

Analysis of the epigenetic status of telomeres by using ChIP-seq data

María I. Vaquero-Sedas¹, Chongyuan Luo² and Miguel A. Vega-Palas^{1,*}

¹Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC-Universidad de Sevilla, IBVF (CSIC-USE), c/ Américo Vespucio n° 49, 41092 Seville, Spain and ²Department of Plant Biology and Pathology, Rutgers University, New Brunswick, NJ, USA

Received June 8, 2012; Accepted July 6, 2012

ABSTRACT

The chromatin structure of eukaryotic telomeres plays an essential role in telomere functions. However, their study might be impaired by the presence of interstitial telomeric sequences (ITSs), which have a widespread distribution in different model systems. We have developed a simple approach to study the chromatin structure of *Arabidopsis* telomeres independently of ITSs by analyzing ChIP-seq data. This approach could be used to study the chromatin structure of telomeres in some other eukaryotes. The analysis of ChIP-seq experiments revealed that *Arabidopsis* telomeres have higher density of histone H3 than centromeres, which might reflect their short nucleosomal organization. These experiments also revealed that *Arabidopsis* telomeres have lower levels of heterochromatic marks than centromeres (H3K9^{Me2} and H3K27^{Me}), higher levels of some euchromatic marks (H3K4^{Me2} and H3K9Ac) and similar or lower levels of other euchromatic marks (H3K4^{Me3}, H3K36^{Me2}, H3K36^{Me3} and H3K18Ac). Interestingly, the ChIP-seq experiments also revealed that *Arabidopsis* telomeres exhibit high levels of H3K27^{Me3}, a repressive mark that associates with many euchromatic genes. The epigenetic profile of *Arabidopsis* telomeres is closely related to the previously defined chromatin state 2. This chromatin state is found in 23% of *Arabidopsis* genes, many of which are repressed or lowly expressed. At least, in part, this scenario is similar in rice.

INTRODUCTION

Telomeres prevent chromosome fusions and degradation by exonucleases and are implicated in DNA repair, homologous recombination, chromosome pairing and segregation. Telomeric DNA usually contains tandem repeats of a

short GC-rich motif, which can also be found at interstitial chromosomal loci (1–5). These interstitial telomeric sequences (ITSs) have a widespread distribution in different model systems, including *Arabidopsis*, and have been related to chromosomal aberrations, fragile sites, hot spots for recombination and diseases caused by genomic instability, although their functions remain unknown (5–8).

Two major chromatin organizations can be found inside the cell nucleus: heterochromatin and euchromatin. Heterochromatic regions are highly condensed in interphase nuclei giving rise to the so-called chromocenters and usually associate with repetitive and silent DNA, although certain level of transcription is required for their establishment and maintenance. By contrast, euchromatic regions have an open conformation and are often related to the capacity to be transcribed. Both kinds of chromatin organizations exhibit defined epigenetic modifications that influence their biochemical behavior. In *Arabidopsis*, chromocenters contain pericentromeric heterochromatin, which associates with the 178-bp satellite repeats (also known as 180-bp repeats) and with other repetitive DNA sequences including mobile elements and ITSs (9–15). *Arabidopsis* heterochromatin is characterized by high levels of cytosine methylation, which can be targeted at CpG, CpNpG or CpNpN residues (where N is any nucleotide), and by H3K9^{Me1,2}, H3K27^{Me1,2} and H4K20^{Me1}. In turn, *Arabidopsis* euchromatin is characterized by H3K4^{Me1,2,3}, H3K36^{Me1,2,3}, H4K20^{Me2,3} and by histones acetylation (16). In addition, many genes that localize in *Arabidopsis* euchromatin are labeled with H3K27^{Me3}, a repressive mark that is thought to regulate tissue-specific gene expression (17–19).

