

Nitrilase and Fhit homologs are encoded as fusion proteins in *Drosophila melanogaster* and *Caenorhabditis elegans*

YURI PEKARSKY*, MANUELA CAMPIGLIO†, ZURAB SIPRASHVILI*, TERESA DRUCK*, YURII SEDKOV*, SERGEI TILLIB*, ALEXANDRA DRAGANESCU*, PETER WERMUTH*, JOEL H. ROTHMAN‡, KAY HUEBNER*, ARTHUR M. BUCHBERG*, ALEXANDER MAZO*, CHARLES BRENNER*, AND CARLO M. CROCE*§

*Kimmel Cancer Institute, Jefferson Medical College, Philadelphia, PA 19107; †Istituto Nazionale Tumori, Milan, Italy, 20133; and ‡University of California, Santa Barbara, CA 93106

Communicated by Sidney Weinhouse, Jefferson Medical College, Philadelphia, PA, May 27, 1998 (received for review April 20, 1998)

ABSTRACT The tumor suppressor gene *FHIT* encompasses the common human chromosomal fragile site at 3p14.2 and numerous cancer cell biallelic deletions. To study *Fhit* function we cloned and characterized *FHIT* genes from *Drosophila melanogaster* and *Caenorhabditis elegans*. Both genes code for fusion proteins in which the *Fhit* domain is fused with a novel domain showing homology to bacterial and plant nitrilases; the *D. melanogaster* fusion protein exhibited diadenosine triphosphate (ApppA) hydrolase activity expected of an authentic *Fhit* homolog. In human and mouse, the nitrilase homologs and *Fhit* are encoded by two different genes: *FHIT* and *NIT1*, localized on chromosomes 3 and 1 in human, and 14 and 1 in mouse, respectively. We cloned and characterized human and murine *NIT1* genes and determined their exon-intron structure, patterns of expression, and alternative processing of their mRNAs. The tissue specificity of expression of murine *Fhit* and *Nit1* genes was nearly identical. Because fusion proteins with dual or triple enzymatic activities have been found to carry out specific steps in a given biochemical or biosynthetic pathway, we postulate that *Fhit* and *Nit1* likewise collaborate in a biochemical or cellular pathway in mammalian cells.

The human *FHIT* gene at chromosome 3p14.2, spanning the constitutive chromosomal fragile site FRA3B, often is rearranged in the most common forms of human cancer and is a candidate tumor suppressor gene (1). The human *FHIT* gene, which is greater than 1 megabase in size, encodes an mRNA of 1.1 kilobases and a protein of 147 amino acids.

The rearrangements most commonly seen are deletions within the gene (1–4). These deletions, often occurring independently in both alleles and resulting in inactivation, have been reported in tumor-derived cell lines and primary tumors of lung (2), head and neck (3), stomach, colon (1), and other organs. In cell lines derived from several tumor types, DNA rearrangements in the *FHIT* locus correlated with RNA and/or *Fhit* protein alterations (4).

Because the inactivation of the *FHIT* gene by point mutations has not been demonstrated conclusively and several reports showed the amplification of aberrant-sized *FHIT* reverse transcription–PCR (RT-PCR) products from normal cell RNA (5, 6), a number of investigators have suggested that the *FHIT* gene may not be a tumor suppressor gene. On the other hand, it recently was reported that re-expression of *Fhit* in lung, stomach, and kidney tumor cell lines lacking endogenous protein suppressed tumorigenicity *in vivo* in four of four cancer cell lines (7). This finding suggests that *FHIT* is indeed a tumor suppressor gene. The same report suggested that *Fhit* enzymatic activity is not required for its tumor suppressor function (7).

Fhit protein is a member of the histidine triad (HIT) superfamily of nucleotide binding proteins (8) and is similar to the *Schizosaccharomyces pombe* diadenosine tetraphosphate (Ap₄A) hydrolase. Barnes *et al.* (9) reported that, *in vitro*, *Fhit* has diadenosine triphosphate (ApppA) hydrolase enzymatic activity.

Neither the *in vivo* function of *Fhit* nor the mechanism of its tumor suppressor activity is known. Nonetheless, genetic, biochemical, and crystallographic analysis suggest that the enzyme-substrate complex is the active form that signals for tumor suppression (10). One approach to investigate function is to study *Fhit* in model organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans*.

Here we describe the isolation and characterization of the *Fhit* gene in these organisms. *Fhit* occurs in a fusion protein, NitFhit, in *D. melanogaster* and *C. elegans* although *FHIT* and *NIT1* are separate genes in mammalian cells. The human and mouse *NIT1* genes are members of an uncharacterized mammalian gene family with homology to bacterial and plant nitrilases, enzymes that cleave nitriles and organic amides to the corresponding carboxylic acids plus ammonia (11).

MATERIALS AND METHODS

Genomic and cDNA Clones. One million plaques of a mouse genomic library (bacteriophage library from strain SVJ129, Stratagene) and 100,000 plaques of a *D. melanogaster* genomic library were screened with corresponding cDNA probes. Clones were purified and DNA was isolated (12). Sequencing was carried out by using Perkin–Elmer thermal cyclers and Applied Biosystems 377 automated DNA sequencers. DNA pools from a human (bacterial artificial chromosome) library (Research Genetics, Huntsville, AL) were screened by PCR with *NIT1* primers (TCTGAACTGCAGTCTGACCTCA and CAGGCACAGCTCCCCTCACTT) according to the supplier's protocol. The DNA from the positive clone, 31K11, has been isolated by using standard procedures (12) and sequenced. Chromosomal localization of the human *NIT1* gene was determined by using a radiation hybrid mapping panel (Research Genetics) according to the supplier's protocol and with the same primers as above. To map murine *Nit1* gene, Southern blot analysis of genomic DNA from progeny of a (*AEJ/Gn-a bp^H/a bp^H* × *Mus spretus*)F₁ × *AEJ/Gn-a bp^H/a*

Abbreviations: RT-PCR, reverse transcription–PCR; HIT, histidine triad; ApppA, diadenosine triphosphate; EST, expressed sequence tag; RACE, rapid amplification of cDNA ends.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF069984–AF069989).

