Ectopic Sequences From Truncated *HMGIC* in Liposarcomas Are Derived From Various Amplified Chromosomal Regions

Leonardo A. Meza-Zepeda,[1] Jeanne-Marie Berner,[1] Jørn Henriksen,[1] Andrew P. South,[2] Florence Pedeutour,[3] Anine B. Dahlberg,[1] Linda H. Godager,[1] Dean Nižetić,[2] Anne Forus,[1] and Ola Myklebost[1]*

[1]Department of Tumor Biology, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo, Norway
[2]Centre for Applied Molecular Biology, School of Pharmacy, University of London, London, U.K.
[3]Laboratoire de Génétique, Hôpital de l'Archet, Nice, France

The *HMGIC* gene codes for an architectural transcription factor frequently rearranged by translocation in lipomas and other benign mesenchymal tumors. In sarcomas, malignant tumors of mesenchymal origin, the gene is also found to be rearranged, but in addition amplified and overexpressed. Here we report the sequence, chromosomal localization, and expression patterns of 11 novel ectopic sequences fused to exons 2 and 3 of *HMGIC* in seven different sarcoma samples. In addition, we identified a number of variant transcripts observed previously in benign tumors. Consistent with the suggested role of *HMGIC* in adipocytic differentiation, most of the novel ectopic sequences were observed in well-differentiated liposarcomas. These tumors are known to have complex marker chromosomes containing amplified segments from several chromosomes. Five novel sequences were derived from 12q14–q15, where *HMGIC* resides, two from 1q24, a region frequently amplified in these types of tumors, two from 11q14, and one from chromosome 2. All except one of the aberrant transcripts encoded truncated proteins with intact DNA-binding domains (AT hooks) but lacking the C-terminal acidic region, a target for constitutive phosphorylation by protein kinase CK2. Some of the ectopic sequences were transcribed in other tissues, and most of the ectopic sequences also showed recurrent amplification in liposarcomas.

INTRODUCTION

The high-mobility group protein HMGIC is characterized by three DNA binding domains (AT hooks) with the ability to bind DNA at AT-rich regions. The protein acts as an architectural transcription factor, modulating DNA conformation (Mantovani et al., 1998; Schwanbeck et al., 2000). Rearrangement of the *HMGIC* gene by reciprocal translocations is frequently observed in benign tumors like lipomas, lung hamartomas, uterine leiomyomas, endothelial polyps, fibroadenomas, and adenolipomas of the breast (Ashar et al., 1995; Kazmierczak et al., 1995, 1996; Schoenmakers et al., 1995). In most of these cases, the translocation breakpoints occur after exon 3, resulting in a chimeric transcript containing exons 1 to 3, coding for the DNA binding domains (AT hooks), fused to various sequences. Because various partner chromosomes are involved, it is thought that the loss of the C-terminal end of HMGIC is important, rather than the various novel sequences added (Fedele et al., 1998). In a functionally related family member, HMGI(Y), the C-terminal part of the protein is a target for constitutive phosphorylation by protein kinase CK2 and it appears to be important for proper conformation and function, since phosphorylation reduces the DNA binding affinity (Wang et al., 1995).

In contrast to the intragenic breakpoints frequently observed in lipomas, a significant number of leiomyomata with t(12;14) contain aberrations outside the transcribed part of the *HMGIC* locus. These rearrangements show breaks 10–100 kb upstream of *HMGIC*, suggesting the disturbance of regulatory sequences, which leads to an abnormal expression of wild-type *HMGIC* (Schoenberg Fejzo et al., 1996). Recent studies of uterine leiomyomas with t(12;14) indicates that upstream rearrangement also can lead to structural changes of the *HMGIC* region, showing preferential fusion of *HMGIC* with the *RAD51B* gene (Schoenmakers et al., 1999). However, in these cases, it seems that the *RAD51B*, involved in DNA repair, may be functionally important, and in these tumors it is the

Supported by: The Research Council of Norway, The Norwegian Cancer Society, The Nordic Cancer Union, Association pour la Recherche contre le Cancer and Ligue Nationale Contre Le Cancer, Région PACA, Comité du Var.

*Correspondence to: Ola Myklebost, Department of Tumor Biology, Institute for Cancer Research, The Norwegian Radium Hospital, N-0310, Oslo, Norway. E-mail: olam@radium.uio.no

Received 28 September 2000; Accepted 21 December 2000

3′ part of the *HMGIC* that is included in fusion transcripts as a result of aberrant splicing.

HMGIC is highly expressed during embryogenesis, in transformed cells and human tumors, but not at significant levels in normal adult tissues (Rogalla et al., 1996; Gattas et al., 1999). Direct evidence for a role for HMGIC protein in cellular transformation was demonstrated by the expression of an antisense *HMGIC* mRNA, preventing the transformation of thyroid cells by *v-mos* and *[v-]Ki-ras* oncogenes (Berlingieri et al., 1995). More recently, it was demonstrated that overexpression of a truncated form of HMGIC protein, containing the three DNA binding domains with or without an ectopic C-terminal peptide, is capable of causing neoplastic transformation of NIH3T3 cells (Fedele et al., 1998).

We have shown that *HMGIC* is frequently amplified and overexpressed in human sarcomas, malignant tumors of mesenchymal origin (Berner et al., 1997). In some cases, we also detected several *HMGIC* transcripts of aberrant size that lack 3′ sequences, and identified a chimeric product with an ectopic sequence fused to the DNA binding domains in one sample. This transcript encoded a truncated protein lacking the C-terminal acidic region, similar to the *HMGIC* fusions from benign mesenchymal tumors.

In this study, we have further analyzed aberrant *HMGIC* transcripts from 13 human sarcomas. We report the cloning and structure of 11 novel ectopic sequences fused to *HMGIC* exons 2 and 3 and describe their chromosomal origin as well as expression and amplification status in sarcoma samples.

