Vol 466|15 July 2010|doi:10.1038/nature09147

LETTERS

Relationship between nucleosome positioning and DNA methylation

Ramakrishna K. Chodavarapu[1]*, Suhua Feng[1,2]*, Yana V. Bernatavichute[1,3], Pao-Yang Chen[1], Hume Stroud[1], Yanchun Yu[4], Jonathan A. Hetzel[1], Frank Kuo[1], Jin Kim[1], Shawn J. Cokus[1], David Casero[1], Maria Bernal[5], Peter Huijser[6], Amander T. Clark[1,7], Ute Krämer[5]†, Sabeeha S. Merchant[3,8], Xiaoyu Zhang[4], Steven E. Jacobsen[1,2,3] & Matteo Pellegrini[1,3]

Nucleosomes compact and regulate access to DNA in the nucleus, and are composed of approximately 147 bases of DNA wrapped around a histone octamer[1,2]. Here we report a genome-wide nucleosome positioning analysis of *Arabidopsis thaliana* using massively parallel sequencing of mononucleosomes. By combining this data with profiles of DNA methylation at single base resolution, we identified 10-base periodicities in the DNA methylation status of nucleosome-bound DNA and found that nucleosomal DNA was more highly methylated than flanking DNA. These results indicate that nucleosome positioning influences DNA methylation patterning throughout the genome and that DNA methyltransferases preferentially target nucleosome-bound DNA. We also observed similar trends in human nucleosomal DNA, indicating that the relationships between nucleosomes and DNA methyltransferases are conserved. Finally, as has been observed in animals, nucleosomes were highly enriched on exons, and preferentially positioned at intron–exon and exon–intron boundaries. RNA polymerase II (Pol II) was also enriched on exons relative to introns, consistent with the hypothesis that nucleosome positioning regulates Pol II processivity. DNA methylation is also enriched on exons, consistent with the targeting of DNA methylation to nucleosomes, and suggesting a role for DNA methylation in exon definition.

To investigate the position of nucleosomes in *Arabidopsis thaliana*, we sequenced micrococcal nuclease (MNase)-digested nucleosomal DNA using an Illumina GAII sequencer to achieve a roughly 68-fold coverage of nucleosomes (see Supplementary Methods). The data are displayed in a modified UCSC genome browser (http://epigenomics.mcdb.ucla.edu/Nuc-Seq/) along with nucleosome density tracks that allow easy visualization of nucleosome positions throughout the genome (Fig. 1a).

To obtain a chromosomal view of nucleosome content, we plotted reads in bins of 100 kilobases tiling the chromosomes. As a control for biases in mapping and sequencing we also sequenced a library of randomly sheared *Arabidopsis* genomic DNA. We then normalized the nucleosome counts by the number of uniquely mapping genomic DNA counts within each bin along the chromosomes. Nucleosome content was relatively uniform throughout the euchromatic regions of chromosomes, but showed significant enrichment in pericentromeric heterochromatin regions, (Fig. 1b, Supplementary Fig. 1). To confirm these results, we performed chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-Seq) using an antibody against unmodified histone H3 (Fig. 1b and Supplementary Fig. 1). These observations indicate that nucleosomes are more densely packed in pericentromeric regions that are transcriptionally silent, are rich in transposons and contain heavily methylated DNA, than in the euchromatic arms[3].

By examining the relationship between reads that map to opposite strands of DNA, we observed a peak of reads on the reverse strand that occurred approximately 145–170 bases downstream of the reads on the forward strand (Supplementary Fig. 2a). This broad peak in the general vicinity of the accepted size of a nucleosome indicates that many nucleosomes are well positioned, leading to the repeated sequencing of both the forward and reverse complement of the corresponding nucleosome regions. Similarly, when we plotted the correlation between positive strand reads with other positive strand reads we saw a progressively decreasing correlation from the start of the nucleosome (Fig. 1c and Supplementary Fig. 2b), with smaller peaks spaced at 174 and 355 base pairs from the starting position of the reference reads indicating some preference for regular spacing of nucleosomes genome-wide. Assuming that nucleosomes are wrapped around 147 base pairs of DNA, the average length of the linkers of regularly spaced nucleosomes is about 30 base pairs. Consistent with the higher content of nucleosomes in pericentromeric heterochromatin regions, we found that the peaks at 174 and 355 were more pronounced in these regions than in the euchromatic arms (Fig. 1c), indicating that pericentromeric nucleosomes have a higher tendency to be in regularly spaced arrays. These results are consistent with earlier evidence for more regular spacing of nucleosomes in *Drosophila* heterochromatin[4].

