

# cfDNA Fragmentation Analysis in ALS and Control Plasma Samples

ML Mansego

## Contents

<b>Introduction</b>	<b>1</b>
<b>1. Metadata Loading</b>	<b>1</b>
<b>2. Insert Size Analysis</b>	<b>2</b>
2.1 Extraction and Preprocessing . . . . .	2
2.2 Visualization of Fragment Size Distribution . . . . .	2
2.3 Fragment Size Differences by Group (ALS vs Control) . . . . .	2
2.4 Fragmentation Differences Between Bulbar- and Spinal-Onset ALS . . . . .	4
<b>Summary and Interpretation</b>	<b>7</b>
<b>Conclusion</b>	<b>8</b>

## Introduction

Cell-free DNA (cfDNA) in plasma originates predominantly from apoptotic and necrotic cellular processes. The resulting fragmentation pattern reflects nucleosomal architecture and chromatin accessibility, providing valuable insights into tissue-specific cell death and disease-associated molecular changes. In the context of amyotrophic lateral sclerosis (ALS), cfDNA has emerged as a potential non-invasive biomarker under active investigation. Fragmentation characteristics—particularly insert size distributions—may capture signals related to disease-specific cell turnover, inflammatory processes, or epigenomic alterations.

The present analysis investigates differences in cfDNA fragmentation patterns between ALS patients and healthy controls, based on insert size distributions derived from EM-seq data. Fragment size profiles are quantified using insert size metrics obtained through Picard tools and compared across groups to identify disease-associated patterns.

## 1. Metadata Loading

The sample metadata, including subject ID, condition (ALS or control), age, gender, and ALS phenotype (bulbar or spinal, for patients only), is imported and used to stratify insert size data.

## 2. Insert Size Analysis

### 2.1 Extraction and Preprocessing

Insert size data were extracted from Picard output files. Only fragment lengths  $\leq 600$  bp were retained for biological relevance. These lengths encompass mono- ( $\sim 147$  bp) and di-nucleosome ( $\sim 320$  bp) structures, typical of apoptotic DNA fragmentation.

### 2.2 Visualization of Fragment Size Distribution

The mean insert size distribution per group is shown below. Shaded ribbons represent  $\pm 1$  standard error. Vertical dashed lines indicate mono- and di-nucleosome boundaries. Additionally, yellow and blue shaded areas highlight the short (110–150 bp) and long ( $\geq 180$  bp) fragment ranges, respectively.

