Review

Cell PRESS

Gene regulation by nucleosome positioning

Lu Bai[1] and Alexandre V. Morozov[2]

[1] The Rockefeller University, New York, NY, 10065, USA
[2] Department of Physics and Astronomy and BioMaPS Institute for Quantitative Biology, Rutgers University, Piscataway, NJ 08854, USA

To achieve high compaction, most genomic DNA in eukaryotes is incorporated into nucleosomes; however, regulatory factors and transcriptional machinery must gain access to chromatin to extract genetic information. This conflict is partially resolved by a particular arrangement of nucleosome locations on the genome. Across all eukaryotic species, promoters and other regulatory sequences are more nucleosome-depleted, whereas transcribed regions tend to be occupied with well-positioned, high-density nucleosomal arrays. This nucleosome positioning pattern, as well as its dynamic regulation, facilitates the access of transcription factors to their target sites and plays a crucial role in determining the transcription level, cell-to-cell variation and activation or repression dynamics.

General view of nucleosomes and transcription

The nucleosome is the fundamental repeating unit of eukaryotic chromatin, consisting of 147 bp of DNA wrapped around a histone core [1]. The histone genes and nucleosome structure are extremely well conserved among eukaryotic species. Importantly, the nucleosome is not just a static building block of chromatin structure. Histones are constantly evicted and reassembled onto the DNA template in a locus-specific fashion, and during such turnover they are sometimes replaced by histone variants. Histones can be marked by a large number of post-translational modifications. These histone variants and covalent modifications affect histone–DNA and histone–histone interactions, as well as interactions between histones and regulatory factors. The accessibility of a binding site for a particular factor is also affected by its position with respect to neighboring nucleosomes: sites in the nucleosome-free linker DNA should be easier to access than those in the middle of a nucleosome. These nucleosome properties, including positioning, turnover and histone variations and modifications, play essential roles in gene regulation by affecting the transcriptional competence of various chromatin regions.

With the recent development of global nucleosome mapping techniques, the research on nucleosome positioning has entered a fast-growing phase. Here, we review the recent progress made in this area, focusing in particular on the relationship between nucleosome positioning and gene regulation.

Corresponding author: Bai, L. (lbai01@rockefeller.edu).

Genome-wide nucleosome distribution

Genome-wide mapping reveals ''canonical'' nucleosome positioning in vivo

The development of microarrays and next-generation sequencing has made it possible to map nucleosome positions on the global scale [2] (Box 1). To date, nucleosome positions have been mapped on many genomes including those of yeast, fly, worm and human, and many more experiments are underway [3–14].

Nucleosome positioning across the genome is far from random. The most striking feature revealed by global mapping is the contrast between nucleosome density in regulatory regions and that in transcribed sequences. In budding yeast, >90% of the promoters contain stretches of DNA with very low nucleosome occupancy [3,4,6,8,11,12]. These nucleosome-depleted regions (NDRs) are on average ~150 bp in length, roughly enough to accommodate a single nucleosome. The NDRs play a crucial role in transcription regulation (see below). The rest of the promoter sequence is assembled into nucleosomes. Some of these nucleosomes have unusual properties, including enrichment in certain histone variants and high turnover rates [15–18], which might also contribute to the regulation of gene expression.

Sequences downstream of the transcription start site (TSS) are usually occupied by well-positioned nucleosomes. Interestingly, when the genome-wide nucleosome density map is aligned with the TSSs of individual genes, nucleosomes in the vicinity of TSSs tend to be located at specific positions, especially the so-called +1 nucleosome (Figure 1a). Further upstream and downstream, nucleosomes gradually lose their phasing. It should be stressed that not all promoters have the nucleosome distribution shown in Figure 1a. For instance, some promoters have nucleosomes in between NDRs and TSSs, such as the *PHO5* and *CLN2* promoters [19,20]. There are also promoters that lack detectable NDRs.

This stereotypical nucleosome configuration is to some extent conserved across all eukaryotic species (Figure 1b). In human cells, nucleosome positioning near TSSs also exhibits the canonical configuration, especially for CpG promoters [9,21–23]. More complicated than in yeast, nucleosome coverage in higher eukaryotes could be cell-type-specific. For instance, nucleosome depletion on the Polycomb response elements of the HOX promoters has been shown to occur during the differentiation of human embryonic stem cells, presumably to facilitate the binding of Polycomb group proteins to achieve gene silencing [24]. The loss of an

0168-9525/$ – see front matter © 2010 Elsevier Ltd. All rights reserved. doi:10.1016/j.tig.2010.08.003 Trends in Genetics xx (2010) 1–8

Box 1. Genome-wide nucleosome mapping methods

The detailed methods used for nucleosome mapping can vary slightly in different labs and for different species but the principle remains the same [2–13]. Live cells or spheroplasts (sometimes crosslinked with formaldehyde) are made permeable and treated with MNase, an endo-exonuclease from *Staphylococcus aureus* that preferentially digests linker DNA (uncovered DNA between neighboring nucleosomes) versus nucleosomal DNA. The reaction is usually carried out to the extent that most of the chromatin is digested to mononucleosomes, with subpopulations of di- and trinucleosomes. The digested chromatin is sometimes purified further by immunoprecipitation with histone antibodies before the crosslink is reversed (if necessary) and the DNA segments are extracted. Mononucleosomal-sized DNA fragments (150–200 bp) are selected by gel purification, and their locations on the genome are mapped by either hybridizing to DNA microarrays or high-throughput sequencing followed by the alignment of sequence tags to the reference genome.

NDR has also been observed in the promoter of the tumor-suppressor gene *MLH1* in cancer cells [25]. Therefore, given the complex chromatin organization and regulatory pathways found in higher eukaryotes, studying nucleosome positioning and its dynamics could provide crucial insights into identifying TSSs, enhancers and other key genomic elements involved in transcriptional reprogramming.

