Evidence against a genomic code for nucleosome positioning

Zhang et al. reply:
It has been proposed that there is a "genomic code for nucleosome positioning"1 in which the pattern of nucleosome positions in vivo is determined primarily by the genomic DNA sequence and can be predicted. As experimental support for such "DNA-encoded nucleosome organization," Kaplan et al.2 generated genome-wide maps of nucleosomes assembled in vitro with purified histones and concluded that these are highly similar to maps of nucleosomes in vivo2. However, in similar experiments, we concluded that "intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo," thereby arguing against a nucleosome positioning code. The originally submitted correspondence of Kaplan et al., to which our response was written, was entitled "a genomic code for nucleosome positioning," and it disputed our analyses and
major conclusion. We are pleased to see that the current correspondence⁴ has now reduced the role of nucleosome sequence preferences from “encoding” to “influencing” in vivo nucleosome organization and leaves the issue of a code “for others to debate.”

The in vitro mapping data in both studies is quite similar, and there is agreement that intrinsic histone-DNA interactions contribute to certain aspects of nucleosome positioning in vivo. The implication that we argue against any biological role of intrinsic histone-DNA interactions is incorrect and indeed inconsistent with our work over the past 25 years⁵–⁸. Nevertheless, we do disagree on the following: (i) the use of nucleosome occupancy measurements to assess nucleosome positioning; (ii) the impact of systematic errors in nucleosome occupancy measurements that overestimate the similarity between in vivo and in vitro samples; (iii) the ability of in vitro assembled nucleosomes to recapitulate the striking in vivo nucleosomal pattern; and (iv) the meaning of a nucleosome code. An independent analysis⁹ of the two key papers²,³ has supported our viewpoint.

The concept that histones have DNA sequence preferences for nucleosome formation was established 25 years ago. In pioneering experiments involving the sequencing of nucleosomal DNA generated by micrococcal nuclease (MNase), the same technique used today, Horace Drew and Andrew Travers showed that nucleosomal DNA in vivo has strong rotational positioning with 10–base pair (bp) helical periodicity that is due to preferences for dinucleotides that face inwards and increases chromatin accessibility in vivo via its intrinsic DNA structure, particularly at yeast promoter regions where these sequences are highly enriched⁵–⁷. Indeed, poly(dA:dT) and (to a lesser extent) dinucleotide frequencies are the most important factors in the algorithm of Kaplan et al.² for predicting nucleosome occupancy.

Prior to the initial paper proposing a nucleosome positioning code, a direct comparison of the location of nucleosomes assembled on the yeast PET56-HIS3-DED1 region in vivo and in vitro with purified histones revealed that both promoter regions intrinsically disfavor nucleosome formation⁸. Furthermore, it was argued that DNA sequence is responsible for nucleosome depletion at most yeast promoter regions in vivo, based on genome-wide occupancy measurements in vivo. Specifically, the relative paucity of nucleosomes at promoter regions with respect to the corresponding coding regions is independent of transcriptional activity and hence is not due to activator– or RNA polymerase II elongation–dependent histone removal⁸. Thus, as the concept and specific aspects of how DNA sequence contributes to nucleosome location in vivo are well established and not at issue, the key disagreement is whether intrinsic histone-DNA interactions have the predominant role in setting up the in vivo pattern and thus constitute a code for nucleosome positioning.

Kaplan et al.² and Zhang et al.³ extend the earlier comparison⁸ of nucleosomes assembled in vivo and in vitro to the entire yeast genome, and at higher (in principle, nucleotide) resolution using high-throughput sequencing. In interpreting the resulting maps, a major conceptual issue concerns the difference between nucleosome ‘occupancy’ and ‘positioning’. Nucleosome occupancy reflects the average histone levels on a given region of DNA in a population of cells, but it does not address where individual nucleosomes are positioned (that is, differentially positioned nucleosomes within a genomic region all contribute to occupancy). In contrast, the translational position of an individual nucleosome refers to the specific 146-bp sequence covered by the histone octamer. On a population basis, positioning can range from perfect (all nucleosomes occupy a specific 146-bp stretch) to random (nucleosomes occupy all possible genomic positions equally). We did not criticize Kaplan et al.² for nucleosome occupancy measurements per se (indeed, we also made this useful measurement; see below for limitations) but rather for using occupancy measurements to infer nucleosome positioning. As acknowledged in their correspondence⁴, Kaplan et al.² did not perform translational positioning analyses in their original paper. In addition, the independent validation experiment mentioned in the correspondence⁴ uses a different method to measure histone occupancy¹⁴, but it does not address translational positioning, the key point of disagreement.

