specified size range. This tailored analysis will be pivotal in evaluating the DNA quality based on a threshold relevant to our specific application, shedding light on potential differences in the efficiency of both conversion methods (Pocernich et al., 2019).

Moreover, the fragment size distribution profiles generated by the system will complement our analysis by illustrating how DNA fragments are distributed across various size ranges. This information allows for characterization of the heterogeneity of DNA fragments within the samples. Femto Pulse, in turn, allows for optimization of this methodology for the accurate analysis of DNA fragmentation patterns in seminal cfDNA.

The significance of this research extends to the academic community, particularly within the Department of Cell Biology and Physiology at BYU. The anticipated results, along with the demonstrated advantages of enzymatic conversion in mitigating biases and reducing DNA fragmentation errors, have the potential to influence methodologies in molecular biology labs beyond BYU. This project sets the stage for a comprehensive exploration of the experimental design, methodologies, and expected outcomes of this error research initiative, ultimately contributing to the broader academic discourse on optimizing DNA sequencing techniques for methylation analysis in complex biological samples.

METHODOLOGY

In this study, seminal cell-free DNA samples from human subjects were subjected to an in-depth investigation to evaluate the efficiency of enzymatic conversion compared to conventional bisulfite conversion techniques in analyzing cytosine methylation. Under Institutional Review Board Approval, human DNA samples were utilized from Utah Fertility Center in the Jenkins Lab at BYU to conduct this experiment. The samples were first processed to extract cell-free DNA, and once cell-free, samples were then selected to either be a control, bisulfite-converted or enzymatic-converted. A total of 60 cfDNA samples were analyzed, with 20 designated for chemical bisulfite conversion, 20 designated for enzymatic conversion, and 20 that were not subjected for methylated cytosine conversion.

The bisulfite conversion process was carried out following a standardized protocol. Initially, DNA strands underwent denaturation to expose their bases, followed by incubation with bisulfite under elevated temperatures. During this incubation, unmethylated cytosine residues were chemically modified to uracil, while methylated cytosines remained unaltered. Subsequent desulfonation removed sulfonated groups, resulting in bisulfite-converted DNA. This conversion was achieved utilizing the EZ DNA Methylation kit from Zymo Research, which offers reliable and reproducible conversion procedures. Additionally, enzymatic conversion procedures were conducted using the Enzymatic Methyl-seq kit to quantify DNA methylation levels. These enzymatic methods were chosen for their ability to minimize biases and reduce errors associated with DNA fragmentation. By employing enzymatic reactions, the enzymatic conversion approach aimed to provide more accurate assessments of DNA methylation.

With hopes of future sequencing producing at least a 300-bp amplicon, the target threshold value for Femto Pulse DNA fragmentation analysis was set to be 300-bp long for this study. DNA fragmentation analysis via Agilent Femto Pulse ensued after pipetting the bisulfite-converted, enzymatic converted, and regular cfDNA into designated wells for the system to process. This enabled assessment of the size distribution and fragmentation patterns of all samples.

Once Femto Pulse data were produced, bioinformatic analyses conducted in R processing software focused on the average base pair calls per sample, DNA fragmentation, and DQN scores. For initial fragmentation comparison between bisulfite-converted and enzymatically converted DNA samples, t-tests, ANOVA, and Tukey's test were conducted on the DQN values of all bisulfite-converted, enzymatic-converted, and control samples. This code can be found in Figure 2. The Femto Pulse system calculated the DQN value, which assessed the fraction of the total measured concentration of the sample that exceeds the specified size threshold (300 base pairs). The DQN score ranged from 0 to 10, with 0 indicating that none of the sample exceeds the threshold and 10 indicating that 100% of the sample lies above the threshold value. This tailored analysis provided a quantitative measure of DNA quality, aiding in the evaluation of conversion techniques based on the user-defined size threshold in preparation for DNA sequencing.

```
# Load necessary libraries
library(ggplot2)
library(dplyr)

# Create a dataframe with the provided data
df <- data.frame(
  Group = c(rep("Control", 2), rep("Enzymatically converted cfDNA", 3), rep("Bisulfite-converted
cfDNA", 2)),
  DQN = c(insert DQN values here separated by commas)
)

# Create color palette
colors <- c("Control" = "lightblue", "Enzymatically converted cfDNA" = "aquamarine3",
"Bisulfite-converted cfDNA" = "coral")

# Reorder factor levels
df$Group <- factor(df$Group, levels = c("Control", "Enzymatically converted cfDNA", "Bisulfite-
converted cfDNA"))

# Create boxplot
ggplot(df, aes(x = Group, y = DQN, fill = Group)) +
  geom_boxplot() +
  scale_fill_manual(values = colors) +
  labs(title = "Comparison of DQN Scores",
       x = "Group",
       y = "DQN") +
  theme_minimal()

# Perform ANOVA
anova_result <- aov(DQN ~ Group, data = df)

# Perform Tukey's post-hoc test
tukey_result <- TukeyHSD(anova_result)

# Print ANOVA summary
summary(anova_result)

# Print Tukey's test results
print(tukey_result)
```