MATERIALS AND METHODS

Specimens

Twenty-five human sarcomas and three normal tissues (breast, placenta, and muscle) were used in this study. The clinical samples chosen had previously been screened for amplification and rearrangement of *HMGIC* (Berner et al., 1997). The sarcoma samples examined were 1 hemangiopericytoma (HP), 1 fibrosarcoma (FS), 1 nonclassified sarcoma (NCS), 2 malignant fibrous histiocytomas (MFH), 2 malignant peripheral nerve sheath tumors (MPNST), 1 osteosarcoma cell line (OSA) (Roberts et al., 1989), 1 rhabdomyosarcoma cell line (RMS13) (Roberts et al., 1989), 2 osteosarcomas (OS), and 14 liposarcomas (LS), including 7 well-differentiated liposarcomas. From LS22, two different tumors were analyzed, LS22-OS and LS22-2a. LS22-OS was from a recurrent tumor, which showed well-differentiated areas with benign ossification. LS22-2a was from a metastatic lesion in the left lung. Clinical specimens were collected immediately after surgery, cut into small pieces, frozen in liquid nitrogen, and stored at −70°C until use. Nine samples tested were grown subcutaneously in immunodeficient mice as xenografts (suffix "x"). All tumors were classified according to the WHO International Histological Classification of Tumors (Schajowicz, 1993; Weiss, 1994).

Rapid Amplification of cDNA Ends (3′ RACE)

RACE amplification of the 3′ end ectopic sequences was performed as previously described (Berner et al., 1997). Briefly, first-strand cDNA synthesis was performed using an adapter primer (3′ RACE System, Gibco-BRL). PCR amplification of the 3′ end of *HMGIC* was achieved using a gene-specific primer and a universal amplification primer. The second round of amplification used a nested gene-specific primer. Cloning of the amplified products was performed using the CloneAmp system (Gibco-BRL). Determination of the sequence content of the amplified fragments was accomplished by cycle sequencing using dye terminator chemistry (Amersham Pharmacia Biotech) and an ALFexpress automatic sequencer (Amersham Pharmacia Biotech).

Southern Blot Analysis

DNA extraction from tumor samples, preparation of blots, hybridization, and a stripping procedure were performed as described previously (Forus et al., 1993). Probes were sequentially hybridized to two identical filters. Scoring of the signal intensities was done visually. Relative signal intensity at least threefold stronger than those obtained from the normal control (peripheral blood leukocytes) was scored as amplification. The signals from specific bands were corrected for unequal sample loading by comparison to an *APOB* (2p24) control probe. Previous work has shown no amplification of this chromosomal region in our sarcoma panel (Forus et al., 1995a, 1995b; Simons et al., 2000); in addition, CGH analysis of liposarcoma from other groups showed no recurrent alterations in 2p24 (Szymanska et al., 1996; Parente et al., 1999).

Northern Blot Analysis

RNA isolation from clinical samples, electrophoresis, transfer to nylon membranes, and hybrid-

ization were performed as described (Forus et al., 1993). Probes were sequentially hybridized, and signals from an 18S rRNA oligonucleotide were used to correct for unequal sample loading.

Screening of Genomic Library

A P1 artificial chromosome library (RPCI-1) was screened by Southern hybridization. The high-density gridded library was initially screened using a pool of radioactive-labeled ectopic sequences as probes. Positive clones were reprobed sequentially using individual ectopic sequences to identify the corresponding PACs for each probe. The first and second screening of the library was performed as previously described (Groet et al., 1998).

Probes

Probes corresponding to each ectopic sequence were obtained by PCR amplification of the desired region from the cloned RACE products. Enzymatic amplification of the ectopic sequences was performed using an *HMGIC*-specific primer complementary to the terminal segment of exon 3 and the SP6 promoter primer present in the pAMP1 cloning vector (Gibco-BRL). As probe for *HMGIC*, we used the 5′ end of the genes, including the complete protein coding region (Schoenmakers et al., 1995; Berner et al., 1997). The control probe *APOB* (clone pB27) gene was kindly provided by D.J. Breslow (Huang et al., 1985).

Fluorescence In Situ Hybridization

Normal metaphase chromosomes from peripheral leukocytes and interphase nuclei from frozen tumor tissues were prepared as described (Forus et al., 1998). PAC DNA was isolated according to standard procedures and labeled with biotin-14d-ATP or digoxygenin-11-dUTP (Roche Biochemical) by nick translation. For each hybridization, 200–500 ng of labeled DNA was prehybridized with a 50- to 100-fold excess of human Cot-1 DNA.

Slides were treated with pepsin (4 mg/ml) for 10 min (metaphases) or 20 min (interphases) at 37°C and washed in PBS, then with 1% formaldehyde in PBS for 10 min at room temperature. Slides were denatured in 70% formamide/2 × SSC for 1.5 min at 70°C (metaphases) or 2 min at 74°C (interphases), and hybridization and washes were as described previously (Forus et al., 1998). For detection of digoxygenin, we used fluorescein isothiocyanate (FITC)-conjugated sheep antidigoxygenin antibody (Roche Biochemicals) followed by FITC- or ALEXA488-labeled donkey antisheep (Molecular Probes, The Netherlands). Detection of biotin-labeled probes was performed using avidin-conjugated Cy3 (Amersham Pharmacia Biotech).

Hybridized slides were examined visually using a Zeiss Axioplan microscope equipped with appropriate single bypass filters for excitation of DAPI, double bypass filters for DAPI/FITC and DAPI/rhodamine (Cy3), and triple bypass filters for excitation of DAPI/FITC/rhodamine (Cy3). For each probe, localization of the signal was evaluated in at least 10 metaphases, and copy numbers were evaluated in at least 200 nuclei. Images were captured using a Leica DMR microscope (Wetzlar, Germany) coupled to a Photometrics camera and software (Photometrics, Tucson, AZ).

RESULTS

Isolation and Characterization of 3′ Ectopic Sequences

Thirteen sarcomas with identified aberrations in *HMGIC* or its mRNA (Berner et al., 1997) were selected for this study (Table 1). From all of these, RACE products could be generated using nested 5′ primers from *HMGIC* exon 1, and further cloned and sequenced (Table 1). Both in the additional clinical samples and in the two normal tissues (breast and placenta), we detected variant species of *HMGIC* mRNA that have been described previously (LM-30.1, Li-538, Li-192) (Schoenmakers et al., 1995; Kottickal et al., 1998). Novel ectopic sequences fused to exon 2 or 3 of *HMGIC* were obtained from seven tumors and were denoted FHSn, for fused to *HMGIC* in sarcomas, as decided previously (Table 1) (Berner et al., 1997).

