Oncogene (2002) 21, 2261–2269

www.nature.com/onc

Positional cloning identifies a novel cyclophilin as a candidate amplified oncogene in 1q21

Leonardo A Meza-Zepeda[1], Anne Forus[1], Birgitte Lygren[1,8], Anine B Dahlberg[1], Linda H Godager[1], Andrew P South[3], Ingo Marenholz[2], Maria Lioumi[4], Vivi Ann Flørenes[6], Gunhild M Mælandsmo[1], Massimo Serra[5], Dietmar Mischke[2], Dean Nizetic[3], Jiannis Ragoussis[4], Maija Tarkkanen[7], Jahn M Nesland[6], Sakari Knuutila[7] and Ola Myklebost*[,1,8]

[1]Department of Tumour Biology, The Norwegian Radium Hospital, Montebello N-0310, Oslo, Norway; [2]Institut für Experimentelle Onkologie un Transplantationsmedizin, Virchow Klinikum, Berlin, Germany; [3]Centre for Applied Molecular Biology, School of Pharmacy, University of London, London, UK; [4]Genomics Laboratory, Division of Medical Genetics UMDS, Guy's Hospital, London, UK; [5]Laboratorio di Ricerca Oncologica, Instituto Ortopedico Rizzoli, Bologna, Italy; [6]Department of Pathology, The Norwegian Radium Hospital, Montebello N-0310, Oslo, Norway; [7]Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki, Finland; [8]Institute of Biochemistry, University of Oslo, Norway

Gains of 1q21-q23 have been associated with metastasis and chemotherapy response, particularly in bladder cancer, hepatocellular carcinomas and sarcomas. By positional cloning of amplified genes by yeast artificial chromosome-mediated cDNA capture using magnetic beads, we have identified three candidate genes (*COAS1*, *-2* and *-3*) in the amplified region in sarcomas. *COAS1* and *-2* showed higher amplification levels than *COAS3*. Most notably, amplification was very common in osteosarcomas, where in particular *COAS2* was highly expressed. *COAS1* has multiple repeats and shows no homology to previously described genes, whereas *COAS2* is a novel member of the cyclosporin-binding peptidyl-prolyl isomerase family, very similar to cyclophilin A. *COAS2* was overexpressed almost exclusively in aggressive metastatic or chemotherapy resistant tumours. Although *COAS2* was generally more amplified than *COAS1*, it was not expressed in well-differentiated liposarcomas, where amplification of this region is very common. All three genes were found to be amplified and over-expressed also in breast carcinomas. The complex nature of the 1q21-23 amplicons and close proximity of the genes make unequivocal determination of the gene responsible difficult. Quite likely, the different genes may give selective advantages to different subsets of tumours.
Oncogene (2002) **21,** 2261–2269. DOI: 10.1038/sj/onc/1205339

Keywords: peptidyl-prolyl isomerase; cyclophilin; breast cancer; sarcoma; amplification; chromosome1

Introduction

Gains or amplifications along the long arm of chromosome 1 are found in leukaemias and solid tumours, and are among the most common chromosomal anomalies in human neoplasia (Alers *et al*., 2000; Kiechle *et al*., 2000; Mairal *et al*., 2000; Malamou-Mitsi *et al*., 1999; Matthews *et al*., 2000; Weber *et al*., 2000a,b). Local gain or high-level amplification affecting mainly 1q21-q23 has been reported for breast cancer cell lines (1q21-q32) (Larramendy *et al*., 2000), hepatocellular carcinomas (1q12-q22) (Guan *et al*., 2000), retinoblastomas (1q21) (Mairal *et al*., 2000) and sarcomas (Forus *et al*., 1995a,b; Szymanska *et al*., 1997; Tarkkanen *et al*., 1995), and is of particular interest since they appear more frequently in aggressive tumours with metastatic potential (Alers *et al*., 2000; Gronwald *et al*., 1997; Tarkkanen *et al*., 1999) and resistance to chemotherapy (Kudoh *et al*., 1999).

We have previously defined the most frequently amplified part of 1q21 using FISH and molecular analyses (Forus *et al*., 1998). This region is localized between the large heterochromatic segment 1q12 and the Epidermal Differentiation Complex (EDC) (Mischke *et al*., 1996). Using available probes we identified a highly and frequently amplified segment covered by a yeast artificial chromosome (YAC) of 1 Mb. Other markers in the region showed mainly low-level amplification and considerable heterogeneity. In order to identify possible target genes we have used this YAC for direct cDNA selection to capture transcripts from an osteosarcoma cell line with high-level amplification of 1q21-q22.

Results

YAC-mediated candidate gene isolation

The CEPH YAC 789f2, representing the highly amplified segment (Forus *et al*., 1998), was labelled with biotin and hybridized to a linkered cDNA pool

*Correspondence: O Myklebost, Department of Tumour Biology, The Norwegian Radium Hospital, Montebello N-0310, Oslo, Norway; E-mail: ola.myklebost@biokjemi.uio.no
Received 30 July 2001; revised 16 January 2002; accepted 18 January 2002

made from the osteosarcoma cell line Saos-2. cDNAs that hybridized to the YAC (or yeast) DNA were isolated using avidin-coated magnetic beads, and the selected population of cDNAs was used to generate a library. Twenty authentic clones were identified based on the absence of hybridization to irrelevant YACs, yeast, or total human DNA, and positive hybridization to YAC 789f2.

Characterization of candidate cDNAs

Sequence analysis revealed that several of the clones were redundant, and represented three candidate genes, none of which have been described previously. The genes were designated *COAS1*, *COAS2* and *COAS3* (for *C*hromosome *O*ne *A*mplified *S*equence). Localization to 1q21 was confirmed by FISH analysis, although minor signals could also be observed in 1p36 with *COAS1* cDNA and *COAS2* cosmid (Figure 1a). This most likely is due to the known duplication of this region (Hardas *et al*., 1994), also detected with YAC 789f2. Preliminary analysis showed all three candidates to be amplified and over-expressed in sarcoma samples with 1q21 amplification.

BLAST analysis (Altschul *et al*., 1997) showed that *COAS1* was identical to segments of partially characterized cDNA sequences, as well as numerous ESTs. One of these was the cDNA KIAA1245, corresponding to a 6.9 kb cDNA isolated from brain tissue, with a predicted open reading frame (ORF) of 892 amino acids. This sequence was also present in PAC RPCI-328E19, partly corresponding to a hypothetical gene predicted as 22 exons, giving rise to a transcript of 3689 bp and an ORF of 921 amino acids. However, no functional leads could be found from the sequence.

