BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **246,** 176–181 (1998)
ARTICLE NO. RC988560

Molecular Cloning and Expression of *Ehf,* a New Member of the *ets* Transcription Factor/Oncoprotein Gene Family

Michelle A. Bochert,[1] Laurie A. Kleinbaum,[1] Li-Yan Sun,[2] and Frank H. Burton[3]
Department of Pharmacology, University of Minnesota, Minneapolis, Minnesota 55455-0347

Received March 28, 1998

The *ets* family is a large multigene family of transcription factors that share a conserved DNA-binding "ETS" domain and include several oncoproteins that induce tumorigenesis when overexpressed. Here we report the cDNA cloning from mouse pituitary somatotroph tumors, sequence characterization and tissue-specific expression pattern in mice of a novel *ets* family gene, "*Ehf*" ("ets homologous factor"). The putative 300 amino acid Ehf protein is a highly divergent *ets* family member, but is most related to the recently identified oncoprotein ESX (36% overall and 84% ETS domain amino acid identity). Thus, Ehf and ESX comprise a new *ets* subfamily. *Ehf* is a single-copy gene, but produces four distinct mRNA transcripts. *Ehf* transcripts are abundant in mouse kidney and lung, less so in muscle and liver, and not detected in brain, spleen or testes. Because of its presence in somatotroph tumors and its relationship to ESX, Ehf may represent a new oncoprotein. © 1998 Academic Press

Key Words: **mouse; pituitary; adenoma; somatotroph; lung; kidney; skeletal muscle; liver; ESX; ESE-1.**

The *ets* family is a class of transcription factors that share homology to one another within an approximately 85 aa long, highly conserved DNA binding domain called the "ETS domain" (1, 2). Although their greatest similarity to each other is within this domain, *ets* factors can also have additional homologous domains including an "acidic domain" (3) and a "pointed domain" (4,5), which may be involved in homodimerization (4), cooperative binding to other transcription factors (3, 5) and transactivation (3, 5). Several members of the *ets* gene family are known oncogenes, or have been shown to be transcriptionally stimulated in proliferating cells, including *ets-1, ets-2, PEA3, ESX, erg, FLI-1, ERP, PU.1,* and *tel* (6–15). Consequently, many *ets* family members may contribute not only to normal development but to abnormal cellular proliferation associated with tumorigenesis.

Although many *ets* factors are expressed specifically in hematopoetic cells, others are expressed in different cell types such as epithelium or even in multiple adult organs, including ets-1, ets-2, elf -1, elf-2, ER81, ERM, ERP, NERF and ESX (3, 12, 16–21). *ESX* (Epithelial-Restricted with Serine boX) is a recently identified novel *ets* family gene which is more divergent than most other members of the *ets* family (9). Interestingly, *ESX* is transcriptionally stimulated in early stage breast cancer (9), suggesting a role in inducing or mediating tumorigenesis of mammary secretory epithelium. This likelihood is supported by evidence that the breast cancer oncogene *HER2/neu* is transcriptionally activated by ESX (9). *ESX* is expressed only in epithelial cells and is induced during terminal differentiation of the epidermis and in a primary human keratinocyte differentiation system (19).

In the pituitary, the presence of an *ets*-responsive DNA element that cooperates with a Pit-1 binding site in the promoter of the prolactin (PRL) gene has led to the prediction that an as yet undiscovered endogenous pituitary *ets* factor may regulate both PRL expression and possibly pituitary hormone-secreting cell development *in vivo* (22, 23). However, the only *ets* gene shown to be expressed in the pituitary to date, *c-ets-1,* is apparently expressed naturally only in angiogenesis-associated vascular endothelial cells, rather than in pituitary hormone-secreting cells (24), although *c-ets-1* has been reported to be expressed in the GH_4 pituitary cell line (22).

Here we report the discovery of a novel member of the *ets* gene family which we call *Ehf* (*ets*-homologous factor). *Ehf* is most closely related to, yet still consider-

[1] The first two authors should be regarded as joint first authors.
[2] Current Address: Hughes Institute, 2657 Patton Road, Roseville, MN 55113.
[3] Corresponding author: Department of Pharmacology, University of Minnesota, 3-249 Millard Hall, 435 Delaware Street S.E., Minneapolis, MN 55455-0347. Fax: (612) 625-8408. E-mail: burto006@maroon.tc.umn.edu.

Abbreviations: aa, amino acid(s); nt, nucleotide(s); bp, base pair(s); kb, kilobase pair(s); ORF, open reading frame; UT, untranslated; cDNA, complementary DNA; PRL, prolactin.

The nucleotide sequence has been deposited in GenBank with the Accession No. AF035527.