Potential problems with the nucleosome mapping method and interpretation

Although high-throughput nucleosome mapping is extremely informative, we should be aware of several potential problems with the current approaches. Most of these studies employ micrococcal nuclease (MNase) digestion to liberate mononucleosome core particles for subsequent sequencing or microarray hybridization (Box 1). It is well known that MNase preferentially cleaves DNA at A/T-rich sites [26–28]. As a result, large nucleosome maps are biased to some extent towards DNA sequences that are more amenable to MNase digestion. Indeed, mononucleosome-sized sequences obtained by the partial MNase digestion of naked genomic DNA exhibit some of the features of nucleosome positioning sequences collected in high-throughput experiments, such as the enrichment of G/C- and depletion of A/T-containing dinucleotides, which is reversed immediately outside the nucleosome border (unpublished observation). In addition, it was recently shown that certain nucleosomes in yeast and human promoters could be especially unstable with respect to salt and MNase [29,30]. Therefore, NDRs previously identified by MNase digestion might not all be nucleosome-free.

Hybridizing nucleosomes to tiling microarrays (which provide a direct nucleosome occupancy readout) introduces a certain loss of resolution: although the latest Affymetrix arrays tile the yeast genome with 4–5 bp steps [4,31], some of the work has employed an earlier customized array with 20 bp steps [3,32]. Although high-throughput sequencing formally provides 1 bp resolution, exact nucleosome positions are still unknown because MNase does not cut DNA precisely at the nucleosome boundary. Furthermore, high-throughput sequencing datasets contain regions of abnormally high and low read coverage, which need to be excluded before genome-mapped read coverage can be interpreted as nucleosome occupancy [12].

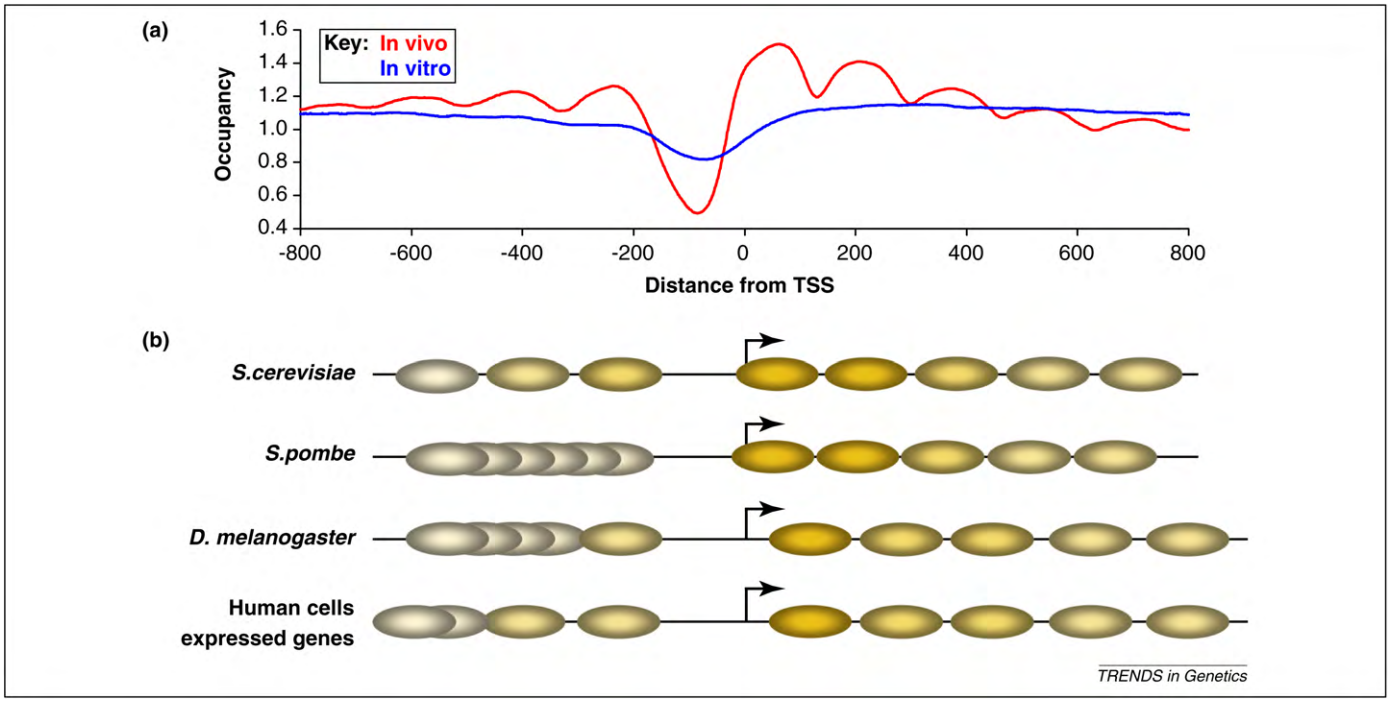

Figure 1. Stereotypical nucleosome positioning in the vicinity of gene transcripts. **(a)** Averaged *in vivo* (red) [12] and *in vitro* (blue) [70] nucleosome occupancy of *Saccharomyces cerevisiae* genes aligned by their TSSs. **(b)** Schematic representation of typical *in vivo* nucleosome positions in *S. cerevisiae* (interpreted from data in **(a)** and in several other species (based on similar measurements [7,9,14]). Arrow: TSS. Yellow oval: nucleosome. The more yellowish the nucleosome, the better it is positioned relative to the TSS. The gray ovals overlapping with each other represent nucleosomes without any phasing. Note that the +1 nucleosome is further downstream in *Drosophila* and human cells compared with yeast. The nucleosomal repeat length (average distance between neighboring nucleosomes) can also be different in different species.

