ACADEMIC PRESS

Biochemical and Biophysical Research Communications 293 (2002) 182–191

BBRC

www.academicpress.com

Chromosomal organization and localization of the human histone deacetylase 9 gene (*HDAC9*)☆

Ulrich Mahlknecht,[a,*] Susanne Schnittger,[b] Jutta Will,[a] Nurten Cicek,[a] and Dieter Hoelzer[a]

[a] *University of Frankfurt Medical Center, Department of Hematology/Oncology, Theodor-Stern-Kai 7, 60590 Frankfurt, Germany*
[b] *University of Munich, Großhadern Medical Center, Department of Medicine III, Marchioninistr. 15, 81377 Munich, Germany*

Received 15 March 2002

Abstract

Epigenetically mediated modulation of gene promoter function through histone acetylation modifying enzymes, which regulate the acetylation state of histone proteins and other promoter-bound transcription factors, is increasingly appreciated as a key component in the regulation of reversible gene expression. While histone acetyltransferases (HATs), which are frequently part of multisubunit coactivator complexes, lead to the relaxation of chromatin structure and transcriptional activation, histone deacetylases (HDACs) tend to associate with multisubunit corepressor complexes, which result in chromatin condensation and transcriptional repression of specific target genes. We have isolated and characterized the human *HDAC9* genomic sequence, which spans a region of 458 kb and which has one single chromosomal locus. Determination of the exon–intron splice-junctions established that *HDAC9* is encoded by 23 exons ranging in size from 22 bp (exon 1) to 264 bp (exon 11). Characterization of the 5′ flanking genomic region revealed that the human *HDAC9* promoter lacks both the canonical TATA and CCAAT boxes; CpG elements are missing. The human *HDAC9* open reading frame is 3036 bp long and encodes a 1011 aa protein with a predictive molecular weight of 111.3 kDa and an isoelectric point of 6.41. Fluorescence in situ hybridization analysis localized the human *HDAC9* gene to chromosome 7p21, a region which has been associated particularly with the pathogenesis of gynecological tumors.

Keywords: Histones; Chromatin; Histone deacetylase; Chromosomes; Genes; Structural; Tumor suppressor

The acetylation and deacetylation of both histones and non-histone proteins occurs through opposing activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs), which are important in the regulation of various cellular processes, which include the alteration of chromatin structure, gene activation and gene silencing, the modulation of meiosis, and mechanisms of aging [1]. While hyperacetylation of histones is generally correlated with transcriptionally active chromatin, hypoacetylation is associated with transcriptional silencing, which in turn is mainly based on the limited access of activation factors [2] and the concurrent binding of transcriptional repressor complexes, of which HDACs themselves are part, to promoter DNA [3]. Based on size and sequence considerations, yeast HDACs and their corresponding mammalian orthologs have been grouped into three categories, class I, class II, and class III HDACs of which yeast class I HDACs include RPD3, HOS1, HOS2, and HOS3 [4], yeast class II HDACs contain HDA1 [4] and yeast class III HDACs comprise the silencing protein SIR2. Likewise, based on their structural similarity with yeast RPD3 mammalian HDAC1 [5], HDAC2 [6], HDAC3 [7], and HDAC8 [8] are referred to as mammalian class I HDACs. Similarly mammalian HDAC4 [9], HDAC5 [9], HDAC6 [9], HDAC7 [10], HDAC9 [11–13], and HDAC10 [14], which are related to yeast HDA1, are designated as mammalian class II HDACs [9]. Finally, on the basis of their homology to the yeast SIR2 protein, mammalian class III HDACs include the SIRT1-7 proteins [15–18].

☆ *Abbreviations:* HDAC, histone deacetylase; HAT, histone acetyltransferase.

* Corresponding author. Fax: +49-69-6301-83046.
E-mail address: Mahlknecht@em.uni-frankfurt.de (U. Mahlknecht).

0006-291X/02/$ - see front matter
PII: S0006-291X(02)00193-6

Histone deacetylase enzymatic activities are associated with both the activation or repression of specific gene products and can be inhibited by histone deacetylase inhibitors [9,19]. Disruption of histone deacetylase genes has been reported to cause histone hyperacetylation and changes in transcription levels of specific gene products [4,20]. In the present study, we report the isolation, characterization, and chromosomal localization of BAC clones 541o10 and 551L10, which contain the human *HDAC9* genomic locus. Only little is known about the functional aspects of HDAC9 and class II HDACs in general. However, it has been reported that HDAC9 (GenBank AAK66821) is preferentially expressed in the brain, heart, and pancreas, while a HDAC9 isoform, which has been described to be truncated at the C-terminal end and which has been termed HDAC9a (GenBank AAK66822), predominates in the lung, liver, and skeletal muscle, while expression of both forms of HDAC9 is equivalent in the placenta and in the kidney [11]. Functional analysis showed that both isoforms have comparable TSA-inhibitable HDAC activity but much less activity than HDAC4 [11]. In addition, HDAC9 has been shown to specifically repress MEF2C-mediated transcriptional activation [11,12]. Misleadingly, a number of sequences, which have been earlier designated as HDAC7B (GenBank XP_004963, NP_ 055522, XP_056945, XP_ 056946, and AAF04254), are identical with the HDAC9 sequences and are clearly distinct from the sequences, which have been deposited at GenBank for HDAC7A (GenBank XP_017202, XP_ 027199, AAF63491, XP_ 27205, NP_056216, XP_0271 98, XP_007047, XP_027206, and NP_057680).

Materials and methods

Identification of the human HDAC9 cDNA. A homology search of the EST database at NCBI (National Center for Biotechnology Information) yielded four positive IMAGE consortium cDNA clones (GenBank AW382651, AW382677, AA701346, and AA488817), of which two clones (GenBank AW382651 and AW382677) were obtained from the Reference Center of the German Human Genome Project (RZPD, Berlin, Germany). Plasmid DNA was prepared according to published protocols [21] and its insert sequence was determined by DNA cycle sequencing.

Identification of BAC genomic clones BAC clones 541o10 and 551L10. An arrayed BAC genomic library (BAC human release II, Incyte Genomics) was screened with clone GenBank NM_014707, which was shown to contain a fragment covering the 3′ end of the human *HDAC9* cDNA. Blots were washed for 20 min in 2× SSPE and prehybridized for 1 h at 55 °C in hybridization buffer (Digene, Beltsville, MD). The *HDAC9* cDNA fragment was excised from the pT7T3D-vector (Pharmacia, Piscataway, NJ) by *Not*I and *Eco*RI digestion, gel-purified and radiolabeled with [α-^{32}P]dCTP (DuPont, Wilmington, DE) using the Multiprime DNA labeling system (Amersham, Arlington Heights, IL), denatured, added to the prehybridization buffer, and allowed to hybridize for at least 16 h at 55 °C. Membranes were then washed twice for 20 min in 2× SSPE-0.1% SDS, twice for 20 min in 0.2× SSPE-0.1% SDS at room temperature, and once for 1 h at 65 °C in 0.2× SSPE-0.1% SDS. Autoradiographic exposures with two intensifying screens were carried out for 1–7 days at −70 °C. Two clones, which were identified as *BAC clones 541o10 and 551L10,* were found to be positive. These clones had an insert of approximately 120 kb and were shown to contain the human *HDAC9* genomic sequence. Bac-DNA was prepared from *clones 541o10 and 551L10* according to published protocols [21].