0006-291X/98 $25.00

ably divergent from, the putative oncogene *ESX*. *Ehf* cDNA was isolated from mouse pituitary somatotroph tumor tissue, indicating a possible role for Ehf in the regulation of somatotroph development, gene expression, tumorigenesis, or tumor angiogenesis. *Ehf* is also expressed in a variety of adult mouse organs in addition to pituitary, including lung, kidney, muscle and liver, but excluding spleen, brain and testes, which suggests that *Ehf* expression is probably not restricted to epithelial cells like *ESX* (19) or to hematopoetic cells like several other *ets* genes. Together, these data indicate that Ehf is potentially a new multi-organ oncoprotein, comprising with ESX a new and divergent subfamily of the *ets* transcription factors.

MATERIALS AND METHODS

cDNA clone isolation. Pituitary somatotroph tumor (25) mRNA from C57Bl/6 × Balb/c hybrid mice was extracted with the QuickPrep Micro mRNA Purification Kit (Pharmacia), and cDNA was synthesized by oligo-T primed reverse transcription, linker ligation, and directional cloning into pGEM11Zf-NotI/EcoRI/BAP vector, using the TimeSaver cDNA Synthesis Kit and Directional Cloning Toolbox (Pharmacia). The cDNA libraries of complexity 10^6 or greater were amplified and library plasmid DNA was prepared (Qiagen Megaprep Kit). Partial cDNA clones were isolated from ~1.5 kb–4.0 kb insert size-selected tumor library DNA by colony lift hybridization with an *Ehf*-specific probe obtained by differential display RT-PCR comparisons of normal vs. tumorous pituitary total RNA (M. Bochert, unpublished data). Plasmid miniprep DNA of the two largest positive cDNA clones as well as several smaller clones were isolated using the Wizard miniprep kit (Promega). Double-stranded plasmid DNA inserts of each clone were single-pass sequenced from both ends and from internal primers based on the sequence of the original *Ehf* differential display cDNA, using the ^{35}S radiolabeled Sanger dideoxy-method (USB Sequenase Version 2.0 DNA Sequencing Kit) followed by autoradiography. These sequences indicated that all smaller cDNA clones represented 5′ truncated subsets of a single cDNA with an invariant internal open reading frame (ORF) sequence which was completely contained within the largest two cDNA clones. Plasmid DNA was prepared from one of the larger *Ehf* cDNA clones by a standard alkaline lysis procedure (26). The complete DNA sequence of this larger *Ehf* cDNA clone (1319 base pair (bp), GenBank accession #AF035527) was obtained on both strands by fluorescence-tagged dideoxy-sequencing on an Applied Biosystems model 373A automated sequencer, in the University of Minnesota Department of Human Genetics Microchemical Facility. This sequence matched a composite sequence compiled from the prior single-pass sequences.

Northern blot analysis A mouse multi-tissue poly(A)$^+$ mRNA northern blot (Clontech) containing 2 μg of purified poly(A)$^+$ RNA from heart, brain, spleen, lung, liver, skeletal muscle, kidney and testes was probed with a ^{32}P-alpha-dCTP radiolabel-incorporated, 230 bp PCR-amplified portion of the ORF region from the *Ehf* cDNA. The Northern blot was exposed to X-ray films (X-OMAT-AR) for an initial 24 hour period followed by various longer exposure times of up to 2 weeks with a "lightning-plus" (Fisher) intensifier screen, to generate both short and long autoradiographs detecting moderate to low level *Ehf* expression in various tissues.

Genomic Southern blot analysis. Tail genomic DNA was isolated from normal Balb/c mice by digesting the tails with 0.3 mg/ml Proteinase K at 60°C overnight, extracting the DNA with chloroform, and precipitating the DNA with ethanol. 10 μg of genomic DNA was restriction-digested to completion with a final concentration of 5U/μg DNA of either EcoRI or HindIII and loaded onto agarose gels for horizontal electrophoresis, Southern blotting onto nitrocellulose membranes, and DNA blot hybridization using a 991 bp radiolabeled *Ehf* probe encompassing the entire *Ehf* ORF region, prepared as described above for the northern blot analysis. Autoradiography was performed using 24- to 72-hour exposures with a "lightning-plus" intensifier screen (Fisher) to detect single-copy *Ehf* gene sequence.