#### Fragment Size Distribution in Plasma cfDNA

Shaded regions represent  $\pm 1$  standard error

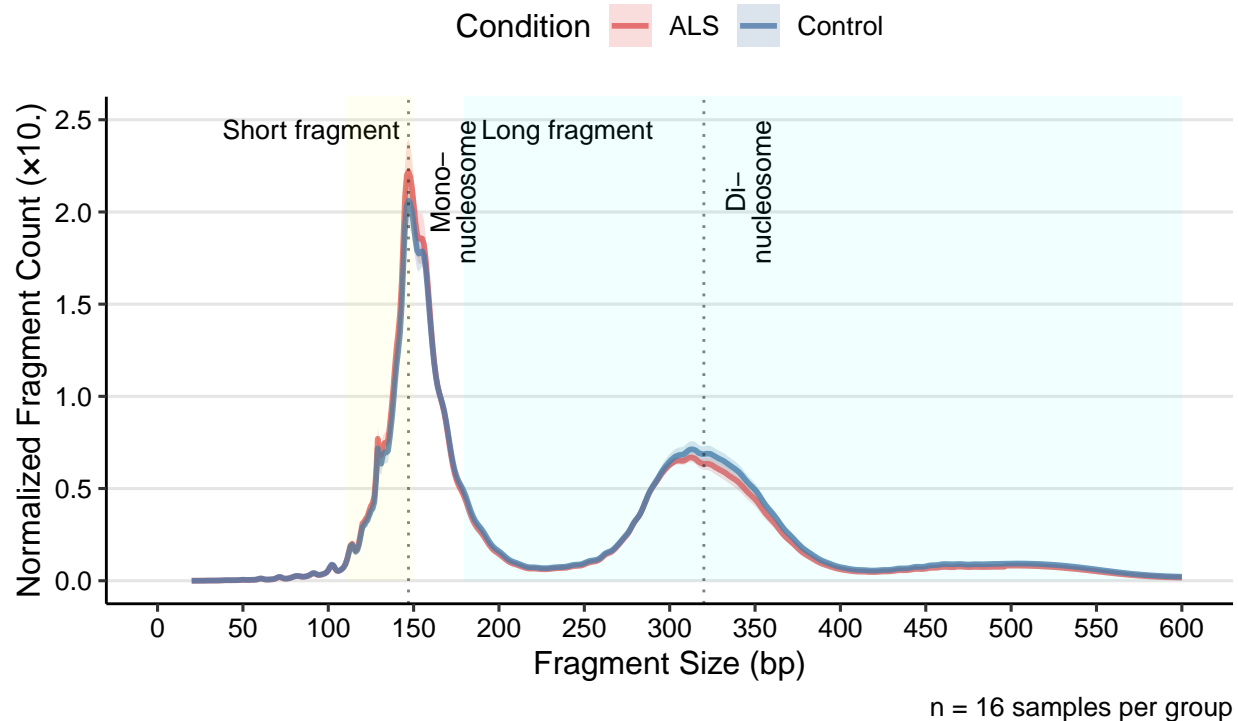


Figure 1: Mean fragment size distribution by group

### 2.3 Fragment Size Differences by Group (ALS vs Control)

To assess differential fragmentation, insert sizes were binned in 5 bp intervals. For each bin, a Wilcoxon rank-sum test was applied to evaluate count differences between ALS and control groups.

#### Significant Bins

The top 10 most differential bins are presented below (sorted by nominal p-value):

Table 1: Top 10 Most Differential Fragment Size Bins

bin	mean_ALS	mean_Control	mean_diff	p.value	bin_start	bin_end	bin_mid
120-125	358431.53	336591.45	21840.075	0.0636	120	125	122.5
130-135	734081.38	682261.20	51820.175	0.1064	130	135	132.5
135-140	1043371.20	962803.15	80568.050	0.1374	135	140	137.5
125-130	585399.57	544443.62	40955.950	0.1460	125	130	127.5
140-145	1670370.43	1541985.35	128385.075	0.1460	140	145	142.5
115-120	223961.20	210958.20	13003.000	0.1527	115	120	117.5
110-115	168327.80	164283.23	4044.575	0.2518	110	115	112.5
145-150	2157997.88	2017578.23	140419.650	0.2770	145	150	147.5
105-110	66905.05	66778.32	126.725	0.2802	105	110	107.5
80-85	23718.70	24564.33	-845.625	0.2900	80	85	82.5

### Differential Fragmentation Plot

The following plot highlights insert size bins where ALS samples differ most from controls. Red bars indicate bins with nominal significance ( $p < 0.15$ ).

