MOLECULAR AND CELLULAR BIOLOGY, Apr. 1989, p. 1587–1593
0270-7306/89/041587-07$02.00/0

Vol. 9, No. 4

Isolation and Sequence Analysis of a Novel Human Tyrosine Kinase Gene

QIAN-LIN HAO, NORA HEISTERKAMP, AND JOHN GROFFEN*

Division of Medical Genetics, Section of Molecular Genetics, Childrens Hospital of Los Angeles, Los Angeles, California 90027

Received 3 October 1988/Accepted 22 December 1988

Using v-*abl* probes, we have identified and cloned a novel *fes/fps*-homologous human cDNA, which we have designated *FER* (pronounced "fair"). This apparently full-length cDNA of 3.0 kilobases has an open reading frame of 2,466 base pairs and the capacity to encode a protein of 94,000 molecular weight. The cDNA contains regions homologous to the highly conserved tyrosine protein kinase domain of other oncogenes and growth factor receptors but lacks a clear transmembrane region, indicating that it encodes a tyrosine kinase of the nonreceptor type. The deduced amino acid sequence of *FER* resembles that of c-*fes/fps*. Our data indicate that the protein product of *FER*, p94^{FER}, corresponds to a previously reported cellular phosphoprotein, NCP94, detected with a v-*fps*-specific antipeptide antiserum.

During the past few years, a number of genes that encode proteins with protein kinase activity specific for tyrosine have been identified. The first tyrosine-specific protein kinase genes, such as *src*, were isolated as viral transforming genes transduced by type C RNA-transforming viruses; subsequently, many viral oncogenes were found to contain a conserved domain encoding a protein with tyrosine kinase activity (12).

Some of the tyrosine kinase-encoding oncogenes are homologous to receptors for growth factors: the epidermal growth factor receptor is homologous to the *erbB* oncogene (27), whereas the *fms* oncogene represents the receptor for colony-stimulating factor 1 (20). Other growth factor receptors, such as those for insulin and the insulinlike growth factor, have no known oncogene homolog but share the same highly conserved tyrosine kinase domain. A different class of tyrosine kinase-encoding oncogenes, although clearly possessing the conserved domain, do not appear to be homologous to receptors, since they lack an extracellular and transmembrane domain. Members of this class include *src*, *fes/fps*, *yes*, and *abl*.

The human *ABL* oncogene encodes a tyrosine kinase of 145,000 molecular weight for which, as for the other tyrosine kinases of its class, the function in unknown. *ABL* is specifically involved in human leukemia: in chronic myelocytic leukemia, the *ABL* gene is translocated to chromosome 22 from its normal location on chromosome 9 (3) and becomes "activated" (22, 25). More recently (10), *ABL* was also found to be involved in a subclass of acute lymphoblastic leukemia.

Because the involvement of *ABL* in certain types of human neoplasia seems well established, it was of interest to examine other *ABL*-related sequences in humans. Previous experiments demonstrated the existence of such sequences in the human genome (8); one of these genes has been partially cloned and has been designated *ARG* (16). In the study reported here, we isolated and characterized a novel *fes/fps*-homologous human cDNA by using a v-*abl* probe under hybridization conditions of relatively low stringency.

* Corresponding author.

MATERIALS AND METHODS

cDNA cloning. Approximately 40,000 recombinants of a previously described human fibroblast cDNA library (9) were plated onto 10 plates of L broth containing ampicillin (50 μg/ml) and grown at 37°C. Colonies from each plate were collected, amplified, and used to isolate 10 plasmid DNAs. Of each plasmid preparation, 10 μg was digested with *Sal*I under conditions recommended by the supplier; the linearized plasmid was run on a 0.6% agarose gel. The presence of v-*abl*-related sequences was analyzed by Southern blotting (8) and hybridization. *ABL*-related sequences detected on the Southern blot were isolated by gel fractionation of a new portion of linearized plasmid DNA on a low-melting-point agarose gel (International Biotechnologies Inc.) followed by purification of the appropriate fraction of plasmid DNA from the agarose (8). Plasmid DNA was recircularized by ligation and used to transform *Escherichia coli* DH5α (Bethesda Research Laboratories). Screening of the resulting size-selected partial cDNA library was done as described elsewhere (9).

Nucleotide sequence analysis. Nucleotide sequences on both DNA strands were determined by the dideoxy-chain termination method (19). The FASTP program (12) was used to compare the deduced amino acid sequence with those of known proteins.

Southern blot and Northern (RNA) blot analyses. High-molecular-weight DNAs or plasmid DNAs were digested with restriction endonucleases and fractionated by agarose gel electrophoresis. Upon blotting to nitrocellulose (Schleicher & Schuell), blots were hybridized essentially as described previously (8). Total RNA was isolated as described elsewhere (25). Ten micrograms of total RNA was electrophoresed on a 1% agarose gel in the presence of formaldehyde; blotting and hybridization were done as described previously (25).

Immunoprecipitation. A 660-base-pair *Eco*RI-*Pvu*II fragment (Fig. 1) corresponding to amino acid residues 148 to 369 (see Fig. 3) was inserted into the appropriate *trpE* expression vector (24), and a 63,000-dalton *trpE* fusion protein was synthesized in *E. coli*. Antiserum (CH-6) was obtained from immunized rabbits (26). Immune complex kinase assays were performed essentially as described elsewhere (18).

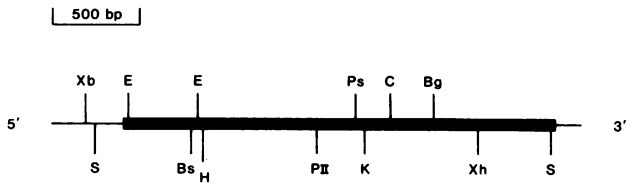

FIG. 1. Restriction enzyme map of *FER* cDNA. The thick line indicates coding sequences, and the thin line represents 5′ and 3′ untranslated regions. Restriction enzymes used were *Bgl*II (Bg), BstEII (Bs), *Cla*I (C), *Eco*RI (E), *Kpn*I (K), *Pst*I (Ps), *Pvu*II (PII), *Sst*I (S), *Xba*I (Xb), and *Xho*I (Xh).