The analysis of telomeric chromatin structure from ChIP, ChIP-on-chip or ChIP-seq experiments might be challenged by the presence of ITSs (20). This problem might also be extensive to other repetitive sequences. Here, we have developed an approach to study the epigenetic modifications of *Arabidopsis* telomeres independently of ITSs by analyzing genome-wide ChIP-seq data. The ChIP-seq experiments revealed that *Arabidopsis* telomeres

*To whom correspondence should be addressed. Tel: +34 954 489574; Fax: +34 954 460065; Email: vega-palas@ibvf.csic.es

have higher density of histone H3 than centromeres. These experiments also revealed that Arabidopsis telomeres have lower levels of heterochromatic marks than centromeres (H3K9^{Me2} and H3K27^{Me}), higher levels of some euchromatic marks (H3K4^{Me2} and H3K9Ac) and similar or lower levels of other euchromatic marks (H3K4^{Me3}, H3K36^{Me2}, H3K36^{Me3} and H3K18Ac). Interestingly, the ChIP-seq data also revealed that Arabidopsis telomeres exhibit higher levels of H3K27^{Me3} than centromeres. At least, in part, this scenario is similar in rice.

MATERIALS AND METHODS

Determination of the relative amounts of (CCCTAAA)₄ sequences at telomeres and ITSs

To analyze the chromatin structure of Arabidopsis telomeres using genome-wide ChIP-seq experiments, we had to define a specific DNA sequence that revealed telomeres but not ITSs. For that purpose, we estimated the number of times that the sequence (CCCTAAA)₄ appears at internal chromosomal loci and at telomeres in the *Arabidopsis thaliana* (Col-0) genome. First, we performed Blast analyses at the Map Viewer web site in National Center for Biotechnology Information (NCBI) to determine the number of times that the sequence (CCCTAAA)₄ appears at internal chromosomal loci (<http://www.ncbi.nlm.nih.gov/mapview>). In the case that a specific ITS contained five perfect tandem telomeric repeats, the Blast analyses revealed two overlapping (CCCTAAA)₄ sequences. If the ITS contained six perfect tandem telomeric repeats, the Blast analyses revealed three overlapping (CCCTAAA)₄ sequences and so on. We found 118 (CCCTAAA)₄ sequences at internal positions in the five chromosomes of Arabidopsis, including subtelomeric regions.

To estimate the number of times that the sequence (CCCTAAA)₄ is found at Arabidopsis telomeres, we assumed that *Arabidopsis thaliana* (Col-0) telomeres are composed of perfect telomeric repeats that spread about 3750 bp (21,22). We estimated that the five Arabidopsis chromosomes should contain about 5350 overlapping (CCCTAAA)₄ sequences at telomeres [(3750/7) × 10]. Therefore, when the frequency of reads containing the sequence (CCCTAAA)₄ is determined in input samples of Arabidopsis ChIP-seq experiments, only 2% of these reads should correspond to ITSs [(118 × 100)/(118 + 5350)]. In consequence, the frequency of the (CCCTAAA)₄ sequence should essentially reveal telomeres in Arabidopsis ChIP-seq experiments. In the case of rice (*Oryza sativa* ssp. japonica cv. Nipponbare), the frequency of the (CCCTAAA)₄ sequence should also reveal telomeres. The rice genome contains a similar number of (CCCTAAA)₄ sequences at internal chromosomal loci than Arabidopsis (127), has more chromosomes (twelve) and similar telomeric length (23).

Determination of enrichment values

We used the Sequence Read Archive database at NCBI to study different epigenetic modifications in Arabidopsis and in rice telomeres. In Arabidopsis, we analyzed all the experiments from study SRP002100 (Gene Expression

Omnibus accession number GSE28398). This ChIP-seq study was performed using aerial tissue (24). We determined the number of reads containing the (CCCTAAA)₄ sequence and the number of reads containing the sequence TTGGCTTTGTATCTTCTAACAAG, which is a conserved region of the 178-bp satellite repeats present at centromeres. This sequence served as heterochromatic reference (12,15). Although a fraction of the 178-bp sequences associates with CENH3 chromatin, surrounding 178-bp repeats associate with H3 chromatin. The conserved sequence that we have selected spans from positions 56 to 79 of the repeats and do not contain motifs specifically associated with CENH3 chromatin (15). Thus, this sequence is present at the centromeric 178-bp repeats that associate with CENH3 chromatin and also at the 178-bp repeats that associate with H3 chromatin. It allowed us to analyze the chromatin organization of Arabidopsis 178-bp satellite repeats as an average, which is known to be heterochromatic (9–11,15,25).

