

DNA の物理的特性の生物学的意義
—真核生物ゲノムの核内収納原理—

Biological significance of physical properties of
DNA: the genome folding principle in eukaryotes

2012 年 12 月

早稲田大学大学院 先進理工学研究科

生命理工学専攻 分子遺伝学研究

木村 元

目次

略語一覧	．．．．．4
序論	．．．．．5
第 1 部 各種真核生物ゲノムの柔軟性地図	
1.1. 緒言	．．．．．10
1.2. 実験方法	
1.2.1. DNA の柔軟性の計算	．．．．．12
1.2.2 SPIKE と rSPIKE の選出	．．．．．14
1.2.3 SPIKE と rSPIKE に含まれるテトラヌクレオチドの出現頻度の解析	．．．．．15
1.2.4 SPIKE と rSPIKE に含まれる繰り返し配列の解析	．．．．．15
1.3. 結果	
1.3.1. ヒトゲノムの柔軟性地図	．．．．．17
1.3.2. 真核生物ゲノムに存在する異常に柔らかい DNA 領域	．．．．．21
1.3.3. SPIKE と rSPIKE の塩基配列の特徴	．．．．．36
1.4. 考察	．．．．．40
第 2 部 ヌクレオソーム DNA に共通した物理的特性	
2.1. 緒言	．．．．．44
2.2. 実験方法	
2.2.1. ヌクレオソーム DNA の柔軟性の計算	．．．．．45
2.3. 結果	
2.3.1. ヌクレオソーム DNA に共通した柔軟性	．．．．．47

2.4. 考察	51
第3部 間期染色体の三次元構造の構築原理	
3.1. 緒言	56
3.2. 実験方法	
3.2.1. シミュレーション方法	58
3.2.2. 細胞核の制限	58
3.2.3. 原子間力顕微鏡（AFM）を用いた DNA 断片の観察	59
3.2.4. FISH 法を用いた 2 点間空間距離の測定	60
3.2.5. 4C データから 2 点間空間距離を導出する方法	62
3.3. 結果	
3.3.1. 間期染色体構造モデルの構築	63
3.3.2. リンカーDNA の持続長の測定	69
3.3.3. 出芽酵母の間期染色体構造のシミュレーション	70
3.3.4. 間期染色体構造モデルの評価	75
3.3.5. 出芽酵母の細胞核におけるクロマチン繊維の態様	84
3.4. 考察	
3.4.1. パラメータの選択が間期染色体構造に与える影響	86
3.4.2. 間期染色体構造の決定因子	89
総括	90
謝辞	91
参考文献	92
研究業績	101

略語一覽

<i>MGAM</i>	maltase-glucoamylase (alpha-glucosidase) gene
<i>SLC1A7</i>	solute carrier family 1 (glutamate transporter), member 7 gene
<i>FAIM2</i>	Fas apoptotic inhibitory molecule 2 gene
<i>NOC4L</i>	nucleolar complex associated 4 homolog (<i>S. cerevisiae</i>) gene
<i>MAPK8IP3</i>	mitogen-activated protein kinase 8 interacting protein 3 gene
<i>AC008132.1</i>	unknown gene
<i>SHANK3</i>	SH3 and multiple ankyrin repeat domains 3 gene
AFM	atomic force microscopy (原子間力顕微鏡)
<i>LEU2</i>	Beta-isopropylmalate dehydrogenase (IMDH) gene
FISH	fluorescence <i>in situ</i> hybridization
4C	chromosome conformation capture-on-chip
PL	persistence length (持続長)

序論

高等真核生物のゲノム DNA には、ある個体が卵から完全な成体へと成熟するために必要なすべての遺伝情報が印されている。すなわち、高等真核生物において、個々の細胞が集合して組織をつくり、種々の組織が同調して器官として働き、ひいては個体が生存するために必要な様々な機能を実現する胚発生の全過程は、予めゲノム DNA にプログラムされている。このように、“生物の設計図”といえるゲノム DNA に印された遺伝情報のすべてを解読することは、生物が普遍的にもつ DNA に基づいて生物を理解することとほぼ同義であることから、生命科学における最も重要な課題のひとつであると考えられる。

2003 年に 30 億塩基対からなるヒトゲノム配列が完全解読された結果、タンパク質のアミノ酸配列をコードする領域は、そのうちのほんの数%の領域を占めるに過ぎないことが明らかになった。それ以降、ENCODE (Encyclopedia of DNA elements) プロジェクトをはじめとする、ゲノム配列に印された新たな遺伝情報を解読することを目的とした研究が著しく発展した。しかしながら、真核生物ゲノムに対する我々の理解は依然として不十分である。そこで、ゲノム DNA に印された遺伝情報のすべてを解読するため、従来の概念にとらわれない新たな視点に基づいたアプローチの提案が急がれている。

近年、DNA の物理的特性に焦点を当てた研究により、DNA に印された新たな遺伝情報を解読することに私が所属している大山研究室が成功した。具体的には、DNA の物理的特性のひとつである DNA の柔軟性 (DNA 分子がもつ硬さ・柔らかさの特性) に、プロモーターの働きを制御する遺伝情報が印されていることが判明した (Fukue *et al.*, *Nucl. Acids Res.* 32, 5834-5840, 2004 ; *ibid.* 33, 3821-3827, 2005)。したがって、従来の“遺伝情報は核酸塩基の配列として符号化されている”という概念を超えて、DNA の物理的特性の生物学的意義を解明することが、上述の課題に対するひとつの解決方策であると考えられる。

本論文は、DNA の物理的特性の生物学的意義を明らかにするために行なった研究の報告である。古くから DNA の物理的特性は、真核生物ゲノムの折りた

たみ機構を解明するための糸口であると考えられてきたが、それを示す直接的な証拠は未だ得られていない。一方、ゲノム DNA が折りたたまれて細胞核内でとる三次元構造は、転写・複製・修復などの様々な核内現象の舞台となる重要な基盤構造である。そこで、DNA の物理的特性が真核生物ゲノムの機能的な折りたたみ機構に果たす役割の解明を目的として、大きく分けて3つの研究を行なった。本論文ではこれらの結果について、それぞれ第1部と第2部、および第3部に分けて述べる。

第1部の研究では、DNA の柔軟性に印された新たな遺伝情報を明らかにする目的で、ヒト、マウス、ショウジョウバエ、線虫のゲノム全域を対象として、柔軟性地図を作成した。その結果、個々の染色体がそれぞれに固有の DNA の柔軟性をもつことなど、様々な特徴が明らかになった。これらの特徴のうち、解析したすべてのゲノムに存在する異常に柔軟な領域 (SPIKE と命名) および異常に硬い領域 (rSPIKE と命名) に焦点を当てた解析を行なったところ、SPIKE は遺伝子のエクソンには存在しないことや、高等な生物ほど、SPIKE が存在する頻度が rSPIKE が存在する頻度と比べて相対的に高くなることが判明した。さらに、SPIKE および rSPIKE は、異常な特性をもつ反復配列と異常な特性をもつ非反復配列で構成されていることが判明した。これらの特異領域は、真核生物ゲノムの折り畳みにおける重要な構造因子である可能性がある。

第2部の研究では、どのような物理的特性をもつ DNA がヌクレオソームを形成しやすいのかを解明することを目的として、ヌクレオソーム DNA に共通する物理的特性について解析を行なった。具体的には、出芽酵母のヌクレオソームに由来する 199 の DNA 配列を対象として、これらの平均的な柔軟性を明らかにした。その結果、極めて硬い領域と極めて柔らかい領域が、それぞれ約 10 bp の間隔で周期的に存在し、両者の間には約 5 bp のずれがあることが判明した。このようなヌクレオソーム DNA に共通した物理的特性は、ヌクレオソームを安定化する重要な物理的特性であると推察される。

第3部の研究では、間期染色体の三次元構造の構築原理を解明することを目的とした。本研究では、出芽酵母 (*Saccharomyces cerevisiae*) の間期染色体の

全長を、ヌクレオソームの配置が識別できる分解能でモデル化した。具体的には、ヌクレオソームの配置、リンカーDNAの物理的特性、および細胞核の大きさの3つの因子を用いてモデルを構築した。予測された間期染色体構造と核内における実際の構造との比較を行なったところ、予測値と実験値とが非常によい一致を示した。従って、モデルの構築に用いた3つの因子（ヌクレオソームの配置、リンカーDNAの物理的特性、細胞核の大きさ）が、間期染色体の三次元構造を規定する本質的な因子であると考えられる。

第 1 部

各種真核生物ゲノムの柔軟性地図

1.1. 緒言

2003年に、30億塩基対からなるヒトゲノム配列が完全解読された。それ以降、ゲノム科学は著しい発展を遂げた [1, 2]。例えば、現在では塩基配列データベースから様々な生物種のゲノム配列を簡単に得ることが可能である [3]。最近、29種の哺乳類を対象とした比較ゲノム研究により、ヒトゲノムの少なくとも5.5%以上の領域が進化的な制約を受けていることが判明した [4]。すなわち、ヒトゲノムの5.5%以上の領域がなんらかの生物学的な機能を担うことが判明した（さらに大きな推定値も報告されている [5-7]）。しかしながら、ヒトゲノムに対する我々の理解は依然として不十分である [8]。現在、ヒトゲノムに印された新たな遺伝情報を明らかにする目的で、ENCODE (Encyclopedia of DNA elements) プロジェクトが進行中である [9]。

ゲノム DNA は、高度に圧縮されて細胞核に納められている。例えば、ヒトの場合、父親のゲノムと母親のゲノムのそれぞれに由来する合計 2 m もの長さの DNA は、長さにして約 20 万倍も圧縮されて、微小な細胞核（直径約 10 μm ）に納められている。古くから、DNA の物理的特性はゲノムの折りたたみ機構を解明するための糸口であると考えられてきた。しかしながら、この機構については未だ不明な点が多い [10-12]。近年、DNA の物理的特性が生物学的な目印として機能していることを、私が所属している大山研究室が明らかにした。具体的には、ヒトの RNA ポリメラーゼ II プロモーター 1,871 種の柔軟性を解析した結果、ヒトの RNA ポリメラーゼ II プロモーターに共通した柔軟性が存在することを明らかにした [13]。さらに、この共通した柔軟性を合成 DNA に賦与すると、この DNA がプロモーターとして機能することを明らかにした [14]。このように、DNA の柔軟性に着目した解析手法は、ヒトゲノムに印された新たな遺伝情報を明らかにするための、有効な方策のひとつであると考えられる。そこで、ゲノム全域を対象として DNA の柔軟性を解析することにした。

本研究では、全 136 種類のテトラヌクレオチドの柔軟性パラメータセットを用いて [15]（表 1.1）、ヒト、マウス、ショウジョウバエ、線虫のゲノム全域

の柔軟性地図を作成した。その結果、解析したすべてのゲノムに、異常な柔軟性をもつ領域が存在することが判明した。このような特異領域を構成する塩基配列の特徴を明らかにするとともに、特異領域が存在する頻度について生物種間の比較を行なった。

1.2 実験方法

1.2.1. DNA の柔軟性の計算

塩基配列データベースである Ensembl ([3] ; <http://www.ensembl.org/index.html>) より、ヒト (*Homo sapiens*; version: GRCh37)、マウス (*Mus musculus*; NCBIM37)、ショウジョウバエ (*Drosophila melanogaster*; BDGP5.25)、ならびに線虫 (*Caenorhabditis elegans*; WS220) のゲノム配列を得た。次に、任意の塩基配列を対象としてその柔軟性を迅速に計算することが可能なコンピュータプログラムを、C 言語および R 言語 [16] を用いて作成した。このコンピュータプログラムによる計算の第一段階において、所与の塩基配列は任意の大きさの解析単位 [この解析単位をウィンドウ (window) と呼ぶ] へと分節化される。この際、隣り合うウィンドウ同士は互いに重なり合うことが許されており、隣り合うウィンドウ間のずれの大きさはウィンドウスライディングステップ (window sliding step) と呼ばれる。計算の第二段階において、各々のウィンドウに属する塩基配列は 4 塩基対 (テトラヌクレオチドステップと呼ぶ) の単位へとさらに分節化される。この際、隣り合うテトラヌクレオチドステップの間には 1 塩基対のずれが必ず存在し、例えば、ウィンドウの大きさが 10,000 bp の場合、この計算によって 9,997 個のテトラヌクレオチドステップが生じる。第三段階において、Packer らが報告している全 136 種類のテトラヌクレオチドの柔軟性パラメータセット [15] (表 1.1) を用いて、各々のテトラヌクレオチドの柔軟性が求められる。そして計算の第四段階において、各々のウィンドウにおける平均値が計算され、この解析単位を代表する値としてプロットされる。なお、約 3.1 Gb にも及ぶ長大なヒトゲノムを俯瞰的にみるために、ウィンドウの大きさを 10 kb、ウィンドウスライディングステップの大きさを 100 bp に設定して計算した (特に記述していない場合、ウィンドウの大きさをこの値に設定して計算した)。また、セントロメア領域ならびにサブセントロメア領域に関しては、その配列情報を塩基配列データベースから得ることができないため、計算の対象から除いた。

表 1.1. Packer らの柔軟性パラメータセット

テトラヌクレオチド	柔軟性 (kJ mol ⁻¹ Å ⁻²) ^{a,b}	CCAG/CTGG	GAGC/GCTC	11.9	CAGC/GCTG	17.1
TACA/TGTA	1.9	CCAG/CTGG	GAGC/GCTC	11.9	CAGC/GCTG	17.1
CATA/TATG	3.2	CGAC/GTGG	CCGA/TCGG	12.1	GACG/CGTC	17.3
TATA/TATA	3.6	AACA/TGTT	CTAC/GTAG	12.1	AACC/GGIT	17.3
TGCA/TGCA	3.8	ACAA/TTGT	GGCA/TGCC	12.1	CAAC/GTTG	17.5
GGCG/GCGC	3.9	ATAC/GTAT	AGCT/AGCT	12.2	AGAA/TTCT	17.7
CACG/CGTG	4.1	AAGT/ACTT	TGAA/TTCA	12.5	CTAG/CTAG	17.8
GAGT/ACTC	4.4	CGCA/TGGG	CGAG/CTCG	12.6	AGAG/CTCT	17.9
CAC/ACTG	4.8	AGCG/CGCT	AATG/CATT	12.7	GGAC/GTCC	18.0
CTAA/TTAG	5.0	AGCA/TGCT	TAGA/TCTA	13.1	TAGT/ACTA	18.1
ATAA/TTAT	6.0	TCAC/GTGA	CCAA/TTGG	13.1	CGAA/TTCC	18.1
TAGG/CGTA	6.2	CGCG/CGCG	GGAG/CTCC	13.3	GGGT/ACCC	18.3
CACA/TGTG	6.6	TCGA/TCGA	ACAT/ATGT	13.5	TAAA/TTTA	18.8
GCAC/GTGC	6.6	TGGC/GCCA	CAGA/TCTG	13.5	GGCC/GGCC	19.1
GCGA/TCGC	7.2	GATA/TATC	GAGG/CCTC	13.6	CGAT/ATCG	19.2
GCAG/CTGC	7.2	AAGG/CCCT	TGAC/GTCA	13.7	GAAA/TTTC	19.3
ACGG/CCGT	7.3	TGGT/ACCA	GGGA/TTCC	13.8	TTAA/TTAA	19.4
CATG/CATG	7.7	GGAA/TTCC	CATC/GATG	13.9	GGGC/GCCC	19.6
CCAC/GTGG	7.9	GTA/TTAC	AAGC/GCTT	14.3	TAAG/CTTA	19.8
CGGA/TCGG	7.9	TCAT/ATGA	CGGC/GCCG	14.5	AGAC/GTCT	20.1
TAGG/CCTA	7.9	AGGG/CCCT	TGAG/CTGA	14.5	CAAG/CTTG	20.2
GTAC/GTAC	8.0	CAGG/CCTG	AGAT/ATCT	14.5	TAAAC/GTTA	20.7
GCAT/ATGC	8.0	TGGG/CCCA	TAAT/ATTA	14.7	GAAT/ATTC	21.3
ACGA/TCGT	8.0	ACAG/CTGT	AATA/TATT	14.8	GAA/CTTC	21.5
GACA/TGTC	8.4	ATAG/CTAT	AACG/CGTT	14.9	AAAT/ATTT	21.7
CGCC/GCGG	8.6	CCGC/GCCG	CGGT/ACCG	15.5	CAAT/ATTG	22.3
TACT/AGTA	8.6	GGAT/ATCC	AACT/AGTT	15.5	GAGA/CTCC	22.7
TGGA/TCGA	8.6	ACGT/ACGT	AGGT/ACCT	15.7	AAAA/TTTT	23.8
ATAT/ATAT	8.6	TAGC/GCTA	AGGC/GCCT	15.9	CAAA/TTTG	24.3
GCAA/TTGC	8.7	TCAA/TTGA	AGCC/GGCT	16.0	AAAG/CTTT	24.5
CACC/GGTG	8.7	CCAT/ATGG	ACAC/GTGT	16.2	GAA/CTTC	24.9
TACC/GGTA	8.8	AGGA/TCCT	GGGG/CCCC	16.4	AAGA/CTTT	25.3
TCAG/CTCA	8.9	ACGC/GCGT	CCGG/CCGG	16.6	AAAT/GATT	26.8
TGAT/ATCA	9.0	CAGT/ACTG	GACT/AGTC	16.9	AAAT/AAAT	26.9
GACC/GGTC	9.1	CGGG/CCCC	GATC/GATC	17.0	AAAC/GTTT	27.2

^a参考文献 15

^b数値が小さいほど柔らかい。

1.2.2. SPIKE と rSPIKE の選出

ヒトゲノムに存在することが明らかになった極めて柔らかい DNA 領域 (SPIKE と命名; 図 1.1) についての詳細な解析を行なうため、特に柔軟な SPIKE (上位 0.05%) を選出し、これらを柔軟性の高さに基づいて 3 つのグループへと分類した。具体的には、最も柔軟な第 1 グループとして、0%分位点から 0.005%分位点までの範囲におさまる柔軟性をもつ SPIKE を分類し、2 番目に柔軟な第 2 グループとして、0.005%分位点から 0.01%分位点までの範囲におさまる柔軟性をもつ SPIKE を分類し、3 番目に柔軟な第 3 グループとして、0.01%分位点から 0.05%分位点までの範囲におさまる柔軟性をもつ SPIKE を分類した。このように分類した方法を以下に述べる。はじめに、第 1 グループの SPIKE を選出するため、ヒトゲノムの柔軟性地図 (ウィンドウの大きさ: 10 kb) から平均値が約 $12 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ 以下のウィンドウを選び出した。すなわち、特に柔軟性が高い上位 0.005%のウィンドウを選び出した [Packer らが報告している柔軟性パラメータセット [15] (表 1.1) においては、値が小さいほど柔軟性が高いことを意味する]。次に、互いに重なり合うウィンドウ同士を各々のグループとし (ひとつのグループを“ウィンドウグループ”と呼ぶ)、各々のウィンドウグループをひとつひとつの SPIKE とした。すなわち、各々の第 1 グループの SPIKE を、各々のウィンドウグループにおける最初のウィンドウの始点から、最後のウィンドウの終点までの DNA 領域として定義した。次に、第 2 グループの SPIKE を選出するため、0.005%分位点 (quantile) である約 $12 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ から、0.01%分位点である約 $12.2 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ までの範囲におさまる柔軟性をもつウィンドウを選び出した。これらのウィンドウから、第 1 グループの SPIKE と重なるウィンドウを除去したあと、上に述べた手法により、第 2 グループの SPIKE を得た。続いて、第 3 グループの SPIKE を選出するため、0.01%分位点である約 $12.2 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ から、0.05%分位点である約 $12.6 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ までの範囲におさまる柔軟性をもつウィンドウを選び出した。これらのウィンドウから第 1 グループおよび第 2 グループの SPIKE と重なるウィンドウを除去したあと、上に述べた手法により、第 3 グループの SPIKE を得た。また、ヒト以外の生物種の SPIKE、ならびに極めて硬い DNA 領域 (rSPIKE と命名) については、本質的に同様の手法を用いて選出したが、グループ分けは行なわなかった

〔SPIKEを選出する際には柔軟性が約 $12.6 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ 以下のウィンドウを選出し、rSPIKEを選出する際には柔軟性が約 $15.6 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ 以上のウィンドウを選出した〕。なお、特に記述していない場合、以降の解析の対象として、ヒトの SPIKE (第 1～第 3 グループ)、他の真核生物の SPIKE、ならびに rSPIKE を用いた。また、以上の計算を行なうコンピュータプログラムを R 言語 [16] を用いて作成した。

1.2.3. SPIKE と rSPIKE に含まれるテトラヌクレオチドの出現頻度の解析

ヒトの SPIKE および線虫の rSPIKE を対象として、はじめに、これらを構成する塩基配列をテトラヌクレオチドステップの単位へと分節化した。この際、隣り合うテトラヌクレオチドステップ間には 1 bp の大きさのずれが必ず存在し、例えば SPIKE または rSPIKE (を構成する DNA 領域) の長さが 10,000 bp の場合、この分節化により 9,997 個のテトラヌクレオチドステップが生じる。次に、全 136 種類のテトラヌクレオチド毎に出現回数を計算したあと、各々のテトラヌクレオチドについて、SPIKE (rSPIKE) を構成する塩基配列 10 kb あたりの出現回数と対照領域 10 kb あたりの出現回数の差 [SPIKE (rSPIKE) を構成する塩基配列 10 kb あたりの出現回数 - 対照領域 10 kb あたりの出現回数] を導いた。なお、対照領域として、各々の SPIKE (rSPIKE) から 100 kb 下流に位置し、SPIKE (rSPIKE) と同じ長さをもつ DNA 領域を用いた。もし、この領域がギャップ領域 (塩基配列データベースから情報を得ることができない領域) に位置する場合や、他の SPIKE (rSPIKE) と重なる場合には、該当する SPIKE (rSPIKE) を解析の対象から除いた。また、以上の計算を行なうコンピュータプログラムを R 言語 [16] を用いて作成した。

1.2.4. SPIKE と rSPIKE に含まれる繰り返し配列の解析

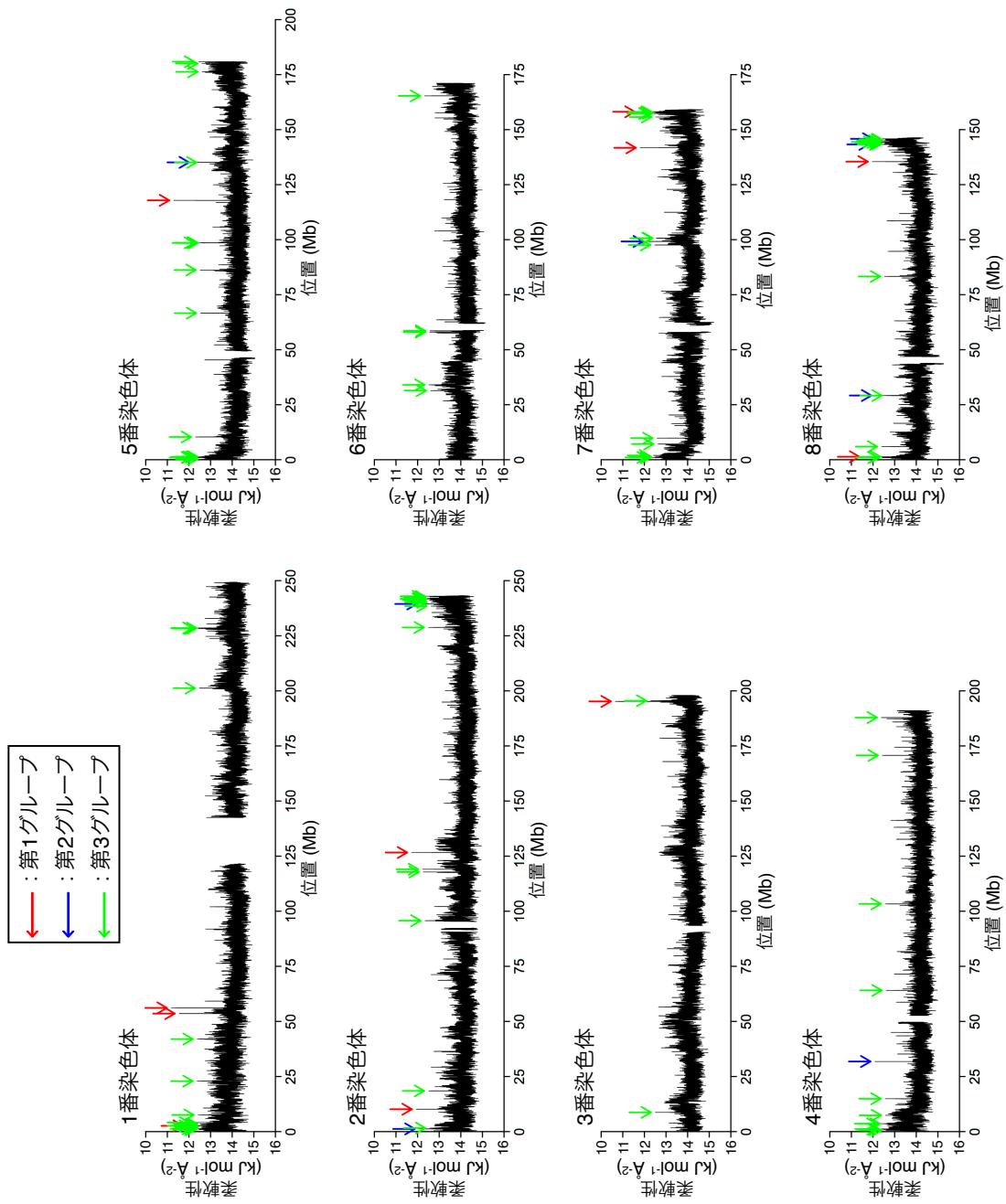
ヒトの SPIKE と線虫の rSPIKE、および対照領域 (“1.2.3. SPIKE と rSPIKE に含まれるテトラヌクレオチドの出現頻度の解析” を参照) がもつ繰り返し配

