270–283 *Nucleic Acids Research, 2012, Vol. 40, No. 1* *Published online 12 September 2011*
doi:10.1093/nar/gkr737

Genomic binding of Pol III transcription machinery and relationship with TFIIS transcription factor distribution in mouse embryonic stem cells

Lucie Carrière[1], Sébastien Graziani[1], Olivier Alibert[2], Yad Ghavi-Helm[1], Fayçal Boussouar[1], Hélène Humbertclaude[1], Sylvie Jounier[1], Jean-Christophe Aude[1], Céline Keime[3], Janos Murvai[1], Mario Foglio[4], Marta Gut[4], Ivo Gut[4], Mark Lathrop[4], Julie Soutourina[1], Matthieu Gérard[1,*] and Michel Werner[1,*]

[1]Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), iBiTec-S, F-91191 Gif-sur-Yvette cedex, [2]CEA, iRCM, F-91057 Evry cedex, [3]Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS, INSERM, Université de Strasbourg, F-67404, Illkirch cedex and [4]CEA, iG, F-91057 Evry cedex, France

Received June 23, 2011; Revised August 22, 2011; Accepted August 24, 2011

ABSTRACT

RNA polymerase (Pol) III synthesizes the tRNAs, the 5S ribosomal RNA and a small number of untranslated RNAs. *In vitro*, it also transcribes short interspersed nuclear elements (SINEs). We investigated the distribution of Pol III and its associated transcription factors on the genome of mouse embryonic stem cells using a highly specific tandem ChIP-Seq method. Only a subset of the annotated class III genes was bound and thus transcribed. A few hundred SINEs were associated with the Pol III transcription machinery. We observed that Pol III and its transcription factors were present at 30 unannotated sites on the mouse genome, only one of which was conserved in human. An RNA was associated with >80% of these regions. More than 2200 regions bound by TFIIIC transcription factor were devoid of Pol III. These sites were associated with cohesins and often located close to CTCF-binding sites, suggesting that TFIIIC might cooperate with these factors to organize the chromatin. We also investigated the genome-wide distribution of the ubiquitous TFIIS variant, TCEA1. We found that, as in *Saccharomyces cerevisiae*, TFIIS is associated with class III genes and also with SINEs suggesting that TFIIS is a Pol III transcription factor in mammals.

INTRODUCTION

In eukaryotes, three nuclear RNA polymerases (Pol) are responsible for the transcription of the genome. Pol I transcribes a single RNA species, the precursor of the large ribosomal RNAs (rRNA) and of the 5.8S rRNA. Pol II transcribes all messenger RNAs and many non-coding RNAs implicated in various processes ranging from splicing, RNAs modification (snoRNAs) or gene regulation (miRNAs). Pol III transcribes the 5S rRNA, the tRNAs and a small number of stable non-coding transcripts (1,2). The U6 snRNA (mRNA splicing), RNAse P RNA (tRNA maturation) and 7SL RNA (signal recognition particle) are produced by Pol III in all eukaryotes examined so far, whereas, the RNase MRP RNA (mitochondrial rRNA maturation) is transcribed by Pol II in *Saccharomyces cerevisiae* and by Pol III in animals. Other short Pol III products of rather poorly defined functions, such as the vault particle, the Y, 4.5S and BC1 RNAs are largely specific to mammals (2). Finally, the short interspersed repeated elements (SINEs) are retrotransposons originating from class III (i.e. Pol III-transcribed) genes

*To whom correspondence should be addressed. Tel: +33 16908 9342; Fax: +33 16908 4712; Email: michel.werner@cea.fr
Correspondence may also be addressed to Matthieu Gérard. Tel: +33 16908 9429; Fax: +33 16908 4712; Email: matthieu.gerard@cea.fr
Present addresses:
Sébastien Graziani, DGA Maîtrise NRBC, Département Evaluation des effets des agents chimiques, ANC/TOGA, 3-5 rue Lavoisier, F-91710 Vert-le-Petit, France.
Yad Ghavi-Helm, Genome Biology Unit, European Molecular Biology Laboratory, D-69117 Heidelberg, Germany.
Fayçal Boussouar, Institut Albert Bonniot, Domaine de la Merci, F-38706 La Tronche cedex, France.
Marta Gut and Ivo Gut, Centre Nacional d'Anàlisis Genòmica, Parc Científic de Barcelona, Torre I, Baldiri Reixac 4, E-08028 Barcelona, Spain.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

present in hundreds of thousands of copies in mammalian genomes (3). SINEs can be transcribed by the Pol III transcription machinery *in vitro* (4).

The promoters of class III genes are divided into three categories depending on their organization and transcription factor dependence (reviewed in refs 5,6; unless stated otherwise, the nomenclature used for the class III machinery subunits is that of the mammalian transcription system according to ref. 5). Type II promoters of tRNA genes harbor intragenic A and B boxes that are recognized by TFIIIC, a six subunits factor (7 and references therein). Once associated with DNA, TFIIIC positions a second factor, called TFIIIB in yeast or TFIIIB-β in mammals (8), upstream of the transcription start site (TSS). Yeast TFIIIB and mammalian TFIIIB-β consist of the TATA box-binding protein (TBP), BDP1 and BRF1, which are related to TFIIB Pol II general transcription factor. Yeast TFIIIB and mammalian TFIIIB-β recruit Pol III through direct protein–protein interactions (9,10). The type I promoter, unique to the 5S rRNA, is located within the transcribed region and requires TFIIIA that acts as an adapter for the binding of TFIIIC. The mammalian U6 promoter, a classical type III promoter, is located upstream of the transcribed region and consists of a proximal sequence element (PSE) recognized by a four-subunit factor variously called the PSE-binding protein (PBP), the PSE transcription factor (PTF), or the snRNA activating protein complex (SNAPc) (11–13), and a TATA box which is bound by TFIIIB-α, in which the BRF1 subunit of TFIIIB-β is replaced by BRF2 (8,14,15).

Since Pol III transcribes several rRNAs and the tRNAs, it plays a central role in determining the translational capacity of the cell (16). The regulation of Pol III transcription plays a critical role in cell proliferation and cancer (17). Indeed, artificially increasing tRNA and 5S rRNA transcription causes increased cell proliferation and oncogenic transformation (18).

We discovered that TFIIS, a Pol II transcription elongation (19) and initiation factor in *S. cerevisiae* (20,21), also functions as a Pol III general transcription factor (22). Indeed, TFIIS binding could be detected on all class III genes and mutations that affected specifically Pol II or Pol III transcription were identified. Biochemical study of TFIIS role in Pol III transcription in yeast indicated that it stimulates faithful transcription initiation *in vitro*. In mouse, three isoforms of TFIIS, encoded by TCEA1, −2 and −3 exist (23–25). TCEA1 is expressed ubiquitously, contrary to TCEA2 and −3, which are expressed in spermatocytes or in the liver and kidney, respectively. Whether or not they are implicated in Pol III transcription is presently not known.

RNAs synthesized by Pol III often originate from repeated genes, be it the repeated 5S gene or multiple copies of some tRNA genes, raising the question whether all copies of a gene are transcribed at a given time. In addition, class III transcripts are difficult to predict bioinformatically. The nature of the class III transcriptome was first investigated in the yeast *S. cerevisiae* by analyzing the genome-wide distribution of the Pol III transcription machinery (26–28). The Pol III transcription machinery was associated with nearly all tRNA genes irrespective of their genomic localization, suggesting that they are transcribed. Only one new class III transcript, snR52 snoRNA, was identified (26–28). The ChIP-Seq method was very recently applied to the analysis of the Pol III transcription machinery in various human cell lines allowing a complete description of the class III transcriptome (29–33). Unlike the situation in *S. cerevisiae* where all tRNA genes are transcribed, only a subset of the tRNA genes is bound by the Pol III machinery in human. The identity of the bound tRNA genes varied from one cell type to another. In addition, these studies allowed the identification of a few dozen new loci bound by Pol III.

In *S. cerevisiae*, TFIIIC was present, independently of Pol III, on a small number of genomic locations, called *ETC* loci for extra TFIIIC (26). These *ETC* loci are conserved among yeast species suggesting that they might have a functional role in chromatin organization. Repressing artificially the expression of histones *in vivo* leads to decreased nucleosome abundance. In such a situation, the expression of the *ETC* loci is induced (34). In *Schizosaccharomyces pombe*, the presence of TFIIIC independently of Pol III and TFIIIB marks several boundaries between euchromatin and heterochromatin domains (35). In *S. pombe*, TFIIIC plays an active role in delimiting the boundaries since *cis*-acting mutations that abolish its binding lead to the spreading of heterochromatic marks in regions that are usually euchromatic and as a consequence, lead to transcriptional silencing. Interestingly, a large number of ETCs was also found in human cell lines (1865 in K562 cells and 307 in HeLa cells; (30,31). Moreover, the ETCs that were highly enriched for TFIIIC were also often associated with CCCTC-binding factor (CTCF) (31), a protein implicated in insulation and chromosome looping and conformation (36), suggesting a role for TFIIIC in defining repressive domains in human.

In this study, we investigated the distribution of Pol III, BRF1, BRF2, TFIIIC and one of the TFIIS homologs, TCEA1, on the genome of mouse embryonic stem (ES) cells using a highly specific ChIP-seq method that entailed tagged ES cell lines. Our work provided a detailed analysis of the active class III genes in mouse ES cells and led to the discovery of new genes transcribed by this enzyme. We found that only a few hundreds SINEs are transcribed by Pol III. Interestingly, the presence of TCEA1 was detected on the majority of the active class III genes suggesting that it acts as a Pol III transcription factor in mammals. We also took advantage of the genome-wide analysis of TCEA1 to show that it is present at similar levels on paused Pol II peaks of active and inactive genes suggesting that it is not TFIIS recruitment that triggers the passage of Pol II into elongation.