The study mentioned earlier in the text focused on nine epigenetic modifications named H3K4^{Me2}, H3K4^{Me3}, H3K9^{Me2}, H3K27^{Me}, H3K27^{Me3}, H3K36^{Me2}, H3K36^{Me3}, H3K9Ac and H3K18Ac. The number of telomeric and centromeric reads corresponding to all these epigenetic marks were determined (Supplementary Table S1) and normalized against the input sample. Then, enrichment values of telomeres versus centromeres were calculated and normalized against histone H3 occupancy. The resulting values are represented in Figure 1. Similar results to those shown in Figure 1 were obtained when a different conserved sequence from the 178-bp satellite repeats was used to estimate relative enrichment values (data not shown). This sequence (CATATTTGACTCCAAAACA CTAA) contains the dinucleotide TG at positions 160–161 of the 178-bp repeats and is not frequent in CENH3 chromatin (15).

The enrichment values for H3K4^{Me3} and H3K9Ac were also calculated by analyzing experiments from a different ChIP-seq study (SRP002650; GSE22276), which was performed using Arabidopsis leaves (26). The enrichment values obtained for these two epigenetic marks were very similar to those shown in Figure 1: 0.7 for H3K4^{Me3} and 3.3 for H3K9Ac.

For rice (*O. sativa* ssp. japonica cv. Nipponbare), enrichment values were calculated by analyzing experiments SRX016118, SRX016122, SRX016126 and SRX016130 from study SRP001788 (GSE19602). These experiments were performed using four-leaf stage seedlings (27). The number of reads containing the (CCCTAAA)₄ sequence and the number of reads containing the sequence CGTTC GTGGCAAAAACACTTCGT, which is part of the CentO-1 centromeric satellite repeat (positions 1–25) (28), were determined for four epigenetic modifications (H3K4^{Me3}, H3K9Ac, H3K27^{Me3} and DNA methylation; see Supplementary Table S2). The CentO repeats, which are known to undergo DNA methylation similarly to the 178-bp repeats from Arabidopsis, served as heterochromatic reference (29–31). Since study SRP001788 did not include input samples, relative enrichment values were calculated. The number of reads corresponding to the

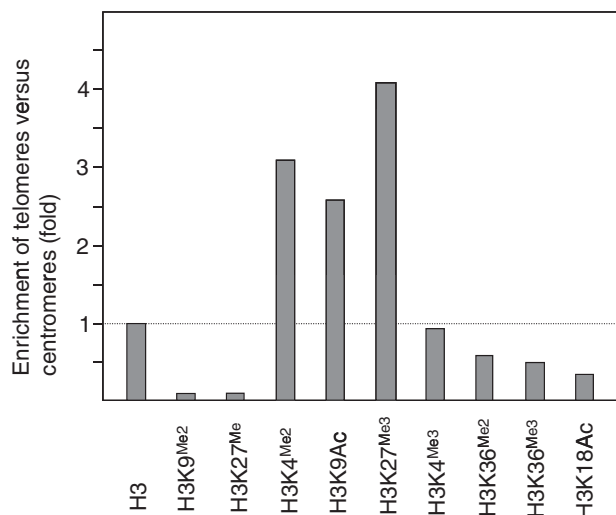


Figure 1. Epigenetic features at Arabidopsis telomeres. Different histone H3 epigenetic modifications were analyzed at telomeres and at centromeres using study SRP002100 (Gene Expression Omnibus accession number GSE28398). Relative enrichment values of telomeres versus centromeres were determined and represented as indicated in the ‘Materials and Methods’ section.

telomeric sequence was divided by the number of CentO-1 reads for every epigenetic modification. Then, relative enrichment values were calculated by normalizing against the resulting value for DNA methylation. To obtain a graphic representation with similar values to those shown in Figure 1, all the relative enrichment values were divided by 10.