Computer analysis of DNA and amino acid sequence. *Ehf* cDNA nucleotide (nt) sequence was searched to EMBL, GenBank, DDBJ, PDB, and EST divisions of EMBL, GenBank and DDBJ nt databases via the NIH-NCBI "BLASTN" search engine. Putative ORF identification and amino acid (aa) sequence translation of the *Ehf* cDNA clone nt sequence was performed using DNASTAR software (DNA*), while the putative Ehf aa sequence was searched to the GenBank CDS translations, PDB, Swiss-PROT and PIR protein databases via the NIH-NCBI "BLASTP" search engine. Subsequent full-length aa sequence alignments and calculations of the extent of similarity between the 300 aa putative Ehf sequence and related database aa sequences discovered by BLASTP were performed by Align Query using sequence data at the GENESTREAM SSEARCH network server, CRBM Montpellier, France.

RESULTS AND DISCUSSION

Isolation and Sequence Analysis of a cDNA Clone Encoding Mouse Ehf

Two plasmid cDNA libraries were constructed from poly(A)$^+$ mRNA of either normal mouse pituitary or pituitary somatotroph tumor tissue, the latter isolated from a transgenic mouse model of G_s-dependent somatotroph tumorigenesis (25). The cDNA libraries were screened with a 230 bp probe representing a part of the ORF of *Ehf* cDNA. This probe was first obtained as a differential display RT-PCR product that was differentially expressed between normal pituitary mRNA and pituitary somatotroph tumor mRNA, whose nt sequence exhibited *ets* family homology (M. Bochert, unpublished data). Probing of both normal and tumorous pituitary cDNA libraries revealed the presence of *Ehf*-positive clones, which were subsequently isolated from the tumor library. Two of six *Ehf*-positive size-selected clones isolated from the tumor library had *Ehf* inserts of approximately 1.4 kb, while the remainder of the clones had smaller inserts. Complete double-stranded sequencing of one of the two largest *Ehf* cDNA clones, whose length was 1319 nt (Figure 1), and partial sequencing of the remaining cDNA clones (not shown) identified that the largest two cDNA clones both contained a long ORF bounded by 5′ and 3′ untranslated (UT) regions. The smaller cDNA clones represented 5′ truncated *Ehf* cDNA inserts presumably created by incomplete reverse transcription. All of the cDNA clones appeared internally primed from an A-rich 3′ UT region not preceded by a canonical poly-A addition site, suggesting that full length *Ehf* mRNA probably contains additional uncloned 3′ UT sequences not detectable with our ORF-specific probe. The absence of any cDNA

```
GGGTGGAAGGGCACGTTCCTTTTTTCCTTTCTTTCTTTTTGTCTTTCTTTCTTTCTTTCCCAAATGAACTCTATGTTCTCCACTC      85
TCCAGTGACATTTGTCTACGAGGCATAGATTTAAGGACTGTTTGATCCCTTGCAGATC                                143

ATG ATT CTG GAA GGA AGT GGT GTA ATG AAT CTC AAC CCA GCC AAC AAC CTC CTT CAC CAG CAA      206
M   I   L   E   G   S   G   V   M   N   L   N   P   A   N   N   L   L   H   Q   Q      21

CCA GCC TGG ACG GAC AGC TAC CCC ACA TGC AAT GTT TCC AGC GGT TTT TTT GGA AGC CAG TGG      269
P   A   W   T   D   S   Y   P   T   C   N   V   S   S   G   F   F   G   S   Q   W      42

CAT GAA ATC CAC CCT CAG TAC TGG ACC AAA TAC CAG GTG TGG GAA TGG CTG CAG CAC CTC CTG      332
H   E   I   H   P   Q   Y   W   T   K   Y   Q   V   W   E   W   L   Q   H   L   L      63

GAC ACC AAC CAG CTA GAC GCT AGC TGC ATC CCT TTC CAG GAG TTC GAC ATT AGC GGA GAA CAC      395
D   T   N   Q   L   D   A   S   C   I   P   F   Q   E   F   D   I   S   G   E   H      84

CTG TGC AGC ATG AGT CTG CAG GAG TTC ACG AGG GCA GCA GGC TCA GCT GGG CAG CTG CTC TAC      458
L   C   S   M   S   L   Q   E   F   T   R   A   A   G   S   A   G   Q   L   L   Y      105

AGC AAC CTA CAG CAT CTC AAG TGG AAC GGC CAA TGC AGC AGT GAC CTT TTC CAG TCC GCA CAC      521
S   N   L   Q   H   L   K   W   N   G   Q   C   S   S   D   L   F   Q   S   A   H      126

AAT GTC ATT GTC AAG ACT GAA CAA ACC GAT CCT TCC ATC ATG AAC ACA TGG AAA GAA GAA AAC      584
N   V   I   V   K   T   E   Q   T   D   P   S   I   M   N   T   W   K   E   E   N      147

TAT CTC TAT GAT CCC AGC TAT GGT AGC ACA GTA GAT CTG TTG GAC AGT AAG ACT TTC TGC CGG      647
Y   L   Y   D   P   S   Y   G   S   T   V   D   L   L   D   S   K   T   F   C   R      168

GCT CAG ATC TCC ATG ACA ACC TCC AGT CAC CTT CCA GTT GCA GAG TCA CCT GAT ATG AAA AAG      710
A   Q   I   S   M   T   T   S   S   H   L   P   V   A   E   S   P   D   M   K   K      189

GAG CAA GAC CAC CCT GTA AAG TCC CAC ACC AAA AAG CAC AAC CCA AGA GGC ACT CAC TTA TGG      773
E   Q   D   H   P   V   K   S   H   T   K   K   H   N   P   R   G   T   H   L   W      210

GAG TTC ATC CGA GAC ATT CTC TTG AGC CCA GAC AAG AAC CCA GGA CTG ATC AAA TGG GAA GAC      836
E   F   I   R   D   I   L   L   S   P   D   K   N   P   G   L   I   K   W   E   D      231

CGT TCG GAA GGC ATC TTC AGG TTC CTG AAG TCA GAA GCT GTG GCT CAG CTG TGG GGG AAA AAG      899
R   S   E   G   I   F   R   F   L   K   S   E   A   V   A   Q   L   W   G   K   K      252

AAA AAT AAC AGT AGC ATG ACA TAC GAG AAG CTC AGC CGG GCT ATG AGA TAT TAC TAC AAA CGA      962
K   N   N   S   S   M   T   Y   E   K   L   S   R   A   M   R   Y   Y   Y   K   R      273

GAA ATC CTG GAA CGT GTG GAT GGA CGA CGG CTG GTC TAT AAG TTT GGG AAG AAT GCT CGT GGA     1025
E   I   L   E   R   V   D   G   R   R   L   V   Y   K   F   G   K   N   A   R   G      294

TGG AGA GAA AAT GAG AAC TGA                                                           1046
W   R   E   N   E   N   .                                                             300

GGCTGCCAGCCCTTGGGACACAAACCAAAACACACAGCAAATGGATTCTGATCAATGAAGAACCGGACGTAAATATCTCAAAGAC    1131
TACTTTTCTGTGATATTTATGTACCATGAAGGGACAAAGAAAATCTACTTCTGACGGGAAGAAGGAACACTACAGTTGATAAAAA    1216
AAAAATTATTTTGTTACTTTGAAGTATGTCCTTTTGTGGGGAACAAATGTACACAGTTTTCTGTGAACTATGAAGCTGTATGTGA    1301
TTGTGAAAAAAAAATTCC                                                                    1319
```

