nature structural & molecular biology

Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells

Zarmik Moqtaderi[1], Jie Wang[2], Debasish Raha[3], Robert J White[4], Michael Snyder[5], Zhiping Weng[2] & Kevin Struhl[1]

Genome-wide occupancy profiles of five components of the RNA polymerase III (Pol III) machinery in human cells identified the expected tRNA and noncoding RNA targets and revealed many additional Pol III–associated loci, mostly near short interspersed elements (SINEs). Several genes are targets of an alternative transcription factor IIIB (TFIIIB) containing Brf2 instead of Brf1 and have extremely low levels of TFIIIC. Strikingly, expressed Pol III genes, unlike nonexpressed Pol III genes, are situated in regions with a pattern of histone modifications associated with functional Pol II promoters. TFIIIC alone associates with numerous *ETC* loci, via the B box or a novel motif. *ETC*s are often near CTCF binding sites, suggesting a potential role in chromosome organization. Our results suggest that human Pol III complexes associate preferentially with regions near functional Pol II promoters and that TFIIIC-mediated recruitment of TFIIIB is regulated in a locus-specific manner.

© 2010 Nature America, Inc. All rights reserved.

RNA polymerase III (Pol III) is responsible for the synthesis of tRNAs and other non–protein-coding RNAs (ncRNAs) in eukaryotes[1–3]. Metazoan Pol III genes are classified into three types, each of which is transcribed by Pol III in concert with other multisubunit general transcription factors. At tRNAs and other Pol III genes categorized as type 2, the recognition factor TFIIIC binds to DNA sequences termed A and B boxes that are typically contained entirely within the body of the structural gene. TFIIIC recruits the initiation factor TFIIIB, which consists of the TATA-binding protein (TBP), Bdp1 and Brf1 subunits. These proteins recruit the polymerase to begin transcription. At the type 1 gene encoding the 5S rRNA, transcription involves a dedicated transcription factor, TFIIIA, which recognizes special elements within the gene and recruits TFIIIC, TFIIIB and Pol III.

Other Pol III–transcribed genes termed type 3 have upstream, gene-external promoters, much like Pol II–transcribed genes. At these loci, a proximal sequence element (PSE) is bound by the SNAPc complex, which is involved in transcription of small nuclear RNAs by either Pol II or Pol III. At Pol III promoters bound by SNAPc, TATA-bound TBP specifies Pol III transcription by recruiting the Brf2 subunit into an alternative TFIIIB comprising TBP, Bdp1 and Brf2, followed by Pol III recruitment[1–3]. *In vitro*, Brf2-dependent Pol III genes can be transcribed in the absence of Brf1 or TFIIIC[3].

Whole-genome chromatin immunoprecipitation (ChIP) studies using microarrays have been performed on subunits of each of these general transcription factors in yeast. In *Saccharomyces cerevisiae*, complete Pol III complexes are found at tRNA genes and at loci encoding a variety of ncRNAs[4–6]. At these regions, the general transcription factors TFIIIC, TFIIIB, and Pol III are present at a fairly constant ratio, suggesting that the amount of TFIIIC binding to promoters is predictive of the level of functional Pol III transcription complexes *in vivo*. However, whereas TFIIIC-mediated recruitment of TFIIIB shows little, if any, promoter specificity, TFIIIB recruitment is generally repressed by Maf1 in response to environmental signals[7–9].

In contrast to conventional Pol III–transcribed genes, eight loci with incomplete transcription complexes consisting only of TFIIIC were identified in *S. cerevisiae* and termed *ETC*, for 'extra TFIIIC'[5]. *ETC* loci are characterized by a sequence motif that contains the B box and an additional three conserved nucleotides, suggesting that TFIIIC's association with this special motif influences its potential to recruit the remainder of the Pol III machinery. The *ETC* loci are remarkably well conserved both structurally and functionally across closely related yeast species[5], indicating that these regions are biologically important. Similar loci were later identified in *Schizosaccharomyces pombe*, where they are more numerous and named 'chromosome organizing clamps' (*COC*s) due to their apparent role in higher-order chromosome organization[10]. It is unknown whether *ETC* loci are present in metazoans.

Here, by analyzing subunits of the three major Pol III transcription factors, we examine the occupancy profile of the Pol III transcription machinery on a genome-wide scale. We identify new Pol III targets, distinct classes of *ETC* loci and distinct binding profiles of Brf1- and Brf2-containing isoforms of TFIIIB. Further, we find binding of the Pol III machinery to target promoters to be strongly linked with (and perhaps dependent on) a nearby functional Pol II promoter and associated histone modifications.

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts, USA. [2]Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [3]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut, USA. [4]Beatson Institute for Cancer Research, Glasgow, UK. [5]Department of Genetics, Stanford University School of Medicine, Stanford, California, USA. Correspondence should be addressed to K.S. (kevin@hms.harvard.edu).

Received 18 December 2009; accepted 25 February 2010; published online 25 April 2010; doi:10.1038/nsmb.1794

RESULTS

Genome-wide occupancy of five Pol III transcription factors

We performed ChIP followed by massively parallel sequencing (ChIP-seq) on three independent biological replicates of K562 cells, using antibodies against five proteins involved in Pol III transcription. These antibodies correspond to subunits of the promoter recognition factor TFIIIC (TFIIIC-110 subunit), the initiation factor, TFIIIB (Bdp1 and Brf1 subunits) and the polymerase itself (Rpc155 subunit). In addition, we used an antibody against Brf2, which is involved in the transcription of Pol III–transcribed genes with gene-external promoter architecture.

The overall quality and reproducibility of the data were high, with correlation coefficients between replicates typically ranging between 0.79 and 0.95. (**Supplementary Fig. 1**). The Brf2 data were lower in overall quality and reproducibility (correlation coefficients between 0.41 and 0.68), but previously known Brf2-bound loci were easily identifiable using high-stringency thresholds. Peak lists for the five proteins are shown in **Supplementary Data 1**.

