

New Problems in RNA Polymerase II Transcription Initiation: Matching the Diversity of Core Promoters with a Variety of Promoter Recognition Factors*

Published, JBC Papers in Press, March 29, 2007, DOI 10.1074/jbc.R700012200

Ferenc Müller^{†1}, Máté A. Demény^{§2}, and László Tora^{§3}

From the [†]Institute of Toxicology and Genetics, Forschungszentrum, Karlsruhe, D-76021 Germany and [§]Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR7104, INSERM U596, Université Louis Pasteur, Illkirch F-67400, France

One of the key events in eukaryotic gene regulation is the assembly of general transcription factors (GTFs)⁴ and RNA polymerase II (Pol II) into a pre-initiation complex (PIC) at promoters. GTFs have been defined biochemically as a set of factors essential for accurate transcription initiation at a TATA box-containing viral promoter *in vitro* (1, 2). The textbook description states that the initiation of mRNA synthesis requires the recruitment and binding of the TATA-binding protein (TBP) as a component of the TFIID complex to a consensus sequence that is found in core promoters and is known as the TATA box (3). This view is now seriously challenged by a series of observations. (i) The TATA box is not a general component of all Pol II core promoters; (ii) not only TBP but different types of TBP-related factors can mediate Pol II transcription initiation; (iii) TBP- or TBP-related factor-independent Pol II transcription has been described; and (iv) TBP binding is not necessarily a prerequisite or even an indicator of promoter activation *in vivo*. The collapse of the old “general” dogma is a result of parallel developments in bioinformatic analysis of promoter sequences and the application of new technologies in studying the biology of transcription in metazoans. Here we review the emerging view suggesting that instead of TBP binding to a TATA box, in metazoans distinct sets of core promoter recognition factors recognize a huge diversity of core promoters (see Fig. 1).

* This minireview will be reprinted in the 2007 Minireview Compendium, which will be available in January, 2008. This work was supported by grants from INSERM, CNRS, ULP, the Fond National de la Science ACI, the Agence Nationale de la Recherche, EU Grants LSHG-CT-2004-502950 and HPRN-CT 00504228 (to L. T.), and by DFG Grant MU 1768/2 and EU Contract 511990 (to F. M.).

¹ To whom correspondence may be addressed. E-mail: Ferenc.Mueller@itg.fzk.de.

² Present address: Dept. of Neuroscience, Institute of Psychiatry, King's College, London, UK.

³ To whom correspondence may be addressed. E-mail: laszlo@igbmc.u-strasbg.fr.

⁴ The abbreviations used are: GTF, general transcription factor; Pol II, RNA polymerase II; PIC, pre-initiation complex; TBP, TATA-binding protein; TSS, transcription start site; TAF, TBP-associated factor; TRF, TBP-related factor; BP, broad peak TSS-containing promoter.

This is an Open Access article under the [CC BY](http://creativecommons.org/licenses/by/4.0/) license.

There Is No Universal Core Promoter Sequence Feature

The core promoter was originally defined biochemically as the minimal DNA fragment that is sufficient to direct correct basal levels of transcription initiation by Pol II *in vitro* on naked DNA templates containing a single well defined transcription start site (TSS). According to this definition, the core promoter typically extends appreciatively 40–50 bp up- and downstream of the TSS and can contain several distinct core promoter sequence elements. An important problem in the understanding of core promoter function at the genomic level has been the difficulty of identifying core promoters in general. This is because they lack easily identifiable universal consensus motifs. Promoter prediction tools are numerous, but because of the lack of general identification features, high false positive and negative rates dampened their efficiency (4). In the past, only a few genes were analyzed sufficiently in depth to identify general core promoter features and TSSs. To get a better grip on the regulation of core promoters in general, a drastic improvement in identifying functional promoter sequences has become necessary.