Zhang et al.³ explicitly examined translational positioning and the relationship between nucleosomes generated in vivo and in vitro, and we disagree with the correspondence⁴ on this point. Specifically, we defined positioned nucleosomes from the in vivo mapping as 20-bp windows centered on the peak position on a gene–by–gene and location basis (+1, +2, etc. with respect to the mRNA initiation site). We then measured the percentage of nucleosome centers within these windows (100% being the value expected for perfect positioning) in the in vitro (and in vivo) data and compared this to randomly positioned nucleosomes (Fig. 4b–d of ref. 3). This analysis is unaffected by nucleosome centers flanking the 20-bp window, and hence the problematic example given in the correspondence⁴ is incorrect and irrelevant (the issue raised does affect the genome-wide measurement of maximal nucleosome positioning degree in Fig. 4a of ref. 3, but this is not relevant to the direct comparison of in vitro and in vivo positions).

We note that our analysis is restricted to nucleosomes that are well positioned in vivo, but the role of intrinsic histone-DNA interactions in setting up the striking in vivo pattern is the key biological issue. The analysis cannot be done on weakly positioned nucleosomes, as their locations are ill defined due to sequencing limitations.

Using in vitro data generated in either paper, we estimated that ~20% of the in vivo positioned nucleosomes are positioned due to intrinsic histone-DNA interactions. As done previously¹⁵ and in contrast to the correspondence⁴, this estimate involved an explicit correction for random chance occurrence. Our estimate is consistent both with the previous observation that 2 out of 7 in vivo positioned nucleosomes in the PET56-HIS3-DED1 region were observed in vitro⁸ and with a previous estimate of ~25% based on computational predictions of positioned nucleosomes¹⁵.

In the correspondence⁴, the authors performed a related positioning analysis using 40-bp windows and obtained a value of 34–41% (perhaps as high as 49% with unspecified data smoothing). However, the calculated values strongly depend on the input parameters and definitions, and the size of the window is particularly important. Indeed, we obtain a value of ~30% when using 40-bp windows (quite similar to that in the correspondence⁴) but only ~15% when using 10-bp windows. Conceptually, a positioned nucleosome has a unique location (1-bp window), and the operational reason for using larger windows is to account for incomplete or excessive trimming of nucleosomes by MNase, which is experimentally unavoidable. Hence, values at smaller window sizes are more meaningful for nucleosome positioning measurements, whereas larger window sizes (for example, 40 bp, or ~25% of all possible positions) begin to approach measurements of nucleosome occupancy (that is, all possible positions), not positioning. In addition, by reporting positioning measurements at each individual base pair (as opposed to restricting such
CORRESPONDENCE

measurements to positioned nucleosomes), the correspondence1 is essentially converting positioning information into nucleosome occupancy.

We agree with Kaplan et al.2 that nucleosome occupancy is an important concept, and indeed the central conclusion of our earlier work8 is that “intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast.” However, we disagree that nucleosome occupancy can be measured simply by counting nucleosome reads. In particular, the use of MNase and Illumina sequencing introduces systematic errors in the measurements and results in overestimates of the similarity between in vivo and in vitro samples. Illumina sequencing shows systematic differences in DNA sequence coverage depending on base composition and causes artifactually high correlations between samples10. Indeed, although Kaplan et al.2 emphasize a correlation of 0.74 between their in vitro and in vivo samples, Stein et al.9 have shown that the correlation is only 0.3 when their in vitro sample is compared to an in vivo sample analyzed by high-resolution microarrays. MNase has well-known DNA sequence specificity17, and this influences both the relative cleavage of linker regions and the relative cleavage of nucleosomal regions as a function of MNase concentration3,18. We agree with Kaplan et al.2 that other parameters, notably sparseness of data, might lead to an underestimation of the correlation, but this issue has not been investigated.

Kaplan et al.2 does not have an explicit control for either DNA sequencing or for sequence specificity of MNase cleavage, making it difficult to determine the extent to which these issues affect the correlation between their in vivo and in vitro samples. Zhang et al.3 analyzed a sonicated control sample to assess DNA sequencing effects and observed a correlation of 0.15–0.2. In a recently performed control, we observed a correlation of 0.3 between MNase-digested naked DNA with all in vitro and in vivo nucleosomal samples, and this may be an underestimate due to sparseness of data. Thus, not only do nucleosome occupancy measurements not address nucleosome positioning, but methodological considerations also significantly reduce the correlation between in vivo and in vitro nucleosome occupancy.