RESULTS

Epigenetic marks at Arabidopsis telomeres

To study the epigenetic modifications present at Arabidopsis telomeres, we analyzed ChIP-seq data. We determined the relative enrichment of telomeres versus centromeres, which served as heterochromatic reference. Previously, we had to find a specific sequence that represents telomeres but not ITSs in ChIP-seq analyses. Since Arabidopsis ITSs are mostly composed of very short stretches of perfect telomeric repeats interspersed with degenerated repeats (7,8,32,33), a short stretch of perfect telomeric repeats might essentially represent telomeres. Blast analyses of the Arabidopsis genome revealed that 98% of the (CCCTAAA)₄ sequences are found at telomeres, whereas only 2% of these sequences localize at ITSs (see ‘Materials and Methods’ section). Therefore, the (CCCTAAA)₄ sequence is essentially found at Arabidopsis telomeres. We chose this sequence to analyze telomeres in ChIP-seq experiments. For centromeres, we selected a conserved sequence from the 178-bp centromeric satellite repeats. It is well known that these repeats are heterochromatic and localize to chromocenters (9–13,15,25).

We analyzed ChIP-seq data from a study that focused on different histone H3 modifications in Arabidopsis. The epigenetic modifications analyzed in this study were two characteristic heterochromatic marks (H3K9^{Me2} and

H3K27^{Me3}) and six marks associated with euchromatin (H3K4^{Me2}, H3K4^{Me3}, H3K36^{Me2}, H3K36^{Me3}, H3K9Ac and H3K18Ac). In addition, this study also focused on unlabeled histone H3 and on H3K27^{Me3}, a repressive mark found in many euchromatic genes (17–19). We determined the enrichment of telomeres versus centromeres and found that the levels of unlabeled histone H3 were 1.6 times higher at telomeres (Supplementary Table S1). In all eukaryotic organisms analyzed, the nucleosomal spacing of telomeres is ~20 bp shorter than the average bulk nucleosome organization (34). This important chromatin feature is probably related to the straight nature of telomeric DNA and might condition the epigenetic behavior of telomeric nucleosomes (32,35–38). The high density of histone H3 determined here at telomeres might reflect the short spacing of telomeric nucleosomes, which has been previously revealed by micrococcal nuclease digestion experiments. Alternatively, we cannot rule out other possibilities including a nucleosomal spacing of the nucleosomes associated with the 178-bp repeats longer than the average bulk nucleosome organization.

We found that the levels of heterochromatic marks analyzed (H3K9^{Me2} and H3K27^{Me3}) were lower at telomeres than at centromeres. On the contrary, the levels of some euchromatic marks were higher at telomeres (H3K4^{Me2} and H3K9Ac; Figure 1). These results indicate that Arabidopsis telomeres exhibit euchromatic features. However, not all the euchromatic marks analyzed were enriched in telomeres versus centromeres. That was the case for H3K4^{Me3}, H3K36^{Me2}, H3K36^{Me3} and H3K18Ac. The levels of these marks were similar or even lower at telomeres than at centromeres. Interestingly, Arabidopsis telomeres were also enriched in H3K27^{Me3} (Figure 1). Since we have previously found that telomeres in Arabidopsis are marked with H4K16Ac (14), we conclude that Arabidopsis telomeres are labeled with H3K4^{Me2}, H3K9Ac, H3K16Ac and H3K27^{Me3}.

Epigenetic marks at rice telomeres

To study whether some of the epigenetic marks that characterize Arabidopsis telomeres are present in other plant species, we analyzed a genome-wide study performed in rice where the levels of H3K4^{Me2}, H3K9Ac, H3K27^{Me3} and DNA methylation were analyzed (27). Since the (CCCTAAA)₄ sequences present at rice ITSs represent ~1% of the total, these sequences also reveal telomeres in rice ChIP-seq experiments (see ‘Materials and Methods’ section). As heterochromatic reference, we analyzed a sequence from CenO, which is a satellite repeat present in rice centromeres. We found that the relative enrichment levels of telomeres versus centromeres were more than one order of magnitude higher for H3K4^{Me2}, H3K9Ac and H3K27^{Me3} than for DNA methylation (Figure 2). These results strongly suggest that rice telomeres are also labeled with H3K4^{Me2}, H3K9Ac and H3K27^{Me3}. However, these results are still compatible with the existence of low levels of DNA methylation at rice telomeres, which can also be present at Arabidopsis telomeres (14).

