Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 6489–6493, July 1995
Medical Sciences

Construction of a 2.8-megabase yeast artificial chromosome contig and cloning of the human methylthioadenosine phosphorylase gene from the tumor suppressor region on 9p21

(chromosomal loss/purine nucleoside phosphorylase/enzyme deficiency/*CDKN2* gene)

OLUFUNMILAYO I. OLOPADE*, HELEN M. POMYKALA, FITSUM HAGOS, LISE W. SVEEN, RAFAEL ESPINOSA III, MARTIN H. DREYLING, SUSAN GURSKY, WALTER M. STADLER, MICHELLE M. LE BEAU, AND STEFAN K. BOHLANDER

Section of Hematology/Oncology, Department of Medicine and The University of Chicago Cancer Research Center, The University of Chicago Pritzker School of Medicine, Chicago, IL 60637

Communicated by Janet D. Rowley, The University of Chicago Pritzker School of Medicine, Chicago, IL, March 10, 1995

ABSTRACT **Many human malignant cells lack methylthioadenosine phosphorylase (MTAP) enzyme activity. The gene (*MTAP*) encoding this enzyme was previously mapped to the short arm of chromosome 9, band p21–22, a region that is frequently deleted in multiple tumor types. To clone candidate tumor suppressor genes from the deleted region on 9p21–22, we have constructed a long-range physical map of 2.8 megabases for 9p21 by using overlapping yeast artificial chromosome and cosmid clones. This map includes the type I *IFN* gene cluster, the recently identified candidate tumor suppressor genes *CDKN2* ($p16^{INK4A}$) and *CDKN2B* ($p15^{INK4B}$), and several CpG islands. In addition, we have identified other transcription units within the yeast artificial chromosome contig. Sequence analysis of a 2.5-kb cDNA clone isolated from a CpG island that maps between the *IFN* genes and *CDKN2* reveals a predicted open reading frame of 283 amino acids followed by 1302 nucleotides of 3′ untranslated sequence. This gene is evolutionarily conserved and shows significant amino acid homologies to mouse and human purine nucleoside phosphorylases and to a hypothetical 25.8-kDa protein in the *pet* gene (coding for cytochrome bc_1 complex) region of *Rhodospirillum rubrum*. The location, expression pattern, and nucleotide sequence of this gene suggest that it codes for the MTAP enzyme.**

Unbalanced translocations or interstitial deletions of 9p are recurring abnormalities in a variety of tumor types including acute lymphoblastic leukemia, glioma, melanoma, non-small cell lung cancer, head and neck cancer, bladder cancer, and mesothelioma (1). Homozygous deletion of DNA sequences on 9p or loss of heterozygosity has now been described in multiple tumor types (2–7). A number of the cell lines and patient samples with 9p gene deletions also lack methylthioadenosine phosphorylase (MTAP) enzyme activity. To our knowledge, the gene encoding MTAP has not been cloned but has been mapped to 9p22–9q13 (8). In a few cases, the deletions that included both the *IFN* gene cluster and the *MTAP* gene were interstitial and submicroscopic, suggesting that these genes or a tumor suppressor gene (TSG) closely linked to them was the target of the 9p deletions. This hypothesis was further supported by the linkage of a gene that confers susceptibility to melanoma (*MLM*) to 9p21 in the region between *D9S126* and the *IFNA* gene cluster (9).

CDKN2 ($p16^{INK4A}$) was recently proposed as a candidate TSG in this locus because the gene has been shown to be rearranged, deleted, or mutated in a majority of tumor cell lines (6, 7). This gene codes for a 16-kDa protein (p16) that inhibits CDK4 and CDK6 by binding in competition with cyclin D (10). In humans, *CDKN2* is adjacent to a gene encoding a similar protein, now called *CDKN2B* ($p16^{INK4B}$), that shares 44% homology with *CDKN2* in the first 50 amino acids and up to 97% homology in the remainder of the protein (11). Whether *CDKN2* is *MLM* remains unclear, because two recent studies (12, 13) provide conflicting evidence. Hussussian *et al.* (12) described six disease-related germ-line mutations in *CDKN2* in 33 of 36 melanoma patients from nine families and suggested that *CDKN2* likely is *MLM*. This is in contrast to 2 of 13 mutations in 9p21-linked families and 0 of 38 familial melanoma patients described by Kamb *et al.* (13). These reports raise the possibility that *CDKN2* may not be the only clinically relevant TSG on 9p and that loss of tumor suppression may involve inactivation of other as yet unidentified genes in the region in certain tumor types. In support of this possibility Cheng *et al.* (14) published their findings of two additional regions of nonoverlapping homozygous deletions on 9p21 in malignant mesothelioma, one telomeric to *CDKN2* and the other centromeric to it.

The coincident loss of MTAP enzyme activity in many tumor cell lines with homozygous *IFN* gene deletions suggests that *MTAP* is closely linked to the *IFN* gene cluster. We have suggested (15) that a 9p TSG should be localized between the *IFN* gene cluster and the *MTAP* locus based on *IFN* gene rearrangements seen in two cell lines and leukemia cells from one patient with deletions on 9p. In the reports published to date, it has been difficult to determine the exact position of the *MTAP* gene in relation to the homozygous deletions on 9p.