§To whom reprint requests should be addressed at: Kimmel Cancer Institute, Thomas Jefferson University, Room 1050 BLSB, 233 S. 10th Street, Philadelphia, PA 19107. e-mail: croce@calvin.jci.tju.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.

© 1998 by The National Academy of Sciences 0027-8424/98/958744-6\$2.00/0
PNAS is available online at <http://www.pnas.org>.

bp^H backcross was performed by using a full-length murine *Nit* cDNA probe (13). This probe detected a unique 2.0-kb *Dra*I fragment in AEJ DNA and a unique 0.75-kb fragment in *M. spretus* DNA. Segregation of these fragments were followed in 180 N2 offspring of the backcross.

Additional Mit markers (*D1Mit34*, *D1Mit35*, and *D1Mit209*) were typed from DNA of 92 mice by using PCR consisting of an initial denaturation of 4 min at 94°C followed by 40 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 30 sec. Linkage analysis was performed by using the computer program SPRETUS MADNESS: PART DEUX. Human and mouse *NIT1* expressed sequence tag (EST) clones were purchased from Research Genetics. The sequences of human and murine *NIT1* genes and cDNAs and *D. melanogaster* and *C. elegans NitFhit* cDNAs have been deposited in GenBank.

In Situ Hybridization. *D. melanogaster* polytene chromosome spreads were prepared from salivary glands of third-instar larvae as described (14). *NitFhit* DNA fragments were labeled with digoxigenin-11-dUTP by using a random-primed DNA labeling kit (Boehringer Mannheim) and were used as probes for the chromosomal *in situ* hybridization. Hybridization was for 20 hr at 37°C in hybridization buffer: 50% formamide, 2× standard saline citrate (SSC), 10% dextran sulfate, and 400 mg/ml of salmon sperm DNA. Antidigoxigenin-fluorescein antibodies (Boehringer Mannheim) were used for detection of hybridizing regions. DNA was counterstained with Hoechst 33258 (Sigma). The slides were analyzed by fluorescence microscopy. For *in situ* hybridization, embryos were fixed and processed as described previously (15), except that single-stranded RNA probes were used. Full-length *NitFhit* cDNA was cloned into BluescriptII KS+ vector and used to synthesize antisense RNA probes with the Genius 4 kit (Boehringer Mannheim).

RT-PCR, Northern, and Rapid Amplification of cDNA Ends (RACE) Analysis. Human and mouse multiple tissue Northern blots (CLONTECH) were hybridized with corresponding *NIT1* cDNA probes and washed by using the supplier's protocol. For the HeLa cell line, total RNA was isolated from 1–5 × 10⁸ cells by using Trizol reagent (GIBCO/BRL). *D. melanogaster* poly(A)⁺ RNA was purchased from CLONTECH. Three micrograms of poly(A)⁺ RNA or 15 μg of total RNA were electrophoresed in 0.8% agarose in a borate buffer containing formaldehyde, transferred to HybondN⁺ membrane (Amersham) by using standard procedures and hybridized as described above. For RT-PCR, 200 ng of poly(A)⁺ RNA or 3 μg of total RNA were treated with DNaseI (amplification grade, GIBCO/BRL) following the manufacturer's protocol. DNaseI-treated RNA was used in RT reactions as follows: 10 nM each dNTP, 100 pmol random hexamers [oligo(dT) priming was used in some cases], DNaseI-treated RNA, and 200 units of murine leukemia virus reverse transcriptase (GIBCO/BRL), in a total volume of 20 μl were incubated at 42° for 1 hr followed by the addition of 10 μg of RNase A and incubation at 37° for 30 min. One microliter of the reaction was used for each PCR. PCRs were carried out under standard conditions by using 10 pmol of each gene-specific primer and 25–35 cycles of: 95° 30", 55–60° 30", 72° 1'. Products were separated on 1.5% agarose gels and sometimes isolated and sequenced or cloned and sequenced. Oligo(dT)-primed double-stranded cDNA was synthesized by using procedures and reagents from the Marathon RACE cDNA amplification kit (CLONTECH); the cDNA was ligated to Marathon adapters (CLONTECH). 3' and 5' RACE products were generated by long PCR using gene-specific primers and the AP1 primer (CLONTECH). To increase the specificity of the procedure, the second PCR was carried out by using nested gene-specific primers and the AP2 primer (CLONTECH). PCRs were performed according to the Marathon protocol by using the Expand long template PCR system (Boehringer Mannheim) and 30 cycles of 94° 30", 60° 30", 68° 4'. RACE

products were electrophoresed, identified by hybridization, and sequenced. Degenerate *Fhit* primers were: GTNGTNC-CNGGNCAAGTNGT and ACRTGNACRTGYTTNACNG-TYTGNCG. *D. melanogaster Fhit* RACE and RT-PCR primers were: GCGCCTTTGTGGCCTCGACTG and CGGTG-GCGGAAGTTGTCTGGT. *C. elegans Fhit* RACE and RT-PCR primers were: GTGGCGGCTGCTCAAAGTGG and TCGCGACGATGAACAAGTCCG. Human *NIT1* RT-PCR primers were: GCCCTCCGGATCCGACCCT (exon1); GACTACTCCCTATCCCGTC (exon 1a); GCTGCGAAGTGC-ACAGCTAAG and AACTGAAGCCTCTTCTCTGAC (exon 1c); TGGGCTTCATCACCAGGCCT and CTGGGCTGAGCACAAAGTACTG (exon 2); GCTTGTCTGGCGT-CGATGTTA (exon 3).