Expected and novel targets of RNA polymerase III

The Rpc155 antibody yielded by far the most robust ChIP signals of the antibodies. Using a stringent threshold, we identified 1,520 Pol III targets, which we then used as the reference list when comparing occupancies by the Pol III transcription factors. To determine the occupancy levels of all five proteins at each Rpc155-associated locus, we counted the sequence reads ('tags') for every protein within a 300–base pair (bp) window centered on the coordinate with maximal Rpc155 occupancy. By determining the relative occupancy of Pol III factors at each genomic region, we obtained information about the nature of the transcription complexes at the Pol III promoters in a manner analogous to that used for Pol II preinitiation complexes[11].

The general transcription factors show a high degree of co-occupancy at Pol III loci (**Fig. 1a–d**). The TFIIIB components Brf1 and Bdp1 were each highly correlated ($R^2 = 0.77$ and 0.80, respectively) with Rpc155 at Pol III loci, validating our Rpc155-associated loci as meaningful targets of the complete Pol III machinery. Fittingly for two subunits of the same factor, Bdp1 and Brf1 showed a correlation ($R^2 = 0.89$) comparable to that between replicate immunoprecipitations of the same protein. At Rpc155 targets, TFIIIC was markedly less well correlated ($R^2 = 0.56$) with polymerase than was TFIIIB. As will be discussed further below, this suggests that the degree of TFIIIC occupancy is not strictly related to the level of Pol III transcription.

We plotted the relative position of each protein with respect to transcription start sites (TSSs) of Pol III genes (**Fig. 1e**,**f**). The peak of Rpc155 covers a region spanning from 5′ of the TSS into the structural gene. TFIIIC occupancy peaks downstream of the TSS and within the gene, reflecting the gene-internal position of the DNA sequences recognized by this factor. The TFIIIB components Brf1, Brf2 and Bdp1 have occupancy peaks upstream of the TSS, in accord with their role in initiation. Thus, the mapping afforded by the sequencing data is of sufficient resolution to distinguish closely spaced binding sites even within small Pol III genes.

As expected[1], the complete Pol III machinery is found at genes encoding tRNAs, 5S rRNA and U6, hY, 7SK and RNase P RNA (**Fig. 2a**), representing all three types of Pol III promoters. In general, TFIIIC, TFIIIB and Pol III association with these loci is robust (>100-fold enrichment of Pol III in many cases). Of the 513 tRNA genes, 392 are targets of Rpc155 by our stringent peak-calling criteria. In addition, we detected Pol III occupancy at 41 out of the 172 tRNA pseudogenes in the human genome. Because of the clustering of many tRNA genes, some Rpc155 peaks contain more than one tRNA target. The percentage of expressed genes for each type of tRNA is variable, with, for example, almost all cysteine tRNA genes but fewer than half of glutamate tRNA genes being occupied by the polymerase (**Fig. 2b**).

In addition to the expected targets, many additional loci are associated with Pol III (**Supplementary Data 2**). These include NF90-associated RNAs encoded by *snaR* loci, which possess B boxes and Pol III termination sequences and are transcribed by Pol III *in vitro*[12]. We do not observe Pol III at any annotated miRNA genes (except for *miR-1975* and *miR-886*, which are unlikely to be relevant as they overlap completely with hY5 and vault RNAs that are known Pol III targets); this includes miRNA genes clustered on chromosome 19 that have been the subject of conflicting reports[13,14]. Small but reproducible amounts of Pol III are associated with *U91*, which overlaps the ncRNA *SCARNA18*, and with *U13* snoRNA. A striking majority (90%) of the otherwise nonannotated Pol III–associated loci are near SINE

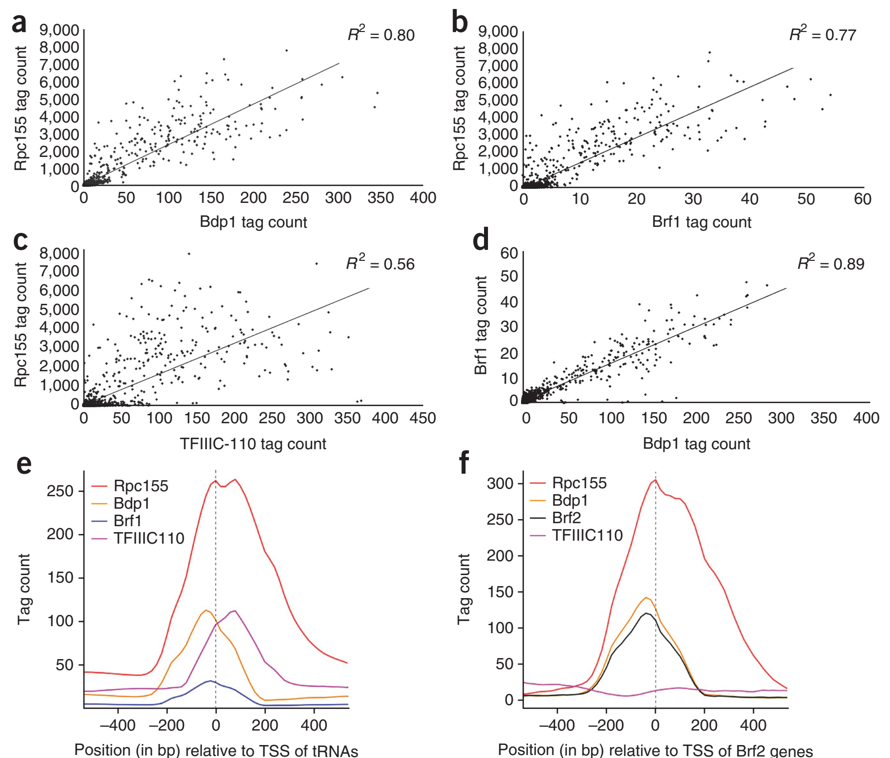

Figure 1 Pairwise comparison of occupancy by proteins at Pol III loci. Each plot compares the average number of sequence reads (tag count) for the indicated proteins within the 300-bp window centered at each Rpc155 summit. (**a**) Rpc155 versus Bdp1. (**b**) Rpc155 versus Brf1. (**c**) Rpc155 versus TFIIIC-110. (**d**) Brf1 versus Bdp1. (**e**) Distribution of Bdp1, Brf1, TFIIIC-110 and Rpc155 association with respect to transcription start sites (TSS) of tRNAs. (**f**) Distribution of Bdp1, Brf2, TFIIIC-110 and Rpc155 association with respect to TSS of Brf2-associated genes.

