

Frequent fusion of the *JAZF1* and *JJAZ1* genes in endometrial stromal tumors

Jason I. Koontz^{*†}, A. Lee Soreng^{*†}, Marisa Nucci[†], Frank C. Kuo^{*†}, Patrick Pauwels[‡], Herman van den Berghe[‡], Paola Dal Cin[†], Jonathan A. Fletcher[‡], and Jeffrey Sklar^{*†§}

^{*}Division of Molecular Oncology, [†]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115; and [‡]Center for Human Genetics, University of Leuven, B-3000 Leuven, Belgium

Communicated by Stanley J. Korsmeyer, Dana-Farber Cancer Institute, Boston, MA, March 16, 2001 (received for review December 19, 2000)

Endometrial stromal tumors are divided into three types: benign stromal nodules, endometrial stromal sarcomas, and undifferentiated endometrial sarcomas. A variety of cytogenetic abnormalities involving chromosome 7 have been reported in endometrial stromal sarcomas, including a recurrent t(7;17)(p15;q21). We have identified two zinc finger genes, which we have termed *JAZF1* and *JJAZ1*, at the sites of the 7p15 and 17q21 breakpoints. Analyses of tumor RNA indicate that a *JAZF1/JJAZ1* fusion is present in all types of endometrial stromal tumors; however, the fusion appears to be rarer among endometrial stromal sarcomas that would be considered high-grade according to certain classification schemes. These findings suggest that the less malignant endometrial stromal tumors may evolve toward more malignant types, but that some endometrial stromal sarcomas with relatively abundant mitotic activity may compose a biologically distinct group.

Endometrial stromal tumors of the uterus constitute a spectrum of neoplasms that exhibit varying degrees of malignancy and can present a number of challenges with respect to diagnosis and classification (1). Stromal nodules lie at the benign end of this spectrum of neoplasms. These nodules consist of well-circumscribed tumors composed of uniform cells resembling those of normal endometrial stroma during the proliferative phase of the menstrual cycle. Endometrial stromal sarcomas (ESSs) are malignant neoplasms that occupy the middle of the spectrum and are histologically similar to stromal nodules except for infiltration of the myometrium and/or vascular invasion. Tumors that depart significantly in histologic appearance from normal endometrial stroma are referred to as undifferentiated endometrial sarcomas (or undifferentiated uterine sarcomas) and represent the most malignant end of this tumor spectrum.

In the older literature, all malignant endometrial stromal tumors were categorized as ESSs and were subclassified into low-grade and high-grade types (2, 3). High-grade ESSs were distinguished from low-grade ESSs by an increased frequency of mitoses (>10 per high power microscopic field) and were generally assumed to have a worse prognosis. More recently, it has been argued that the number of mitoses within ESSs is largely irrelevant to outcome, which is said to be almost exclusively a function of stage at diagnosis (4–6).

Diagnosis and classification of endometrial stromal tumors has until now been based primarily on histologic criteria. In recent years, specific genetic alterations identified in different types of human tumors have provided useful diagnostic markers, led to insights into the basic biology of both neoplastic and normal tissues, and increasingly contributed to the development of rational forms of cancer therapy (7–9). With these considerations in mind, we have characterized the DNA and genes surrounding the breakpoints of a recurrent chromosomal translocation, the t(7;17)(p15;q21), reported in several cases of low-grade ESS (10–13). We have found that recombination at these breakpoints results in fusion of two previously unknown genes, which we have termed *JAZF1* and *JJAZ1*. Fusion of these genes appears to be common in low-grade ESSs but is not limited to these neoplasms and can be found in other types of endome-

trial stromal tumors as well. However, the incidence of this fusion appears to be reduced among ESSs that might be classified as high-grade.

Materials and Methods

Tissues and Cells. Initial molecular genetic studies were performed on four cases of low-grade ESS found to have the t(7;17)(p15;q21) (see Table 1 for the full karyotypes). In case BWH-42, the tissue was obtained from a metastatic retroperitoneal mass in a 58-yr-old woman who had undergone a hysterectomy for a low-grade ESS 16 years earlier; in case BWH-665, from a metastatic vaginal mass in a 69-yr-old woman who had undergone a hysterectomy for a low-grade ESS 18 years earlier; and in both case LU-550 and case LU-965, from primary uterine tumors in 41-yr-old women. Cells cultured directly from the tumors were grown in RPMI medium 1640 supplemented with 20% FCS, Mito + Serum Extender (Becton Dickinson), and bovine pituitary extract (Becton Dickinson) under standard conditions.

Yeast Artificial Chromosomes (YACs), Bacterial Artificial Chromosomes (BACs), and cDNAs. YACs and BACs were purchased from Research Genetics (Huntsville, AL). Methods for growth of yeast and bacteria, as well as procedures for isolation of YAC and BAC DNA, were as previously described (14, 15). cDNAs were identified in lambda bacteriophage libraries prepared from human fetal brain (Stratagene) and human umbilical vein endothelial cells (HUVEC) (16) by hybridization of probes to plaque lifts on nylon membranes. Phage DNA was isolated from clones in the HUVEC library according to standard methods (17), and phagemid cDNA clones were rescued from the brain library according to the supplier's instructions. Plasmids containing expressed sequence tag cDNAs (ESTs) were purchased from Genome Systems (St. Louis). Plasmid DNA was purified from bacteria following standard protocols (18).

Fluorescence *in Situ*, Southern Blot, and Northern Blot Hybridization.

DNA probes for fluorescence *in situ* hybridization (FISH) were labeled by nick translation with biotin-tagged nucleoside triphosphates by using the BioPrime labeling system (Life Technologies). Hybridization was performed on slides of metaphase chromosomes prepared according to methods of Fletcher and colleagues (15, 19). Procedures for Southern blot and Northern blot hybridization have been described (15).

Abbreviations: ESS, endometrial stromal sarcoma; YAC, yeast artificial chromosome; BAC, bacterial artificial chromosome; EST, expressed sequence tag; FISH, fluorescence *in situ* hybridization; RT-PCR, reverse transcription-PCR; STS, sequence-tagged site; I-PCR, inverse PCR.

[§]To whom reprint requests should be addressed at: Harvard Medical School, Brigham and Women's Hospital, Department of Pathology, 75 Francis Street, Thorn 605, Boston MA 02115. E-mail: jsklar@rics.bwh.harvard.edu.

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.