Effect of nucleosome positioning on gene expression

Nucleosome configuration versus transcription on inducible promoters

A clear example illustrating how nucleosomes function in transcriptional regulation is provided by the phosphate-regulated *PHO5* promoter (*PHO5pr*). At high phosphate concentration, the repressed *PHO5pr* is assembled into an array of well-positioned nucleosomes interrupted by a ~80 bp NDR. The NDR is constitutive (independent of the transcriptional status of *PHO5pr*), and contains one of the binding sites of the activator Pho4 [19] (Figure 2a). Upon activation, the nucleosomes –1 to –4 are disassembled from the promoter, exposing an additional Pho4 binding site, as well as allowing transcriptional machinery to access the TATA box. On the *GAL1-10* promoter, four binding sites of the activator Gal4 are located in an unstable, partially unwound nucleosome bound by the nucleosome-remodeling enzyme RSC, which apparently allows efficient Gal4 binding [33] (Figure 2b). Under activating conditions, nucleosomes flanking the upstream activation sequence are rapidly removed from the promoter and the Gal gene is activated [34]. Similar changes in nucleosome configuration are observed on *PHO8pr* and *CHA1pr* [35,36].

Besides genes responding to environmental cues, nucleosome configuration changes through the cell cycle at many cell-cycle-regulated promoters [37]. *CLN2pr*, for example, contains three nucleosomes and a ~300 bp NDR (Figure 2c). During the cell cycle, the NDR remains nucleosome-depleted, whereas the occupancy of –1 and –2 nucleosomes fluctuates. The timing of nucleosome eviction coincides with the activation of *CLN2* [20]. The activator-binding and nucleosome-depletion process can even happen in tandem: on *HOpr*, the binding of activator Swi5 in the upstream NDR causes the eviction of neighboring nucleosomes and eventually leads to a second "wave" of downstream nucleosome depletion, which exposes multiple SBF (activator) binding sites and allows efficient SBF binding and transcriptional activation [38] (Figure 2d). This complicated pathway serves to implement cell-type regulation (Swi5 is only active in mother cells) on top of SBF cell cycle regulation.

