

製品ハイライト:

- ChIP-seqとHi-Cデータを1つのライブラリにまとめてキャプチャ
- 8種類の抗体で検証済み
- モノヌクレオソームレベルの最高解像度でクロマチン相互作用をマッピング

はじめに

タンパク質因子に媒介される3次元(3D)クロマチンコンフォメーションは、遺伝子制御に寄与することが知られています。クロマチンコンフォメーションキャプチャー(3C)法は、ゲノムトポロジーの高次構造を解析するために一般的に用いられており、メガベース(Mbp)離れた遠い位置の遺伝子座がどのように共調節されているかを説明するのに役立ちます。今日では、ゲノムの3次元構造を媒介するタンパク質因子を理解するために、ChIP-seqはゲノム全体のDNA結合因子部位のマッピングに欠かせないツールとなっています。しかし、ChIP-seqは全ゲノムのアプローチでタンパク質結合部位の新規発見を可能にしていますが、2つ以上のDNA結合因子のタンパク質-タンパク質相互作用によって促進される遠い位置の相互作用については、ほとんど観測することができません。

Dovetail™ HiChIP MNaseキットは、ChIP-seqとHi-Cの利点を組み合わせたもので、標準的なイルミナのペアエンドシーケンシングを用いて長距離の相互作用を捕捉する近接ライゲーション法をさらに発展させたものです。目的のDNA結合因子をエンリッチすることで、結果として得られるライブラリは、従来のHi-Cアプローチよりも低いシーケンシング深度でタンパク質のトポロジカルな特徴を表示することができます。

サンプルからシーケンシングライブラリまでのシンプルな3日間のワークフロー

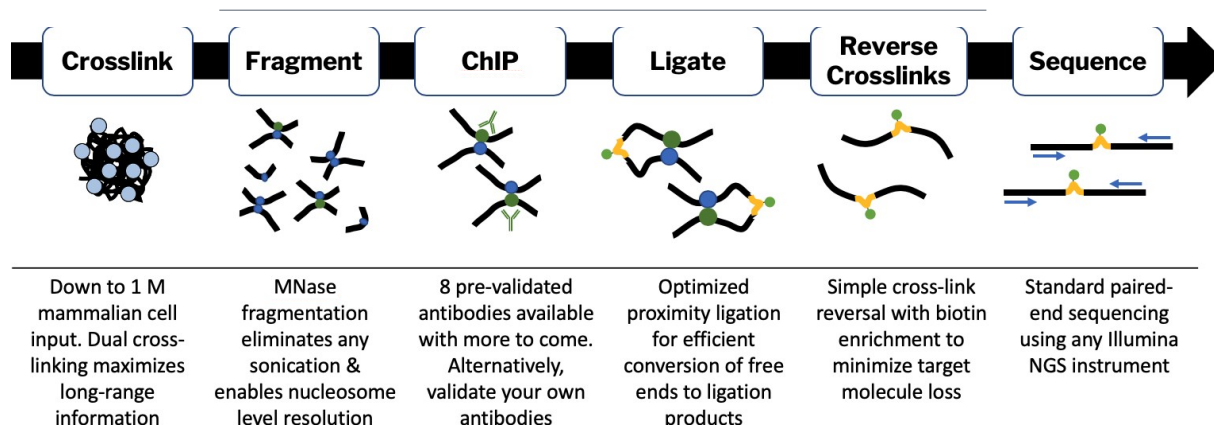
Dovetail HiChIPワークフロー(図1)は、他の方法に比べていくつもの優れた点があります。

- クロマチンの酵素的かつ均一な断片化
- 近接ライゲーションによる長距離情報のエンリッチ
- 100万細胞インプットまでのプロトコルの高い堅牢性と再現性
- 一般的なタンパク質に対する検証済み抗体のリストと使用ガイドライン

Micrococcal nuclease (MNase)を採用したDovetail™ Micro-Cと、Dovetail™ HiChIPのユニークな組み合わせにより、クロマチンを繊細な**超音波処理を用いず**に、シーケンスバイアスなく均一に断片化することが可能となり、クロマチン相互作用を最高の解像度(モノヌクレオソームレベル)で提供することができます。