Table 1. Karyotypes of low-grade ESSs initially studied

Case	Karyotype
BWH-42	46, XX, t(7;13)(p15;p13), t(7;17)(p15;q21)
BWH-665	46, XX, t(7;17)(p15;q21)
LU-550	46, XX, t(7;17)(p15;q21)
LU-954	45, XX, -7, t(7;17)(p15;q21)

Inverse PCR (I-PCR) and Reverse Transcription (RT)-PCR. I-PCR was performed following protocols previously described (15). For RT-PCR, total RNA was transcribed into cDNA with Superscript II (Life Technologies), using either 7AntisenseOuter or 17AntisenseOuter as a primer (see below). The resulting cDNA was subjected to two rounds of PCR using first the “Outer” and then the “Inner” set of primers. The *JAZF1* primers were 7SenseOuter 5'-CCACAGCAGTGGGAAGCCTTA-3', 7AntisenseOuter 5'-GCTTCTCTTCCCTCCATTCAT-3', 7SenseInner 5'-ATCACCCCCTCCTTTCATT-3', and 7AntisenseInner 5'-GGACTCATCGCTGTCCGACT-3'. The *JJAZ1* primers were 17SenseOuter 5'-GTTACTCGGCCTCCTCCTCCTC-3', 17AntisenseOuter 5'-GGTTCAAATTCATTACTGGAACTGC-3', 17SenseInner 5'-GAGCTTTTCCTCCAGGCCTTTG-3', and 17AntisenseInner 5'-CCGGGTTTTGTTTGATTGAGG-3'. Specific primers for glyceraldehyde 3-phosphate dehydrogenase (5'-CACATCGCTCAGACACCATG-3' and 5'-GCCATGGAATTTGCCATGGG-3') were used to assess the quality of the input RNA.

For RT-PCR of RNA from formalin-fixed, paraffin-embedded tissue, 10 serial 10- μ m tissue sections were deparaffinized with three washes in 10 ml of xylene, washed twice with 100% EtOH, and then digested for 16 h at 60°C in 2 ml 1 \times digestion solution (20 mM Tris-HCl, pH 8.0/20 mM Na₂EDTA/2% SDS/2.5 mg/ml proteinase K). Eight milliliters of Trizol was added, and the RNA was isolated according to the supplier's instructions. First strand synthesis and nested PCR were performed with the following primers: FusionOutF 5'-CACGCCACAGCAGTGGGAAGC-3', FusionOutR 5'-TTTGTCTGGAGTTTCGATGAGACA-3', FusionInnerF 5'-CCCACCCATCACCCCCTCCT-3', and FusionInnerR 5'-GGTGCTATGAGATCCGAGTTCGAAGA-3'. A 10- μ m tis-

sue section adjacent to those used for extraction of RNA was examined histologically.

Determination of Nucleotide Sequence and Computer-Assisted Analysis of Derived Sequence. Sequence data were obtained by using the Amplicycle Sequencing Kit (Perkin-Elmer) and analyzed by using the National Center for Biotechnology Information (NCBI) BLAST server and the GCG software package (Genetics Computer Group, Madison, WI). Iterated sequence was suppressed in analyses for genes by use of the REPEATMASKER2 software program available at the MRC Mammalian Genetics Unit WWW server (<http://www.mgu.har.mrc.ac.uk/>). Peptides were analyzed by using software available at the Expert Protein Analysis System (ExPASy) WWW server (<http://www.expasy.ch/>).

Results

Identification of a YAC Spanning the Chromosome 7 Breakpoint. FISH was performed on metaphase preparations from the tumor of case BWH-42, using as probes YACs from the Whitehead Centre d'Étude du Polymorphisme Humain (CEPH) megaYAC contig 7.3 in the region of the *HOXA* gene cluster (Fig. 1a). FISH with YAC y830D2 produced hybridization signals limited to both derivative 7 (der7) chromosomes in this tumor, whereas the analysis with YAC y816F12 demonstrated hybridization to the der13 and der17 chromosomes. FISH with YACs y910G4 and y929H6 demonstrated hybridization signals on the der17 and the der7 of the t(7;13). These results indicated that the chromosome 7 breakpoints of both the t(7;17) and the t(7;13) of BWH-42 map within the 7.3 contig between YACs y830D2 and y816F12 but that the breakpoints are separated by more than 1 megabase. Furthermore, the chromosome 7 breakpoint in the t(7;17) lies between DNA contained within y830D2 and y910G4.

YAC y908B12, which contains DNA between clones y830D2 and y910G4, was then used as a probe and showed hybridization to both der7 chromosomes and the der17 chromosome of BWH-42, thereby demonstrating that y908B12 spans the chromosome 7 breakpoint of the t(7;17) (Fig. 1b). Additionally, analysis with YAC y908B12 of interphase nuclei from a second case of low-grade ESS, BWH-665, produced three hybridization signals (Fig. 1c).

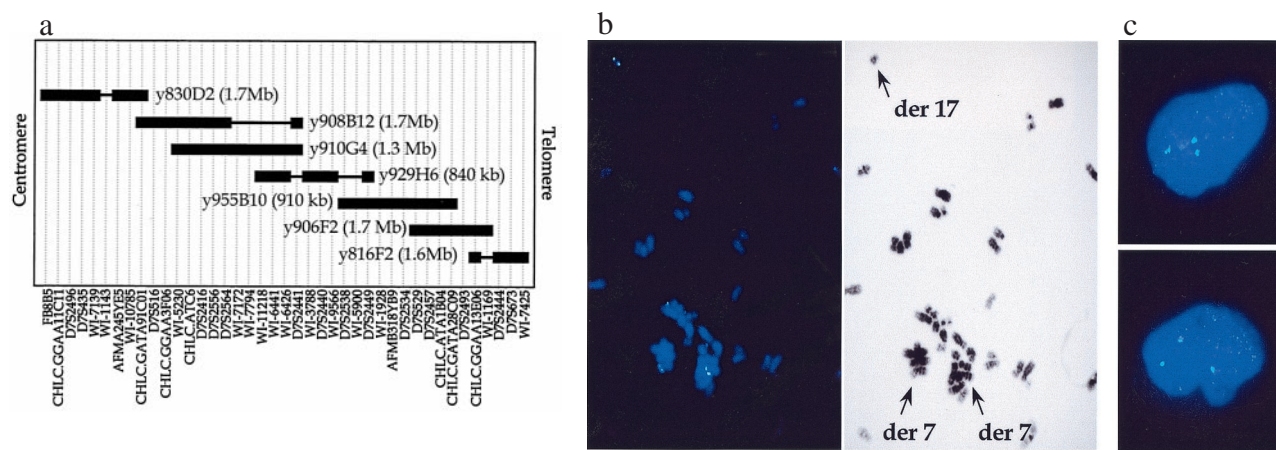


Fig. 1. Analysis of YACs in the vicinity of the 7p15 breakpoint. (a) Map of YAC clones at the 7p15 site. STSs are shown at the top without reference to the genetic distances between them. Solid, boxed areas indicate the presence of the STS within the DNA of the YAC; horizontal lines indicate the absence of the STS, presumably because of a deletion in the YAC insert. (b) FISH analysis of tumor BWH-42 with YAC probes derived from chromosome 7. Chromosomes and cells were hybridized with biotin-labeled DNA, and hybridization was detected with avidin-conjugated FITC. (Left) FISH was performed with YAC y908B12 and the chromosomal DNA counterstained with 4',6-diamidino-2-phenylindole (DAPI). The hybridization signal is split between two derivative 7 chromosomes—that of the t(7;17) to the left and of the t(7;13) to the right—and one derivative 17 chromosome, as identified in the contrast-enhanced image of the DAPI stain on the right. (c) FISH performed with the y908B12 probe on interphase nuclei isolated from paraffin-embedded tumor in case BWH-665.

