GENOMICS **39,** 192–197 (1997)
ARTICLE NO. GE964502

cDNA Cloning, Expression Analysis, and Chromosomal Localization of a Gene with High Homology to Wheat eIF-(iso)4F and Mammalian eIF-4G

JOHN D. SHAUGHNESSY JR., NANCY A. JENKINS, AND NEAL G. COPELAND[1]

The Mammalian Genetics Laboratory, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, P.O. Box B, Frederick, Maryland 21702

Received September 4, 1996; accepted November 7, 1996

A novel mammalian gene, *Eif4g2,* with a high degree of homology to the p82 subunit of the wheat germ eukaryotic translation initiation factor eIF-(iso)4F and mammalian eIF-4G has been isolated. Zoo blot analysis indicates that *Eif4g2* is a single-copy gene that is highly conserved among vertebrates. Northern blot analysis shows that *Eif4g2* is ubiquitously expressed at high levels in all human and mouse tissues examined. The 3810-nucleotide *Eif4g2* cDNA contains a 907-amino-acid open reading frame that codes for a polypeptide with a predicted molecular mass of 102 kDa. The *Eif4g2* polypeptide exhibits an overall similarity to wheat p82 of 52%. A 248-amino-acid segment at the amino-terminal end of both peptides exhibits 63% similarity and contains conserved potential RNA binding domains and a phosphorylation site. The *Eif4g2* polypeptide contains multiple potential N-linked glycosylation sites as well as protein kinase C and casein kinase II phosphorylation sites. Southern blot analysis of DNA from interspecific backcross mice shows that *Eif4g2* is localized to distal mouse chromosome 7 in a region syntenic with human chromosome 11p15.

INTRODUCTION

The initiation of protein synthesis requires over 10 initiation factors, messenger RNA, ATP, GTP, and the 40S and 60S ribosomal subunits (for review see Merrick, 1990). All eukaryotic cellular mRNAs are blocked at their 5′ termini with m^7GTP, referred to as the cap (Shatkin, 1976). The initiation of translation is mediated by the specific recognition of the cap structure by polypeptides of the eIF-4 group. The eukaryotic initiation factor eIF-4F recognizes and binds the m^7GTP cap and catalyzes the ATP-dependent unwinding of the 5′-untranslated region of mRNA prior to the binding of the 40S ribosomal subunit (for review see Thach, 1992). The mammalian eIF-4F contains three subunits: the m^7GTP cap binding protein, eIF-4E, the ATP-dependent RNA helicase, eIF-4A, and a molecule of unknown function that forms complexes with other eIF factors, eIF-4G (Merrick, 1990; Thach, 1992). An eIF-4F complex has also been identified in wheat germ and contains only two subunits, p28 and p220, homologues of eIF-4E and eIF-4G, respectively (Lax *et al.,* 1986).

An isozyme form of wheat eIF-4F, eIF-(iso)4F, has also been identified (Browning *et al.,* 1987, 1990; Allen *et al.,* 1992). The eIF-(iso)4F complex is made up of two factors, p26 and p82, which have been shown to be the homologues of p28 and p220, respectively. The eIF-(iso)4F complex is functionally equivalent to, but distinct from, eIF-4F (Lax *et al.,* 1985, 1986; Browning *et al.,* 1987; Abramson *et al.,* 1988; Allen *et al.,* 1992). Although eIF-(iso)4F has been shown to exist in other higher plants, such as maize and cauliflower (Browning *et al.,* 1992), it has not been identified in other types of eukaryotic cells. Here we describe the cloning and characterization of a highly conserved vertebrate gene whose protein product has a high degree of homology to the p82 subunit of the wheat germ eIF-(iso)4F and to mammalian eIF-4G.

MATERIALS AND METHODS

Exon trapping. A non-chimeric 650-kb yeast artificial chromosome, YLA77E11, spanning the *Mrvi1* locus, was obtained from Research Genetics (J.Shaughnessy, unpublished data). Yeast chromosomes were isolated according to the manufacturer's protocol (BIO101). The yeast chromosomes were loaded on a 1% gel in a running buffer of 0.5× TAE. The chromosomes were separated by pulsed-field gel electrophoresis using a CHEF MAPPER (Bio-Rad). The gel was stained with 1 mg/ml ethidium bromide and visualized under ultraviolet illumination. The YAC DNA was excised from the gel, purified, digested with both *Bam*HI and *Bgl*II, and ligated into the *Bam*HI site of the exon trapping vector sPL3 (Nissan *et al.,* 1994). The remainder of the exon trapping protocol was performed according to the manufacturer's protocol (Gibco-BRL).

cDNA cloning. Approximately 750,000 recombinant phage clones from a mouse brain oligo(dT)-primed cDNA library (Stratagene) were

Sequence data of *Eif4g2* have been deposited with the EMBL/GenBank Data Libraries under Accession No. U63323.

