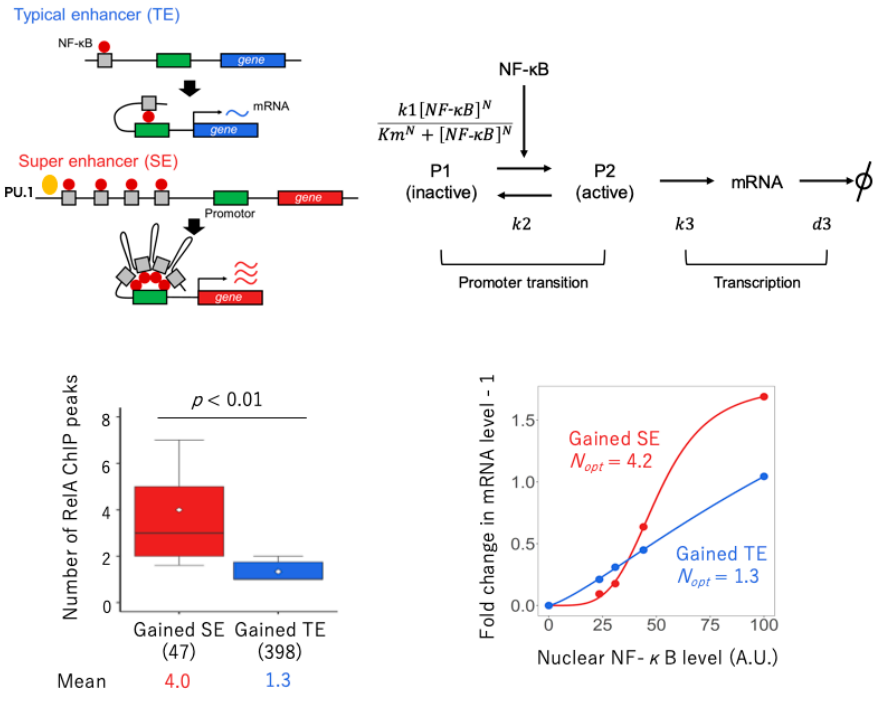
76　情報統合解析による細胞時間発展の分子機序の解明	岡田　眞里子

【目的】本研究では、細胞核内のクロマチン状態の変化と、それに連動して起こる細胞の不可逆な遷移過程をデータ駆動的に捕捉することにより、細胞個性獲得のための分子制御機構を明らかにすることを目指す。クロマチン状態、転写因子結合、エンハンサー活性、遺伝子発現の定量的関係を明らかにし、数理モデル化することにより、さまざまな疾患解明に応用可能な数理情報基盤を構築する。

【方法】免疫 B 細胞の分化における NF-κB の転写制御を対象とし、エンハンサーと呼ばれる遺伝子の制御領域とその標的遺伝子発現との関係に注目した情報統合解析を行った。マウス脾臓から B 細胞を単離し、anti-IgM 刺激有無の条件下で、細胞試料を作製した。Super enhancer（SE）と Typical enhancer （TE）というタイプの異なるエンハンサーの同定を目的とした H3K27ac 免疫沈降シーケンス（acetylated histone H3 lysine 27 ChIP-seq）、NF-κB（RelA）の DNA 結合領域を同定するための RelA ChIP-seq、クロマチン開口領域を同定するための ATAC-seq（an assay for transposase-accessible chromatin through sequencing）を行った。さらに、エンハンサーの標的遺伝子の発現量を求めるために、一細胞 RNA シーケンス（RamDA-seq）を行った。これらの次世代シーケンスデータをもとに、統合情報解析を行った。抗原刺激前後での SE および TE における各シグナルの定量的比較、クロマチン開口領域のモチーフ解析による SE と TE に特徴的な結合転写因子の予測を行った。最終的に、エンハンサーのクロマチン開口領域における NF-κB 結合数と各遺伝子発現量の関係を数理モデルを用いて解析した。