Construction of a BAC Contig Across the Chromosome 7 Breakpoint.

An arrayed BAC library was screened by PCR using primers for the four sequence-tagged sites (STSs) CHLC.GATA91C01, D7S516, CHLC.GGAA3F06, and WI-5230, which lie at the centromeric end of the y908B12 YAC insert. This screen identified BACs b60B19, b319D14, b390A17, and b332C6. To determine overlaps and find additional BACs in the region, the sequences at the ends of the inserts of the BACs were determined. A search of the GenBank database indicated that DNA within a PAC clone, p881H5, included the telomeric sequence within b390A17. DNA of this PAC clone had been sequenced by the Genome Sequencing Center (GSC) at Washington University (<http://genome.wustl.edu/est/esthmpg.html>), and further examination of chromosome 7 sequence data on the GSC web site using the end sequences of the BAC inserts resulted in construction of the contig shown in Fig. 2a. FISH analyses on cases BWH-42 and BWH-665 using pooled probes from BACs b60B19 through b459N13 in this contig demonstrated an identical hybridization pattern to that seen with YAC y908B12 (data not shown). These results indicated that the BAC contig spans the chromosome 7 breakpoint in the t(7;17) of case BWH-42.

Identification of a Gene in the Chromosome 7p15 BAC/PAC Contig and Detection of the Breakpoint by Southern Blot Analysis.

The EST database was queried for matches to the genomic sequence available for this region. Probes for Southern blot analysis of tumor DNA were then constructed by PCR amplification of repeat-free segments of genomic DNA adjacent to or included in the expressed regions (Fig. 2b). Such a probe derived from the region of overlap between BAC b459N13 and PAC p881H5 (probe 459-10) detected non-germline bands in both *Bam*HI and *Hind*III digests of tumor DNA from case BWH-42 but not in similar digests of DNA from normal fibroblasts (Fig. 2c). The probe 459-10 also detected a non-germline band in a *Bam*HI digest of tumor DNA from case BWH-665.

The breakpoint region was further narrowed by hybridizing Southern blots of tumor DNA with two probes, 459-BP8 and 459-BP9, amplified from sequence lying 3-kb telomeric to 459-10 and separated from each other by 798bp. Probe 459-BP9 detected bands identical to those seen with probe 459-10 in the two cases. In both *Hind*III and *Bam*HI digests of DNA from BWH-42, probe 459-BP8 detected non-germline bands distinct from those in case BWH-665 whereas, in case BWH-665, probe 459-BP8 detected the non-germline band seen with probe 459-10, as well as a new non-germline band. Taken together, these results indicated that the chromosome 7 breakpoint in case BWH-42 lies in DNA between that contained in probes 459-BP8 and 459-BP9 and that the breakpoint in case BWH-665 lies within DNA sequence contained in probe 459-BP8.

Detection of a Tumor-Specific Transcript with Hybridization Probes for DNA Near the Chromosome 7 Breakpoint.

The chromosome 7 breakpoint in cases BWH-42 and BWH-665 mapped to the region of putative intronic sequence predicted by the alignment of the EST AA431106 with the sequences of b459N13 and p881H5. A series of overlapping ESTs presumably derived from a single gene were found in the database, and a hybridization probe consisting of the three most 3' exons of the gene was constructed. This probe was hybridized to Northern blots containing polyadenylated RNA purified from normal human fibroblasts, normal human endometrium, and four cases of low-grade ESSs known to contain a t(7;17) (Fig. 3). A band corresponding to a transcript of ≈ 3.2 kb appeared in analyses of RNA from both normal tissues and two tumors, and a second band corresponding to a length of about 4.5 kb appeared in analyses of RNA only from the four tumors.

To determine which part of the normal chromosome 7 gene hybridized to the 4.5-kb transcript found in tumors, probes were

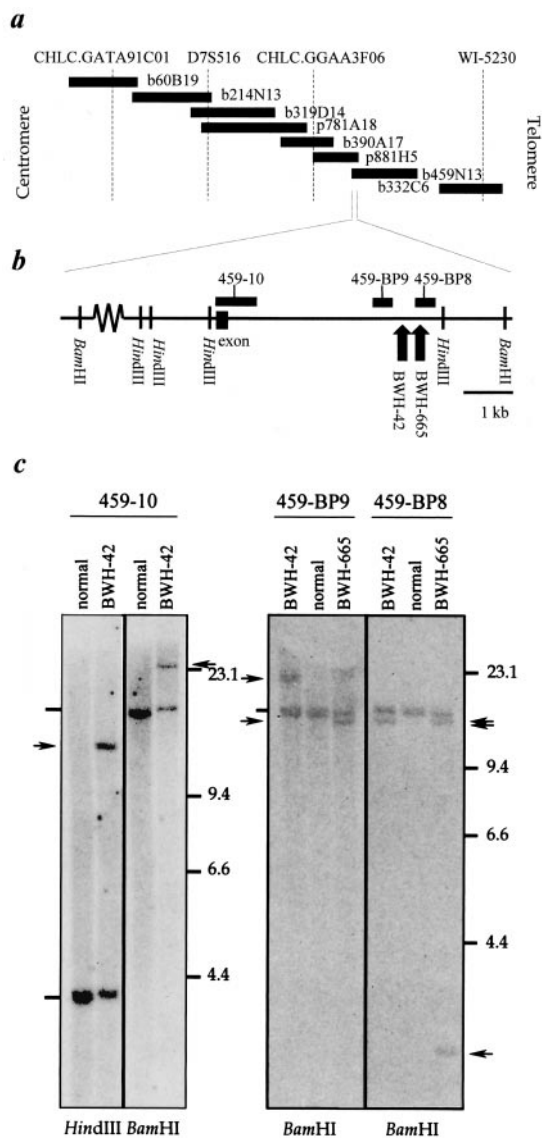


Fig. 2. Localization of the 7p15 breakpoint to the centromeric end of BAC b459N13. (a) Map of the BAC and PAC clones in the centromeric region of YAC y908B12. STSs are shown at the top of the map at relative positions based on DNA sequence available for the region. (b) Position of Southern blot hybridization probes within DNA near the centromeric end of BAC b459N13. The approximate positions of chromosome 7p15 breakpoints in cases BWH-42 and -665, as determined in the Southern blots in c, are illustrated with upward arrows near the telomeric end of the region shown. (c) Southern blot analyses of DNA derived from the tumors in cases BWH-42 and -665. Normal DNA was extracted from fibroblasts of the patient in case BWH-42. The probes used for each blot are indicated above the autoradiograms; the restriction enzymes used are shown below. Arrowheads indicate the positions of rearranged restriction fragments; dashes the positions of the germline restriction fragments.