Promoter identification has improved drastically in the last several years because of the technological development in two areas of genomics: large scale sequencing of accurately determined 5'-ends of cDNAs (such as Cap analysis of gene expression, CAGE (5)) and the availability of numerous genome sequences from yeast to human, which together allowed significant improvement in identifying and annotating TSSs on a large scale (6–8). The CAGE tag analysis of the mammalian promoterome provided the first truly global insight into the sequence structure of core promoters and the expression dynamics associated with core promoters. Strikingly, several classes of promoter types were identified on the basis of the differential usage of TSSs (8). At the extreme ends of the spectrum, two significant classes of promoters can be defined: (a) single peak promoters, which have a relatively tightly defined TSS position within a few base pairs and (b) broad peak promoters, which show a random distribution of many TSSs in a 100-bp window. Strikingly, single peak promoters were found to be more likely associated with TATA boxes and to possess a higher frequency of transcription factor binding sites than broad peak promoters. The latter often carry CpG islands, are TATA-less, and are associated with less consensus binding sites. Moreover, it seems that TATA box-associated single peak promoters are involved in tight regulation of genes, which are often tissue-specific. In contrast, the broad peak promoters are often associated with ubiquitously expressed genes (8, 9). In addition, from these studies it became clear that the TATA box, which was once believed to be a general feature of core promoters, is only present in less than 10% of all human Pol II promoters (10, 11). The existence of TATA-less promoters regulating tissue-specific genes (12) and TATA-containing promoters driving transcription of ubiquitously expressed genes (13) demonstrates that the regulatory logic suggested by the above genome-wide tendencies cannot be universally applied to gene activation. Moreover, even when a TATA box is present in a core promoter, it may not be critical for defining promoter activity, and throughout the genome a consensus TATA element does not

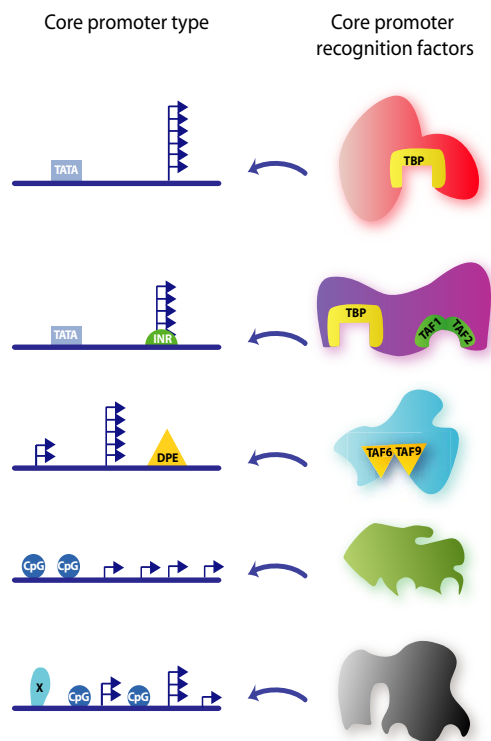


FIGURE 1. A variety of hypothetical core promoters and a schematic representation of protein complexes that may bind to them. Sequence specificity of the distinct promoter recognition factors (on the right) results in a variety of choices of transcription initiation sites (blue arrowheads) on the different core promoters (on the left). The number of blue arrowheads represents the strength of the given initiation site on the promoter. The different consensus sequence elements (TATA box, INR, DPE, CpG island, XCPE1 (X)) provide binding sites for subunits of promoter recognition factors (8, 16, 19, 32, 37). None of these elements are present in all promoters, and the interactions described may not fully represent all the possible protein-DNA interactions on the specified elements or may only be hypothetical.

appear to be a major determinant of either TBP binding or gene expression in yeast (14, 15).

Recent bioinformatic studies of vertebrates failed to identify a single core promoter element that would universally cluster close to the TSS for all, or at least the majority, of the analyzed core promoters. Nevertheless, a series of core promoter elements has been shown to exist in smaller subsets of core promoters with positional constraints in relation to the TSS and with specific transcription factors binding to them. These include the INR (initiator), BRE (TFIIB recognition element), DCE (downstream core element), DPE (downstream core promoter element), and MTE (mutation ten element) (Fig. 1) (see Refs. 3, 16, and 17 for reviews). Interestingly, the INR with its degenerate sequence YYANWYY (18) does not seem to cluster at the TSSs (10). These elements can positively or negatively correlate with the presence of other consensus core promoter sequences; however the regulatory significance of these correlations is still unclear. For example the DPE-containing promoters are often, but not exclusively, TATA-less (3), and the newly described XCPE1 element (the X-gene core promoter element 1) has also been associated with TATA-less promoters (19).