【結果】H3K27ac データにより、SE および TE を同定し、この情報統合解析を行ったところ、SE において、クロマチン開口、NF-κB 結合が連動して起こり、遺伝子発現を高く、また閾値様に誘導することが明らかになった。さらに、このデータを定量的に解析することにより、SE と TE の違いは、DNA の長さとそれに伴う NF-κB 結合数の増加にあることがわかった。詳細な解析により、DNA の長さは遺伝子の倍数変化に影響し、NF-κB の結合数は、閾値変化に重要であることがわかった。また、モチーフ解析の結果から、この閾値応答には、パイオニア因子 PU.1 と NF-κB の共在が重要であることがわかった。SE により発現誘導される遺伝子のひとつに免疫細胞分化に重要な転写因子 IRF4 が見出されたことから、SE を介した転写因子の発現の時間発展が、細胞の不可逆的遷移に関与することが示唆された。

数理モデルを用いたエンハンサーによる遺伝子発現制御機構の同定



76. 情報統合解析による細胞時間発展の分子機序の解明

岡田　眞里子

大阪大学 蛋白質研究所 細胞システム研究室

Key words：転写制御，エピゲノム，クロマチン，シグナル伝達，数理モデル

緒　言

次世代シーケンス法と情報解析手法の開発により、細胞や組織における遺伝子やエピゲノムの情報が容易に取得・解析可能になった。そして、さまざまな大規模シーケンス解析により、ヒトの組織や細胞の網羅的カタログ、それぞれの状態を説明する遺伝子が同定された。しかし、一方で、このような解析は統計解析を主に用い、細胞や組織の違いを分類することができても、分子制御機構の同定は困難なことが多い。そこで、筆者らは、シーケンス解析に数理モデリングの手法を適用することで、細胞における環境依存的な転写制御機構の解明とその時間遷移解析を目指している。筆者らが特に注目するのは、NF-κB 転写因子の制御機構である。NF-κB は、細胞の分化、増殖、死などの多彩な現象に関わる転写因子で、細胞と環境の組み合わせに応じた遺伝子発現の制御を行うことから、この脱制御は、がん、炎症、免疫不全など広範な疾患に関与する。NF-κB はサイトカインなどの受容体シグナル伝達系により細胞質から核内に移行し、DNA に結合し、遺伝子発現を誘導するが、その活性化特性のひとつに、閾値応答が知られている［1］。閾値機構は、外的あるいは内的な因子濃度のわずかな量の違いといったアナログ情報を、有か無のデジタルな情報に変換させる分子機構で、細胞の個性の決定、不可逆性、記憶にも関与し、生体の恒常性維持や疾患機序としても重要である。筆者らはこれまで免疫 B 細胞における核内 NF-κB 活性の閾値が、シグナル伝達系内のフィードバック制御により制御されることを実験と数理モデルを組み合わせて見出した［2］。しかし、これにより活性化された NF-κB が遺伝子発現をどのように制御するのか、アナログ的なのかデジタル的なのかといった問題には、研究者の間でさまざまな議論があり、明快な答は得られていない。そこで本研究では、NF-κB の量的な遺伝子発現機構を明らかにすることを目的に研究を行った。本研究により、免疫 B 細胞において、抗原刺激による細胞内の NF-κB シグナル伝達経路の活性化により、シグナル依存的に形成された Super enhancer（SE）と呼ばれる長いエンハンサー［3］のクロマチン開口領域に通常のエンハンサー（Typical enhancer；TE）よりも NF-κB が多数箇所結合すること、そして、そのことにより、標的遺伝子の発現が閾値応答を示すことがわかった［4］。このことにより、発現誘導される遺伝子のひとつ IRF4 は免疫細胞分化に重要な転写因子であり、SE を介した転写因子の発現の時間発展が、細胞の不可逆的遷移に関与することが示唆された。

方　法

本研究では、免疫細胞におけるヒストン修飾、クロマチン開閉、 NF-κB 免疫沈降、遺伝子発現などの次世代シーケンスデータを取得し、エンハンサーと呼ばれる遺伝子発現の制御領域に注目し、データの統合情報解析を行った。抗原刺激前後での定量的比較、モチーフ解析による SE と TE に特徴的な結合転写因子の予測を行い、最終的には、エンハンサー領域における NF-κB 結合数と各遺伝子発現量の関係を数理モデルを用いて解析した。具体的な方法を以下に示す。

