IMPACT OF MNase DIGESTION PROFILES ON CONFORMATION ANALYSIS IN MICRO-C LIBRARIES

Key Take Homes

- Micro-C reactions representing a broad range of digestion profiles all detect more loops than restriction-enzyme-based Hi-C approaches.
- Chromatin digestion optimization is not required to generate high-quality contact matrices for a vast majority of chromatin conformation studies.
- Ultra-high-resolution studies (requiring >3 billion read pairs for sub-1 kbp contact matrices) benefit from an optimized MNase digestion profile.

Introduction

Chromatin conformation analysis plays a pivotal role in understanding gene regulation and genome organization. Micro-C has emerged as a leading assay for studying chromatin conformation due to its:

a. Superior signal-to-noise ratio compared to other Hi-C techniques.
b. Greater proportion of topologically informative read pairs.
c. Ability to achieve nucleosome-level resolution.

The success of Micro-C is thought to be dependent on efficient digestion of chromatin into predominantly mononucleosome-size fragments (ideally >40% mononucleosomes). However, achieving an optimal digestion profile can be challenging as digestion kinetics can vary based on input material. Sample specific digestion optimization can be performed, however, this is challenged by sample availability and adds additional assay time and steps.
This white paper investigates the importance of sample digestion and the impact of under digested chromatin in Micro-C libraries on chromatin conformation analysis compared to traditional restriction enzyme (RE) based Hi-C approaches.

Methods

Micro-C libraries were prepared using the Dovetail® Micro-C Kit according to the manufacturer's specifications. Traditional Hi-C approaches using single and multiple restriction enzyme reactions were also generated for comparison. Micrococcal nuclease (MNase) fragmentation was varied , employing different digestion times, to generate digestion profiles representing a range of mononucleosome content (Figure 1A).

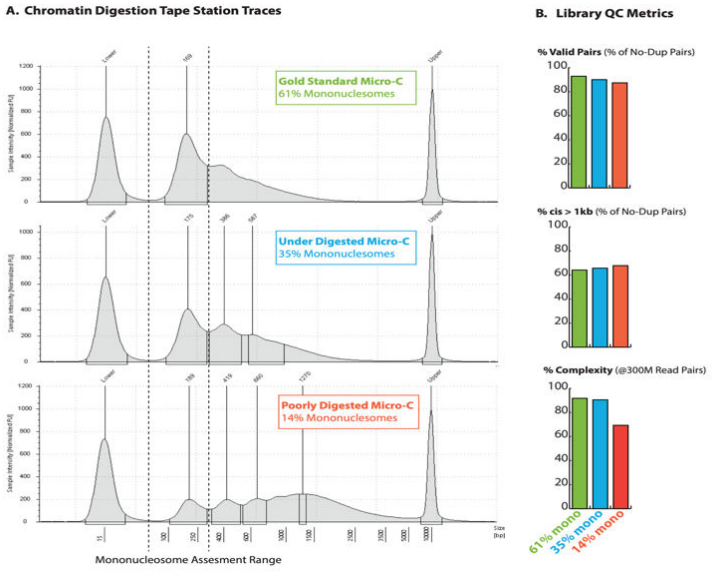

Figure 1 – Fragmentation traces and QC results. A) Tapestation digestion profiles from Stage 1 demonstrate the percentage of mononucleosome content. Optimal fragmentation, > 40% (green), under digested, 40%>X>20% (blue) and poorly digested, <20% (orange). B) Key QC metrics from the libraries are listed below.

Dovetail GENOMICS
Part of Cantata Bio, LLC

Cantata Bio
100 Enterprise Way, A101
Scotts Valley, CA 95066
831-713-4465

- Gold standard sample: 61% mononucleosomes
- Under digested sample: 35% mononucleosomes
- Poorly digested sample: 14% mononucleosomes

All libraries were sequenced to 800M total read pairs and processed following recommended guidelines[1] for primary and secondary analysis.

Results

Micro-C QC Results

All resulting Dovetail Micro-C libraries passed quality control sequencing metrics, including the percentage of no-duplicate read pairs, percentage of valid pairs, and long-range enrichment (Figure 1A and B). The only exception was the poorly digested sample, which did not meet the complexity threshold. To overcome this, an additional library was utilized to achieve the required number of no-duplicate read pairs without increasing PCR duplicates.

Chromatin Topology Analysis

All Dovetail Micro-C libraries consistently outperformed RE-based Hi-C approaches in capturing chromatin topology, including intra-TAD dynamics (Figures 2A and 2B). The contact density within TADs and loops were enriched in the Micro-C libraries compared to RE-based Hi-C, with fewer non-topology informative contacts. These results highlight the improved signal-to-noise ratio of Micro-C libraries, even when poorly-digested, compared to RE-based approaches.