The nucleotide sequence of the novel fusion transcripts revealed that all except FHS4 contain exons 1, 2, and 3 of *HMGIC* fused to ectopic sequences. These sequences coded for novel amino acids replacing the acidic domain at the C-terminal end of the HMGIC protein (Tables 1 and 2). FHS4 was fused to exon 2 of *HMGIC* and added 49 new amino acids C-terminally (Table 2). Interestingly, the FHS4 sequence was part also of FHS3, which in addition to exon 3 contained 386 bp 5′ to FHS4.

When the ectopic sequences were compared to Genbank, FHS2 showed identity at the nucleotide level to the human *NOT2* gene, a negative transcriptional regulator (Albert et al., 2000). The deduced open reading frame for the fused transcript included the N-terminal portion of *HMGIC* (exons 1–3) fused in frame to the C-terminal domain of *NOT2* (Table 2). No homology to previously described genes was identified for the remaining 10

TABLE 1. Sequences Fused to *HMGIC* Transcripts

	RACE results			*HMGIC*[b]	
Sample[a]	Breakpoint[c]	Sequence[d]	Size[e]	Amplification	Rearrangement
HP1x	Exon 4	LM-30.1	174 bp	A	R
LS2	Exon 3	**FHS2**	85 bp	A	
	Exon 3	**FHS3**	533 bp		
	Exon 2	**FHS4**	147 bp		
LS6	Exon 3	Li-192	77 bp	A	
LS9	Exon 3	Li-192	77 bp	N	
	Exon 3	**FHS5**	195 bp		
	Exon 4	LM-30.1	174 bp		
LS11	Exon 3	Li-192	77 bp	A	R
	Exon 3	**FHS6**	150 bp		
LS13	Exon 3	**FHS7**	153 bp	N	
	Exon 3	**FHS8**	130 bp		
LS22-OS	Exon 3	Li-192	77 bp	A	R
	Exon 3	LM-30.1	174 bp		
	Exon 3	**FHS9**	198 bp		
LS22-2a	Exon 3	Li.-538	383 bp	A	Not determined
	Exon 3	Li-192	77 bp		
	Exon 3	**FHS10**	200 bp		
	Exon 4	LM-30.1	174 bp		
LS28x	Exon 3	**FHS11**	25 bp	A	R
	Exon 3	**FHS12**	206 bp		
LS35x	Exon 3	Li-192	77 bp	N	
	Exon 4	LM-30.1	174 bp		
LS38	Exon 3	Li-192	77 bp	A	
	Exon 4	LM-30.1	174 bp		
LS41	Exon 3	**FHS1**	157 bp	A	
LS42	Exon 4	LM-30.1	174 bp	A	
Breast	Exon 4	LM-30.1	174 bp	N	
Placenta	Exon 4	LM-30.1	174 bp	N	

[a]HP, hemangiopericytoma; LS, liposarcoma.
[b]Previously described in Berner et al. (1997).
[c]Ectopic sequence fused to *HMGIC* exon 2, 3, or 4.
[d]LM-30.1, Li-538, and Li-192 have previously been described (Schoenmakers et al., 1995). FHS1 was described previously (Berner et al., 1997).
[e]Size of the cloned ectopic sequence fused to *HMGIC*.

ectopic sequences, but 3 ectopic sequences, FHS1, FHS3, and FHS4, showed identity to expressed sequence tags (ESTs) (accession numbers AA084073, AI916807, and AI655157, respectively, among others). Sequences corresponding to FHS5 (from LS9), FHS7 (from LS13), and FHS8 (from LS13) identified sequences derived from P1 artificial chromosome clones (PACs) from 1q24, thus giving immediate chromosomal localization (Table 2). In the same way, different genomic clones from chromosome 2, 11q14, 12q, and 15, were identified for other ectopic sequences by sequence homology search (Table 2).

Chromosomal Localization of Origin of Ectopic Sequences

Specific genomic clones for seven ectopic sequences were identified by screening of a P1 artificial chromosome library (Table 2). The isolated PACs were used in FISH experiments to determine the chromosomal origin of the corresponding ectopic sequences (Table 2, Fig. 1). Most of the sequences were derived from 12 q13–q15, where the *HMGIC* gene is located. Three ectopic sequences, FHS5, FHS7, and FHS8, generated from two different clinical samples, were localized to 1q24, and for the first two, the localization by FISH could be confirmed by sequence search in genome databases (Table 2). The PAC (2M17) detected with FHS6 was localized by FISH to chromosome band 12q14. On the other hand, the sequence of FHS6 was contained in BAC RP11-366L20, supposedly localized to chromosome 15 (accession number AC022608). However, further analysis of RP11-366L20 revealed that it also contained the *HMGIC* gene, indicating that the localization to chromosome 15 was an error.

TABLE 2. Liposarcoma-Associated Fused Sequences: Deduced Amino Acid Sequence and Chromosomal Origin