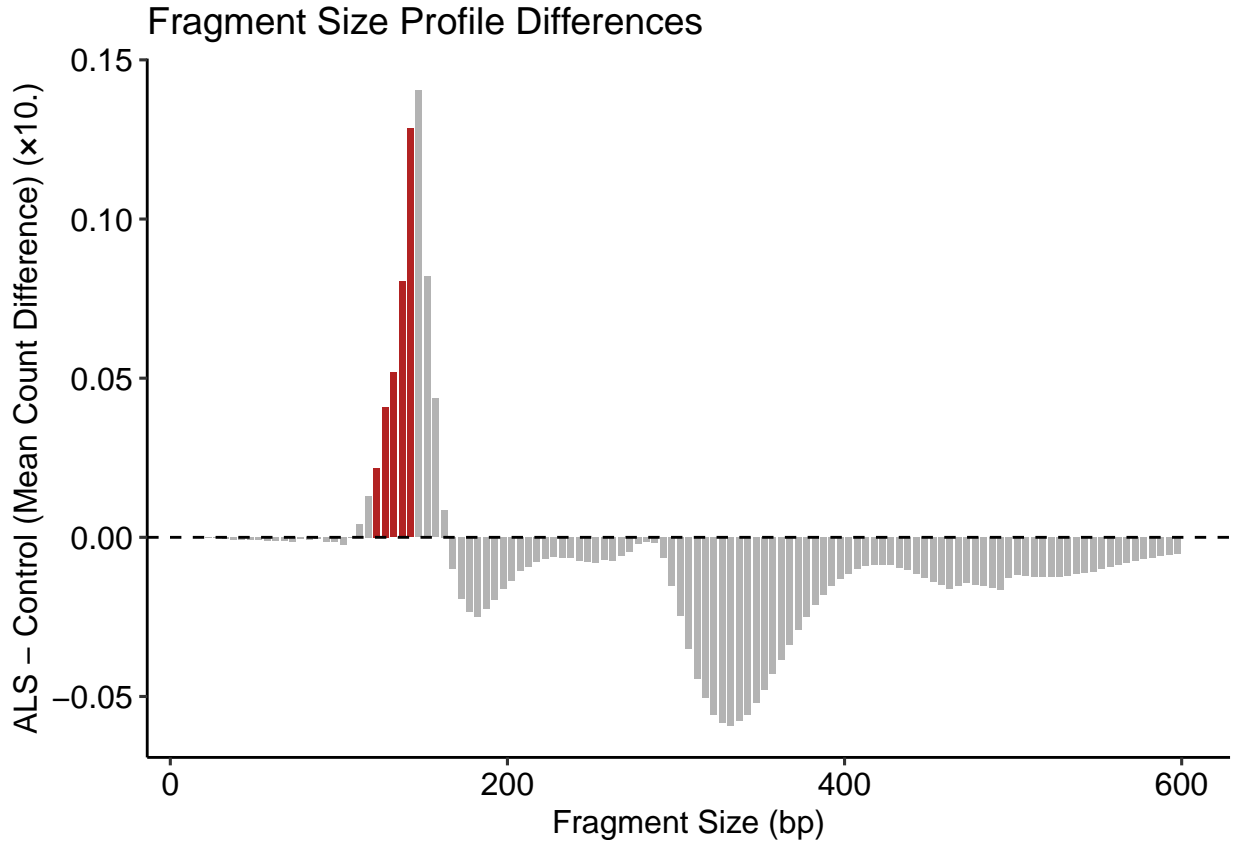


Figure 2: Differential fragment size profile

### Short-to-Long Fragment Ratio Analysis

Fragment size distributions in cfDNA are shaped by nucleosomal protection and degradation dynamics. Short fragments (~110–150 bp) typically arise from apoptotic cleavage, while longer fragments (~200–600

bp) may reflect alternative forms of cell death or reduced chromatin degradation. The relative abundance of these fragment classes can provide insight into underlying biological processes.

In this section, the abundance of short and long fragments was quantified per sample. The short-to-long fragment ratio was calculated and compared between ALS and control groups. All values were normalized per million fragments.