MATERIALS AND METHODS

Construction of the mouse ES cell lines

We used the recombineering technology (37) to generate the targeting vectors that introduce the triple affinity tag in the subunits of Pol III, TFIIIB-α or -β, TFIIIC and TFIIS. The 46C ES cell line (38) was transfected by

electroporation with each targeting vector. Cells were plated in D15 medium as described (39) and selected with G418. Individual ES cell colonies were collected 7 days after electroporation, amplified and genotyped by Southern blotting, in order to identify the clones that underwent a homologous recombination event. In these clones, a sequence encoding a 6 Histidine-Flag-HA tag, followed by a neomycin resistance marker flanked by loxP sites, was inserted just after the last codon of the gene encoding the protein to be tagged. For TCEA1, the tag was inserted just after the start codon. The sequence of the insertion cassette is given in the Supplementary Table S2. The integration of the cassette at the right loci was verified by Southern blotting using three or four different restriction enzymes and DNA-polymerase chain reaction (PCR). Transient transfection with a Cre recombinase expression vector was used to remove the selection cassette. The karyotypes of all cell lines were verified. The genes that were modified in the cell lines encoded RPC1 (MGI:2681836), RPC4 (MGI:1914315), BRF1 (MGI:1919558), BRF2 (MGI:1913903), TFIIIC220 (MGI:107887), TFIIIC110 (MGI:1919002), TFIIIC90 (MGI:2138937) and TCEA1 (MGI:1196624), respectively. The expression of the tagged versions of the proteins was verified by western blotting using HA7 anti-HA antibodies (Sigma; Supplementary Figure S2).

ES cell lines culture

ES cells were cultured on embryonic fibroblast feeder cells blocked with mitomycin C in D15 medium [Dulbecco's Modified Eagle's Medium High Glucose supplemented with 15% fetal bovine serum, 2 mM L-glutamine, 50 U/ml penicillin, 50 μg/ml streptomycin, 0.1 mM non-essential amino acids (all from GIBCO), 0.1 mM 2-mercaptoethanol (Sigma) and 1000 U/ml LIF]. ES cells were maintained at 37°C, 5% carbon dioxide, fed with fresh media daily, and transferred to new plates after trypsinization.

Chromatin immunoprecipitation

Typically $(150–200) \times 10^6$ cells were collected for each ChIP experiment. The cellular proteins and DNA were cross-linked by the addition of formaldehyde (0.4% final concentration). The plates were incubated for 10 min at room temperature and then the reaction was stopped by the addition of glycine (0.125 M final concentration). The cells were washed twice with 10 ml chilled phosphate buffered saline (PBS). Each plate was scraped with 2 ml PBS with protease inhibitors (Roche complete, 10 mM PMSF dissolved in ethanol), pelleted by centrifugation at 2500 rpm and resuspended in 1 ml per dish of FA/SDS buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS) with protease inhibitors and incubated on ice for 15 min. The subsequent operations were performed at 4°C. Chromatin was collected by centrifugation for 20 min at 12 000 rpm, resuspended in the same amount of FA/SDS and put on a rotating wheel for 1 h, centrifuged and resuspended in 0.4 volume of FA/SDS with protease inhibitors. The 400 μl batches of chromatin were fragmented with a sonicator (Diagenode) to generate DNA fragments of 300–400 bp mean size. The chromatin was collected by centrifugation at 14 000 rpm for 10 min and was kept at −80°C until needed.

The immunoglobulin-coupled magnetic beads (Dynal) used in the chromatin immunoprecipitation experiments were prepared essentially as described previously (40). Anti-HA antibody was HA7 H-3663 (Sigma) and anti-Flag anti-body was F-1804 (Sigma). For each ChIP experiment, 240 μg DNA (RPC1, RPC4, 46C) or 750 μg DNA (BRF1, BRF2) of sonicated chromatin were incubated with 150 μl beads for 2 h at room temperature. The beads were magnetically pelleted and the supernatant was removed. The beads were washed with 1 ml of low salt buffer (pH 8.0, 20 mM Tris–HCl, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.1% SDS), then with the same volume of high salt buffer (pH 8.0, 20 mM Tris–HCl, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS), then with low LiCl buffer (pH 8.0, 10 mM Tris–HCl, 1 mM EDTA, 250 mM LiCl, 1% NP-40, 1% deoxycholate) and finally twice with TE buffer (pH 8.0, 10 mM Tris–HCl, 1 mM EDTA). The first ChIP was performed using the anti-HA antibody. The bound chromatin was eluted using FA/SDS buffer containing the HA peptide (0.5 mg/ml; Ansynth Service) for 4 h at 16°C then overnight at 4°C. For the second ChIP, the supernatant was incubated with 50 μl beads coupled to anti-Flag antibodies which were treated as above except for the elution that was performed by incubation with a buffer containing 1% SDS and 0.1 M $NaHCO_3$. Reversal of the cross-links and DNA purification were performed as described previously (41). TFIIIC subunits and TCEA1 ChIPs were fragmented using MNase I. The protocol will be described in detail elsewhere (M. Gérard, manuscript in preparation).

The immunoprecipitated DNA was analyzed by quantitative real time PCR on an ABI Prism 7000 or 7300 machine (Applied Biosystem; 40). Relative quantification using a standard curve method was performed and the occupancy level for a specific fragment was defined as the ratio of immunoprecipitated DNA over total DNA.

Sequencing of the immunoprecipitated DNA and analysis of the regions bound by the RNA pol III

For each chromatin immunoprecipitation, the DNA was sequenced on a single Solexa genome analyzer GS or GA IIx channel using the procedures recommended by the manufacturer (Illumina). The characteristics of each sequencing experiment are indicated in Supplementary Table S3. The bound regions were identified using Quest version 2.3 (42) and an in-house program that used Quest method for peak calling. The data have been deposited to the ArrayExpress database under accession number E-MTAB-767.

RNA preparation

Total RNAs were isolated from 46C wild-type ES cells, using 1 ml of Trizol (Invitrogen) per 10 cm-dish, as indicated by the manufacturer. Total RNAs were prepared following the manufacturer's protocol, except that the

RNA pellet were washed with 80% ethanol, without shaking. The RNAs were suspended in diethylpyrocarbonate treated water, at around 2 mg/ml and stored at −80°C until needed. Integrity of RNA was tested on 1.5% agarose gel. Total RNA concentration and purity were verified using a NanoDrop spectrophotometer ND-1000 (Thermo Scientific), measuring absorbance at OD 260/280. Total RNA samples were treated with 1 U of RQ1 DNase (Promega) per μg of RNA, for 1 h at 37°C. DNase was inactivated by the addition of 20 mM ethylene glycol tetraacetic acid, pH 8.0 (Stop solution) and heated at 65°C for 10 min. The RNA was then precipitated with isopropanol.

cDNA synthesis

cDNA synthesis was performed using SuperScript II reverse transcriptase with random hexamer primers (Invitrogen) according to the manufacturer's instructions. In brief, 5 μg of total RNA, 2 μl of random hexamer primers (10 μM), 1 μl of dNTP mix (5 mM each) to 11 μl in total, were incubated at 65°C for 5 min. After chilling on ice for 2–3 min and brief centrifugation, 5 μl of first-strand synthesis buffer (5×, containing 250 mM Tris–HCl [pH 8.3], 375 mM KCl, 15 mM MgCl2), 2 μl of 0.1 M DTT, were added and the tubes were incubated at 25°C for 1 min. Then, 1 μl of (200 U/μl) of SuperScript II reverse transcriptase was added and the reaction was first incubated at 25°C for 10 min, followed by incubation at 42°C for 1 h. Reverse transcriptase activity was terminated by incubation at 70°C for 15 min. 1 μl RNase H (Invitrogen, 2U/μl) was added and further incubated at 37°C for 20 min. Samples were stored at −20°C until needed. The cDNA solution was diluted 10-fold before use in reverse transcriptase-PCR (RT–PCR). PCR were performed using specific primers covering the region of interest. The reactions contained 25 ng of cDNA, and primers at a final concentration of 150 nM. RT-PCR reactions were analyzed by gel electrophoresis. Further, total RNA samples were analyzed for the possible presence of DNA contamination by PCR using RNA not reverse-transcribed.

RESULTS

Setting up a high-specificity tandem affinity chromatin immunoprecipitation method in mouse ES cells

To identify the regions bound by the Pol III transcription machinery, we developed a tandem affinity chromatin immunoprecipitation method in mouse ES cells. Two different specific subunits of Pol III, RPC1 and RPC4, the BRF1 and BRF2 subunits of TFIIIB-β or TFIIIB-α, respectively and the TFIIIC220, TFIIIC110 and TFIIIC90 subunits of TFIIIC were tagged (Supplementary Table S1). To investigate whether or not TFIIS is a Pol III transcription factor, TCEA1 was also tagged. We introduced a cassette encoding consecutively six histidines, one Flag and one HA epitope just after the last sense codon of the Pol III, BRF and TFIIIC genes. A neomycin marker flanked by loxP sites follows the tag cassette. The construction was introduced in 46C mouse ES cell genome using the recombineering method (Supplementary Figure S1 and Supplementary Table S2) (37). The neomycin marker was removed by expressing the cre recombinase. The correct integration of the cassette at the endogenous locus and the excision of the neomycin marker were verified by Southern blotting and PCR. The expression and correct size of the C-terminally tagged proteins were verified by western blotting with a monoclonal anti-HA antibody (Supplementary Figure S2).

Chromatin from the RPC1-, RPC4-, BRF1- or BRF2-tagged ES cell lines was prepared. ChIP experiments in RPC1 or RPC4 ES lines with anti-HA antibodies enriched strongly a tRNA-val gene and the H1 gene (from 40- to 110-fold) but not the ARBP gene, which is transcribed by Pol II (Figure 1A). We then tested whether genes that depend on BRF1 or BRF2 could be distinguished. BRF1 ChIP enriched specifically two tRNA genes. Conversely, BRF2, but not BRF1, was associated with H1 gene, encoding the RNA subunit of the RNase P, and U6 (Figure 1B) (5,6).

To improve the specificity of the ChIP experiments, tandem immunoprecipitations were performed. The chromatin was first ChIPed with the anti-HA antibody as above, and then eluted by competition with an HA peptide. The enriched chromatin was submitted to a second round of ChIP with a monoclonal anti-Flag antibody. The second ChIP further improved the enrichment ratio above background up to 8-fold depending on the protein and gene considered (Figure 1C and Supplementary Figure S3). The protocol used here is thus, highly specific for ChIP experiments of tagged proteins expressed from their native locus in mouse ES cells. It is also generic since it does not depend on the generation of high specificity antibodies.

Analysis of the regions bound by RNA polymerase III

DNA from the chromatin associated with Pol III was sequenced using a Solexa Genome Analyzer. The comparison of ChIP-Seq experiments using tagged RPC1 and RPC4 ES cell lines allowed us to cross-validate the Pol III-binding sites. We considered as bound regions, only those that showed co-occupancy by RPC1 and RPC4. These regions were validated if the number of tags mapping within the defined interval in the ChIP-Seq experiments with RPC1 and RPC4 were both 5-fold higher than in the untagged ES cell line, which was used as a negative control. Regions that satisfied only one of the two conditions were visually inspected using the UCSC genome browser and rejected if the 46C track showed a high background around the bound region. The protein-bound regions were annotated using the mm9 UCSC mouse database. The regions that were associated with more than one annotation were visually inspected and eventually split to associate one region with each annotation. If one annotation was associated with two or more regions, these were inspected and eventually fused. Figure 2 and Supplementary Figure S4 show examples of tag density profiles on various regions representative of several class III genes. Data concerning the bound regions can be found in Supplementary Table S4.