V8 protease digestion. V8 protease cleavage and analysis of the digestion products were performed as described elsewhere (28). Briefly, in vitro phosphate-labeled polypeptides were separated on a 7.5% sodium dodecyl sulfate-polyacrylamide gel, identified by autoradiography, excised, and applied to the wells of a 12.5% sodium dodecyl sulfate-polyacrylamide gel.

Phosphoamino acid analysis of p94^{FER}. In vitro ^{32}P-labeled proteins were eluted from gel slices, precipitated with trichloroacetic acid, and subjected to acid hydrolysis in 6 N HCl at 110°C for 90 min as described elsewhere (26). Hydrolysates were applied to thin-layer cellulose plates (10 by 10 cm) and electrophoresed toward the anode at the top for 3.0 h at 400 V in TLE buffer (pyridine, acetic acid, water [5:50:945]). Phosphoamino acid standards were detected by ninhydrin staining. Phosphoamino acid composition of the hydrolysates was analyzed by autoradiography.

RESULTS

Isolation and sequencing of a cDNA clone for the *FER* gene. To isolate *ABL*-homologous sequences, we selected a previously described (8) v-*abl* 1.6-kilobase (kb) *Pst*I fragment as a probe, since this fragment contains the domain well conserved among members of the tyrosine kinase gene family. Of the numerous v-*abl*-hybridizing fragments present in the cDNA library, two large clones were selected for further characterization; the cDNA clones were isolated as described in Materials and Methods. Restriction enzyme maps were constructed for both clones; one clone has an insert of 5.0 kb, and the other had an insert of 3.0 kb. Upon comparison of the restriction enzyme maps of these clones with those published previously, it became apparent that the largest (5.0-kb) cDNA clone was essentially identical to a cDNA encoding the insulinlike growth factor receptor (17); partial sequence analysis of this clone confirmed its identity (results not shown). The 3.0-kb cDNA clone, however, was distinct on the restriction enzyme level from human *ABL* and from any other previously isolated human proto-oncogene or growth factor receptor cDNA clone (Fig. 1). These results suggest that this clone represents a novel *ABL*-related gene, which we have named *FER* (pronounced "fair"). Upon hybridization of the entire *FER* cDNA to a Southern blot with human DNA, numerous restriction enzyme fragments were detected (results not shown) after relatively stringent posthybridization washings (0.1× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate] at 65°C). After very stringent washings (0.03× SSC at 65°C), only four hybridizing bands, which probably represented *FER* gene genomic sequences, remained visible in a *Bam*HI digest. The presence of other hybridizing bands after relatively stringent washings (0.1× SSC at 65°C) indicates that the human genome may contain *FER*-related sequences; in view of the limited relatedness of *FER* and *ABL* (see below) and the stringency of washing, these related sequences cannot belong to the *ABL* locus. Further characterization of the *FER* locus will have to await molecular cloning of its genomic sequences. To further characterize the *FER* gene, individual restriction enzyme fragments of the cDNA were inserted into M13 mp18 and mp19, and the cDNA was subjected to complete nucleotide sequence analysis for both strands (Fig. 2).

Product of the human *FER* gene. The cDNA was 3,002 base pairs in length; the longest open reading frame started with an ATG codon at position 385 in the nucleotide sequence, and the sequence flanking it was in agreement with the consensus sequence of Kozak (15). Multiple stop codons in all three reading frames were found immediately 5′ of the proposed translational start ATG codon at position 385. The predicted coding sequence of 2,466 base pairs is flanked by 5′ and 3′ untranslated sequences of 384 and 142 base pairs, respectively. This large open reading frame has the capacity to encode a polypeptide of 822 amino acids with a calculated molecular weight of 94,600. Characteristic sequences common to tyrosine protein kinase genes, such as the consensus sequence for the ATP binding site Gly-X-Gly-X-X-Gly-X15-K (12, 14), were conserved in the FER protein and were found at amino acid residues 571 to 592 (Fig. 3). Moreover, the Tyr-416 phosphorylation site of pp60^{src} (23) was also conserved in the FER protein at amino acid residue position 714. A stretch of hydrophobic amino acids typical for a transmembrane region was not found in the FER protein, suggesting that this protein does not represent a transmembrane receptor. As in pp60^{src}, the first methionine of the FER protein was immediately followed by a glycine; in pp60^{src} this glycine has been found to be myristylated and is thought to anchor the protein to the cell membrane (2, 13). By analogy, the FER protein may be bound to the plasma membrane. However, the glycine present in the FER protein does not necessarily indicate that myristylation and subsequent plasma membrane binding occur in vivo: the c-*fps/fes* protein contains an N-terminal glycine; nonetheless, 60 to 90% of the protein is found in the soluble fraction (29).