The above described examples in which core promoters are considered as a combination of distinct motifs representing putative transcription factor binding sites may not be the only

way to define the functionally relevant structure of core promoters. Chargaff's second parity rule states that the number of As equals the number of Ts and the number of Cs equals the number of Gs in a single strand over regions of sufficient size (20). This rule seems not to be valid in the immediate neighborhood of TSSs (11, 21). Furthermore, on the basis of mononucleotide distribution around TSSs four possible promoter classes can be defined in which the GC/AT distribution around the TSSs is significantly diverged (11). However, at present it is not clear how these distinct classes of promoters with different GC/AT distribution can be recognized by transcription factors that direct transcription initiation site selection.

Thus, the variations of sequence motif combinations and other yet unexplored sequence features in core promoters and the corresponding binding of diverse GTFs may function together in the determination of the transcriptional initiation site.

Matching the Core Promoter Elements with a Large Diversity of Promoter Recognition Factors

The regulatory significance of the above described diversity of core promoters is not yet resolved. In the simplest scenario, distinct core promoter elements would create binding sites for different recognition factors (Fig. 1) that would allow the assembly of PICs with distinct factor composition. Thus, not all the GTFs will be present in every PIC, suggesting that several of the GTFs may not be appropriately called "general" (22). This suggestion is in good agreement with the observation that certain GTFs are not expressed in every cell type of a given organism (Refs. 22–25 and references therein). In the next paragraphs we will summarize findings about the known diversity of core promoter binding factors in support of the novel view that many structurally and functionally different PICs will assemble on a wide diversity of Pol II core promoters to direct the initiation of transcription.

The TFIID complex, composed always of TBP and all the 13 or 14 TBP-associated factors (TAFs) (Ref. 26 and references therein), has long been viewed as the only core promoter recognition factor in Pol II transcription. However, TFIID complexes containing or lacking certain TAFs exist, generating a spectrum of variants of the prototype complex (reviewed in Refs. 22 and 27). How do these different TFIID complexes form? Several experiments indicated that a subset of TAFs (TAF4, TAF5, TAF6, TAF8, TAF9, TAF10, and TAF12) forms a minimal TFIID core complex that lacks TBP and other TAFs (28–30). This stable TFIID core is suggested to associate with combinations of other TAFs and TBP (28, 30) to form a large variety of different complexes.

In agreement with a model according to which different submodules of TFIID complexes recognize distinct core promoter elements, several TAFs have been implicated in the recognition of the promoter DNA around the TSS (31). TAF1 and TAF2 can bind the INR (32). TFIID complexes with or without TAF2 have been described (33, 34), suggesting that only complexes that contain TAF1 and TAF2 will be recruited to the INR-containing class of core promoters. TAF1 was also shown to bind DCE-containing promoters (35). Two domains of TAF1 are important for regulating largely non-overlapping sets of genes *in vivo* (36), suggesting that even different domains of distinct TAFs may bind different core promoter regions. The TAF6/TAF9 and the TAF4b/TAF12 heterodimers have also been shown to bind DNA, and the TAF6/

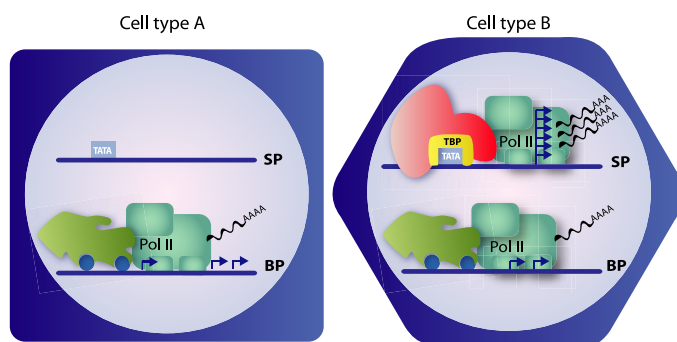


FIGURE 2. Models for core promoter-dependent regulation of transcription initiation. Arrowheads indicate transcriptional start sites. Depending on promoter motif composition single peak (SP) and broad peak (BP) promoters bind different sets of promoter recognition factors and may be differentially active in TSS usage and/or overall transcriptional activity depending on the cell type (8). Cell types A and B express different sets of promoter recognition factors. Pol II represents the set of GTFs that are necessary for the correct initiation of transcription on a given SP- or BP-type promoter and in a given cell type. The number of wavy lines represents the degree of transcription initiation from the promoter.