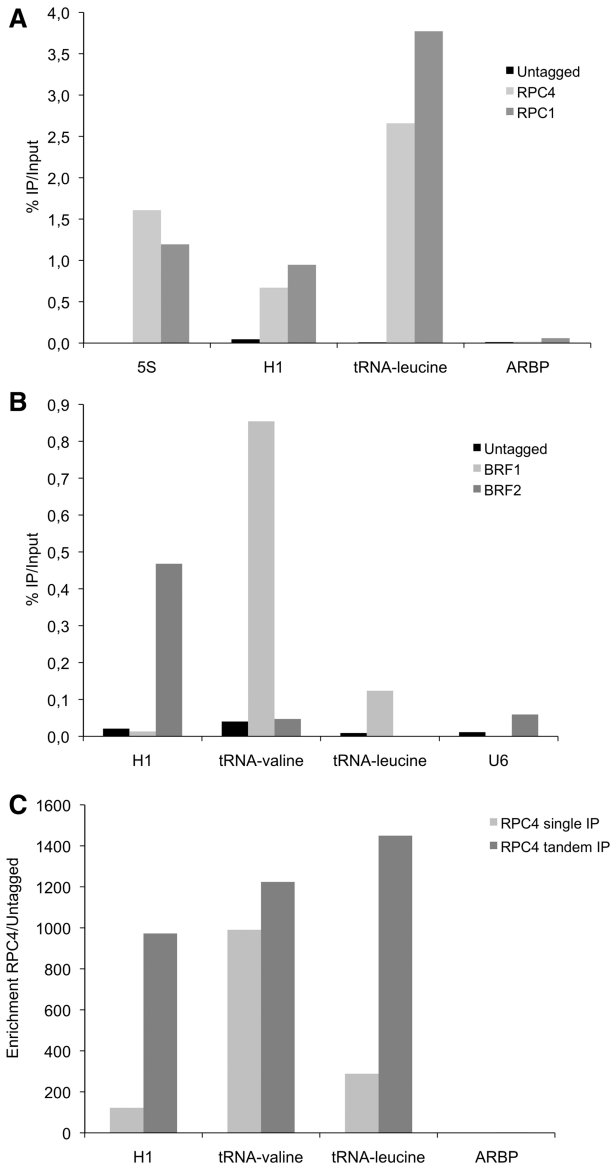

Figure 1. Chromatin immunoprecipitation of the Pol III transcription machinery in mouse ES cell lines. ChIP experiments were performed with the HA7 H-3663 anti-HA antibody (**A** and **B**) as described in 'Materials and Methods' section. (A) Chromatin extracts from RPC1 or RPC4 cell lines were used to show Pol III enrichment on different types of class III genes but not on the ARBP class II gene. (B) Extracts from BRF1 or BRF2 cell lines were used to demonstrate specific enrichment of TFIIIB-β or TFIIIB-α on genes with type 2 (tRNAs) or type 3 promoters (H1, U6), respectively. An untagged cell line was used as a negative control. (**C**) Tandem ChIPs improved signal to noise ratios. Single round ChIP (RPC4 single IP) were performed as above with HA7 H-3663 anti-HA antibody and were compared with experiments in which a second round (RPC4 tandem IP) of immunoprecipitation was done with F-1804 anti-Flag antibody after elution of proteins with an HA peptide. Enrichment-folds of sequences from an RPC4 ES cell line chromatin extract relative to an untagged cell line chromatin are indicated.

tRNAs. A total of 284 tRNA genes were bound by Pol III, including one selenocysteine tRNA gene and one possible suppressor. Of these, 271 were predicted by Coughlin *et al.* (43) and 281 by the Genomic tRNA Database (GtRNADb). Coughlin *et al.* predict that 461 tRNA genes exist based on the sequence of the expressed tRNAs. GtRNADb, which relies upon the tRNAscan-SE program (44), predicts the existence of 433 tRNA genes (including two selenocysteine tRNAs and one possible suppressor) in the mouse genome. Altogether, 526 tRNAs are annotated on the mouse genome by these two databases. Totally, 59 and 65% of the genes predicted by Coughlin *et al.* and GtRNADb, respectively, were bound by Pol III in mouse ES cells. This observation raised the possibility that some tRNAs were not detected because they cannot be identified by unique tags. The percentage of nucleotides that could not be mapped (NM score) within 150 nt on either side or within the tRNA genes that were actually bound by Pol III was computed. For each tRNA, we considered the lowest NM score from the three regions. We found that the worst NM score for a bound tRNA was 75%. We figured out that among the 242 tRNA genes predicted by Coughlin *et al.* or GtRNAdb that were not bound by Pol III, only 14 had an NM score >75%. Hence, we did not underestimate the number of bound tRNA genes by >13.2%. We performed independent ChIP experiments for three bound tRNA genes and two mappable unbound tRNA genes, which agreed with the ChIP-Seq experiments (Supplementary Figure S5). The genomic distribution of bound tRNA genes and other class III genes is shown in Supplementary Figure S6.

snRNAs. In addition to tRNA genes, we also found binding of Pol III transcription machinery on the U6, 5S, 7SK, 7SL, 4.5S, BC1, HY1 and HY3 genes. However, the number of bound genes was very small compared to that predicted in Repbase (Table 1). For example, Repbase predicts 1269 U6 genes and Ensembl, 617 when we actually found only 5 that were bound. As for the tRNA genes, we performed an independent ChIP experiment and verified for two genes of each 7SL, 7SK, U6 and HY3, one bound in the ChIP-Seq experiment and one mappable but unbound, that Pol III was present or not as expected (Supplementary Figure S5).

New Pol III transcripts. About 30 regions were bound by both RPC1 and RPC4, but were not annotated as Pol III-transcribed genes or SINEs and could thus correspond to new Pol III transcripts. Of note, these regions were neither annotated as transcribed by Pol II, nor corresponded to miRNAs. We verified by an independent ChIP experiment for nine of those regions that they were indeed bound by RPC1 and RPC4 (Supplementary Figure S7A). Furthermore, we looked for the presence of an RNA associated with the Pol III-bound region by RT–PCR (Supplementary Figure S7B). An RNA was transcribed from 16 regions out of the 18 that were examined. Most of these regions were conserved in the rat genome but only one was highly conserved in human (Supplementary Figure S7C and Supplementary Table S5). The latter region, which was also associated with Pol III in human, had a typical organization for class III genes with conserved A-box and B-box and terminator. Altogether, these results strongly suggest that we have identified new class III transcripts.

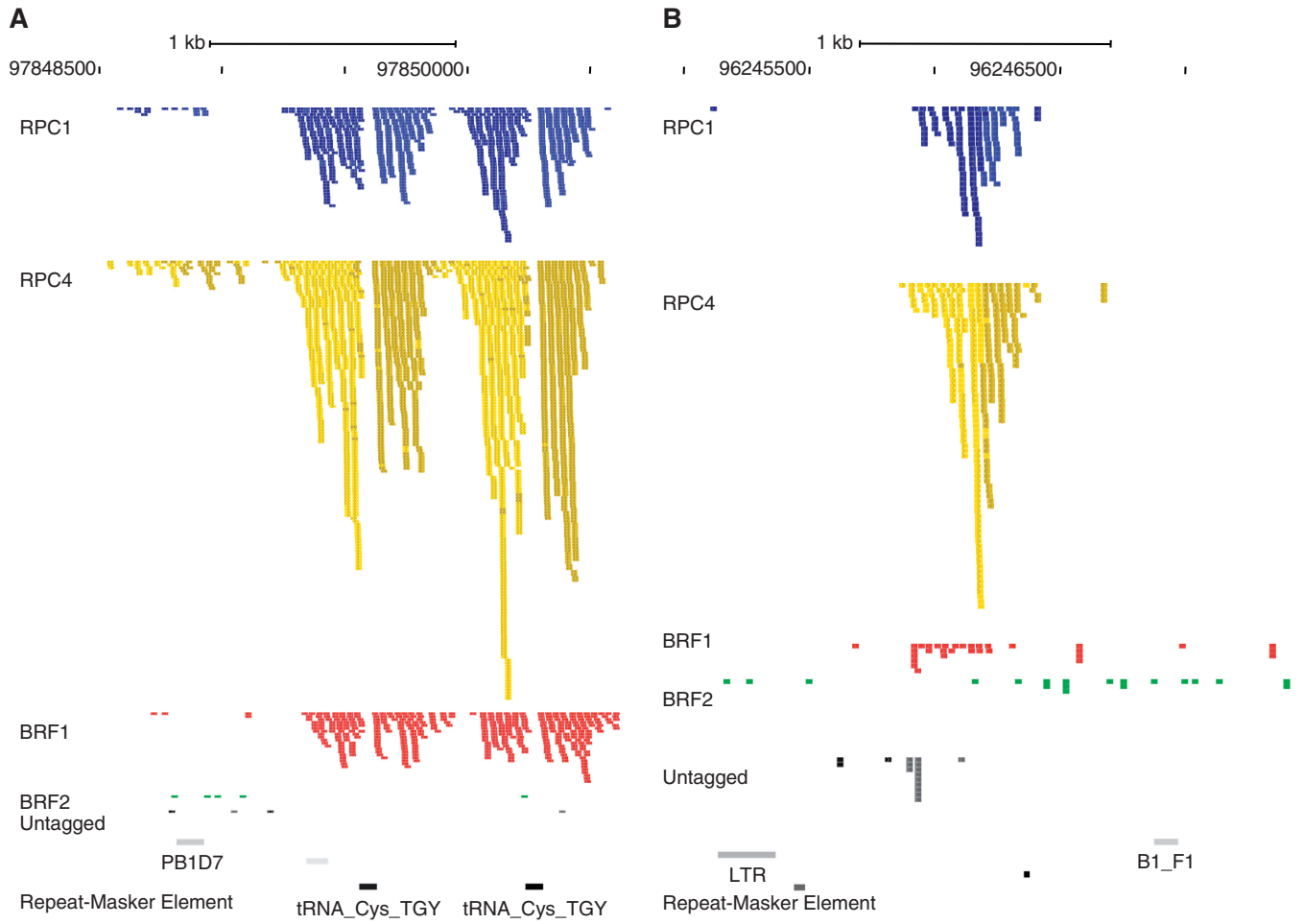

Figure 2. Representative examples of sequence reads that map on bound regions. (**A**) Binding of Pol III and BRF1 to two tRNA genes and a SINE. The sequences that map to chromosome 11 are displayed using the UCSC genome browser. A 1 kb scale is shown at the top of the Figure. The sequence tags are represented by colored rectangles. Tags identified in the ChIP-Seq experiments performed with the RPC1-tagged, RPC4-tagged, BRF1-tagged or BRF2-tagged ES cell lines and 46C negative control cell line are colored in blue, yellow, red, green and gray, respectively. Light colors indicate the sequences that match the top strand. Dark colors indicate the tags that match the bottom strand. The location of repeated sequences, according to RepBase, is indicated at the bottom of the Figure by gray or black boxes. The two black boxes refer to two tRNA-Cys-TGY genes. The bound SINE is a PB1D7 repeat of the Alu family. (**B**) Binding of Pol III transcription machinery to an unannotated region on chromosome 3.

