FROM NUCLEOSOME TO CHROMOSOME: A DYNAMIC ORGANIZATION OF GENETIC INFORMATION

P. Fransz\(^1\) and J.H. de Jong\(^2\)

\(^1\)Nuclear Organization Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, the Netherlands
\(^2\)Laboratory of Genetics, Wageningen University and Research Center, Wageningen, The Netherlands

ABSTRACT

Gene activity is controlled at different levels of chromatin organization, which involve genomic sequences, nucleosome structure, chromatin folding and chromosome arrangement. These levels are interconnected and influence each other. At the basic level nucleosomes occlude DNA sequence from interacting with DNA binding proteins. Evidently, nucleosome positioning is a major factor in gene control and chromatin organization. Understanding the biological rules that govern deposition and removal of the nucleosomes to and from the chromatin fibre is the key to understanding gene regulation and chromatin organization. In this review we describe and discuss the relationship between the different levels of chromatin organization in plants and animals.

INTRODUCTION

The development of a multicellular organism involves constant changes in gene activity in a development- and tissue specific manner. Instrumental to this process is the ability of the nucleus to modify its program by switching on or off different sets of genes and changing its protein and RNA composition. As a result the cell can adequately respond to different stimuli and, if necessary, change its identity. Under certain conditions the cell can even turn into a less differentiated state with increased potency to develop different cell types, which illustrates the plasticity of the nuclear program in eukaryote cells. Understanding the biological mechanisms that underlie nuclear
(re)programming is a major challenge and one of the key objectives in contemporary medical science and biotechnology. Chromatin is the main component in these processes, not only because it accommodates the DNA sequence with genetic information, but also because it comprises RNA and protein complexes. A crucial property of chromatin is that it is precisely and efficiently packed in order to retrieve the genetic information in an accurate manner at the right time and in the correct cell. Specific folding configurations of the chromatin fiber are exemplified by enhancer-promoter interactions (Hadjur et al., 2009, Mishiro et al., 2009) and transcription factories (Mitchel and Fraser, 2008) found in human cells and by the formation of heterochromatic domains such as chromocenters in mouse (Guenatrie et al., 2004) and Arabidopsis (Fransz et al., 2002). We distinguish three major levels of chromatin organization: (1) the nucleosome, where DNA is wrapped around the histone core; (2) the 30 nm fiber, with an estimated linear compaction of 100-200 kb/µm; (3) higher folding levels, which position the chromosome fiber within a chromosome territory and enable the high compaction of metaphase chromosomes. These higher folding levels are poorly understood. Over the last two decades a number of models have been presented to describe higher order folding of chromatin fibers. These include random walk and loop models, which differ in frequency and combination of the intra-chromosomal interactions (Mateos-Langerak et al., 2009). The models help us to understand the physical properties of the chromatin polymer and contribute to our knowledge of chromosome folding. Yet, we are still ignorant of how chromosome organization is established. How do the different folding levels contribute to chromatin compaction? Which factors control chromosome folding? Here we present an overview of chromosome organization in plant nuclei. Since most of the information on eukaryote chromosomes is generated in mammalian cell systems we will also include recent data from human and animal chromosome organization.

**CHROMATIN STRUCTURE**

*Low level chromatin folding models*

The eukaryote chromosome consists of a long chromatin fiber, that is folded and compacted in such a way that the genomic sequence is efficiently stored, but can be retrieved upon request for transcription, replication or other genomic activities. The general picture of the chromatin fiber is a polymer of nucleosomes, each of which is formed by an octamere of the canonical histones H2A, H2B, H3 and H4, around which a 147 bp DNA helix (Fig.1a) is wrapped 1.7 times (Luger et al., 1997). The lowest level of chromatin organization is the ‘beads-on-the-string’ configuration, which has been observed under non-physiological conditions. However, it is not likely that this configuration of nucleosomes exist in a living cell, since under physiological conditions the string of nucleosomes
can instantly form a 30 nm structure (van Holde 1988, Bednar et al., 1998). The 30 nm fiber undergoes complex folding steps generating higher levels of chromatin compaction up to the extremely condensed metaphase chromosome. Despite years of extensive chromosome studies we are still ignorant of the internal chromatin organization above the 30 nm level. Even the organization of the 30 nm fiber remains a matter of debate (Maeshima et al., 2010). The current view of the 30-nm fiber is a helical arrangement with a compaction degree of 70-160 kb/µm (Robinson et al., 2006), for which two basic models have been proposed (Wu et al., 2007). The one-start helix (solenoid) is linearly arranged, whereas the two-start helix (zigzag) forms two nucleosomal rows in such a way that following nucleosomes alternate from one to the other row (Fig. 1b-d) and second nearest neighbour nucleosomes become interacting partners. Recent experimental data, including crystal structure analysis, reveal two stacks of nucleosomes, which favors the zigzag model (Dorigo et al., 2004, Schalch et al., 2005, Grigoryev et al., 2009). Within the two-start zigzag model two submodels are distinguished, the ‘helical ribbon’ and the ‘crossed-linker’ type, each explaining different characteristics of the 30 nm fiber, depending on the length of the linker DNA, which is the stretch of DNA connecting two adjacent nucleosomes (Dorigo et al., 2004, Wu et al., 2007).