Homology to other tyrosine protein kinase genes. To determine the extent of relatedness of *FER* to other tyrosine protein kinases, its deduced amino acid sequence was compared by using the FASTP program (17). Surprisingly, the highest degree of homology observed was to the chicken c-*fps/fes* (11) proto-oncogene: although the *FER* cDNA was isolated by using a v-*abl* probe, computer homology searches revealed that *FER* has significantly higher homology to c-*fes/fps*; therefore, we propose the name *FER* for this *fps/fes*-related gene. The c-*fes/fps* proto-oncogene is the cellular homolog of v-*fes* and v-*fps* as found in Sneider-Theilen feline sarcoma virus from cats (7) and Fujinami sarcoma virus from chickens (21). Previously, it was demonstrated that both v-*fes* and v-*fps* have a common cellular homolog in humans, the *FES/FPS* proto-oncogene (6). The 822-amino-acid sequence deduced from the cDNA specifies a protein closely related to, but distinct from, the tyrosine protein kinase encoded by c-*fes/fps* (51% homology) (Fig. 3). The homology is most pronounced at the carboxy-terminal region of the protein. Since the tyrosine kinase activity is encoded within the carboxy-terminal domain of pp60^{src}, *fps*, and other related oncogenes, the FER protein would be expected to exhibit tyrosine kinase activity. The relatively low degree of homology found at the amino terminus suggests that this part of the protein is responsible for the specific functions of the protein. Table 1 summarizes these

```
GTCACACCCTCGAATAATGACGCATACCTATCCTACTGTTACTGATCACCTAGTAATAATCTTGTAGTTCACATTACTCATTTTTCACCAAATTCTTTTGGTGAAGGACGCTTCAGAAAC  120

GGCCATCACTGAAGAGCAGACCCGTTTGGGTTCTCCACGCATTCTAGACTCCCGAAGAGCTCATGTTTTTTGGCTAGACCTATGACCATTTTCGCTAGACTTCACTGCACGTTTTCTCAA  240

GTATCTTCTTTGTCCCTAATGTGTGACACCTCATCATGGACACGCTACTTTAGCTAAGGCATGACCAGCAATGAACAGTAGTAAGATATGTGCTGATTAGAAGGCTCACTTGTGCAGTGT  360

GGAGGATAACCAGTGCCTTACAAAATGGGGTTTGGGAGTGACCTGAAGAATTCACATGAAGCAGTGTTAAAATTGCAAGACTGGGAATTACGGTTACTGGAAACAGTAAAGAAATTTATG  480
                          M  G  F  G  S  D  L  K  N  S  H  E  A  V  L  K  L  Q  D  W  E  L  R  L  L  E  T  V  K  K  F  M
GCCCTGAGAATAAAAAGTGATAAAGAATATGCATCTACTTTACAGAACCTTTGTAATCAAGTTGATAAGGAAAGTACTGTCCAAATGAATTATGTCAGCAACGTATCCAAGTCTTGGCTA  600
 A  L  R  I  K  S  D  K  E  Y  A  S  T  L  Q  N  L  C  N  Q  V  D  K  E  S  T  V  Q  M  N  Y  V  S  N  V  S  K  S  W  L
CTTATGATTCAGCAGACAGAACAACTTAGTAGGATAATGAAGACACATGCAGAGGACTTGAACTCTGGACCTTTACACAGGCTCACCATGATGATTAAGGACAAGCAGCAGGTGAAGAAA  720
 L  M  I  Q  Q  T  E  Q  L  S  R  I  M  K  T  H  A  E  D  L  N  S  G  P  L  H  R  L  T  M  M  I  K  D  K  Q  Q  V  K  K
AGTTACATAGGTGTTCATCAGCAGATAGAGGCAGAGATGATCAAGGTTACCAAAACAGAATTGGAGAAGTTAAAATGCAGCTATAGACAATTAATAAAAGAAATGAATTCTGCCAAAGAG  840
 S  Y  I  G  V  H  Q  Q  I  E  A  E  M  I  K  V  T  K  T  E  L  E  K  L  K  C  S  Y  R  Q  L  I  K  E  M  N  S  A  K  E
AAATATAAAGAAGCTTTAGCTAAAGGGAAGGAAACTGAAAAGGCCAAGGAACGATACGACAAAGCCACAATGAAACTTCATATGTTGCACAATCAGTATGTATTGGCGTTGAAAGGGGCA  960
 K  Y  K  E  A  L  A  K  G  K  E  T  E  K  A  K  E  R  Y  D  K  A  T  M  K  L  H  M  L  H  N  Q  Y  V  L  A  L  K  G  A
CAGCTCCATCAGAATCAGTATTATGATATCACACTTCCCCTGCTTCTGGACTCCTTACAAAAGATGCAAGAAGAAATGATAAAAGCACTCAAAGGTATATTTGATGAATACAGCCAGATA 1080
 Q  L  H  Q  N  Q  Y  Y  D  I  T  L  P  L  L  L  D  S  L  Q  K  M  Q  E  E  M  I  K  A  L  K  G  I  F  D  E  Y  S  Q  I
ACCAGTCTTGTCACAGAGGAAATAGTGAATGTCCATAAAGAGATTCAAATGTCGGTTGAACAGATAGATCCTAGTACAGAATACAATAATTTCATAGATGTTCACAGAACAACGGCTGCT 1200
 T  S  L  V  T  E  E  I  V  N  V  H  K  E  I  Q  M  S  V  E  Q  I  D  P  S  T  E  Y  N  N  F  I  D  V  H  R  T  T  A  A