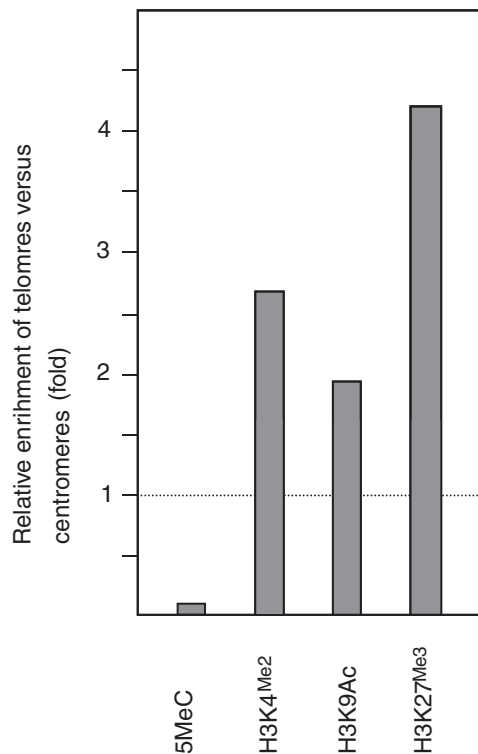


Figure 2. Epigenetic features at rice telomeres. Different histone epigenetic modifications were analyzed at telomeres and at centromeres using study SRP001788 (GSE19602). Relative enrichment values of telomeres versus centromeres were determined and represented as indicated in the 'Materials and Methods' section.

DISCUSSION

ChIP-seq data confirm that Arabidopsis telomeres exhibit euchromatic features

We have previously analyzed the chromatin structures of Arabidopsis telomeres and ITSs by ChIP, digestion with the restriction endonuclease Tru9I, which digested ITSs, and hybridization with a telomeric probe. These experiments revealed that, although subtelomeric regions and ITSs associate with heterochromatic marks, Arabidopsis telomeres exhibit euchromatic features (14). More specifically, we found that telomeres have lower levels of heterochromatic marks than ITSs (H3K9^{Me2}, H3K27^{Me} and DNA methylation) and higher levels of euchromatic marks (H3K4^{Me2}, H3K9Ac and H4K16Ac). Here, we have shown by analyzing ChIP-seq data that telomeres have lower levels of heterochromatic marks than centromeres (H3K9^{Me2} and H3K27^{Me}) and higher levels of some euchromatic marks (H3K4^{Me2} and H3K9Ac). Thus, these data confirm our previous results and validate the methodological approach presented here.

The experimental approach that we have used here to study the chromatin structure of Arabidopsis telomeres independently of ITSs could be applied to other eukaryotes. It will depend on how well genomes are assembled, including subtelomeric regions, and also on the structure of ITSs, including their length and number of perfect tandem repeats.

Arabidopsis telomeres are labeled with H3K27^{Me3}

We have shown that Arabidopsis telomeres are labeled with H3K27^{Me3}, a repressive mark that usually do not colocalize with DNA methylation. Recently, ~4400 genes have been found to be labeled with this epigenetic modification. These genes are often expressed in a tissue-specific manner and many of them might be involved in plant development. H3K27^{Me3} is established by the Polycomb Repressive Complex 2. This complex is conserved along evolution and is present in Arabidopsis, where its disruption causes serious developmental problems (18,19,39–42). It will be interesting to ascertain whether some of these problems are related to telomeres biology.

