Doctoral Dissertation

Academic Year 2015

Analysis of chromatin structure and dynamics with single nucleosome imaging

Graduate School of Media and Governance

Keio University

Tadasu Nozaki

Doctoral Dissertation
Academic Year 2015

Analysis of chromatin structure and dynamics with single nucleosome imaging

Abstract

DNA is the media of hereditary information of living organisms on the earth. In eukaryotes, long genomic DNA is wrapped around core histones to form a nucleosome fiber. In addition, nucleosome fiber is organized three dimensionally as chromatin and folded within a tiny space such as a nucleus or chromosomes. In cellular functions based on DNA transaction including transcription and DNA replication, the cells must precisely execute the readout of information superimposed on genomic DNA. In this situation, a protein as molecular machinery needs to move around such complicated folded DNA and interrogate its target site, and the chromatin becomes a huge obstacle. Thus, chromatin structure and dynamics seem to be deeply related to diverse cellular functions but not merely packaging, and play a fundamental role in epigenetic regulation. Therefore, making the detailed observation and description of chromatin is essential to understand the various life phenomena in eukaryotes. The objective of this thesis is to contemplate the relationship between chromatin and epigenetics. To achieve this objective, multiple methods including genome-wide analysis of epigenetic regulation and live cell imaging of chromatin were performed. First, the epigenetic regulation and promoter types were analyzed based on genome-wide ChIP-Seq and DeepCAGE data. It was found that, in humans, broad promoters but not peak promoters had significant associations with histone modification and nucleosome position. Secondly, single nucleosome imaging was newly developed. To observe single nucleosomes in living mammalian cells, the histone was fused with photoactivatable (PA)-green fluorescent protein (GFP) and expressed in mammalian cells at a very low level. For single nucleosome imaging, an oblique illumination was used to illuminate a limited thin area within the cell. It was found that a small fraction of histone-PA-GFP was spontaneously activated without laser stimulation. Lastly, based on single nucleosome imaging, the chromatin structure and dynamics at single nucleosome level in a living mammalian cell and their variation related to different types of histone modification and chromatin associated proteins were successfully observed. Furthermore, numerous chromatin domains were identified throughout the cell cycle. These results will provide a basis of dynamic and flexible nature of chromatin toward understanding the function of chromatin in eukaryotic cell and allow new insight and perspective on the relationships between chromatin and epigenetics.

Keywords: chromatin, epigenetics, histone modification, single molecule imaging, super resolution imaging

論文題目

一分子ヌクレオソームイメージング法を用いたクロマチン構造と動態の解析

論文要旨

DNA は超高密度の情報媒体である．長大な真核細胞のゲノム DNA は，ヒストンタンパク質に巻き付くことで，ヌクレオソームを形成する．ヌクレオソームは，さらに三次元的に折りたたまれることで，クロマチン構造を形成し，核や染色体のような微小な空間に折り畳まれている．転写やDNA複製において，生命の基本機能に関わるゲノムDNA情報の読み出しを，細胞は正確に実行しなければならない．分子機械であるタンパク質は，情報読み出しのために, 核酸やタンパク質で混雑している核内を移動し，標的 DNA サイトを検索するが，その際クロマチンは障害となる．そのため，DNA に関連する様々な細胞機能にクロマチンの構造と動態は関与し，その結果，エピジェネティック制御の根幹を担う．さらに，分化を伴う多細胞生物は，クロマチン構造と動態を変化させることによって，その細胞機能に応じて検索可能な情報を選択・限定していると思われる．したがって，クロマチンの詳細な観察と記述は，真核生物における様々な生命現象を理解するために必須である．本論文の目的は，クロマチンに関する詳細な観察と記述を行い，クロマチンとエピジェネティクスの関係性を考察することである．その目的を達成するために，エピジェネティック制御のゲノムワイド解析とクロマチンのライブセルイメージングを行った．まず，ヒストン修飾によるエピジェネティック制御とプロモーターの関係性を明らかにするために，ヒストンの修飾とポジションに関する ChIP-Seq データの解析をゲノムワイドに行った．その結果，プロモータータイプとエピジェネティック制御の関係性について明らかにした．次に，斜光照明法によって，自発的に活性化した光活性化タンパク質を観察することで，生細胞におけるヌクレオソームの一分子イメージングを成功させた．最後に，生細胞におけるクロマチンの構造と動態を一分子レベルで，網羅的に観察した．その結果，クロマチン環境の多様性はヒストン修飾，もしくはクロマチン関連タンパク質と関係していることを明らかにした. さらに，細胞周期を通して，類似したクロマチン構造が観察されることを明らかにした．これらの結果は，クロマチンの動的で柔らかい性質を明らかにし，クロマチン環境のエピジェネティック制御における重要性を示唆している．

キーワード: クロマチン，エピジェネティクス，ヒストン修飾，一分子イメージング，超解像イメージング

Contents

Figure List

Chapter 1

Introduction

1.1. DNA as information media of cell

DNA is the media of hereditary information of living organisms and consists of four types of nucleotides: A (adenine), T (thymine), G (guanine), and C (cytosine). Living organisms are characterized through the expression of genes that are the units of cell function. DNA satisfies the complexity of gene functions and genome encompasses the entire set of genetic information required to characterize a species. The gene coded on DNA is transcribed into the RNA, which is then translated to the protein. Through such information processing, also known as "central dogma", the information of gene coded on DNA is translated and physically transformed into a protein, which acts as a molecular machinery in a cell (Crick, 1970). This information processing occurs in all living organisms. Moreover, to proliferate the number of cells and duplicate the genomic information, DNA is replicated and the genomic information is precisely transmitted from a cell to daughter cells during cell division. Additionally, DNA is stable to thermal perturbation, and DNA replication keeps a low error rate because of the complementarity and DNA repair system.

In terms of macro-scale, the cells, which are the basic units of life, are organized into a tissue, the tissues are organized into an organ, and the organs are organized into an organism (Simon, 1991). In terms of micro-scale, the nucleus, cell membrane, cytoskeleton, endoplasmic reticulum, and mitochondria can be found as the functional subcomponents of a cell. Each cell is comprised of subsystems including metabolic pathway, transcriptional system, DNA replication system, DNA repair system, immune system, signaling network, and so on. Thus, a living organism is comprised of hierarchical and complicated systems (Simon, 1991). All the information for creating and maintaining living organisms are certainly coded on DNA.

Moreover, DNA has a physical aspect; therefore, it has a three-dimensional entity including shape and length. In eukaryotes, the huge length of linear genomic DNA is packed into a tiny space. For instance, the total length of genomic DNA in a human diploid cell reaches approximately 2 m, but the diameter of the nucleus into which the DNA is packed is just about 10 μm (a volume of only ~100 fL to 1 pL). The folding of DNA is a complicated process because of its negative charges with the phosphate groups and repulsion derived on them. Such a complicated folding of DNA is achieved by specialized proteins, which neutralize the negative charges of DNA, and the complex of DNA and binding proteins is called "chromatin". Nevertheless, the detail of the folding process has been unclear (Figure 1.1).

To execute the program superimposed on DNA, the proteins, for instance, transcriptional machinery, must interrogate their target sites on such complicated folded DNA. But how is it possible? We know that a diploid cell has only two loci of the target gene and a small number of the transcription factor, which does not have "the map of nucleus leading to the target", have to find their target sites in limited time and with less energy. In our life, we usually use many efficient algorithms, e.g., SSEARCH based on Smith-Waterman algorithm, to perform pair-wise alignment of DNA sequences (Smith and Waterman, 1981a, b) and Google based on Page Rank algorithm to search information in the vast and infinite internet (Brin and Page, 1998). However, how does the cell efficiently search and find the information on DNA? We are not even sure of "the algorithm in cell" at all. The proteins in the nucleus search their target sites by Brownian motion. When a protein searches its target, the chromatin structure becomes a big obstacle (Bancaud et al., 2009; Gorisch et al., 2005); therefore, the chromatin structure is considered to be deeply related to diverse cellular functions including transcriptional regulation, DNA replication, DNA repair, and DNA recombination, and not merely packaging. Moreover, the variation in the readout of genetic information from identical genomes in a multi-cellular organism has been considered to be controlled by chromatins. Thus, for the investigation of the search algorithm in a cell, it is important to make detailed observation and description of chromatin.

Although the study on chromatin began in the 19th century and chromatin is one of the oldest research subjects in cell biology, we have understood only a part of the chromatin structure and function. Understanding the detail of chromatin can lead us to understand "the algorithm in cell" and the various cellular functions. Thus, I begin this chapter by describing the history of chromatin research and then present the concept of epigenetics. Finally, I will present the objective of this thesis.

1.2. Chromatin and heredity

Microscopy is the essential technology to capture the biological phenomena. The invention and development of microscopy have enabled us to observe the invisible micro-world and led us to the cellular world. The observation of "happy" living cells (but not fixed or dead) is vital to understand the dynamic mechanism in the organism (Figure 1.2). Thus, the limitation of observable micro-world is defined by the power of microscopy, and the new technology that makes the invisible visible is important for the progress of biological study. In the 17th century, Hooke first observed myriads of small rooms in cork by his own microscopy, and he termed the small room "cell" as an empty vessel (Hooke, 1665). Then, in the 19th century, Schleiden (Schleiden, 1838) and Schwann (Schwann, 1839) showed the importance of cells in plants and animals, respectively. They proposed that the bodies of all species of plants and animals are composed of cells. Subsequently, it was observed that cell division and daughter cells are produced from the duplication of a mother cell (Meyen, 1830; Mohl, 1835; Nägeli, 1842; Remak, 1852, 1855), and this was summarized as Virchow's rule "Omnis cellula e cellula" (Virchow, 1858). Therefore, microscopy led us to the cellular world and gave us two big rules; the complicated body of an organism consists of cells, and the cells are created from a cell.

Although the importance of a cell and inheritance in life was shown, the researchers had guessed that if the cell was the unit of a living organism, then it might have information coded on the subject that was inherited from parents to a child. However, they did not know about them at that time. Later, the eukaryotic nucleus that has a densely stained region was discovered, and the dense structure in the nucleus was termed "chromatin" (Brown, 1833; Flemming, 1878, 1879). During the same period, it was discovered that chromatic elements in the nucleus showed varying forms, such as rod-like structures, during cell division (Flemming, 1882; Nägeli, 1842, 1844; Strasburger, 1880; Waldeyer, 1888). Flemming identified that chromatin in the nucleus was correlated with rod-like structures during nuclear division, and such chromatic elements having rod-like structures during cell division were termed chromosomes (Waldeyer, 1888).

Such behavior of chromosomes during cell division implied the role of chromosomes in heredity. In 1900, the importance of Mendel's study was rediscovered and recognized by several researchers simultaneously (Correns, 1900; De Vries, 1900; Tschermak, 1900). Mendel had proposed the mathematical logic to the inheritance of phenotypic traits before the discovery of mitotic chromosome (Mendel, 1866). Immediately after the rediscovery of Mendel's law, Sutton and Boveri proposed chromosome theory, which represented that chromosomes carry the genetic information and genes are located on chromosomes (Sutton, 1902). This was later

validated by Morgan (Morgan, 1910; Morgan, 1911). It was revealed that gene is the unit of cell function and that genome is assembled of genes and contains the totality of information to characterize the species. The generalization and abstraction of chromosome behavior in heredity by Mendel seem marvelous.

It appeared that chromatin is the carrying trait of genes and that the chromosome is a variant shape of chromatin appearing at mitosis. The relationship between genetic phenomena and physical chromosome behavior was established, and subsequent experiments showed that the DNA is the minimal component of genes (Watson and Crick, 1953).

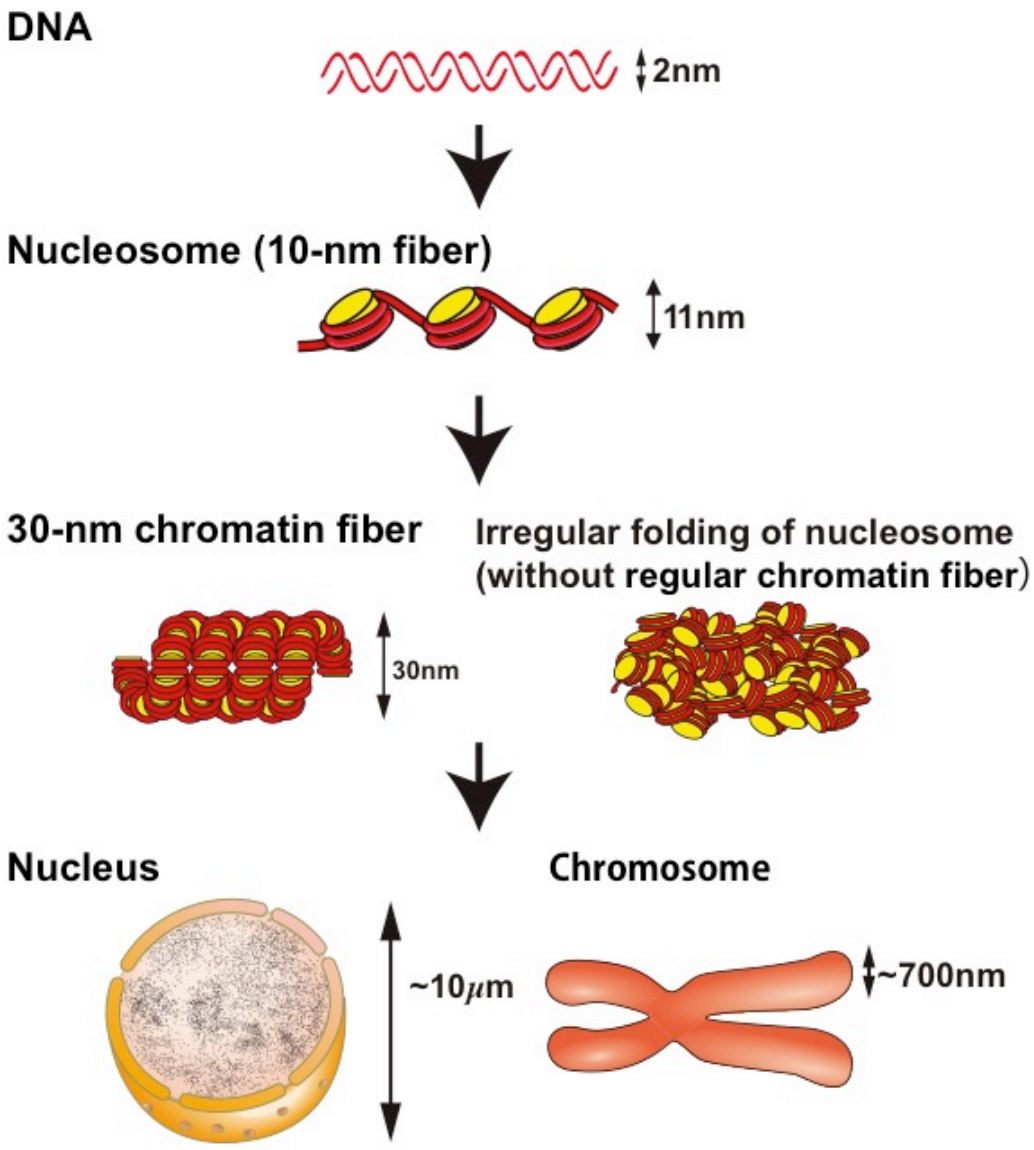

Figure 1.1 From DNA to chromatin

In eukaryotes, long genomic DNA is packed into a tiny space of nucleus or chromosomes. DNA is wrapped around histone proteins comprised at two copies of each H2A, H2B, H3, and H4 protein and forms nucleosome. Histone proteins have positively charged amino acids containing the lysine and arginine, and they neutralize a part of negative charges on DNA. Nucleosomes have been assumed to be folded into regular 30-nm chromatin fiber. However, currently, nucleosomes have been considered to be folded irregularly and form chromatin domains that do not possess the regular size (figure from (Maeshima et al., 2014)).

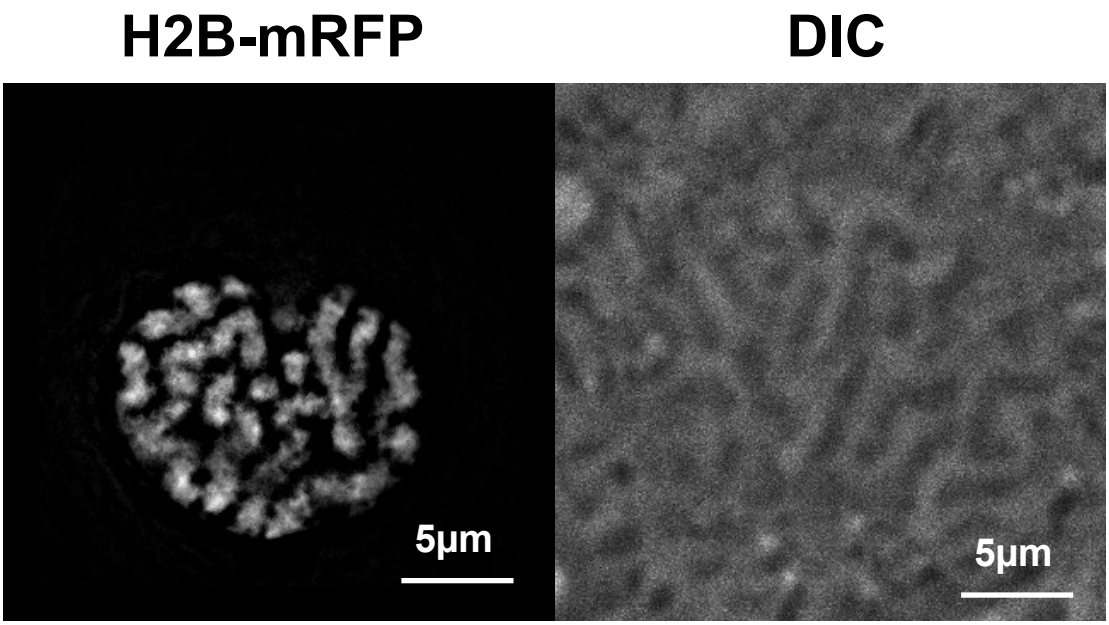

Figure 1.2 Imaging of chromosomes with microscopy

Fluorescent microscopy has a big power when observing target proteins or target objects in a living cell. H2B-mRFP shows the clear rod-like structure of chromosomes in a living cell. In contrast, non-fluorescent microscopy DIC (differential interference contrast microscopy) also shows chromosomes. Fluorescent microscopy is very useful when the targets are already determined, but it is important to remember that it reduces much information. Non-fluorescent microscopy is also important for cell biology.

1.3. Chromatin structure and dynamics

After the discovery of the double-helical structure of DNA and central dogma, researchers continued to study of components in chromatin and chromosome although it had been established that the genetic information is coded in DNA. In 1974, a number of interphase nuclei were isolated, and the "beads on string" structure in chromatin was visualized by electron microscopy (EM) (Olins and Olins, 1974). Additionally, the 200 bp repetitive structure was investigated by nuclease digestion (Kornberg, 1974; Kornberg and Thomas, 1974) and was coined "nucleosome" (Oudet et al., 1975). Presently, the structure of nucleosome is well known at the atomic resolution (1.9 A) (Davey et al., 2002). It was revealed as the detail of nucleosome structure that 147 bp of DNA is wrapped in 1.7 left-handed superhelical turns around a core histone octamer, which consists of two copies of the four histones H2A, H2B, H3, and H4. Each nucleosome structure is connected by linker DNA (20-80 bp) and forms repetitive units of ~200 bp (Olins and Olins, 2003). This nucleosome fiber is also known as 10-nm fiber (Figure 1.1).

To further investigate the detail of chromatin, Finch and Klug found that purified nucleosome fiber with linker histone H1 or Mg^{2+} ions were folded into fibers with a diameter 30-nm by electron microscopy (Finch and Klug, 1976). This fiber was named "30-nm chromatin fiber" (Figure 1.1). The 30-nm chromatin fiber has been studied extensively using various techniques, including biochemistry, biophysics, X-ray crystallography, conventional EM, cryo-EM and small angle X-ray scattering (Bordas et al., 1986; Bystricky et al., 2004; Dorigo et al., 2004; Finch and Klug, 1976; Gilbert et al., 2004; Hansen, 2002; Kruithof et al., 2009; Langmore and Paulson, 1983; Robinson et al., 2006; Robinson and Rhodes, 2006; Schalch et al., 2005; Widom and Klug, 1985; Woodcock et al., 1984). Several other models have also been proposed to describe the structure of higher order chromatin, and the "hierarchical helical folding model" suggests that a 30-nm chromatin fiber is folded progressively into larger fibers, including ~100-nm and then ~200-nm fibers, to form large interphase chromatin fibers or mitotic chromosomes (Belmont et al., 1989; Belmont and Bruce, 1994; Sedat and Manuelidis, 1978). In contrast, the "radial loop model" assumes that a 30-nm chromatin fiber folds into rapidly oriented loops to form mitotic chromosomes (Laemmli et al., 1978; Marsden and Laemmli, 1979; Paulson and Laemmli, 1977)

Therefore, 10-nm nucleosome fiber has long been assumed to form a 30-nm chromatin fiber and a further helically folded larger fiber. Nevertheless, several lines of evidence, including recent studies using cryo-EM and X-ray scattering analyses, detected no 30-nm structure, but detected 11- and 6-nm structure derived from nucleosome in interphase chromatin

and mitotic chromosomes (Dubochet et al., 1986; Eltsov et al., 2008; Joti et al., 2012; Nishino et al., 2012). Furthermore, ultra-small-angle X-ray scattering showed that there were no regular periodic structures between ~30- and 1,000-nm in interphase nuclei and mitotic chromosomes (Joti et al., 2012; Nishino et al., 2012). These recent findings suggest the common structural feature in interphase chromatin and mitotic chromosome: compact and irregular folding of nucleosome fibers occurs without a 30-nm chromatin structure (i.e., a polymer melt-like structure or chromatin domain) (Eltsov et al., 2008; Maeshima et al., 2010) (Figure 1.1).

Similar models of chromatin domain were also proposed in previous papers. The chromatin domains were originally identified as DNA replication foci by pulse labeling (Albiez et al., 2006; Berezney et al., 2005; Nakamura et al., 1986; Schermelleh et al., 2001), and they were shown to maintain stability during subsequent cell generations (Jackson and Pombo, 1998; Ma et al., 1998; Zink et al., 1999). Additionally, the chromosome territory-interchromatin compartment (CT-IC) model was proposed as the other model in which each chromosome territory was built from a series of interconnected, 1 Mb-sized chromatin domains (Cremer and Cremer, 2001; Cremer et al., 2000). Furthermore, recent high-throughput studies such as Hi-C and 5C have also proposed the physical packaging of genome DNA (Dekker, 2003, 2008; Lieberman-Aiden et al., 2009), which has been termed “topologically associating domains (TADs)” (Nora et al., 2012), “topological domains” (Dixon et al., 2012), or “physical domains” (Sexton et al., 2012). Some biological implications of TADs were proposed that they correspond to functional domains including DNA replication domain and lamina-associated domains (LADs) (Guelen et al., 2008). These findings suggest that condensed chromatin domains play a crucial role in compartmentalization of chromatin function and maintenance of genomic integrity.

Compared with the 30-nm chromatin fiber and further hierarchical regular structures proposed previously, irregular folding of nucleosome fibers implies a less physically constrained state that could be locally dynamic. Chromatin is prone to thermal agitation; therefore, it has a flexible and dynamic structure and is fluctuating in the nucleus. Photobleaching and photoactivation techniques enabled us to visualize turnover of nuclear proteins and measured dynamics of chromatin. Moreover, the single DNA site within chromatin labeled by LacO-array/LacI-GFP system showed large displacement of specific chromatin regions that encompass 20-50 nucleosomes in various cells and organisms (Belmont et al., 1999; Chubb et al., 2002; Hajjoul et al., 2013; Heun et al., 2001; Levi et al., 2005; Straight et al., 1996; Vazquez et al., 2001). Especially, by induction of DNA damage, an increase in chromatin

dynamics was visualized and observed (Mine-Hattab and Rothstein, 2012). What is the function of these extensive chromatin dynamics? There are three hypotheses: one is to reach the certain locus in the nucleus (e.g., transcriptional factory), the other is to facilitate the exposure of genomic DNA to the surface of chromatin domain, and the third is to assist the movement of proteins in the chromatin domain. However, the answer has not been clear yet.

The chromatin observed by Flemming was derived from the complex of DNA and proteins. Is the chromatin merely for packaging DNA into a tiny space? Do chromatin structure and dynamics have a cellular function? Recent studies show that chromatin structure is strongly related to not only DNA packing but also cell function involving transcription, replication, and DNA repair. In this thesis, the mechanism of regulating cellular functions by chromatin is shown.

1.4. From epigenesis to epigenetics

How does a single cell give rise to such a complicated body? Now, it is known that the cells of the eye and those of the skin have identical genome DNA in each individual, so their differences are assumed to be derived from the combination of expressed and unexpressed genes. But, how? How does a cell control gene expression in accurate time and place? How does a cell count the replication times from the fertilized egg? How does a cell recognize its position in an organism? In the case of a mistake in gene control, the organism would immediately go into a worse state. For a long time, all these questions have been unsolved and have deeply fascinated researchers.

In ancient Greece, Aristotle had observed many kinds of animals and described their morphology and developmental processes. He further explained that an organism is not fully formed in an egg; an embryo is gradually formed in a typical order (Peck, 1943). Due to these observations, Aristotle is recognized as the founder of "epigenesis" (Peck, 1943). Although now it is known that the germ cell contains no pre-formed embryo, the "preformation theory" had been assumed as the principle of the developmental process of the organism instead of epigenesis for about 2,000 years (Bowler, 1971).

Epigenesis and preformation are two competing theories that describe and explain the relevant embryonic development. Preformationism is a theory that organisms develop from the miniature versions of themselves, fully formed in a fertilized egg. Epigenesis is the eukaryotic developmental process through a sequence of steps when cells differentiate and organs form from a fertilized egg. Depending on the preformationist view, all organs of the adult were prefigured in a miniature within a gamete. In other words, this theory mainly discussed that the miniature of the organism is located in sperm or ovum.

However, Harvey and Wolff suspected the validity of the preformation theory and restored the epigenesis theory by the redescription of the developmental process of a chicken egg (Harvey, 1651; Wolff, 1759). Epigenesis gradually received support from researchers, and currently, it stands for all the processes that implement genetic instructions contained within a fertilized egg.

Epigenesis required the principle that could genetically explain its system through development. To account for epigenesis in biological development, the concept of "epigenetics", which connects epigenesis with genetics, was introduced by Waddington (Waddington, 1942). He coined "epigenetics" as "the branch of biology that studies the causal interactions between genes and their products which bring the phenotype into being" (Waddington, 1968).

Waddington broadly defined epigenetics as the mechanism that unfolds genomic information through development, and he did not use a specific definition for epigenetics. According to Holliday, what Waddington had in mind was "all those events which lead to the unfolding of the genetic program for development" (Holliday, 2006).

After Waddington's works, many researchers redefined the word "epigenetics"; therefore, there are many definitions of epigenetics. Nanney used the term epigenetics to distinguish between different types of cellular control systems, and he proposed that genetic components were responsible for maintaining and perpetuating a library of genes, both expressed and unexpressed, through a template replicating mechanism (Deans and Maggert, 2015; Nanney, 1958). Furthermore, Nanney insisted that the expression states could persist through cell division (Nanney, 1958). Nanney's contemplation of the stability of cellular expression states was an important addition to Waddington's ideas, which had significant impacts on the future direction of epigenetics (Deans and Maggert, 2015; Haig, 2004; Nanney, 1958).

However, for 30 years, the term "epigenetics" was little utilized in developmental biology, and during the 1980s and 1990s, the definition of epigenetics withdrew the developmental processes and became more generalized. After the discovery of genomic imprinting in mammals, it was apparent that this was due to genetic information coded on DNA that could be reversed at meiosis or during gametogenesis. Thus, it was clear that there was a new type of inheritance that was not based on changes in DNA sequence and emphasized the importance of genetic and non-genetic factors in controlling gene expression. Medawar and Medawar proposed the modern definition of epigenetics; "In the modern usage epigenesis stands for all the processes that go into the implementation of the genetic instructions contained within the fertilized egg" (Medawar and Medawar, 1983). Subsequently, Hall described the definition of epigenetics as "the sum of the genetic and non-genetic factors acting upon cells to control selectively the gene expression that produces increasing phenotypic complexity during development" (Hall, 2012). Holliday suggested another definition of epigenetics (Holliday, 1994):

(i) The study of the changes in gene expression which occur in organisms with differentiated cells, and the mitotic inheritance of given patterns of gene expression.

(ii) Nuclear inheritance which is not based on changes in DNA sequence.

This definition was closely related to the definition by Smith; he described the dual inheritance theory including the genetic and epigenetic system (Smith, 1990).

(i) The familiar system, depending on DNA sequence, used in transmitting information between sexual generations.

(ii) An epigenetic inheritance system (EIS), responsible for cellular inheritance during ontogeny —for example, fibroblast give rise to fibroblast, epithelial cells to epithelial cells, and Drosophila wing disc continue to be wing discs in serial transfer.

Therefore, from epigenesis, the concept of epigenetics was born, which is currently the comprehensive principle that explains morphogenesis, pattern formation, and foundation of gene regulatory system in eukaryotes. Epigenetics is considered as a bridge crossing between genotype and phenotype.

1.5. Epigenetics and chromatin

Currently, because it is difficult to capture the essence of epigenetics, I have been described the history of epigenetics. The current concept of epigenetics is mainly based on Holliday's definition, and such epigenetic regulation that accurately unfolds or folds genetic information has been assumed to be implemented by histone modifications and DNA methylation (Holliday, 1994). Furthermore, Allis has defined epigenetics as "the sum of the alterations to the chromatin template that collectively establish and propagate different patterns of gene expression (transcription) and silencing from the same genome" (Allis et al., 2006; Jenuwein and Allis, 2001; Strahl and Allis, 2000). Epigenetics can switch the readable and unreadable genomic region and increase the combination patterns executed from identical genomic DNA. The regulation of the decoding of the information on DNA is attributed to the accessibility of proteins to the target sites, and thus, the chromatin structure plays a critical role. Histone proteins can be subjected to various covalent modifications, e.g., acetylation of lysine, methylation of lysine, and phosphorylation of serine. These modifications are reversible and can cause the change in the chromatin environment by *cis*-effects and *trans*-effects (Allis et al., 2006; Jenuwein and Allis, 2001; Strahl and Allis, 2000). Then the histone code hypothesis was proposed; many of the combinations of histone modifications appeared to have a specific meaning for the cell and connect to the epigenetic regulation (Strahl and Allis, 2000). The *cis*-effects are brought about by the changes in the physical properties of the modified histone tails. Such a modulation in the tail structure or electric charge alters the nucleosome-nucleosome interactions (Allis et al., 2006). For instance, acetylation of lysine removes the positive charges on histone proteins and changes the chromatin structure because about half of the negative charges in the DNA are neutralized by positively charged lysine and arginine residues of histone proteins (Maeshima et al., 2014). Histone modifications may also elicit the recruitment of modification-binding proteins to the target chromatin as *trans*-effects. This can be viewed as "reading" a particular mark of a modified histone in a context-dependent fashion (Allis et al., 2006). For instance, H3K9me3 is recognized by chromodomain, which is located on HP1 protein, which makes the chromatin denser and represses the gene expression (Maison and Almouzni, 2004). Through both *cis*-regulation and *trans*-regulation, the histone modification has been considered to change the chromatin environment (structure and dynamics) and switch on/off the gene expression.

Understanding histone modifications, which determine how and when genes will be expressed, as well as other biological functions by altering the chromatin structures around

specific genes and the accessibility of proteins are big challenges. The form of chromatin structure can be simply classified into two distinct types: the open form of chromatin called "euchromatin" allows access of the transcriptional machinery to the binding site, whereas the closed form called "heterochromatin" does not allow such access. Thus, euchromatin is transcriptionally more active than heterochromatin. The combination of researches for epigenetics and chromatin structures helps to understand the details of a living eukaryotic organism.