Figure 2: Example code for creating a DQN data frame, calculating boxplots, ANOVA, and Tukey's test

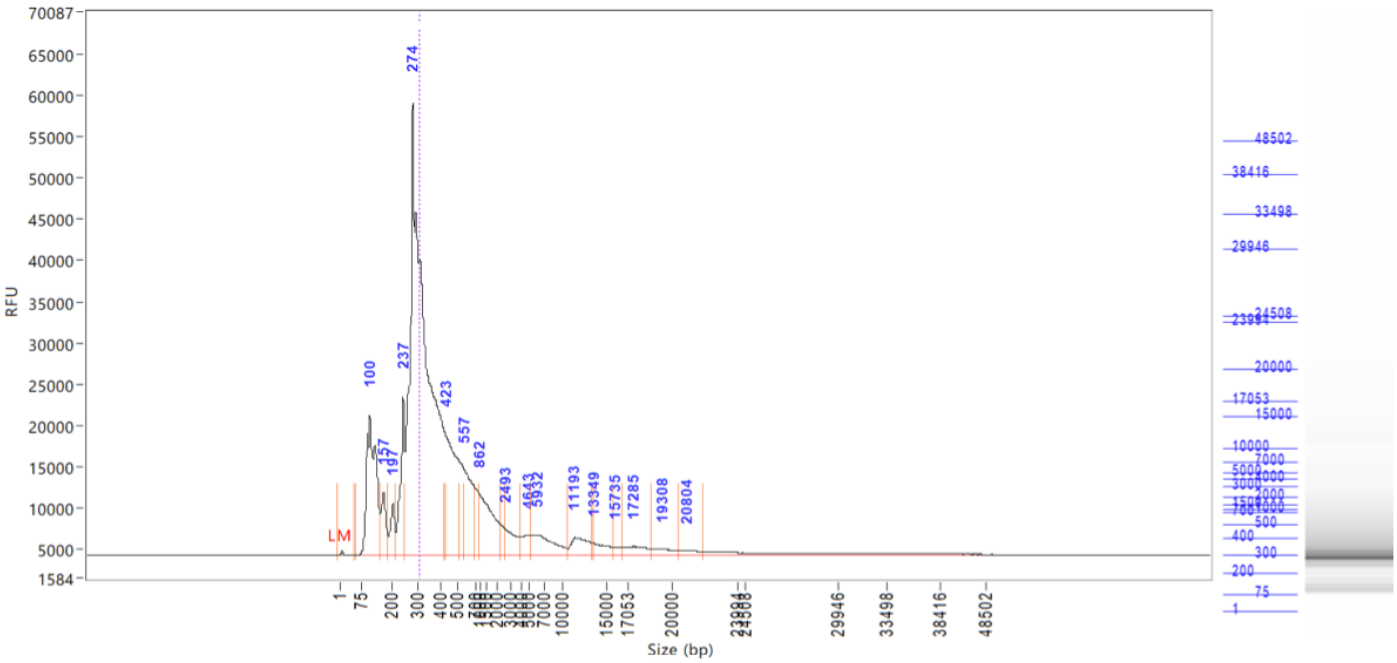
Sample data were then organized into a data frame entitled “honors” with three columns: “basepair length,” “RFU” associated with that base pair length, and its associated “sample” number. Different sample numbers indicated whether the sample was bisulfite-converted, enzymatic-converted, or a control. Utilizing R-processing, a new data frame was made to include the average base pair lengths found that corresponded to $\geq 75^{th}$ percentile of RFU scores per sample. The cfDNA sizes associated with the top 25% RFU values were chosen to analyze due to the vast amount of base pair lengths found by Femto Pulse analysis, as well as stronger Femto Pulse RFU signals corresponding to a greater presence of DNA. Statistical analyses were performed, including analysis of variance (ANOVA) and a Tukey's test. These analyses aimed to compare the average base pair length among the different cfDNA conversion methods. Sample code for this process can be found in Figure 3.

```
# Load necessary libraries
library(dplyr)
library(ggplot2)

# Assuming you already have your "honors" dataframe loaded

# Step 1: Filter data to include only those rows with RFU values >= 75th percentile within each sample group
new_df <- honors %>%
  group_by(sample) %>%
  filter(RFU >= quantile(RFU, 0.75)) %>%
  select(sample, bp_length)

#The following code is an example of if you had samples 24 and 25 being enzymatic-converted cfDNA, as well as 27 and 28 being bisulfite-converted DNA

# Step 2: Combine samples 24 and 25, and samples 27 and 28
new_df$group <- ifelse(new_df$sample %in% c(24, 25), "Enzymatically Converted cfDNA", "Bisulfite-converted cfDNA")

# Step 3: Create boxplot
ggplot(new_df, aes(x = group, y = bp_length, fill = group)) +
 geom_boxplot() +
 scale_fill_manual(values = c("Enzymatically Converted cfDNA" = "seagreen3", "Bisulfite-converted cfDNA" = "lightcoral")) +
 labs(title = "Comparison of bp Length",
      x = "Sample Group",
      y = "bp Length")

# Step 4: Perform ANOVA and Tukey's test
anova_result <- aov(bp_length ~ group, data = new_df)
tukey_result <- TukeyHSD(anova_result)

# Print ANOVA summary
summary(anova_result)

# Print Tukey's test results
tukey_result
```