Ectopic sequence	Accession number ESTs[a]	C-terminal sequence[b]	Genomic clone		Chromosomal origin	
			BAC or PAC	Accession number	By sequence	By FISH
FHS1	**U73513**[c]	AQRRGHVRTW	238G16			12q14–15
(LS41)	AA084073		RP11-401I10	AC025157	12q	
	AL040600					
FHS2	**BE352844**	EFHLEYDKLEERPHLPSTFNYNPAQ	273J24			12q14.3
(LS2)		QAF	RP11-588H23	AC025569	12q	
			RP11-792C16	AC011612	12q	
FHS3[d]	**BE352845**	GFRQAVSPRPV				
(LS2)	AI916807		RP11-762L8	AP001148	11q14	
	AI073989		RP11-185J12	AC018915	11	
FHS4[d]	**BE352846**	AAQGLGMGSSFIMAAGVLLTQPC	RP11-762L8	AP001148	11q14	
(LS2)	A1655157	CCSMTLCTKLFSSVMNSGNNKYFISA	RP11-185J12	AC018915	11	
	A1218035		RP11-625M1	AC021168	11	
FHS5	**BE352847**	I	94B2			1q24
(LS9)			304J10			1q24
			RP11-58A9	AL031285	1q24.1–q24.3	
			135I12	AC003115	1	
FHS6	**BE352848**	CDKAEHPRDMLQVS	2M17			12q14
(LS11)			RP11-366L20	AC022608	15	
			RP11-677M24	AC021472		
FHS7	**BE352849**	HCSQ	50C7			1q24
(LS13)			RP1-79C4	Z97200	1q24	1q24
FHS8[d]	**BE352850**	EKNLFILSVT				
(LS13)			RP11-545I10	AL162399	1q24	
FHS9[d]	**BE352851**	ACEILNKEPTKFCPDFSLTEQ				
(LS22-OS)			RP11-389F18	AC021781	2	
FHS10	**BE352842**	PGS	41C6			12q14
(LS22-2a)			RP11-677M24	AC021472		
FHS11[d,e]		QGLAGPGS				
(LS28x)						
FHS12	**BE352843**	STCPSAVNVMRGPCWEHVLLGI	142B23			12q13–14
(LS28x)		SPSLCPVPTPPRRCSCSEPTPHLS				
		EATQRAPCGQSWILALNKKSESWEG				

[a]Accession number of novel ectopic sequences (**bold**) and ESTs.
[b]Ectopic amino acid sequence, continuation of the ORF of *HMGI-C* after exon 2 or 3.
[c]Previously described (Berner et al., 1997).
[d]Not screened for genomic clone.
[e]No accession number or chromosomal localization assigned due to short length of the sequence.

Copy Number Determination by Southern Blot and FISH Analysis

We have further analyzed whether the eight ectopic sequences were amplified in a panel of 19 sarcomas by Southern blot analysis (Figs. 2 and 3). Seven out of eight FHS sequences analyzed were found to be amplified in the sample from which they were cloned, whereas the last sample, FHS7, could not be analyzed in the original tumor. However, all eight were amplified in additional sarcoma samples (Fig. 3). Ectopic probes generated from LS11 (FHS6), LS22-2a (FHS10), and LS28x (FHS12; Figs. 2 and 3), all originating from 12q14, detected amplification in a larger number of clinical samples than those from LS41 and LS13 (FHS1 and FHS7; Fig. 3). However, all probes detected amplification in multiple samples (Fig. 3), indicating that they represent recurrently amplified regions.

A selected group of PACs representing FHS sequences with 12q14 and 1q24 origin were used for interphase FISH analyses in one pleomorphic and four well-differentiated liposarcomas (Figs. 1 and 4). Increased numbers of signals were detected in more than 60% of the nuclei for almost all samples. Thus, the corresponding genomic regions were consistently overrepresented in this group of tumors.

Northern Blot Analysis

In several cases, tumors with amplification expressed a population of *HMGIC* mRNAs (Fig. 2), some of which were much larger than the wild-type

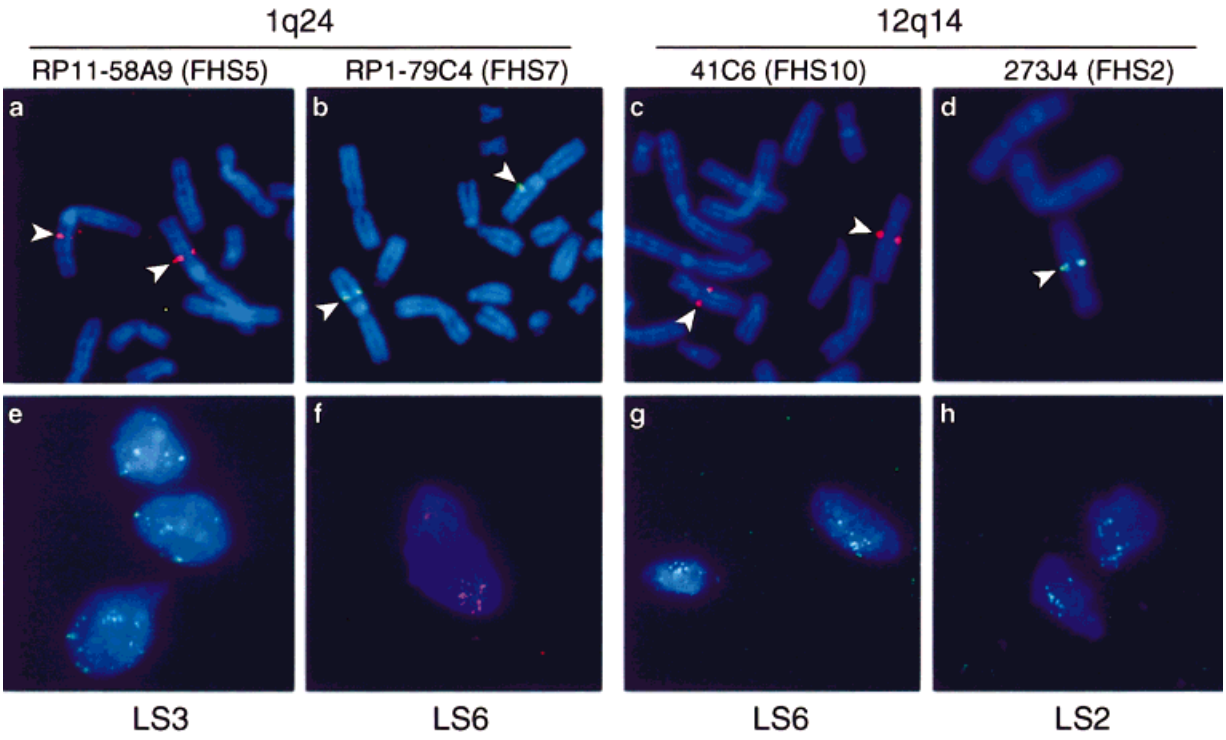

Figure 1. Chromosomal localization and detection of ectopic sequence amplification by FISH analysis. **a** and **b:** Cy3 (red)- and FITC (green)-labeled PAC RP11-58A9 and RP1-79C4 were hybridized to normal metaphase chromosomes. A strong signal was detected on 1q24, revealing the chromosomal origin of FHS5 and FHS7 (arrowheads). **c** and **d:** PACs 41C6 and 273J4 were localized to 12q14 (arrowheads). PACs were also hybridized to interphase nuclei from sarcoma samples. **e** and **f:** Amplification detected with PACs RP11-58A9 and RP1-79C4 in interphase nuclei from LS3 and LS6, respectively, and (**g** and **h**) with PACs 41C6 and 273J4 in LS6 and LS2, respectively. LS, liposarcoma.