The recent development of ChIP-Seq technology can help detect and quantify the genome-wide histone modification patterns and the relationships between them and gene expression (Barski et al., 2007; Johnson et al., 2007; Mikkelsen et al., 2007). Presently, much knowledge on histone modifications and their functions in the cell have been listed and are considered to be complicated. However, only the correlation of histone modification patterns and cellular function has been observed, and no direct process as to how the histone modification changes the chromatin structure and accessibility of protein to DNA have been observed. It is important to understand not only "what" but also "how" and start the distinct approaches separately.

Furthermore, the current researches have failed to include the part of inheritance as mentioned in Holliday's definition of epigenetics (Holliday, 1994). How many percentages of the accurate epigenetic marks are inherited to daughter cells through mitotic division? Is it sufficient for the accuracy of information in the cell? Is it reconstructed based on chromatin structure after mitosis? Do epigenetic marks act as cell memory and time arrow? Epigenetics was first proposed for development, so it is required to reconsider the relationships between time and epigenetics.

Chromatin structure and dynamics are connected to epigenetic regulation in both aspects of *cis*- and *trans*-regulation. Hence, the detail observation of chromatin is essential for understanding the various cellular functions in eukaryotes. Additionally, how do the chromatin structure and dynamics contribute to the protein behavior? How does each histone modification change the chromatin environment? For such kind of research, it is required to develop a new method to investigate the direct relationship between chromatin and epigenetics and uncover the black box of chromatin function.

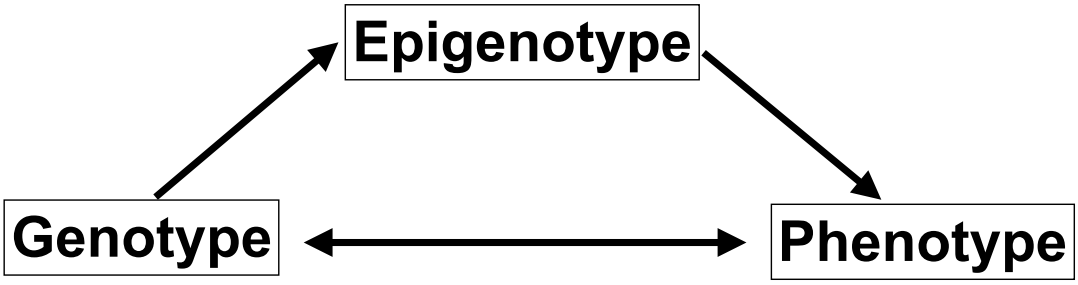

Figure 1.3 From genotype to phenotype

In eukaryotes, epigenotype has the role of the bridge crossing from genotype to phenotype. Since chromatin is the main component of epigenetics, the research of epigenetics and chromatin is related to many biological questions.

1.6. Objective

As mentioned earlier, chromatin consists not only DNA but also a complex of nucleosomes and their binding proteins; moreover, it has important roles in epigenetics. The objective of this thesis is to reveal the relationships between chromatin and epigenetics. For this, multiple methods including genome-wide analysis and live cell imaging were considered. Then, the aim of the study was to reveal the relationships between epigenetic regulation and promoter types by bioinformatics approach and examine the chromatin structure and dynamics at nano-scale by the newly developed imaging method.

Previously, the genome-wide detection of histone modifications superimposed on chromatin was difficult. However, the recent progress in high-throughput technologies has made it possible to collect a variety of "omics" data on the epigenetic behaviors of the histone modifications and transcripts. The ChIP-Seq method has been developed for analyzing genome-wide protein biding pattern including specific histone modifications. Currently, many sets of data on histone modifications and gene expression patterns are available. The ChIP-Seq data analysis was connected to the promoter analysis to understand the epigenetic regulation in eukaryotes (Chapter 2).

The range of observable world is determined by the power of microscopy. Due to the light diffraction limit, it was impossible to observe objects in the range of 10-200 nm even though the size of chromatin domain is in such range. Although this resolution of microscopy impedes the progress of chromatin research, the newly developed single molecule imaging and super-resolution imaging make it possible to smash the light diffraction limit and observe the details of chromatin. Chromatin environment: chromatin structure, dynamics and histone modifications are crucial components that are involved in the epigenetic regulation, and thus, the study of the chromatin environment will provide new information towards understanding the various cellular systems in eukaryotes. In this thesis, to observe the chromatin below the diffraction limit in a living cell, the single nucleosome imaging method was first developed that helped to observe chromatin at a single molecule level (Chapter 3). We further applied this single nucleosome imaging to super-resolution imaging and chromatin dynamics analysis of the overall nucleus. The super-resolution imaging PALM (photoactivated localization microscopy) revealed the formation of numerous chromatin domains in a living human nucleus and mitotic chromosomes and demonstrated the formation system of the chromatin domains. Furthermore, based on single nucleosome imaging, the chromatin dynamics was successfully measured at the single nucleosome level, and the heat map of dynamics of chromatin domain was also created.

This analysis showed the difference in chromatin dynamics at the local region of the nucleus and revealed the control system of chromatin dynamics. Finally, after the histone modification patterns were changed or the chromatin-associated proteins were depleted, the changes in chromatin environment were observed (Chapter 4). It is believed that the results of this thesis will contribute to the investigation of epigenetics.

Chapter 2

Tight Associations between Transcription Promoter Type and Epigenetic Variation in Histone Positioning and Modification

2.1. Introduction

Recent progress in high-throughput technologies has made it possible to collect a variety of "omics" data on transcripts and on the epigenetic behaviors of the histones that are often associated with these transcripts (Barski et al., 2007; Carninci et al., 2005; Schones et al., 2008; Suzuki et al., 2009; Wang et al., 2008).

Cap analysis of gene expression (CAGE) is a high-throughput method that enables large-scale identification of transcription start sites (TSSs) of eukaryotic species. This method measures gene expression levels simultaneously with TSS identification by counting the sequenced 5' ends of full-length cDNAs, termed CAGE tags (Carninci et al., 2005; Shiraki et al., 2003). With the development of deep sequencing methods, more high-throughput, and high resolution "tag depth" measurements have become available (DeepCAGE, nanoCAGE and CAGEscan) (Plessy et al., 2010; Suzuki et al., 2009). Such recent whole-cell-level pictures of quantitative transcriptomes have revealed the complex transcriptional network of mammalian species (Carninci et al., 2005; Shiraki et al., 2003; Suzuki et al., 2009). According to recent CAGE-based analyses of human TSSs, the human "promotome" can be classified into two types of promoters by the degree of imprecision of their transcription initiation sites (Carninci et al., 2006). One is the peak promoter, which initiates transcription strictly from a narrow genomic region (within a distance of 1-4 bp), and the other is the broad promoter, which initiates transcription from wide-ranging positions (>4 bp) (Carninci et al., 2006; Kratz et al., 2010). The

peak promoters are suggested to be closely associated with the presence of the TATA box (which enables proper control of gene expression by binding with transcription factors) and with tissue-specific gene expression. The broad promoters have been observed in the presence of CpG islands and drive relatively ubiquitous expression of the genes they control (Carninci et al., 2006; Frith et al., 2008; Kawaji et al., 2006; Ponjavic et al., 2006). The CpG-rich broad promoters are considered evolutionarily new and more likely to be controlled by epigenetic mechanisms, including DNA methylation and sense–antisense regulation, than the peak promoters (Carninci et al., 2006; Coleman and Pugh, 1995; Kawaji et al., 2006). These differences between broad and peak promoters raise questions of how these promoter types are associated with chromosomal structures and modifications and of how their difference confers cellular function.

In eukaryotic species, chromosomal DNA is packed into nucleosomes, each of which comprises approximately 147 bp wrapped around a histone protein octamer consisting of two copies of each of the four core histones, H2A, H2B, H3, and H4 (Davey and Richmond, 2002; Luger et al., 1997). Two biologically important aspects of these histones are their modifications and positions, and it has been shown that these factors regulate transcription initiation (Kornberg and Lorch, 1999; Lister et al., 2009; Wyrick et al., 1999). Several methodologies have rapidly been developed for high-throughput identification of histone positions and modifications. ChIP-chip identifies the histone-binding positions of genomic DNA by using a combination of chromatin immunoprecipitation and tiling array (Yuan et al., 2005). Although ChIP-chip used to be a widely-used method, today, with the growing demand to develop high-throughput sequencing, the ChIP-Seq method has been developed as a promising alternative to the tiling array-based approach in analyzing genome-wide nucleosome positioning (Jiang and Pugh, 2009; Park, 2009). These methodologies have revealed several insights into the intertwining of gene expression with nucleosome position and histone modification. For example, the degree of eviction of nucleosomes from the upstream regions of TSSs is correlated with gene expression patterns in yeasts (Mavrich et al., 2008; Yuan et al., 2005) and humans (Ozsolak et al., 2007; Schones et al., 2008; Smith et al., 2009). Moreover, the methylated histone H3 at lysine 4 (H3K4me1, -2, and -3) and acetylated histone H3 at lysine 9 (H3K9ac), located around TSSs, are linked to gene activation (Barski et al., 2007; Bernstein et al., 2005; Karlic et al., 2010; Roh et al., 2006; Vakoc et al., 2005), whereas H3K27me3 and H3K9me3 are linked to gene repression (Bannister and Miska, 2000; Barski et al., 2007; Roopra et al., 2004; Vakoc et al., 2005). These modifications and related gene regulatory behaviors support the

"histone code" hypothesis (Strahl and Allis, 2000), i.e. that multiple histone modifications specify unique downstream functions. However, the detailed mechanisms underlying transcriptional regulation by these histone behaviors are still obscure.

H3K9ac has recently been frequently observed around broad promoters (Kratz et al., 2010). This implies that histone behavior is associated with promoter architecture, although this association has so far been found only in the case of H3K9ac, and the extent of such associations is unclear. In this study, we systematically analyzed the relationships between histone behaviors and promoter architecture types by using information about (1) modified/unmodified histones; (2) their genomic positions relative to TSSs; (3) their positional stabilities on the genome under two cellular conditions; and (4) gene expression. The results showed that promoter architecture type and gene expression are tightly associated with the modification pattern and genomic positional stability of the histones forming nucleosomes. They provide new insights into the epigenetic mechanisms of transcriptional regulation in terms of histone behavior.

2.2. Materials and methods

2.2.1. Nucleosome position detection and dataset

Nucleosome-resolution (MNase digestion) ChIP-Seq Solexa tags for histone H3 were obtained by (Schones et al., 2008). The genomic positions of the methylated histones and the histone variant H2A.Z were obtained from (Barski et al., 2007), and those of acetylated histones were from the study by (Wang et al., 2008). All of these data were obtained in human resting CD4+ T cells. To determine the genomic positions of nucleosomes according to the ChIP-Seq data, we used the software published in (Zhang et al., 2008). Human genome hg18 was used.

2.2.2. Transcription start site detection and dataset

TSSs were detected by DeepCAGE data obtained by the FANTOM 4 project (Suzuki et al., 2009); 10,971 TSSs of broad promoters and 3,621 TSSs of peak promoters were detected by applying the methods used in FANTOM 3 (Carninci et al., 2006; Kratz et al., 2010) to the FANTOM 4 dataset (Kratz et al., 2010). We used only those promoters for which the corresponding probes were clustered on the genome (level 3 promoters; (Kratz et al., 2010; Suzuki et al., 2009)), and for each promoter the neighboring position that had the highest density of overlapping CAGE tags was determined as the position of the TSS. Promoters containing TATA boxes within 50 bp upstream of TSSs were determined by using position-specific weight matrices from JASPAR4 (with a confidence score of more than 75%) (Sandelin et al., 2004), and promoters containing CpG islands within 200 bp upstream of TSSs were obtained from the UCSC Genome Browser database (http://genome.ucsc.edu/). Alternative dataset of CpG islands were obtained from (Bock et al., 2007).

2.2.3. Distribution of nucleosome positions around TSSs

As described above, the genomic positions of nucleosomes as well as TSSs for both broad and peak promoters were determined. The distributions of nucleosomes within the genomic regions from –5 kb to 5 kb with respect to TSSs were calculated by dividing the number of nucleosomes at each position by the number of TSSs. Genomic positions from –15 bp to 15 bp with respect to the central positions of the nucleosomes were assumed as the genomic positions where nucleosomes existed. The distributions of nucleosomes near broad and peak promoters were calculated separately.

2.2.4. Distribution of Sp1 binding sites and other transcription factor binding sites

Sp1, MAZ, and PU.1 binding sites were obtained from FANTOM 4 (http://fantom.gsc.riken.jp/4/download/GenomeBrowser/hg18/TFBS_CAGE/allsites_cage_tfbs_feb09_latest.gff.gz) (Suzuki et al., 2009). The distributions of these transcription factor binding sites around TSSs (from –500 bp to 500 bp) were calculated by dividing the number of these sites for each position by the number of TSSs used for the analysis.

2.2.5. Stability of nucleosome positions under different cellular conditions

We compared the nucleosome positions obtained in human resting CD4+ T cells with those obtained in human activated CD4+ T cells. We calculated the minimum distance between each nucleosome in resting T cells and the closest nucleosome in activated cells. This distance was considered to denote how far each nucleosome moved along the genome in response to the change in cellular condition (from resting to activated). The distributions of these distances were calculated by dividing the number of nucleosomes that moved specified distances (from X bp to X + 15 bp) by the total number of nucleosomes. The average absolute minimum distance between each nucleosome in resting T cells and the closest nucleosome in activated cells was also calculated.

2.2.6. Relative abundances of peak promoters and broad promoters

The abundance of peak promoters relative to that of broad promoters at position j was calculated by $(B_j - P_j) / \sum_i B_i$, where B_j and P_j denote the proportions of nucleosomes at position j for broad and peak promoters, and $\sum_i B_i$ denotes the sum of proportions of nucleosomes around the TSS (from –2,000 bp to 2,000 bp), respectively.

2.2.7. Gene expression in human resting CD4+ T cells

The gene expression profile in human resting CD4+ T cells was obtained from the Gene Expression Omnibus (GSE10437) (Schones et al., 2008). We used genes (total number 8,007, with 7,591 associated with broad promoters and 416 associated with peak promoters) annotated with Entrez gene IDs in FANTOM 4 and with expression present in Present/Absent calls generated by the Affymetrix microarray platform. Nineteen types of methylated histones and 18 types of acetylated histones obtained in CD4+ T cells were used (Barski et al., 2007; Wang et al., 2008). Acetylated histones located around TSSs are linked only to gene activation. To investigate the upregulation of genes associated with histone acetylation and their dependence

on promoter type, we made two groups of histones: one having modified histones (18 types of acetylated histone) around TSSs (from –500 bp to 500 bp) and the other having no modified histones. In contrast to acetylated histones, methylated histones located around TSSs are linked to both gene activation and repression. Furthermore, the functions of many methylated histones are still unknown. Therefore, for histone methylation, we made the following two groups, one having H3K4me1, -2 or -3, which are known to upregulate downstream genes, and the other having no modified histones. Distributions of gene expression levels were represented as box plots. *P* values for evaluating the significance of gene expression changes were calculated by the Wilcoxon rank sum test.

To compare the distributions of nucleosomes that had H3K4me3, were located upstream of TSSs (positions from –150 to –100 bp), and were associated with either broad promoters or peak promoters in cases where the downstream genes showed similar expression levels, we selected 1,788 genes associated with broad promoters and 138 associated with peak promoters that had expression levels in the range of 250 to 750 (Figure 2.9). The chi-squared test was applied to assess the difference in nucleosome distribution between these two types of promoter.

2.3. Results

2.3.1. Promoter architecture and nucleosome positioning

We first focused on differences in nucleosome distribution around the two different types of transcription promoters (i.e. peak and broad promoters). We used human promoter positions for which information about the degree of transcription start imprecision had been obtained in a previous study (Kratz et al., 2010), as well as nucleosome positions defined as the genomic positions of histone H3 proteins in the resting condition in human CD4+ T cells (Schones et al., 2008). We mapped them on human genomic sequences. (See Materials and methods for details of data manipulations.) We then calculated the ratio of nucleosomes located at each genomic position relative to each peak and broad promoter. We found that the nucleosome positions associated with broad promoters had markedly aligned and periodic patterns compared with those of peak promoters (Figure 2.1 A). More strikingly, only in broad promoters, the first nucleosomes immediately downstream of the promoter were likely to be located in similar positions and those immediately upstream of the promoter were depleted (see the magnified view in Figure 2.1 A). This was contrary to our expectation; previous studies have reported that, in general, nucleosomes are distributed evenly around the promoter region (Ramirez-Carrozzi et al., 2009; Tolstorukov et al., 2009). We had therefore expected that the nucleosome positions would be spread around the broad promoter and well aligned around the peak promoter, because TSSs are widely spread in the broad promoter region but narrowly spread in the peak promoter region. However, our results show that the broad promoter was specifically associated with a more aligned pattern of nucleosomes than the peak promoter.

H2A.Z is a histone variant of H2A that is highly conserved among lower and higher eukaryotes. Enrichment of H2A.Z around the promoter region has been also reported in yeast (Albert et al., 2007) and humans (Jin et al., 2009). In terms of promoter architecture, we performed a similar analysis to the one of H3 shown in Figure 2.1 A of the positions of human nucleosomes harboring the histone variant H2A.Z in human resting CD4+ T cells (Barski et al., 2007). H2A.Z was highly enriched around broad promoters but not peak promoters (Figure 2.1 B). For example, the statistical significance of the enrichment was $P < 1.0 \times 10^{-25}$(chi-squared test) for positions +100 to +130 with respect to the TSS. Moreover, the distribution patterns of H2A.Z were similar to those of H3; the positions of H2A.Z were markedly aligned around broad promoters but not around peak promoters.

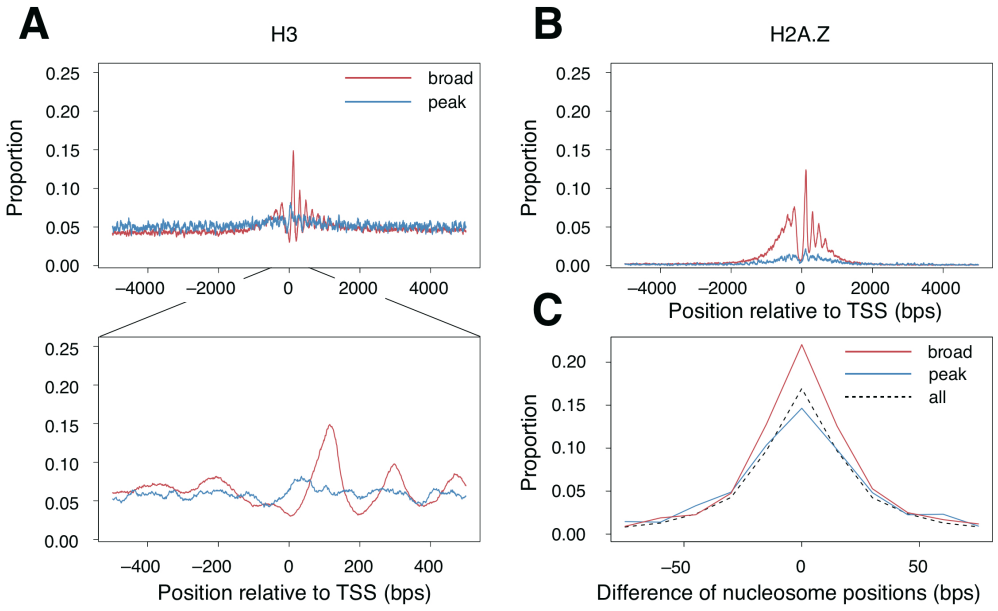

Figure 2.1 Distributions of nucleosome positions around transcription start sites (TSSs)

(A) Distributions of the central positions of histone H3 around broad and peak promoters. The *x*-axis shows genomic positions with respect to TSSs (from –5 kb to 5 kb, upper panel; and from –500 bp to 500 bp, lower panel). The central positions of nucleosomes are defined as the positions from –15 bp to 15 bp with respect to the center of the nucleosome. (B) Distributions of nucleosomes containing the histone variant H2A.Z around TSSs (from –5 kb to 5 kb). H2A.Z around TSSs associated with broad promoters are highly enriched, unlike those associated with peak promoters. (C) Distributions of minimum distances from each of the nucleosomes in human resting T cells compared with those in activated T cells. The *x*-axis shows the minimum distances and the *y*-axis shows the proportions of nucleosomes with the specified minimum distances. Proportions within every 15 bp were averaged. Minimum distances were calculated for all nucleosomes on the genome (dashed line), for those associated with broad promoters (red line), and for those associated with peak promoters (blue line).

2.3.2. Accessibility of transcription factor Sp1

The two promoter architectures are associated with characteristic sequence contexts: the peak promoter is located close to a TATA box and the broad promoter close to CpG islands (Carninci et al., 2006). Using the genomic positions of putative TATA box sites predicted by a position-specific weight matrix and the positions of CpG islands obtained from the UCSC Genome Browser database (Bucher, 1990; Karolchik et al., 2003), we confirmed that TATA boxes were overrepresented in peak promoters and that broad promoters were highly associated with the presence of CpG islands (Figure 2.2).

It is possible that the aligned patterns of nucleosome positions around broad promoters are due to the accessibility of transcription factors to DNA. For instance, in the absence of the TATA box, the ubiquitous transcription factor Sp1 can recruit TATA-binding proteins to initiate transcription (Butler and Kadonaga, 2002). It has already been reported that consensus Sp1 sites with high overall GC contents are overrepresented among broad promoters, and the positions of these sites for individual transcription units are less precise than those of TATA boxes (Carninci et al., 2006). Consequently, we investigated the possibility that the nucleosomes around a broad promoter align in a more orderly fashion than those around the peak promoter because of the need to create a nucleosome-free region upstream of the TSS to confer DNA accessibility of transcription factor proteins. We superimposed the distribution of putative Sp1 sites (Suzuki et al., 2009) around broad promoters onto that of the nucleosome positions (see Materials and methods), and we observed increased proportions of Sp1 sites about –50 bp upstream of the broad promoter, where the nucleosome distribution was markedly depleted (Figure 2.3). We conducted the same analysis for peak promoters. The inverse relationship between Sp1 site and nucleosome abundance around the broad promoter was much higher than that around the peak promoter, suggesting the plausibility of the DNA accessibility model. Furthermore, we conducted a similar analysis for the binding sites of two other transcription factors, PU.1 and MAZ, as a previous study (FANTOM 4) had analyzed the binding sites of these two factors in detail (Suzuki et al., 2009). The binding sites of both PU.1 and MAZ were distributed on nucleosome-free regions around broad promoters, whereas no such trends were observed around peak promoters (Figure 2.4). These results support the strong connection between the nucleosome-free region and the accessibility of transcription factors, which was specific to broad promoters.

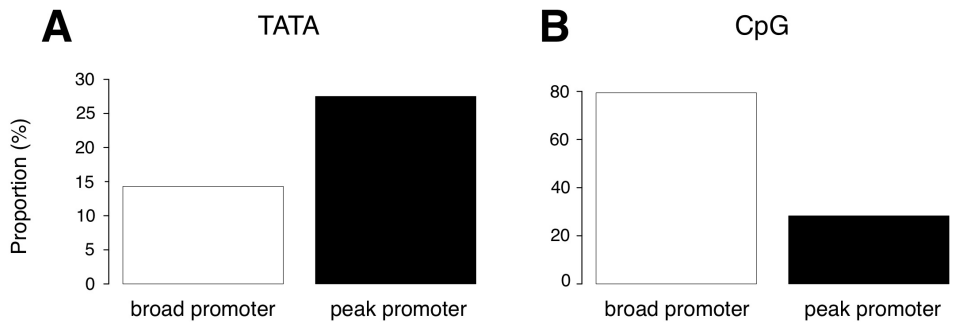

Figure 2.2 Frequencies of occurrence of TATA boxes and CpG islands around transcription start sites (TSSs)

Frequencies of the characteristic sequence patterns associated with promoters are shown by bar charts. The *y*-axis shows the proportions of broad and peak promoters that have TATA boxes (A) and CpG islands (B).

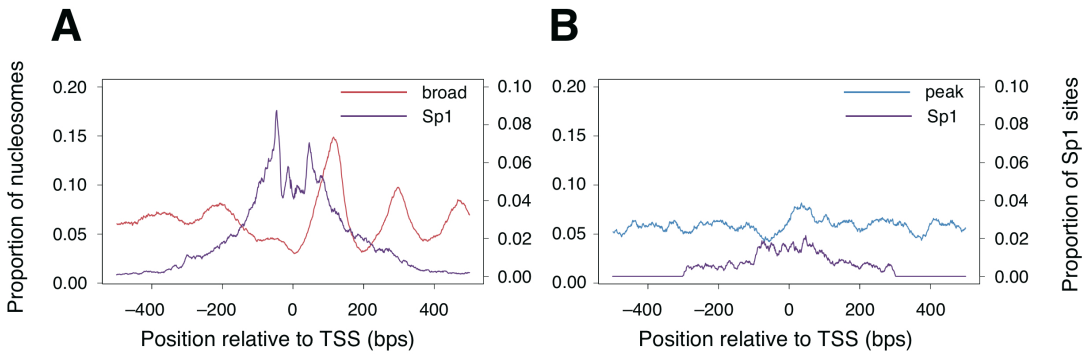

Figure 2.3 Distributions of Sp1 sites around transcription start sites (TSSs)

Distributions of nucleosome regions and Sp1 sites around TSSs associated with broad (A) and peak (B) promoters are shown. Nucleosome position is defined as the center position of the nucleosome. The *x*-axis shows genomic positions with respect to TSSs (from –500 bp to 500 bp). Sp1 sites were obtained by the FANTOM 4 project.

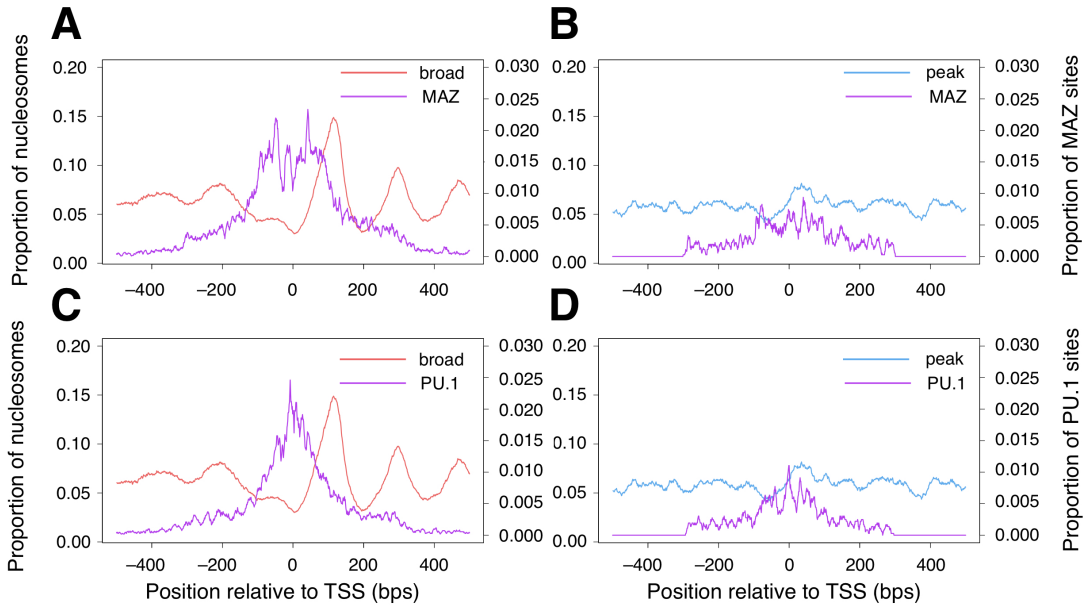

Figure 2.4 Distributions of the binding sites of two transcription factors (MAZ and PU.1) around transcription start sites (TSSs)

Distributions of nucleosome regions and transcription factor binding sites around TSSs associated with broad (A: MAZ, C: PU.1) and peak (B: MAZ, D: PU.1) promoters are shown. Nucleosome position is defined as the center position of the nucleosome. The *x*-axis shows genomic positions with respect to TSSs (from –500 bp to 500 bp). Both MAZ and PU.1 sites were obtained by the FANTOM 4 project.

2.3.3. Positional stability of nucleosomes around broad promoters

If nucleosome positioning around broad promoters confers DNA accessibility for the binding of transcription factors, then the nucleosome positions around broad promoters should be more stable throughout different cellular conditions than those around peak promoters, because broad promoters are usually associated with ubiquitously expressed gene (in contrast, peak promoters are associated with tissue- and condition-specific expressed gene) (Carninci et al., 2006; Frith et al., 2008; Kawaji et al., 2006; Ponjavic et al., 2006) and the genomic positions of transcription factor binding sites are fixed. We analyzed the positional stability of nucleosomes located within positions +1 to +200 with respect to each promoter under "resting" and "activated" conditions of human CD4+ T cells (Schones et al., 2008) (see Materials and methods). For each nucleosome position in the resting condition, we calculated the distance to the nearest nucleosome position in the activated condition in order to assess the positional stabilities of single nucleosomes under the two different cellular conditions. The overall minimum distance was markedly shorter for broad promoter-associated nucleosomes than for peak promoter-associated ones (Figure 2.1 C). In fact, the average absolute minimum distance in the case of the broad promoter (20.70 bp) was significantly shorter than that for the peak promoter (25.08 bp) ($P = 4.83 \times 10^{-7}$; *t*-test. Note that we did not take into account nucleosomes for which a minimum distance longer than 100 bp was found between the two conditions, because these were more likely to be different or neighboring nucleosomes rather than those that moved along the DNA with the change in conditions.) These results demonstrated that the positions of nucleosomes around the broad promoters were more stable than those of nucleosomes around the peak promoters.

2.3.4. Distribution of nucleosomes containing modified histones

It has been suggested that not only nucleosome position, but also nucleosomal histone modification, can regulate transcription (Barski et al., 2007; Bernstein et al., 2005; Roh et al., 2006; Vakoc et al., 2005). For instance, histone methylation is associated with either gene activation or repression, depending on the methylation site and state on the histone protein; in particular, methylation of histone H3 (H3K4me1, -2, and -3) in nucleosomes around the transcription promoter are well known to regulate gene expression (Barski et al., 2007; Bernstein et al., 2005; Karlic et al., 2010; Roh et al., 2006; Vakoc et al., 2005). To investigate the differences in positional distribution of nucleosomes containing methylated histones around the two different types of promoter, we obtained nucleosome positions corresponding to each of three methylation types (H3K4me1, -2, and -3) in human CD4+ T cells from a previous study (Barski et al., 2007), and we mapped these onto genomic sequences with the broad and peak promoter positions. Similar to the result for histone H3, nucleosomes having H3K4me1, -2, and -3 were all highly enriched and well aligned around broad promoters, whereas they were depleted around peak promoter regions (Figure 2.5 A-C). However, the alignment pattern of nucleosome positions differed depending on the type of methylation. Within the region downstream of the broad promoter, the first frequency peak of nucleosomes having H3K4me1 and -2 occurred in the +700 to +730 region (Figure 2.5 A and B), whereas those having H3K4me3 occurred in the +100 to +130 region (Figure 2.5 C; this was similar to the result for histone H3, perhaps because the majority of H3K4 were trimethylated.) For each methylation type, the difference in frequency of occurrence of nucleosomes with each type of modified histone in these regions between the peak and broad promoters was significant ($P < 1.0 \times 10^{-10}$ for H3K4me1 and -2, and $P < 1.0 \times 10^{-50}$ for H3K4me3; chi-squared test). Note that the values on the *y*-axes in Figure 2.5 are not influenced by the absolute numbers of nucleosomes in each type of promoter, as they indicate the proportion of nucleosome-harboring TSSs for each type of TSS. In addition to methylation, acetylation may control gene expression (Barski et al., 2007; Bernstein et al., 2005; Karlic et al., 2010; Roh et al., 2006; Vakoc et al., 2005). We further analyzed nucleosome positioning corresponding to histone acetylation (H3K9ac) in human CD4+ T cells and observed results similar to those for H3K4me3 ($P < 1.0 \times 10^{-50}$ for +100 to +130 region; chi-squared test; Figure 2.5 D). For each of H3, H2A.Z, H3K4me3, and H3K9ac, we estimated the abundance of nucleosomes associated with peak promoters relative to that of nucleosomes associated with broad promoters (Figure 2.6; see Materials and methods). Compared with nucleosomes carrying histone H3, the relative abundances of nucleosomes

carrying the modified histones or the histone variant were large, suggesting that the presence of histone modifications or a histone variant was highly associated with the broad promoter but not the peak promoter.