Figure 3: Example code to create a new data set including only the data associated with the top 25% of RFU values for each sample, and how to code ANOVA and Tukey's test in R

RESULTS

Results from Femto Pulse include visualized graphs of DQN scores, lengths of DNA fragments found in each sample, DNA concentration, and the relative fluorescence units (RFU) as a quantity of the intensity of fluorescence emitted by nucleic acids during electrophoresis. RFU values referred to the amount of DNA present in a sample based on the intensity of fluorescence detected by the instrument. An example of the fragmentation graphs and charts produced by Femto Pulse is found in Figure 4. An example DQN graph is shown in Figure 5.



Peak	Size (bp)	Conc. (pg/uL)	From (bp)	To (bp)	Avg. Size (bp)	CV%	RFU	Corr. Peak Area
1	1 (LM)	2.9922	0	44	3	186.96	589	4.933
2	100	3780.4036	51	144	108	16.17	16898	623.259
3	157	895.0121	144	178	158	5.50	7635	147.557
4	197	610.3443	178	207	193	3.91	6276	100.625
5	237	1728.7268	207	243	230	3.93	19131	285.008
6	274	18315.4316	243	418	314	14.38	54776	3019.589
7	423	3162.1001	423	515	464	5.45	14498	521.322
8	557	1621.0182	557	688	620	6.21	10238	267.250
9	862	1784.1108	862	2218	1499	25.73	7342	294.139
10	2493	557.9460	2493	3800	3051	12.53	3160	91.986
11	4643	351.6090	3800	5001	4396	8.28	2395	57.968
12	5932	815.8538	5001	10345	7072	20.81	2461	134.506
13	11193	499.3502	10483	13166	11823	6.14	2125	82.326
14	13349	274.5244	13349	15547	14421	4.51	1386	45.260
15	15735	103.1042	15547	16451	15983	1.66	988	16.998
16	17285	285.8683	16451	18588	17530	3.31	1007	47.130
17	19308	177.9359	18588	20318	19409	2.55	735	29.336
18	20804	113.4852	20366	21827	21051	1.97	541	18.710