In this manuscript, we describe the construction of a long-range physical map around the *IFN* gene cluster that covers a distance of 2.8 megabases (Mb), as determined by pulsed-field gel electrophoresis (PFGE) and also the isolation of the *MTAP* gene cDNA.† We have localized all of the known genes and several CpG islands on this map. Restriction sites and PFGE fragment sizes are clearly delineated on the resultant map, which extends further proximally than the one presented by Weaver-Feldhaus *et al.* (16). Also, several additional markers are described and localized on this map. The approximate location of the shortest region of overlap of 9p deletions in gliomas, melanomas, lung cancer, leukemia, mesothelioma, head and neck cancer, and bladder cancers in relation to this map is discussed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: YAC, yeast artificial chromosome; TSG, tumor suppressor gene; Mb, megabase(s); STS, sequence tagged site; FISH, fluorescence *in situ* hybridization; SRO, shortest region of overlap; PNP, purine nucleoside phosphorylase; PFGE, pulsed-field gel electrophoresis.

*To whom reprint requests should be addressed at: The University of Chicago Medical Center, 5841 South Maryland Avenue, MC 2115, Chicago, IL 60637-1470.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. U22233).

MATERIALS AND METHODS

Cell Lines. The clinical and cytogenetic characterization of the tumor cell lines used in this study has been published (4, 5). Cell lines were used for the deletion mapping because they provide an indefinite supply of DNA and the deletions present in cell lines were similar to the deletions observed in primary leukemias and gliomas (2–4).

Analysis of Yeast Artificial Chromosome (YAC) Clones. YAC clones corresponding to *IFNA1* (17, 18) and *D9S966* (19) sequence tagged sites (STSs) were isolated from the Centre d'Etude Polymorphisme Humain (CEPH) MegaYAC library. High molecular weight YAC DNA was isolated digested, electrophoresed, and blotted as described (17). To detect the left and right YAC vector arms, 346-bp *Hin*dIII–*Bam*HI and 276-bp *Bam*HI–*Sal* I restriction fragments from the plasmid pBR322 were used, respectively. Other probes used are as described in Fig. 1. The DNA probes were labeled with [α-^{32}P]dATP by using the random-primer labeling technique (20). PCR of the STSs was performed with 50 ng of YAC or human genomic DNA as template.

Fluorescence *in Situ* Hybridization (FISH) Analysis. YAC clones were purified by PFGE and the DNA was amplified by using a sequence-independent amplification technique (21). The amplification products were then labeled with biotin-11-dUTP or directly labeled nucleotides (Spectrum-Orange, Vysis), and FISH was performed as described (22).

Cloning YAC-End-Specific Clones. To obtain YAC-end-specific probes, YAC-end rescue was performed as described by Hermanson *et al.* (23). Single-copy fragments from the YAC-end-rescued inserts were used as probes for Southern blot hybridization to DNA of the different YACs and to a panel of tumor cell lines. Each end-clone probe was also used to screen a copy of the Lawrence Livermore Laboratory chromosome 9 flow-sorted cosmid library. The overlap between YACs was identified by comparing Southern blots after hybridization to the *IFNA2* gene probe that cross hybridizes to a large number of *IFN* genes and after hybridization to the different end-clone probes. YACs were also aligned by comparison of their long-range restriction maps.

Screening for Expressed Sequences. The cosmids obtained by using probes from this ordered YAC contig were used for several different strategies to detect expressed sequences including direct screening of cDNA libraries, exon trapping, and a cDNA selection protocol based on the capture of sequence independently amplified cosmid fragments by biotinylated cDNA (24, 25). Exon trapping was performed according to the manufacturer's protocol by using the Exon Trapping kit (BRL/Life Technologies, Gaithersburg, MD). Each exon-trapped product or cDNA-selected fragment from the cosmids was hybridized to a multiple-tissue Northern blot (Clontech), a somatic cell hybrid panel (Oncor), "zoo blots" (Bios, New Haven, CT), and Southern blots of tumor cell lines. Products that detected transcripts on Northern blot analysis were used to screen cDNA libraries from human adult brain, fetal brain, and fibroblast constructed in λgt10 or -11 (Clontech). In each case, 7.5×10^5 plaques were screened. Positive clones were subcloned into the pBluescript vector (Stratagene) for sequencing.

Nucleotide Sequencing and Generation of STSs. Sequencing was performed on an Applied Biosystems model 373A DNA sequencing system with the Prism Ready Reaction Dye-Deoxy Terminator Cycle sequencing kit (Applied Biosystems). cDNA clones were sequenced entirely on both strands by using double-stranded templates. The DNA sequence and the predicted open reading frame were compared with GenBank data bases (on September 13, 1994) by using the BLASTN and BLASTP programs (26).

RESULTS

Construction of the Physical Map. Fifteen YAC clones were identified after screening the CEPH MegaYAC library with the *IFNA* and *D9S966* STSs. Five clones (33%) were found to be nonchimeric by FISH analysis and were further analyzed. Two YAC clones (YAC 802B11 and 886F9) contain the entire *IFN* gene cluster whereas the remaining three YAC clones (883G5, 942A3, and 807E4) contain *D9S966.* YACs A73B12 and A88E10 were obtained with consensus *IFN* STS from the St. Louis YAC library (17). The YACs were digested with the