## Fragment Size Distribution

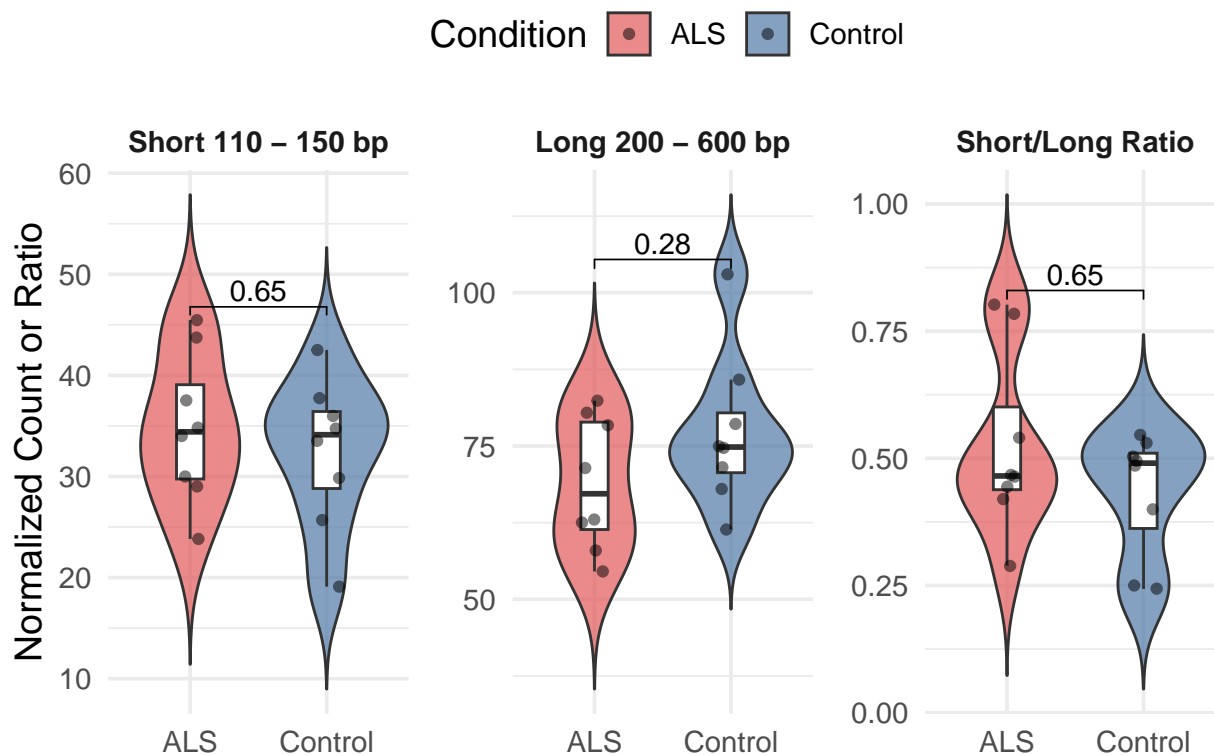


Figure 3: Short-to-long fragment ratio by condition

**Distribution of Short-to-Long Ratios** ALS samples showed a non-significant trend toward higher abundance of short fragments and reduced levels of long fragments compared to controls. Consequently, the short-to-long ratio appeared slightly elevated in the ALS group. While these tendencies are biologically plausible—potentially reflecting increased apoptotic cfDNA release—they did not reach statistical significance and should be interpreted with caution.

## 2.4 Fragmentation Differences Between Bulbar- and Spinal-Onset ALS

Within the ALS group, patients can present distinct onset phenotypes: bulbar or spinal. These phenotypes may reflect different neurodegenerative mechanisms and progression patterns. Here, we examine whether cfDNA fragmentation differs between bulbar- and spinal-onset ALS patients.