Instrumental methods. Dye terminator cycle sequencing was performed using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq DNA polymerase (Perkin–Elmer, Branchburg, NJ) and analyzed with an ABI PRISM 310 Genetic Analyzer which utilizes the four-color, single-lane sequencing chemistry.

PCR methods. The *HDAC9* sequence was determined by primer walking on both strands using a direct sequencing strategy. Sequencing reactions were performed using 0.6 μg cDNA and 20–30 mer oligonucleotide primers (GENSET, La Jolla, CA). Sequencing reactions were set up in a volume of 20 μl containing 10 pmol of the sequencing primer, 4 μl BigDye Terminator Cycle Sequencing Ready Reaction Mix (Perkin–Elmer, Norwalk, CT), DNA as indicated, and ddH_2O added up to a final volume of 20 μl. The thermal cycling profile for the sequencing of the cDNA-clones was as follows: denaturation at 95 °C for 30 s, annealing at 50 °C for 15 s, extension at 60 °C for 4 min (25 cycles), and storage at 4 °C.

Sequence analysis and computer database searches. DNA sequence analysis was performed using the MacVector program (Oxford Molecular Group PLC). Computer-aided chromosomal mapping was carried out online using the Genome Database (GDB), hosted by The Hospital for Sick Children Toronto (Ontario, Canada) and the UniGene and LocusLink programs at the National Center for Biotechnology Information (NCBI). Sequence comparisons were done with the BLAST algorithm of the GenBank and EMBL databases [22]. Protein similarity scores were calculated with the CLUSTAL W Multiple Alignment Program Version 1.7 [23]. Protein motifs were identified online at the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB) with the PROSITE program and double-checked using the MotifFinder software hosted by the GenomeNet WWW server at Institute for Chemical Research, Kyoto University (Japan), but still remain to be experimentally confirmed. Motifs in unaligned multiple sequences were identified with the MEME software on the HUSAR server hosted by the Biocomputing Service Group at the German Cancer Research Center (DKFZ, Heidelberg). Potential transcription factor binding sites were identified with the TFSEARCH program [24]. Sequence similarities were calculated with the GAP software, which considers all possible alignments and gap positions between two sequences and creates a global alignment that maximizes the number of matched residues and minimizes the number and size of gaps on the HUSAR server [25]. Repetitive elements were identified on the Repeat Masker Server at the University of Washington and CpG elements were found with the CPG software hosted by the European Bioinformatics Institute (EMBL outstation).

Phylogenetic analysis. Phylogenetic trees were constructed from known human HDAC7 and HDAC9 histone deacetylase sequences, which were obtained from a protein sequence similarity search with the human HDAC9 (GenBank NM_014707) protein and nucleotide sequences using the BLAST 2.0 program at NCBI database (non-redundant GenBank CDS: translations + PDB + SwissProt + SPupdate + PIR). Progressive multiple sequence alignments were performed with the CLUSTAL W Multiple Alignment Program Version 1.7 [26]. Trees were calculated and drawn with the CLUSTREE software, which computes a phylogenetic tree according to the Neighbor-Joining Method. While vertical numbers along branches represent percentage values for bootstrap statistical support, horizontal values indicate the