TAF9 pair was also found to contact the DPE-containing class of core promoters *in vivo* (3, 37) (Fig. 1).

The core TFIID model might imply that satellite TAFs are less broadly used than those that form the core. In agreement with the idea that TAF10 is part of the TFIID core, TAF10-lacking early mouse embryos seem to have an altered TFIID structure and have undetectable levels of Pol II transcription (23). In contrast, later in development many genes are transcribed normally without TAF10, suggesting that TAF10-containing TFIID or other TFIID-like complexes are only required for the transcription of a subset of promoters at certain developmental stages or cell types (38) (Fig. 2).

To elucidate the contribution of TAFs to promoter selectivity *in vivo* a genome-wide promoter occupancy analysis has been performed (39). A tight correlation between promoter occupancy by TAF1, probably representing the binding of a TFIID subpopulation, and the activity of the corresponding gene was detected for only 75% of the genes tested. In this category of genes TFIID-containing PICs on promoters may lead to transcription as the classic model suggested. However, the remaining 25% of gene promoters either (i) did not bind TAF1 while being expressed or (ii) were bound by TAF1, but no transcription was detected (39). In the former case it is possible that TAF1 could bind alone without other components of the TFIID or that TAF1-containing TFIID only primed these genes for future activation of transcription (40, 41). The genes that were transcribed without TAF1 binding may represent the activity of a variant TFIID core complex (28) or that of a TBP-free transcription complex (see also below). In agreement with the idea that not every TAF is present at every active promoter, knock-out or silencing of *TAF* genes in different metazoan cells influences the expression of only a small subset of genes (42, 43), further suggesting that TFIID subunits are not generally involved in gene regulation at every promoter.

TBP-independent Transcription Initiation Mechanisms

Although there is a clear link between promoter occupancy by TBP and gene expression in yeast (44), this apparent

correlation is not always true in metazoans. Overexpression of TBP increases transcription from TATA-containing promoters although it rather represses transcription from TATA-less promoters (45, 46), suggesting differential (even opposite) roles for TBP in Pol II transcription initiation at distinct promoters. Similarly, the TATA binding activity of TBP is not required for the functional recruitment of TFIID to certain natural promoters (47). Also, no direct correlation between promoter occupancy by TBP and activation of a promoter by a strong viral enhancer in chicken DT40 cells was detected (48). Thus, TBP may be present at the promoters but may not contribute to promoter recognition and/or transcription initiation events. TBP-free transcription initiation has also been suggested at zygotic developmental regulator genes, which were not affected by loss of TBP in vertebrate embryos (Ref. 22 and references therein), and type I interferon responsive genes can be independently transcribed of TBP in human HeLa cells (49). A recently identified set of TBP-related factors (TRFs) has been suggested to function similarly to TBP in Pol II promoter recognition and may indicate alternative mechanisms for TBP-free transcription initiation. *Drosophila*-specific TRF1 and vertebrate-specific TRF3/TBP2 were shown to bind the TATA box, interact with GTFs, and mediate Pol II transcription initiation from a limited set of promoters (50–53). Thus, the specificity of TBP, TRF1, and TRF3/TBP2 to promoters may be determined by factors with which they form a complex and also by their differential expression patterns in the distinct cell types of a given organism. In contrast, the different DNA binding specificity of TRF2/TLF may explain its affinity to distinct set of promoters (Refs. 22, 46, and 54 and references therein).

Taken together, the recognition of promoters by multiple TBP-related factors or other TAF-containing complexes, lacking TBP, may represent an additional regulatory level acting on core promoters and may explain the diversity of core promoters observed in vertebrates.