1.　細胞試料の調製

マウス脾臓から B 細胞を単離し、anti-IgM 刺激（1 時間）有無条件下で、細胞試料を作製した［2］。

2.　次世代シーケンス解析

Super enhancer（SE）と Typical enhancer（TE）の同定を目的として、H3K27ac 免疫沈降シーケンス（acetylated

histone H3 lysine 27 ChIP-seq）を行った。NF-κB（RelA）の DNA 結合領域を同定するため、RelA ChIP-seq を行った。クロマチン開口領域を同定するために、ATAC-seq（an assay for transposase-accessible chromatin through sequencing）を行った。エンハンサーの標的遺伝子の発現量を求めるために、一細胞 RNA シーケンス（RamDA-seq）［5］を行った。

3. データ情報解析

得られたシーケンス配列をマウスゲノム配列にマッピングし、リード数を正規化した。SE と TE の同定と分類は H3K27ac データを用い、ROSE アルゴリズムを用いて行った［4］。標的遺伝子とエンハンサーの対応は、同染色体における転写開始点とエンハンサーの最短距離の指標を用いて行った。モチーフ解析は HOMER を用いて行った。

4. 数理モデル解析

数理モデルは先に報告されている常微分方程式［6］を改変して用い、エンハンサー活性（図中にも記した）は、以下の式で示した。

$$\frac{k1[NF\kappa B]^N}{Km^N + [NF\kappa B]^N}$$

N：クロマチンが開いたエンハンサー領域における NF-kB の結合数、Km: NF-kB と DNA の親和定数、$k1$：プロモーター活性の速度定数を示す。

結　果

1. エンハンサー領域におけるクロマチン開口、NF-κB 結合、DNA 長との関係について

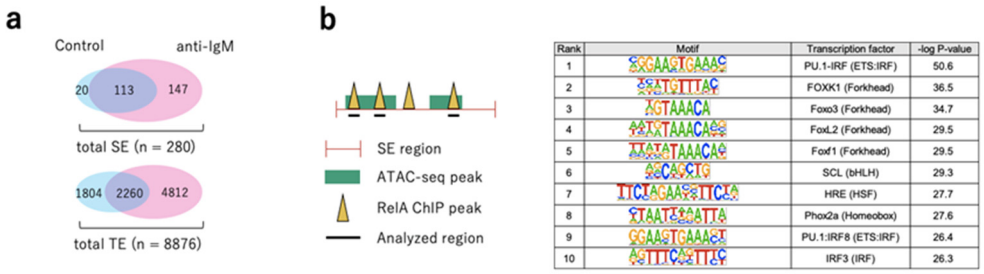
H3K27ac ChIP-seq により、SE および TE を同定し、各エンハンサー領域に対して、ATAC-seq および NF-κB ChIP-seq のシグナル強度を調べたところ、SE 領域において、H3K27ac とこれらのシグナル変化が高い相関を示すことがわかった。一方で、TE ではそのような相関が見られなかった。また、DNA 長との関係を調べたところ、SE、SE ともに、H3K27ac、ATAC-seq、NF-κB ChIP-seq のシグナルと DNA がおよそ相関係数 0.9 以上で相関することがわかった。

2. エンハンサーを介したシグナル依存的な遺伝子発現制御について

同定されたエンハンサーを H3K27ac シグナルを指標に、anti-IgM 依存的に新しく形成された SE/TE、消失した SE/TE、変化しない SE/TE の 3 群に分類し、標的遺伝子発現の倍数変化を調べたところ、SE/TE ともに、遺伝子の増加、減少、不変の変化が見られた。しかし、その変化量は、SE により制御される遺伝子のほうが大きいことがわかった。