AAAGAACAAGAAATAGAGTTTGATACTTCCTTACTGGAAGAAAATGAAAATCTTCAGGCAAATGAGATCATGTGGAATAACTTAACAGCAGAAAGTTTGCAAGTAATGTTGAAAACGTTA 1320
 K  E  Q  E  I  E  F  D  T  S  L  L  E  E  N  E  N  L  Q  A  N  E  I  M  W  N  N  L  T  A  E  S  L  Q  V  M  L  K  T  L
GCGGAAGAACTTATGCAAACACAGCAGATGCTTTTAAACAAGGAGGAGGCTGTTTTGGAGTTAGAGAAGAGAATTGAAGAATCTTCTGAAACTTGTGAGAAGAAGTCTGATATTGTGCTT 1440
 A  E  E  L  M  Q  T  Q  Q  M  L  L  N  K  E  E  A  V  L  E  L  E  K  R  I  E  E  S  S  E  T  C  E  K  K  S  D  I  V  L
CTGCTAAGCCAAAAACAGGCACTTGAAGAACTGAAACAGTCAGTCCAGCAGCTGAGATGCACTGAAGCAAAGTTTTCAGCACAGAAGGAATTACTAGAGCAAAAAGTGCAAGAAAATGAT 1560
 L  L  S  Q  K  Q  A  L  E  E  L  K  Q  S  V  Q  Q  L  R  C  T  E  A  K  F  S  A  Q  K  E  L  L  E  Q  K  V  Q  E  N  D
GGGAAAGAGCCACCTCCAGTAGTAAATTATGAAGAAGATGCACGATCAGTTACATCTATGGAAAGAAAGGAGAGGCTATCCAAATTTGAATCTATTCGTCATTCAATTGCTGGAATAATT 1680
 G  K  E  P  P  P  V  V  N  Y  E  E  D  A  R  S  V  T  S  M  E  R  K  E  R  L  S  K  F  E  S  I  R  H  S  I  A  G  I  I
AGGTCTCCAAAATCTGCAGTGGGCTCTTCAGCACTTTCTGATATGATCTCCATCAGTGAGAAGCCTTTGGCAGAACAGGACTGGTACCATGGTGCAATTCCCAGAATAGAAGCTCAAGAA 1800
 R  S  P  K  S  A  V  G  S  S  A  L  S  D  M  I  S  I  S  E  K  P  L  A  E  Q  D  W  Y  H  G  A  I  P  R  I  E  A  Q  E
CTGTTAAAAAAACAAGGAGACTTTTTGGTGCGAGAGAGTCATGGGAAACCTGGTGAATATGTCCTTTCTGTATATTCTGATGGACAGAGGAGACATTTTATCATACAATATGTTGATAAC 1920
 L  L  K  K  Q  G  D  F  L  V  R  E  S  H  G  K  P  G  E  Y  V  L  S  V  Y  S  D  G  Q  R  R  H  F  I  I  Q  Y  V  D  N
ATGTATCGATTCGAGGGCACTGGGTTTTCAAACATTCCTCAACTTATAGATCATCACTATACAACAAAACAGGTCATCACTAAGAAATCAGGTGTAGTTCTGCTGAATCCTATTCCTAAG 2040
 M  Y  R  F  E  G  T  G  F  S  N  I  P  Q  L  I  D  H  H  Y  T  T  K  Q  V  I  T  K  K  S  G  V  V  L  L  N  P  I  P  K
GACAAGAAATGGATTCTCAGTCATGAAGATGTCATATTGGGAGAATTACTGGGCAAGGGAAATTTTGGTGAAGTATATAAGGGCACATTAAAGGATAAAACTTCTGTTGCTGTTAAAACA 2160
 D  K  K  W  I  L  S  H  E  D  V  I  L  G  E  L  L  G  K  G  N  F  G  E  V  Y  K  G  T  L  K  D  K  T  S  V  A  V  K  T
TGTAAAGAAGATCTTCCTCAGGAATTGAAAATAAAATTTTTACAAGAAGCCAAAATTCTCAAGCAATATGATCATCCCAATATTGTCAAACTTATAGGAGTTTGCACACAAAGACAGCCT 2280
 C  K  E  D  L  P  Q  E  L  K  I  K  F  L  Q  E  A  K  I  L  K  Q  Y  D  H  P  N  I  V  K  L  I  G  V  C  T  Q  R  Q  P
GTCTACATCATTATGGAACTGGTTTCAGGAGGTGATTTCCTCACCTTTCTGAGAAGGAAGAAGGATGAACTAAAACTCAAACAGTTAGTGAAATTTTCATTAGACGCTGCTGCTGGTATG 2400
 V  Y  I  I  M  E  L  V  S  G  G  D  F  L  T  F  L  R  R  K  K  D  E  L  K  L  K  Q  L  V  K  F  S  L  D  A  A  A  G  M
TTGTATCTCGAGAGTAAAAACTGTATACACAGGGACCTTGCTGCAAGAAACTGCCTGGTAGGTGAAAATAATGTTCTGAAAATCAGTGACTTTGGAATGTCTCGTCAAGAGGATGGTGGA 2520
 L  Y  L  E  S  K  N  C  I  H  R  D  L  A  A  R  N  C  L  V  G  E  N  N  V  L  K  I  S  D  F  G  M  S  R  Q  E  D  G  G
GTGTATTCATCTTCTGGCTTAAAGCAGATTCCCATTAAATGGACAGCACCGGAAGCTCTTAATTATGGGAGATACAGTTCAGAGAGTGACGTGTGGAGCTTTGGCATCCTTCTCTGGGAG 2640
 V  Y  S  S  S  G  L  K  Q  I  P  I  K  W  T  A  P  E  A  L  N  Y  G  R  Y  S  S  E  S  D  V  W  S  F  G  I  L  L  W  E
ACCTTCAGCTTAGGGGTTTGTCCGTACCCTGGAATGACAAATCAGCAAGCAAGAGAGCAAGTAGAAAGAGGATACCGGATGTCAGCTCCCCAGCACTGTCCAGAGGATATTTCCAAAATC 2760
 T  F  S  L  G  V  C  P  Y  P  G  M  T  N  Q  Q  A  R  E  Q  V  E  R  G  Y  R  M  S  A  P  Q  H  C  P  E  D  I  S  K  I
ATGATGAAGTGTTGGGATTATAAACCTGAAAATCGCCCTAAGTTCAGTGAACTTCAGAAAGAGCTCACTATCATCAAGAGAAAACTCACATAGTGACAGGATGGCGCCAAACTCAGCCTT 2880
 M  M  K  C  W  D  Y  K  P  E  N  R  P  K  F  S  E  L  Q  K  E  L  T  I  I  K  R  K  L  T  -

CAGGACTCTGTCCTCCAGCAGAGTAACATTATTGTTCTCATTAACAATGAATTTATACCACATTACCTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3000
```