A Moving Target, Transcription Initiation from Broad TSS-containing Promoters

According to our present understanding the ultimate purpose of the binding of a PIC to the core promoter is to mark precisely the TSS for Pol II transcription initiation. However, a whole class of promoters, the broad peak TSS-containing promoters (BP) (8), do not fit into this model, and their regulation may be fundamentally different. BPs are much more abundant than single peak promoters (55), yet the mechanisms underlying the initiation of transcription from them remain obscure. On these promoters multiple TSSs are distributed over a region of 100 bp. It is possible that on these usually TATA-less BPs (i) only one PIC assembles at a time and the same PIC “slides” to different sites in the same region for separate initiation events; (ii) distinct consecutively arriving PICs bind to slightly different sites; (iii) several PICs can be formed at the same time at different sites in these regions; or (iv) the distinct TSSs represent single or a low number of TSSs specific to each cell, but they appear as multiple TSSs in a cell population. This latter scenario has been suggested by a recent study, which indicated different TSS utilization at BPs in different cell types (56) (Fig. 2). The fact that on these BPs many initiation sites have been detected suggests that the PICs binding to these regions do not have strong

sequence specificity. Several studies suggested that active promoters are devoid of nucleosomes in a ~150-bp to ~200-bp region upstream of the TSSs (57). Thus, it is conceivable that PICs on the multiple TSS-containing promoters assemble at the nucleosome-free regions “non-specifically,” because the BPs do not seem to contain well defined core promoter motifs. However, the fact that initiation from BPs seems to show a certain tissue specificity (56) suggests that initiation from BPs may not be random involving some sequence motif-specific targeting mechanisms.

Flexibility of Transcription Initiation by Alternative Promoters

In addition to multiple TSSs within a promoter region, an important development in the exploration of gene regulation at the core promoter level is the discovery of the widespread usage of alternative promoters of protein coding genes. Genome-wide promoter analyses in vertebrates suggest that 20–52% of the genes carry alternative promoters where each promoter is clearly separated by a genomic space of several hundred base pairs (7, 39). The prevalence of alternative promoters is evident; however, their biological significance is largely unexplored with the exception of a small number of genes where alternative promoters were shown to receive differential signaling information to generate different protein isoforms (e.g. p53, p63, and p73 family (58)).

Alternative promoter usage may provide a mechanism for tissue and developmental stage-specific activity of genes. For example, TATA box-containing promoters were shown to be utilized after embryonic development, and TATA-less promoters of the same genes were shown to be active during early embryo development (59–61). A recent global analysis of mammalian promoters concluded that alternative promoters mostly play a role at highly regulated developmental genes and that single promoter genes are more likely involved in general cellular processes active in a broad range of tissues (62). Thus, PICs assembled at the alternative promoters may vary in the developmental stage or signaling pathway-specific manner.

The above results also raise the issue of specificity of communication between core promoters and cis-regulatory elements through protein-protein interactions. The apparent specificity in promoter-enhancer interactions was demonstrated elegantly in *Drosophila* by the description of the differential ability of a DPE-containing promoter to interact with a set of enhancers (63) as compared with a TATA-containing promoter. The specificity of promoter-enhancer interactions enables promoters to distinguish between enhancers and the information displayed by them (64). This may be necessary when overlapping gene loci with multiple promoters are targeted by a single enhancer or when several enhancers may target a single promoter through long distance interactions.

The molecular basis for the specificity of core promoter-enhancer interactions is unknown but likely involves different protein-protein interactions directed by variant complexes forming on different core promoters.

The Core Promoter as a Target of Signaling Input

The biological significance of the diversity of the core promoters and the binding of the corresponding complexes to them remains the ultimate question to be answered (Fig. 1).

Analysis of molecular interactions between activator proteins and core promoter context-dependent PICs may argue for a model in which the core promoter is involved in processing information received on the cis-regulatory level of a gene. For example, different TATA box-containing promoters respond differentially to activators, suggesting that they do not bind the same TFIID and/or PIC complexes (65, 66). *c-Fos* and *Elf-1* differentially activate transcription from TATA- and INR-containing core promoters (67, 68), whereas optimal induction by papilloma virus E2 and the artificial Gal-VP16 activator requires both TATA and INR elements (69). As a parallel example, p53 specifically represses the TATA box but not TATA-less INR-containing promoters (70). On the other hand, DPE-containing IRF-1 and TAF7 promoters, but not the DCE-containing cyclin D promoter, require protein kinase CK2 and positive coactivator 4 (PC4) for transcription (71).