Further analysis of the *COAS1* and KIAA1245 sequences revealed the presence of a repeat unit of 866 bases within the 3′ untranslated region (UTR). This repeat is unique to *COAS1*, which contains three complete copies, two with 100% identity and the third differing in only one base pair. A truncated fourth copy was found 5′ to the others. However, each of these 866 repeat are spliced together from separate exons containing four smaller subrepeats of 175 bp, with different degrees of identity ranging from 100 to 85%. Further analysis revealed that such 175 bp repeats are found both in 1q21 and 1p35-ptel, most likely contributing to the FISH signal observed in 1p36 (Figure 1a). In addition, sequences with 98% similarity to the 175 bp subrepeats were also found in two PACs localized to chromosomes 3 and 6.

The complete coding sequence of *COAS2* was found in one exon, and the corresponding sequence was cloned from YAC 789f4. There was discrepancy between the sequence of the region from the Sanger centre and from Celera, and three independent clones were sequenced from each of two independent PCR reactions to verify the sequence of the YAC-derived gene. The predicted ORF showed 84% amino acid identity to cyclophilin A (*PPIA*, chromosome 7p13), belonging to the cyclophilin family of peptidyl-prolyl cis-trans isomerases (Haendler *et al*., 1987). The similarity to other known cyclophilins was less; 66% to CycF, 61% to CycE, 59% to CycC, 55% to CycB, and 53% to CycD. Figure 2 illustrates the similarities at the amino acid level of COAS2 with other members of the cyclophilin family. The active site of human cyclophilin A is formed by the residues R55, I57, F60, Q63, A101, N102, Q111, F113, L122, H126 and R148 (Spitzfaden *et al*., 1992), all perfectly conserved in COAS2. In addition, the cis-proline binding pocket formed by the amino acids F60, M61, F13 and L122 is also conserved.

Although the initial captured *COAS3* cDNA showed homology to Profilin 1, localized to chromosome 17p13.3, only a small region of homology to Profilin was found in the available genomic sequence from 1q21. A larger *COAS3* cDNA of 1338 bases (DKFZp418G1157) was isolated from a mammary gland library (DKFZ Lib. No. 418). Sequence analysis confirmed its location within the genomic clone RPCI-565E6, in 1q21, but the similarity to Profilin 1 was only in the 5′ end. The longest predicted open reading frame showed no homology to the Profilin protein family or any other described amino acid sequence.

Physical mapping of the clones

To further characterize the order and precise location of *COAS1*, *-2* and *-3*, the isolated cDNA clones were hybridized to filters containing fragmented YACs from a 6 Mb contig covering the region of interest (Lioumi *et al*., 1998). The three genes were situated between the genetic markers *D1S442* and *WI7969*. *COAS3* was the most centromeric sequence and *COAS2* the most distal (Figure 1b).

Searches of the High Throughput Genomic Sequence Database showed that *COAS1* and *COAS2* were located approximately 60 kb apart in the genomic clone RPCI-328E19, oriented head to tail. RPCI-328E19 overlaps with RPCI-565E6, containing *COAS3*, placing *COAS3* about 11 kb centromeric from *COAS1* (Figure 1b). The PACs containing *COAS1*, *-2* and *-3* were reported by the Sanger Centre to be localized in 1q12-22.2, thus confirming our results. Figure 1 shows the chromosomal localization of the genes, as well as a map of the genomic region.

Analysis of amplification in clinical samples

Gene copy numbers were determined by FISH to interphase nuclei from selected soft tissue sarcomas, osteosarcoma xenografts, and cell lines with amplification of 1q21-q22 (Forus *et al*., 1998, 1995b; A Forus *et al*., unpublished), as well as additional unselected osteosarcomas and breast carcinomas.

There was considerable heterogeneity in copy numbers in the clinical samples, as observed by others (Ambros *et al*., 2001; Forus *et al*., 1998). However, with the exception of MFH25, used as a negative control because it lacks detectable 1q21-q23 amplifica-

