ARTICLE

doi:10.1038/nature11142

A map of nucleosome positions in yeast at base-pair resolution

Kristin Brogaard[1], Liqun Xi[2], Ji-Ping Wang[2] & Jonathan Widom[1,3]‡

The exact positions of nucleosomes along genomic DNA can influence many aspects of chromosome function. However, existing methods for mapping nucleosomes do not provide the necessary single-base-pair accuracy to determine these positions. Here we develop and apply a new approach for direct mapping of nucleosome centres on the basis of chemical modification of engineered histones. The resulting map locates nucleosome positions genome-wide in unprecedented detail and accuracy. It shows new aspects of the *in vivo* nucleosome organization that are linked to transcription factor binding, RNA polymerase pausing and the higher-order structure of the chromatin fibre.

Nucleosomes distort and occlude the genomic DNA from access to most DNA-binding proteins, and their exact positions affect the structure of the chromatin fibre[1]. Even single base-pair shifts of the position of a nucleosome can change chromatin configurations[2] and protein binding kinetics[3–5]. A map of nucleosome positions at single-base-pair resolution is necessary to understand fully a wide range of biological processes including RNA polymerase activity[6–8], transcription-factor binding kinetics[4,5,9,10], DNA replication[11,12], centromere structure[13–15] and gene splicing[16–18].

The most commonly used method for nucleosome mapping relies on treatment of chromatin with micrococcal nuclease (MNase) to digest DNA sequences not protected by the histone core. The nucleosome positions are indirectly inferred by the centres of undigested DNA sequence fragments. Although this method has provided insights into our understanding of nucleosomes, it is imprecise because of large variability of read lengths, transient unwrapping of nucleosomes, MNase sequence preferences and interference of other DNA-binding proteins[19–23]. Thus, a different approach is required to measure nucleosome locations with greater accuracy.

A chemical approach to mapping nucleosomes

We developed a new, genome-wide mapping approach that directly determines nucleosome centre positions with single-base-pair resolution. It derives from earlier work in which localized hydroxyl radicals were used to map reconstituted mononucleosomes centre positions[24,25]. In *Saccharomyces cerevisiae* we introduced a unique cysteine into histone H4 at position 47 (H4S47C) to covalently attach a sulphhydryl-reactive, copper-chelating label, *N*-(1,10-phenanthroline-5-yl)iodoacetamide. This cysteine mutation positions the label in close proximity to the DNA backbone and at sites symmetrically flanking the nucleosome centre (Fig. 1a). Log-phase yeast was collected and permeabilized immediately[26–29]. The label was subsequently introduced to the cells, where it reacts with the cysteines present in histone H4. Addition of copper and hydrogen peroxide catalyses the formation of short-lived hydroxyl radicals that react with and cleave the DNA backbone (Supplementary Methods). This strategy leads to a highly specific cleavage of the DNA backbone at sites that symmetrically flank the nucleosome centre[24,25]. On an agarose gel, a DNA banding pattern forms only when the reaction includes the sulphhydryl-reactive label, copper, H_2O_2 and the H4S47C mutant yeast, where the smallest fragment corresponds to DNAs between the centres of neighbouring nucleosomes (Fig. 1b).

We purified and sequenced the lowest molecular weight DNA band from the agarose gel (Fig. 1b) in six independent experiments (four with single-end and two with paired-end parallel sequencing), and produced a map of 105 million cleavages on each DNA strand (Figs 1c and 2a). The cleavage peaks on opposite DNA strands show two important correlations: a separation by either +2 or −5 nucleotides (with the 5′ to 3′ direction defined as positive) (Fig. 2b). This pattern is expected if cleavage occurs mainly at positions −1 (primary site) and +6 (secondary site) relative to the nucleosome centre (Fig. 1c), similar to that described for reconstituted mononucleosomes[24,25] (Supplementary Methods). The dominance and sharpness of these peaks show that the chemical maps contain high-resolution nucleosome positioning information.

We defined the nucleosome centres by identifying the characteristic pattern of chemical cleavages genome-wide (Supplementary Methods). We first trained a single-template model and then a four-template model to represent the average cleavage pattern at eight nucleotide positions centred around the primary and secondary sites (Supplementary Table 1). The four-template model accounts for four principal distinct cleavage patterns associated with nucleotide composition at positions −3/+3 and outperforms the single-template model (Supplementary Fig. 1a–c). The templates were built into a Bayesian deconvolution algorithm to calculate the nucleosome centre positioning (NCP) score at every genomic location. A larger NCP score means more cleavages observed in positions that conform to primary and secondary site configuration, thus indicating more well-positioned nucleosomes. Nucleosome centres were defined on the basis of the magnitude of the NCP score-to-noise ratio (Fig. 2a and Supplementary Fig. 1d).

Comparing two independent chemical mapping experiments, the resulting nucleosome positions are extremely consistent (Fig. 2c). Hence we combined the six independent data sets and produced a final unique map of 67,543 nucleosomes (Supplementary Methods), allowing two neighbouring nucleosomes to overlap by no more than 40 base pairs (bp) (occupying 79.9% of the genome, Supplementary Table 2); and a redundant map of 351,264 nucleosomes (allowing nucleosomes to overlap arbitrarily) by including all positions whose

[1]Department of Molecular Biosciences, Northwestern University, Evanston, Illinois 60208, USA. [2]Department of Statistics, Northwestern University, Evanston, Illinois 60208, USA. [3]Department of Chemistry, Northwestern University, Evanston, Illinois 60208, USA.
‡Deceased.