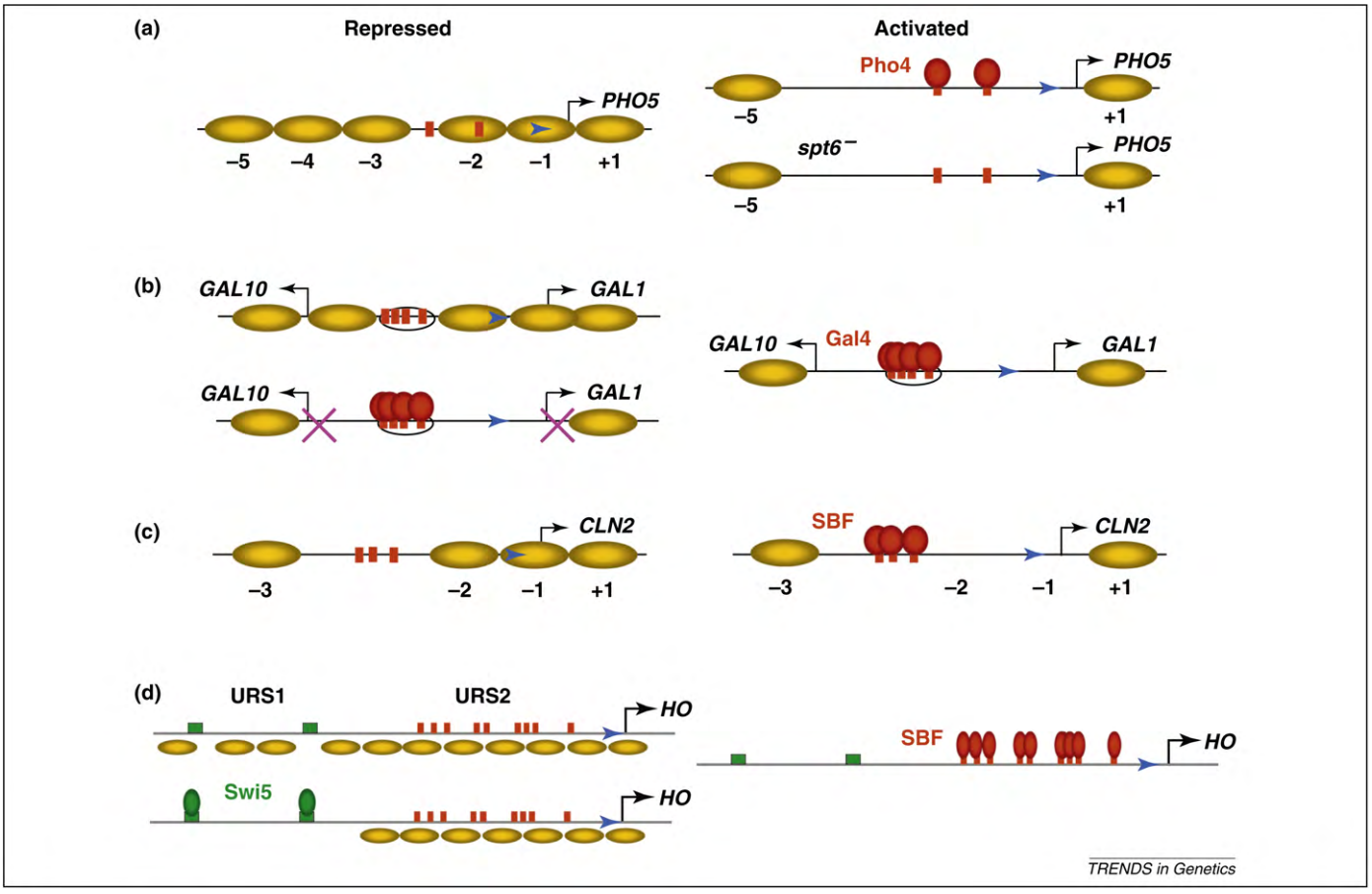

Figure 2. Nucleosome configurations of several *S. cerevisiae* promoters in their transcriptionally repressed and activated states. In all panels, the arrows represent the TSS; blue triangles: TATA boxes; red or green rectangles: activator binding sites; red or green ovals: activators; yellow ovals: nucleosomes. **(a)** The repressed *PHO5pr* contains a well-positioned nucleosome array with an NDR in the middle. During activation, binding of the transcription factor Pho4 leads to the disassembly of nucleosomes –1 to –4, exposing all the functional sites and allowing the transcription machinery to bind [19]. If the nucleosomes fail to reassemble on *PHO5pr* (e.g. by the deletion of the histone chaperone *SPT6*), transcription can persist even after Pho4 dissociation, indicating that the nucleosomes might be solely responsible for *PHO5pr* repression [40]. **(b)** The *GAL1-10* promoter contains a RSC–nucleosome complex covering four Gal4 (activator) binding sites, which apparently allows the efficient binding of the Gal4 protein [33]. Similar to *PHO5pr*, nucleosomes are rapidly removed from *GAL1-10pr* upon its activation. However, in certain conditions *GAL1-10pr* can remain repressed and nucleosome-free, indicating that nucleosomes are not the only mechanisms for gene repression [34]. C) *CLN2pr* is a cell cycle-regulated promoter that has three SCBs located in an NDR. The NDR is constitutive during the cell cycle, whereas the downstream nucleosomes dissociate during *CLN2* activation and reassemble during repression [20]. D) The activation of *HOpr* follows two steps: first, Swi5 (green) binds on the upstream NDR located around the upstream repressing sequence 1 (URS1) and causes the eviction of neighboring nucleosomes. Second, the "wave" of nucleosome depletion extends downstream in the URS2, allowing the efficient binding of SBF and transcriptional activation [38].

Activators require coactivators to change nucleosome configuration, including nucleosome remodelers, histone modifiers, histone chaperones and FACT (for facilitating chromatin transcription) complexes [38–41]. In many cases, coactivators load onto the promoter in a sequential manner, indicating that they can be recruited by each other [42]. Different promoters seem to require different sets of coactivators. Such specificity could be partially explained by the binding preferences of different activators [43]. It could also be affected by the nucleosome configuration on the promoter. For instance, the disruption of Sin3p–Rpd3p histone deacetylase complex, which targets *HOpr*, would increase the nucleosome acetylation level on *HOpr* and reduce the need of Gcn5 for *HO* activation [44,45]. Finally, the activator binding site occupancy, which is affected by activator concentration, binding strength and nucleosome density over the binding site, has been found to significantly affect the requirement of coactivators [46–48].

What happens during repression? In most cases, repression is accompanied by chromatin reassembly. On *PHO5pr*, histone deposition is mediated by the histone chaperone Spt6 [49]. In the absence of Spt6, *PHO5pr* remains nucleosome-free and transcriptionally active after the dissociation of the activator Pho4 [50]. This result indicates that nucleosome coverage could be solely responsible for *PHO5pr* repression, and nucleosome-free *PHO5pr* is sufficient for supporting transcriptional activity. However, under certain conditions, Gal4 can continue to bind on the *GAL1-10pr* and recruit factors that maintain the promoter nucleosome-free while transcription is repressed [34]. Therefore, nucleosome coverage of promoter and repression is not always tightly coupled.

Nucleosome configuration versus transcription on constitutive promoters

The promoters discussed above are so-called “inducible promoters”. They are more likely to contain a TATA box [51] and, unlike the canonical nucleosome configuration (Figure 1a), the regions immediately upstream of the TSS on these promoters tend to be nucleosome-rich under repressive conditions. In this way, the TATA box and TSSs are covered by nucleosomes, keeping the TATA box inaccessible to the TATA-binding protein (TBP) [52] and basal transcription at a low level, which can help achieve a larger dynamic range in expression.

By contrast, constitutive promoters are often TATA-less and more likely to adopt the canonical nucleosome configuration [11,53]. Some of them, such as *AKY2* and *RIO1*, lack “traditional” activating signals and only contain sequence elements that antagonize nucleosome formation [54,55]. In these cases, the NDRs immediately upstream of the TSS are apparently sufficient to support transcription, presumably by allowing the transcription machinery to bind without competing with nucleosomes. Some constitutive promoters bind transcription factors to modulate the level of transcription, albeit within a smaller dynamic range. For instance, the expression level of some constitutive promoters varies when a certain environmental condition is changed (e.g. a switch of carbon source). However, unlike inducible promoters, such transcriptional regulation is usually not accompanied by a significant rearrangement of promoter nucleosomes [31].

Nucleosome effects on transcription factor binding

Many functional activator binding sites are located in NDRs. Even on inducible promoters, which tend to have higher nucleosome coverage, some activator binding sites are usually pre-exposed (Figure 2). For the transcription factor SBF, among 50 SBF-bound promoters, 49 have at least one of their candidate SBF binding sites (SCBs) in constitutive NDRs [20]. The only exception is the *HO* promoter, where the NDR over SCBs is created right before activation (see above). The prediction of Leu3 binding sites *in vivo* was found to be significantly improved by considering its preference to binding sites located inside NDRs [56]. The correlation between NDRs and binding sites could even be exploited to identify new regulatory elements; for example, the NDRs near human *HOX* genes have revealed targeting sequences of Polycomb proteins [24]. These results strongly indicate the importance of the NDR localization of regulatory sites.

However, NDRs are not always essential for activator binding. Some factors, such as NF-κB p50, can apparently bind to nucleosomal DNA with the same affinity as free DNA by accommodating the nucleosome structure in a ternary complex including DNA, nucleosome and transcription factor [57]. Many other factors can access their nucleosomal binding sites but with a lower binding affinity [58]. This process is likely to be promoted by nucleosome dynamics including the spontaneous “wrapping–unwrapping” of nucleosomal DNA under thermal fluctuations [59], and histone turnover that might result in temporarily released naked DNA [17,18]. The occupancy of activators on its nucleosomal site should be affected by the activator concentration, intrinsic binding affinity, stability of DNA on the nucleosome surface, location and orientation of the binding sites relative to the nucleosome core and presence of histone variants and modifications. In the presence of multiple binding sites in a single nucleosome, one bound factor can secure partially unwrapped nucleosomal DNA and facilitate the binding of a second factor, inducing apparent binding cooperativity [60]. The factors bound on the nucleosomal sites can then recruit coactivators, remodel nucleosomes (including the ones covering their own binding sites) and activate transcription. For instance, Gal4, Pho4 and SBF have all been reported to be able to access their embedded binding sites, disrupting local nucleosomes and driving gene expression to higher levels [20,33,48,61,62].