© 2010 Nature America, Inc. All rights reserved.

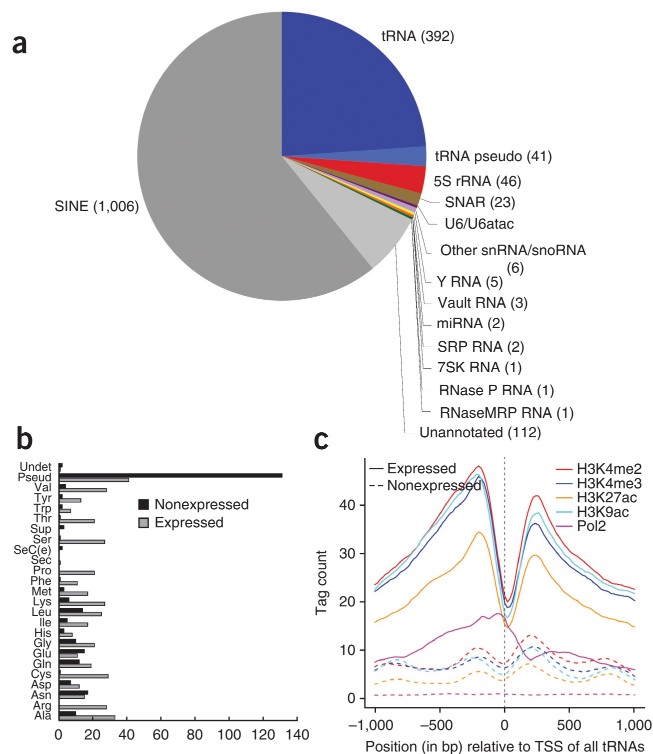

Figure 2 Properties of Pol III targets. (**a**) Pie chart depicting the categories of Pol III–occupied loci. (**b**) Occupancy of tRNA genes by isotype. (**c**) Positions of histone modifications and Pol II occupancy relative to expressed (that is, Pol III–occupied, solid lines) or nonexpressed (dashed lines) tRNA genes.

elements. In general, these SINE-linked loci have considerably lower levels of Pol III association than observed at tRNA genes.

Active Pol III genes are near Pol II–dependent histone marks

Here, we define expressed and nonexpressed tRNAs by virtue of Pol III association. Although it is formally possible that Pol III association might not result in RNA synthesis, this is unlikely because tRNA levels are very high. We analyzed the proximity of expressed (that is, Pol III–occupied) and nonexpressed tRNAs with regard to nearby histone modifications. There is a marked dip in all histone marks corresponding exactly to the TSS of expressed tRNAs, suggesting a strong and localized nucleosome depletion over the TSS (**Fig. 2c**, solid lines). This apparent nucleosome depletion is similar to that observed at tRNA genes in *S. cerevisiae*[15,16], and it is likely due to nucleosome exclusion by the extremely high levels of transcription at Pol III genes.

Although Pol III peaks are often in proximity to Pol II peaks[17], the functional relationship of this observation is unclear. Strikingly, for the 392 expressed Pol III genes, there is a high correlation with a pattern of histone marks typical of Pol II TSS regions and with occupancy by Pol II itself. This proximity is directional, as Pol II occupancy peaks 5′ of expressed Pol III genes. In contrast, we observed no such pattern for the nonexpressed subset of tRNA genes (**Fig. 2c**, dashed lines). Together, these results suggest that the histone modification pattern at Pol II TSS regions is an important determinant of Pol III expression.

Brf2 sites lack Brf1 and are very low in TFIIIC

Brf2 occupies U6 and several other loci with TATA-containing promoters located upstream of the sequences encoding mature RNA. *In vitro*, Brf2 is recruited by TBP in a TFIIIC-independent manner to these promoters, forming a complex containing Bdp1 but lacking Brf1[3]. In accord with this biochemical observation, Brf1 and Brf2 occupancies are uncorrelated at Pol III–occupied loci (**Fig. 3a**). Furthermore, Brf1 occupancy at Brf2 sites is very low or absent (**Fig. 3b**), indicating that Brf1 is not required for transcription of Brf2-dependent genes *in vivo*. TFIIIC association is also very low at Brf2 targets, although it is not entirely absent (~2% of the level observed at Brf1-associated loci), suggesting that TFIIIC is unnecessary for Pol III recruitment to Brf2 loci *in vivo* (**Fig. 3b**). Notably, the Brf2-occupied locus with the highest TFIIIC binding (though still only a small fraction of the TFIIIC level at Brf1 targets) is *tRNASec*, which in *Xenopus laevis* uses both a gene-internal B box and the gene-external promoter structure typical of type 3 genes[18].

