

## 製品ハイライト:

- 最高解像度の Hi-C データ
- ゲノム全体にわたる均一なカバレッジ
- ヌクレオソーム間の相互作用を検出

## はじめに

クロマチンコンフォメーションキャプチャ(3C)法は、クロマチンループ、トポロジカルアソシエイトドメイン(TAD)、A/Bコンパートメントなど、クロマチンの高次構造の研究を可能にします。特にHi-Cは、ゲノム全体のスケールで、最大でメガベースまで離れたゲノムの位置間の相互作用の頻度を評価するために広く採用されています。相互作用点をマッピングできる解像度は、クロマチン断片化アプローチに関連する2つの要因に依存しています。

- フラグメント間の距離
- フラグメントの長さの均一性

**制限酵素(RE)**はHi-Cプロトコルのクロマチン断片化に広く使用されていますが、サイトはゲノム全体に均一に分布していないため、非常に不均一な断片を生成します。この不均一性は、カバレッジの偏りに影響し、クロマチンコンタクトの解像度を低下させます。

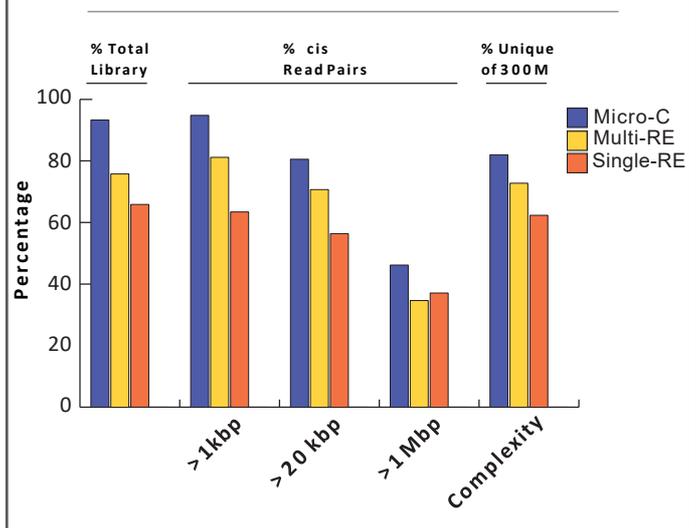
Dovetail Genomicsは、Dovetail™ Micro-C Proximity Ligation Assayでクロマチン断片化に **Micrococcal Nuclease (MNase)**を採用しています。結果として得られる非常に均一な短い断片は、理論上の最大分解能であるクロマチンコンタクトのモノヌクレオソームレベルの高解像度を可能にしました。ここでは、Dovetail™ Micro-Cによって生成された超高解像度クロマチンコンフォメーション情報をREベースのHi-Cデータと比較し、そのヌクレオソームの高分解能を明らかにしています。

## Dovetail Micro-Cライブラリは、望ましいHi-C特性が豊富です

MNaseは、配列に依存しないエンドヌクレアーゼと

図1 - Dovetail™ Micro-Cライブラリには高品質な近接ライゲーション特性があり、有効な長距離情報リードが豊富に含まれています。

100万個のGM12878細胞から生成されたDovetail™ Micro-C(青色)と、マルチ(黄色)およびシングル(オレンジ)REベースのHi-Cを比較しています。すべてのライブラリは1Mリードペアにサブサンプリングされ、Dovetail Genomics QCパイプラインで処理されました。ライブラリ全体の有効リードの割合(trans+cis>1kbp)、インサートサイズが1kbp以上、20kbp以上、1Mbp以上のcisリードをプロットしています。ライブラリの複雑さは、300Mリードペアにおけるユニークリードの割合として、Preseq library complexity toolを使用して推定したものです。



エキソヌクレアーゼの両方の活性を持っているため、生成されるフラグメントはモノヌクレオソーム長(146bp)です。プロトコルの近接ライゲーション部分は、長距離の相互作用を最大化するように最適化されています。Dovetail™ Micro-Cライブラリは、3億リードで95%以上の有効リードペアを生成し、75%のユニークな分子を生成しました。これはシングルまたはマルチREアプローチよりも大幅に高い数値です(図1)。さらに、すべてのシス染色体コンタクトを考慮すると、92%が>1kbp、42%が>1Mbpでした。これらのデータは、高品質なHi-Cライブラリに求められるライブラリ特性が、Dovetail™ Micro-Cアッセイで大幅に改善されていることを示しています。

