

Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome

Gangning Liang^{*†}, Joy C. Y. Lin[‡], Vivian Wei[‡], Christine Yoo[‡], Jonathan C. Cheng[‡], Carvell T. Nguyen[‡], Daniel J. Weisenberger[‡], Gerda Egger[‡], Daiya Takai[‡], Felicidad A. Gonzales[‡], and Peter A. Jones^{*†‡}

Departments of ^{*}Urology and [‡]Biochemistry and Molecular Biology, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 1441 Eastlake Avenue, MS 8302L, Los Angeles, CA 90089-9181

Edited by Mark T. Groudine, Fred Hutchinson Cancer Research Center, Seattle, WA, and approved March 31, 2004 (received for review March 17, 2004)

Almost 1–2% of the human genome is located within 500 bp of either side of a transcription initiation site, whereas a far larger proportion ($\approx 25\%$) is potentially transcribable by elongating RNA polymerases. This observation raises the question of how the genome is packaged into chromatin to allow start sites to be recognized by the regulatory machinery at the same time as transcription initiation, but not elongation, is blocked in the 25% of intragenic DNA. We developed a chromatin scanning technique called ChAP, coupling the chromatin immunoprecipitation assay with arbitrarily primed PCR, which allows for the rapid and unbiased comparison of histone modification patterns within the eukaryotic nucleus. Methylated lysine 4 (K4) and acetylated K9/14 of histone H3 were both highly localized to the 5' regions of transcriptionally active human genes but were greatly decreased downstream of the start sites. Our results suggest that the large transcribed regions of human genes are maintained in a deacetylated conformation in regions read by elongating polymerase. Common models depicting widespread histone acetylation and K4 methylation throughout the transcribed unit do not therefore apply to the majority of human genes.

Most of our knowledge on the relationships between histone modifications and transcription comes from elegant studies using yeast as a model organism (1–3). The active transcription of a gene occurs in two fundamentally different processes involving the formation of an active initiation complex, which is then followed by elongation (4). Initiation is associated with histone acetylation and methylation, whereas elongation in yeast is associated with not only histone acetylation and methylation but also the recruitment of elongation factors. In yeast, histone H3-K9/14 acetylation and H3-K4 methylation are associated not only with the promoter regions but also with coding regions, suggesting that these histone modifications may also play an important role in transcriptional elongation (5–9).

These concepts have provided a great deal of background for studies in mammalian cells in which the same histone modifications have been associated with transcriptional initiation at the 5' regions of human genes (10). However, the average *Saccharomyces cerevisiae* gene size is 2 kb, and the yeast genome is much more compact than the human genome, wherein the average gene size is ≈ 27 kb, mainly because of the presence of large introns that include interspersed repeats (11, 12). The regions within 500 bp of either side of the 5' start sites of human genes represent a distinct minority (≈ 1 –2%) of the total genome so that concepts generated by using the highly compact yeast genome might not have general applicability, particularly with regard to the nature of chromatin in the transcribed regions. Indeed, recent studies have shown that there are significant differences in H3-K4 methylation patterns at several loci in chicken and yeast (13, 14). This promoted us to ask whether the patterns of histone modifications at the start sites and the far more prevalent elongated regions were the same in the human genome.

To answer this question, we developed a genome scanning technique called ChAP, which couples the chromatin immunoprecipitation (ChIP) assay with the scanning capabilities of arbitrarily primed PCR (AP-PCR) to develop a fingerprinting method to assess patterns of modifications across native chromatin structures in an unbiased fashion. This screen differs from the so called “Chip on Chip” assays (15) in that it is not limited by the sequences applied to the microarray, which are mostly coding regions; thus, it allows for an analysis of a random sample of the entire mammalian genome, which is composed of $\approx 99\%$ non-protein-coding DNA (12). We found that methylated lysine 4 (K4) and acetylated K9/14 of histone H3 were both highly localized to the 5' regions of transcriptionally active human genes but were greatly decreased downstream of the start sites, suggesting that mammalian promoters have similar chromatin configurations to those described in yeast. However, most transcribed DNA is not associated with these modifications, even on active genes.

Materials and Methods

Cell Lines and ChAP Assay. The two cell lines analyzed, T24 and LD419, were cultured and maintained as described in ref. 16. The ChAP assay couples the ChIP assay and AP-PCR genomic screening techniques.