Similar to findings in animal and fungal high-throughput nucleosome sequencing studies[5–8], we found 10-base periodicities in WW (W = A or T) dinucleotides, and SS (S = G or C) dinucleotides that were 5 bases out of phase with the WW dinucleotides (Fig. 1d and Supplementary Figs 3–5). WW dinucleotides are favoured at sites where the minor groove faces the histone core, whereas SS dinucleotides are favoured at sites where the minor groove faces away from the histone core[9,10]. The out of phase peaks in the frequencies of these dinucleotides leads to optimal bending of DNA, as A/T nucleotides cause a negative base roll and G/C nucleotides cause a positive base roll[9,10].

To study the relationship of DNA methylation with nucleosome positions, we used our single-nucleotide resolution whole-genome bisulphite-sequencing data[3]. *Arabidopsis* cytosines are methylated by

[1]Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California 90095, USA. [2]Howard Hughes Medical Institute, University of California Los Angeles, Los Angeles, California 90095, USA. [3]Molecular Biology Institute, University of California Los Angeles, Los Angeles, California 90095, USA. [4]Department of Plant Biology, University of Georgia, Athens, Georgia 30602, USA. [5]University of Heidelberg, BIOQUANT 23, R. 645, Im Neuenheimer Feld 267, D-69120 Heidelberg, Germany. [6]Department of Molecular Plant Genetics, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany. [7]Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, Los Angeles, California 90095, USA. [8]Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, USA. †Present address: Department of Plant Physiology, University of Bochum, Universitaetsstrasse 150, D-44801 Bochum, Germany.
*These authors contributed equally to this work.