Additional cellular signaling events that target TFIID components may provide the means for regulatory inputs to be transduced to the core promoter. An example of such a signaling mechanism is the observation that apoptotic stimuli can lead to alternative splicing of TAF6, resulting in a TFIID-type complex containing TAF6 δ but lacking TAF9 (72). This alternative TFIID complex results in altered transcription programs. The fact that an inducible TAF isoform is capable of selectively altering TFIID composition and consequently gene expression programs underlines the biological significance of TFIID variants regulating gene activities at the core promoter level.

Conclusions and Future Directions

In conclusion, a wealth of data gathered in recent years has established the diversity of core promoters and a variety of protein complexes that may bind to them. This diversity argues for a combined model for transcription regulation in multicellular organisms; differential input signals target both types of cis-regulatory elements such as enhancers as well as the core promoter. The resulting transcriptional output is dependent on the integration of both sets of signals. The input signals targeting either the enhancers or the core promoters may be qualitatively different and represent distinct aspects of cellular stimuli, thus necessitating the plasticity of core promoters and that of the complexes forming on them. Thus, the diversity of core promoters will facilitate additional plasticity in transcriptional outputs. Yet the regulatory logic that underlies the interactions between core promoters and the binding factor complexes remains mostly unexplored. It represents the next challenge in transcription biology, which will, without doubt, bring us many more surprises. This challenge is expected to be addressed by developing novel detection techniques to measure promoter occupancy by transcription factors in real time and detect core promoter activity and TSSs at a single cell level as well as to develop new bioinformatics tools to decipher the code embedded in core promoter sequences.

Acknowledgments—We are grateful to S. Bour for illustrations and to B. Bell, D. Devys, Z. Nagy, and M. E. Torres Padilla for critically reading the manuscript.