Aside from the technical issues raised above, both studies agree that in vitro assembled nucleosomes do not show the striking in vivo pattern in which the +1 nucleosome centered just downstream from the mRNA initiation site is highly positioned, with more downstream nucleosomes arrayed in the coding region becoming gradually less positioned19,20. This pattern is the hallmark of ‘statistical positioning’ of nucleosomes from a fixed barrier such as a DNA-binding protein21 or perhaps a nucleosome-free region22. Kaplan et al.2 correctly argue that the low and nonphysiological level of histones in their in vitro assembly reaction is unsuitable for forming nucleosome arrays and hence observing statistical positioning, but this issue does not apply to Zhang et al.3, where the histone:DNA ratio was physiological and nucleosome arrays clearly evident. The use of limiting histone concentrations by Kaplan et al.2 is advantageous for measuring intrinsic affinities of different genomic regions. In this regard, differences in nucleosome positioning and occupancy between the two studies are of potential interest, although they do not affect the key issues discussed here.

The mechanism by which the +1 nucleosome is positioned is the key to understanding how the in vivo nucleosomal pattern is generated. In vitro, correctly localized +1 nucleosomes are formed only to a limited extent, and unlike the situation in vivo, the +1 nucleosome behaves similarly to all other nucleosomes (to +10) with respect to the degree of localization. In contrast, as shown by Zhang et al.3, the position of the +1 nucleosome in vivo is strikingly linked to the location of the mRNA initiation site and preinitiation complex in both yeast and flies, arguing for a transcription-based mechanism. A transcription-based mechanism for positioning the +1 (and more downstream) nucleosomes is further supported by the observation that the barrier for the in vivo pattern of statistical positioning occurs specifically at promoters (as opposed to terminator regions that also appear to be depleted of nucleosomes) and is unidirectional (only in the downstream direction)3. Lastly, the loss of RNA polymerase II significantly alters nucleosome positioning to more closely match in vitro preferences, arguing for an important role of transcription in determining nucleosome positioning in vivo18. These observations are in striking contrast to the transcription-independent depletion of nucleosomes at yeast promoter regions with respect to their corresponding coding regions5. Further, these observations are inconsistent with the idea that intrinsic histone-DNA interactions are central to establishing where nucleosomes are actually positioned (as opposed to being absent) in vivo, and the correspondence1 does not address these inconsistencies.

Lastly, we do not agree with the use of the terms “nucleosome code” and “DNA-encoded nucleosome organization” to describe the experimental observations, and indeed, these terms are not clearly defined9. In common parlance, a code involves a system of words, letters or symbols that convey definite meanings. The genetic code, by which nucleic acid sequence is translated into protein sequence with high accuracy, clearly fits this definition. In contrast, 15–40% similarity (depending on definitions and methods, and we believe that the lower values are more relevant for positioning) between in vitro and in vivo nucleosome positions clearly does not convey a definite meaning for DNA sequence. More generally, ‘preferences’ are conceptually different from “codes.” Thus, although intrinsic histone-DNA interactions contribute, they are not the major determinant of nucleosome positions in vivo. As such, the proposed nucleosome code is not supported.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Yong Zhang1, Zarmik Moqtaderi2, Barbara P Ratnner3, Ghia Euskirchen4, Michael Snyder4, James T Kadonaga1, X Shirley Liu5 & Kevin Struhl2

1Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Massachusetts, USA. 2Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA. 3Section of Molecular Biology, University of California, San Diego, La Jolla, California, USA. 4Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut, USA.

e-mail: kevin@hms.harvard.edu

To the Editor:

‘Occupancy’ is a measure of histone or nucleosome density. Occupancy is typically measured on a genomic scale using microarrays or through deep sequencing (Fig. 1). Kaplan et al.1 were correct in that the underlying DNA sequence has a predominant influence on occupancy levels in vivo. However, this and related work1–3 often interchanged the term ‘occupancy’ with ‘positioning’, which is confusing. ‘Positioning’ is a measure of the extent to which a population of nucleosomes resists deviating from its consensus location along the DNA and can be thought of in terms of a single reference point on the nucleosome, like its dyad (Fig. 1)4. A low standard deviation means high positioning. Zhang et al.5,6 were correct in that the underlying DNA sequences are not widespread determinants of nucleosome positioning in vivo, although they are major determinants at some positions. An important question now is how nucleosomes become uniformly spaced and precisely positioned in vivo.

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

B Franklin Pugh
Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania, USA.
E-mail: bfp2@psu.edu


A preoccupied position on nucleosomes

Figure 1. Illustration of how nucleosome occupancy and positioning differ. The upper panel shows a cross-section of a nucleosome, in which occupancy is distinguished from positioning. The lower panel shows how the two are measured. Occupancy is the area under the curve and reflects the local density of nucleosomes in a population, as illustrated by the column of spheres. Positioning or fuzziness is reflected in the standard deviation of the curve and is illustrated by how well the spheres are aligned in a column. The position of a nucleosome relative to some standard is indicated by how closely two peaks are separated. Comparing peaks of curves having high standard deviations is not likely to be meaningful because both peak locations have very high uncertainty.