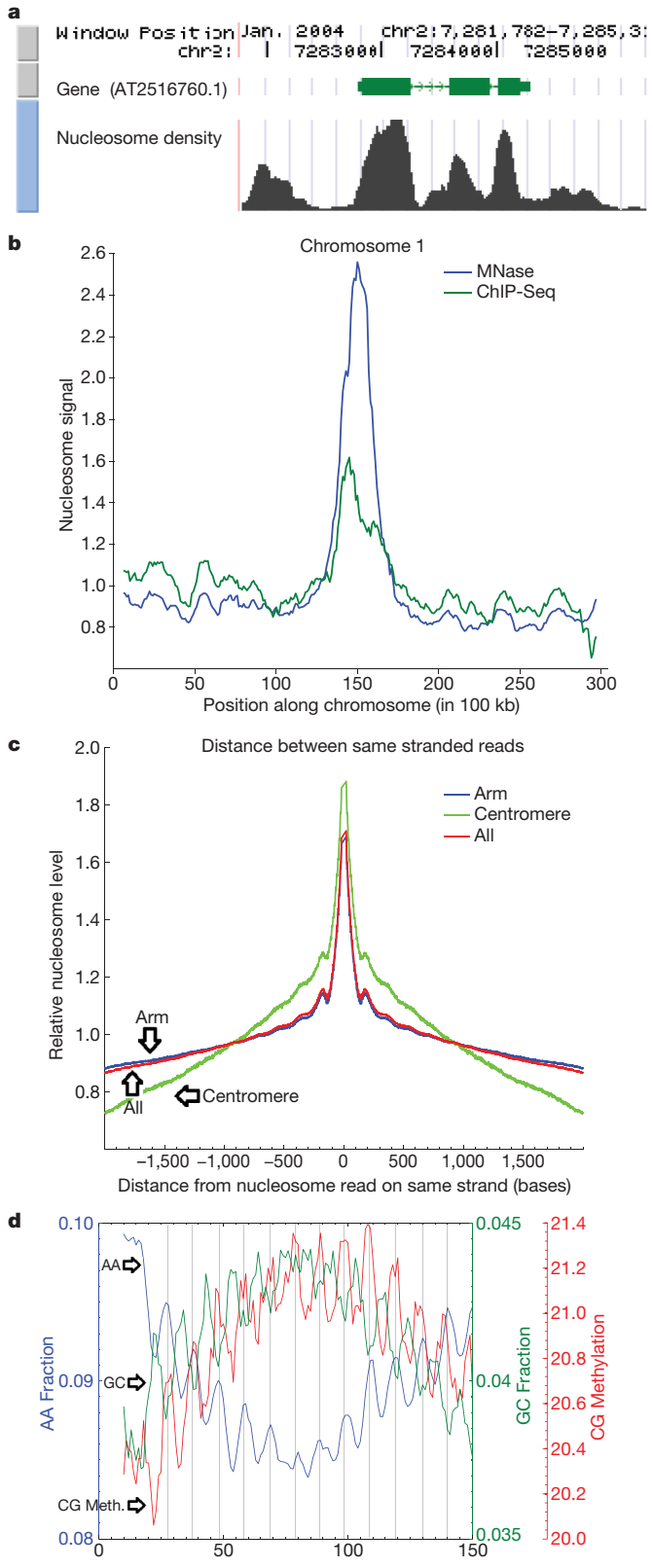

Figure 1 | Characterization of *A. thaliana* nucleosome data. **a**, Representative UCSC Browser screenshot of a gene showing calculated nucleosome densities. **b**, Chromosomal view of nucleosome read counts and ChIP-seq read counts in 100-kb tiles along chromosome 1 shows nucleosome enrichment in pericentromeric regions. **c**, Autocorrelation of positive stranded reads shows local peaks at positions 174 and 355 bases. **d**, AA and GC dinucleotide and CG methylation profiles show a 10-base periodicity over nucleosome-bound DNA, with DNA methylation profiles in phase with WW dinucleotide profiles, and out of phase with SS dinucleotide profiles.

one of three different DNA methyltransferases depending on their sequence context. CG sites are methylated by MET1, and CHG sites (where H is A, C or T) are methylated by CMT3. Finally, CHH sites are methylated by DRM2, a *de novo* methyltransferase that is targeted by small RNAs[11]. Notably, all three types of methylation showed a 10-base periodicity on nucleosomal DNA, which was in phase with the WW dinucleotides and out of phase with the SS dinucleotides (Figs 1d, 2a–c and Supplementary Fig. 6). These methylation preferences were not correlated with preferences for CG, CHG or CHH sequences at these locations (Supplementary Fig. 7). Because DNA methyltransferases access the major groove, this methylation would be on DNA that is on the outside of the nucleosome (minor groove facing the histones) and thus more accessible to the DNA methyltransferases. This in turn indicates that DNA might be in part methylated as the DNA is still bound to nucleosomes, leading to the observed 10-base pair periodicity. It has been proposed that chromatin remodelling enzymes are required for DNA methyltransferases to gain access to the DNA, and indeed DRD1, DDM1 and LSH1 are remodellers known to be important factors controlling DNA methylation[12]. However, our data indicate that nucleosomal DNA may also be a substrate for DNA methyltransferases *in vivo*, demonstrating a prominent role of the nucleosome in determining methylation patterning throughout the genome.