FIG. 2. Nucleotide and amino acid sequence analysis of *FER*. The sequence has been deposited in the EMBL/GenBank data base (accession no. J03358).

similarities, which are most pronounced within the domain containing the enzymatic activity of tyrosine protein kinases.

mRNA expression of *FER*. Total RNA was isolated from various human cell lines. Blot-hybridization analysis of these RNAs with a 0.4-kb *Eco*RI fragment (Fig. 1) as the probe revealed the presence of a 3.0-kb transcript in the cell lines K562 (chronic myelogenous leukemia) and A172 (human glioblastoma) (Fig. 4), in a bladder carcinoma cell line (Scaber), and in an acute lymphocytic leukemia cell line (CCRF-HSB_2) (data not shown). Surprisingly, in addition to the 3.0-kb *FER* mRNA, strong hybridization signals around

```
Fer          1 MGFGSDL--KNSHEAVLKLQDWELRLLETVKKFMALRIKSDKEYASTLQNLCNQVDKEST
               ::: : :     :      : ::::::   : :: : ::: :::  :     :
c-Fps/Fes    1 MGFSSELCSPQGHGVLQQMQEAELRLLEGMRKWMAQRVKSDREYAGLLHHMSLQDSGGQS

Fer         59 VQMNYVSNVSKSWLLMIQQTEQLSRIMKTHAEDLNSGPLHRLTMMIKDKQQVKKSYIGVH
                 : : ::    ::: :::    ::::::::::  :   :   ::  : :
c-Fps/Fes   61 RAISPDSPISQSWAEITSQTEGLSRLLRQHAEDLNSGPLSKLSLLIRERQQLRKTYSEQW

Fer        119 QQIEAEMIKVTKTELEKLKCSYRQLIKEMNSAKEKYKEALAKGKETEKAKERYDKATMKL
               ::   : :      :::: :: :      :: :: ::  : :   ::: :     ::
c-Fps/Fes  121 QQLQQELTKTHSQDIEKLKSQYRALARDSAQAKRKYQEA-SKDKDRDKAKDKYVRSLWKL

Fer        179 HMLHNQYVLALKGAQLHQNQYYDITLPLLLDSLQKMQEEMIKALKGIFDEYSQITSLVTE
                 :: :::    ::::        :: :: :::  :::   :: :  ::  : :::
c-Fps/Fes  180 FAHHNRYVLGVRAAQLHHQHHHQLLLPGLLRSLQDLHEEMACILKEILQEYLEISSLVQD

Fer        239 EIVNVHKEIQMSVEQIDPSTEYNNFIDVHRTTAAKEQEIEFDTSLLEENENLQANEIMWN
               : :  : :      : :  ::  :                :: ::::: : :   :  :
c-Fps/Fes  240 EVVAIHREMAAAAARIQPEAEYQGFLRQYGSAPDVPPCVTFDESLLEEGEPLEPGELQLN

Fer        299 NLTAESLQVMLKTLAEELMQTQQMLLNKEEAVLELEKRIEESSETCEKKSDIVLLLSQKQ
               :: :: :  :    ::     :    : :  :          :        : ::   :
c-Fps/Fes  300 ELTVESVQHTLTSVTDELAVATEMVFRRQEMVTQLQQELRNEEENTHPR-ERVQLLGKRQ

Fer        359 ALEELKQSVQQLRCTEAKFSAQKELLEQKVQENDGKEPPPVVNYEEDARSVTSME--RK-
               : :  :  :   :  :: :: :::  :       :::::     :  :  : :  :
c-Fps/Fes  359 VLQEALQGLQVALCSQAKLQAQQELLQTKLEHLGPGEPPPVLLLQDDRHSTSSSEQEREG

Fer        416 ERLSKFESIRHSIAGIIRSPKSAVGSSALSDMISISEKPLAEQDWYHGAIPRIEAQELLK
               :    :     : :: : ::       :   :    ::: :: ::::::::: :  :::
c-Fps/Fes  419 GRTPTLEILKSHISGIFR-PKFSL-PPPL-QLIPEVQKPLHEQLWYHGAIPRAEVAELLV

Fer        476 KQGDFLVRESHGKPGEYVLSVYSDGQRRHFIIQYVDNMYRFEGTGFSNIPQLIDHHYTTK
                 :::::::: ::  :::::::  ::  :::::::  :: :: :: ::  :: ::::   :
c-Fps/Fes  476 HSGDFLVRESQGKQ-EYVLSVLWDGLPRHFIIQSLDNLYRLEGEGFPSIPLLIDHLLSTQ

Fer        536 QVITKKSGVVLLNPIPKDKKWILSHEDVILGELLGKGNFGEVYKGTL-KDKTSVAVKTCK
               :  ::::::::    ::: :: : :::  :::  : :::::::  : :  : : :::: :
c-Fps/Fes  535 QPLTKKSGVVLHRAVPKD-KWVLNHEDLVLGEQIGRGNFGEVFSGRLRADNTLVAVKSCR

Fer        595 EDLPQELKIKFLQEAKILKQYDHPNIVKLIGVCTQRQPVYIIMELVSGGDFLTFLRRKKD
               : ::  :: :::::: ::::: ::::: :::::::: :: :: :::: :::::::::
c-Fps/Fes  594 ETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIVMELVQGGDFLTFLRTEGA

Fer        655 ELKLKQLVKFSLDAAAGMLYLESKNCIHRDLAARNCLVGENNVLKISDFGMSRQEDGGVY
               :  : :     :::::: ::::: :::::::::::::: : :::::::::::: :  :::
c-Fps/Fes  654 RLRVKTLLQMVGDAAAGMEYLESKCCIHRDLAARNCLVTEKNVLKISDFGMSREEADGVY

Fer        715 -SSSGLKQIPIKWTAPEALNYGRYSSESDVWSFGILLWETFSLGVCPYPGMTNQQAREQV
                 : :: : : :::::::::::::::::::::::::::::::::  :::   ::: :: :
c-Fps/Fes  714 AASGGLRQVPVKWTAPEALNYGRYSSESDVWSFGILLWETFSLGASPYPNLSNQQTREFV

Fer        774 ERGYRMSAPQHCPEDISKIMMKCWDYKPENRPKFSELQKELTIIKRKLT*
               : : :   :  ::     :  :: : :  :: ::    ::  :
c-Fps/Fes  774 EKGGRLPCPELCPDAVFRLMEQCWAYEPGQRPSFSTIYQELQSIRKRHR*
```

FIG. 3. Comparison of the amino acid sequences of *FER* and *FES/FPS*. Symbols: :, amino acid residues showing complete identity; ★, termination codon at the carboxy terminus of the protein. The underlined amino acids in *FES/FPS* refer to a synthetic peptide used to generate antisera (23).