[1] To whom correspondence should be addressed. Telephone: (301) 846-1260. Fax: (301) 846-6666.

0888-7543/97 $25.00

screened with the probe YLAET2512/9 according to conventional methods. Sixteen phage plaques were purified to homogeneity by repeated dilution and rescreening. cDNA inserts were rescued from phage by *in vivo* excision according to the manufacturer's protocol (Stratagene). cDNA insert sizes were determined by restriction digestion and gel electrophoresis. 5′ end confirmation was carried out by nested 5′ rapid amplification of cDNA ends (5′ RACE) using mouse brain Marathon Ready cDNA and the gene-specific primers EIF1 (5′-ACCGAAGACTCAGCAGCTGCCACCG-3′) and 5′-EIF1NEST (5′-GCCTCCTCTGGATCCGGTCGTCGGGG-3′) and the manufacturer-supplied adaptor primers AP1 and AP2 (Clontech).

DNA extraction and zoo blot hybridizations. High-molecular-weight genomic DNAs were extracted from various species by conventional methods. Restriction endonuclease digestions, agarose gel electrophoresis, Southern blot transfers, hybridizations, and washes were performed as previously described (Jenkins *et al.*, 1982). Hybridizations were performed at 65°C in 5× SSC, and a final wash was performed at 65°C in 1× SSC. The blot was then exposed to X-ray film for 10 days at −70°C with an intensifying screen.

Northern blot hybridizations. Northern blots containing 2 mg of twice selected poly(A)$^+$ RNA (Clontech) were prehybridized and hybridized with ExpressHyb according to the manufacturer's protocol (Clontech). Blots were washed to a final stringency of 0.1× SSC at 65°C. Blots were then exposed to X-ray film at −70°C with an intensifying screen.

DNA sequencing. DNA sequencing was performed using the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing kit (Perkin–Elmer) on the ABI Model 373A DNA Sequencer (Applied Biosystems). Sequence primers were either the T3, T7 sequencing primers or synthetic oligomers derived from previously determined sequence.

Chromosomal mapping. Interspecific backcross progeny were generated by mating (C57BL/6J × *Mus spretus*) F_1 females and C57BL/6J males as described (Copeland and Jenkins, 1991). A total of 204 N2 mice were used to map *Eif4g2*. DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer, and hybridization were performed as described (Jenkins *et al.*, 1982). Probes were labeled with [α-^{32}P]dCTP by the method of random priming (Stratagene). The probe YLAET2512/9 detected a 2.3-kb fragment in *Pvu*II-digested C57BL/6J DNA and a 6.6-kb fragment in *Pvu*II-digested *M. spretus* DNA. The presence or absence of the 6.6-kb *Pvu*II *M. spretus*-specific fragment was followed in backcross mice. The *Nup98* probe, pAmp-4-9, detected a major band of 7.4 kb in *Sca*I-digested C57BL/6J DNA and a major band of 5.5 kb in *Sca*I-digested *M. spretus* DNA. The segregation of the two *M. spretus*-specific bands recognized by each probe was followed in 204 backcross mice. Recombination distances were calculated as described (Green, 1981) using the computer program SPRETUS MADNESS.

Probes. The probe YLAET2512/9 was a 230-bp *Xba*I–*Eco*RI mouse genomic fragment representing an exon of the mouse *Eif4g2* gene. The probe pAmp-4-9 was a 410-bp *Not*I–*Eco*RI fragment derived from a cDNA clone of the human *NUP98* gene.

Computer analysis of DNA and protein sequences. DNA sequence analysis was performed by using the software package of the Genetics Computer Group (Devereux *et al.*, 1984). The sequence homology search was conducted at the protein level using the National Center for Biotechnology Information and the BLAST network service. The sequence alignment between the mouse p82 homologue and the wheat p82 was performed and analyzed based on the progressive sequence alignment program (Feng and Doolittle, 1987). Protein subsequence motifs were identified using the program Mac Vector (Kodak).