Chromatin structure depends on DNA linker length

The average length of DNA associated with one nucleosome is the nucleosome repeat length (NRL), of which the linker DNA is the variable component. DNA linker lengths are usually in steps of 10 bp, such as 20, 30 up to 90 bp (Wong et al., 2007), corresponding the number of basepairs within one rotational turn of the DNA helix (Widom, 1992). Together with the nucleosomal DNA (147 bp) they give rise to NRLs of 167, 177 up to 237 bp, respectively, which corresponds to the repeat unit length of many satellite DNA repeat classes. The average DNA linker length shows tissue and species specificity: 20 bp in yeast (Lee et al., 2007), 30 bp in Drosophila (Mavrich et al., 2008), 40 in mammal (Schones et al., 2008) and 30 bp in Arabidopsis (Chodavarapu et al., 2010). The binding of linker histones (H1, H5) to linker DNA in the interior of the fiber induces a further compaction and stabilization of the 30 nm fiber, in particular after deacetylation of core histones and in support of ATP-dependent nucleosome remodeling complexes (Li et al., 2010, Robinson et al., 2008). However, a comparative study between chromatin arrays with different NRLs, 167 and 197, revealed that the dependency on linker histones for compaction is different between short and longer DNA linkers (Routh et al., 2008). Longer linker DNA stretches require linker histones for more compaction and regular nucleosome arrangement. In contrast, chromatin stretches with short DNA linker are less affected by linker histone and give rise to a less compact chromatin fiber (Grigoryev et al., 2009, Routh et al., 2008). Evidently, long-linker chromatin is associated with repressed chromatin state, whereas chromatin with short linkers corresponds to a transcriptionally active state. The highly active yeast chromosomes have short linker DNA stretches, while long DNA linkers have been found in
inactive chromatin such as in spermatid nuclei (Athey et al., 1990). DNA linker length is therefore an important factor in chromatin structure and activity.

**Nucleosome positioning is directly related to chromatin activity**

The average DNA linker length is reflected in the nucleosome positioning, defined as the probability that a nucleosome starts at a given base pair within the genome (Segal and Widom, 2009). Since nucleosomes occlude DNA sequence from interaction with DNA binding proteins, such as activators and repressors, nucleosome positioning plays an important role in activation and repression of chromatin regions (Wyrick et al., 1999). Consequently, factors that regulate nucleosome positioning control chromatin activity. A number of factors have been proposed to direct nucleosome positioning. An experimental and computational study demonstrated that genomes encode an intrinsic nucleosome organization, pointing at a nucleosome positioning code in the DNA sequence (Segal et al., 2006, Kaplan et al., 2009). However, the genomic sequence may not be the primary determinant for nucleosome positioning (Zang et al., 2009). There is evidence that DNA methylation via protein interactions in the major groove affects DNA flexibility and hence nucleosome positioning (Pennings et al., 2005). Also histone modifications are likely to affect the local position of nucleosomes. Modelling studies of oligonucleosome structures revealed that histone tails, in particular of histone H4, play an important role in the higher order compaction of chromatin (Arya and Schlick 2006, 2009). Indeed hyperacetylation of histones increases the accessibility of nucleosomal DNA (Manohar et al., 2009, Anderson et al., 2001). Even a single acetylation event at H4K16 can prevent formation of compact chromatin fibers *in vitro* (Shogren-Knaak et al., 2006). Collectively, the histone modifications together with DNA methylation, transcription factors, chromatin remodelers and the DNA sequence direct the position of nucleosomes along the fiber and regulate the activity of chromatin regions.

Until few years ago only limited gene regions could be examined for nucleosome positioning in relation to sequence and their effects on transcription. It appeared that promoters show a specific arrangement of nucleosomes. The transcription start site (TSS) is preceded by a nucleosome free region (NFR), which is flanked by conserved positioning of relatively stable nucleosomes that contain the histone variant H2A.Z (Lee et al., 2004, Raisner et al., 2005, Yuan et al., 2005). Due to the rapid development of high throughput processing technologies large numbers of chromatin profiling data have been generated, including the global nucleosome positioning. High-resolution genome-wide maps of nucleosome distribution in yeast (Lee et al., 2007) and human (Schones et al., 2008) showed on a genome-wide scale that specific and overall patterns of nucleosome occupancy correlate with transcript abundance and transcription rate. The reports confirm the occurrence of the NFR in promoters, whereas the gene bodies, especially exons, shows high density of nucleosomes (Andersson et al., 2009). The NFR has been found especially in constitutive genes, for example genes encoding glycolysis proteins, and is flanked by stable nucleosomes at the +1 and -1 position. The +1
nucleosome lies at or near the transcription start site (TSS). Less active promoters such as those from stress genes, e.g. the PHO gene in yeast, have a diminished NFR or are completely covered with nucleosomes (Shivaswamy et al., 2008). They require chromatin remodelers and activators to increase accessibility of the promoter. A recent paper has demonstrated that the NFR of active promoters is likely to be occupied to a considerable extent by unstable H3.3/H2A.Z-containing nucleosomes (Jin et al., 2009). Their purification protocol enabled to isolate chromatin fragments with unstable nucleosomes, which appeared to be enriched for H3.3 and H2A.Z. In this context the term "unstable nucleosome regions" is perhaps more appropriate.

Plants show similar characteristics and patterns with respect to nucleosome positioning. Genome-wide nucleosome positioning in relation to DNA methylation was recently studied in Arabidopsis (Chodavarapu et al., 2010). Using high throughput sequencing of nucleosome-associated fragments in combination with their bisulphate-sequencing data at single nucleotide resolution the authors discovered a 10 bp periodicity over nucleosome-bound DNA in the methylated cytosines profiles, irrespective of the sequence context of the cytosine (CG, CHG, CHH). This means that methylation likely takes place to nucleosomal DNA rather than to linker DNA. Indeed, gene body regions, in particular exons, are higher DNA methylated and contain more nucleosomes than flanking regions (Chodavarapu et al., 2010, Anderson et al., 2009). Evidently, DNA methyltransferases preferentially target nucleosome-bound DNA, which suggests that nucleosome positioning controls DNA methylation. Similar patterns were observed in human nucleosomal DNA indicating that nucleosome-controlled DNA methylation is conserved between plant and human. In fact, gene body methylation has been demonstrated to be a highly conserved feature in plants and humans (Feng et al., 2010, Zemach et al., 2010).

