Cell, Vol. 84, 587–597, February 23, 1996, Copyright ©1996 by Cell Press

The *FHIT* Gene, Spanning the Chromosome 3p14.2 Fragile Site and Renal Carcinoma–Associated t(3;8) Breakpoint, Is Abnormal in Digestive Tract Cancers

Masataka Ohta,* Hiroshi Inoue,*†
Maria Grazia Cotticelli,*† Kumar Kastury,*†
Raffaele Baffa,*† Juan Palazzo,*
Zurab Siprashvili,* Masaki Mori,‡
Peter McCue,* Teresa Druck,*
Carlo M. Croce,* and Kay Huebner*
*Kimmel Cancer Center
Jefferson Medical College
Philadelphia, Pennsylvania
‡Department of Surgery
Medical Institute of Bioregulation
Kyushu University
Beppu 874
Japan

Summary

A 200–300 kb region of chromosome 3p14.2, including the fragile site locus *FRA3B*, is homozygously deleted in multiple tumor-derived cell lines. Exon amplification from cosmids covering this deleted region allowed identification of the human *FHIT* gene, a member of the histidine triad gene family, which encodes a protein with 69% similarity to an S. pombe enzyme, diadenosine 5′, 5‴ P^1, P^4-tetraphosphate asymmetrical hydrolase. The *FHIT* locus is composed of ten exons distributed over at least 500 kb, with three 5′ untranslated exons centromeric to the renal carcinoma–associated 3p14.2 breakpoint, the remaining exons telomeric to this translocation breakpoint, and exon 5 within the homozygously deleted fragile region. Aberrant transcripts of the *FHIT* locus were found in ~50% of esophageal, stomach, and colon carcinomas.

Introduction

In 1979, an Italian-American family was observed to be transmitting a constitutional reciprocal t(3;8)(p14.2;q24) chromosome translocation (Cohen et al., 1979; Wang and Perkins, 1984), which segregated in the family with early onset, bilateral, and multifocal clear cell renal carcinoma (RCC). It was suggested that the translocation affects expression of a tumor suppressor gene (Cohen et al., 1979), and several investigators have sought candidate suppressor genes. We had suggested the protein-tyrosine phosphatase γ gene (*PTPRG*) as a candidate tumor suppressor gene (LaForgia et al., 1991) and showed that the 5′ end of the *PTPRG* gene maps near but centromeric to the 3p14.2 break (Kastury et al., 1996b) and that the majority of clear cell RCCs exhibit loss of heterozygosity of a 0.5 Mb region flanking the translocation (Lubinski et al., 1994; Druck et al., 1995), although we did not observe aberrations in the remaining *PTPRG* gene. The 3p14.2 region is also included in deletions in numerous other tumor types, including nasopharyngeal carcinomas (Lo et al., 1994).

The t(3;8) translocation breakpoint was cloned, and a 3 kb transcript of a candidate tumor suppressor gene was detected using a probe from near the breakpoint (Boldog et al., 1993); further details concerning this transcript have not been reported. Cytogenetic studies in several of the familial tumors demonstrated that the tumors had lost the derivative 8 chromosome carrying the translocated 3p14-pter region; consequently, the tumors were homozygous for all loci telomeric to the 3p14.2 break (Li et al., 1993), suggesting that the putative suppressor gene should be telomeric to the 3p14.2 break. It has also been suggested that there may not be a suppressor gene at 3p14.2, that in fact the t(3;8) translocation was a mechanism for losing the von Hippel–Lindau gene, a tumor suppressor gene at 3p25 (Gnarra et al., 1994).

Another cytogenetic landmark in chromosome region 3p14.2 is the most common of the constitutive aphidicolin inducible fragile sites, *FRA3B*, which is cytogenetically indistinguishable from the t(3;8) translocation (Glover et al., 1988). Fragile sites are regions of the human genome that reveal cytogenetically detectable gaps when exposed to specific reagents or culture conditions (for review, see Sutherland, 1991); several folate sensitive, heritable, X-linked, and autosomal fragile sites have been localized to unstable CCG or CGG repeats (Yu et al., 1991; Kremer et al., 1991; Verkerk et al., 1991; Fu et al., 1991), and one of these, the *FRA11B* at 11q23.3, is associated with Jacobsen (11q-) syndrome, showing a direct link between a fragile site and in vivo chromosome breakage (Jones et al., 1994). Because the induced fragile sites resemble gaps or breaks in chromosomes, it has frequently been speculated that fragile sites could be sites of chromosomal rearrangement in cancer (Yunis and Soreng, 1984), perhaps by increasing the incidence of chromosome deletion during tumorigenesis.

The *FRA3B* region has been delineated by studies of several groups using rodent–human hybrids; hybrid cells retaining human chromosome 3 or 3 and X on a hamster background were treated with aphidicolin or 6-thioguanine (to select hybrids that had lost the X chromosome), and subclones were selected. Subclones retaining portions of chromosome 3 with apparent breaks in region 3p14-p21 were characterized for loss or retention of specific 3p markers to determine the position of 3p14–21 breaks (LaForgia et al., 1991, 1993; Paradee et al., 1995). Markers derived from yeast artificial chromosome (YAC) clones covering the t(3;8) break and surrounding regions were used to demonstrate that the 3p14.2 breaks representing the *FRA3B* were within a single YAC, *850A6*, which thus covered the *FRA3B*, the t(3;8) break, and the 5′ end of the *PTPRG* gene (Paradee et al., 1995; Kastury et al., 1996b). Using a 3p-linked probe isolated by representational difference analysis (RDA) (Lisitsyn et al., 1995), which detected homozygous deletions in colon cancer-derived cell lines, we have recently further mapped this RDA probe to 3p14.2 within the region defined as the *FRA3B* and have defined the homozygous deletions in a range of types of cancer–derived cell lines (Kastury et al., 1996a). We have now

† These authors contributed equally to this work.

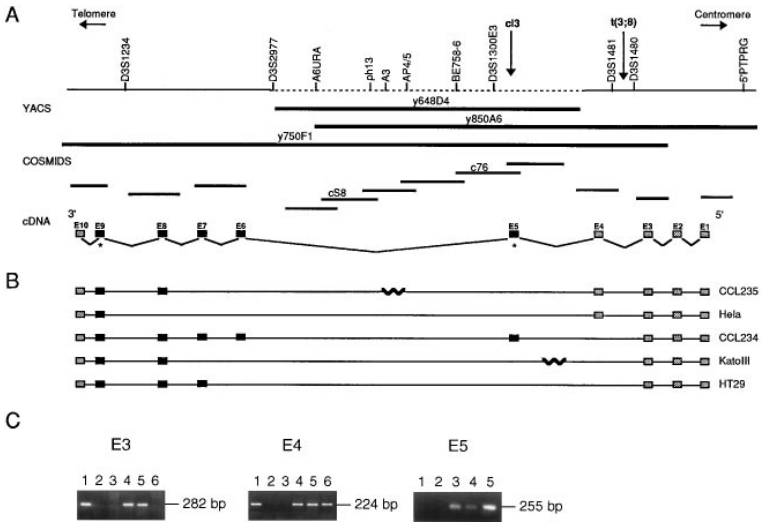

Figure 1. Organization of the *FHIT* Gene Relative to the 3p14.2 *FRA3B* and Translocation Sites

(A) A scheme of the normal 3p14.2 region is shown with the chromosomal region (not to scale) represented by the top line, with positions of STS markers (position of *D3S1234* relative to the gene is not known), the *FRA3B* represented by the hybrid cl3 break, and the t(3;8) translocation break point shown. The dashed portion represents the region involved in the homozygous deletions in tumor cell lines. Three of the YAC clones used in developing the above markers, map, and cosmid contig are shown with the cosmid contig below and the distribution of exons in the *FHIT* transcript shown below the contig. Closed and stippled boxes represent coding and noncoding exons, respectively, E1–E10; asterisks indicate exons with start and stop codons.

(B) Organization of types of aberrant transcripts from tumor cell lines are illustrated, with zigzag regions representing insertions of new sequence, usually repetitive, into the aberrant transcripts. With CCL235 RNA as template, apparently normal and aberrant RT–PCR products were amplified, with the aberrant product resulting from splicing of exon 4 to exon 8 with a repetitive insert of 140 bp (contributing an in-frame Methione codon) between E4 and E8. RT–PCR amplification of RNA from HeLa cells revealed normal and aberrant-sized products, the smallest product resulting from splicing of exon 4 to exon 9. RT–PCR amplification of RNA from KatoIII, a gastric carcinoma-derived cell line with discontinuous deletions leaving all *FHIT* exons intact, resulted in an aberrant-sized product that is missing exons 4 through 7, with an 86 bp repeat, inserted downstream of exon 3, contributing an in-frame Methionine codon. Amplification of the RT product from HT29, a colon carcinoma–derived cell line with a large deletion that included exon 5, gave only an aberrant-sized product resulting from splicing of exon 3 to exon 7. Numerous other tumor-derived cell lines from lung carcinoma (1 of 3 tested), osteosarcoma (1 of 2), NPC (3 of 3), ovarian carcinoma (2 of 2), and hematopoietic (4 of 5) tumors, exhibited aberrant *FHIT* transcription products (data not shown).