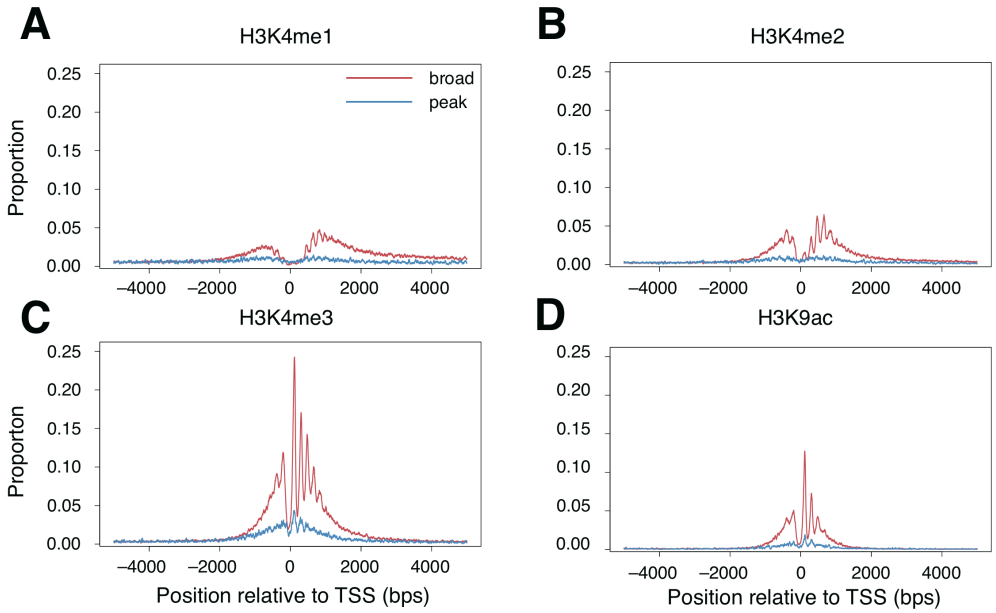

Figure 2.5 Distribution of modified histones around transcription start sites (TSSs)

Distributions of nucleosomes containing methylated and acetylated histones. (A) H3K4me1, (B) H3K4me2, (C) H3K4me3, and (D) H3K9ac around TSSs are shown. All of the modified histones were highly enriched around the TSSs associated with broad promoters, unlike those associated with peak promoters. The *x*-axis shows the genomic positions with respect to the TSSs (from –5 kb to 5 kb).

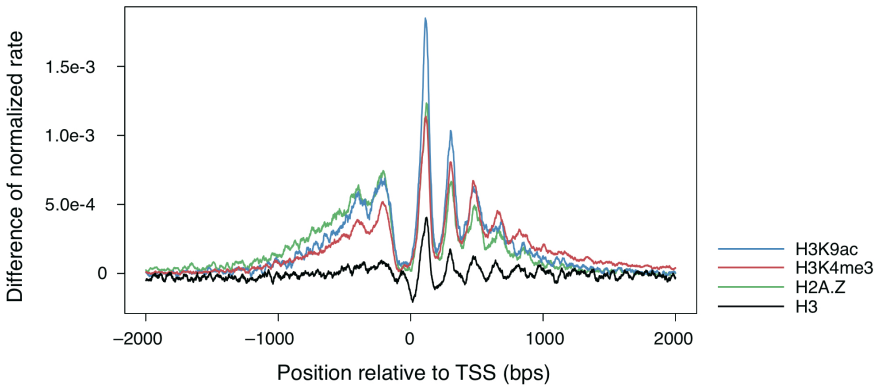

Figure 2.6 Relative abundance in histone distributions

Normalized differences in histone distributions (H3, H3K4me3, H3K9ac, and H2A.Z) between broad and peak promoters (from –2 kb to 2 kb) at each position are shown. The *y*-axis shows the normalized differences in histone distributions between broad and peak promoters. H3K4me3, H3K9ac, and H2A.Z had larger differences than H3.

2.3.5. Analysis of another genomic element that potentially influences histone behavior

Methylation of CpG islands is tightly associated with the expression of downstream genes; a number of studies have therefore been conducted to analyze CpG islands at a genome-wide level (Ioshikhes and Zhang, 2000; Wang and Hannenhalli, 2006). As described above, broad promoters are strongly associated with CpG islands (Figure 2.2). Therefore, it is possible that the enrichment of histone modifications and histone variants in the broad promoter region is derived merely from the effect of CpG islands and is independent of promoter architecture. In fact, it has been shown that promoters with many CpG islands are more likely to harbor modified histones than promoters with fewer CpG islands (Bhandare et al., 2010). To address this issue, we analyzed the positions of nucleosomes having histone H3 and those having H3K4me3 around broad and peak promoters with and without CpG islands (Figure 2.7). We found that, in the case where promoters were associated with CpG islands, nucleosomes with histone H3K4me3 were likely to be well aligned even around peak promoters. However, broad promoter-associated nucleosomes were significantly more enriched than peak promoter-associated nucleosomes, especially in the region downstream of the promoter (Figure 2.7 A; $P < 1.0 \times 10^{-16}$ for +100 to +130 region; chi-squared test). (Note, however, that the set of "peak promoters" used in this study may have included "broad promoters", and that this may have affected the highly aligned nature of H3K4me3 around "peak promoters". This was because the definition of promoter architecture thus far was whether there was a cluster of TSSs located within a narrow genomic region or whether the TSSs were dispersed, and low TSS coverage increased the possibility of promoters being classified as "peak promoters".)

In contrast, when we focused only on promoters without CpG islands, nucleosomes having H3K4me3 were well aligned and enriched only around broad promoters (Figure 2.7 B); the difference in the frequencies of downstream nucleosomes (from +100 to 130) potentially resulting from the difference in the alignment were significant ($P < 1.0 \times 10^{-56}$, chi-squared test). Broad promoters with CpG islands had an aligned pattern of nucleosomes carrying H3, whereas no clear alignment was observed for peak promoters (Figure 2.7 C). Broad promoters without CpG islands still showed an aligned pattern of nucleosomes having H3 (although the pattern was less clear than in those with CpG islands), whereas peak promoters had little alignment in the pattern (Figure 2.7 D). These results show that the enrichment of nucleosomes having certain histones around a broad promoter is independent of the existence of CpG islands.

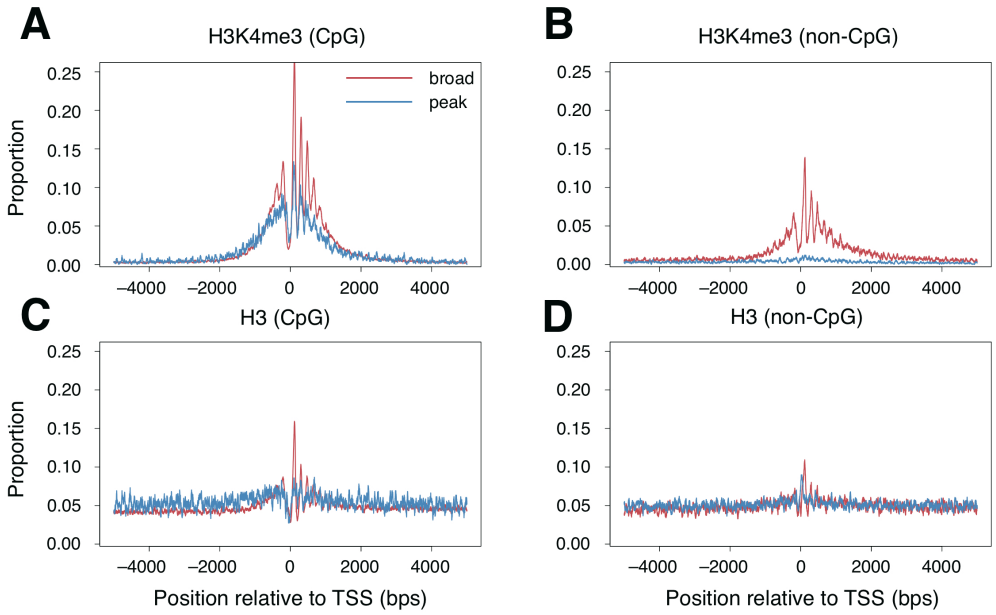

Figure 2.7 Distributions of nucleosomes around transcription start sites (TSSs) with and without CpG islands

Distributions of nucleosomes containing H3K4me3 (A, B) and H3 (C, D) around broad and peak promoters are shown. The analyses were conducted separately for TSSs that were associated with CpG islands (A, C) and those that were not (B, D). Broad promoters had aligned patterns of nucleosomes containing H3 and H3K4me3, regardless of the existence of CpG islands, and were enriched in H3K4me3. In contrast, peak promoters had little alignment of the H3 pattern, regardless of the presence of CpG islands. The proportion of nucleosomes containing H3K4me3 associated with peak promoters was lower than that associated with broad promoters, particularly in the absence of CpG islands.

2.3.6. Effect of histone modification on gene expression

To explore whether histone modification around the promoter affects gene expression, we analyzed the difference in expression levels of RNAs transcribed from peak and broad promoters in terms of the existence of modified/unmodified histones in their surrounding regions. We compared data sets of methylated/unmethylated histones and acetylated/unacetylated histones measured under resting conditions in human CD4+ T cells (Barski et al., 2007; Wang et al., 2008). Gene expression data for resting CD4+ T cells were obtained from a previous study (Schones et al., 2008); we used only those genes for which the expression levels had been measured. We classified promoters having at least one methylated/acetylated histone within the region from –500 to +500 as "promoters with methylated/acetylated histones" and all others as ones with unmethylated/unacetylated histones (see Materials and methods). Expression levels of genes associated with broad promoters that had methylated histones were significantly higher than those of genes associated with broad promoters with unmethylated histones ($P < 9.1 \times 10^{-11}$, U-test; Figure 2.8). Conversely, the expression levels of genes associated with peak promoters having only unmethylated histones were as high as those of genes associated with peak promoters with methylated histones, and thus no significant difference was observed ($P = 0.97$, U-test; Figure 2.8). Likewise, in the comparison between acetylated and unacetylated histones, the expression levels of genes associated with broad promoters that had acetylated histones were significantly higher than those of genes associated with broad promoters with no acetylated histones ($P < 2.2 \times 10^{-16}$, U-test; Figure 2.8), but the expression levels of genes associated with peak promoters that had acetylated histones did not differ markedly from those of genes associated with peak promoters with only unacetylated histones ($P = 0.69$, U-test; Figure 2.8). These results suggest that the regulation of gene expression levels by histone modification is specific to broad promoter-associated genes.

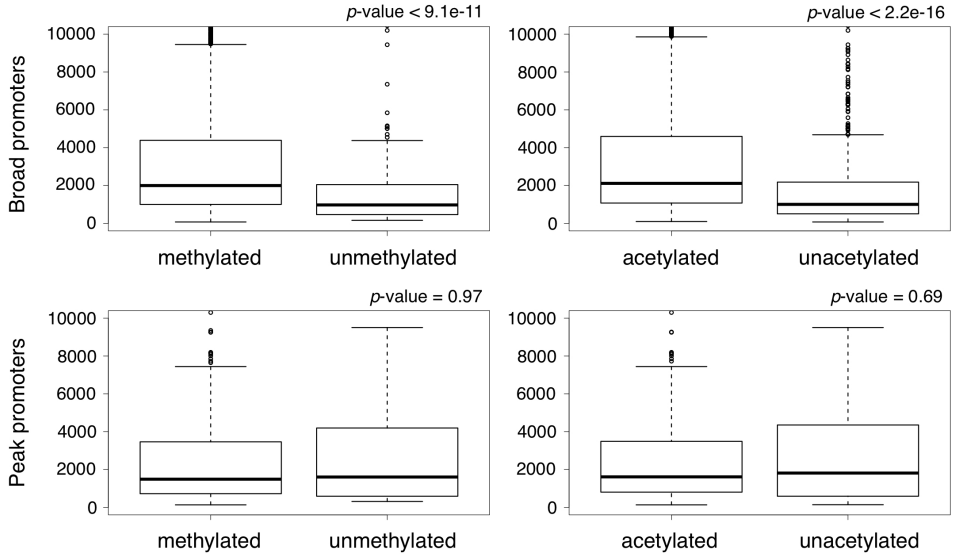

Figure 2.8 Box plots of gene expression in human resting CD4+ T cells

The box plots represent the distributions of gene expression levels. Distributions of the four groups of genes are drawn separately, i.e. those with broad or peak promoters, each of which was further associated with modified histones in activated cells or with unmodified histones. The *y*-axis shows the microarray intensities of the gene sets in each category.

2.4. Discussion

We analyzed the global landscape of epigenetic relationships between histone modifications and transcription initiation by investigating genome-wide ChIP-Seq data and DeepCAGE data. The results presented here show differences in the architecture of the broad and peak promoters that regulate gene expression. Especially, we revealed that the broad promoters were strongly associated with histones immediately downstream of the TSS and they were frequently modified, presumably to regulate gene expression levels.

In previous studies, aligned patterns of nucleosome positions around TSSs have been identified in yeasts and humans (Mavrich et al., 2008; Ramirez-Carrozzi et al., 2009; Segal et al., 2006; Tolstorukov et al., 2009). However, we confirmed this alignment only for regions around TSSs derived from broad promoters, not for those around TSSs derived from peak promoters. Broad promoters have an aligned pattern of nucleosome positions around TSSs and have large nucleosome-free regions immediately upstream of TSSs. Studies in yeasts have validated the model of "open promoters", which have large, nucleosome-free regions immediately upstream of the TSS and are often associated with TATA-less promoters and poly (dA:dT)-rich tracts, the sequences of which are unbendable and unstable for histone binding (Cairns, 2009). The broad promoter characteristics that we found in humans are consistent with this model, because in humans the sequence patterns in CpG islands located upstream of TSSs, in contrast to the yeast poly (dA:dT)-rich tracts, have been shown to be unstable (Ramirez-Carrozzi et al., 2009).

Our data indicate that the nucleosomes that are immediately downstream of TSSs and associated with broad promoters are positioned in specific regions. We suggest that broad promoters have these aligned patterns of nucleosome positions around TSSs because the nucleosome position has a stronger impact on broad promoters than on peak promoters on the determination of TSSs by transcription factors in the cell.

As an example of transcription factors that target broad promoters, we investigated the Sp1 binding sites around TSSs. Sp1 recognizes binding region of DNA via its zinc finger domain whereas TBP recognizes TATA box via its DNA binding domain. Sp1 binding sites were enriched in the regions upstream of TSSs corresponding to the nucleosome-free regions. We observed similar tendencies for the binding sites of two transcription factors, PU.1 and MAZ. Although biological experiments are necessary to investigate molecular mechanism behind this observation, we speculate that the nucleosome-free regions serve as "landing sites" for transcription factors, including Sp1, which have less precise binding motifs (which are overrepresented among broad promoters) than the TATA box (Illingworth and Bird, 2009).

In addition to histone H3, we also analyzed the positions of the histone H2A variant H2A.Z, which is enriched around TSSs (Tirosh and Barkai, 2008), and we obtained similar results. In contrast, peak promoters did not have aligned patterns of nucleosome positions. One might suspect that the observation is due to high expression of genes associated with broad promoters, and low expression of those associated with peak promoters. However, even after we limited the analysis to broad and peak promoters both of which are associated with highly expressed genes, we still observed the preferences of H3 for broad promoters (region +100 to +130 bp with respect to TSSs) compared to peak promoters ($P < 1.0 \times 10^{-9}$, chi-squared test, data not shown). Although TSSs for TATA promoters are often fixed to single positions, our results suggest that such strictly controlled positions of TSSs are not regulated by nucleosome position. However, there is some evidence that the nucleosomes around TATA promoters have regulatory roles in gene expression. In yeasts, the TATA promoter is one type of "covered promoter", and expression of the genes associated with such promoters is more likely to be inhibited by the presence of nucleosomes than expression of the genes associated with "open promoters", which are located in nucleosome-free regions (Cairns, 2009); in covered promoters, nucleosomes often cover transcription factor binding sites to repress the expression of downstream genes. It is also possible that, in humans, peak promoters associated with the TATA box belong to one type of "covered promoter" where the expression of downstream genes is repressed by the presence of nucleosomes. Therefore, we speculate that transcription factor binding is controlled by nucleosome position in the case of peak promoters.

In our analysis of epigenetic control by histone modification, we uncovered a difference between broad and peak promoters. H3K4me1, -2, and -3 and H3K9ac, which are associated with gene activation, were more highly enriched around TSSs associated with broad promoters than around those associated with peak promoters. Thus broad promoters appeared to be under stronger epigenetic control than peak promoters. We found a trend that further supported this hypothesis: the expression levels of genes associated with broad promoters that had modified histones had higher expression levels than genes associated with broad promoters without modified histones. In contrast, peak promoters appeared to be under weaker epigenetic control, because far fewer of them harbored modified histones. Furthermore, there were no significant differences in the expression levels of genes associated with peak promoters that harbored or did not harbor modified histones.

It has been shown that promoters with many CpG islands are more likely to harbor modified histones than promoters with fewer CpG islands (Bhandare et al., 2010). However,

even after we limited our analysis to promoters having CpG islands, number of broad promoters harboring H3K4me3 was still statistically higher than that of peak promoters. Even more remarkable differences were observed after we limited our analysis to promoters without CpG islands. Although these results may depend on the dataset of CpG islands we used, enrichment of H3K4me3 in downstream region (+100 to +130 bp) of broad promoters were still observed in the analysis using different dataset of CpG islands (Bock et al., 2007) ($P < 1.0 \times 10^{-20}$ for CpG-related genes, $P < 1.0 \times 10^{-30}$ for CpG-unrelated genes).

Genes associated with broad promoters tend to be expressed ubiquitously, whereas those associated with peak promoters are likely to be expressed in specific tissues and may show low expression levels in most tissue types (Carninci et al., 2006). Therefore, if high levels of gene expression are directly associated with histone modifications around TSSs, then we may observe spurious correlations between promoter type and histone modification. In fact, H3K4me3 is known to upregulate the expression of downstream genes. We therefore compared the distribution patterns of nucleosomes containing H3K4me3 around broad and peak promoters in cases where the downstream genes showed similar expression levels (Figure 2.9). We found that the broad promoters also harbored more nucleosomes containing H3K4me3 in cases where the downstream genes showed similar expression levels (data not shown); the difference in the distributions of H3K4me3 around the broad and peak promoters was statistically significant (all positions from +100 to +130 showed significant differences; $P < 1.0 \times 10^{-3}$, chi-squared test), suggesting that promoter type was indeed associated with differences in epigenetic regulation by histone modifications.

Peak promoters containing the TATA box are regulated at their transcription initiation step, generally by the assembly of a pre-initiation complex with three additional components: the TATA-associated factors, the so-called mediator complexes, and positive and negative cofactors. We presume that peak promoters containing no TATA box are regulated in a similar way. This transcription system is widely used in various species, and our results suggest that it is unlikely to use epigenetic controls. Thus, broad and peak promoters have distinct systems to regulate gene expression.

Throughout this work, we employed widely-accepted definition of peak promoters, i.e. those which initiate transcription within the range of 4 bp. Changing this threshold to 10 bp did not have much effect on the distribution patterns of nucleosomes around broad and peak promoters as shown by Pearson's correlation coefficients between histone distribution pattern around broad promoters (–5,000 to 5,000 bp with respect to TSS) defined by >4 bp threshold

and that defined by >10 bp threshold. For H3 distribution patterns, correlation coefficients were 0.99 and 0.94 for broad and peak promoters, respectively. For H3K4me3 distribution patterns, the correlation coefficients were 0.99 for both broad and peak promoters. These results suggest the robustness of the relationships between the imprecision of TSS and patterns of histone distributions.

TATA boxes are used in a wide range of organisms, including prokaryotes, and are thought to be part of an ancient transcriptional system. In contrast, broad promoters are thought to be newly evolved (Carninci et al., 2006) and have incorporated histone modification systems. Our results showed that peak promoters, which are frequently associated with such ancient TATA boxes, have not incorporated histone modification systems.

By using a computational approach, we discovered the general relationships between the two types of promoter architecture and histone behavior, including positioning and modification (Figure 2.10). We first showed that the positions of histones around broad promoters were highly aligned and stable compared with those around peak promoters. Furthermore, we suggest that marked numbers of transcription initiations related to broad promoters are under the control of histone methylation and acetylation and are associated with gene expression level, whereas this is not the case with peak promoters. These results indicate that the expression of genes associated with broad promoters, but not peak promoters, is highly associated with histone position and modification. We believe that our study is a step in uncovering the general mechanisms underlying transcriptional systems and inferring how these systems have evolved. This should eventually help us to understand the complexity of mammalian transcription.

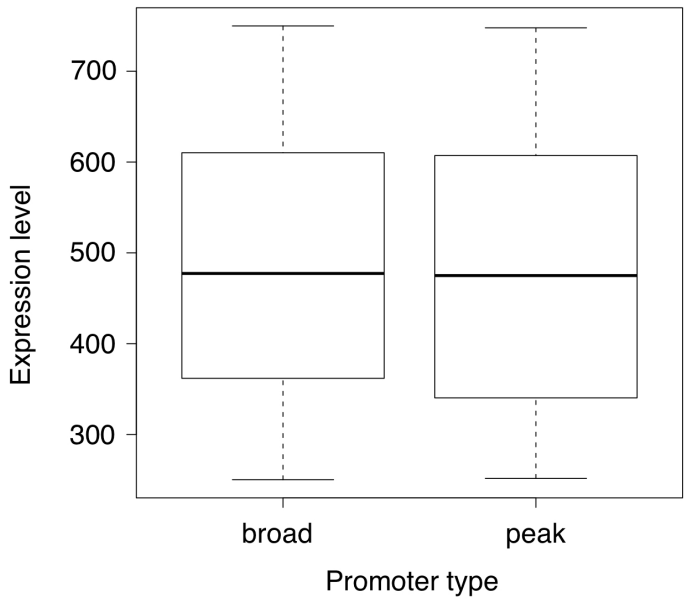

Figure 2.9 Distributions of expression levels of genes selected for comparison of broad and peak promoters associated with similar downstream gene expression

The box plots represent the distributions of the microarray intensities of the gene sets that were selected from among those associated with broad and peak promoters and that had similar expression levels (from 250 to 750).

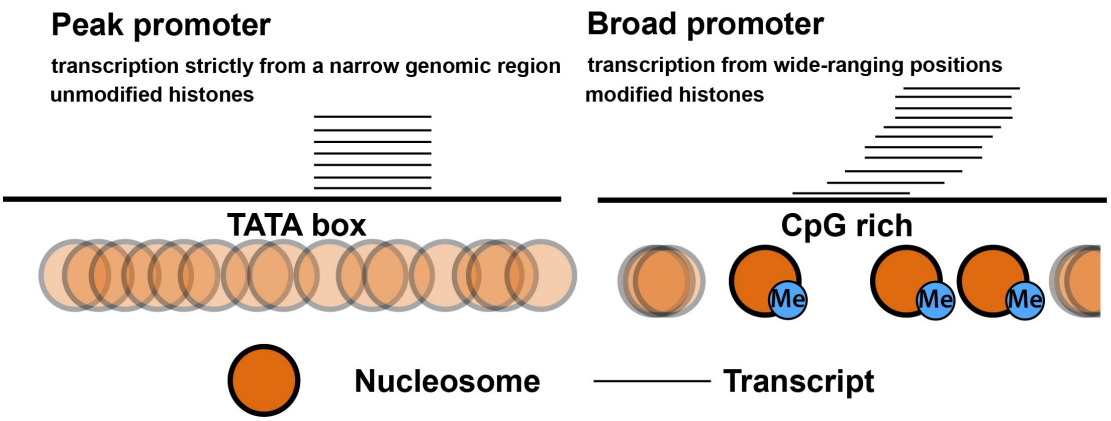

Figure 2.10 Summary of promoter types and epigenetic regulation

We found that, in humans, broad promoters (but not peak promoters) generally had significant associations with nucleosome positioning and histone modification. Specifically, around broad promoters histones were aligned in an orderly fashion. This feature was more evident with histones that were methylated or acetylated; moreover, the nucleosome positions around the broad promoters were more stable than those around the peak ones. More strikingly, the overall expression levels of genes associated with broad promoters (but not peak promoters) with modified histones were significantly higher than the levels of genes associated with broad promoters with unmodified histones.

Chapter 3

Single Nucleosome Imaging in Living Mammalian Cells

3.1. Introduction

Approximately 2 m of human genomic DNA is organized in the cell. DNA is wrapped around core histones and forms a nucleosome fiber known as “beads-on-a-string”. The nucleosome fiber has long been assumed to be folded into a regular 30-nm chromatin fiber and larger fibers folded into helical structures. However, several lines of evidence, including recent studies using cryo-EM and synchrotron X-ray scattering analyses (Eltsov et al., 2008; Joti et al., 2012; Nishino et al., 2012) demonstrated no regular structure >11 nm in interphase chromatin and mitotic chromosomes. These findings suggest that interphase chromatin and mitotic chromosomes consist of compact and irregularly folded nucleosome fibers without a 30-nm chromatin structure (i.e., a polymer melt-like structure) (Eltsov et al., 2008; Maeshima et al., 2010). This structure implies a less physically constrained and more locally dynamic state. Nucleosome fibers may thus be constantly moving and rearranging at the local level, which may be crucial for protein factors in the search for genomic DNA. How can we observe such local behavior of the nucleosome in living human cells? The obvious question was whether nucleosome fluctuation could be detected in living mammalian cells. For this purpose, we newly developed and performed single particle imaging of nucleosomes in living cells.

3.2. Materials and methods

3.2.1. Microscopy setup

Nikon laser TIRF microscope system Ti was used to observe the distribution of single H4-PA-GFP molecules expressed in DM (Indian Muntjac) cells. 488-nm laser (100 mW and ND 50%) was introduced into the microscope through an optical path installed on a vibration insulation table. Through an objective lens (100× PlanApo TIRF, NA 1.49; Nikon), DM cells grown on a glass base dish (Iwaki) were exposed to the excitation light. The incident angle of the laser beam to the specimen plane was chosen so as to obtain a oblique illumination (Tokunaga et al., 2008). Collected fluorescence from the cells was focused on the sCMOS ORCA-Flash 4.0 camera (Hamamatsu Photonics). Movie of sequential 1,000 frames was acquired by MetaMorph software (Molecular Device) at a frame rate of 30 ms under continuous illumination. The observation stage was kept at a constant 37°C, 5% CO_2, and humidity (Tokai-hit live-cell chamber and GM-8000 digital gas mixer).

3.2.2. Detection of single fluorescent particle

Subpixel accuracy positions of the PA-GFP dots were determined using the image-processing software u-track (Jaqaman et al., 2008; Rogers et al., 2007) and the accuracy for determining the position of fluorescent dots was estimated using the FIONA method (Ober et al., 2004; Thompson et al., 2002; Yildiz et al., 2003). With this procedure, the trajectory of each fluorescent dot was obtained. We calculated the displacement and the mean square of displacement (MSD) of fluorescent nucleosomes from the tracking data. The originally calculated MSD was in two dimensions. To obtain the 3D value, the two-dimensional value was multiplied by 1.5 ($6Dt/4Dt$). Using KaleidaGraph (Synergy Software, USA), histograms of the displacement were prepared.

Histone H4-PA-GFP has some flexible regions, including the linker and histone tail, which is maximally 50 amino acid residues, corresponding to a length of ~17 nm. The previous study observed by FCS rapid movement of free GFP in the chromosomes at 15 mm^2/s (Hihara et al., 2012). If PA-GFP is rapidly mobile within a restricted area, like a “dog on a leash”, we consider that the effect of the flexible region on the nucleosome position determination is negligible.

3.3. Results and discussion

To observe single nucleosomes and analyze their local dynamics in living human cells, histone H4 was fused with photoactivatable (PA)-green fluorescent protein (GFP) and expressed in DM cells at a very low level (Figure 3.1 A). For single nucleosome imaging, an oblique illumination microscopy (Figure 3.1 B) was used to illuminate a limited thin area within the cell (Nikon laser TIRF microscope system Ti with sapphire 488-nm laser). Generally, PA-GFP exhibits green fluorescence only after activation by a 405-nm laser. However, we unexpectedly found that a small fraction of H4-PA-GFP in DM cells was spontaneously activated without laser stimulation, and were observed as dots (Figure 3.1 A and C).

Then, we observed single-step photobleaching of the H4-PA-GFP dots (Figure 3.1 D), although photobleaching of multiple H4-PA-GFP complexes should occur in multiple steps or gradually. Thus, each bright dot in the nucleus represents a single H4-PA-GFP in a single nucleosome (Figure 3.1 C).

The estimated number of H4-PA-GFP was less than 5% of endogenous H4 (Kimura and Cook, 2001). In addition, since free H4-PA-GFP should have a much larger *D* value, we could trace only the H4-PA-GFP fraction that was stably incorporated into nucleosomes. To show the displacement distribution, we measured approximately 20,000 signal points with single nucleosome tracking with H4-PA-GFP from seven DM cells at a video-rate (~30 ms/frame) as a movie. For single nucleosome tracking, we used u-track software in MATLAB (Jaqaman et al., 2008; Rogers et al., 2007). The dots were fitted to an assumed Gaussian point spread function to determine the precise centers of the dots with higher resolution. We were able to analyze the behaviors of the nucleosomes over short periods of time, ranging from 0 to 0.18 s.

Local nucleosome fluctuation was observed (~50 nm/30 ms), presumably caused by Brownian motion. Three representative trajectories of fluorescently tagged single nucleosomes are shown (Figure 3.1 E; scale bar = 100 nm). Recently, the McNally group also published single-nucleosome tracking data using H2B-EGFP (Mazza et al., 2012), which appear to be consistent with our single nucleosome tracking with H4-PA-GFPs.

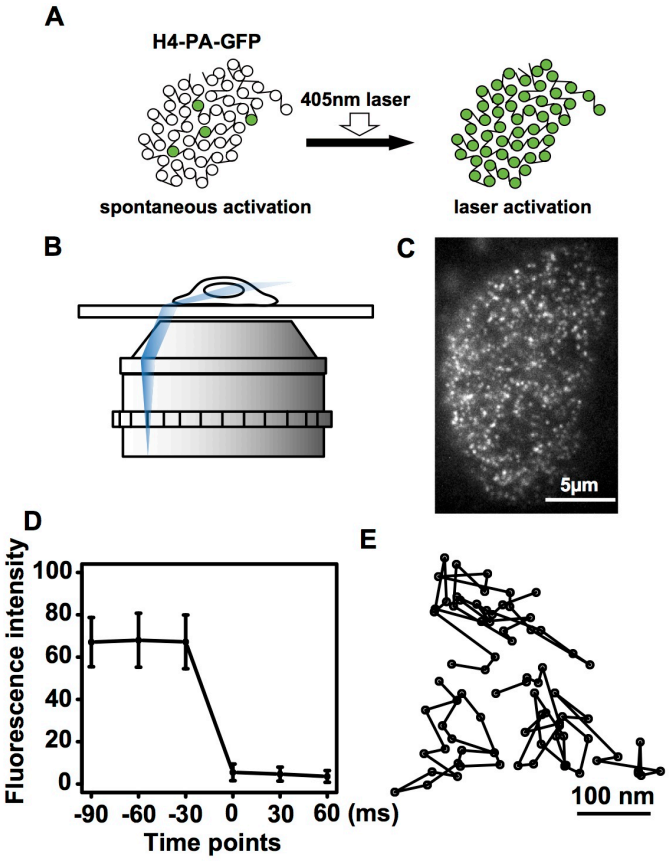

Figure 3.1 Single nucleosome imaging

(A) Generally, PA-GFP shows GFP fluorescence only after activation by a 405-nm laser (right). However, a small fraction of H4-PA-GFP was spontaneously activated without laser activation (left) and was used for our single-nucleosome imaging. (B) A schematic representation of the oblique illumination microscopy (Tokunaga et al., 2008). We used Nikon laser TIRF microscope system Ti with Sapphire 488-nm laser (Coherent). Using a sheet light (blue), a limited area in the cell is illuminated. (C) Single-nucleosome image of a DM cell nucleus expressing H4-PA-GFP observed as a bright dot using the oblique illumination microscopy system. For single-nucleosome tracking, the software u-track, was used. The dots were fitted to an assumed Gaussian point spread function to determine the precise center of signals with higher resolution. (D) Single-step photobleaching of the H4-PA-GFP dots. (E) Representative trajectories of fluorescently tagged single nucleosome.