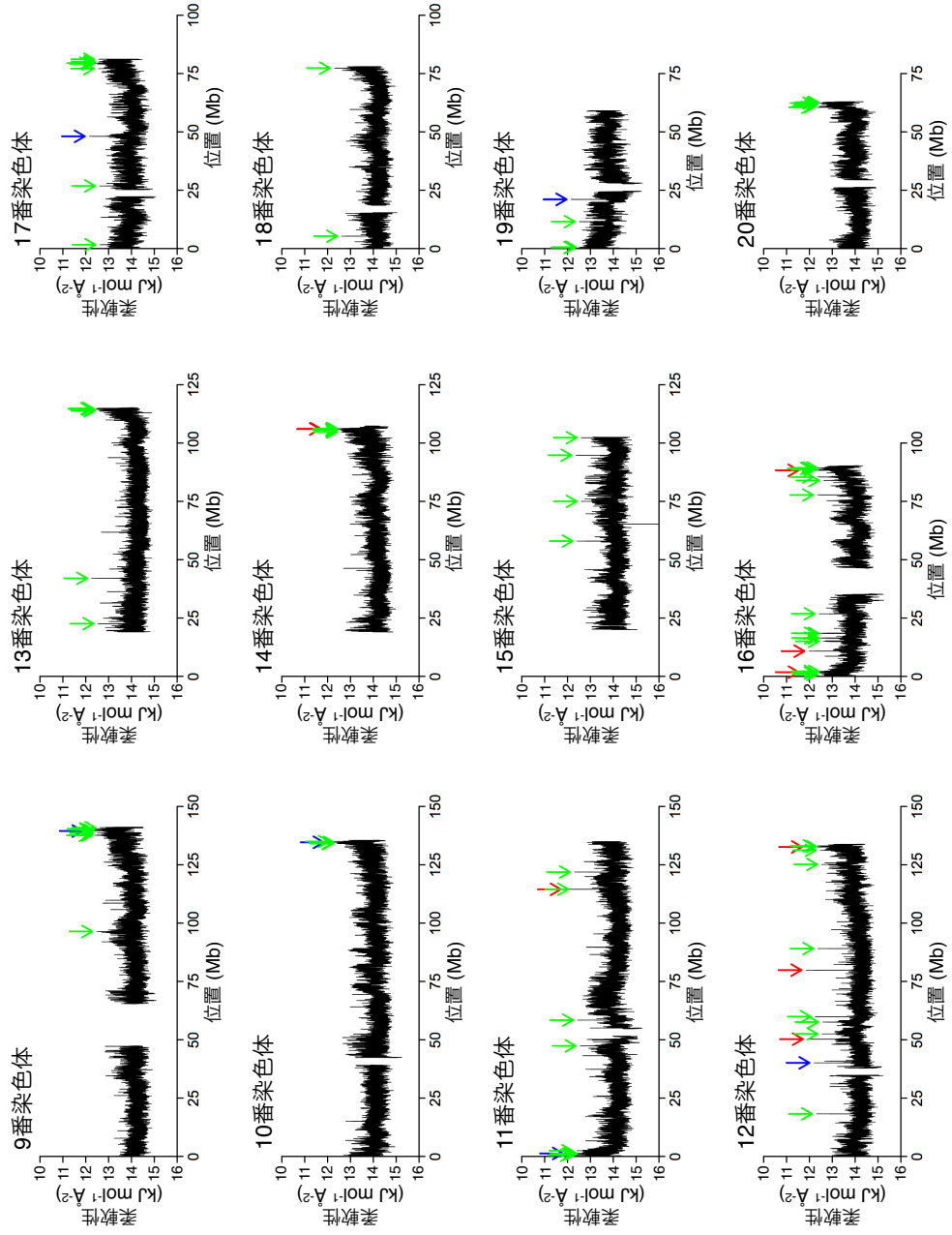
列を、塩基配列から繰り返し配列を発見するコンピュータプログラムである RepeatMasker [17] を用いて同定した。

1.3. 結果

1.3.1. ヒトゲノムの柔軟性地図

約 3.1 Gb の長さをもつヒトゲノム全域の柔軟性を解明するために、DNA の柔軟性を迅速に計算することが可能なコンピュータプログラムを作成した（“1.2.1. DNA の柔軟性の計算” に詳述）。このコンピュータプログラムを用いて作成した、ヒトゲノムの柔軟性地図を図 1.1 に示す。なお、セントロメア領域ならびにサブセントロメア領域に関しては、その配列情報を塩基配列データベースから得ることができないため、計算の対象から除いている。ヒトゲノムの柔軟性地図を鳥瞰すると、全体的な DNA の柔軟性は、染色体間で一定の範囲に落ち着くものの、個々の染色体はそれぞれに固有の DNA の柔軟性をもつことが分かる。大変興味深いことに、“異常に柔らかい DNA 領域”がすべての染色体において発見された。この点以外にも、今後のさらなる解析を通して重要な知見を得られる可能性があるが、本研究においては、この極めて柔らかい DNA 領域に焦点を当てて解析を行なった。





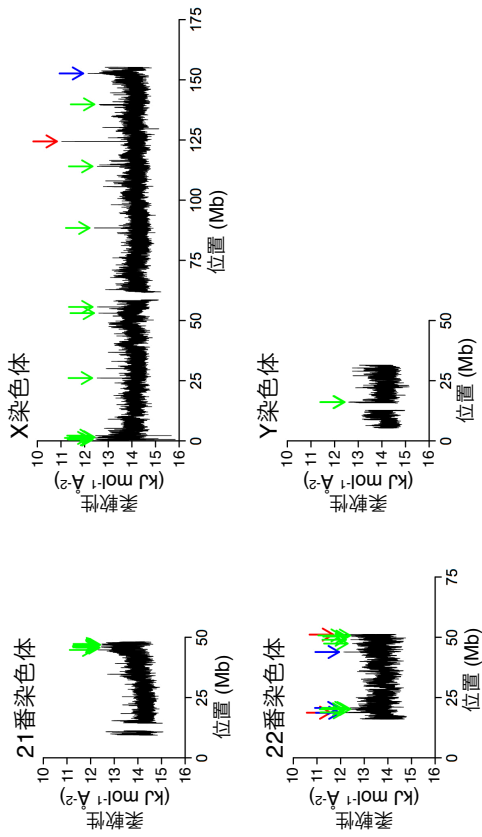


図 1.1. ヒトゲノムの柔軟性地図

1 番染色体から Y 染色体までの全 24 本の染色体の DNA の柔軟性を示す。Packer らが報告している柔軟性パラメータセット [15] (表 1.1) を用いて、DNA の柔軟性を計算した。この柔軟性パラメータセットにおいては、値が小さいほど柔軟性が高いことを意味する。ウィンドウの大きさを 10 kb、ウィンドウサイズインデックスの大きさを 100 bp に設定して計算した。各々のウィンドウの平均値を、各々のウィンドウの始点の位置にプロットした。なお、セントロメア領域ならびにサブセントロメア領域に関しては、その配列情報を塩基配列データベースから得ることができなため、計算の対象から除いている。極めて柔らかい DNA 領域 (SPIKE) のうち、第 1 グループに属する SPIKE (0%分位点から 0.005%分位点までの範囲におさまる柔軟性をもつ SPIKE) を赤の矢印で、第 2 グループに属する SPIKE (0.005%分位点から 0.01%分位点までの範囲におさまる柔軟性をもつ SPIKE) を青の矢印で、第 3 グループに属する SPIKE (0.01%分位点から 0.05%分位点までの範囲におさまる柔軟性をもつ SPIKE) を緑の矢印で示した。

1.3.2. 真核生物ゲノムに存在する異常に柔らかい DNA 領域

ヒトゲノムには、極めて柔らかい DNA 領域 (SPIKE と命名) が存在することが明らかになった (図 1.1)。SPIKE についての詳細な解析を行なうため、特に柔軟な SPIKE (上位 0.05%) を選出し、これらを柔軟性の高さに基づいて 3 つのグループへと分類した (“1.2.2. SPIKE と rSPIKE の選出” に詳述)。最も柔軟な第 1 グループには、0%分位点から 0.005%分位点までの範囲におさまる柔軟性をもつ SPIKE が属する [Packer らが報告している柔軟性パラメータセット [15] (表 1.1) においては、値が小さいほど柔軟性が高いことを意味する]。2 番目に柔軟な第 2 グループには、0.005%分位点から 0.01%分位点までの範囲におさまる柔軟性をもつ SPIKE が属する。3 番目に柔軟な第 3 グループには、0.01%分位点から 0.05%分位点までの範囲におさまる柔軟性をもつ SPIKE が属する。表 1.2 に示したように、第 1 グループの SPIKE (全 23 本) は、1 番染色体から X 染色体までの計 12 本の染色体にみられた。多くの場合において、SPIKE (を構成する DNA 領域) は 10 ~ 15 kb の長さをもつ。また、各々の SPIKE の周辺領域 (0.5 Mb) に存在する遺伝子 [3] を表 1.2 に示したが、これらの遺伝子に共通した特徴はみられなかった。

次に、SPIKE と遺伝子との位置関係について解析したところ、第 1 グループに属する全 23 本の SPIKE のうち、14 本が遺伝子間領域に位置し、7 本が少なくともひとつのエキソンと重なり、残る 2 本がイントロンに位置することが明らかになった (表 1.2)。ここで、ヒトゲノムの柔軟性地図を作成するために用いた 10 kb というウィンドウの大きさ (“1.2.1. DNA の柔軟性の計算” を参照) は、約 3.1 Gb にも及ぶ長大なヒトゲノムを俯瞰的にみるという目的と合致する一方で、注目する DNA 領域を詳細に解析するには大き過ぎると考えられる。そこで、第 1 グループに属する SPIKE のうち、エキソンと重なる 7 本の SPIKE を対象として、より分解能の高い柔軟性地図 (ウィンドウの大きさ: 1 kb) を作成した (図 1.2)。その結果、分解能を高めた場合に、極めて柔らかい DNA 領域はエキソンと重ならないことが判明した。例えば、 α -グルコシダーゼをコードする *MGAM* 遺伝子周辺の柔軟性をみると (図 1.2.A)、極めて柔らかい DNA

領域はエクソンと重ならないことが分かる（他の 6 本の SPIKE についても同様のことが言える）。

また、第 2 グループの SPIKE の位置、長さ、遺伝子との位置関係、および SPIKE の周辺領域に存在する遺伝子を表 1.3 に示した。第 1 グループの SPIKE と第 2 グループの SPIKE との違いは、主に柔軟性の大きさのみにある。一方、第 3 グループの SPIKE は、テロメア領域に高頻度に存在する傾向がみられた。具体的に述べると、第 3 グループの SPIKE は、1 番染色体の 0 – 4 Mb の領域、2 番染色体の 238 – 243 Mb の領域、4 番染色体の 0 – 4 Mb の領域、5 番染色体の 0 – 2 Mb および 176 – 181 Mb の領域、7 番染色体の 0 – 2 Mb および 155 – 159 Mb の領域、8 番染色体の 143 – 146 Mb の領域、9 番染色体の 137 – 141 Mb の領域、10 番染色体の 134 – 135 Mb の領域、11 番染色体の 0 – 3 Mb の領域、12 番染色体の 131 – 134 Mb の領域、13 番染色体の 113 – 115 Mb の領域、14 番染色体の 104 – 107 Mb の領域、16 番染色体の 0 – 3 Mb および 88 – 90 Mb の領域、17 番染色体の 77 – 82 Mb の領域、20 番染色体の 60 – 63 Mb の領域、21 番染色体の 44 – 48 Mb の領域、22 番染色体の 44 – 51 Mb の領域、ならびに X 染色体の 0 – 3 Mb の領域において、特に高頻度に存在していた（図 1.1）。

さらに、比較的規則的に配置された“小さな SPIKE”がヒトゲノムに存在することが明らかになった。例として、21 番染色体の DNA の柔軟性を図 1.3 に示す。この小さな SPIKE は、21 番染色体において約 4 Mb の間隔で存在していた。このような比較的規則的な SPIKE は、その間隔や規則性がつづく範囲の大きさが一定ではないものの、他のいくつかの染色体においてもみられた。

続いて、SPIKE について生物種間の比較を行なうために、マウス (*M. musculus*)、ショウジョウバエ (*D. melanogaster*)、ならびに線虫 (*C. elegans*) のゲノムを対象として、柔軟性地図を作成した。図 1.4、図 1.5 および図 1.6 に示したように、ヒト以外の真核生物のゲノムにも SPIKE が存在することが明らかになった。ヒトの SPIKE と同様に、これらの生物種の SPIKE の周辺領域に存在する遺伝子に共通した特徴はみられなかった。しかし、これらの生物種の柔軟性地図とヒトゲノムの柔軟性地図を比較することで、いくつかの興味深いことが明らか

になった。第一に、本研究で解析の対照とした真核生物ゲノムの中で、ヒトゲノムにおいて最も高頻度に SPIKE が存在することが判明した (図 1.7.A)。第二に、線虫のゲノムに、極めて硬い DNA 領域 [rSPIKE (rigid SPIKE) と命名] が高頻度に存在することが判明した (図 1.6)。興味深いことに、rSPIKE は線虫のゲノムに限らずショウジョウバエやマウスのゲノムにも比較的高頻度に存在しており、これらのゲノムに存在する頻度はヒトゲノムに存在する頻度よりも高かった (図 1.7.B; ヒトの rSPIKE については図 1.8 を参照)。さらに、rSPIKE が存在する頻度と進化的な位置づけとの間には、負の相関関係があることが明らかになった (図 1.7.B)。第三に、SPIKE と rSPIKE の総数に対する SPIKE の数の割合と、進化的な位置づけとの間に、正の相関関係があることが判明した (図 1.7.C)。

表 1.2. 第1グループの SPIKE の位置とその周辺に存在する遺伝子

第1グループ SPIKE 番号	染色体 番号	始点 (bp)	終点 (bp)	長さ (bp)	遺伝子との 位置関係	SPIKE の周辺領域 (0.5 Mb) に存在する遺伝子 ^a
1	1	2,605,701	2,633,000	27,300	イントロン	PLCH2, PANK4, HES5, TNFRSF14, C1orf93, MMEL1, TTC34
2	1	53,555,501	53,571,700	16,200	エキソンと重複	ECHDC2, SCP2, PODN, SLC1A7 ^b , CPT2, C1orf123, MAGOH, LRP8, ZYG11A
3	1	56,121,901	56,138,900	17,000	遺伝子間領域	
4	2	10,146,501	10,158,900	12,400	遺伝子間領域	TAF1B, GRHL1, KLF11, CYS1, AC104794.4, RRM2, C2orf48
5	2	126,642,401	126,657,600	15,200	遺伝子間領域	
6	3	195,195,101	195,237,000	41,900	遺伝子間領域	ACAP2, PPP1R2, APOD, MUC20, C3orf21
7	5	117,868,201	117,885,500	17,300	遺伝子間領域	
8	7	141,735,601	141,751,800	16,200	エキソンと重複	PRSS37, MGAM ^b , OR9A4, CLEC5A, TAS2R38, RP11-1220K2.2, PRSS58
9	7	158,119,101	158,136,900	17,800	イントロン	
10	8	1,340,001	1,358,100	18,100	遺伝子間領域	DLGAP2
11	8	135,472,001	135,482,700	10,700	遺伝子間領域	ZFAT
12	11	114,471,701	114,482,600	10,900	遺伝子間領域	C11orf71, RBM7, REXO2, FAM55A, FAM55D, FAM55B
13	12	502,65,001	50,281,500	16,500	エキソンと重複	FMNL3, TMBIM6, NCKAP5L, BCDIN3D, FAIM2 ^b , AQP2, AQP5, AQP6, RACGAP1, ACCN2, SMARCD1, GPD1, C12orf62, AC074032.1, PRPF40B
14	12	79,849,501	79,863,200	13,700	遺伝子間領域	PAWR, SYTI
15	12	132,632,801	132,650,300	17,500	エキソンと重複	PUS1, EP400, EP400NL, DDX51, NOC4L ^b , GALNT9, ULK1
16	14	106,027,301	106,039,700	12,400	遺伝子間領域	TEX22, MTAI, CRIP2, CRIP1, C14orf80, TMEM121, PACS2
17	14	106,147,501	106,159,800	12,300	遺伝子間領域	CRIP2, CRIP1, C14orf80, TMEM121, AB019441.33, MTAI
18	16	1,780,201	1,795,700	15,500	エキソンと重複	TELO2, IFT140, TMEM204, CRAMP1L, HNIL, MAPK8IP3 ^b , NME3, MRPS34, EME2, SPSB3, NUBP2, IGFALS, HAGH, FAHD1, C16orf73, HS3ST6, SEPX1, RPL3L, NDUFB10, RPS2, RNF151, TBL3, NOXO1, GFER, PTX4
19	16	10,792,101	10,802,500	10,400	遺伝子間領域	EMP2, TEK15, NUBP1, FAM18A, CIITA, DEXI, CLEC16A ATF7IP2
20	16	88,304,201	88,320,900	16,700	遺伝子間領域	AC134312.1, ZNF469, ZFPFMI, BANP
21	22	18,711,001	18,725,800	14,800	エキソンと重複	PEX26, TUBA8, USP18, AC008132.1 ^b , AC008132.13, AC008103.1, DGCR6, PRODH, XXbac-B461K10.4, MICAL3
22	22	51,111,401	51,124,000	12,600	エキソンと重複	SBF1, ADM2, MIOX, LMF2, NCAPH2, SCO2, TYMP, ODF3B, KLHDC7B, C22orf41, CPT1B, CHKB, MAPK8IP2, ARSA, SHANK3 ^b , ACR, RABL2B, PPP6R2
23	X	124,402,401	124,418,300	15,900	遺伝子間領域	

^a 参考文献 3^b 図 1.2

表 1.3. 第 2 グループの SPIKE の位置とその周辺に存在する遺伝子

第 2 グループ SPIKE 番号	染色体 番号	始点 (bp)	終点 (bp)	長さ (bp)	遺伝子との 位置関係	SPIKE の周辺領域 (0.5 Mb) に存在する遺伝子 ^a
1	2	1,218,401	1,229,500	11,100	イントロン	TPO, SNTG2
2	2	239,441,201	239,453,000	11,800	遺伝子間領域	TRAF3IP1, ASB1, PER2
3	4	31,810,701	31,820,900	10,200	遺伝子間領域	
4	5	135,122,201	135,132,600	10,400	遺伝子間領域	CXCL14, SLC25A48, IL9, LECT2, TGFB1
5	7	99,183,801	99,196,100	12,300	遺伝子間領域	ARPC1B, PDAP1, BUD31, PTCDD1, ATP5J2-PTCD1, CPSF4, AC073063.1, ATP5J2, ZNF789, ZNF394, ZKSCAN5, FAM200A, ZNF655, ZNF498, CYP3A5, CYP3A7, CYP3A4, CYP3A43, ARPC1A
6	8	29,175,701	29,188,200	12,500	遺伝子間領域	DUSP4, KIF13B
7	8	143,298,701	143,313,300	14,600	エキソンと重複	TSNARE1, BAI1
8	8	145,819,701	145,831,900	12,200	エキソンと重複	GSI-393G12.10, GPR172A, FBXL6, ADCK5, CPSF1, SLC39A4, VPS28, TONSL, CYHR1, CTD-2517M22.16, KIFC2, FOXH1, PPP1R16A, CTD-2517M22.14, GPT, MFSD3, LRRRC14, LRRRC24, C8orf82, ARHGAP39, AF186192.1, AF186192.2, ZNF251, ZNF34, RPL8, ZNF517, ZNF7, COMMD5
9	9	139,471,101	139,484,700	13,600	遺伝子間領域	DNLZ, CARD9, SNAPC4, SDCCAG3, PMPCA, INPP5E, SEC16A, C9orf163, NOTCH1, EGFL7, AGPAT2, FAM69B, LCN6, LCN10, LCN8, LCN15, TMEM141, KIAA1984, C9orf86, GPSM1
10	10	134,630,301	134,644,100	13,800	イントロン	NKX6-2, C10orf93, GPR123, INPP5A
11	11	1,260,101	1,276,200	16,100	エキソンと重複	MUC2, MUC5AC, MUC5B, TOLLIP, BRSK2, AC091196.1, MUC6
12	12	40,103,201	40,113,400	10,200	エキソンと重複	ABCD2, C12orf40, SLC2A13
13	17	48,096,601	48,108,300	11,700	遺伝子間領域	KAT7, TAC4, DLX4, DLX3, ITGA3, PDK2, SAMD14, SGCA, COL1A1, TMEM92
14	19	21,049,301	21,069,300	20,000	遺伝子間領域	ZNF85, ZNF430, ZNF714, ZNF626
15	19	21,081,701	21,099,300	17,600	遺伝子間領域	ZNF85, ZNF430, ZNF714, ZNF431, ZNF626
16	22	18,680,201	18,694,000	13,800	遺伝子間領域	PEX26, TUBA8, USP18, AC008132.1, AC008132.13, AC008103.1, DGCR6, PRODH, XXbac-B461K10.4, MICAL3
17	22	20,683,701	20,697,600	13,900	遺伝子間領域	RIMBP3, USP41, ZNF74, SCARF2, KLHL22, MED15
18	22	43,887,001	43,901,600	14,600	エキソンと重複	MPPED1, EFCAB6, SCUBE1
19	X	152,647,601	152,664,700	17,100	遺伝子間領域	MAGEA1, ZNF275, TREX2, HAUS7, U82695.9, BGN, ATP2B3, FAM58A

^a 参考文献 3

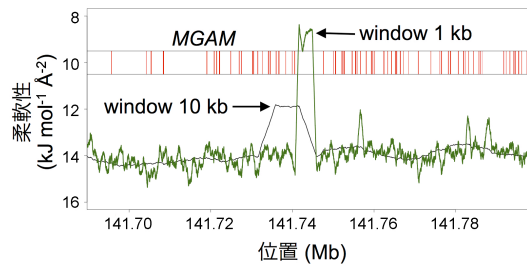
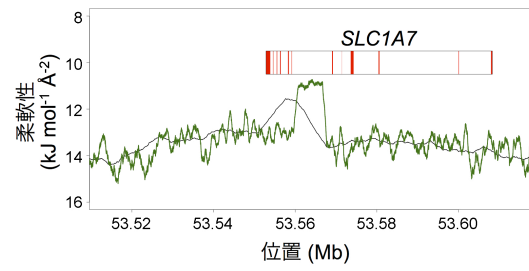
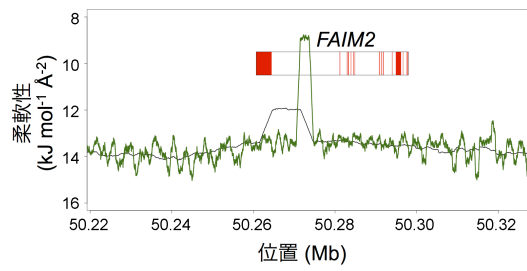
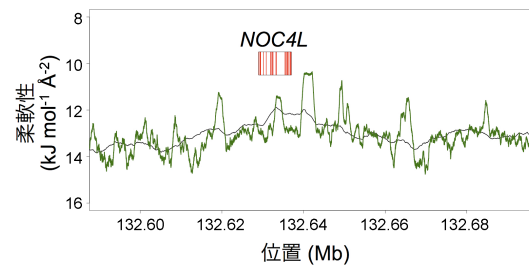
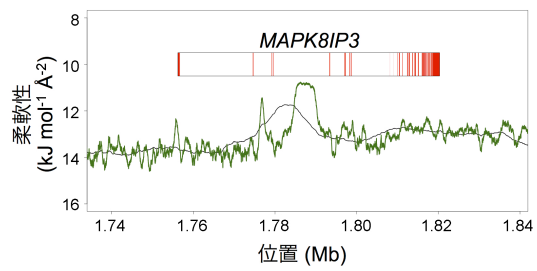
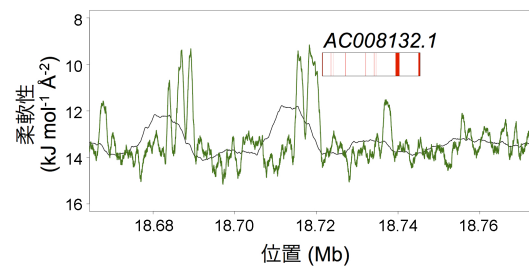
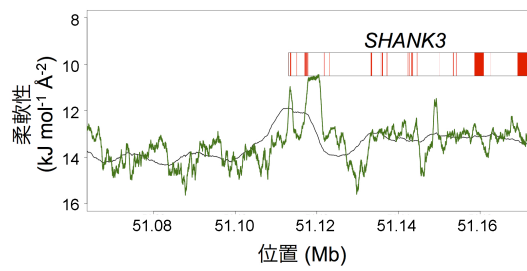
A**B****C****D****E****F****G**

図 1.2. SPIKE とエキソンの関係

第 1 グループに属する SPIKE のうち、エキソンと重なる 7 本の SPIKE を対象として、より分解能の高い柔軟性地図（ウィンドウの大きさ：1 kb）を作成した。*MGAM* 遺伝子 (A)、*SLC1A7* 遺伝子 (B)、*FAIM2* 遺伝子 (C)、*NOC4L* 遺伝子 (D)、*MAPK8IP3* 遺伝子 (E)、*AC008132.1* 遺伝子 (F)、ならびに *SHANK3* 遺伝子 (G) の各々の周辺領域の柔軟性（ウィンドウの大きさ：1 kb）を緑色の線で示した。また、比較のために、ウィンドウの大きさを 10 kb としたときの柔軟性を黒線で示した。四角は各々の遺伝子を表わし、赤線はエキソンを表わす。なお、解析に関連しない遺伝子は示していない。