(C) Amplification of specific exons from YAC, hybrid, and tumor DNA templates shows that exon 3 (amplified using primers *iex3F*, 5′-AGGGTGATACTAGCTGCTTT-3′, *iex3R*, 5′-TGACTTTAGCCAGTGGCA) is present in YAC *850A6* (E3, lane 1), absent in hamster and mouse DNA (E3, lanes 2 and 3), present in human and hybrid cl3 DNA (E3, lanes 4 and 5), and absent in hybrid clone 5L.8 DNA (E3, lane 6). Since hybrid cl3 retains chromosome 3 region 3p14.2 to 3qter with the break in the fragile region as shown above (see Figure 1A) and is positive for E3, while hybrid 5L.8 retains 3p14.2 to 3pter, the portion of 3 translocated to chromosome 8 in the t(3;8) family, and is negative for E3, exon 3 must be centromeric to the t(3;8) break. Similarly, the E4 panel with identical DNA templates, *850A6* (lane 1), hamster, and mouse DNA (lanes 2 and 3), human (lane 4), cl3 (lane 5), and 5L.8 (lane 6) DNAs, from which exon 4 was amplified (using primers iex4F, 5′-TTGTACCTAGAGCCATCTGG-3′; *iex4R*, 5′-GGATACTCACAGCAGGTCAA-3′), shows that exon 4 is present in both hybrids cl3 and 5L.8, indicating that it must be telomeric to the t(3;8). Exon 4 is not present in the *648D4* cosmid contig (data not shown) and so must be between this cosmid contig and the t(3;8) break. Other exons and markers were similarly mapped. Exon 5 (amplified using primers X2, 5′-TGAGGACATGTCGTTCAGATTT-3′; G1, 5′-CTGGTGTCTCCGAAGTGGGAGGG-3′), was absent in tumor cell lines Hela and HT29 (E5, lanes 1 and 2) and present in tumor cell lines LS180, KatoIII, and YAC *850A6* (E5, lanes 4–6).

developed a cosmid contig covering the homozygous deletions and have characterized a gene that is a target of deletions in uncultured tumors of esophagus, stomach, colon, and probably other organs.

Results

The Cosmid Contig

From the 648D4 cosmid library, clones were selected initially using the *BE758-6*, *A6URA*, *A3*, and *1300E3* probes, which were distributed across the homozygously deleted region as shown in Figure 1A. Cosmid end clones were then isolated and used for the next round of cosmid screening. The cosmid map was assembled by polymerase chain reaction (PCR) amplification of the starting sequence-tagged sites (STSs), and new ones developed from cosmid ends, using cosmid DNA templates. Additionally, each new STS was tested against the YAC contig (also shown in Figure 1A) against cell lines with homozygous deletions and rodent–human hybrids retaining portions of chromosome 3 (LaForgia et al., 1993; Druck et al., 1995). Six cosmids were assembled into a contig that covered the homozygously deleted region.

To define more precisely the homozygously deleted region, which we will refer to as the fragile region, 42 STS markers, spanning the chromosomal region from the *PTPRG* locus to *D3S1234*, derived from cosmid walking and exon trapping, were tested by PCR amplification for presence in eleven cancer-derived cell lines that had been tested previously with a subset of markers (Lisitsyn et al., 1995; Kastury et al., 1996a).

Colon carcinoma–derived LoVo, HT29, SW480, and gastric carcinoma–derived AGS cell lines showed similar large deletions, as depicted by the dotted portion of the top line in Figure 1A. Colon carcinoma–derived LS180 and breast carcinoma–derived MDA–MB436 cells exhibited discontinuous deletions covering this same region, with most markers lost but some retained. The gastric KatoIII cells appeared to have lost the *D3S1481* marker and the telomeric portion of the fragile region, from *AP4/5* to *D3S2977* (see Figure 1A). The HKI cells, derived from a nasopharyngeal carcinoma (NPC), had lost the

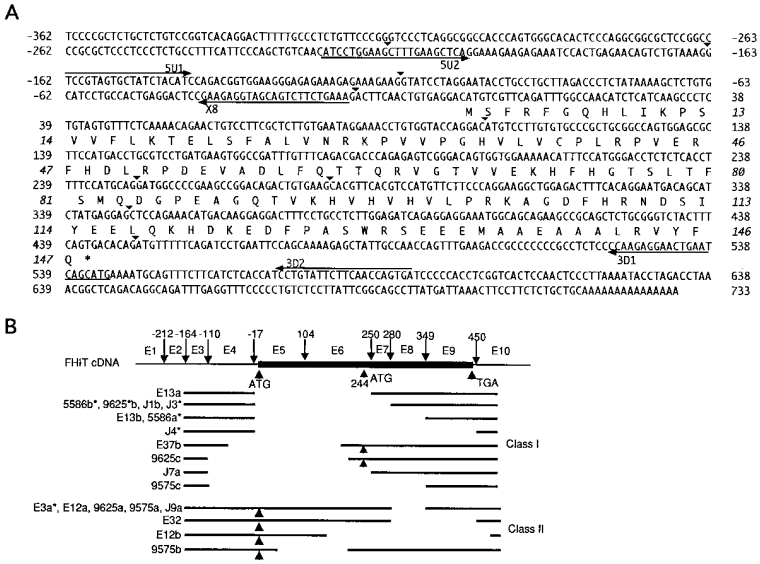

Figure 2. Structure of Normal and Aberrant *FHIT* cDNAs

(A) The nucleotide and predicted amino acid sequences of the *FHIT* gene are shown with positions of exons indicated by arrowheads above the sequence and positions of primers used in nested PCR and RACE reactions indicated by arrows below the sequence.

(B) A schematic presentation of some of the aberrant transcripts observed in uncultured tumor tissues of digestive organs is shown. Only transcripts that showed deletion of coding sequence in Table 2 are presented. The top line shows the intact *FHIT* cDNA map. The thick and thin bars show the coding and untranslated regions, respectively. The positions of splice sites are shown by downward arrows, according to the nucleotide numbers shown above in (A). The class I transcripts lack exon 5, while class II transcripts retain exon 5 but generally lose exon 8. In the transcripts with asterisks, insertions of various lengths were observed downstream of exon 4. E1–E10 indicate exons 1–10.

region between *D3S1481* and the *AP4/5* marker, while CNE2, another NPC-derived cell line, had discontinuous deletions that included a region near the t(3;8) and the region between *D3S1481* and *D3S2977*. HeLa cells also exhibited discontinuous deletions, with one deleted region near the t(3;8) and one between *D3S1481* and *AP4/5*. The NPC-derived CNE1 cells were tested with most markers without detection of a deletion. Thus, there are many different tumor-associated 3p14.2 chromosome breakpoints surrounding the t(3;8), the *FRA3B* locus, and the homozygously deleted region covered by the cosmid contig.

Isolation of cDNAs

The six cosmids covering the homozygous deletion, shown in Figure 1A, were used in exon-trapping experiments aimed at identifying genes within the deleted region. Sequences of trapped exons were compared against nucleotide sequence databases, and one exon, trapped from a cosmid 76 subclone (c76, Figure 1A), matched a number of cDNA sequences from breast (GenBank accession numbers R53187 and R86313) and fetal liver and spleen (R11128) libraries submitted by the Washington University–Merck expressed sequence tag (EST) Project. A 23 bp oligonucleotide primer designed from this sequence (primer *X8*, Figure 2A) was used in primer extension to obtain a 5′ extended product of the cDNA by a rapid amplification of cDNA ends (RACE) reaction. The longest product (370 bp) from the RACE reaction detected a ubiquitously expressed 1.1 kb mRNA by Northern blot analysis of mRNAs from various normal tissues. The size was similar to the length of the largest cDNA clone isolated from a normal colon cDNA library, using the same DNA fragment as a probe. The DNA sequence analysis of this full-length clone (Figure 2A) revealed a 5′ untranslated region of more than 350 bp, followed by an initial methionine codon with surrounding sequence fitting Kozak's rule, an open reading frame (ORF) of 147 amino acids, a 3′ untranslated region, a polyadenylation consensus sequence, and a poly(A) tail.

This cDNA was translated and ORF amino acid sequences (Figure 2A) compared with the protein databases. The full-length cDNA probe was then hybridized to Northern blots carrying mRNA from a spectrum of tissues. As shown in Figure 3A, the cDNA detected the ubiquitously expressed 1.1 kb transcript.