Distribution of BRF1 and BRF2 on class III genes

We also tagged the BRF1 or BRF2 subunits of TFIIIB-β or TFIIIB-α, respectively, to investigate which genes depend on either form of the Pol III factor. Type I and II promoters require the presence of BRF1 in TFIIIB-β transcription factor for their transcription while type III uses BRF2 as a subunit of TFIIIB-α. As expected, BRF1 was exclusively associated with types I and II promoters (Figures 2A and 3A–C), whereas, BRF2 was bound to type III promoters only (Figure 3D and Supplementary Figure S4), indicating that no class III gene could use both TFIIIB variant. BRF2 was associated with a small number of snRNAs that included U6, 7SK, the H1 and MRP RNA genes, U6atac, the $tRNA^{sec}$ and two of the Y RNA genes, HY1 and HY3. Apart from these genes that were already predicted to depend on BRF2, no new region was found. The new genes were associated with low, or sometimes background levels of BRF1 (Figure 3E). This situation probably stems from the fact that the number of tags associated with the new genes is often low even for the Pol III ChIP. The number of tags associated with class III genes in experiments with the BRF1 cell line is around 10-fold smaller than with the RPC4 cell line (Supplementary Table S4, compare RPC4.SUM_DENSITY to BRF1.SUM_DENSITY). A thorough assessment of BRF1 and/or BRF2 presence on the new genes would thus require a 10-fold deeper sequencing.

We wondered where the two forms of TFIIIB were positioned, relative to the mature RNA 5′-end (the TSS of mouse class III genes have not been systematically determined). On tRNA genes, BRF1 peak density was located 10 nt upstream of RNA 5′-ends (Figure 3A), which is within a few nucleotides of the position determined by *in vitro* footprinting experiments. A similar situation was found for BRF2 on genes with type III promoters with a peak of tag density located 14 nt

upstream RNA 5′-ends (Figure 3D). Using data from Kagey *et al.* (45), we also looked at the distribution of TBP relative to the class III genes. TBP was positioned 17 or 12 nt upstream of RNA 5′-ends for class II or class III promoters, respectively.

Table 1. Pol III-associated regions

RNA	Number bound	Number predicted[a]	Percent bound
tRNA	284	526	53.8
4.5S	13	1475	0.81
5S	3	984	0.30
7SK	1	665	0.15
7SL	2	276	0.72
BC1	2	6351	0.03
HY1	1	37	2.70
HY3	1	15	6.67
HY4	0	2	0.00
HY5	0	1	0.00
U6	5	1269	0.39

[a]The number of predicted class III transcripts are extracted from GtRNAdb and Coughlin *et al.* (43) for the tRNAs and from Repbase for the other ones.

Transcription of SINEs

Pol III *in vitro* transcription systems can drive the transcription of SINEs (4). We wondered whether some of the SINEs could be bound by Pol III, and thus most probably be transcribed *in vivo*. Indeed, we found 241 locations on the mouse genome that were associated with both RPC1 and RPC4 and had a SINE annotation according to Repeatmasker. We first looked at the distribution of the sequence divergence of the SINEs (Supplementary Figure S8A). The distribution is bi-modal with maxima at 7 and 27% divergence and a minimum at 13%. The SINEs were separated in two classes according to their divergence using 13% as threshold. Notably, most (80%) of the bound SINEs belonged to the highly conserved category (Supplementary Figure S8B).

We wondered if the small number of observed bound SINEs in our ChIP-Seq experiments could stem from the low mappability of the regions encompassing the SINEs. To answer this question, we used an approach that was similar to the one applied to the tRNA genes. The bound SINEs were divided in two populations depending on whether their divergence was ≥13%. We determined the lowest NM score for bound SINEs of each category. The

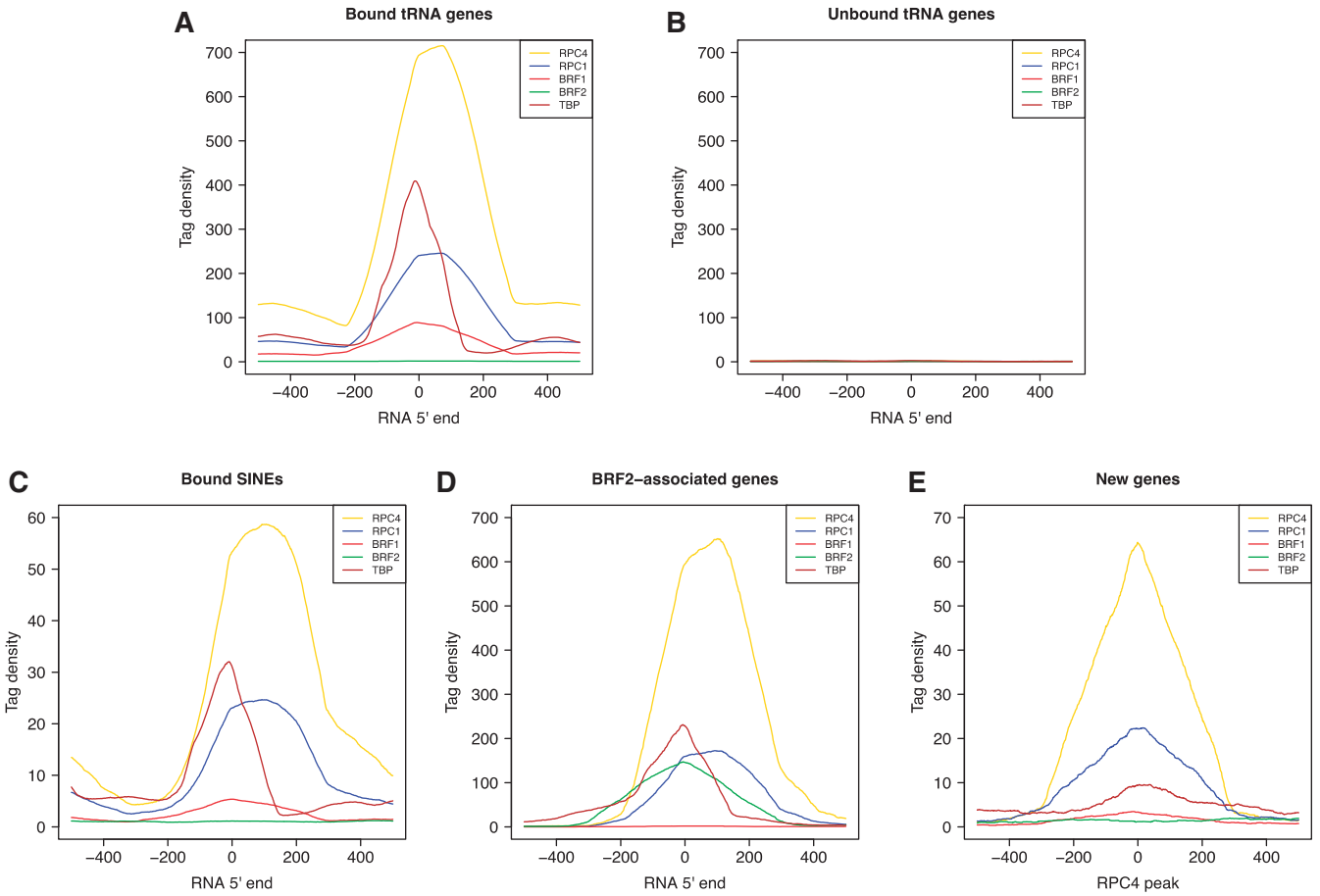

Figure 3. Distribution of Pol III, TBP, BRF1 and BRF2 relative to class III genes. Tag densities are shown for RPC1, RPC4, BRF1, BRF2 and TBP (45) relative to the location of the RNA 5′-end for (**A**) tRNA genes, (**B**) unbound tRNA genes, (**C**) SINEs, (**D**) BRF2-associated genes and (**E**) new class III genes. For these, neither the transcription orientation, nor the TSS or end is known. Hence, the distribution is shown relative to the RPC4 density peak.

worst NM score within the SINE or the 150 nt flanking-regions was 65% for the conserved category and 50% for the diverged category. We then computed the NM score for all the SINEs predicted by Repeatmasker for each of the two divergence classes and determined the number of SINEs present in the mouse genome that had their two flanking regions with NM scores above those observed in bound SINEs. We postulated that the proportion of bound SINEs in the non-mappable population was similar to that in the mappable population. We could thus estimate that only around 10 non-mappable SINEs might be bound.

We searched in the bound SINEs for the presence of the TRGYTYARTGG and GTTCRAWTC sequences, which correspond to the A- and B-boxes in human (30). Of them, 64% contained an A-box and 87% a B-box (at *P*-value 10^{-3}), indicating that most of them had highly conserved type II promoters. Pol III-associated SINEs were found throughout the mouse genome and were not particularly associated with other categories of class III genes (Supplementary Figure S6B). Altogether, these observations indicate that a limited set of the SINEs direct Pol III transcription *in vivo*.

Chromatin environment of class III genes

We looked for the presence of various chromatin marks around the bound and unbound class III genes, comparing the distribution of H3K4me1, -me2, -me3, which are considered as active chromatin marks, and H3K27me3, that is an inactive chromatin mark (46–48). We also analyzed the distribution of H3K9me3, which are deposited at many euchromatic loci in mouse ES cells (49). Interestingly, both Pol III bound tRNA genes and BRF2-dependent genes were surrounded by peaks of H3K4me3 located around 400 nt on either side of the RNA 5′-ends (Supplementary Figure S9). H3K4me3 mark was absent from unbound tRNA genes (data not shown). The euchromatic mark H3K9me3 might be slightly enriched on either side of active class III genes. On the contrary, H3K27me3, which is typically heterochromatic was completely absent. This pattern is extremely similar to that of Pol II-transcribed genes.