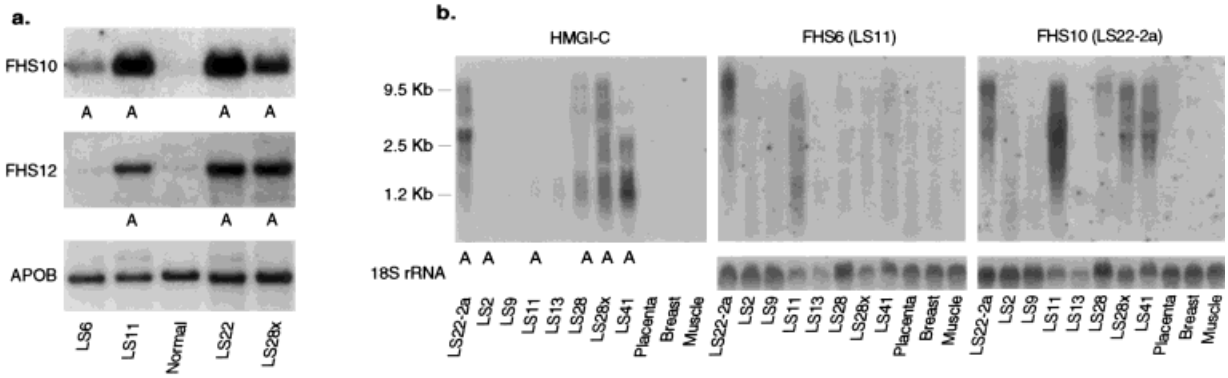

Figure 2. Southern and Northern blot hybridization showing amplification and overexpression of ectopic sequences. **a:** Representative Southern blot hybridizations demonstrating amplification of FHS10 and FHS12. A weak signal was observed on the normal leukocyte DNA (Normal) for each of the ectopic sequences tested. Different degrees of amplification were seen in clinical samples. Hybridization with a probe for *APOB* was used to correct for unequal sample loading. An "A" below the panel indicates that the corresponding ectopic sequence is amplified more than threefold. **b:** Expression of *HMGIC*, FHS6, and FHS10 in tumor and normal tissue was analyzed by Northern blots. *HMGIC* detects transcripts of aberrant size in different tumor samples. An "A" below the *HMGIC* blot indicates more than threefold amplification of the gene. All probes failed to detect transcripts in normal tissues. Hybridization to an 18S rRNA oligonucleotide probe is shown as loading control. *HMGIC* and FHS10 were sequentially hybridized to the same filter and FHS6 to a duplicate. LS, liposarcoma.

transcript, and by RACE we probably selected only the 5′ part of some representatives of these populations. We thus investigated whether probes from the ectopic parts of the fusions could detect the same bands as did the *HMGIC* probe. Eight ectopic sequences were investigated by Northern hybridization in seven patient samples, one xenograft, and three normal tissues, and compared to the results obtained with an *HMGIC* cDNA probe. Three of the eight ectopic probes tested (FHS3,

Tumour	*HMGIC*		FHS1 LS41		FHS2 LS2		FHS3 LS2		FHS5 LS9		FHS6 LS11		FHS7 LS13		FHS10 LS22-2a		FHS12 LS28x	
	DNA*	RNA	DNA	RNA	DNA	RNA	DNA	RNA	DNA	RNA	DNA	RNA	DNA	RNA	DNA	RNA	DNA	RNA
FS2	N		N		N		N		■		N		■		N		■	
HP1x	■		N		■		N		N		■		N		■		N	
LS2	■	1.4	■	-	■	2.5	■	-	N	-	■	-	N	-	■	-	N	-
LS6	■		N		N		N		■		■		■		■		N	
LS9	N	3.1 [a]	N	-	N	2.5	N	-	■	-	■	-	N	-	N	-	N	-
LS11	■	3.0 [a]	N	-	N	-	N	-	N	-	■	1.4	N	-	■	9.5 [d]	■	-
LS13	N	1.4		-		-		-		-		-		-		-		-
LS22-2a	■	3.3 [b]	■	-	N	-	N	-	N	-	■	9.5	N	-	■	9.5 [d]	■	-
LS28	■	1.4		-		2.5		-		-		-		-		9.5		1.4
LS28x	■	1.4	N	-	N	2.5	■	-	N	-	N	-	N	-	■	9.5 [d]	■	1.4
LS38	■		N		N		N		N		■		N		■		■	
LS41	■	1.2 [c]	■	1.2 [c]	■	2.5	N	-	N	-	■	1.4	N	-	■	9.5 [d]	■	-
MFH3x	N		N		N		N		N		N		N		N		N	
MFH25	N		N		■		N		N		N		N		N		N	
MPNST2x	■		N				N		N		■		N		■		N	
MPNST8x	N		N		N		N		N		N		N		N		■	
NCS2x	■		N		N		N		N		■		N		■		■	
OS6x	N		N		N		N		N		N		N		N		■	
OS11x	■		N		■		N		N		■		N		■		■	
OSAcl	■		N		N		N		N		N		N		N		■	
RMS13cl	N		N		N		N		N		N		N		N		■	

Not available N Normal copy number ■ Amplified more than 3-fold Transcript detected

[a] Additional band of 1.4 kb
[b] Additional bands of 6.0 and 9.5 kb
[c] Additional bands of 1.4 and 2.5 kb
[d] Additional band of 3.3 kb
* Described previously (Berner et al., 1997)
- : Not expressed

FS, fibrosarcoma; HP, hemangiopericytoma; LS, liposarcoma; MFH malignant fibrous histiocytoma; MPNST, malignant peripheral nerve sheath tumour; NCS, non-classified sarcoma; OS, osteosarcoma; RMS, rhabdomyosarcoma

Figure 3. Summary of estimated copy number and expression of fused sequences.