Numerous *ETC* loci in human cells

TFIIIC peaks vastly outnumber the peaks for the other transcription factors; our TFIIIC target list consists of 5,474 loci, in contrast with 1,520 at Rpc155, even though fold enrichments, and hence assay sensitivity, are higher for Rpc155. Unlike TFIIIB loci, TFIIIC targets are highly variable in the extent to which they are occupied by the other components of the Pol III machinery (**Fig. 4a**), with only a limited correlation ($R^2 = 0.37$) between Rpc155 and TFIIIC at TFIIIC targets. To identify human *ETC* loci, we first restricted our TFIIIC peak list to an even higher level of stringency (see Online Methods). As the median ratio of TFIIIC to Rpc155 at tRNA genes is 0.06, we defined an *ETC* locus as having a TFIIIC/Rpc155 ratio higher than 2.04, 3 s.d. above this median ratio. *ETC*s by this definition do not need to be completely lacking in Pol III occupancy. We validated

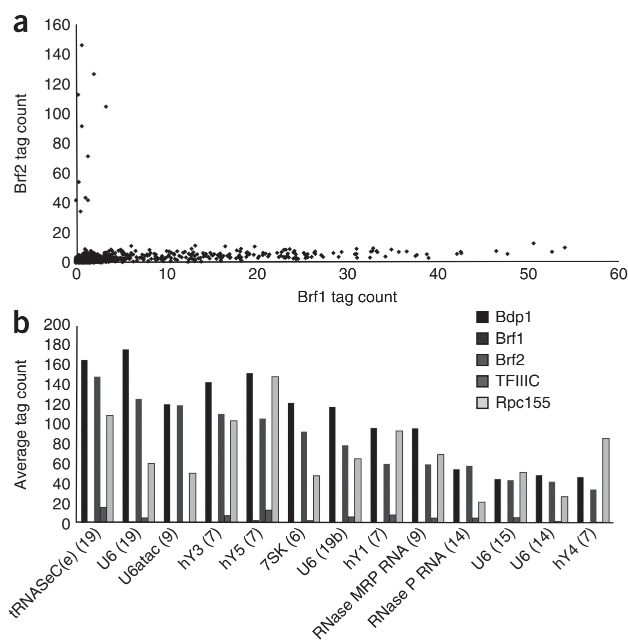

Figure 3 Distinctions between TFIIIB isoforms. (**a**) Pairwise comparison of Brf1 and Brf2 occupancy at Rpc155 peaks. The average number of sequence reads for Brf1 and Brf2 are compared for the 300-bp window centered at each Rpc155 summit. (**b**) Tag counts for all five proteins at Brf2 targets. Tag counts are shown for each protein within a 300-bp window centered at the summit of every indicated Brf2 peak (the chromosome number of each locus is indicated in parentheses). We show only those peaks with above-background levels of Rpc155. Rpc155 tag counts were divided by 25 to allow use of a single scale for viewing all proteins.

© 2010 Nature America, Inc. All rights reserved.

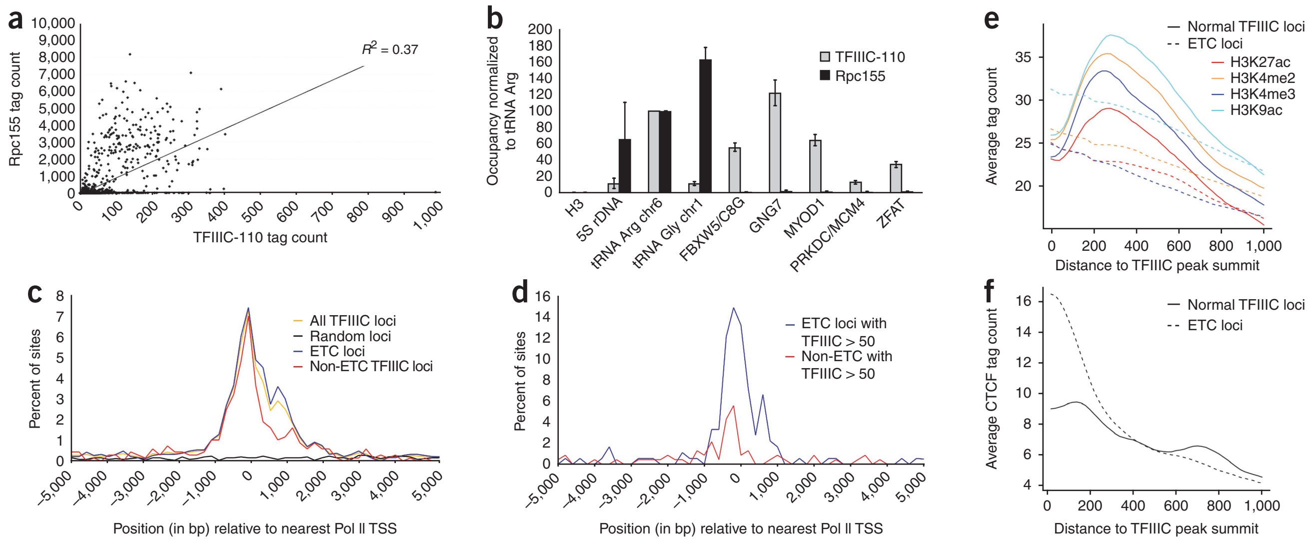

Figure 4 Relationship of TFIIIC occupancy to Pol III occupancy and identification of *ETC* loci. (**a**) Correlation of Rpc155 and TFIIIC-110 occupancy at TFIIIC-110 peaks. Comparison of the number of Rpc155 and TFIIIC-110 sequence reads within the 300-bp window centered on each TFIIIC-110 peak summit. (**b**) Pol III and TFIIIC-100 association (±s.d.) at typical (left) and *ETC* loci (right) by quantitative PCR in real time. (**c**) *ETC* loci are near the TSS of Pol II genes. Shown is the frequency distribution of TFIIIC-110 peak summits with respect to the nearest Pol II TSS. (**d**) Frequency distribution of *ETC* and non-*ETC* TFIIIC-110 peaks with the highest TFIIIC levels (average tag counts > 50 in the 300-bp window around the peak summit) with respect to Pol II TSS. (**e**) Positions of histone modifications relative to TFIIIC-110 peak summits at non-*ETC* (solid lines) or *ETC* (dashed lines). (**f**) CTCF tag count as a function of distance from TFIIIC-110 peak summits for non-*ETC* (solid lines) or *ETC* (dashed lines) loci.

