

Product Highlights:

- Highest resolution Hi-C data
- Uniform coverage across the genome
- Detect nucleosome-to-nucleosome interactions

Introduction

Chromatin conformation capture (3C) methods enable the study of higher-order features of chromatin organization such as chromatin loops, topologically associating domains (TADs), and A/B compartments. In particular, Hi-C has been widely adopted to assess the frequency of interaction between distant genomic locations up to megabases apart on a genome-wide scale. The resolution at which the interaction point can be mapped is reliant on two factors associated with the chromatin fragmentation approach:

- Distance between fragmentation sites
- Uniformity of fragmentation length

While restriction enzymes (RE) are widely used for chromatin fragmentation in Hi-C protocols, they are not uniformly distributed across the genome thereby generating fragments of highly variable length. This non-uniformity contributes to variable coverage and reduces achievable chromatin contact resolution.

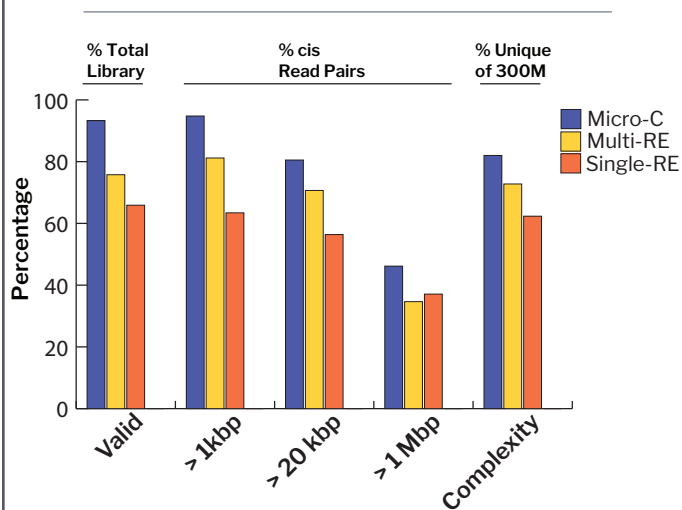
Dovetail Genomics has employed micrococcal nuclease (MNase) in our Dovetail™ Micro-C Proximity Ligation Assay. The resulting highly uniform, short fragments enable nucleosome-level resolution of chromatin contacts, a theoretical resolution maximum. Here, we compare the ultra-high-resolution chromatin conformation information generated by Dovetail™ Micro-C to RE-based Hi-C data and highlight its nucleosome-resolution capabilities.

Dovetail Micro-C libraries are enriched for desirable Hi-C properties

MNase enzyme possesses both sequence-independent endonuclease and exonuclease

Figure 1 – Dovetail™ Micro-C libraries contain proximity ligation properties and are enriched in long-range informative reads.

Dovetail™ Micro-C (blue) is compared to single (yellow) and multi (orange) RE-based Hi-C generated from 1 million GM12878 cells. All libraries were subsampled to 1 million read-pairs and processed through the Dovetail Genomics QC pipeline. Library characteristics are compared across libraries. Percentages of valid reads (trans + cis > 1 kbp) of the total library, cis reads with insert sizes > 1 kbp, > 20 kbp, and > 1 Mbp are plotted. Library complexity, as percent unique reads at 300 million read pairs, is extrapolated using the preseq library complexity tool.



activities so generated fragments are of mononucleosome length (146bp). The proximity ligation portion of the protocol is optimized to maximize long-range interactions. Our Dovetail™ Micro-C libraries produced > 95% of valid read-pairs with 75% unique molecules at 300 million reads. Notably, this is significantly higher than single- or multi-RE approaches (**Figure 1**). Moreover, considering all cis-chromosome contacts, 92% were > 1 kbp and 42% > 1 Mbp. These data indicate that the library characteristics desired in high quality Hi-C libraries are significantly improved in the Dovetail™ Micro-C Assay.