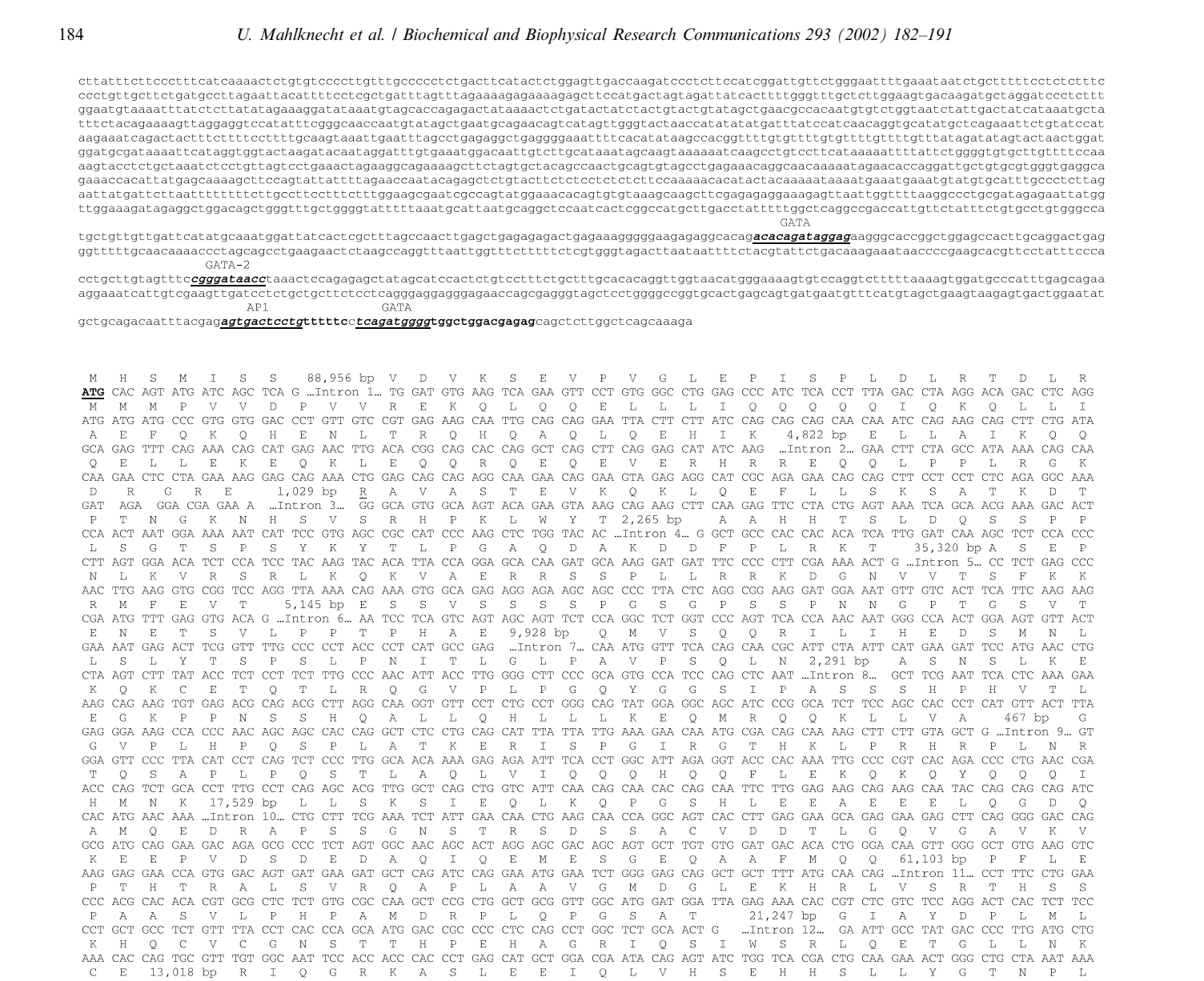

Fig. 1. The complete sequence of *HDAC9* cDNA together with the predicted amino acid sequence is shown with the location of each intron with respect to the cDNA sequence. The *HDAC9* cDNA has an open reading frame of 3036 bp, which yields a 1011 aa protein. Approximately 2 kb of the 5′ upstream promoter region is indicated in small letters. Putative transcription factor binding sites are underlined. The translational start (ATG) and stop codons (TGA) are bold and underlined.

percent divergence figures between two pairs of sequences, which are used by the Neighbor-Joining Method to graphically display the phylogenetic tree [27] (Fig. 5).

Chromosomal localization of clone BAC clones 541o10 and 551L10 by fluorescence in situ hybridization (FISH). FISH analysis with BAC clones 541o10 and 551L10 was performed according to previously

published protocols [28]. For chromosome identification the procedure was repeated with the same hybridization mix containing an additional μl of a chromosome 7 centromeric probe (CEP7, Vysis, Bergisch Gladbach, Germany).

Results

Identification and cloning of cDNAs encoding human HDAC9

Homology searches of the dbEST at NCBI (National Center for Biotechnology Information) [22] using the 4238 bp human *HDAC9* cDNA sequence (GenBank NM_014707) yielded four positive IMAGE consortium cDNA clones (GenBank AW382651, AW382677, AA701346, and AA488817), of which two clones (GenBank AW382651 and AW382677) were obtained from the Reference Center of the German Human Genome Project (RZPD, Berlin, Germany). The authenticity of their insert was confirmed by DNA cycle sequencing (Fig. 1). Amino acid sequence alignments of human class II HDACs, which were calculated with the CLUSTAL W Multiple Alignment algorithm, identified several truncated isoforms of human HDAC9, which have been deposited at GenBank XP_004963, NP_055522, XP_056945, XP_056946, and AAF04254 and which have been misleadingly named HDAC7B (Fig. 2), and which are clearly distinct from the sequences, which have been deposited at GenBank for HDAC7A (GenBank XP_017202, XP_027199, AAF63491, XP_27205, NP_056216, XP_027198, XP_007047, XP_027206, and NP_057680) both at the DNA and the protein levels.

Identification of BAC genomic clones BAC clones 541o10 and 551L10

An arrayed BAC genomic library (Genome Systems) was screened with an [α-^{32}P]dCTP-radiolabeled ~1000 bp fragment of the 3′ human *HDAC9* cDNA sequence (GenBank NM_014707) by hybridization at high stringency [19]. Two clones with an average insert size of 120 kb were found to be positive, identified as BAC clones

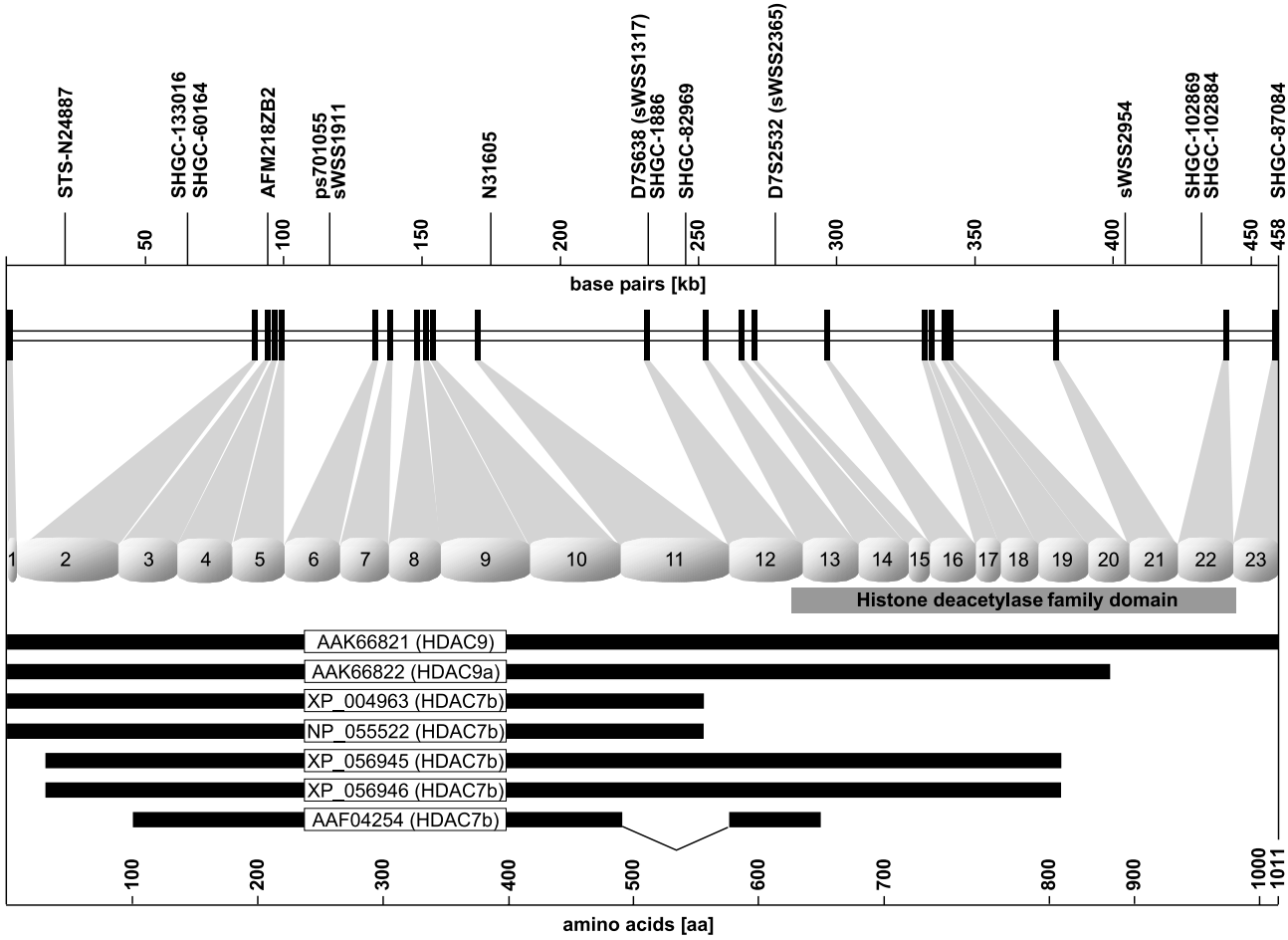

Fig. 2. Genomic organization of the human *HDAC9* gene. The genomic organization of the *HDAC9* gene, which includes the relative position of exons and introns, is shown. The relative positions of overlapping STS markers and the histone deacetylase family domain are also indicated.