© 2010 Nature America, Inc. All rights reserved.

binding of Rpc155 and TFIIIC at several normal and *ETC* targets using quantitative PCR analysis in real time (**Fig. 4b**). Under these strict criteria (only 3 tRNAs pass), we identified 1,865 *ETC*s (listed in **Supplementary Data 3**) and suspect that there are several thousand in the genome.

*ETC*s are often near Pol II genes and CTCF-bound regions

The distribution of TFIIIC-occupied loci of both the *ETC* and the non-*ETC* types reveals a positional bias toward the TSSs of Pol II genes (**Fig. 4c**). *ETC* loci show a wide dynamic range in the amount of TFIIIC binding, with some sites emerging as especially pronounced *ETC*s. The 181 *ETC* loci with the highest levels of TFIIIC occupancy (>50 sequence reads per site) are strikingly well correlated with the TSSs of Pol II genes, with 68% being located within 1 kilobase of a Pol II TSS (**Fig. 4d**). This is reminiscent of the *S. cerevisiae ETC*s, which are ~200–300 bp upstream of a neighboring Pol II–transcribed gene. Gene ontology categories of Pol II genes adjacent to *ETC* loci show significant ($P = 1.9 \times 10^{-9}$) overrepresentation of nuclear-localized proteins. In addition, *ETC*s are significantly overrepresented in regions upstream of closely spaced, divergently transcribed pairs of Pol II genes. Of the 1,431 shared upstream sequences for Pol II genes (defined as having TSSs less than 1 kilobase apart), 86 harbor *ETC* loci ($P < 10^{-300}$).

As expected from the proximity of strong *ETC*s to mRNA initiation sites, the *ETC* population shows histone modifications associated with functional Pol II promoters, with the highest levels of these modifications near the TFIIIC peaks (**Fig. 4e**). This pattern of histone modifications at *ETC* loci (dashed lines) differs from the pattern at expressed Pol III genes (solid lines), in which histone modifications appear very low near the promoter due to nucleosome loss accompanying high levels of Pol III transcription. Lastly, in accord with the observation that TFIIIC targets in yeast play a role in genome organization[10,19,20] and act as transcriptional insulator elements[21–23], TFIIIC association at *ETC* loci is highly correlated with association of CTCF (**Fig. 4f**), a protein that interacts with cohesins and is involved in insulation, looping and chromosome conformation[24,25]. This observation suggests that TFIIIC may play a role in chromosome organization in humans.

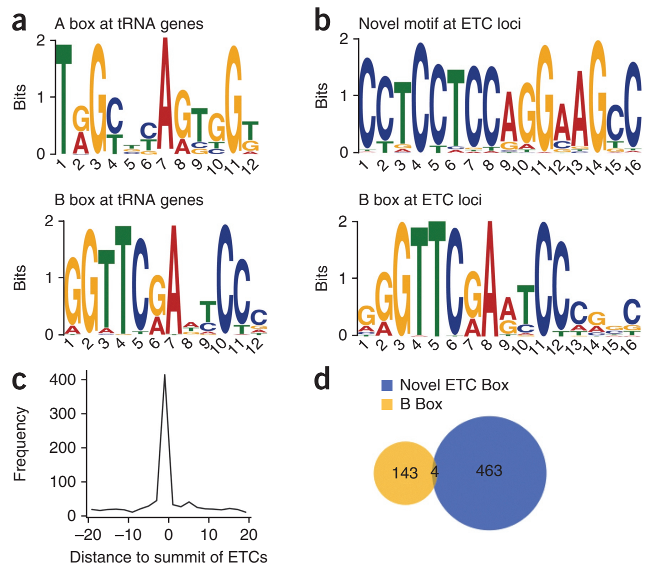

Figure 5 *ETC* motifs. (**a**) Known A- and B-box motifs found by *de novo* motif searching of tRNA genes. (**b**) Motifs found by *de novo* searching of *ETC* loci are a novel motif (top) and the B-box motif (bottom). (**c**) Frequency plot of the novel *ETC* motif position as a function of distance from the TFIIIC-110 peak summit. (**d**) Venn diagram illustrating the number of *ETC*s possessing a B box (yellow) or the *ETC* motif (blue) with motif scores $P < 10^{-5}$. Only four *ETC* loci have both motifs.

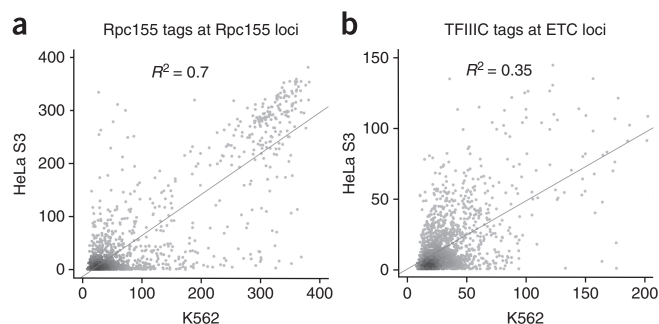

Figure 6 Cell specificity of *ETC* loci. (**a**) Correlation of Rpc155 tag counts in K562 and HeLa cells at Rpc155-occupied loci. (**b**) Correlation of TFIIIC-110 tag counts in K562 and HeLa cells at *ETC* loci.

A novel sequence motif at a subset of *ETC* loci

De novo motif searching revealed that, in contrast with tRNAs, which have A and B boxes, the *ETC* loci are characterized by one of two significantly overrepresented sequence motifs: the standard B box or a novel motif that may be loosely related to a motif for Ets transcription factors (**Fig. 5a**,**b**). The *ETC* motif coincides with the summit of TFIIIC peaks at many *ETCs* (**Fig. 5c**). *ETC* and B-box motifs are mostly mutually exclusive, with only four *ETC* loci having both motifs (**Fig. 5d**). The existence of an *ETC*-specific motif is analogous to the situation in *S. cerevisiae*, where *ETC* loci possess a novel motif extending the canonical B box. In both organisms, TFIIIC could associate with these *ETC*-specific motifs in a manner structurally unconducive to the assembly of complete Pol III complexes.