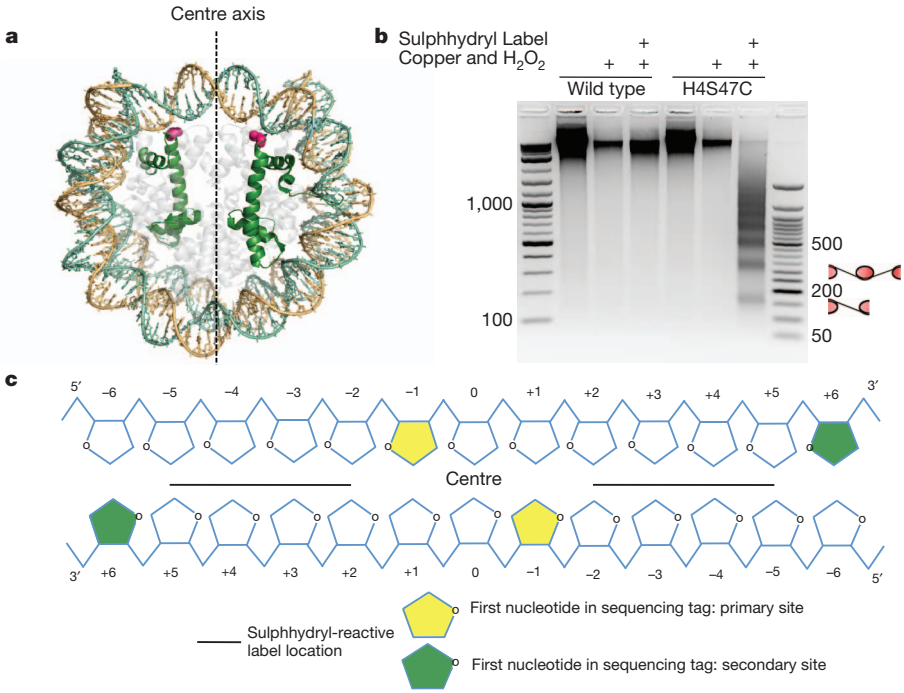


Figure 1 | Mapping nucleosome centres by site-specific chemical cleavage. **a**, Nucleosome structure, highlighting histone H4 (green) and residue serine 47 (pink), which is mutated to a cysteine where the sulphhydryl-reactive label covalently binds. **b**, Ethidium-bromide-stained agarose gel showing the chemical mapping results in a DNA banding pattern, which occurs only when the reaction includes (indicated by '+') the sulphhydryl-reactive label, copper, H_2O_2 and the H4S47C mutant yeast. The cartoons adjacent to the agarose gel illustrate mapping successive nucleosome centres produced the banding pattern. **c**, Locations of dominant hydroxyl radical cleavages relative to the nucleosome centre (base pair 0).

NCP score-to-noise ratio exceeded the smallest ratio value from the unique map (Supplementary Table 3). The chemical map shows general accordance with published MNase maps (MNase 1 (ref. 30) and MNase 2 (ref. 31)) of nucleosome positions, establishing that the chemical mapping technique does map nucleosome centres (Fig. 2a, d and Supplementary Fig. 2a, b). All analyses (unless indicated) presented in the main text use the unique set of nucleosomes. Important conclusions drawn from the unique map were confirmed using the redundant map, and results are presented separately in Supplementary Fig. 12.

The current chemical data show unbalanced cleavage patterns between the two strands at nucleosomes flanking long linkers including nucleosome-free regions upstream of transcription start sites (TSSs) (Supplementary Fig. 3a). As only the densest portion of the lowest molecular mass DNA fragment (~125–200 bp) was isolated from the agarose gel, DNA fragments spanning longer linkers were selected against (Fig. 1b). We confirmed this by isolating and sequencing DNA fragments of higher molecular mass (~200–500 bp) produced from one chemical mapping experiment (Supplementary Methods). The resulting data recovered the expected symmetric cleavage pattern at nucleosomes flanking the nucleosome-free regions (Supplementary Fig. 3b), and confirmed that failure to map nucleosomes because of consecutive long linkers is not a concern in the current data (Supplementary Fig. 3c). The cleavage asymmetry is not problematic for defining nucleosome centres, as one strand can still give a local maximum positioning score in our Bayesian deconvolution approach.

Nucleosome overlap and base composition

The chemical map shows detail of competing preferential nucleosome positions genome-wide. The redundant map suggests that in a population of cells nucleosomes often maintain several overlapping positions, creating closely clustered cleavage sites (Supplementary Fig. 4a, b). These sites were deconvoluted to yield locations and relative abundance of overlapping nucleosomes. Strikingly, overlapping nucleosomes are predominantly spaced by integral multiples of the DNA helical repeat (~10 bp) (Fig. 2e and Supplementary Fig. 4b). This result accords with ~10-bp periodic cleavage endpoints in MNase maps[32,33] and clarifies that MNase cleavage periodicities reflect several nucleosome positions instead of an artefact of the MNase itself[32]. This overlapping positioning of nucleosomes could be a consequence of the population or dynamic average, or even the co-existence of overlapping nucleosomes on the same DNA molecule[34]. By shifting in 10-bp steps (one helical turn), nucleosomes can maintain the local chromatin three-dimensional structure, while sampling several positions.