FIG. 1. Nucleotide sequence of a 1319 nt mouse *Ehf* partial cDNA containing the complete Ehf protein coding region (GenBank accession #AF035527). Putative aa sequence is shown below the nt sequence within the ORF. Numbering of nt or aa (bold) is shown at right. Regions homologous to known *ets* family domains are underlined as follows: pointed domain homology (nt 282-401); ETS domain homology (nt 750-1007). In-frame terminator codons within the 5′ and 3′ UT regions that bound the ORF are also shown (boldface).

clones containing both ORF and complete 3′ UT sequence in our library screening suggests such clones may be infrequent, perhaps due to a large 3′ UT region. This is supported by our preliminary finding that human *Ehf* cDNA has a very long 3′ UT region (L. Kleinbaum, unpublished data). In addition to being incomplete at their 3′ UT ends, the two largest *Ehf* cDNA clones differed slightly in their 5′ endpoints, suggesting that these clones may also have an incomplete 5′ UT region. Nevertheless, the ORF sequence in all examined *Ehf* clones was invariant, suggesting that this ORF represents the complete *Ehf* coding sequence. This 900 nt ORF is bounded at its 5′ end by appropriate eukaryotic translation initiation sequences (27) and by a stop codon at its 3′ end, and is further surrounded on both 5′ and 3′ sides by UT regions containing multiple stop codons (Figure 1).

```
PU.1:   GSKKKIRLYQFLLDLLRSGDMKDSIWWVDKDKGTFQFSSKHKEALAHRWGIQKGNRKKMTYQKMARALRNYGKTGEVKKV-KKKLTYQF   (31.8%)
ESX:    NPRG-THLWEFIRDILIHPELNEGLMKWENRHEGV-FKFLRSEAVAQLWG-QKKKNSNMTYEKLSRAMRYYYKREILERVDGRRLVYKF   (83.7%)
Ehf:    NPRG-THLWEFIRDILLSPDKNPGLIKWEDRSEGI-FRFLKSEAVAQLWG-KKKNNSSMTYEKLSRAMRYYYKREILERVDGRRLVYKF
elf-1:  GNT--IYLWEFLLALLQDKATCPKYIDWT-QREKGIFKLVDSKAVSRLWG-KHKNKPDMNYETMGRALTYYYQRGILAKVEGGRLVYQF   (44.0%)
elk-1:  -PSV-T-LWQFLLQ-LLREQGNGHIISWTSR-DGGEFKLVDAEEVARLWG-LRKNKTNMNYDKLSRALRYYYDKNIIRKVSGQKFVYKF   (47.6%)
ets-1:  GSGP-IQLWQFLLE-LLTDKSCQSFISWT--GDGWEFKLSDPDEVARRWG-KRKNKPKMNYEKLSRGLRYYYDKNIIHKTAGKRYVYRF   (39.8%)
```