## トポロジー構造の検出力の向上

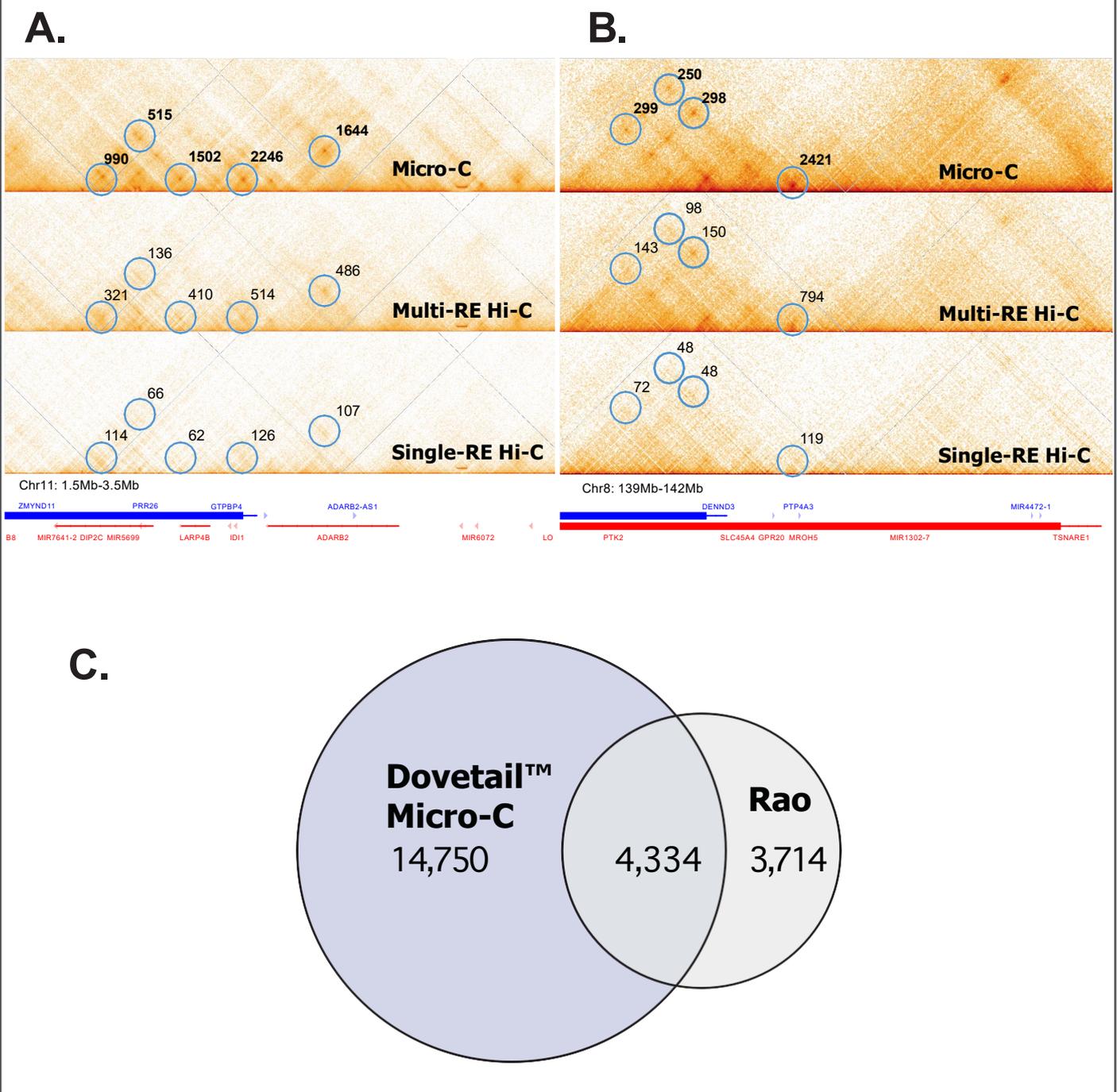
近接ライゲーションデータからクロマチンループなどの高次構造を検出できるかどうかは、クロマチン相互作用の頻度をキャプチャするための長距離情報リードのカバレッジに依存しています。