We previously reported a 10-nucleotide periodicity in CHH methylation data when performing autocorrelation analysis on the methylation pattern of the whole genome[3]. Our previous interpretation was that the structure of the DRM2 enzyme might be responsible for this pattern, because the orthologous Dnmt3 enzymes in mammals are known to act as heteromeric complexes in which two methyltransferase active sites have a spacing equivalent to roughly 10 nucleotides of DNA[13]. However, the current data in which we see 10-base pair periodicities for all types of methylation indicate a more general explanation: that nucleosomes are to some extent dictating access to the DNA and therefore setting the register of methylation for all DNA methyltransferases. Nucleosomal preferences could also partially explain the sequence preferences that we observed previously for CHG and CHH methylation[3]. Highly methylated cytosines tended to be followed immediately by A/T but not C, consistent with our finding that DNA methylation is out of phase with CC dinucleotides (Supplementary Fig. 6).

On a larger scale, we also observed that levels of all three types of DNA methylation were higher in nucleosome-spanning DNA than in flanking DNA, indicating that nucleosome-bound DNA is enriched for DNA methylation (Fig. 2 and Supplementary Fig. 8). This finding supports the view that nucleosomes are preferentially targeted by DNA methyltransferases. In the case of CMT3, it is predicted that this enzyme is recruited or activated by histone H3 lysine 9 dimethylation, because a knockout of histone H3 lysine 9 methyltransferase mimics a knock out of CMT3, and because the CMT3 chromodomain can bind to methylated histones[14]. However, our data show that all types of methylation are enriched on nucleosome-bound DNA, indicating that nucleosomes or histone modifications may assist in the recruitment of all of the *Arabidopsis* DNA methyltransferases.

To test whether the patterns of DNA methylation on nucleosomal DNA are also found in humans, we used previously published MNase nucleosome sequencing data[15] together with our single-nucleotide resolution bisulphite sequencing data on the human embryonic stem cell line HSF1 (see Supplementary Methods). We found that human nucleosome-bound DNA also showed a 10-base periodicity in its CG, CHG and CHH methylation status (Fig. 3). Furthermore, as in *Arabidopsis*, the overall level of methylation was higher on nucleosome-bound DNA than in flanking regions. This indicates that nucleosome architecture has a role in shaping DNA methylation patterning in the human genome, and is consistent with the recent finding of stable anchoring of Dnmt3 DNA methyltransferases to mammalian chromatin[16] and more generally with the conservation of DNA methyltransferase function in plants and animals[17]. We also