541o10 and 551L10, and confirmed to contain the *HDAC9* genomic clone both by Southern blot hybridization and cycle sequencing. BAC clones 541o10 and 551L10 were then used as a probe for fluorescence in situ hybridization studies and for the determination of the genomic organization of *HDAC9* (Fig. 4).

HDAC9 is a single-copy gene

Hybridization screening of an arrayed human genomic DNA library using a *HDAC9* cDNA fragment as the probe yielded two positive clones (BAC clones 541o10 and 551L10) with an insert size of approximately 120 kb in the 7.4 kb vector pBeloBAC 11. Both partial sequencing and results obtained by electronic PCR of these two clones identified a series of STS markers (STS-N24887, SHGC-133016, SHGC-60164, AFM218ZB2, ps701055, sWSS1911, N31605, D7S638, SHGC-1886, SHGC-82969, D7S2532, sWSS2954, SHGC-102869, SHGC-102884, SHGC-87084) (Fig. 4) to overlap with the genomic sequence of human *HDAC9*, which allowed the orientation of the transcriptional unit with its 5′ extremity being telomeric with regard to its 3′ extremity. These data, together with the results obtained from fluorescence in situ hybridization studies, which revealed one single site of hybridization on human metaphase chromosomes, indicate that *HDAC9* is present in the genome as a single-copy gene (Fig. 4).

Characterization of the structure of the human HDAC9 genomic locus

Partial sequencing of BAC clones 541o10 and 551L10 helped us to identify clone RP11-273F8 (GenBank NT_007918) with the BLAST algorithm. This clone was fully sequenced by the Washington University, Genome Sequencing Center and contained the complete *HDAC9* genomic sequence, which we have then used for the determination of *HDAC9* introns and exon/intron boundaries (Table 1). Human *HDAC9* spans a region of ~458 kb. Determination of the exon–intron splice-junctions established that the gene is encoded by 23 exons ranging in size from 22 bp (exon 1) to 264 bp (exon 11). A series of STS markers has been identified, of which D7S1990 is located ~7 kb upstream of the 5′ extremity of the *HDAC9* ATG start codon. STS-marker N24887 is located within intron 1, upstream of the overlapping markers SHGC-133016 and SHGC-60164, which are also situated within intron 1. AFM218ZB2 has been found within exon 4 and ps701055 has been identified within intron 5, sWSS1911 has also been targeted to intron 5, but a little more distally. Marker N31605 was mapped to intron 11 and the overlapping markers D7S638 and SHGC-1886 were localized within intron 12. SHGC-82969 is also located within intron 12, but more distally than the other two markers. D7S2532 has been identified within intron 15 and sWSS2954 within intron 20. SHGC-102869 and SHGC-102884

Table 1
Exon/intron splice-junctions of the human *HDAC9* gene

Exon no.	Exon size	5′ Splice donor	Intron no.	Intron size	3′ Splice acceptor
1	22	ATCAGCTCAG**gt**aagatcctct	1	88.956	ctggttctttt**ag**TGGATGTGAA
2	242	GCATATCAAG**gt**agcaaatgct	2	4.822	aagttgcaac**ag**GAACTTCTAG
3	142	GGACGAGAAA**gt**aagaggcacc	3	1.029	tgtgtatttc**ag**GGGCAGTGGC
4	127	TCTGGTACAC**gt**atgttcagtg	4	2.265	tgtcttttct**ag**GGCTGCCCAC
5	122	CGAAAAACTG**gt**aagttggttt	5	35.320	ctcaatcccc**ag**CCTCTGAGCC
6	132	GAGGTGACAG**gt**aattgaggac	6	5.145	aatatttttc**ag**AATCCTCAGT
7	116	TCATGCCGAG**gt**aagaccctta	7	9.928	ttttttaac**ag**CAAATGGTTT
8	123	CCAGCTCAAT**gt**aagtcattgc	8	2.991	ttctcaacac**ag**GCTTCGAATT
9	214	CTTGTAGCTG**gt**aattcattat	9	467	tttttttttc**ag**GTGGAGTTCC
10	218	CATGAACAAA**gt**aagcctccaa	10	17.529	actctcttct**ag**CTGCTTTCGA
11	264	TATGCAACAG**gt**aataggcaaa	11	61.103	tcttggcaac**ag**CCTTTCCTGG
12	178	TCTGCAACTG**gt**aggaatccct	12	21.247	cttgtcttaa**ag**GAATTGCCTA
13	134	TAAATGTGAG**gt**aatccagaat	13	13.018	attttcttgc**ag**CGAATTCAAG
14	121	ATACTCCTAG**gt**ctgtacgggc	14	4.828	cttactgtat**ag**GTGATGACTC
15	50	TGGACTTGGG**gt**aagtacaagt	15	26.189	ctgtttgctc**ag**GTGGACAGTG
16	108	AGAGCTGAAG**gt**gaggtccggg	16	35.708	ttgttttcac**ag**AATGGGTTTG
17	56	CCACAGCCAT**gt**aagtaccagg	17	244	tctattccgc**ag**GGGGTTCTGC
18	88	TGTAGATCTG**gt**atgtattcct	18	5.918	atttccctgt**ag**GATGTTCACC
19	120	CCCAAATGAG**gt**tcggtttatt	19	313	ttctcttccc**ag**GTTGGAACAG
20	98	AAGCATTCAG**gt**tggtacttct	20	38.480	tttactgtgc**ag**GACCATCGTG
21	119	ACGGCAAAAT**gt**aagtacctct	21	61.212	gtattatggt**ag**GTTTTGGTCA
22	134	AGGAAATGAG**gt**aaaaaagtaa	22	18.203	ctattcttgc**ag**CTGGAGCCAC
23	108				

Exon sequences are given in uppercase letters and intron sequences are given in lowercase letters. The sizes of the single exons and introns are indicated. Consensus splice donor and splice acceptor sequences are given in bold and are underlined.

have also been found to overlap within intron 21 and finally SHGC-87084 has been found within exon 23, the last exon of human *HDAC9* (Fig. 2). Characterization of ~2 kb of the 5′ upstream promoter region failed to identify any CpG elements. Additionally, the human *HDAC9* promoter lacks both the canonical TATA and CCAAT boxes, which could be an indication that *HDAC9* may be a housekeeping gene. Several putative optimal transcription factor binding sites have been identified, however, their biological relevance awaits still to be investigated experimentally (Fig. 1). Translational stop codons in all reading frames precede the human *HDAC9* open reading frame. The human *HDAC9* mRNA encodes a 1011 aa protein with a predictive molecular weight of 111.3 kDa and an isoelectric point of 6.41. The predictive HDAC9 histone deacetylase domain is encoded by exons 12–22 and is highly conserved within the category of mammalian class II HDACs (Fig. 2 and 3).