ChIP assay. The detailed method for the ChIP assay is described in refs. 17 and 18. Ten milliliters of anti-methyl CpG binding protein 2 (MeCP2), 10 μ l of antiacetylated H3-K9/14, 10 μ l of antimethylated H3-K4 (Upstate Biotechnology, Lake Placid, NY), 10 μ l of antitrimethylated H3-K4 (Abcam, Cambridge, U.K.), or 10 μ l of normal mouse IgG as negative control (Santa Cruz Biotechnology) was used.

AP-PCR. Five microliters each of ChIP DNA was amplified by AP-PCR with a combination of three or four random primers (GC-rich or GC-poor). Isolation of fragments of interest was performed as described in refs. 19 and 20. The resulting nucleotide sequences were compared with the GenBank sequences by using the BLAST program (www.ncbi.nlm.nih.gov/blast), the University of California at Santa Cruz Human Genome Browser (<http://genome.cse.ucsc.edu>), and the CPG ISLAND SEARCHER program (www.uscnorris.com/cpgislands) (21).

PCR Analysis of Immunoprecipitated DNA. Amplification was achieved by using Expand DNA polymerase (Roche Molecular Biochemicals) with 5 μ l of immunoprecipitated DNA, 5 μ l of nonspecific antibody negative control (NAC), or 1 μ l of input

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ChIP, chromatin immunoprecipitation; AP-PCR, arbitrarily primed PCR; MeCP2, methyl CpG binding protein 2; NAC, nonspecific antibody negative control.

[†]To whom correspondence may be addressed. E-mail: jones.p@ccnt.hsc.usc.edu or gliang@usc.edu.

© 2004 by The National Academy of Sciences of the USA

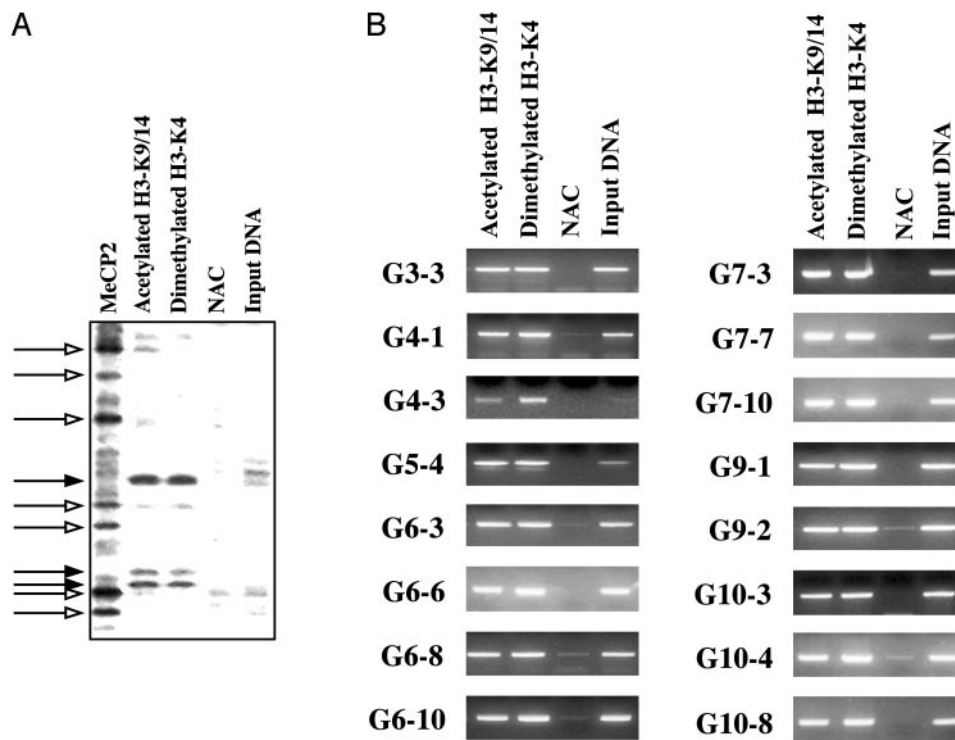


Fig. 1. Typical ChAP assay and validation by conventional ChIP assay. T24 cells were treated with formaldehyde and immunoprecipitated with antibodies specific for MeCP2 as a control (inactive marker), acetylated H3-K9/14, and dimethylated H3-K4. The final precipitated DNA was used for AP-PCR and conventional ChIP-PCR. (A) ChAP assay (which combines ChIP and AP-PCR). The ChAP fingerprinting assay gel was generated by using radioactively labeled AP-PCR products. Arrows represent bands with potentially interesting patterns, which were subsequently isolated from the gel and sequenced. Solid arrows represent bands precipitated by antibodies specific to active chromatin markers (acetylated H3-K9/14 and dimethylated H3-K4), whereas open arrows represent bands precipitated by the inactive marker antibody, MeCP2. (B) Confirmation of bands with a standard ChIP assay. Sixteen randomly chosen bands obtained from the ChAP assay were examined by using a conventional ChIP assay. Input DNA, sample representing total input chromatin (0.2%) for each experiment.

chromatin (17). PCR products were electrophoresed on 2% agarose gels. PCR conditions and primers used for these conventional ChIP PCRs are available upon request.