Given the observations above, what could be the advantage of having activator-binding sites localized in NDRs? First, activation via nucleosome-covered sites might require higher activator concentration and have elevated dependence on nucleosome remodeling complexes [20,48]. Second, the binding competition between the activator and nucleosomes can delay the transcriptional response to the activating signal, a case demonstrated by Gal4 activation [33]. Third, nucleosome-embedded activator-binding sites could lead to higher transcriptional variability. For instance, with buried SCBs, SBF activation becomes bimodal: strong activation is observed in some cell cycles but is undetectable in others. By contrast, nucleosome-free SCBs lead to reliable

activation, once per cell cycle [20]. Consistently, there is a genome-wide positive correlation between nucleosome density on activator binding sites and transcriptional noise [11,53,63]. All these observations are potentially related: the weak binding of the activator might elevate the need for nucleosome remodelers [46,47]; the extra remodeling step could delay the activation onset; and a slower response to the activation signal could result in no activation at all when the activator is only available transiently (as in the case of SBF). The requirement for higher activator concentration, elevated dependence on the nucleosome remodeler, slower activation response and unreliable activation are likely to entail a fitness cost, as has been demonstrated directly using *CLN2pr* variants [20].

Because the activators access their nucleosomal versus exposed binding sites with different affinities and timescales, promoters sometimes fine-tune their nucleosome coverage to achieve certain activation dynamics. On Pho4-activated promoters, it was found that the affinity of the NDR-localized Pho4 binding site is the primary determinant of the minimal Pho4 concentration for activation, whereas the nucleosome-embedded site contributes significantly more to the final expression level once the nucleosomes are removed [62].

Nucleosome effect on transcription initiation and TSS selection

In contrast to activators, which are able to access nucleosome-covered-binding sites, PolII assembly on promoters is more strictly anticorrelated with nucleosome occupancy. On all promoters mentioned above, PolII initiation sites are either exposed upon activation or located in constitutive NDRs. Consistently in the human genome, nucleosome occupancy immediately upstream of TSSs is decreased in a PolII-dependent manner [21,22]. The transcription initiation complex is bulky and some of its components (such as TBP) severely bend DNA, which could explain why it is incompatible with the nucleosome structure. Indeed, TBP binding has been confirmed to be strongly inhibited when the TATA box is buried in a nucleosome [50,64].

Since correlation does not always imply causality, one can ask if nucleosome removal is a pre-requisite for PolII binding, or its consequence. On a *PHO5pr* with a mutated TATA box, PolII cannot load and activate transcription, but nucleosome loss occurs normally [65]. If nucleosome removal is blocked (for instance, by *Asf1* deletion), *PHO5pr* cannot be activated [40]. When an NDR over *RNR3pr* is created artificially by the insertion of sequences unfavorable for nucleosome formation (see below), PolII is able to bind and drive transcription in the absence of normally required activating signals [66]. Artificially recruited TBP cannot remodel nucleosomes near the TATA box [67]. In addition, if we look at the sequence of events on *HOpr*, the nucleosome disruption starts shortly after being released from a metaphase block, clearly preceding the PolII binding [38]. All this evidence indicates that at least the initial PolII binding is dictated by nucleosome positioning, not the other way around.

However, once PolII is bound and initiated, it could have a role in stabilizing the NDR as well as defining the NDR boundary. Supporting this view, the inactivation of *Rpb1* (PolII large subunit) in yeast results in an increase in nucleosome density upstream of the TSS and a somewhat surprising downstream shift of +1 nucleosome [30]. In *Drosophila* cells, eliminating PolII paused at the promoter–proximal regions also increases local nucleosome occupancy [68].

The narrow distribution of TSSs relative to the +1 nucleosome and systematic shift of this distance in different species (the upstream edge of nucleosome +1 relative to the TSS is –10 to –15 bp in budding yeast and around +60 bp in humans and flies) indicate a co-evolution of the TSS and the +1 nucleosome positioning [69]. Again, the causality between them could go both ways: either the nucleosome positioning limits the selection of the TSS by restraining the assembly site of the transcriptional machinery or the initiating transcription machinery interacts with and actively repositions the +1 nucleosome [70]. Overall, the mechanism of TSS selection is unclear, and the interplay between nucleosome positioning and the TSS provides an interesting avenue of research. In addition to transcription initiation, NDRs have been found on the 3′ end of many genes, and might have some function in transcription termination [8].

Nucleosome effect on transcription elongation and promoter–proximal pausing

The elongating polymerase has to go through a nucleosome barrier. *In vitro* under physiological salt concentrations, nucleosomes represent a formidable block for elongation [71]. On highly transcribed genes, such as the *Drosophila Hsp70* gene after heat shock, nucleosomes in the entire transcribed region are disrupted in a transcription-independent manner, accommodating a high density of elongating polymerase [72]. However, most of the genes, even those with reasonably high transcription levels, are assembled into nucleosome arrays. A large group of factors is involved in facilitating PolII passage through the nucleosomes as well as maintaining the chromatin structure in this process [73].

Nucleosomes might also play an active role in regulating elongation. A significant fraction of genes in *Drosophila* and human cells have initiated but stalled PolII within their promoter–proximal regions [74,75]. Interestingly, PolII often stalls in both sense and antisense directions at around +50 bp (sense) and −250 bp (antisense) relative to the TSS. These two locations are close to the edge of the −1 and +1 nucleosomes, respectively. It is therefore tempting to speculate that nucleosomes play a role in the pausing of PolII.