designed to represent the portions of the gene on the centromeric and telomeric sides of the chromosome 7 breakpoint. Northern blot analysis with a centromeric probe yielded a pattern of bands identical to that seen with the original probe, and analysis with a telomeric probe yielded only the 3.2-kb band (data not shown). These findings indicated that the chromosome 7 gene is expressed in normal endometrium and that disruption of the penultimate intron of the gene in at least some tumors with the t(7;17) results in the production of an abnormal transcript that contains only the centromeric portion of the gene.

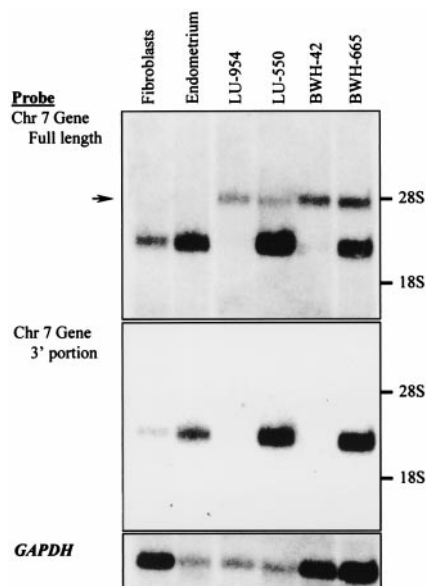


Fig. 3. Northern blot analyses of poly(A)⁺ RNA from ESSs and normal control tissues. The source of normal fibroblasts was the same as that in Fig. 2, whereas normal endometrium was obtained from an unrelated patient. (a) Analysis performed with a probe for the 5' end of EST AA431106 (nucleotides 74 to 1561 of *JAZF1* in Fig. 5a). The arrow indicates the position of a band corresponding to a ≈4.5-kb transcript seen only in tumor samples. (b) Analysis performed with a probe for the 3' end of EST AA431106 (nucleotides 1120 to 2883 of *JAZF1* in Fig. 5a). (c) Analysis with a probe for the *GAPDH* gene used as a control for the quality and amount of RNA on the blot.

Identification of a BAC Clone Spanning the Breakpoint on Chromosome 17. Two sets of nested PCR primers were constructed to be oriented in a tail-to-tail configuration and complementary to sequence 450 bp apart on each side of the telomeric *Hind*III site in Fig. 2b. Template for I-PCR with these primers was prepared by *Hind*III digestion and subsequent self-ligation of genomic DNA from tumor tissue in case BWH-42 and from normal fibroblasts of the same patient. I-PCR amplification of these templates produced a predicted, unrearranged 3.5-kb product from both tumor and normal DNA and a predicted, rearranged 1.6-kb product only from the tumor DNA (data not shown).

The 1.6-kb I-PCR product was ligated into the pCR2.1Topo vector, and the sequences at the ends of the insert of the resulting clone were determined. One end sequence corresponded to chromosome 7 sequence up to a *Hind*III site and then immediately diverged into non-chromosome 7 sequence. When this new sequence was used to search the GenBank database, it was found that a BAC, b307A16, from chromosome 17 contained matching sequence. These results indicated that the breakpoint at chromosome 17q21 in case BWH-42 was contained in DNA of b307A16.

Identification of a Gene in the Chromosome 17 Breakpoint Region. Analysis of DNA within b307A16 revealed that about 3.8 kb of a 4.44-kb EST, KIAA0160, map within genomic DNA beginning less than 20 kb away from the chromosome 17 breakpoint; however, the first 647 bp of this EST is not present anywhere in b307A16 or the overlapping BAC b542B22. This observation suggested that a deletion had occurred in the genomic DNA contained in b307A16 and that the predicted gene extends across the chromosome 17 breakpoint. If this gene were disrupted by the chromosome 17 breakpoint, the orientation of the gene would be consistent with a fusion of this gene with the gene disrupted by the breakpoint in chromosome 7. To test this possibility, RNA derived from tumor BWH-42 and from control

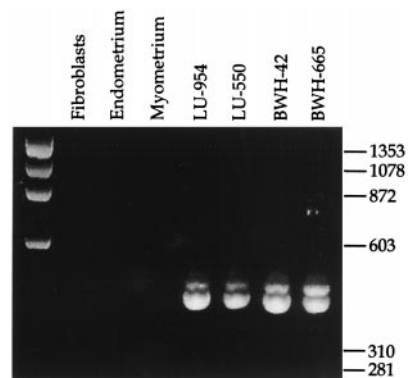


Fig. 4. RT-PCR analysis of the RNA sequence transcribed across the site of recombination in the t(7;17)(p15;q21). Ethidium bromide-stained agarose gel of RT-PCR products amplified from total RNA of normal tissues and the four low-grade ESSs known to carry the translocation. The source of fibroblasts and normal endometrium was the same as that in Fig. 3; the normal myometrium was obtained from the same specimen as the endometrium.

fibroblasts was analyzed by RT-PCR using primers complementary to sense sequence at the 5' end of both the 7 and 17 genes and anti-sense sequence located at the 3' end of both genes, with the primers paired in all four possible ways. In these experiments, the primer pairs for the intact chromosome 17 gene amplified products from the RNA of both the normal tissue and the tumor, but the recombinant 5' chromosome 7–3' chromosome 17 pair amplified a product of about 440 bp only from the tumor RNA (Fig. 4). The reciprocal recombinant primer pair failed to amplify a product from either of the RNA samples. The 5' chromosome 7–3' chromosome 17 primer pair used in an RT-PCR assay of RNA derived from the three other tumors analyzed previously in Northern blots yielded similar findings. These results demonstrated that the chromosome 17 gene was disrupted by the translocation and that a fusion transcript containing the 5' end of the chromosome 7 gene and the 3' end of the chromosome 17 gene had been created. Sequence analysis of the RT-PCR products confirmed that RNA sequence transcribed from chromosome 7 had been joined to RNA sequence transcribed from chromosome 17 at the sites indicated in Fig. 5.