Relationship of the cDNA to the Genomic Map of the Region

Oligonucleotide primers from the initially trapped exon were used to generate intron sequences from cosmid 76; these sequences were used in turn to prepare primers and probes to map the exon (E5 in Figure 1A) on the cosmids, YACs, and DNA from cancer cell lines with deletions, as illustrated in Figures 1A and 1C. Using cDNA as template, oligonucleotide primer pairs bracketing the exons upstream and downstream of exon 5 were then used to amplify cDNA fragments to serve as probes for mapping the 5′ and 3′ flanking exons on the cosmid contig; these probes demonstrated that the cDNA sequences 5′ and 3′ of exon 5 were not within the *648D4* cosmid contig covering the homozygous deletions. Thus, cosmid libraries from YACs *850A6* and *750F1*, which extend centromeric and telomeric to the fragile region deletions, respectively, as shown in Figure 1A, were prepared and screened with the 5′ and 3′ cDNA probes flanking exon 5. Cosmids containing the remaining exons were then used to derive intron sequences using cDNA primers, and the structure of the gene was determined as shown in Figure 1A. The cDNA consisted of ten exons that were distributed among three YAC clones (Figure 1A); exons 1 through 4 mapped to YAC clone *850A6*, exon 5 was present in all three YAC clones, and exons 6 through 10 mapped to YAC clone *750F1*. Only exon 5 fell within the region of homozygous deletion in tumor-derived cell lines, i.e., within YAC clone *648D4*, as illustrated for some cell lines in

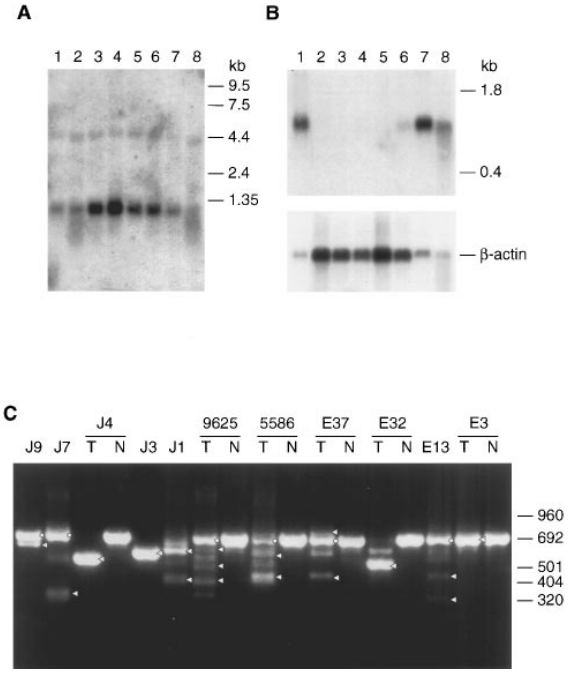

Figure 3. Expression of the *FHIT* Gene in Normal Tissues and Tumors

Northern blot (A and B) and RT–PCR analysis (C) of normal and tumor-derived *FHIT* mRNA.

(A) Normal mRNAs (2 μmg/lane) from spleen (lane 1), thymus (lane 2), prostate (lane 3), testis (lane 4), ovary (lane 5), small intestine (lane 6), colon (mucosal lining) (lane 7), and peripheral blood leukocytes (lane 8), hybridized with the *FHIT* cDNA probe.

(B) mRNAs (2 μg/lane) from normal small intestine (lane 1) and mRNAs from tumor-derived cell lines: KatoIII (lane 2), HK1 (lane 3), LoVo (lane 4), CNE2 (lane 5), CNE1 (lane 6), Colo320 (lane 7), LS180 (lane 8), hybridized with the *FHIT* cDNA probe (top). The same blot was hybridized with a β-actin cDNA probe (bottom).

(C) Amplified products observed after nested RT–PCR amplification of mRNAs from matched uncultured tumor (T) and normal (N) tissues of the same patients (J4, 9625, 5586, E37, E32, and E3), or mRNAs from tumor tissues only. Arrowheads show the positions of amplified products with abnormal DNA sequence. The details of the DNA sequences of corresponding transcripts are shown in Table 2 and Figure 2B. White dots in the tumor lanes show the position of transcripts with normal DNA sequence.

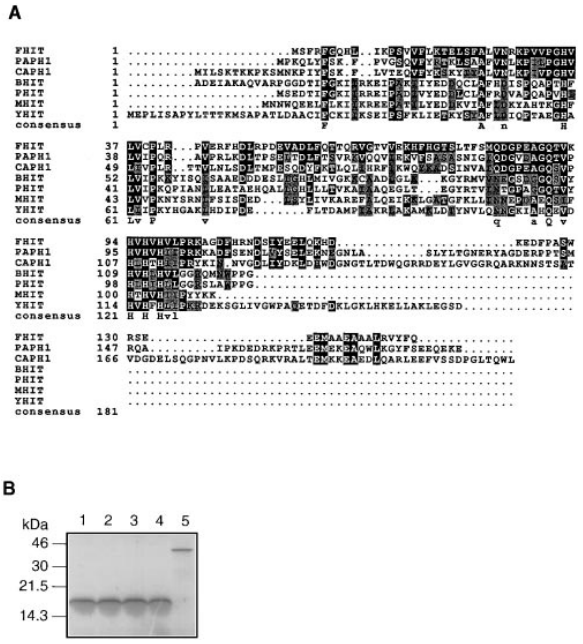

Figure 4. Alignment of Amino Acid Sequences of HIT Family Proteins and Translation of *FHIT* cDNAs

Alignment was performed using BOXSHADE version 3.0. White type on a black background indicates two or more identical residues at a position; stippling indicates similarity. PAPH1 (accession number U32615) designates the S. pombe *aph1* gene, and CAPH1 (accession number U28374), designates a potential S. cerevisiae *aph1*. PHIT indicates the HIT family member from the cyanobacterian Synechococcus Sp. (accession number P32084), BHIT, indicates the protein kinase C inhibitor from Bovine taurus (accession number P16436)), MHIT from Mycoplasma hyorhinis (accession number M37339), YHIT from S. cerevisiae (accession number Q04344); the FHIT protein is 69% similar to the S. pombe (*PAPH1*) gene over a length of 109 amino acids.

(B) In vitro translation products from recombinant plasmids containing different alleles of the *FHIT* gene: pFHIT1 with a deletion of noncoding exon 4 (lane 1); pFHIT2 with an insertion of 72 bp between exons 4 and 5 (lane 2); pFHIT3 with a wild-type *FHIT* lacking exon 1 (lane 3); the pFHIT full-length wild-type gene in Bluescript (lane 4); control reaction, in vitro translation from the pBCAH vector, carrying a portion of the extracellular region of the *PTPRG* gene (predicted molecular weight of 40 kDa) (lane 5).

Figure 1C. The coding region of the ORF began in exon 5 and ended in exon 9, as shown in detail in Figures 2A and 2B.

Most interestingly, the first three exons (E1, E2, and E3) of the gene mapped centromeric to the t(3;8) break, between the t(3;8) break, and the 5′ end of the *PTPRG* gene, as determined by amplification of these exons from the YAC DNAs and DNAs derived from hybrids carrying portions of chromosome 3, derived from the t(3;8) break and a *FRA3B* break, as illustrated in Figure 1C. Thus, this gene became a strong candidate for involvement in initiation of the familial RCCs, because one copy of the gene is disrupted by the translocation.

The homology search in amino acid sequence databases showed a significant homology to a group of proteins that have a histidine triad motif, designated HIT proteins (Seraphin, 1992). The predicted amino acid sequence of the cDNA for the human gene, designated the Fragile Histidine Triad gene or the *FHIT* gene, is shown in Figure 4A compared to the other members of the *HIT* family. The highest homology of the FHIT protein (∼50% identity) is to the yeast Schizosaccaromyces pombe diadenosine 5′,5′′′-P^1, P^4 (Ap_4A), asymmetrical hydrolase gene (*aph1*), shown in Figure 4A as PAPH1 (Huang et al., 1995) and a potential *aphI* gene, CAPH1, from Saccaromyces cerevisiae. The consensus sequence for the *HIT* family proteins is shown below the amino acid sequences in Figure 4A.

To recapitulate, the *FHIT* gene, which may be the human cognate gene for the S. pombe Ap_4A hydrolase gene, spans a >500 kb region that includes the t(3;8), the *FRA3B*, and a tumor cell–specific commonly deleted region.