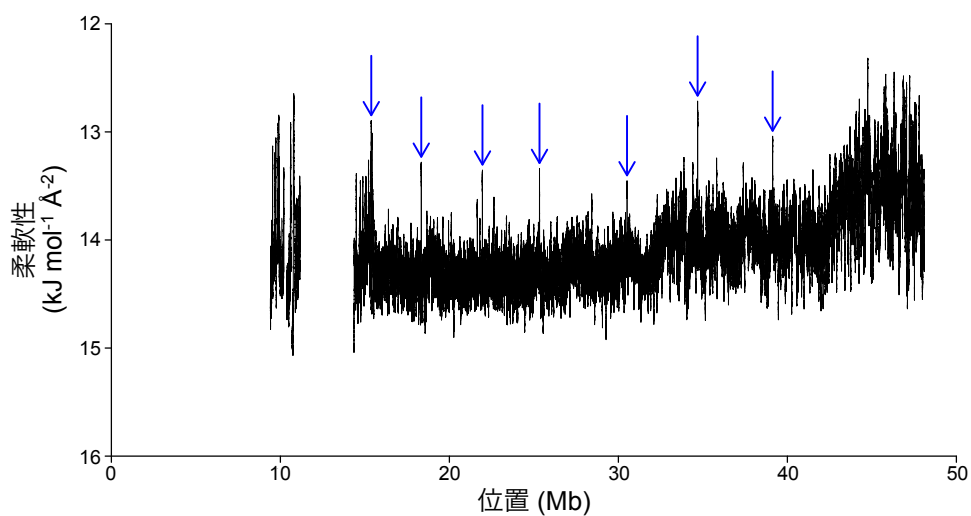
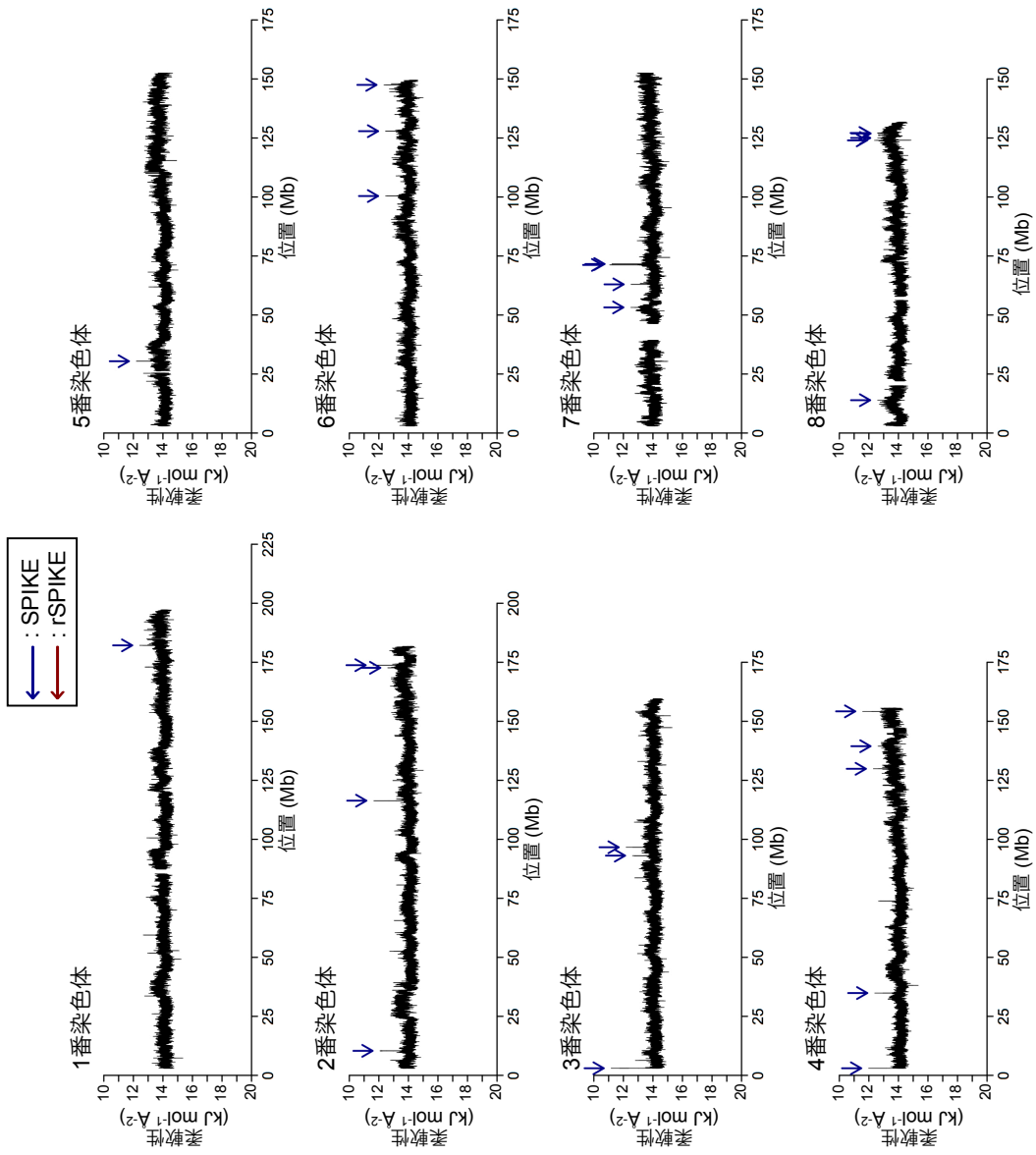
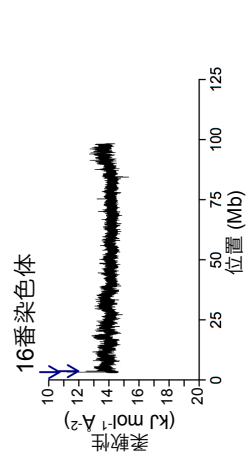
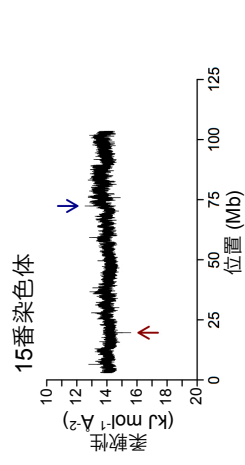
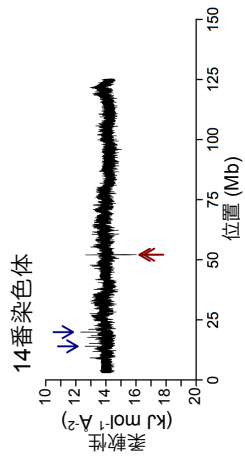
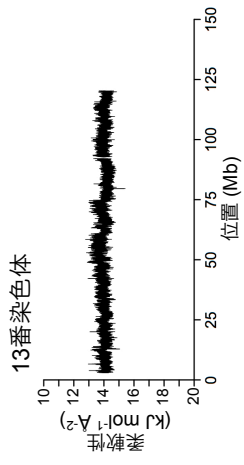
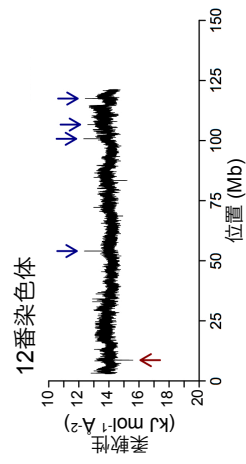
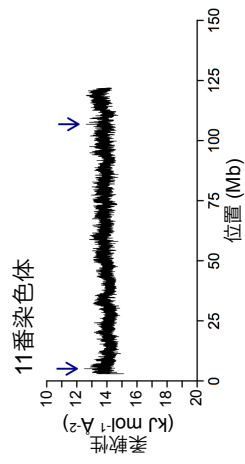
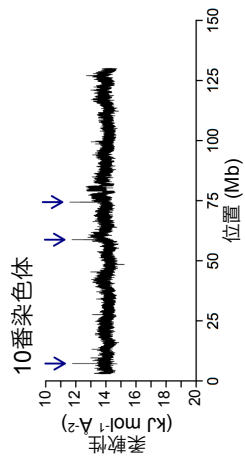
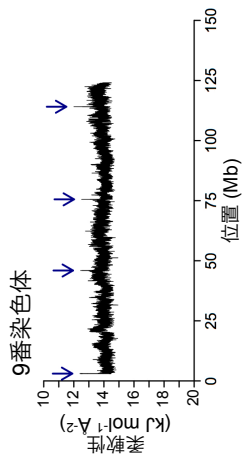


図 1.3. ヒト 21 番染色体の DNA の柔軟性

極めて柔軟な DNA 領域を矢印で示した。これらの領域も SPIKE に含まれるが、第 1、第 2 および第 3 グループに属する SPIKE と比べて、これらの領域がもつ柔軟性は低い。これらの SPIKE は、ヒト 21 番染色体において約 4 Mb の間隔で存在する。ウィンドウの大きさを 10 kb、ウィンドウスライディングステップの大きさを 100 bp に設定して計算した。





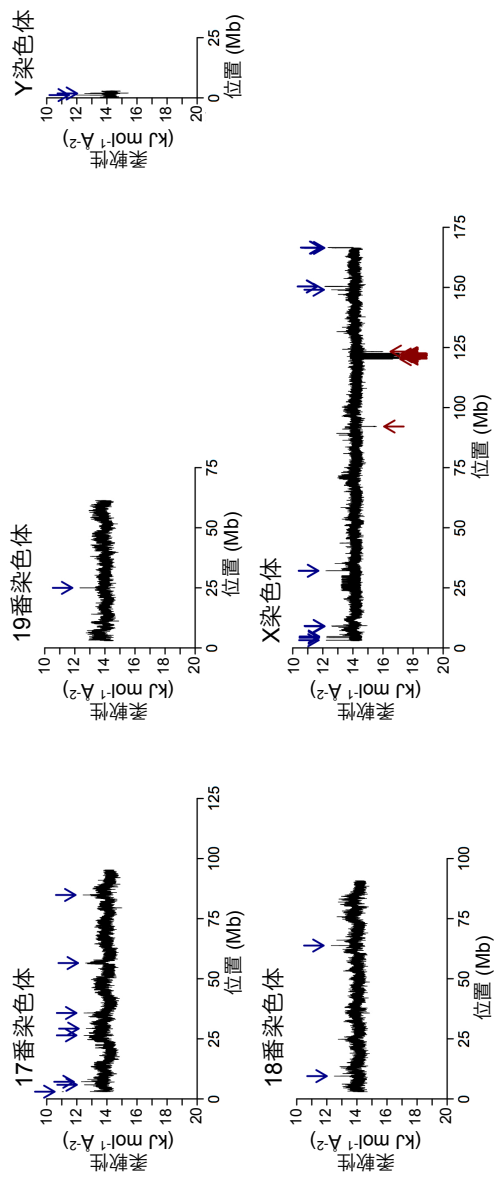


図 1.4. マウスゲノムの柔軟性地図

1 番染色体から Y 染色体までの全 21 本の染色体の DNA の柔軟性を示す。極めて柔らかい DNA 領域 (SPIKE) を濃紺の矢印で、極めて硬い DNA 領域 (rSPIKE) を濃赤の矢印で示した。ウインドウの大きさを 10 kb、ウインドウスライディングステップの大きさを 100 bp に設定して計算した。

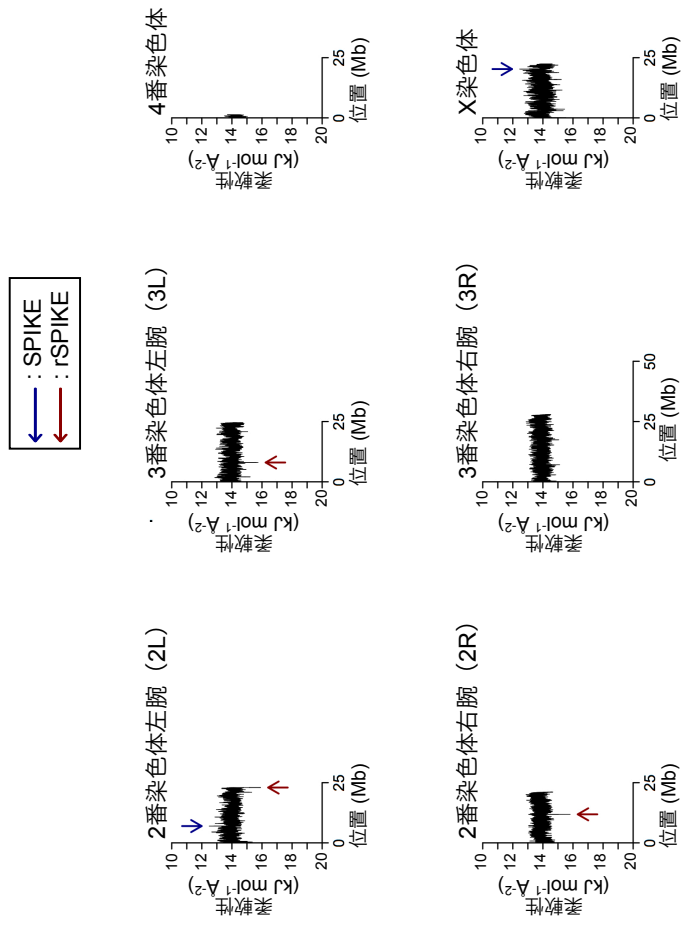


図 1.5. ショウジョウバエゲノムの柔軟性地図

1 番染色体から X 染色体までの全 6 本の染色体の DNA の柔軟性を示す。極めて柔らかい DNA 領域 (SPIKE) を濃紺の矢印で、極めて硬い DNA 領域 (rSPIKE) を濃赤の矢印で示した。ウインドウの大きさを 10 kb、ウインドウスライディングステップの大きさを 100 bp に設定して計算した。

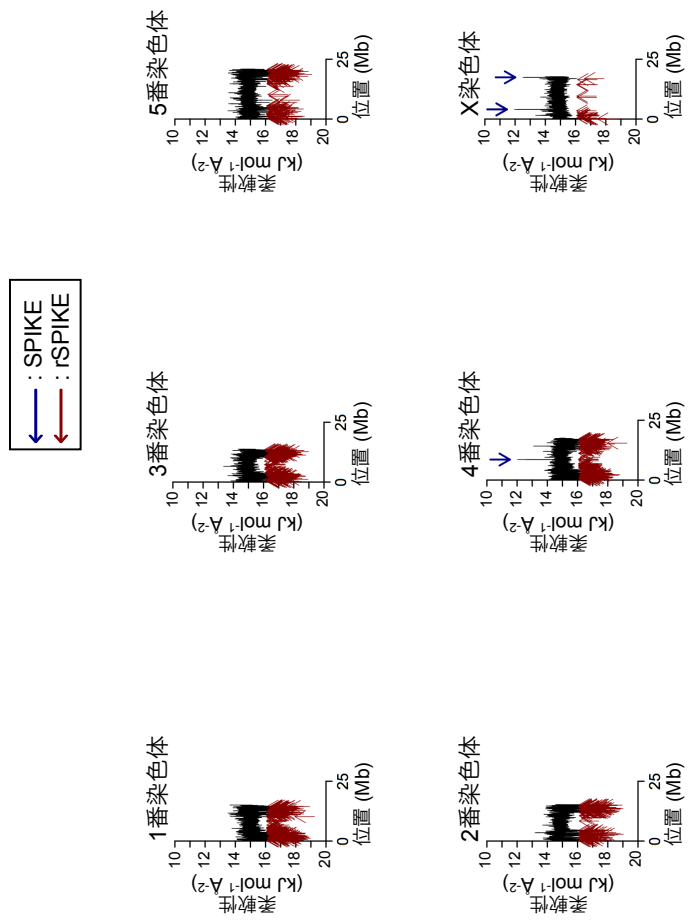


図 1.6. 線虫ゲノムの柔軟性地図

1 番染色体から X 染色体までの全 6 本の染色体の DNA の柔軟性を示す。極めて柔らかい DNA 領域 (SPIKE) を濃紺の矢印で、極めて硬い DNA 領域 (rSPIKE) を濃赤の矢印で示した。ウインドウの大きさを 10 kb、ウインドウスライディングステップの大きさを 100 bp に設定して計算した。

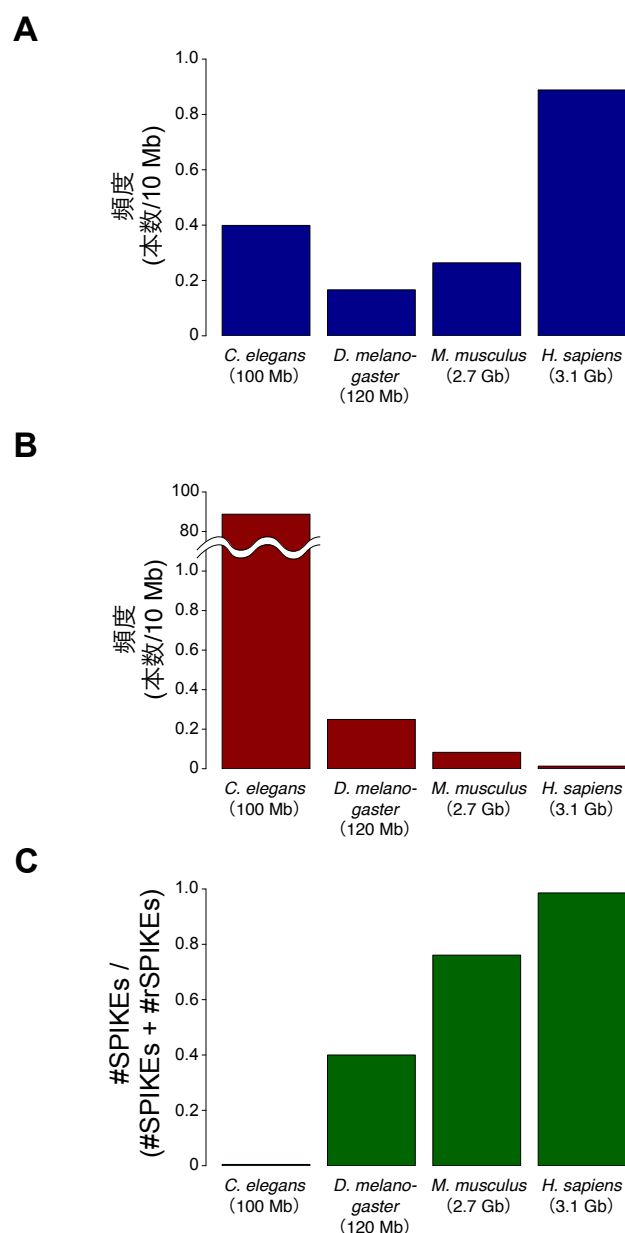


図 1.7. 真核生物ゲノムにおいて SPIKE と rSPIKE が存在する頻度の比較

(A, B) ヒト、マウス、ショウジョウバエ、線虫のゲノムにおいて SPIKE (A) および rSPIKE (B) が存在する頻度の比較

10 Mb あたりの SPIKE の数 (A) または 10 Mb あたりの rSPIKE の本数 (B) を示した。

(C) SPIKE と rSPIKE の総数に対する SPIKE の数の割合

図中の #SPIKE は SPIKE の総数を表わし、#rSPIKE は rSPIKE の総数を表わす。なお、各々の生物種のゲノムの大きさを括弧の中に示した。

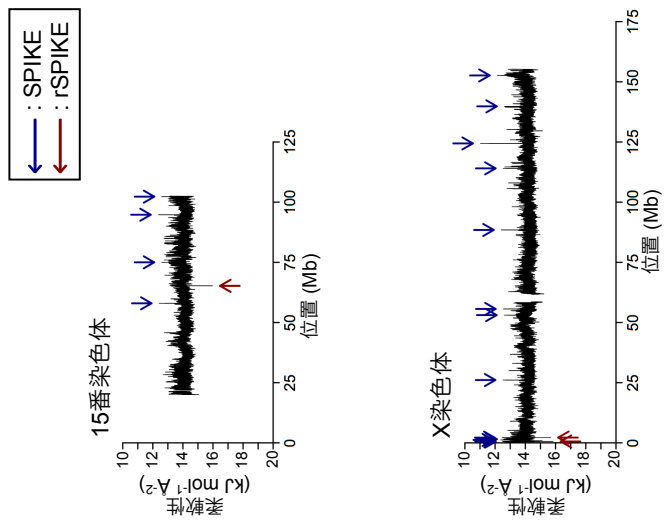


図 1.8. rSPIKE をもつヒト染色体の DNA の柔軟性

極めて柔らかい DNA 領域 (SPIKE) を濃紺の矢印で、極めて硬い DNA 領域 (rSPIKE) を濃赤の矢印で示した。ウインドウの大きさを 10 kb、ウインドウスライディングステップの大きさを 100 bp に設定して計算した。

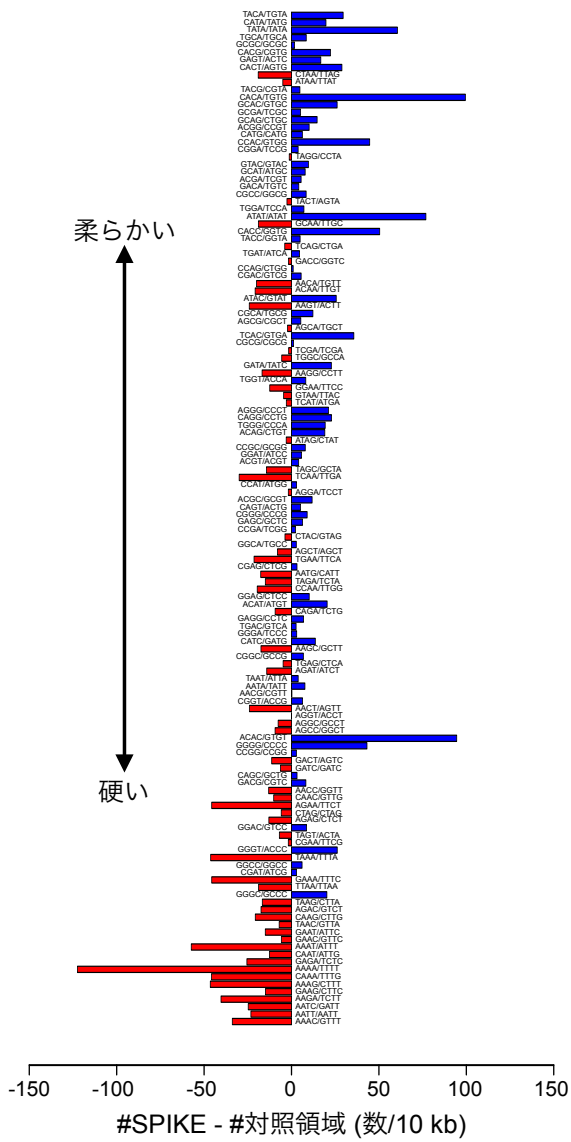
1.3.3. SPIKE と rSPIKE の塩基配列の特徴

SPIKE を構成する塩基配列の特徴を明らかにするため、はじめに、ヒトの SPIKE を対象として、各テトラヌクレオチド（全 136 種類）の出現頻度を解析した。その結果、SPIKE は、TATA/TATA、CACA/TGTG、ATAT/ATAT、ならびに ACAC/GTGT [全てのテトラヌクレオチド（全 136 種類）における各配列の柔軟性は、それぞれ 3 番目、12 番目、28 番目、98 番目に高い [15]（表 1.1）] を、対照領域（非 SPIKE 領域）に比べて高い頻度でもつことが判明した（図 1.9.A）。一方でまた SPIKE は、AAAA/TTTT [8 番目に硬いテトラヌクレオチド [15]（表 1.1）] を、対照領域に比べて低い頻度でもつことが判明した。興味深いことに、SPIKE が高い頻度でもつ TATA/TATA、CACA/TGTG、ATAT/ATAT、ACAC/GTGT はすべて、ジヌクレオチドの繰り返しから成るテトラヌクレオチドである。このことから、SPIKE は柔軟性の高いマイクロサテライトを豊富にもつことが示唆される。そこで、塩基配列から繰り返し配列を発見するコンピュータプログラムである RepeatMasker [17] を用いて、SPIKE に含まれる繰り返し配列を同定した。その結果、予想通り、SPIKE は (CA/TG) n や (TA/TA) n のような柔らかいマイクロサテライトを豊富にもつことが明らかになった。また、Packer らが報告しているテトラヌクレオチドの柔軟性パラメータセット [15]（表 1.1）を用いて計算すると、サテライト配列である SATR1 や SATR2、ならびにレトロウイルス由来の塩基配列である HERVL-int は、非常に柔軟な塩基配列であることが分かる（図 1.9.B）。以上の解析から、柔軟な繰り返し配列が SPIKE を構成していることが判明した（柔軟な繰り返し配列が SPIKE の形成に寄与する程度については、“1.4. 考察”を参照）。

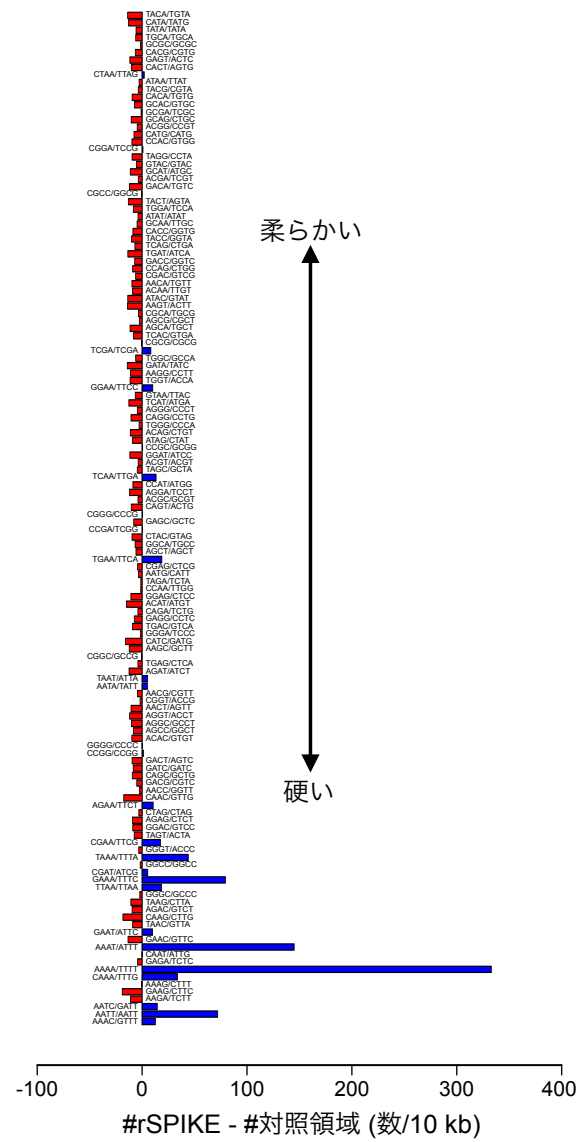
次に、線虫の rSPIKE を対象として同様の解析を行なった。その結果、rSPIKE は、AATT/AATT、AAAA/TTTT、AAAT/ATTT、ならびに GAAA/TTTC [全てのテトラヌクレオチド（全 136 種類）における各配列の柔軟性は、それぞれ 135 番目、129 番目、126 番目、117 番目 [15] に高い（表 1.1）] を、対照領域（非 rSPIKE 領域）に比べて高い頻度でもつことが明らかになった（図 1.9.C）。また、RepeatMasker [17] を用いた解析の結果、rSPIKE が、アデニン（A）またはチミン（T）に富んだ塩基配列 [極めて硬いテトラヌクレオチドである

AAAA/TTTT を豊富に含むと推察される [15] (表 1.1)]、トランスポゾンである HelitronY1A_CE、機能が未知の CERP3、ミニサテライトである MINISAT1_CE、ならびにトランスポゾン関連配列である CeRep5 を豊富にもつことが判明した (図 1.9.D)。以上の解析から、硬い繰り返し配列が rSPIKE を構成していることが判明した (硬い繰り返し配列が rSPIKE の形成に寄与する程度については、“1.4. 考察”を参照)。