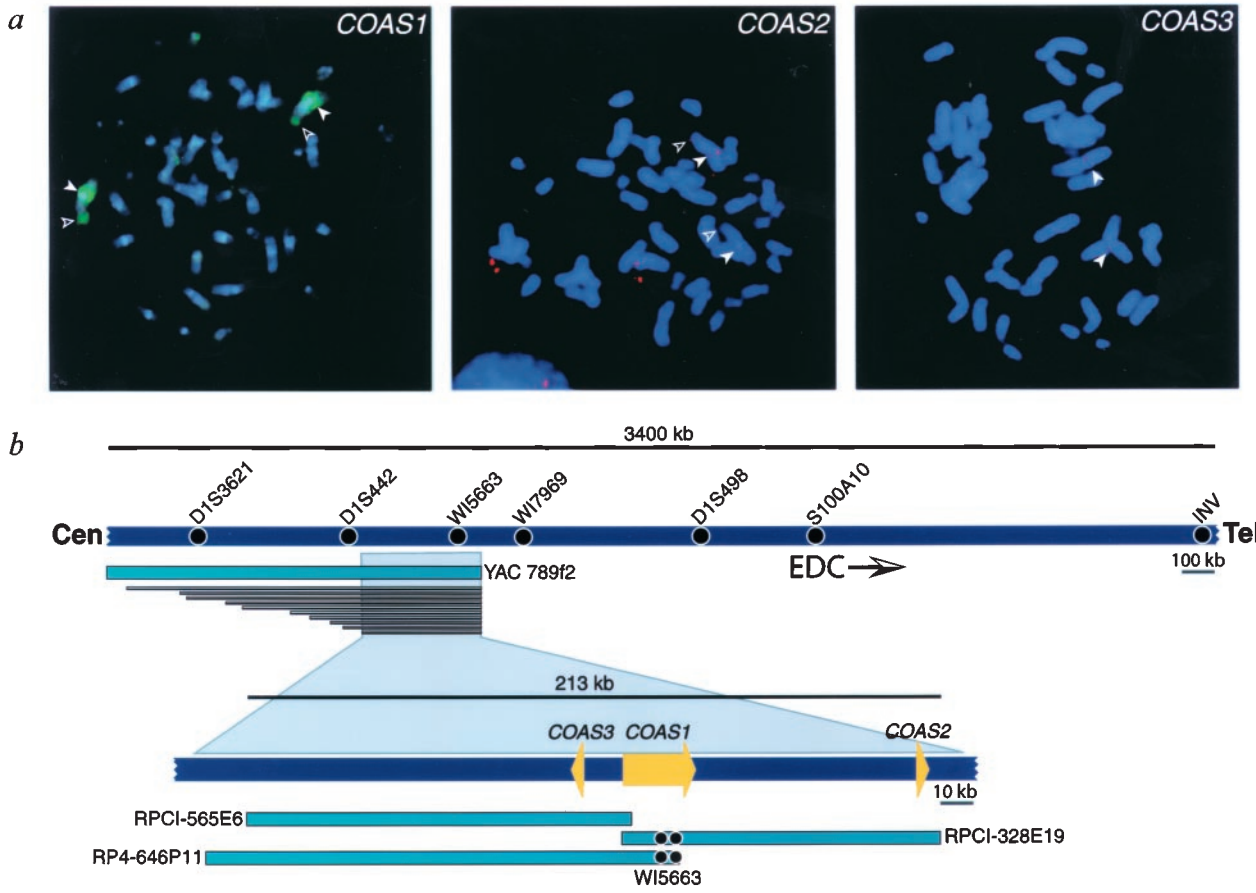

Figure 1 Chromosomal localization and physical mapping of the *COAS1*, *-2*, and *-3*. (**a**) Biotin labelled *COAS1* was hybridized to normal metaphase chromosomes, amplified by tyramid signal amplification (TSA) and detected by FITC conjugated streptavidin (green). A strong signal was observed on 1q21 (arrow head) and a weaker signal on 1p36 (open arrow head). Cosmid ICRFc112N1180 and ICRFc112I1772 specific for *COAS2* and *COAS3* respectively, were labelled with biotin and hybridized to metaphase chromosomes. Cosmids were detected using Cy3 (red) conjugated avidin. A clear signal was observed on 1q21 (arrow head) and for ICRFc112N1180 an additional weak signal on 1p36 was seen (open arrow head). (**b**) A physical map of the *COAS1*, *-2* and *-3* was designed by hybridization of the different genes to a filter of fragmented YACs from a 6 Mb contig covering the region of interest and sequence analysis. PAC RPCI565E6 and RPCI 328E19 show an overlap of approximately 2.2 kb with 99% of identity. A third PAC, RP4-646P11, partially overlapped the two previously described PACs. *COAS3* was the most centromeric gene and *COAS2* the most telomeric. The orientation of transcription is shown with an arrow. All the genes were located between the genetic markers D1S3621 and WI7969, proximal to the epidermal differentiation complex (EDC)

tion, all the samples, even those from unselected cases, showed increased copy numbers of at least one of the three candidate genes (Figure 3a,b). In most of the soft tissue sarcomas with 1q amplification, a relatively large fraction of nuclei had high-level amplification of all three genes, but *COAS1* and *COAS2* were present in higher copy numbers than *COAS3* (Figure 3b). Five of the nine samples had high-level amplification of at least one of the genes in more than 40% of the nuclei. In three osteosarcoma cell lines, almost all nuclei showed high-level amplification for all three genes. In the OS patient and xenograft panel, a majority of the samples had high-level amplification of *COAS1* and/or *COAS2* in more than 40% of the nuclei, whereas *COAS3* in general was moderately amplified.

In the breast tumours and cell lines, the three genes showed mostly moderate amplification levels, but some samples had high-level amplification, although only in 20–37% of the nuclei. In these cases *COAS2* and *COAS3* showed higher copy numbers than *COAS1* (Figure 3b).

Expression studies

The expression levels of the three genes were analysed in clinical samples and cell lines from sarcomas, mammary carcinomas and 15 normal tissues (Figure 4). The *COAS1* transcript was expressed in most of the normal tissues tested, with highest levels in brain, ovary, mammary gland, skin and adipose tissue (Figure 4). Due to the large size of the transcript, the signals were somewhat degraded and spread out. Similarly, *COAS2* was expressed in the majority of normal tissues tested, with the highest transcript levels detected in brain, and the lowest in skeletal muscle (Figure 4). *COAS3* was also expressed in normal tissue, with high levels in adipose tissue, intestine, and spleen (Figure 4).

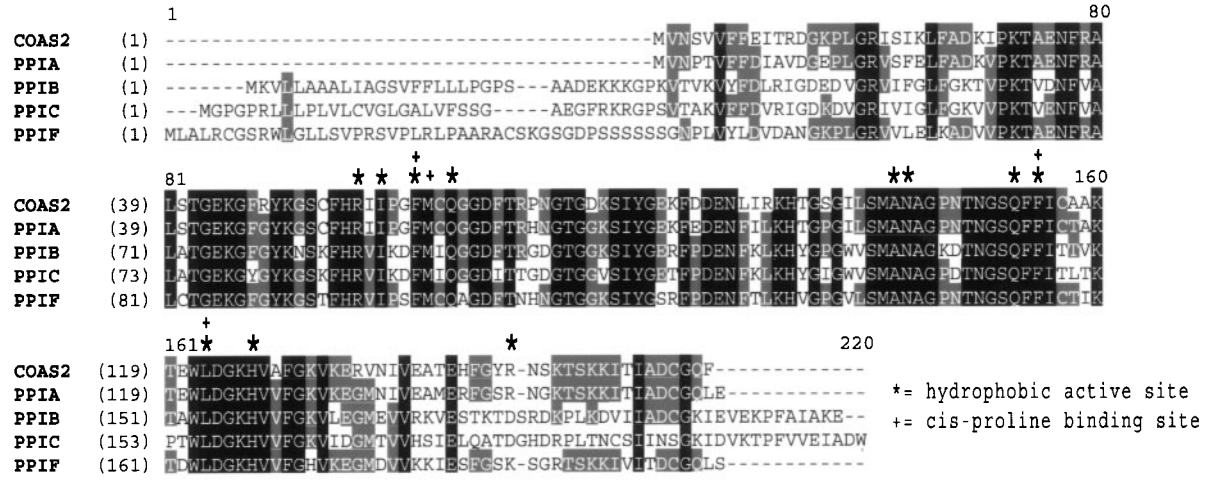

Figure 2 Homology comparison between different members of the human cyclophilin family. Comparison between the *COAS2* deduced amino acid sequence and cyclophilin A (PPIA), cyclophilin B (PPIB), cyclophilin C (PPIC) and Cyclophilin F (PPIF). Perfectly conserved amino acids are shown in black, and partially conserved shaded grey. (*) Indicates conserved hydrophobic amino acids in the active site and (+) indicate the amino acids of the cis-proline binding pocket