Figure 4: An example of one fragmentation analysis produced by Femto Pulse on enzymatic-converted cfDNA

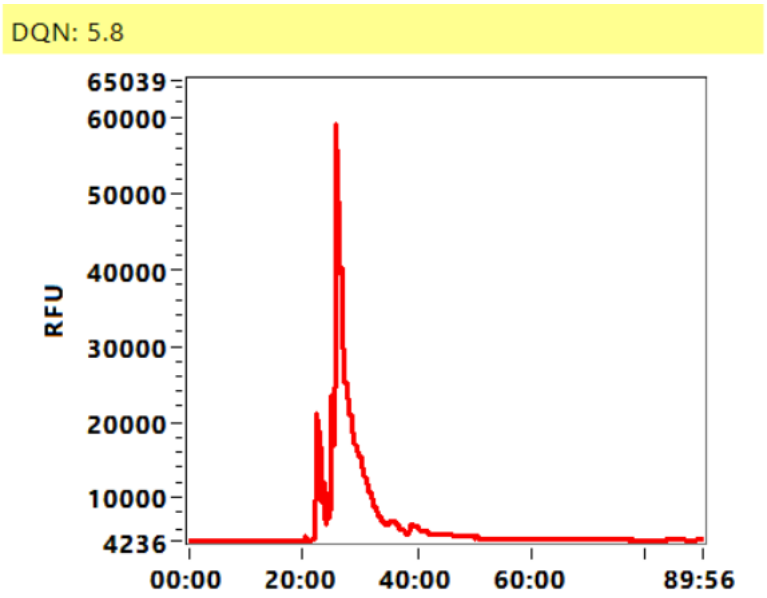


Figure 5: An example of one DQN analysis produced by Femto Pulse on enzymatic-converted cfDNA with the y-axis being time in minutes.

Initial insights derived from the Femto Pulse analysis provide a preliminary understanding of DNA quality number values in the context of different cfDNA conversion methods. The observed trend in DQN values indicated a progressive decrease from non-converted cfDNA to enzymatically converted cfDNA and further to bisulfite-converted DNA. Specifically, non-converted cfDNA exhibited the highest DQN values, averaging around 9 DQN. In contrast, enzymatically converted cfDNA displayed an intermediate decrease, averaging approximately 4.5 DQN. Notably, bisulfite-converted DNA exhibited the lowest DQN values, with a range around 1.8 DQN. These initial findings are found in Figure 6.