HDAC9 is localized in the 7p21 region

The chromosomal localization of human *HDAC9* was determined by FISH using BAC clones 541o10 and 551L10 as a probe. Hybridization signals were visualized on the short arm of chromosome 7 in the p21 region with no secondary sites of hybridization. Human *HDAC9* is located within the chromosomal interval D7S1990 and SHGC-87084 (position of interval: 29.6 cM from the top of chromosome 7—this position has been determined for marker N31605 with the genemap software at NCBI [http://www.ncbi.nlm.nih.gov/genemap/]) (Fig. 4).

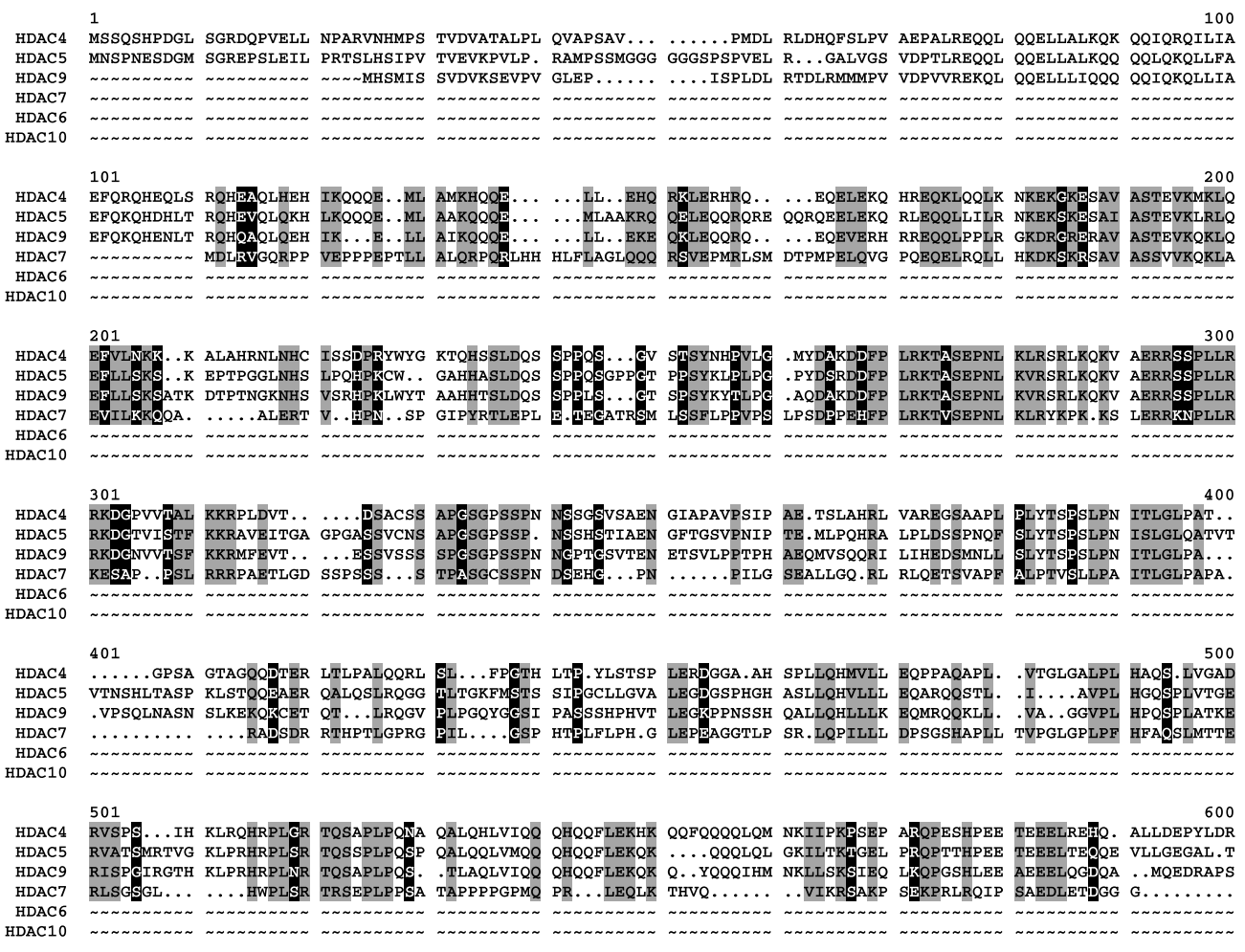

Fig. 3. Amino acid sequence alignment of human class II HDACs. Accession numbers of the sequences used in this alignment: HDAC4 (GenBank AAD29046), HDAC5 (GenBank AAD29047), HDAC6 (GenBank AAD29048), HDAC7 (GenBank AAF04254), HDAC9 (GenBank AAK66821), and HDAC10 (GenBank AAL30513). ClustalW Colors mark similarities in *Protein sequences*. Bright grey/black letters: All amino acids of a column are identical. Dark grey/black letters: More than half of the amino acids of a column are identical or belong to one of the *strong groups* (amino acids with strong similarities). Black/white letters: More than half of the amino acids of a column belong to one of the *weak groups* (amino acids with weak similarities).