図1 - Dovetail™ HiChIP MNase Kit 3日間のワークフローの概要。

10⁶から10⁷の細胞から開始します。1日目は、クロスリンク、断片化、ChIP抗体を添加してO/Nインキュベーション、で構成されています。2日目には、クロマチン免疫沈降と近接ライゲーションが含まれており、ライゲーション接合部にビオチン化されたブリッジを組み込むことでキメラ分子をマークします。3日目に、クロスリンクを解放し、キメラ分子を濃縮することで、NGSライブラリを調製する準備が完了します。



近接ライゲーション部分は、遊離末端からライゲーションプロダクトへの変換が非常に効率的に行われるように最適化されており、最終的なライブラリ内の長距離情報を最大限に取り込むことができます。

一例として、CTCFのHiChIPライブラリでは、82%の有効なリードペアが生成されました(表1)。この値は、Dovetail™ Micro-Cで生成された高品質なHiChIPライブラリと一致しています。さらに、HiChIPライブラリのS/N比からも明らかのように、シーケンシングリードはタンパク質因子結合部位に非常に富みます。

他のHiChIPアッセイに共通する制限は、許容できる再現性とS/N比を達成するために必要な細胞のインプット量です。Dovetail™ HiChIPアッセイでは、抗体免疫沈降に必要な細胞数はわずか100万個(表2)で、再現性が高く、100万~1,000万個の細胞で免疫沈降を行った際に、クロマチンコンタクトとタンパク質結合シグナルの両方で高いオーバーラップが見られることから明らかです(図2)。

Dovetailでは、導入を容易にするために、現在、一般的に使用されている8種類の抗体と免疫沈降ガイドラインの検証済みセットを用意しており、今後も抗体の選択肢を増やしていく予定です(表2)。

表 2 - Dovetail™ HiChIPで検証済みの抗体。

Dovetail™ HiChIP MNase キットでは、8種類の抗体が検証済みですが、近日中に7種類の抗体の追加検証が予定されています。プロトコルには、抗体のカタログ番号と100万個の細胞までのサンプルインプットでの使用ガイドラインが記載されています。

Validated Antibodies	Future Antibodies
CTCF	Sox2
H3K27ac	Nanog
H3K4me3	Oct4
H3K27me3	KLF4
H3K36me3	PRC
H3K4ac	Suz12
PoII	Smc1
Smc3	

タンパク質結合部位と長距離タンパク質指向性相互作用のマッピング

Dovetail™ HiChIPは、タンパク質を中心とした3次元クロマチン相互作用の研究を可能にします。ChIP-seqと同様に、Dovetail™ HiChIPデータは目的のタンパク質に直接結合した配列を捕捉します(次ページ図3)。同じ抗体を用いて作成したCTCFのChIP-seqデータとHiChIPデータを比較すると、ChIP-seqデータで濃縮された配列がHiChIPデータでも同様に捕捉されていることが確認されました(それぞれのベッドグラフで示されています)。

しかし、CTCF因子によって媒介され、異なる部位で結合した長距離相互作用は、HiChIPデータでのみ弧線で見視化されます。ChIP-seqと比べHiChIPが優れている部分であり、クロマチン断片化にMNaseを使用することで、ヌクレオソームレベルの高分解能が提供されることも示しています。

CTCF結合部位の拡大図に示される、周期的なパターンはヌクレオソームの位置を反映しており、ピークはヌクレオソームの中心、トラフはヌクレオソームリンカーDNAを示しています。また同時に、ヌクレオソームレベル分解能でのタンパク質媒介クロマチンコンタクトの捕捉を示しています。

表 1 - Dovetail™ HiChIPライブラリは、近接ライゲーション情報リードが豊富に含まれています。

Dovetail™ HiChIP Mnase (88Mリードペア)とDovetail™ Micro-C (1.7Bリードペア)のライブラリ統計の比較。CTCFに対するDovetail™ HiChIPは、CTCFターゲットサイトに富み(S/N比2.5)、長距離情報を含んでいます(有効リードペア数82%)。S/N = ターゲットサイトのカバー率の平均(上位25%)/非ターゲットサイトのカバー率の平均(上位25%)、IgG HiChIPに正規化したもの。有効なリードの割合は、全アラインメントのtrans+cis > 1kbpと定義されています。