The mechanism of nucleosome positioning

Because nucleosome positioning plays a crucial role in transcriptional regulation, it is important to understand how this positioning is established. Here, we briefly summarize some recent key developments in this rapidly developing field.

Nucleosome positioning and phasing based on intrinsic histone–DNA affinity

DNA is severely bent in the nucleosome structure [1], and DNA flexibility strongly affects intrinsic histone–DNA

affinity [76]. For instance, GC-rich sequences are believed to facilitate nucleosome formation by increasing DNA flexibility [77–79], whereas relatively rigid poly-AT sequences disfavor nucleosome assembly [8,11,12]. In addition, because DNA bends differently in different directions, AA/TT/TA dinucleotides occur preferentially where the minor groove faces the histone octamer, whereas GC/CC/GG dinucleotides tend to occur where the minor groove points away [80]. *In vitro*, there is clear evidence that DNA sequences can position nucleosomes both translationally and rotationally (translational positioning refers to the 147 bp sequence covered by a histone octamer and rotational positioning refers to the 10–11 bp periodic orientation of the DNA helix in the histone–DNA complex), and it is appealing to think that nucleosome positions *in vivo* are also largely controlled by the underlying DNA sequence.

Recently, considerable effort has been devoted to testing this hypothesis. Of special interest are genome-wide comparisons between *in vivo* and *in vitro* nucleosome maps. In the latter case, nucleosomes are assembled on genomic DNA using salt dialysis, so that nucleosome positions are affected solely by intrinsic histone–DNA interactions and steric exclusion [12,70]. Nucleosomes assembled *in vitro* were found to be depleted around the TSS (Figure 1a) and termination site, probably because the corresponding sequences are more poly-AT rich (although observed nucleosomes could be biased towards such sequences by MNase digestion, see above; unpublished observation). In addition, *in vivo* and *in vitro* nucleosomal DNA showed similar 10–11 bp periodicities of dinucleotide distributions, although the amplitude of the periodic change was more prominent in the latter case [12]. These observations lead to the proposal that nucleosome positions are controlled primarily by DNA sequence in living cells (the nucleosome code hypothesis) [11,12,81].

However, the importance of intrinsic histone–DNA interactions for *in vivo* nucleosome positioning remains controversial because there are substantial differences between *in vitro* and *in vivo* nucleosome maps in the vicinity of coding regions: NDRs are less pronounced in the former and the nucleosomes downstream of the NDRs are not phased (Figure 1a) [70]. Therefore, it has been argued that DNA sequence is not a major determinant of *in vivo* nucleosome positioning [70]. Because these differences are observed for *in vitro* chromatin assembled at the 1:1 histone-to-DNA mass ratio, which corresponds to *in vivo* levels of nucleosome occupancy, they cannot be attributed simply to the difference in total histone concentration. Rather, it has been proposed that the +1 nucleosome is localized *in vivo* through yet unknown interactions with the components of transcription initiation machinery, dictating downstream nucleosome positioning by steric exclusion (which might be aided by regularly spaced nucleosome positioning sequences).

Factors of multiple families can affect nucleosome positioning

Non-histone factors must play a role in nucleosome distribution because nucleosome positioning *in vivo* can be only partially predicted from DNA sequence and, moreover, is not static but dynamically regulated. Some DNA-binding factors in yeast, such as Abf1, Reb1 and Rap1, have been proposed to have "chromatin reorganizing" activity, either by directly competing with histones for DNA binding or by recruiting chromatin remodelers to displace neighboring nucleosomes [32]. Consistently, the binding sites of Abf1 and Reb1 have large discrepancies between their *in vivo* and *in vitro* nucleosome occupancies [12]. An essential remodeler RSC can directly bind to specific sequences of DNA, and the deletion of its subunits with DNA-binding or ATPase activity affects a significant fraction of NDRs [32,82]. Moreover, once NDRs are formed, they can serve as "barriers" that constrain the locations of nearby nucleosomes by the so-called "statistical positioning" mechanism [8,83].

It should be noted that some NDRs observed *in vivo* do not contain poly-AT or the binding sites of the factors mentioned above (unpublished observation), and not all NDRs are affected by the deletion of these factors [32,82]. In addition, the NDRs responding to the factor deletion tend to shrink instead of completely disappear [32]. Therefore, the NDR formation mechanism requires further elucidation.

Box 2. Outstanding questions

- Why do some nucleosome-depleted promoters remain transcriptionally repressed?
- What determines the range and directionality of nucleosome remodeling?
- What is the relationship between NDRs and divergent transcription?
- How is nucleosome positioning established and maintained *in vivo*?
- Is there cell-to-cell variation in nucleosome positioning and how does it affect the variation in transcriptional activity?
- How is nucleosome positioning inherited through DNA replication?
- How does nucleosome positioning evolve among species?

Concluding remarks and future perspectives

Nucleosome positioning is an important chromatin feature that regulates gene expression. In particular, nucleosome depletion in the promoter and the precise positioning of the downstream nucleosomes play crucial roles in determining the transcription level, cell-to-cell variation, activation and repression dynamics, and might also function in defining the start and end points of transcribed regions. Nucleosomes affect transcription mostly by modulating the accessibility of regulatory factors and the transcriptional machinery to the underlying DNA sequence.

There are still many remaining questions in this field (Box 2). However, there are recent publications that provide significant insights into some of these questions and more are to follow [49,84–86].

Acknowledgements

The authors acknowledge Dr Fred Cross and Dr Eric Siggia for advice on the manuscript, and thank Dr Denis Tolkunov for help with the figures. L.B is supported by a Damon Runyon Cancer Research Fellowship, and A.V.M by an NIH grant HG004708 and an Alfred P. Sloan Research Fellowship.