Table 1 Histopathological characteristics of the tumours analysed

Sample	*Sample from*	*Metastatic disease*	*Histological grade*	*Status*	*Survival (months)*
LP45	Recurrence	–			
LS2-WDLPS	Primary	–	I	CR	63
LS3x	Primary	–	IV	D-UK	7
LS6-WDLPS	Primary	–	I	D-UK	30
LS9-WDLPS	Primary	–	I	CR	98
LS13-WDLPS	Recurrence	–	I	CR	77
LS28	Primary	+	IV	D	10
LS42	Recurrence	–	IV	D-OD	36
LS43-WDLPS	Primary	–	I	CR	nd
MFH25	Primary	–	IV	CR	96
MFH26x	Primary	+	IV	D	76
MFH36	Primary	+	IV	D	14
OS1x	Primary	+	IV	D	76
OS3x	Primary	+	IV	D	44
OS4x	Primary	+	IV	D	37
OS5x	Primary	+	IV	D	4
OS6x	Primary	–	IV	D	33
OS7x	Primary	–	IV	CR	99
OS8x	Metastasis	+	IV	D	19
OS9x	Primary	+	IV	D	13
OS11x	Primary	+	IV	D	194
OS12x	Metastasis	+	IV	D	49
OS13x	Metastasis	+	III	CR	174
OS14	Primary	+	IV	D	13
OS16x	Primary	+	IV	D	7
OS29	Primary	+	IV	D	4
OS31	Primary	+	IV	D	3
Ma10	Primary	+	III	D-OD	4
Ma15	Primary	+	II	D	37
Ma40	Primary	+	II	A	122
Ma41	Primary	+	II	D	39
Ma51	Primary	+	II	D	42
Ma95	Primary	+	II	D	116
Ma201	Primary	+	III	D-OD	68
Ma215	Primary	+	II	D	64

LP, lipoma; LS, liposarcoma; WDLPS, Well-differentiated liposarcoma; MFH, Malignant fibrous histiocytoma; OS, Osteosarcoma; Ma, Mammary carcinomas, CR, complete remission; D-UK, dead unknown cause, D, dead of cancer; D-OD, dead of other disease; A, alive with disease

COAS1 was expressed in most of the tumour samples and cell lines tested, showing very high expression levels in the osteosarcomas that had high-level amplification of the gene (Figure 4). In soft tissue sarcomas, expression levels were generally moderate, but in several cases much higher than in normal tissue (compare with samples from mammary gland and adipose tissue on the same filter). Two of the mammary carcinomas, Ma201 and Ma215, showed very high expression levels of the gene, whereas mRNA levels in the remaining samples were moderate (Figure 4).

Although *COAS2* was amplified in all the liposarcomas (Figure 3b), elevated expression was detected only in pleomorphic and undifferentiated liposarcomas (LS3, LS28x and LS42, data not shown for LS28x and LS42), being almost undetectable in well-differentiated cases (Figure 4). The gene showed relatively high expression levels in the OS samples with amplification, and in four additional OS samples (OS1x, OS2x, OS11x and OS12x) that were not included in the FISH analyses. All the osteosarcoma cell lines showed much higher expression levels of *COAS2* than the normal tissues. *COAS2* was expressed in all the mammary carcinomas and cell lines, and expression levels were four to seven times higher in the cell lines than in normal mammary gland (Figure 4).

COAS3 was expressed in all mammary carcinomas and cell lines, and most of the soft-tissue sarcomas, but generally at low levels in the OS samples and cell lines (Figure 4).

Discussion

Using a YAC-based gene capture technique, we have isolated and characterized three new amplified genes

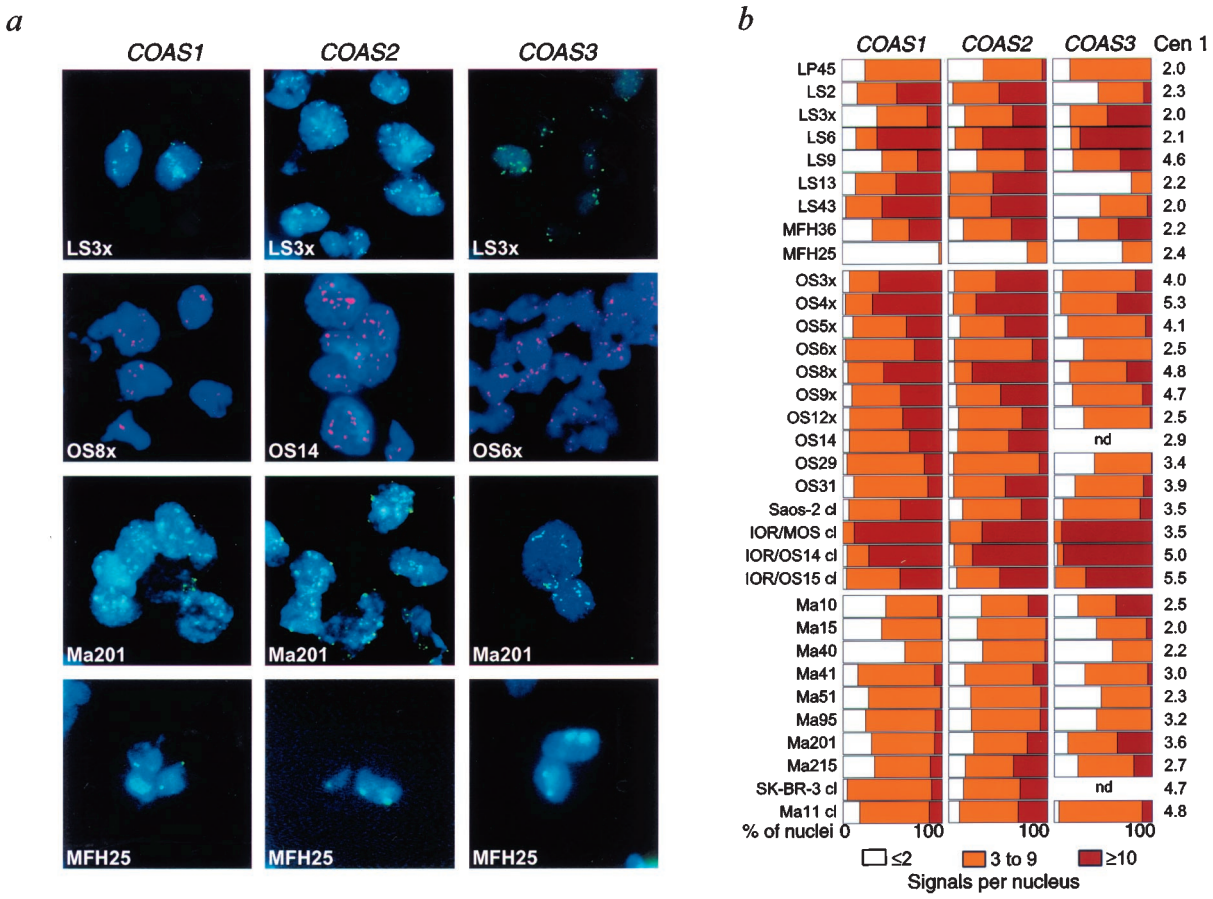

Figure 3 DNA copy number by FISH analysis on interphase nuclei. (**a**) Cosmid ICRFc112K0437, ICRFc112N1180 and ICRFc112I1772, specific for *COAS1*, *COAS2* and *COAS3*, respectively, were labelled and hybridized to interphase nuclei from different liposarcomas, malignant fibrous histiocytomas, osteosarcomas and mammary carcinomas. (**b**) Summary of DNA copy number determined by interphase FISH. The colour of the shading indicates the range of signals observed per nucleus, and the area the percentage of each group among the scored nuclei. For each sample an average of centromere 1 signals is shown. LP, lipoma; LS, liposarcoma; MFH, malignant fibrous histiocytoma; OS, osteosarcoma; Ma, mammary carcinomas; cl, cell line; nd, not determined; suffix 'x', xenograft