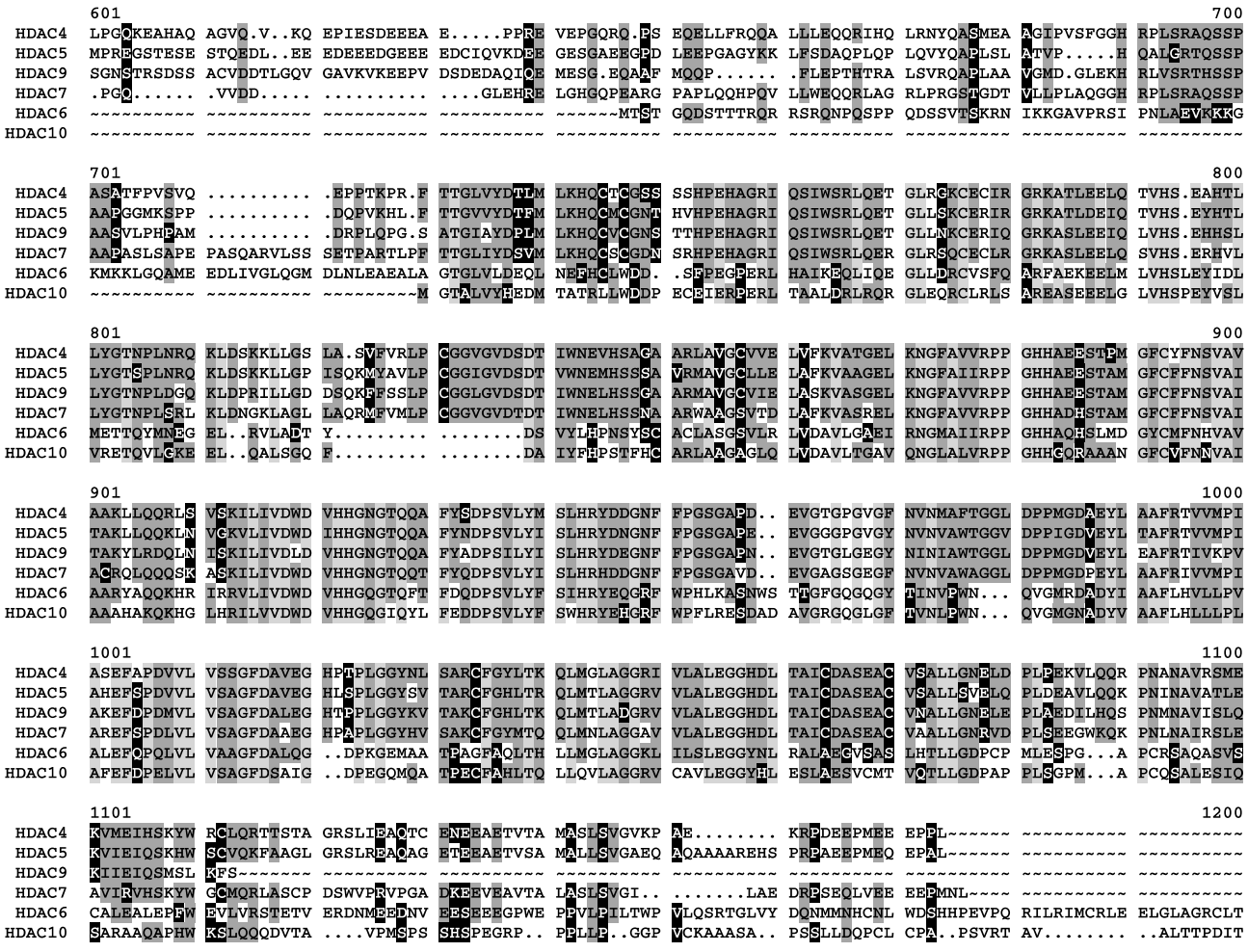

Fig. 3. *(continued)*

Phylogenetic analysis

We have screened the expressed sequence tag database (NCBI) with the human HDAC9 protein sequence (GenBank AAK66821) and identified a number of related human HDACs, which may be divided into three classes of proteins depending on their similarity to the corresponding yeast ancestor proteins RPD3, HDA1, and SIR2. A consensus evolutionary tree was constructed after bootstrapping on the basis of an alignment of human HDAC9 related proteins as outlined under Materials and methods (Fig. 5, Table 2). Class I HDACs are homologs of the yeast RPD3 (GenBank P32561) protein and include the human histone deacetylase enzymes HDAC1 (GenBank Q13547), HDAC2 (GenBank Q92769), HDAC3 (GenBank O15379), and HDAC8 (GenBank AAF73428). Class II HDACs are descendants of yeast HDA1 (GenBank P53073) and comprise HDAC4 (GenBank AAD29046), HDAC5 (GenBank AAD29047), HDAC6 (GenBank AAD29048), HDAC7 (GenBank AAF04254), HDAC9 (GenBank AAK66821), and HDAC10 (GenBank AAL30513). Finally, class III HDACs, which consist of the human proteins SIRT1 (GenBank AAD40849), SIRT2 (GenBank AAD40850), SIRT3 (GenBank AAD40851), SIRT4 (GenBank AAD40852), SIRT5 (GenBank AAD40853), SIRT6 (GenBank AAF43432), and SIRT7 (GenBank AAF43431) are orthologs of the yeast protein SIR2 (GenBank CAA25667) (Fig. 5).

Discussion

Modifiers of chromatin structure are gaining increasing attention as potential targets in the treatment of cancer. HAT and HDAC enzymatic activities are known to be involved both in the pathogenesis as well as in the suppression of cancer. Some of the genes encoding these enzymes have been shown to be rearranged in the context of chromosomal translocations in human acute leukemias and solid tumors, where fusions of regulatory and coding regions of a variety of transcription factor genes result in completely new gene products, which may interfere with regulatory cascades that control cell