Target	CTCF HiChIP	Micro-C
Input Cells	1M	1M
Total Read-Pairs	88 M	1.3B
PCR Duplicates	2%	19%
% cis Read-Pairs	60%	68%
% Valid Read-Pairs	82%	94%
Signal-To-Noise	2.5	0.9

図2 細胞のインプット量が変化しても安定した性能を発揮します。

10⁶から10⁷のGM12878細胞のインプット量を、4種類の抗体(CTCF、H3K4me3、H3K27ac、H3K36me3)でテストしました。データを比較すると、ベッドグラフのリードの分布で示されるように、テストしたすべての入力でほぼ同等であることが確認されました。

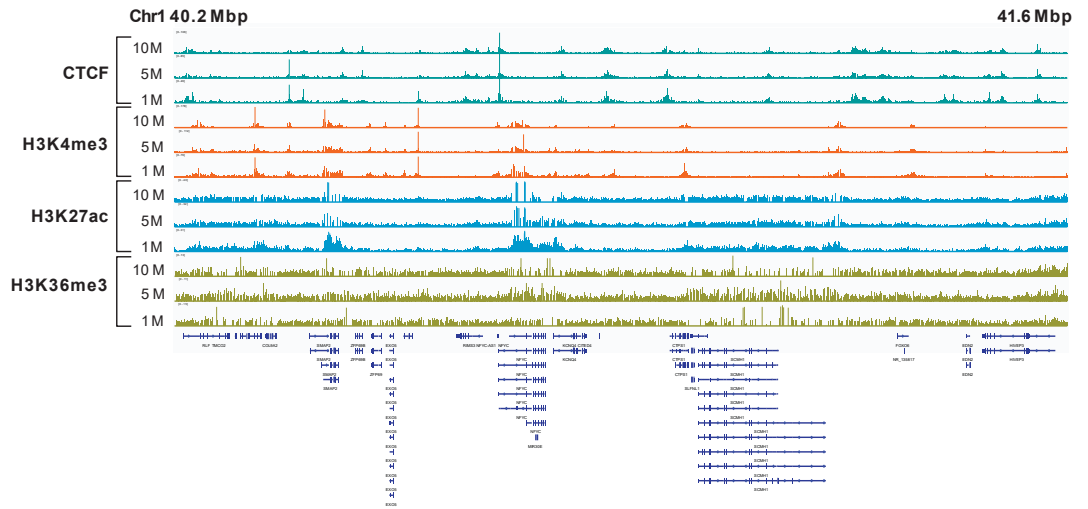
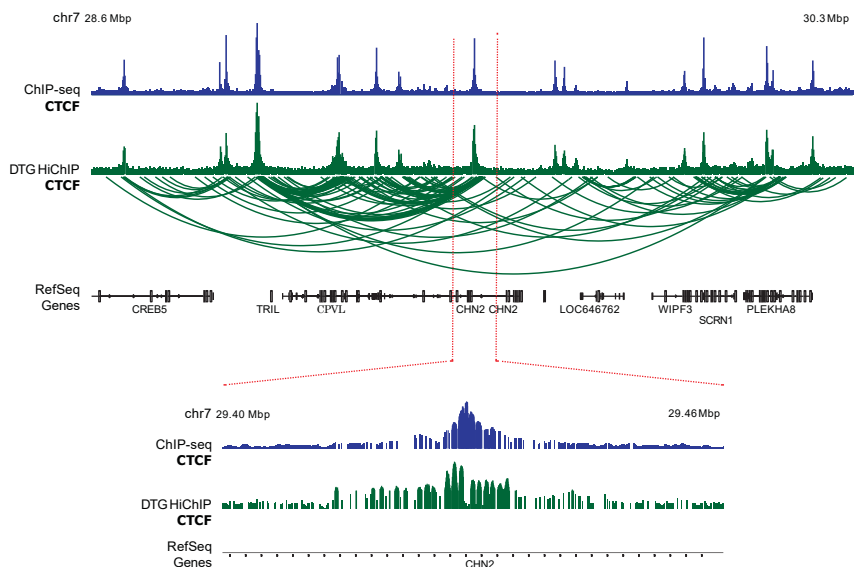