Distribution of TFIIIC

TFIIIC is required for the transcription of class III genes that are under the control of types I and II promoters, i.e. 5S and tRNA genes. However, in *S. pombe*, *S. cerevisiae* and human, TFIIIC binds some regions independently of Pol III (26,30,31,35). We explored the genome-wide distribution of TFIIIC in mouse ES cell lines where one subunit of TFIIIC, TFIIIC220, −110 or −90, was tagged. Since the standard ChIP protocol that used sonication for DNA shearing did not give satisfactory results, it was modified by the inclusion of a MNase I DNA digestion step (M.G., manuscript in preparation). Tandem ChIP experiments allowed the identification of 2652 regions that were considered as bound since they were consistently enriched in ChIP experiments with the three TFIIIC-tagged cell lines but not in an untagged control cell line (Supplementary Table S6).

Of the 283 tRNA genes bound by Pol III (excluding the selenocysteine tRNA which has a type III promoter), 261 were also associated with significant levels of all three TFIIIC subunits. The three subunits displayed similar distributions downstream of BRF1 on the body of the tRNA genes (Figure 4A). The distribution of the tags on tRNA genes was very similar for the three subunits tested, showing an extended association with tRNA genes. TFIIIC was absent from the tRNA genes that were not bound by Pol III or BRF1 (data not shown). In line with their promoter organization, Pol III-bound SINEs were also associated with TFIIIC and BRF1, but not BRF2 (Figure 4B). As expected from previous *in vitro* experiments (5), the BRF2-associated genes were completely devoid of TFIIIC (Figure 4C). Interestingly, the three TFIIIC subunits were generally associated with the new class III genes while no BRF2 could be found suggesting that they depend on TFIIIC for their transcription (Figure 4D).

Several recent studies pointed at the presence in human cell lines of numerous ETC loci where TFIIIC is present but devoid of Pol III or TFIIIB-α or -β (26,30,31,35). Similarly, 2233 ETC loci bound by the three TFIIIC subunits were detected by our ChIP-Seq experiments. TFIIIC has been shown to play a role in the organization of chromatin in *S. cerevisiae* and *S. pombe* and to act as a barrier in the extension of heterochromatin into euchromatic territories (35,50,51). It has been shown that CTCF protein plays a similar insulator role in mammals (36). Moreover, CTCF interacts with cohesins to position it at many sites (52). A correlation between CTCF and ETC sites in human cells has been found (30,31). To investigate the potential correlation between the ETCs- and CTCF-binding sites in mouse ES cells, the ETCs were ordered according to their distance relative to the CTCF-binding sites (data taken from ref. 53). In parallel, an identical number of randomly selected regions were ordered using the CTCF-binding site distance criterion. Heat maps of the tag density for TFIIIC220 and CTCF were generated (Figure 5A) and compared with the heat maps of the randomly selected regions (Figure 5B). In agreement with the observations made in human cells, we found that 85% of ETCs also had CTCF-binding sites within 20 kb (Figure 5A). The distribution of CTCF was clearly different when regions were selected randomly (Figure 5B; compare the sigmoid curves in the two panels). The CTCF-binding sites were previously shown to be associated with the Smc1A and Smc3 subunits of cohesin (45). Heat maps were drawn for Smc1 and Smc3 and confirmed the clear co-occupancy of these proteins with CTCF. Intriguingly, in addition to being closely associated with CTCF, we observed that Smc1A and, to a lesser extent, Smc3 were enriched at the ETCs themselves, as indicated by the increased tag density centered on the ETCs (Figure 5A) which is absent from the maps of the randomly selected regions (Figure 5B).

Enrichment of TCEA1 isoform of TFIIS at class II and class III genes

TFIIS is a transcription elongation factor that stimulates Pol II transcription elongation and initiation. Moreover,

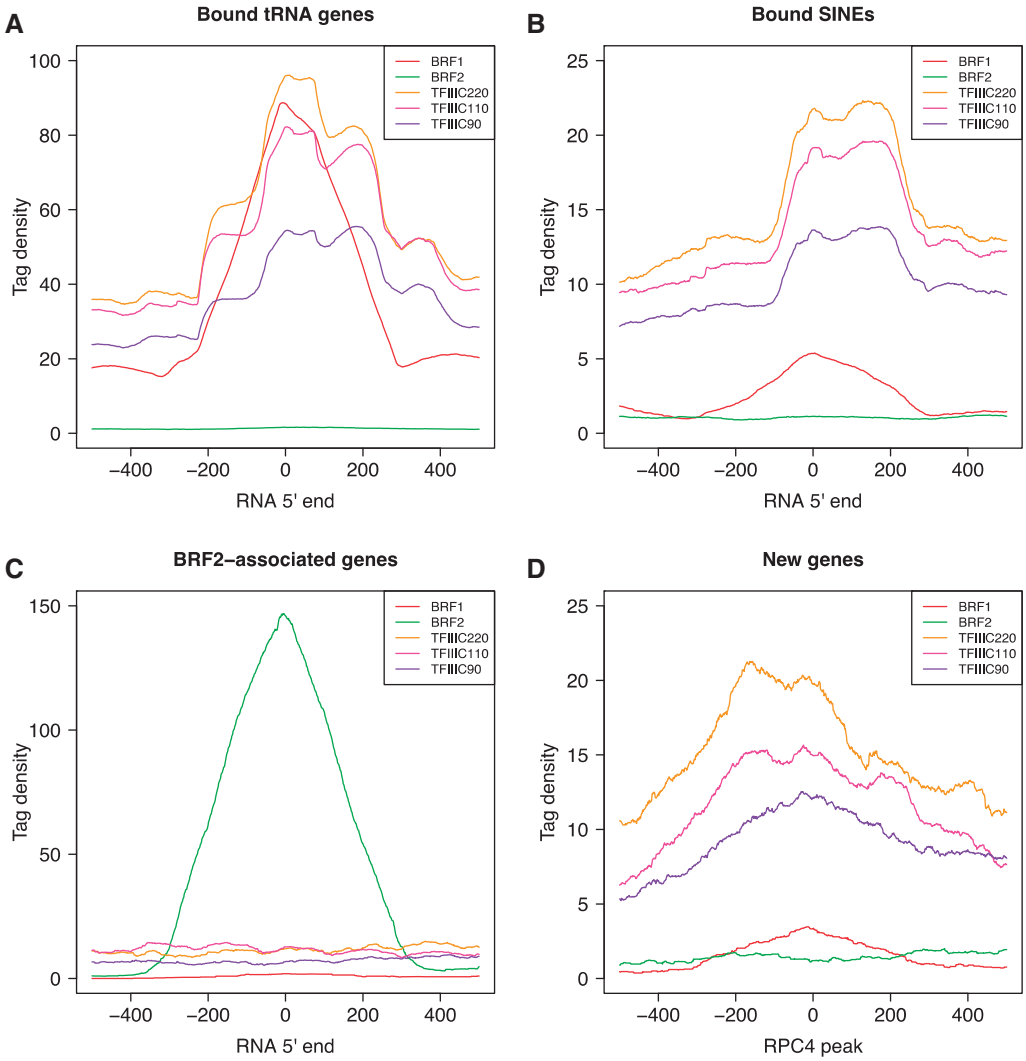

Figure 4. Distribution of TFIIIC relative to class III genes. Tag densities are shown for TFIIIC220, −110, −90, BRF1 and BRF2 relative to the location of the RNA 5′-end for (**A**) tRNA, (**B**) SINEs, (**C**) BRF2-associated. For the new class III genes (**D**) the distribution is shown relative to the RPC4 density peak.

we have shown that it acts positively on Pol III transcription in *S. cerevisiae* (22). The distribution of TFIIS on the genome of mouse ES cells was thus investigated. Three variants of TFIIS are encoded by the mouse genome. The ubiquitously expressed variant TCEA1 was tagged at its N-terminus because C-terminal tagging abolishes TFIIS function in yeast. Chromatin immunoprecipitation was performed using the same protocol that was used for TFIIIC.

We found that 57.6% of the tRNA genes and 50% of the BRF2-dependent genes that were associated with Pol III were also bound by TCEA1 (Figure 6 and Supplementary Figure S10) using a 3-fold signal to background threshold. About 41.1% of the bound SINEs and 16.7% of the new class-III genes were also bound by TCEA1 using the same criterion. We wondered if the presence of TCEA1 on class III genes does require Pol II (data from ref. 54). On tRNA genes and active SINEs, only very low levels of hypophosphorylated Pol II is found upstream of the genes (Figure 6A and B). Low levels of hypophosphorylated Pol II was also present downstream of the SINEs. Pol II phosphorylated on serine 2, 5 or 7 of the CTD was not significantly enriched on the regions surrounding active tRNA genes or SINEs. The distribution of TCEA1 on tRNAs and SINEs resembled that of Pol III but was shifted upstream toward the 5′-end in line with its possible role in transcription initiation by Pol III (22). Intriguingly, the hypophosphorylated, S7P and S5P forms of Pol II were all significantly present upstream of BRF2-associated genes (Figure 6C). The level of unphosphorylated Pol II was around 5-fold more abundant on BRF2-associated genes, than on active tRNA genes. However, two peaks of TCEA1 were found on BRF2-dependent genes, one associated with Pol II, the other with Pol III. These observations suggest that TFIIS plays a role in Pol III transcription.