References

1 Luger, K. *et al.* (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389, 251–260

2 Tolkunov, D. and Morozov, A.V. (2010) Genomic studies and computational predictions of nucleosome positions and formation energies. *Adv. Protein Chem. Struct. Biol.* 79, 1–57

3 Yuan, G.C. *et al.* (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* 309, 626–630
4 Lee, W. *et al.* (2007) A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* 39, 1235–1244
5 Barski, A. *et al.* (2007) High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–837
6 Shivaswamy, S. *et al.* (2008) Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* 6, e65
7 Mavrich, T.N. *et al.* (2008) Nucleosome organization in the *Drosophila* genome. *Nature* 453, 358–362
8 Mavrich, T.N. *et al.* (2008) A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* 18, 1073–1083
9 Schones, D.E. *et al.* (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132, 887–898
10 Valouev, A. *et al.* (2008) A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* 18, 1051–1063
11 Field, Y. *et al.* (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* 4, e1000216
12 Kaplan, N. *et al.* (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458, 362–366
13 Westenberger, S.J. *et al.* (2009) Genome-wide nucleosome mapping of Plasmodium falciparum reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. *BMC Genomics* 10, 610–626
14 Lantermann, A.B. (2010) *Schizosaccharomyces pombe* genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*. *Nat. Struct. Mol. Biol.* 17, 251–257
15 Albert, I. *et al.* (2007) Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* 446, 572–576
16 Jin, C. *et al.* (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* 41, 941–945
17 Dion, M.F. *et al.* (2007) Dynamics of replication-independent histone turnover in budding yeast. *Science* 315, 1405–1408
18 Mito, Y. *et al.* (2007) Histone replacement marks the boundaries of *cis*-regulatory domains. *Science* 315, 1408–1411
19 Svaren, J. and Horz, W. (1997) Transcription factors vs. nucleosomes: regulation of the PHO5 promoter in yeast. *Trends Biochem. Sci.* 22, 93–97
20 Bai, L. *et al.* (2010) Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle. *Dev. Cell* 18, 544–555
21 Ozsolak, F. *et al.* (2007) High-throughput mapping of the chromatin structure of human promoters. *Nat. Biotechnol.* 25, 244–248
22 Boyle, A.P. *et al.* (2008) High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132, 311–322
23 Tillo, D. *et al.* (2010) High nucleosome occupancy is encoded at human regulatory sequences. *PLoS One* 5, e9129
24 Woo, C.J. *et al.* (2010) A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* 140, 99–110
25 Lin, J.C. *et al.* (2007) Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. *Cancer Cell* 12, 432–444
26 Wingert, L. and Von Hippel, P.H. (1968) The conformation dependent hydrolysis of DNA by micrococcal nuclease. *Biochim. Biophys. Acta* 157, 114–126
27 Horz, W. and Altenburger, W. (1981) Sequence specific cleavage of DNA by micrococcal nuclease. *Nucleic Acids Res.* 9, 2643–2658
28 Dingwall, C. *et al.* (1981) High sequence specificity of micrococcal nuclease. *Nucleic Acids Res.* 9, 2659–2673
29 Jin, C. *et al.* (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat. Genet.* 41, 941–945
30 Weiner, A. *et al.* (2010) High-resolution nucleosome mapping reveals transcription-dependent promoter packaging. *Genome Res.* 20, 90–100
31 Zawadzki, K.A. *et al.* (2009) Chromatin-dependent transcription factor accessibility rather than nucleosome remodeling predominates during global transcriptional restructuring in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 20, 3503–3513
32 Hartley, P.D. and Madhani, H.D. (2009) Mechanisms that specify promoter nucleosome location and identity. *Cell* 137, 445–458
33 Floer, M. *et al.* (2010) A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell* 141, 407–418
34 Bryant, G.O. *et al.* (2008) Activator control of nucleosome occupancy in activation and repression of transcription. *PLoS Biol.* 6, 2928–2939
35 Barbaric, S. *et al.* (1992) Activation of the weakly regulated PHO8 promoter in *S. cerevisiae*: chromatin transition and binding sites for the positive regulatory protein PHO4. *Nucleic Acids Res.* 20, 1031–1038
36 Moreira, J.M. and Holmberg, S. (1998) Nucleosome structure of the yeast CHA1 promoter: analysis of activation-dependent chromatin remodeling of an RNA-polymerase-II-transcribed gene in TBP and RNA pol II mutants defective *in vivo* in response to acidic activators. *EMBO J.* 17, 6028–6038
37 Hogan, G.J. *et al.* (2006) Cell cycle-specified fluctuation of nucleosome occupancy at gene promoters. *PLoS Genet.* 2, e158
38 Takahata, S. *et al.* (2009) FACT and Asf1 regulate nucleosome dynamics and coactivator binding at the HO promoter. *Mol. Cell* 34, 405–415
39 Naar, A.M. *et al.* (2001) Transcriptional coactivator complexes. *Annu. Rev. Biochem.* 70, 475–501
40 Adkins, M.W. *et al.* (2004) Chromatin disassembly mediated by the histone chaperone Asf1 is essential for transcriptional activation of the yeast PHO5 and PHO8 genes. *Mol. Cell* 14, 657–666
41 Korber, P. *et al.* (2006) The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. *J. Biol. Chem.* 281, 5539–5545
42 Cosma, M.P. *et al.* (1999) Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* 97, 299–311
43 Korzus, E. *et al.* (1998) Transcription factor-specific requirements for coactivators and their acetyltransferase functions. *Science* 279, 703–707
44 Krebs, J.E. *et al.* (1999) Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. *Genes Dev.* 13, 1412–1421
45 Sternberg, P.W. *et al.* (1987) Activation of the yeast HO gene by release from multiple negative controls. *Cell* 48, 567–577
46 Dhasarathy, A. and Kladde, M.P. (2005) Promoter occupancy is a major determinant of chromatin remodeling enzyme requirements. *Mol. Cell Biol.* 25, 2698–2707
47 Koutroubas, G. *et al.* (2008) Bypassing the requirements for epigenetic modifications in gene transcription by increasing enhancer strength. *Mol. Cell Biol.* 28, 926–938
48 Burns, L.G. and Peterson, C.L. (1997) The yeast SWI-SNF complex facilitates binding of a transcriptional activator to nucleosomal sites *in vivo*. *Mol. Cell Biol.* 17, 4811–4819
49 Adkins, M.W. and Tyler, J.K. (2006) Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Mol. Cell* 21, 405–416
50 Ohsawa, R. *et al.* (2009) Epigenetic inheritance of an inducibly nucleosome-depleted promoter and its associated transcriptional state in the apparent absence of transcriptional activators. *Epigenetics Chromatin* 2, 11
51 Basehoar, A.D. *et al.* (2004) Identification and distinct regulation of yeast TATA box-containing genes. *Cell* 116, 699–709
52 Imbalzano, A.N. *et al.* (1994) Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature* 370, 481–485
53 Tirosh, I. and Barkai, N. (2008) Two strategies for gene regulation by promoter nucleosomes. *Genome Res.* 18, 1084–1091
54 Angermayr, M. *et al.* (2002) Transcription initiation *in vivo* without classical transactivators: DNA kinks flanking the core promoter of the housekeeping yeast adenylate kinase gene, AKY2, position nucleosomes and constitutively activate transcription. *Nucleic Acids Res.* 30, 4199–4207
55 Angermayr, M. *et al.* (2003) A nucleosome-free dG-dC-rich sequence element promotes constitutive transcription of the essential yeast RIO1 gene. *Biol. Chem.* 384, 1287–1292
56 Liu, X. *et al.* (2006) Whole-genome comparison of Leu3 binding *in vitro* and *in vivo* reveals the importance of nucleosome occupancy in target site selection. *Genome Res.* 16, 1517–1528
57 Angelov, D. *et al.* (2004) The histone octamer is invisible when NF-kappaB binds to the nucleosome. *J. Biol. Chem.* 279, 42374–42382