Expression of the *FHIT* Gene

Our previous study placed the *BE758–6* locus, first identified by RDA (Lisitsyn et al., 1995), and the microsatellite

marker, *D3S1300*, within the region of common loss in a variety of tumor-derived cell lines, and our loss of heterozygosity (LOH) study of gastric and colon tumors detected a high frequency of allelic deletion, often involving *D3S1300* (Kastury et al., 1996a), in the region between the t(3;8) and the *D3S1234* locus (see Figure 1A). Now, the localization of both the *BE758–6* and *D3S1300* loci within the *FHIT* gene locus, close to the first coding exon, exon 5, suggested that the *FHIT* gene was the target of deletion in uncultured tumors, as well as tumor-derived cell lines. To begin an analysis of *FHIT* transcripts in tumor-derived cells, poly(A)$^+$ mRNAs from tumor-derived cell lines and normal tissues was studied by Northern blot analysis.

A low level of expression of the *FHIT* gene occurred in all human tissues tested, (as shown in Figure 3A), for spleen, thymus, prostate, testis, ovary, small intestine, colon, and peripheral blood lymphocytes. The major transcript was ~1.1 kb with a longer transcript at ~5 kb, which was barely detectable on some blots; this longer transcript may represent an incompletely processed precursor RNA. Similar transcripts were seen in mRNA from brain, heart, lung, liver, skeletal muscle, kidney, and pancreas.

mRNAs from tumor-derived cell lines with known homozygous deletions in the fragile region exhibited varying levels of *FHIT* transcripts, from barely detectable (KatoIII, HKI, and LoVo mRNA; Figure 3B, lanes 2–4, respectively) to an almost normal level (LS180, lane 8), relative to normal small intestine mRNA (lane 1).

Note that the NPC cell lines with (CNE2 and HK1; Figure 3B, lanes 5 and 3) and without (CNE1; Figure 3B, lane 6) documented homozygous deletions (Kastury et al., 1996a) expressed barely detectable *FHIT* mRNA. The NPC-derived cell line CNE2 exhibited a possible smaller transcript (Figure 3B, lane 5), while Colo320, a colorectal carcinoma–derived cell line without a deletion, exhibited an apparently normal-sized *FHIT* transcript (Figure 3B, lane 7), although it should be noted that size alone does not imply presence of a wild-type transcript. One conclusion of the Northern blot analysis is that there was no direct relationship between size or abundance of transcript and detection of homozygous deletions in specific tumor-derived cell lines, suggesting that there may be small deletions in some tumor cell lines that have not been detected with the available markers.

Reverse Transcription–Polymerase Chain Reaction and cDNA Sequence Analysis of Tumor-Derived mRNA

To look for abnormalities in *FHIT* transcripts from deleted and nondeleted tumor cell lines, we reverse-transcribed mRNAs with $(dT)_{17}$ primer, amplified the cDNA with 5′ and 3′ primers, and then reamplified using primers inside the original primers (nested PCR), as described in Experimental Procedures. Positions of the primers are shown in Figure 2A. The amplified products were separated on agarose gels, and fragments were cut from the gels and sequenced (examples of aberrant bands are shown for mRNAs of uncultured tumors in Figure 3C; reverse transcription–polymerase chain reaction (RT–PCR) products from tumor cell lines were very similar). The tumor-derived cell lines exhibited a pattern of products ranging from one apparently normal-sized amplified transcript to numerous aberrant bands without a normal-sized band. The sequencing of the aberrant bands revealed numerous abnormal products, some of which are illustrated in Figure 1B. Colon tumor-derived CCL235 and CCL234 cell lines did not show deletion of the STS markers tested, but both showed aberrant transcripts, as illustrated, with CCL235 exhibiting a normal-sized product in addition. HT29 and KatoIII cell lines both showed homozygous deletion (Kastury et al., 1996a), but the KatoIII cell line exhibited a deletion of the telomeric portion of the homozygously deleted region and not the region containing exons 4 and 5, nor the region of exon 6, exons that are all missing in the aberrant RT–PCR product, as illustrated in Figure 1B.

Ten cases of uncultured esophagus, nine of stomach, and eight of colon tumors were also analyzed, and aberrant transcripts were observed in five, five, and three cases, respectively (summarized in Tables 1 and 2 and illustrated in Figures 2B and 3C). The sequence analyses of the aberrant cDNAs revealed absence of various regions between exons 4 and 9 (Table 2; Figure 2B), while the RT–PCR and cDNA sequence analyses of normal tissue mRNAs from the same organs did not exhibit alterations of the coding region sequence (E3, E12, E32, E37, J1, J4, J9, 9625, 5586, and 9575). In 8 of 13 cases with aberrant transcripts, normal-sized transcripts were also observed (E3, E12, E13, E37, 9625, 9575, J7, and J9, Figure 3C; E12 and 9575, data not shown), while in 5 of 13 cases normal-sized transcripts were not detected (J3 and J4; Figure 3C), or were barely detected (E32, 5586, and J1; Figure 3C). In most of the aberrant transcripts, the beginning and the end of the deleted portions of the transcripts coincided with splice sites (Figure 2B), suggesting that the cDNA deletions resulted from the loss of genomic regions containing or surrounding the relevant lost *FHIT* exons. The aberrant transcripts can be classified into two groups (class I and II, Figure 2B): class I transcripts lack exon 5, which has the initial methionine codon of the *FHIT* ORF, resulting in the loss of the intact ORF; class II transcripts have an intact initial methionine codon but do not include exon 8, except for 9575b, which exhibited a frameshift after exon 6. Thus, in all the class II transcripts, the wild-type ORF of exon 8, the histidine triad-containing domain, is not present. Moreover, some of the class II transcripts exhibited loss only of exon 8 (E3a, E12a, 9625a, 9575a, and J9a; Figure 2B), suggesting that exon 8 was the target of deletion. Since exon 8 encodes the histidine triad motif, it is likely that neither class I nor class II transcripts, constituting the major fraction of aberrant transcripts, can encode a fully functional protein. However, there is an in-frame methionine codon in exon 6 (see Figure 2B), and in some cases, insertions contribute an in-frame Methione (data not shown); thus, the majority of aberrant transcripts could encode partial proteins, with or without the HIT domain, as indicated in Table 2. Insertions of various lengths of DNA not derived from the *FHIT* gene were observed in some transcripts, but only downstream of exon 4 (5586a, 5586b, 9625, J3, and J4; Table 2). A minor group of aberrant transcripts retained intact full-length ORFs, but

Table 1. Derivation of FHIT RT–PCR-Amplified Products and cDNA Sequences from Uncultured Tumors of Gastric Organs

Origin of Tumors	Number of Cases Analyzed	Cases with Aberrant Transcripts	
		Number of Cases	Codes[1]
Esophagus	10	5	E3*, E12*, E13 E32*, E37*
Stomach	9	5	J1*, J3, J4*, J7, J9*
Colon	8	3	9625*, 5586*, 9575*

[1] In cases with asterisks, normal tissues from the same organs were analyzed and did not exhibit alterations in the coding region sequences.

were missing exon 4 (J1a; Table 2), or had insertion of 72 bp of DNA sequence in the 5′ noncoding region between exon 4 and 5 (E37a, Table 2; Figure 1B).

To determine whether the wild-type *FHIT* cDNA and various cDNAs derived from tumor-specific transcripts, which retained the entire coding region, could be translated in vitro, several recombinant plasmids were constructed, each containing a *FHIT* gene downstream from the T7 promoter and lacking the first noncoding exon. The pFHIT1 plasmid carried a cDNA missing exon 4 from the CCL234 colon cancer cell line. Plasmid pFHIT2 carried a cDNA from esophageal tumor E37 with an insertion of 72 bp between exons 4 and 5. The pFHIT3 plasmid contained the wild-type *FHIT* gene. The constructs were used for in vitro translation and analysis of translation products (Figure 4B), which showed the predicted 16.8 kDa protein translated from each cDNA construct.

The FHIT Protein

The protein sequence predicted by the *FHIT* cDNA is very similar (57 of 109 amino acid identities; 76 of 109 or 69%, similarities, calculated by the National Center for Biotechnology Information (NCBI) BLAST server) to the S. pombe *aph1* gene (Huang et al., 1995), as shown in the amino acid alignment in Figure 4A, where PAPH1 represents the S. pombe sequence.