FIG. 2. Amino acid alignment of the ETS domain between mouse Ehf and PU.1 (accession #P17947), ESX (accession #U66894), elf-1 (accession #P32519), elk-1 (accession #P19419) and ets-1 (accession #P14921). The *ets* proteins to which Ehf is aligned show essentially zero divergence in their ETS domain between rodents and humans (data not shown). Percent aa identity to Ehf is shown at right. Dashes represent spaces inserted for optimal alignment.

Homology of Ehf to the ets Transcription Factor/ Oncoprotein Family

Ehf is a novel cDNA sequence, revealing no exact or close nt matches in the non-redundant GenBank, EMBL, DDBJ, and PDB nt sequence databases, including the dbEST, STS and GSS divisions, and only one nearly exact match to an unpublished 253 nt mouse partial expressed sequence tag (EST) in the dbEST database (accession #AA466217). The complete and invariant 900 nt *Ehf* ORF putatively encodes a 300 aa polypeptide (Figure 1). This putative Ehf aa sequence shows no exact or close matches in the PIR, PDB, GenBank CDS translations or Swiss-PROT aa databases. However, both the *Ehf* ORF DNA sequence and putative Ehf aa sequence reveal significant homology to other *ets* family genes and their encoded proteins (Figures 2, and 3), and mostly to ESX, a recently identified divergent *ets* factor that is epithelial cell-specific and has been variously called ESE (19), Elf3 (accession # AF016294), ERT (accession # AF017307), and jen (accession # U97156) in more recent mouse and human cDNA cloning studies. The similarity of mouse Ehf aa sequence to human ESX is 36% aa identity overall, with 84% aa identity within the ETS domain (Figure 2). Ehf is equally similar (39% overall and 84% ETS domain aa identity) to the more recently reported mouse homolog of ESX, Elf3 (accession # AF016294), confirming that *Ehf* is not mouse *ESX* but is a separate gene distantly related to *ESX*. The similarity of Ehf to ESX is greater than Ehf's similarity to other *ets* factors like ets-1 (20% overall and 40% ETS domain aa identity) (28) and elk-1 (14% overall and 48% ETS domain aa identity) (29) (Figure 2). Thus, Ehf is not only a novel *ets* factor, but, along with its distant relative ESX, constitutes a new and divergent *ets* subfamily (Figure 3). Based on ETS domain divergence, this "ESX-Ehf" subfamily is the second most divergent clade of known mammalian *ets* factors, behind the most divergent mammalian *ets* protein known to date, the oncoprotein PU.1 (Figure 3).

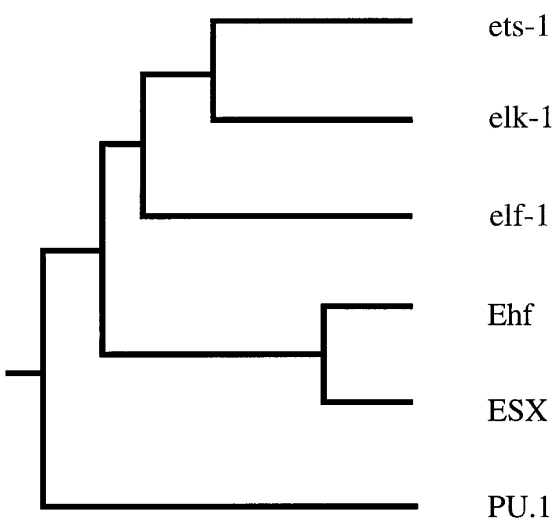

FIG. 3. Dendrogram showing the evolutionary relationship of the ETS domains in Ehf and other representative *ets* family members. The dendrogram was derived from the distance matrix of the percent aa difference between the *ets* factor alignments shown in Figure 2. PU.1, ESX, elf-1, elk-1 and ets-1 were chosen for alignment because each represents a major mammalian *ets* family clade (2,9), with the PU.1 branch containing only the PU.1 gene (2); the ESX branch containing only ESX and Ehf; the elf-1 branch containing only elf-1 (2); and the elk-1 and ets-1 branches each containing many divergent to highly homologous *ets* factors comprising the remainder of the mammalian *ets* family (2).