Real-Time PCR Analysis of Immunoprecipitated DNA. Quantitative PCR was performed with a DNA Engine Opticon System (MJ Research, Cambridge, MA) using AmpliTaq Gold DNA polymerase (Applied Biosystems) with 5 μ l of immunoprecipitated DNA, 5 μ l of NAC sample, or 1 μ l of input sample (0.2%). Fluorescently labeled TaqMan probes were synthesized by Biosearch. The primer and probe sequences are available upon request. All PCRs were carried out under the same conditions: 95°C for 15 s and 59°C for 1 min for 45 cycles. With each set of PCR, titrations of known amounts of DNA were included as a standard for quantitation. DNA from the ChIP samples immunoprecipitated with antiacetylated H3-K9/14, antidimethylated H3-K4, antitrimethylated H3-K4 and from the ChIP samples immunoprecipitated with nonspecific antibody (NAC) were included in each PCR set. The fraction of immunoprecipitated DNA was calculated as (amount of immunoprecipitated sample with antibody – amount of NAC)/(amount of input DNA (1%) – amount of NAC).

Results and Discussion

The human bladder cancer cell line, T24, was treated *in situ* with formaldehyde and subjected to standard ChIP with antibodies directed against MeCP2, H3-K9/14 acetylation, and H3-K4 dimethylation. DNAs from immunoprecipitates were amplified by AP-PCR using random GC-rich or non-GC-rich 10-mer primer sets and radioactive products displayed on polyacrylamide gels (Fig. 1A). Autoradiographs showed bands present in

the lanes from immunoprecipitates that were absent in the lane from normal mouse IgG as a nonspecific antibody negative control (NAC). These bands were also present in a diluted sample of genomic DNA (Input DNA). Informative bands were considered those present in the immunoprecipitated lanes and not present in the NAC lane. We used an antibody to MeCP2 as a marker for inactive chromatin (17, 18). The bands generated from the active markers (acetylated H3-K9/14 and dimethylated H3-K4) were coupled but distinct from those derived from the MeCP2. This result suggests that the active and inactive chromatin markers were mutually exclusive on the AP-PCR gels and demonstrated the utility and validity of the ChAP assay (Fig. 1A). In these experiments, we used three to four random primers (GC-rich or GC-poor) for each AP-PCR. In general, by choosing different combinations of primers for each AP-PCR, we obtained \approx 10–20 informative bands per reaction. We analyzed a total of 20 AP-PCRs (10 each for GC-rich and GC-poor primers) and obtained 288 informative bands precipitated by acetylated H3 and dimethylated H3-K4 antibodies. These DNA species were either weakly or not precipitated by MeCP2, and \approx 85% of the bands showed a strong concordance between H3-K9/14 acetylation and H3-K4 dimethylation (data not shown). In most cases, informative bands also arose weakly in the 0.2% input control lane. These results suggested that H3-K9/14 acetylation and H3-K4 dimethylation were most often present at the same DNA sequences in the human genome.

Fifty-seven of the 288 informative bands were randomly excised from the gels, cloned, and sequenced. Conventional ChIP analyses were performed on 16 randomly chosen fragments to confirm the copresence of both of the active markers on the isolated chromatin fragments (Fig. 1B), and RT-PCR