The histone code hypothesis involves a complex interplay of writing, reading and erasing activities that regulate the levels and the output of different epigenetic modifications (43). We have previously found that Arabidopsis telomeres are labeled with H3K4^{Me2}, H3K9Ac and H4K16Ac. Here, we have confirmed these results and have also shown that Arabidopsis telomeres are labeled with H3K27^{Me3}. The analysis of telomeres in mutants affected in the establishment, recognition and/or erasing of all these epigenetic marks open new avenues for future telomeric studies.

The chromatin organization of Arabidopsis telomeres shares similarities with the previously reported chromatin state 2

Four main chromatin states (CS1–CS4) have been recently defined in Arabidopsis based on the genome-wide distribution of 12 epigenetic marks analyzed by ChIP-on-chip (44–46). Since six of these marks were also analyzed here (H3K4^{Me2}, H3K4^{Me3}, H3K9^{Me2}, H3K27^{Me}, H3K27^{Me3} and H3K36^{Me3}), we compared both studies and found that the chromatin organization of Arabidopsis telomeres is most similar to chromatin state 2 (CS2). This chromatin state is associated with 23% of Arabidopsis genes, being many of them repressed or lowly expressed. CS2 is characterized by the presence of H3K4^{Me2}, H3K27^{Me2} and, above all, H3K27^{Me3}. In addition, CS2 is also characterized by the absence of H3K4^{Me3}, H3K9^{Me2}, H3K9^{Me3}, H3K27^{Me}, H3K36^{Me3}, H3K56Ac, H4K20^{Me}, H2Bub and ⁵MeC (45).

ACCESSION NUMBERS

GSE28398, GSE22276 and GSE19602.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2.

ACKNOWLEDGEMENTS

We want to thank the laboratories of Eric Lam, Jeffrey Chen and Xing-Wang Deng for making available their ChIP-seq data at the public databases. We are especially

grateful to Eric Lam for allowing us to use their data prior to publication. We also thank NCBI for providing the tools that we have used to analyze the ChIP-seq data.

FUNDING

The Spanish Ministry of Education and Science [grant BFU2008-02497/BMC] and by FEDER funds. Funding for open access charge: Waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