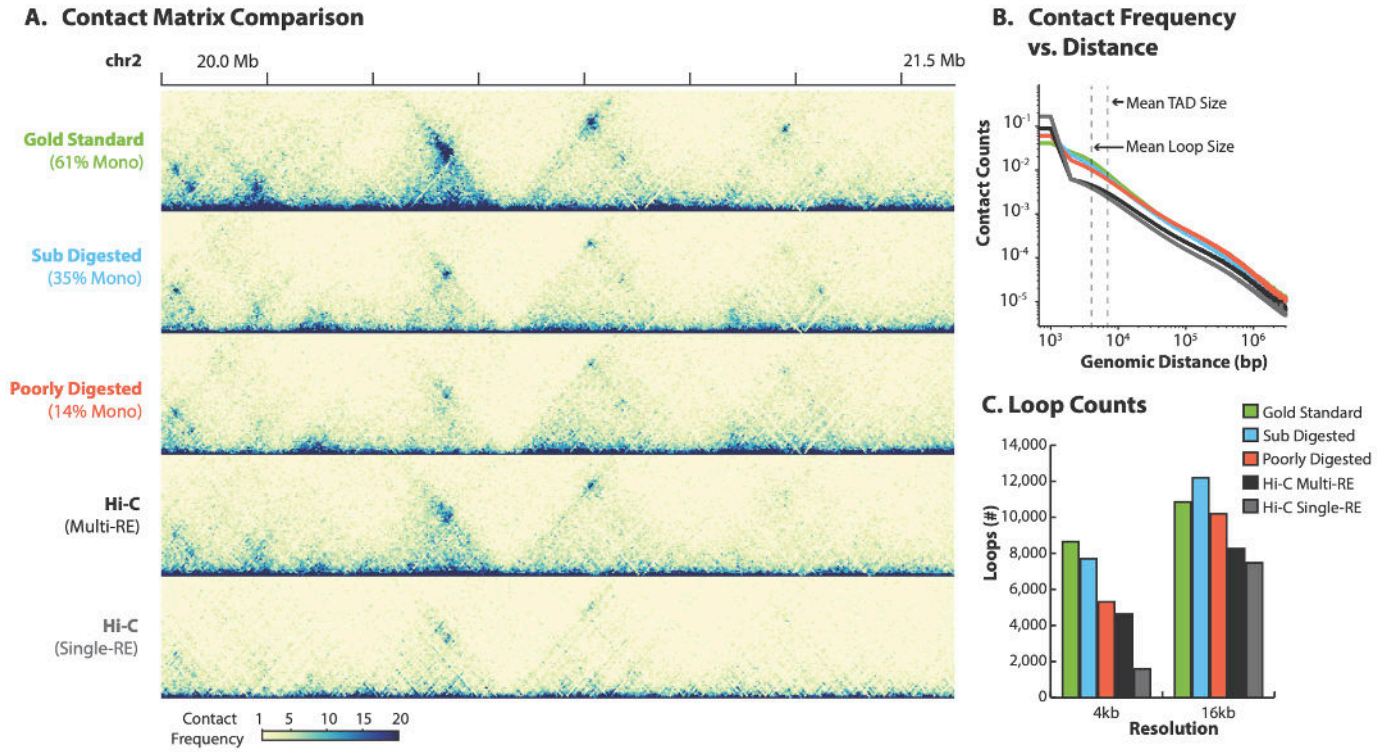

Figure 2– All Micro-C digestion profiles outperform traditional Hi-C in chromatin conformation analysis. A) Matrix tracks highlight the high quality of Micro-C libraries compared to RE-based Hi-C. Tracks were plotted at 5kb resolution using the R package plotGardener[3]. B) Contact counts as a function of distance captures the improved signal-to-noise of all Micro-C digestion profiles were generated with HiCExplorer[4] at 1kb resolution. C) TAD and loop call counts demonstrates the superior loop discovery rate inherit to all Micro-C libraries regardless of digestion profile. Loops were generated from Mustache at both 4 and 16kb resolution.

Dovetail GENOMICS
Part of Cantata Bio, LLC

Cantata Bio
100 Enterprise Way, A101
Scotts Valley, CA 95066
831-713-4465

Loop calling performance at standard Hi-C resolutions (16kb and 4kb) was not significantly impacted by mononucleosome content (Figure 2C). No significant difference between contact frequency vs. distance was discernable across the three Micro-C datasets. Moreover, both under-digested Micro-C libraries exhibited superior loop discovery rates compared to RE-based Hi-C approaches across all resolutions. These findings underscore the ability of Micro-C to call topological features at >1kb resolution, surpassing the performance of traditional Hi-C even with severely under-digested libraries.