Protein Expression and Enzymatic Characterization. The *NitFhit* cDNA was amplified with primers TGACGTGCA-CATATGTCAACTCTAGTTAATACCACG and TGGG-TACCTCGACTAGCTTATGTCC, digested with *Nde*I and *Kpn*I, and cloned into plasmid pSGA02 (16) as a *Nde*I-*Kpn*I fragment. *Escherichia coli* strain SG100 (16) transformants were grown in Luria-Bertani medium with 100 μg/ml of ampicillin and 15 μg/ml of chloramphenicol at 15°C. When the culture reached an optical density (600 nm) of 0.25, isopropyl β-D-thiogalactoside was added to a final concentration of 200 μM. *NitFhit* protein was purified from inclusion bodies as described (17). Briefly, the cell pellet from a 1-liter culture was resuspended in 50 ml of 20 mM Tris-HCl (pH 7.5), 20% sucrose, 1 mM EDTA and repelleted. Outer cell walls were lysed by resuspension in ice water. Spheroplasts were pelleted, resuspended in 140 mM NaCl, 2.7 mM KCl, 12 mM Na₂P₀₄ (pH 7.3), 5 mM EDTA, 500 mM phenylmethylsulfonyl fluoride, 1 μg/ml of leupeptin, and 20 μg/ml of aprotinin, and sonicated. The resulting inclusion body preparation was washed and solubilized in 5 M guanidinium hydrochloride, 50 mM Tris-HCl (pH 8.0), 5 mM EDTA. Soluble *NitFhit* protein was added dropwise to 250 ml of 50 mM Tris-HCl (pH 8.0), 1 mM DTT, 20% glycerol at 40°C. After a 14-hr incubation, the 13-kg supernatant was concentrated 100-fold with a Centricon filter. A 1-liter culture yielded approximately 200 μg of partially purified, soluble *NitFhit*. ApppA hydrolase activity was assayed at 30°C in 20 μl of 50 mM Na-Hepes (pH 7.5), 10% glycerol, 0.5 mM MnCl₂, 4 mM ApppA, 1 μM *NitFhit*. TLC plates were developed as described (9).

RESULTS

Cloning and Characterization of *D. melanogaster* and *C. elegans Fhit* Homologs. To obtain *D. melanogaster Fhit* sequences, degenerate primers were designed to amplify sequences in the conserved regions of exons 5 and 7 of human *FHIT*. RT-PCR experiments with these primers and *D. melanogaster* RNA resulted in an ≈200-bp product whose translated sequence predicts a protein that showed ≈50% identity to human *Fhit* protein. This amplified sequence was used to design specific *D. melanogaster Fhit* primers. 5' and 3' RACE with these primers yielded an ≈1.5-kb full-length cDNA [including polyadenylation signal and poly(A) tail] encoding a 460-aa protein with a 145-aa C-terminal portion homologous to human *Fhit* (40% identity and 47% similarity) and a 315-aa N-terminal extension (Fig. 1). Northern analysis (Fig. 2C) showed a single band of ≈1.5 kb in both embryo and adult *D. melanogaster*, implying that the full-length cDNA had been cloned.

The 460-aa predicted protein sequence was used in a BLASTP search. Of the top 50 scoring alignments, 22 aligned with the 145-residue C-terminal segment (*Fhit*-related sequences) and 28 aligned with the 315-residue N-terminal segment. The 28 sequences aligning with the N terminus were led by an uncharacterized gene from chromosome X of *Saccharomyces cerevisiae* (*P*-value of 1.4 × 10⁻⁴⁵), followed by uncharacterized