FHS5, FHS7) failed to give detectable signals, indicating that they may represent minor components of the corresponding mRNA population, or that these ectopic mRNAs do not have a discrete size, due to inaccurate splicing or termination (Fig. 3). The remaining five detected transcripts in the sample from which they were isolated, and in some cases (FHS2, FHS6, and FHS10) also in additional samples (Fig. 3). A Northern blot using *HMGIC*, FHS6, and FHS10 as probes is shown in Figure 2. In most of these cases, the probes recognized transcripts of the same molecular weight as did the *HMGIC* probe (Fig. 2b). Strong signals were obtained from samples with amplification of the gene, but no ectopic or *HMGIC* transcripts could be detected in any of the normal tissues evaluated. FHS6 detected a transcript in LS11, the sample it was derived from, and an additional high-molecular-weight signal was observed in LS22-2a. FHS10 recognized transcripts in LS22-2a, from where it originated, and was also expressed in LS28, LS28x, and LS41. A summary of the expression results is presented in Figure 3.

DISCUSSION

We have previously reported recurrent amplification and overexpression of *HMGIC* in human sarcomas, as well as rearrangement of the gene (Berner et al., 1997). In several cases, we found *HMGIC* transcripts of aberrant size, lacking the 3′ sequences of the gene. This situation resembles that in benign tumors, where a truncated HMGIC protein is generated, with the DNA-binding AT hooks intact, but lacking the C-terminal acidic domain (Schoenmakers and Van de Ven, 1997). It seems that the truncated protein induces neoplastic growth in the benign tumors, where it is expressed at low levels, and we suggested that additional overexpression due to amplification could promote malignant progression. Such oncogenic potential is supported by recent reports showing lipomatosis and a high incidence of lipomas in transgenic mice with truncated HMGIC (Battista et al., 1999; Arlotta et al., 2000) and neoplastic transformation of murine fibroblasts upon transfection (Fedele et al., 1998). In both of these cases, the genes were artificially overexpressed.

In the present study, we have identified and characterized a number of novel tumor-associated rearranged *HMGIC* transcripts and determined the chromosomal localization of their ectopic component. The aim was to determine the structural basis of these rearrangements both at the amino acid–coding and chromosomal level. Northern blot anal-

Sample	Signals per nuclei	1q24 origin			12q14-15 origin	
		94B2	RP1-79C4	RP11-58A9	41C6	273J24
LS2	≤2	39 %	85 %		3 %	36 %
	3-9	19 %	15 %	74 %	42 %	41 %
	≥10	42 %		26 %	55 %	23 %
LS3x	≤2	16 %	57 %	33 %	65 %	69 %
	3-9	38 %	43 %	23 %	35 %	31%
	≥10	46 %		44 %		
LS6	≤2	35 %	15 %	25 %	20 %	63 %
	3-9	15 %	15 %	17 %	39 %	37 %
	≥10	50 %	70 %	58 %	41 %	
LS9	≤2	35 %	37 %	29 %	31 %	27 %
	3-9	34 %	56 %	55 %	63 %	55 %
	≥10	31 %	7 %	16 %	6 %	18 %
LS43	≤2	29 %	63 %	73 %	29 %	29 %
	3-9	51 %	23 %	27 %	60 %	36 %
	≥10	20 %	14 %		11 %	35 %

More than 60% of the nuclei with copy number increases

LS, liposarcoma

Figure 4. Summary of DNA copy number by FISH analysis on interphase nuclei.

ysis revealed aberrant mRNAs for several tumors, and in a number of cases strong smears were observed, most likely due to imprecise splicing, transcription termination, and heterogeneity between amplified gene copies.

Using RACE analysis, we have identified novel chimeric and variant *HMGIC* transcripts, some of which were also found in normal tissues. Whereas the majority of the cloned transcripts had novel sequences replacing exons 4–5 of *HMGIC*, in LM-30.1 only exon 5 was replaced. This is apparently a normal alternate splicing variant or intermediate, as it was found also in normal tissues. A second class of variant transcripts (Li-538, Li-192) is recurrently observed in lipomas, although the translocation partner of *HMGIC* varies (Schoenmakers et al., 1995). These transcripts were also observed in samples with amplification, in some cases in addition to bona fide ectopic fusions (e.g., in LS9), but not in normal tissues. Most likely, these variants have been aberrantly spliced because the normal 3′ splice site is lost as a result of translocation or rearrangement, and adjacent sequence, probably from the large intron 3, is recruited for splicing. In all other cases, with one exception, the novel ectopic sequences were spliced onto exon 3, consistent with the proposed oncogenic hypothesis. The exception was FHS4, which had also lost the third exon of *HMGIC*. Interestingly, two other ectopic transcripts containing exon 3 were observed in this tumor, and the sequence of FHS4 was also found at the 3′ end of one of them (FHS3). Thus, in this case, it seems likely that two different splice variants from the same genomic rearrangement have been cloned, and in one of them exon 3 has been spliced out. It is possible that the overrepresentation of rearrangements within intron 3 rather than 2 is caused simply by its much larger size (140 kb). Thus, we cannot conclude from the present data that exon 3 is needed for oncogenic activity. The ectopic peptides reported here have no common pattern, and one ectopic sequence codes for a single isoleucine (FHS5, Table 2). The many different ectopic sequences involved, in many cases originating from nontranscribed sequences, indicate that reciprocal fusion transcripts are of no importance.