Improved calling of topological features

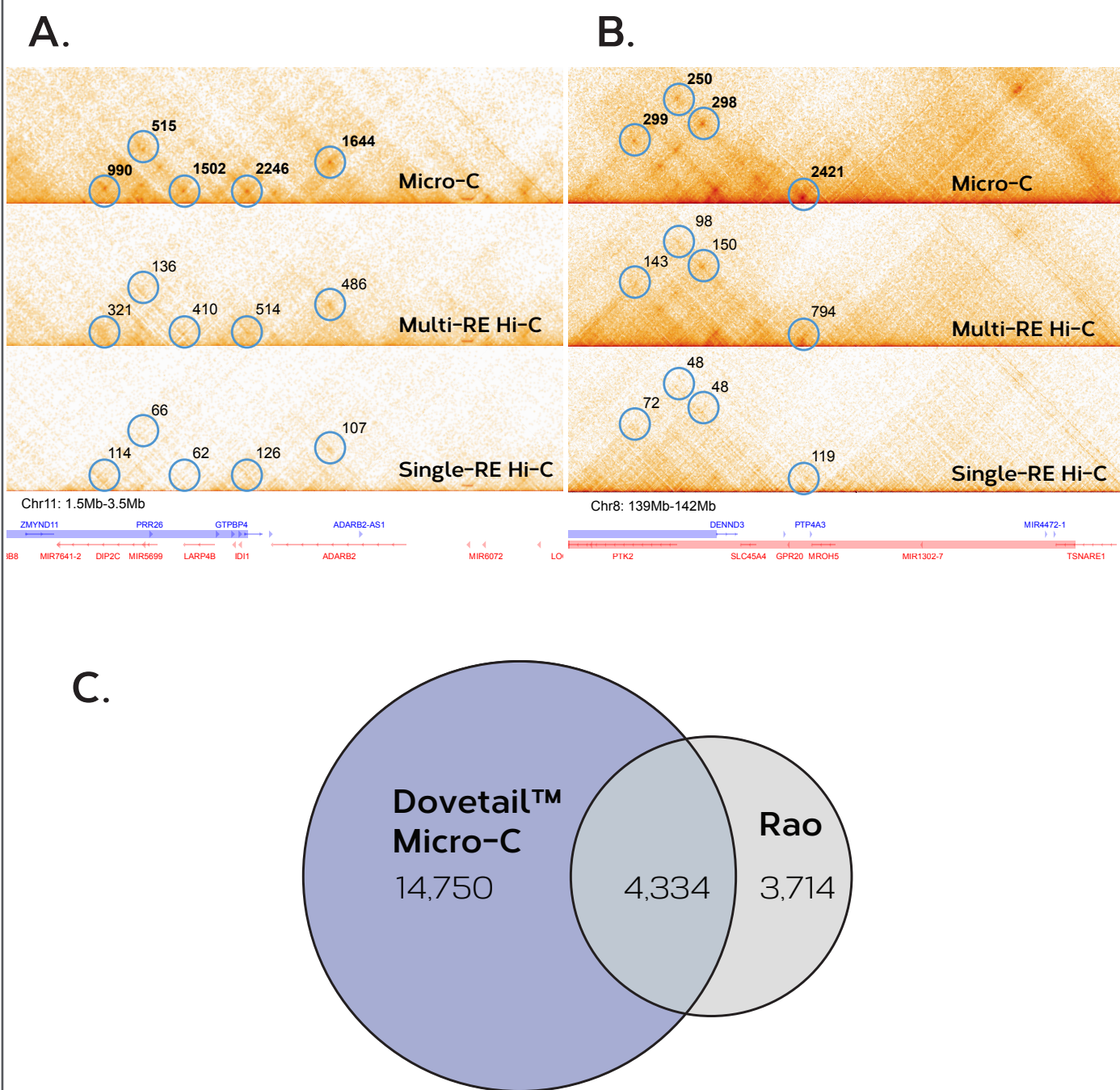
The ability to detect higher-order features, such as chromatin loops, in proximity ligation data is dependent on enriching long-range informative reads to capture chromatin interaction frequency. The increased chromosome

conformation informative reads combined with ultra-high-resolution improves loop calling compared to RE-based methods (**Figure 2A**), as demonstrated for previously enumerated loops (*Rao et al.*, 2014). Furthermore, *de novo* loop calling on the GM12878 cell line using Juicer (*Durand et al.*) demonstrates that the Dovetail

Micro-C based contact matrix contained 19,084 chromatin loops with 4,334 overlapping with *Rao et al.* Overlapping loop calls between Dovetail Micro-C and *Rao et al.* are similar in number to other such comparisons.

Figure 2 – Dovetail™ Micro-C contact matrices display improved signal-to-noise enabling superior detection of higher-order features of chromatin organization.

A&B) Contact matrix images from Dovetail™ Micro-C, single-, and multi-RE Hi-C are shown at 4 kbp resolution from GM12878 cell lines. Known loops from *Rao et al.* are outlined with circles and the number of supporting reads from a 800 M total read depth for each library type is indicated. **C)** Venn diagram comparing Juicer detected loops between *Rao et al.*, 2014 and Dovetail Micro-C.



Dovetail Micro-C uniquely captures nucleosome positioning

Chromatin digested with MNase reveals a genome-wide nucleosome map that is visible in the Dovetail Micro-C libraries. To highlight this quality, a metagene analysis of high occupancy CTCF sites was performed. The result displays coverage periodicity relative to the CTCF anchor (**Figure 3A**) in which peaks indicate DNA that is protected by the nucleosome and troughs represent intervening DNA that is accessible to MNase digestion. This oscillation occurs at a frequency of ~146 bp (the length of DNA wrapped around a mono-nucleosome). This feature is unique to Micro-C data as it is absent from RE-based approaches.