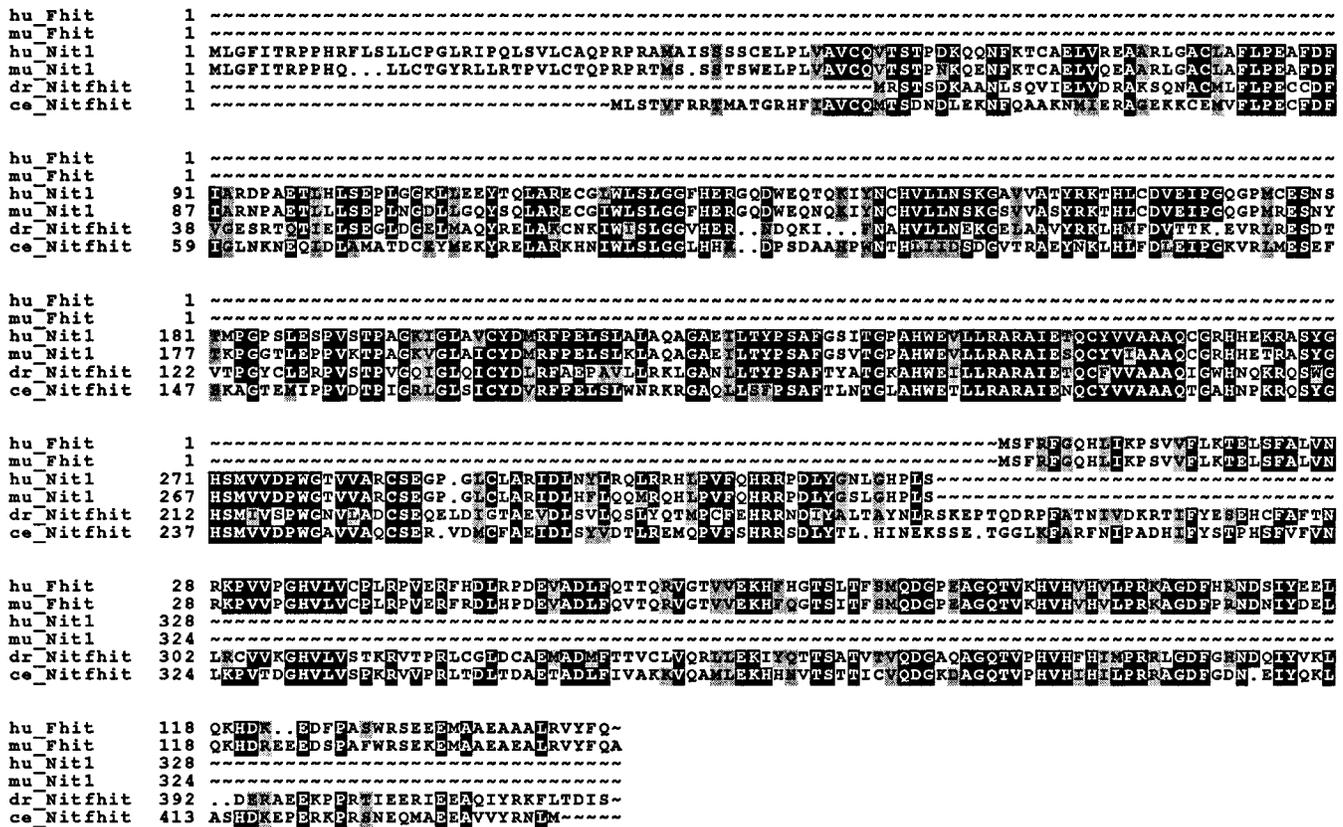


FIG. 1. Sequence comparison of human, murine, *D. melanogaster*, and *C. elegans* Nit1 and Fhit proteins. Identities are shown in black boxes, similarities are shown in shaded boxes. For human and mouse *FHIT* GenBank accession nos. are U46922 and AF047699, respectively.

ORFs of many bacterial genomes and a series of enzymes from plants and bacteria that have been characterized as nitrilases and amidases (11). Thus, the 460-aa predicted protein contains a N-terminal nitrilase domain and a C-terminal Fhit domain and was designated NitFhit.

The *D. melanogaster* NitFhit cDNA probe was used to screen a *D. melanogaster* lambda genomic library. Sequencing of positive clones revealed that the gene is intronless and, interestingly, the 1.5-kb NitFhit gene is localized within the 1.6-kb intron 1 of the *D. melanogaster* homolog of the murine glycerol kinase (*Gyk*) gene (18). The direction of transcription of the NitFhit gene is opposite to that of the *Gyk* gene (Fig. 3A). It

is not known whether such localization affects transcriptional regulation of these two genes.

The cytological position of the NitFhit gene was determined by *in situ* hybridization to salivary gland polytene chromosomes. These experiments showed that there is only one copy of the sequence, which is localized to region 61A, at the tip of the left arm of chromosome 3 (not shown). Digoxigenin-labeled RNA probes were hybridized to whole-mount embryos to determine the pattern of expression during development. NitFhit RNA was found to be uniformly expressed throughout the embryo (not shown), suggesting that NitFhit protein could be important for most of the embryonic cells.

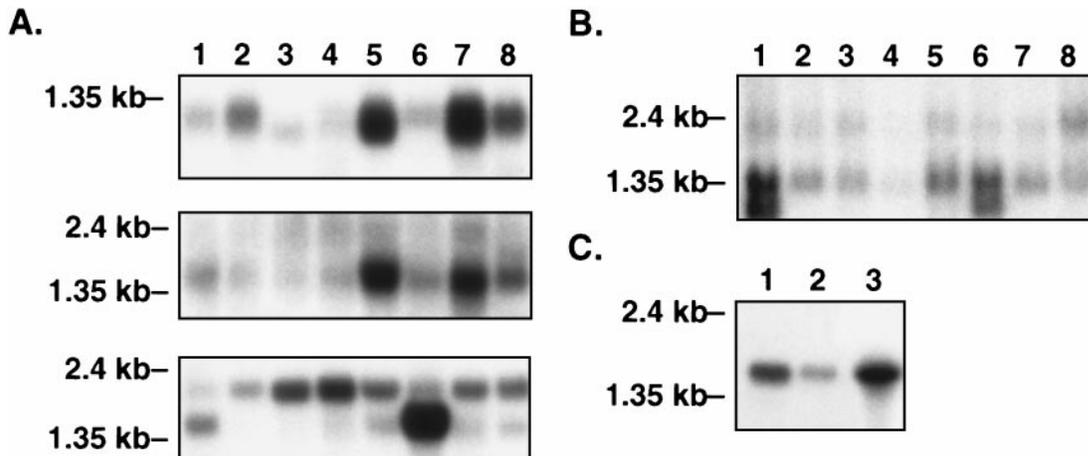


FIG. 2. Expression of Nit1 and Fhit mRNAs in murine and human tissues. (A) Mouse multiple tissues Northern blot. Lanes 1–8: heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis. (Top) Fhit probe; (Middle) Nit1 probe; (Bottom) actin probe. (B) Human blot, NIT1 probe. Lanes 1–8: heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. (C) Lanes 1 and 2: *D. melanogaster* adult; *D. melanogaster* embryo; *D. melanogaster* NitFhit probe. Lane 3: *C. elegans* adult; *C. elegans* NitFhit probe.