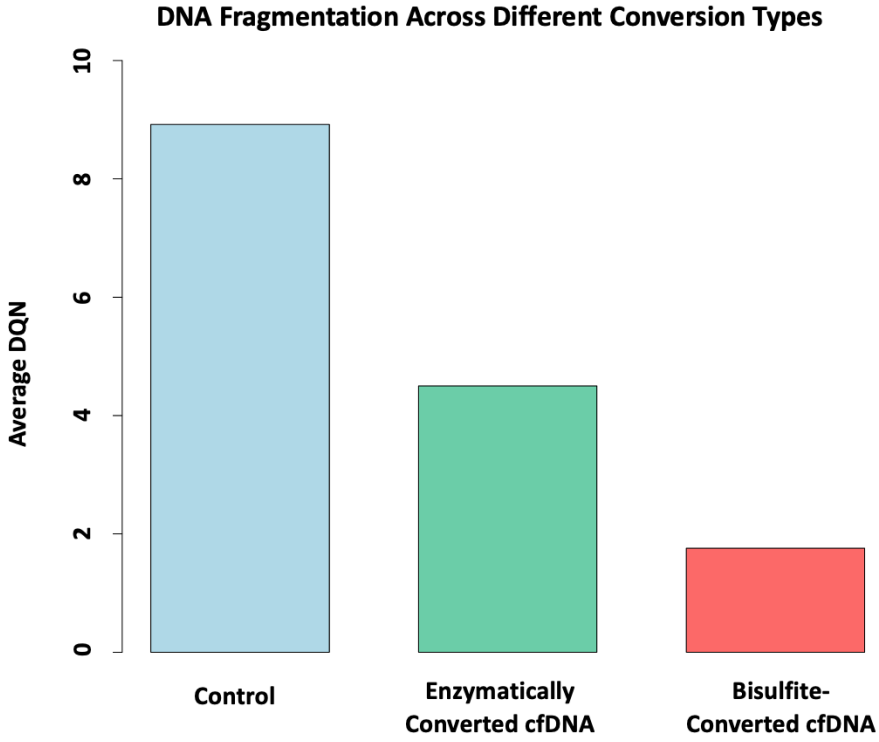


Figure 6: A comparison of average DQN values between non-converted, enzymatic-converted, and bisulfite-converted cfDNA is shown following Femto Pulse DNA fragmentation analyses. The average control DQN value was ~9, while enzymatically converted cfDNA had an average DQN of 4.5 and bisulfite-converted cfDNA had an average DQN of ~1.8.

The ANOVA analysis revealed a statistically significant difference in DQN values between enzymatically converted cfDNA and bisulfite-converted cfDNA samples ($p = 0.048$). The p-value reached conventional significance levels ($p < 0.05$), suggesting a notable trend in DQN values that warrants further exploration. Post hoc analysis using Tukey's test was conducted to discern specific differences between enzymatically converted, bisulfite-converted, and control cfDNA DQN values. The mean difference between enzymatic-converted cfDNA and bisulfite-converted DNA was -2.74, with a 95% confidence interval ranging from -5.47 through -0.017. A statistically significant difference was found between the control and bisulfite-converted cfDNA DQN values ($p=0.0000006$). The mean difference between the control and bisulfite-converted cfDNA

was -7.16 with a 95% confidence interval ranging from -9.46 through -4.87. A statistically significant difference was also found between enzymatic-converted cfDNA and the control DQN values (p=0.00085). The mean difference between the control and enzymatic-converted cfDNA was -4.42 with a 95% confidence interval ranging from -6.97 through -1.87. These results can be found in Figure 7.

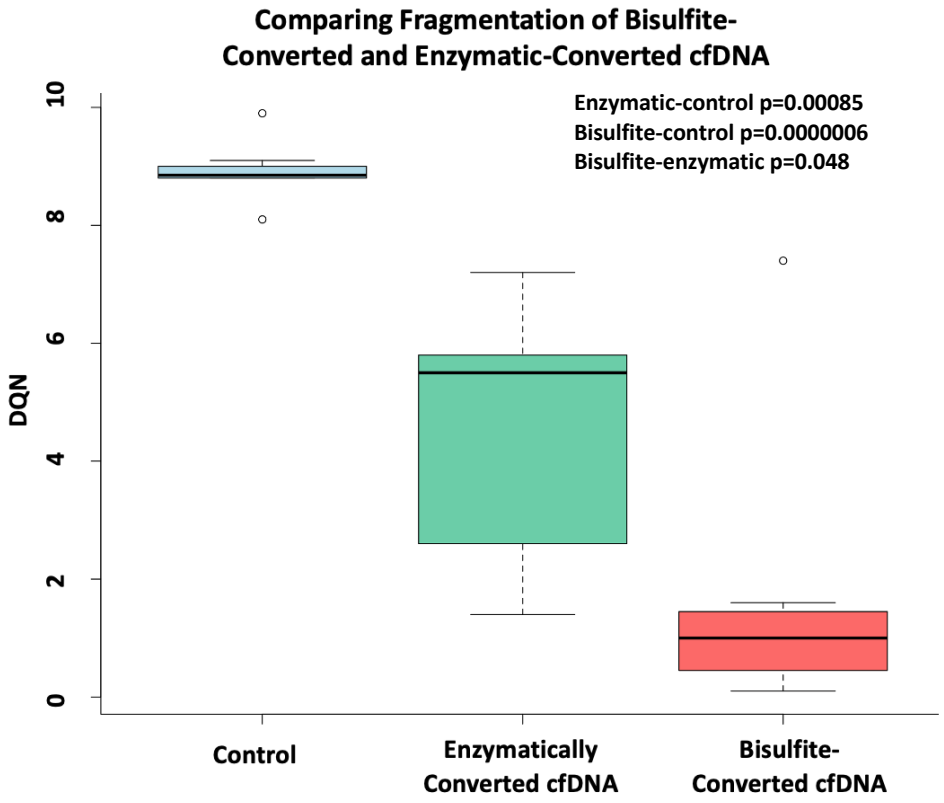


Figure 7: ANOVA results indicating the statistically significant difference between DQN values of the control group, enzymatic-converted cfDNA, and bisulfite-converted cfDNA is shown. Tukey's test results are in the upper right corner.