The regularly spaced MNase fragmentation sites resulting from the naturally occurring nucleosome distribution produces uniform read coverage across the genome. To understand this further, we focused on an 8 kbp CTCF-mediated region on chr1. Micro-C genomic coverage across this region displays both uniform sequence distribution and the ability to map nucleosome position (**Figure 3B**).

In contrast, the single-RE Hi-C dataset for the same region exhibits a coverage drop off, likely due to an area of low RE-site density. As a result, chromatin contacts are only mappable at low resolution. These data demonstrate that Dovetail Micro-C data not only identifies nucleosomes, but also provides more uniformly distributed fragmentation sites over other Hi-C approaches.

Dovetail Micro-C data contains nucleosome-to-nucleosome contacts

The combined genome-wide nucleosome positioning and ultra-resolution chromosome topology enabled by Dovetail Micro-C facilitates mapping from nucleosome-to-nucleosome chromatin contacts. To demonstrate this feature, Hi-C contacts within 1,000 bp of all strong CTCF sites were aggregated at a 1 bp contact map resolution (**Figure 4**). This reveals both nucleosome phasing around CTCF (as a 2-D version of **Figure 3A**) and nucleosome-to-nucleosome interactions (observed off the diagonal). This capability is unique to Micro-C data as RE-based Hi-C lacks the resolution

Figure 3 – Dovetail™ Micro-C libraries are enriched in nucleosome-protected fragments enabling genome-wide nucleosome resolution.

A) Metagene analysis of relative coverage across ~20,000 high occupancy CTCF regions in GM12878 cells. Dovetail™ Micro-C (blue) is compared to single (yellow) and multi (orange) RE-based Hi-C. For the Micro-C dataset, peaks in coverage represent DNA protected by the nucleosome and troughs are DNA accessible to MNase during chromatin fragmentation. These data are the average of normalized (by read depth) coverage. **B)** Genomic coverage at a CTCF occupied site in chr1. Libraries were subsampled to 40X coverage and plotted along an 8 kbp region associated with the CTCF occupied gene, ACAP3. In contrast to RE-based approaches, Micro-C libraries display nucleosome periodicity and fragment length uniformity. Regions of low coverage in the RE-based Hi-C coverage are regions with low RE-site density.