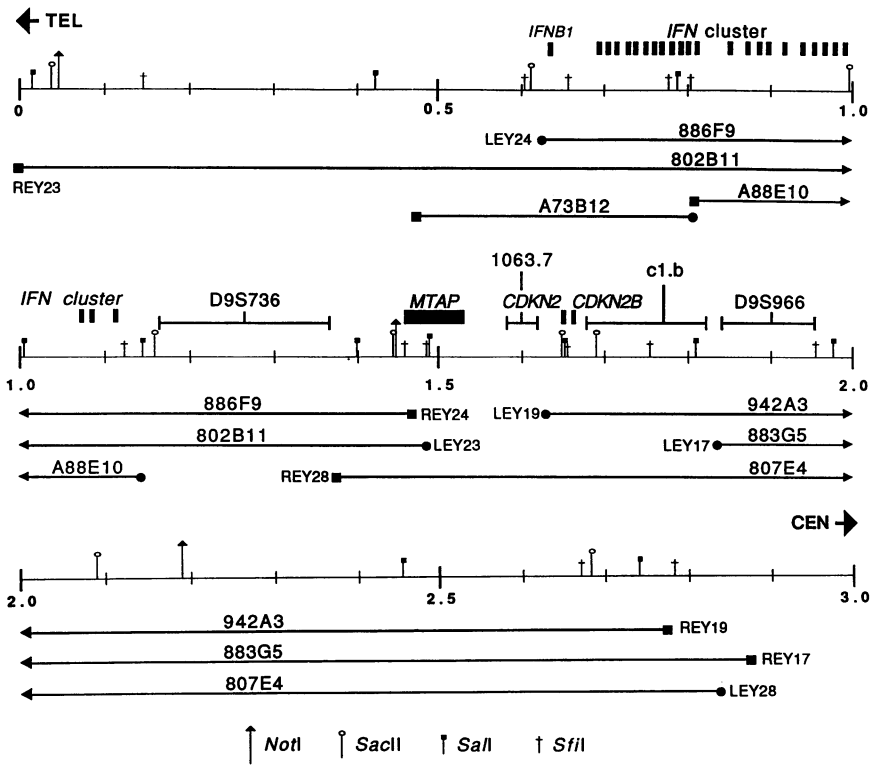

FIG. 1. PFGE map of the 2.8-Mb YAC contig on 9p21. The individual YACs are aligned to the long-range restriction map above each overlapping segment. Restriction sites are designated as shown. This map does not show every restriction site for the enzymes *Sal* I, *Sac* II, and *Sfi* I because sites that are further away from the probes used were not detected. The *IFNB1* and *IFNA* genes and pseudogenes and the *MTAP*, *CDKN2*, and *CDKN2B* genes are represented by solid vertical bars. Broad locations for the markers *D9S736*, *D9S966*, 1063.7, and c1.b are shown. REY and LEY designate the right and left YAC vector arms, respectively. YACs A73B12 and A88E10 were derived with *IFN* STS from the St. Louis YAC library (27). Cosmids and end-rescued plasmid clones are not shown. Distances are drawn to scale; marked distances are in megabases.

Table 1. Homozygous loss of 9p markers in tumor cell lines

Cell type (*n*)	% cell lines test showing homozygous deletions							
	IFNB1	*IFNA*	*D9S736*	M1.4	*CDKN2*	*CDKN2B*	*D9S966*	*D9S171*
Leukemia (18)	39	44	ND	65	89	78	44	6
Melanoma (18)	0	0	0	ND	45	45	15	0
Glioma (26)	27	42	42	63	69	65	42	12
Bladder (16)	0	31	ND	50	50	50	44	ND
Head and neck (8)	0	0	0	0	38	25	0	ND
Lung (58)	6	8	ND	34	34	29	5	2
Mesothelioma (5)	0	0	0	100	100	100	40	20

Homozygous deletion of these markers was detected by Southern blot analysis or STS–PCR. The location of the markers is shown in Fig. 1. ND, not done; M1.4, 1.4-kb fragment from the 3′ untranslated portion of the *MTAP* gene.

rare-cutting restriction endonucleases *Not* I, *Sac* II, *Sal* I, and *Sfi* I. After PFGE and Southern blot transfer, the blots were hybridized to a battery of probes including the *IFNA*, *IFNB1*, the left and right vector arm probes, *D9S966*, six end probes, and the *CDKN2* cDNA probe. The resulting map is shown in Fig. 1. The *IFN* genes contained within YACs 802B11 and 886F9 were identified and aligned with our previous map of the *IFN* gene cluster (17). Except for YAC 886F9, none of the YACs demonstrated any unusual deletions or rearrangements, as determined by STS content. YAC 886F9, described by Weaver-Feldhaus *et al.* (16), is larger and extends further centromeric than the clone isolated in our laboratory, suggesting that this YAC may have undergone an internal deletion. However, the STS content of the remaining human insert was consistent with the other *IFN*-derived YACs in our laboratory. To characterize the YAC clones further, single-copy DNA fragments obtained from the YAC-end clones were used as probes on Southern blot hybridization. The results are included in Fig. 1. Each end-clone probe mapped back to the respective YACs and to chromosome 9 by FISH analysis. This map does not show every restriction site for the enzymes *Sal* I, *Sac* II, and *Sfi* I because sites that are further away from the probes used were not detected. However, several CpG islands can readily be identified on this map.