The S. pombe *aph1* enzyme was cloned by purification of the enzyme, amino acid sequencing of the N-terminus, and design of primers to amplify a partial cDNA; the full-length genomic region and a cDNA of 1.2 kb were then cloned, sequenced, and translated (Huang et al., 1995). By similar methods, a human Ap_4A hydrolase (*APH1*) has been cloned, sequenced, and translated (Thorne et al., 1995) and, surprisingly, does not resemble the S. pombe *aph1* gene nor the *FHIT* gene. Since higher eukaryotes appear to possess a single 16–21 kDa Ap_4A

Table 2. Aberrant Transcripts Observed in Uncultured Tumors

Tumor-Derived Transcript[1]		Deletion (position[2])	Insertion[3]		Effect	Putative Protein[4] Coded in Frame
			Size (bp)	Homology		
*E3	a	280–348	72	NS	Ex 8 loss	HIT(−)
*E12	a	280–348	—	—	Ex 8 loss	HIT(−)
	b	122–516	—	—	FS after Ex 6	HIT(−)
*E13	a	−17 to 249	—	—	Ex 5 and 6 loss	—
	b	−17 to 348	—	—	Ex 5–8 loss	—
E32	—	280–449	—	—	Ex 8 and 9 loss	HIT(−)
*E37	a	—	72	NS	None	intact
	b	−73 to 173	—	—	Ex 5 loss	HIT(+)
*9625	a	280–348	—	—	Ex 8 loss	HIT(−)
	b	−17 to 279	87	Alu	Ex 5–7 loss	—
	c	−110 to 204	—	—	Ex 4 and 5 loss	HIT(+)
*5586	a	−17 to 349	135	Alu	Ex 5–8 loss	—
	b	−17 to 279	37	NS	Ex 5–7 loss	—
*9575	a	280–348	—	—	Ex 8 loss	HIT(−)
	b	60–181	—	—	FS after Ex 5	HIT(−)
	c	−110 to 348	—	—	Ex 4–8 loss	—
J1	a	−110 to −17	—	—	None	Intact
	b	−17 to 279	—	—	Ex 5–7 loss	—
J3	—	−17 to 279	173	Alu	Ex 5–7 loss	HIT(+)
J4	—	−17 to 457	305	Alu	Ex 5–9 loss	—
*J7	a	−110 to 249	—	—	Ex 4–6 loss	—
*J9	a	280–348	—	—	Ex 8 loss	HIT(−)

All the aberrant transcripts that involve the coding sequence of the *FHIT* gene are shown in Figure 3B.
[1] In tumors with asterisks, normal transcripts without alteration of coding region sequence were also observed.
[2] The positions of the first and last nucleotides of the deletions are shown according to the nucleotide numbers in Figure 2.
[3] The position of all insertions was downstream of exon 4.
[4] Putative protein coded in frame with the FHIT protein is shown: HIT(+), protein with HIT motif; HIT(−), protein without HIT motif; −, no protein in frame.
Abbreviations: Alu, Alu repeat; FS, frameshift; NS, no significant homology; Ex, exon.

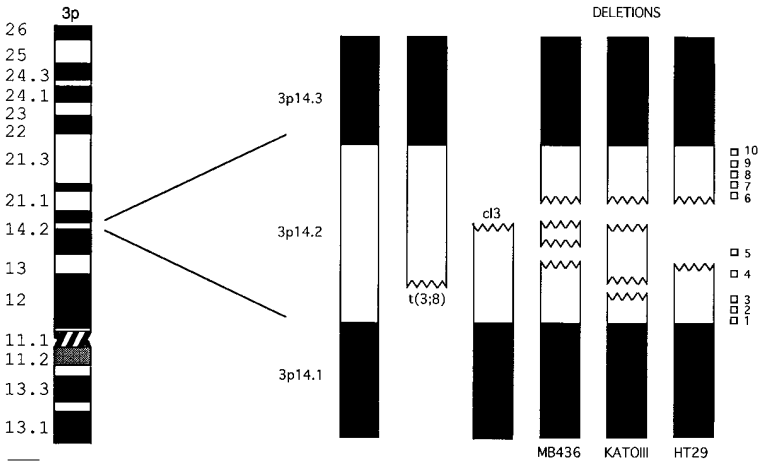

Figure 5. Organization of the *FHIT* Gene Relative to Documented Chromosome Breaks in the 3p14.2 Fragile Region

One *FHIT* allele is disrupted in all the translocation carriers of the t(3;8) family, with exons 1, 2, and 3 remaining on the derivative 3 chromosome and exons 4–10, including the entire coding region, being translocated to the derivative 8 chromosome, as illustrated above. The hybrid cell line cl3 with a de novo *FRA3B* break just telomeric to exon 5 has lost most of the *FHIT* coding region. The KatoIII cells apparently retain all *FHIT* exons, but encode only an abnormal transcript that lacks exons 4–7 and thus cannot produce FHIT protein. The MB436 and HT29 cells have both lost exon 5 through deletion of different segments of the fragile region.

asymmetrical pyrophosphohydrolase (cited in Thorne et al., 1995), it is thus not clear if the *FHIT* gene is a human *APH1* enzyme, although it may be a human cognate of the S. pombe *aph1* enzyme.

The *FHIT* gene is also very similar to a S. cerevisiae gene (CAPH1 in Figure 4A) that may be a cerevisiae *aph1* gene, based upon the homology to PAPH1 shown in Figure 4A. The other proteins in Figure 4A are members of this *HIT* gene family, a family of proteins present in prokaryotes, yeast, and mammals, described by Seraphin (1992). The signature feature of the family is the histidine triad (Figure 4A; amino acids 94–98 of the FHIT protein), which for the case of *BHIT* (Figure 4A), the bovine inhibitor of protein kinase C (*PKCI1*), has been shown to be a zinc-binding site (Pearson et al., 1990; Mozier et al., 1991). The FHIT protein product is only 39% similar to the bovine PKCI1 protein over FHIT amino acids 12–100, as calculated by NCBI BLAST. Thus, the *FHIT* gene is not likely to be the human *PKCI1* gene. Functions of the other *HIT* genes are not known.

Discussion

The meaning of fragile sites for cancer has been a subject of speculation for years, and the near coincidence of the chromosomal position of the *FRA3B* and the t(3;8) translocation at 3p14.2 has been especially intriguing. The *FRA3B* is constitutive; that is, after treatment of peripheral blood lymphocytes with $\sim$0.4 μM aphidicolin, which interferes with the action of DNA polymerase α, the characteristic gaps in chromosome region 3p14.2 are observed in $\sim$70% of metaphases from all individuals. So the structural basis for the induction of gaps is present in all individuals. It is also known that within the 3p14.2 band, some of the induced gaps represent chromosome breaks, which occur possibly at several sites in the chromatin of an $\sim$200–300 kb region (Paradee et al., 1995). Thus, the sequences involved in gaps and breaks may occur in more than one site within the fragile region. Perhaps the *FRA3B* appears to be the most common fragile site because it actually represents a collection of different fragile sites in a small chromosomal region. The specific sequences responsible for the breaks at *FRA3B* in hybrid cells have not been described, but we have observed that many tumor-derived cell lines exhibit apparent discontinuous homozygous deletions. Figure 5 diagrams the relationship between the various types of chromosome breaks in 3p14.2 and the organization of the *FHIT* gene relative to the breaks. Note that in Figure 5, the chromosome breaks and deletions in the KatoIII gastric carcinoma-derived cells leave the coding region intact, but we have observed only an aberrant *FHIT* transcript in this cell line. Thus, inapparent chromosomal abnormalities must account for the lack of normal transcription in KatoIII and other tumor cells. Some cancer-derived cell lines and uncultured tumors showed transcripts with alterations to noncoding regions of the *FHIT* transcript. These transcripts were translated into full length protein in vitro (Figure 4B), but perhaps in the tumor cells from which they were derived, the lack of exon 4, or insertion of new sequences, would affect expression of the FHIT protein. Another puzzle, if the *FHIT* gene acts as a classical suppressor gene with inactivation of both alleles, is the presence of normal-sized transcripts along with aberrant products in the RT–PCR amplification products of tumor-derived cell lines such as CCL235 (colon), A549 (lung), and HeLa (cervical). It is possible that the aberrant transcripts, which in most cases might encode partial FHIT proteins, could interfere with the function of a normal FHIT protein. A number of the uncultured tumors also exhibited aberrant and normal-sized products, and sequencing showed that some of these normal-sized products were indeed wild-type products. In these cases, the normal transcripts could have derived from admixed normal cells. We have not yet observed point mutations within the coding region of any *FHIT* transcripts, perhaps suggesting that aberrant *FHIT* genes usually are the result of deletions.

Aphidicolin, which inhibits the action of DNA polymerase α, induces the gaps and breaks observed in the *FRA3B* region in normal metaphases; thus, in the digestive tract tumors and tumor cell lines we have studied, the genomic deletions resulting in aberrant transcription and loss of functional FHIT protein could have been induced by exposure of these organs to other agents that interfere with DNA replication, such as nicotine, caffeine, alcohol, or other known carcinogens. Interestingly, zinc deficiency is associated with a high frequency of esophageal tumors in human (Yang, 1980) and rat

(Fong et al., 1978); zinc deficiency may cause proliferation of the epithelial cells lining the esophagus (Yang et al., 1987), so perhaps zinc deficiency mimics loss of the FHIT protein, which may require bound zinc for its function. It is, therefore, interesting that *FHIT* gene exon 8, carrying the HIT motif, the presumptive zinc-binding site, is a target of deletion in numerous digestive tract tumors.