3. パイオニア因子の関与について

SE および TE 領域におけるモチーフ解析を行ったところ、細胞刺激前の SE のクロマチン開口領域において、パイオニア因子 PU.1 と NF-κB 共在のエンリッチメントが有意に見られた（図 1）。また、PU.1 と NF-κB が共在する SE のほうがそうでない SE よりも強い閾値様パターンを示した。このことから、この共在が、SE 標的遺伝子の閾値様発現を引き起こすことが考えられ、パイオニア因子が SE を既定しているといった可能性が示された。



Rank	Motif	Transcription factor	-log P-value
1		PU.1-IRF (ETS:IRF)	50.6
2		FOXK1 (Forkhead)	36.5
3		Foxo3 (Forkhead)	34.7
4		Fox.L2 (Forkhead)	29.5
5		Foxf1 (Forkhead)	29.5
6		SCL (bHLH)	29.3
7		HRE (HSF)	27.7
8		Phox2a (Homeobox)	27.6
9		PU.1:IRF8 (ETS:IRF)	26.4
10		IRF3 (IRF)	26.3

図 1. B 細胞におけるエンハンサーと転写因子の結合モチーフ

a) anti-IgM 刺激前後における SE と TE の同定数

b) SE 領域における結合転写因子の予測（Welch's t-test）

4. NF-κB の協同的な結合による遺伝子の閾値様発現について

前述の SE 領域において、PU.1 パイオニア因子と NF-κB の共在が有意に見られたことから、この共在が、SE 標的遺伝子の閾値様発現に必要なことが考えられた。PU.1 は B 細胞の系列決定因子であり、この因子がクロマチンの開口領域に NF-κB を留め、シグナル依存的に核内量が増加した NF-κB をプロモーター領域で協同的に結合させるという仮説を立て、NF-κB の結合個数を N とした数理モデル（図 2）を構築した。このモデルにより、SE および TE の遺伝子発現動態を説明できたことから、NF-κB のエンハンサー領域の結合数が標的遺伝子を閾値様に発現させるのに重要であると考えられた。

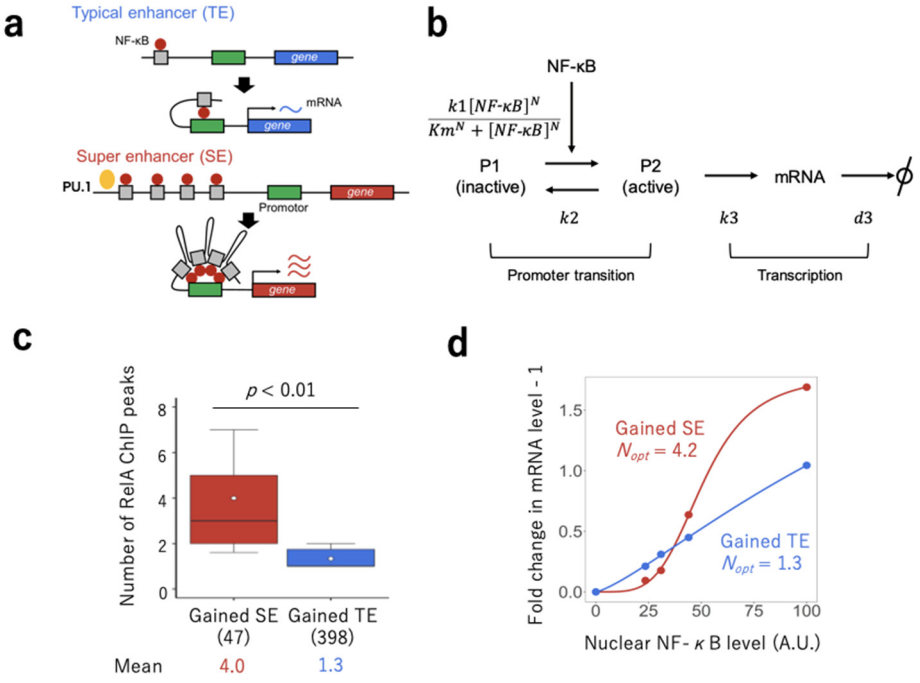


図 2. 数理モデルを用いたエンハンサーを介した遺伝子発現制御の解析

a) SE と TE 領域における NF-κB の結合様式。

b) 数理モデルの概念図。P はプロモーター、*k2*、*k3*、*d3* はそれぞれ、プロモーターの不活性化定数、転写速度定数、mRNA 分解定数を示す。

c) シグナル依存的に活性化された SE および TE のクロマチン開口領域における NF-κB ChIP のピーク数の数。検定は Welch's t-test を用いた。