Deletion Mapping Analysis. Each unique fragment from the end clones and additional STSs were tested on our panel of cell lines to refine the deletion map. The results are summarized in Table 1. Homozygous deletion of at least one marker derived from this YAC contig was detected in 69% of the glioma cell lines, 45% of the melanoma cell lines, 50% of the bladder cancer cell lines, 89% of the leukemias, 100% of the mesotheliomas, 38% of the head and neck cancer, and 34% of the lung cancer cell lines. The majority of the cell lines had large homozygous deletions that overlapped around *CDKN2/CDKN2B*. We have shown (27) that the deletion in Hs294T, a melanoma cell line, could not be complemented by introducing a chromosome 9 derived from the T98G cell line by microcell chromosome transfer. However, introducing a normal short arm of chromosome 9 derived from a human fibroblast cell line induced senescence in Hs294T (27). The region deleted in Hs294T is flanked by *D9S736* and *D9S966*. In T98G, the homozygous deletion is flanked by *MTAP* and *CDKN2B* (data not shown). Therefore, a region of ≈100 kb was defined by the overlapping homozygous deletions in these two cell lines. Thus, we were able to define a shortest region of overlap (SRO) of these 9p deletions between the 3′ end of *MTAP* and *CDKN2B*. From Table 1, it is apparent that the pattern and percentage of 9p homozygous deletions differ in different tumor types. For example, in melanomas, mesotheliomas, and head and neck cancers, the deletions rarely extend into the *IFN* gene cluster, whereas the *IFN* genes are included in 27–44% of the deletions in leukemias, bladder cancer, and gliomas. Moreover, *MTAP* is homozygously deleted with the same frequency as *CDKN2* in some tumor types.

The following markers were not present in our YAC contig: *D9S3*, *D9S126*, *D9S171*, *D9S162*, *D9S962* (MDS10), *D9S963* (MDS36), and an STS from the *D9S171* YAC that maps at least 500 kb telomeric of *D9S171* (refs. 19 and 28 and Joseph Testa, personal communication). We were able to localize *D9S736* within YACs 802B11 and 886F9 in a 170-kb *Sal* I–*Sfi* I fragment centromeric to the *IFN* gene cluster and close to the right end of YAC 886F9; 1063.7 was present in YAC 807E4 only and c1.b was in YACs 942A3 and 807E4 (Fig. 1) (7, 16). Because the distance from the *IFN* gene cluster to the centromeric end of this YAC contig is 1.8 Mb, we estimate that *D9S171* should be a minimum distance of 2.3 Mb from the centromeric end of the *IFN* gene cluster and *D9S736* would be at least 2.0 Mb from *D9S171*. This is consistent with previous estimates. *D9S736* has been estimated to be 2 centimorgans from *D9S171* (16), whereas *D9S126* was estimated to be at a minimum distance of 1.0 Mb from the *IFN* gene cluster (29).

Expressed Sequences Within and Around the SRO. An 85-bp exon-trapped product obtained from a cosmid that maps in the CpG island at the right end of YAC 886F9 was used to screen a cDNA library. One of the clones, a 2.5-kb cDNA, detects two major transcripts of ≈2.3 kb and 6.0 kb, as shown in Fig. 2. This gene is expressed to various degrees in all tissue types and is conserved in all mammalian species, as judged by zoo-blot hybridization (data not shown). The nucleotide sequence reveals an open reading frame coding for 283 amino acids that included the initiator methionine codon (Fig. 3*a*). The protein sequence shows homology to the human, mouse, and bacteria purine nucleoside phosphorylase (PNP) gene and to a hypothetical 25.8-kDa protein in the *pet* genes (coding for cytochrome bc_1 complex) region of *Rhodospirillum rubrum*, and also to a recently described open reading frame from *Saccharomyces cerevisiae* (Fig. 3*b*). MTAP is a PNP but has different substrate specificity than the PNPs that have been cloned to date. As shown in Fig. 3*b*, the region of homology to

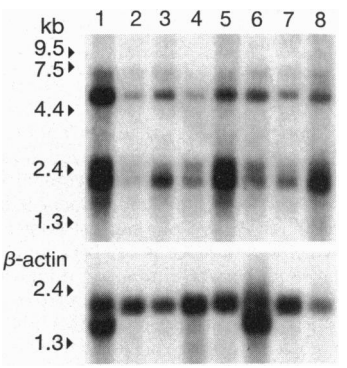

FIG. 2. (*Upper*) Northern blot of RNA from multiple human tissues hybridized with the 3′ 1.4-kb fragment of the *MTAP* cDNA clone. (*Lower*) The blot was reprobed with β-actin cDNA. Lanes: 1, heart; 2, brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; 8, pancreas.