Nucleosomes *in vivo* are enriched for particular DNA sequence motifs at particular points along the nucleosomal DNA, most notably ~10-bp periodic occurrences of AA/TT/AT/TA dinucleotides that are favoured when the DNA backbone (minor groove) faces inwards towards the histone core, and CC/CG/GC/GG dinucleotides when the DNA backbone faces outwards[27,35,36]. These signals are seen in the chemical map with a period of 10.3 bp, and are far stronger than what was seen from MNase maps (Figs 3a, b and Supplementary Fig. 5a), suggesting that imprecision in the MNase maps degrades the sequence alignment. The ~0.2-bp difference between the dinucleotide periodicity and the classical helix repeat length (that is, 10.5 bp (ref. 37)) implies that DNAs probably over-stretch by 1–2 bp to accommodate the packing constraints of nucleosome structures in the chromatin fibre[1]. The strength of dinucleotide signal increases with the NCP score-to-noise ratio, suggesting that stronger dinucleotide sequence features are associated with positioning of more stable nucleosomes (Fig. 3a).

These dinucleotide signals are average features of the aggregated nucleosomes in the unique set, and of the overlapping nucleosomes (Supplementary Fig. 5b). The unique nucleosomes show an average of 40–50% occurrence rate of AA/TT/AT/TA at peak positions while lacking a dominant AA/TT/AT/TA bias at the edges of the nucleosomes observed in MNase maps (Fig. 3b)[19,21]. The preferences of nucleotide five-base polymers from the chemical map agree well with those measured by MNase, including the strong negative preference of nucleosomes for A/T-rich pentanucleotides[38,39]. However, both the

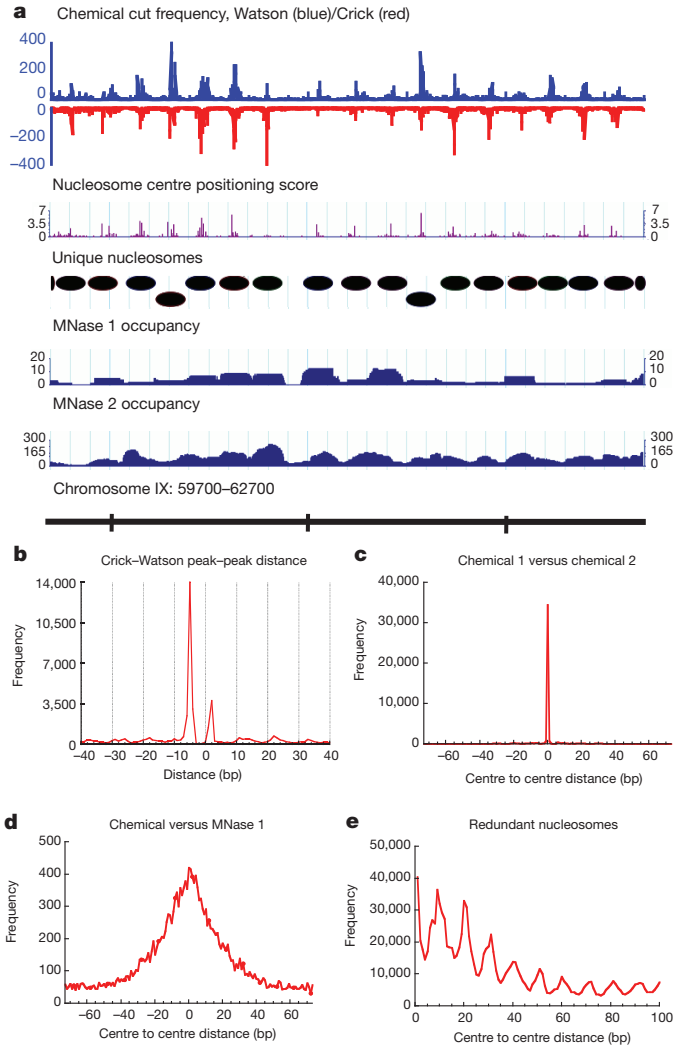


Figure 2 | Raw data, defined nucleosomes and pairwise comparison of nucleosome maps. **a**, Screenshot of raw data, nucleosome centre positioning score (NCP score), defined unique nucleosomes and two MNase maps (MNase 1 (ref. 30) and MNase 2 (ref. 31)). **b**, Crick–Watson cleavage peak–peak distances showing two dominant distances: +2 and −5 nucleotides. **c**, Distances between the nearest nucleosome centres in two independent chemical maps and **d**, in the combined chemical map compared with the MNase 1 map. **e**, Distance between nucleosome centres in the redundant map showing preferential spacing of multiples of ~10 bp.

nucleosome and linker regions are less strongly biased against A/T pentanucleotides in the chemical map than in the MNase maps (Fig. 3c), indicating that MNase preferentially degrades A/T-rich nucleosomes.