図3 - Dovetail™ HiChIPキットは、ChIP-seqデータとHi-Cの長距離情報を組み合わせたものです。

同じ抗CTCF抗体を使用したHiChIPとChIP-seqデータの比較。HiChIPデータセットには、ChIP-seqデータに加えて、弧線で示される長距離CTCF相互作用が捕捉されています。相互作用するゲノムピンを結ぶ弧線は、Rパッケージ「Sushi」で可視化されたもので、円弧の高さは相互作用の距離に依存します。50kbp以上の距離で、相互作用頻度が1以上のシスコンタクトのみを考慮しています。プロットウィンドウを超えた相互作用は表示されていません。断片化にMNaseを使用することで、Dovetail™ HiChIPとDovetail™ Micro-Cキットの両方でヌクレオソームの分解能を向上させることができます。下部パネルの拡大表示では、ヌクレオソームを中心としたピークの周期性を表示しています。



少ないシーケンス量で高解像度なコンタクトマップを作成

タンパク質指向性クロマチン構造のエンリッチにより、より少ないリードデプスで高解像度コンタクトマップの作成が可能となります。6B以上のリードペアを用いて作成された高解像度制限酵素(RE)ベースのHi-C (Rao et al.)と比較して、Dovetail™ HiChIPデータは、ループやクロマチン相互作用などの高次クロマチン構造を、少ないリードデプスで可視化することができます(CTCFでは88M、H3K4me3では160Mリードペア、図4)。

また、CTCF HiChIPデータは、2つのCTCF結合部位の対角線上の交点(破線の円で囲まれた部分)でのリードの濃度によって可視化された、CTCF因子間の相互作用を示します。

興味深いことに、これとは対照的に、転写活性の指標であるH3K4me3は、遺伝子転写方向のプロモーターから発せられる一時的な相互作用を示唆する、より拡散的なパターンを生成します。

図5は、クロマチンループ検出のためのCTCFエンリッチの有用性と、それに伴うコスト削減をさらに強調したもので、ループ構造を中心として、Dovetail HiChIPとマルチREベースのHi-Cコンタクトマップの比較を示しています。Dovetail™ HiChIPライブラリでは、このCTCFが媒介するクロマチンループでシグナルがエンリッチされ、他の場所ではシグナルが減少していることがすぐにわかります。この比較は、近接ライゲーションとタンパク質濃縮の両方でS/N比を向上させることができるDovetail™ HiChIPキットの有用性を示しています。

図 4 - 高解像度なコンタクトマトリクスはCTCF/H3K4me3 HiChIPで観測されるコンタクトをハイライトします。

GM12878細胞におけるH3K4me3 (160Mリードペア)、CTCF (88Mリードペア)、IgG (72 Mリードペア、ネガティブコントロール)のDovetail™ HiChIP解析により得られたコンタクトマトリクス。CTCF HiChIPデータは、2つのCTCF結合部位の対角線上の交点(破線の円で囲んだ部分)でのリードの濃度によって示されるように、CTCF因子間の相互作用を示している。対照的に、転写活性の指標であるH3K4me3では、遺伝子の転写方向にあるプロモーターから発せられる一時的な相互作用を示唆する、より拡散したパターンが見られます。