a

```
                                                    GAATTCCGCTCCGCACTGCTCAC
   TCCCGCGCAGTGAGGTTGGCACAGCCACCGCTCTGTGGCTCGCTTGGTTCCCTTAGTCCCGAGCG
   CTCGCCCACTGCAGATTCCTTTCCCGTGCAGAC ATG GCC TCT GGC ACC ACC ACC ACC
                                      M   A   S   G   T   T   T   T 8
   GCC GTG AAG ATT GGA ATA ATT GGT GGA ACA GGC CTG GAT GAT CCA GAA
 9 A   V   K   I   G   I   I   G   G   T   G   L   D   D   P   E 24
   ATT TTA GAA GGA AGA ACT GAA AAA TAT GTG GAT ACT CCA TTT GGC AAG
25 I   L   E   G   R   T   E   K   Y   V   D   T   P   F   G   K 40
   CCA TCT GAT GCC TTA ATT TTG GGG AAG ATA AAA AAT GTT GAT TGC ATC
41 P   S   D   A   L   I   L   G   K   I   K   N   V   D   C   I 56
   CTC CTT GCA AGG CAT GGA AGG CAG CAC ACC ATC ATG CCT TCA AAG GTC
57 L   L   A   R   H   G   R   Q   H   T   I   M   P   S   K   V 72
   AAC TAC CAG GCG AAC ATC TGG GCT TTG AAG GAA GAG GGC TGT ACA CAT
73 N   Y   Q   A   N   I   W   A   L   K   E   E   G   C   T   H 88
   GTC ATA GTG ACC ACA GCT TGT GGC TCC TTG AGG GAG GAG ATT CAG CCC
89 V   I   V   T   T   A   C   G   S   L   R   E   E   I   Q   P 104
   GGC GAT ATT GTC ATT ATT GAT CAG TTC ATT GAC AGG ACC ACT ATG AGA
105 G   D   I   V   I   I   D   Q   F   I   D   R   T   T   M   R 120
   CCT CAG TCC TTC TAT GAT GGA AGT CAT TCT TGT GCC AGA GGA GTG TGC
121 P   Q   S   F   Y   D   G   S   H   S   C   A   R   G   V   C 136
   CAT ATT CCA ATG GCT GAG CCG TTT TGC CCC AAA ACG AGA GAG GTT CTT
137 H   I   P   M   A   E   P   F   C   P   K   T   R   E   V   L 152
   ATA GAG ACT GCT AAG AAG CTA GGA CTC CGG TGC CAC TCA AAG GGG ACA
153 I   E   T   A   K   K   L   G   L   R   C   H   S   K   G   T 168
   ATG GTC ACA ATC GAG GGA CCT CGT TTT AGC TCC CGG GCA GAA AGC TTC
169 M   V   T   I   E   G   P   R   F   S   S   R   A   E   S   F 184
   ATG TTC CGC ACC TGG GGG GCG GAT GTT ATC AAC ATG ACC ACA GTT CCA
185 M   F   R   T   W   G   A   D   V   I   N   M   T   T   V   P 200
   GAG GTG GTT CTT GCT AAG GAG GCT GGA ATT TGT TAC GCA AGT ATC GCC
201 E   V   V   L   A   K   E   A   G   I   C   Y   A   S   I   A 216
   ATG GCG ACA GAT TAT GAC TGC TGG AAG GAG CAC GAG GAA GCA GTT TCG
217 M   A   T   D   Y   D   C   W   K   E   H   E   E   A   V   S 232
   GTG GAC CGG GTC TTA AAG ACC CTG AAA GAA AAC GCT AAT AAA GCC AAA
233 V   D   R   V   L   K   T   L   K   E   N   A   N   K   A   K 248
   AGC TTA CTG CTC ACT ACC ATA CCT CAG ATA GGG TCC ACA GAA TGG TCA
249 S   L   L   L   T   T   I   P   Q   I   G   S   T   E   W   S 264
   GAA ACC CTC CAT AAC CTG AAG AAT ATG GCC CAG TTT TCT GTT TTA TTA
265 E   T   L   H   N   L   K   N   M   A   Q   F   S   V   L   L 280
   CCA AGA CAT TAA AGTAGCATGGCTGCCCAGGAGAAAAGAAGACATTCTAATTCCAGTCA
281 P   R   H
   TTTTGGGAATTCCTGCTTAACTTGAAAAAAATATGGGAAAGACATGCAGCTTTCATGCCCTTGCC
   TATCAAAGAGTATGTTGTAAGAAAGACAAGACATTGTGTGTATTAGAGACTCCTGAATGATTTAG
   ACAACTTCAAAATACAGAAGAAAAGCAAATGACTAGTAAACATGTGGGAAAAAATATTACATTTT
   AAGGGGGAAAAAAAAAAACCCCACCATTCTCTTCTCCCCCTATTAAATTTGCAACAATAAAGGGTG
   GAGGGTAATCTCTACTTTCCTATACTGCCAAAGAATGTGANGGAAGAAATGGGACTCTTTGGTTA
   TTTATTGATGCGACTGTAAATTGGTACAGTATTTCTGGAGGGCAATTTGGTAAAATGCATCAAAA
   GACTTAAAAATACGGACGTCCTTTGGTGCTGGGAACTCTACATCTAGCAATTTCTCTTTAAAACC
   ATATCAGAGATGCATACAAAGAATTATATATAAAGNAAGGGTGTTTAATAATGATAGTTATAATA
   ATAAATAATTGAAACAATCTGAATCCCTTGCAATTGGAGGTAAATTATGTCTTAGTTATAATCTA
   GATTGTGAATCAGCCAACTGAAAATNCCTTTTTGCATATTTCAATGTCCTAAAAAGACACGGTTG
   CTCTATATATGAAGTGAAAAAAGGATATGGTAGCATTTTATAGTACTAGTTTTGCTTTAAAATGC
   TATGTAAATATACAAAAAAACTAGAAAGAAATATATATAACCTTGTTATTGTATTTGGGGGAGGG
   ATACTGGGATAATTTTTATTTTCTTTGAATCTTTCTGTGTCTTCACATTTTTCTACAGTGAATAT
   AATCAAATAGTAAAGGGCCGTAAAAATAAAAGTGGATTTAGAAAGATCCAGTTCTTGAAAACACT
   GTTTCTGGTAATGAAGCAGAATTTAAGTTGGTAATATTAAGGTGAATGTCATTTAAGGGAGTTAC
   ATCTTTATTCTGCTAAAGAAGAGGATCATTGATTTCTGTACAGTCAGAACAGTACTTGGGTGTGC
   AACAGCTTTCTGAGAAAAGCTAGGTGTATAATAGTTTAACTGAAAGTTTAACTATTTAAAAGACT
   AAATGCACATTTTATGGTATCTGATATTTTAAAAAGTAATGTGAGCTTCTCCTTTTTATGAGTTA
   AATTATTTTATACGAGTTGGTAATTTGTGCCTTTTAATAAAGTGGAAGCTTGCTTTTTAAAAAAA
   AAAAAAAAGCGGAATTC
```