Table 1. Description of 57 identified fragments from the ChAP assay

	Fragments	Size, bp	Identification	CGI
5' region	P6-2	330	EST (AK05602)	N
	P9-2	440	<i>TCEB2</i> gene	N
	P10-2	280	<i>LAMB3</i> gene	N
	P1-1	275	EST (AK098790)	N
	P4-6	236	EST (BC040647)	N
	P7-1	264	EST (BG548159)	N
	P10-1	273	EST (BF674163)	N
	P10-5	187	EST (AW978993)	N
	P10-3	271	<i>BDNF</i> gene	N
	G10-3	612	<i>BDNF</i> gene	N
	G5-4	220	<i>GEMIN6</i> gene	N
	G8-4	202	EST (BQ189524)	N
	G8-1	437	EST (AI633959 and BE669903), bid.	N
	P1-3	166	<i>VPS33B</i> gene EST (AK027754), bid.	Y
	G4-3	244	EST (L02867)	Y
	G6-8	165	<i>HTATIP2</i> gene	Y
	G7-18	130	<i>PTPRU</i> gene	Y
	G9-1	576	ESTs (AL050022 and BI459794), bid.	Y
	G3-2	398	<i>BHD</i> gene	Y
	G7-7	389	<i>DBL</i> gene	Y
	G7-10	175	<i>FLJ20534</i> gene	Y
	G7-15	185	<i>HNRU</i> gene	Y
	G8-5	245	<i>AXOT</i> gene and EST (T86526), bid.	Y
	G10-4	362	<i>GFI1</i> gene	Y
	G2-4	191	EST (BG527537)	Y
	G3-3	188	<i>CPSF1</i> gene	Y
	G7-3	476	<i>MAN1</i> gene	Y
	G7-11	207	<i>ARL7</i> gene (BG990421)	Y
	G8-3	284	ESTs (AA158879 and BG830482), bid.	Y
	G9-2	285	EST (X13294)	Y
	G10-1	193	EST (BC010012)	Y
	G10-6	299	EST (AB033030)	Y
	G10-8	299	EST (AL449869)	Y
Body of genes	P1-2	350	<i>CAPN11</i> gene	N
	P10-4	289	EST (AB011106) with LTR	N
	P4-3	413	<i>C50rf7</i> gene	N
	P6-1	315	<i>HUMPPA</i> gene	N
	P4-4	268	EST (AW236867) with LTR	N
	P4-5	245	<i>BDH</i> gene	N
	P5-1	254	EST (BC027456)	N
	G6-1	250	<i>FLJ22402</i> gene	N
	G2-1	176	EST (AW070489)	N
	G6-6	294	<i>C11orf17</i> gene	N
	GW6-1	158	EST (BX640978)	N
	G4-1	238	<i>MCM2</i> gene	N
	G6-10	750	EST (AW028743)	N
	G7-9	262	EST (AK127336)	Y
	G6-3	469	<i>DKFZp761B1514</i> gene	Y
G10-9	262	<i>NOTCH2</i> gene	Y	
Other	G6-4	434	Chromosome 18 (Sine and Line)	N
	G6-7	451	No match	N
	G5-5	159	Chromosome 10	N
	G2-5	381	No match	N
	G5-3	227	Chromosome 6	N
	G2-3	209	Chromosome 1	Y
	G3-1	451	Chromosome 1 with LTR	Y
	G7-14	317	No match	Y

The 5' region is located within 500 bp of a transcriptional start site or the 5' end of an EST. Body of genes, any region 500 bp or farther downstream of the transcriptional start site; Other, nongene regions. P and G indicate fragments produced by GC-poor and GC-rich primers, respectively. CGI, CpG island; N, non-CpG island; Y, CpG island of $\geq 55\%$ GC, with a proportion of observed to expected CpG of ≥ 0.65 and a length of ≥ 500 bp (21); bid., a potentially bidirectional promoter.