ETC loci show cell-type specificity

To determine whether Pol III occupancy varies with cell type, we examined the genome-wide association of Pol III factors in a second cell line, HeLa S3. Notably, whereas binding of Pol III factors at tRNA loci is highly correlated between cell lines (**Fig. 6a**), there are differences in TFIIIC occupancy at *ETC* loci between HeLa S3 and K562 cells (**Fig. 6b**). This observation suggests that TFIIIC binding and/or TFIIIB recruitment at *ETC* loci, but not tRNA genes, is influenced by cell type–specific factors.

DISCUSSION

Genome-wide profiling reveals ~1,500 Pol III–associated loci in K562 cells. In general, the expected 400–500 Pol III–transcribed genes encoding tRNAs and ncRNAs show very high levels of all three Pol III factors. In contrast, ~1,000 Pol III–associated loci have not been previously described, and the vast majority of these (90%) are located near SINEs. The SINE-associated loci show much lower levels of Pol III factors compared to tRNA genes, although these levels are clearly above the background. The transcriptional products and biological functions of these Pol III–associated loci near SINEs are unknown. The prevalence of SINEs near these Pol III–associated loci might reflect DNA sequences resembling B or A blocks in a subset of SINEs.

Most, but not all, tRNA genes are occupied by the complete Pol III transcription machinery, and Pol III association at expressed versus nonexpressed genes differs by a factor of 100. Strikingly, the genomic regions in the vicinity of expressed and nonexpressed Pol III genes are different. Expressed Pol III genes are located close to regions that have histone modifications characteristic of functional Pol II promoters and Pol II itself. In contrast, this distinctive chromatin signature is absent from nonexpressed Pol III genes. These observations suggest that Pol III factors bind preferentially to genomic regions with a histone modification pattern generated by a functional Pol II promoter. We think it unlikely that nonexpressed Pol III genes have defective TFIIIC-recognition sites, because they generally possess high-quality B box sequences. Conversely, TFIIIC associates with <2% of the B-box motifs in the human genome, indicating that the presence of a B box alone is insufficient for binding *in vivo*. In addition, the fact that *ETC* loci are located near Pol II promoters and associated histone modification indicates that TFIIIC association depends on Pol II–generated chromatin regions.

The characteristic pattern of histone modifications at Pol II promoters depends on Pol II preinitiation complexes but not on extensive elongation, because Pol II is often paused just downstream of many promoters in a manner that precludes any appreciable transcription[26,27]. This suggests that TFIIIC binding to regions near functional Pol II promoters is largely, and perhaps completely, independent of Pol II transcriptional activity (that is, mRNA synthesis). In accord with this suggestion, treatment of human cells with α-amanitin, an inhibitor that blocks Pol II transcription after preinitiation-complex formation, has limited effects on Pol III transcription[17]. For these reasons, we speculate that TFIIIC binding is enhanced by the chromatin structure generated by nearby functional Pol II promoters. In principle, TFIIIC binding might be increased by promoter accessibility due to histone acetylation reducing histone-DNA contacts, direct interactions with modified histones via an effector domain(s) in a TFIIIC subunit or associated protein, or a histone variant (for example, H2AZ) near Pol II promoter regions.

However, accessibility to the recognition factor TFIIIC is not the only determinant of Pol III association in human cells. In *S. cerevisiae*, TFIIIC levels are highly correlated with Pol III occupancy, and recruitment of the complete Pol III machinery shows a one-to-one correspondence with TFIIIC binding. In contrast, the human TFIIIC/Pol III ratio is considerably more variable than it is in yeast, and a wide range of TFIIIC levels may recruit a given amount of polymerase. The presence of a Brf2 mechanism for Pol III recruitment in human cells is one obvious alternative pathway, allowing transcription with little or no TFIIIC. At the opposite extreme are the *ETC* loci, which recruit little or no polymerase despite having high occupancy by TFIIIC. Even at standard type 2 Pol III genes, the ratio of TFIIIC to polymerase varies over a wide range, in contrast to the TFIIIB/Pol III ratio, which is relatively consistent.

The different patterns of histone marks near *ETC*-type and transcriptionally active TFIIIC sites suggest that TFIIIC's ability to recruit TFIIIB to type 2 genes might depend on a particular histone-modification pattern. Indeed, myc-dependent activation of Pol III transcription is associated with targeted histone acetylation and increased association of TFIIIB[28]. In a similar vein, transcription of type 3 genes is influenced by the CHD8 chromatin-modifying protein[29]. Notably, TFIIIC either possesses histone acetylase activity[30,31] and/or recruits the p300 histone acetylase[32] such that it could actively participate in generating a chromatin state appropriate for TFIIIB recruitment. Association of the initiator TFIIIB is an important rate-limiting step in Pol III gene expression in humans, because the excellent correlation of TFIIIB and polymerase levels suggests that Pol III occupancy follows linearly from TFIIIB recruitment. In this regard, TFIIIB is a target of regulation Maf1 (ref. 33), Rb[34], p53 (ref. 35), Erk[36] and myc[28].

The presence of TFIIIC at many loci without TFIIIB or Pol III, much like similar loci in yeast, suggests a role for human TFIIIC beyond its function in Pol III transcription. TFIIIC bound to the intergenic region of closely spaced, divergently transcribed Pol II genes might

© 2010 Nature America, Inc. All rights reserved.

be important for the regulation and separate expression of the two genes. In yeast, TFIIIC-bound loci function as heterochromatin barriers and insulators[10,21,22] and also participate in higher-order chromosome organization[10,37]. The correlation of TFIIIC peaks with CTCF is noteworthy and suggests the possible involvement of human TFIIIC in chromosome organization.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Accession codes. The DNA sequencing datasets described in this work are available at http://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg18&g=wgEncodeYaleChIPseq

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank T. Liu and X. Liu for initial comments about the sequencing data, J. Mellor for helpful advice on programming in Python, N. Lamarre-Vincent, X. Fan, H. Hirsch and M. Lindahl Allen for discussion, G. Euskirchen, H. Monahan, M. Shi and P. Lacroute for help with DNA sequencing and M. Wilson for help with database submission. Funding for this work was provided by US National Institutes of Health grants GM 30186 (K.S.), HG 4558 (M.S. and K.S.) and HG 4695 (Z.W.).