RESULTS

In an effort to identify novel disease genes in the region surrounding retroviral integration sites found in myeloid leukemias of BXH2 mice, we trapped a 103 base pair (bp) exon, YLAET2512/9, from a 650 kilobase (kb) yeast artificial chromosome that spans a site that

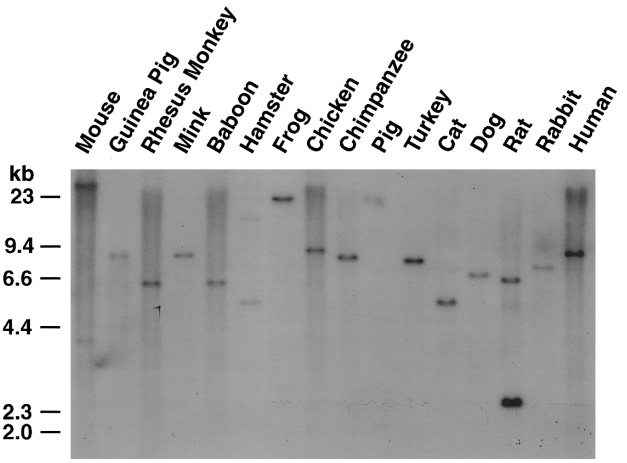

FIG. 1. A zoo blot of various vertebrate DNAs hybridized with the probe YLAET2512/9. Molecular weight size markers are indicated along the left side.

we have termed *Mrvi1* (J. Shaughnessy *et al.*, manuscript in preparation). To determine the relative conservation of YLAET2512/9 we radioactively labeled the DNA fragment and used it as a probe on Southern blots of various vertebrate DNAs (Fig. 1). These results showed that the probe recognized a single band in all samples with the exception of mouse, rat, and hamster, in which a major band and minor light hybridizing band were observed. The hybridization and wash stringencies used in the experiment were relatively high, with hybridization performed at 65°C with a wash stringency of 65°C and 1× SSC. This high stringency of hybridization and the recognition of a single band in most species suggested that the exon was likely part of a highly conserved, single-copy gene in vertebrates.

To analyze the expression pattern of the YLAET2512/9 associated gene and to determine the size of the mRNA, we hybridized YLAET2512/9 to Northern blots containing poly(A)$^+$ RNA from mouse embryo and various adult mouse and human tissues (Fig. 2). A major tanscript of approximately 4.6 kb was observed in all tissues tested in both mouse and human. An additional minor band of approximately 6.8 kb was observed in mouse. Hybridization of a Northern blot of whole mouse embryos constituting Days 7, 11, 13, and 15 of gestation showed that the gene is expressed at all stages of development. The expression level of this gene was found to be exceedingly high in most tissues with levels approximating those seen with the housekeeping genes used to normalize Northern blot RNA loading.

To a isolate a full-length YLAET2512/9 cDNA we screened an oligo(dT)-primed mouse brain cDNA library using YLAET2512/9 as a probe. From a screen of approximately 750,000 recombinants we isolated 16 positive clones. Eight of the 16 clones contained inserts of approximately 4.0 kb. The complete 3810-nucleotide sequence of one of the full-length clones is shown in Fig. 3. The sequence of this cDNA terminates at a consensus polyadenylation signal located 14 nucleotides upstream of a poly(A) tract. Since the 3810-nucleotide