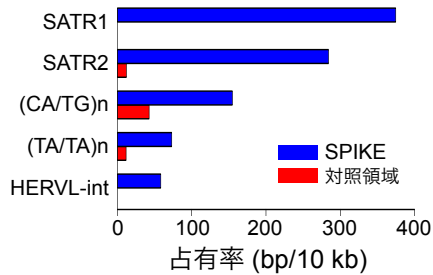
A



C



B



D

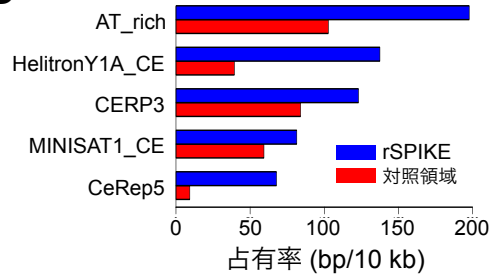


図 1.9. SPIKE および rSPIKE を構成する塩基配列の特徴

(A) ヒトの SPIKE における各々のテトラヌクレオチドの出現頻度

SPIKE に含まれる各々のテトラヌクレオチドの出現回数と SPIKE の長さを基に、各々のテトラヌクレオチドについて、SPIKE を構成する塩基配列 10 kb あたりの出現回数と対照領域 10 kb あたりの出現回数の差 [SPIKE を構成する塩基配列 10 kb あたりの出現回数 - 対照領域 10 kb あたりの出現回数] を求めた。対照領域については、“1.2.3. SPIKE と rSPIKE に含まれるテトラヌクレオチドの出現頻度の解析” に詳述した。柔軟性の高いテトラヌクレオチドを上方に、柔軟性の低いテトラヌクレオチドを下方に示した [15] (表 1.1)。

(B) ヒトの SPIKE に高い割合で存在する繰り返し配列

上位 5 つの繰り返し配列を示した。SPIKE に含まれる繰り返し配列を、RepeatMasker [17] を用いて同定した。SPIKE を構成する塩基配列のうち、これら上位 5 つの繰り返し配列が占める割合は、10 kb あたり約 950 塩基対であった (SPIKE に存在する全ての繰り返し配列が占める割合は、SPIKE を構成する塩基配列 10 kb あたり約 2,880 塩基対であった)。

(C) 線虫の rSPIKE における各々のテトラヌクレオチドの出現頻度

rSPIKE に含まれる各々のテトラヌクレオチドの出現回数と rSPIKE の長さを基に、各々のテトラヌクレオチドについて、rSPIKE を構成する塩基配列 10 kb あたりの出現回数と対照領域 10 kb あたりの出現回数の差 [rSPIKE 10 kb を構成する塩基配列あたりの出現回数 - 対照領域 10 kb あたりの出現回数] を求めた。対照領域については、“1.2.3. SPIKE と rSPIKE に含まれるテトラヌクレオチドの出現頻度の解析” に詳述した。柔軟性の高いテトラヌクレオチドを上方に、柔軟性の低いテトラヌクレオチドを下方に示した [15] (表 1.1)。

(D) 線虫の rSPIKE に高い割合で存在する繰り返し配列

上位 5 つの繰り返し配列を示した。rSPIKE に含まれる繰り返し配列を、RepeatMasker [17] を用いて同定した。rSPIKE を構成する塩基配列のうち、これら上位 5 つの繰り返し配列が占める割合は、10 kb あたり約 600 塩基対であった (rSPIKE に存在する全ての繰り返し配列が占める割合は、rSPIKE を構成する塩基配列 10 kb あたり約 2,090 塩基対であった)。

1.4. 考察

本研究では、ヒト、マウス、ショウジョウバエ、線虫のゲノム全域を対象として柔軟性地図を作成した。その結果、これらのゲノムの大部分の領域がもつ柔軟性は約 $13 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ から約 $15 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ の範囲におさまるものの、各々の生物種のゲノムはそれぞれに固有の柔軟性をもつことが明らかになった (図 1.1、図 1.4、図 1.5、図 1.6)。また、解析したすべての生物種のゲノムが、極めて柔らかい DNA 領域である SPIKE をもつことが判明した。SPIKE は、マウス、ショウジョウバエ、線虫のゲノムと比較して、ヒトゲノムに最も高頻度に存在した (図 1.7A)。また、ヒトの SPIKE は、対照領域 (非 SPIKE 領域) に比べて柔軟性の高いテトラヌクレオチドを豊富にもつとともに、柔軟性の高いサテライト配列、マイクロサテライト、ならびにレトロウイルス由来の配列を豊富に含んでいた (図 1.9A, B)。このような繰り返し配列は、当然 SPIKE の形成に寄与するものと考えられるが、その寄与の程度は限られていると考えられる。実際に、SPIKE に存在する全ての繰り返し配列がもつ柔軟性の平均値は $12.49 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ であり、SPIKE の選出に用いた基準値である $12.6 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ を僅かに下回る程度であった [Packer らが報告している柔軟性パラメータセット [15] (表 1.1) においては、値が小さいほど柔軟性が高いことに注意]。さらに、SPIKE を構成する塩基配列の大部分は、非反復性の塩基配列であることを考慮に入れると (SPIKE に存在する全ての繰り返し配列が占める割合は、SPIKE を構成する塩基配列 10 kb あたり約 2,880 塩基対である)、柔軟性の高い非反復性の塩基配列が、SPIKE を形成する主要因であると考えられる。

現在のところ、SPIKE がなんらかの生物学的機能をもつのか否かは不明である。しかしながら、SPIKE が存在する頻度と進化的な位置づけとの間に相関関係があること、ならびにヒトゲノムが非常に多くの SPIKE をもつことを考慮に入れると (図 1.7A)、SPIKE はゲノムの折りたたみ機構に関与している可能性が考えられる。例えば、ヒトゲノムは長さにして約 20 万倍も圧縮されて細胞核に納められていることから、SPIKE は、ゲノムの折りたたみにおける折り返し点として機能することが推察される。したがって、上述した繰り返し配列および柔軟性の高い非反復性の塩基配列は、このような機能を担うことを目的とし

て進化したとも解釈され得る。さらに、SPIKE が間期染色体および分裂期染色体の基本骨格を形成する構造因子として機能することや、SPIKE が隣り合うクロマチンドメイン間の境界構造として機能することが考えられる。一般的に、ひとつのクロマチンドメインは数 Mb の DNA 領域から構成されるが [18-22]、大変興味深いことに、この大きさは隣り合う SPIKE の間隔と非常に類似している。また、テロメア領域に高頻度に存在する第 3 グループの SPIKE も、ゲノムの折りたたみ機構へ関与している可能性が考えられる。

極めて硬い DNA 領域である rSPIKE は、線虫のゲノムに最も高頻度に存在する一方で、ヒトゲノムに最も低い頻度で存在した。また、線虫の rSPIKE は、トランスポゾンやトランスポゾン関連配列を含む、硬い繰り返し配列を豊富に含んでいた (図 1.9D)。このような繰り返し配列は、当然 rSPIKE の形成に寄与するものと考えられるが、その寄与の程度は限られていると考えられる。実際、rSPIKE に存在する全ての繰り返し配列がもつ柔軟性の平均値は $16.11 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ であり、rSPIKE の選出に用いた基準値である $15.6 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ を上回るものの [Packer らが報告している柔軟性パラメータセット [15] (表 1.1) においては、値が小さいほど柔軟性が高いことに注意]、rSPIKE を構成する塩基配列の大部分は非反復性の塩基配列であることに加えて (rSPIKE に存在する全ての繰り返し配列が占める割合は、rSPIKE を構成する塩基配列 10 kb あたり約 2,090 塩基対である)、線虫のゲノムがもつ柔軟性の平均値が $15.03 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ であることを考慮に入れると、非反復性の硬い塩基配列が、rSPIKE を形成する主要因であると考えられる [例えば、10 kb の長さをもつ rSPIKE が、2,090 塩基対の硬い繰り返し配列と、平均的な柔軟性をもつ 7,910 塩基対の非反復性の塩基配列から成ると仮定した場合、この DNA 領域がもつ柔軟性の平均値は、 $(16.11 \times 2,090 + 15.03 \times 7,910) / 10,000 = \text{約 } 15.26 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ となり、rSPIKE の選出に用いた基準値である $15.6 \text{ (kJ mol}^{-1} \text{ \AA}^{-2})$ を下回る]。

第 2 部

ヌクレオソーム **DNA** に共通した物理的特性

2.1. 緒言

ゲノム DNA は高度に圧縮されて細胞核に納められている。例えば、ヒトの場合、父親のゲノムと母親のゲノムのそれぞれに由来する合計 2 m もの長さにあふ DNA は、長さにして約 20 万倍も圧縮されて微小な細胞核 (直径約 10 μm) に納められている。ゲノム科学の著しい発展にも拘らず、このようなゲノムの収納機構については、依然として不明な点が多い。真核生物ゲノムの核内収納機構を解明する糸口を得るために、クロマチン基盤構造を規定する DNA の物理的特性を明らかにする必要がある。

クロマチンの基本構造は、ゲノム DNA とヒストン八量体からなる 10 nm クロマチン繊維 (beads-on-a-string structure) である。特に、ヌクレオソームは、クロマチン繊維の重量の 95% を占めることから [23]、クロマチン繊維を構成する最も基本的な単位構造であると言える。ヌクレオソームは、146 塩基対の DNA と、ヒストン H2A、H2B、H3、H4 タンパク質からなるヒストン八量体によって構成されており、また、ヌクレオソームを形成しやすい塩基配列が存在する一方で、ヌクレオソームを形成しづらい塩基配列が存在することが知られている [24]。しかしながら、どのような物理的特性をもつ DNA がヌクレオソームを形成しやすいのかは不明である。

これまでに、ヌクレオソームを形成しやすい塩基配列には、AT、TT、または TA が約 10 塩基対の間隔で周期的に存在し、さらに、これとは 5 塩基対だけ位相をずらして、GC が約 10 塩基対の間隔で周期的に存在することが知られている [25, 26]。このような周期的に存在するジヌクレオチドは、ヒストン八量体の表層において DNA の鋭い屈曲を可能とすることにより、ヌクレオソームの安定性に寄与していると考えられている [25]。しかしながら、このようなジヌクレオチドの周期性とヌクレオソーム DNA の柔軟性との間に、どのような関係があるかは不明である。そこで、ヌクレオソーム DNA がもつ平均的な柔軟性を解析した。

2.2 実験方法

2.2.1. ヌクレオソーム DNA の柔軟性の計算

出芽酵母 (*S. cerevisiae*) のヌクレオソームに由来する 199 の DNA 配列を解析の対象とした [26]。また、Packer らによって報告されているテトラヌクレオチドの柔軟性パラメータセット [15] (表 1.1) および Brukner らによって報告されているトリヌクレオチドの柔軟性パラメータセット [27] (表 2.1) を用いて解析を行なった。なお、DNA の柔軟性の計算およびデータのプロットの方法については、Fukue らの方法に従った [13, 14]。

表 2.1. Brukner らの柔軟性パラメータセット

トリヌクレオチド	柔軟性 ($\ln p$) ^{a,b}
AAT/ATT	-0.28
AAA/TTT	-0.274
CCA/TGG	-0.246
AAC/GTT	-0.205
ACT/AGT	-0.183
CCG/CGG	-0.136
ATC/GAT	-0.11
AAG CTT	-0.081
CGC/GCG	-0.077
AGG/CCT	-0.057
GAA/TTC	-0.037
ACG/CGT	-0.033
ACC/GGT	-0.032
GAC/GTC	-0.013
CCC/GGG	-0.012
ACA/TGT	-0.006
CGA/TCG	-0.003
GGA/TCC	0.013
CAA/TTG	0.015
AGC/GCT	0.017
GTA/TAC	0.025
AGA/TCT	0.027
CTC/GAG	0.031
CAC/GTG	0.04
TAA/TTA	0.068
GCA/TGC	0.076
CTA/TAG	0.09
GCC/GGC	0.107
ATG/CAT	0.134
CAG/CTG	0.175
ATA/TAT	0.182
TCA/TGA	0.194

^a参考文献 27

^b数値が大きいほど柔らかい。計算式に対数が含まれるため負の数値も現れるが、符号自体に特別な意味はない。

2.3. 結果

2.3.1. ヌクレオソーム DNA に共通した柔軟性

ヌクレオソームに由来する 199 の DNA 配列を対象として解析を行なった。ヒストン八量体に約2周巻き付いている 146塩基対のヌクレオソーム DNA は、ヌクレオソーム DNA の中心に位置する二回対称軸に関して対称な構造をとっている。従って、ヌクレオソーム DNA が極性をもたないことを考慮に入れると、この二回対称軸を境界として左側の DNA (1-73 bp) と右側の DNA (74-146 bp) は、ほぼ同じ柔軟性をもつと考えられる。実際、左右の DNA がもつ柔軟性は、予想通り、二回対称軸に関して鏡像の関係を示した。そこで、解析に用いる塩基配列の数を増やすために、両者を区別しないで解析を行なった。その結果、興味深いことに、ヌクレオソームを形成しやすい DNA には、極めて硬い領域と極めて柔らかい領域が交互に存在し、それぞれが約 10 塩基対の間隔で周期的に存在していることが判明した (図 2.1)。なお、DNA の柔軟性の解析には、Packer らによって報告されているテトラヌクレオチドの柔軟性パラメータセット [15] (表 1.1)、および Brukner らによって報告されているトリヌクレオチドの柔軟性パラメータセット [27] (表 2.1) を用いた。前者は理論的なアプローチにより求められた柔軟性パラメータセットであり、後者は実験的なアプローチにより求められた柔軟性パラメータセットである。すなわち、これらは互いに独立に求められたパラメータセットである。それにも拘らず、両者を用いて計算したヌクレオソーム DNA の柔軟性プロファイルは、互いによく類似していた (図 2.1)。すなわち、ここで明らかにしたヌクレオソーム DNA の柔軟性が、ヌクレオソーム DNA に普遍的な性質であることを意味している。

これまでに、ヌクレオソームを形成しやすい塩基配列には、約 10 塩基対の間隔で AA/TT が高頻度に存在することが報告されている [28]。このような知見と図 2.1 の結果を合わせて考えると、本研究でみられた極めて硬い領域の位置が、AA/TT が高頻度に存在する位置と一致する可能性が推察される。さらに、AA/TT を含むトリヌクレオチドおよびテトラヌクレオチドが硬い (柔軟性が低い) ことも、このような可能性を支持している [15, 27] (表 1.1、表 2.1)。そこで、極めて硬い領域の位置と、AA/TT が高頻度に存在する位置との比較を行

なった (図 2.2) 。その結果、報告されている通り AA/TT が約 10 塩基対の間隔で周期的に存在することが確認されたことに加えて [28] 、極めて硬い領域の位置と、AA/TT が高頻度に存在する位置とがほぼ一致することが判明した。

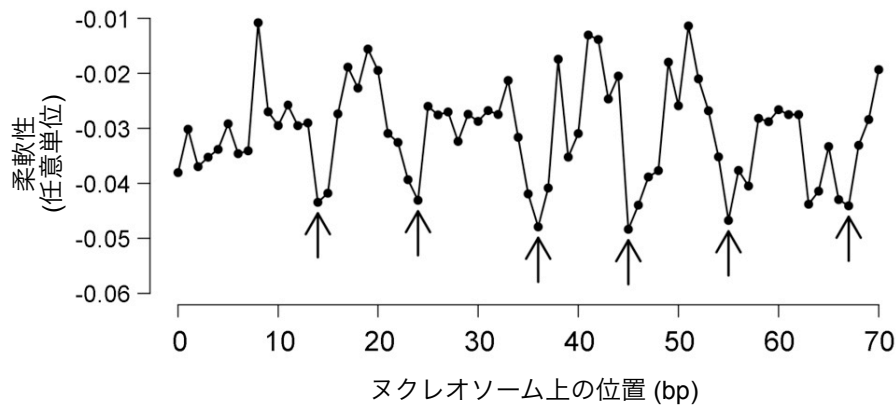
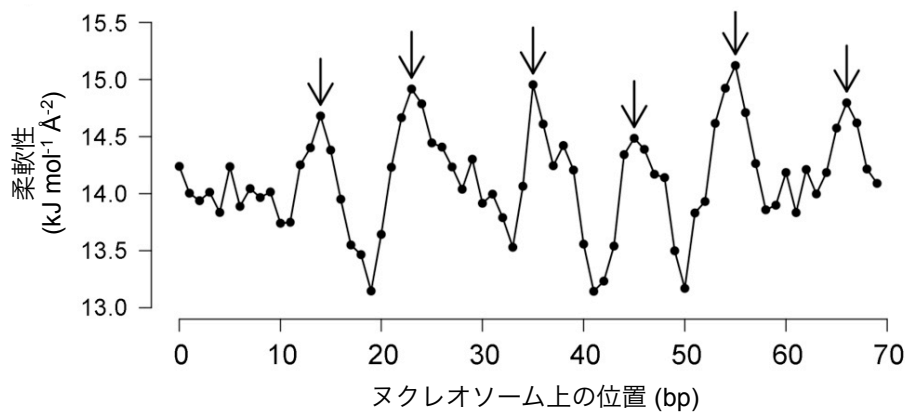
A**B**

図 2.1. ヌクレオソームに由来する 199 の DNA 配列がもつ平均的な柔軟性 DNase I を用いて得られた柔軟性パラメータセット [27] (A) およびポテンシャルエネルギー曲面モデル (potential energy surface model) を用いて得られた柔軟性パラメータセット [15] (B) を用いて、ヌクレオソームに由来する 199 の DNA 配列がもつ平均的な柔軟性を計算した。なお、(B) においては値が小さいほど柔軟性が高いことを意味する。また、ヌクレオソーム DNA は二回対称軸に関して対称な構造をとるため、二回対称軸を境界にして左側の DNA と右側の DNA を区別せずに解析を行なった ($n=398$)。図中において、位置 0 (bp) が二回対称軸の位置に相当する。(A) においては、極めて硬い領域が位置 14、24、36、45、55、67 (bp) に存在し、(B) においては、極めて硬い領域が位置 15、24、36、46、56、67 (bp) に存在した (位置を矢印で示す)。

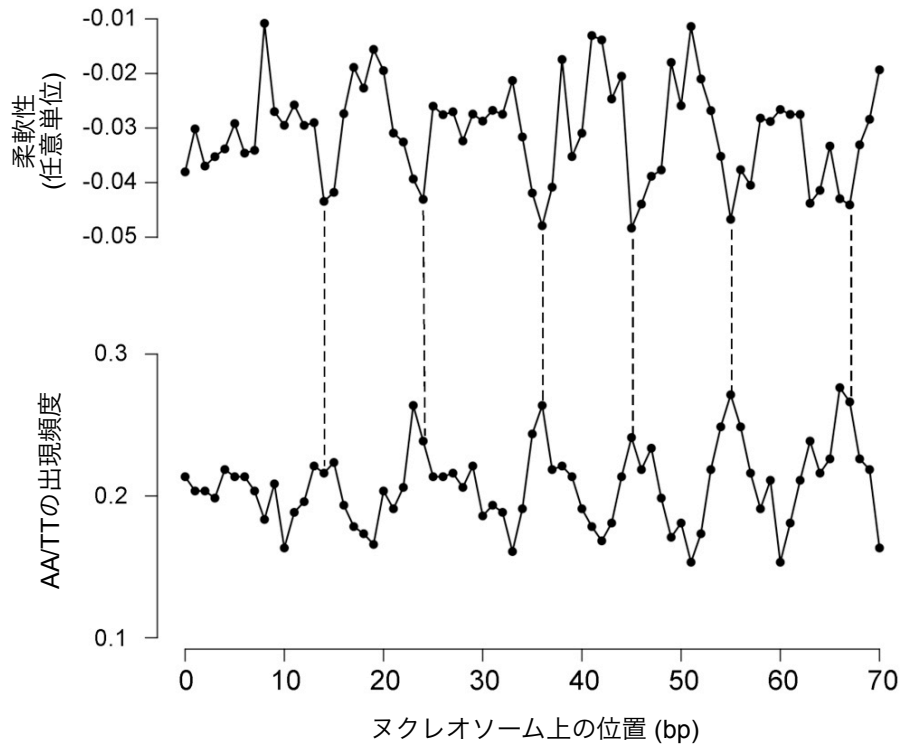


図 2.2. ヌクレオソーム DNA がもつ平均的な柔軟性と AA/TT の出現頻度の比較

ヌクレオソームに由来する 199 の DNA 配列を解析の対象とした (n=398)。比較のために、図 2.1.A で示した柔軟性プロファイルを再び示した。なお、AA/TT の出現頻度を、すべてのジヌクレオチドの出現回数に対する AA/TT の出現回数の割合として示した。

2.4. 考察

本研究の解析から、ヌクレオソームを形成しやすい DNA には、極めて硬い領域と極めて柔らかい領域が交互に存在し、それぞれが約 10 塩基対の間隔で周期的に存在していることが判明した。以下では、このようなヌクレオソーム DNA に共通した物理的特性の生物学的意義について考察する。

これまでに、ヌクレオソーム DNA に周期的に存在する AA/TT が、ヒストン八量体に対して内側に位置することが報告されている [29]。一方、本研究で明らかにしたように、極めて硬い領域の位置と AA/TT が高頻度に存在する位置がほぼ一致することから (図 2.2)、極めて硬い領域がヒストン八量体に対して内側に位置することが予想される。そこで、各々の極めて硬い領域と二回対称軸との距離を基にして (図 2.1)、極めて硬い領域がヌクレオソーム DNA に占める位置を明らかにした (図 2.3)。その結果、予想通り、極めて硬い領域がヒストン八量体に対して内側に位置すること (すなわち、ヒストン八量体に最も近い位置に存在すること) が確認された。このように、極めて硬い領域の位置とヌクレオソーム構造との間に相関がみられることから、周期的に存在する極めて硬い領域が、ヌクレオソームの形成機構に関与することが示唆される。

ヌクレオソーム DNA に周期性に存在する AA/TT については、ヌクレオソームの形成機構におけるその役割が明らかにされている。一般に、ヒストン八量体に対して内側に位置する DNA 配列は、副溝側が極めて狭い (副溝側に圧縮された) 構造をとる必要がある [29]。一方で、AA/TT は、副溝側が極めて狭い構造を好んでとることが知られている [30]。つまり、AA/TT のように生来的に副溝側が極めて狭い DNA 配列が、ヌクレオソーム DNA に対して内側に位置することで、エネルギー的に非常に安定した状態をつくることができる。これが、AA/TT がヌクレオソームの形成機構において担う役割である。では、本研究で発見した周期的に存在する極めて硬い領域は、ヌクレオソームの形成機構においてどのような役割を担うのだろうか。ここで、DNA の柔軟性とエントロピーの関係についての Travers らの議論 [31] を参考にして推論を進めると、本研究で発見した極めて硬い領域は、エントロピーの観点からみて、ヌクレオ

ソームの形成を促進する役割を担うことが帰結される。具体的に述べると、例えば、一般に、極めて硬い領域をもつ DNA 配列がとり得るかたちの自由度は低い、その一方で、全体的に柔軟な DNA 配列がとり得るかたちの自由度は非常に高い。すなわち、エントロピーの観点からみると、前者の極めて硬い領域をもつ DNA の方が、ヒストン八量体に巻き付くために必要なコストは小さくて済む。要するに、本研究で発見した周期的に存在する極めて硬い領域は、このような機構により、ヌクレオソームを安定化する役割を担うと考えられる。それでは、本研究で発見した極めて硬い領域が、ヒストン八量体に対して内側に位置するのはなぜだろうか。興味深いことに、極めて硬い領域の位置と、AA/TT が高頻度に存在する位置が一致することは (図 2.2)、AA/TT がもつ、副溝側に鋭く圧縮された構造を好む性質と、AA/TT を含む DNA 領域が極めて硬いことが両立する (少なくとも矛盾しない) ことを意味している。これを出発点として推論を進めると、上の問いに対する答えがみえてくる。具体的に述べると、ヌクレオソーム DNA は、ヒストン八量体に巻き付くために全体的に鋭く屈曲する必要があることから、高い柔軟性をもつことが要請される。しかし、その一方で、上に述べたように、全体的に柔軟な DNA は、エントロピーの観点からはヌクレオソームの安定性において極めて不利である。そこで、唯一、ヒストン八量体に対して内側に位置する DNA 領域のみが、極めて硬い特性をもつことが許されているのではないか。言い換えると、本研究で発見した極めて硬い領域がヒストン八量体に対して内側の領域に位置することは、ヌクレオソームの熱力学的な安定性を総合的に追求した結果であり、理に適っている。

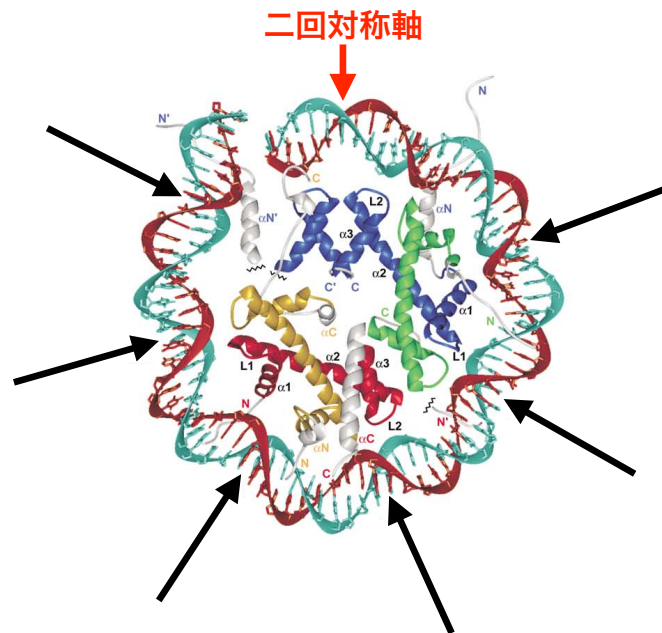


図 2.3. 極めて硬い領域はヒストン八量体に対して内側に位置する
 二回対称軸と各々の極めて硬い領域との距離を基にして、極めて硬い領域がヌクレオソーム DNA に占める位置を求めた。極めて硬い領域の位置を黒の矢印で示す。極めて硬い領域が、ヒストン八量体に対して内側（すなわちヒストン八量体に最も近い位置）に存在することが分かる。なお、ヌクレオソームの構造図については、参考文献 32 より引用した。

第 3 部

間期染色体の三次元構造の構築原理

3.1. 緒言

クロマチン繊維が核内でとる三次元構造は、転写・複製・修復などの様々な核内現象の舞台となる重要な基盤構造である [33-35]。従って、クロマチン繊維の構築機構を解明することは、核内現象の分子機構を理解するための鍵となると考えられる。これまでに、クロマチン繊維の構築機構に対する洞察を得ることを目的として、様々なクロマチン繊維構造モデルが報告されてきた。これらのモデルは、アプローチの違いに基づいて大きく 2 種類に分類することができる。ひとつは、数十から百程度のヌクレオソームからなるクロマチン繊維を対象として構築されるメソスケールモデル (mesoscale model) であり [36-40]、もうひとつは間期染色体の全域を対象として構築されるフルスケールモデル (full-scale model) である [41-47]。前者のメソスケールモデルの特色は、微視的な観点からクロマチン繊維を特徴づけるパラメータを用いて、モデルを構築する点にある。このようなパラメータの例として、ヌクレオソームリピートの長さ、イオン濃度、ヒストンテール、ヌクレオソーム間の静電的相互作用、ヌクレオソームとリンカーDNA 間の静電的相互作用、ヌクレオソーム前後のリンカーDNA がなす角度 (entry-exit angle) などが挙げられる [36, 38, 40]。