Structure of the Genes Identified at 7p15 and 17q21. Fig. 5a presents the sequence of the chromosome 7 gene, as deduced from the GenBank and EST databases plus an additional 176 bp at the 5' end acquired through screening of two cDNA libraries derived from human brain and human umbilical vein endothelial cells (HUVECs). The entire sequence of the chromosome 17q21 gene is contained in the EST KIAA0160 (Fig. 5b). The true 5' ends of these transcripts remain unknown because the ORFs continue to the 5' ends of the sequences assembled so far.

Conceptual translation of the ORF of the chromosome 7 gene revealed an N-proximal zinc finger domain as well as a nearly tandem pair of zinc finger domains close to the C terminus. Similar analysis of the chromosome 17 gene sequence revealed a zinc finger domain in the C-terminal third of the coding sequence. All four of these zinc finger domains are of the Cys₂His₂ class. Other than a possible bipartite nuclear localization sequence in the chromosome 17 gene, these domains represented the only identifiable motifs in the coding sequences and the only significant similarity to other known genes. In light of these structural features, we refer to the chromosome 7 gene by the acronym *JAZF1*, for Juxtaposed with Another Zinc Finger gene, and the chromosome 17 gene by *JJAZ1*, for Joined to *JAZF1*.

Analysis of Endometrial Stromal Tumors for the *JAZF1*/*JJAZ1* Fusion Transcript. RT-PCR was performed on RNA extracted from paraffin-embedded, formalin-fixed tissues of five archival cases

a

CCGCCGGGCGCCGTCAGAGGGGAGACGAGGGGCGCTCGGAGTAGACACATGACAGGCCATCGGCCGCCCTCTCTTCCCAATCC 90
M T G I A A A S E F S
TGGCGATTGGGGGGTGGGGACTCCACTTCCCACTCCCTGGGAGAGCTCATGAGGACACATGAGGACACACCAATGATACAGCATCGAGG 180
R F F G C L L H F L F L D L I E H I E D N I D T D P R
GTTTTAAAGAAAAGGAAATACAGCAGCCACCTATGTTCOCCTGAGTACATAAATAGATTGATGACATGATGCTGCGCGCGAGAGAGAG 270
V L E K Q E L Q P T Y V A L S Y I N R F M T D A A R R E Q
GAGTCCCAAAAGAGAGATCAGCAGAGCTCTCCTGAGCTCTCCAGCAGTGTCTCAGAGGAAATGTGTCCACTCCCACGGCCAGC 360
E S L K K K I Q P K L S L T L S S V S R G N V S T P P R H
AGCATGGAGAGCTTACTCCCGCTGACCCACCACATCCCGCCCTCTCATCTCCGACAGCAGCCTCCGACAGCAGCAGGAGATGATC 450
S S G S L T P P V T P P T P P I T P S S F R S S S T P T G S E Y D
GAGGAGAGGTGGACTATGAGGAGTGGACAGCGATGATGGCTGACAGGAGATGATCCACTGAGTCCGAGAGCATCTCCAGCTCCCTG 540
E E E V D Y E E S D S D E S W T E S A I S S E A I L S S M
TGCATGATGGAGGGAGAGAGAGCCTTGGCTCCGAGCTCTCGAGTGTAAAAGAGATACAGAAATGTGAATGACATAAAGATGATCC 630
C M N G G E E K P A C P V F G C K R Y K N V N G I R Y
GCTAAGATGTCACAGACAGCAGATGTGTCCGCAACATTCAGTGTGCTGGAGAGATGTCAGAGCAGCTCAGGCCCTCAGGCCCTGG 720
A K N G R R T Q I R V R K P F K R C R C G K S L Y T A Q G L R
CACACCAATCAATTCCATCCCGGCTGGCGTGGATATCAGGAGAGTGGCAATACATAGCTGGTCAATACTGTGCCAAGA 810
E T I N P P P V S A E I I R K M Q
CCTCCACGAGCTGCTGATTGTAAAAGAGCAGCTCTTTTCGAGGGAGCATTGAGCAACCTTTAAGAAAAGAGATTAATGAT 900
TCTTAAATTTTTCTGATTTTGAAGAGATGTATCTTTGAGAGTAAAGTATTGACATTTGACATGATGATATCATACACATCC 990
TGACTTTGATAGAT 1080
AGATTTTGAATTTGATGCGTGGTATGAT 1170
CATTTCACACATCAACATCACAATATGTCAAAACAAAATAATACCTTATCAAACTGTTACAGGCTTACAAATGGATGAT 1260
TATTGTAGTAGAAGAGCTTAGTAAAGAAATAAAAGAAGAAATGTTAACAATAATGCAAAATATTTTCAATTTAAAT 1350
AGCATAGTAAGGTGCTTCTGTGTTTATTATCTTGAAATCTTTGCCCTTAAAGAGTCAAAATCTGCATCTGACCTA 1440
CTACTGATTTAGTATTAATAAGTCTTTGTAGAGCTCTCTGTTCCTCTTCAAGCTAATTTTACTTTTTTAAAGAGT 1530
TAGGCAAGAGGATCAATTTGAT 1620
TGCAGCTTGAATCTTCAGTCAAGGTGTGTGTGAACTTTAGTATATCTCTCATCTCAGATGAAATTTTAAAGACCTTGTAG 1710
TTCCTTATGCAACATTTATGTTATGTGTCACATTTGCAACCTGCTGCAACAGCTGGTGGTGCAAACTGAGACAT 1800
TATAGAGCCTTTCAAGTACACTTTGTAAGACTCTTTGCTCTCTTCAAGCTAATTTTACTTTTAAAGAGT 1890
TCTTCATCAATCTGATAGAGTGTGATGTAATGAAATGTTGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
AGCCCTATGCAATTAAGTCTTTCGACAGGTGTCAAGGAAAGCAGTCAATAGAGCTATTAGAGTAACTTACATCTGCGCC 2070
ATCTTAAGTTCCTCCAGCAGTACCACTGTGCTGTGTCTTTCCCAAGCTATTCCTTAACAGCAGGATTTACCA 2160
TGCCTAATATGATTAATAATGCTTTGATAGTCACTGCTCAATTAAGCAGTCAATCTTCAATGATTTTCCCTCATTG 2250
CATTCTGTAAGAGTCTTACATCAATCTGAGGACAGGCTCTGCTCATCTTTTCGAGAGTGTGAAGAGTACAGCTCTG 2340
TGTGTATGATGAACTCTTATTCTAATGAACTTAAATCTCCTATTAGGAGGATCAAGATTTTAAAGACAGCTGGCCATA 2430
AATTTCCCTCAGATGATGCTTGGCTCGAATATACAGTGTATTCTATTTTAAAGAGGTTGATCCAGAGTGAAGGCAAT 2520
GTGTGTGATGAAAGAGCCTCCAGTTTCATATCTACATAATGTACCAATTAATGCAAAATTTTAAAGCAGTATTTATTTAG 2610
TCCCTCTCTGATGTCGAGAGTCTGATGAGTCAAGCTCAACATGAAGATCAATTAATGATGATGATGATGATGATGATGATGATGAT 2700
TCTCAAGGTGTCGAGAGTCAATTTCAATTTCTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2790
ATATCTGTGAT 2880
TAAGCTGAAAGCAGGAGGTTGTTTGAAGTGAAGCAATTAATCAATCTCTGTCAAAAGAGTCTTCAATTAATGATGATGAT 2970
AATCAGATTTAAATTAAGCTGCTGACCCAAAATA 3010