Amino Acid Composition Analysis of Ehf

Amino acid composition analysis of the putative Ehf polypeptide sequence indicates that Ehf contains a signature ETS domain (Figure 1), indicating that it is a transactivator of other genes. Ehf also contains a "pointed domain" (Figure 1), an approximately 100 aa sequence which is highly conserved within some members of the *ets* protein family including ets-1, ets-2, erg, erg-B (30) and ESX (9). The "pointed domain" has been postulated to be involved in homodimerization (4) or transcriptional activation (5). As with ESX, the functional significance of Ehf's "pointed domain" homology is not known, especially given that both Ehf and ESX lack a conserved MAP Kinase substrate site found in the "pointed domain" of ets-1, ets-2, and the *Drosophila ets* protein, Pointed P2 (9).

Tissue Specific Expression of Ehf mRNA in Mouse

Northern blot analysis of 2 μg of poly(A)$^+$ RNA extracted from multiple tissues of adult mice showed four transcripts of 1.4 kb, 2.0 kb, 2.3 kb and 3.8 kb in length

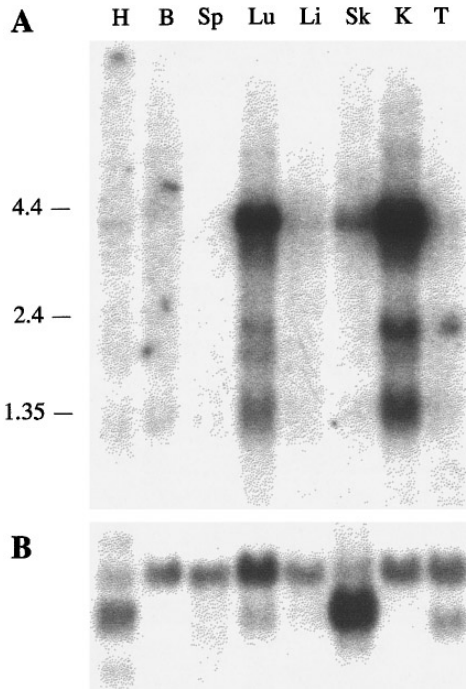

FIG. 4. (A) Northern analysis of *Ehf* mRNA levels in mouse tissues. RNA size marker locations are shown at left. Lane abbreviations: H (heart), B (brain), Sp (spleen), Lu (lung), Li (liver), Sk (skeletal muscle), K (kidney), T (testes). (B) Beta actin re-probing of blot to confirm equivalent RNA loading per lane. The different banding pattern evident in skeletal muscle and heart (cardiac muscle) reflect different beta-actin mRNA transcripts in muscle tissues.

that were variably detected in certain tissues using an *Ehf*-specific probe from within the *Ehf* ORF region (Figure 4). The major *Ehf* mRNA transcript is 3.8 kb, indicating most *Ehf* cDNA clones should be primed from this large mRNA species. However, all our cDNA clones were smaller than 1.4 kb and were primed internally at an A-rich 3′ UT sequence. This suggests the predominant *Ehf* transcript contains an additional long 3′ UT region not detectable with our ORF probe, a conclusion supported by our recent detection of a very large 3′ UT region in a human *Ehf* cDNA clone (L. Kleinbaum, unpublished data). Because every *Ehf* cDNA clone examined contains an invariant ORF, the most likely form of differential processing resulting in the four *Ehf* mRNA species is alternative 3′ UT polyadenylation or UT splicing, rather than alternative ORF splicing.

Ehf mRNA transcripts were differentially expressed in kidney, lung, skeletal muscle, heart (cardiac muscle) and liver, in order of decreasing abundance, and were not detected in brain, spleen, or testes. Re-probing of the commercial northern blot with a beta-actin specific probe confirmed approximately equal loading of mRNA. In addition to these above tissues, *Ehf* mRNA is also present in adult pituitary tissue and pituitary somatotroph tumor tissue. This is based on our isolation of multiple *Ehf* clones from cDNA libraries made from pituitary and pituitary tumor mRNA, as well as our detection of *Ehf* transcripts in total RNA extracts of these tissues by differential mRNA display (M. Bochert, unpublished data).