The majority of the ectopic sequences were derived from 12q14–15 (Table 2) and probably arise from local rearrangement within this region during amplification. Amplification patterns around *HMGIC* are discontinuous (Berner et al., 1997). Exons 4–5, being approximately 140 kb away from the first three exons, may easily be lost due to

random chromosomal breakage during amplification and subsequently become enriched in the tumor cell population due to positive selection (Berner et al., 1997). Such a mechanism is supported by the observation that the ectopic sequences were frequently found to be amplified together with the truncated *HMGIC* also when they are not fused to *HMGIC* (Figs. 1–4). FISH and sequence analyses showed that the remaining ectopic sequences originated from other chromosomes, indicating considerable complexity of the amplicons. The amplification-associated rearrangement of *HMGIC* is typical of well-differentiated liposarcomas. These tumors are characterized by giant rod or ring marker chromosomes, containing amplified sequences of various chromosomal origins, and in all our cases including parts of the 12q13–15 region and the *MDM2* gene (Pedeutour et al., 1994, 1999). Our finding of *HMGIC* transcripts fused to sequences from the long arm of chromosome 1 (FHS5, FHS7, and FHS8) is consistent with this region also being frequently amplified in these tumors (Forus et al., 1995a, 1998; Pedeutour et al., 1999). Sequences from chromosomes 1 and 12 have been shown to be intermingled in the markers (Dal Cin et al., 1993; Pedeutour et al., 1994), revealing zebra-striped patterns upon FISH analysis (Gisselsson et al., 1998; Pedeutour et al., 1999). It is interesting to note that in LS2, *HMGIC* was fused to various ectopic sequences of different chromosomal origins (chromosomes 11 and 12).

Most sequences did not show similarity to previously described genes, with one exception. FHS2 codes for *NOT2*, the human ortholog of the yeast *Not2* gene (Albert et al., 2000), localized to 12q. This protein is part of the Ccr4-NOT protein complex that globally regulates basal transcription by RNA polymerase II. The deduced amino acid sequence of FHS2 corresponds to the last 28 amino acids of the C-terminal domain, which is highly conserved throughout species and might well confer some gain of function to HMGIC.

Searching in the EST database revealed similarity of FHS1, FHS3, and FHS4 to previously described ESTs, indicating that these sequences corresponded to transcribed parts of the genome. In addition, some other probes (e.g., FHS6, FHS10) detected tumor transcripts not coinciding with those detected with the *HMGIC* probe, indicating that they also originated from genes, although no expression was detected in the normal samples (Fig. 2).

In sarcomas with 12q13–15 amplification, *MDM2* and/or *CDK4* are always included in the amplicon (Berner et al., 1997). Although truncated *HMGIC* alone may transform murine cells in culture, our results show that in human malignant tumors *HMGIC* was always amplified together with *MDM2* (Berner et al., 1997), which may be a prerequisite to bypass *TP53*-mediated tumor suppression and thus for malignant progression.

It is clear that amplification and rearrangement of *HMGIC* alone are not sufficient for malignant transformation of human cells. However, the frequent truncation and coamplification of *HMGIC* in liposarcomas points to a key role for this protein in adipocytic tumor development.

ACKNOWLEDGMENTS

We are grateful to Dr. J. Høie and Dr. P. Helgerud for supplying clinical samples and Dr. A.E. Stenwig and Dr. B. Bjerkehagen for the histological classification of tumors. We also thank Dr. T.J. Meza for critical reading of the article and S. Ibrahim Salem for her assistance in image capturing. Additional thanks are due to the Sanger Centre for providing PACs identified by sequence homology during this study.

REFERENCES

Albert TK, Lemaire M, van Berkum NL, Gentz R, Collart MA, Timmers HT. 2000. Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits. Nucl Acids Res 28:809–817.

Arlotta P, Tai AK, Manfioletti G, Clifford C, Jay G, Ono SJ. 2000. Transgenic mice expressing a truncated form of the high mobility group I-C protein develop adiposity and an abnormally high prevalence of lipomas. J Biol Chem 275:14394–14400.

Ashar HR, Fejzo MS, Tkachenko A, Zhou X, Fletcher JA, Weremowicz S, Morton CC, Chada K. 1995. Disruption of the architectural factor HMGI-C: DNA-binding AT hook motifs fused in lipomas to distinct transcriptional regulatory domains. Cell 82:57–65.

Battista S, Fidanza V, Fedele M, Klein-Szanto AJ, Outwater E, Brunner H, Santoro M, Croce CM, Fusco A. 1999. The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis. Cancer Res 59:4793–4797.

Berlingieri MT, Manfioletti G, Santoro M, Bandiera A, Visconti R, Giancotti V, Fusco A. 1995. Inhibition of HMGI-C protein synthesis suppresses retrovirally induced neoplastic transformation of rat thyroid cells. Mol Cell Biol 15:1545–1553.

Berner JM, Meza-Zepeda LA, Kools PF, Forus A, Schoenmakers EF, Van de Ven WJ, Fodstad O, Myklebost O. 1997. HMGIC, the gene for an architectural transcription factor, is amplified and rearranged in a subset of human sarcomas. Oncogene 14:2935–2941.

Dal Cin P, Kools P, Sciot R, De Wever I, Van Damme B, Van de Ven W, Van den Berghe H. 1993. Cytogenetic and fluorescence in situ hybridization investigation of ring chromosomes characterizing a specific pathologic subgroup of adipose tissue tumors. Cancer Genet Cytogenet 68:85–90.

Fedele M, Berlingieri MT, Scala S, Chiariotti L, Viglietto G, Rippel V, Bullerdiek J, Santoro M, Fusco A. 1998. Truncated and chimeric HMGI-C genes induce neoplastic transformation of NIH3T3 murine fibroblasts. Oncogene 17:413–418.

Forus A, Florenes VA, Maelandsmo GM, Meltzer PS, Fodstad O, Myklebost O. 1993. Mapping of amplification units in the q13-14 region of chromosome 12 in human sarcomas: some amplica do not include MDM2. Cell Growth Differ 4:1065–1070.

Forus A, Weghuis DO, Smeets D, Fodstad O, Myklebost O, van Kessel AG. 1995a. Comparative genomic hybridization analysis of human sarcomas. I. occurrence of genomic imbalances and identification of a novel major amplicon at 1q21–q22 in soft tissue sarcomas. Genes Chromosomes Cancer 14:8–14.

Forus A, Weghuis DO, Smeets D, Fodstad O, Myklebost O, Geurts van Kessel A. 1995b. Comparative genomic hybridization analysis of human sarcomas. II. identification of novel amplicons at 6p and 17p in osteosarcomas. Genes Chromosomes Cancer 14:15–21.

Forus A, Berner JM, Meza-Zepeda LA, Saeter G, Mischke D, Fodstad O, Myklebost O. 1998. Molecular characterization of a novel amplicon at 1q21–q22 frequently observed in human sarcomas. Br J Cancer 78:495–503.