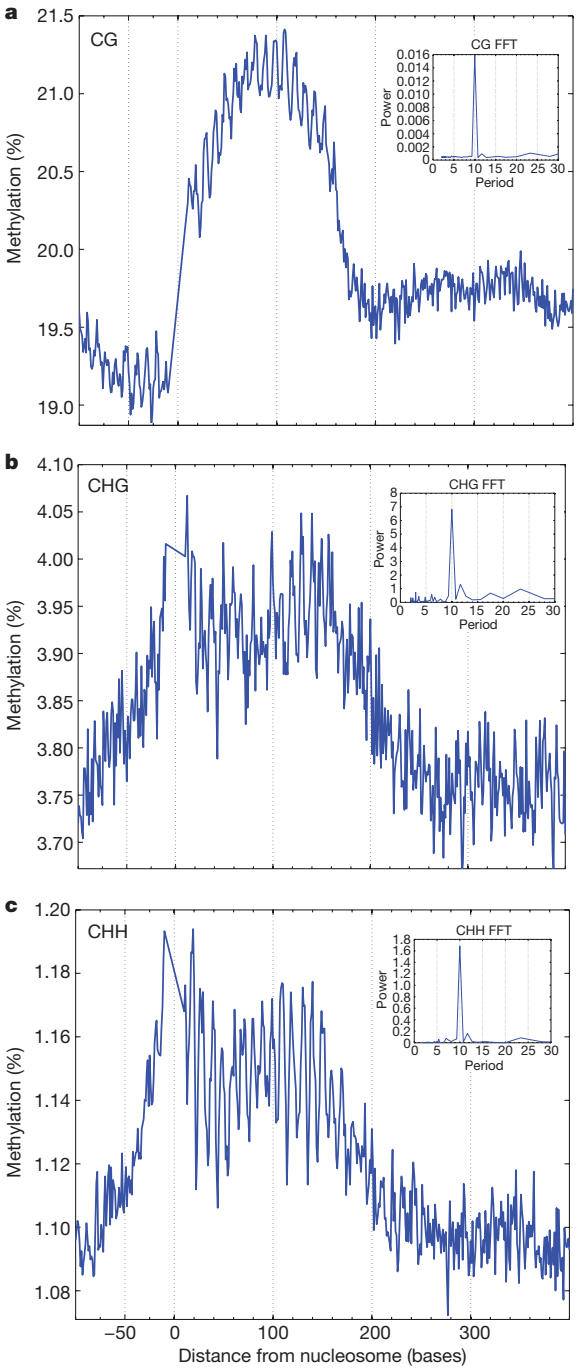

Figure 2 | DNA methylation profiles of nucleosome-bound DNA in *Arabidopsis*. The weighted average percent DNA methylation was calculated and plotted at each distance from nucleosome start sites (0). **a**–**c**, CG methylation (**a**), CHG methylation (**b**) and CHH methylation (**c**) each show a 10-base periodicity over nucleosome-bound DNA (1–147 bases). Fast Fourier transforms (FFTs) can be used to deconstruct inherent frequencies in a complex signal. The FFTs calculated over the region of the nucleosome (insets) demonstrate this periodicity in CG (**a**), CHG (**b**) and CHH (**c**) contexts.

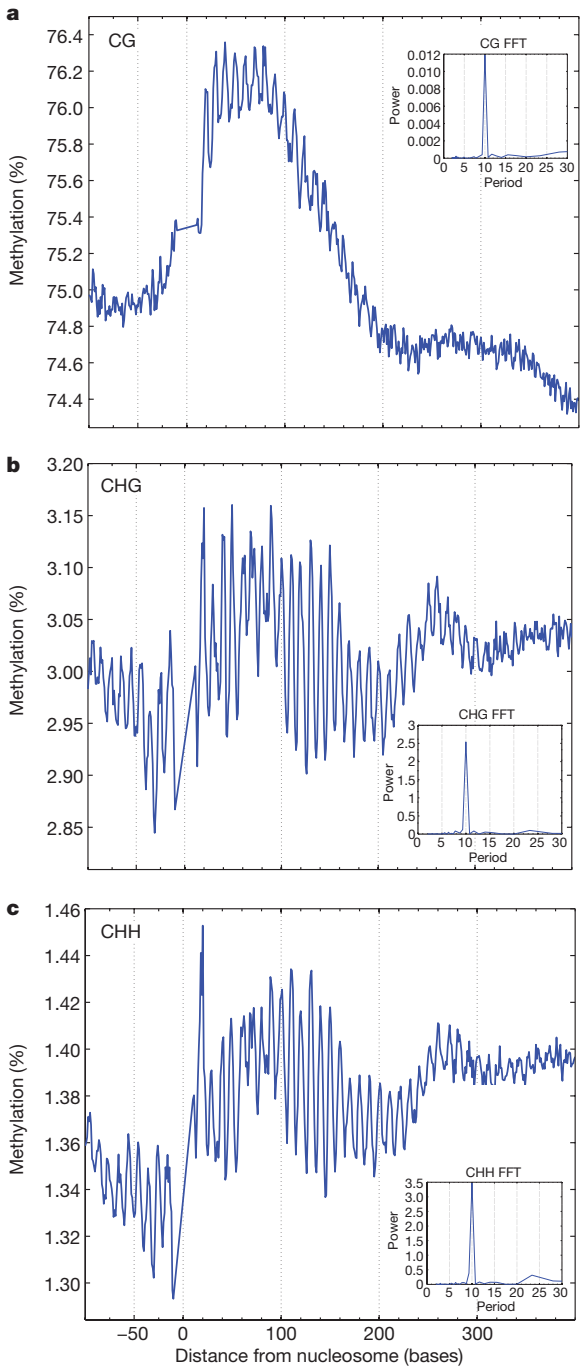

Figure 3 | DNA methylation profiles of nucleosome-bound DNA in humans. The weighted average percent DNA methylation was calculated and plotted at each distance from nucleosome start sites (0). **a**–**c**, CG methylation (**a**), CHG methylation (**b**) and CHH methylation (**c**) each show a 10-base periodicity over nucleosome-bound DNA (1–147 bases). The FFTs calculated over the region of the nucleosome (insets) demonstrate this periodicity in CG (**a**), CHG (**b**), and CHH (**c**) contexts.