in 1q21. The *COAS1*, *-2* and *-3* genes were highly expressed in both sarcomas and breast cancer. They show recurrently high amplification levels in sarcomas and moderate in breast cancer, and represent the first candidate target genes for this very common amplification. Because of the clustering of the genes within about 70 kb, amplification patterns are very similar, although *COAS3*, which is very close to *COAS1*, is in most cases less amplified than the other two. Based on the amplification frequencies and expression pattern, *COAS1*, which has no homology to any known gene, and *COAS2*, which shows high homology to cyclophilin A, are the most interesting candidates.

The very large *COAS1* mRNA contains several unique almost perfect repeats at the 3′ end. However, due to the lack of structural or functional data, we cannot suggest a role for this gene in tumour biology. *COAS1* is expressed in most of the tumour samples, also in the almost benign well-differentiated liposarcomas (WDLPS), suggesting a more general function in tumorigenesis, rather than being a marker for tumour progression.

For *COAS2*, on the other hand, the homology to cyclophilins suggests many possible roles. At least eight different human cyclophilins have been described, as well as a number of pseudogenes, one of which has been shown to be amplified in breast cancer (Collins *et al*., 1998). Cyclophilins are a ubiquitous group of peptidyl-prolyl cis-trans isomeres (PPIs) that are highly conserved during evolution (Ozaki *et al*., 1996), and bind the immunosuppressive drug Cyclosporin A (CsA) with high affinity (Liu *et al*., 1991). PPIs may modulate the activity of other proteins through the direct structural effect of peptide bond isomerization, but also through reciprocal modulation of protein phosphorylation (Lu, 2000).

Increased cyclophilin activity in tumours has previously been reported (Koletsky *et al*., 1986), and Rey *et al*. (1999) have recently shown an increase in cyclophilin A expression during the transition of human keratinocytes from an immortal to a malignant phenotype. Another novel member of the cyclophilin family (*PPIE*) was amplified in small cell lung cancer cell lines with L-myc amplification (Kim *et al*., 1998).

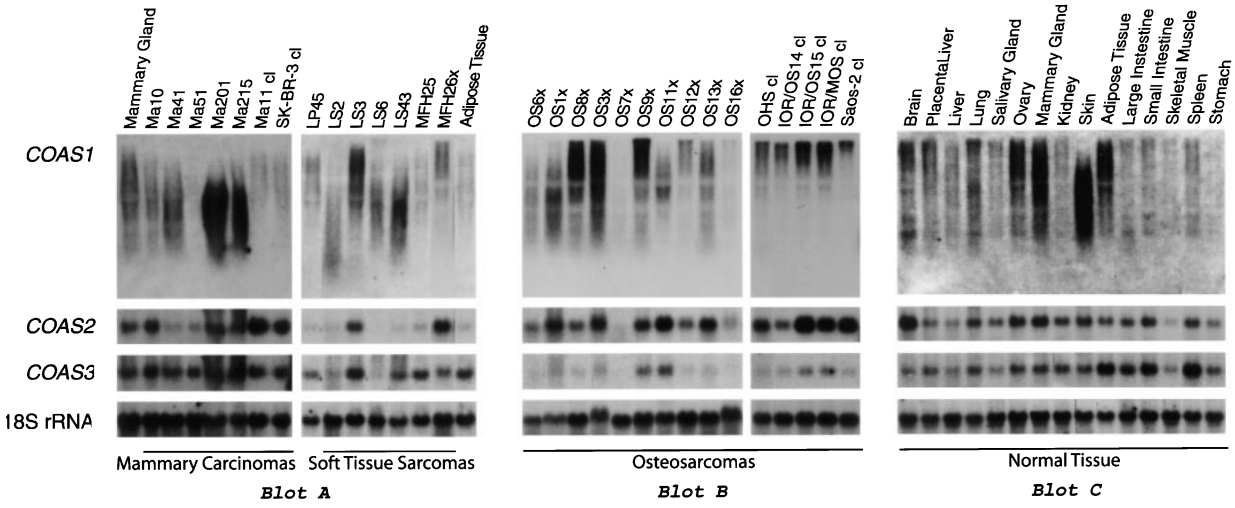

Figure 4 Expression analysis. Expression of *COAS1*, *-2* and *-3* in tumour and normal tissue. COAS1, -2 and -3 were sequentially hybridized to three different Northern blots (**a**,**b**,**c**), containing total RNA from soft tissue sarcomas, breast carcinomas, and cell lines, osteosarcoma xenografts and cell lines, and a panel of different normal tissues. Hybridization to an 18 S rRNA oligonucleotide probe is shown as loading control. LP, lipoma; LS, liposarcoma; MFH, malignant fibrous histiocytoma; OS, osteosarcoma; Ma, mammary carcinomas; cl, cell line; suffix 'x', xenograft

Furthermore, cyclophilin A has been suggested to be secreted and have growth factor activity under oxidative stress (Jin *et al*., 2000), and exposure of hepatocytes to peroxisome proliferation inducers upregulates cyclophilin A expression (Corton *et al*., 1998).

In addition to the cyclophilins, two other families of PPIs have been described, the FK506-binding proteins (FKBPs) (Harding *et al*., 1989; Siekierka *et al*., 1989) and parvulins (Rahfeld *et al*., 1994), both structurally unrelated to the cyclophilins. In spite of this, it has been shown that overexpression of cyclophilin A can rescue a lethal parvulin mutant in yeast (Arevalo-Rodriguez *et al*., 2000). Thus, amplification and overexpression of one PPI might also affect pathways regulated by structurally unrelated PPIs.

Other PPIs have functions that are highly relevant to the phenotypes of tumours with 1q amplification. cFKBP/SMAP is important in smooth muscle cell differentiation (Fukuda *et al*., 1998), and FKBP51 is highly expressed during the clonal expansion phase of adipocyte differentiation (Yeh *et al*., 1995), implicating these proteins in mesenchymal differentiation. The metazoan parvulin, *Pin1*, is involved in chromatin condensation and regulation of the G2/M transition (Lu *et al*., 1996). Furthermore, FKBP12 activity has been shown to be important for multidrug resistance through the P-glycoprotein pathway in yeast (Hemenway and Heitman, 1996), perhaps relevant to the poor chemotherapy response associated with 1q21-22 amplification in ovarian carcinomas (Kudoh *et al*., 1999). In this context it is notable that the osteosarcomas with their very frequent amplification and over-expression of *COAS2* routinely receive preoperative chemotherapy.