The distribution of TCEA1 on class II genes was also analyzed. When ordered according to RNA steady-state levels transcribed from a given gene (measured by RNA-seq; C. Keime and M. Gérard, manuscript in

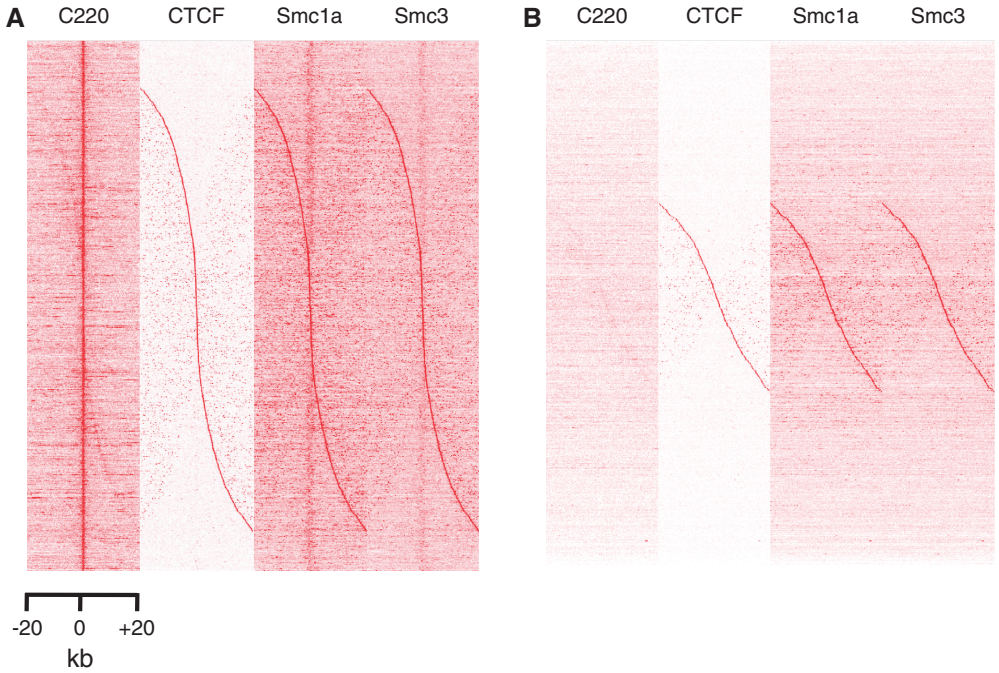

Figure 5. Distribution of CTCF, Smc1a, Smc3 relative to ETC-bound TFIIIC220. (**A**) Distribution of CTFC, Smc1a and Smc3, relative to TFIIIC220. The ETC regions bound by TFIIIC220 (C220) were ordered relative to their distance to CTFC-binding sites using seqMINER (63). The windows span 20 kb upstream and downstream the center of TFIIIC bound regions. The distribution of Smc1a and Smc3 on the same regions is shown on the two right panels according to the same order. (**B**) Distribution of CTCF, Smc1a and Smc3 relative to randomly selected regions. Randomly selected regions were sorted according to the distance of CTCF-binding site to the center of the region. Smc1a- and Smc3-binding sites are shown according to the same in the two right panels.

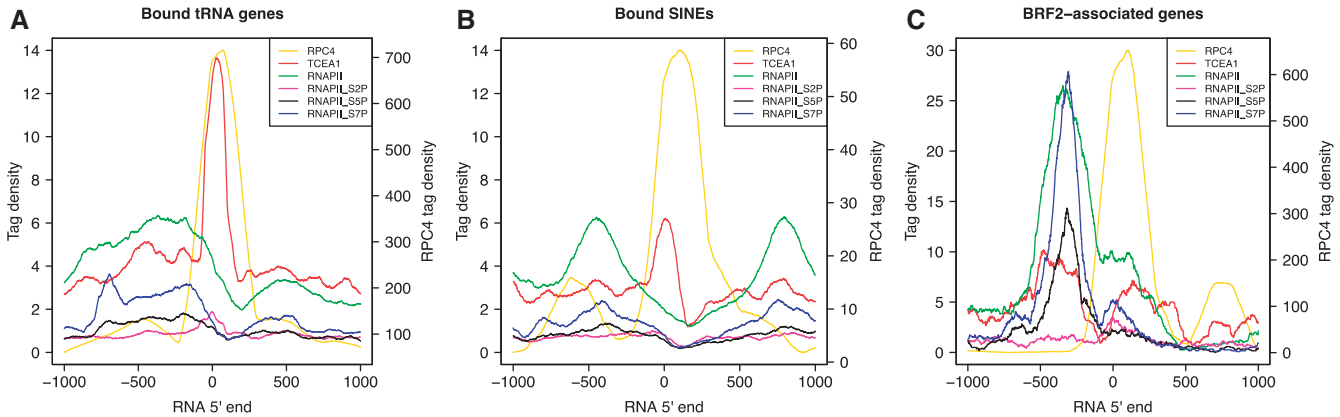

Figure 6. Distribution of TCEA1 and Pol II relative to class III genes. Tag densities for RPC4, TCEA1 (this work) and the hypophosphorylated (RNAPII) (64), serine 2 (RNAPII_S2P) (54), serine 5 (RNAPII_S5P) and serine 7 (RNAPII_S7P) phosphorylated forms of Pol II, are shown relative to the location of the RNA 5′-end for bound (**A**) tRNA, (**B**) SINEs and (**C**) BRF2-associated genes. The tag density scale for RPC4 is indicated on the right side of the plots.

preparation), Pol II and TCEA1 distributions were similar (Figure 7). Indeed for the 10 000 most highly expressed genes, the Spearman correlation coefficient between Pol II and TCEA1 occupancies on class II genes was 0.77. A large fraction of the genes, active or inactive, have paused Pol II just after the TSS. We wondered if TCEA1 occupancy might correlate with the transition from pausing to elongation. This is not the case since the levels of Pol II and TCEA1 on the pausing regions (defined here as the 300 bp after the TSS) of actively transcribed or non-productive genes with paused Pol II were similar, with Spearman correlation coefficient of 0.59 and 0.54, respectively (see Figure 7B and C for examples of Pol II and TCEA1 distributions). The Spearman correlation coefficient between TCEA1 and Pol II occupancies downstream of the pausing region were 0.75 and 0.76, respectively, for the two classes of genes.

DISCUSSION

The genome-wide distribution of Pol III, TFIIIB-β, its variant form TFIIIB-α and TFIIIC has been established

in mouse ES cells using a highly specific ChIP-Seq procedure. Only 284 tRNA genes out of 526 predicted genes were indeed associated with Pol III and, hence, were likely to be transcribed. The fact that Pol III was found only on a very small number of SINEs suggests that mechanisms exist to prevent their transcription. Additionally, 30 sites on the mouse genome were associated with Pol III, but lacked any annotation, most of them being transcribed. The regions encoding the new transcripts were usually conserved in the rat genome but except in one case, this conservation did not extend to man. The chromatin environment of class III genes resembles strongly that of class II genes with high levels of H3K4me3 upstream and downstream of the gene and low levels around the transcription start site. Studies in human have indicated that Pol II is associated with class III transcription (29,30,32). However, we found different situations depending on whether we looked at BRF1- or BRF2-associated genes. Upstream of tRNA genes, only very low levels of hypophosphorylated Pol II were observed. On the contrary, Pol II was present upstream of BRF2-associated genes. Finally, we observed that TCEA1 isoform of TFIIS was associated with classes II and III genes. The distribution pattern of TCEA1 suggests that it

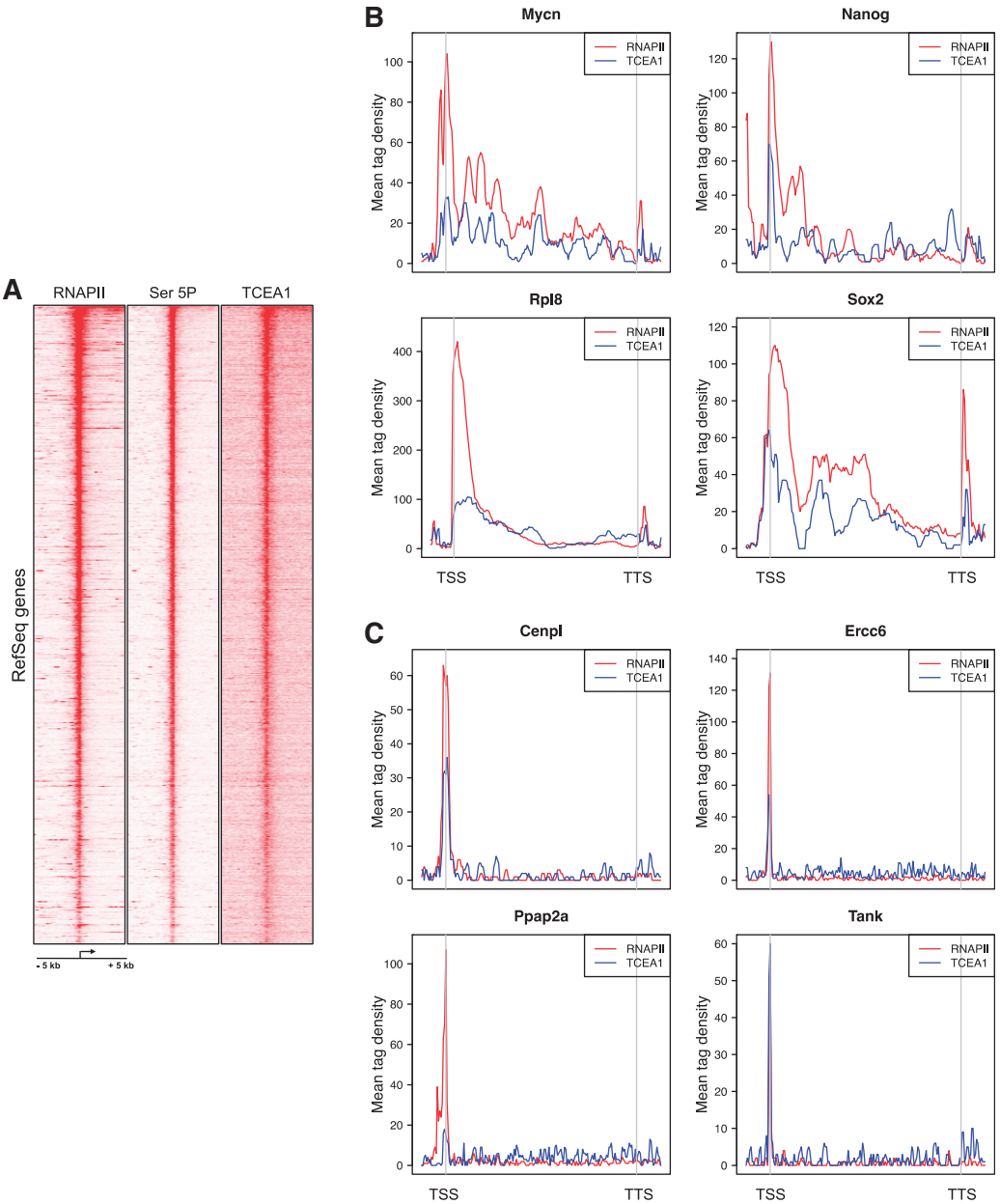

Figure 7. Distribution of TCEA1 on class II genes. (**A**) Hypophosphorylated Pol II (RNAPII) and TCEA1 have similar distribution on genes. The distribution of Pol II, S5P Pol II and TCEA1 5 kb upstream and downstream of the TSS is plotted for the 10 000 most expressed genes in ES cells based on RNA-Seq (M.G., manuscript in preparation). The genes were ordered according to their expression levels using seqMINER. (**B**) Examples of Pol II and TCEA1 profiles on highly transcribed genes. TTS: Transcription termination site. (**C**) Examples of Pol II and TCEA1 profiles on paused unproductive genes.

is not sufficient to stimulate the transition of paused Pol II into elongation. Moreover, it indicates that TFIIS could, as in yeast, be a Pol III general transcription factor.