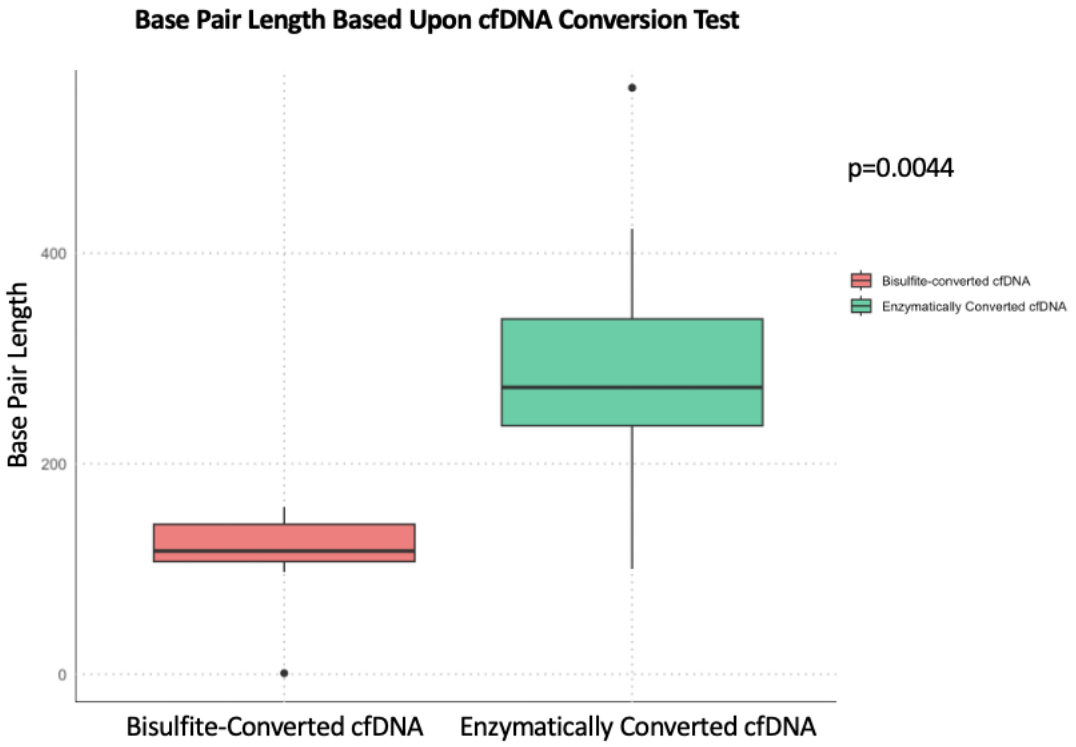


Figure 8: T-tests indicating the statistically significant difference in base pair lengths among cfDNA samples that were enzymatic-converted versus bisulfite-converted. Green indicates longer fragment length. Only base pair lengths of ≥75th percentile of RFU values were considered.

The analysis on the top 25% of found base pair lengths revealed that enzymatically converted cfDNA exhibited notably longer base pair lengths compared to bisulfite-converted cfDNA. Specifically, t-tests conducted on the base pair lengths with ≥75th percentile of RFU values yielded a p-value of 0.0044, indicating a statistically significant difference between the two conversion methods (Figure 8).

DISCUSSION

The statistical analyses conducted on DNA quality number values in this study have provided valuable insights into the fragmentation disparities observed between bisulfite-converted and enzymatically converted cell-free DNA. ANOVA results met conventional thresholds of significance and the findings from Tukey's test indicate a potentially meaningful distinction between the two conversion methods. However, it is crucial to exercise caution in interpreting these results, especially given the adjusted p-values and small sample sizes, which underscore the need for nuanced interpretation. Consequently, further investigation, potentially involving larger sample sizes and perhaps exploring additional statistical approaches, is warranted to fully elucidate the nature and significance of the observed differences in values between the two conversion methods. It is crucial to understand that DQN only measures the amount of DNA available during Femto Pulse that reach the targeted user-defined threshold value for base pair length. ANOVA tests of DQN data between DNA conversion tests suggest a significant difference between the tests' ability to produce the selected threshold value of 300 base pairs. ANOVA results comparing base pair lengths also suggest that there is a statistically significant difference in the ability of conversion tests to produce longer lengths of cfDNA.