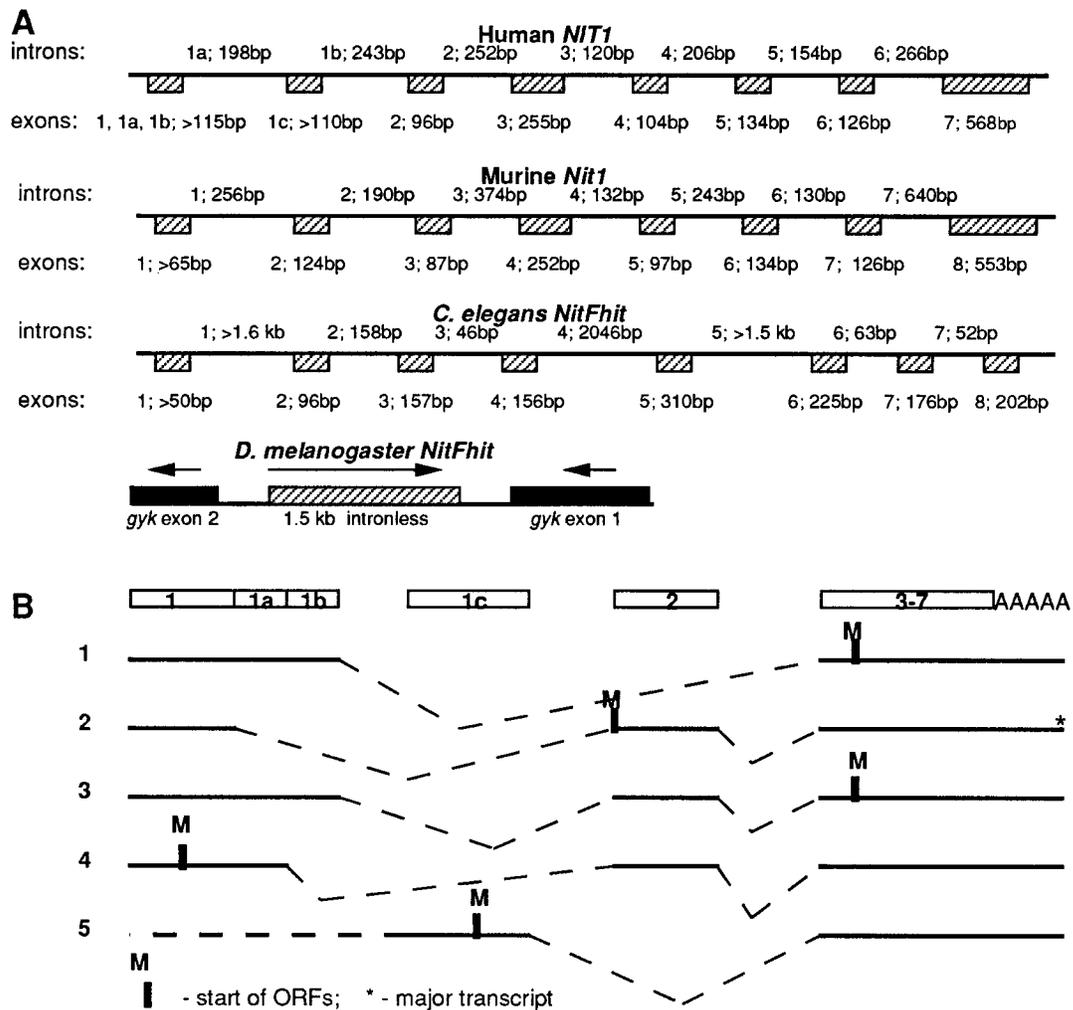


FIG. 3. Genomic organization of human and murine *Nit1* genes and *D. melanogaster* and *C. elegans NitFhit* genes. (A) Exon-intron structure of the genes. (B) Alternative processing of human *NIT1* gene.

Because human Fhit protein and the *D. melanogaster* Fhit domain of NitFhit were only 40% identical, to show that we cloned the authentic *D. melanogaster* Fhit homolog, we tested its enzymatic activity. Fig. 4 shows that recombinant *D.*

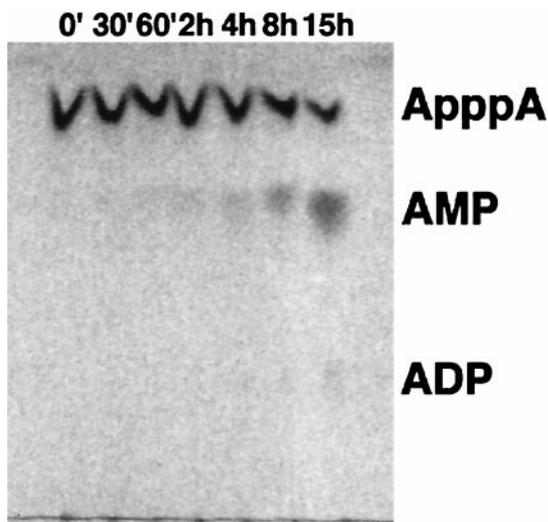


FIG. 4. Cleavage of ApppA by *D. melanogaster* NitFhit. At indicated times of incubation, samples were spotted on TLC plates with appropriate nucleotide standards.

melanogaster NitFhit is capable of cleaving ApppA to AMP and ADP and therefore possesses ApppA hydrolase activity.