b

CTCTGAGGAGACATTTTTTCTCCT 90
CTGCGTCCCGGAGCCTCTGCGGAGGCTGTGATTTGGAGGCGTCTGCTCTCCGAGGCGCAGCCAGCGGTAGCTGCGGGGAGG 180
AGGCGAAAGGCGGCTCTGCGGAGGCTGTGATTTGGAGGCGTCTGCTCTCCGAGGCGCAGCCAGCGGTAGCTGCGGGGAGG 180
M A P D K H G E G G G S G S G G G P
GGGTTGGGAGGCTGGGAGGCGGCGTCTGCGGAGGCGCAGCCAGCGGCGGAGCTGGGAGGAGCTGGGAGGCTGGGAGGCTGGG 360
G S A A V A A A T A S G G K S G G G S G G G V S A S
CCT 450
S S B A A A A A A G A A V L F V K K P K H N H Y Q D H E
TCTCAGGCTTTGAGAGTCAACAGCTATAGATTTCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 540
L Q A F E H P T Q I Y R F L R T R N L A F I F L H R T
CTCAGCTCTCCT 630
Y M S H R N S R T N I R K R T P K V D H L S K V E K N K G
GAGGACAGATCTCAGTCTGAGCTCATGCTCAGCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 720
E Q S H S L S A H L Q L T T E F F H K N D K P S F N S E
AAATGAAACAAATCTGTTACCTCGGAGCTGCTGTGAAAGTGTGCAAAAGAGGAGATGATGATGATGATGATGATGATGATGATGAT 810
N E Q S V T L E V L L V G H C H K R K D V S C P I R Q V
TCCACAGTAAAGAGGAGGCTTGTCTCTGAGTCAATCAACAAAGGCGGAAATTTCCGCTCCTGATGATTTCCGATGAT 900
P T G K K Q V P L I P D L K T F G V N F F S L A V S N E
AAITTTGACTAGTACAGCAGTGTGATGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990
F E P S N S H M V K S Y S L L R V R T R P G R R E F N G M I
TAAITGAGAAACCAATAAATTTGAT 1080
N G E T N E I D V N E L P A R R K N R E D D E K T F V
TTCAGGAGTCAAGTATTTGATTAACAGGCGCTTACAGTCTTAAAGTGGGAAATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1170
A Q M T V F D K N R L Q L L D G E V E A M Q E M E E C P
CAATAGCAGAAAGAGGCAATGGAGGATTTCTGAGGAGAGGCTGCTCCTTCAAGCAATTTTCCAGGAGCTCAGTTGG 1260
I S K R A T W E T I L D G K R L P P F F E T F S O G P T L G
AGTTCACTCTTCTGGACAGGAGCAAGTAAATACAGTCTCCTTGGCAAGCTGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1350
T L R T M G E T N D K S T A P I A K P L A T R N S E L H
ATCAGAAACAGGCTGTGATTAACCTACTCAACTGCTTTAAAGTAAATCATTAGCAGATCTCAACAGAGAAAGAA 1440
Q E N K P G S V K P T O T I A V K E S L T T D L Q T R K E
AGATATCTCATAAGAAACAGCAAAATTAAGATATTTATCAGTCTTCTATACACAAATCAAGGCGCAAACTGAGGCAAG 1530
D T P N E N R Q K L R I F Y Q F L Y N N N T R Q O T E A R D
ATGACCTCAGTCCCTTGTGTACTCTGAACTGCAAACTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
D H C P M O T L N C R A K Y S L L K B L K L C S R F I
TCAACTGTTTATCATCCAAAGGCTGAGGATGATGATTTCTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1710
N Y Y H P K G A R I D V S I N E Y D G S Y A G N P Q D I
TTCATCCCAAACCTGATGCTTTTGTGCGAGCGGAGCTTAAAGAGACACTCAACACATTTCTGTCGAGCGGCAAAAGCAG 1800
H R Q P G F A F S R N G P K E T P I T I L V C R P K R
CAAAAGCAAGATGCTGAACTTGTGATGAGTGGGAGTGAAGCAAGGAGGAGATAGTAGTGGCCAGCATCTCTGCTGAT 1890
K A S M S E F L E S E D G E V E Q R T Y S S G H N R L Y
TCCATGAT 1980
H S D T C L P L R P Q E M E V D S E D E K D P E M L R E K
CCATACACAAATGAGAGTTCCTGATGATTAAGGAGGAGAAAGGATGATGAACTCGGATCTCCATGCTCATGAGAGCTGGT 2070
I T Q I E E F S D V N E G E V M L K W L N L H V M K H G F
TTATTGTCAGCAATCAATGAT 2160
I A D N Q M N H A C M L F V E N Y G Q K I I K N L C R N F
TCACTGTCTCAGTCAAGCAGGAT 2250
M L H L V S M H D F T A I S I M S I D K A V T K L R E M Q Q
AAATATGAAAGAGGAGGATGCTCCT 2340
K L E K E S A S P A N E I T E Q N G T A N G F S E I N
ACTCAAAAGGAGGAGGATGCTCAGGAGTGTTCAGGAGGCGCAAAAGCAAAACTCTGAAAAGCTCAACCCAT 2430
S K E K A L E T D S V S G V S K Q S K K Q K L
TGTATGAGACAACTGAAATGATGATTTAGGAGTCAATCCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
GTATGATGAGGAGGAT 2610
GTATGATCAAAACTCTGATGCTTAAAGCAAGAGCAATTTCAAGTATTTTAACTCAACAATGATCATCAATGATGATGATGATGAT 2700
TGTCTGATGAAATTTTATCATAATAAGCAGATTTTTCGATTTGAAAGGCGCTTTTCTACTTTGATTTGGGAGCATTCT 2790
TGGGAGGAGGAGGATGAGGAGTTCCTTTTAGAAATTTGAGTGTCTCATATGTCACATACAGAAAGGAAAGAAATTTGATGAG 2880
AGGATTTTAAAGATATATAGTACTTATTCGATCAAACTTGAATCTGTTGTTGAAATCTTTGCAATTTGGAGATGATGATGAT 2970
TAAATGATTTTCTCTGAT 3060
TCAAGTGTGACAGTGAAGTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3150
GTATGTTTATTTTAACTGAACATTTTCTGTTTGTGTTTTTAAATAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
CTTTAAATTTGAT 3330
CAAGAACTGAT 3420
AATCATTTGAGGAGGAT 3510
TGCACATTTGACAGGAGGAGTGTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
GTATGATTTTAACTGAGATTTTCTGTTTGTGTTTTTAAATAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3690
CAGTGTGATTTTCTTATTTGAGGAGAT 3780
ACTGCTTTAAAGGAT 3870
TTTTGTTTAAAGGAGTGAAGGAGGAGACTGCTTTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960
AGTAACCTGATTTTCTGCTGAT 4050
AAAAAAGAGATTTTAGGAGGAT 4140
AGGATGATGATTTTGGCAAAAGGAT 4230
TAATCTGTAAGGAT 4320
GATGTAAGACATTTGAT 4410
AAATGATTTTGTACATAAGTCTCAATAAT 4410