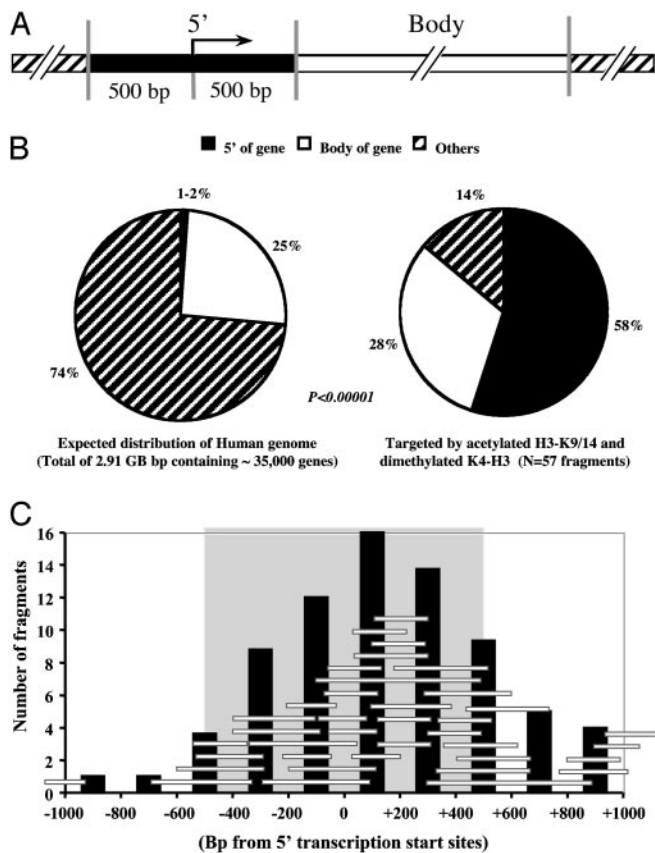


Fig. 2. (A) Map of the 5' region (located within 500 bp of either side of a transcriptional start site or 5' end of an EST gene) and the body of a gene (any region 500 bp or farther downstream of the transcriptional start site in the gene). (B) Comparison of the 5' region and body of genes in bands targeted by acetylated H3-K9/14 and dimethylated H3-K4 and the expected distribution of these regions in the whole human genome. Others, nongene regions. (C) Sizes and locations of the sequences with respect to the 5' start sites of 38 regions immunoprecipitated by acetylated H3-K9/14 and dimethylated H3-K4 antibodies within 1 kb of the nearest transcription start site. Each horizontal bar represents an individual band cloned and sequenced from the AP-PCR gels and positioned with respect to the start site in the GenBank database. Vertical bars represent the number of sequences in each 200-bp window at the indicated position relative to the start site. The gray box represents the region within 500 bp of either side of a transcriptional start site.

experiments showed that all 16 genes were expressed (data not shown). Strong bands were evident after ChIP-PCR of the DNA precipitated by the active markers. Similar results were seen concordantly in the immunoprecipitates of all 16 bands analyzed, demonstrating that the ChAP assay reliably reflected these modification patterns. Approximately 15% (43 of 288) of the fragments, such as G7-7, G4-1, G6-8, and G10-8 (Fig. 1B), showed differences in intensities in the AP-PCR gel. Although these differences were not manifest by the conventional ChIP assay, quantitative differences could be documented by real-time PCR analysis (see Figs. 3 and 4).

Fifty-seven bands immunoprecipitated by the antibodies to the active markers showed that nearly all of the bands were associated with genes (Table 1), and, remarkably, 33 of 57 (58%) fragments were located within 500 bp of either side of the transcriptional start site of a known gene or an EST in the database (Fig. 2A). Only 16 of 57 (28%) fragments analyzed were located within the body of a gene (defined as any region 500 bp or farther downstream of the transcription start site), whereas 8 of 57 (14%) were in nongene regions consisting mainly of

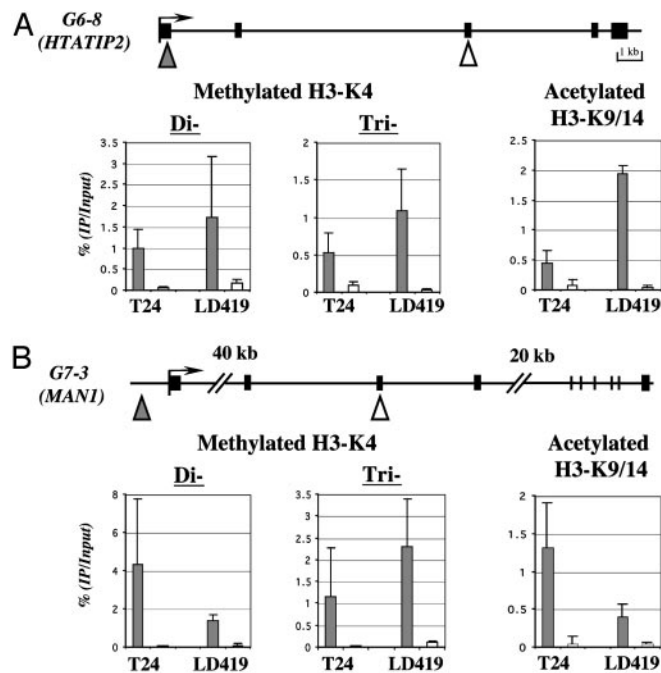


Fig. 3. Comparison of the levels of acetylated H3 and dimethylated H3-K4 in the 5' regions versus the body of genes *HTATIP2* (A) and *MAN1* (B). Results were obtained from three separate real-time PCRs of three independent ChIP assays. (A and B Upper) Solid boxes indicate exons, and arrows indicate potential transcriptional start sites. Triangles (filled triangle for 5' region of the gene and open triangle for body of gene) indicate regions analyzed by quantitative real-time PCR. (A and B Lower) Error bars indicate the standard deviation obtained from three independent experiments.