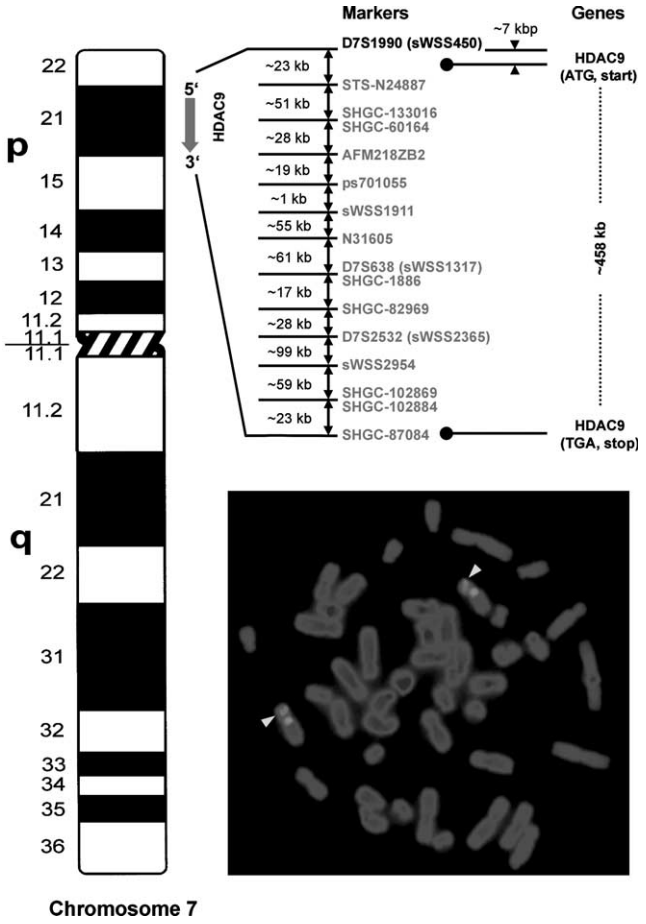

Fig. 4. Chromosomal mapping of the human *HDAC9* gene. Lower right panel: Fluorescence in situ hybridization of BAC clone 551L10 (labeled with Texas Red) to human chromosome 7p21. Centromeric region of chromosome 7 labeled in green. Left panel: Chromosome 7 idiogram according to the International System for Cytogenetic Nomenclature (ISCN 1995), which illustrates the chromosomal position of BAC clone 551L10 within the interval D7S1990 to SHGC-87084 (position of interval: 29.6 cM from the top of chromosome 7). Neighboring markers are also indicated. The chromosomal orientation of *HDAC9* is shown (arrow, upper right panel).

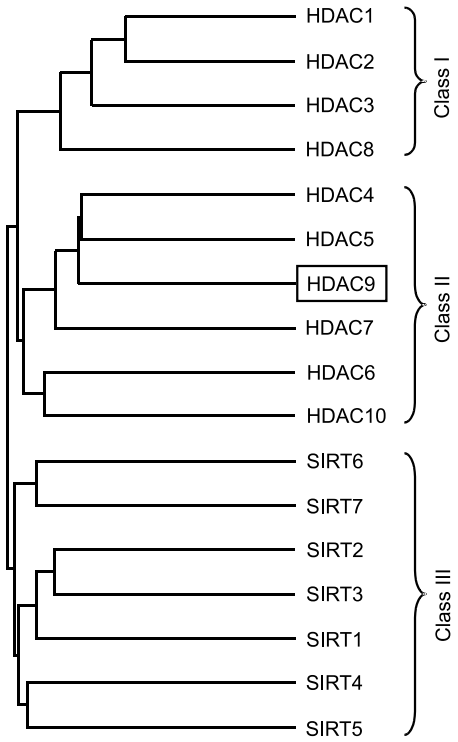

Fig. 5. HDAC phylogenetic tree. This tree is based on the human HDAC9 protein sequence and includes human orthologs of the yeast RPD3, HDA1, and SIR2 families of HDACs.

growth and differentiation [29–31]. On the other hand, some histone acetylation modifying enzymes have been located within chromosomal regions that are particularly prone to chromosomal breaks. In these cases gains and losses of chromosomal material may affect the availability of functionally active HATs and HDACs, which in turn disturbs the tightly controlled equilibrium of histone acetylation [32].

In the present study, we report the identification, cloning, and mapping of *HDAC9* on the genomic level. Human *HDAC9* is a single-copy gene that spans a region of approximately 458 kb. It is composed of 23 exons ranging in size from 22 (exon 1) to 264 bp (exon 11). Characterization of the 5′ flanking genomic region revealed that the human *HDAC9* promoter lacks both the canonical TATA and CCAAT boxes; CpG elements are missing. The human *HDAC9* mRNA encodes a 1011

Table 2
Sequence comparison between human class II HDAC protein sequences

Similarity	Identity						
	Yeast HDA1	Human HDAC4	Human HDAC5	Human HDAC6	Human HDAC7	Human HDAC9	Human HDAC10
Yeast HDA1		37	36	32	36	34	38
Human HDAC4	46		63	36	54	58	37
Human HDAC5	47	69		34	54	58	34
Human HDAC6	41	45	42		25	41	52
Human HDAC7	36	60	59	31		44	51
Human HDAC9	51	71	71	56	55		36
Human HDAC10	61	53	53	68	37	53	

Numbers to the upper right represent percentages of sequence identity, while numbers to the lower left stand for the percentage of similarity between two protein sequences.

aa protein with a predictive molecular weight of 111 kDa and an isoelectric point of 6.41. Fluorescence in situ hybridization analysis localized the human *HDAC9* gene to chromosome 7p21, a region which has been associated particularly with the pathogenesis of gynecological tumors. According to promoter analyses the human *HDAC9* is unlikely to be a housekeeping gene, since it lacks CpG-rich elements in the promoter and characteristic binding sites for the zinc-finger transcription factor Sp1 about 300 bp upstream of the 5′ end of the *HDAC9* transcriptional start site [33], even though it lacks the canonical TATA and CCAAT boxes [34]. The predictive HDAC9 histone deacetylase domain is encoded by exons 12–22 and is highly conserved within the category of mammalian class II HDACs. Human *HDAC9* is located on chromosome 7p21 in a region, which is characterized by frequent gains and losses of chromosomal material. While chromosomal gains in the 7p21 region have been observed particularly in primary colorectal cancer [35], in biliary tract carcinomas [36], hepatocellular carcinoma [37], non-small cell lung carcinoma [38], and sarcomatoid renal cell carcinomas [39], loss of heterozygosity has been observed in thymic epithelial tumors [40], in sporadic Wilms' tumor [41], and ovarian cancer [42], while rearrangements of 7p21 have been associated with low-grade endometrial stromal sarcomas [43]. In view of the fact that the steady states of histone acetylation and deacetylation play a key role in the regulation of transcription, deletion of HDAC9 would most probably shift the steady state toward acetylation at the level of specific genes targeted by HDAC9 and either upregulate or downregulate transcriptional events [20,31]. Such a dysregulation might represent a critical event in the multistep pathway leading to full cellular transformation and malignancy. Accordingly, the *HDAC9* gene and its product could play a key role in the pathogenesis of solid tumors.

Acknowledgments

The work was partly supported by the German National Science Foundation (Deutsche Forschungsgemeinschaft, MA 257/2) and institutional funds from the University of Frankfurt Medical Center. The sequence data of human clone RP11-273F8 were produced by the Washington University, Genome Sequencing Center and can be obtained from http://www.ncbi.nlm.nih.gov/Entrez (GenBank NT_007918). We also thank the Reference Center of the German Human Genome Project (RZPD, Berlin, Germany) for kindly providing IMAGE consortium cDNA clones (GenBank AW382651 and AW382677).