Intriguingly we noted a high frequency of occurrences (~0.6) of an A at position −3, and its symmetry-related T at position +3 relative to the nucleosome centre (Supplementary Figs 6 and 7). This high frequency has been observed in previous *in vitro* nucleosome positioning studies, with a frequency of 0.5 at these two locations[40]. Nucleosomes lacking both A(−3) and T(+3) (10% of the unique set) show similar but slightly weaker dinucleotide preference patterns compared with those having one or two of these nucleotides (Supplementary Fig. 5c). However, we cannot exclude the possibility that the chemical map might present a minor bias favouring an A(−3) or a T(+3) on top of the true preference of these nucleotides.

Global features of the nucleosome map

The high-resolution chemical map allows for more accurate conclusions about global nucleosome features. We defined the nucleosome

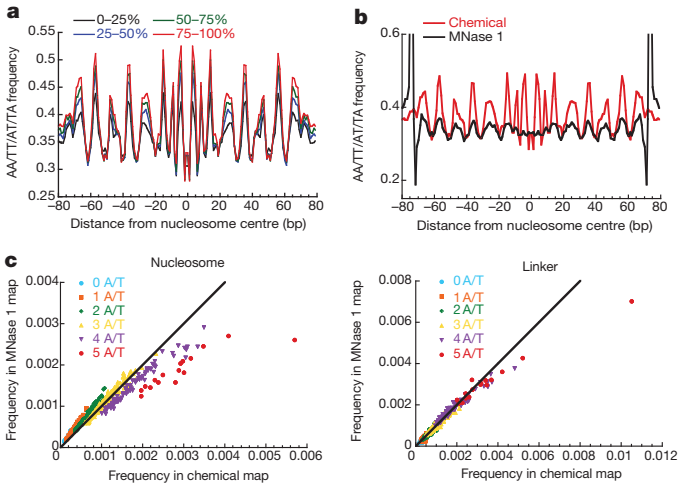


Figure 3 | Nucleosome sequence preferences. **a**, Position-dependent frequencies of AA/TT/AT/TA dinucleotides in the chemical map, by quartiles of NCP score-to-noise ratio. **b**, Comparison of dinucleotide frequency between the chemical map and the MNase 1 map. The chemical map shows significantly stronger AA/TT/AT/TA dinucleotide signals, suggesting an improved centre alignment of nucleosomal DNA. **c**, Pentanucleotide preferences of nucleosomes (left) and linker DNA regions (right) in the chemical map compared with the MNase 1 map, coloured by numbers of A or T nucleotides.

occupancy score at any given location as the total NCP scores in the ±73-bp region after correction for cleavage asymmetry on the two strands and controlling for signal-to-noise ratio as done in the redundant map (Supplementary Methods). The chemical map reproduces the well-known strong depletion of nucleosomes immediately upstream of the TSS[41] exposing the binding sites of regulatory proteins, with ordered nucleosomes extending downstream over the protein coding sequence, and much weaker ordering upstream of the TSS (Fig. 4a). The average occupancy curve within 1,000 bp of TSS is highly consistent with that from the MNase map ($r = 0.98$, Fig. 4a), suggesting the chemical map can provide a reasonable measure of nucleosome occupancy. Substantial nucleosome depletion is shown at the ends of the open reading frames (Fig. 4a), implying that nucleosome depletion at 3′ gene ends[20] is not an artefact of MNase digestion, but a real phenomenon. For transfer RNA (tRNA) genes, which are highly transcribed by RNA polymerase III and have their promoters inside the gene bodies[42], the upstream nucleosomes are strongly ordered whereas the gene body regions are nucleosome-depleted (Fig. 4a). No particular ordering of nucleosomes was observed relative to exon–intron or intron–exon boundaries (Supplementary Fig. 8a). We also observed strong ordering of nucleosomes on both sides of autonomously replicating sequences (ARS), and at centromere start (start of centromere DNA element I (CDEI)) and end (end of centromere DNA element III (CDEIII)) sites (Supplementary Fig. 8a, b). For the centromere sites, the nucleosome centres are highly localized at several specific positions, as shown by the sharp spikes of NCP score. Thus, the locations of the unique centromeric nucleosomes seem to be positioned relative to the exact locations of other key centromere-specific DNA-binding proteins.

An important unresolved problem concerns the role of DNA sequence features in positioning nucleosomes in protein-coding regions. Nucleosomes can be classified, with +1 referring to a nucleosome covering the start of the coding sequence, and +2 … +n for each successive nucleosome downstream. Nucleosomes of all classes (not only evident in the −1 and +1 class[43]) show strong AA/TT/AT/TA patterns with periodicity ranging from 10.20 to 10.26 bp, at which the amplitude of transform differs by no more than 5% between any two classes (Fig. 4b, Supplementary Fig. 9 and Supplementary Methods). Separate analysis using MNase data shows a similar but weaker periodic pattern, comparable across different classes (Fig. 4b