REFERENCES

- Orphanides, G., Lagrange, T., and Reinberg, D. (1996) *Genes Dev.* **10**, 2657–2683
- Roeder, R. G. (2003) *Nat. Med.* **9**, 1239–1244
- Butler, J. E., and Kadonaga, J. T. (2002) *Genes Dev.* **16**, 2583–2592
- Bajic, V. B., Tan, S. L., Suzuki, Y., and Sugano, S. (2004) *Nat. Biotechnol.* **22**, 1467–1473
- Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T., Fukuda, S., Sasaki, D., Podhajska, A., Harbers, M., Kawai, J., Carninci, P., and Hayashizaki, Y. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15776–15781
- Cooper, S. J., Trinklein, N. D., Anton, E. D., Nguyen, L., and Myers, R. M. (2006) *Genome Res.* **16**, 1–10
- Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T., and Sugano, S. (2006) *Genome Res.* **16**, 55–65
- Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A., Taylor, M. S., Engstrom, P. G., Frith, M. C., Forrest, A. R., Alkema, W. B., Tan, S. L., Plessy, C., Kodzius, R., Ravasi, T., Kasukawa, T., Fukuda, S., Kanamori-Katayama, M., Kitazume, Y., Kawaji, H., Kai, C., Nakamura, M., Konno, H., Nakano, K., Mottagui-Tabar, S., Arner, P., Chesi, A., Gustincich, S., Persichetti, F., Suzuki, H., Grimmond, S. M., Wells, C. A., Orlando, V., Wahlestedt, C., Liu, E. T., Harbers, M., Kawai, J., Bajic, V. B., Hume, D. A., and Hayashizaki, Y. (2006) *Nat. Genet.* **38**, 626–635
- Schug, J., Schuller, W. P., Kappen, C., Salbaum, J. M., Bucan, M., and Stoekert, C. J., Jr. (2005) *Genome Biol.* **6**, R33
- FitzGerald, P. C., Shlyakhtenko, A., Mir, A. A., and Vinson, C. (2004) *Genome Res.* **14**, 1562–1574
- Bajic, V. B., Tan, S. L., Christoffels, A., Schonbach, C., Lipovich, L., Yang, L., Hofmann, O., Kruger, A., Hide, W., Kai, C., Kawai, J., Hume, D. A., Carninci, P., and Hayashizaki, Y. (2006) *PLoS Genet.* **2**, e54
- Latinkic, B. V., Umbhauer, M., Neal, K. A., Lerchner, W., Smith, J. C., and Cunliffe, V. (1997) *Genes Dev.* **11**, 3265–3276
- Quitschke, W. W., DePonti-Zilli, L., Lin, Z. Y., and Paterson, B. M. (1989) *Mol. Cell. Biol.* **9**, 3218–3230
- Kim, J., and Iyer, V. R. (2004) *Mol. Cell. Biol.* **24**, 8104–8112
- Basehoar, A. D., Zanton, S. J., and Pugh, B. F. (2004) *Cell* **116**, 699–709
- Juven-Gershon, T., Hsu, J. Y., and Kadonaga, J. T. (2006) *Biochem. Soc. Trans.* **34**, 1047–1050
- Gross, P., and Oelgeschlager, T. (2006) *Biochem. Soc. Symp.* **73**, 225–236
- Smale, S. T., and Kadonaga, J. T. (2003) *Annu. Rev. Biochem.* **72**, 449–479
- Tokusumi, Y., Ma, Y., Song, X., Jacobson, R. H., and Takada, S. (2007) *Mol. Cell. Biol.* **27**, 1844–1858
- Chargaff, E. (1951) *Fed. Proc.* **10**, 654–659
- Aerts, S., Thijs, G., Dabrowski, M., Moreau, Y., and De Moor, B. (2004) *BMC Genomics* **5**, 34
- Muller, F., and Tora, L. (2004) *EMBO J.* **23**, 2–8
- Mohan, I. W., Scheer, E., Wendling, O., Metzger, D., and Tora, L. (2003) *Mol. Cell. Biol.* **23**, 4307–4318
- Hiller, M., Chen, X., Pringle, M. J., Suchorolski, M., Sancak, Y., Viswanathan, S., Bolival, B., Lin, T. Y., Marino, S., and Fuller, M. T. (2004) *Development* **131**, 5297–5308
- Xiao, L., Kim, M., and DeJong, J. (2006) *Gene Expr. Patterns* **6**, 409–419
- Tora, L. (2002) *Genes Dev.* **16**, 673–675
- Bell, B., and Tora, L. (1999) *Exp. Cell Res.* **246**, 11–19
- Wright, K. J., Marr, M. T., 2nd, and Tjian, R. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12347–12352
- Leurent, C., Sanders, S. L., Demeny, M. A., Garbett, K. A., Ruhlmann, C., Weil, P. A., Tora, L., and Schultz, P. (2004) *EMBO J.* **23**, 719–727
- Demény, M., Soutoglou, E., Nagy, Z., Scheer, E., Janoshazi, A., Richardot, M., Argentini, M., Kessler, P., and Tora, L. (2007) *PLoS ONE* **2**, e316
- Oelgeschlager, T., Chiang, C. M., and Roeder, R. G. (1996) *Nature* **382**, 735–738
- Chalkley, G. E., and Verrijzer, C. P. (1999) *EMBO J.* **18**, 4835–4845
- Kaufmann, J., Ahrens, K., Koop, R., Smale, S. T., and Muller, R. (1998) *Mol. Cell. Biol.* **18**, 233–239