b

```
MTAP :   12 IGIIGGTGLDDPEILBGRTEKYVDTPFGKPSDALILGKIKNVDCILLARHGRQHTIMPSK 71
RRUORF:      G.IGG.G..D...LBG.....V..PFG..SD....G.........L.RHGR.H...PS.

MTAP :   72 VNYQANIWALKEBGCTHVIVTTACGSLREEIQPGDIVIIDQFIDRTTMRPQSFYDGSHSC 131
RRUORF:     VNY.ANI.ALK..G.T......A.GSL.E...PG..VI.DQFIDRT..R..SF
HSAPNP:                  G.....VT.A.G.L......GDI..I...I.......Q....G....
MMUPNP:                  G.....VT.A.G.L......GDI..I...I.......Q....G....
SCE8167:                      IVT.A.G......Q..D...I...........Q....G....

MTAP : 132 ARGVCHIPMAEPFCPKTREVLIETAKKLGLRCHSKGTMVTIEGPRFSSRAESFMFRTWGA 191
RRUORF:
HSAPNP:     ..G.....M........R.....T.K..G......GT.V...GP.F...AE.......GA
MMUPNP:     ..GV....M........R.......K..G......GT.V...GP.F...AES......GA
SCE8167:    ..G..............R..L....K.L.......GT.....GP.F..RAES.M.R..G.
BSUPNP:                                      KG......GP.....AE....RT.G.
MLEPNP:                                       G......GP.....AE..M.R..GA

MTAP : 192 DVINMTTVPEVVLAKEAGICYASIAMATD 220
RRUORF:                                  (54/114 identities)
HSAPNP:     D...M.TVPEV..A...G.........T  (39/135 identities)
MMUPNP:     D...M.TVPEV..A...G.........T  (40/135 identities)
SCE8167:    D...M.TVPEV..A...G.........T  (37/130 identities)
BSUPNP:     D...M.TVPEV..A..AG.....I     ( 20/50 identities)
MLEPNP:     D...M.TV.E...A..AG.........T  ( 18/53 identities)
```

FIG. 3. (*a*) Cloning of *MTAP*. The nucleotide sequence of the *MTAP* cDNA is shown along with the deduced amino acid sequence of the MTAP protein. (*b*) Protein sequence comparison of the MTAP protein with the highest scoring protein sequences found in a BLASTP search of the protein databases. The highest homology was detected with RRUORF, which shows 47% (54 of 114) amino acid identity. The function of this protein from *R. rubrum* is not known. The homology detected with other PNPs was lower but extended over a slightly larger region in eukaryotic proteins (HSAPNP, MMUPNP, and SCE8167). Only identical amino acids are shown. Dots denote nonidentical amino acids; no gaps were introduced into the alignment. RRUORF, hypothetical 25.8-kDa protein in cytochrome c_1 (*petC*) 3′ region from *Rhodospirillum rubrum* (Swiss-Prot accession no. P23139); HSAPNP, human PNP (EC 2.4.2.1; GenBank accession no. K02574); MMUPNP, murine PNP (GenBank accession no. L11292); SCE8167, L8167.19 gene product from *S. cerevisiae* (GenBank accession no. U14913);

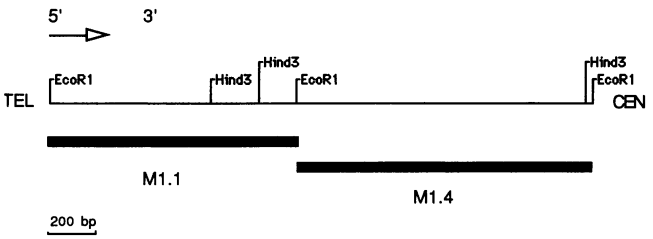

FIG. 4. Map of *MTAP* cDNA clone. Not all restriction sites are indicated.

the 25.8-kDa protein is distinct from the region of homology to the PNPs with only a minor overlap.

We correlated the presence or absence of a 1.4-kb subclone of the cDNA (probe M1.4, Fig. 4) with the presence or absence of MTAP enzyme activity in previously characterized cell lines (4, 5). This 1.4-kb probe is deleted in every cell line that lacks MTAP enzyme activity and is present in all cell lines with MTAP enzyme activity. When the 1.1-kb 5′ fragment of the cDNA probe (M1.1, Fig. 4) was used for Southern blot analysis of *Pst* I-digested human genomic DNA, four bands were seen (data not shown). One of these bands was always seen in cell lines with homozygous deletions of the 9p21 region. The M1.1 probe was then used on a somatic cell hybrid panel and found to hybridize only to human chromosomes 9 and 3. Thus, it appears that another gene or pseudogene homologous to *MTAP* maps to human chromosome 3.

DISCUSSION

The *MTAP* gene was previously localized (8) by using somatic cell hybrids. We refined the location by using information obtained from PFGE and cell lines (2, 15). Because there was no probe available for the *MTAP* gene, we concluded that the putative TSG must lie between the *IFN* gene cluster and the *MTAP* gene. By using similar reasoning, Coleman *et al.* (30) placed the SRO in melanomas centromeric to the *MTAP* gene and suggested that the SRO in melanoma was distinct from the SRO in gliomas, leukemias, and lung cancers. Barring any complex rearrangements in both T98G (glioma) and Hs294T (melanoma), we believe that the position of the TSG should be within the region defined by the homozygous deletions in these two cell lines. This region maps centromeric to *MTAP* and the *IFN* gene cluster but distal to *D9S966* and includes *CDKN2* (Fig. 1). This region corresponds to the only critical region that we have defined by using primary samples from patients with gliomas and leukemias (ref. 31 and unpublished data). The region overlaps the *MLM* locus because it maps in the 2-centimorgan region between *D9S736* and *D9S171* (9, 28). These data are consistent with recent results published by Jen *et al.* (32) who found a high frequency of homozygous deletions of *CDKN2* and *CDKN2B* in primary glioma samples. No point mutations of either gene were observed in primary gliomas.