超高解像度と組み合わせられた染色体コンフォメーションの有効リードの増加は、REベースの方法(図2A)と比較してループ検出力を向上することが、以前に報告されているループ(Rao et al., 2014)により実証されました。Juicer(Durand et al., 2016)を用いたGM12878細胞株のde novoループ検出では、

Dovetail™ Micro-Cベースのコンタクトマトリクスが19,084のクロマチンループを検出し、そのうちの4,334がRaoらの報告とオーバーラップすることを示しています。

図 2 - Dovetail™ Micro-Cのコンタクトマトリクスは、S/N比を向上させ、クロマチン組織の高次構造の優れた検出を可能にします。

A&B) Dovetail™ Micro-C、シングルRE、マルチRE Hi-Cからのコンタクトマトリクス画像をGM12878細胞株からの4kbp分解能で示しています。Raoらの報告した既知のループを丸で囲み、各ライブラリタイプの800Mリードデプスからのサポートリード数を示しています。C) Rao et al., 2014とDovetail Micro-Cの間でJuicerが検出したループを比較するベン図。



## Dovetail™ Micro-Cは、ヌクレオソームのポジションを独自にキャプチャします

MNaseで消化されたクロマチンは、ゲノム全体のヌクレオソームマップを明らかにし、それはDovetail™ Micro-Cライブラリで可視化されています。これを検証するために、高占有CTCF領域のメタジーン解析を行いました。その結果、ピークはヌクレオソームによって保護されているDNA、トラフはMNaseがアクセス可能なDNAを示し、CTCFアンカー(図3A)に対するカバレッジの周期性が示されます。この周期性は、146bp(モノヌクレオソームに巻かれたDNAの長さ)で発生し、REベースのアプローチでは見られない、Micro-Cデータに特有のものです。

自然に発生するヌクレオソームの分布に起因するMNaseの断片化サイトは一定の間隔で配置されているため、ゲノム全体で均一なリードカバレッジが得られます。このことをさらに理解するために、私たちは1番染色体上の8kbpのCTCFが介在する領域に着目しました。この領域のMicro-Cゲノムカバレッジは、均一な配列分布とヌクレオソーム位置のマッピング能力の両方を示しています(図3B)。

対照的に、同じ領域のシングルRE Hi-Cデータセットでは、REサイトの密度が低い領域のため、カバレッジが低下しています。その結果、クロマチンコンタクトは低い分解能でしかマッピングできませんでした。これらのデータは、Dovetail™ Micro-Cデータがヌクレオソームを同定するだけでなく、他のHi-Cアプローチよりもフラグメントサイトがより均一に分布していることを示しています。

## Dovetail™ Micro-Cデータには、ヌクレオソーム-ヌクレオソーム間のコンタクト情報が含まれています

Dovetail™ Micro-Cは、ゲノム全体のヌクレオソーム位置と高解像度クロモソームトポロジーを組み合わせることで、ヌクレオソームからヌクレオソームへのクロマチンコンタクトのマッピングを可能にしています。これを検証するために、すべての強いCTCFサイトから1,000bp以内のHi-Cコンタクトを1bpの解像度で集約しました(図4)。これにより、CTCF周辺のヌクレオソームのフェージング(図3Aの2D版として)とヌクレオソーム-ヌクレオソーム間の相互作用の両方を明らかにしています。

図3 - Dovetail™ Micro-Cライブラリには、ヌクレオソームで保護されたフラグメントが豊富に含まれており、ゲノムワイドなヌクレオソーム分解能を実現します。

A) GM12878細胞における約20,000の高占有CTCF領域の相対カバレッジのメタジーン解析。Dovetail™ Micro-C(青色)と、マルチ(黄色)およびシングル(オレンジ色)REベースのHi-Cを比較しています。Micro-Cデータセットでは、カバレッジのピークはヌクレオソームで保護されたDNAを表し、トラフはクロマチン断片化の際にMNaseがアクセス可能なDNAを表しています。これらのデータは、(リードデプスで)正規化されたカバレッジの平均値です。B) 1番染色体のCTCF占有部位でのゲノムカバレッジ。ライブラリを40xカバレッジにサブサンプリングし、CTCF占有遺伝子ACAP3に関連する8kbpの領域に沿ってプロットしたものです。REベースのアプローチとは対照的に、Micro-Cライブラリはヌクレオソームの周期性とフラグメント長の均一性を示しています。REベースのHi-Cカバレッジの低い領域は、REサイト密度が低い領域です。