後者のフルスケールモデル (full-scale model) は、ポリマー理論に基づくモデルである。具体的には、ランダムウォークモデル (random-walk model)、みみず鎖モデル (worm-like chain model)、ビーズ・スプリングポリマーモデル (bead-spring polymer model) などを基にして構築したモデルであり [41-47]、ヒトやショウジョウバエ、出芽酵母の染色体を対象とすることが多い [42-47]。このモデルの特色は、クロマチン繊維を染色体全域にわたり均一な物性をもつ架空の高分子として扱うことと、この架空の高分子を記述する各種のパラメータを、シミュレーション結果と実験データとが合致するように調整するトップダウン・アプローチを用いることにある。また、フルスケールモデルは、メソスケールモデルにみられる微視的な観点に基づいたパラメータを用いておらず、その代わりに、いわゆる 30 nm クロマチン繊維の持続長を用いることが多い [44, 45, 47]。しかしながら、30 nm クロマチン繊維が実際に存在するか否かについては議論の余地がある [48-50]。なお、どちらのモデルにおいても、モンテカ

ルロ (Monte Carlo) 法やブラウンダイナミクス (Brownian dynamics) 法を用いてシミュレーションを行うことが多い。

これまでに、出芽酵母を対象とした間期染色体構造モデルとして、ビーズ・スプリングモデル [41] を用いた二つのモデルが報告されている [45, 47]。しかしながら、これらのモデルは、ゲノム全域にわたるヌクレオソームの配置のデータを用いてモデルを構築する代わりに、直径 30 nm のビーズ (ひとつのビーズが 3,000 塩基対の DNA に相当する) を連結させた鎖として、クロマチン繊維をモデル化している。従って、これらのモデルを用いて、注目するクロマチン領域が含むヌクレオソームの数や、この領域のリンカーDNA がとる三次元構造などに言及することは原理的に不可能である。

本研究では、真核生物ゲノムの核内収納原理を明らかにするために、ボトムアップ・アプローチを用いて間期染色体構造のモデルを構築した。なお、出芽酵母において、ゲノム全域にわたるヌクレオソームの配置が明らかにされているため [26, 51-53]、モデルを構築する対象として、出芽酵母の染色体を選択した。具体的には、ヌクレオソームの配置 [26, 51, 53]、リンカーDNA の物理的特性、および細胞核の大きさ (一倍体細胞) の 3 つの因子を用いてモデルを構築し、間期核における全 16 本の染色体構造のスナップショットを、ヌクレオソームの配置が識別できる分解能で示した。予測された間期染色体構造と、核内における実際の構造との比較を行なったところ、予測値と実験値とが非常によい一致を示した。以上の解析に基づいて、DNA の物理的特性と細胞核の大きさが、出芽酵母の間期染色体の基盤構造を決定する因子であるという仮説を提唱する。

3.2 実験方法

3.2.1. シミュレーション方法

モンテカルロ法を用いて、間期染色体構造のシミュレーションを行なった(モデルの構築については“3.3.1. 間期染色体構造モデルの構築”を参照)。なお、本研究のシミュレーション方法は、無作為に選んだベクトルを伸長中の鎖に次々と付加する点において、Wall と Erpenbeck が報告している鎖伸長法 (chain-enrichment procedure) [54] と本質的に類似している。特に記述していない場合、各々の染色体の一端を起点としてシミュレーションを開始した。また、隣り合うヌクレオソーム同士が特に近接して存在する場合 (リンカーDNAの長さが 12 nm 以下の場合) に、ヌクレオソーム同士の衝突が原因で、シミュレーションを完遂できない場合があった。このような場合は、該当するヌクレオソームについて、DNA の巻き数を基本となる 1.64 周から 1.35 周へと減らしてリンカーDNA の長さを増すことで、シミュレーションを完遂させた。この際、ヌクレオソーム前後のリンカーDNA がなす角度 (entry-exit angle) の一種である角度 α (図3.2.A 参照) を、巻き数の変化に連動させて -54° とした。他には、既に配置された鎖の部分領域との衝突が原因で、シミュレーションを完遂できない場合があった。このような場合は、鎖の伸長が停止した領域の直前に位置するヌクレオソームの空間配置を決定するところから再計算した [引き返し再開法 (back-and-resume procedure)]。なお、以上の計算を行なうコンピュータプログラムを C 言語および R 言語 [16] を用いて作成した。

3.2.2. 細胞核の制限

まず、一倍体の出芽酵母の G1 期における細胞核を、直径 2 μm の球とみなした [55] (表 3.3)。次に、上述の手法により得られた各々の間期染色体構造のサンプルに固有な最小包含球を、Welzl によって報告されているアルゴリズム [56] を用いて求め、直径 2 μm の球に納まる間期染色体構造のサンプルをスクリーニングした。なお、以上の計算を行なうコンピュータプログラムを C 言語および R 言語 [16] を用いて作成した。

3.2.3. 原子間力顕微鏡 (AFM) を用いた DNA 断片の観察

まず、出芽酵母のゲノム DNA および次頁に示したプライマーを用いて、5種類の DNA 断片を PCR 法により調製した。続いて、得られた PCR 産物の単離・精製を、1.5%アガロースゲルおよび illustra GFX PCR DNA and Gel Band Purification Kit (GE ヘルスケア社) を用いて行なった。

次に、10 mM MgCl₂ 処理したマイカ基盤上に、8 ng の DNA と 10 mM MgCl₂ を含む水溶液 20 μl を滴下して 5 分間放置したあと、さらに 50 ml の純水で 2 回リンスを行い、乾燥させた。そして、シリコンカンチレバー (AC160TS; オリパス社) および NanoWizard (JPK Instruments AG 社) を用いて、タッピングモードにて観察像 (1 μm × 1 μm; 512 × 512 ピクセル) を得た。なお、すべての操作を室温で行なった。続いて、観察像に対する画像処理および解析を、それぞれ JPK SPM 画像処理ソフト (JPK Instruments AG 社) および ImageJ (NIH) を用いて行なった。持続長を測定した 5 種類の DNA 断片の柔軟性はそれぞれ、11.6 (kJ mol⁻¹ Å⁻²) (n = 191)、12.0 (n = 233)、13.1 (n = 206)、13.6 (n = 336)、および 14.4 (n = 266) であった。

表 3.1. PCR に用いたプライマー

プライマー	塩基配列 (5'→3')	対象とした領域	柔軟性 (kJ mol ⁻¹ Å ⁻²)
II-flx14.4-F	GGACCACTCGCGTACTTAAAAC	2 番染色体; 24,491-24,990	14.4
II-flx14.4-R	CCGCGTTATGACGAAATAGTGC		
II-flx13.6-F	GCAAAGGTCCCAATCGAAAT	2 番染色体; 44,502-45,001	13.6
II-flx13.6-R	TCTGACAACAGGGTGGAAAC		
IV-flx13.1-F	ATATACCGCTTACCACCTACCG	4 番染色体; 1,385,500-1,385,991	13.1
IV-flx13.1-R	AACTTATGCTTCCCTCCATAC		
VII-flx12.0-F	ATGCGACTACCAACTCCAG	7 番染色体; 1,088,985-1,089,484	12.0
VII-flx12.0-R	CTTATACGACTCTTTGTTAATGTCG		
VIII-flx11.6-F	AGTATGGCATGTGGTGAGGT	8 番染色体; 556,501-557,000	11.6
VIII-flx11.6-R	AGTGACATATATATACACAC		

3.2.4. FISH 法を用いた 2 点間空間距離の測定

一倍体株である BY4741 株 (*MATa his3-1 leu2 met15 ura3*) を、YPAD 培地 (30°C) を用いて OD₆₀₀ が 0.4 となるまで培養した。次に、Gotta ら [57] の手法を修飾して、FISH (fluorescence *in situ* hybridization) を行なった。まず、DNA プローブを、出芽酵母のゲノム DNA と次頁に示したプライマーを用いて PCR 法により調製した。続いて、ニックトランスレーション法により、これらのプローブを Alexa488-5-dUTP (Molecular Probes 社) または digoxigenin-11-2'-dUTP (Roche 社) にて標識した。そして、標識プローブ (62.5 mg/ml; 50%ホルムアミド溶液)、10%デキストラン硫酸、2× SSC、40 mM リン酸ナトリウム (pH 7.0)、1× デンハルト試薬、0.5 mM EDTA、およびサケの精子由来の一本鎖 DNA (0.12 mg/ml) を用いて、37°C で 16 時間のハイブリダイゼーションを行なった。なお、ジゴキシゲニン標識プローブの免疫検出を、ローダミン結合抗ジゴキシゲニン Fab 断片 (Roche 社) を用いて行なった。さらに、備え付けのソフトウェア (FV-10; ver 2.1) および 対物レンズ (×100 UPlan SApo) を用いて、共焦点レーザー顕微鏡 (Olympus FluoView FV1000-D) により観察像を得た (ピクセルの大きさが 100 nm となるように調整)。この際、Z 軸に沿って 0.2 μm 毎に、計 16 枚の観察像を得た。また、検出された 2 つのシグナル間の空間距離を、最も輝度の高いピクセル間の空間距離として測定した。

表 3.2. PCR に用いたプライマー

プライマー	塩基配列 (5'→3')	プローブ
IV-100-F	CTTTGGTTGGTGATGGATCC	IV-100
IV-100-R	TTTACCTTGTCCCGGGTCT	
IV-300-F	AGAACGGCGATGATATGGAC	IV-300
IV-300-R	CAGAACCGAGAAGCTCCAAC	
IV-600-F	TCGGGGAAGTTTACAGGAAG	IV-600
IV-600-R	CGGATGATAGGTTTGCACAG	
IV-1,200-F	AGTTACGCCGCAAATACCTG	IV-1,200
IV-1,200-R	AGCTTGGCTTTTTCTGGTGA	
IV-1,400-F	TTCAGCAGACACACCAGCTT	IV-1,400
IV-1,400-R	GACACGATTGGTCGTGAATG	
V-100-F	AGGTCAAACATCGGTGGAAG	V-100
V-100-R	CCCGGTAAAGCACTGTTCTC	
V-200-F	ATTACAGACCCACCCAGCAG	V-200
V-200-R	AGTTCCAACAGCGCTTGAAT	
V-300-F	CTGGTGAATTAGCCCGTGTT	V-300
V-300-R	ATAACCCTGTTGGCTGTTTCG	
V-400-F	GATTTGACTACGGCCTCTG	V-400
V-400-R	ACGAGCCCTGTTGCTAAGAA	
V-500-F	CTTTCCTCCGCTTTGTCAAC	V-500
V-500-R	TCCTTTGGTTGGTTCTCTGC	
VII-100-F	CAACCAATACCGCAAGTCCT	VII-100
VII-100-R	ACTCCCGAATTGTGGTCTCA	
VII-500-F	CCAATTTGCAACCTGTCTT	VII-500
VII-500-R	CACTTCTGGGCTTGGAGAAG	
VII-700-F	CGTGGAAGAACCCTGATGTT	VII-700
VII-700-R	AGGGTAAAGAAGGCGACGAT	
VII-1,000-F	CGTCTTTCTTCCCTCGTCTG	VII-1,000
VII-1,000-R	AGGAGGCAAACCCTTGAAGT	

3.2.5. 4C データから 2 点間空間距離を導出する方法

FISH 法を用いて明らかにした結果を、他の手法を用いて確認するために、染色体上の 2 点間の相対的空間距離を検出する実験手法である 4C (chromosome conformation capture-on-chip)を用いて得られたデータ [58] を用いた。まず、4C を用いて得られたデータから、図 3.7.B で示した合計 29 組の測定箇所に対応するデータを抽出した。次に、4C を用いて得られた 2 点間の相対的空間距離 (2 領域の接触頻度) と、2 点間空間距離を関係づける関係式 [58] (図 3.1) を用いて、各々の測定箇所の 2 点間空間距離を計算した (図 3.7.C)。

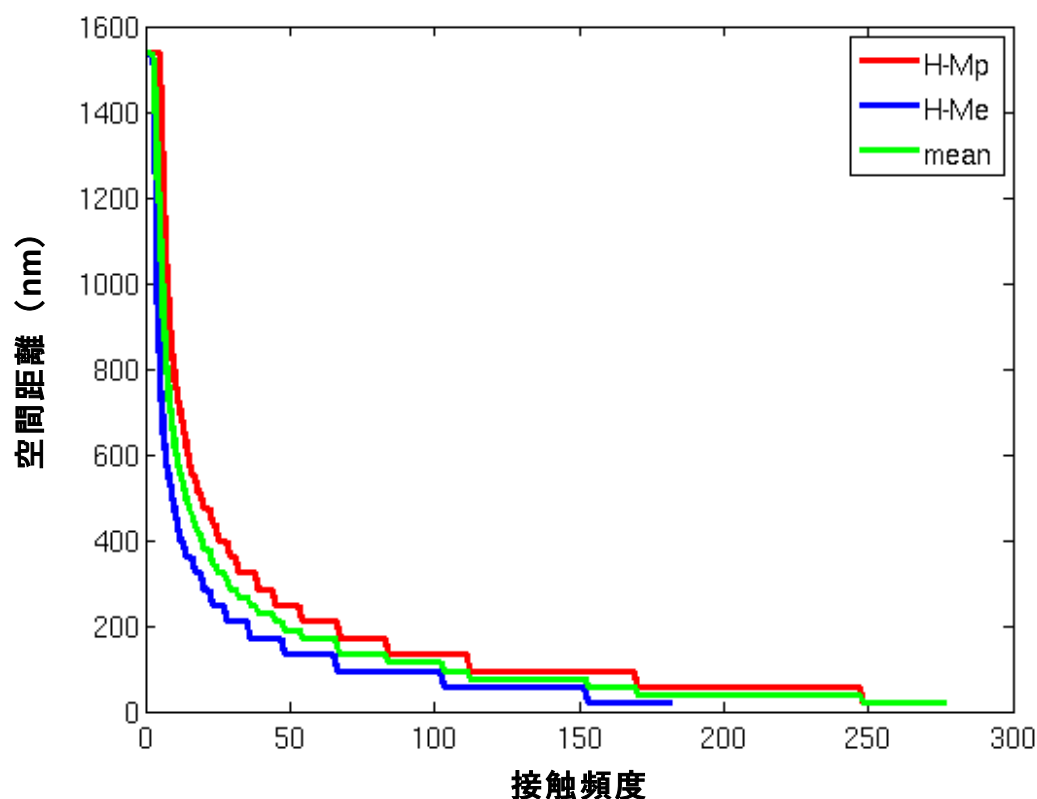


図 3.1. 接触頻度と空間距離の関係

赤線は H-Mp を用いて測定を行なった場合の関係を表わし、青線は H-Me を用いて測定を行った場合の関係を表わす。緑色の線は 2 つの平均を表わす。参考文献 58 に掲載された “Supplementary Figure 17” を示した。

3.3. 結果

3.3.1. 間期染色体構造モデルの構築

出芽酵母の間期染色体の全長を対象として、ヌクレオソームの配置が識別できる分解能でモデルを構築した。はじめに、DNA を、DNA 二本鎖の太さに等しい直径 2 nm のビーズ（ひとつのビーズが 6 塩基対に相当する）を連結させた、数珠状の鎖としてモデル化した（図 3.2.A）。なお、隣り合う二つのヌクレオソーム間の回転的配向を定めるために、隣接するビーズの間にねじれ角 ρ [ρ : 約 206° ($360^\circ \times 6/10.5$)] を導入した（図 3.2.A、右の挿入図）。すなわち、このようなねじれ角 ρ の和として、隣り合う二つのヌクレオソーム間の回転的配向を求めた。次に、ヌクレオソームについては、円盤（円柱）状のヌクレオソーム構造（直径 5.7 nm、高さ 11 nm）を忠実に反映するように、直径 5.7 nm の球と、この球に左巻きに巻き付いたビーズの鎖としてモデル化した（図 3.2.A）。すなわち、ヌクレオソームを、直径 2 nm のビーズの鎖が架空の円盤に 1.64 周巻き付いた構造としてモデル化し、この架空の円盤の代わりに球を配置した。なお、必要な場合に限り、ビーズの鎖が巻き付く回数を、基本となる 1.64 周から変更した（“3.2.1. シミュレーション方法”を参照）。

リンカーDNA の三次元的な軌道は、排除体積を考慮に入れた“みみず鎖モデル（worm-like chain model）”を用いて記述した〔なお、自由回転鎖としてみみず鎖モデルを具現化した [59]（図 3.2.A、左の挿入図）〕。従って、 i 番目のリンカーDNA の軌道は、以下の関係式を用いて求める結合角 θ_i (i 番目のリンカーDNA に固有な結合角) の大きさに依存する。

$$\frac{a}{1 - \cos \theta_i} \approx PL_i \quad [60]$$

ここで、“ a ” は繰り返し単位の長さ（本研究においては 2 nm）を表わし、 PL_i は i 番目のリンカーDNA の持続長を表わす。なお、 PL_i を、図 3.3.F に示した関係式と、 i 番目のリンカーDNA の柔軟性に基づいて求めた。このようにして i 番目のリンカーDNA の軌道を決定したあと、このリンカーDNA の末端を i 番目のヌクレオソーム DNA に接続させた（図 3.2.B）。なお、ヌクレオソーム前後

のリンカーDNA がなす角度 (entry-exit angle) である角度 α および角度 β として、それぞれ 49° および 4.6° を与えた (図 3.2.A)。最後に、得られた間期染色体構造のサンプルから、出芽酵母の細胞核の大きさである直径 $2 \mu\text{m}$ の球におさまる間期染色体構造のサンプルをスクリーニングした (図 3.2.C)。

出芽酵母の間期染色体構造モデルの構築に用いたパラメータを表 3.3 に示す。また、既に報告されているクロマチン繊維構造モデルが用いるパラメータと、本研究で構築したモデルで用いたパラメータの比較を行なった (表 3.4)。なお、既存のモデルを、アプローチの違いに基づいて大きく 2 種類に分類した上で比較を行なった。ひとつは、数十から百程度のヌクレオソームからなるクロマチン繊維を対象として構築されたメソスケールモデル (mesoscale model) であり、もうひとつは間期染色体の全域を対象として構築されたフルスケールモデル (full-scale model) である。

表 3.3. 出芽酵母の間期染色体構造モデルの構築に用いたパラメータ

リンカーDNA	値 (データ)
配置	Kaplan ら (2009) ^a ; Brogaard ら (2012) ^b
持続長	25.8 - 57.1 nm
ビーズの直径 ^c	2 nm
ねじれ角 ρ ^c	206°
ヌクレオソーム	
配置	Kaplan ら (2009) ^a ; Brogaard ら (2012) ^b
角度 α (entry-exit angle α) ^c	49°
角度 β (entry-exit angle β) ^c	4.6°
細胞核	
直径	$2.0 \mu\text{m}$ ^d

^a 参考文献 51

^b 参考文献 53

^c 図 3.2.A を参照。

^d 参考文献 55

表 3.4. クロマチン繊維構造モデルで用いられる代表的なパラメータの比較

パラメータ	メソスケールモデル (mesoscale model) ^a	本研究で構築した フルスケールモデル (full-scale model)	既存の フルスケールモデル (full-scale model) ^b
排除体積効果	○	○	○ (多くの場合)
イオン濃度	○	○ ^c	×
静電的相互作用	○	×	×
ゲノム全域にわたるヌクレオソームの配置	×	○	×
ゲノム全域にわたるリンカーDNAの配置	×	○	×
ヌクレオソームリポートの長さ (固定値)	○	×	×
ヌクレオソームの配置のゆらぎ	○ (参考文献 39 のみ)	○ ^d	×
リンカーDNAの持続長	○ (50 nm)	○ (塩基配列依存的)	×
リンカーDNAのねじれ	○	○ (図 3.2.A を参照)	×
角度 α (entry-exit angle α ; 図 3.2.A を参照)	○	○	×
角度 β (entry-exit angle β ; 図 3.2.A を参照)	×	○ (多くの場合)	×
ヒストンテール	○ (多くの場合)	×	×
リンカーヒストン	○ (多くの場合)	×	×
細胞核の大きさ (直径)	×	○	○ (参考文献 43, 45, 47 のみ)
30 nm クロマチン繊維の持続長	×	×	○ (参考文献 44, 45, 47 のみ)
クロマチンループ	×	×	○ (参考文献 44, 46 のみ)
クロマチンの圧縮率 (compaction ratio)	×	×	○ (参考文献 44, 45, 47 のみ)

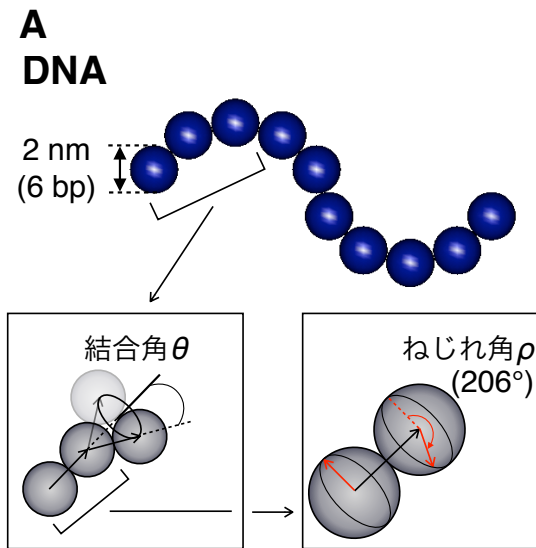
^a 参考文献 36-39

^b 参考文献 42-47

^c リンカーDNAの持続長に反映されている。

^d 図 3.4 を参照。

^e 出芽酵母 (*S. cerevisiae*) はリンカーヒストンをもたない。



ヌクレオソーム

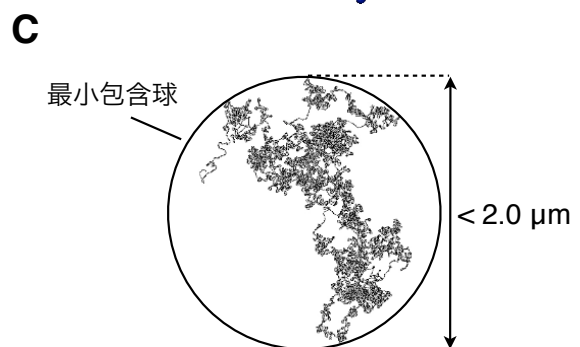
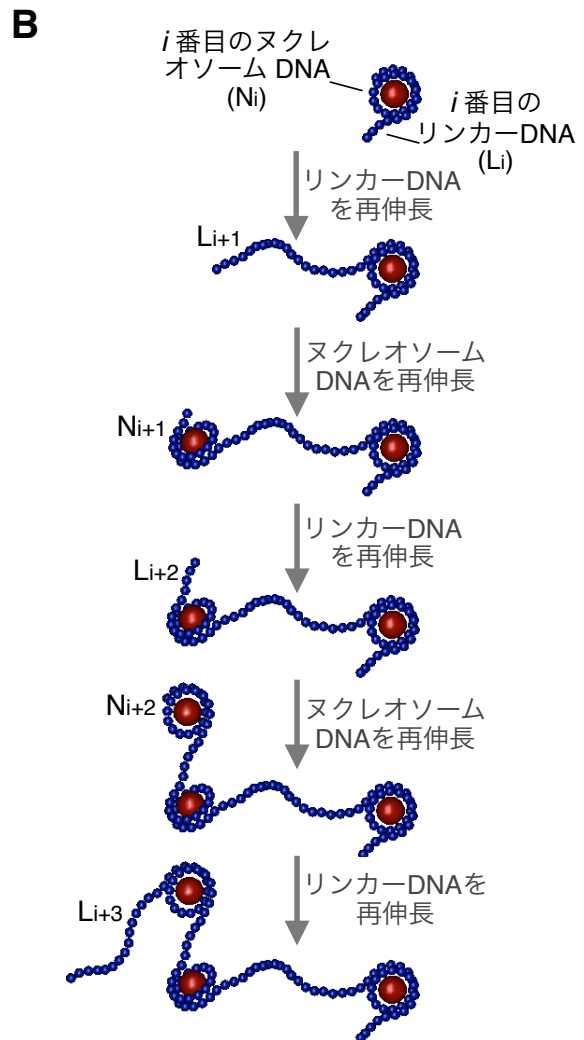
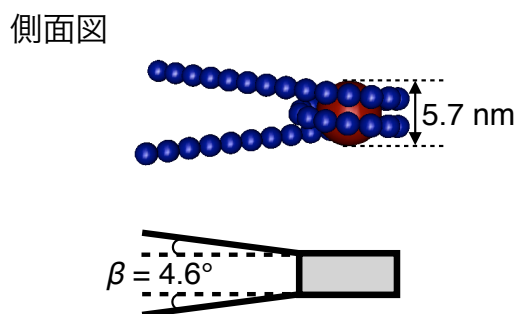
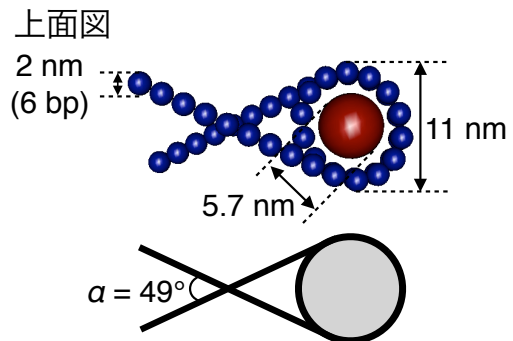


図 3.2. 間期染色体構造のシミュレーション方法

(A) DNA、ヌクレオソーム、リンカーDNA の三次元軌道を記述するパラメータ

DNA を、直径 2 nm のビーズ（ひとつのビーズが 6 塩基対に相当する）を連結させた数珠状の鎖としてモデル化した。また、リンカーDNA の三次元的な軌道を Håkansson と Elvingson らのみみず鎖モデル[59]を用いて記述した。この際、結合角 θ を、各々のリンカーDNA に固有な持続長および繰り返し単位の長さ（本研究においては 2 nm）に基づいて計算した。そして、ねじれ角 ρ として 206° ($360^\circ \times 6/10.5$) を与え、ヌクレオソーム前後のリンカーDNA がなす角度（entry-exit angle）である角度 α および角度 β に、それぞれ 49° および 4.6° を与えた。さらに、円柱状のヌクレオソーム構造（直径 5.7 nm、高さ 11 nm）を、直径 5.7 nm の球と、この球に左巻きに 1.64 周巻き付いたビーズの鎖としてモデル化した。なお、必要な場合に限り、このビーズの鎖が巻き付く回数を、基本となる 1.64 周から変更した（“3.2.1. シミュレーション方法”を参照）。

(B) クロマチン繊維のシミュレーション

リンカーDNA とヌクレオソーム DNA を交互に再伸長させる手法を通して、クロマチン繊維構造のサンプルを得た。なお、この手法は Wall と Erpenbeck が報告している鎖伸長法（chain-enrichment procedure）[54]を修飾した方法である（“3.2.1. シミュレーション方法”を参照）。