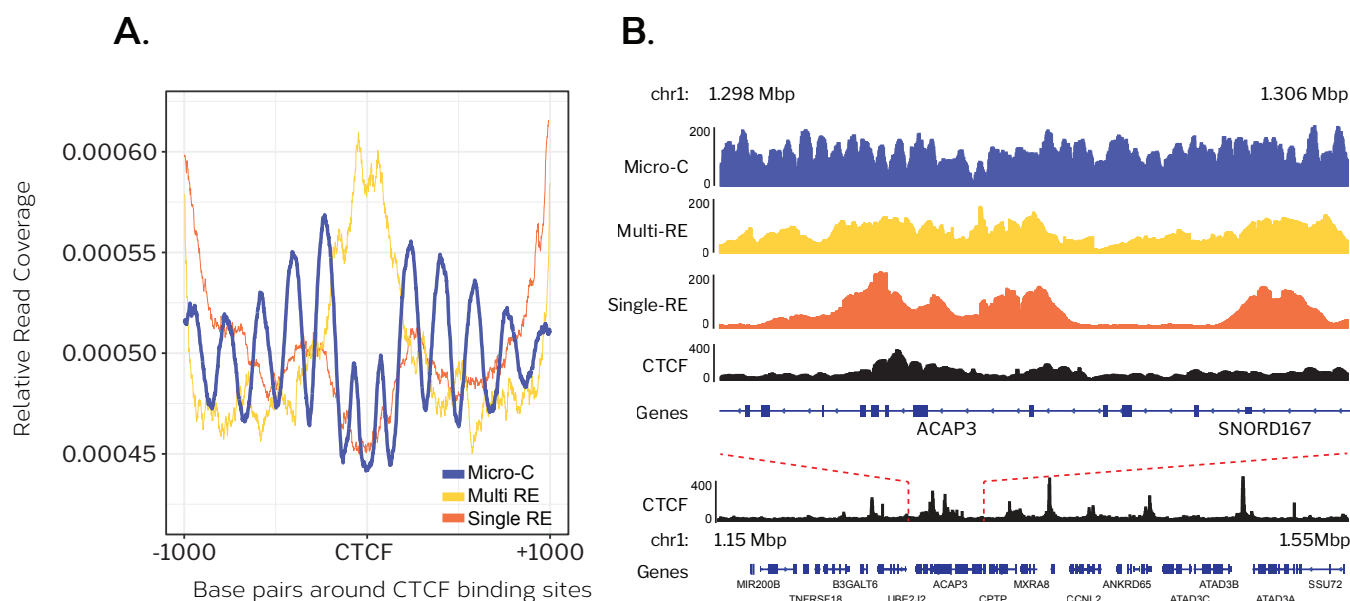
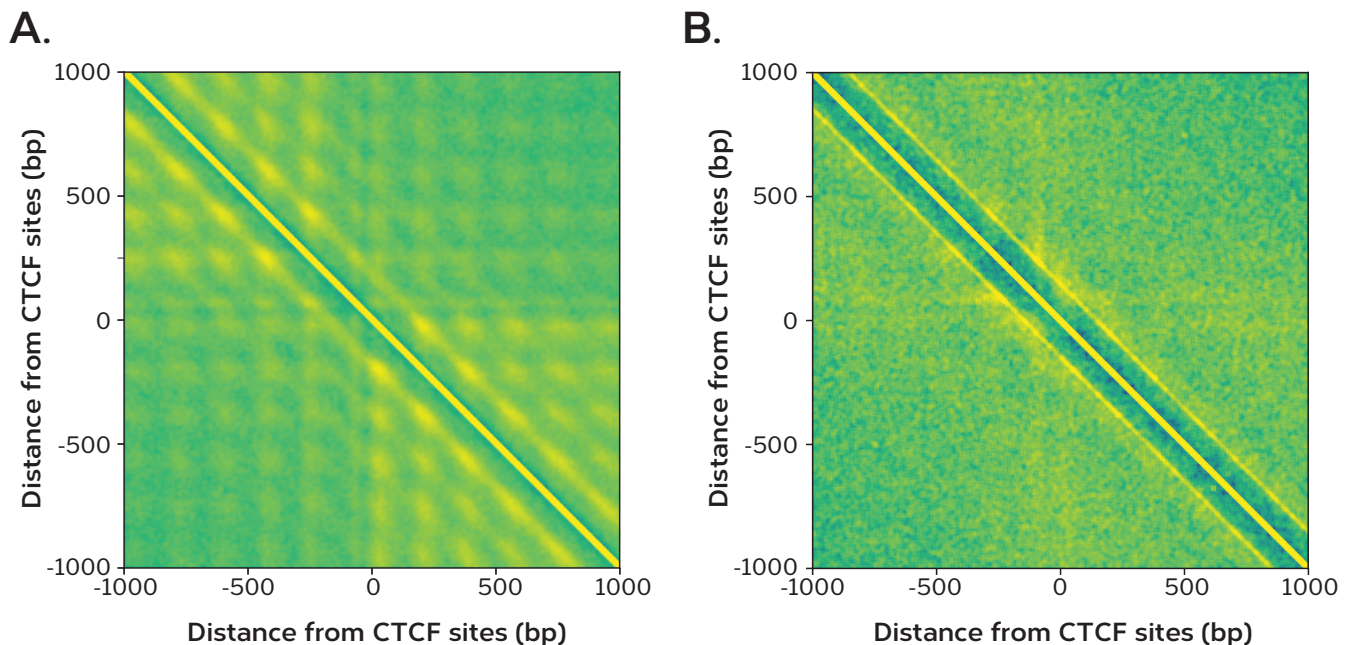


Figure 4 – Nucleosome-level resolution of chromatin contacts generated with Dovetail™ Micro-C.

Hi-C contacts within 1,000 bp of all strong CTCF sites were aggregated at a 1 bp contact map resolution for Micro-C (A) and Multi-RE Hi-C (B). This reveals both nucleosome phasing around CTCF (as a 2-D version of Figure 3A) and nucleosome-to-nucleosome interactions (observed off the diagonal) in the Micro-C library that is not present in RE-based Hi-C. The decrease in interaction across the CTCF regions indicate that CTCF acts as an insulator preventing chromatin interactions. This matrix was generated from 1.3 B paired-end Dovetail Micro-C reads from the human cell line GM12878.



to capture nucleosome interactions. This highlights that Dovetail™ Micro-C can be used to detect and explore higher-order features at nucleosome-level resolution around a genomic region of interest or, if coupled with the Dovetail™ HiChIP MNase Assay, for a protein factor of interest.

Summary

Here we showcase Dovetail Micro-C, a proximity-ligation assay with superior genomic resolution of chromatin contacts and higher order structures involved in chromatin topology. Dovetail Micro-C has been optimized to produce the highest signal-to-noise data with both enrichment of long-range informative reads and nucleosome protected fragments. These unique properties enable nucleosome-level resolution of chromatin contacts.

References

- Durand, NC *et al.* (2016) Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* 3(1): 95–98. doi:10.1016/j.cels.2016.07.002.
- Rao, SSP *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 159:1665-1680. DOI:https://doi.org/10.1016/j.cell.2014.11.021