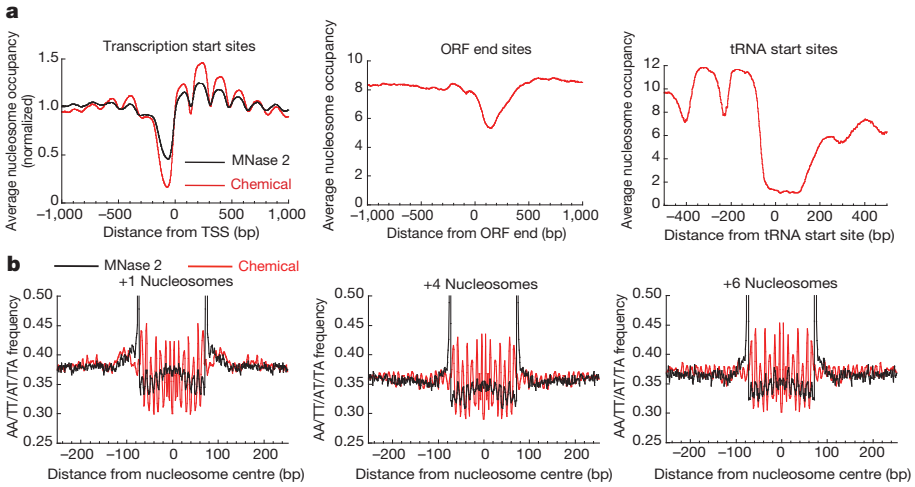
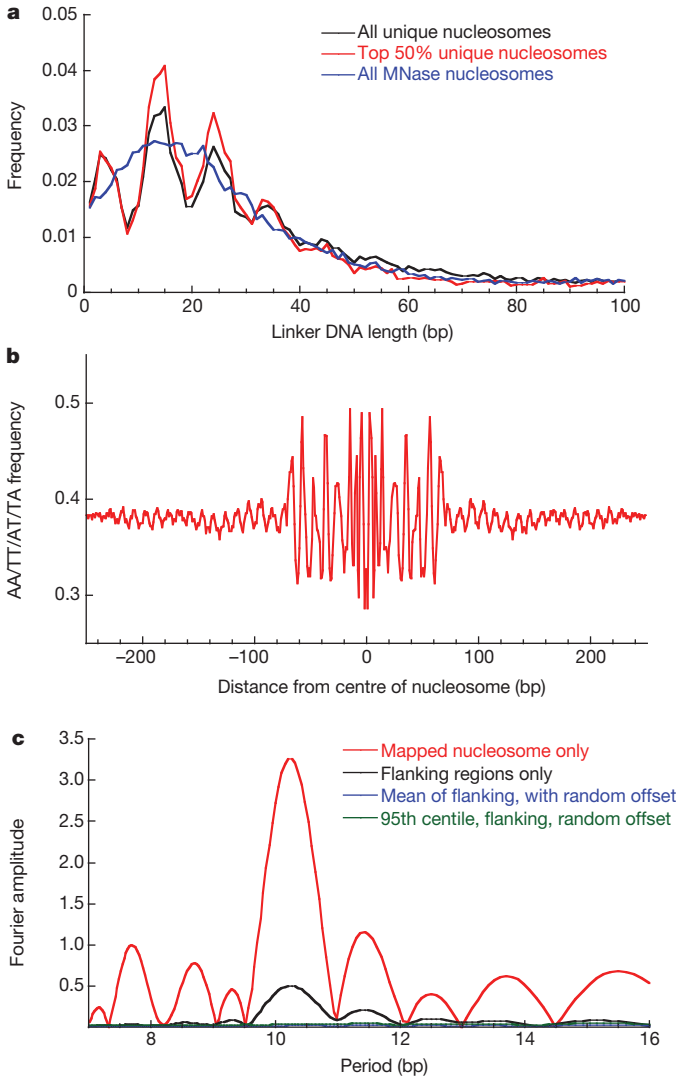


Figure 4 | Genome-wide features of nucleosome positions. **a**, Genome-wide average nucleosome occupancy across all sites in the chemical map plotted relative to TSSs (3,017 in total[41]), open reading frame (ORF) ends (centre) and tRNA start sites (right). **b**, AA/TT/AT/TA dinucleotide frequency within nucleosomes for nucleosome classes +1, +4 and +6 from the chemical map and the MNase 2 map.

and Supplementary Fig. 9). This suggests that DNA sequence features have an important role in detailed positioning of nucleosomes. A recent *in vitro* study[44] postulates an ATP-facilitated positioning, suggesting that chromatin complexes may use ATP to override DNA-intrinsic positioning and drive directional packaging of nucleosomes against a 5′ barrier at the 5′ ends of genes and to a lesser extent at the 3′ ends. Future studies are needed to understand the exact relationship of DNA sequence features and ATP in nucleosome positioning *in vivo*.



Chromatin structure and nucleosome spacing

The length of free DNA between two neighbouring nucleosomes (linker length) is a crucial determinant in the structure and compaction of the chromatin fibre and is directly related to the regulation of DNA-associated cellular processes[2,45,46]. The linker length distribution has a median of 23 bp (for linkers ≤ 100 bp) and shows a non-random pattern with a set of strong peaks at lengths equal to 3, 15, 24, 34 and 44 bp, differing by multiples of the DNA helical repeat (Fig. 5a and Supplementary Fig. 10). If we add 1 or 2 bp uniformly to each linker length to account for over-stretching of nucleosome DNAs, the linker length will follow a form of $10.41n + 4.6$ bp or $10.47n + 5.5$ bp (for integer n) respectively by Fourier analysis (for linker lengths ≤ 60 bp). The actual over-stretching varies across nucleosomes, which may alter details of the linker length distribution.