Analysis of Ehf Gene Copy Number

Genomic Southern blot analysis of EcoRI-restricted or HindIII-restricted tail DNA extracted from normal Balb/c mice and hybridized with an *Ehf* ORF-specific probe detected a single genomic DNA band in EcoRI-digested DNA and two bands in HindIII-digested DNA (Figure 5). The intensity of these bands was equivalent to that of a single-copy control marker, indicating that the probable copy number of the *Ehf* gene is one per haploid genome. Thus, *Ehf* is probably a single gene rather than one of a group of closely related genes. Moreover, this supports the contention that the four distinct *Ehf* mRNA transcripts (Figure 4) are most likely made from one *Ehf* gene, rather than being separate transcripts made from multiple *Ehf* genes.

Conclusion

Our data indicate that *Ehf* is a new member of the *ets* transcription factor family. *Ehf* comprises a new and divergent *ets* subfamily along with its closest homologue, *ESX*. *ESX* is transcriptionally up-regulated in early-stage breast cancer, and in turn induces the transcription of the *HER2/neu* oncogene (9). Although *Ehf* is expressed in other tissues besides those which express *ESX*, *Ehf* may act similarly to *ESX*, as well as to many other *ets* family oncogenes (6-15), inducing proliferation in the distinct tissues where it is expressed. *Ehf* is the first *ets* family gene known to be expressed in pituitary somatotroph tumors, as well as in a variety of normal organs. The only other *ets* gene

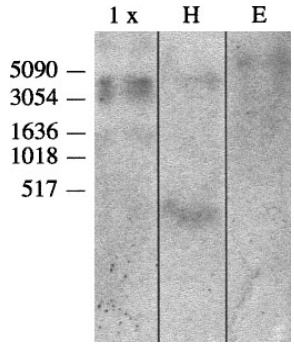

FIG. 5. Southern analysis of *Ehf* gene copy number in the mouse genome. DNA size marker locations are shown at left. 1x, single-copy *Ehf* control DNA diluted in salmon sperm DNA; H, HindIII-digested Balb/c mouse genomic DNA; E, EcoRI-digested Balb/c mouse genomic DNA. Signals in each lane were autoradiographically detected after hybridization with a PCR-radiolabeled *Ehf* ORF-specific fragment. The detection of a single band exhibiting 1x-equivalent intensity in EcoRI-digested mouse DNA indicates that *Ehf* is most likely a single-copy gene.

known to be expressed in pituitary, *c-ets-1*, was shown to be naturally expressed only in vascular cells during embryonic pituitary development, and is assumed to be involved with angiogenesis rather than pituitary hormone-secreting cell development, proliferation or function (24). The expression of *Ehf* in both adult pituitary and pituitary tumor tissues is thus of interest because unidentified *ets* protein(s) have been assumed to regulate pituitary-specific gene expression and possibly pituitary development and physiology. For example, the PRL gene, which is expressed in somatotroph-derived mammosomatotrophs and lactotrophs, contains *ets*-binding enhancer elements that can be transactivated in cultured cells by transfecting the cells with exogenous *ets* factors (22, 23). These exogenous *ets* factors enhance PRL transcription by acting cooperatively with the POU domain protein Pit-1, a known regulator of PRL and pituitary development. The endogenous pituitary *ets* factor(s) responsible for transcriptional regulation of PRL expression *in vivo,* as well as other aspects of pituitary somatotroph cell development, physiology and tumorigenesis, has not yet been identified (23). Thus, Ehf is a possible candidate because of its expression in the pituitary and in somatotroph cell tumors *in vivo.* In the future, overexpressing Ehf protein in Ehf-positive cell types of the pituitary (25) or other organs may allow a direct test of *Ehf*'s tumorigenic potential as well as its role in normal cell development and gene expression.

ACKNOWLEDGMENTS

We thank Jason Fink and Eileen Matro for assistance in preparing the manuscript. This work was supported by NIH training grant T32GM07994 to L.A.K. and by research grants to F.H.B. from the American Cancer Society, the Milheim Foundation, the Elsa U. Pardee Foundation, the Minnesota Medical Foundation, and the University of Minnesota Graduate School Grant-in-Aid Program.