Assessing Bias in Topology Calls

To investigate the possibility that under-digested samples may exhibit bias towards chromatin state (e.g. heterochromatin vs. euchromatin), we investigated the calling of A/B compartments in heterochromatic regions. Our findings demonstrated that A/B compartment calls showed low variability and high concordance across the different methods and digestion profiles, indicating that under-digested Micro-C libraries do not introduce significant bias (Figure 3).

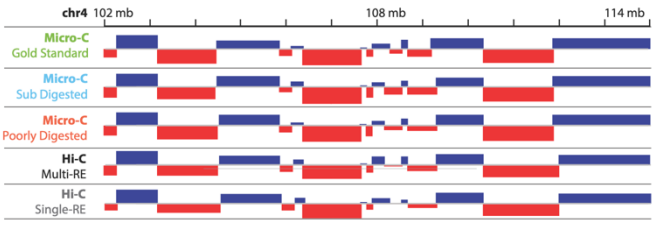

Figure 3 – MNase digestion profiles does not introduction bias in topology calls. A/B compartments in a dynamic region of chr4 demonstate that A/B compartment calls show low variability and high concordance across the different methods and digestion profiles. Compartments calls were generated from the compartment function in FanC[5], A compartments, or euchromatin, are illustrated by positive values in blue, B compartments which indicate heterochromatin are captured in negative values in red blocks.

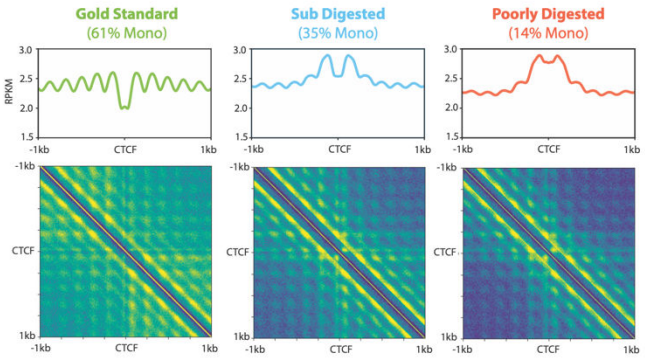

Figure 4 – Nucleosome Positioning. The ability for the Micro-C libraries to position nucleosomes in the linear enrichment was assessed through deepTools[2] compute matrix with +/- 1kb of CTCF binding sites. Aggregated contact matrices (1 bp resolution) of the same window with respect to CTCF bound DNA demonstrates the impact of chromatin digestion on the ability to map nucleosome-nucleosome contacts.

Ultra-high Resolution and Micro-C Nucleosome Positioning

Linear positioning and matrix positioning of nucleosomes were assessed in all Micro-C libraries. We observed the efficiency of nucleosome positioning decreased with under-digestion of chromatin, as evidenced by diminished coverage signal intensity and reduced nucleosome-nucleosome contacts (Figure 4). This and other datasets generated at Cantata Bio strongly support that ultra-high resolution analyses, that is the detection of topological features at sub-1kb resolutions, benefits from optimally digested libraries in which the mononucleosome content exceeds 40%.

Conclusions

Optimally digested samples were required to enable ultra-high resolution and maximize nucleosome phasing, information unique to the Micro-C chemistry. Thus, researchers aiming to unlock the full potential of Micro-C assays are encouraged to invest time in optimizing the digestion profile prior to proximity ligation.

On the other hand, under-digested Micro-C libraries perform exceptionally well in standard Hi-C applications, such as TAD

Dovetail GENOMICS
Part of Cantata Bio, LLC

Cantata Bio
100 Enterprise Way, A101
Scotts Valley, CA 95066
831-713-4465

and loop calling, outperforming traditional Hi-C approaches for all digestion profiles tested. Researchers seeking to generate high-quality contact matrices at 5 and 10kb resolutions can leverage the high signal-to-noise features even when working with Micro-C libraries containing only 14% mono-nucleosome content, surpassing the performance of traditional Hi-C approaches. However, keep in mind that severely underdigested libraries (our poorly digested profile, displaying only 14% mononucleosome content) can also be utilized but may require additional libraries or samples to overcome complexity limitations. Furthermore, this study demonstrated that fragmentation profiles do not introduce bias towards heterochromatin or euchromatin.

For standard loop and TAD calling experiments, our data indicates that optimizing MNase digestion is not a requirement.

References:

[1] https://micro-c.readthedocs.io/en/latest/
[2] https://deeptools.readthedocs.io/en/develop/
[3] https://phanstiellab.github.io/plotgardener/articles/introduction_to_plotgardener.html
[4] https://hicexplorer.readthedocs.io/en/latest/
[5] https://fan-c.readthedocs.io/en/latest/

Cantata Bio
100 Enterprise Way, A101
Scotts Valley, CA 95066
831-713-4465