C. elegans. Fhit genomic sequences were identified from the Sanger database (contig Y56A3) by using BLAST searches. 5' and 3' RACE with *C. elegans Fhit* specific primers yielded a 1.4-kb cDNA [including polyadenylation signal and poly(A) tail] coding for a 440-aa protein (Fig. 1). Northern analysis (Fig. 2C) showed a single band of a similar size in adult worms. Similarly to *D. melanogaster*, the *C. elegans* protein contains an N-terminal nitrilase domain and a C-terminal Fhit domain (Fig. 1) with 50% identity and 57% similarity to human Fhit. Comparison between *C. elegans NitFhit* cDNA and genomic sequences from the Sanger database revealed that the *C. elegans NitFhit* gene comprises eight exons and is more than 6.5 kb in size (Fig. 3A); the nitrilase domain is encoded by exons 1–6, and the Fhit domain is encoded by exons 6–8. *D. melanogaster* and *C. elegans* NitFhit proteins are 50% identical and 59% similar and exhibit several conserved domains (Fig. 1).

Cloning and Characterization of Human and Murine *NIT1* cDNAs and Genes. Because Fhit and nitrilase domains are part of the same polypeptides in *D. melanogaster* and *C. elegans*, it is reasonable to suggest that they may be involved in the same biochemical or cellular pathway(s) in these organisms. Because we found that nitrilase homologs are conserved in animals, the mammalian nitrilase homologs were cloned as candidate Fhit-interacting proteins.

To obtain human and murine *Nit1* sequences we used the *D. melanogaster* nitrilase domain sequence in BLAST searches of

the GenBank EST database. We found numerous partially sequenced human and murine *Nit1* homologous ESTs. All mouse *Nit1* ESTs were identical as were all human *NIT1* ESTs, suggesting the presence of a single *NIT1* gene in mouse and human. To obtain the full-length human and mouse cDNAs we completely sequenced several human and mouse ESTs and human 5' and 3' RACE products. This resulted in the isolation of a ≈ 1.4 -kb full-length human sequence encoding 327 amino acids and a ≈ 1.4 -kb mouse full-length sequence coding for 323 amino acids (Fig. 1), although several alternatively spliced products were detected in both cases (see below and Fig. 3B). Both cDNAs are polyadenylated, but lack polyadenylation signals, although AT-rich regions are present at the very 3' end of each cDNA. Mouse and human *Nit1* amino acid sequences are 90% identical; the human *Nit* amino acid sequence is 58% similar and 50% identical to the *C. elegans* nitrilase domain and 63% similar and 53% identical to the *D. melanogaster* nitrilase domain (Fig. 1).

Murine lambda and human bacterial artificial chromosome genomic libraries were screened with the corresponding *NIT1* cDNA probes, yielding one mouse lambda clone and one human BAC clone containing the *NIT1* genes. The human and murine *NIT1* genomic regions were sequenced and compared with the corresponding cDNA sequences. The genomic structure of human and mouse *NIT1* genes is shown in Fig. 3A. Both genes are small: the human gene is ≈ 3.2 kb in size and contains seven exons; the murine gene is ≈ 3.6 kb in size and contains eight exons. Southern analysis confirmed that both human and mouse genomes harbor a single *NIT1* gene (not shown).

A radiation hybrid mapping panel (GeneBridge 4) was used to determine the chromosomal localization of the human *NIT1* gene. By analysis of PCR data at the Whitehead/Massachusetts Institute of Technology database (<http://www-genome.wi.mit.edu>), the *NIT1* gene was localized to 6.94 cR from the marker CHLC.GATA43A04, which is located at 1q21–1q22.

A full-length murine *Nit1* cDNA probe was used to determine the chromosomal location of the murine gene by linkage analysis. Interspecific backcross analysis of 180 N₂ mice demonstrated that the *Nit1* locus cosegregated with several previously mapped loci on distal mouse chromosome 1. The region to which *Nit1* maps was further defined by PCR of genomic DNA from 92 N₂ mice using the markers *D1Mit34*, *D1Mit35*, and *D1Mit209* (Research Genetics). The following order of the genes typed in the cross and the ratio of recombinants to N₂ mice was obtained: centromere - *D1Mit34* - 7/78 - *D1Mit35* - 8/90 - *Nit1* - 11/91 - *D1Mit209* - telomere. The genetic distances given in centiMorgans (\pm SE) are as follows: centromere - *D1Mit34* - 9.0 ± 3.2 - *D1Mit35* - 8.9 ± 3.0 - *Nit1* - 12.1 ± 3.4 - *D1Mit209* - telomere. This region of mouse chromosome 1 (1q21–1q23) is syntenic to human chromosome 1q and is therefore consistent with the localization of the human ortholog of *Nit1*.

Expression and Alternative Splicing of Human and Murine *NIT1* Genes. For the human gene, Northern analysis revealed two major transcripts of ≈ 1.4 kb and ≈ 2.4 kb in all adult tissues and tumor cell lines tested. A third band of ≈ 1.2 kb was observed in adult muscle and heart (Fig. 2B). Our longest cDNA (≈ 1.4 kb) corresponds to the ≈ 1.4 -kb transcript observed on Northern blots. The 1.2-kb band corresponds to transcript 1 on Fig. 3B (see below). It is not known whether the ≈ 2.4 -kb RNA represents an additional transcript or an incompletely processed mRNA. No significant variation in human *NIT1* mRNA levels was observed in different tissues (Fig. 2B). On the contrary, different mouse tissues showed different levels of expression of *Nit1* mRNA (Fig. 2A). The highest levels of *Nit1* mRNA were observed in mouse liver and kidney (Fig. 2A Middle, lanes 5 and 7). Interestingly, the pattern of *Nit1* expression was similar to the pattern of expression of *Fhit* (Fig. 2A Top and Middle), supporting the hypothesis that the proteins may act in concert or participate in the same pathway.