- Blackburn,E. (2005) Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett.*, **579**, 859–862.
- Cech,T. (2004) Beginning to understand the end of the chromosome. *Cell*, **43**, 405–413.
- Greider,C. (2006) Telomerase RNA levels limit the telomere length equilibrium. *Cold Spring Harb. Symp. Quant. Biol.*, **71**, 225–229.
- Meyne,J., Baker,R., Hobart,H., Hsu,T., Ryder,O., Ward,O., Wiley,J., Wurster-Hill,D., Yates,T. and Moyzis,R. (1990) Distribution of non-telomeric sites of the (TTAGGG)_n telomeric sequence in vertebrate chromosomes. *Chromosoma*, **99**, 3–10.
- Richards,E., Goodman,H. and Ausubel,F. (1991) The centromere region of *Arabidopsis thaliana* chromosome 1 contains telomere-similar sequences. *Nucleic Acids Res.*, **19**, 3351–3357.
- Lin,K. and Yan,J. (2008) Endings in the middle: current knowledge of interstitial telomeric sequences. *Mutation Res.*, **658**, 95–110.
- Regad,F., Lebas,M. and Lescure,B. (1994) ITSs within the *Arabidopsis thaliana* genome. *J. Mol. Biol.*, **239**, 163–169.
- Uchida,W., Matsunaga,S., Sugiyama,R. and Kawano,S. (2002) Interstitial telomere-like repeats in the *Arabidopsis thaliana* genome. *Genes Genet. Syst.*, **77**, 63–67.
- Johnson,L., Cao,X. and Jacobsen,S. (2002) Interplay between two epigenetic marks: DNA methylation and histone H3 lysine 9 methylation. *Curr. Biol.*, **12**, 1360–1367.
- Lindroth,A., Cao,X., Jackson,J., Zilberman,D., McCallum,C., Henikoff,S. and Jacobsen,S. (2001) Requirement of *CHROMOMETHYLASE3* for maintenance of CpXpG methylation. *Science*, **292**, 2077–2080.
- Lindroth,A., Shultis,D., Jasencakova,Z., Fuchs,J., Johnson,L., Schubert,D., Patnaik,D., Pradhan,S., Goodrich,J., Schubert,I. et al. (2004) Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with *CHROMOMETHYLASE3*. *EMBO J.*, **23**, 4286–4296.
- Martínez-Zapater,J., Estelle,M. and Somerville,C. (1986) A high repeated DNA sequence in *Arabidopsis thaliana*. *Mol. Gen. Genet.*, **204**, 417–423.
- Nagaki,K., Talbert,P., Zhong,C., Dawe,R., Henikoff,S. and Jiang,J. (2003) Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional element of *Arabidopsis thaliana* centromeres. *Genetics*, **163**, 1221–1225.
- Vaquero-Sedas,M., Gámez-Arjona,F. and Vega-Palas,M. (2010) *Arabidopsis thaliana* telomeres exhibit euchromatic features. *Nucleic Acids Res.*, **39**, 2007–2017.
- Zhang,W., Lee,H., Koo,D. and Jiang,J. (2008) Epigenetic modification of centromeric chromatin: Hypomethylation of DNA sequences in the CENH3-associated chromatin in *Arabidopsis thaliana* and maize. *Plant Cell*, **20**, 25–34.
- Fuchs,J., Demidov,D., Houben,A. and Schubert,I. (2006) Chromosomal histone modification patterns-from conservation to diversity. *Trends Plant Sci.*, **11**, 199–208.
- Fischer,A., Hofmann,I., Naumann,K. and Reuter,G. (2006) Heterochromatin proteins and the control of heterochromatic gene silencing in *Arabidopsis*. *J. Plant Physiol.*, **163**, 358–368.
- Johnson,L., Mollah,S., Garcia,B., Muratore,T., Shabanowitz,J., Hunt,D. and Jacobsen,S. (2004) Mass spectrometry analysis of *Arabidopsis* histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res.*, **32**, 6511–6518.
- Zhang,X., Clarenz,O., Cokus,S., Bernatavichute,Y., Pellegrini,M., Goodrich,J. and Jacobsen,S. (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.*, **5**, e129.
- Vaquero-Sedas,M. and Vega-Palas,M. (2011) On the chromatin structure of eukaryotic telomeres. *Epigenetics*, **6**, 1055–1058.
- Richards,E. and Ausubel,F. (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell*, **53**, 127–136.
- Shakirov,E. and Shippen,D. (2004) Length regulation and dynamics of individual telomere tracts in wild-type *Arabidopsis*. *Plant Cell*, **16**, 1959–1967.
- Ohmido,N., Kijima,K., Ashikawa,I., de Jong,J. and Fukui,K. (2001) Visualization of the terminal structure of rice chromosomes 6 and 12 with multicolor FISH to chromosomes and extended DNA fibers. *Plant Mol. Biol.*, **47**, 413–421.
- Luo,C., Sidote,D., Zhang,Y., Kerstetter,R., Michael,T. and Lam,E. (2012) Integrative analysis of chromatin states in *Arabidopsis* identified potential regulatory mechanisms for Natural Antisense Transcript production. (under review).
- Fransz,P., Armstrong,S., Alonso-Blanco,C., Fischer,T., Torres-Ruiz,R. and Jones,G. (1998) Cytogenetics for the model system *Arabidopsis thaliana*. *Plant J.*, **13**, 867–876.
- Ha,M., Ng,D., Li,W. and Chen,Z. (2011) Coordinated histone modifications are associated with gene expression variation within and between species. *Genome Res.*, **21**, 590–598.
- He,G., Zhu,X., Elling,A., Chen,L., Wang,X., Guo,L., Liang,M., He,H., Zhang,H. and Chen,F. (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell*, **22**, 17–33.
- Cheng,Z., Dong,F., Langdon,T., Ouyang,S., Buell,C., Gu,M., Blattner,F. and Jiang,J. (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell*, **14**, 1691–1704.
- Lee,H., Neumann,P., Macas,J. and Jiang,J. (2006) Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. *Mol. Biol. Evol.*, **23**, 2505–2520.
- Yan,H., Kikuchi,S., Neumann,P., Zhang,W., Wu,Y., Chen,F. and Jiang,J. (2010) Genome-wide mapping of cytosine methylation revealed dynamic DNA methylation patterns associated with genes and centromeres in rice. *Plant J.*, **63**, 353–365.
- Zhang,W., Yi,C., Bao,W., Liu,B., Cui,J., Yu,H., Cao,X., Gu,M., Liu,M. and Cheng,Z. (2005) The transcribed 165-bp CentO satellite is the major functional centromeric element in the wild rice species *Oryza punctata*. *Plant Physiol.*, **139**, 306–315.
- Gámez-Arjona,F., López-López,C., Vaquero-Sedas,M. and Vega-Palas,M. (2010) On the organization of the nucleosomes associated with telomeric sequences. *Biochim. Biophys. Acta*, **1803**, 1058–1061.
- Vannier,J., Depeiges,A., White,C. and Gallego,M. (2009) ERCC1/XPF protects short telomeres from homologous recombination in *Arabidopsis thaliana*. *PLoS Genet.*, **5**, e1000380.
- Pisano,S., Galati,A. and Cacchione,S. (2008) Telomeric nucleosomes: Forgotten players at chromosome ends. *Cell. Mol. Life Sci.*, **65**, 3553–3563.
- Fajkus,J., Kovari,K.A., Kralovics,R. and Bezdek,M. (1995) Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. *Mol. Gen. Genet.*, **247**, 633–638.
- Mechelli,R., Anselmi,C., Cacchione,S., De Santis,P. and Savino,M. (2004) Organization of telomeric nucleosomes: atomic force microscopy imaging and theoretical modeling. *FEBS Lett.*, **566**, 131–135.
- Pisano,S., Pascucci,E., Cacchione,S., De Santis,P. and Savino,M. (2006) AFM imaging and theoretical modeling studies of sequence-dependent nucleosome positioning. *Biophys. Chem.*, **124**, 81–89.
- Pisano,S., Marchioni,E., Galati,A., Mechelli,R., Savino,M. and Cacchione,S. (2007) Telomeric nucleosomes are intrinsically mobile. *J. Mol. Biol.*, **369**, 1153–1162.