analysed DNA methylation patterning on nucleosomes in different regions of the *Arabidopsis* and human genomes, including genes, promoters, pericentromeric regions and euchromatic arm regions (Supplementary Figs 9–17) and found that the 10-nucleotide periodicity was found in all cases, indicating that the relationship between nucleosome positioning and DNA methylation is general.

We observed that nucleosomes were much more abundant in exons than in introns (see Fig. 1a for an example), consistent with findings from several recent nucleosome positioning studies in other organisms[18–23]. The number of nucleosomes per base pair in introns was only 63% of the level found in exons. Furthermore, we found a strong peak of nucleosome start sites at intron–exon junctions and at exon–intron junctions (Fig. 4, Supplementary Fig. 18). The enrichment of nucleosomes in exons was not solely owing to higher G+C content (Supplementary Fig. 19a–c), or because of consensus splice site sequences (Supplemental Fig. 19d–f), and was confirmed using independent chromatin immunoprecipitation methods (Supplementary Fig. 20). Longer exons contained a higher number of

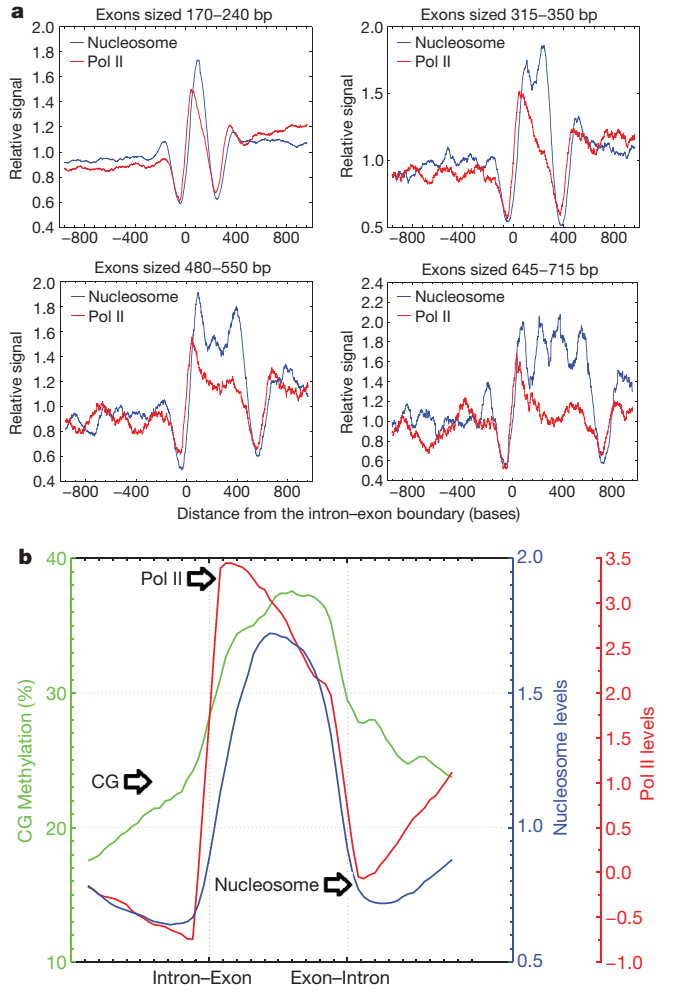

Figure 4 | Nucleosome and Pol II levels in exons. We performed chromatin immunoprecipitation with an antibody against RNA Pol II, and hybridized the resulting DNA to a whole genome Affymetrix microarray. We normalized these data to randomly sheared genomic DNA to control for probe efficiencies. **a**, Nucleosomes are phased in exons. Exon size limits were selected such that the exon could house no more than 1, 2, 3 or 4 nucleosomes. Nucleosome midpoints are plotted over these intron–exon boundaries. **b**, Each exon was divided into 25 equal sized bins and the nucleosome midpoints, Pol II and CG methylation levels are plotted over the exon and flanking introns.