(C) 細胞核に納まる間期染色体構造のスクリーニング

Welzl によって報告されているアルゴリズム [56]を用いて、各々の間期染色体構造のサンプルに固有な最小包含球を求め、出芽酵母の細胞核の大きさである直径 2 μm の球におさまる間期染色体構造のサンプルをスクリーニングした（“3.2.2. 細胞核の制限”を参照）。

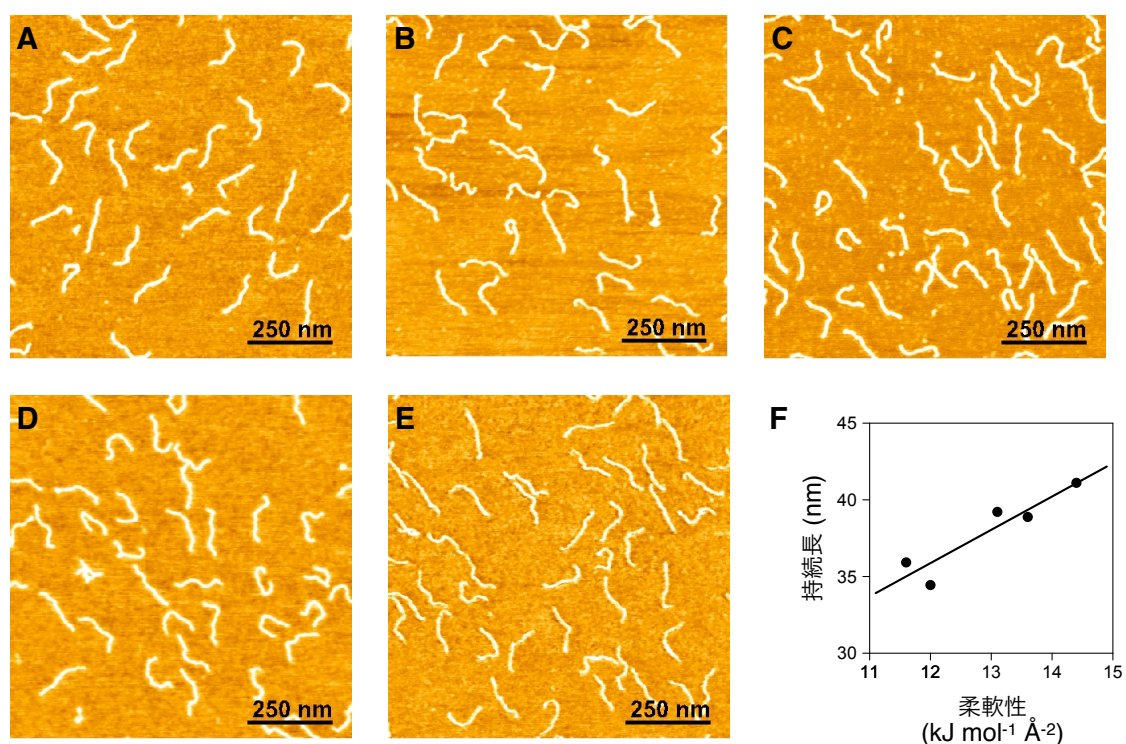


図 3.3. DNA の持続長と柔軟性の関係

(A-E) 5 種類の異なる柔軟性をもつ DNA 断片の AFM 観察像

10 mM の MgCl_2 存在下で、各々の DNA 断片の形状を原子間力顕微鏡 (AFM) で観察した。観察像の例を示す。なお、各々の DNA 断片の柔軟性 ($\text{kJ mol}^{-1} \text{Å}^{-2}$) は、14.4 (A)、13.6 (B)、13.1 (C)、12.0 (D)、および 11.6 (E) であった [Packer らが報告している柔軟性パラメータセット [15] (表 1.1) を用いて計算した; このパラメータセットにおいては、値が小さいほど柔軟性が高いことを意味する]。

(F) DNA の持続長と柔軟性の関係

持続長を、Rivetti らが報告している関係式 [61] :

$$\langle R^2 \rangle_{2D} = 4qL \left(1 - \frac{2q}{L} \left(1 - e^{-\frac{L}{2q}} \right) \right)$$

を用いて計算した。ここで、“ q ” は持続長を表わし、“ L ” は DNA の長さを表わし、“ $\langle R^2 \rangle_{2D}$ ” は平面的な観察像 (二次元の観察像) における平均二乗末端間距離を表わす。直線は最小二乗法で求めた近似直線である。なお、各々の DNA 断片について、190 以上の分子の解析を行なった。

3.3.2. リンカーDNAの持続長の測定

第1部および第2部の研究で述べたように、柔軟性パラメータセットを用いることにより [15, 27] (表 1.1、表 2.1)、任意の DNA 断片がもつ柔軟性を、その塩基配列から導くことが可能である [13, 14]。一方、DNA の持続長と柔軟性は、一枚のコインの表と裏に相当する相補的な概念であると考えられる。すなわち、任意の DNA 断片の持続長もまた、DNA の柔軟性を媒介変数にすることで、塩基配列から導くことが可能であると考えられる。しかしこれまで、持続長と塩基配列との関係についてはほとんど注意が払われてこなかった。また、このように持続長と塩基配列の関係に着眼することで、これまでに測定された持続長の間隔 [62] を、測定に用いた DNA の塩基配列の違い（ひいては測定に用いた DNA の柔軟性の違い）として説明できる可能性がある。

DNA の持続長と柔軟性の関係を明らかにするために、まず、異なる柔軟性をもつ 5 種類の DNA 断片を原子間力顕微鏡 (AFM) で観察した (図 3.3 A-E)。次に、Rivetti らが報告している関係式 [61] を用いて、各々の DNA 断片がもつ持続長を明らかにし、この結果を図 3.3.F に示した。なお、任意の塩基配列の三次元構造を予測するコンピュータプログラムである DIAMOD [63] を用いて、5 種類の DNA 断片が、ほぼ真っ直ぐな軌道をもつことを確認している。したがって、図 3.3.F における測定値の間隔は、各々の DNA 断片がもつ異方的な柔軟性によりもたらされたと考えられる。このようにして明らかにした DNA の持続長と柔軟性の関係の他に、ゲノム配列と、ゲノム全域にわたるヌクレオソームの配置のデータを用いて、出芽酵母がもつすべてのリンカーDNA の持続長を計算した。

3.3.3. 出芽酵母の間期染色体構造のシミュレーション

各々のヌクレオソームは一定の自由度をもって配置されている [64, 65]。すなわち、各々のリンカーDNA の配置および長さも、一定のゆらぎをもつ。そこでまず、このようなヌクレオソームの配置や長さの変化がリンカーDNA の柔軟性に与える影響を明らかにするため、出芽酵母の研究に頻繁に用いられる *LEU2* 遺伝子を対象とした以下の解析を行なった (図 3.4)。はじめに、*LEU2* 遺伝子周辺のヌクレオソームの初期位相として、Segal らによって報告されているデータ [26] を用いた。次に、kono 初期位相を基準として、前後 10 塩基対の範囲で各々のヌクレオソームを 1 塩基対ずつ移動させた。この結果として、各々のヌクレオソームは、全部で 21 通りの配置をとる (初期位相の前方に 1 塩基対ずつ移動させると 10 通り、初期位相の後方に 1 塩基対ずつ移動させると 10 通りの配置を生じる。ゆえに、初期位相と合わせて、 $1 + 10 + 10 = 21$ 通り)。従って、各々のリンカー領域は、各々のリンカー領域の前後に位置するヌクレオソームがそれぞれ 21 通りの配置をとるため、441 (21×21) 通りの配置 (長さ) をとる。続いて、441 通りのリンカー領域の配置 (長さ) の各々について、リンカーDNA の柔軟性を計算した。その結果、計 6 箇所 of リンカー領域の全てにおいて、初期位相におけるリンカーDNA の柔軟性と、全 441 通りのリンカーDNA がもつ柔軟性の平均値は、非常に近いことが判明した。すなわち、本研究においては、ヌクレオソームのゆらぎが DNA の柔軟性に与える影響を無視できると考えられる。

出芽酵母のゲノム全域にわたるヌクレオソームの配置のデータとして、Kaplan らが報告しているデータ [51] および Brogaard らが報告しているデータ [53] を用いて、出芽酵母の間期染色体構造のシミュレーションを行なった。図 3.5 に、シミュレーション結果の一例を示した。これらの構造は、動的な間期染色体構造のある瞬間をとらえたスナップショットに相当する。すなわち、図 3.5 の個々の画像は、熱運動のために常に構造を変化させている間期染色体のある時間断面における構造に相当する。なお、12 番染色体に関しては、rDNA 領域の塩基配列を塩基配列データベースから得ることができないため、計算の対象から除いている。

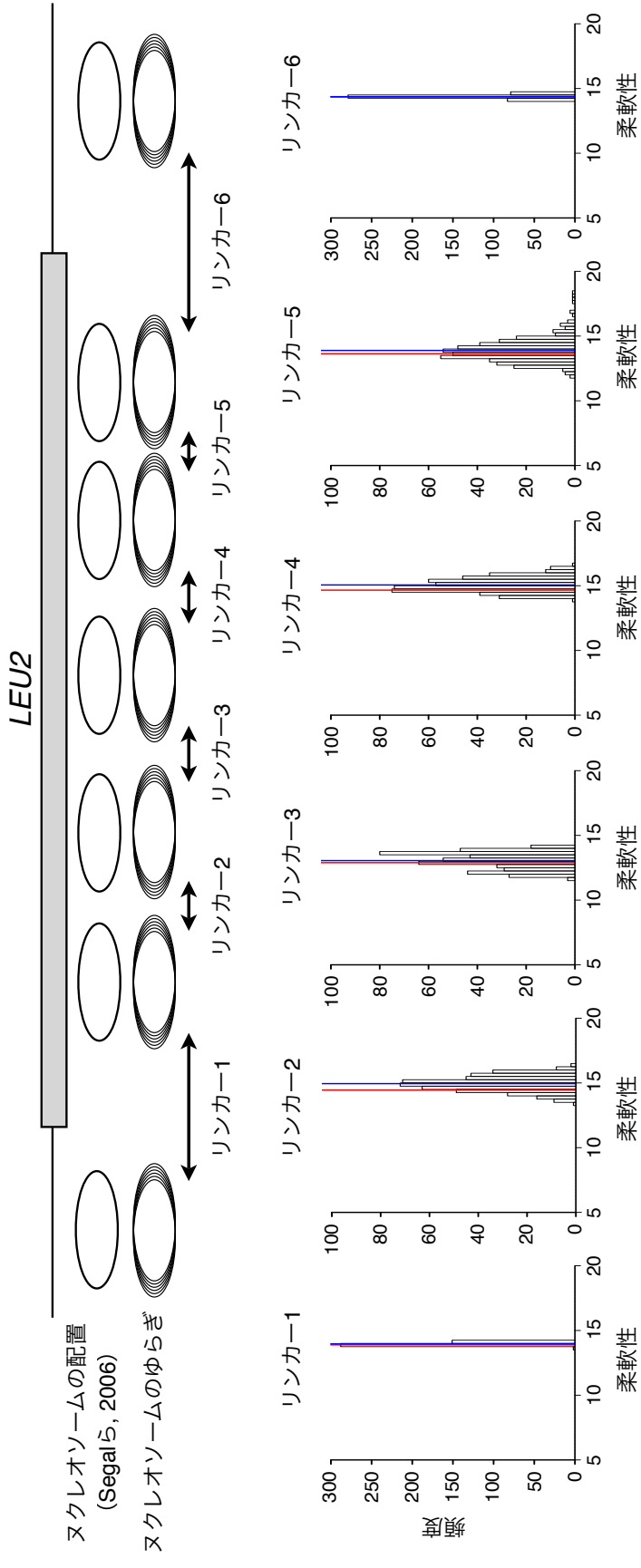
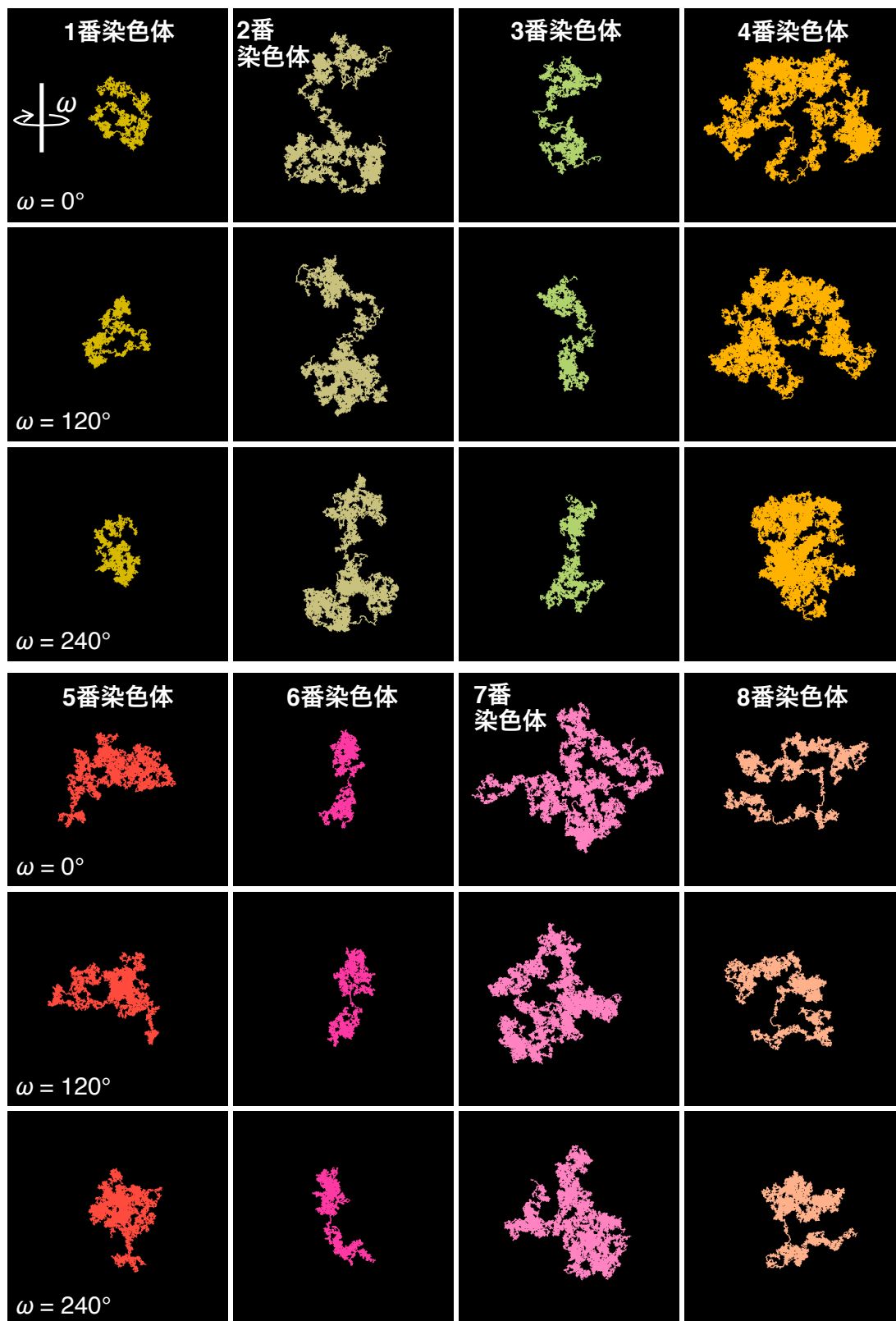


図 3.4. ヌクレオソームのゆらぎがリンカーDNAの柔軟性に与える影響

LEU2 遺伝子を例にして検討を行なった。なお、ヌクレオソームの初期位相として、Segal らによって報告されているデータ [26] を用いた。まず、初期位相を基準として、前後 10 塩基対の範囲で各々のヌクレオソームを 1 塩基対ずつ移動させた [各々のヌクレオソームがとり得る配置は 21 通りあり、従って各々のリンカー領域がとり得る配置 (長さ) は 441 (21×21) 通りある (本文に詳述)]。続いて、441 通りのリンカー領域の配置 (長さ) の各々について、リンカーDNAの柔軟性を計算し、その分布をヒストグラムに示した。初期位相におけるリンカーDNAの柔軟性を赤線で示し、全 441 通りのリンカーDNAの平均値を青線で示した。



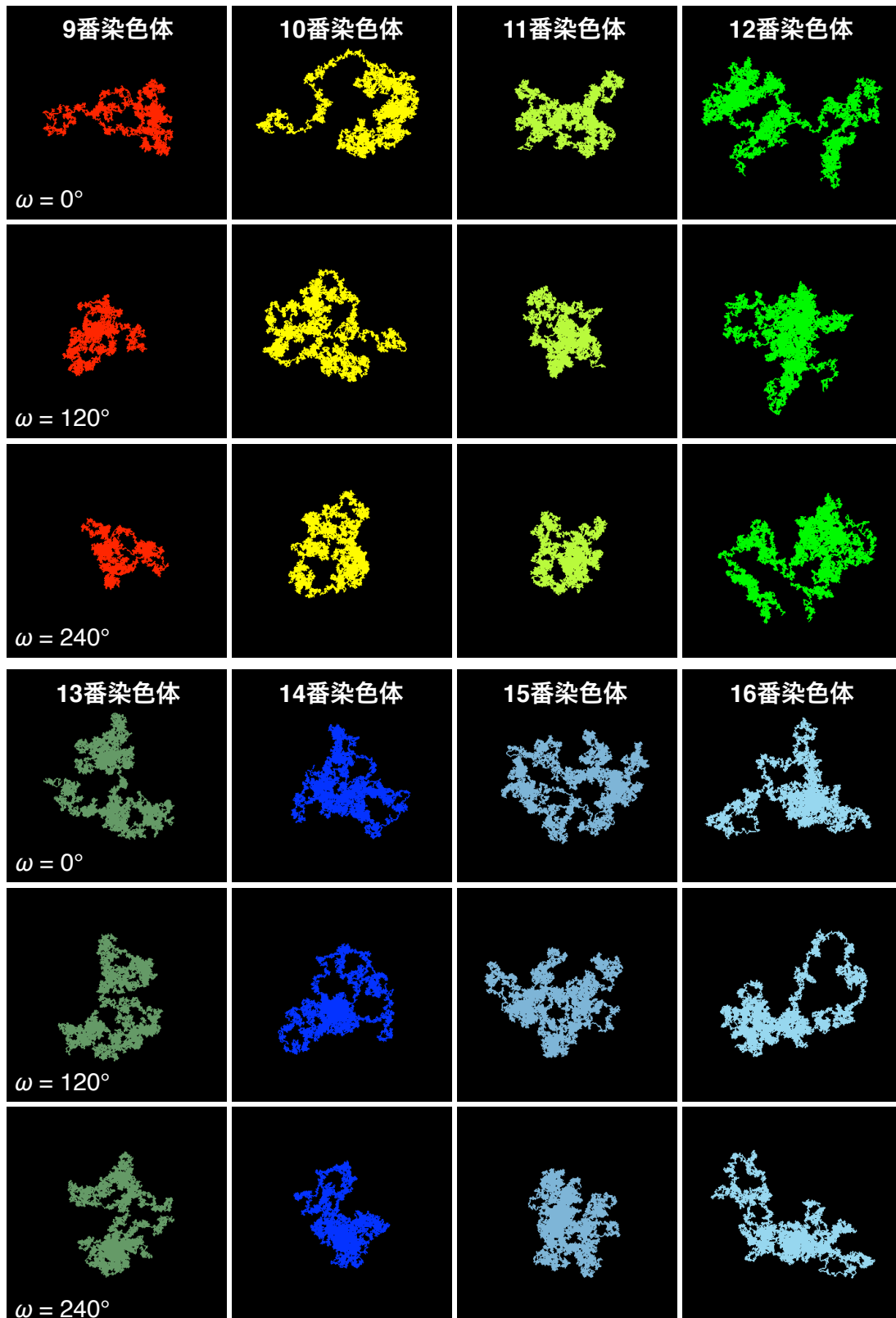


図 3.5. 出芽酵母の間期染色体構造のスナップショット

シミュレーション結果の例を示す。ゲノム全域にわたるヌクレオソームの配置 [51] と、すべてのリンカーDNA の持続長を基にして、出芽酵母の間期染色体構造のシミュレーションを行なった。異なる 3 つの視座から間期染色体構造を示した。なお、12 番染色体に関しては、rDNA 領域の塩基配列を塩基配列データベースから得ることができないため、計算の対象から除いている。また、ここに示した間期染色体構造はすべて直径 2 μm の球に納まる。

3.3.4. 間期染色体構造モデルの評価

本研究で構築したモデルにおいて、間期染色体構造における各々のパラメータ（表 3.3）の重要性を定量的に評価するために、一倍体の出芽酵母の G1 期における、染色体上の 2 点間の空間距離に関する実験データを用いた。しかしながら、既に報告されているデータは限られていたため [66, 67]、FISH (fluorescence *in situ* hybridization) 法を用いて、2 点間空間距離に関する体系的な測定を行なった（図 3.6、図 3.7.A）。この結果を図 3.7.B に示す。なお、縦軸は平均二乗 2 点間空間距離 (μm^2) を、横軸は、2 点がゲノム上でどれほど離れているか (kb; 以下、2 点間ゲノム距離と呼ぶ) を表わしている。図 3.7.B に示したように、2 点間空間距離と 2 点間ゲノム距離との間に、250-500 kb を境界とする、2 相の関係が存在することが判明した。まず、250-500 kb の境界よりも 2 点間ゲノム距離が短い領域については、以下の関係式（スケーリング因子 $\nu \approx 0.5$ [12, 46]）で定義される冪乗則に従うことが明らかになった。

$$\langle R^2 \rangle \sim L^{2\nu}$$

ここで、 $\langle R^2 \rangle$ は平均二乗 2 点間空間距離を表わし、 L は 2 点間ゲノム距離を表わす。解析の結果、スケーリング因子 ν が 0.5 にほぼ等しいことが判明した。従って、距離が 250-500 kb 以下の領域において、間期染色体はランダムウォークポリマー (random-walk polymer) にほぼ等しいことを意味する。一方、250-500 kb の境界よりも 2 点間ゲノム距離が長い領域については、2 点間空間距離は約 1 μm に収束した。すなわち、250-500 kb の境界よりも距離が短い領域に見られたクロマチン繊維の乱雑な動きが、なんらかの制限を受けていることが示唆された。次に、FISH 法を用いて明らかにした以上の結果を、他の手法を用いて確認するために、染色体上の 2 点間の相対的空間距離を検出する実験手法である 4C (chromosome conformation capture-on-chip) を用いて得られたデータ [58] から、図 3.7.B で示した合計 29 組の測定箇所に対応するデータを抽出した（図 3.7.C）。その結果、4C を用いて測定した場合にも、2 点間空間距離と 2 点間ゲノム距離との間に、FISH を用いた場合と同様の関係が存在することが確認された。ちなみにヒトの間期核においては、2 点間空間距離と 2 点間ゲノム距離との間に、約 2 Mb を境界として、2 相の関係が存在することが報告されている [68]。Yokota

らはこの観察結果から、ヒトの間期核において、染色体は異なる2種類の折りたたみ機構を経て収納されていると推論している。

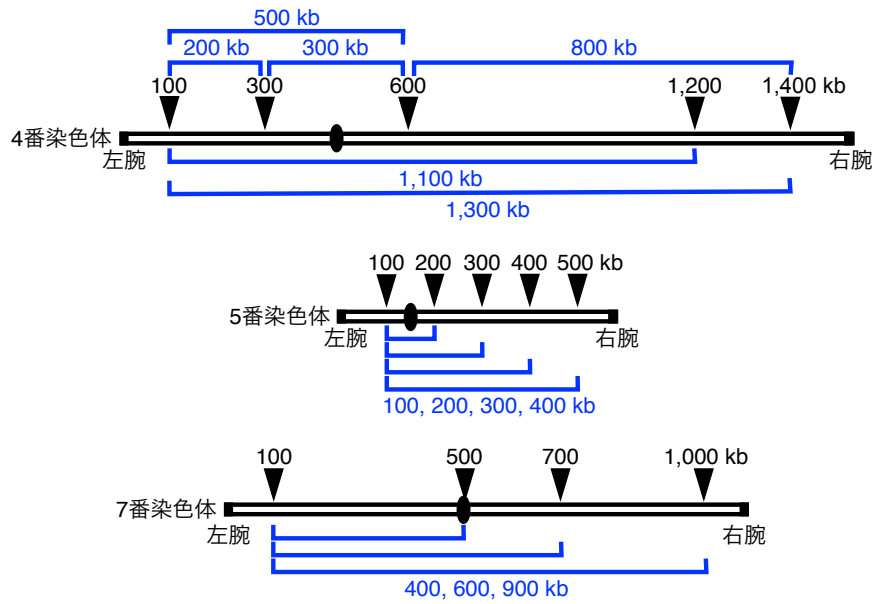
以上の FISH および 4C を用いた解析により判明した、2 点間空間距離と 2 点間ゲノム距離の間の 2 相の関係をもたらした因子は、細胞核であると考えられる。この理由を以下に述べる。第一に、細胞核の存在を考慮に入れてシミュレーションを行なった場合、実験データと同様に、2 点間空間距離と 2 点間ゲノム距離との間に 250-500 kb を境界とした 2 相の関係がみられた (図 3.7 D, E)。なお、モンテカルロ法を用いたシミュレーションでは、各々の染色体は、試行の度に異なる構造をとる (“3.2.1. シミュレーション方法” を参照)。そこで、全 16 本の染色体の各々について 35,000 回以上のシミュレーションを行ない、細胞核 (直径 2 μm の球) に納まる間期染色体構造のサンプルのみを対象として、図 3.7.B で示した合計 29 組の測定箇所の 2 点間空間距離を計算した。その結果、極めて重要なことに、2 点間ゲノム距離が短い領域においても長い領域においても、シミュレーション結果と実験データとの間に非常によい一致がみられた (図 3.7 D, E; “3.4.1. パラメータの選択が間期染色体構造に与える影響” および “3.4.2. 間期染色体構造の決定因子” を参照)。第二に、細胞核の存在を考慮に入れずにシミュレーションを行なった場合、2 点間ゲノム距離が短い領域においても長い領域においても、2 点間空間距離と 2 点間ゲノム距離との間に冪乗則 ($\langle R^2 \rangle \sim L^{2\nu}$; スケーリング因子 $\nu \approx 0.5$) が成り立つことが判明した。すなわち、距離が 250-500 kb 以下の領域において、シミュレーション結果と実験データが一致するものの、距離が 250-500 kb 以上の領域において両者は一致しなかった (図 3.7.F)。第三に、細胞核の直径を 1.8 μm と仮定してシミュレーションを行なった場合、細胞核の直径を本来の 2.0 μm [55] (表 3.3) と仮定した場合 (図 3.7 D, E) と比べて、シミュレーション結果と実験データが一致する度合いが少し低くなった (図 3.7.G)。以上の検討結果から、250-500 kb の境界よりも距離が長い領域において、2 点間空間距離を約 1 μm に収束させる要因、すなわちクロマチン繊維の乱雑な動きを制限する因子は、細胞核である可能性が極めて高いと考えられる。