It is striking that although *COAS2* is the most amplified gene also in liposarcomas, it is hardly expressed in the well-differentiated liposarcomas (WDLPS), as opposed to high-grade samples. It thus seems likely that in WDLPS, *COAS2* is less important, and *COAS1* is a more likely target gene. This resembles the situation observed for 12q13-15 (Berner *et al*., 1996), where amplification always includes *CDK4* and is associated with metastasis and tumour progression in osteosarcomas. In the less malignant WDLPS, on the other hand, amplification is associated with complex marker ring or rod chromosomes that always include the *MDM2* gene. These marker chromosomes frequently also contain segments from 1q (Pedeutour *et al*., 1999), and in several cases transcripts fusing the *HMGIC* gene in 12q15 to sequences from 1q have been found (Meza-Zepeda *et al*., 2001).

In general, expression of *COAS2* was associated with clinical behaviour in our preliminary analysis of a small number of cases. In addition to the correlation mentioned for liposarcomas, the only osteosarcoma analysed that did not express any of the *COAS* genes (OS7x) was from one of two patients who did not develop metastasis, but showed complete remission (Table 1). Interestingly, five out of five xenografts with expression of the gene respond poorly to chemotherapy using methotrexate, cisplatin or ifosfamide, and two of these were resistant to all three drugs (K Breistøl, unpublished).

COAS1 and *-2* represent the first likely target genes for the 1q21 amplifications. Further studies need to be done to determine their possible role in the development or progression of different cancer subtypes.

Materials and methods

Specimens

Eight soft tissue sarcomas of various subtypes, 10 osteosarcoma samples or xenografts, four osteosarcoma cell lines, Saos-2 (ATCC, HTB-85), IOR/MOS, IOR/OS14 and IOR/OS15 (established at the Istituti Ortopedici Rizzoli, Bologna, Italy), eight ductal breast carcinomas and two breast cancer cell lines, Ma11 and SK-BR-3 (ATCC, HTB-30), were analysed by FISH. All soft-tissue sarcomas, except for one (MFH25), as well as the OS cell lines and two of the xenografts (OS4x and OS9x) had previously identified amplification of 1q21-q22 (Forus *et al*., 1995a,b, 1998; A Forus *et al*., unpublished). Since RNA was not available for all the OS samples analysed by FISH, six additional OS samples and the OHS cell line were included in the expression analyses. OHS was established from the same patient sample as OS5x. Clinical specimens were collected immediately after surgery, cut into small pieces, frozen in liquid nitrogen and stored at -70°C until use. All tumours were classified according to the WHO International Histological Classification of Tumours.

Cloning and characterization of cDNAs

Poly(A)$^+$ RNA was isolated from 75 μg of total RNA from Saos-2 cells, using the mRNA direct kit (Dynal, Norway) as described by the supplier. One microgram of poly(A)$^+$ RNA was used as substrate for the first and second strand cDNA synthesis, using the Marathon cDNA amplification kit (Clontech, Palo Alto, CA, USA). Adapters were ligated to the cDNAs, and amplification of the library was performed by touchdown Polymerase Chain Reaction (PCR). An aliquot of the diluted linkered cDNAs was amplified using adapter-specific primers and the Advantage cDNA polymerase mix (Clontech, Palo Alto, CA, USA), according to the suppliers manual. Direct cDNA selection was performed as described previously (Lovett *et al*., 1991) with some modifications. Briefly, 6 μg of the cDNA library was mixed with 10 μg of Cot-1 DNA (Gibco–BRL), and precipitated with ethanol. The DNA pellet was washed and resuspended in 100 μl of prehybridization solution (2 × SSC, 0.1% SDS, 1 × Denhardt), denatured for 5 min at 95°C and prehybridized at 68°C for 4 h.

Total yeast DNA containing CEPH megaYAC 789f2 was purified by standard procedures, and labelled with biotin 14-dATP by nick translation at 16°C (Bionick, Gibco–BRL). Two micrograms of biotinylated YAC DNA was mixed with 30 μg of total yeast DNA, ethanol precipitated and washed and pre-hybridized as described for the cDNA. Thereafter, cDNA and YAC DNA was pooled, precipitated with ethanol, dissolved in 100 μl of hybridization solution (6 × SSC, 0.1% SDS, 1 × Denhardt), and hybridized at 68°C for 48 h. Two milligrams of streptavidin-coated magnetic beads (Roche Molecular Biochemicals) were washed three times with 1 × washing/binding buffer (0.1 × SSC, 0.1% SDS), resuspended in 2 × binding buffer and added to the YAC/cDNA hybridization mix. The YAC/cDNA complex was allowed to anneal to the streptavidine-coated beads for 30 min at room temperature, and the bead/DNA complex was collected using a magnet. Beads were resuspended and washed 2 × 15 min at RT and then 3 × at 65°C, in 0.1 × SSC, 0.1% SDS. Selected cDNAs were eluted in 100 μl of dH_2O by 10 min incubation at 80°C and recaptured as a bead/genomic DNA complex. Amplification of the captured cDNAs was adapter-specific as previously described. The selected and amplified clones correspond to the Saos-2 selected library described later.

Amplified cDNAs were purified and incubated in 1 × AmpliTaq Buffer in the presence of AmpliTaq Gold (Perkin Elmer Co), 2 mM Mg^{2+} and dATP for 30 min at 72°C. An aliquot of the adenosine extended PCR products was ligated to the pCR 2.1 vector (Invitrogen, The Netherlands) as described by the supplier. Two microliters of the ligation mix was used to transform chemically competent cells (TOP10, Invitrogen) by standard heat shock procedure. Transformed cells were plated in LB ampicillin plates containing IPTG and X-Gal.

DNA from the 96 selected clones was isolated by standard miniprep protocols (Qiagen) and amplified by PCR, and together with DNA from the YACs (789f2 and the adjacent 934g9, 950e2, 692c1, 955e11 and 764a1) were digested by *Bam*HI before electrophoresis and blotting. The membrane with cDNAs was hybridized with total human DNA, YAC 789f2 and an irrelevant YAC (883H6), and the YAC membrane was hybridized with the captured cDNA clones for *COAS1*, *COAS2* and *COAS3*.