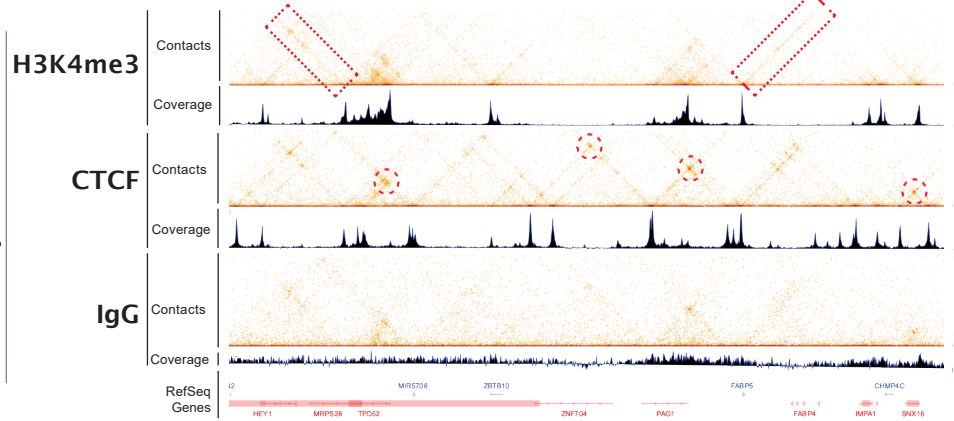
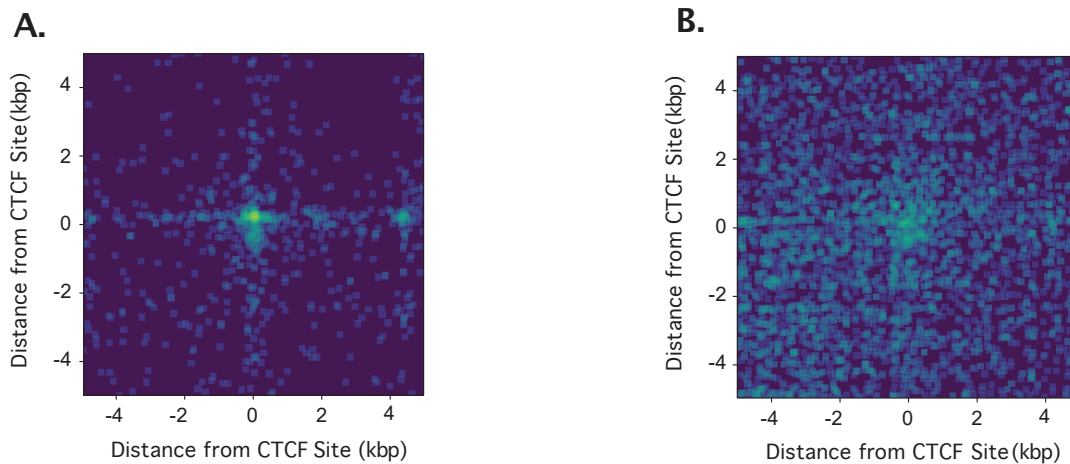


図 5 - 高いS/B比をもつクロマチン相互作用マップにより、より低いシーケンシングコストで高次構造を可視化できます。

Raoらの報告したループ位置を中心としたコンタクトマップの比較。A) CTCFに対するDovetail™ HiChIP MNase Kitデータ、B) GM12878細胞におけるマルチRE Hi-Cデータ。Dovetail™ HiChIPを用いたCTCF免疫沈降は、全てのクロマチン相互作用をバックグラウンドとするマルチRE Hi-Cと比べ、CTCFを介したループなどのクロマチン相互作用のみのシグナルを増加させます。このコンタクトシグナルはDovetail HiChIP抗CTCFデータセットでは150Mリードペア、ゲノム全体のREベースのHi-Cでは800Mリードペアから作成されたもので、例えばIllumina NovaSeq 6000SP - 300サイクルフローセルでのシーケンシングでは、約2,700ドルのコスト差があります。



まとめ

Dovetail™ HiChIPキットは、標準的なChIP-seqと同様のタンパク質DNA結合部位のゲノムワイドマッピングに加えて、クロマチン構造や遺伝子制御に関与するタンパク質指向性の長距離相互作用を同時に捕捉することができます。

また、Dovetail™ HiChIPのワークフローは、分子生物学的に改良された近接ライゲーションとタンパク質濃縮を組み合わせたもので、検証済みの抗体を使用することができ、超音波破碎を必要としない使いやすい方法になっています。この新しいツールは、転写因子結合部位と遺伝子プロモーターを近接した位置に配置する役割を果たすTADやループなどの、タンパク質指向性のトポロジー構造の発見とマッピングを可能にします。

参考文献

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>