Table 2. Summary of RT-PCR results for JAZF1/JJAZ1 fusion RNA in archival tissue specimens

Tumor type	Total cases	Positive cases
Stromal nodule	3	3
Low-grade ESS	5	5
High-grade ESS*	7	3
Normal endometrium	10	0

*Includes two cases classifiable as undifferentiated endometrial sarcoma, of which one was positive.

of low-grade ESS, for which cytogenetic analysis had not been done, of three stromal nodules, and of seven high-grade ESSs. RT-PCR performed with sense primers for both *JAZF1* and *JJAZ1* and nested antisense *JJAZ1* primers demonstrated amplifiable template from each case, as judged by the generation of *JJAZ1* product. Each of the three stromal nodules was also positive for the fusion transcript. Among the ESSs, all low-grade cases were positive in this assay, but only three of seven high-grade cases were positive (Table 2). Two of the high-grade tumors would be considered undifferentiated endometrial sarcomas in the more recently proposed classification system, and of these, one was positive for the fusion and the other not.

Discussion

Chromosomal translocations generally affect the malignant phenotype of cells that contain them through various kinds of alterations in genes at or near the breakpoints in DNA of one or both of the participating chromosomes (20). The t(7;17)(p15;q21) appears to be an example of a chromosomal translocation resulting in a gene fusion, in which the 5' end of one gene, termed here *JAZF1*, on chromosome 7 is joined to the 3' end of a second gene, termed *JJAZ1*, on chromosome 17. In this respect, the t(7;17)(p15;q21) resembles the majority of chromosomal translocations so far analyzed in soft tissue sarcomas, most of which contain fusions of *EWS* or a related gene to a second gene encoding a DNA-binding protein (21, 22). Whether detection of the *JAZF1/JJAZ1* fusion always represents the presence of the t(7;17)(p15;q21), or may be found in cells without a translocation identifiable by cytogenetic analysis, remains to be determined.

The cDNAs for *JAZF1* and *JJAZ1* appear to specify the synthesis of proteins composed of 243 and 739 aa, respectively. Determination of the amino acid sequences encoded by the genes is complicated by the fact that large ORFs are continuous with the 5' ends of the cDNA sequences assembled for the two genes; however, several features of the cDNA sequences for both genes support the conclusion that the first methionine codons represent the actual beginning of the coding sequences. For instance, the total length of the cDNA sequences for the two genes coincides with the estimated size of mRNA in bands identified by Northern blot hybridization. Furthermore, in RNase protection assays, probes generated from genomic DNA overlapping sequence at the 5' end of the *JAZF1* cDNA protected a fragment of RNA from liver and lung cells corresponding to a size consistent with the cDNA representing the full

length of the transcript for this gene (comparable experiments for *JJAZ1* were not possible because genomic sequence at the 5' end of this gene has not yet been identified). Also, the nucleotides surrounding the putative initiation codon in both genes (CACCATGA in *JAZF1* and CGCGATGG in *JJAZ1*) conform well to the consensus rules proposed for vertebrate translational start sites (23). Finally, the mouse ortholog of *JAZF1* (contained in the overlapping murine ESTs AI595264, AA061309, and AI428135) shows striking homology to the human cDNA over the entire presumptive ORF but diverges from the human cDNA 5' of the position of the codon provisionally assigned as the first methionine. Similarly, the *JJAZ1* *Drosophila* ortholog (encoding a conceptually translated protein AAF49094) in genomic fly DNA contains termination codons in all three reading frames shortly upstream of the position corresponding to the putative first methionine in the human gene.

The specific functions of *JAZF1* and *JJAZ1* and the reason for their involvement in a gene fusion associated with neoplasia are not directly apparent from the sequences of the cDNAs. The only recognizable regions within the two cDNAs resembling sequences within known genes in humans and other species encode zinc finger motifs, as often found in DNA binding proteins (24, 25). The two C-proximal zinc finger domains encoded by *JAZF1* are particularly similar in structure and relative spacing to two zinc fingers in the yeast protein Sfp1p, although the homology between these two proteins is limited to these two paired zinc fingers. This homology, albeit slight, might seem noteworthy in a cancer-related gene because Sfp1p negatively regulates the G₂/M transition in the cell cycle of *Saccharomyces cerevisiae* by serving as a transcriptional activator of the gene *PDS1* (26, 27), the product of which inhibits Esp1p (28, 29), a protein that dissociates the cohesin complexes responsible for holding together sister chromatids during G₂ (30). However, in transfected cultured cell lines, *JAZF1* protein appears to have no significant effect on transcription of reporter genes driven by the promoter of *PTTG1* (31, 32), the human ortholog of *PDS1* (data not shown).

Another finding possibly relevant to the role of *JAZF1* in oncogenesis is the lack of *JAZF1* RNA in tumors BWH-42 and

LU-954. The reason for the absence of this RNA is different in the two cases: tumor LU-954 is monosomic for chromosome 7, whereas tumor BWH-42 contains unrearranged DNA for *JAZF1* on the allele not broken by the t(7;17), as demonstrated by Southern blot analysis. Failure to detect RNA in the latter case may therefore be due to transcriptional silencing of the normal copy of *JAZF1*, as has been noted in the unrearranged alleles of *MYC* and *BCL2* in some tumors bearing chromosomal translocations with breakpoints in or near these genes (33, 34). Loss of expression for normal versions of *JAZF1* in multiple tumors suggests a possible role of this gene as a tumor suppressor, similar to the situation observed for *TEL*, which is involved in chromosomal translocations and has tumor suppressor activity in some acute lymphoblastic leukemias (35–37).