AUTHOR CONTRIBUTIONS

Z.M. designed the experiments, performed and analyzed the ChIP and validation experiments, performed bioinformatics analysis, constructed figures and wrote the paper; J.W. performed bioinformatics analysis, discussed the results and constructed figures; D.R. supervised the library construction and Illumina sequencing of the immunoprecipitated DNA; R.J.W. provided the antibodies used in the ChIPs, discussed the results and suggested modifications to the text; M.S. supervised the library construction and Illumina sequencing of the immunoprecipitated DNA; Z.W. supervised the bioinformatics analysis and participated in substantial discussion of the bioinformatics and figures; K.S. participated in substantial discussion of all experiments, results and figures and contributed to the text.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

© 2010 Nature America, Inc. All rights reserved.

1. Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M. & Pagano, A. The expanding RNA polymerase III transcriptome. *Trends Genet.* **23**, 614–622 (2007).
2. Geiduschek, E.P. & Kassavetis, G.A. The RNA polymerase III transcription apparatus. *J. Mol. Biol.* **310**, 1–26 (2001).
3. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev.* **16**, 2593–2620 (2002).
4. Harismendy, O. *et al.* Genome-wide location of yeast RNA polymerase III transcription machinery. *EMBO J.* **22**, 4738–4747 (2003).
5. Moqtaderi, Z. & Struhl, K. Genome-wide occupancy of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol. Cell. Biol.* **24**, 4118–4127 (2004).
6. Roberts, D.N., Stewart, A.J., Huff, J.T. & Cairns, B.R. The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc. Natl. Acad. Sci. USA* **100**, 14695–14700 (2003).
7. Oficjalska-Pham, D. *et al.* General repression of RNA polymerase III transcription is triggered by protein phosphatase type 2A-mediated dephosphorylation of Maf1. *Mol. Cell* **22**, 623–632 (2006).
8. Roberts, D.N., Wilson, B., Huff, J.T., Stewart, A.J. & Cairns, B.R. Dephosphorylation and genome-wide association of Maf1 with Pol III-transcribed genes during repression. *Mol. Cell* **22**, 633–644 (2006).
9. Upadhya, R., Lee, J. & Willis, I.M. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Mol. Cell* **10**, 1489–1494 (2002).
10. Noma, K., Cam, H.P., Maraia, R.J. & Grewal, S.I. A role for TFIIIC transcription factor complex in genome organization. *Cell* **125**, 859–872 (2006).
11. Kuras, L., Kosa, P., Mencia, M. & Struhl, K. TAF-containing and TAF-independent forms of transcriptionally active TBP *in vivo*. *Science* **288**, 1244–1248 (2000).
12. Parrott, A.M. & Mathews, M.B. Novel rapidly evolving hominid RNAs bind nuclear factor 90 and display tissue-restricted distribution. *Nucleic Acids Res.* **35**, 6249–6258 (2007).
13. Borchert, G.M., Lanier, W. & Davidson, B.L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13**, 1097–1101 (2006).
14. Bortolin-Cavaille, M.-L., Dance, M., Weber, M. & Cavaille, J. C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts. *Nucleic Acids Res.* **37**, 3464–3473 (2009).
15. Lee, W. *et al.* A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **39**, 1235–1244 (2007).
16. Mavrich, T.N. *et al.* A barrier nucleosome model for statistical positioning of nucleosome throughout the yeast genome. *Genome Res.* **18**, 1073–1083 (2008).
17. Raha, D. *et al.* Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc. Natl. Acad. Sci. USA* (in the press), 107 (2010).
18. Carbon, P. & Krol, A. Transcription of the *Xenopus laevis* selenocysteine tRNA(Ser)Sec gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes. *EMBO J.* **10**, 599–606 (1991).
19. D'Ambrosio, C. *et al.* Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes Dev.* **22**, 2215–2227 (2008).
20. Haeusler, R.A., Pratt-Hyatt, M., Good, P.D., Gipson, T.A. & Engelke, D.R. Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. *Genes Dev.* **22**, 2204–2214 (2008).
21. Simms, T.A. *et al.* TFIIIC binding sites function as both heterochromatin barriers and chromatin insulators in *Saccharomyces cerevisiae*. *Eukaryot. Cell* **7**, 2078–2086 (2008).
22. Valenzuela, L., Dhillon, N. & Kamakaka, R.T. Transcription independent insulation at TFIIIC-dependent insulators. *Genetics* **183**, 131–148 (2009).
23. Donze, D., Adams, C.R., Rine, J. & Kamakaka, R.T. The boundaries of the silenced HMR domain in *Saccharomyces cerevisiae*. *Genes Dev.* **13**, 698–708 (1999).
24. Wallace, J.A. & Felsenfeld, G. We gather together: insulators and genome organization. *Curr. Opin. Genet. Dev.* **17**, 400–407 (2007).
25. Parelho, V. *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
26. Core, L.J. & Lis, J.T. Transcription regulation through promoter-proximal pausing of RNA polymerase. *Science* **319**, 1791–1792 (2008).
27. Fuda, N.J., Ardehali, M.B. & Lis, J.T. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**, 186–192 (2009).
28. Kenneth, N.S. *et al.* TRRAP and GCN5 are used by c-Myc to activate RNA polymerase III transcription. *Proc. Natl. Acad. Sci. USA* **104**, 14917–14922 (2007).
29. Yuan, C.-C. CHD8 associates with human Staf and contributes to efficient U6 RNA polymerase III transcription. *Mol. Cell. Biol.* **27**, 8729–8738 (2007).
30. Hsieh, Y.J., Kundu, T.K., Wang, Z., Kovelman, R. & Roeder, R.G. The TFIIIC90 subunit of TFIIIC interacts with multiple components of the RNA polymerase III machinery and contains a histone-specific acetyltransferase activity. *Mol. Cell. Biol.* **19**, 7697–7704 (1999).
31. Kundu, T.K., Wang, Z. & Roeder, R.G. Human TFIIIC relieves chromatin-mediated repression of RNA polymerase III transcription and contains an intrinsic histone acetyltransferase activity. *Mol. Cell. Biol.* **19**, 1605–1615 (1999).
32. Mertens, C. & Roeder, R.G. Different functional modes of p300 in activation of RNA polymerase transcription from chromatin templates. *Mol. Cell. Biol.* **28**, 5764–5776 (2008).
33. Rollins, J., Veras, I., Cabarcas, S., Willis, I. & Schramm, L. Human Maf1 negatively regulates RNA polymerase III transcription via the TFIIB family members Brf1 and Brf2. *Int. J. Biol. Sci.* **3**, 292–302 (2007).
34. Sutcliffe, J.E., Brown, T.R., Allison, S.J., Scott, P.H. & White, R.J. Retinoblastoma protein disrupts interactions required for RNA polymerase III transcription. *Mol. Cell. Biol.* **20**, 9192–9202 (2000).
35. Crighton, D. *et al.* p53 represses RNA polymerase III transcription by targeting TBP and inhibiting promoter occupancy by TFIIIB. *EMBO J.* **22**, 2810–2820 (2003).
36. Felton-Edkins, Z.A. *et al.* The mitogen-activated protein (MAP) kinase ERT induces tRNA synthesis by phosphorylating TFIIIB. *EMBO J.* **22**, 2422–2432 (2003).
37. Hiraga, S., Botsios, S. & Donaldson, A.D. Histone H3 lysine 56 acetylation by Rtt1090 is crucial for nucleosome positioning. *J. Cell Biol.* **183**, 641–651 (2008).