nucleosomes. Examination of exons in the size ranges of 170–240, 315–350, 480–550 and 645–715 base pairs revealed peaks of one, two, three or four well-positioned nucleosomes respectively (Fig. 4a), indicating that nucleosomes are phased within exons.

Strong nucleosome enrichment on exons was detected both for genes that are highly expressed and for those not expressed (Supplementary Fig. 21 and Supplementary Table 1), suggesting that DNA sequences position nucleosomes in the absence of active transcription and splicing. Consistent with this hypothesis, using a nucleosome positioning prediction algorithm[24] we found similar patterns of nucleosome positioning in introns and exons between the experimental and theoretical data sets (Supplementary Fig. 22). Similarly, using theoretically predicted nucleosome positions, we observed similar patterns of DNA methylation on nucleosome-bound DNA (Supplementary Fig. 23), as well as enrichment of predicted nucleosomes in pericentromeric regions (Supplementary Fig. 24).

Because nucleosomes present a barrier to RNA Pol II transcription, we tested for Pol II occupancy in exons using a chromatin immunoprecipitation microarray approach. We observed significant enrichment of Pol II in exons relative to introns, consistent with the hypothesis that Pol II is paused on exonic DNA (Fig. 4). One possibility is that Pol II stalling on exons could enhance accurate splicing of upstream introns, thus reducing exon skipping and aiding in the fidelity of exon definition[20]. This is consistent with the finding of Pol II enrichment on human exons and indicates that Pol II enrichment on exons might be a common eukaryotic feature[20]. Furthermore, particular histone modifications have been recently shown to recruit splicing regulators, providing additional possible mechanisms for the regulation of splicing by nucleosome positioning[25].

Because of the enhancement of DNA methylation over nucleosomal DNA, and the enrichment of nucleosomes on exons, a prediction is that DNA methylation should be enhanced on exons relative to introns. Indeed, we found depletion of DNA methylation in introns and enrichment in exons in a pattern that was similar to nucleosome occupancy (Fig. 4b). This is consistent with recent findings of enhanced exonic methylation in other plant species, as well as in the *Chlamydomonas*, honeybee, *Ciona* and human genomes[26,27]. This suggests the possibility that DNA methylation, which frequently exists in the transcribed regions of active genes[26,28], could have a conserved role in exon definition or splicing regulation. These findings also reinforce the view that nucleosomal positions have an important role in shaping the methylation landscape of eukaryotic genomes.

METHODS SUMMARY

Approximately 300 ng of MNase-digested mononucleosome DNA were used for Illumina library generation following manufacturer instructions. Libraries were sequenced using an Illumina Genome Analyzer II following manufacturer instructions. Resulting reads were mapped to the TAIR7 annotation of the *Arabidopsis* genome and reads that mapped to multiple sites were eliminated. A total of 24.2 million reads were mapped to the forward strand and 23.9 million reads to the reverse strand, generating a 68-fold coverage of nucleosome space.

Pol II ChIP-chip was performed as described previously[29]. Briefly, crosslinked chromatin was extracted, sonicated and used in ChIP with an antibody against Pol II CTD (ab817; Abcam). Pol II-bound DNA and input genomic DNA were extracted, amplified, labelled and hybridized to Affymetrix tiling microarrays as described previously[29]. Three biological replicates were performed and the $\log_2$ ratios of Pol II-bound DNA over input DNA were calculated using the Tiling Analysis Software (Affymetrix).

Received 5 March; accepted 7 May 2010.
Published online 30 May 2010.

1. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
2. Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285–294 (1999).
3. Cokus, S. J. *et al.* Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**, 215–219 (2008).
4. Sun, F. L., Cuaycong, M. H. & Elgin, S. C. Long-range nucleosome ordering is associated with gene silencing in *Drosophila melanogaster* pericentric heterochromatin. *Mol. Cell. Biol.* **21**, 2867–2879 (2001).
5. Albert, I. *et al.* Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**, 572–576 (2007).
6. Mavrich, T. N. *et al.* A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18**, 1073–1083 (2008).
7. Valouev, A. *et al.* A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* **18**, 1051–1063 (2008).
8. Mavrich, T. N. *et al.* Nucleosome organization in the *Drosophila* genome. *Nature* **453**, 358–362 (2008).
9. Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* **34**, 269–324 (2001).
10. Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442**, 772–778 (2006).
11. Henderson, I. R. & Jacobsen, S. E. Epigenetic inheritance in plants. *Nature* **447**, 418–424 (2007).
12. Kobor, M. S. & Lorincz, M. C. H2A.Z and DNA methylation: irreconcilable differences. *Trends Biochem. Sci.* **34**, 158–161 (2009).
13. Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. & Cheng, X. Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. *Nature* **449**, 248–251 (2007).
14. Lindroth, A. M. *et al.* Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with *CHROMOMETHYLASE3*. *EMBO J.* **23**, 4146–4155 (2004).
15. Schones, D. E. *et al.* Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898 (2008).

16. Jeong, S. *et al.* Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol. Cell. Biol.* **29,** 5366–5376 (2009).
17. Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Rev. Genet.* **11,** 204–220 (2010).
18. Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C. & Komorowski, J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* **19,** 1732–1741 (2009).
19. Tilgner, H. *et al.* Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* **16,** 996–1001 (2009).
20. Schwartz, S., Meshorer, E. & Ast, G. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* **16,** 990–995 (2009).
21. Kolasinska-Zwierz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature Genet.* **41,** 376–381 (2009).
22. Spies, N., Nielsen, C. B., Padgett, R. A. & Burge, C. B. Biased chromatin signatures around polyadenylation sites and exons. *Mol. Cell* **36,** 245–254 (2009).
23. Ponts, N. *et al.* Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res.* **20,** 228–238 (2010).
24. Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458,** 362–366 (2009).
25. Luco, R. F. *et al.* Regulation of alternative splicing by histone modifications. *Science* **327,** 996–1000 (2010).
26. Feng, S. *et al.* Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl Acad. Sci. USA* **107,** 8689–8694 (2010).
27. Laurent, L. *et al.* Dynamic changes in the human methylome during differentiation. *Genome Res.* **20,** 320–331 (2010).
28. Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* advance online publication doi:10.1126/science.1186366 (15 April 2010).
29. Zhang, X. *et al.* Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* **5,** e129 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank members of our laboratories for input and guidance, and the Department of Energy, Institute for Genomics and Proteomics for support. R.K.C. is supported by NIH training grant GM07104. S.F. is a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation. P.-Y.C. is funded by Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research at UCLA. M.B. was supported by an Alexander-von-Humboldt Fellowship and U.K. by a Heisenberg Fellowship. X.Z. was supported by a Faculty Research Grant (JR-040) from the University of Georgia. Research was supported by an Innovation Award from the Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research at UCLA, NIH grants GM60398 and GM42143, and NSF Plant Genome Research Program #0701745. S.E.J. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions S.E.J. and M.P. designed the research. S.F., Y.V.B., H.S., Y.Y., J.A.H., M.B., P.H., A.T.C., U.K., S.S.M. and X.Z. performed experiments. R.K.C., P.-Y.C, F.K., J.K., S.J.C and D.C. analysed data. R.K.C., S. F., S.E.J. and M.P. wrote the paper.

Author Information All sequencing files have been deposited in GEO under accession code GSE21673. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.E.J. (jacobsen@ucla.edu) or M.P. (matteop@mcdb.ucla.edu).