8.0 kb were observed in all the cell lines. This transcript must be highly related to *FER*, as it remained visible after high-stringency posthybridization washings. As discussed above, Southern blot analysis with *FER* cDNA probes indicated that the human genome contains sequences highly related to the *FER* gene. The 8.0-kb mRNA might represent the transcript of a gene highly homologous to *FER*. However, we cannot completely eliminate the possibility that it represents an authentic transcript from the *FER* locus; transcription of other tyrosine kinase genes, such as *ABL*, can result in two different-sized mRNAs from one locus (1).

Identification of the p94^{FER} protein. Since *FER* is homologous to other tyrosine kinase genes, many of which exhibit intrinsic autophosphorylation activity in vitro, we sought to determine whether the FER protein could be detected by an in vitro kinase assay. Protein from unlabeled cell extracts from human cell lines or the mouse cell line NIH 3T3 were immunoprecipitated by rabbit anti-TrpE-FER fusion protein antiserum (CH-6; see Materials and Methods), and immunoprecipitates were assayed for protein kinase activity. The region of FER (amino acids 148 to 369 [Fig. 3]) used to obtain the TrpE fusion protein corresponded to a domain which

TABLE 1. Similarities among tyrosine protein kinases

FER protein		No. of amino acids overlapping	% Identical amino acids	
Gene	Source		Overall	In kinase domain[a]
FPS/FES	Human	822	49	69
c-*fps*	Chicken	822	51	70
v-*fps*	Fujinami sarcoma virus	822	51	70
v-*fes*	Feline sarcoma virus	535	53	63
v-*abl*	Abelson murine leukemia virus	378	39	51
c-*src*	Chicken	533	26	49
HIR	Human	ND[b]	ND	38

[a] Amino acids 561 to 808 of the FER protein.
[b] ND, Not determined.

was less closely related to c-FPS/FES (Fig. 3); therefore we expected little or no cross-reactivity of the CH-6 antiserum with the p92 c-FPS/FES protein. Immunoprecipitates of the FER protein were able to catalyze the transfer of $^{32}PO_4$ from [γ-^{32}P]ATP to FER in an apparent autophosphorylation reaction: a phosphoprotein with a molecular weight of approximately 94,000 was immunoprecipitated in all cell lines tested (Fig. 5). Upon treatment of the gel with KOH (5), the p94^{FER} protein remained visible, whereas most of the nonspecifically precipitated proteins disappeared (data not shown). The specificity of the CH-6 antiserum for the p94 protein was demonstrated in experiments performed simultaneously with two distinct, irrelevant antisera, and p94 was not detected in cell lines A498 and K562 (results not shown). The level of expression of p94 varied in different cell lines; a small amount of p94 was detected in HL-60 (Fig. 5, lane 5; longer exposure of the gel confirmed the presence of the p94), a promyelocytic leukemia cell line, whereas p94^{FER} could be more readily detected in a human B-lymphocyte cell line (DAKIKI) (lane 1) and in K562 (lane 4). Examination of the human fibroblastic cell line GM 3377A (Fig. 5, lane 2) and the mouse cell line NIH 3T3 (lane 3) revealed relatively high levels of expression. We also found high levels of p94^{FER} in some tumor cell lines, such as human kidney carcinoma A498 (Fig. 5, lane 6) and glioblastoma A172 (data not shown).

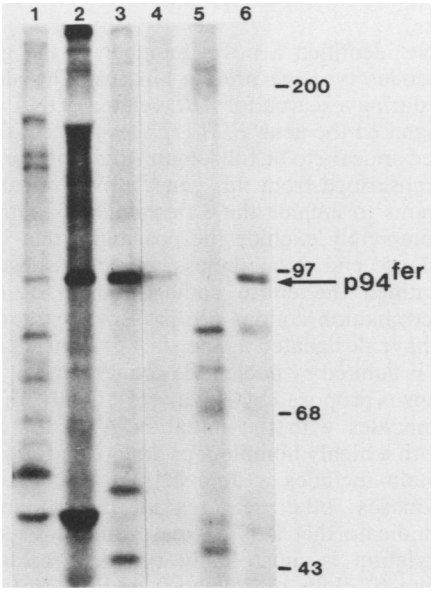

FIG. 5. Identification of the p94^{FER} protein in human and mouse cell lines. Lane 1, DAKIKI; lane 2, GM3377A; lane 3, NIH 3T3 mouse fibroblast; lane 4, K562; lane 5, HL-60; lane 6, A498. Cells were lysed and immunoprecipitated with CH-6 antiserum. Immune complexes were incubated with [γ-^{32}P]ATP, and in vitro-phosphorylated proteins were identified by electrophoretic separation and autoradiography. Exposure time, 4 h. Molecular size markers are in kilodaltons.

Phosphoamino acid analysis of p94^{FER}. The identity of the amino acids serving as phosphate acceptors in the p94^{FER} kinase assay was determined by acid hydrolysis of in vitro-phosphorylated p94^{FER} followed by electrophoresis through thin-layer cellulose plates. The amino acid acceptor of phosphate in p94^{FER} was found to be tyrosine (Fig. 6). Traces of phosphothreonine and phosphoserine were observed upon longer exposure of the plates (data not shown).

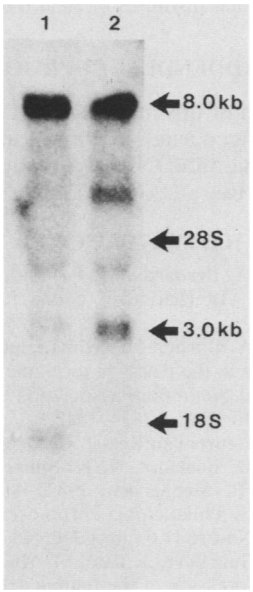

FIG. 4. Northern blot analysis of *FER* mRNA. Lane 1, A172 RNA; lane 2, K562 RNA. The positions of 28S and 18S rRNAs are indicated. Exposure time, 48 h.