Thanks to our highly specific tandem ChIP-Seq protocol and cross-validation by experiments that investigated independently the distribution of several subunits belonging to Pol III or TFIIIC, the mouse Pol III transcriptome is now precisely defined. As in human cell lines (29–33), the Pol III transcription machinery is associated with only a subset of the tRNA genes. This situation also holds true for the BRF2-associated genes. We estimated that <5% of the tRNA genes might have escaped our analysis. A total of 30 unannotated regions were associated with significant levels of Pol III. Several arguments indicate that they are indeed transcribed by Pol III. First, we could verify the presence of Pol III in independent experiments. Second, in addition to Pol III, most of these regions were associated with TBP and TFIIIC. In contrast, BRF2 was present only on regions that were already known to require TFIIIB-α for their transcription *in vitro*. Third, an RNA was found in >80% of the new regions that we tested. Finally, none of these regions was annotated. Experiments have suggested that a small number of miRNAs is transcribed by Pol III in human (55,56). This view was later challenged by other experiments (57). In mouse ES cells, none of the newly identified regions overlapped with miRNA annotations suggesting that the transcription of this class of RNA by Pol III is very limited. The function, if any, of the RNA transcribed from the new Pol III-associated regions will await further studies.

In vitro transcription experiments have demonstrated that Pol III is able to transcribe the SINEs (4). It is also known that SINEs, which number in hundreds of thousands in mammalian genomes, result from the insertion of retrotranscribed Pol III transcripts (3). We thus wondered if we could estimate how many of them are indeed associated with the Pol III transcription machinery. Unexpectedly, only 241 SINEs were bound, usually at rather low levels compared with tRNAs. The transcribed SINEs were distributed throughout the genome independently of tRNA or other class III genes. The small number of transcribed SINEs suggests that a general mechanism might repress their expression. This situation might, however, be specific to ES cells.

In human cells, class II genes are closely associated with class III genes. Pol II transcribed genes are often present upstream of class III genes (29–32). Even though Pol II transcription inhibition has only a modest effect on the transcription of adjacent class III genes, it has been argued that Pol II might regulate Pol III transcription (32). Our observations in mouse indicated that, at actively transcribed tRNA genes, the level of Pol II is extremely low. Moreover, while the hypophosphorylated non-elongating form of Pol II was present upstream of the bound tRNA genes at very low levels, the other forms were virtually absent. In contrast, hypophosphorylated, S7P and S5P Pol II were all present upstream of BRF2-associated genes. This observation suggests that, in mouse, the chromatin organization around transcribed tRNA genes, allows low levels of hypophosphorylated Pol II to bind but prevents the transition to elongation. On the other hand, the promoters of BRF2-associated genes, which are gene external and resemble those of Pol II transcribed genes, would allow the association of class II transcription machinery and divergent transcription relative to Pol III. The observation of different behaviors for tRNA genes and BRF2-associated genes strongly suggests that, on tRNA genes, Pol II presence has probably limited functional significance in mouse. Type III promoters are gene external and their transcription by Pol III is determined by the presence of a TATA box which, if absent, leads to transcription by Pol II (58,59). It has also been shown that U6 transcription is decreased ~2-fold in human cell lines when Pol II transcription is inhibited (60). It is thus possible that, as in human, the presence of Pol II upstream of BRF2-associated genes in mouse might stimulate their transcription by Pol III.

The binding pattern of TFIIIC was clearly distinct from that of Pol III. Whereas the factor was present on tRNA genes and on the new Pol III-associated regions, more than 2200 TFIIIC-binding sites were located far away from class III genes. This number is similar to that of the ETC loci in K562 human cells where 1865 such sites were found (31). The ETC sites have first been detected in *S. cerevisiae* and *S. pombe* and have been shown to play a role in the organization of chromatin (26,35,51). Even though the ChIP-Seq experiments were performed with different protocols, antibodies and in different cell lines, the ETC site number in human is remarkably similar to that found in mouse, both organisms having comparably sized genomes (3.1 Gb and 2.9 Gb for man and mouse, respectively). CTCF has been shown to have an enhancer-blocking activity (36). In mouse ES cell lines, CTCF is loosely associated with ETC loci, a relationship that might be indicative of an interdependence and/or common role in silencing transcription of adjacent sequences. Cohesins are required for the localization of CTCF at boundary elements (52). In line with previous observations, in mouse and human, Smc1a and Smc3 distributions closely follow that of CTCF. Remarkably, Smc1a and Smc3 were also present on the TFIIIC peak. In *S. cerevisiae* mutations in Smc1 or Smc3 affects the boundary function of tRNA genes (35,50,51). Moreover, a recent report has shown that cohesins connect enhancers and promoters through DNA loops via interactions with the Mediator, stimulating Pol II transcription (45). These observations raise the intriguing possibility that the presence of cohesins at ETCs and CTFC-binding sites might stimulate the formation of DNA loops, promoting an enhancer-blocking activity of TFIIIC and shape the organization of chromatin.

In *S. cerevisiae*, we showed that TFIIS occupies nearly all class III genes. In addition, some TFIIS mutations specifically affect Pol II or Pol III transcription *in vitro* and *in vivo* demonstrating that TFIIS plays a key role in Pol III transcription in yeast (22). The observation that, in mouse ES cells, TCEA1 is associated with the majority of class III genes independently of Pol II, strongly suggests that TFIIS is also a Pol III transcription factor in mammals. Intriguingly, its distribution on class III genes is skewed toward the TSS when compared to that of Pol III, in line

with a putative role in transcription initiation as we proposed in yeast.

Previous studies in *Drosophila* have shown that TFIIS plays a role in stimulating paused Pol II to enter active elongation (61,62), a situation that could also hold in mouse ES cells. We wondered if the presence or absence of TFIIS could be related to pausing on class II genes. TFIIS occupancy on class II genes followed that of Pol II both on elongating genes and on non-productive genes that have high levels of paused Pol II, indicating that the pause is not the consequence of the absence of TFIIS. Our observations thus raise the intriguing possibility that TFIIS activity, but not its recruitment, might be somehow stimulated upon the entry into elongation.

In summary, this study provides a high-resolution map of class III genes active in mouse ES cells and it shows that the TFIIIC transcription factor is present at around 2000 sites on the genome independently of Pol III and its other factors TFIIIB-α or -β, possibly playing a role in chromatin organization. Finally, we provide evidence that support a role for TFIIS Pol II elongation factor in class III transcription.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank I. Davidson for inspiring us to start this project. We thank P. Thuriaux and H. Neil-Bernet for critical reading of the manuscript. We thank T. Ye and A. Krebs for providing seqMINER before publication, M. de Dieuleveult, R. Fenouil, Y Duffour and N. Naouar for help with bioinformatics, I. Hmitou for help with some experiments, U. Rogner and A. Smith for the gift of the 46C ES cell line.

FUNDING

The Agence Nationale de la Recherche (ANR-05-BLAN-0396); the Association pour la Recherche sur le Cancer (3164); the Association Française contre les Myopathies (MNM2 2008-13630); the Ile-de-France Region (grant 2745 to L.C.); the Fondation pour la Recherche Médicale (grant 382-2010 to L.C.). Funding for open access charge: Commissariat à l'Energie Atomique et aux Energies Alternatives.

Conflict of interest statement. None declared.

REFERENCES

1. Paule,M.R. and White,R.J. (2000) Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res.*, **28**, 1283–1298.
2. Dieci,G., Fiorino,G., Castelnuovo,M., Teichmann,M. and Pagano,A. (2007) The expanding RNA polymerase III transcriptome. *Trends Genet.*, **23**, 614–622.
3. Kramerov,D.A. and Vassetzky,N.S. (2005) Short retroposons in eukaryotic genomes. *Int. Rev. Cytol.*, **247**, 165–221.
4. White,R.J. (1998) *RNA Polymerase III Transcription*, 2nd edn. Springer, Berlin.
5. Schramm,L. and Hernandez,N. (2002) Recruitment of RNA polymerase III to its target promoters. *Genes Dev.*, **16**, 2593–2620.
6. Geiduschek,E.P. and Kassavetis,G.A. (2001) The RNA polymerase III transcription apparatus. *J. Mol. Biol.*, **310**, 1–26.
7. Dumay-Odelot,H., Marck,C., Durrieu-Gaillard,S., Lefebvre,O., Jourdain,S., Prochazkova,M., Pflieger,A. and Teichmann,M. (2007) Identification, molecular cloning, and characterization of the sixth subunit of human transcription factor TFIIIC. *J. Biol. Chem.*, **282**, 17179–17189.
8. Teichmann,M. and Seifart,K.H. (1995) Physical separation of two different forms of human TFIIIB active in the transcription of the U6 or the VAI gene in vitro. *EMBO J.*, **14**, 5974–5983.
9. Brun,I., Sentenac,A. and Werner,M. (1997) Dual role of the C34 subunit of RNA polymerase III in transcription initiation. *EMBO J.*, **16**, 5730–5741.
10. Wang,Z. and Roeder,R.G. (1997) Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. *Genes Dev.*, **10**, 1315–1326.
11. Sadowski,C.L., Henry,R.W., Lobo,S.M. and Hernandez,N. (1993) Targeting TBP to a non-TATA box cis-regulatory element: a TBP-containing complex activates transcription from snRNA promoters through the PSE. *Genes Dev.*, **7**, 1535–1548.
12. Murphy,S., Yoon,J.B., Gerster,T. and Roeder,R.G. (1992) Oct-1 and Oct-2 potentiate functional interactions of a transcription factor with the proximal sequence element of small nuclear RNA genes. *Mol. Cell Biol.*, **12**, 3247–3261.
13. Waldschmidt,R., Wanandi,I. and Seifart,K.H. (1991) Identification of transcription factors required for the expression of mammalian U6 genes in vitro. *EMBO J.*, **10**, 2595–2603.
14. Teichmann,M., Wang,Z. and Roeder,R.G. (2000) A stable complex of a novel transcription factor IIB- related factor, human TFIIIB50, and associated proteins mediate selective transcription by RNA polymerase III of genes with upstream promoter elements. *Proc. Natl Acad. Sci. USA*, **97**, 14200–14205.
15. Schramm,L., Pendergrast,P.S., Sun,Y. and Hernandez,N. (2000) Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes Dev.*, **14**, 2650–2663.
16. Warner,J.R. (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.*, **24**, 437–440.
17. White,R.J. (2008) RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet.*, **24**, 622–629.
18. Marshall,L., Kenneth,N.S. and White,R.J. (2008) Elevated tRNA(iMet) synthesis can drive cell proliferation and oncogenic transformation. *Cell*, **133**, 78–89.
19. Wind,M. and Reines,D. (2000) Transcription elongation factor SII. *Bioessays*, **22**, 327–336.
20. Kim,B., Nesvizhskii,A.I., Rani,P.G., Hahn,S., Aebersold,R. and Ranish,J.A. (2007) The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. *Proc. Natl Acad. Sci. USA*, **104**, 16068–16073.
21. Guglielmi,B., Soutourina,J., Esnault,C. and Werner,M. (2007) TFIIS elongation factor and Mediator act in conjunction during transcription initiation in vivo. *Proc. Natl Acad. Sci. USA*, **104**, 16062–16067.
22. Ghavi-Helm,Y., Michaut,M., Acker,J., Aude,J.C., Thuriaux,P., Werner,M. and Soutourina,J. (2008) Genome-wide location analysis reveals a role of TFIIS in RNA polymerase III transcription. *Genes Dev.*, **22**, 1934–1947.
23. Ito,T., Seldin,M.F., Taketo,M.M., Kubo,T. and Natori,S. (2000) Gene structure and chromosome mapping of mouse transcription elongation factor S-II (Tcea1). *Gene*, **244**, 55–63.
24. Ito,T., Xu,Q., Takeuchi,H., Kubo,T. and Natori,S. (1996) Spermatocyte-specific expression of the gene for mouse testis-specific transcription elongation factor S-II. *FEBS Lett.*, **385**, 21–24.
25. Taira,Y., Kubo,T. and Natori,S. (1998) Molecular cloning of cDNA and tissue-specific expression of the gene for SII-K1, a novel transcription elongation factor SII. *Genes Cells*, **3**, 289–296.
26. Moqtaderi,Z. and Struhl,K. (2004) Genome-wide occupancy profile of the RNA polymerase III machinery in Saccharomyces