d) 核内 NF-κB 量に対する遺伝子発現量のシミュレーション結果。*Nopt* は遺伝子発現量から推定された NF-κB 結合数（N）を示す。

考　察

本研究により、SE と TE の違いは、DNA の長さとそれに伴う NF-κB 結合数の増加にあることがわかった。詳細な解析により、DNA の長さは遺伝子の倍数変化に影響し、NF-κB の結合数は、閾値変化に重要である可能性が示された。また、この閾値応答には、パイオニア因子 PU.1 と NF-κB の共在が重要なことが示されたことから、パイオニア因子が SE を事前に規定していることが明らかになった。このように、SE に注目したシーケンスデータの統合解析から、転写の分子制御機構を定量的に明らかにできることがわかった。特に、クロマチンの開閉、ヒストン修飾、遺伝子発現などの公共データなどから容易に得ることのできるスナップショットのシーケンスデータと、モチーフ解析を組み合わせることにより、転写の機構が動的に説明できる可能性が示された。筆者らは、SE とクロマチンの開閉の時間変化が、転写因子の結合量の時間変化を引き起こし、遺伝子発現の変化をもたらすこと、このような核内の時間変化が生体内の細胞や周辺組織に、条件依存的な不可逆的時間発展をもたらすのではないかと考えている。今後は、クロマチン状態を含む SE の時間発展と、それに連動して起こる細胞の不可逆な遷移過程をデータ駆動的に捕捉することにより、細胞個性の獲得のための分子制御機構を明らかにするとともに、 シーケンスデータから細胞個性を予測する数理モデルの構築を目指していきたいと考えている。このような研究を行うことで、分子機構解明を目指した生命情報解析の考え方や手法を開拓し、さまざまな疾患の解明に利用していきたい。

共同研究者・謝辞

本研究の共同研究者は、東京大学大学院新領域創成科学研究科の鈴木穣教授、理化学研究所生命機能科学研究センターバイオインフォマティクス研究開発チームの二階堂愛チームリーダーである。

文　献

1) Tay S, Hughey JJ, Lee TK, Lipniacki T, Quake SR, Covert MW. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. Nature. 2010 Jul 8;466(7303):267-71. doi: 10.1038/nature09145.

2) Shinohara H, Behar M, Inoue K, Hiroshima M, Yasuda T, Nagashima T, Kimura S, Sanjo H, Maeda S, Yumoto N, Ki S, Akira S, Sako Y, Hoffmann A, Kurosaki T, Okada-Hatakeyama M. Positive feedback within a kinase signaling complex functions as a switch mechanism for NF-κB activation. Science. 2014 May 16;344(6185):760-4. doi: 10.1126/science.1250020.

3) Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013 Apr 11;153(2):307-19. doi: 10.1016/j.cell.2013.03.035.

4) Michida H, Imoto H, Shinohara H, Yumoto N, Seki M, Umeda M, Hayashi T, Nikaido I, Takeya K, Suzuki Y, Okada-Hatakeyama M. The number of transcription factors at an enhancer determine switch-like gene expression. In revision

5) Hayashi T, Ozaki H, Sasagawa Y, Umeda M, Danno H, Nikaido I. Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. Nat Commun. 2018 Feb 12;9(1):619. doi: 10.1038/s41467-018-02866-0.

6) Hao N, O'Shea EK. Signal-dependent dynamics of transcription factor translocation controls gene expression. Nat Struct Mol Biol. 2011 Dec 18;19(1):31-9. doi: 10.1038/nsmb.2192.