Whether this region of 3p14.2 contains repeated CCG or CGG triplets is not yet known, but because there are differences between the rare, inherited folate-sensitive fragile sites that have been characterized and the common, constitutive, aphidicolin fragile sites, perhaps a different basis for the fragility should be expected. Thus far, we have noted that there are many *Alu* repeats in the telomeric portion of the fragile region, and there is a $(TAA)_{15}$ repeat in this same commonly deleted region, for which the number of repeats is highly variable (Kastury et al., 1996a). Also in ~9 kb pairs of sequenced portions of the cosmid S8 (telomeric portion of the fragile region, see Figure 1A), several *Alu* repeats and a long interspersed repetitive element (LINE) were encountered; the nucleotide content of the sequenced region was 57.4% A and T residues, while the *FHIT* cDNA nucleotide content was 48% A and T. A high A and T content is characteristic of some origins of DNA replication, especially in yeast, and in fact it has been speculated that *Alu* repeats may be connected with replication in higher eukaryotes. Another notable feature of the *FHIT* gene is that nearly all the exons end with the sequence AG, the usual sequence for splice acceptor sites. Based on our observation of frequent aberrant splicing in this fragile region, it is tempting to speculate that the region is especially rich in sequences resembling splice acceptor sites.

Studies of *FHIT* gene RT–PCR products from RNA of numerous cell lines suggested that *FHIT* gene abnormalities could be important not only in airway and digestive tract tumors such as nasopharyngeal, esophageal, stomach, and colorectal carcinomas, but possibly also in ovarian, cervical, and lung tumors, osteosarcoma, and some leukemias; also a bladder and breast carcinoma cell line exhibited homozygous deletions in the fragile region (Lisitsyn et al., 1995; Kastury et al., 1996a). Thus, uncultured tumors of these types should be tested for *FHIT* gene abnormalities.

Clear cell RCCs might also be expected to involve *FHIT* gene aberrations because the *FHIT* gene is disrupted by the familial RCC translocation break in 3p14.2 and the translocation/*FRA3B* region is the target of allelic loss in most sporadic clear cell RCCs (Druck et al., 1995). Since the *FHIT* ORF is contained in exons 5 through 9, translocated to chromosome 8 in the t(3;8) family, it is possible that both alleles could still be expressed in some or all tissues. If the *FHIT* gene disruption is the first "hit" to a suppressor gene in this family, then the second *FHIT* allele should be altered in the t(3;8) tumors. Since we have not yet detected point mutations in the *FHIT* gene, the best way to look for alterations of the *FHIT* gene in t(3;8) RCCs would be to amplify the *FHIT* reverse transcript, as done for uncultured tumors in this study. We have done this experiment for RNA from two RCC cell lines and two uncultured RCCs, all from sporadic tumors, and have observed normal-sized products, which have not yet been cloned and sequenced. Nor have we yet observed homozygous deletions in RCCs using a subset of STS markers in the fragile region. Nevertheless, it would be surprising if the *FHIT* gene is not involved in some sporadic RCCs.

Since the *FHIT* gene is probably ubiquitously expressed, it may not be surprising if it can serve as a tumor suppressor gene for specific tissues of many different organs, perhaps predominantly of the digestive tract, or maybe predominantly organs with epithelial cell linings. Another common denominator of the types of tumor exhibiting aberrant *FHIT* alleles might be that they are predominantly organs directly exposed to environmental carcinogens; some of the types of tumors exhibiting *FHIT* gene aberrations occur very frequently in restricted regions of the globe, NPC in China, gastric cancer in southeast Asia, and often there are environmental factors at play. A possible role for Epstein-Barr virus (EBV) in promotion of Chinese NPCs might be through viral DNA integration into the *FRA3B* region, suggested by the previous experiments of Rassool et al. (1992), showing apparent preferential integration of exogenous DNA into induced fragile sites in cultured cells. Similarly, human papillomaviruses associated with cervical carcinomas might promote induction of the *FRA3B*, contributing to the loss of heterozygosity on 3p in cervical cancers (Yokota et al., 1989) and possibly to inactivation of the *FHIT* gene.

The observation of strong similarity of the *FHIT* gene to S. pombe Ap_4A hydrolase does not supply obvious clues to the in vivo function of the FHIT protein, nor to the mechanism of its function as a candidate tumor suppressor gene. Specific roles for the diadenosine Ap_4A have not been defined (Huang et al., 1995), and it is not clear that the Ap_4A hydrolase activity is the only, or even the major, in vivo function of these proteins. Very little is known of the function of the other members of the *HIT* family. If indeed the *FHIT* gene is the cognate of the S. pombe *aph1* gene, then the strong conservation (69% similarity) between the yeast and human genes suggests important functions. Whether the *FHIT* gene does or does not encode an Ap_4A hydrolase, it is likely that the study of S. pombe *aph1* knockouts and other S. pombe *aph1* mutations will be useful in understanding the functions of the FHIT protein.

There is some suggestion that as an intracellular regulatory molecule, Ap_4A may regulate ability of cells to adapt to metabolic stress such as heat, oxidation, and DNA damage; thus, deviation from a normal level of Ap_4A may result in the inability of cells to adapt to environmental stresses imposed by carcinogens or viruses that cause genetic damage.

Experimental Procedures

Tissues and Cell Lines

Matched normal and cancerous tissues from patients with primary esophageal, colon, and stomach carcinomas were obtained immediately after surgery, and tumors were dissected to eliminate normal tissue before preparation of DNA. Many cell lines were obtained from the ATCC. The RC kidney cell lines were provided by E. Lattime.

RNA Extraction and Reverse Transcription

Total and poly(A)$^+$ mRNA was extracted from cell lines and tissues using the RNAzol kit (TelTest, Incorporated) or the FastTrack Kit

(Invitrogen), respectively. To obtain mRNA from tissues, fresh specimens were frozen immediately after excision and stored at −85°C or in liquid nitrogen until extraction of mRNA. RNA was stored as a pellet under ethanol or solubilized in RNase-free water and kept at −70°C. Reverse transcription was performed in 30 μl final volume of 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 2 μM dNTPs, 500 ng oligo(dT), 600 units of MMLV-RT (Gibco-BRL), 40 Units RNasin (Promega), and 2 μg of RNA. This reaction was incubated at 37°C for 90 min and boiled for 5 min.

DNA Sequence Analysis

cDNA, genomic clones, and putative exons were sequenced using primers specific for vector-flanking sequences and various synthetic oligonucleotides. RT–PCR products were directly sequenced after isolation of bands from low melt agarose and purification by column chromatography (QIAGEN). Sequencing of double-stranded plasmids, PCR products, and phage or cosmid genomic clones was performed using Taq DyeDeoxy Terminator Cycle Sequencing Kits, Applied Biosystems, Incorporated; reaction products were electrophoresed and recorded on the 373 or 377 DNA sequencer, Applied Biosystems, Incorporated. Sequences were analyzed using GCG, BLAST, and GRAIL software.

PCR Amplification

The oligonucleotides for generating probes, PCR products, and RT–PCR products were designed using the computer program Oligo 4.0 (National Biosciences). Hybridization probes were produced by PCR amplication using various *FHIT*-specific primers, as indicated in Results. Sequences and positions of some primers are shown in Figure 2A.

PCR reactions were carried out in 12.5 or 25 μl final volume with 1–100 ng of template, 20–40 ng of primers, 10 mM tris–HCl (pH 8.3), 50 mM KCl, 0.1 mg/ml gelatin, 15 mM $MgCl_2$, 200–600 μM dNTPs, and 0.5–2.5 Units of Taq polymerase, Applied Biosystems, Incorporated. The amplifications were performed in a Perkin–Elmer Cetus thermal cycler for 30 cycles of 94°C for 30 s (for denaturation), 60°C (varied for specific primer pairs) for 30 s (for annealing), and extending at 72°C for 30–45s. The PCR products were visualized in ethidium bromide–stained low melting agarose gels.

Identification of YACs

We and others (Boldog et al., 1993, 1994; Wilke et al., 1994; Michaelis et al., 1995; Kastury et al., 1996b) had previously identified the *850A6* clone from the Genethon mega YAC library as containing the *D3S1300* and *D3S1481* markers (Roche et al., 1994). Overlapping YACs were identified by analysis of the GenBank database information.