The statistically significant difference between the DQN of the control group and bisulfite-converted cfDNA as well as the statistically significant difference between the DQN of the control group and enzymatic-converted cfDNA suggest that any kind of unmethylated cytosine conversion causes fragmentation to cfDNA samples, and the availability of 300-bp long genomic material is different between those groups. This

fragmentation was further elucidated by comparing the average base pair lengths that Femto Pulse registered for each sample.

Investigation uncovering the statistically significant difference ($p=0.0044$) between base pair lengths found in bisulfite-converted cfDNA compared to enzymatically converted cfDNA further allowed quantification of the extent by which the size of cfDNA is impacted by bisulfite conversion. It is important to note that this study only analyzed the top 25% lengths of base pairs found corresponding to the 75^{th} percentile of RFU values and beyond, so this difference may not be fully representative of the average length of base pairs found in samples.

Methylation analyses, integral to our understanding of epigenetic mechanisms, rely heavily on the accurate representation of DNA fragments to decipher methylation patterns at specific genomic sites (Umer & Herceg, 2013). The disparities observed in samples that underwent bisulfite conversion underscore the critical importance of meticulously assessing the impact of conversion methods on cfDNA integrity when interpreting methylation data. Moreover, common challenges encountered in sequencing, such as uneven coverage and sequencing errors, further emphasize the pivotal role of preserving cfDNA integrity for accurate downstream analysis. Addressing these challenges requires a comprehensive understanding of the effects of different conversion methods on DNA fragmentation, including potential sources of bias like bisulfite conversion techniques explored in this study.

Researchers may believe this study's data suggest that enzymatic conversion is the best methylated cytosine converter for cfDNA; yet, it is pertinent to consider the financial implications associated with each conversion kit, as this may significantly

influence the selection of conversion method in both research and clinical settings. An average Enzymatic Methyl-seq kit costs around $904 for 24 reactions or $3,394 for 96 reactions (Biolabs). On the other hand, a bisulfite-conversion kit through Zymo Research costs $170.60 for 50 reactions, or $587.10 for 200 reactions (EZ DNA). While enzymatic conversion may offer advantages in preserving longer DNA fragments, the higher associated cost compared to bisulfite conversion could pose a barrier, particularly for studies with limited budgets. For this reason, exploring cost-effective alternatives or optimizations to mitigate financial constraints associated with enzymatic conversion is essential for ensuring accessibility and affordability in research and clinical practice. Laboratories must consider whether they value less cfDNA fragmentation or less financial constraints when determining to use chemical or enzymatic cfDNA bisulfite conversion methodologies. Oxford Nanopore sequencing strategies may also be a well sought-after alternative to traditional bisulfite conversion technologies due to the bypassing of chemical bisulfite conversion and recognition of methyl groups on DNA during sequencing.

Future initiatives could include comparative studies evaluating fragmentation post-sequencing using nanopore or Illumina platforms, given their potential differential sensitivity to DNA fragmentation. Additionally, comparative analysis of fragmentation patterns across various biological sources of DNA, such as neuronal, blood, and sperm cfDNA, could yield valuable insights into the impact of tissue-specific factors on DNA integrity. By addressing these considerations with meticulousness, we can advance our understanding of epigenetic mechanisms and their implications in diverse biological

contexts, ultimately contributing to advancements in both basic and translational research endeavors.

CONCLUSION

In choosing between enzymatic and bisulfite conversion methods, researchers must carefully weigh the specific requirements of their study. While bisulfite conversion is widely employed for methylation studies, these results suggest potential drawbacks in terms of DNA fragmentation. Enzymatic conversion, on the other hand, emerges as a promising alternative for preserving DNA integrity, especially when longer fragments are essential for downstream analyses.