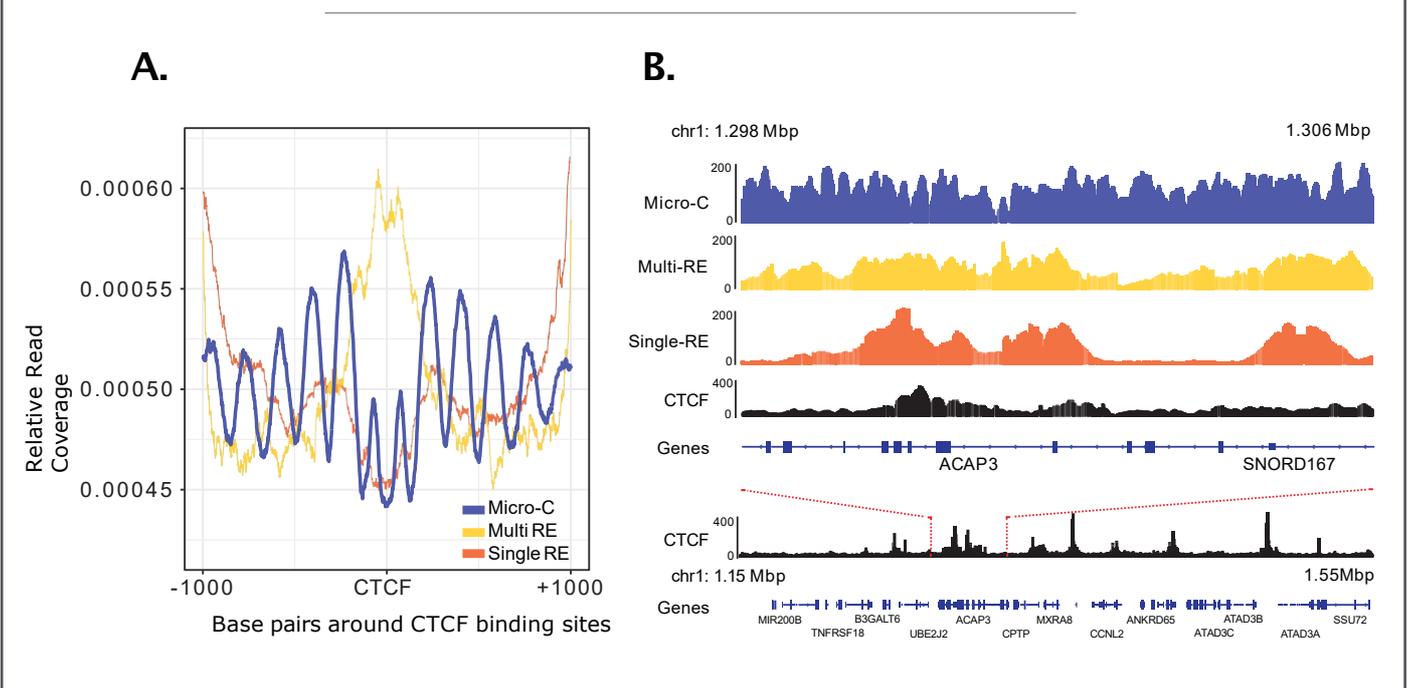
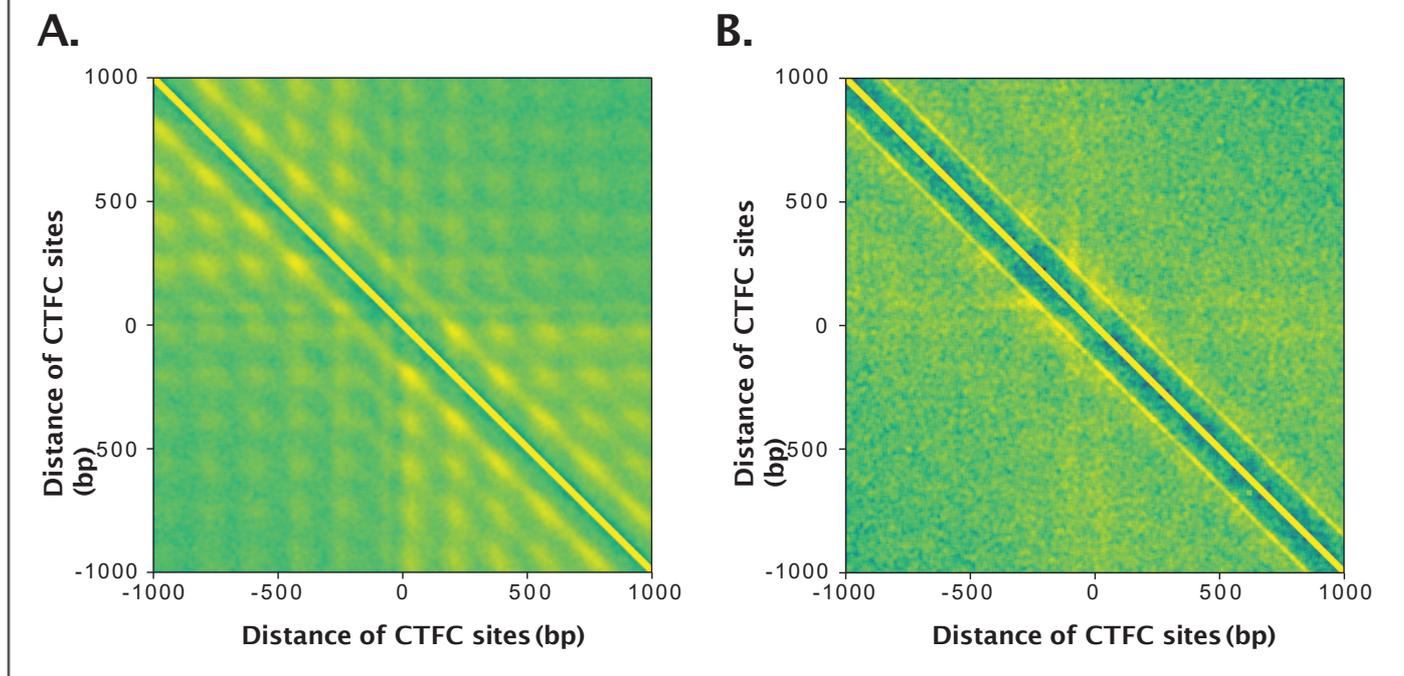