The non-random spacing of nucleosomes is also explicitly shown in the genomic DNA sequence. The periodicity of AA/TT/AT/TA motifs extends from the aligned unique nucleosomes to flanking regions beyond one nucleosome repeat length (Fig. 5b) with a period of 10.26 bp and a phase offset that accords with nucleosome spacing of $\sim 10n + 5$ bp, confirming the direct analyses above and previous findings[47] (Fig. 5c and Supplementary Methods). This suggests that given a nucleosome somewhere in the genome, the next nucleosome down the chain probably starts at approximately the opposite face of the double helical DNA, creating an open zigzag intrinsic chromatin structure, consistent with the two-start architecture of the 30 nm fibre[48].

Nucleosomes and DNA-binding proteins

We also examined the spatial relationships between nucleosomes and other high-resolution genomic features critical to genetic regulation.

Figure 5 | Nucleosome spacing and higher-order chromatin structures. **a**, Linker length distribution for all ~67,000 unique nucleosomes (black), for the top 50% of nucleosomes with highest NCP score-to-noise ratio (red) and for ~45,000 MNase 2 unique nucleosomes (blue). **b**, AA/TT/AT/TA frequency in the genomic DNA flanking the nucleosomes. **c**, Fourier transform of AA/TT/AT/TA signals in the unique nucleosomes (red), in downstream flanking regions (black), and the mean (blue) and 95th centile (green) of transform in the flanking regions after permuting the phase offset between the flanking region and the mapped nucleosome (Supplementary Methods).

Functional transcription-factor-binding sites, defined as sites that are conserved and occupied *in vivo*[49], were frequently contained inside mapped nucleosomes (Figs 6a, b, top and Supplementary Fig. 11a). This observation signifies that in a population of cells at any moment, some cells lack a nucleosome and have a bound factor, whereas other cells lack the bound transcription factor and have the nucleosome instead, or some cells might have transcription factor bound to the edge of the nucleosomes while the latter are partly unwrapped. More remarkably, within the nucleosome, transcription factor sites are strongly enriched at highly specific sites where the nucleosome sequence preferences best match those of the transcription factor; thus they can be predicted, given a position weight matrix for the transcription factor (Fig. 6a). Similarly, for the Hermes transposase, another sequence-specific DNA-binding protein, the transposon integration sites occur periodically through the nucleosome (Fig. 6b; bottom, and Supplementary Fig. 11b) at locations predictable by the transposase sequence specificity, demonstrating a type of multiplexing of genomic information.

Nucleosomes also influence elongation by RNA polymerase II (RNAP), with a recent genome-wide analysis showing enhanced pausing by RNAP predominantly within the first half of the nucleosome[8]. The chemical map shows enhanced pause sites across the full length of nucleosomes, suggesting that DNA downstream of the RNAP can remain wrapped even after RNAP has progressed far inside the nucleosome. However, we cannot exclude the possibility that this observation is complicated by the population average of nucleosomes from different locations within clusters. Notably, the RNAP pause sites do not occur at random locations in the nucleosome: rather, they occur periodically, at the DNA helical repeat (Fig. 6c and Supplementary Fig. 11b), indicating that RNAP stalls preferentially at specific locations relative to the nucleosome rotational position. The locations at which RNAP pauses are best reflected in a *dst1-Δ* strain that lacks the ability for RNAP to backtrack. Strikingly, in this mutant, RNAP pause sites again appear periodically, but at locations that are systematically displaced by ~5 bp: that is, by half of a DNA helical turn relative to the wild-type pause sites (Supplementary Fig. 11c). In the *dst1-Δ* strain, the RNAP has the greatest probability of pausing at locations where the DNA backbone faces out, away from the histones. This suggests that RNAP pauses at sites where further forward motion requires rotation around the DNA backbone, creating a steric clash between the polymerase and the histones. Conversely, the backtracking RNAP is most likely to pause at locations in the nucleosome at which the backbone faces inwards towards the histones (with maximal steric clash between histone and RNAP). In this situation, at the initiation of transcription, the first few forward steps will reduce this clash, as the polymerase rotates away from the histone surface.

