

# New evidence that DNA encodes its packaging

Sevinc Ercan & Jason D Lieb

**New experimental approaches combined with statistical models show that DNA sequence strongly influences how the genome is packaged into nucleosomes. The studies predict that genes regulated by fundamentally different mechanisms have distinct nucleosome positioning signatures encoded in their DNA.**

Nowadays, even most schoolchildren are familiar with the concept that DNA encodes genetic information. But many unanswered questions remain. How much information is encoded directly by the nucleotide sequence of genome, and in what form? Two studies, one by Ioshikhes *et al.*<sup>1</sup> on page 1210 of this issue and one by Segal *et al.*<sup>2</sup>, present strong evidence for a newly discovered type of information encoded in DNA, one that controls how the genomic DNA is packaged into molecular spools called nucleosomes.

## A diffuse 'code'

A unifying characteristic of all organisms with a DNA-based genome is the genetic code, in which nucleotide triplets can be translated into protein sequence using fairly rigid and straightforward rules. However, genome sequences also contain diffusely encoded information that is more difficult to approach experimentally. Familiar examples include the DNA sequence motifs bound by regulatory proteins which, unlike the genetic code, are flexible such that many variations of a motif may be bound by a given protein<sup>3</sup>. The newly discovered set of sequence-based rules for nucleosome positioning is an extreme example of diffusely encoded information, which is part of the reason their discovery comes over 40 years after elucidation of the genetic code.

Most genomic DNA in eukaryotes is packaged into nucleosomes, each of which is composed of 147 bp of DNA wrapped around an octamer of histone proteins. The nucleosomes are separated from each other by 'linker' DNA that is typically between 10–50 bp in length. Sharp bending of DNA is required to form nucleosomes<sup>4</sup>, and DNA sequence affects the formation and stability of nucleosomal DNA *in vitro*<sup>5</sup>. A long-standing question is whether DNA sequence governs the positioning of nucleosomes at most genomic locations in living cells. The position of a nucleosome is important because nucleosomes

can regulate how the information at a genomic location is used<sup>6</sup>.

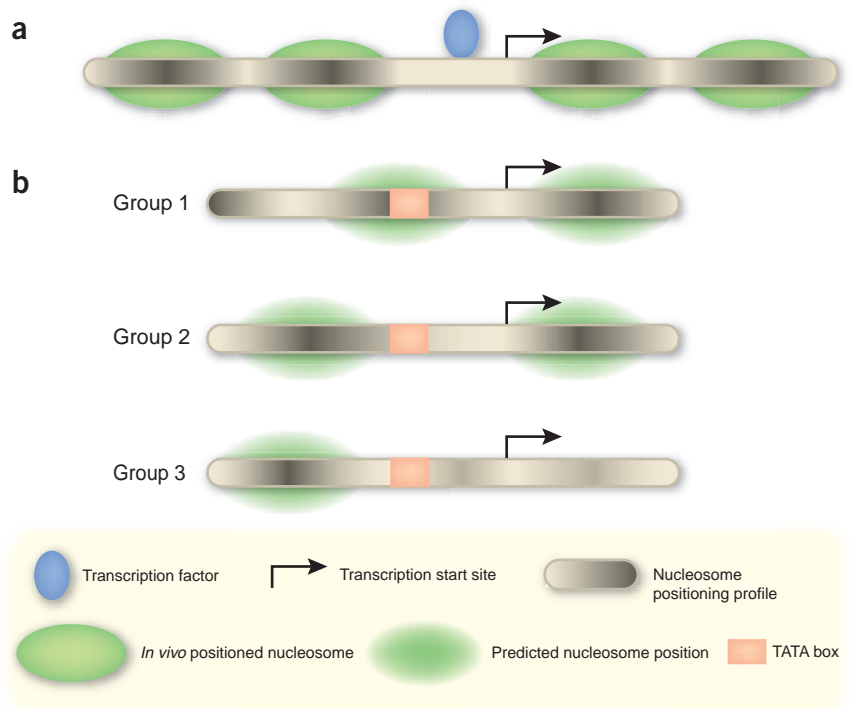
To explore this question, Segal *et al.*<sup>2</sup> identified DNA sequences that were stably incorporated into nucleosomes and used those sequences to construct a dinucleotide-based model for nucleosome positioning. Similarly, Ioshikhes *et al.*<sup>1</sup> used the occurrence of periodically distributed AA and TT dinucleotides, which had been previously shown to promote stable nucleosome formation, to define what they termed a 'nucleosome positioning sequence'<sup>7</sup>. Both groups then used their models to predict nucleosome positions across the entire yeast genome and compared the predictions with experimentally determined *in vivo* nucleosome locations<sup>8</sup>.