**Histone variant H2A.Z counteracts DNA methylation**

A prominent mark of promoters in euchromatin is the histone variant H2A.Z, which has been found in promoter regions of nearly all genes, active and inactive (Raisner et al., 2005). The histone variant is generally located in the nucleosomes that flank the NFR. It has been proposed that H2A.Z serves as an antisilencing factor to antagonize the switch from active euchromatin to inactive heterochromatin (Meneghini et al., 2003). However, H2A.Z may also be involved in heterochromatin silencing for example in combination with the heterochromatin protein HP1-α (Fan et al., 2004). This suggests that H2A.Z has opposite functions in gene regulation. Using a high (300 bp) resolution Chip-chip approach, Guillemette et al. (2005) demonstrated that H2A.Z preferentially associates with promoters of inactive yeast genes and conclude that the H2A.Z incorporation induces the chromatin to poised genes for transcriptional activation.

Plants are not different from yeast and human with respect to H2A.Z function. Arabidopsis H2A.Z is deposited, like in other organisms, in the +1 nucleosome (Zilberman et al., 2008), underscoring the idea of a conserved nucleosome at or near the TSS. The incorporation is carried out by the ATPase
chromatin remodeler SWR1 and promotes transcriptional competence of the gene. Remarkably, H2A.Z was absent from methylated gene bodies and methylated transposons. Strikingly, there is an opposite pattern of DNA methylation along the genome compared to H2A.Z distribution. Apparently, the two chromatin marks are mutually exclusive epigenetic traits. By analyzing patterns of H2A.Z incorporation and methylated DNA in specific mutants, involved in either chromatin feature, the authors conclude that DNA methylation excludes H2A.Z incorporation, while H2A.Z in turn protects DNA from methylation and hence from gene silencing. This confirms the observation that H2A.Z is found in active promoters, or promoters that are poised to activity, while the gene body contains methylated DNA (Cokus et al., 2008, Feng et al., 2010, Zemach et al., 2010, Maunakea et al., 2010).

CHROMOSOME ORGANIZATION IN THE INTERPHASE NUCLEUS

Position and orientation of the chromosome territory depend on many factors
The efficient packing of long DNA fibers into a small nuclear space and the accurate retrieval of genetic information via local unfolding of chromatin is crucial for an organism to manage the genome. Nucleosome density affects the stiffness of the chromatin fiber and consequently changes the folding properties of the chromosome. From a naked DNA fiber to the 30 nm fiber means ~40 fold compaction and a 3-fold decrease in stiffness (Ostashevsky, 2002). Folding and bending of the chromatin fiber will be more limited towards higher order organization levels of the chromosome, with the metaphase chromosome as the ultimate state of chromosome compaction. The biological principles underlying nucleosome positioning are likely to rule the higher order organization of chromatin fibers and the folding pattern of entire chromosomes. How chromosomes are spatially organized in the cell nucleus depends on a multitude of factors, such as chromosome size, distribution of heterochromatin along the linear chromosome, nuclear space, DNA sequence, gene activity, but also genetic background and external (a)biotic factors (see below).

Chromosomes occupy distinct nuclear subdomains, known as chromosome territories (CTs). The first concept of a chromosome territory was proposed about 100 year ago by Boveri (1909), but was established only two decades ago in human cells (Lichter et al., 1988, Manuelides et al., 1988) using FISH with chromosome specific DNA probes. In plants the chromosome specific domains were first observed in monosomic addition lines using genomic in situ hybridization (reviewed in Schubert et al., 2001). Individual plant chromosomes were initially visualized in Arabidopsis using chromosome specific BAC DNA probes (Lysak et al., 2001, Pecinka et al., 2004) and later in related crucifer species (Lysak et al., 2006, Berr et al., 2006). In members of other plant families, such as medicago, rice and tomato, chromosome-specific BAC-FISH was successfully applied to detect small chromosomal subregions (Kulikova et al., 2001, Nagaki et al., 2004, Szinay et al., 2010).
Inside the chromosome territory the centromere and telomere regions occupy specific positions. While centromere regions are generally located near the nuclear periphery, the position of the telomeres varies not only between different species, but also between the stages of the cell cycle. For example, in human cells (3000 Mb) the centromeres are either attached to the nuclear membrane or more internal in the nucleus, depending on the cell cycle phase (Solovei et al., 2004), while telomeres are dynamic in the interior of the nucleus (Molenaar et al., 2003). Plant species with a large genome display a Rabl configuration of their chromosomes with centromeres and telomeres positioned facing opposite poles of the nucleus as a relic from the late anaphase of the preceding mitosis (Fig. 2, Rawlins et al., 1991, Aragón-Alcaide et al., 1997). However, comparative studies with several plant species suggest that a minimum chromosome size rather than the genome size accounts for the Rabl orientation (Dong et al., 1998). For example, in Crepis capillaris (1C = 2000 Mb, n=3) the chromosomes have a Rabl configuration, whereas chromosomes in Zea mays (1C = 2540 Mb, n=10) do not, despite the larger size of its genome. The most plausible explanation for this discrepancy is that Crepis chromosomes (~670 Mb) are significantly larger than the average size of a maize chromosome (~250 Mb).

Comparing chromosome size and genome size in different plant species now suggests that the critical chromosome size for a Rabl configuration is around 500 Mb. Chromosome size, however, is not the only factor that controls the orientation and position of chromosomes. In species with relatively small chromosomes various telomere dispositions have been reported (Fig. 2). In Arabidopsis (150 Mb) the telomeres are around the nucleolus, while centromeres localize in conspicuous heterochromatin domains, known as chromocenters, that are located at the nuclear periphery (Fransz et al., 2002). The telomeres in tomato (950 Mb), however, are at the edge of peripheral heterochromatic islands that accommodate the centromere (Fransz, 2004), whereas in rice (490 Mb) the telomeres and centromeres are dispersed around the nuclear periphery (Prieto et al., 2004). In bakers yeast (16 Mb) centromeres are also at the periphery opposite to the nucleolus, while telomeres are at the periphery near the nuclear pores (Jin et al., 2000, Saez-Vasquez and Gadal, 2010). Some species with small chromosomes even showed Rabl oriented chromosomes. The chromosomes in budding yeast display a Rabl-like orientation with a typical V-shape and clustered telomere ends. The telomeres, however, are not necessarily at the pole opposite to centromeres (Bystricky et al., 2005). The Rabl orientation of Drosophila chromosomes was investigated in atypical nuclei: giant polytene nuclei (Hochstrasser et al., 1986) and embryonic cells. The first group consists of endoreduplicated chromosomes, while the latter lacks a G1 phase. In other Drosophila cells the chromosomes display a non-Rabl orientation (Csink and Henikoff, 1998). Similarly, some cells in rice, such as xylem vessel cells, show a Rabl configuration of chromosomes. This may be explained by endoreduplication of xylem cells (Prieto et al., 2004), resembling the polytenic organisation of the salivary gland nuclei in Drosophila.