The displacement data of purified GFP fixed on a glass surface were obtained (Figure 3.2 C and D). The average displacement values were 12.8 ± 0.2 nm/30 ms, which was much smaller than those of H4-PA-GFP observed in living cells (59.0 ± 0.2 nm/30 ms in Figure 3.2 A). To evaluate the contribution of whole cell or nuclear movements to the single nucleosome displacements, we again calculated the centroid movements for many nucleosomes in the same time flame (illustrated in Figure 3.3 A). The centroid movements (5.5 ± 0.1 nm/30 ms) (Figure 3.3 B) are much smaller than those in Figure 3.2 A and 3.2 B, suggesting that the detected nucleosome movement is not derived from the global motion of cells or nuclei.

To analyze local nucleosome movement in DM cells further, the MSD values (μm^2) of nucleosomes in the DM cells and fixed EGFP were plotted with their standard errors (Figure 3.3 C). The plots for the nucleosomes fitted well with the exponential equation, $MSD = 0.022t^{0.36}$ (Figure 3.3 C). These results support the restricted nucleosome movement model.

It was reported that movements of large chromatin domains tagged with LacO-array/LacI-GFP (~100 kb and above) were on the 10 and 100 nm length scale for 30 ms (e.g., (Levi et al., 2005)). We could not exclude the possibility that movements of larger domains might contribute to the local nucleosome fluctuation that we observed. Their relationship would be an intriguing issue for further study. This study revealed local nucleosome fluctuation in living mammalian cells (Figure 3.4). Monte Carlo computer simulations suggested that nucleosome fluctuation facilitates the mobility of diffusing proteins in the environment (Hihara et al., 2012). In addition, such nucleosome fluctuation can expose DNA sequences to the surface of chromatin domains more often, while static regular folding structures, such as 30-nm chromatin fibers, can hide or mask most DNA sequences (Figure 3.4). It should be emphasized that both facilitation of protein mobility and DNA exposure lead to an increase in chromatin accessibility. This is in good agreement with our finding that tight cross-linking of nucleosomes blocked antibody accessibility and targeting in dense chromatin regions (Hihara et al., 2012).

Recently, some papers suggested that interphase chromatin forms numerous compact chromatin domains resembling "chromatin liquid drops" in interphase cells (Joti et al., 2012; Maeshima et al., 2010). This chromatin liquid drops view is in line with predictions of the chromosome territory-interchromatin compartment (CT-IC) model (Cremer and Cremer, 2001; Cremer et al., 2000). In the CT-IC model, each CT is built up from a series of interconnected, megabase-sized chromatin domains, which were originally identified using a pulse labeling as the DNA replication foci (Albiez et al., 2006; Berezney et al., 2005; Schermelleh et al., 2001). Recent high-throughput 3C studies, such as Hi-C and 5C, have also proposed physical packaging of genomic DNA, termed "topologically associating domains" (Nora et al., 2012), "topological domains" (Dixon et al., 2012) or "physical domains" (Sexton et al., 2012). Local nucleosome fluctuation would play an especially important role in such compact chromatin domains because the effect of nucleosome fluctuation on the facilitation of chromatin accessibility is more significant in a compact chromatin environment (Hihara et al., 2012).

Many biological processes, including transcription, DNA repair, replication, and recombination, are based on "scanning genomic DNA and targeting" (Figure 3.4). For example, in transcriptional regulation, the dynamic movement of nucleosome fibers assists with the targeting of transcription factors and complexes. Dynamic irregular folding can easily form loops, facilitating interaction between promoters and enhancer sequences. Since the flexible and dynamic nature of nucleosome fibers in living mammalian cells drives such biological processes, regulation of local nucleosome dynamics, possibly by histone modification and/or specific proteins, will be an important factor in their regulation.

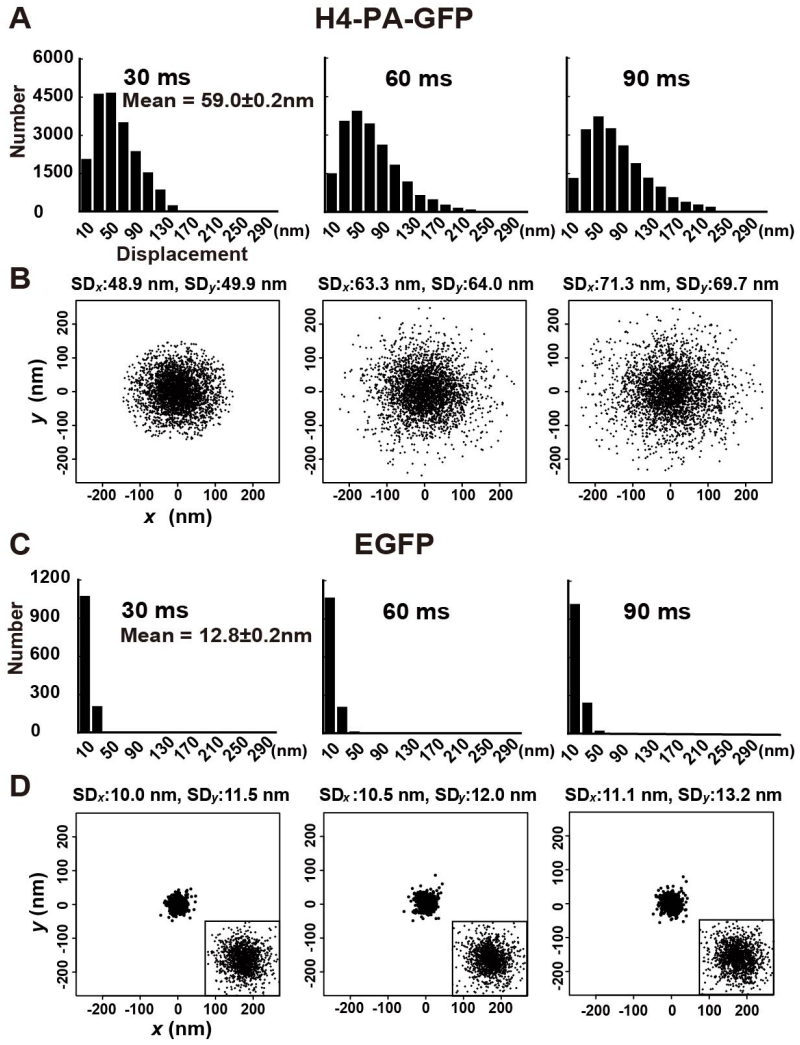

Figure 3.2 Single nucleosome analysis of H4-PA-GFP

(A) Displacement (movement) histograms of single nucleosomes in interphase DM cells for 30 ms (left), 60 ms (center), and 90 ms (right; n = 20,000 in seven cells). The mean displacement ± standard error for the 30 ms time point was indicated. (B) Distribution of nucleosome displacements in the *x*-*y* plane for 30 ms (left), 60 ms (center), and 90 ms (right). n = 3,000 in seven cells. Standard deviations (SD_x and SD_y) were shown. In panels (A) and (B), the originally calculated displacement were in two dimensions. To obtain their 3D values, the two-dimensional values of displacement were multiplied by $\sqrt{1.5}$ ($\sqrt{(6Dt/4Dt)}$). (C) Displacement histograms of single EGFP on a glass surface for 30 ms (left), 60 ms (center) and 90 ms (right; n = 1,300). The mean ± standard error for the 30 ms time point was indicated. (D) Distribution of EGFP displacements in the *x*-*y* plane for 30 ms (left), 60 ms (center), and 90 ms (right; n = 1,300). Standard deviations (SD_x and SD_y) were shown. The insets are 3-fold magnified images of the distributions.

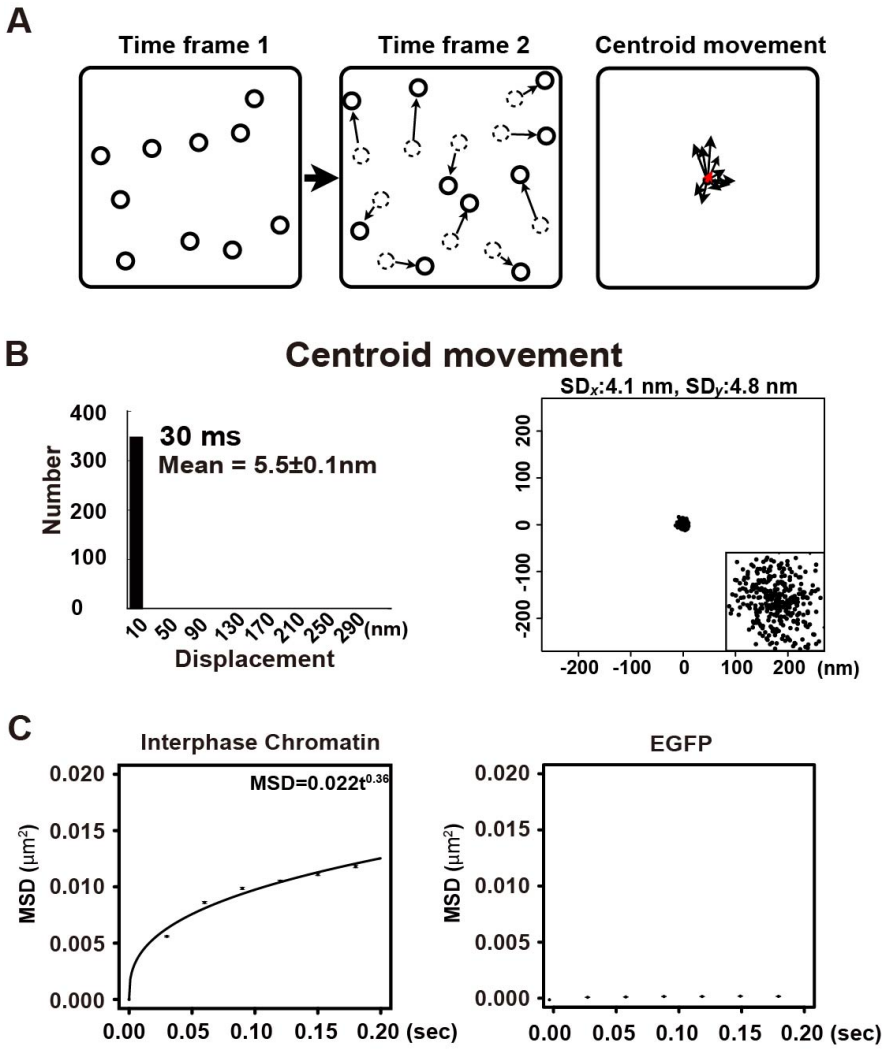

Figure 3.3 Local nucleosome fluctuation in living mammalian cells

(A) Schematic representation for calculating centroid movement for many nucleosomes in the same time frame. The black arrows represent the movements of nucleosomes and the small red arrow represents centroid movement in the frame. (B) Displacement histogram (left) and distribution in the *x*-*y* plane (right) of centroid movement in the frame for 30 ms (n = 350 in seven cells). The inset is 10-fold magnified image of the distribution. (C) Plots of the mean square displacements (MSDs) with standard errors of single nucleosomes in interphase chromatin (left) and single EGFPs on a glass surface (right) from 0 to 0.18 s. The plots for single nucleosomes were fitted as an anomalous diffusion, suggesting that nucleosome movement supports a restricted diffusion model. Error bars represent the standard errors of the mean value. In the panel of MSD for single nucleosomes, the originally calculated MSD were in two dimensions. To obtain their 3D values, the two-dimensional values of MSD were multiplied by 1.5 ($6Dt/4Dt$).

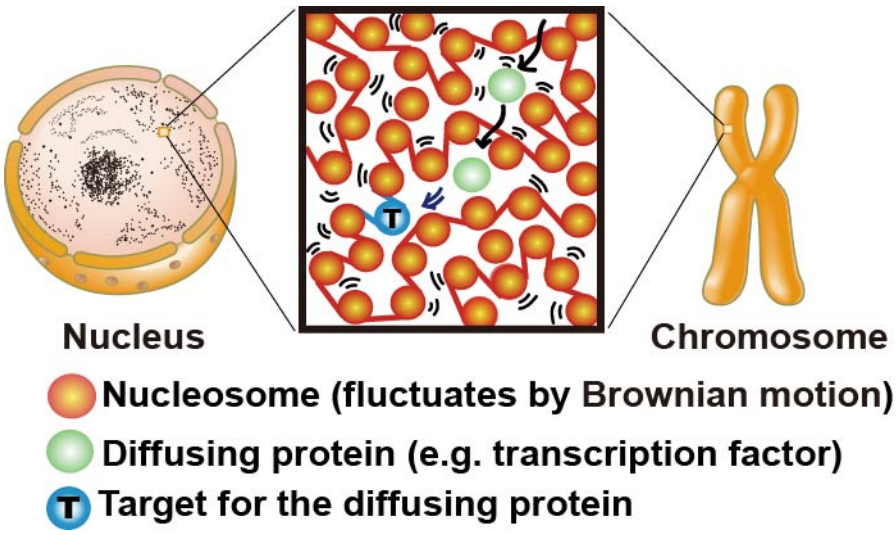

Figure 3.4 Flexible and dynamic nucleosome fiber in living mammalian cells

We demonstrated the flexible and dynamic nature of the nucleosome fiber in living mammalian cells with the newly developed single nucleosome imaging. Since the flexible and dynamic nature of nucleosome fibers in living mammalian cells drives biological processes, regulation of local nucleosome dynamics, possibly by histone modification and/or specific proteins, will be an important factor in their regulation.

Chapter 4

Dynamic Chromatin Domains Are Organized by Nucleosome-nucleosome Interactions and Cohesin in Living Mammalian Cells

4.1. Introduction

Genomic DNA is wrapped around core histones and forms a nucleosome fiber (10-nm fiber) (Kornberg and Lorch, 1999). Recent evidence has demonstrated that the nucleosome fiber is irregularly folded in the cell as chromatin (Eltsov et al., 2008; Fussner et al., 2012; Nishino et al., 2012). For higher order chromatin structure, the chromosome conformation capture and related methods (Dekker et al., 2013) have suggested that genomic DNA forms numerous chromatin domains such as "topologically associating domains" (TADs) (Dixon et al., 2012; Nora et al., 2012; Rao et al., 2014; Sexton et al., 2012) that are of hundreds kilobases in size and are considered to be as functional units of genome for RNA transcription and DNA replication. However, since such domains were identified using chemically fixed cells, how such chromatin domains form, distribute and behave in the living cells remains unclear. Here, by combining two imaging methods based on single nucleosome imaging (Chapter 3): super-resolution imaging PALM (photoactivated localization microscopy) (Betzig et al., 2006; Rust et al., 2006) and single nucleosome tracking, we developed a novel nuclear imaging system, which allows us to integrate information for chromatin domains and dynamics in living mammalian cells. We created a chromatin domain map and also its "heat map" in the nucleus, revealing that the more heterochromatic regions are, the less chromatin movement (or "colder"). Our study demonstrated that the formation of the chromatin domains is orchestrated by electrostatic force (Maeshima et al., 2014) and molecular crowding effect (Marenduzzo et al., 2006) as well

involving the cohesin complex (Nasmyth and Haering, 2005; Shintomi and Hirano, 2010) and nucleosome-nucleosome interactions (Kalashnikova et al., 2013). Notably, the chromatin domains were observed throughout the cell cycle. We provide a novel mechanistic insight into mitotic chromosome assembly with keeping epigenetic information: the chromatin domains formed in interphase cells are assembled together to create a rod-like chromosomal shape during mitosis to function as building blocks of chromosomes.

4.2. Materials and methods

4.2.1. PALM imaging

PALM imaging was carried out using an inverted Nikon Eclipse Ti microscope with 100 mW Sapphire 561-nm laser (Coherent) and sCMOS ORCA-Flash 4.0 camera (Hamamatsu Photonics). Cells were exposed to the excitation laser through an objective lens (100× PlanApo TIRF, NA 1.49; Nikon). The images were taken by oblique illumination system with TIRF unit (Nikon) to illuminate a limited thin area in the cell nucleus. Movie of sequential 1,000 frames was acquired by MetaMorph software (Molecular Device) at a frame rate of 50 ms under continuous illumination. To maintain cell-culture condition (37°C, 5% CO_2, and humidity) during the imaging, Tokai-hit live-cell chamber and GM-8000 digital gas mixer were used.

4.2.2. Plasmid construction

Construction of pEF1α-H2B-PA-mCherry-FRT is the following: PA-mCherry sequence was amplified with the addition of *Bam*HI and *Not*I sites on the ends from pPA-mCherry-N1 (Clontech) using the following primer pair:

5’-CGCGGATCCACCGGTCGCCACCATGGTGAGCAAGGG-3’

and

5’-AAGGAAAAAAGCGGCCGCTTACTTGTACAGCTCGTCCA-3’.

The amplified PA-mCherry fragment was replaced with PA-GFP region of pH2B-PA-GFP vector (EUROSCARF, Ellenberg lab) via *Bam*HI and *Not*I sites. H2B-PA-mCherry sequence was then amplified using the following PCR primer pair:

5’-CTAGCTAGCATGCCAGAGCCAGCGAAGTC-3’

and

5’-AAGGAAAAAAGCGGCCGCTTACTTGTACAGCTCGTCCA-3’.

This fragment was inserted into *Eco*RV site of pEF5/FRT/VS-DEST Gateway Vector (Invitrogen) to obtain pEF1α-H2B-PA-mCherry-FRT.

To construct pEF1α-H2B-PA-mCherry in PiggyBac vector (pPB-EF1α-H2B-PA-mCherry-PGKneo), H2B-PA-mCherry sequence was amplified from pEF1α-H2B-PA-mCherry using the following primer pair:

5’-AAAGATATCGGTCTTGAAAGGAGTGCCTCG-3’

and

5’-AAAGATATCAAGCCATAGAGCCCACCGCAT-3’.

The amplified fragment was digested with *Eco*RV and then inserted into *Eco*RV site of

pPB-PGKneo vector.

4.2.3. Cell culture and isolation of stable cell lines

The cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (HeLa cell) and 15% (DM cell) fetal bovine serum at 37°C in 5% CO_2. ES E14Tg2a cells were cultured in Glasgow minimum essential medium (GMEM) supplemented with 10% fetal bovine serum, 1% non-essential amino acids, 1 mM sodium pyruvate solution, 2 mM L-glutamine, 1 mM 2-mercaptoethanol, and 100 U/ml leukemia inhibitory factor (LIF). For LIF(-) condition, ESCs were cultured for 5 days after LIF withdrawal. For the establishment of HeLa or DM cells stably expressing H2B-PA-mCherry, Flp-In system (Invitrogen) was used as described previously with Effectene Transfection Reagent (QIAGEN) and then transformants were selected with 200 μg/ml Hygromycin. For the establishment of ESCs stably expressing H2B-PA-mCherry, PiggyBac transposon system was used. pPB-EF1α-H2B-PA-mCherry-PGKneo and pCMV-hyPBase were transfected into the cells with Effectene Transfection Reagent (QIAGEN) and then transformants were selected with 600 μg/ml neomycin G418. For PALM imaging, all types of cells were plated onto glass bottom dishes (Iwaki) treated with poly-lysine. Before the microscopy imaging the medium was replaced by DMEM (no Phenol Red and 15% FBS).

4.2.4. Chemical treatment

To increase histone H4 tail acetylation, cells were treated with 500 nM tricostatin A (TSA) (Sigma-Aldrich) for 8 hours. To deplete ATP, cells were incubated in the medium supplied with 10 mM sodium azide (Sigma-Aldrich) and 50 mM 2-Deoxy-Glucose (Sigma-Aldrich) for 30 minutes. For transcription inhibition, cells were cultured in the medium with 100 μM DRB (Sigma-Aldrich) for 2-3 hours. For hypotonic treatment, cells were incubated in the medium including 1 ml DMEM and 1 ml MilliQ water for 2 hours.

4.2.5. Cell cycle synchronization for PALM imaging

HeLa cells were synchronized by 0.08 μg/ml of nocodazole (Wako) for 4 hours, and mitotic cells were harvested by shake off. Cells were washed with PBS(-) and plated on the glass base dishes.

4.2.6. Conventional and correlative immunostaining

Immunostaining was carried out as described previously (Chereji and Morozov, 2011;

Ghirlando and Felsenfeld, 2013; Hiratani and Gilbert, 2010; Koch et al., 1987; Meluzzi and Arya, 2013; Rust et al., 2006). Primary antibodies were mouse anti-Rad21 (Millipore, 05-908), rabbit anti-CTCF (Millipore, 07-729), mouse anti-NIPBL (Santa Cruz, sc-374625), rat anti-CAP-H2 (5F2G4), rabbit anti-Sox2 (abcam, ab97959), mouse H3K9me3 (a kind gift from Prof. Hiroshi Kimura), and mouse Polymerase II Ser5ph (a kind gift from Prof. Hiroshi Kimura). Images were obtained with Delta Vision microscopy (Applied precision). For DNA staining in living cells, Hoechst 33342 (500 ng/ml) was added to the cells for 30 minutes and washed by PBS(-).

For correlative immunostaining, cells were plated on a glass base dish with a grid (Matsunami) coated with poly-lysine. After live PALM imaging, cells were fixed with 2% formaldehyde, followed by the conventional immunostaining. After the staining, the same cell was sought based on coordinates of a grid and the images were obtained with Delta Vision microscopy or Nikon Eclipse Ti microscope.

4.2.7. RNA interference

The siRNA transfection was performed with Lipofectamine RNAiMAX (Invitrogen) according to manufacturer's instruction. siRNA oligos were used as reported previously: Rad21 (Wendt et al., 2008); CTCF (Wendt et al., 2008); CAP-H2 (Ono et al., 2003); CAP-G2 (Ono et al., 2003); NIPBL (Zuin et al., 2014b). Low GC content oligo (Invitrogen, 45-2002) was used as control. For double treatments of Rad21-KD and TSA, cells were cultured for 48 hours after Rad21 siRNA transfection and then treated with TSA (500 nM) for 4 hours.

4.2.8. EU and EdU labeling

EU and EdU incorporations were performed with Click-iT RNA Imaging kits (Invitrogen) using Alexa Fluor 594 according to manufacturer's instructions.

4.2.9. DNA replication foci formation

To make DNA replication foci with Cy3-dCTP labeling, the scratch method was used as (Schermelleh, 2005). With 200 nM of Cy3-dCTP (GE), cells were scratched 200 times with a G27 fine needle.

4.2.10. Biochemical fractionation of nuclei from the cells expressing H2B-PA-mCherry

Nuclei were isolated from HeLa cells expressing H2B-PA-mCherry as described previously (Rust et al., 2006). The nuclei were incubated on ice for 15 minutes with a series of the buffers: HE (10 mM Hepes-NaOH pH 7.5, 1 mM EDTA and 0.1 mM PMSF), HE + 100 mM NaCl, HE + 500 mM NaCl, HE + 1 M NaCl, HE + 2 M NaCl. After incubation with salt, centrifugation was performed to separate the nuclei solutions into supernatant and pellet fractions. The proteins in supernatant fractions were precipitated with 17% TCA and cold acetone. Both pellets were suspended in SDS-PAGE buffer and subjected to 12.5% SDS-PAGE and subsequent CBB staining and Western blot with anti-H2B (Millipore) and anti-mCherry (RFP) (MBL) antibodies.

4.2.11. Data analysis for PALM imaging and single nucleosome tracking

Sequential microscopy images were converted to 8-bit gray scale and the background signals were subtracted with ImageJ. The nuclear regions in the images were extracted. Following this step, the centroid of each fluorescent dot in every image was determined and its trajectory was tracked by u-track (MATLAB package). To make PALM image based on the data, the nucleosome positions were mapped using R software (65 nm/pixel), and then a Gaussian blur (sigma = 1 pixel) was added to obtain a smoother rendering with ImageJ.

For single nucleosome movement analysis, the displacement and MSD of fluorescent signals were calculated based on the u-track data. The originally calculated MSD was in two dimensions. To obtain the 3D value, the two-dimensional value was multiplied by 1.5 ($4Dt \rightarrow 6Dt$). Using KaliedaGraph (Synergy Software, USA), histograms of the displacement were prepared.

To make the heat map of domain dynamics, the median of movements of nucleosomes (during 50 ms) in 3 × 3 pixels (65 nm/pixel) was calculated and plotted with a color scale (from blue to red) using R.

4.2.12. Clustering analyses of nucleosomes in PALM images

The two-dimensional radial distribution function (RDF) is given by

$$g_{2D}(r)=\left[\left(\frac{S}{N-1}\right)\frac{1}{\pi\left(2r\Delta r+\Delta r^{2}\right)}\right]\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{i\neq j}\delta\left(r-r_{i,j}\right)\right],$$

where Δr=10 nm is the binning width and $(N-1)/S$ is the average particle density of the S, which is the square of the total area. N is the total number of particles which area contains. The factor $\pi(2r\Delta r+\Delta r^{2})$ is just the area of a ring of width Δr with a mean radius of $r+\Delta r$. Delta function is given by

$$\delta\left(r-r_{i,j}\right)=\begin{cases}1, & (r-r_{i,j})\leq\Delta r\\ 0, & others\end{cases},$$

$r_{i,j}$ is the distance between r_i and r_j.

Ripley's K function is given by

$$K(r)=\left(\frac{S}{N-1}\right)\left[\frac{1}{N}\sum_{i=1}^{N}\sum_{i\neq j}\delta\left(r-r_{i,j}\right)\right],$$

where $(N-1)/S$ is the average particle density of the S, which is the square of the total area. N is the total number of particles which area contains. Delta function is given by

$$\delta\left(r-r_{i,j}\right)=\begin{cases}1, & r_{i,j}\leq r\\ 0, & r_{i,j}>r\end{cases},$$

$r_{i,j}$ is the distance between r_i and r_j.

L function is given by

$$L(r)=\sqrt{\frac{K(r)}{\pi}}.$$

Area (S) of the total region of nucleus was estimated using Fiji plugin Weka and the area of the whole region was measured by Analyze Particles.

4.2.13. Computer modeling of sparse and crowded chromatin domains

To qualitatively verify existence of domain structures in the mitotic chromosome, we performed computer simulations using a coarse-graining model. Here we represent a chromatin fiber as a polymer with single chain structure. Neighboring monomers are connected by harmonic bonds with a potential; $E_b = k_b(b - b_0)^2/2$, where k_b, b and b_0 are a force constant, a distance between neighboring monomers and a targeting distance, respectively. All monomers except neighbors interact via a Lennard-Jones potential; $E_{LJ} = e(r_0^{12}/r^{12} - 2r_0^6/r^6)$, where e, r and r_0 are a depth of the potential well, a distance between the monomers and the distance where E_{LJ} has the minimum value, respectively. To mimic domain structures in the chromatin fiber, we add a constraint potential; $U_d = k_d|\mathbf{r}_a - S\mathbf{r}_i/N_d|^2/2N_d$, to each monomer *a*, where k_d, $\mathbf{r}_i$ and N_d are a force constant, the positional vector of a monomer *i* and the number of monomers within the domain that includes monomer *a*, respectively. The summation in the potential energy formula, U_d, is taken for monomers in the domain. We defined the region of domains arbitrary in this qualitative model simulation. To model chromatin structure that satisfies the contact map by Hi-C, we also add a constraint potential; $U_{exp} = k_{exp}(r - r_{exp})^2$, for all monomer pairs to be contacted. k_{exp} and r_{exp} are a force constant and the distance that mimics contact in Hi-C map, respectively. The number of monomers is 4,000, k_b=100kT, b_0=3.8 nm, e=0.25kT, k_d=0.1kT, k_{exp}=0.001kT and r_{exp}=19 nm in this manuscript. To simulate a band-like structure in Hi-C map, two monomers within 400 units along the polymer chain are defined as "contacted" here. To simulate a stable conformation of the chromatin fiber, we conducted a simulated annealing method. We started a molecular dynamics simulation of the polymer model with kT=1,000, then temperature was gradually decreased after every 5,000 time steps until it reaches kT=0.1. We adopted a constant temperature MD simulation with the Gaussian constraint method.

4.3. Results

4.3.1. Single nucleosome imaging in a living HeLa cell

To get an entire view of chromatin domain structure and dynamics in living mammalian cells, we utilized a live-cell PALM (photoactivated localization microscopy) that allows obtaining images with a resolution beyond the diffraction limit (Betzig et al., 2006; Rust et al., 2006). We fused histone H2B with photoactivatable (PA)-mCherry (Subach et al., 2009), which acquires fluorescence upon laser stimulation (Figure 4.1 A), and expressed the fusion protein at a low level in human HeLa cells (Figure 4.2 A). The histone H2B is considered to be incorporated into the nucleosomes throughout the genome by frequent histone replacement (Kimura and Cook, 2001). For the chromatin domain imaging, we used an oblique illumination microscopy, allowing us to illuminate a thin optical layer within a single nucleus (Figure 4.1 B) (Tokunaga et al., 2008). Using this microscopy imaging, we found that a relatively small number (~100/frame) of H2B-PA-mCherry in the cells was spontaneously and continuously activated without the laser stimulation (Figure 4.1 A for the scheme), and observed as clear dots (Figure 4.1 C). Notably, these dots showed single-step photobleaching (Figure 4.1 D and E), suggesting that each dot represents a single H2B-PA-mCherry molecule in a single nucleosome. In addition, a stepwise salt washing of the isolated nuclei from the expressed cells confirmed that ectopically expressed H2B-PA-mCherry behaved similarly to endogenous H2B, suggesting that the H2B-PA-mCherry molecule was properly incorporated into the nucleosome structure in the cells (Figure 4.2 A and B).

We recorded the nucleosome dots in the interphase chromatin at 50 ms/frame (~1,000 frames, totally 50 sec total) as a movie and then fitted the dots in each image frame to an assumed Gaussian point spread function to determine the precise center of dots with a higher resolution below the diffraction limit. Note that our microscopy camera was able to detect the PA-mCherry signals only in a thin optical layer with ~250 nm thickness, which were in focus, as shown by the point spread function (PSF) of the H2B-PA-mCherry signal measured in the cells (Figure 4.1 F).