図 4 - Dovetail™ Micro-Cで生成されたクロマチンコンタクトのヌクレオソームレベルの分解能。

Micro-C(A)とマルチRE Hi-C(B)について、すべての強いCTCFサイトから1,000bp以内のHi-Cコンタクトを1bpのコンタクトマップ分解能で集約しました。Micro-Cライブラリでは、CTCFの周りのヌクレオソームのフェージング(図3Aの2D版として)とヌクレオソーム-ヌクレオソーム間の相互作用(対角線から外れて観測される)の両方を観測できますが、REベースのHi-Cでは観測できません。CTCF領域にわたる相互作用の減少は、CTCFがクロマチン相互作用を妨げるインシュレーターとして作用することを示しています。このマトリクスは、ヒト細胞株GM12878からの1.3BペアエンドDovetail™ Micro-Cリードから作成されたものです。



REベースのHi-Cにはヌクレオソームの相互作用をキャプチャできる解像度がないたため、この特徴はMicro-Cデータに固有のもので、これは、Dovetail™ Micro-Cを使用して、目的のゲノム領域周辺のヌクレオソームレベルの解像度で、またはDovetail™ HiChIP MNaseアッセイと組み合わせ、目的のタンパク質因子の高次構造を検出および探索できることを示しています。これにより、Dovetail™ Micro-Cは、目的のゲノム領域周辺をヌクレオソームレベルの解像度で、またはDovetail™ HiChIP MNaseと組み合わせることで、目的タンパク質因子の高次構造を検出し、探索するために使用できることが明らかになりました。

## まとめ

ここでは、クロマチンコンタクトとクロマチントポロジーに關与する高次構造のゲノム分解能に優れた近接ライゲーションアッセイであるDovetail™ Micro-Cを紹介し、Dovetail™ Micro-Cは、長距離コンタクト情報リードとヌクレオソームが保持されたフラグメントの両方がエンリッチされ、最高のS/N比データが得られるように最適化されています。これらのユニークな特性により、ヌクレオソームレベルでのクロマチンコンタクトの分解能を実現しています。

## 参考文献

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