Apparently, apart from chromosome size there are more factors that affect the large-scale organization of chromosomes.
Steric constraints due to limited space have been proposed to play a major role in directing overall chromatin organization during interphase. Using a novel in vivo labeling technique, involving photoactivation, Muller et al. (2010) monitored an individual human chromosome in living cells and demonstrated a rapid overall decondensation of the mitotic chromosome upon entrance into G1 by a factor 2.5. The bulk of chromosome shape and territory was settled within 2 hours. This was attributed to a balance between expansion of the chromosome and spatial constraints imposed by surrounding chromosomes and nuclear envelope. Similarly, morphological features, such as the nucleolar organizing regions containing ribosomal genes (NOR), nuclear shape, endopolyploidy level and nuclear volume were proposed to direct chromosome arrangement in Arabidopsis (Berr et al., 2007, Schubert and Oud, 1997). These data were further substantiated by computational simulations of Arabidopsis chromosomes, showing that nuclear space affects organization and position of chromosomes (De Nooijer et al., 2009). Nuclear size has also been proposed to play a role in a most special case of chromosome positioning. Chromosomes in the retina of nocturnal mammals display an inverted pattern compared to other cells (Solovei et al., 2009). Instead of having a peripheral position the heterochromatin segments localize in the nuclear center, where they probably support the function of the retina cells in transmitting the light, by transforming the inverted nucleus into collecting lenses.

**Chromosome territory arrangement reflects the linear organization of the chromosome**

Based on the nuclear position of telomere and centromere we can infer folding patterns of a chromosome assuming that intermingling of different chromosome territories is absent or limited to the periphery of the territory (Cremer et al., 2006, Branco et al., 2006). The CT of a Rabl oriented chromosome has a tube-like shape in which the chromosome axis runs from a telomere at one pole to the opposite pole with the centromere and back to the telomere pool (Fig. 2). Both arms of a chromosome are running parallel in a V-shape manner with their telomeres in relatively close vicinity. In tomato nuclei both telomeres and centromere are on a heterochromatin island. This is probably due to the association of the heterochromatic blocks around the centromeres and in the subtelomeric region (Ramanna and Prakken, 1967, Zhong et al., 1998). In Arabidopsis the CT lies between the nuclear envelope and the nucleolus. The two arms are either in V-shape parallel or locate at opposite sides of the centromere. This is well illustrated in a CT study, where the two arms of chromosome 2 are differently painted (Berr et al., 2007). Interestingly, close examination of the images reveals no intermingling of the two arm territories, resembling the situation in mammalian cells where different parts of the same chromosome show no or little mixing (Goetze et al., 2007, Dietzel et al., 1998). Rice contains small chromomereres scattered along the chromosome arms (Cheng et al., 2001). This may be the reason why rice does not display discrete chromocenters during interphase (de Jong, unpublished).

*Homologous repeat regions rather than homologous chromosomes associate during interphase*
The development of chromosome painting in *Arabidopsis* has enabled to monitor association frequencies of chromosomes. Using differential painting of all five chromosomes Pecinka *et al.*, (2004) reported that side-by-side arrangement of heterologous chromosome territories occurs randomly for chromosomes 1, 3 and 5, while chromosomes 2 and 4 associate with each other more than expected on the basis of random grouping. This is probably due to the presence of the nucleolus organizing region (NOR) which is often closely connected with the nucleolus. These data corroborate the non-random association of chromocenters that contain the NOR (Fransz *et al.*, 2002) and supports the idea of “sticky” heterochromatin. The heterochromatic NORs at chromosome 2 and 4, which are mapped close to the distal telomeres, indeed colocalize with the pericentric heterochromatin that flank their respective centromeres forming a chromocenter. Moreover, homologous tandemly repetitive transgenes preferentially associate with each other and with heterochromatin rather than with normal euchromatin regions (Pecinka *et al.*, 2005). The association frequencies, however, may vary depending on construct, chromosomal insertion position, cell type and the number and repetitiveness of inserts (Jovtchev *et al.*, 2008). In all cases heterochromatin association has been shown an important factor in the higher level organization of chromosomes.

**Heterochromatin in small genome plant species is often clustered into chromocenters**

Chromocenters are discrete heterochromatin domains. They are well distinguishable under the microscope, due to the sharp transition to surrounding euchromatin, which probably reflects the organization of heterochromatin and euchromatin domains along the linear chromosome (De Jong *et al.*, 1999). In a study of 67 plant species from different families Ceccarelli *et al.* (1998) observed discrete chromocenters of which the number and size varied between plants and during development. In most species the maximum number of chromocenters corresponds to the chromosome number, pointing at one major heterochromatic block per chromosome. Since many plants contain several blocks of heterochromatin along the linear chromosome arm, we can conclude that these repeat-rich regions are associated with each other during interphase. Moreover, the average number of chromocenters is always less than the chromosome number, suggesting that chromocenters of different chromosomes are associated. This is well illustrated for Arabidopsis, which has on average eight distinct and well-defined chromocenters per nucleus (Fransz *et al.*, 2002). Since diploid cells in Arabidopsis have ten chromosome (n=5), three to four chromocenters are generally associated. Intriguingly, in 88% of the associated chromocenters the two centromeres of the chromosomes show one FISH signal at the core of the chromocenter instead of two separate spots (Pavlova *et al.*, unpublished data). This points at fusion of the two chromocenters and suggests intimate contact between the corresponding chromosome territories. Chromocenters in Arabidopsis contain the majority of the genomic repeats, including transposons, the tandemly arrayed centromeric 180 bp repeat and ribosomal genes (Fransz *et al.*, 2002). Consequently, chromocenters are heavily decorated with typical
epigenetic marks for inactive chromatin such as DNA methylation and histone H3K9 dimethylation and monomethylation (Soppe et al., 2002, Jasencakova et al., 2003, Probst et al., 2003) but also with H3K9me1, H3K27me1, H3K27me2 and H4K20me1 (Lindroth et al., 2005, Mathieu et al., 2005, Nauman et al., 2005). The repressive mark H3K27me3 is not in chromocenters, but in the surrounding gene-rich euchromatin, since it is involved in the repression of genes, especially genes associated with development.