Captured cDNAs were sequenced on an ALF express automatic sequencer (Amersham Pharmacia Biotech, UK) using vector primers and the Thermo Sequenase cycle sequencing kit (Amersham Pharmacia Biotech, UK). Sequence assembly and analysis was performed using the DNAstar (DNASTAR Inc.) analysis package, the Vector NTI suite (Informax Inc.) and the BLAST search engines at NCBI.

Screening of cosmid libraries

A flow-sorted chromosome 1 cosmid library (ICRFc112) (Nizetic *et al*., 1994) was screened by Southern hybridization. The library was initially screened using a pool of the three captured cDNAs as probe (Groet *et al*., 1998). The following cosmids were selected and used for FISH analysis: ICRFc112K0437 (*COAS1*), ICRFc112N1180 (*COAS2*) and ICRFc112I1772 (*COAS3*).

Fluorescence in situ *hybridization (FISH)*

Normal metaphase chromosomes from peripheral leukocytes and interphase nuclei from frozen tumour tissues were prepared as described previously (Forus *et al*., 1998). PCR products and cosmid DNA was labelled with biotin-14-dATP or digoxygenin-11-dUTP (Roche Biochemical) by nick translation. For each hybridization, 200–500 ng of labelled DNA was prehybridized with 50–100-fold excess of human Cot-1 DNA. Pretreatment of slides, hybridization and washes were done as described (Forus *et al*., 1998, 2001). For detection of digoxygenin we used fluorescein isothiocyanate (FITC) conjugated sheep anti-digoxygenin antibody (Roche Biochemicals) followed by FITC or ALEXA 488 labelled donkey anti sheep (Molecular Probes, The Netherlands). Detection of biotin-labelled cosmid probes was detected by avidin-conjugated Cy3 (Amersham Pharmacia Biotech, UK), whereas biotin-labelled PCR products were detected using the thyramide amplification procedure (TSA-Indirect (ISH), NEN Life Science Products, Boston, MA, USA). The antibodies used were avidin-Cy3 (Amersham Pharmacia Biotech, UK), FITC-labelled streptavidin (NEN) and biotinylated anti-avidin D (Vector Laboratories). Hybridized slides were examined visually using a Zeiss Axioplan microscope with filters for excitation of DAPI, DAPI/FITC, DAPI/Rhodamine (Cy3), and DAPI/FITC/Rhodamine (Cy3). For each probe localization of the signal was evaluated in at least 10 metaphases, and copy numbers evaluated in at least 150 nuclei. Amplification levels were grouped as follows: normal (two signals); moderate (3–9 signals); and high (10 or more signals).

Northern blot analysis

RNA purification, electrophoresis, transfer to nylon filter and hybridization were as previously described (Forus *et al*., 1993). Quantification of the signal intensities was performed by two-dimensional densitometry on a Molecular Dynamics laser densitometer. The net signal values from specific bands were corrected for unequal sample loading by calibration relative to the signal obtained from an 18S ribosomal RNA probe.

Different cDNAs and PCR products were used as probe for Northern blot hybridizations. As probe for *COAS1* was used a 1 kb cDNA fragments containing the 866 bp repeat unit; for *COAS2* a 495 bp PCR product enclosing the complete open reading frame; and for *COAS3* a 1338 bp cDNA fragment isolated from the library screening.

Acknowledgements
We are grateful to J Høie and P Helgerud for supplying clinical samples and AE Stenwig and B Bjerkehagen for the histological classification of tumours. Additional recognition to The Sanger Center for providing PACs identified by sequence homology during this study, and to Phil Rye and TJ Meza for critical reading of the manuscript. This work was supported by The Norwegian Research Council, the Norwegian Cancer Society, The Nordic Cancer Union, the European Union (through the Biomedicine and Health grant BMH4-CT96-0319), the Cancer Research Campaign, UK (SP 2475/0101) (Nizetic, South), Leiras Research Foundation (Knuutila, Tarkkanen) and Associazione Italiana per la Ricerca sul Cancro (AIRC) and of the Istituti Ortopedici Rizzoli (Ricerca Corrente and Ricerca Finalizzata) (Serra). The Celera Human Genome Sequence was kindly made available by Celera (www.celera.com) (Venter *et al*. (2001) *Science*, **291**, 1304–1351).

Accession numbers
COAS1, BG154169; *COAS2*, GI:4826471; *COAS3*, AF345651; KIAA1245, AB033071; Hypothetical gene spanning *COAS1*, AL117237.

References

Alers JC, Rochat J, Krijtenburg PJ, Hop WC, Kranse R, Rosenberg C, Tanke HJ, Schroder FH and van Dekken H. (2000). *Lab. Invest.,* **80,** 931–942.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman D. (1997). *Nucleic Acids Res.,* **25,** 3389–3402.

Ambros PF, Ambros IM, Kerbl R, Luegmayr A, Rumpler S, Ladenstein R, Amann G, Kovar H, Horcher E, De Bernardi B, Michon J and Gadner H. (2001). *Med. Ped. Oncol.,* **36,** 1–4.

Arevalo-Rodriguez M, Cardenas ME, Wu X, Hanes SD and Heitman J. (2000). *EMBO J.,* **19,** 3739–3749.

Berner JM, Forus A, Elkahloun A, Meltzer PS, Fodstad O and Myklebost O. (1996). *Genes Chrom. Cancer,* **17,** 254–259.

Collins C, Rommens JM, Kowbel D, Godfrey T, Tanner M, Hwang SI, Polikoff D, Nonet G, Cochran J, Myambo K, Jay KE, Froula J, Cloutier T, Kuo WL, Yaswen P, Dairkee S, Giovanola J, Hutchinson GB, Isola J, Kallioniemi OP, Palazzolo M, Martin C, Ericsson C, Pinkel D and Gray JW. (1998). *Proc. Natl. Acad. Sci. USA,* **95,** 8703–8708.

Corton JC, Moreno ES, Merritt A, Bocos C and Cattley RC. (1998). *Cancer Lett.,* **134,** 61–71.

Forus A, Florenes VA, Maelandsmo GM, Meltzer PS, Fodstad O and Myklebost O. (1993). *Cell Growth Differ.,* **4,** 1065–1070.

Forus A, Weghuis DO, Smeets D, Fodstad O, Myklebost O and Geurts van Kessel A. (1995a). *Genes Chrom. Cancer,* **14,** 15–21.

Forus A, Weghuis DO, Smeets D, Fodstad O, Myklebost O and van Kessel AG. (1995b). *Genes Chrom. Cancer,* **14,** 8–14.

Forus A, Berner JM, Meza-Zepeda LA, Saeter G, Mischke D, Fodstad O and Myklebost O. (1998). *Br. J. Cancer,* **78,** 495–503.

Forus A, Larramendy ML, Meza-Zepeda LA, Bjerkehagen B, Godager LH, Dahlberg AB, Saeter G, Knuutila S and Myklebost O. (2001). *Cancer Genet. Cytogenet.,* **125,** 100–111.