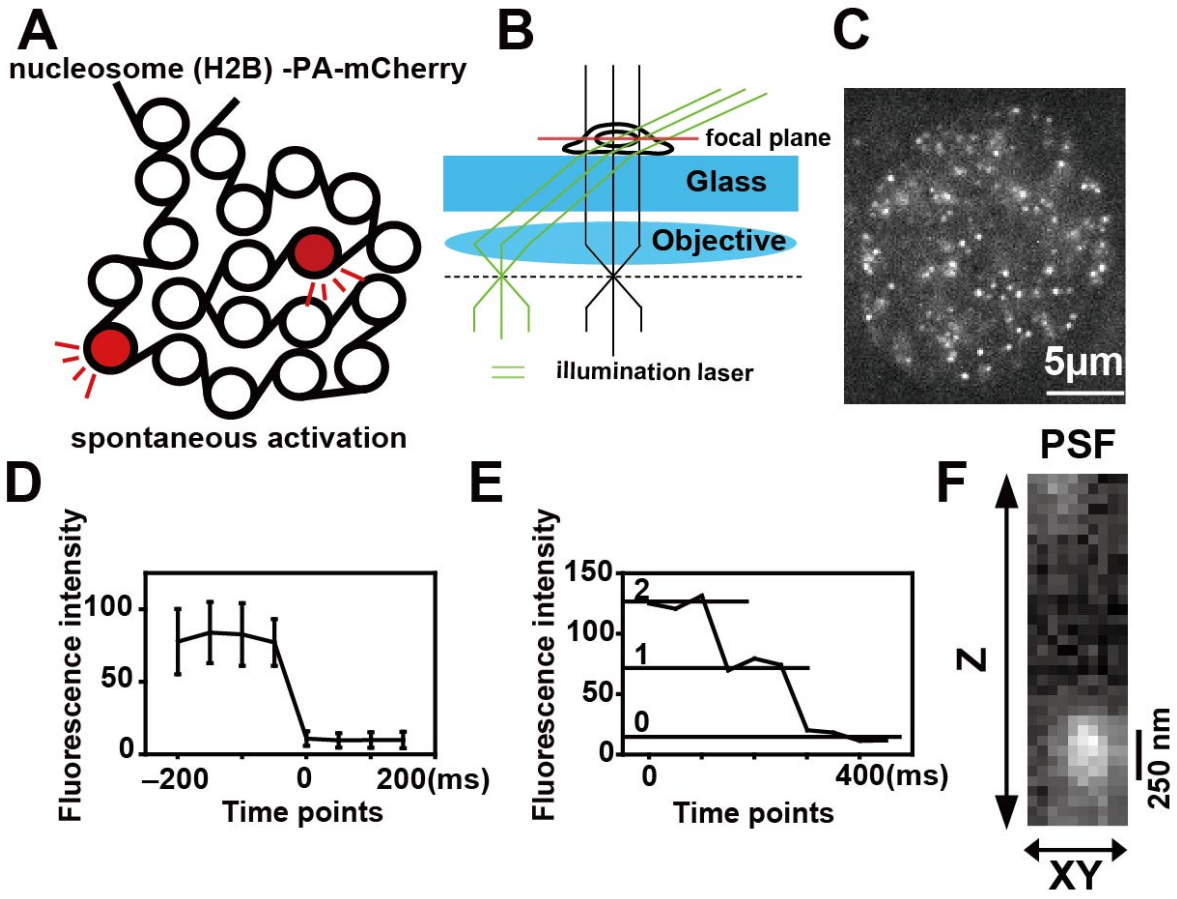

Figure 4.1 Single nucleosome imaging

(A) A small fraction of H2B-PA-mCherry was activated spontaneously without 405-nm laser activation and was used for live PALM imaging and single nucleosome tracking. (B) Schematic representation of the oblique illumination microscopy. We used Nikon laser TIRF microscope system Ti with Sapphire 564-nm laser. Using a sheet light (green), only a thin optical layer within the nucleus (red) is illuminated with very low background. (C) Single-nucleosome (H2B-PA-mCherry) image of a living HeLa nucleus. Scale bar shows 5 μm. (D) Single-step photobleaching of the H2B-PA-mCherry dots. The vertical axis represents the fluorescence intensity of each H2B-PA-mCherry dot. The horizontal axis is the tracking time series (photobleaching point is set as time 0; n = 100). Due to the clear single-step photobleaching

profile of the H2B-PA-mCherry dots, each dot represents a single H2B-PA-mCherry molecule in a single nucleosome. (E) A rare example of multiple (two)-step photobleaching of the H2B-PA-mCherry dots. The bright dot with such bleaching pattern was assumed to be derived from multiple H2B-PA-mCherry molecules. (F) PSF of observed H2B-PA-mCherry in a fixed HeLa cell. The *xy*-*z* kymograph shows that H2B-PA-mCherry can be detected in an optical layer with ~250 nm thickness. The vertical axis represents the depth of optical layer (*z*-plane) (50 nm/pixel), and the horizontal axis represents the *xy*-plane (65 nm/pixel). Scale bar shows 250 nm.

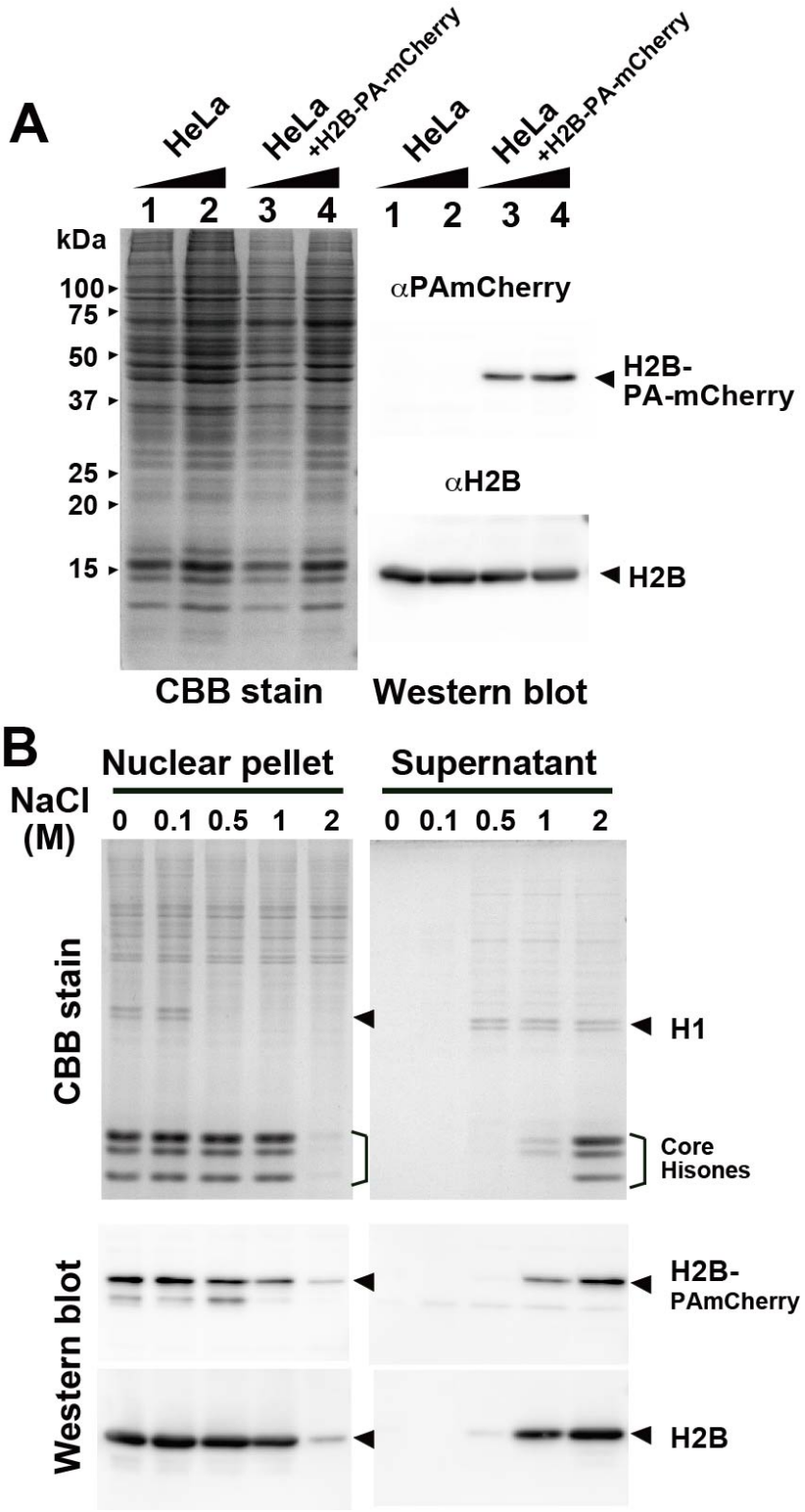
A
HeLa
HeLa +H2B-PA-mCherry
HeLa
HeLa +H2B-PA-mCherry
1 2 3 4 1 2 3 4
kDa
100
75
50
37
25
20
15
αPAmCherry
H2B-PA-mCherry
αH2B
H2B
CBB stain
Western blot
B
Nuclear pellet
Supernatant
NaCl (M)
0 0.1 0.5 1 2 0 0.1 0.5 1 2
CBB stain
H1
Core Hisones
Western blot
H2B-PAmCherry
H2B

Figure 4.2 H2B-PA-mCherry molecule was properly incorporated into the nucleosome structure

(A) Lysates of the cells expressing H2B-PA-mCherry were subjected to SDS-PAGE and subsequent CBB staining and western blot with anti-mCherry (RFP) and anti-H2B antibodies. Expression of H2B-PA-mCherry was detected by the anti-mCherry antibody. We could not detect the H2B-PA-mCherry by anti-H2B antibody, most likely because the expression level of H2B-PA-mCherry is estimated at less than ~5% of endogenous H2B. (B) The nuclei isolated from the HeLa cells expressing H2B-PA-mCherry were washed with indicated buffers including various concentrations of NaCl. The resultant nuclear pellets (left) and supernatants (right) were electrophoresed. The gel was subjected to CBB stain (upper) and western blot (bottom) with anti-mCherry and anti-H2B antibodies. Positions of histone H1 and core histones in CBB stain and H2B-PA-mCherry and H2B in western blot are indicated with arrows. Note that H2B and H2B-PAmCherry started to dissociate from chromatin with 1 M NaCl and went to supernatant fraction, suggesting that H2B-PA-mCherry was incorporated into nucleosome structures as endogenous H2B.

4.3.2. Live PALM imaging of chromatin structure

From the movie data, which contains information for nucleosome position in a thin optical layer, we extracted the position information and mapped them to obtain live PALM image. The number of mapped nucleosomes are around 80,000/nucleus in a thin layer (~1,300/mm^2). This roughly corresponds to about 5% of the total nucleosomes in the region.

The resultant live PALM image showed that the nucleosomes seem to be highly clustered (Figure 4.3 and 4.4 A), reminiscent of the chromatin domains including topologically associating domains (TADs) identified by 5C or Hi-C related method (Dixon et al., 2012; Nora et al., 2012; Rao et al., 2014; Sexton et al., 2012). Nuclear periphery and nucleoli edges, which are heterochromatic regions, appear as dense nucleosome clusters (Figure 4.4 B). This characteristic was also supported by a correlative Hoechst (DNA) staining of the same cells after the live PALM imaging (Figure 4.3). To quantitatively verify the nucleosome clustering, we first performed the radial distribution function (RDF) analysis, which describes how nucleosome density varies as a function of distance from a reference point (Figure 4.5 and 4.6 A) (Bohn et al., 2010). RDF of a random distribution is nearly equal to 1 and clustering tendency gives higher than 1. Note that the plots of chromatin are higher than 1 in the wide range from 0 to ~250 nm (control in Figure 4.6 A). This analysis demonstrated that the nucleosome distribution is not at random and indeed forms many compact chromatin domains in the living cells (control in Figure 4.6 A).

To further investigate the property of chromatin domains, we used L-function [L(r)-r versus r plot; r, correlation distance], which is a similar statistics to RDF to detect clustering, but can visualize clustered domains in a more quantitative manner (Figure 4.5 and 4.6 B) (Kiskowski et al., 2009). A control plot in Figure 4.6 B shows a curve with a peak at ~115 nm (i.e. ~230 nm diameter) while the plot of the random distribution model with the same density dots as the PALM image (Figure 4.5 A) is almost zero (dot line in Figure 4.6 B). This calculated domain size is similar to that of the domains recently identified by high resolution Hi-C method (~185 kb, ~181 nm diameter) (Rao et al., 2014), suggesting that both methods observed similar domain structures.

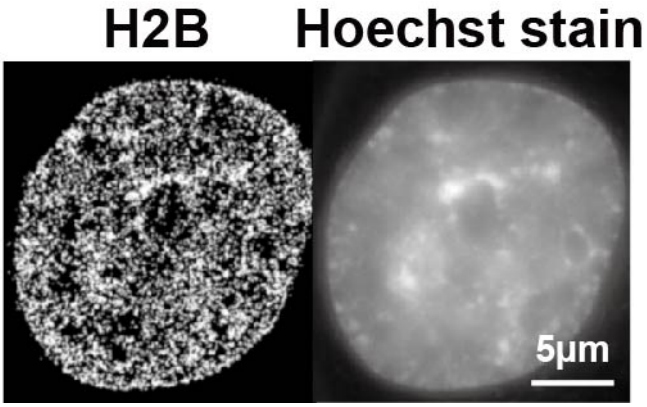

Figure 4.3 PALM image of histone H2B and Hoechst staining of DNA in a living cell

PALM image of histone H2B (left) and correlative Hoechst 33342 DNA staining of the same cell (right). Scale bar shows 5 µm.

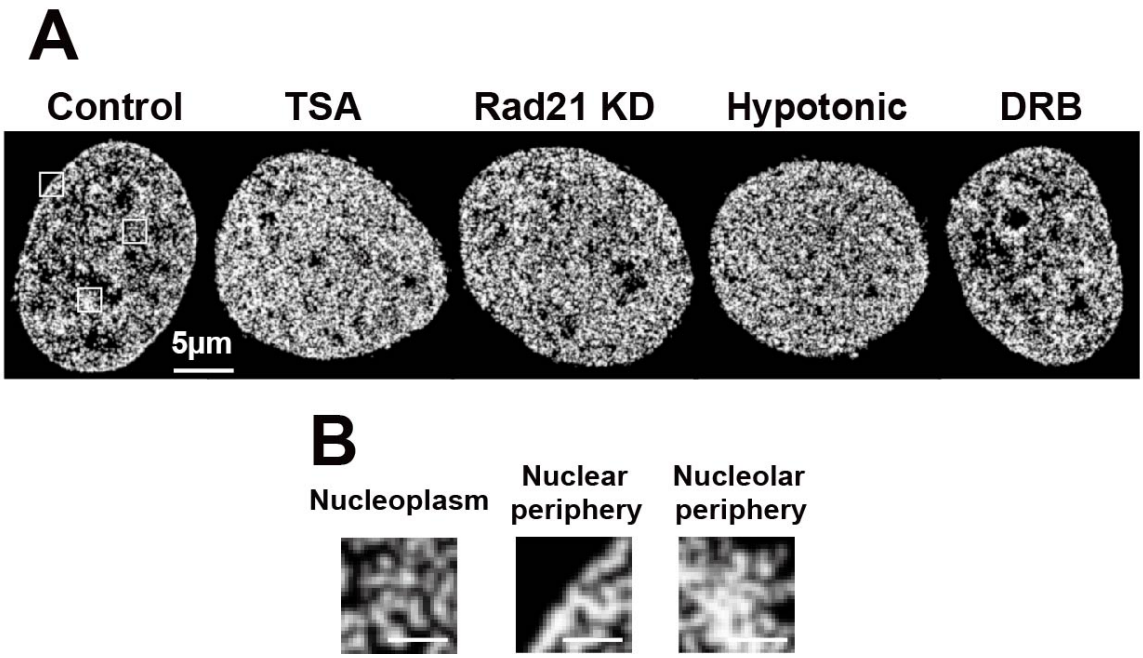

Figure 4.4 PALM images of interphase chromatin based on H2B-PA-mCherry

(A) PALM images of interphase chromatin based on H2B-PA-mCherry in living cells: from left to right, a control HeLa cell, TSA-treated cell, Rad21-KD cell, hypotonic treated cell and DRB treated cell. Scale bar shows 5 μm. (B) Magnified images from boxed regions in the control PALM image in (A): nucleoplasm region, nuclear periphery, and nucleolar periphery. Scale bar shows 1 μm.

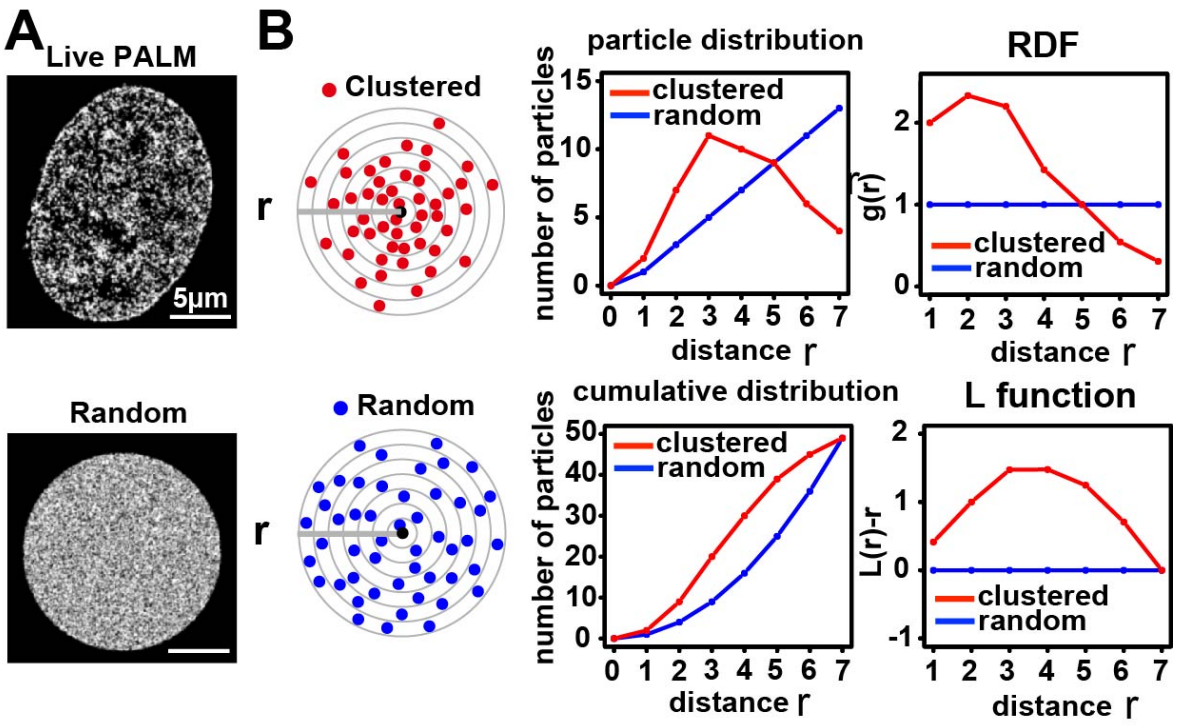

Figure 4.5 Description of RDF and L-function

(A) The examples of actual clustered distribution (top) and random distribution (bottom) of H2B are shown. Random distribution of dots was artificially generated at the same density to actually observed H2B in the nucleus. (B) A simplified scheme for RDF and L-function analyses. Total 49 particles were set in a clustered (red spheres, left top) or random manner (blue spheres, left bottom) around the origin point (black circle). Particles in the shell (middle top) and cumulative particles in the circle (middle bottom) were counted. The RDF (right top) and L-function (right bottom) plots are essentially obtained by taking the ratio between actual density and random density from particles in the shell and cumulative particles in the circle respectively. Horizontal axis represents correlation distance (r). RDF and L-function plots of a random pattern (blue) are nearly equal to 1 and 0, respectively.

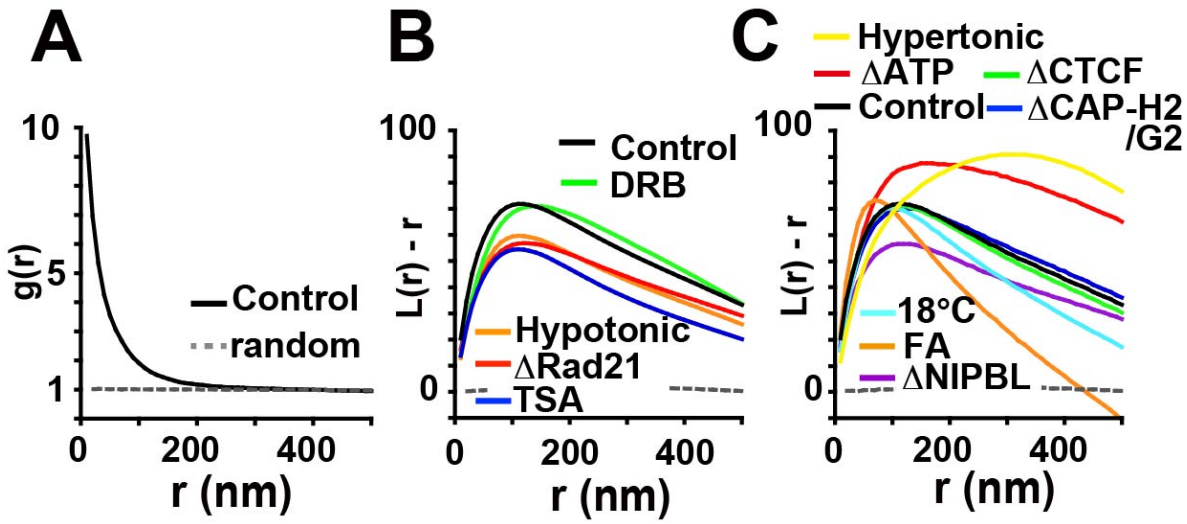

Figure 4.6 RDF plots and L-function plots of interphase chromatin

(A) RDF plots of interphase chromatin (black) are higher than 1 in the wide range from 0 to ~250 nm while the random distribution plots (gray dot line) show nearly equal to 1 in the same range, indicating that chromatin indeed forms compact domains. n = 75 cells. (B) L-function plots of chromatin with the same conditions as panel (Figure 4.4 A). The cells with TSA treatment (blue), Rad21-KD (red), and hypotonic treatment (orange) declined L-function plots, compared to the control cells (black), indicating decondensation of chromatin domains. In contrast, L-function plot of DRB treated cells (green) is similar to control. For each condition, n = 25-75 cells. (C) L-function plots of chromatin in hypertonically treated cells (yellow), ATP-depleted cells (red), CAP-H2/G2-KD cells (blue), CTCF-KD cells (green), cells cultured in room temperature (18°C) (cyan), formaldehyde-fixed cells (orange), and NIPBL-KD cells (purple). For each condition, n = 25-30 cells.

4.3.3. Chromatin domains are organized by nucleosome-nucleosome interactions and cohesin

We then wondered what kind of factors are involved in the domain formation. We first examined the effect of histone deacetylase (HDAC) inhibitor tricostatin A (TSA) on the domain formation (Gorisch et al., 2005; Yoshida et al., 1990). TSA treatment increases histone-tail acetylation including Lysine 16 of Histone H4 (H4K16) (Figure 4.7 A), which can inhibit the nucleosome-nucleosome interaction by preventing binding of the H4 tail to the acidic patch of the neighbor nucleosomes (Kalashnikova et al., 2013). TSA treatment led to domain decondensation in the cells (Figure 4.4 A and 4.6 B), suggesting that nucleosome-nucleosome interactions are required for chromatin domain formation. DNA staining by DAPI showed decondensed chromatin texture (Gorisch et al., 2005) and good agreement with the L-function analysis (Figure 4.7 B). This finding also implies that the chromatin domain organization can be controlled by histone modifications.

We next investigated if cohesin complex is involved in the formation of chromatin domains. Cohesin can capture chromatin fibers in its ring structure, and form loops and subsequent higher-order chromatin structure (Nasmyth and Haering, 2005; Shintomi and Hirano, 2010). Knock-down (KD) of a cohesin subunit Rad21 after 60 hours siRNA (Wendt et al., 2008) decondensed the domains (Figure 4.4 A and 4.6 B). The reduction of Rad21 protein was confirmed by immunostaining with anti-Rad21 antibody (Figure 4.8 A). KD of a cohesin loader NIPBL by siRNA (Zuin et al., 2014b) also decondensed the domains with a NIPBL-reduction Figure 4.6 C, 4.8 B, and 4.10), again supporting a critical cohesin function in the chromatin domain organization. In addition, double treatments of TSA and cohesin-KD further decondensed the domains (Figure 4.9 A and B), suggesting their coordination in the process. On the other hand, KD of CCCTC-binding factor (CTCF) (Wendt et al., 2008), which is also involved in loop formation, partly together with cohesin, did not change the L-function plots (Figure 4.6 C and 4.10) although the CTCF protein was not detectable (Figure 4.8 C). This finding implies their functional difference. Double KD of CAP-H2 and G2 in the condensin II complex, which locates in the interphase nuclei and functions in sister chromatid resolution during S phase (Ono et al., 2013), almost completely suppressed the protein level (Figure 4.8 D), but did not alter domain formation as cohesin (Figure 4.6 C and 4.10). Taken together, our results demonstrate the key role of cohesin in the chromatin domain organization (Figure 4.24), in agreement with previous reports (Mizuguchi et al., 2014; Sofueva et al., 2013; Zuin et al., 2014a).

Since macromolecular crowding force and cations play an important role in chromatin organization in the cells, we examined them by changing osmotic pressure. Hypoosmotic conditions with diluted medium (~140 mOsm, instead of normal condition, ~290 mOsm) (Albiez et al., 2006), which can lower the cations and crowding effect, decondensed the domains (Figure 4.4 A and 4.6 B): the molecular crowding force and cations can contribute to the domain formation. On the other hand, as expected, a hypertonic treatment (~570 mOsm) had an opposite effect: chromatin hyper-condensation (Figure 4.6 C and 4.10).

Furthermore, to test the role of the transcriptional process in the domain formation, we treated the cells with 6-Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB), which is a selective inhibitor of transcription elongation by RNA polymerase II in eukaryotic cells and dissociates the transcriptional complexes (Kimura et al., 2002). Although the treatment severely suppressed the global RNA synthesis in the cells (Figure 4.8 E), it was not effective in the domain formation (Figure 4.4 A and 4.6 B). This suggests that transcription and its complexes are not directly involved in the domain formation. Consistent with this finding, a correlative immunostaining of the same cell after the live PALM imaging with active RNA Pol II marker anti-phosphorylated serine 5 antibody revealed that the active RNA Pol II clusters often locate outside the chromatin domains (Figure 4.11 and 4.24) in accordance with some previous reports (Markaki et al., 2010; Niedojadlo et al., 2011).

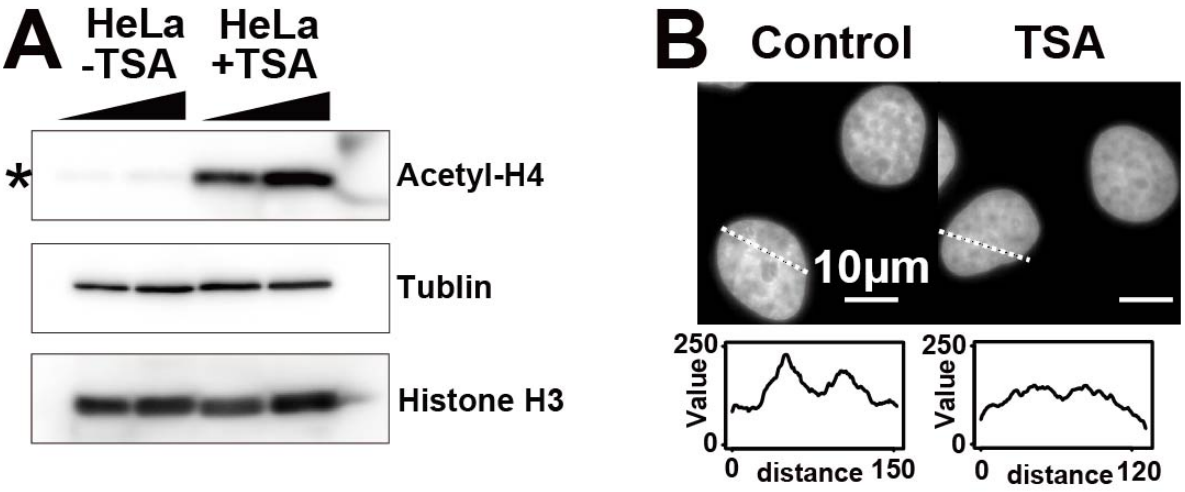

Figure 4.7 Western blotting and imaging of TSA-treated cells

(A) Detection of histone H4 acetylation in the cell lysates for TSA-treated cells (right) and untreated cells (left) by western blotting with anti-acetylated histone H4 tail antibody including acetylated K16. Note the specific increase in histone H4 acetylation in the TSA-treated cells. As controls, blotting results using anti-tublin and anti-histone H3 antibodies are shown. (B) TSA-treated cells (right) and untreated cells (left) by Hoechst DNA staining are shown. The corresponding scan plots to the dot lines are also shown at the bottom. Note the chromatin decondensation apparent in the TSA-treated cells.

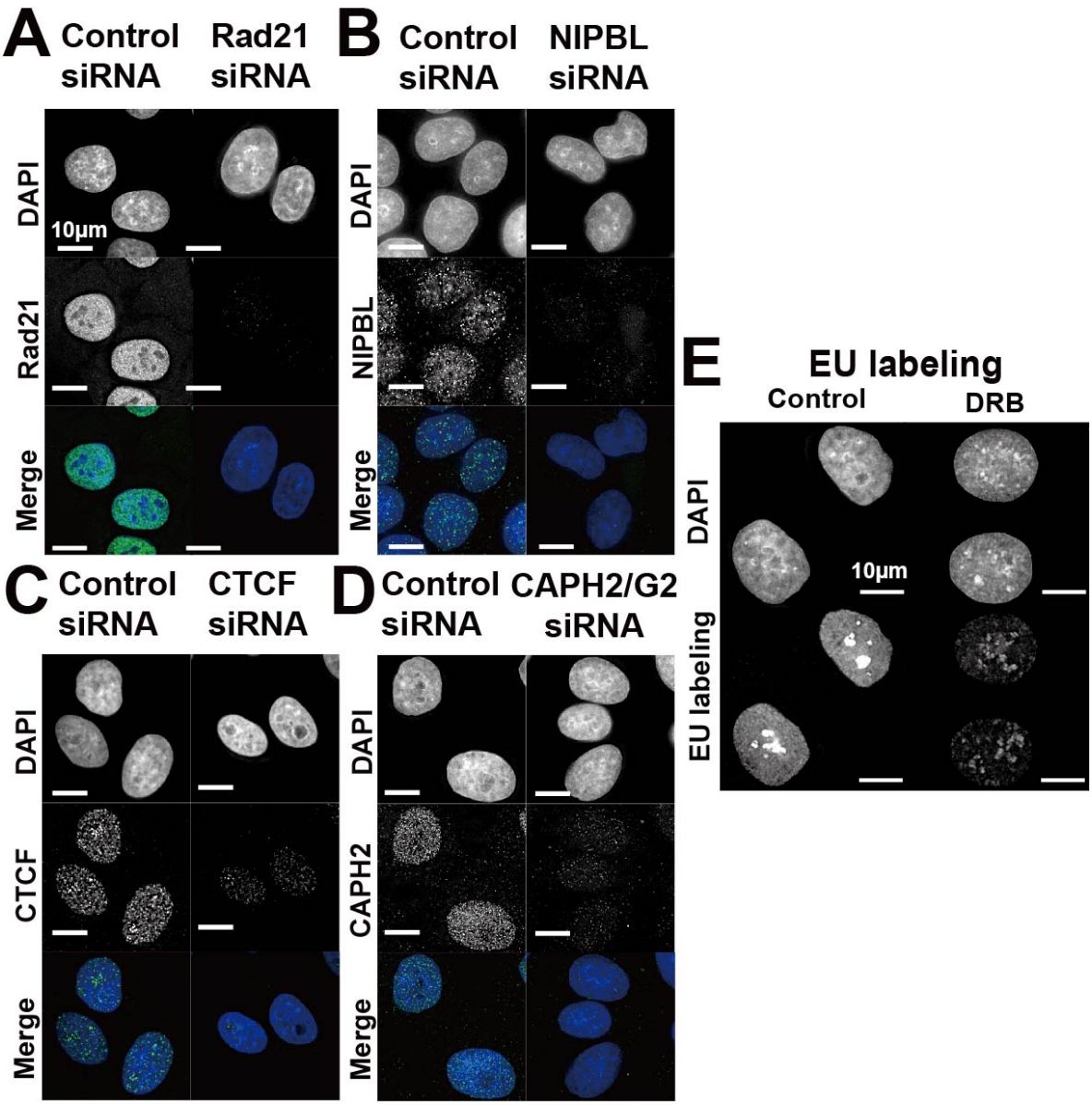

Figure 4.8 Immunostaining of protein knock down or drug treated cells

(A-D) Verification of KD for the indicated chromatin associated proteins by immunostaining: A, Rad21; B, NIPBL; C, CTCF; D, CAP-H2/G2. The images show significant reductions of targeted proteins, compared to the control siRNA cells. Scale bar shows 10 µm. (E) EU labeling of newly synthesized RNA in control and DRB-treated cells. DRB-treatment drastically decreased the efficiency of EU labeling.