REFERENCES

1. Karim, F. D., Urness, L. D., Thummel, C. S., Klemsz, M. J., McKercher, S. R., Celada, A., VanBeveren, C., Maki, R. A., Gunther, C. V., Nye, J. A., and Graves, B. J. (1990) *Genes Dev.* **4,** 1451–1453.
2. Macleod, K., Leprince, D., and Stehelin, D. (1992) *Trends Biochem.* **17,** 251–256.
3. Monté, D., Coutte, L., Baert, J. L., Angeli, I., Stéhelin, D., and de Launoit, Y. (1995) *Oncogene* **11,** 771–779.
4. Mclean, T. W., Ringold, S., Neuberg, D., Stegmaier, K., Tantravahi, R., Ritz, J., Koeffler, H. P., Takeuchi, S., Janssen, J. W., Seriu, T., Bartram, C. R., Sallan, S. E., Gilliland, D. G., and Golub, T. R. (1996) *Blood* **88,** 4252–4258.
5. Siddique, H. R., Ras, V. N., Lee, L., and Reddy, E. S. (1993) *Oncogene* **8,** 1751–1755.
6. Seth, A., and Papas, T. S. (1990) *Oncogene* **5,** 1761–1767.
7. Seth, A., Watson, D. K., Blair, D. G., and Papas, T. S. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7833–7837.
8. Trimble, M. S., Xin, J.-H., Guy, C. T., Muller, W. J., and Hassell, J. A. (1993) *Oncogene* **8,** 3037–3042.
9. Chang, C. H., Scott, G. K., Kuo, W. L., Xiong, X., Suzdaltseva, Y., Park, J. W., Sayre, P., Erny, K., Collins, C., Gray, J. W., and Benz, C. C. (1997) *Oncogene* **14,** 1617–1622.
10. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., DeJong, P., Rouleau, G., Aurias, A., and Thomas, G., (1992) *Nature* **359,** 162–165.
11. Sorensen, P. H. B., Lessnick, S. L., Lopezterrada, D., Liu, X. F., Triche, T. J., and Denny, C. T. (1994) *Nature Genet.* **6,** 146–151.
12. Lopez, M., Oettgen, P., Akbarali, Y., Dendorfer, U., and Libermann, T. A. (1994) *Mol. Cell. Biol.* **14,** 3292–3309.
13. Moreau-Gachelin, F., Tavitian, A., and Tambourin, P. (1988) *Nature* **331,** 277–280.
14. Moreau-Gachelin, F., Ray, D. Mattei, M. G., Tambourin, P., and Tavitian, A. (1989) *Oncogene* **4,** 1449–1456.
15. Golub, T. R., Barker, G. F., Lovett, M., and Frampton, J. (1994) *Cell* **77,** 307–316.
16. Bhat, N., Fisher, R. J., Fujiwara, S., Ascione, R., and Papas, T. S. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3161–3165.
17. Monté, L., Baert, J. L., Defossez, P. A., and de Launoit, Y., and Stéhelin, D. (1994) *Oncogene* **9,** 1397–1406.
18. Oettgen, P., Akbarali, Y., Boltax, J., Best, J., Kunsch, C., and Libermann, T. (1996) *Mol. Cell. Biol.* **16,** 5091–5106.
19. Oettgen, P., Alani, R. M., Barcinski, M., Brown, L., Akbarali, Y., Boltax, J., Kunsch, C., Munger, K., and Libermann, T. (1997) *Mol. Cell. Biol.* **17,** 4419–4433.
20. Wilkinson, D. A., Neale, G. A., Mao, S., Naeve, C. W., and Goorha, R. M. (1997) *Leukemia* **11,** 86–89.
21. Maroulakou, I. G., Papas, T. S., and Green, J. E. (1994) *Oncogene* **9,** 1551–1565.
22. Bradford, A. P., Conrad, K. E., Wasylyk, C., Wasylyk, B., and Gutierrex-Hartmann, A. (1995) *Mol. Cell. Biol.* **15,** 2849–2857.
23. Howard, P. W., and Maurer, R. A. (1995) *J. Biol. Chem.* **270,** 20930–20936.
24. Laurent-Huck, F. M., Egles, C., Kienlen, P., Stoeckel, M. E., and Felix, J. M. (1996) *Brain Research. Dev. Brain Research.* **97,** 107–117.
25. Burton, F. H., Hasel, K. W., Bloom, F. E., and Sutcliffe, J. G. (1991) *Nature* **350,** 74–77.
26. Maniatis, T., Fritch, E. F., and Sambrook, J. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor University Press, Cold Spring Harbor, NY.
27. Kozak, M. (1984) *Nucleic. Acids Res.* **12,** 857–872.
28. Watson, D. K., McWilliams, M. J., Lapis, P., Lautenberger, J. A., Schweinfest, C. W., and Papas, T. S. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7862–7866.
29. Rao, V. N., Huebner, K., Isobe, M., Ar-Rushdi, A., Croce, C. M., and Reddy, S. P. (1989) *Science* **244,** 66–70.
30. Lautenberger, J. A., Burdett, L. A., Gunnell, M. A., Shengmei, Q. ,Watson, D. K., O'Brian, S. J., and Papas, T. S. (1992) *Oncogene* **7,** 1713–1719.