Future investigations should explore how bisulfite conversion effects different biological sources of cfDNA, such as blood. Additionally, exploring the impact of these conversion methods on specific genomic regions of interest will provide more targeted insights into their utility in the study of cytosine methylation.

This study suggests that enzymatically converted cfDNA maintain a higher mean base pair length, indicating a potential advantage in preserving longer DNA fragments compared to the bisulfite conversion technique. These findings underscore the importance of methodological considerations in the study of cfDNA, offering valuable insights for researchers seeking to optimize their experimental approaches and decrease DNA fragmentation in methylation studies.

REFERENCES

Biolabs, N. E. (n.d.). *NEBNext® Enzymatic Methyl-seq Kit*. NEB. https://www.neb.com/en-us/products/e7120-nebnext-enzymatic-methyl-seq-kit

Darst, R. P., Pardo, C. E., Ai, L., Brown, K. D., & Kladde, M. P. (2010). Bisulfite sequencing of DNA. *Current protocols in molecular biology*, *Chapter 7*, Unit–7.9.17. https://doi.org/10.1002/0471142727.mb0709s91

EZ DNA Methylation Kit. ZYMO RESEARCH. (n.d.). https://www.zymoresearch.com/collections/ez-dna-methylation-kits/products/ez-dna-methylation-kit

Golenberg, E. M., Bickel, A., & Weihs, P. (1996). Effect of highly fragmented DNA on PCR. Nucleic Acids Research, 24(24), 5026–5033. https://doi.org/10.1093/nar/24.24.5026

Grunau, C., Clark, S. J., & Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Research, 29(13), e65. https://doi.org/10.1093/nar/29.13.e65

Kint, S., De Spiegelaere, W., De Kesel, J., Vandekerckhove, L., & Van Criekinge, W. (2018). Evaluation of bisulfite kits for DNA methylation profiling in terms of DNA fragmentation and DNA recovery using digital PCR. *PloS one*, *13*(6), e0199091. https://doi.org/10.1371/journal.pone.0199091

Lambert, N., Robertson, A., Srivas, R., Peterman, N., Close, J., Wilson, T., George, P., Wood, H., Wong, B., Tezcan, A., & Tezcan, H. (2019). 47P - Comparison of enzymatic-and bisulfite conversion to map the plasma cell-free methylome in cancer. *Annals of Oncology, 30*, v13. 10.1093/annonc/mdz238.045

Li Y, Tollefsbol TO. DNA methylation detection: bisulfite genomic sequencing analysis. Methods Mol Biol. 2011;791:11-21. doi: 10.1007/978-1-61779-316-5_2. PMID: 21913068; PMCID: PMC3233226.

Pocernich, C., Uthe, J., Pike, W., & Wong, K. (2019). Best Quantification Practices with the Agilent 5200 Fragment Analyzer System. *Genetic Engineering & Biotechnology News.* https://www.agilent.com/cs/library/brochures/FA_eBook_Agilent.pdf

Rubenstein, D. R., & Solomon, J. (2023). Target-enriched enzymatic methyl sequencing: Flexible, scalable and inexpensive hybridization capture for quantifying DNA methylation. *PLOS ONE*, *18*(3), e0282672. https://doi.org/10.1371/journal.pone.0282672

Umer, M., & Herceg, Z. (2013). Deciphering the epigenetic code: an overview of DNA methylation analysis methods. *Antioxidants & redox signaling*, *18*(15), 1972–1986. https://doi.org/10.1089/ars.2012.4923

Vaisvila, R., Ponnaluri, V. K. C., Sun, Z., Langhorst, B. W., Saleh, L., Guan, S., Dai, N., Campbell, M. A., Sexton, B. S., Marks, K., Samaranayake, M., Samuelson, J. C., Church, H. E., Tamanaha, E., Corrêa, I. R., Jr, Pradhan, S., Dimalanta, E. T., Evans, T. C., Jr, Williams, L., & Davis, T. B. (2021). Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome research*, *31*(7), 1280–1289. Advance online publication. https://doi.org/10.1101/gr.266551.120

Wang, T., Loo, C. E., & Kohli, R. M. (2022). Enzymatic approaches for profiling cytosine methylation and hydroxymethylation. *Molecular metabolism*, *57*, 101314. https://doi.org/10.1016/j.molmet.2021.101314