References

[1] R.A. Frye, Biochem. Biophys. Res. Commun. 273 (2000) 793–798.
[2] A.P. Wolffe, D. Pruss, Cell 84 (1996) 817–819.
[3] T. Heinzel, R.M. Lavinsky, T.M. Mullen, M. Soderstrom, C.D. Laherty, J. Torchia, W.M. Yang, G. Brard, S.D. Ngo, J.R. Davie, E. Seto, R.N. Eisenman, D.W. Rose, C.K. Glass, M.G. Rosenfeld, Nature 387 (1997) 43–48.
[4] S.E. Rundlett, A.A. Carmen, R. Kobayashi, S. Bavykin, B.M. Turner, M. Grunstein, Proc. Natl. Acad. Sci. USA 93 (1996) 14503–14508.
[5] J. Taunton, C.A. Hassig, S.L. Schreiber, Science 272 (1996) 408–411.
[6] W.M. Yang, C. Inouye, Y. Zeng, D. Bearss, E. Seto, Proc. Natl. Acad. Sci. USA 93 (1996) 12845–12850.
[7] W.M. Yang, Y.L. Yao, J.M. Sun, J.R. Davie, E. Seto, J. Biol. Chem. 272 (1997) 28001–28007.
[8] E. Hu, Z. Chen, T. Fredrickson, Y. Zhu, R. Kirkpatrick, G.F. Zhang, K. Johanson, C.M. Sung, R. Liu, J. Winkler, J. Biol. Chem. 275 (2000) 15254–15264.
[9] C.M. Grozinger, C.A. Hassig, S.L. Schreiber, Proc. Natl. Acad. Sci. USA 96 (1999) 4868–4873.
[10] H.Y. Kao, M. Downes, P. Ordentlich, R.M. Evans, Genes Dev. 14 (2000) 55–66.
[11] X. Zhou, P.A. Marks, R.A. Rifkind, V.M. Richon, Proc. Natl. Acad. Sci. USA 98 (2001) 10572–10577.
[12] D.B. Sparrow, E.A. Miska, E. Langley, S. Reynaud-Deonauth, S. Kotecha, N. Towers, G. Spohr, T. Kouzarides, T.J. Mohun, Embo. J. 18 (1999) 5085–5098.
[13] T. Nagase, K. Ishikawa, M. Suyama, R. Kikuno, N. Miyajima, A. Tanaka, H. Kotani, N. Nomura, O. Ohara, DNA Res. 5 (1998) 277–286.
[14] H.Y. Kao, C.H. Lee, A. Komarov, C.C. Han, R.M. Evans, J. Biol. Chem. 24 (2001) 24.
[15] R.A. Frye, Biochem. Biophys. Res. Commun. 260 (1999) 273–279.
[16] G. Afshar, J.P. Murnane, Gene 234 (1999) 161–168.
[17] J.C. Tanny, D. Moazed, Proc. Natl. Acad. Sci. USA 98 (2001) 415–420.
[18] K.G. Tanner, J. Landry, R. Sternglanz, J.M. Denu, Proc. Natl. Acad. Sci. USA 97 (2000) 14178–14182.
[19] S. Emiliani, W. Fischle, C. Van Lint, Y. Al-Abed, E. Verdin, Proc. Natl. Acad. Sci. USA 95 (1998) 2795–2800.
[20] C. Van Lint, S. Emiliani, E. Verdin, Gene Expr. 5 (1996) 245–253.
[21] H.C. Birnboim, J. Doly, Nucleic Acids Res. 7 (1979) 1513–1523.
[22] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Nucleic Acids Res. 25 (1997) 3389–3402.
[23] W.J. Wilbur, D.J. Lipman, Proc. Natl. Acad. Sci. USA 80 (1983) 726–730.
[24] T. Heinemeyer, E. Wingender, I. Reuter, H. Hermjakob, A.E. Kel, O.V. Kel, E.V. Ignatieva, E.A. Ananko, O.A. Podkolodnaya, F.A. Kolpakov, N.L. Podkolodny, N.A. Kolchanov, Nucleic Acids Res. 26 (1998) 362–367.
[25] S.B. Needleman, C.D. Wunsch, J. Mol. Biol. 48 (1970) 443–453.
[26] J.D. Thompson, D.G. Higgins, T.J. Gibson, Nucleic Acids Res. 22 (1994) 4673–4680.
[27] N. Saitou, M. Nei, Mol. Biol. Evol. 4 (1987) 406–425.
[28] U. Mahlknecht, S. Schnittger, O.G. Ottmann, C. Schoch, M. Mosebach, W. Hiddemann, D. Hoelzer, Biochim. Biophys. Acta 1493 (2000) 342–348.
[29] T. Taki, M. Sako, M. Tsuchida, Y. Hayashi, Blood 89 (1997) 3945–3950.
[30] K. Ida, I. Kitabayashi, T. Taki, M. Taniwaki, K. Noro, M. Yamamoto, M. Ohki, Y. Hayashi, Blood 90 (1997) 4699–4704.
[31] U. Mahlknecht, D. Hoelzer, Mol. Med. 6 (2000) 623–644.
[32] U. Mahlknecht, O.G. Ottmann, D. Hoelzer, Mol. Carcinog. 27 (2000) 268–271.
[33] M. Holler, G. Westin, J. Jiricny, W. Schaffner, Genes Dev. 2 (1988) 1127–1135.

[34] B. Lewin, Regulation of Transcription, Oxford University Press, New York, 1997.
[35] H. Aragane, C. Sakakura, M. Nakanishi, R. Yasuoka, Y. Fujita, H. Taniguchi, A. Hagiwara, T. Yamaguchi, T. Abe, J. Inazawa, H. Yamagishi, Int. J. Cancer 94 (2001) 623–629.
[36] K. Shiraishi, K. Okita, T. Harada, N. Kusano, T. Furui, S. Kondoh, A. Oga, S. Kawauchi, Y. Fukumoto, K. Sasaki, Cancer 91 (2001) 570–577.
[37] D.B. Zimonjic, C.L. Keck, S.S. Thorgeirsson, N.C. Popescu, Hepatology 29 (1999) 1208–1214.
[38] C. Luk, M.S. Tsao, J. Bayani, F. Shepherd, J.A. Squire, Cancer Genet. Cytogenet. 125 (2001) 87–99.
[39] F. Jiang, H. Moch, J. Richter, C. Egenter, T. Gasser, L. Bubendorf, R. Gschwind, G. Sauter, M.J. Mihatsch, J. Pathol. 185 (1998) 382–388.
[40] R. Zhou, A. Zettl, P. Strobel, K. Wagner, H.K. Muller-Hermelink, S. Zhang, A. Marx, P. Starostik, Am. J. Pathol. 159 (2001) 1853–1860.
[41] R.G. Grundy, J. Pritchard, P. Scambler, J.K. Cowell, Oncogene 17 (1998) 395–400.
[42] V.M. Wasenius, A. Jekunen, O. Monni, H. Joensuu, S. Aebi, S.B. Howell, S. Knuutila, Genes Chromosom. Cancer 18 (1997) 286–291.
[43] L.R. Donner, Cancer Genet. Cytogenet. 78 (1994) 115–126.