Does chromosome territory organization correspond with genome activity?

The nuclear organization of the CT has implications for the subnuclear location of gene activity and the organization of other components in the nucleus. For example, a species like barley has large chromosomes that mainly consist of heterochromatin segments and a small subtelomere region accommodating of the most genes. Consequently, active chromatin and related molecular factors are located towards the telomere pole, whereas heterochromatin-related factors are positioned towards the opposite pole. Replication in barley starts in the nucleolar region and proceeds from telomere pole to centromere pole, following the euchromatin-heterochromatin axis (Jasencakova et al., 2001). The displacement of replication activity in the nucleus is different in Arabidopsis where the CT surrounds the peripheral chromocenter (Jasencakova et al., 2003).

Also the distribution of repressive epigenetic marks such as DNA methylation and H3K9me is different between species that differ in genome and/or chromosome size, reflecting the different nuclear positions of active and inactive chromatin. A comparative study on the distribution of epigenetic marks in 24 plant species resulted in two different patterns of histon H3K9me2: (1) strong H3K9me2 restricted to constitutive heterochromatin domains and (2) uniform nuclear distribution of H3K9me2 (Houben et al., 2003). The distribution pattern of H3K9me2 appears to correspond with genome size. Species with a genome size up to ~500 Mb have a type 1 pattern whereas a type 2 is observed in species with a larger genome. However, the linear distribution of repeat sequences also affects the H3K9me2 pattern, since rice (490 Mb) with dispersed repeats along the chromosome arms, shows a random distribution, whereas Ricinus communis (515 Mb) with discrete heterochromatin segments shows distinct domains of H3K9me2.

In Arabidopsis gene-rich regions show epigenetic marks for active chromatin such as H3K4me3, histone acetylation (Soppe et al., 2002, Jasencakova et al., 03) and histone variant H2A.Z (Deal et al., 2007) but also epigenetic marks for silent genes such as H3K27me3 (Nauman et al., 2005) and the heterochromatin protein LHP1 (Libault et al., 2005). Indeed chromosome painting with gene-rich BACs covering almost the entire genome showed only labeling outside chromocenters, indicating that the majority of the Arabidopsis genes, active and silent, are not in chromocenters (Fransz et al., 2006). Whether euchromatin regions in other species are also empty for transposons remains to be investigated.
Chromosomes have a loop organization and display long-range interactions

A generally accepted characteristic of chromosome organization is the formation of loop structures. All published chromosome models include chromatin loop configurations at one or the other level of chromosome organization (Munkel et al., 1999, Mateos-Langerak et al., 2009). Chromatin loops have been associated with several genome features, in particular in the context of gene activity. Chromatin regions can loop out of a condensed region when genes become active, such as the hox genes in mammals (Morey et al., 2009, Chambeyron et al., 2004). Small loops up to 100 kb are formed when enhancer and promoter physically interact to stimulate the expression of genes such as the β-globin locus in mammals (Tolhuis et al., 2002). The formation of loop structures has also been reported in repressive situations and may play an important role in gene silencing. For example, Polycomb group (PcG) complexes maintain the silent state of developmental gene regions. Via long-range interactions distant PcG targets can cluster into domains of repressed chromatin while the regions in between form loops (Tiwari et al., 2008). Similarly, genomic loci can interact at the nuclear lamina, a fibrillar network at the inner side of the nuclear membrane in animals, giving rise to a repressive environment for chromatin (Guelen et al., 2008).

Chromatin loops are difficult to visualize, with the exception of lampbrush chromosomes in amphibian oocytes. Only with FISH technology and sufficient resolution is it possible to visualize a chromatin loop structure (Fransz et al., 2002, Byrd et al., 2003, Heng et al., 2004). In general, the presence of loops is deduced via indirect evidence. For example, in Arabidopsis the subtelomeric nucleolar organizing region (NOR) in chromosome arms 2S and 4S colocalize with their pericentric heterochromatin regions in chromocenters, while the gene-rich regions in between are largely outside the chromocenter. Hence, the short arms form a loop structure of approximately 2 Mb if intermediate sequences are outside the chromocenter. In many cells, however, interstitial loci also associate with chromocenters giving rise to multiple smaller loops (Fransz et al., 2002). Based on these data the chromocenter-loop model was proposed to describe the organization of a chromosome territory in Arabidopsis (Fig 3d): a repeat-rich chromocenter from which euchromatin loops emanate. This organization reflects the distribution of genes and repeats along the linear chromosome arms. The chromosome arms are gene-rich (20-35 per 100 kb) and contain very few transposon elements (less than 5 per 100 kb), while the pericentromeric regions are gene-poor (less than 5 per 100 kb) (Fig. 4). The loop organization implies anchor points of euchromatin loops to the chromocenter. In this view it is tempting to speculate that the few transposons along the arms associate with pericentric repeats and anchor the euchromatic loops to heterochromatin domains. If this is true then the distribution of transposon elements along the chromosome arms suggests chromatin loop sizes spanning 100 kb to 2 Mb, which corresponds to the FISH data on interphase nuclei of Arabidopsis (Fransz et al., 2002).