次に、間期染色体構造におけるリンカーDNAの持続長の重要性を評価した。具体的には、このパラメータをモデルから除いた場合、すなわち、リンカーDNAの三次元的な軌道を記述する“みみず鎖”の結合角 θ に 0° を与えた場合に、間期染色体構造が受ける影響を評価した。この際、生来的にらせん軸の曲がったDNA (intrinsic curved DNA structure) の存在を無視した (生来的にらせん軸の曲がったDNA領域は、多くの場合、ヌクレオソームDNAとしてヌクレオソームに取り込まれていると考えられるため [69])。このような単純化を行なうと、すべてのリンカー領域は真っ直ぐな軌道をとる。このような条件の下でシミュレーションを始めた場合、計算を開始してすぐに、伸長中の鎖と、既に配置された鎖の部分領域との間で衝突が生じた。また、この条件の下では、“引き返し再開法 (back-and-resume procedure; “3.2.1. シミュレーション方法”を参照)”は意味をなさない。そこで、排除体積効果を無視することにより (鎖同士が互いに重なり合うことを許すことにより)、シミュレーションを完遂させた。しかしながら、結果として、3番染色体、4番染色体、12番染色体、15番染色体は、出芽酵母の細胞核に相当する直径 $2\ \mu\text{m}$ の球に納まらなかった。以上から、高い精度で間期染色体構造をモデル化するためには、リンカーDNAの持続長をパラメータとして用いる必要があることが示された。

続いて、ゲノムの核内収納におけるヌクレオソームの重要性を検討した。具体的には、すべてのヌクレオソームを除去した場合、すなわち全てのヌクレオソームDNAが巻き戻されて裸のDNAとなった場合に、ゲノムが核内に収納されるか否かについて解析を行なった。方法としては、巻き戻されたヌクレオソームDNAがとる軌道を、その持続長から求め、構造上のずれやねじれがないようにしてリンカーDNAへ接続させてその結果がどのようになるかを調べた。具体的には、2番目に短い6番染色体と3番目に短い3番染色体を対象としてシミュレーションを行ない、細胞核 (直径 $2\ \mu\text{m}$ の球) に納まる間期染色体構造のサンプルのみを対象として、図3.7.Bで示した合計29組の測定箇所の2点間空間距離を計算した。その結果、2点間空間距離と2点間ゲノム距離との間に2相の関係がみられるものの、その境界は50-100 kbに存在し、この境界よりも2点間ゲノム距離が長い領域においては、2点間空間距離は図3.7 D, Eと同様に約 $1\ \mu\text{m}$ に収束した (図3.7.H)。すなわち、裸のDNAの持続長だけで

シミュレーションを行なった場合の、50-100 kb の境界における 2 点間空間距離と、10 nm クロマチン繊維を折り畳むことでシミュレーションを行なった場合の、250-500 kb の境界における 2 点間空間距離とが等しいことが判明した。換言すると、ゲノム DNA は、ヌクレオソームの形成によって、約 1/5 倍の大きさに圧縮されていることが分かった。この圧縮の程度は、よく知られている圧縮率 (packing ratio) と合致する。さらに、図 3.7.H の結果は、核内に、ヌクレオソームを持たないゲノム DNA のみが存在すると仮定した場合にも、2 点間空間距離は (2 点間ゲノム距離が長い領域において) 細胞核の半径に等しい 1 μm に収束することを示している。

A



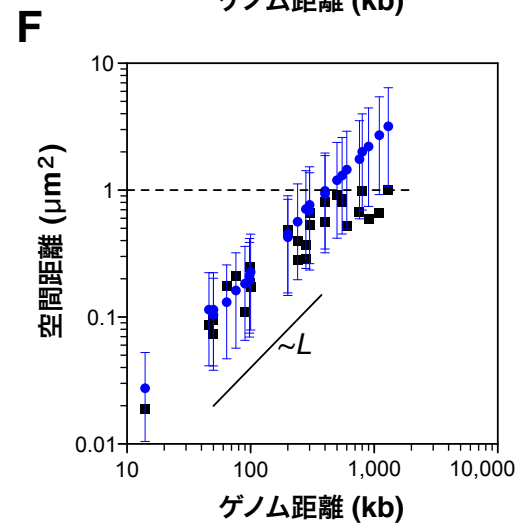
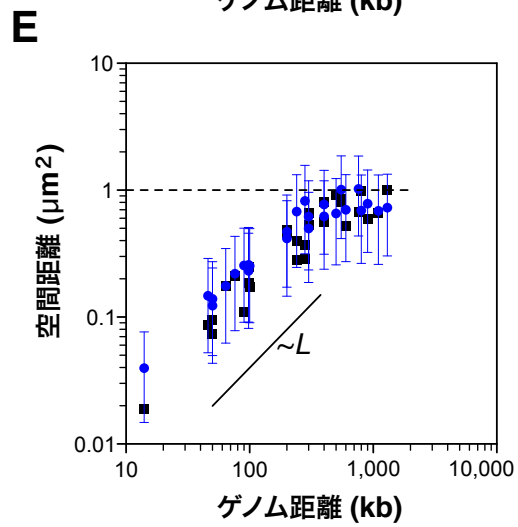
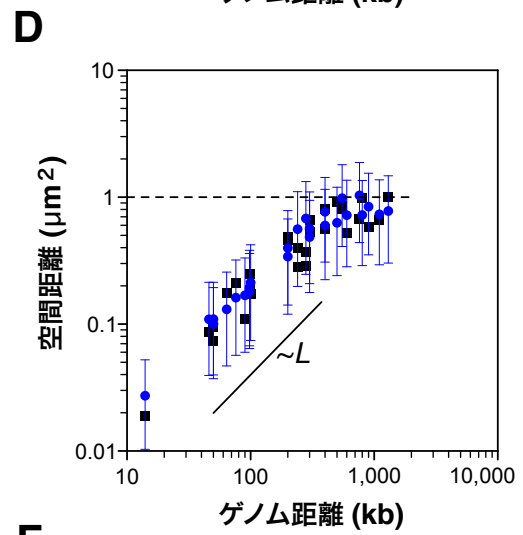
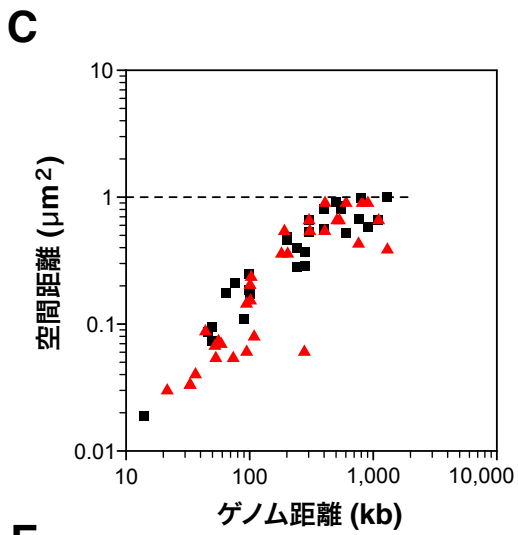
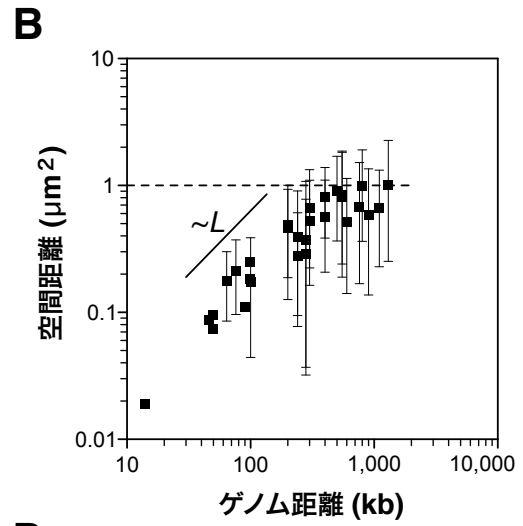
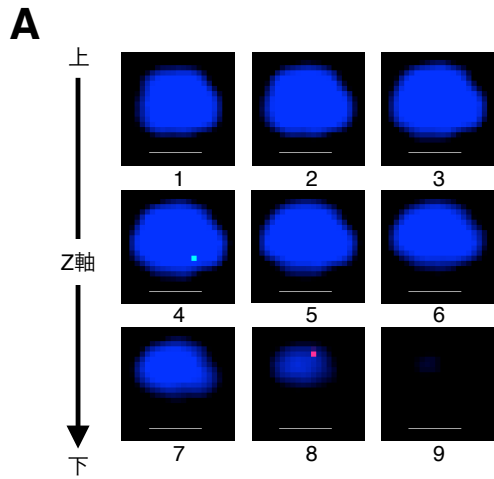
B

プローブ名	染料	染色体	2点間 ゲノム距離 (kb)	観察像	2点間空間距離 (平均値 ± 標準偏差; nm)
IV-100	ローダミン	4番	200		699 ± 266 (n=92)
IV-300	Alexa 488				
IV-300	Alexa 488	4番	300		726 ± 322 (n=76)
IV-600	ローダミン				
IV-100	ローダミン	4番	500		953 ± 349 (n=132)
IV-600	ローダミン				
IV-600	ローダミン	4番	800		991 ± 389 (n=122)
IV-1400	ローダミン				
IV-100	Alexa 488	4番	1,100		813 ± 335 (n=60)
IV-1200	ローダミン				
IV-100	Alexa 488	4番	1,300		1098 ± 550 (n=152)
IV-1400	ローダミン				
V-100	Alexa 488	5番	100		416 ± 206 (n=54)
V-200	ローダミン				
V-100	Alexa 488	5番	200		678 ± 323 (n=47)
V-300	ローダミン				
V-100	Alexa 488	5番	300		813 ± 340 (n=51)
V-400	ローダミン				
V-100	Alexa 488	5番	400		752 ± 297 (n=61)
V-500	ローダミン				
VII-100	Alexa 488	7番	400		898 ± 278 (n=35)
VII-500	ローダミン				
VII-100	Alexa 488	7番	600		720 ± 345 (n=57)
VII-700	ローダミン				
VII-100	Alexa 488	7番	900		766 ± 396 (n=63)
VII-1000	ローダミン				

図 3.6. FISH 法を用いた 2 点間空間距離の測定

(A) 2 点間空間距離を測定した箇所と 2 点間ゲノム距離

(B) 共焦点レーザー顕微鏡による観察像の例と、2 点間空間距離の測定値
緑のシグナルが Alexa488 標識プローブの位置を表わし、赤のシグナルがローダ
ミン標識プローブの位置を表わす。DAPI 染色は、細胞核の確認のために用い
た。スケールバーは 1 μm に相当する。



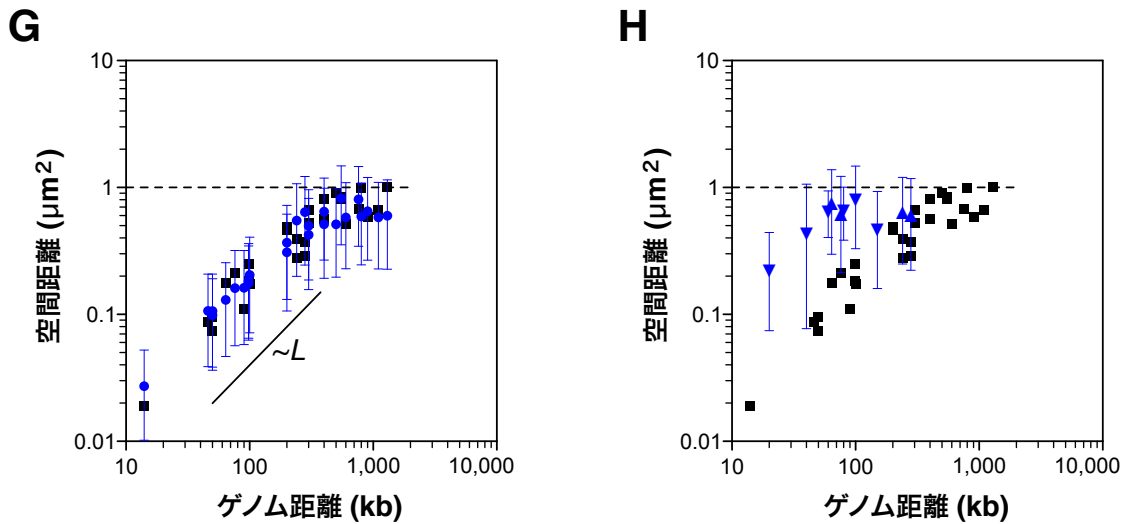


図 3.7. モデルと実験の比較

(A) FISH 法を用いた 2 点間空間距離の測定

Z 軸方向に沿って 0.2 μm 毎に、観察像を得た。FISH 法により可視化した 2 点の間の空間距離を、各々の点の座標および以下の数式を用いて求めた。

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

ここで、 (x_1, y_1, z_1) および (x_2, y_2, z_2) は、2 つのシグナルの三次元空間座標を表わす。なお、全 16 枚の観察像のうち、9 枚の観察像を示した。

(B) FISH 法を用いて測定した空間距離とゲノム距離の関係

平均値 (■) 土標準偏差を示した。Bystricky らのデータ [67] (n=48 - 160)、Bystricky らのデータ [66] (n=31 - 127; 標準偏差なし)、および本研究で FISH 法を用いて測定したデータ (図 3.6; n=35 - 152) を示した。比較のために、以下の図にもこれらのデータ (平均値のみ) を示した。“L” については、本文を参照。

(C) 4C により測定された空間距離とゲノム距離の関係

図 3.7.B で示した合計 29 組の測定箇所に対応する距離データを、染色体上の 2 点間の相対的空間距離を検出する実験手法である 4C を用いて得られたデータ [58] から抽出し、赤三角で示した (“3.2.5. 4C データから 2 点間空間距離を導出する方法” を参照)。

(D, E) モデルと実験の比較

Kaplan ら [51] (D) および Brogaard ら [53] (E) によって報告されたヌクレオソームの配置のデータと、すべてのリンカーDNA の持続長を用いて、間期染色体構造のシミュレーションを 35,000 回以上行なった。シミュレーションにより得られた間期染色体構造のサンプルから、細胞核（直径 2 μm の球）に納まる間期染色体構造のサンプルをスクリーニングし、これらを対象として、(B) で示した合計 29 組の測定箇所 の 2 点間空間距離を計算した (D, $n = 3,065 - 39,557$; E, $n = 1,194 - 38,304$)。平均値（青丸） \pm 標準偏差を示した。

(F, G) 細胞核の存在を考慮に入れずにシミュレーションを行なった場合 (F) および細胞核の直径を 1.8 μm と仮定してシミュレーションを行なった場合 (G) について、(B) で示した合計 29 組の測定箇所 の 2 点間空間距離を計算した (F, $n \geq 35,663$; G, $n = 1,017 - 39,163$)。平均値（青丸） \pm 標準偏差を示した。

(H) ヌクレオソームの存在を考慮に入れずにシミュレーションを行なった場合に間期染色体構造に与える影響

3 番染色体および 6 番染色体を対象として、(D) および (E) と同様にしてシミュレーションを行なった (3 番染色体, $n = 43$; 6 番染色体, $n = 189$)。 (B) で示した合計 29 組の測定箇所のうち、3 番染色体または 6 番染色体に存在する 2 点間の空間距離を青三角で示し、3 番染色体上で無作為に選んだ 2 点の間の空間距離を青逆三角で示した。

3.3.5. 出芽酵母の細胞核におけるクロマチン繊維の態様

図 3.5 に示した全 16 本の間期染色体の構造は、各々の染色体を、他の 15 本の染色体の存在（および核小体の存在）を考慮に入れずにシミュレーションを行って得た構造である。それにも拘らず、2 点間空間距離において、シミュレーション結果と実験データは非常によく合致した（図 3.7 D, E）。この結果は、各々の染色体の構造が、互いに、他の 15 本の染色体（や核小体）の影響を受けないことを示唆している。そこで、出芽酵母の核に相当する直径 $2\ \mu\text{m}$ の球が、全 16 本の染色体や、他の核内因子を納めるために十分な空間であることを確認するため、同一の球内部にて、全ての染色体構造のシミュレーションを行なった。

図 3.8 に示した結果は、全ての間期染色体を内包した核の状態の一例である。これまでに、出芽酵母の間期染色体は、核内において比較的自由に混ざり合っており、少なくとも植物や哺乳類と比べて、より乱雑に配置されていることが報告されている [70, 71]。このような観察事実は、シミュレーション結果と整合する。さらに、シミュレーションにおいて各々の染色体は、ある染色体の窪んだ部分に、別の染色体の突出した部分を配置するなど、限られた空間を効率的に利用している。このように、出芽酵母に相当する直径 $2\ \mu\text{m}$ の球は、効率的に全ての間期染色体を収納できることが判明した。

また、興味深いことに、シミュレーションされた全 16 本の染色体の全体的な構造は、実際の核内構造と非常によく似ていた。すなわち、染色体が占有する領域の他にも、隙間の大きさや形が、生細胞において核小体が占める領域の大きさや形と酷似していた [55, 72]。この隙間に、12 番染色体の rDNA が構成するクロマチン領域が配置されて、核小体が構築されと考えられる。さらに、染色体内部や染色体間に、微小な隙間が存在する。このような微小な空間に、核酸やタンパク質、ならびにこれらの複合体などを収納していると推察される。

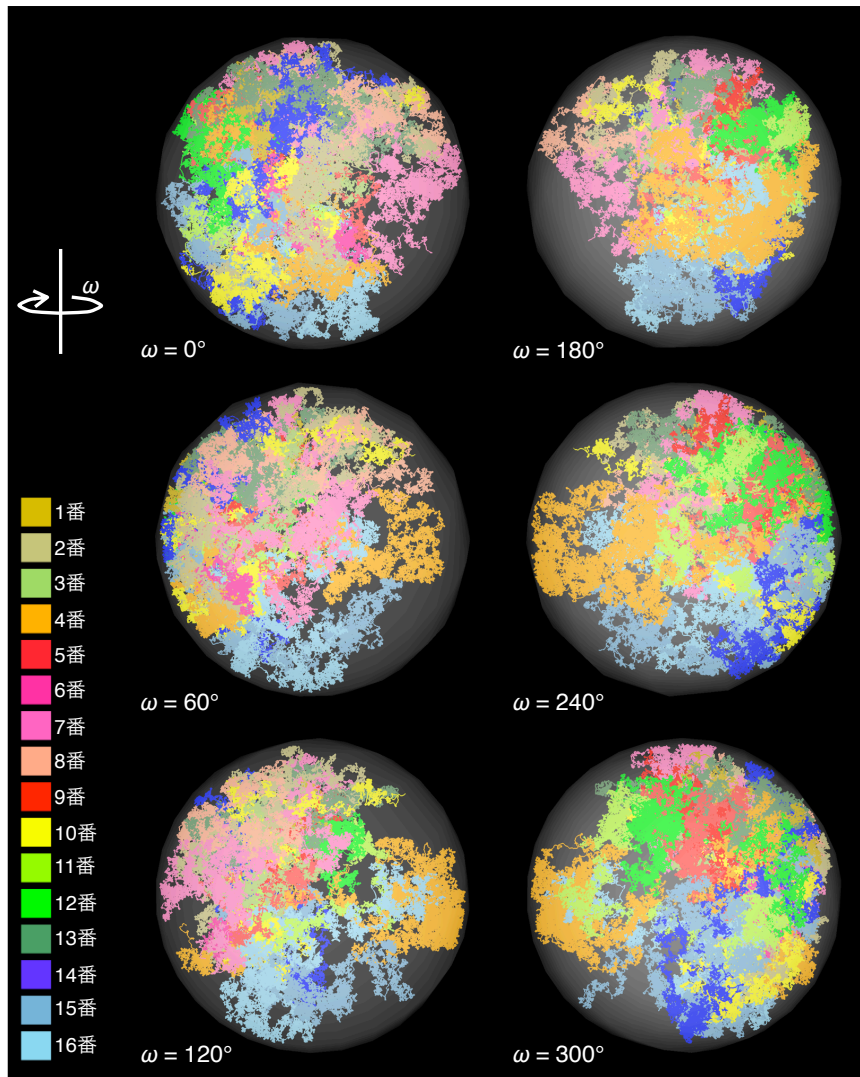


図 3.8. 全 16 本の出芽酵母間期染色体の核内構造シミュレーション
 細胞核を直径 $2 \mu\text{m}$ の球で示し、各々の染色体を図 3.5 と同じ色で示した。はじめに、Duan らが報告しているデータ [58] に基づいて、各々の染色体のセントロメア領域を核内に配置した。次にセントロメアを起点として、左腕および右腕のシミュレーションを行なった。この際、最も長い 4 番染色体のシミュレーションを最初に行い、続いて、長い染色体から順にシミュレーションを行なった (シミュレーションを行なう順序を変えると、全 16 本の染色体の全体的な構造は変化するが、ここではその一例を示した)。なお、12 番染色体に関しては、rDNA 領域の塩基配列を塩基配列データベースから得ることができないため、計算の対象から除いている。また、ヌクレオソームの配置として、Kaplan らが報告しているデータ [51] を用いた。6 つの異なる視座から構造を示した。

3.4. 考察

3.4.1. パラメータの選択が間期染色体構造に与える影響

間期染色体の全長を対象としてモデルを構築するために、3つの因子、すなわち生細胞におけるゲノムワイドなヌクレオソームの配置、リンカーDNAの持続長、細胞核の大きさを用いた。なお、リンカーDNAの持続長に関しては、生理的イオン濃度下で（10 mMの Mg^{2+} 存在下で）原子間力顕微鏡による測定を行なった（図3.3）。本研究で構築した間期染色体構造モデルの最大の特徴は、ボトムアップ・アプローチを用いて構築されていることであり、この点において、従来のトップダウン・アプローチを用いて構築された間期染色体構造モデル（フルスケールモデル）とは大きく異なる（後者においては、クロマチン繊維を染色体全域にわたり均一な物性をもつ架空の高分子として扱うことが多い。また、この架空の繊維を記述する各種のパラメータを、シミュレーション結果と実験データとが合致するように調整する手法をしばしば用いる。ここでは、このような手法をトップダウン・アプローチと呼んでいる）。

極めて重要なことに、本研究で構築した間期染色体構造モデルは、2点間空間距離において、実験データを非常によく説明した（図3.7 D, E）。従って、本研究で用いた3つのパラメータ（ヌクレオソームの配置、リンカーDNAの持続長、細胞核の大きさ）が、間期染色体の三次元構造を規定する本質的な因子であると考えられる。また、この3つの決定因子のひとつであるリンカーDNAの持続長の意義に対する洞察を得るために、すべてのリンカーDNAの持続長を一定の値（50 nm）とした条件の下で、シミュレーションを行なった。その結果、2点間ゲノム距離が長い領域においては、リンカーDNAの持続長を一定の値（50 nm）とした場合と、リンカーDNAの持続長を配列依存的な値とした場合との間に、顕著な違いはみられなかった（図3.9）。この結果は、一般的に予測される通り、マクロな系においては、配列依存的な持続長の効果を無視できることを示している。しかしながら、2点間ゲノム距離が短い領域においては、最大で約12%の違いがみられた（図3.9）。すなわち、リンカーDNAの持続長として、配列依存的な値の代わりに一定の値（50 nm）を用いた場合、詳細なクロマチン構造を再現する能力が大きく低下することが確認された。さらに、

ここで強調すべき点は、本研究が、リンカーDNA の持続長をひとつのパラメータとして用いることにより、その存在自体に議論の余地のある 30 nm クロマチン繊維 [48-50] の持続長を用いずに、モデルの構築に成功している点である。

また、本研究で構築した間期染色体構造モデルは、ヒストンテールに媒介されるヌクレオソーム間の相互作用 [38]、ヒストン修飾がクロマチン繊維の凝縮度合いに変化を与える効果 [73, 74]、ヌクレオソーム間の静電的相互作用、ヌクレオソーム-リンカーDNA 間の静電的相互作用[36]などを無視している。これらの因子は、生細胞において、局所的なクロマチン構造を凝縮させる効果をもつと推察される。事実、図 3.7.E にみられるように、シミュレーション結果は、2 点間ゲノム距離が短い領域において、実験データに比べて僅かに大きい値を示した。この僅かな違いは、モデルと実際のクロマチン繊維との違いを反映していると考えられる。従って、これらの因子をモデルに導入することにより、局所的なクロマチン構造を再現する能力を向上できると考えられる。

一方、DNA らせん軸の屈曲を生じるタンパク質や、クロマチン上の 2 領域間の相互作用を引き起こすタンパク質など、ヒストン八量体以外のタンパク質の存在を無視したことは、間期染色体構造を再現する能力を低下させなかったようである（このようなタンパク質として、クロマチン関連タンパク質や、転写・複製・組換えなどの基礎的な遺伝現象に関わるタンパク質などが挙げられる）。このような結果は、“時間平均 (time average)” または“集合平均 (ensemble average)” の観点から説明することができる（間期染色体は熱運動のために常に構造を変化させており、時間平均または集合平均という概念は、このような性質を議論するために適している）。すなわち、ヒストン八量体と比較して、他のタンパク質がゲノム DNA に滞在する時間は圧倒的に短い。従って、平均すると、他のタンパク質が間期染色体構造に及ぼす影響は非常に小さく、その帰結として、生細胞における平均 2 点間空間距離にはほとんど影響を与えないものと考えられる。