## Fragment Size Distribution in Plasma cfDNA in Bulbar vs Spinal A

Shaded regions represent  $\pm 1$  standard error

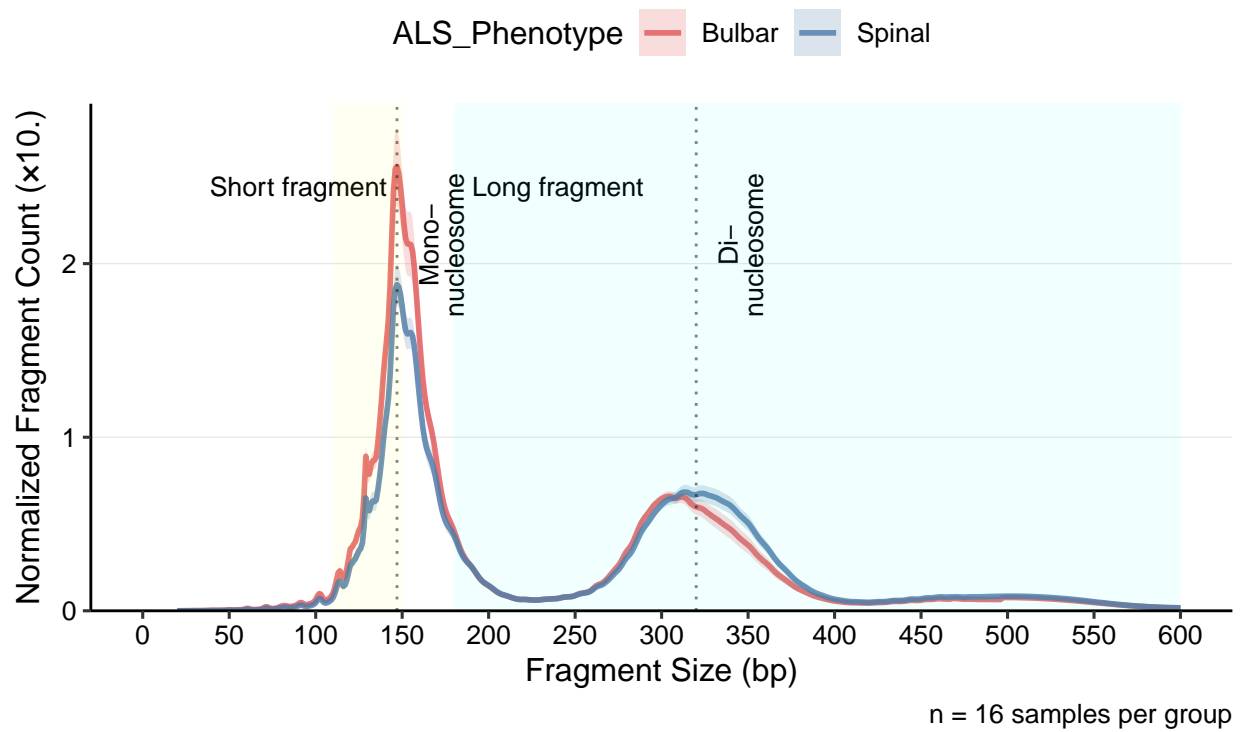


Figure 4: Fragment size distribution in bulbar- vs spinal-onset ALS

## Fragment Size Distribution by ALS Phenotype

### Statistical Comparison

Wilcoxon tests were conducted for each fragment size bin (5 bp resolution) to identify bins with significant differences between bulbar and spinal-onset ALS patients.

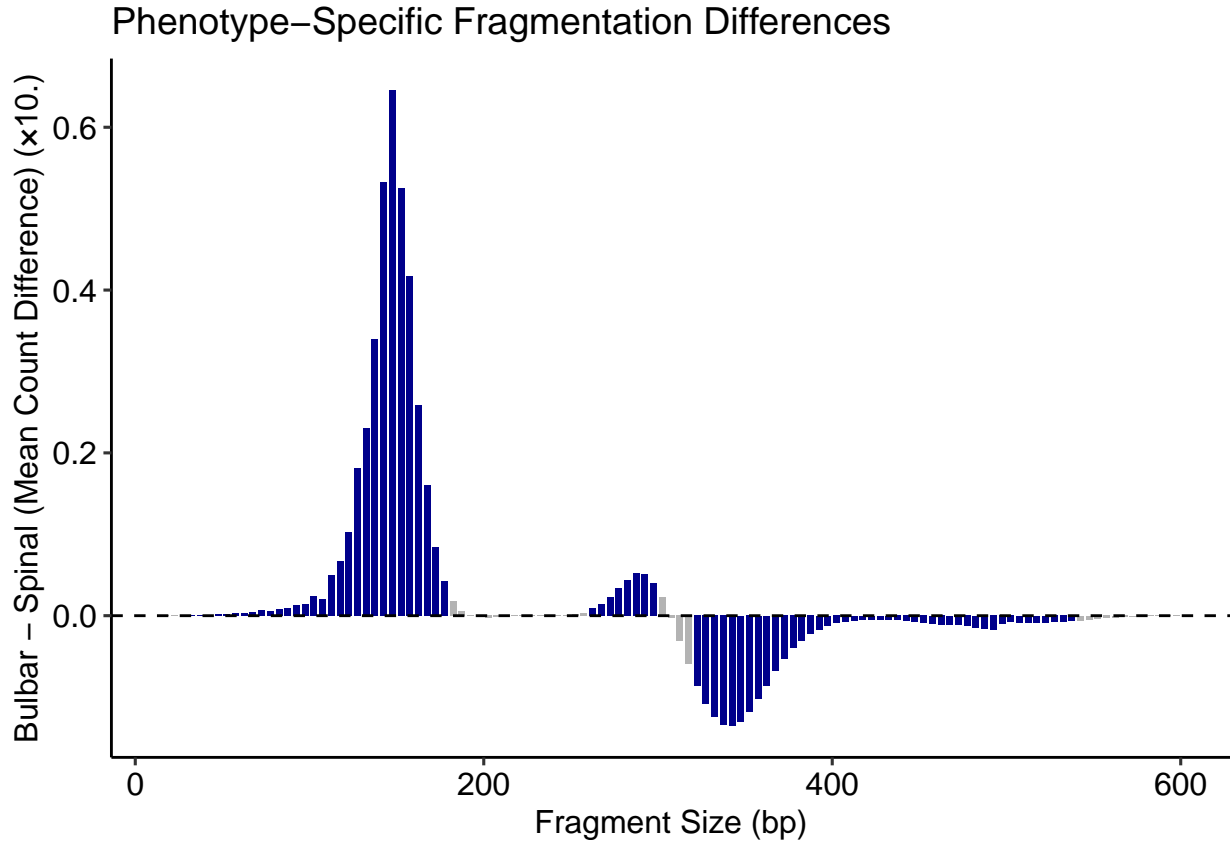
### Significant Bins

The top 10 most differential bins are presented below (sorted by nominal p-value):