Loop structures imply intimate association between distant loci without involvement of intervening sequences. The chromosome conformation capture (3C) technology is a powerful tool to study
chromatin loop formation, since it identifies physical associations between distant loci throughout the genome (Dekker et al., 2002, Tolhuis et al., 2002). The technique is based on crosslinking interacting or adjacent chromatin protein complexes by formaldehyde-fixation, followed by cutting the DNA strands with restriction enzymes and ligating their free ends. The sequence of ligated DNA fragments represents 2 loci that are physically associated in the protein complex. The reads include proximal and distant genomic loci. The 3C-based techniques have been widely adopted to investigate not only loops and long-range interactions but also the large-scale folding patterns of chromatin (Palstra et al., 2009, Fulwood et al., 2009, Lieberman-Aiden et al., 2009).

Recently, the 3C technique has been applied successfully in plants to reveal chromatin looping at the \textit{b1} locus in maize (Louwers et al., 2009). The \textit{b1} gene, involved in red pigmentation of husk leaves and other tissue, is regulated by a set of seven tandem repeats located \(\sim\)100 kb upstream of the \textit{b1} gene. The repeats function as a transcriptional enhancer. However, the molecular mechanism underlying the effect of the repeats on the \textit{b1} gene was not clear. Using 3C technology the authors demonstrated a physical interaction between the regulatory repeats and the \textit{b1} gene. In addition, an intermediate sequence also appeared to be involved in the regulation of \textit{b1}. The results denote a multiloop structure required to increase the expression of the \textit{b1} gene.

How chromatin sites are tethered to one place and held together forming a loop structure is still an unanswered question in cell biology. However, strong evidence is now available pointing at the nuclear matrix and associated sequences to play a major role in loop formation. The nuclear matrix has been proposed as a highly dynamic network of proteins that interacts with chromatin to facilitate domains of local activity or repression. Biochemically, the nuclear matrix is defined as the fraction that remains after high salt extraction, while the nucleosome structure is disrupted. The sequences that co-purify with matrix proteins are AT-rich and are termed matrix-attachment-region (MAR) or scaffold-attachment-region (SAR). MARs/SARs co-localize with insulators such as the gypsy element (Nabirochkin et al., 1998) and CTCF-binding sequences (Dunn et al., 2003), both of which mediate genomic interactions via association with distant sequences and form loop structures. As a consequence the region inside the loop is shielded from the influence of flanking chromatin sequences. CTCF is a multi zinc finger transcription factor that is known to block enhancer activity on human genes by binding to target sequences in between the enhancer and the gene. Recently, CTCF has been shown to recruit cohesin, which is required for sister chromatid cohesion, to target chromatin sites (Wendt et al., 2008, Nativio et al., 2009, Hadjur et al., 2009). It is proposed that the CTCF-cohesin complex connects two DNA sequences and stabilizes a loop structure.

A blast search did not reveal CTCF-like proteins in Arabidopsis (Heger et al., 2009). However, the action of the plant transcription factors ASYMMETRIC LEAVES1 (AS1) and AS2 in silencing the KNOX gene resembles the CTCF action in human cells. The AS1-AS2 complex binds to two specific sequences in the KNOX promoter to form a loop structure that represses KNOX expression during leaf development (Guo et al., 2008). The interaction involves the chromatin-remodelling factor HIRA.
Similarly, CTCF action involves ATP-dependent chromatin remodeling factors (Fu et al., 2008). Here CTCF provides an anchor point for nucleosome positioning, thus linking two levels of chromatin organization (nucleosomes and loops). Since CTCF-binding is sensitive to DNA methylation and DNA methylation is directed by nucleosome occupancy, it follows that chromatin looping is controlled by nucleosome positioning, which in turn is regulated by DNA sequence and epigenetic modifications.

**DYNAMICS IN CHROMATIN COMPACTION**

*Decondensation and recondensation of heterochromatin domains*

Chromocenters represent the domains of one or more heterochromatin blocks and can be visualized with DAPI and C-banding. According to Brown (1966) they contain constitutive heterochromatin, which remains condensed throughout the cell cycle. Arabidopsis chromocenters are the C-band positive domains (Ambros and Schweizer 1976) and therefore considered constitutive heterochromatin. However, when differentiated mesophyll cells are transformed into protoplasts and brought into culture, large-scale decondensation of heterochromatin domains occurs in the majority of nuclei after few hours (Tessadori et al., 2007a). The process is accompanied by decondensation of all major repeat arrays with a decompaction rate up to 14 times. Even the tandemly arranged centromeric 180 bp repeat regions undergo drastic decondensation. The phenomenon is reminiscent of the ‘short lasting dispersion phase’ where all heterochromatin disappears in the nuclei (Barlow et al., 1976). These data suggest that constitutive heterochromatin is not as permanently condensed as we assumed (Fig. 3). If we consider a compaction degree for heterochromatin domains of ~1Mb/µm (Fransz et al., 2002), then in protoplasts it can unfold to about 71 kb/µm, which falls within the compaction range of the 30 nm fiber (70-160 kb/µm). A similar level of compaction was monitored for some euchromatin regions in leaf nuclei. Apparently, all chromatin in a cell can adopt different degrees of compaction from the highly condensed metaphase (16Mb/µm ) down to the level of the 30 nm fiber in decondensed interphase chromatin of protoplasts (70 kb/µm ). Large-scale decompaction of chromatin is however not restricted to protoplasts, since similar observations of heterochromatin decondensation although less pronounced were found during developmental changes such as seedling development (Mathieu et al., 2003) and the floral transition (Tessadori et al., 2007b), but also in mutants that affect chromatin remodelling such as the DNA methylation mutants *ddm1*, an ATP-dependent chromatin remodeler, and *met1*, a DNA methyltransferase, (Soppe et al., 2002) and *hda6*, a histone deacetylase, (Probst et al., 2004). The *ddm1*-5 mutant, which lacks detectable levels of *DDM1* transcripts displays a severe chromatin phenotype, in which all pericentric heterochromatin including the 180 bp centromere repeat is decondensed (Mittelsten-Scheid et al., 2002, Probst et al., 2003). It indicates the important role of the chromatin remodeler DDM1 in the formation of heterochromatin domains.
Recently, chromatin decondensation has been observed under biotic and abiotic stress situations such as bacterial infection (Pavet, et al., 2006) low light (Tessadori et al., 2009, van Zanten et al., 2010a) and high temperatures (Pecinka et al., 2010). The data suggest that large-scale chromatin decompaction accompanies changes in the nuclear program in response to a developmental transition or (a)biotic stress conditions. Moreover, chromatin decompaction is a reversible process. Chromocenter reconstitution occurs within a few days or upon reversion to the original situation. This is well illustrated in a low-light experiment where three weeks old plants were deprived from normal light levels (200 µmol/m².s⁻¹), by transferring to 15 µmol/m².s⁻¹ for 4 days and returned to normal light conditions (van Zanten et al., 2010a). The percentage of nuclei with normal heterochromatin content decreased from 80% to 10%, in favor of a reduced heterochromatin phenotype, and returned back to 80% after increasing the light to normal conditions. Apparently, large-scale decondensation is a transient state of chromatin in response to changing conditions. Interestingly, reformation of chromocenters occurs via a process of sequential steps, in which the longest arrays of tandem repeats (45S rDNA genes) are first condensed, followed by smaller tandem arrays (centromere repeats and 5S rDNA genes) and finally the dispersed transposon elements (Tessadori et al., 2007a). It suggests that the repeat length is an important factor in chromocenter formation.