Identification of Region-Specific STSs

The *A6URA* marker was from the 850A6 URA end, and A3 from an *Alu*vectorette–amplified fragment of 850A6 (Kastury et al., 1996b); *BE758-6*, and *D3S1480*-amplified fragments were used as probes to select phage genomic clones from which end sequences were obtained and sequence tagged. A phage genomic clone for *D3S1300* was selected from the 850A6 (Kastury et al., 1996b) phage library, and end clone *D3S1300E3* was isolated. Other *D3S* and *WI* marker primer pairs were obtained from Research Genetics or were synthesized from sequences provided in the WI database. The $(TAA)_{15}$ trinucleotide repeat, designated locus *ph13*, and the *AP4/5* STS were previously described (Kastury et al., 1996a).

Cosmid Mapping

High molecular weight YAC containing yeast DNA in agarose plugs was partially restricted with the Sau3AI enzyme and subcloned into a cosmid vector as described (Ohta et al., 1995). This cosmid library was initially screened with DNA probes derived from STSs previously mapped to this region. The ends of the insert DNAs flanking the cosmid vector were sequenced to find new STSs, which were used as probes to rescreen the cosmid libraries.

Exon Trapping and cDNA Cloning

The cosmid DNAs were partially restricted with Sau3AI enzyme and run on a 1.0% agarose gel, and fragments larger than 2 kb were cut out and subcloned into the pSPL vector and transfected into COS7 cells, according to the instructions of the manufacturer. The DNA inserts trapped between the splice sites of the vector were sequenced by a primer supplied with the vector (GIBCO/BRL). The cDNA was extended in the 5′ direction by PCR amplification of a total human fetal brain cDNA using an exon-specific primer, *X8*, and a RACE reaction kit (Clontech). The normal colon cDNA library was purchased from Clontech.

***FHIT* Exon Mapping**

The genomic sequences of exon-intron junctions of the *FHIT* gene were determined by sequencing the positive cosmids with primers derived from the cDNA. Localization of exons of the *FHIT* gene was determined by PCR amplification using primers derived from each exon, with YAC and chromosome 3 hybrid DNAs as templates. The primer sequences used to obtain cDNA probes flanking exon 5 were: 5′-TCTGCTCTGTCCGGTCACA-3′ (nucleotides −355 to −337) with primer X8 (shown in Figure 2A) for 5′ flanking exons; 5′-ATGTCCTTG TGTGCCCGCT-3′ (nucleotides 105 to 123) with 3D2 (see Figure 2A) for 3′ flanking exons.

Northern Blot and Hybridization

We electrophoresed 2 μg of mRNAs through a 1.5% agarose gel in 2.2 M formaldehyde and 1× MOPS buffer and blotted to a positively charged membrane by standard procedures. Northern blot filters of multiple normal tissue mRNAs were purchased (Clontech). The *FHIT* cDNA probe for hybridization was obtained using the *FHIT* cDNA as template for PCR amplification with the following primer pair: 5′-TGAGGACATGTCGTTCAGATTTGG-3′, nucleotides −7 to 17; 5′-CTGTGTCACTGAAAGTAGACC-3′, nucleotides 449 to 429. Probes were labeled by random priming with [^{32}P]dCTP, and 2×10^6 cpm/ml was hybridized to each filter using standard methods.

Nested RT–PCR and Sequencing of cDNAs

First strand cDNAs were synthesized, and 1 μl of each product was subjected to a first round of PCR amplification with 30 cycles of 95°C for 20 s, 60°C for 30 s, and 72°C for 1 min, with 5% dimethylsulfoxide and 0.5 mM spermidine in 10 μl reaction volume under standard conditions, using primers 5U2 and 3D2, indicated in Figure 2A. We subjected 1 μl of the reaction products, after 20-fold dilution, to a second round of PCR amplification using nested primers 5U1 and *3D1* (shown in Figure 2A), under the conditions noted above, except the reaction volume was 30 μl. The PCR products were run on 1.5% agarose gels, stained with ethidium bromide, and purified, and 2.5 ng was sequenced using the 5U1 primer.

In Vitro Transcription and Translation

Three different fragments of DNA containing the *FHIT* gene were obtained by PCR, using oligonucleotides UR5 (5′-CTGTAAAGGTCC GTAGTG-3′, nucleotides −171 to −154 in Figure 2A) and O6 (5′CTGTGTCACTGAAAGTAGACC3′, the reverse complement of nucleotides 429-449). Amplifications were performed in 100 μl final volume of 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM deoxynucleotide triphosphates, 10 ng of RT–PCR products, and 2.5 U of Taq polymerase using an Omni Gene Thermal Cycler for 25 cycles, consisting of 94°C for 1 min, 52°C for 1 min, and an extension step at 72°C for 45 s. PCR products were inserted in a PCRII plasmid using the TA cloning system (Invitrogen). Recombinant vectors, containing the normal *FHIT* and aberrant genes under the control of the T7 promoter, were sequenced and used for in vitro transcription and translation.

The in vitro transcription and translation reactions were performed by TnT-coupled reticulocyte systems (Promega) in a final volume of 50 μl containing 1/2 vol of rabbit reticulocyte lysate, 1 μg of recombinant plasmid DNA, 10 U of T7 polymerase, 20 μM amino acid mixture without methionine, 40 μM [^{35}S] methionine (Amersham), and 40 U of RNasin ribonuclease inhibitor. Reactions were carried out for 90 min at 30°. The synthesized proteins were analyzed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and autoradiography.

Acknowledgments

Correspondence should be addressed to K. H. This work was supported by United States Public Health Service grants CA51083, CA39860, and CA21124 and a gift from R. R. M. Carpenter III and M. K. Carpenter; National Cancer Institute Cancer Center grant CA56336 supported the Kimmel Cancer Center shared research facilities that expedited the study. We thank C. Williams for collection of tissues and Drs. L. Barnes and P. Garrison for helpful discussions of diadenosine tetraphosphate hydrolases. We also thank M. Shimizu for expert technical assistance and A. Mathis for preparation of the manuscript.

Received January 5, 1996; revised January 19, 1996.

References

Boldog, F.L., Gemmill, R.M., Wilke, C.M., Glover, T.W., Nilsson, A.S., Chandrasekharappa, S.C., Brown, R.S., Li, F.P., and Drabkin, H.A. (1993). Positional cloning of the hereditary renal carcinoma 3;8 chromosome translocation breakpoint. Proc. Natl. Acad. Sci. USA *90*, 8509–8513.

Boldog, F.L, Waggoner, B., Glover, T.W., Chumakov, I., Paslier, D., Cohen, D., Gemmill, R.M., and Drabkin, H.A. (1994). Integrated YAC contig containing the 3p14.2 hereditary renal carcinoma 3;8 translocation breakpoint and the fragile site FRA3B. Genes Chromosom. Cancer *11*, 216–221.

Cohen, A.J., Li, F.P., Berg, S., Marchetto, D.J., Tsai, S., Jacobs, S.C., and Brown, R.S. (1979). Hereditary renal-cell carcinoma associated with a chromosomal translocation. N. Engl. J. Med. *301*, 592–595.

Druck, T., Kastury, K., Hadaczek, P., Podolski, J., Toloczko, A., Sikorski, A., Ohta, M., LaForgia, S., Lasota, J., McCue, P., Lubinski, J., and Huebner, K. (1995). Loss of heterozygosity at the familial RCC t(3;8) locus in most clear cell renal carcinomas. Cancer Res. *55*, 5348–5355.

Fong, L.Y.Y., Sivak, A., and Newberne, P.M. (1978). Zinc deficiency and methylbenzylnitrosaamine-induced esophageal cancer in rats. J. Natl. Cancer Inst. *61*, 145–150.

Fu, Y-H., Kuhl, D.P.A., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M.H., Holden, J.J.A., Fenwick, R.G., Warren, S. T., Oostra, B.A., Nelson, D.L., and Caskey, C.T. (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell *67*, 1047–1058.

Glover, T.W., Coyle-Morris, J.F, Brown, R.S., Berger, C.S., Gemmill, R.M., and Hecht, F. (1988). Translocation t(3;8)(p14.2;q24) in renal cell carcinoma affects expression of the common fragile site at 3p14 (FRA3B) in lymphocytes. Cancer Genet. Cytogenet. *31*, 69–73.

Gnarra, J.R., Tory, K., Weng, Y., Schmidt, L., Wei, M.H., Li, H., Latif, F., Liu, S., Chen, F., Duh, F.-M., Lubensky, I., Duan, D.R., Florence, C., Pozzatti, R., Walther, M.M., Bander, N.H., Grossman, H.B., Brauch, H., Pomer, S., Brooks, J.D., Isaacs, W.B., Lerman, M.I., Zbar, B., and Linehan, W.M. (1995). Mutations of the VHL tumour suppressor gene in renal carcinoma. Nature Genet. *7*, 85–90.