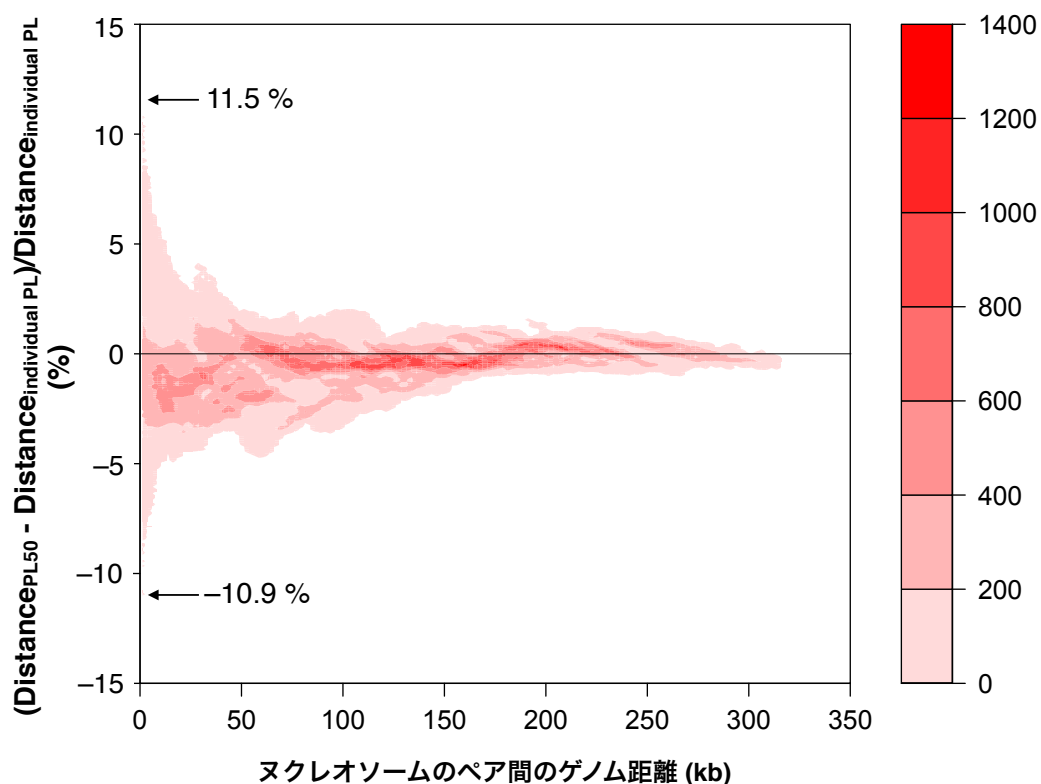


図 3.9. 全リンカーDNA の持続長を一定の値 (50 nm) とした場合に 2 点間空間距離に与える影響

リンカーDNA の持続長として、一定の値 (50 nm) を用いた場合および塩基配列依存的な値を用いた場合のシミュレーションを行い、それぞれの場合について、37,744 個の間期染色体構造のサンプルを得た。解析の対象とした 3 番染色体は、1,491 個のヌクレオソーム [51] をもつため、計 1,110,795 通りのヌクレオソームのペアをつくることができる。これら全てのヌクレオソームのペアに対して、 $\left[(\text{一定の値を用いた場合の平均空間距離}) - (\text{塩基配列依存的な値を用いた場合の平均空間距離}) \right] / (\text{塩基配列依存的な値を用いた場合の平均空間距離})$ (単位は%) を求めた。これらの値を、ヌクレオソームのペア間のゲノム距離 (kb) に対してプロットした。なお、データの密度勾配については右方を参照。また、図中では、一定の値を用いた場合の平均空間距離を $\text{Distance}_{\text{PL50}}$ 、塩基配列依存的な値を用いた場合の平均空間距離を $\text{Distance}_{\text{individual PL}}$ と表わした。なお、PL は persistence length (持続長) を表わす。

3.4.2. 間期染色体構造の決定因子

以上で述べたように、主に3つの因子(ヌクレオソームの配置、リンカーDNAの持続長、細胞核の大きさ)を用いて構築したモデルは、間期染色体構造を非常に高い精度で再現した。ここで、3つの因子のひとつであるリンカーDNAの持続長がDNAの物理的特性のひとつであることに加えて、ヌクレオソームの配置もまた、DNAの物理的特性により決定されると考えられている。一般的に、ヌクレオソームは、約10 bpの間隔で周期的に存在するATおよび約10 bpの間隔で周期的に存在するGCをもつ塩基配列に形成され易い。言い換えれば、ヌクレオソームを形成しやすい塩基配列は、約5bpおきに、主溝が広い領域と主溝が狭い領域とを交互にもち[26, 28]、このような構造は、DNAとヒストン八量体間の相互作用を安定化させる。実際、理論的な手法を用いてヌクレオソームの配置を予測することが可能であるとともに[26]、このような塩基配列に基づいた予測と、生細胞における実際のヌクレオソームの配置はほぼ一致する[51, 53]。以上の議論から、DNAの物理的特性と細胞核の大きさが、出芽酵母の間期染色体の基盤構造の決定因子であると考えられる。

最後に、シミュレーションにより得られた間期染色体構造それ自体の有用性について簡潔に述べる。例えば、予測された構造を用いて、注目する2点間の空間距離や遺伝子間距離などの情報を得られる他に、任意の領域の詳細なクロマチン環境や、遺伝子同士の空間的な配置などの情報を得ることが可能である。さらに、シミュレーションにより得られた間期染色体構造を詳細に調べることで、クロマチン構造と機能に関する、新たな知見や洞察を得ることができると考えられる。

総括

本研究は、DNA の物理的特性の生物学的意義を明らかにすることを目的として行われた。第 1 部では、各種の真核生物ゲノムの柔軟性地図を作成し、解析した結果について報告した。この研究では、極めて柔軟な DNA 領域を発見した。これらの領域は、ゲノムの折りたたみにおける重要な構造因子であると考えられる。第 2 部では、ヌクレオソームを形成しやすい DNA が、約 10 塩基対の間隔で周期的な物理的特性を共通にもつことを発見して報告した。第 3 部では、間期染色体構造のモデルを構築して様々な解析を行なった。そして、DNA の物理的特性と細胞核の大きさのふたつの因子だけで、間期染色体の三次元構造が決定されているという仮説を提唱した。このように、ゲノムの核内収納機構について、俯瞰的な視点から解析を行なった第 1 部、基本的かつ普遍的な構造について解析を行なった第 2 部、そして DNA の物理的特性が直接的に染色体構造に関与することを示した第 3 部を通して、DNA の物理的特性がゲノム DNA の核内収納機構に果たす役割を明らかにした。

謝辞

本研究を遂行し博士論文をまとめるにあたり、早稲田大学 教育・総合科学学術院 大山隆教授には、大学4年生時に研究室へ配属されてから7年間という長きにわたり御指導を賜りました。大山教授には、研究の進め方、考え方、議論の仕方などの研究の基礎から、独創性を追求する態度や「研究を楽しむことが第一」という姿勢まで、研究者として生きるために必要となる多くのことを御教授頂きました。いつも寛大でフランクに接して下さる大山教授と、大きな夢を共有し、ひとつひとつの研究結果に一喜一憂した毎日は、一人前になるための長い鍛錬の日々であるとともに、私にとっての青春でもありました。このような、最高の勉学の機会を与えて下さった大山教授に、心より感謝申し上げます。

学位論文審査において、貴重な御指導を頂いた早稲田大学 教育・総合科学学術院 東中川徹名誉教授、同 並木秀男教授、同 園池公毅教授に、深く感謝致します。特に、東中川徹名誉教授には、大学2,3年生時に分子生物学の基礎を御教授頂いたことに加えて、進路を決めるにあたり有益な御助言を頂きました。

研究の実施にあたり、西川純一博士、下岡保俊氏、三浦理氏の御協力なしには、本研究を完成させることは不可能でした。特に、西川純一博士には多くの重要な御助言を頂きました。また、早稲田大学 大山研究室の一期生と二期生として、共に新規研究課題の立ち上げに奮闘した、景山大氏、古屋美香氏と交わした議論の数々が、本研究を遂行するための貴重な糧となりました。

早稲田大学 理工学部 村田昇教授には、コンピュータプログラミングの基礎を御教授頂きました。慶應義塾大学 医学部 清水信義名誉教授には、ゲノム配列データベースの使い方について御教授頂きました。京都産業大学 理学部 山田修司教授には、第3部の研究の遂行にあたり、お力を貸して頂きました。食品総合研究所 ナノバイオ工学ユニット長 杉山滋博士には、原子間力顕微鏡を用いた測定についてお力を貸して頂きました。深く感謝致します。

最後に、私の研究生生活を支えてくれた家族に、心より感謝します。

参考文献

1. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.
2. Venter JC, Adams MD, Myers EW, *et al.* (2001) The Sequence of the Human Genome. *Science* 291:1304-1351.
3. Birney E, Andrews TD, Bevan P, *et al.* (2004) An Overview of Ensembl. *Genome Res.* 14:925-928.
4. Lindblad-Toh K, Garber M, Zuk O, *et al.* (2011) A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478:476-482.
5. Asthana S, Noble WS, Kryukov G, Grant CE, Sunyaev S, and Stamatoyannopoulos JA (2007) Widely distributed noncoding purifying selection in the human genome. *Proc. Natl Acad. Sci. USA* 104:12410-12415.
6. Pheasant M and Mattick JS (2007) Raising the estimate of functional human sequences. *Genome Res.* 17:1245–1253.
7. Ponting CP and Hardison RC (2011) What fraction of the human genome is functional?. *Genome Res.* 21:1769–1776.
8. Blaxter M (2010) Revealing the dark matter of the genome. *Science* 330:1758–1759.
9. The ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57-74.

10. Misteli T (2007) Beyond the Sequence: Cellular Organization of Genome Function. *Cell* 128:787–800.
11. Emanuel M, Radja NH, Henriksson A, and Schiessel H (2009) The physics behind the larger scale organization of DNA in eukaryotes. *Phys. Biol.* 6:025008.
12. Tark-Dame M, van Driel R, and Heermann DW (2011) Chromatin folding – from biology to polymer models and back. *J. Cell Sci.* 124:839-845.
13. Fukue Y, Sumida N, Nishikawa J, and Ohshima T (2004) Core promoter elements of eukaryotic genes have a highly distinctive mechanical property. *Nucleic Acids Res.* 32:5834-5840.
14. Fukue Y, Sumida N, Tanase J, and Ohshima T (2005) A highly distinctive mechanical property found in the majority of human promoters and its transcriptional relevance. *Nucleic Acids Res.* 33:3821-3827.
15. Packer MJ, Dauncey MP, and Hunter CA (2000) Sequence-dependent DNA structure: tetranucleotide conformational maps. *J. Mol. Biol.* 295:85-103.
16. R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (ISBN 3-900051-07-0; <http://www.R-project.org/>).
17. Smit AFA, Hubley R, and Green P (1996–2004) RepeatMasker Open-3.3.0 (<http://www.repeatmasker.org>).
18. Okada TA and Comings DE (1979) Higher order structure of chromosomes. *Chromosoma* 72:1-14.

19. Ferreira J, Paoletta G, Ramos C, and Lamond AI (1997) Spatial Organization of Large-Scale Chromatin Domains in the Nucleus: A Magnified View of Single Chromosome Territories. *J. Cell Biol.* 139:1597–1610.
20. Vogel MJ, Guelen L, de Wit E, Peric-Hupkes D, Lodén M, Talhout W, Feenstra M, Abbas B, Classen AK, and van Steensel B (2006) Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res.* 16:1493–1504.
21. Cremer T, Cremer M, Dietzel S, Müller S, Solovei I, and Fakan S (2006) Chromosome territories – a functional nuclear landscape. *Curr. Opin. Cell Biol.* 18:307–316.
22. de Wit E and van Steensel B (2009) Chromatin domains in higher eukaryotes: insights from genome-wide mapping studies. *Chromosoma* 118:25–36.
23. Wagner RP, Maguire MP, and Stallings RL (1993) *Chromosomes-a synthesis-* (Wiley-Liss, New York).
24. Wolffe A, *Chromatin-structure & function- 3rd ed.* (Academic Press, San Diego).
25. Travers AA and Klug A (1990) in *DNA TOPOLOGY AND ITS BIOLOGICAL EFFECTS*, Cozzarelli, N.R. and Wang, F.C. eds (CSH, New York), pp.57-107.
26. Segal E, et al. (2006) A genomic code for nucleosome positioning. *Nature* 442:772-778.

27. Brukner I, Sanchez R, Suck D, and Pongor S (1995) *EMBO J.* 14:812-1818.
28. Satchwell SC, Drew HR, Travers AA (1986) Sequence periodicities in chicken nucleosome core DNA. *J Mol Biol* 191:659-675.
29. Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A (1984) Structure of the nucleosome core particle at 7 Å resolution. *Nature* 311:532-537.
30. Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B (2009) The role of DNA shape in protein–DNA recognition. *Nature* 461:1248-1253.
31. Travers A, Hiriart E, Churcher M, Caserta M, Mauro ED (2010) The DNA Sequence-dependence of Nucleosome Positioning *in vivo* and *in vitro*. *J. Biomol. Struct. Dyn.* 27:713-724.
32. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002) Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution. *J. Mol. Biol.* 319:1097-1113.
33. Ransom M, Dennehey BK, and Tyler JK (2010) Chaperoning histones during DNA replication and repair. *Cell* 140:183-195.
34. Svejstrup JQ (2010) The interface between transcription and mechanisms maintaining genome integrity. *Trends Biochem Sci* 35:333-338.
35. Weake VM, Workman JL (2010) Inducible gene expression: diverse regulatory mechanisms. *Nat Rev Genet* 11:426-437.
36. Sun J, Zhang Q, Schlick T (2005) Electrostatic mechanism of nucleosomal array folding revealed by computer simulation. *Proc Natl Acad Sci USA* 102:8180-8185.

37. Kepper N, Foethke D, Stehr R, Wedemann G, Rippe K (2008) Nucleosome Geometry and Internucleosomal Interactions Control the Chromatin Fiber Conformation. *Biophys J* 95:3692-3705.
38. Arya G, Schlick T (2009) A Tale of Tails: How Histone Tails Mediate Chromatin Compaction in Different Salt and Linker Histone Environments. *J Phys Chem A* 113:4045-4059.
39. Perišić O, Collepardo-Guevara R, Schlick T (2010) Modeling studies of chromatin fiber structure as a function of DNA linker length. *J Mol Biol* 403:777-802.
40. Routh A, Sandin S, Rhodes D (2008) Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc Natl Acad Sci USA* 105:8872-8877.
41. Kremer K, Grest GS (1990) Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J Chem Phys* 92:5057.
42. van den Engh G, Sachs R, Trask BJ (1992) Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. *Science* 257:1410-1412.
43. Hahnfeldt P, Hearst JE, Brenner DJ, Sachs RK, Hlatky LR (1993) Polymer models for interphase chromosomes. *Proc Natl Acad Sci USA* 90:7854-7858.
44. Münkler C, Langowski J (1998) Chromosome structure predicted by a polymer model. *Phys Rev E* 57:5888-5896.

45. Rosa A, Everaers R (2008) Structure and dynamics of interphase chromosomes. *PLoS Comput Biol* 4:e1000153.
46. Mateos-Langerak J, et al. (2009) Spatially confined folding of chromatin in the interphase nucleus. *Proc Natl Acad Sci USA* 106:3812-3817.
47. Tokuda N, Terada TP, Sasai M (2012) Dynamical Modeling of Three-Dimensional Genome Organization in Interphase Budding Yeast. *Biophys J* 102:296-304.
48. Eltsov M, Maclellan KM, Maeshima K, Frangakis AS, Dubochet J (2008) Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc Natl Acad Sci USA* 105:19732-19737.
49. Fussner E, Ching RW, Bazett-Jones DP (2011) Living without 30 nm chromatin fibers. *Trends Biochem Sci* 36:1-6.
50. Nishino Y, et al. (2012) Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *EMBO J* 31:1644-1653.
51. Kaplan N, et al. (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458:362-366.
52. Cole HA, Howard BH, Clark DJ (2011) Activation-induced disruption of nucleosome position clusters on the coding regions of Gcn4-dependent genes extends into neighbouring genes. *Nucleic Acids Res* 39:9521-9535.
53. Brogaard K, Xi L, Wang JP, Widom J (2012) A map of nucleosome positions in yeast at base-pair resolution. *Nature* 486:496-501.

54. Wall FT, Erpenbeck JJ (1959) New Method for the Statistical Computation of Polymer Dimensions. *J Chem Phys* 30:634
55. Berger AB, et al. (2008) High-resolution statistical mapping reveals gene territories in live yeast. *Nat Methods* 5:1031-1037.
56. Welzl E (1991) Smallest enclosing disks (balls and ellipsoids). *Lecture Notes in Computer Science* 555:359-370.
57. Gotta M, Laroche T, Formenton A, Maillet L, Scherthan H, Gasser SM (1996) The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *J Cell Biol* 134:1349-1363.
58. Duan Z, et al. (2010) A three-dimensional model of the yeast genome. *Nature* 465:363-367.
59. Håkansson C, Elvingson C (1994) Semiflexible chain molecules with nonuniform curvature. 1. Structural properties. *Macromolecules* 27:3843-3849.
60. Yamakawa H (1971) *Modern Theory of Polymer Solutions* (Harper & Row, Publishers, Inc., New York).
61. Rivetti C, Guthold M, Bustamante C (1996) Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J Mol Biol* 264:919-932.

62. Ullner M (2008) "Polyelectrolytes. Physicochemical aspects and biological significance." in *DNA Interactions with Polymers and Surfactants*, eds Dias RS, Lindman B (John Wiley & Sons, Inc., New Jersey), pp 1-39.
63. Dlakic M, Harrington RE (1998) DIAMOD: display and modeling of DNA bending. *Bioinformatics* 14:326-331.
64. Verdone L, Camilloni G, Di Mauro E, Caserta M (1996) Chromatin remodeling during *Saccharomyces cerevisiae ADH2* gene activation. *Mol Cell Biol* 16:1978-1988.
65. Travers A, et al. (2009) Nucleosome positioning-what do we really know? *Mol Biosyst* 5:1582-1592.
66. Bystricky K, Heun P, Gehlen L, Langowski J, Gasser SM (2004) Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. *Proc Natl Acad Sci USA* 101:16495-16500.
67. Bystricky K, Laroche T, van Houwe G, Blaszczyk M, Gasser SM (2005) Chromosome looping in yeast: telomere pairing and coordinated movement reflect anchoring efficiency and territorial organization. *J Cell Biol* 168:375-387.
68. Yokota H, van den Engh G, Hearst JE, Sachs RK, Trask BJ (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. *J Cell Biol* 130:1239-1249.

69. Ohshima T (2005) "The role of unusual DNA structures in chromatin organization for transcription." in *DNA Conformation and Transcription*, ed Ohshima T (Springer, New York), pp 177-188.
70. Haber JE, Leung WY (1996) Lack of chromosome territoriality in yeast: promiscuous rejoining of broken chromosome ends. *Proc Natl Acad Sci USA* 93:13949-13954.
71. Meaburn KJ, Misteli T (2007) Chromosome territories. *Nature* 445:379–381.
72. Sàez-Vàsquez J, Gadal O (2010) Genome organization and function: a view from yeast and Arabidopsis. *Mol Plant* 3:678-690.
73. Lachner M, Jenuwein T (2002) The many faces of histone lysine methylation. *Curr Opin Cell Biol* 14:286-298.
74. Trojer P, Reinberg D (2007) Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* 28:1-13.

研究業績

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
論文 ○	<p>1. <u>Kimura H</u>, Kageyama D, Furuya M, Sugiyama S, Murata N, Ohyama T Regions with Unusually High Flexibility Occur Frequently in Human Genomic DNA. <i>Biosci Biotechnol Biochem</i>, in press.</p> <p>○ 2. <u>Kimura H</u>, Ohyama T (2008) A common mechanical property shared by yeast nucleosomal DNAs. <i>J Adv Sci</i> 20:37-40.</p> <p>3. Udagawa K, <u>Kimura H</u>, Tanabe H, Ohyama T (2011) Nuclear localization of reporter genes activated by curved DNA. <i>J Biosci Bioeng</i> 113:431-437.</p>
総説	<p>1. 大山 隆、木村 元、下岡 保俊. ゲノムDNAはいかにして折り畳まれるか—DNA物性とゲノム収納 (2009) 実験医学増刊 細胞核 27, 2715-2722.</p>
講演	<p>1. Ohyama T, <u>Kimura H</u>, Nishikawa J, Shimooka Y, Miura O, Sugiyama S, Yamada S THE GENOME FOLDING PRINCIPLE AS REVEALED BY MODELING OF YEAST INTERPHASE CHROMOSOMES. The 22nd Wilhelm Bernhard Workshop 2011 Aug. Riga, Latvia</p> <p>2. <u>Kimura H</u>, Shimooka Y, Nishikawa J, Miura O, Sugiyama S, Yamada S, Ohyama T THE GENOME FOLDING PRINCIPLE IN YEAST. The 22nd Wilhelm Bernhard Workshop 2011 Aug. Riga, Latvia</p> <p>3. <u>Kimura H</u>, Shimooka Y, Nishikawa J, Miura O, Yamada S, Ohyama T The prime determinants of interphase chromosome structures. International Symposium on the Physicochemical Field for Genetic Activities 2011 Jan. Awaji Island, Japan</p> <p>4. <u>Kimura H</u>, Miura O, Arakawa J, Seki M, Yamada S, Ohyama T DNA Flexibility Maps of Eukaryotic Genomes. 50th ASCB ANNUAL MEETING 2010 Dec. Philadelphia, USA</p> <p>5. <u>Kimura H</u>, Shimooka Y, Nishikawa J, Miura O, Yamada S, Ohyama T Three-dimensional architecture of chromatin fibers in the interphase nucleus. 50th ASCB ANNUAL MEETING 2010 Dec. Philadelphia, USA</p> <p>6. <u>Kimura H</u>, Shimooka Y, Sugiyama S, Ohyama T 3D STRUCTURE OF 10 NM CHROMATIN FIBER DEDUCED FROM DNA FLEXIBILITY. 2nd HOPE Meeting 2009 Sep. Hakone, Japan</p> <p>7. <u>Kimura H</u>, Shimooka Y, Sugiyama S, Ohyama T 3D STRUCTURE OF 10 NM CHROMATIN FIBER DEDUCED FROM DNA FLEXIBILITY. The Wilhelm Bernhard Work shop 21st International Workshop on the cell Nucleus 2009 Sep. Ustron, Poland</p>

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
	<p>8. Ohyama T, <u>Kimura H</u>, Shimooka Y, Kageyama D, Furuya M, Udagawa K, Tanase J How genomic DNA is functionally folded in a nucleus. The Wilhelm Bernhard Workshop 21st International Workshop on the cell Nucleus 2009 Sep. Ustron, Poland</p> <p>9. Ohyama T, <u>Kimura H</u>, Kageyama D, Furuya M The DNA flexibility maps of eukaryotic genomes. 8th EMBL Transcription meeting 2009 Aug. Heidelberg, Germany</p> <p>10. Ohyama T, <u>Kimura H</u>, Kageyama D, Furuya M HOW GENOMIC DNA IS FUNCTIONALLY FOLDED IN A NUCLEUS. International Symposium on Chromosome Dynamics in Ise 2008 May Ise, Japan</p> <p>11. Ohyama T, Tanase J, <u>Kimura H</u> GENETIC INFORMATION CARRIED IN DNA CONFORMATION AND PROPERTIES. Embo Conference Series on Nuclear Structure and Dynamics 2007 Sep. Montpellier, France</p> <p>12. Ohyama T, Sumida N, Inoue S, Tanase J, <u>Kimura H</u>, Miura M HOW GENOMIC DNA IS FUNCTIONALLY FOLDED IN A NUCLEUS. Functional Organization of the Nucleus 2007 Jan. Awaji Island, Japan</p> <p>13. <u>木村 元</u>、西川 純一、下岡 保俊、三浦 理、杉山 滋、山田 修司、大山 隆。 The principle in structuring 3D architecture of interphase chromosomes. 第 34 回日本 分子生物学会年会 2011 年 12 月 横浜</p> <p>14. <u>木村 元</u>、下岡 保俊、西川 純一、三浦 理、山田 修司、大山 隆。出芽酵母 間期染色体の三次元構造とゲノム折り畳み原理. 第 28 回染色体ワークショップ 2011 年 1 月 山代温泉</p> <p>15. <u>木村 元</u>、下岡 保俊、西川 純一、三浦 理、大山 隆。間期クロマチン三次元 構造の構築原理。第 33 回日本分子生物学会年会 2010 年 12 月 神戸</p> <p>16. 大山 隆、<u>木村 元</u>、下岡 保俊、西川 純一、三浦 理、荒川 潤、横尾 岳大、 関 瑞穂。間期クロマチンの三次元構造とゲノム収納の原理。第 33 回日本分子 生物学会年会 2010 年 12 月 神戸</p> <p>17. <u>木村 元</u>、下岡 保俊、西川 純一、三浦 理、大山 隆。間期クロマチン三次元 構造の構築原理。「細胞を創る」研究会 3.0 2010 年 11 月 東京大学 生産技術 研究所</p> <p>18. <u>木村 元</u>、下岡 保俊、西川 純一、三浦 理、杉山 滋、大山 隆。間期クロマ チン構造の <i>in silico</i> モデリング. 第 27 回染色体ワークショップ 2010 年 1 月 御 殿場</p>

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
	<p>19. 大山 隆、古屋 美香、荒川 潤、木村 元、下岡 保俊、西川 純一、浅野 士郎、杉山 滋. ゲノム収納の原理とクロマチン繊維の基本構造. 第 27 回染色体ワークショップ 2010 年 1 月 御殿場</p> <p>20. 木村 元、下岡 保俊、三浦 理、杉山 滋、大山 隆. 間期クロマチン構造の <i>in silico</i> モデリング. 第 32 回日本分子生物学会年会 2009 年 12 月 横浜</p> <p>21. 大山 隆、木村 元、古屋 美香、荒川 潤、下岡 保俊、杉山 滋. ゲノム収納の原理とクロマチン繊維の基本構造. 第 32 回日本分子生物学会年会 2009 年 12 月 横浜</p> <p>22. 木村 元、下岡 保俊、杉山 滋、大山 隆. 間期細胞核におけるクロマチン繊維の構造. 第 82 回日本生化学会大会 2009 年 10 月 神戸</p> <p>23. 木村 元、下岡 保俊、古屋 美香、景山 大、杉山 滋、大山 隆. クロマチン繊維の自己組織化. 第 26 回染色体ワークショップ 2009 年 1 月 姫路</p> <p>24. 木村 元、景山 大、古屋 美香、村田 昇、下岡 保俊、大山 隆. 真核生物ゲノムの DNA 物性とクロマチン基盤構造. 第 31 回日本分子生物学会年会 2008 年 12 月 神戸ポートアイランド</p> <p>25. 大山 隆、木村 元、景山 大、古屋 美香、深川 竜郎、浅野 士郎. ゲノムの階層的折り畳みの基盤となる DNA 物性. 第 31 回日本分子生物学会年会 2008 年 12 月 神戸ポートアイランド</p> <p>26. 木村 元、古屋 美香、景山 大、深川 竜郎、大山 隆. エピジェネティクスの DNA 基盤—ヒトゲノムの物理的特性に印された遺伝情報—. 第 2 回エピジェネティクス研究会 2008 年 5 月 東レ総合研修センター (静岡県三島市)</p> <p>27. 木村 元、景山 大、古屋 美香、大山 隆. ヒトゲノムの機械的特性と染色体構造. 第 25 回染色体ワークショップ 2008 年 1 月 湯河原</p> <p>28. 木村 元、古屋 美香、景山 大、村田 昇、大山 隆. ヒトゲノムの機械的特性地図. 第 30 回日本分子生物学会年会 2007 年 12 月 横浜</p>