Analysis of mouse *Nit1* ESTs revealed that some transcripts lack exon 2 and encode a 323-aa protein. An alternative transcript containing exon 2 encodes a shorter, 290-aa protein starting with the methionine 34 (Fig. 1).

Analysis of human ESTs and 5' RACE products from HeLa and testis also suggested alternative processing. To investigate this possibility, a series of RT-PCR experiments was carried out. Fig. 5 shows the results obtained from HeLa RNA (similar results were obtained by using RNAs from the MDA-MB-436 breast cancer cell line and adult liver). The alternatively spliced transcripts are shown on Fig. 3B. Transcript 1, lacking exon 2, was represented by several ESTs in the GenBank EST database. This transcript probably corresponds to the ≈ 1.2 -kb transcript observed on Northern blots in adult muscle and heart. Transcript 2 encoding the 327-aa *Nit1* protein (Fig. 1) is a major transcript of human *NIT1* at least in the cell lines tested. This transcript lacks exons 1a and 1b. Transcript 3 has exons 1a and 1b; transcript 4 has exon 1a but lacks exon 1b (Fig. 3B). It is not known whether transcript 5 (lacking exon 2) starts from exon 1 or 1c.

The alternative initiating methionines of different transcripts are shown on Fig. 3B. Future immunoblot experiments will determine which of these methionines are used *in vivo*. Our preliminary data suggest that at least in COS-7 cells transfected with a construct containing transcript 2, the methionine in exon 3 (shown in transcripts 1 and 3, Fig. 3B) initiates more efficiently than the methionine in exon 2 (Fig. 3B, transcript 2).

DISCUSSION

Although the frequent loss of *Fhit* expression in several common human cancers is well documented (19–21), and results supporting its tumor suppressor activity have been reported (7), the role of *Fhit* in normal and tumor cell biology and the mechanism of its action *in vivo* are unknown. The AppA hydrolytic activity of *Fhit* (9) seems not to be required for its tumor suppressor function (7), and it has been suggested that the enzyme-substrate complex is the active form of *Fhit* (10). To facilitate an investigation of *Fhit* function, we initiated an analysis of the gene in model organisms by cloning and characterizing *D. melanogaster* and *C. elegans* *Fhit* genes.

Surprisingly, in flies and worms, *Fhit* is expressed as a fusion protein with the *Fhit* domain fused to a "Nit" domain showing homology to plant and bacterial nitrilases (11). We further isolated human and murine *NIT1* genes. *Nit* and *Fhit* are expressed as separate proteins in mammals but, at the mRNA level, are coordinately expressed in mouse tissues.

In several eukaryotic biosynthetic pathways multiple steps are catalyzed by multifunctional proteins containing two or more enzymatic activities. The same steps in prokaryotes frequently are carried out by monoenzymatic proteins that are homologs of each domain of the corresponding eukaryotic protein (22). For example, Gars, Gart, and Airs are domains of the same protein in *D. melanogaster* and mammals. These

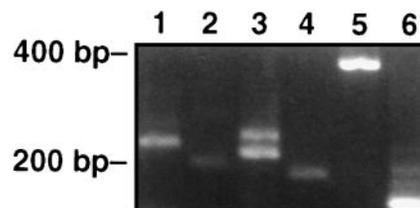


Fig. 5. Analysis of alternative transcripts of human *NIT1* by RT-PCR. RT-PCR of HeLa RNA was performed with primers in different exons. Lanes 1–6: exons 1 and 3 (transcript 2); exons 1C and 3 (transcript 5); exons 1A and 3 (transcripts 3, upper band and 4, lower band); exons 2 and 3 (transcripts 2–4); exons 1 and 1C (transcript 5); and exons 1 and 2 (transcript 2).

domains catalyze different steps in *de novo* synthesis of purines. In yeast, Garts homolog (Ade8) is a separate protein, and Gars and Airls homologs (Ade5 and Ade7) are domains of a bienzymatic protein; in bacteria, all three homologs (PurM, PurN, and PurD) are separate proteins (22). *De novo* pyrimidine biosynthesis illustrates a similar case (23). Recently, a fusion protein of a lipoyxygenase and catalase, both participating in the metabolism of fatty acids, has been identified in corals (24). In all of these examples, if domains of a multienzymatic protein in some organisms are expressed as individual proteins in other organisms, the individual proteins participate in the same pathways. This observation and the fact that Fhit and Nit1 exhibit similar expression patterns in murine tissues suggest that Fhit and Nit1 participate in the same cellular pathway in mammalian cells.