- Martinez, E., Ge, H., Tao, Y., Yuan, C. X., Palhan, V., and Roeder, R. G. (1998) *Mol. Cell. Biol.* **18**, 6571–6583
- Lee, D. H., Gershenzon, N., Gupta, M., Ioshikhes, I. P., Reinberg, D., and Lewis, B. A. (2005) *Mol. Cell. Biol.* **25**, 9674–9686
- O'Brien, T., and Tjian, R. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 2456–2461
- Shao, H., Revach, M., Moshonov, S., Tzuman, Y., Gazit, K., Albeck, S., Unger, T., and Dikstein, R. (2005) *Mol. Cell. Biol.* **25**, 206–219
- Indra, A. K., Mohan, W. S., 2nd, Frontini, M., Scheer, E., Messaddeq, N., Metzger, D., and Tora, L. (2005) *Dev. Biol.* **285**, 28–37
- Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D., and Ren, B. (2005) *Nature* **436**, 876–880
- Christova, R., and Oelgeschlager, T. (2002) *Nat. Cell Biol.* **4**, 79–82
- Marr, M. T., 2nd, Isogai, Y., Wright, K. J., and Tjian, R. (2006) *Genes Dev.* **20**, 1458–1469
- Chen, Z., and Manley, J. L. (2000) *Mol. Cell. Biol.* **20**, 5064–5076
- Frontini, M., Soutoglou, E., Argentini, M., Bole-Feysot, C., Jost, B., Scheer, E., and Tora, L. (2005) *Mol. Cell. Biol.* **25**, 4638–4649
- Kuras, L., and Struhl, K. (1999) *Nature* **399**, 609–613
- Colgan, J., and Manley, J. L. (1992) *Genes Dev.* **6**, 304–315
- Chong, J. A., Moran, M. M., Teichmann, M., Kaczmarek, J. S., Roeder, R., and Clapham, D. E. (2005) *Mol. Cell. Biol.* **25**, 2632–2643
- Martinez, E., Chiang, C. M., Ge, H., and Roeder, R. G. (1994) *EMBO J.* **13**, 3115–3126
- Chen, Z., and Manley, J. L. (2003) *Mol. Cell. Biol.* **23**, 7350–7362
- Paulson, M., Press, C., Smith, E., Tanese, N., and Levy, D. E. (2002) *Nat. Cell Biol.* **4**, 140–147
- Holmes, M. C., and Tjian, R. (2000) *Science* **288**, 867–870
- Bartfai, R., Balduf, C., Hilton, T., Rathmann, Y., Hadzhiev, Y., Tora, L., Orban, L., and Muller, F. (2004) *Curr. Biol.* **14**, 593–598
- Jallow, Z., Jacobi, U. G., Weeks, D. L., Dawid, I. B., and Veenstra, G. J. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101**, 13525–13530
- Persengiev, S. P., Zhu, X., Dixit, B. L., Maston, G. A., Kittler, E. L., and Green, M. R. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 14887–14891
- Dantone, J. C., Wurtz, J. M., Poch, O., Moras, D., and Tora, L. (1999) *Trends Biochem. Sci.* **24**, 335–339
- Carninci, P. (2006) *Trends Genet.* **22**, 501–510
- Kawaji, H., Frith, M. C., Katayama, S., Sandelin, A., Kai, C., Kawai, J., Carninci, P., and Hayashizaki, Y. (2006) *Genome Biol.* **7**, R118
- Barrera, L. O., and Ren, B. (2006) *Curr. Opin. Cell Biol.* **18**, 291–298
- Muller, M., Schleithoff, E. S., Stremmel, W., Melino, G., Krammer, P. H., and Schilling, T. (2006) *Drug Resist. Updat.* **9**, 288–306
- Davis, W., Jr., and Schultz, R. M. (2000) *Dev. Biol.* **218**, 275–283
- Duan, Z. J., Fang, X., Rohde, A., Han, H., Stamatoyannopoulos, G., and Li, Q. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5509–5514
- Majumder, S., and DePamphilis, M. L. (1994) *Mol. Cell. Biol.* **14**, 4258–4268
- Baek, D., Davis, C., Ewing, B., Gordon, D., and Green, P. (2007) *Genome Res.* **17**, 145–155
- Butler, J. E., and Kadonaga, J. T. (2001) *Genes Dev.* **15**, 2515–2519
- Kulkarni, M. M., and Arnosti, D. N. (2003) *Development* **130**, 6569–6575
- Simon, M. C., Fisch, T. M., Benecke, B. J., Nevins, J. R., and Heintz, N. (1988) *Cell* **52**, 723–729
- Wefald, F. C., Devlin, B. H., and Williams, R. S. (1990) *Nature* **344**, 260–262
- Metz, R., Kouzarides, T., and Bravo, R. (1994) *EMBO J.* **13**, 3832–3842
- Ernst, P., Hahm, K., Trinh, L., Davis, J. N., Roussel, M. F., Turck, C. W., and Smale, S. T. (1996) *Mol. Cell. Biol.* **16**, 6121–6131
- Ham, J., Steger, G., and Yaniv, M. (1994) *EMBO J.* **13**, 147–157
- Mack, D. H., Vartikar, J., Pipas, J. M., and Laimins, L. A. (1993) *Nature* **363**, 281–283
- Lewis, B. A., Sims, R. J., 3rd, Lane, W. S., and Reinberg, D. (2005) *Mol. Cell Biol.* **18**, 471–481
- Bell, B., Scheer, E., and Tora, L. (2001) *Mol. Cell* **8**, 591–600