Gattas GJ, Quade BJ, Nowak RA, Morton CC. 1999. HMGIC expression in human adult and fetal tissues and in uterine leiomyomata. Genes Chromosomes Cancer 25:316–322.

Gisselsson D, Hoglund M, Mertens F, Mitelman F, Mandahl N. 1998. Chromosomal organization of amplified chromosome 12 sequences in mesenchymal tumors detected by fluorescence in situ hybridization. Genes Chromosomes Cancer 23:203–212.

Groet J, Ives JH, South AP, Baptista PR, Jones TA, Yaspo ML, Lehrach H, Potier MC, Van Broeckhoven C, Nizetic D. 1998. Bacterial contig map of the 21q11 region associated with Alzheimer's disease and abnormal myelopoiesis in Down syndrome. Genome Res 8:385–398.

Huang LS, Bock SC, Feinstein SI, Breslow JL. 1985. Human apolipoprotein B cDNA clone isolation and demonstration that liver apolipoprotein B mRNA is 22 kilobases in length. Proc Natl Acad Sci USA 82:6825–6829.

Kazmierczak B, Wanschura S, Rosigkeit J, Meyer-Bolte K, Uschinsky K, Haupt R, Schoenmakers EF, Bartnitzke S, Van de Ven WJ, Bullerdiek J. 1995. Molecular characterization of 12q14–15 rearrangements in three pulmonary chondroid hamartomas. Cancer Res 55:2497–2499.

Kazmierczak B, Rosigkeit J, Wanschura S, Meyer-Bolte K, Van de Ven WJ, Kayser K, Krieghoff B, Kastendiek H, Bartnitzke S, Bullerdiek J. 1996. HMGI-C rearrangements as the molecular basis for the majority of pulmonary chondroid hamartomas: a survey of 30 tumors. Oncogene 12:515–521.

Kottickal LV, Sarada B, Ashar H, Chada K, Nagarajan L. 1998. Preferential expression of HMGI-C isoforms lacking the acidic carboxy terminal in human leukemia. Biochem Biophys Res Comm 242:452–456.

Mantovani F, Covaceuszach S, Rustighi A, Sgarra R, Heath C, Goodwin GH, Manfioletti G. 1998. NF-kappaB mediated transcriptional activation is enhanced by the architectural factor HMGI-C. Nucl Acids Res 26:1433–1439.

Parente F, Grosgeorge J, Coindre JM, Terrier P, Vilain O, Turc-Carel C. 1999. Comparative genomic hybridization reveals novel chromosome deletions in 90 primary soft tissue tumors. Cancer Genet Cytogenet 115:89–95.

Pedeutour F, Suijkerbuijk RF, Forus A, Van Gaal J, Van de Klundert W, Coindre JM, Nicolo G, Collin F, Van Haelst U, Huffermann K. 1994. Complex composition and co-amplification of SAS and MDM2 in ring and giant rod marker chromosomes in well-differentiated liposarcoma. Genes Chromosomes Cancer 10:85–94.

Pedeutour F, Forus A, Coindre JM, Berner JM, Nicolo G, Michiels JF, Terrier P, Ranchere-Vince D, Collin F, Myklebost O, Turc-Carel C. 1999. Structure of the supernumerary ring and giant rod chromosomes in adipose tissue tumors. Genes Chromosomes Cancer 24:30–41.

Roberts WM, Douglass EC, Peiper SC, Houghton PJ, Look AT. 1989. Amplification of the gli gene in childhood sarcomas. Cancer Res 49:5407–5413.

Rogalla P, Drechsler K, Frey G, Hennig Y, Helmke B, Bonk U, Bullerdiek J. 1996. HMGI-C expression patterns in human tissues: implications for the genesis of frequent mesenchymal tumors. Am J Pathol 149:775–779.

Schajowicz F. 1993. Histological typing of bone tumours. Berlin: Springer Verlag.

Schoenberg Fejzo M, Ashar HR, Krauter KS, Powell WL, Rein MS, Weremowicz S, Yoon SJ, Kucherlapati RS, Chada K, Morton CC. 1996. Translocation breakpoints upstream of the HMGIC gene in uterine leiomyomata suggest dysregulation of this gene by a mechanism different from that in lipomas. Genes Chromosomes Cancer 17:1–6.

Schoenmakers EF, Van de Ven WJ. 1997. From chromosome aberrations to the high mobility group protein gene family: evidence for a common genetic denominator in benign solid tumor development. Cancer Genet Cytogenet 95:51–58.

Schoenmakers EF, Wanschura S, Mols R, Bullerdiek J, Van den Berghe H, Van de Ven WJ. 1995. Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. Nat Genet 10:436–444.

Schoenmakers EF, Huysmans C, Van de Ven WJ. 1999. Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas. Cancer Res 59:19–23.

Schwanbeck R, Manfioletti G, Wisniewski JR. 2000. Architecture of high mobility group protein I-C.DNA complex and its perturbation upon phosphorylation by Cdc2 kinase. J Biol Chem 275: 1793–1801.

Simons A, Schepens M, Jeuken J, Sprenger S, van de Zande G, Bjerkehagen B, Forus A, Weibolt V, Molenaar I, van den Berg E, Myklebost O, Bridge J, van Kessel AG, Suijkerbuijk R. 2000. Frequent loss of 9p21 (p16(INK4A)) and other genomic imbalances in human malignant fibrous histiocytoma. Cancer Genet Cytogenet 118:89–98.

Szymanska J, Tarkkanen M, Wiklund T, Virolainen M, Blomqvist C, Asko-Seljavaara S, Tukiainen E, Elomaa I, Knuutila S. 1996. Gains and losses of DNA sequences in liposarcomas evaluated by comparative genomic hybridization. Genes Chromosomes Cancer 15:89–94.

Wang DZ, Ray P, Boothby M. 1995. Interleukin 4-inducible phosphorylation of HMG-I(Y) is inhibited by rapamycin. J Biol Chem 270:22924–22932.

Weiss SW. 1994. Histological typing of soft tissue tumours. Berlin: Springer Verlag.