**Light affects chromatin compaction**

Light-induced chromatin (de)compaction is controlled via cryptochrome (CRY) and phytochrome (PHY) photoreceptors, since their mutants display no or limited reduction of heterochromatin under low-light stress (van Zanten et al., 2010a). The authors suggest that stabilization of CRY2 in low light triggers chromatin decondensation. A similar role for CRY2 was proposed during the floral transition (Tessadori et al., 2007b). Both observations suggest a relatively short signalling pathway, since the blue light photoreceptor is constitutively located inside the nucleus (Guo et al., 1999). Moreover, CRY2 is associated with chromatin (Cutler et al., 2000) and interacts with CIB1, a transcription factor, which binds to promoter sequences (Liu et al., 2008). In addition, CRY2 represses the COP1/DET/FUS photomorphogenesis complex via the RING finger ubiquitin ligase COP1 (Wang et al., 2001). Interestingly, the DET1 component binds to the non-acetylated N-terminal tail of the core histone H2B and is supposed to facilitate a condensed state of chromatin (Benvenuto et al., 2002). Based on these findings it is suggested that the photoreceptors control chromatin compaction via a chromatin protein complex (Tessadori et al., 2007b, van Zanten et al., 2010a), that may contain members of E3 ligase complexes such as COP1, which interacts with nuclear photoreceptors (Yi et al., 2005). The chromatin protein complex may also contain HDA6, which is also involved in light-regulated chromatin compaction (Tessadori et al., 2009).

**Is chromatin compaction independent of epigenetic marks?**

Changes in chromatin are generally accompanied by epigenetic changes of histone marks or DNA
methylation. Surprisingly, no large-scale alterations in H3K9me2 or 5-methyl cytosine was observed during chromatin decondensation in protoplasts (Tessadori et al., 2007a) suggesting that chromatin compaction and decondensation can occur independent of these marks. Recently, three studies from different labs show that heat stress to Arabidopsis plants reactivates silent transcriptional elements with no or little alteration of epigenetic marks for silencing (Tittel-Elmer et al., 2010, Lang-Mladek et al., 2010, Pecinka et al., 2010). In all cases the activation of transgenes and endogenous repeat elements occurred without loss of DNA methylation. In addition, H3K9me2, H3K27me2 and H3K27me3 remained unaffected (Tittel-Elmer et al., 2010), while histone acetylation increased. Strikingly, the increase of gene activity by heat stress was accompanied by loss of chromocenter organization and decondensation of chromatin. Moreover, a dramatic loss of histone H3 and H4 was observed indicating depletion of nucleosomes (Pecinka et al., 2010, Lang-Mladek et al., 2010). Similar to the low light response the heat-induced state is reversible, since recovery from the heat stress results in restoration of nucleosome loading and gene silencing. These data suggest that environmental conditions can overrule epigenetic control of gene activity, at least transiently. The heat stress response may be considered as a molecular mechanism with direct control over nucleosome positioning without significant interference of epigenetic regulatory factors.

Recently it has been shown that temperature regulated gene expression is controlled by the histone variant H2A.Z (Kumar et al., 2010). At low temperature (17°C) H2A.Z nucleosomes are bound just downstream of the TSS in many active and inactive genes of Arabidopsis. In this situation the transcriptional activity of a gene remains unchanged. Raising the temperature to 27°C results in thermal instability at the H2A.Z nucleosome leading to a decrease in H2A.Z occupancy at the TSS and a change in gene activity. This occurs in temperature responsive genes such as the FT gene, which is involved in flowering time. Plants deficient in ARP6, a component of the SWR1 remodelling complex, which establishes H2A.Z deposition, are unable to respond to temperature changes, because the nucleosomes contain the canonical histone H2A. In the same study similar results were obtained with budding yeast. The data imply that H2A.Z nucleosomes mediate the thermosensory response in both plants and yeast. Compared to DNA at nucleosomes containing H2A, the DNA at H2A.Z nucleosomes is more tightly wrapped, but in a temperature dependent way. This allows a direct mechanism for chromatin to respond to temperature fluctuations.

It is not clear if there is a causal relationship between the loss of nucleosomes from individual transcription elements and the large-scale decondensation of chromatin involving entire heterochromatin domains. The two levels of chromatin organization are physically connected, but not necessarily causally related. The recovery of the heat shock did not lead to heterochromatin recompaction (Pecinka et al., 2010), while light-induced decondensation of chromocenters did not change nucleosome density in repeat regions (Van Zanten et al., 2010b).