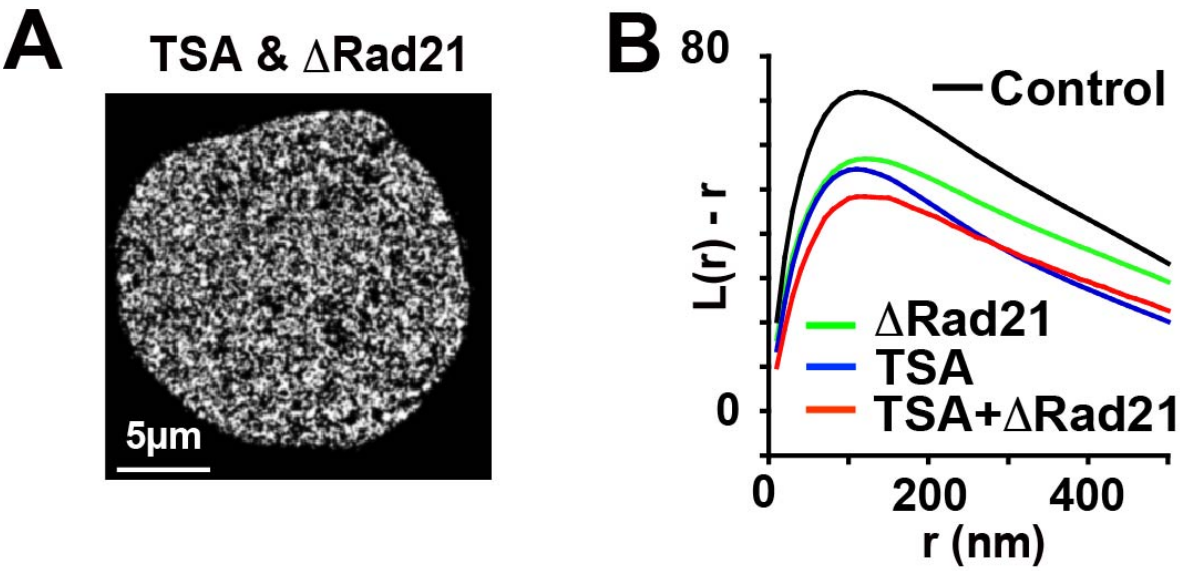

Figure 4.9 PALM images of interphase chromatin in the cell with both TSA treatment and Rad21-KD

(A) PALM images of interphase chromatin in the cell with both TSA treatment and Rad21-KD. Scale bar shows 5 μm. (B) L-function plots of chromatin in the cells (red, n = 20 cells). Double treatment of TSA and Rad21-KD further declines L-function plots and domain decondensation, compared to TSA treatment (blue) or Rad21-KD (green) alone.

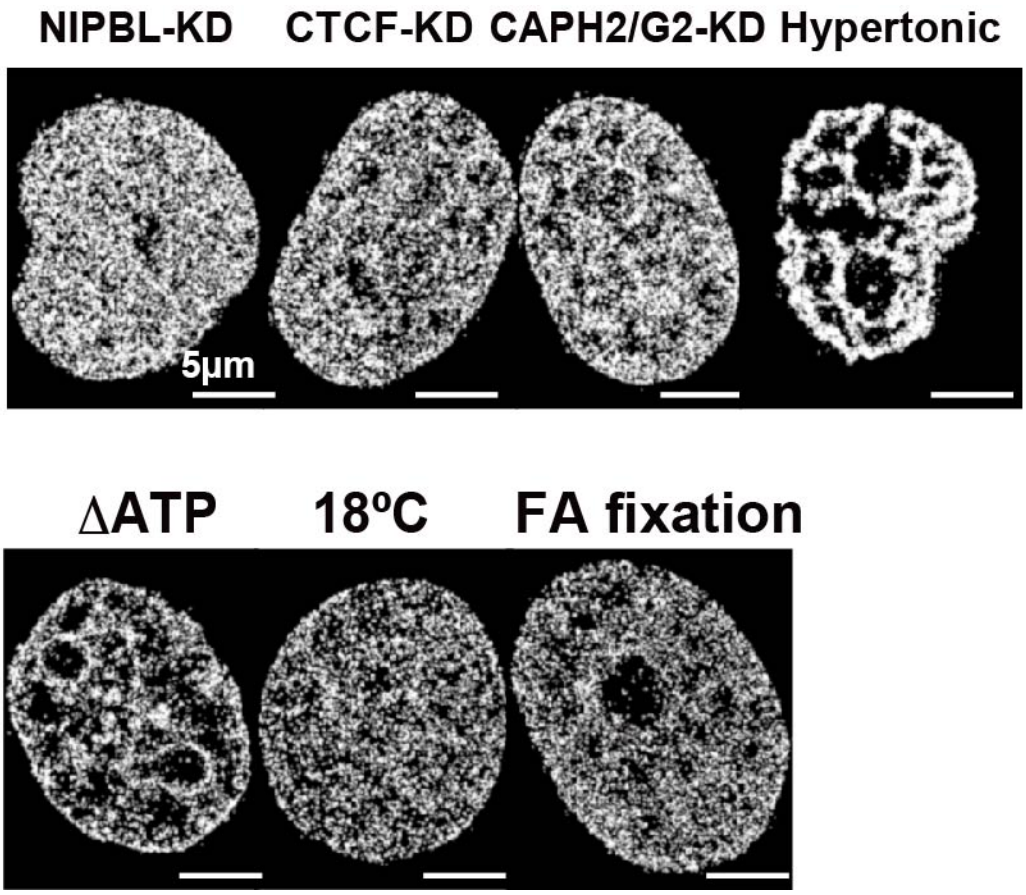

Figure 4.10 PALM images of chromatin in various conditions

PALM images of chromatin in NIPBL-KD cell, CTCF-KD cell, CAP-H2/G2-KD cell, hypertonically treated cell, ATP-depleted cell, cell cultured in room temperature (18°C), and formaldehyde fixed cell. Scale bar shows 5 μm.

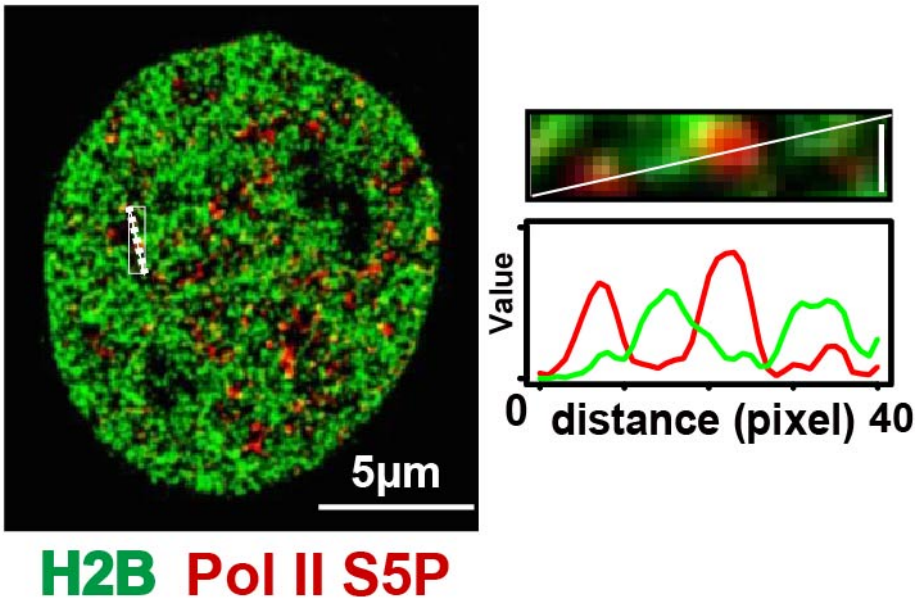

Figure 4.11 Correlative immunostaining with active RNA Pol II after PALM imaging

Correlative immunostaining with active RNA Pol II marker, anti-RNA Pol II phospho-Ser5 antibody, on the same cell after PALM imaging. (left) Overlayed images of PALM (green) and Pol II phospho-Ser5 staining (red). Active RNA Pol II clusters seem to localize outside the chromatin domains. Scale bar shows 5 μm. (right) Magnified image from the boxed region in the left image. Scale bar shows 500 nm. Mutually exclusive pattern of chromatin (green) and active Pol II (red) on the white line is also indicated by intensity line-scan.

4.3.4. Chromatin domains were observed throughout the cell cycle

To investigate the behavior of the chromatin domains during cell cycle, a time course experiment for PALM imaging was performed using the synchronized cells that were released from mitotic phase (Figure 4.12 A and B). We observed similar chromatin domains from G1, S, to G2 phases, which are in a good agreement with the finding on TADs by Hi-C technique (Naumova et al., 2013). However, differently from the case of TADs (Naumova et al., 2013), we found that mitotic chromosomes still retain chromatin domain structures: The PALM image demonstrated highly clustered nucleosome signals on whole mitotic chromosomes (Figure 4.13 A). L-function plot shows that the mitotic chromosomes have a notable peak (Figure 4.13 B). The RDF plot is also higher than 1 (Figure 4.13 C). The both plots indicate existence of chromatin domain structures (Figure 4.13 B and C). This observation became more prominent when used Indian Muntjac DM cells, which have large chromosomes and are advantageous for this analysis (Figure 4.14 A-C) (Hihara et al., 2012). Taken together, our results suggest that the chromatin domains are retained throughout cell cycle (Figure 4.24).

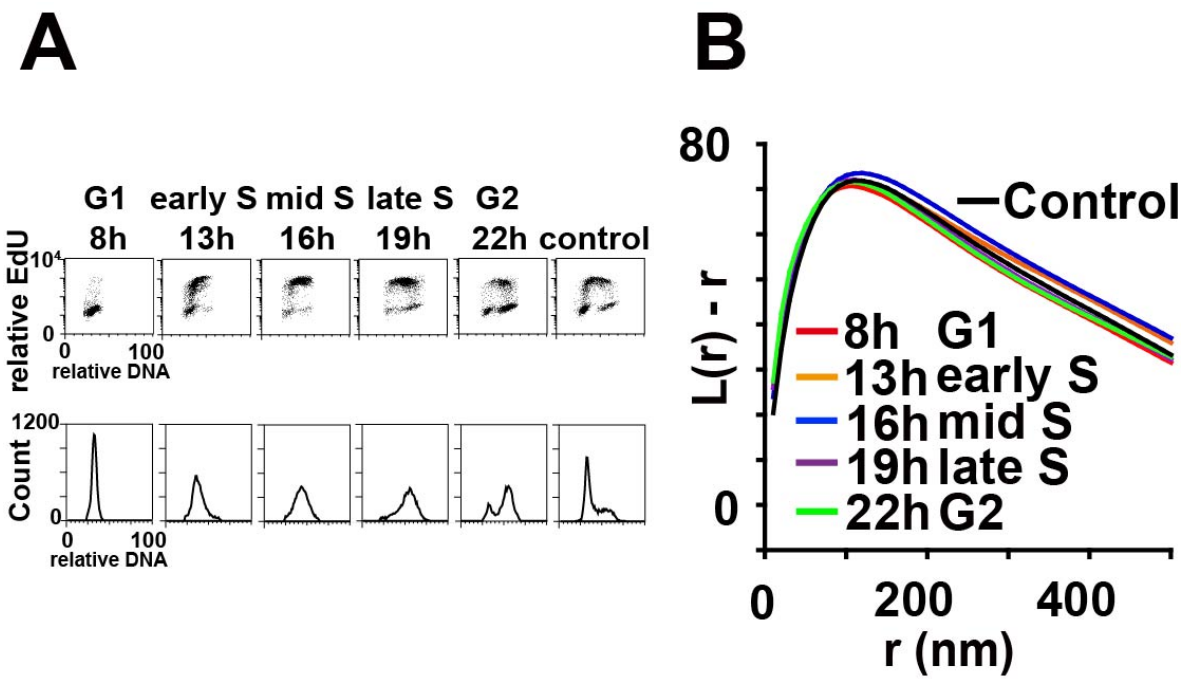

Figure 4.12 Chromatin structure during cell cycle

(A) To obtain the synchronized cells, mitotic cells were collected by shake off using nocodazole and released. Each panel shows FACS profiles of synchronized cells at a certain time point after the release. To verify the cell cycle stages, labeling intensity with BrdU to visualize DNA replication (vertical axis) and Hoechst DNA staining (horizontal axis) to determine DNA content are shown. (B) L-function plots of chromatin in various cell-cycle stages (8 hours after mitotic release, red; 13 hours, orange; 16 hours, blue; 19 hours, purple; 22 hours; control, black). L-function shows that chromatin structure does not significantly change throughout interphase.

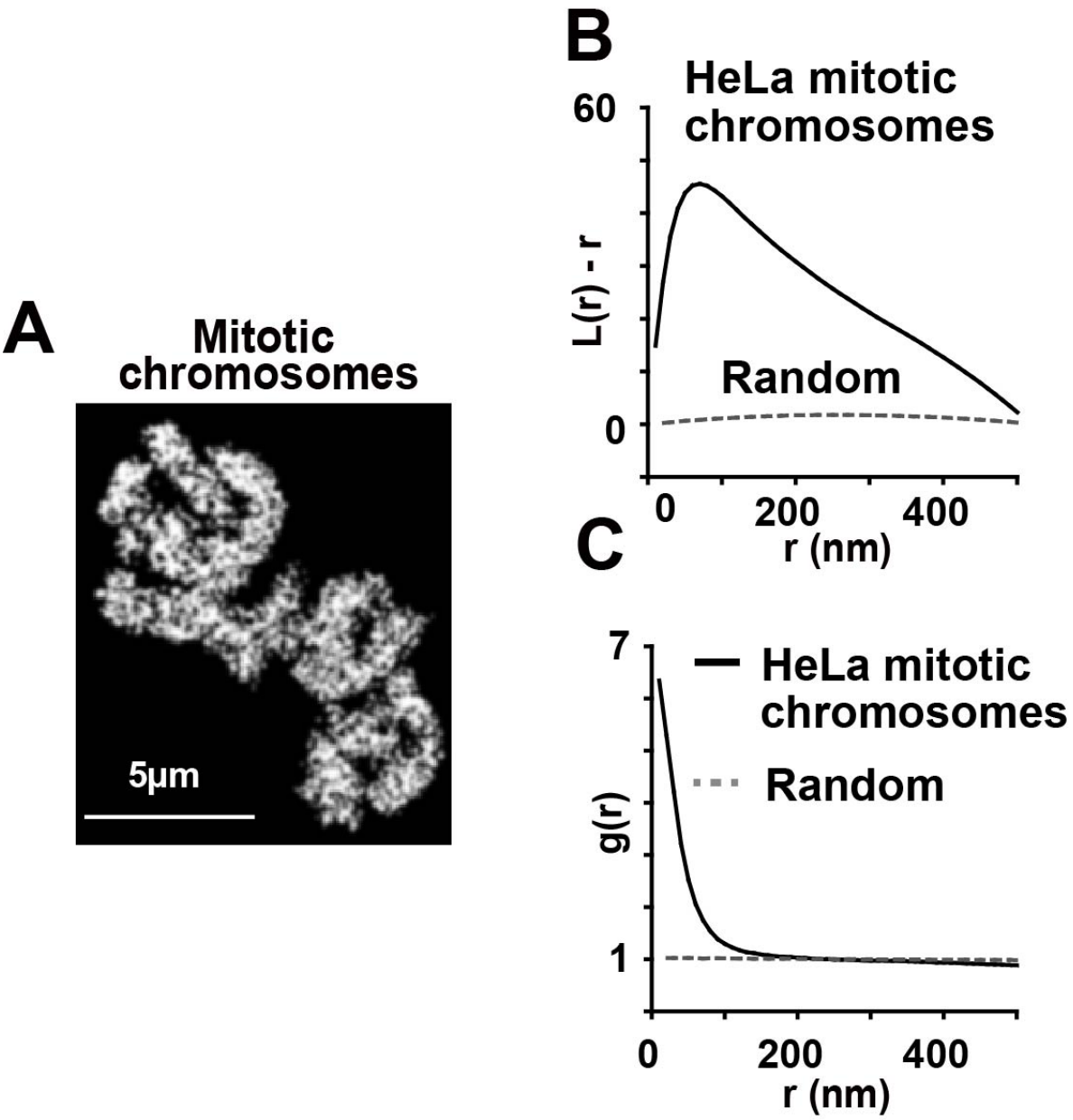

Figure 4.13 PALM image of mitotic chromosomes in a HeLa cell

(A) PALM image of mitotic chromosomes in a HeLa cell. Scale bar shows 5 μm. The domain structures are notable. (B) L-function plots of chromatin in mitotic chromosomes (n = 20 cells), suggesting that mitotic chromosomes have similar chromatin domains to those of interphase chromatin. (C) RDF plots for mitotic HeLa cells demonstrate that the nucleosomes indeed form compact chromatin domains even during mitosis.

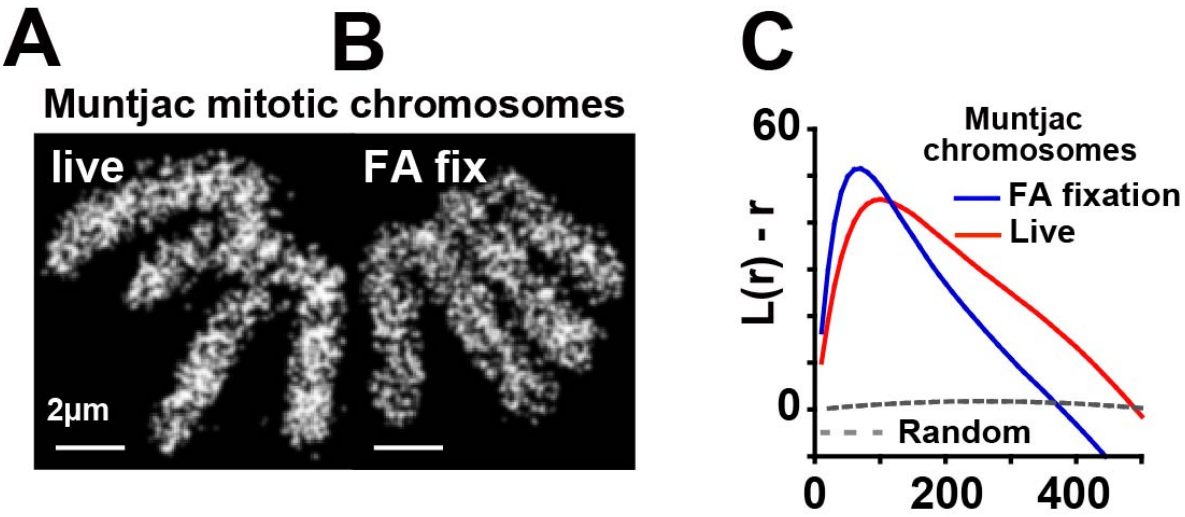

Figure 4.14 PALM images of mitotic chromosomes in a DM cell

PALM images of mitotic chromosomes in living (A) and formaldehyde fixed (B) DM cells. Scale bar shows 2 μm. (C) L-function plots for mitotic DM cells demonstrate that the nucleosomes indeed form compact chromatin domains during mitosis. In (C), for living and fixed cells analyzed, n = 11 cells and 16 cells, respectively.

4.3.5. Chromatin dynamics in a living cell

Since our movie data (20 frames/sec) also contains information for nucleosome movements in a thin optical layer, we next pursued dynamics of the chromatin domains based on movements of individual nucleosomes (Figure 4.15 A). We first tracked each nucleosome movement from 0 to ~0.5 sec (~11 frames) by a tracking software u-track (Figure 4.15 B) (Jaqaman et al., 2008). About 60 nm of nucleosome movement for 50 msec was observed (Figure 4.15 C), consistent with the result of Chapter 3. The plots of calculated mean square displacement (MSD) were well fitted to an anomalous diffusion model (control in Figure 4.16 A). Chemical fixation of the cells with disuccinimidyl glutarate (DSG) and formaldehyde (FA) to crosslink nucleosomes severely suppressed the movements (Figure 4.16 A), demonstrating that majority of the observed movement derived from real nucleosome movements in living cells.

We then asked whether the observed movement of the individual nucleosomes is related to the domain dynamics or not. To address this question, we measured the movements of DNA replication foci, which were composed of a megabase-sized genomic DNA labeled by Cy3-dCTP (Albiez et al., 2006) and reported to cover a large area of TADs (Figure 4.17) (Pope et al., 2014), and compared them to the nucleosome movements (Figure 4.16 B). Interestingly the MSD plot of the nucleosome movement is similar to that of DNA replication foci, suggesting that a substantial fraction of the observed nucleosome movements could be derived from those of chromatin domains to which the nucleosomes belong (Figure 4.16 B). Indeed using a dual color labeling and imaging of the nucleosomes and domains (Figure 4.18 A), we found that some of them move similarly (Figure 4.18 B and C), suggesting a correlative movement between the nucleosome and domain. We thus approximately used the nucleosome movement as a proxy for the domain dynamics.

In addition to MSD analysis, which calculates ensemble average of the domain movement, we also integrated the movement data on a 2D plane to visualize magnitude of chromatin domain dynamics as a 2D "heat map": more domain movement is in more red (or "hot") pixels and less movement in more blue (or "cold") (Figure 4.19 A). This heat map provides a spatial domain dynamics in a whole nucleus of the living cell (Figure 4.19 B). On the heat map, the nuclear periphery and nucleoli edges, which are presumably heterochromatin-rich regions, showed more blue, suggesting that the domain dynamics is suppressed in the heterochromatic region (Figure 4.19 C). The inner nucleoplasmic regions with less heterochromatin appear more red (Figure 4.19 C).

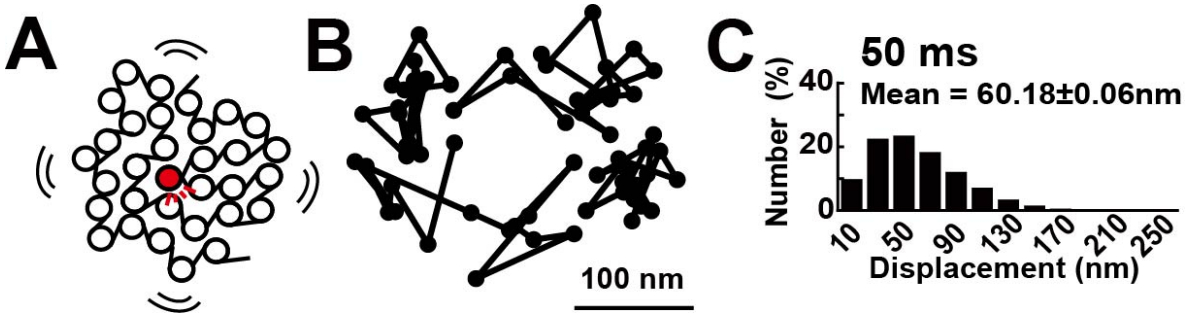

Figure 4.15 Scheme for visualization of chromatin domain dynamics

(A) Scheme for visualization of chromatin domain dynamics based on the single nucleosome tracking. Following the movement of chromatin domain, nucleosomes move around. (B) Representative tracked trajectories of single nucleosomes. Scale bar shows 100 nm. (C) Displacement (movement) distributions of single nucleosomes in interphase chromatin of living HeLa cells for 50 ms. n = 75 cells.

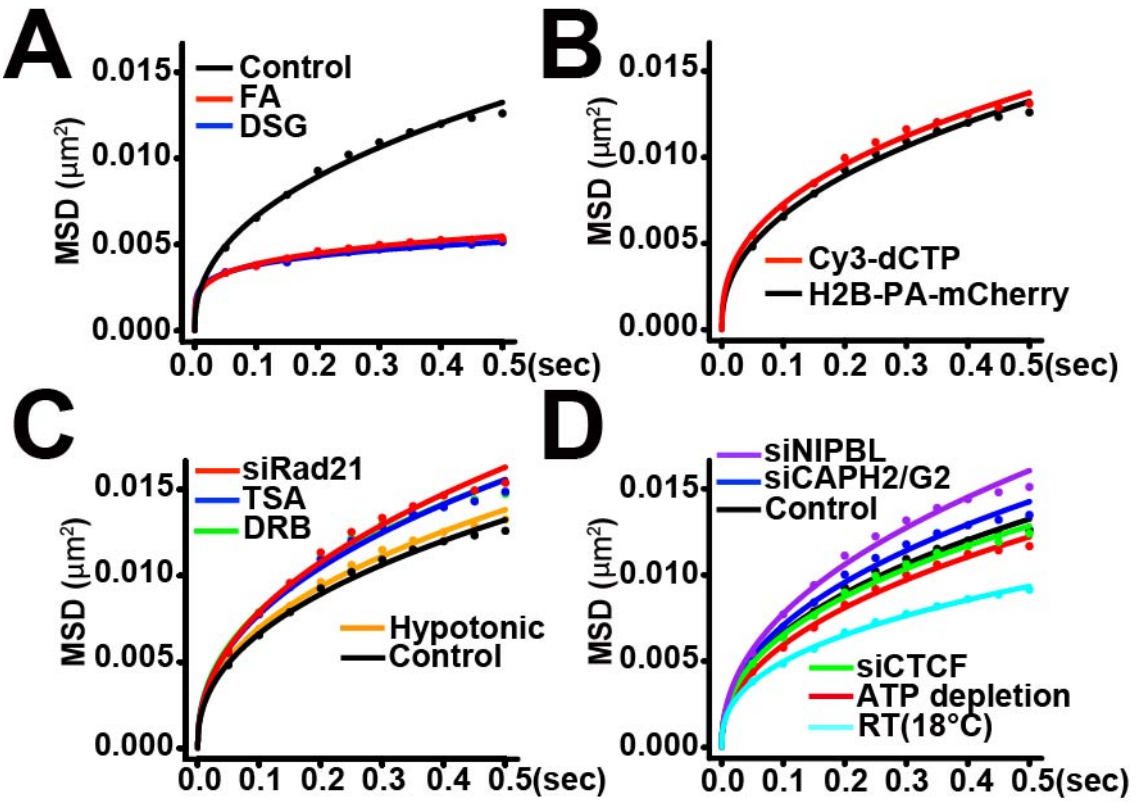

Figure 4.16 Mean square displacements plots of single nucleosomes in interphase chromatin of living HeLa cells

(A) Mean square displacements (MSDs) plots of single nucleosomes in interphase chromatin of living HeLa cells (black), formaldehyde fixed (red), and DSG fixed cells (blue) from 0 to 0.5 s. The plots for single nucleosomes were fitted to an anomalous diffusion model. For each sample, n = 25-75 cells. In the MSD analyses for single nucleosomes, the originally calculated MSD was in two dimensions. To obtain their 3D values, the original values of MSD were multiplied by 1.5 ($6Dt/4Dt$). (B) MSD plots of DNA replication foci labeled by Cy3-dCTP (red, n = 30 cells) compared to those of H2B-PA-mCherry (black, n = 75 cells) from 0 to 0.5 s. Nucleosome movement seems to be similar to that of DNA replication foci, and a substantial fraction of the nucleosome dynamics could be derived from those of chromatin domains to which the nucleosomes belong. (C) MSD plots of the domains in Rad21-KD cells (blue), TSA treated cells (red), DRB treated cells (green), Hypotonic treated cells (orange), and control cells (black) from 0 to 0.5 s. For each condition, n = 20-75 cells. (D) MSD plots for NIPBL-KD cells (purple), in CAP-H2/G2-KD cells (blue), CTCF-KD cells (green), ATP-depleted cells (red), cells cultured in room temperature (18°C) (cyan), and control cells (black) from 0 to 0.5 s.

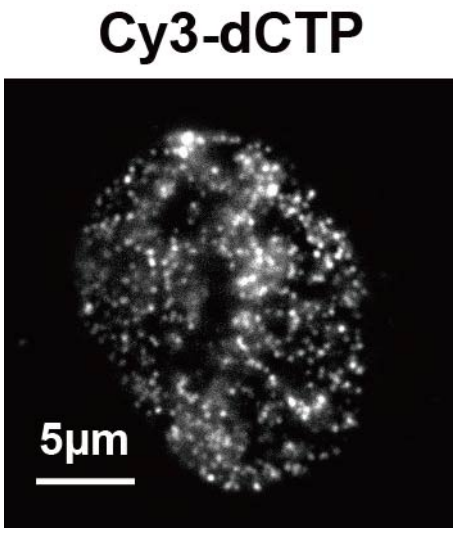

Figure 4.17 Imaging of DNA replication foci

Image of DNA replication foci labeled by Cy3-dCTP in a living HeLa nucleus. Scale bar shows 5 µm.

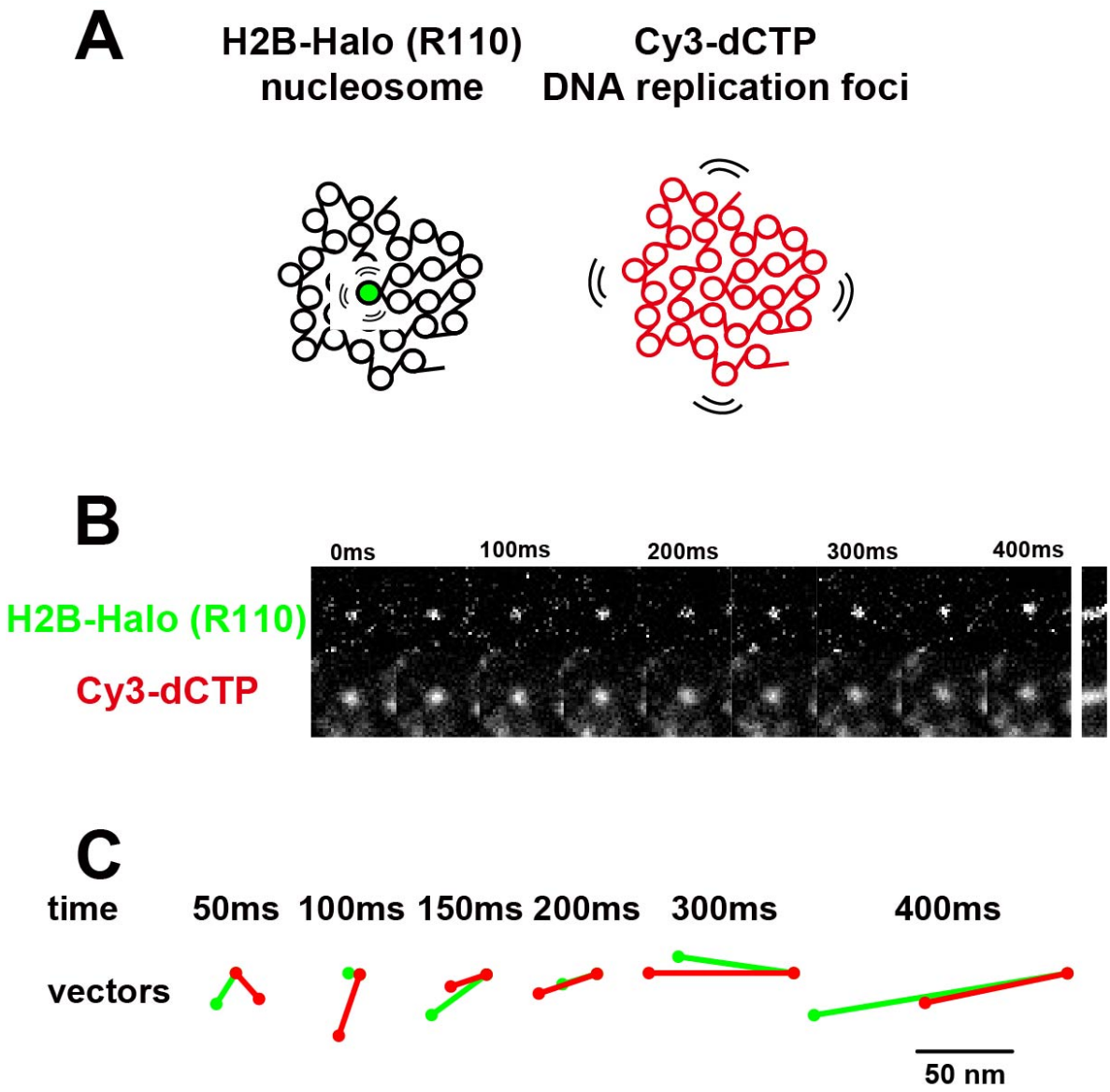

Figure 4.18 Scheme for a dual color labeling and imaging of the nucleosomes and DNA replication domain

(A) Scheme for a dual color labeling and imaging of the nucleosomes with H2B-Halo (R110) and Cy3-incorporated DNA replication domains. For this, the cells expressing H2B-Halo was used and labeled with 1 nM of R110 fluorescent dye. A dual color imaging was performed through W-VIEW GEMINI (Hamamatsu Photonics). (B) An example that nucleosome and replication domains move similarly. (C) A correlative movement between the nucleosome and domain.