The results are astonishing: both models predicted the *in vivo* locations of nearly half of the nucleosomes. This suggests that the DNA

sequence of the genome itself partly determines the locations of nucleosomes. Significantly, constructing the computational model using data from experiments performed with chicken DNA or with synthesized DNA also predicted nucleosome positioning in yeast, suggesting that the general properties of DNA governing nucleosome positioning are fundamental and applicable to a wide range of genomes<sup>2</sup>. Ioshikhes *et al.*<sup>1</sup> leveraged this conservation by averaging predicted profiles from six related yeast species, thereby amplifying the encoded signal, which may be difficult to detect in any single genome.

## Functional consequences

The discovery of rules that influence nucleosome positioning have immediate implications because the exact placement of nucleosomes could mean big changes in how other genomic



**Figure 1** Nucleosome positioning profiles determined by DNA sequence predict *in vivo* nucleosome locations. (a) A typical nucleosome profile in yeast. (b) Access to the TATA box may be regulated by a nucleosome positioning profile.

Sevinc Ercan and Jason D. Lieb are in the Department of Biology and Carolina Center for the Genome Sciences, 202 Fordham Hall, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA. e-mail: jlieb@bio.unc.edu

Kim Caesar

signals are used. For example, the promoters of almost all yeast genes harbor a nucleosome-free region very near the transcription start site<sup>8</sup>. Moving a nucleosome a few bases in one direction or the other could obscure (or allow access) to a transcription factor binding site. Indeed, both studies<sup>1,2</sup> found that functional sites for transcription factor binding are located in regions where nucleosome formation is not predicted to be favorable. This observation indicates an important role for DNA sequence at this nucleosome-free region and is consistent with a model in which the genome directs transcription factors to their functionally relevant sites by keeping them depleted of nucleosomes, thereby directly contributing to regulation of transcription at the chromatin level.

In addition to affecting transcription factor binding, a small difference in positioning could regulate assembly of the general transcription machinery or control transcriptional initiation itself. This raises the question of whether the nucleosome positioning profiles of genes regulated by fundamentally different mechanisms show different nucleosome positioning patterns. In yeast, about 20% of promoters contain a regulatory element called a TATA

box upstream of the transcription start site, and regulation of these genes differs mechanistically from genes that lack the TATA box<sup>9</sup>. Ioshikhes *et al.*<sup>1</sup> predict that the nucleosome positioning sequence profiles of these two groups of genes differ. TATA-less promoters have uniform DNA sequence-based positioning profiles that are very similar to experimentally determined stereotypic nucleosome positions (Fig. 1a). In contrast, TATA-containing promoters can be divided into at least three distinct classes of nucleosome organization and may be activated in genetic backgrounds that impair nucleosome function (Fig. 1b)<sup>1</sup>. This suggests that these promoters may be specially poised for regulation by subtle changes in nucleosome positioning, which could be mediated by *trans*-acting factors like the ATP-dependent nucleosome remodeling complexes<sup>10</sup>.

### Challenges and prospects

Broadly speaking, the papers are in close agreement, but many important issues require further research. Both models are relatively simple, and improved sequence-based models or integration of structural information about the nucleosome and DNA may allow

for more accurate predictions. Certainly, DNA sequence is not the sole determinant of nucleosome positioning. At promoters regulated by chromatin remodeling enzymes, the specific mechanisms used to move nucleosomes to alternate DNA-encoded positions under different conditions remain unresolved. The good news is that nucleosomes and the biophysical properties of DNA are nearly identical in all eukaryotic species, so the DNA-based rules that influence positioning are likely to apply, with slight modification, to more complex genomes.