Cells in mammals are less responsive to environmental changes compared to plants. Yet, we may find an equivalent of chromocenter reformation in cell differentiation during major developmental
switches. For example, following the fertilization of the mouse egg there is a dynamic reorganization of heterochromatin domains with complete absence of chromocenters. The changed configuration involves the major and minor satellite repeats, which map to the pericentromere and centromere region, respectively (Probst et al., 2007). Reformation of chromocenters starts in the two-cell stage embryo. Intriguingly, recondensation of the pericentric heterochromatin is associated with transcriptional activity of the major satellite repeats (Probst et al., 2010). Moreover, transcription of the repeats occurs exactly when chromocenters are being formed and are essential for proper development of the mouse embryo.

Conclusions and perspectives

Eukaryote organisms have their genetic and epigenetic information stored in a polymer construction of nucleosomes that is continuously meeting numerous, different protein complexes. These interact with the genomic sequence to protect, repair and replicate the DNA or to retrieve the correct information at the proper time and in the right place. The basic components such as histones, chromatin modifiers and other chromatin proteins are highly conserved among all eukaryotes. The same holds true for the molecular mechanisms that rule the accessibility of the DNA sequence. This is exemplified by the nucleosome structure, nucleosome positioning and higher order folding patterns of chromatin such as loop formation, but also the nucleosome-controlled DNA methylation in Arabidopsis and human cells. The precise biological meaning of large-scale chromatin decompaction in plants remains to be elucidated, but is likely related with the reorganization of chromatin and the rearrangement of DNA accessibility in response to a developmental switch or a stress situation. Since plants are sessile organisms and cannot escape from changing environmental conditions, they must respond in a rapid and adequate way by the nuclear program. Chromatin decondensation may be part of a mechanism that facilitates such a response.

Differences in nuclear organization of chromosomes between organisms are determined by genomic sequence, but also by genome size and the distribution of repeats and genes along the linear sequence. Chromatin is highly dynamic at different levels of organization. How the different configurations, such as loop structures, relate with genome activity is currently investigated with different techniques. The development of ChIP-seq, 3C and related methods, in combination with massive parallel sequencing technology enables us to analyse at a genome-wide scale chromatin profiles long range chromatin interactions and chromosome folding patterns in large genome organisms (Lieberman-Aiden et al., 2009). The huge datasets require extensive bioinformatic tools in order to efficiently transform data into information and subsequently into models in order to fully understand the structure-function relation of chromosomes.
Acknowledgements

We wish to acknowledge Dóra Szinay, Xianwen Ji, Laksana Kantama, Penka Pavlova and Federico Tessadori for providing the images of barley, tomato, rice and Arabidopsis.
References


Science, 295, 1306-1311.


Fan, J.Y., Rangasamy, D., Luger, K. and Tremethick, D.J. (2004) H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. Mol Cell, 16, 655-661


Ramanna M.S. and Prakken, R. (1967) Structure of and homology between pachytene and somatic metaphase chromosomes of the tomato. *Genetica, 38*, 115-133


Legends

Figure 1.
Structure and arrangement of nucleosomes. (a) Top view of nucleosome showing the DNA helix (grey) wrapped around an octamere of histones (coloured). (b) Top view of the one start helical solenoid model. Nucleosomes form a spiral arrangement. (c,d) Side view of the 2-start helix configuration. Histone tails (grey lines) protrude from the nucleosome (c) and following nucleosomes display a zig-zag structure (d). (Adapted from Dorigo et al., 2004, Wong et al., 2007)

Figure 2.
Diagrammatic representations of heterochromatin distribution along the linear chromosome, chromosome territory organization and DAPI stained nuclei in plants with different chromosome sizes. All plants have their centromeres (red) towards the nuclear periphery, surrounded by pericentric heterochromatin (dark grey), while telomeres (green) occupy different positions. The chromosomal axes (waved line) run from centromere to telomere. Plants with large chromosomes, such as Hordeum vulgare, show a typical Rabl orientation with centromeres and telomeres at opposite poles of the nucleus. The major part of the CT is occupied by heterochromatin. Plants with a tomato type arrangement have their telomeres at the edge of the heterochromatic islands. The Arabidopsis type has the telomeres close to the nucleolus. The rice type has more diffuse heterochromatin segments along the arms and telomeres are towards the nuclear periphery.

Figure 3.
Chromocenter-loop organization changes in Arabidopsis nuclei with decondensed chromatin in cultured protoplasts. The images (a,b,e,f,i,j) show DAPI stained nuclei and FISH signals of 45S rDNA (green) and 5S rDNA (red) genes. (a-d) Condensed situation with 45S rDNA and 5S rDNA colocalized in chromocenter #4. Chromosome territory (d) shows a typical chromocenter-loop arrangement of repeats (colored) and euchromatin loops (light grey). (e-h) Decondensed situation with less and smaller chromocenters. The subtelomeric 45S rDNA and pericentric 5S rDNA loci of chromosome 4 are no longer colocalizing. Consequently, the 45S rDNA does not form a loop anymore. (i-l) Fully decondensed chromocenters. Only the long repeat arrays of 45S rDNA genes of chromosomes 2 and 4 remain partially condensed and colocalize to one domain. blue, centromere repeat, green, 45S rDNA, red, 5S rDNA, purple, TEs.

Figure 4.
Scatter plot showing the distribution of genes (blue) and transposons (TEs, pink) along the linear chromosome #4 of Arabidopsis. The ribosomal genes at the subtelomeric end of the short arm are not included (data from ftp.arabidopsis.org).
Figure 1
Fransz & de Jong
Figure 2
Fransz & de Jong

Hordeum vulgare

Solanum lycopersicum

Arabidopsis thaliana

Oryza sativa
Figure 3
Fransz & de Jong
Figure 4
Fransz & de Jong