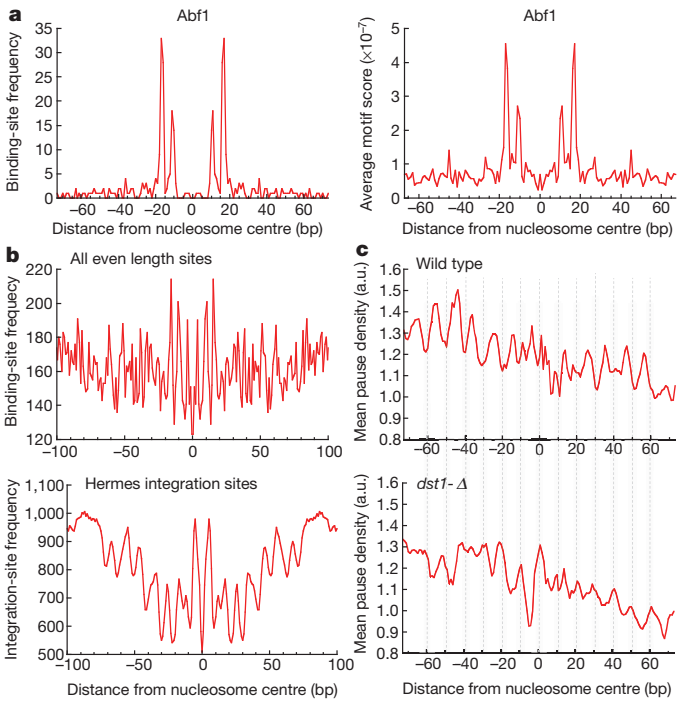


Figure 6 | High-resolution nucleosome centre positions relative to genomic features. **a**, Frequency of functional Abf1 sites (left) and the average motif score calculated using position weight matrix for Abf1 (right) relative to nucleosome centres. **b**, Frequency of all even-length functional transcription-factor-binding sites (top) and Hermes transposon integration sites (bottom) relative to the nucleosome centres. **c**, Frequencies of RNA polymerase II pause sites mapped inside nucleosomes +2 to +4 relative to the nucleosome centres, for wild-type yeast cells (top) and *dst1-Δ* yeast cells (lacking polymerase backtracking ability; bottom).

Conclusions and prospects

We developed a new method for mapping nucleosome centres in yeast with unprecedented consistency and accuracy. The resulting map shows significantly stronger dinucleotide signals in nucleosome DNAs than MNase maps. These signals are present and comparable in nucleosomes spanning entire protein-coding regions, demonstrating an important role of the DNA sequence features in detailed positioning and spacing of nucleosomes genome-wide. The observed linker length distribution suggests that the genome prefers an organized local chromatin structure. The exact locations of nucleosomes are associated with diverse biological phenomena, including the exact locations of transcription-factor-binding sites, transposon integration sites and RNAP pausing. We anticipate that this map will shed more light on many other genetic processes, including DNA replication, recombination, mutation, repair and genome evolution, because of the intimate relationship between nucleosome locations and these diverse processes.

METHODS SUMMARY

We genetically engineered *S. cerevisiae* to contain a cysteine at position 47 in histone H4. Cells grown to mid-log phase were collected, permeabilized and labelled with *N*-(1,10 phenanthroline-5-yl)iodoacetamide. The label covalently bound to the cysteine and allowed for copper chelation. Copper chloride, mercaptoproprionic acid and hydrogen peroxide were added sequentially, creating hydroxyl radicals that cleaved the nucleosomal DNA at sites flanking the centre. After the mapping reaction, the genomic DNA was purified from the cells and run on an agarose gel. The lowest molecular mass DNA fragment (~150–200 bp) was purified and prepared for high-throughput parallel sequencing.

Received 23 September 2011; accepted 11 April 2012.
Published online 3 June 2012.

1. Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423,** 145–150 (2003).
2. Koslover, E. F., Fuller, C. J., Straight, A. F. & Spakowitz, A. J. Local geometry and elasticity in compact chromatin structure. *Biophys. J.* **99,** 3941–3950 (2010).
3. Li, G. & Widom, J. Nucleosomes facilitate their own invasion. *Nature Struct. Mol. Biol.* **11,** 763–769 (2004).
4. Raveh-Sadka, T., Levo, M. & Segal, E. Incorporating nucleosomes into thermodynamic models of transcription regulation. *Genome Res.* **19,** 1480–1496 (2009).
5. Mao, C., Brown, C. R., Griesenbeck, J. & Boeger, H. Occlusion of regulatory sequences by promoter nucleosomes *in vivo*. *PLoS ONE* **6,** e17521 (2011).
6. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128,** 707–719 (2007).
7. Petesch, S. J. & Lis, J. T. Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* **134,** 74–84 (2008).
8. Churchman, L. S. & Weissman, J. S. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469,** 368–373 (2011).
9. Owen-Hughes, T. & Workman, J. L. Experimental analysis of chromatin function in transcription control. *Crit. Rev. Eukaryot. Gene Expr.* **4,** 403–441 (1994).
10. Li, G., Levitus, M., Bustamante, C. & Widom, J. Rapid spontaneous accessibility of nucleosomal DNA. *Nature Struct. Mol. Biol.* **12,** 46–53 (2005).
11. Lipford, J. R. & Bell, S. P. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell* **7,** 21–30 (2001).
12. Dorn, E. S. & Cook, J. G. Nucleosomes in the neighborhood: new roles for chromatin modifications in replication origin control. *Epigenetics* **6,** 552–559 (2011).