4.3.6. Chromatin dynamics are organized by nucleosome-nucleosome interactions and cohesin

To test functional relevance of these domain dynamics, we performed a series of perturbation experiments. KD of a cohesin subunit Rad21 increased the dynamics in MSD (Figure 4.16 C and 4.19 B), suggesting that release of the two sister chromatin and/or loops folded by cohesin facilitates the domain dynamics (Figure 4.20 and 4.24). Increased histone acetylation by TSA treatment, which can break nucleosome-nucleosome interactions and decondense the chromatin domains (Figure 4.4 A and 4.6 B), led to an upregulation of the dynamics (Figure 4.16 C and 4.19 B), presumably because decondensation of the chromatin domain makes chromatin more flexible and mobile (Figure 4.20). ATP-depletion of the cells also decreased the dynamics (Figure 4.16 D). By contrast, hypo-osmotic conditions with diluted medium (~140 mOsm), which can lower the cations and crowding effect, did not change the domain dynamics (Figure 4.16 C, 4.19 B, and 4.20).

Very interestingly, treatment of DRB to dissociate the transcription machinery increased the domain dynamics (Figure 4.16 C and 4.19 B) although it was not so effective in the domain formation (Figure 4.4 A and 4.6 B). Considering that the transcriptional complexes often locate outside the domains (Figure 4.11), this finding suggests that the domains are tethered to each other through the transcriptional machinery or factories (Figure 4.24). Dissociation of the transcriptional machinery by DRB treatment would lead to release of constraint of the domains and increase the domain dynamics (Figure 4.20). On the other hand, reduction of the temperature of cells from 37ºC to 18ºC caused a dramatic chromatin domain slow-down (Figure 4.16 D), although the domain organization did not change significantly (Figure 4.6 C). This finding is well consistent with the notion that the domain dynamics is partly driven by Brownian motion.

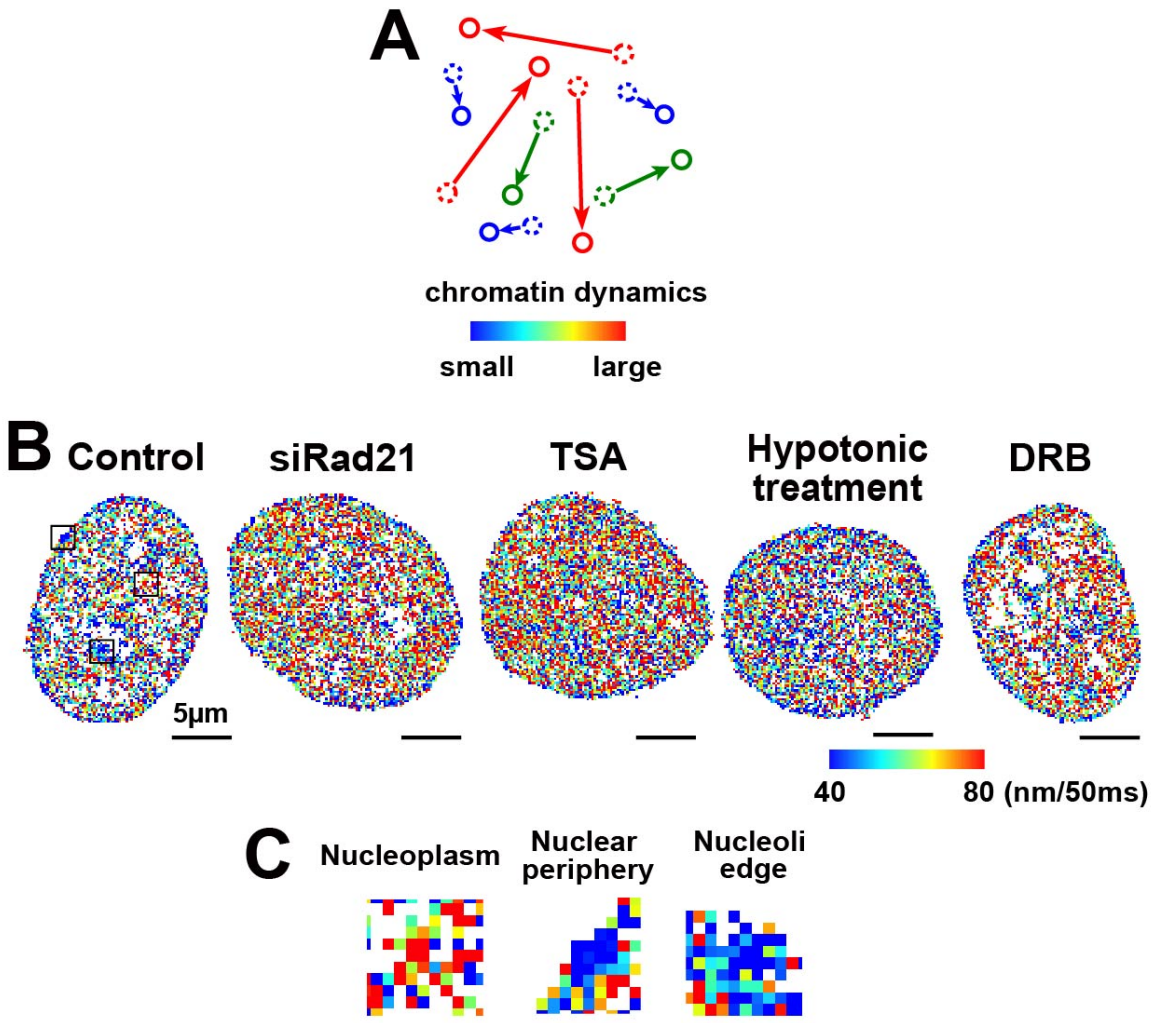

Figure 4.19 Heat maps of chromatin dynamics

(A) Schematic representation of chromatin heat map. The color of heat map represents the size of nucleosome movement. Small movements are colored in blue and large are in red. (B) Chromatin heat maps for 50 ms in a living HeLa cell (control), Rad21-KD cell, TSA treated cell, hypotonic treated cell, and DRB treated cell. Blue dots represent smaller movement and red dots represent larger movement. Color bar shows the scale of nucleosome movement in the heat map. Scale bar shows 5 μm. (C) Magnified images of the nucleoplasm, the nuclear periphery and nucleolar edge marked with boxes in (B). Domain dynamics near the nuclear periphery and the nucleoli edge are suppressed as compared to the nucleoplasm region.

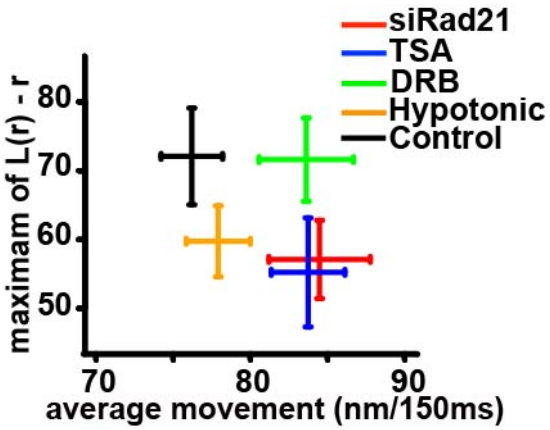

Figure 4.20 Plots of domain dynamics versus domain size

Plots of domain dynamics (movement/150 ms) versus domain size (maximum values in L-function): control cells, black; DRB treated cells, green; hypotonically treated cells, orange; Rad21-KD cells, red; TSA treated cells, blue. Data are mean ± S.D.

4.3.7. Property of chromatin of ES cells during differentiation

Since we so far focused on the HeLa cells expressing H2B-PA-mCherry, to apply the established imaging method to another type of cells, we established mouse embryonic stem cells (ESCs) expressing H2B-PA-mCherry. Interestingly, L-function plot in ESCs is a rather flat (Figure 4.21 A and B), suggesting that ESCs have more decondensed chromatin than HeLa cells, the domain structure might not be well defined and the dynamics is greater (Figure 4.21 A-C), which is consistent with previous reports (Mattout and Meshorer, 2010; Ricci et al., 2015). Chromatin heat map analysis showed that generally the domain dynamics seem upregulated in ESCs (Figure 4.21 A). On the other hand, in the nuclear periphery and the chromocenters (peri-centromeric heterochromatin), which were confirmed by the correlative immunostaining of histone H3K9me3, a heterochromatin marker, the domain dynamics slow down (Figure 4.22).

Since the rather ambiguous chromatin domains and hot chromatin property of the ESCs might be related to their pluripotency, we induced embryoid body cells by depletion of LIF the ESCs (Meshorer et al., 2006). The pluripotent marker Sox2 was not detectable in the cells (ESC(-)LIF in Figure 4.21 A). After the differentiation, the nuclei became larger (Figure 4.21 A) and L-function plots showed a sharper peak (ESC(-)LIF in Figure 4.21 B), indicating that the chromatin domain becomes more defined. MSD analysis also showed that the domain dynamics decrease with the heat map becoming more blue, suggesting a hot property of chromatin in pluripotent cells (Figure 4.21 A and C). So far chromatin behavior in ESC had been mainly studied in indirect ways such as FRAP experiment for histone H2B replacement (Mattout and Meshorer, 2010). As far as we know, this is first direct visualization of chromatin dynamics in a whole nucleus. Our finding strengthens the notion of chromatin plasticity in pluripotent cells and would provide a powerful tool to investigate genome-wide chromatin structure and dynamics during differentiation.

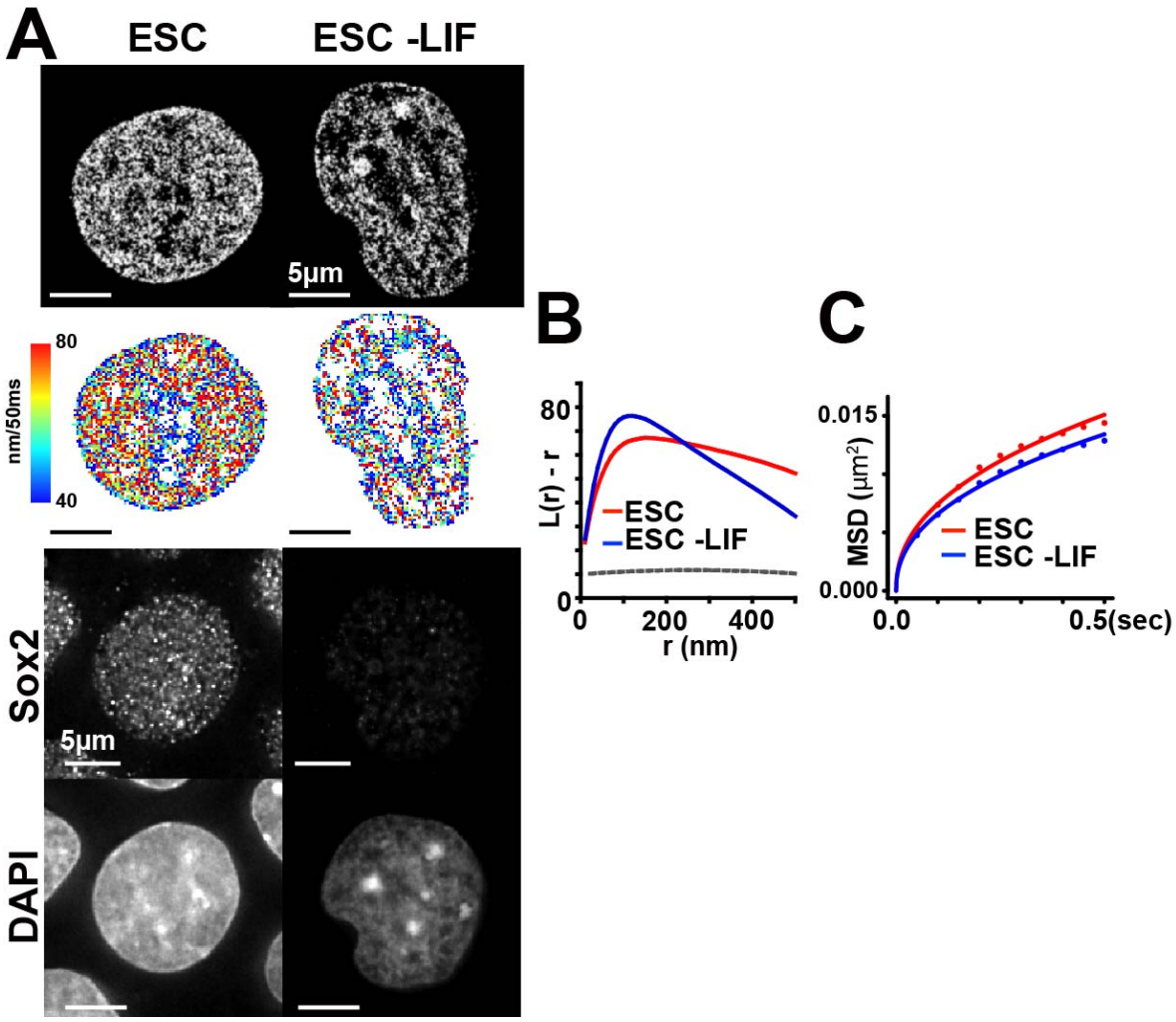

Figure 4.21 PALM images and chromatin heat maps for ESCs

(A) PALM images and chromatin heat maps for ESCs and ESCs cultured in medium without differentiation inhibitory factor LIF for 5 days. Immunostaining with anti-Sox2 antibody and DAPI staining are also shown. Scale bar shows 5 μm. (B) L-function plots of chromatin in ESCs (red) and ESCs 5 days after LIF withdrawal (blue). n = 35-40 cells. During culture with LIF-withdrawal, chromatin structure was condensed. (C) MSD plots of nucleosomes in ESCs (red) and ESCs 5 days after LIF withdraw (blue) from 0 to 0.5 s. Five days culture with LIF-withdrawal declines the domain dynamics compared to ESCs. n = 35-40 cells.

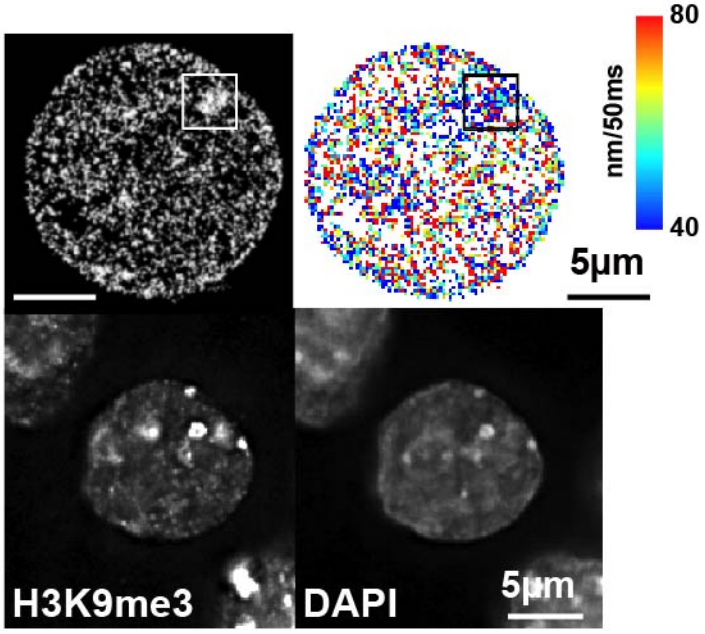

Figure 4.22 Live PALM image and chromatin heat map in chromocenter

Live PALM image (top left) and chromatin heat map (top right) in chromocenter (box) of an ESC, showing chromatin condensation and the decrease in domain dynamics in the region. The chromocenter region was confirmed by correlative immunostaining against anti-H3K9me3 antibody (bottom left) and DAPI staining (bottom right) after the PALM imaging. Scale bar shows 5 μm.

4.4. Discussion

In this study, we developed a novel nuclear imaging system, which allows us to integrate information for chromatin domains and dynamics in living mammalian cells. Subsequent correlative immunostaining of the same cell after the live PALM imaging provided further functional information at high resolution (Figure 4.11, 4.21 A, and 4.22).

Our study demonstrated that the formation of chromatin domains is orchestrated by both protein factor and physical forces: cohesin complex and nucleosome-nucleosome interactions, as well as cations (electrostatic force) and molecular crowding effect (Figure 4.4 A and 4.6 B). Interestingly upon transcription, the domains appear tethered by their machinery and become stable (Figure 4.24). This finding agrees with previous reports that a certain genomic loci became stable upon their transcription (Ochiai et al., 2015). The domain dynamics seems to be temperature-dependent, suggesting that the dynamics are essentially driven by Brownian motion. Temperature appeared the largest parameter affecting domain dynamics. On the other hand, although we observed that ATP-depletion decreased the domain dynamics, we cannot conclude energy-dependence of the movement because simultaneously chromatin condensation was also observed (Figure 4.6 C and 4.16 D), presumably due to the reported rapid rise in Ca^{2+} upon ATP-depletion (Martin et al., 2007). Systematic KDs of ATP-dependent chromatin proteins such as remodelers will provide a clue on this issue.

We found that the chromatin domain structures are retained throughout the cell-cycle including mitosis (Figure 4.24) although Hi-C analysis did not show notable TADs structures in human mitotic chromosomes. A possible reason for this discrepancy is the following: In the mitotic chromosomes, the nucleosome concentration is much higher (>0.5 mM) than that in interphase nuclei (~0.05 mM) (Hihara et al., 2012). This highly crowding environment generally produces higher contact probability in Hi-C and higher background, which might make it more difficult to identify the chromatin domains, especially by the Hi-C method. Indeed, our computational modeling study suggested that certain modeled TADs structures can diminish in the distance map when chromatin domains are highly compacted like mitotic chromosomes (Figure 4.23). In addition, it is known that the DNA replication foci with a megabase-sized genomic DNA have been also observed using pulse labeling (Schermelleh et al., 2001) and were retained stably during the cell cycle including mitosis and subsequent cell generations (Jackson and Pombo, 1998; Ma et al., 1998; Zink et al., 2003), which is in agreement with our finding.

We consider that retention of the chromatin domains throughout cell cycle has a couple of advantages for the genome functions. First, since the chromatin domains function as building

blocks of chromosomes, chromosome assembly and disassembly processes become smoother: The chromatin domains formed in the interphase are assembled together, presumably condensin, topoisomerase IIα and other factors, to create a rod-like chromosomal shape during mitosis (Figure 4.24). Second, since the chromatin domains are functional units of the genome, the memories of epigenetic marks in these units could be easily kept throughout cell cycle.

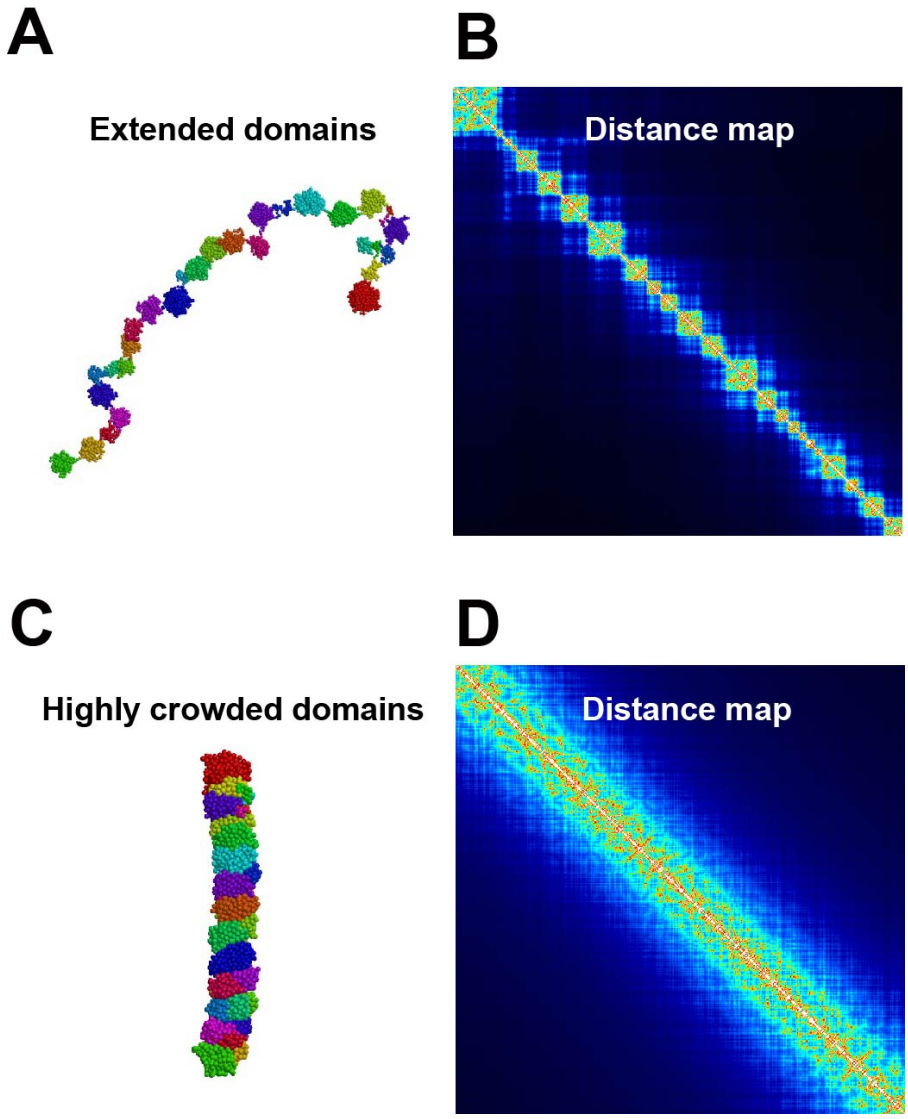

Figure 4.23 Computational simulation of chromatin domains

(A) A model for extended chromatin domains, corresponding to interphase chromatin domains. (B) Distance map shows clear domain structures on a diagonal line. (C) A model for highly crowded chromatin domains for mitotic chromosomes. (D) The distance map of these highly crowded chromatin domains shows no clear domain structures, suggesting that TADs structures can diminish in the distance map in case that they are highly compacted.

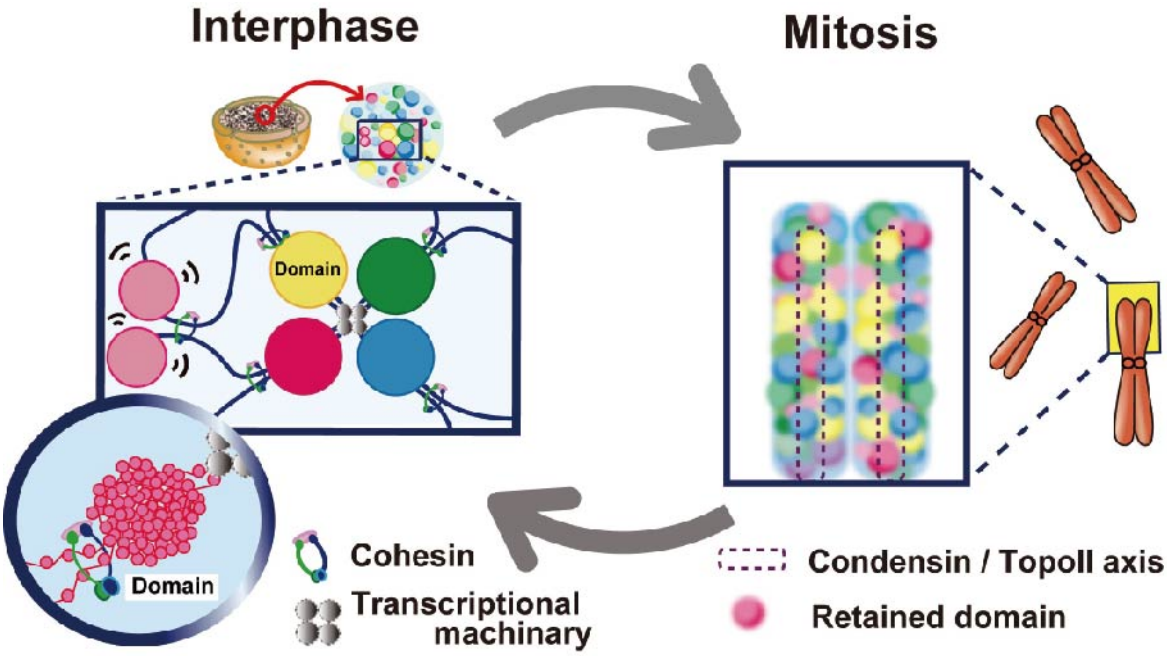

Figure 4.24 Various domains are formed by several factors including cohesin and nucleosome-nucleosome interactions

Summary Figure. In interphase (left), various domains are formed by several factors including cohesin and nucleosome-nucleosome interactions. Cohesin folds the domain itself (enlarged domain in the circle), possibly via loop formation, and also the sister domains (two pink domains in the square). For transcription, the chromatin domains are tethered by transcriptional machinery (gray spheres). The release of cohesin and transcriptional machinery increase domain dynamics. During mitosis (right), the chromatin domains are assembled together, presumably by condensin (and topoisomerase IIα) and other forces, to make a rod-like shape. The retained domains are visualized.

Chapter 5

Conclusions

5.1. Summary

In this thesis, I have mainly discussed chromatin and epigenetics. First, the epigenetic regulation and promoter types were analyzed based on genome-wide ChIP-Seq and DeepCAGE data. Secondly, a single nucleosome imaging method that can be implemented in a living nucleus was developed. Lastly, the chromatin structure and dynamics at the single nucleosome level and their variation related to different types of histone modifications and chromatin-associated proteins were successfully observed.

Currently, the tool "ChIP-Seq", which detects the histone modification patterns genome-widely, can be utilized to tackle the epigenetic phenomena. In the study in this thesis, a systematic analysis of transcription promoters and gene expression was performed. In addition, the epigenetic histone behaviors containing several types of histone modification and genomic positioning were also systematically analyzed. It was found that, in humans, the broad promoters, which initiate transcription from a wide-ranging region, had significant associations with histone modification and nucleosome position. However, this was not the case in peak promoters, which initiate transcription from a narrow genomic region. Specifically, the histones were highly distributed and aligned in an orderly fashion around broad promoters; this feature was more evident in the case of methylated or acetylated histones. Moreover, the nucleosome positions around the broad promoters were more stable than those around the peak promoters. The findings of nucleosome positioning and transcriptional factor binding support the strong connection between nucleosome-free region and accessibility of transcription factors, which is specific to broad promoters. More strikingly, the overall expression levels of genes associated with broad promoters (but not peak promoters) and modified histones were significantly higher than the levels of genes associated with broad promoters and unmodified histones. These results

shed light on how epigenetic regulatory networks of histone modifications are associated with promoter architecture (Chapter 2).

ChIP-Seq analysis is adequate for the correlation analysis of epigenetic regulation and gene expression but is insufficient to investigate chromatin environment and the process of epigenetic regulation. Therefore, there is a requirement for a tool that can directly observe the chromatin structure and dynamics. Thus, in the study in this thesis, a single nucleosome imaging method was developed to visualize chromatin structure and dynamics at single nucleosome level. The simultaneous activation of PA-GFP or PA-mCherry without 405-nm laser activation and the oblique illumination enabled us to observe the single molecules in the nucleus. The mean square displacement of nucleosomes showed abnormal diffusion, suggesting that the movement of nucleosome was restricted by DNA. Furthermore, the validity of single nucleosome imaging methods was carefully evaluated compared to EGFP or the movement of the nucleus. These results show the flexible and dynamic structure of chromatin in the nucleus (Chapter 3).

Although previous studies have proposed static chromatin nature, our results suggested the dynamic nature of chromatin. The newly developed imaging method of single nucleosomes in a living nucleus helped reconstruct a super-resolution image of chromatin that showed the clustered pattern of nucleosomes. The statistic analysis showed the existence of a chromatin domain but not its random distribution. This result supports the irregular folding model, and the average size of radius ~115 nm (i.e. ~230 nm diameter) is consistent with the previous X-ray scattering analysis (Joti et al., 2012). This calculated domain size is similar to that of the domains recently identified by high-resolution Hi-C method (~185 kb, ~181 nm diameter) (Rao et al., 2014). This suggested that both methods observed similar domain structures. Since TADs or chromatin domains are functional units of the genome, keeping these units throughout the cell cycle will prove to be efficient for the cells. This point is also the case for the memory of epigenetic marks on the chromatin domains during the cell cycle. Subsequently, our results propose a novel mechanistic insight into mitotic chromosome assembly: the chromatin domains formed in the interphase are assembled to create a rod-like chromosomal shape during mitosis and the chromatin domains function as building blocks of chromosomes. Furthermore, a “heat map” was successfully created that revealed the chromatin movement in the local position of the nucleus. It was found that the more the extent of heterochromatic regions, the less was the chromatin movement. The heat map and MSD analysis showed that the chromatin domain fluctuated by the alteration of histone modification and hypotonic condition and the depletion of chromatin-associated proteins. Especially, TSA, which inhibits histone deacetylation and

increases acetylated marks, showed a decondensation of the chromatin domain and an increase in chromatin mobility. Additionally, the heterochromatin region enriched with methylated histones showed the condensed pattern and low mobility of chromatin, and through the differentiation of ESCs, the chromatin becomes more condense and experiences more decrease in its mobility. These results suggested that the change in chromatin environment at a single molecule level is related to histone modification patterns (Chapter 4).

5.2. Future directions

The experimental techniques such as ChIP-Seq, Hi-C, and single nucleosome imaging were collected for the study of chromatin and epigenetics. However, there remain many issues that need to be studied:

1. The direct function of chromatin environment
2. The relationships between chromatin environment and various histone modification types
3. Chromatin environment on specific genomic loci
4. Histone modification imaging in a living cell
5. Genotype, epigenotype, and phenotype

In this thesis, the variation of chromatin structure and dynamics at the euchromatic region and heterochromatic region was revealed, suggesting that the difference in chromatin environments is directly related to the cellular function. For the first issue mentioned above, it is required to investigate how the chromatin environment interacts with proteins and has a function through the combined single nucleosome imaging with single molecule imaging of proteins that interrogate the target site.

It was observed in the present study that the methylated or acetylated histones contributed to the chromatin environment; however, currently, numerous histone modification types are assumed to interact with the chromatin environment (Strahl and Allis, 2000). Thus, the second issue mentioned above can be resolved by performing knock-out or knock-down experiments on histone modifier or binding proteins on modified histones.

The chromatin environment in the total nucleus was observed in this thesis (Chapter 4); therefore, to further understand the detail of chromatin, it is important to observe the chromatin structure and dynamics on specific genomic loci, especially around the target gene. To resolve the third issue mentioned above, the TALEN and CRISPR systems (Chen et al., 2013; Miyanari

et al., 2013) can be considered, but at present, it is not sufficient and needs improvement in the background level.

Furthermore, it is hard to observe the only the specifically modified histones in a living cell. Thus, for the fourth issue mentioned above, it is important to improve the method for observing specific histone modification based on Fab or other techniques (Stasevich et al., 2014).

The biggest question is to reveal the relationships among "genotype, epigenotype, and phenotype". Epigenotype is the bridge between genotype and phenotype; thus, it is crucial to carry out more research on chromatin and epigenetics.