Huang, Y., Garrison, P.N., and Barnes, L.D. (1995). Cloning of the *Schizosaccharomyces pombe* gene encoding diadenosine 5′,5′′′-P^1, P^4-tetraphosphate (Ap_4A) asymmetrical hydrolase: sequence similarity with the histidine triad (HIT) family. Biochem. J. *312*, 925–932.

Jones, C., Slijepcevic, P., Marsh, S., Baker, E., Langdon, W.Y., Richards, R.I., and Tunnacliffe, A. (1994). Physical linkage of the fragile site FRA11B and a Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. Hum. Mol. Genet. *3*, 2123–2130.

Kastury, K. Baffa, R., Druck, T., Ohta, M., Cotticelli, M.G., Inoue, H., Negrini, M., Rugge, M., Huang, D., Croce, C.M., Palazzo, J., and Huebner, K. (1996a). Potential gastrointestinal tumor suppressor locus at the 3p14.2 FRA3B site identified by homozygous deletions in tumor cell lines. Cancer Res., in press.

Kastury, K., Ohta, M., Lasota, J., Moir, D., Dorman, T., LaForgia, S., Druck, T., and Huebner, K. (1996b). Structure of the human receptor tyrosine phosphatase γ gene (PTPRG) and relation to the familial RCC t(3;8) chromosome translocation. Genomics, in press.

Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T., Schlessinger, D., Sutherland, G.R., and Richards, R.I. (1991). Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence $p(CCG)_n$. Science *252*, 1711–1714.

LaForgia, S., Morse, B., Levy, J., Barnea, G., Cannizzaro, L.A., Li, F., Nowell, P.C., Boghosian-Sell, L., Glick, J., Weston, A., Harris, C.C., Drabkin, H., Patterson, D., Croce, C.M., Schlessinger, J., and Huebner, K. (1991). Receptor protein-tyrosine phosphatase γ is a candidate tumor suppressor gene at human chromosome region 3p21. Proc. Natl. Acad. Sci. USA *88*, 5036–5040.

LaForgia, S., Lasota, J., Latif, F., Boghosian-Sell, L., Kastury, K., Ohta, M., Druck, T., Atchison, L., Cannizzaro, L., Barnea, G., Schlessinger, J., Modi, W., Kuzmin, I., Tory, K., Zbar, B., Croce, C. M., Lerman, M., and Huebner, K. (1993). Detailed genetic and physical map of the 3p chromosome region surrounding the familial RCC chromosome translocation, t(3;8)(p14.2;q24.1). Cancer Res. *53*, 3118–3124.

Li, F.P., Decker, H.J.H., Zbar, B., Stanton, V.P., Kovacs, G., Seizinger, B.R., Aburatani, H., Sandberg, A.A., Berg, S., Hosoe, S., and Brown, R.S. (1993). Clinical and genetic studies of renal cell carcinomas in a family with a constitutional chromosome 3;8 translocation. Ann. Intern. Med. *118*, 106–111.

Lisitsyn, N.A., Lisitsina, N.M., Dalbagni, G., Barker, P., Sanchez, C.A., Gnarra, J., Linehan, W.M., Reid, B.J., and Wigler, M.H. (1995). Comparative genomic analysis of tumors: detection of DNA losses and amplification. Proc. Natl. Acad. Sci. USA *92*, 151–155.

Lo, K.-W., Tsao, S.-W., Huang, D.P., and Lee, J.C.K. (1994). Detailed deletion mapping on the short arm of chromosome 3 in nasopharyngeal carcinomas. Int. J. Oncol. *4*, 1359–1364.

Lubinski, J., Hadaczek, P., Podolski, J., Toloczko, A., Sikorski, A., McCue, P., Druck, T., and Huebner, K. (1994). Common regions of deletion in chromosome regions 3p12 and 3p14.2 in primary clear cell renal carcinomas. Cancer Res. *54*, 3710–3713.

Michaelis, S.C., Bardenheuer, W., Lux, A., Schramm, A., Gockel, A., Siebert, R., Willers, C., Schmidtke, K., Todt, B., van der Hout, A.H., Buys, C.H.C.M., Heppell-Parton, A.C., Rabbitts, P.H., Ungar, S., Smith, D., LePaslier, D., Cohen, D., Opalka, B., and Schütte, J. (1995). Characterization and chromosomal assignment of yeast artificial chromosomes containing human 3p13-p21-specific sequence tagged sites. Cancer Genet. Cytogenet. *81*, 1–12.

Mozier, N.M, Walsh, M.P., and Pearson, J.D. (1991). Characterization of a novel zinc binding site of protein kinase C inhibitor-1. FEBS Lett. *279*, 14–18.

Ohta, M., Berd, D., Shimizu, M., Nagai, H., Mastrangelo, M., Shields, J.A., Shields, C.L., Croce, C.M., and Huebner, K. (1995). Distinct common regions of deletion within chromosome region 9p21 in melanomas. Int. J. Cancer, in press.

Paradee, W., Mullins, C., He, Z., Glover, T., Wilke, C., Opalka, B., Schutte, J., and Smith, D. (1995). Precise localization of aphidicolin-induced breakpoints on the short arm of human chromosome 3. Genomics *27*, 358–361.

Pearson, J.D., DeWald, D.B., Mathews, W.R., Mozier, N.M., Zurcher-Neely, H.A., Heinrikson, R.L., Morris, M.A., McCubbin, W.D., McDonald, J.R., Fraser, E.D., Vogel, H.J., Kay, C.M., and Walsh, M.P. (1990). Amino acid sequence and characterization of a protein inhibitor of protein kinase C. J. Biol. Chem. *265*, 4583–4591.

Rassool, F.V., LeBeau, M.M., Neilly, M.E., Van Melle, E., Espinosa, R., and McKeithan, T.W. (1992). Increased genetic instability of the common fragile site at 3p14 after integration of exogenous DNA. Am. J. Hum. Genet. *50*, 1243–1251.

Roche, J., Whisenant, E., Boldog, F., Loeb, D., Vance, J.M., and Drabkin, H. (1994). Dinucleotide repeats flanking the renal cell carcinoma breakpoint at 3p14.2. Hum. Mol. Genet. *3*, 215.

Seraphin, B. (1992). The HIT protein family: a new family of proteins present in prokaryotes, yeast and mammals. DNA Seq. *3*, 177–179.

Sutherland, G.R. (1991). Chromosomal fragile sites. Genet. Anal. Tech. Appl. *8*, 1616–166.

Thorne, N.M.H., Hankin, S., Wilkinson, M.C., Nunez, C., Barraclough, R., and McLennan, A.G. (1995). Human diadenosine 5′5′′′-P^1,

P^4-tetraphosphate pyrophosphohydrolase is a member of the MutT family of nucleotide pyrophosphatases. Biochem. J. *311*, 717–721.

Verkerk, A.J.M.H., Pieretti, M., Sutcliffe, J.S., Fu, Y.-H., Kuhl, D.P.A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B.E., van Ommen, G.J.B., Blonden, L.A.J., Riggins, D.L., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A., and Warren, S.T. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell *65*, 905–914.

Wang, N., and Perkins, K.L. (1984). Involvement of band 3p14 in t(3;8) hereditary renal carcinoma. Cancer Genet. Cytogenet. *11*, 479–481.

Wilke, C.M., Guo, S.-W., Hall, B.K., Boldog, F., Gemmill, R.M., Chandrasekharappa, S.C., Barcroft, C.L., Drabkin, H.A., and Glover, T.W. (1994). Multicolor FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the t(3;8) associated with hereditary renal cell carcinoma. Genomics *22*, 319–326.

Yang, C.S. (1980). Research on esophageal cancer in China. Cancer Res. *40*, 2633–2644.

Yang, G., Lipkin, M., Yang, K., Wang, G., Li, J., Yang, C., Winawer, S., Newmark, H., Biot, W.J., and Fraumeni, J.F., Jr. (1987). Proliferation of esophageal cells among residents of Linxian, People's Republic of China. J. Natl. Cancer Inst. *79*, 1241–1246.

Yokota, J., Tsukada, Y., Nakajima, T., Gotoh, M., Shimosato, Y., Mori, N., Tsunokawa, Y., Sugimura, T., and Terada, M. (1989). Loss of heterozygosity on the short arm of chromosome 3 in carcinoma of the uterine cervix. Cancer Res. *49*, 3598–3601.

Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, J.C., Warren, S.T., Schlessinger, D., Sutherland, G.R., and Richards, R.I. (1991). Fragile X genotype characterized by an unstable region of DNA. Science *252*, 1179–1181.

Yunis, J.J., and Soreng, A.L. (1984). Constitutive fragile sites and cancer. Science *226*, 1199–1204.

GenBank Accession Number

The accession number for the sequence described in this paper is U46922.