Our long-range map covers 2.8 Mb including the *IFN* gene locus but does not reach *D9S126* or *D9S171*. There are now two reports (14, 33) of homozygous deletions on 9p that do not extend into the *CDKN2* locus. In fact, these two reports suggest that one of the other 9p TSGs is telomeric to *CDKN2*. To date, all the data available in primary tumors and tumor cell lines suggest that the preferred mechanism for gene inactivation on 9p is homozygous deletion rather than point mutations. We know of no other chromosomal region with such a high frequency of homozygous deletions. It is rather intriguing that all of the genes (the *IFN* gene cluster, *MTAP*, *CDKN2*, and

BSUPNP, *Bacillus subtilis* PNP (Protein Information Resource accession no. A42708); MLEPNP, *Mycobacterium leprae* PNP (GenBank accession no. U00022).

CDKN2B) identified thus far in this region could have some significant biological role in cancer. The most efficient way to inactivate all of these genes if they are biologically important would be by a large enough deletion. Alternatively, these genes may have been deleted as "innocent bystanders" because intrinsic fragility or recombinogenicity around the TSG may make the region a hot spot for illegitimate recombination.

It has been proposed that the inclusion of *MTAP* gene in these deletions may present an opportunity to use this phenomenon in drug development (8, 34, 35). MTAP is involved in the purine salvage pathway in which methylthioadenosine is recycled to the purine nucleotide pool. MTAP deficiency interferes with this salvage pathway. MTAP deficiency in human malignancy may permit the development of chemotherapeutic approaches in which MTAP-negative cancer cells will be selectively killed with drugs causing the depletion of purine nucleotides. This major difference between normal and malignant cells might be used to design more effective chemotherapy approaches in gliomas, lung cancer, and other solid tumors where there are currently no effective therapies. Thus, further characterization of this gene may have both diagnostic and therapeutic implications.

The data presented here support the fact that *CDKN2* is the most frequently homozygously deleted marker within 9p21 in gliomas and leukemias. However, it is apparent that it is not the only clinically relevant gene for every tumor type. Thus, it will be necessary to identify and map other candidate genes in this region. The reagents described herein should be useful in further characterization of the 9p tumor suppressor locus.

This work has benefited from discussions with many colleagues, including Drs. Janet Rowley, John Minna, Michael Lerman, and Manuel Diaz. We are grateful to Drs. Denis Le Paslier and Daniel Cohen for screening the CEPH MegaYac library; to Dr. Cannon-Albright for the *D9S736* primers information prior to publication; to Dr. Joseph Testa for the *D9S171* end-clone information; to Oncor for gridding the chromosome 9 library and for sequencing the *MTAP* cDNA clone; to Dr. Peter De Jong and the Lawrence Livermore Laboratory for the chromosome 9 flow-sorted library; to Dr. Marion Buckwalter for the pICL and pLUS vectors; to Drs. Sandra Bigner, Meenhard Herlyn, and Alan Linnenbach for providing us with tumor cell lines; and to Dr. David Beach for the *CDKN2* and *CDKN2B* cDNA probes. This research was supported by funds from the J. S. McDonnell foundation and the Cancer Center Core Grant from the National Cancer Institute (CA14599-19 to O.I.O.), U.S. Public Health Service Grant CA40046 to M.M.L.B., and Grant CA42557 and Department of Energy Grant DE-FG02-86ER60408 to Janet D. Rowley. S.G. was supported by National Institutes of Health Training Grant 5-T32GMO7183-20. M.H.D. was supported by a training grant from the Deutsche Forschungsgemeinschaft.