repetitive elements (12). The bands precipitated by the markers for active chromatin were highly localized to the 5' region of genes, regardless of the GC contents of the primers. Thus, 23 of 40 of the fragments for GC-rich primers (G series) and 10 of 17 of the fragments for GC-poor primers (P series) were localized to 5' regions. Twenty of the 33 fragments located in the 5' regions of genes fulfilled the criteria of CpG islands as expected for human 5' regions (12, 21). These results show that the distribution of bands depends on the antibodies and is not due to a biased amplification by the GC-poor or -rich AP-PCR primers.

We next compared the distribution of chromatin fragments with respect to the expected frequency in the human genome (Fig. 2B). For this purpose, we used the completed sequences of the human genome, which contain 2.91 gigabases of DNA and $\approx 35,000$ genes (12). The 5' regions of genes represent a minority of the genome (1–2%), with the assumption of an average of two transcription start sites per gene. Thus, the 33 of 57 (58%) fragments immunoprecipitated by acetylated H3-K9/14 and dimethylated H3-K9 antibodies were represented at least 30-fold more frequently than anticipated on a random basis ($P < 0.0001$). The preferential location of 38 of the bands near the 5' start site for each gene is clearly visible, as depicted in Fig. 2C. The peak number of sequences immunoprecipitated by the antibodies of histone acetylated H3-K9/14 and methylated H3-K4 was in the 200-bp window including and just downstream of the transcription start sites.

The ChAP assay, which is presumably unbiased, provided convincing evidence that the markers of active chromatin were preferentially located near transcription start sites, where they were easily recognized by the antibodies used for ChIP. We next validated the data, using the quantitative capabilities of real-time PCR by investigating the levels of H3-K9/14 acetylation and dimethylation of H3-K4 at the transcription start sites and

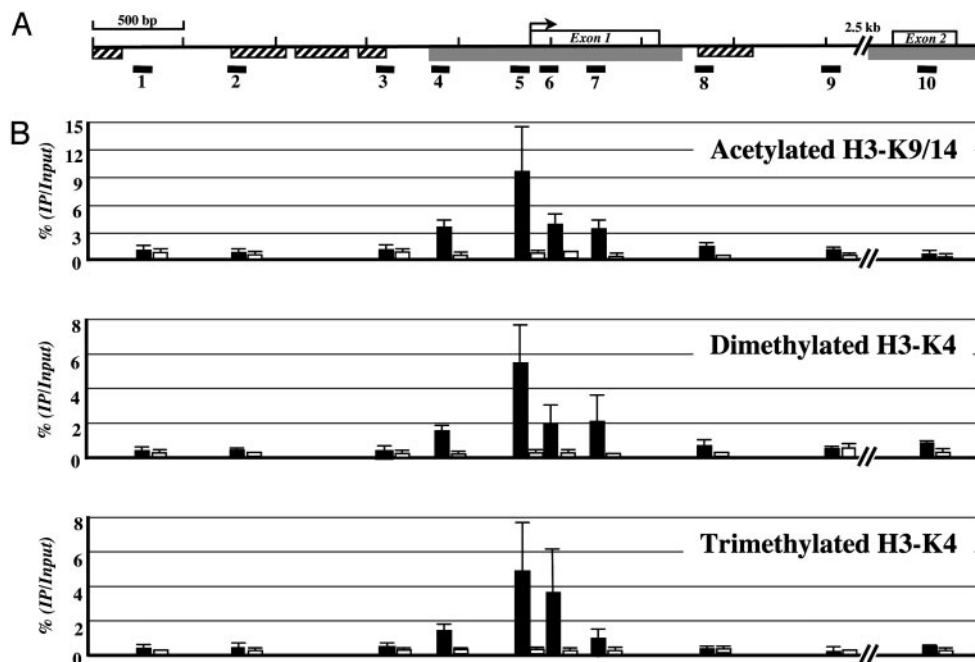


Fig. 4. (A) Map of the *p16* gene locus spanning the region upstream of the promoter through exon 2. The arrow indicates the transcription start site. The gray boxes indicate CpG islands, and the hatched boxes indicate repetitive elements. The horizontal bars below the map indicate the location of the 10 DNA regions (≈ 78 –112 bp in size) amplified by quantitative real-time PCR. (B) Map of histone modifications along the *p16* locus (regions 1–10) immunoprecipitated with antibodies for acetylated H3-K9/14, dimethylated H3-K4, trimethylated H3-K4, respectively, in the LD419 fibroblast line, which expresses *p16* (filled bars), and the T24 cancer cell line, which does not express *p16* (open bars). Error bars indicate the standard deviation obtained from three to five independent experiments.