58 Owen-Hughes, T. and Workman, J.L. (1994) Experimental analysis of chromatin function in transcription control. *Crit. Rev. Eukaryot. Gene Expr.* 4, 403–441
59 Li, G. *et al.* (2005) Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* 12, 46–53
60 Polach, K.J. and Widom, J. (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* 254, 130–149
61 Morse, R.H. (1993) Nucleosome disruption by transcription factor binding in yeast. *Science* 262, 1563–1566
62 Lam, F.H. *et al.* (2008) Chromatin decouples promoter threshold from dynamic range. *Nature* 453, 246–250
63 Choi, J.K. and Kim, Y.J. (2009) Intrinsic variability of gene expression encoded in nucleosome positioning sequences. *Nat. Genet.* 41, 498–503
64 Godde, J.S. *et al.* (1995) The amino-terminal tails of the core histones and the translational position of the TATA box determine TBP/TFIIA association with nucleosomal DNA. *Nucleic Acids Res.* 23, 4557–4564
65 Barbaric, S. *et al.* (2007) Redundancy of chromatin remodeling pathways for the induction of the yeast PHO5 promoter *in vivo*. *J. Biol. Chem.* 282, 27610–27621
66 Zhang, H. and Reese, J.C. (2007) Exposing the core promoter is sufficient to activate transcription and alter coactivator requirement at RNR3. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8833–8838
67 Ryan, M.P. *et al.* (2000) Artificially recruited TATA-binding protein fails to remodel chromatin and does not activate three promoters that require chromatin remodeling. *Mol. Cell Biol.* 20, 5847–5857
68 Gilchrist, D.A. *et al.* (2008) NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev.* 22, 1921–1933
69 Jiang, C. and Pugh, B.F. (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* 10, 161–172
70 Zhang, Y. *et al.* (2009) Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nat. Struct. Mol. Biol.* 16, 847–852
71 Izban, M.G. and Luse, D.S. (1992) Factor-stimulated RNA polymerase II transcribes at physiological elongation rates on naked DNA but very poorly on chromatin templates. *J. Biol. Chem.* 267, 13647–13655
72 Petesch, S.J. and Lis, J.T. (2008) Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* 134, 74–84
73 Li, B. *et al.* (2007) The role of chromatin during transcription. *Cell* 128, 707–719
74 Core, L.J. *et al.* (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848
75 Nechaev, S. *et al.* (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* 327, 335–338
76 Morozov, A.V. *et al.* (2009) Using DNA mechanics to predict *in vitro* nucleosome positions and formation energies. *Nucleic Acids Res.* 37, 4707–4722
77 Peckham, H.E. *et al.* (2007) Nucleosome positioning signals in genomic DNA. *Genome Res.* 17, 1170–1177
78 Chung, H.R. and Vingron, M. (2009) Sequence-dependent nucleosome positioning. *J. Mol. Biol.* 386, 1411–1422
79 Tillo, D. and Hughes, T.R. (2009) G+C content dominates intrinsic nucleosome occupancy. *BMC Bioinformatics* 10, 442
80 Travers, A. *et al.* (2009) Nucleosome positioning–what do we really know? *Mol. Biosyst.* 5, 1582–1592
81 Segal, E. *et al.* (2006) A genomic code for nucleosome positioning. *Nature* 442, 772–778
82 Badis, G. *et al.* (2008) A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol. Cell* 32, 878–887
83 Fedor, M.J. *et al.* (1988) Statistical positioning of nucleosomes by specific protein-binding to an upstream activating sequence in yeast. *J. Mol. Biol.* 204, 109–127
84 Jessen, W.J. *et al.* (2006) Active PHO5 chromatin encompasses variable numbers of nucleosomes at individual promoters. *Nat. Struct. Mol. Biol.* 13, 256–263
85 Tirosh, I. *et al.* (2010) Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. *Mol. Syst. Biol.* 6, 365
86 Xu, Z. *et al.* (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature* 457, 1033–1037