The critical questions are, what is this pathway and how is it altered in human cancers? Nothing is known about the function of nitrilase homologs in mammalian cells, and there is no obvious connection between the ability to bind and hydrolyze ApppA and the hydrolysis of nitrile-containing compounds. Enzymatic activity has yet to be demonstrated for animal nitrilase homologs. At this point we can only hypothesize about possible connections between the two enzymes. The best known nitrilases are plant enzymes responsible for synthesis of the plant growth factor auxin (25). One possibility is that an active form of Fhit stimulates or inhibits a nitrilase in production of growth-regulating compounds. Some of the plant nitrilases are tightly associated with plasma membranes (25). If this is true for the mammalian nitrilases and they can interact (directly or indirectly) with Fhit, they could be a factor that places Fhit in the proximity of the plasma membrane, similarly to the recruitment of Raf-1 by Ras (26). Genetic and biochemical analysis of the *Nit/Fhit* genes should help to reveal their roles in tumor progression.

We thank Becky Cusick and Jean Letofsky for technical assistance. This work was supported by National Institutes of Health Grants CA39880, CA51083, CA21124, and T32CA09678 (C.M.C.), AG13736 (J.H.R.), and R01CA75954 (C.B.)

1. Ohta, M., Inoue, H., Coticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., *et al.* (1996) *Cell* **84**, 587–597.
2. Sozzi, G., Veronese, M. L., Negrini, M., Baffa, R., Coticelli, M. G., Inoue, H., Tornielli, S., Pilotti, S., DeGregorio, L., Pastorino, V., *et al.* (1996) *Cell* **85**, 17–26.
3. Virgilio, L., Schuster, M., Gollin, S. M., Veronese, M. L., Ohta, M., Huebner, K. & Croce, C. M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 9770–9775.

4. Druck, T., Hadaczek, P., Fu, T.-B., Ohta, M., Siprashvili, Z., Baffa, R., Negrini, M., Kastury, K., Veronese, M. L., Rosen, D., *et al.* (1997) *Cancer Res.* **57**, 504–512.
5. Panagopoulos, I., Thelin, S., Mertens, F., Mitelman, F. & Aman, P. (1997) *Genes Chromosomes Cancer* **19**, 215–219.
6. Van den Berg, A., Draaijers, T. G., Kok, K., Timmer, T., Van der Veen, A. Y., Veldhuis, P. M. J. F., de Leij, L., Gerhartz, C. D., Naylor, S. L., Smith, D. I. & Buys, C. H. (1997) *Genes Chromosomes Cancer* **19**, 220–227.
7. Siprashvili, Z., Sozzi, G., Barnes, L. D., McCue, P., Robinson, A. K., Eryomin, V., Sard, L., Tagliabue, E., Greco, A., Fusetti, L., *et al.* (1997) *Proc. Natl. Acad. Sci. USA* **94**, 13771–13776.
8. Brenner, C., Garrison, P., Gilmour, J., Peisach, D., Ringe, D., Petsko, G. A. & Lowenstein, J. M. (1997) *Nat. Struct. Biol.* **4**, 231–238.
9. Barnes, L. D., Garrison, P. N., Siprashvili, Z., Guranowski, A., Robinson, A. K., Ingram, S. W., Croce, C. M., Ohta, M. & Huebner, K. (1996) *Biochemistry* **35**, 11529–11535.
10. Pace, H. C., Garrison, P. N., Robinson, A. K., Barnes, L. D., Draganescu, A., Rosler, A., Blackburn, M. G., Siprashvili, S., Croce, C. M., Huebner, K. & Brenner, C. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 5484–5489.
11. Bork, P. & Koonin, E. V. (1994) *Protein Sci.* **3**, 1244–1246.
12. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
13. Marini, J. C., Nelson, K. K., Battey, J. & Siracusa, L. D. (1993) *Genomics* **15**, 200–202.
14. Fauvarque, M. O. & Dura, J. M. (1993) *Genes Dev.* **7**, 1508–1520.
15. Tautz, D. & Pfeifle, C. (1989) *Chromosoma* **98**, 81–85.
16. Ghosh, S. & Lowenstein, J. M. (1996) *Gene* **176**, 249–255.
17. Lin, K.-H. & Cheng, S.-Y. (1991) *BioTechniques* **11**, 748–754.
18. Huq, A. H., Lovell, R. S., Sampson, M. J., Decker, W. K., Dinulos, M. B., Distech, C. M. & Craig, W. J. (1996) *Genomics* **36**, 530–534.
19. Greenspan, D. L., Connolly, D. C., Wu, R., Lei, R. Y., Vogelstein, J. T., Kim, Y. T., Mok, J. E., Munoz, N., Bosch, F. X., Shah, K. & Cho, K. R. (1997) *Cancer Res.* **57**, 4692–4698.
20. Sozzi, G., Tornielli, S., Tagliabue, E., Sard, L., Pezzella, F., Pastorino, U., Minoletti, F., Pilotti, S., Ratcliffe, C., Veronese, M. L., *et al.* (1997) *Cancer Res.* **57**, 5207–5212.
21. Hadaczek, P., Siprashvili, Z., Markiewsky, M., Domagala, W., Druck, T., McCue, P., A., Pekarsky, Y., Ohta, M., Huebner, K. & Lubinski, J. (1998) *Cancer Res.*, in press.
22. Davidson, J. N. & Peterson, M. L. (1997) *Trends Genet.* **13**, 281–285.
23. Davidson, J. N., Chen, K. C., Jamison, R. S., Musmanno, L. A. & Kern, C. B. (1993) *BioEssays* **15**, 157–164.
24. Koljak, R., Boutaud, O., Shieh, B., Samel, N. & Brash, A. R. (1997) *Science* **277**, 1994–1996.
25. Bartling, D., Seedorf, M., Schmidt, R. C. & Weiler, E. W. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6021–6025.
26. Marais, R., Light, Y., Paterson, H. F. & Marshall, C. J. (1995) *EMBO J.* **13**, 3136–3145.