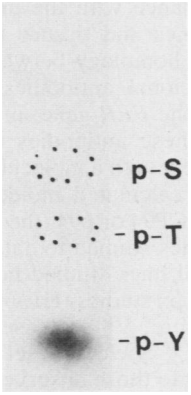

FIG. 6. Phosphoamino acid analysis of p94^{FER}. Abbreviations: p-S, phosphoserine; p-T, phosphothreonine; p-Y, phosphotyrosine. Exposure time, 24 h.

DISCUSSION

We have identified a novel member of the gene family which encodes tyrosine protein kinases. The identification occurred during a search for *ABL*-related cDNA clones. We have designated the gene *FER* (*fps/fes*-related). The cDNA we isolated appears to be full length when compared with the mRNA transcribed from this gene. However, although the cDNA seems to include the 3′ terminus of the mRNA, we cannot completely exclude the possibility that it is incomplete at the 5′ end. Nonetheless, our data suggest that we have identified the entire coding domain of *FER*: first, several termination codons are present upstream from the ATG we have designated as the start of translation; second, this ATG is flanked by nucleotide sequences that conform to the consensus proposed by Kozak (15); finally, this initiation codon coincides with the initiation ATG for c-*fps/fes*, a protein with a highly homologous amino acid sequence. The FER protein includes a domain that is characteristic of protein kinases; other features in the protein sequence strongly indicate that it is a kinase and that its target for phosphorylation is most likely tyrosine residues. These features are (i) close resemblance of FER to the c-*fps/fes* protein demonstrated to be a tyrosine-specific kinase, (ii) a tyrosine residue at position 714 that is characteristic for tyrosine protein kinases and is subject to autophosphorylation, (iii) a stretch of amino acids that signal the location of a binding site for ATP, and (iv) an amino-terminal sequence similar to that required for myristylation and membrane anchoring of *src*-related proteins. Despite its resemblance to other tyrosine protein kinases, a FER cDNA has not previously been identified. However, it is not unlikely that the FER protein is identical to a recently detected cellular phosphoprotein with a molecular weight of 94,000 (p94, or NCP94) (4, 18); apart from the similarity in molecular weight, this protein also exhibits associated tyrosine-specific protein kinase activity; moreover, it was detected by antibodies directed against an amino acid sequence (K-Q-I-P-V-K-W-T-A-P-E-A) of the conserved domain of the v-*fps* transforming protein (Fig. 3). In addition to immunoprecipitating the c-*fes/fps* protein, this antiserum immunoprecipitated a phosphorylated 94-kilodalton protein, which appears to be present in all cells. Because this antiserum detects the c-*fes/fps* protein (which contains the amino acid sequence R-Q-V-P-V-K-W-T-A-P-E-A), it seems not unlikely that the same antibodies could recognize the sequence K-Q-I-P-I-K-W-T-A-P-E-A of the FER protein. Moreover, the expression of FER mRNA in four cell lines of different origins is in concordance with the universal expression of NCP94 in all cell lines and tissues tested (5). To further analyze the possible homology between these proteins, we have obtained polyclonal antibodies directed against the protein encoded by the *FER* gene; in an in vitro autophosphorylation assay, these antibodies immunoprecipitated a phosphorylated protein with a molecular weight of 94,000, in agreement with the calculated molecular weight of FER. Moreover, as with NCP94 or p94, the $p94^{FER}$ protein can be detected in all cell lines studied to date, and the levels of its expression in the cell lines studied here are comparable to those of p94 in previous studies (HL-60, K562, and NIH 3T3 cells) (18). Furthermore, V8 protease digestion of in vitro-phosphorylated $p94^{FER}$ yielded a set of cleavage products (Fig. 7) highly similar to those observed by MacDonald et al. (18) after digestion of p94 with V8 protease. Finally, the finding that $p94^{FER}$ was phosphorylated in vitro on tyrosine with traces of phosphothreonine and phosphoserine leads us to conclude that the *FER* gene isolated here encodes the previously identified NCP94 or p94 protein. Although the *FER* gene apparently has not been transduced by retroviruses, further experiments will have to be performed to determine its oncogenic potential and its possible involvement in human cancer.

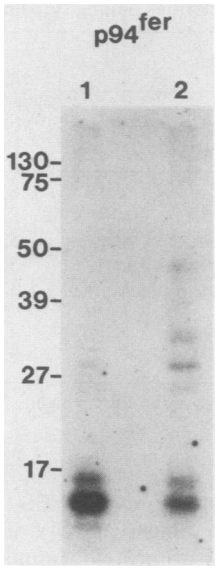

FIG. 7. V8 protease digestion of $p94^{FER}$. In vitro-phosphorylated protein of cell line A498 was identified by gel electrophoresis and autoradiography, incubated with 200 ng (lane 1) or 50 ng (lane 2) of V8 protease, and then subjected to electrophoresis through a 12.5% sodium dodecyl sulfate-polyacrylamide gel. Exposure time, 28 h. Molecular size markers are in kilodaltons.

ACKNOWLEDGMENTS

We thank John A. Barranger for critical review of the manuscript and Karen Betz and Olga Torrejon for typing of the manuscript.

This work was supported by Public Health Service grant CA 47456 from the National Institutes of Health to J.G.

ADDENDUM IN PROOF

After submission of this manuscript, T. Pawson (personal communication; Oncogene, in press) informed us of the isolation of a partial rat cDNA (*flk*) that could represent the rat homolog of FER.