1. Mitelman, F. (1994) *Catalog of Chromosome Aberrations in Cancer* (Wiley/Liss, New York), 5th Ed.
2. Díaz, M. O., Zieman, S., Le Beau, M. M., Pitha, P., Smith, S. D., Chilcote, R. R. & Rowley, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5259–5263.
3. Díaz, M. O., Rubin, C. M., Harden, A., Ziemin, S., Larson, R. A., Le Beau, M. M. & Rowley, J. D. (1990) *New Engl. J. Med.* **322,** 77–82.
4. Olopade, O. I., Jenkins, R. B., Ransom, D. T., Malik, K., Pomykala, H., Nobori, T., Cowan, J. M., Rowley, J. D. & Diaz, M. O. (1992) *Cancer Res.* **52,** 2523–2529.
5. Olopade, O. I., Buchhagen, D. L., Malik, K., Sherman, J., Nobori, T., Bader, S., Nau, M. M., Gazdar, A. F., Minna, J. D. & Díaz, M. O. (1993) *Cancer Res.* **53,** 2410–2415.
6. Nobori, T., Miura, K., Wu, D. J., Lois, A. Takabayashi, K. & Carson, D. A. (1994) *Nature (London)* **368,** 753–756.
7. Kamb, A., Gruis, N. A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S. V., Stockert, E., Day, R. S. I., Johnson, B. E. & Skolnick, M. H. (1994) *Science* **264,** 436–440.
8. Carerra, C. J., Eddy, R. L., Shows, T. B. & Carson, D. A. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2665–2668.
9. Cannon-Albright, L. A., Goldgar, D. E., Meyer, L. J., Lewis, C. M., Anderson, D. E., Fountain, J. W., Hegi, M. E., Wiseman, R. W., Petty, E. M., Bale, A. E., Olopade, O. I., Diaz, M. O., Kwiatkowski, D. J., Piepkorn, M. W., Zone, J. J. & Skolnick, M. H. (1992) *Science* **265,** 1148–1152.
10. Serrano, M., Hannon, G. J. & Beach, D. (1993) *Nature (London)* **366,** 704–707.
11. Hannon, G. J. & Beach, D. (1994) *Nature (London)* **371,** 257–259.
12. Hussussian, C. J., Struewing, J. P., Goldstein, A. M., Higgins, P. A. T., Ally, D. S., Sheahan, M. D., Clark, W. H., Jr., Tucker, M. A. & Dracopoli, N. C. (1994) *Nature Genet.* **8,** 15–19.
13. Kamb, A., Shattuck-Eidens, D., Eeles, R., Liu, Q., Gruis, N. A., Ding, W., Hussey, C., Tran, T., Miki, Y., Weaver-Feldhaus, J., McClure, M., Aitken, J. F., Anderson, D. E., Bergman, W., Frants, R., Goldgar, D. E., Green, A., Maclennan, R., Martin, N. G., Meyer, L. J., Youl, P., Zone, J. J., Skonick, M. H. & Cannon-Albright, L. A. (1994) *Nature Genet.* **8,** 22–26.
14. Cheng, J. Q., Jhahwar, J. S., Klein, W. M., Bell, D. W., Lee, W.-C., Altomare, D. A., Nobori, T., Olopade, O. I., Buckler, A. J. & Testa, J. R. (1994) *Cancer Res.* **54,** 5547–5551.
15. Olopade, O. I., Bohlander, S. K., Pomykala, H., Maltepe, E., Nobori, T., Le Beau, M. M. & Diaz, M. O. (1992) *Genomics* **14,** 437–443.
16. Weaver-Feldhaus, J., Gruis, N. A., Neuhausen, S., Le Paslier, D., Stockert, E., Skolnick, M. H. & Kamb, A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7563–7567.
17. Diaz, M. O., Pomykala, H. M., Bohlander, S. K., Maltepe, E., Malik, K., Brownstein, B. & Olopade, O. I. (1994) *Genomics* **22,** 540–552.
18. Kwiatkowski, D. J. & Diaz, M. O. (1992) *Hum. Mol. Genet.* **1,** 658.
19. Bohlander, S. K., Dreyling, M. H., Hagos, F., Sveen, L., Olopade, O. I. & Diaz, M. O. (1994) *Genomics* **24,** 211–217.
20. Feinberg, A. P. & Vogelstein, B. (1983) *Ann. Biochem.* **132,** 6–13.
21. Bohlander, S. K., Espinosa, R., Fernald, A. A., Rowley, J. D., Le Beau, M. M. & Diaz, M. O. (1994) *Cytogenet. Cell Genet.* **65,** 108–110.
22. Rowley, J. D., Díaz, M. O., Espinosa, R., Patel, Y. D., van Melle, E., Ziemin, S., Taillon-Miller, P., Lichter, P., Evans, G. A., Kersey, J. H., Ward, D. C., Domer, P. H. & Le Beau, M. M. (1990) *Proc. Natl. Acad. Sci. USA* **97,** 9358–9362.
23. Hermanson, G. G., Hoekstra, M. F., McElligot, D. L. & Evans, G. A. (1991) *Nucleic Acids Res.* **19,** 4943–4948.
24. Buckler, A. J., Chang, D. D., Graw, S. L., Brook, D. J., Haber, D. A., Sharp, P. A. & Housman, D. E. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 4005–4009.
25. Lovett, M., Kere, J. & Hinton, L. M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 9628–9632.
26. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
27. Porterfield, B. W., Diaz, M. O., Rowley, J. D. & Olopade, O. I. (1992) *Proc. Am. Assoc. Cancer Res.* **33,** 72 (abstr.).
28. Povey, S. & the Utah Genome Center Genetic Marker and Mapping Group (1994) *Ann. Hum. Genet.* **58,** 177–250.
29. Fountain, J. W., Karayiogu, M., Taruscio, D., Graw, S. L., Buckler, A. J., Ward, D. C., Dracopoli, N. C. & Housman, D. E. (1992) *Genomics* **14,** 105–112.
30. Coleman, A., Fountain, J. W., Nobori, T., Olopade, O. I., Robertson, G. & Housman, D. E. (1994) *Cancer Res.* **54,** 344–348.
31. Dreyling, M. H., Bohlander, S. K., Adeyanju, M. O. & Olopade, O. I. (1995) *Cancer Res.* **55,** 984–988.
32. Jen, J., Harper, W., Bigner, S. H., Bigner, D. D., Papadoupoulos, N., Markowitz, S., Willson, J. K. V., Kinzler, K. W. & Vogelstein, B. (1994) *Cancer Res.* **54,** 6353–6358.
33. Lydiatt, W. A., Murty, V. V. V. S., Davidson, B. J., Xu, L., Dyomina, K., Sacks, P. G., Schantz, S. P. & Chaganti, R. S. K. (1994) *Genes Chromosomes Cancer*, in press.
34. Della Ragione, F., Palumbo, R., Russo, G. L., Gragnaniello, V. & Zappia, V. (1992) *Biochem. J.* **281,** 533–538.
35. Nobori, T., Szinai, I., Amox, D., Parker, B., Olopade, O. I., Buchhagen, D. L. & Carson, D. A. (1993) *Cancer Res.* **53,** 1098–1101.