downstream regions for two examples of genes isolated from AP-PCR gels, HIV-1 Tat interactive protein 2 gene (*HTATIP2*) and integral inner nuclear membrane protein gene (*MAN1*). The quantitation was extended to include an analysis of the state of histone modification in human LD419 normal bladder fibroblasts as well as in T24 bladder cancer cells. We also measured the levels of trimethylated H3-K4 in both cell lines, as this histone modification was recently shown to be localized exclusively to active genes in yeast (6, 8) (Fig. 3). The quantitative results from real-time PCR once again demonstrated the presence of all three markers near the transcription start sites of both genes and confirmed that these markers were 6 to 122 times more enriched at the start sites relative to downstream regions (Fig. 3). Quantitative differences in the levels of the markers between the start sites of normal and cancer cells were detected, but virtually none of these modifications were apparent in the downstream regions of the two genes, which were expressed in both cell types (data not shown).

We extended our study to include a detailed analysis of histone modifications in a 7-kb region of the transcriptional unit of the *p16* gene in both cell types (Fig. 4A). We have previously characterized the *p16* promoter in detail, so that we were able to study a well characterized promoter region (22) and not simply a 5' start site. Previous work from our laboratory showed nearly complete methylation of CpG sites in the *p16* promoter in T24 cells, and this methylation was associated with histone deacetylation and H3-K4 hypomethylation (17, 18). Quantitative real-time PCR analysis showed the presence of a "bubble" of acetylated H3-K9/14 and H3-K4 di- and trimethylation, respectively, around the transcription start site (regions 4–7) in LD419 fibroblasts, which actively express the gene (Fig. 4B) (17, 18). This region is critical for transcriptional activity (22), and both active chromatin markers were ≈ 40 - to 51-fold enriched in LD419 cells compared with T24 cells at this region. Di- and trimethylated K4-H3 and acetylated K9/14-H3, however, were substantially decreased in regions 3 and 8, located on either side

of the transcription start site in LD419 cells. Because the distance between regions 3 and 8 is ≈ 2 kb (approximately the size of an average yeast gene), the enrichment of these histone modifications indicated that these histone modifications are localized to a bubble of ≈ 12 –14 nucleosomes. This distance may be the reason that histone H3 acetylation and H3-K4 methylation have not been shown to be substantially decreased in the compact coding regions of yeast. As found earlier for *HTATIP2I* and *MAN1*, which were identified by the ChAP approach (Fig. 3), the levels of the two active chromatin modifications were substantially decreased at regions away from the promoter. These findings suggest that in human genes these histone modifications, potentially involved in initiation of transcription and/or the transition between initiation and elongation, may be confined to promoter regions, as has been previously described in yeast (1, 4). However, human genes, with their large intronic structures, have a different chromatin organization with respect to transcriptional elongation. Our findings also provide more global support for previous studies that have shown quite different histone modification patterns in chicken and yeast (13, 14).

Our results, obtained by ChAP analysis and confirmed by more quantitative real-time PCR, suggest at least two potential mechanisms for transcription initiation and elongation in human genes. H3-K9/14 acetylation and H3-K4 methylation may be required for transcription initiation and the transition stage between initiation and elongation but may not track with RNA polymerase II (pol II) throughout transcribed regions. This possibility was suggested by our earlier studies showing that regions containing heavy CpG methylation, MeCP2 binding, and nuclease inaccessibility do not block transcriptional elongation (17, 18). Alternatively, the H3-K9/14 acetylation may be "reset" to the unmodified state soon after the progression of human pol II, because it has been shown that Hos2, a histone deacetylase, is associated with the coding region of active genes and deacetylated histones H3 and H4 in yeast (2, 23). In contrast to acetylation, histone methylation is not known to be easily

reversible, and it seems likely that nucleosomes without di- and trimethylated H3-K4 are traversed by human pol II. The transcribed regions of human genes, with much larger transcriptional units than yeast, may be maintained in a deacetylated conformation to avoid inappropriate transcript initiation from cryptic promoters, including those associated with transposable elements. Recently, studies have indicated that histone H3.3 deposition is enriched on active chromatin (24, 25). Histone H3.3 also has a relatively higher enrichment of histone modifications, which include di- and trimethylated K4 and acetylated K9/14, than H3 (26). The antibodies against di- and trimethylated H3-K4 and acetylated H3-K9/14 we have used in this study target not only H3 but also H3.3. Therefore, our results suggest that the 5' regions of active human genes possibly have higher levels of histone H3.3 than transcribed regions.