LITERATURE CITED

1. **Ben-Neriah, Y., A. Bernards, M. Paskind, G. Q. Daley, and D. Baltimore.** 1986. Alternative 5′ exons in c-*abl* mRNA. Cell **44**:577–586.
2. **Cross, F. R., E. A. Garber, D. Pellman, and H. Hanafusa.** 1984. A short sequence in the $p60^{src}$ N terminus is required for $p60^{src}$ myristylation and membrane association and for cell transformation. Mol. Cell. Biol. **4**:1834–1842.
3. **de Klein, A., A. Geurts van Kessel, G. Grosveld, C. R. Bartram, A. Hagemeijer, D. Bootsma, N. K. Spurr, N. Heisterkamp, J. Groffen, and J. R. Stephenson.** 1982. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. Nature (London) **300**:765–767.
4. **Feldman, R., J. Gabrilove, J. Tam, M. Moore, and H. Hanafusa.** 1985. Specific expression of the human cellular *fps/fes*-encoded protein NCP92 in normal and leukemic myeloid cells. Proc. Natl. Acad. Sci. USA **82**:2379–2383.
5. **Feldman, R. A., J. P. Tam, and H. Hanafusa.** 1986. Antipeptide

antiserum identifies a widely distributed cellular tyrosine kinase related to but distinct from the c-*fps*/*fes*-encoded protein. Mol. Cell. Biol. **6**:1065–1073.

6. **Groffen, J., N. Heisterkamp, M. Shibuya, H. Hanafusa, and J. R. Stephenson.** 1983. Transforming genes of avian (*v-fps*) and mammalian (*v-fes*) retroviruses correspond to a common cellular locus. Virology **125**:480–486.
7. **Hampe, A., I. Laprovotte, F. Galibert, A. Fedele, and C. J. Sherr.** 1982. Nucleotide sequences of feline retroviral oncogenes (*v-fes*) provide evidence for a family of tyrosine-specific protein kinase genes. Cell **30**:775–785.
8. **Heisterkamp, N., J. Groffen, and J. R. Stephenson.** 1983. The human v-*abl* cellular homologue. J. Mol. Appl. Genet. **2**:57–68.
9. **Heisterkamp, N., K. Stam, J. Groffen, A. de Klein, and G. Grosveld.** 1985. Structural organization of the *bcr* gene and its role in the Ph′ translocation. Nature (London) **314**:758–761.
10. **Hermans, A., N. Heisterkamp, M. von Lindern, S. van Baal, D. Meijer, D. van der Plas, L. M. Wiedemann, J. Groffen, D. Bootsma, and G. Grosveld.** 1987. Unique fusion of *bcr* and *c-abl* genes in Philadelphia chromosome positive acute lymphoblastic leukemia. Cell **51**:33–40.
11. **Huang, C.-C., C. Hammond, and J. M. Bishop.** 1985. Nucleotide sequence and topography of chicken *c-fps*: genesis of a retroviral oncogene encoding a tyrosine-specific protein kinase. J. Mol. Biol. **181**:175–186.
12. **Hunter, T., and J. A. Cooper.** 1985. Protein-tyrosine kinases. Annu. Rev. Biochem. **54**:897–930.
13. **Kamps, M., J. Buss, and B. Sefton.** 1985. Mutation of NH_2-terminal glycine of $p60^{src}$ prevents both myristylation and morphological transformation. Proc. Natl. Acad. Sci. USA **82**:4625–4629.
14. **Kamps, M. P., S. S. Taylor, and B. M. Sefton.** 1984. Direct evidence that oncogenic tyrosine kinases and cyclic AMP-dependent protein kinases have homologous AMP-binding sites. Nature (London) **310**:589–592.
15. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNA's. Nucleic Acids Res. **12**:857–872.
16. **Kruh, G., C. King, M. Kraus, N. Posescu, S. Amsbaugh, W. McBride, and S. Aaronson.** 1986. A novel human gene closely related to the *abl* proto-oncogene. Science **234**:1545–1548.
17. **Lipman, D. F., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
18. **MacDonald, I., J. Levy, and T. Pawson.** 1985. Expression of the mammalian c-*fes* protein in hematopoietic cells and identification of a distinct *fes*-related protein. Mol. Cell. Biol. **5**:2543–2555.
19. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Sherr, C. J., C. W. Rettenmier, R. Sacca, M. F. Roussel, A. T. Look, and E. R. Stanley.** 1985. The *c-fms* proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor, CSF-1. Cell **41**:665–676.
21. **Shibuya, M., and H. Hanafusa.** 1982. Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses. Cell **30**:787–795.
22. **Shtivelman, E., B. Lifshitz, R. P. Gale, and E. Cannaani.** 1985. Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukemia. Nature (London) **315**:550–554.
23. **Smart, J. E., H. Oppermann, A. P. Czernilofsky, A. F. Purchio, R. L. Erikson, and J. M. Bishop.** 1981. Characterization of sites for tyrosine phosphorylation in the transforming protein of Rous sarcoma virus ($pp60^{v\text{-}src}$) and its normal cellular homologue ($pp60^{c\text{-}src}$). Proc. Natl. Acad. Sci. USA **78**:6013–6017.
24. **Spindler, K. R., D. S. E. Rosser, and A. J. Berk.** 1984. Analysis of adenovirus transforming proteins from early regions 1A and 1B with antisera to inducible fusion antigens produced in *Escherichia coli*. J. Virol. **49**:132–141.
25. **Stam, K., N. Heisterkamp, G. Grosveld, A. de Klein, R. S. Verma, M. Coleman, H. Dosik, and J. Groffen.** 1985. Evidence of a new chimeric *bcr*/*c-abl* *m*RNA in patients with chronic myelocytic leukemia and the Philadelphia chromosome. N. Engl. J. Med. **313**:1429–1433.
26. **Stam, K., N. Heisterkamp, F. H. Reynolds, Jr., and J. Groffen.** 1987. Evidence that the *phl* gene encodes a 160,000-dalton phosphoprotein with associated kinase activity. Mol. Cell. Biol. **7**:1955–1960.
27. **Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg.** 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature (London) **309**: 418–425.
28. **Weinmaster, G., E. Hinze, and T. Pawson.** 1983. Mapping of multiple phosphorylation sites within the structural and catalytic domains of the Fujinami avian sarcoma virus transforming protein. J. Virol. **46**:29–41.
29. **Young, J. C., and G. S. Martin.** 1984. Cellular localization of c-*fps* gene product NCP98. J. Virol. **52**:913–918.