Table 2: Top 10 Most Differential Fragment Size Bins

bin	mean_bulbar	mean_spinal	mean_diff	p.value	bin_start	bin_end	bin_mid
130-135	849443.50	618719.25	230724.2	0e+00	130	135	132.5
150-155	2141833.50	1616836.35	524997.2	0e+00	150	155	152.5
145-150	2480680.20	1835315.55	645364.7	0e+00	145	150	147.5
120-125	409544.10	307318.95	102225.1	0e+00	120	125	122.5
135-140	1213400.85	873341.55	340059.3	0e+00	135	140	137.5
335-340	482622.60	616887.40	-134264.8	1e-04	335	340	337.5
90-95	41544.85	28781.25	12763.6	1e-04	90	95	92.5
345-350	394811.25	524730.15	-129918.9	1e-04	345	350	347.5
340-345	439860.20	574870.65	-135010.5	1e-04	340	345	342.5
155-160	1837638.50	1421006.90	416631.6	1e-04	155	160	157.5

The following plot highlights insert size bins showing the greatest differences between bulbar-onset and spinal-onset ALS samples. Blue bars indicate bins with nominal statistical significance ( $p < 0.05$ ).



This plot illustrates the mean count differences in cfDNA fragment sizes between bulbar- and spinal-onset ALS patients. A prominent peak is observed around 150–170 bp, indicating a higher abundance of short mononucleosomal fragments in the bulbar group. In contrast, a relative depletion of longer fragments (approximately 320–400 bp) is apparent in bulbar-onset cases.

### Short-to-Long Fragment Ratio Analysis

As described in Section 2.3, the short-to-long fragment ratio was also evaluated within the ALS group to compare ALS phenotypes, aiming to explore potential differences in fragmentation patterns associated with clinical presentation.

In the comparison between ALS phenotypes, bulbar-onset samples showed a significantly higher short-to-long fragment ratio compared to spinal-onset cases ( $p = 0.029$ ). This difference was driven by both an increased abundance of short fragments ( $p = 0.057$ ) and a decreased abundance of long fragments ( $p = 0.029$ ) in the bulbar group.

These results suggest a potential shift in cfDNA fragmentation profiles associated with ALS clinical subtype. However, given the limited sample size ( $n = 4$  per group), findings should be interpreted with caution and require validation in larger cohorts.

## Summary and Interpretation

The cfDNA fragmentation analysis between ALS patients and healthy controls showed consistent nucleosomal fragmentation patterns in both groups. ALS samples exhibited a non-significant trend toward a higher abundance of short fragments and lower abundance of long fragments, resulting in an increased short-to-long fragment ratio. These trends align with the hypothesis of enhanced apoptotic cfDNA release in ALS, although no statistically significant differences were detected at the group level.

Within the ALS cohort, the comparison between bulbar- and spinal-onset phenotypes revealed more pronounced differences. Bulbar-onset patients showed significantly lower long fragment abundance and higher short-to-long ratios compared to spinal-onset cases. This suggests potential phenotype-specific differences in cfDNA release dynamics, which may reflect underlying pathophysiological processes. However, given the limited sample size ( $n = 4$  per group), these findings must be interpreted with caution and require validation in larger cohorts.

## Conclusion

This study provides a preliminary characterization of cfDNA fragmentation patterns in ALS plasma samples using EM-seq data. While no statistically significant differences were observed between ALS and control groups, phenotype-specific trends within the ALS cohort suggest that cfDNA fragmentation metrics, particularly the short-to-long ratio, may offer biologically relevant insights. These findings support further investigation in larger, independent cohorts to assess their robustness and potential utility as biomarkers in ALS.