ONLINE METHODS

Cell culture and formaldehyde crosslinking. We cultured three separate preparations of K562 cells in RPMI supplemented with 10% (v/v) fetal bovine serum to a density of 2×10^7 cells per ml and cross-linked the cells with 1% (v/v) formaldehyde for 10 min before harvesting and freezing them.

Chromatin immunoprecipitation and quantitative PCR. We performed chromatin preparation and chromatin immunoprecipitations as previously described[38] on material derived from three independent biological replicates of 2×10^7 cells each. The antibodies used were 2663 for Bdp1 (ref. 39), 128 for Brf1 (ref. 40), 4286 for TFIIIC110 (ref. 34) and 1900 for Rpc155 (ref. 39). Antibody 4295 against Brf2 was raised by immunizing rabbits with keyhole limpet hemocyanin coupled to synthetic peptides VSRSQQRGLRRVRDLC and SDSEIEQYLRTPQEVR, corresponding to human Brf2 residues 66–80 and 385–400, respectively. Two-thirds of each immunoprecipitation was used for Illumina sequencing. We performed quantitative PCR in real time as previously described[5]. We obtained occupancy values for genomic loci by subtracting the value obtained for the control H3 locus and then expressing the results relative to the level at the arginine tRNA locus on chromosome VI that was arbitrarily defined as 100.

Illumina sequencing. We subjected input control and immunoprecipitated DNA to amplification and sequencing on an Illumina Genome Analyzer as previously described[41]. The immunoprecipitated DNA from each replicate was amplified separately and sequenced in different lanes, yielding an average read count of more than 15–19 million.

Data analysis and peak identification. We analyzed only those matches aligning to unique positions in the genome. For most factors, we obtained 9–10 million uniquely aligned reads per K562 replicate and 14–15 million per HeLa replicate. TFIIIC-110 gave a lower percentage of uniquely mappable reads (approximately 4 million per replicate), most likely because TFIIIC recognizes a DNA region with a high degree of sequence identity across tRNAs. We used the MACS peak-finding software[42] to analyze the sequencing data and identify targets of each proteins. For each factor, sequence data from the three replicates were analyzed both individually and as a single merged dataset. We considered a peak to be a target if it had a P value in the combined dataset of 1×10^{-7} or better, with the additional requirement that two or more replicates individually yield P values of 1×10^{-5} or lower. When looking for *ETC*s, we increased the stringency of the cutoff used to define TFIIIC sites by requiring that a TFIIIC peak be defined by an overall P value of 1×10^{-7} or better and a P value of 1×10^{-5} or better in all three replicates.

We visualized peaks using the Affymetrix Integrated Genome Browser[43]. We obtained genome information from the UCSC Genome Browser[44] and the Genome tRNA database[45]. Data for histone modifications and CTCF came from the Broad Institute database, and Pol II ChIP-sequencing data was from the Yale ENCODE sequencing project data on the UCSC genome database. We identified sequence motifs using the MEME suite[46].

38. Miotto, B. & Struhl, K. HBO1 histone acetylase is a co-activator of the replication licensing factor Cdt1. *Genes Dev.* **22**, 2633–2638 (2008).
39. Fairley, J.A., Scott, P.H. & White, R.J. TFIIIB is phosphorylated, disrupted and selectively released from tRNA promoters during mitosis *in vivo*. *EMBO J.* **22**, 5841–5850 (2003).
40. Cairns, C.A. & White, R.J. p53 is a general repressor of RNA polymerase III transcription. *EMBO J.* **17**, 3112–3123 (1998).
41. Zhang, Y. *et al.* Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 847–852 (2009).
42. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
43. Nicol *et al.* The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730–2731 (2009).
44. Karolchik, D. *et al.* The UCSC Genome Browser database. *Nucleic Acids Res.* **31**, 51–54 (2003).
45. Chan, P.P. & Lowe, T.M. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* **37**, D93–D97 (2009).
46. Bailey, T.L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).

npg © 2010 Nature America, Inc. All rights reserved.