cerevisiae reveals loci with incomplete transcription complexes. *Mol. Cell Biol.*, **24**, 4118–4127.
27. Roberts,D.N., Stewart,A.J., Huff,J.T. and Cairns,B.R. (2003) The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc. Natl Acad. Sci. USA*, **100**, 14695–14700.
28. Harismendy,O., Gendrel,C.G., Soularue,P., Gidrol,X., Sentenac,A., Werner,M. and Lefebvre,O. (2003) Genome-wide location of yeast RNA polymerase III transcription machinery. *EMBO J.*, **22**, 4738–4747.
29. Barski,A., Chepelev,I., Liko,D., Cuddapah,S., Fleming,A.B., Birch,J., Cui,K., White,R.J. and Zhao,K. (2010) Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat. Struct. Mol. Biol.*, **17**, 629–634.
30. Oler,A.J., Alla,R.K., Roberts,D.N., Wong,A., Hollenhorst,P.C., Chandler,K.J., Cassiday,P.A., Nelson,C.A., Hagedorn,C.H., Graves,B.J. *et al.* (2010) Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat. Struct. Mol. Biol.*, **17**, 620–628.
31. Moqtaderi,Z., Wang,J., Raha,D., White,R.J., Snyder,M., Weng,Z. and Struhl,K. (2010) Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat. Struct. Mol. Biol.*, **17**, 635–640.
32. Raha,D., Wang,Z., Moqtaderi,Z., Wu,L., Zhong,G., Gerstein,M., Struhl,K. and Snyder,M. (2010) Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc. Natl Acad. Sci. USA*, **107**, 3639–3644.
33. Canella,D., Praz,V., Reina,J.H., Cousin,P. and Hernandez,N. (2010) Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res.*, **20**, 710–721.
34. Guffanti,E., Percudani,R., Harismendy,O., Soutourina,J., Werner,M., Iacovella,M.G., Negri,R. and Dieci,G. (2006) Nucleosome depletion activates poised RNA polymerase III at unconventional transcription sites in Saccharomyces cerevisiae. *J. Biol. Chem.*, **281**, 29155–29164.
35. Noma,K., Cam,H.P., Maraia,R.J. and Grewal,S.I. (2006) A role for TFIIIC transcription factor complex in genome organization. *Cell*, **125**, 859–872.
36. Wallace,J.A. and Felsenfeld,G. (2007) We gather together: insulators and genome organization. *Curr. Opin. Genet. Dev.*, **17**, 400–407.
37. Liu,P., Jenkins,N.A. and Copeland,N.G. (2003) A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.*, **13**, 476–484.
38. Ying,Q.L., Stavridis,M., Griffiths,D., Li,M. and Smith,A. (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotech.*, **21**, 183–186.
39. Tessarollo,L. (2001) Manipulating mouse embryonic stem cells. *Methods Mol. Biol.*, **158**, 47–63.
40. Esnault,C., Ghavi-Helm,Y., Brun,S., Soutourina,J., Van Berkum,N., Boschiero,C., Holstege,F. and Werner,M. (2008) Mediator-dependent recruitment of TFIIH modules in preinitiation complex. *Mol. Cell*, **31**, 337–346.
41. Kuras,L., Borggrefe,T. and Kornberg,R.D. (2003) Association of the Mediator complex with enhancers of active genes. *Proc. Natl Acad. Sci. USA*, **100**, 13887–13891.
42. Valouev,A., Johnson,D.S., Sundquist,A., Medina,C., Anton,E., Batzoglou,S., Myers,R.M. and Sidow,A. (2008) Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat. Methods*, **5**, 829–834.
43. Coughlin,D.J., Babak,T., Nihranz,C., Hughes,T.R. and Engelke,D.R. (2009) Prediction and verification of mouse tRNA gene families. *RNA Biol.*, **6**, 195–202.
44. Lowe,T.M. and Eddy,S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, **25**, 955–964.
45. Kagey,M.H., Newman,J.J., Bilodeau,S., Zhan,Y., Orlando,D.A., van Berkum,N.L., Ebmeier,C.C., Goossens,J., Rahl,P.B., Levine,S.S. *et al.* (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**, 430–435.
46. Meissner,A., Mikkelsen,T.S., Gu,H., Wernig,M., Hanna,J., Sivachenko,A., Zhang,X., Bernstein,B.E., Nusbaum,C., Jaffe,D.B. *et al.* (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, **454**, 766–770.
47. Mikkelsen,T.S., Ku,M., Jaffe,D.B., Issac,B., Lieberman,E., Giannoukos,G., Alvarez,P., Brockman,W., Kim,T.K., Koche,R.P. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553–560.
48. Marson,A., Levine,S.S., Cole,M.F., Frampton,G.M., Brambrink,T., Johnstone,S., Guenther,M.G., Johnston,W.K., Wernig,M., Newman,J. *et al.* (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*, **134**, 521–533.
49. Bilodeau,S., Kagey,M.H., Frampton,G.M., Rahl,P.B. and Young,R.A. (2009) SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.*, **23**, 2484–2489.
50. Donze,D. and Kamakaka,R.T. (2001) RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. *EMBO J.*, **20**, 520–531.
51. Valenzuela,L., Dhillon,N. and Kamakaka,R.T. (2009) Transcription independent insulation at TFIIIC-dependent insulators. *Genetics*, **183**, 131–148.
52. Parelho,V., Hadjur,S., Spivakov,M., Leleu,M., Sauer,S., Gregson,H.C., Jarmuz,A., Canzonetta,C., Webster,Z., Nesterova,T. *et al.* (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, **132**, 422–433.
53. Chen,X., Xu,H., Yuan,P., Fang,F., Huss,M., Vega,V.B., Wong,E., Orlov,Y.L., Zhang,W., Jiang,J. *et al.* (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, **133**, 1106–1117.
54. Rahl,P.B., Lin,C.Y., Seila,A.C., Flynn,R.A., McCuine,S., Burge,C.B., Sharp,P.A. and Young,R.A. (2010) c-Myc regulates transcriptional pause release. *Cell*, **141**, 432–445.
55. Borchert,G.M., Lanier,W. and Davidson,B.L. (2006) RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.*, **13**, 1097–1101.
56. Ozsolak,F., Poling,L.L., Wang,Z., Liu,H., Liu,X.S., Roeder,R.G., Zhang,X., Song,J.S. and Fisher,D.E. (2008) Chromatin structure analyses identify miRNA promoters. *Genes Dev.*, **22**, 3172–3183.
57. Bortolin-Cavaille,M.L., Dance,M., Weber,M. and Cavaille,J. (2009) C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts. *Nucleic Acids Res.*, **37**, 3464–3473.
58. Lobo,S.M. and Hernandez,N. (1989) A 7 bp mutation converts a human RNA polymerase II snRNA promoter into an RNA polymerase III promoter. *Cell*, **58**, 55–67.
59. Mattaj,I.W., Dathan,N.A., Parry,H.D., Carbon,P. and Krol,A. (1988) Changing the RNA polymerase specificity of U snRNA gene promoters. *Cell*, **55**, 435–442.
60. Listerman,I., Bledau,A.S., Grishina,I. and Neugebauer,K.M. (2007) Extragenic accumulation of RNA polymerase II enhances transcription by RNA polymerase III. *PLoS Genet.*, **3**, e212.
61. Nechaev,S., Fargo,D.C., dos Santos,G., Liu,L., Gao,Y. and Adelman,K. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. *Science*, **327**, 335–338.
62. Adelman,K., Marr,M.T., Werner,J., Saunders,A., Ni,Z., Andrulis,E.D. and Lis,J.T. (2005) Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. *Mol. Cell*, **17**, 103–112.
63. Ye,T., Krebs,A.R., Choukrallah,M.A., Keime,C., Plewniak,F., Davidson,I. and Tora,L. (2010) seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.*, **39**, e35.
64. Seila,A.C., Calabrese,J.M., Levine,S.S., Yeo,G.W., Rahl,P.B., Flynn,R.A., Young,R.A. and Sharp,P.A. (2008) Divergent transcription from active promoters. *Science*, **322**, 1849–1851.