1. Ioshikhes, I.P., Albert, I., Zanton, S.J. & Pugh, B.F. *Nat. Genet.* **38**, 1210–1215 (2006).
2. Segal, E. *et al. Nature* **442**, 772–778 (2006).
3. Pabo, C.O. & Nekludova, L. *J. Mol. Biol.* **301**, 597–624 (2000).
4. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. *Nature* **389**, 251–260 (1997).
5. Widom, J. *Q. Rev. Biophys.* **34**, 269–324 (2001).
6. Lu, Q., Wallrath, L.L. & Elgin, S.C. *J. Cell. Biochem.* **55**, 83–92 (1994).
7. Ioshikhes, I., Bolshoy, A., Derenshteyn, K., Borodovsky, M. & Trifonov, E.N. *J. Mol. Biol.* **262**, 129–139 (1996).
8. Yuan, G.C. *et al. Science* **309**, 626–630 (2005).
9. Huisinga, K.L. & Pugh, B.F. *Mol. Cell* **13**, 573–585 (2004).
10. Vignali, M., Hassan, A.H., Neely, K.E. & Workman, J.L. *Mol. Cell. Biol.* **20**, 1899–1910 (2000).

## A WNK in the kidney controls blood pressure

Thomas M Coffman

**Precise control of sodium excretion by the kidney is critical for maintaining fluid and electrolyte homeostasis. A new mouse model provides compelling evidence that a kinase, WNK4, provides key signals for regulating blood pressure and potassium balance by controlling the structure and function of the distal convoluted tubule.**

Hypertension is one of the most common diseases of the modern world. Despite its high prevalence, the specific causes of high blood pressure cannot be discerned in the majority of affected individuals. More than 30 years ago, Guyton and colleagues first suggested that defective handling of sodium by the kidney is a requisite, final common pathway in hypertension pathogenesis<sup>1</sup>. Powerful support for this hypothesis has emerged from the genetic studies of Lifton and colleagues showing that virtually all mendelian disorders with major effects on blood pressure are caused by gene variants affecting salt reabsorption within a circumscribed portion of

the distal nephron<sup>2</sup>. On page 1124 of this issue, Lalioti *et al.*<sup>3</sup> use a mouse model to explore the genetic basis of one of these disorders, pseudo-hypoaldosteronism type II (PHAII), identifying a novel pathway regulating solute excretion in the distal convoluted tubule (DCT) (Fig. 1).

### WNKs and ion transport

PHAII is a rare, autosomal dominant disease characterized by hypertension and elevated serum potassium levels, with normal renal function<sup>4</sup>. PHAII has been associated with mutations in genes encoding WNK1 and WNK4 (ref. 5), which belong to a small family of kinases characterized by the absence of a conserved lysine found in the catalytic domain of all other known serine-threonine kinases (hence the name ‘with no K’ (lysine))<sup>6</sup>. WNK1 and WNK4 are expressed in the distal nephron, specifically in the DCT and collecting duct<sup>5</sup>. In studies dissecting their physiological role in blood pressure regulation,

attention has focused on the potential interactions of WNKs with a specific transporter, the sodium-chloride cotransporter (NCC), located in the DCT<sup>7</sup> (Fig. 1). NCC is the molecular target of thiazide diuretics, a widely used class of antihypertensive drugs. These agents effectively reverse hypertension and hyperkalemia in PHAII (ref. 4). Furthermore, the PHAII phenotype of hypertension, hyperkalemia and hypercalciuria is the virtual mirror image of the low blood pressure, hypokalemia and hypocalciuria that are the major features of Gitelman syndrome, which is caused by loss-of-function mutations in the gene encoding NCC (*SLC12A3*)<sup>2</sup>.

A direct link was established when coexpression of WNK4 with NCC in *Xenopus laevis* oocytes was found to cause dramatic inhibition of NCC function owing to impaired expression of NCC proteins in the cell membrane<sup>8–10</sup>. To examine the functions of WNK4 *in vivo*, Lalioti *et al.* generated transgenic mice using BAC clones

Thomas M. Coffman is in the Division of Nephrology, Department of Medicine, Duke University and the Durham Veterans Administration Medical Centers, Durham, North Carolina, 27710 USA.  
e-mail: [tcoffman@duke.edu](mailto:tcoffman@duke.edu)