Fukuda K, Tanigawa Y, Fujii G, Yasugi S and Hirohashi S. (1998). *Development,* **125,** 3535–3542.

Groet J, Ives JH, South AP, Baptista PR, Jones TA, Yaspo ML, Lehrach H, Potier MC, Van Broeckhoven C and Nizetic D. (1998). *Genome Res.,* **8,** 385–398.

Gronwald J, Storkel S, Holtgreve-Grez H, Hadaczek P, Brinkschmidt C, Jauch A, Lubinski J and Cremer T. (1997). *Cancer Res.,* **57,** 481–487.

Guan XY, Fang Y, Sham JS, Kwong DL, Zhang Y, Liang Q, Li H, Zhou H and Trent JM. (2000). *Genes Chrom. Cancer,* **29,** 110–116.

Haendler B, Hofer-Warbinek R and Hofer E. (1987). *EMBO J.,* **6,** 947–950.

Hardas BD, Zhang J, Trent JM and Elder JT. (1994). *Genomics,* **21,** 359–363.

Harding MW, Galat A, Uehling DE and Schreiber SL. (1989). *Nature,* **341,** 758–760.

Hemenway CS and Heitman J. (1996). *J. Biol. Chem.,* **271,** 18527–18534.

Jin ZG, Melaragno MG, Liao DF, Yan C, Haendeler J, Suh YA, Lambeth JD and Berk BC. (2000). *Circ. Res.,* **27,** 789–796.

Kiechle M, Hinrichs M, Jacobsen A, Luttges J, Pfisterer J, Kommoss F and Arnold N. (2000). *Am. J. Pathol.,* **156,** 1827–1833.

Kim JO, Nau MM, Allikian KA, Makela TP, Alitalo K, Johnson BE and Kelley MJ. (1998). *Oncogene,* **17,** 1019–1026.

Koletsky AJ, Harding MW and Handschumacher RE. (1986). *J. Immunol.,* **137,** 1054–1059.

Kudoh K, Takano M, Koshikawa T, Hirai M, Yoshida S, Mano Y, Yamamoto K, Ishii K, Kita T, Kikuchi Y, Nagata I, Miwa M and Uchida K. (1999). *Clin. Cancer Res.,* **5,** 2526–2531.

Larramendy ML, Lushnikova T, Bjorkqvist AM, Wistuba II, Virmani AK, Shivapurkar N, Gazdar AF and Knuutila S. (2000). *Cancer Genet. Cytogenet.,* **119,** 132–138.

Lioumi M, Olavesen MG, Nizetic D and Ragoussis J. (1998). *Genomics,* **49,** 200–208.

Liu J, Chen CM and Walsh CT. (1991). *Biochemistry,* **30,** 2306–2310.

Lovett M, Kere J and Hinton LM. (1991). *Proc. Natl. Acad. Sci. USA,* **88,** 9628–9632.

Lu KP, Hanes SD and Hunter T. (1996). *Nature,* **380,** 544–547.
Lu KP. (2000). *Prog. Cell Cycle Res.,* **4,** 83–96.
Mairal A, Pinglier E, Gilbert E, Peter M, Validire P, Desjardins L, Doz F, Aurias A and Couturier J. (2000). *Genes Chrom. Cancer,* **28,** 370–379.
Malamou-Mitsi VD, Syrrou M, Georgiou I, Pagoulatos G and Agnantis NJ. (1999). *J. Exp. Clin. Cancer Res.,* **18,** 357–361.
Matthews CP, Shera KA and McDougall JK. (2000). *Proc. Soc. Exp. Biol. Med.,* **223,** 316–321.
Meza-Zepeda LA, Berner JM, Henriksen J, South AP, Pedeutour F, Dahlberg AB, Godager LH, Nizetic D, Forus A and Myklebost O. (2001). *Genes Chrom. Cancer,* **31,** 264–273.
Mischke D, Korge BP, Marenholz I, Volz A and Ziegler A. (1996). *J. Invest. Dermatol.,* **106,** 989–992.
Nizetic D, Monard S, Young B, Cotter F, Zehetner G and Lehrach H. (1994). *Mamm. Genome,* **5,** 801–802.
Ozaki K, Fujiwara T, Kawai A, Shimizu F, Takami S, Okuno S, Takeda S, Shimada Y, Nagata M, Watanabe T, Takaichi A, Takahashi E, Nakamura Y and Shin S. (1996). *Cytogenet. Cell Genet.,* **72,** 242–245.
Pedeutour F, Forus A, Coindre JM, Berner JM, Nicolo G, Michiels JF, Terrier P, Ranchere-Vince D, Collin F, Myklebost O and Turc-Carel C. (1999). *Genes Chrom. Cancer,* **24,** 30–41.
Rahfeld JU, Rucknagel KP, Schelbert B, Ludwig B, Hacker J, Mann K and Fischer G. (1994). *FEBS Lett.,* **352,** 180–184.
Rey O, Baluda MA and Park NH. (1999). *Oncogene,* **18,** 827–831.
Siekierka JJ, Hung SH, Poe M, Lin CS and Sigal NH. (1989). *Nature,* **341,** 755–757.
Spitzfaden C, Weber HP, Braun W, Kallen J, Wider G, Widmer H, Walkinshaw MD and Wuthrich K. (1992). *FEBS Lett.,* **300,** 291–300.
Szymanska J, Virolainen M, Tarkkanen M, Wiklund T, Asko-Seljavaara S, Tukiainen E, Elomaa I, Blomqvist C and Knuutila S. (1997). *Cancer Genet. Cytogenet.,* **99,** 14–18.
Tarkkanen M, Karhu R, Kallioniemi A, Elomaa I, Kivioja AH, Nevalainen J, Bohling T, Karaharju E, Hyytinen E and Knuutila S. (1995). *Cancer Res.,* **55,** 1334–1338.
Tarkkanen M, Elomaa I, Blomqvist C, Kivioja AH, Kellokumpu-Lehtinen P, Bohling T, Valle J and Knuutila S. (1999). *Int. J. Cancer,* **84,** 114–121.
Weber RG, Pietsch T, von Schweinitz D and Lichter P. (2000a). *Am. J. Pathol.,* **157,** 571–578.
Weber T, Weber RG, Kaulich K, Actor B, Meyer-Puttlitz B, Lampel S, Buschges R, Weigel R, Deckert-Schluter M, Schmiedek P, Reifenberger G and Lichter P. (2000b). *Brain Pathol.,* **10,** 73–84.
Yeh WC, Li TK, Bierer BE and McKnight SL. (1995). *Proc. Natl. Acad. Sci. USA,* **92,** 11081–11085.