13. Tachiwana, H. *et al.* Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476,** 232–235 (2011).
14. Cole, H. A., Howard, B. H. & Clark, D. J. The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proc. Natl Acad. Sci. USA* **108,** 12687–12692 (2011).
15. Dalal, Y., Furuyama, T., Vermaak, D. & Henikoff, S. Structure, dynamics, and evolution of centromeric nucleosomes. *Proc. Natl Acad. Sci. USA* **104,** 15974–15981 (2007).
16. Beckmann, J. S. & Trifonov, E. N. Splice junctions follow a 205-base ladder. *Proc. Natl Acad. Sci. USA* **88,** 2380–2383 (1991).
17. Schwartz, S., Meshorer, E. & Ast, G. Chromatin organization marks exon-intron structure. *Nature Struct. Mol. Biol.* **16,** 990–995 (2009).
18. Tilgner, H. *et al.* Nucleosome positioning as a determinant of exon recognition. *Nature Struct. Mol. Biol.* **16,** 996–1001 (2009).
19. Dingwall, C., Lomonossoff, G. P. & Laskey, R. A. High sequence specificity of micrococcal nuclease. *Nucleic Acids Res.* **9,** 2659–2673 (1981).
20. Fan, X. *et al.* Nucleosome depletion at yeast terminators is not intrinsic and can occur by a transcriptional mechanism linked to 3′-end formation. *Proc. Natl Acad. Sci. USA* **107,** 17945–17950 (2010).
21. Hörz, W. & Altenburger, W. Sequence specific cleavage of DNA by micrococcal nuclease. *Nucleic Acids Res.* **9,** 2643–2658 (1981).
22. Chung, H. R. *et al.* The effect of micrococcal nuclease digestion on nucleosome positioning data. *PLoS ONE* **5,** e15754 (2010).
23. Zhang, Y. *et al.* Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nature Struct. Mol. Biol.* **16,** 847–852 (2009).
24. Flaus, A., Luger, K., Tan, S. & Richmond, T. J. Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc. Natl Acad. Sci. USA* **93,** 1370–1375 (1996).
25. Flaus, A. & Richmond, T. J. Base-pair resolution mapping of nucleosome positions using site-directed hydroxy radicals. *Methods Enzymol.* **304,** 251–263 (1999).
26. Yager, T. D., McMurray, C. T. & van Holde, K. E. Salt-induced release of DNA from nucleosome core particle. *Biochemistry* **28,** 2271–2281 (1989).
27. Widom, J. Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* **34,** 269–324 (2001).
28. Thåström, A., Bingham, L. M. & Widom, J. Nucleosomal locations of dominant DNA sequence motifs for histone–DNA interactions and nucleosome positioning. *J. Mol. Biol.* **338,** 695–709 (2004).
29. Teif, V. B. & Rippe, K. Predicting nucleosomes positions on the DNA: combining intrinsic sequence preferences and remodeler activities. Predicting nucleosome positions on the DNA. *Nucleic Acids Res.* **37,** 5642–5655 (2009).
30. Field, Y. *et al.* Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* **4,** e1000216 (2008).
31. Floer, M. *et al.* A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell* **141,** 407–418 (2010).
32. Cockell, M., Rhodes, D. & Klug, A. Location of the primary sites of micrococcal nuclease cleavage on the nucleosome core. *J. Mol. Biol.* **170,** 423–446 (1983).
33. Albert, I. *et al.* Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446,** 572–576 (2007).
34. Engeholm, M. *et al.* Nucleosomes can invade DNA territories occupied by their neighbors. *Nature Struct. Mol. Biol.* **16,** 151–158 (2009).
35. Wang, J. P. & Widom, J. Improved alignment of nucleosome DNA sequences using a mixture model. *Nucleic Acids Res.* **33,** 6743–6755 (2005).
36. Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442,** 772–778 (2006).
37. Travers, A. A. & Klug, A. The bending of DNA in nucleosomes and its wider implications. *Phil. Trans. R. Soc. Lond. B* **317,** 537–561 (1987).
38. Segal, E. & Widom, J. What controls nucleosome positions? *Trends Genet.* **25,** 335–343 (2009).
39. Segal, E. & Widom, J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **19,** 65–71 (2009).
40. Thåström, A., Lowary, P. T. & Widom, J. Measurement of histone-DNA interaction free energy in nucleosomes. *Methods* **33,** 33–44 (2004).
41. David, L. *et al.* A high-resolution map of transcription in the yeast genome. *Proc. Natl Acad. Sci. USA* **103,** 5320–5325 (2006).
42. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* **16,** 2593–2620 (2002).
43. Mavrich, T. N. *et al.* A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18,** 1073–1083 (2008).
44. Zhang, Z. *et al.* A packaging mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332,** 977–980 (2011).
45. Lohr, D. & van Holde, K. E. Organization of spacer DNA in chromatin. *Proc. Natl Acad. Sci. USA* **76,** 6326–6330 (1979).
46. Routh, A., Sandin, S. & Rhodes, D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl Acad. Sci. USA* **105,** 8872–8877 (2008).
47. Wang, J. P. *et al.* Preferentially quantized linker DNA lengths in *Saccharomyces cerevisiae*. *PLoS Comput. Biol.* **4,** e1000175 (2008).
48. Dorigo, B. *et al.* Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306,** 1571–1573 (2004).
49. MacIsaac, K. D. *et al.* An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics* **7,** 113 (2006).

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

Acknowledgements We dedicate this paper to J.W., who guided this project. We thank R. Holmgren, A. Matouschek, B. Meyer, R. Phillips, E. Segal and O. Uhlenbeck for comments and discussions. We are grateful to Northwestern University's Genomic Core for all sequencing completed for this project. The work was supported by National Institutes of Health grants T32GM00806 (to K.B.), R01GM058617 (to J.W.), R01GM075313 (to J.-P.W.) and U54CA143869 (to J.W.).

Author Contributions K.B. did all the experimental work. L.X. and J.-P.W. developed the algorithm and performed the analyses. K.B., J.-P.W. and J.W. wrote the paper. J.W. directed the project.

Author Information Sequence data are deposited in National Center for Biotechnology Information Gene Expression Omnibus database under accession number GSE36063. 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 J.-P.W. (jzwang@northwestern.edu).