39. Feng,S. and Jacobsen,S. (2011) Epigenetic modifications in plants: an evolutionary perspective. *Curr. Opin. Plant Biol.*, **14**, 179–186.
40. Hennig,L. and Derkacheva,M. (2009) Diversity of Polycomb group complexes in plants: same rules, different players? *Trends Genet.*, **25**, 414–423.
41. Pien,S. and Grossniklaus,U. (2007) Polycomb group and trithorax group proteins in Arabidopsis. *Biochim. Biophys. Acta*, **1769**, 375–382.
42. Zhang,K., Sridhar,V., Zhu,J., Kapoor,A. and Zhu,J. (2007) Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS One*, **2**, e1210.
43. Gardner,K., Allis,C. and Strahl,B. (2011) Operating on chromatin, a colorful language where context matters. *J. Mol. Biol.*, **409**, 36–46.
44. Turck,F., Roudier,F., Farrona,S., Martin-Magniette,M., Guillaume,E., Buisine,N., Gagnot,S., Martienssen,R., Coupland,G. and Colot,V. (2007) Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.*, **3**, e86.
45. Roudier,F., Ahmed,I., Bérard,C., Sarazin,A., Mary-Huard,T., Cortijo,S., Bouyer,D., Caillieux,E., Duvernois-Berthet,E., Al-Shikhley,L. *et al.* (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J.*, **30**, 1928–1938.
46. Vaughn,M., Tanurdzić,M., Lippman,Z., Jiang,H., Carrasquillo,R., Rabinowicz,P., Dedhia,N., McCombie,W., Agier,N. and Bulski,A. (2007) Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol.*, **5**, e174.