5.3. Conclusion

Chromatin led the biological research in the 20th century, and epigenetics is currently one of the most popular areas for biological research in the 21th century. Although the majority of epigenetic research is related to histone modification and DNA methylation, the research that makes the bridge between chromatin and epigenetics is not sufficient. Epigenetics is one of the most important concepts to understand the highly complicated eukaryotic body, but the types of histone modifications and related regulations are diverse and complicated. Although there are diverse aspects of living organisms, it is crucial to describe the more generalized phenomena for understanding chromatin, which is the vessel of epigenetic components. The results of imaging and bioinformatics approaches in this thesis will provide a basis of dynamic and flexible nature of chromatin toward the better understanding of the chromatin function in the eukaryotic epigenetic system and stimulating the mind of researchers.

Acknowledgements

Through my career as a researcher, I took three of fortune. The first fortune is that I belonged to Prof. Masaru Tomita's Lab in Keio University. When I was a high school student, I was thinking about going to the different campus of Keio University. However, one day, a teacher Mr. Ryowa Tsurunaga told me about the study of Prof. Masaru Tomita on the day I was sleeping alone in the clubroom. As that got my interest, I joined the program of summer course of molecular biology for high school students held in Tsuruoka City that summer and had the chance to meet Prof. Tomita. The discussion with Prof. Tomita and member of his lab was so exciting that I immediately decided to proceed to Keio University Shonan Fujisawa Campus and join Prof. Tomita's lab. I'd like to thank Prof. Tomita who providing me with an opportunity to start the research of life science, an optimal working environment, and the freedom to follow my own research interests. The second fortune is that I met with chromatin research when I was an undergrad student and took the class of Dr. Saito. In that class, I joined the project of Mr. Ogawa (currently Dr. Ogawa) and in group predicted of the 3D structure of non-coding RNA with the algorithm based on machine learning. By chance, Dr. Yachie came to the class to give us the comment and he rejected our plan because of possibility. We remade the plan and proposed the development of the detection method of nucleosome positions from ChIP-Seq data based on the idea of Ogawa, and it was accepted. Based on the result of its class, we described two papers of chromatin using bioinformatics approach. I would like to appreciate Dr. Rintaro Saito, Dr. Nozomu Yachie, and Dr. Ryu Ogawa. The third fortune is meeting with Prof. Kazuhiro Maeshima in National Institute of Genetics. When I was a master course student, I was not filled with the research of chromatin using a bioinformatics and hoped to start the research of chromatin structure. On twitter, I captured the tweet of Prof. Maeshima's research for the irregular folding model of chromatin. Immediately, I went to National Institute of Genetics and started the collaboration research at that summer. In Prof. Maeshima's lab, I could take some impressive data for single molecule imaging. I would like to thank Prof. Maeshima gratefully for making me aware of the happiness of imaging and observing. I also thank Prof. Takeharu Nagai in Osaka University, Dr. Yasushi Okada in RIKEN QBiC, Dr. Tomomi Tani in MBL, Dr. Koichi Takahashi in RIKEN QBiC, Dr. Kazunari Kaizu in RIKEN QBiC, and Dr. Yasumasa Joti in JASRI for a technical advice of single molecule imaging and chromatin

simulation.

I would like to thank Prof. Mitsuhiro Itaya in Keio University who first taught me genetic experiments when I was an undergraduate student and support my Doctoral thesis and Associate Prof. Yasuhiro Naito in Keio University who gave me a physical perspective in biology in his lecture and support my Doctoral thesis. I was supported by the earnest discussions and joyful conversations with Prof. Akio Kanai, Prof. Tomoyoshi Soga, Project Associate Prof. Kazuharu Arakawa, Dr. Yosuke Ozawa, Dr. Motomu Matsui, Mr. Shinya Murata, Dr. Nobuaki Kono, Mr. Satoshi Tamaki, Mr. Hiroyuki Nakamura, Ms. Tomoko Kikuchi, Dr. Kentaro Hayashi, Dr. Kiyofumi Hamashima, Mr. Keita Ikegami, Mr. Toshinori Tsuboi, Mr. Keisuke Morita, Mr. Kazuki Oshita, Mr. Yuki Shindo, Mr. Yasufumi Yamashita, Mr. Yuka Hirose, Mr. Soh Ishiguro, Mr. Gakuto Makino, Mr. Kenta Morimoto, and Mr. Takahiro Masuda in Keio University. I would also like to thank the members of Maeshima's lab in National Institute of Genetics, Dr. Ichiro Hiratani, Assistant Prof. Satoru Ide, Assistant Prof. Kayo Hibino, Dr. Saera Hihara, Ms. Sachiko Tamura, Mr. Ryosuke Imai, Ms. Mai Tanbo, Ms. Asuka Sasaki, and Mr. Ryosuke Nagashima.

Finally, I would like to thank my family, especially my parents Keizo and Chie, my brother Hiroki, and my wife Keiko for their continuous support and encouragement throughout my student life.

The studies in this thesis were supported by research funds from the Yamagata Prefectural Government and Tsuruoka City, and by a Grant-in-Aid from the Japan Society for the Promotion of Science (JSPS), Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Science and Technology Agency (JST) CREST.

References

Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C., and Pugh, B.F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. Nature *446*, 572-576.

Albiez, H., Cremer, M., Tiberi, C., Vecchio, L., Schermelleh, L., Dittrich, S., Kupper, K., Joffe, B., Thormeyer, T., von Hase, J.*, et al.* (2006). Chromatin domains and the interchromatin compartment form structurally defined and functionally interacting nuclear networks. Chromosome Res *14*, 707-733.

Allis, C.D., Jenuwein, T., and Reinberg, D. (2006). Epigenetics (Cold Spring Harbor laboratory Press).

Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., and Ellenberg, J. (2009). Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. EMBO J *28*, 3785-3798.

Bannister, A.J., and Miska, E.A. (2000). Regulation of gene expression by transcription factor acetylation. Cell Mol Life Sci *57*, 1184-1192.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Belmont, A.S., Braunfeld, M.B., Sedat, J.W., and Agard, D.A. (1989). Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro. Chromosoma *98*, 129-143.

Belmont, A.S., and Bruce, K. (1994). Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. J Cell Biol *127*, 287-302.

Belmont, A.S., Li, G., Sudlow, G., and Robinett, C. (1999). Visualization of large-scale chromatin structure and dynamics using the lac operator/lac repressor reporter system. Methods Cell Biol *58*, 203-222.

Berezney, R., Malyavantham, K.S., Pliss, A., Bhattacharya, S., and Acharya, R. (2005). Spatio-temporal dynamics of genomic organization and function in the mammalian cell nucleus. Adv Enzyme Regul *45*, 17-26.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R.*, et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. Science *313*, 1642-1645.

Bhandare, R., Schug, J., Le Lay, J., Fox, A., Smirnova, O., Liu, C., Naji, A., and Kaestner, K.H. (2010). Genome-wide analysis of histone modifications in human pancreatic islets. Genome Res *20*, 428-433.

Bock, C., Walter, J., Paulsen, M., and Lengauer, T. (2007). CpG island mapping by epigenome prediction. PLoS Comput Biol *3*, e110.

Bohn, M., Diesinger, P., Kaufmann, R., Weiland, Y., Muller, P., Gunkel, M., von Ketteler, A., Lemmer, P., Hausmann, M., Heermann, D.W.*, et al.* (2010). Localization microscopy reveals expression-dependent parameters of chromatin nanostructure. Biophys J *99*, 1358-1367.

Bordas, J., Perez-Grau, L., Koch, M.H., Vega, M.C., and Nave, C. (1986). The superstructure of chromatin and its condensation mechanism. I. Synchrotron radiation X-ray scattering results. Eur Biophys J *13*, 157-173.

Bowler, P.J. (1971). Preformation and pre-existence in the seventeenth century: A brief analysis. J Hist Biol *4*, 221-244.

Brin, S., and Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. In Seventh International World-Wide Web Conference (WWW 1998) (Brisbane, Australia).

Brown, R. (1833). On the Organs and Mode of Fecundation in Orchidea and Asclepiadea. Transactions of the Linnean Society of London *16*, 685-738.

Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J Mol Biol *212*, 563-578.

Butler, J.E., and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev *16*, 2583-2592.

Bystricky, K., Heun, P., Gehlen, L., Langowski, J., and Gasser, S.M. (2004). Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. Proc Natl Acad Sci U S A *101*, 16495-16500.

Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. Nature *461*, 193-198.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R.,

Ravasi, T., Lenhard, B., Wells, C.*, et al.* (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559-1563.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A., Taylor, M.S., Engstrom, P.G., Frith, M.C.*, et al.* (2006). Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet *38*, 626-635.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S.*, et al.* (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479-1491.

Chereji, R.V., and Morozov, A.V. (2011). Statistical Mechanics of Nucleosomes Constrained by Higher-Order Chromatin Structure. J Stat Phys *144*, 379-404.

Chubb, J.R., Boyle, S., Perry, P., and Bickmore, W.A. (2002). Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol *12*, 439-445.

Coleman, R.A., and Pugh, B.F. (1995). Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. J Biol Chem *270*, 13850-13859.

Correns, K. (1900). Ueber den Einfluss, welchen die Zahl der zur Bestäubung verwendeten Pollenkörner auf die Nachkommenschaft hat. Berichte der Deutschen Botanischen Gesellschaft Bd *18*, 422-435.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292-301.

Cremer, T., Kreth, G., Koester, H., Fink, R.H., Heintzmann, R., Cremer, M., Solovei, I., Zink, D., and Cremer, C. (2000). Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. Crit Rev Eukaryot Gene Expr *10*, 179-212.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

Davey, C.A., and Richmond, T.J. (2002). DNA-dependent divalent cation binding in the nucleosome core particle. Proc Natl Acad Sci U S A *99*, 11169-11174.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol *319*, 1097-1113.

De Vries, H. (1900). Das Spaltungsgesetz der bastarde (Borntraeger).

Deans, C., and Maggert, K.A. (2015). What do you mean, "epigenetic"? Genetics *199*, 887-896.

Dekker, J. (2003). A closer look at long-range chromosomal interactions. Trends Biochem Sci *28*, 277-280.

Dekker, J. (2008). Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. J Biol Chem *283*, 34532-34540.

Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet *14*, 390-403.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R., and Richmond, T.J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science *306*, 1571-1573.

Dubochet, J., Adrian, M., Schultz, P., and Oudet, P. (1986). Cryo-electron microscopy of vitrified SV40 minichromosomes: the liquid drop model. EMBO J *5*, 519-528.

Eltsov, M., Maclellan, K.M., Maeshima, K., Frangakis, A.S., and 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 U S A *105*, 19732-19737.

Finch, J.T., and Klug, A. (1976). Solenoidal model for superstructure in chromatin. Proc Natl Acad Sci U S A *73*, 1897-1901.

Flemming, W. (1878). Zur Kenntnis der Zelle in ihrer Teilung-Erscheinungen. Schr Naturwiss Ver Schlesw-Holst *3*, 26.

Flemming, W. (1879). Ueber das Verhalten des Kerns bei der Zelltheilung, und über die Bedeutung mehrkerniger Zellen. Virchows Archiv *77*, 1-29.

Flemming, W. (1882). Zellsubstanz, kern und zelltheilung (Vogel).

Frith, M.C., Valen, E., Krogh, A., Hayashizaki, Y., Carninci, P., and Sandelin, A. (2008). A code for transcription initiation in mammalian genomes. Genome Res *18*, 1-12.

Fussner, E., Strauss, M., Djuric, U., Li, R., Ahmed, K., Hart, M., Ellis, J., and Bazett-Jones, D.P. (2012). Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. EMBO Rep *13*, 992-996.

Ghirlando, R., and Felsenfeld, G. (2013). Chromatin structure outside and inside the nucleus. Biopolymers *99*, 225-232.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. (2004). Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell *118*, 555-566.

Gorisch, S.M., Wachsmuth, M., Toth, K.F., Lichter, P., and Rippe, K. (2005). Histone

acetylation increases chromatin accessibility. J Cell Sci *118*, 5825-5834.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W.*, et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Haig, D. (2004). The (dual) origin of epigenetics. Cold Spring Harb Symp Quant Biol *69*, 67-70.

Hajjoul, H., Mathon, J., Ranchon, H., Goiffon, I., Mozziconacci, J., Albert, B., Carrivain, P., Victor, J.M., Gadal, O., Bystricky, K.*, et al.* (2013). High-throughput chromatin motion tracking in living yeast reveals the flexibility of the fiber throughout the genome. Genome Res *23*, 1829-1838.

Hall, B.K. (2012). Evolutionary developmental biology (Springer Science & Business Media).

Hansen, J.C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu Rev Biophys Biomol Struct *31*, 361-392.

Harvey, W. (1651). Exercitationes de generatione animalium.

Heun, P., Laroche, T., Shimada, K., Furrer, P., and Gasser, S.M. (2001). Chromosome dynamics in the yeast interphase nucleus. Science *294*, 2181-2186.

Hihara, S., Pack, C.G., Kaizu, K., Tani, T., Hanafusa, T., Nozaki, T., Takemoto, S., Yoshimi, T., Yokota, H., Imamoto, N.*, et al.* (2012). Local nucleosome dynamics facilitate chromatin accessibility in living mammalian cells. Cell Rep *2*, 1645-1656.

Hiratani, I., and Gilbert, D.M. (2010). Autosomal lyonization of replication domains during early Mammalian development. Adv Exp Med Biol *695*, 41-58.

Holliday, R. (1994). Epigenetics: an overview. Dev Genet *15*, 453-457.

Holliday, R. (2006). Epigenetics: a historical overview. Epigenetics *1*, 76-80.

Hooke, R. (1665). Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses, with observations and inquiries thereupon.

Illingworth, R.S., and Bird, A.P. (2009). CpG islands--'a rough guide'. FEBS Lett *583*, 1713-1720.

Ioshikhes, I.P., and Zhang, M.Q. (2000). Large-scale human promoter mapping using CpG islands. Nat Genet *26*, 61-63.

Jackson, D.A., and Pombo, A. (1998). Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol *140*, 1285-1295.

Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S.L., and Danuser, G.

(2008). Robust single-particle tracking in live-cell time-lapse sequences. Nat Methods *5*, 695-702.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet *10*, 161-172.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet *41*, 941-945.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497-1502.

Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., and Maeshima, K. (2012). Chromosomes without a 30-nm chromatin fiber. Nucleus *3*, 404-410.

Kalashnikova, A.A., Porter-Goff, M.E., Muthurajan, U.M., Luger, K., and Hansen, J.C. (2013). The role of the nucleosome acidic patch in modulating higher order chromatin structure. J R Soc Interface *10*, 20121022.

Karlic, R., Chung, H.R., Lasserre, J., Vlahovicek, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A *107*, 2926-2931.

Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J.*, et al.* (2003). The UCSC Genome Browser Database. Nucleic Acids Res *31*, 51-54.

Kawaji, H., Frith, M.C., Katayama, S., Sandelin, A., Kai, C., Kawai, J., Carninci, P., and Hayashizaki, Y. (2006). Dynamic usage of transcription start sites within core promoters. Genome Biol *7*, R118.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J Cell Biol *153*, 1341-1353.

Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. J Cell Biol *159*, 777-782.

Kiskowski, M.A., Hancock, J.F., and Kenworthy, A.K. (2009). On the use of Ripley's K-function and its derivatives to analyze domain size. Biophys J *97*, 1095-1103.

Koch, M.H., Vega, M.C., Sayers, Z., and Michon, A.M. (1987). The superstructure of chromatin and its condensation mechanism. III: Effect of monovalent and divalent cations X-ray solution scattering and hydrodynamic studies. Eur Biophys J *14*, 307-319.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285-294.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

Kratz, A., Arner, E., Saito, R., Kubosaki, A., Kawai, J., Suzuki, H., Carninci, P., Arakawa, T., Tomita, M., Hayashizaki, Y.*, et al.* (2010). Core promoter structure and genomic context reflect histone 3 lysine 9 acetylation patterns. BMC Genomics *11*, 257.

Kruithof, M., Chien, F.T., Routh, A., Logie, C., Rhodes, D., and van Noort, J. (2009). Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. Nat Struct Mol Biol *16*, 534-540.

Laemmli, U.K., Cheng, S.M., Adolph, K.W., Paulson, J.R., Brown, J.A., and Baumbach, W.R. (1978). Metaphase chromosome structure: the role of nonhistone proteins. Cold Spring Harb Symp Quant Biol *42*, 351-360.

Langmore, J.P., and Paulson, J.R. (1983). Low angle x-ray diffraction studies of chromatin structure in vivo and in isolated nuclei and metaphase chromosomes. J Cell Biol *96*, 1120-1131.

Levi, V., Ruan, Q., Plutz, M., Belmont, A.S., and Gratton, E. (2005). Chromatin dynamics in interphase cells revealed by tracking in a two-photon excitation microscope. Biophys J *89*, 4275-4285.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M.*, et al.* (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315-322.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Ma, H., Samarabandu, J., Devdhar, R.S., Acharya, R., Cheng, P.C., Meng, C., and Berezney, R. (1998). Spatial and temporal dynamics of DNA replication sites in mammalian cells. J Cell Biol *143*, 1415-1425.

Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? Curr Opin Cell Biol *22*, 291-297.

Maeshima, K., Imai, R., Tamura, S., and Nozaki, T. (2014). Chromatin as dynamic 10-nm fibers. Chromosoma *123*, 225-237.

Maison, C., and Almouzni, G. (2004). HP1 and the dynamics of heterochromatin maintenance. Nat Rev Mol Cell Biol *5*, 296-305.

Marenduzzo, D., Finan, K., and Cook, P.R. (2006). The depletion attraction: an underappreciated force driving cellular organization. J Cell Biol *175*, 681-686.

Markaki, Y., Gunkel, M., Schermelleh, L., Beichmanis, S., Neumann, J., Heidemann, M., Leonhardt, H., Eick, D., Cremer, C., and Cremer, T. (2010). Functional nuclear organization of transcription and DNA replication: a topographical marriage between chromatin domains and the interchromatin compartment. Cold Spring Harb Symp Quant Biol *75*, 475-492.

Marsden, M.P., and Laemmli, U.K. (1979). Metaphase chromosome structure: evidence for a radial loop model. Cell *17*, 849-858.

Martin, R.M., Gorisch, S.M., Leonhardt, H., and Cardoso, M.C. (2007). An unexpected link between energy metabolism, calcium, chromatin condensation and cell cycle. Cell Cycle *6*, 2422-2424.

Mattout, A., and Meshorer, E. (2010). Chromatin plasticity and genome organization in pluripotent embryonic stem cells. Curr Opin Cell Biol *22*, 334-341.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res *18*, 1073-1083.

Mazza, D., Abernathy, A., Golob, N., Morisaki, T., and McNally, J.G. (2012). A benchmark for chromatin binding measurements in live cells. Nucleic Acids Res *40*, e119.

Medawar, P., and Medawar, J. (1983). Aristotle to Zoos (Cambridge, MA: Harvard University Press).

Meluzzi, D., and Arya, G. (2013). Recovering ensembles of chromatin conformations from contact probabilities. Nucleic Acids Res *41*, 63-75.

Mendel, G. (1866). Versuche über Pflanzenhybriden. Verhandlungen des Naturforschenden Vereines in Brünn 4 *44*, 3-47.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell *10*, 105-116.

Meyen, F. (1830). Mikroskopische Abbildungen (auf vierzehn Kupfertafeln) zur Phytotomie (und Pflanze. Abbildungen botanische/zellbiologische mikroskopische Beobachtungen) (Haude

und Spenersche Buchhandlung).

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., and Koche, R.P. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Mine-Hattab, J., and Rothstein, R. (2012). Increased chromosome mobility facilitates homology search during recombination. Nat Cell Biol *14*, 510-517.

Miyanari, Y., Ziegler-Birling, C., and Torres-Padilla, M.E. (2013). Live visualization of chromatin dynamics with fluorescent TALEs. Nat Struct Mol Biol *20*, 1321-1324.

Mizuguchi, T., Fudenberg, G., Mehta, S., Belton, J.M., Taneja, N., Folco, H.D., FitzGerald, P., Dekker, J., Mirny, L., Barrowman, J.*, et al.* (2014). Cohesin-dependent globules and heterochromatin shape 3D genome architecture in S. pombe. Nature *516*, 432-435.

Mohl, H.v. (1835). Sur la structure et les formes des grains de pollen. Ann Sci Nat Bot *2*, 304-346.

Morgan, T.H. (1910). Sex Limited Inheritance in Drosophila. Science *32*, 120-122.

Morgan, T.H. (1911). Random segregation versus coupling in Mendelian inheritance. Science *34*, 384-384.

Nägeli, K.W. (1842). Zur entwicklungsgeschichte des pollens bei den phanerogamen (Orell Füssli).

Nägeli, K.W. (1844). Zellenkerne, Zellenbildung und Zellenwachsthum bei den Pflanzen (Verlag von Meyer und Zeller).

Nakamura, H., Morita, T., and Sato, C. (1986). Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. Exp Cell Res *165*, 291-297.

Nanney, D.L. (1958). Epigenetic control systems. Proc Natl Acad Sci U S A *44*, 712.

Nasmyth, K., and Haering, C.H. (2005). The structure and function of SMC and kleisin complexes. Annu Rev Biochem *74*, 595-648.

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B.R., Mirny, L.A., and Dekker, J. (2013). Organization of the mitotic chromosome. Science *342*, 948-953.

Niedojadlo, J., Perret-Vivancos, C., Kalland, K.H., Cmarko, D., Cremer, T., van Driel, R., and Fakan, S. (2011). Transcribed DNA is preferentially located in the perichromatin region of mammalian cell nuclei. Exp Cell Res *317*, 433-444.

Nishino, Y., Eltsov, M., Joti, Y., Ito, K., Takata, H., Takahashi, Y., Hihara, S., Frangakis, A.S., Imamoto, N., Ishikawa, T.*, et al.* (2012). Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. EMBO J *31*,

1644-1653.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berlum, N.L., Meisig, J., Sedat, J.W.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Ober, R.J., Ram, S., and Ward, E.S. (2004). Localization accuracy in single-molecule microscopy. Biophys J *86*, 1185-1200.

Ochiai, H., Sugawara, T., and Yamamoto, T. (2015). Simultaneous live imaging of the transcription and nuclear position of specific genes. Nucleic Acids Res *43*, e127.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. Nat Rev Mol Cell Biol *4*, 809-814.

Ono, T., Losada, A., Hirano, M., Myers, M.P., Neuwald, A.F., and Hirano, T. (2003). Differential contributions of condensin I and condensin II to mitotic chromosome architecture in vertebrate cells. Cell *115*, 109-121.

Ono, T., Yamashita, D., and Hirano, T. (2013). Condensin II initiates sister chromatid resolution during S phase. J Cell Biol *200*, 429-441.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281-300.

Ozsolak, F., Song, J.S., Liu, X.S., and Fisher, D.E. (2007). High-throughput mapping of the chromatin structure of human promoters. Nat Biotechnol *25*, 244-248.

Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet *10*, 669-680.

Paulson, J.R., and Laemmli, U.K. (1977). The structure of histone-depleted metaphase chromosomes. Cell *12*, 817-828.

Peck, A.L. (1943). Aristotle generation of animals (Brill Archive).

Plessy, C., Bertin, N., Takahashi, H., Simone, R., Salimullah, M., Lassmann, T., Vitezic, M., Severin, J., Olivarius, S., Lazarevic, D.*, et al.* (2010). Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. Nat Methods *7*, 528-534.

Ponjavic, J., Lenhard, B., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., and Sandelin, A. (2006). Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters. Genome Biol *7*, R78.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K.*, et al.* (2014). Topologically associating domains are stable units of

replication-timing regulation. Nature *515*, 402-405.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell *138*, 114-128.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Remak, R. (1852). Ueber extracellulare Entstehung thierischer Zellen und über Vermehrung derselben durch Theilung. Archiv für Anatomie, Physiologie und wissenschaftliche Medizin, 47-57.

Remak, R. (1855). Untersuchungen über die Entwickelung der Wirbelthiere (G. Reimer).

Ricci, M.A., Manzo, C., Garcia-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145-1158.

Robinson, P.J., Fairall, L., Huynh, V.A., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proc Natl Acad Sci U S A *103*, 6506-6511.

Robinson, P.J., and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. Curr Opin Struct Biol *16*, 336-343.

Rogers, S.S., Waigh, T.A., Zhao, X., and Lu, J.R. (2007). Precise particle tracking against a complicated background: polynomial fitting with Gaussian weight. Phys Biol *4*, 220-227.

Roh, T.Y., Cuddapah, S., Cui, K., and Zhao, K. (2006). The genomic landscape of histone modifications in human T cells. Proc Natl Acad Sci U S A *103*, 15782-15787.

Roopra, A., Qazi, R., Schoenike, B., Daley, T.J., and Morrison, J.F. (2004). Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. Mol Cell *14*, 727-738.

Rust, M.J., Bates, M., and Zhuang, X.W. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods *3*, 793-795.

Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res *32*, D91-94.

Schalch, T., Duda, S., Sargent, D.F., and Richmond, T.J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. Nature *436*, 138-141.

Schermelleh, L. (2005). In vivo DNA Replication Labeling. In Cell Biology, Four-Volume Set: A Laboratory Handbook, 301-303.

Schermelleh, L., Solovei, I., Zink, D., and Cremer, T. (2001). Two-color fluorescence labeling of early and mid-to-late replicating chromatin in living cells. Chromosome Res *9*, 77-80.

Schleiden, M.J. (1838). Beiträge zur phytogenesis (publisher not identified).

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887-898.

Schwann, T. (1839). Mikroskopische Untersuchungen ueber die Uebereinstimmung in der Struktur und dem Wachsthum der Thiers und Pflanzen (GL Reimer).

Sedat, J., and Manuelidis, L. (1978). A direct approach to the structure of eukaryotic chromosomes. Cold Spring Harb Symp Quant Biol *42 Pt 1*, 331-350.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thastrom, A., Field, Y., Moore, I.K., Wang, J.P., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature *442*, 772-778.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. Cell *148*, 458-472.

Shintomi, K., and Hirano, T. (2010). Sister chromatid resolution: a cohesin releasing network and beyond. Chromosoma *119*, 459-467.

Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T.*, et al.* (2003). Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A *100*, 15776-15781.

Simon, H.A. (1991). The architecture of complexity (Springer).

Smith, A.E., Chronis, C., Christodoulakis, M., Orr, S.J., Lea, N.C., Twine, N.A., Bhinge, A., Mufti, G.J., and Thomas, N.S. (2009). Epigenetics of human T cells during the G0-->G1 transition. Genome Res *19*, 1325-1337.

Smith, J.M. (1990). Models of a dual inheritance system. J Theor Biol *143*, 41-53.

Smith, T.F., and Waterman, M.S. (1981a). Comparison of biosequences. Adv Appl Math *2*, 482-489.

Smith, T.F., and Waterman, M.S. (1981b). Identification of common molecular subsequences. J Mol Biol *147*, 195-197.

Sofueva, S., Yaffe, E., Chan, W.C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H.,

Pollard, S.M., Schroth, G.P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J *32*, 3119-3129.

Stasevich, T.J., Hayashi-Takanaka, Y., Sato, Y., Maehara, K., Ohkawa, Y., Sakata-Sogawa, K., Tokunaga, M., Nagase, T., Nozaki, N., McNally, J.G.*, et al.* (2014). Regulation of RNA polymerase II activation by histone acetylation in single living cells. Nature *516*, 272-275.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Straight, A.F., Belmont, A.S., Robinett, C.C., and Murray, A.W. (1996). GFP tagging of budding yeast chromosomes reveals that protein-protein interactions can mediate sister chromatid cohesion. Curr Biol *6*, 1599-1608.

Strasburger, E. (1880). Zellbildung und Zelltheilung (Fischer).

Subach, F.V., Patterson, G.H., Manley, S., Gillette, J.M., Lippincott-Schwartz, J., and Verkhusha, V.V. (2009). Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. Nat Methods *6*, 153-159.

Sutton, W.S. (1902). On the morphology of the chromosome group in. Brachystola magna. The Biological Bulletin *4*, 24-39.

Suzuki, H., Forrest, A.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J.*, et al.* (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat Genet *41*, 553-562.

Thompson, R.E., Larson, D.R., and Webb, W.W. (2002). Precise nanometer localization analysis for individual fluorescent probes. Biophys J *82*, 2775-2783.

Tirosh, I., and Barkai, N. (2008). Two strategies for gene regulation by promoter nucleosomes. Genome Res *18*, 1084-1091.

Tokunaga, M., Imamoto, N., and Sakata-Sogawa, K. (2008). Highly inclined thin illumination enables clear single-molecule imaging in cells. Nat Methods *5*, 159-161.

Tolstorukov, M.Y., Kharchenko, P.V., Goldman, J.A., Kingston, R.E., and Park, P.J. (2009). Comparative analysis of H2A.Z nucleosome organization in the human and yeast genomes. Genome Res *19*, 967-977.

Tschermak, E. (1900). Ueber künstliche Kreuzung bei Pisum sativum (E. Tschermak).

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell *19*, 381-391.

Vazquez, J., Belmont, A.S., and Sedat, J.W. (2001). Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. Curr Biol *11*, 1227-1239.

Virchow, R.L.K. (1858). Die Cellularpathologie: in ihrer Begründung auf physiologische und pathologische Gewebelehre, Vol 1 (Verlag von August Hirschwald, Unter den Linden No. 68).

Waddington, C.H. (1968). The basic ideas of biology. Nature *218*, 525-527.

Waddington, C.H. (1942). The epigenotype. Endeavour *1*, 18-20.

Waldeyer, W. (1888). Ueber Karyokinese und ihre Beziehungen zu den Befruchtungsvorgängen. Archiv für mikroskopische Anatomie *32*, 1-122.

Wang, J., and Hannenhalli, S. (2006). A mammalian promoter model links cis elements to genetic networks. Biochem Biophys Res Commun *347*, 166-177.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q.*, et al.* (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet *40*, 897-903.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids. Nature *171*, 737-738.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Widom, J., and Klug, A. (1985). Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. Cell *43*, 207-213.

Wolff, C.F. (1759). Theoria generationis. Cum 2 tab. aen (Hendelius).

Woodcock, C.L., Frado, L.L., and Rattner, J.B. (1984). The higher-order structure of chromatin: evidence for a helical ribbon arrangement. J Cell Biol *99*, 42-52.

Wyrick, J.J., Holstege, F.C., Jennings, E.G., Causton, H.C., Shore, D., Grunstein, M., Lander, E.S., and Young, R.A. (1999). Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. Nature *402*, 418-421.

Yildiz, A., Forkey, J.N., McKinney, S.A., Ha, T., Goldman, Y.E., and Selvin, P.R. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. Science *300*, 2061-2065.

Yoshida, M., Kijima, M., Akita, M., and Beppu, T. (1990). Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. J Biol Chem *265*, 17174-17179.

Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-scale identification of nucleosome positions in S. cerevisiae. Science *309*, 626-630.

Zhang, Y., Shin, H., Song, J.S., Lei, Y., and Liu, X.S. (2008). Identifying positioned nucleosomes with epigenetic marks in human from ChIP-Seq. BMC Genomics *9*, 537.

Zink, D., Bornfleth, H., Visser, A., Cremer, C., and Cremer, T. (1999). Organization of early and late replicating DNA in human chromosome territories. Exp Cell Res *247*, 176-188.

Zink, D., Sadoni, N., and Stelzer, E. (2003). Visualizing chromatin and chromosomes in living cells. Methods *29*, 42-50.

Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van Ijcken, W.F.*, et al.* (2014a). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc Natl Acad Sci U S A *111*, 996-1001.

Zuin, J., Franke, V., van Ijcken, W.F., van der Sloot, A., Krantz, I.D., van der Reijden, M.I., Nakato, R., Lenhard, B., and Wendt, K.S. (2014b). A cohesin-independent role for NIPBL at promoters provides insights in CdLS. PLoS genetics *10*, e1004153.

"The History of the Earth is recorded in the Layers of its Crust;
The History of all Organisms is inscribed in the Chromosomes."

Hitoshi Kihara (1946)