The ChAP assay is therefore a rapid and robust comparative approach to analyze the distribution and changes of the histone

code along the genome in an unbiased way. Importantly, it allowed us to examine the whole genome rather than being constrained by the availability of coding sequences that make up a very small proportion of the whole human genome. The increasing availability of specific antibodies to modified histones and other chromatin proteins makes ChAP an ideal method to compare the relative distributions of these modifications, not only in individual genomes but also between different cell types. The strong association of H3-K9/14 acetylation and H3-K4 methylation with the 5' regions of genes suggests that these regions may stand out as "beacons" against the background of the rest of the genome, allowing them to be easily identified by the ChAP assay and presumably by transcription initiation factors.

This work was supported by National Institutes of Health Grant CA 82422 and Training Grant T32 CA 09659 (to D.J.W.).

1. Hartzog, G. A. (2003) *Curr. Opin. Genet. Dev.* **13**, 119–126.
2. Kurdistani, S. K. & Grunstein, M. (2003) *Nat. Rev. Mol. Cell Biol.* **4**, 276–284.
3. Gerber, M. & Shilatifard, A. (2003) *J. Biol. Chem.* **278**, 26303–26306.
4. Pokholok, D. K., Hannett, N. M. & Young, R. A. (2002) *Mol. Cell* **9**, 799–809.
5. Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. & Schreiber, S. L. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 8695–8700.
6. Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J. & Kouzarides, T. (2002) *Nature* **419**, 407–411.
7. Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., Hall, M. C., Borchers, C. H. & Strahl, B. D. (2003) *Genes Dev.* **17**, 654–663.
8. Ng, H. H., Robert, F., Young, R. A. & Struhl, K. (2003) *Mol. Cell* **11**, 709–719.
9. Schaft, D., Roguev, A., Kotovic, K. M., Shevchenko, A., Sarov, M., Neugebauer, K. M. & Stewart, A. F. (2003) *Nucleic Acids Res.* **31**, 2475–2482.
10. Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D. & Felsenfeld, G. (2001) *Science* **293**, 2453–2455.
11. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., *et al.* (1996) *Science* **274**, 546–563–567.
12. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., *et al.* (2001) *Science* **291**, 1304–1351.
13. Myers, F. A., Evans, D. R., Clayton, A. L., Thorne, A. W. & Crane-Robinson, C. (2001) *J. Biol. Chem.* **276**, 20197–20205.
14. Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C. & Kouzarides, T. (2004) *Nat. Cell Biol.* **6**, 73–77.
15. Shannon, M. F. & Rao, S. (2002) *Science* **296**, 666–669.
16. Liang, G., Gonzales, F. A., Jones, P. A., Orntoft, T. F. & Thykjaer, T. (2002) *Cancer Res.* **62**, 961–966.
17. Nguyen, C. T., Gonzales, F. A. & Jones, P. A. (2001) *Nucleic Acids Res.* **29**, 4598–4606.
18. Nguyen, C. T., Weisenberger, D. J., Velicescu, M., Gonzales, F. A., Lin, J. C., Liang, G. & Jones, P. A. (2002) *Cancer Res.* **62**, 6456–6461.
19. Liang, G., Salem, C. E., Yu, M. C., Nguyen, H. D., Gonzales, F. A., Nguyen, T. T., Nichols, P. W. & Jones, P. A. (1998) *Genomics* **53**, 260–268.
20. Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W. & Jones, P. A. (2002) *Mol. Cell Biol.* **22**, 480–491.
21. Takai, D. & Jones, P. A. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 3740–3745.
22. Gonzalzo, M. L., Hayashida, T., Bender, C. M., Pao, M. M., Tsai, Y. C., Gonzales, F. A., Nguyen, H. D., Nguyen, T. T. & Jones, P. A. (1998) *Cancer Res.* **58**, 1245–1252.
23. Wang, A., Kurdistani, S. K. & Grunstein, M. (2002) *Science* **298**, 1412–1414.
24. Ahmad, K. & Henikoff, S. (2002) *Mol. Cell* **9**, 1191–1200.
25. Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. (2004) *Cell* **116**, 51–61.
26. McKittrick, E., Gafken, P. R., Ahmad, K. & Henikoff, S. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 1525–1530.