The distribution of the *JAZF1/JJAZ1* fusion among uterine stromal tumors carries a number of implications for the biology of these neoplasms. For example, the finding of the fusion in stromal nodules raises the possibility that malignant endometrial stromal tumors can develop from stromal proliferations that are initially benign. Additionally, of the seven high-grade ESSs studied, only three showed evidence of the fusion, suggesting that some high-grade ESSs may arise by a different pathogenetic mechanism than low-grade ESSs, despite the current tendency to consider all of these tumors as clinically homogeneous. Alternatively, this discrepancy between high and low grade tumors could be due to the relatively small sample of cases studied or to the fact that we analyzed tumors for only one form of the fusion and that functionally equivalent genetic recombinations between *JAZF1* and *JJAZ1* at intragenic sites that would not have been detected by us may be present in some tumors. These explanations notwithstanding, if, after study of more high-grade ESSs, the fusion is confirmed to be present only within a subset of such tumors, it will be important to correlate possible differences in clinical behavior with the presence or absence of this genetic lesion.

We thank Megan Smith for valuable technical assistance. This work was supported by the National Foundation for Cancer Research and National Institutes of Health-National Cancer Institute Grant CA-85995.

- Zaloudek, C. & Norris, H. J. (1994) in *Blaustein's Pathology of the Female Genital Tract*, ed. Kurman, R. J. (Springer, New York), pp. 487–528.
- Norris, H. J. & Taylor, H. B. (1966) *Cancer* **19**, 755–766.
- Tavassoli, F. A. & Norris, H. J. (1981) *Histopathology* **5**, 1–10.
- Evans, H. L. (1982) *Cancer* **50**, 2170–2182.
- Chang, K. L., Crabtree, G. S., Lim-Tan, S. K., Kempson, R. L. & Hendrickson, M. R. (1990) *Am. J. Surg. Pathol.* **14**, 415–438.
- Hendrickson, M. R., Longacre, T. A. & Kempson, R. L. (1999) in *Diagnostic Surgical Pathology*, ed. Sternberg, S. S. (Lippincott Williams & Wilkins, Philadelphia), pp. 2203–2305.
- Pegram, M., Hsu, S., Lewis, G., Pietras, R., Beryt, M., Sliwkowski, M., Coombs, D., Baly, D., Kabbinar, F. & Slamon, D. (1999) *Oncogene* **18**, 2241–2251.
- Heise, C., Sampson-Johannes, A., Williams, A., McCormick, F., Von Hoff, D. D. & Kirn, D. H. (1997) *Nat. Med.* **3**, 639–645.
- le Coutre, P., Mogni, L., Cleris, L., Marchesi, E., Buchdunger, E., Giardini, R., Formelli, F. & Gambacorti-Passerini, C. (1999) *J. Natl. Cancer Inst.* **91**, 163–168.
- Sreekantiah, C., Li, F. P., Weidner, N. & Sandberg, A. A. (1991) *Cancer Genet. Cytogenet.* **55**, 163–166.
- Dal Cin, P., Aly, M. S., De Wever, I., Moerman, P. & van den Berghe, H. (1992) *Cancer Genet. Cytogenet.* **63**, 43–46.
- Pauwels, P., Dal Cin, P., Van de Moosdijk, C. N., Vrints, L., Sciort, R. & van den Berghe, H. (1996) *Histopathology* **29**, 84–87.
- Hennig, Y., Caselitz, J., Bartnitzke, S. & Bullerdiek, J. (1997) *Cancer Genet. Cytogenet.* **98**, 84–86.
- Philippson, P., Stotz, A. & Scherf, C. (1991) *Methods Enzymol.* **194**, 169–182.
- Morgan, J. A., Yin, Y., Borowsky, A. D., Kuo, F., Nourmand, N., Koontz, J. I., Reynolds, C., Soreng, L., Griffin, C. A., Graeme-Cook, F., et al. (1999) *Cancer Res.* **59**, 6205–6213.
- Ginsburg, D., Handin, R. I., Bonthron, D. T., Donlon, T. A., Bruns, G. A., Latt, S. A. & Orkin, S. H. (1985) *Science* **228**, 1401–1406.
- Lech, K., Reddy, K. J. & Sherman, L. A. (1990) in *Current Protocols in Molecular Biology*, eds. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (Wiley, New York), Vol. 1, pp. 1.13.11–11.13.10.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Fletcher, J. (1994) in *Current Protocols in Human Genetics*, eds. Dracopoli, N. C., Haines, J. L., Korf, B. R., Moir, D. T., Morton, C. C., Seidman, C. E., Seidman, J. G. & Smith, D. R. (John Wiley & Sons, New York), Vol. 1, pp. 10.13.11–10.13.10.
- Rabbitts, T. H. (1994) *Nature (London)* **372**, 143–149.
- Ladanyi, M. (1995) *Diagn. Mol. Pathol.* **4**, 162–173.
- Dei Tos, A. P. & Dal Cin, P. (1997) *Virchows Arch.* **431**, 83–94.
- Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148.
- Berg, J. M. & Shi, Y. (1996) *Science* **271**, 1081–1085.
- Clarke, N. D. & Berg, J. M. (1998) *Science* **282**, 2018–2022.
- Xu, Z. & Norris, D. (1998) *Genetics* **150**, 1419–1428.
- Nasmyth, K., Peters, J.-M. & Uhlmann, F. (2000) *Science* **288**, 1379–1384.
- Yamamoto, A., Guacci, V. & Koshland, D. (1996) *J. Cell Biol.* **133**, 99–110.
- Yamamoto, A., Guacci, V. & Koshland, D. (1996) *J. Cell Biol.* **133**, 85–97.
- Ciosk, R., Zachariae, W., Michaelis, C., Shevchenko, A. Mann, M. & Nasmyth, K. (1998) *Cell* **93**, 1067–1076.
- Zou, H., McGarry, T. J., Bernal, T. & Kirschner, M. W. (1999) *Science* **285**, 418–421.
- Kakar, S. S. (1999) *Gene* **240**, 317–324.
- Bentley, D. L. & Groudine, M. (1986) *Mol. Cell. Biol.* **6**, 3481–3489.
- Cleary, M. L., Smith, S. D. & Sklar, J. (1986) *Cell* **47**, 19–28.
- Golub, T. R., Barker, G. F., Lovett, M. & Gilliland, D. G. (1994) *Cell* **77**, 307–316.
- Golub, T. R., Barker, G. F., Stegmaier, K. & Gilliland, D. G. (1996) *Curr. Top. Microbiol. Immunol.* **211**, 279–288.
- Takeuchi, S., Seriu, T., Bartram, C. R., Golub, T. R., Reiter, A., Miyoshi, I., Gilliland, D. G. & Koefler, H. P. (1997) *Leukemia* **11**, 1220–1223.