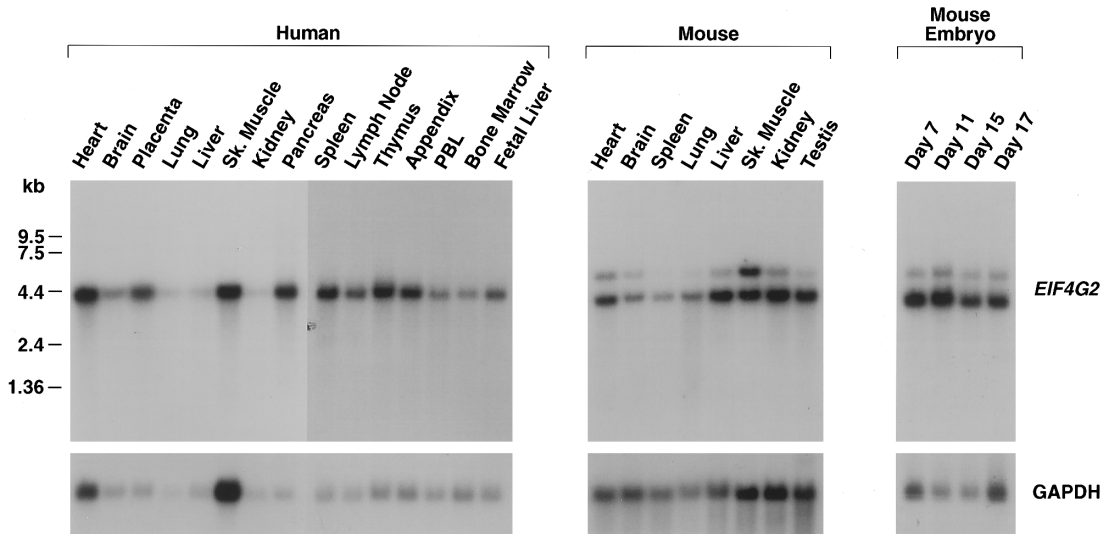

FIG. 2. Northern blot analysis of various adult human (**left**) and adult mouse (**center**) tissues and whole mouse embryos at various stages of gestation (**right**) hybridized with the probe YLAET2512/9. Each lane contains 2 μg of poly(A)$^+$ mRNA from various tissues, indicated across the top. Each autoradiogaph represents a 1-h exposure, except the embryo blot, which was exposed 1/2 h. All blots were hybridized with the *GAPDH* probe to control for RNA loading. Sizes in kilobases are indicated along the left side.

cDNA was shorter than the length predicted from Northern blots, we designed nested oligonucleotide primers complementary to sequences from the known cDNA sequence and performed 5′ RACE. Using a mouse brain Marathon Ready cDNA library (Clontech) as the template, we generated 5′ RACE fragments that were subcloned and sequenced. The analysis of the 5′ ends of these clones failed to extend the 5′ end beyond the point already established by the conventional cDNA clones (data not shown). We conclude that the sequence presented here represents the full-length cDNA. The cause of the discrepancy between the size of the messenger RNA observed in Northern blots and the cDNA sequence has not been resolved.

An analysis of the cDNA sequence indicted that the polypeptide could initiate at a GTG codon located within the context of a Kozak consensus sequence (Kozak, 1986). This GTG appears to be the initiation codon, as site-directed mutations within this codon in the human gene result in a loss of polypeptide synthesis in *in vitro* assays (N. Sonenberg, Montreal, pers. comm., 1996). The open reading frame ends with an in-frame termination codon at nucleotide 3006. The polypeptide produced by this mRNA would consist of 907 amino acids with a predicted molecular mass of 102 kDa. A search for protein subsequence motifs revealed five potential N-linked glycosylation sites at amino acids 98, 210, 474, 790, and 873 (Fig. 3). Two potential protein kinase C phosphorylation sites were identified at amino acids 87 and 565, and two potential casein kinase II sites were located at amino acids 302 and 313 (Fig. 3).

A computer aided comparison of the 907-amino-acid open reading frame to the database of protein sequences maintained at the National Center for Biotechnology Information using the BLASTX program (Devereux *et al.*, 1984) revealed that the predicted polypeptide has a high degree of homology the p82 subunit of wheat germ eIF-(iso)4F (Allen *et al.*, 1995) and to the p220 subunit of mammalian eIF-4G (Yan *et al.*, 1992). A more refined comparison using the BESTFIT program showed that the predicted polypeptide has a 52% similarity over 870 amino acids to the wheat p82 peptide and a 49% similarity over 1032 amino acids to the human EIF4G peptide. The highest degree of similarity among the three proteins occurs in the amino terminus. A 248-amino-acid stretch at the amino-terminal end of the gene described here (amino acids 111 to 408) exhibits 63% similarity to p82 (amino acids 211 to 497) (Fig. 4) and p220 (data not shown). Within this region a highly conserved putative RNA binding domain is conserved among all three peptides. A second potential RNA binding domain within the region is conserved between the gene described here and wheat p82 (Fig. 4). There also exist within this region a potential protein kinase C phosphorylation site in p82 (amino acids 397 to 400) and a casein kinase II phosphorylation site in the protein described here (amino acids 302 to 306) that is conserved in position with respect to the RNA binding domains of both polypeptides.

The mouse chromosomal location of the gene described here (designated *Eif4g2* for eukaryotic initiation factor 4 gamma homologue 2) was determined by interspecific backcross analysis using progeny derived from a mating of [(C57BL/6J $\times$ *M. spretus*)F_1 $\times$ C57BL/6J] mice (Copeland and Jenkins, 1991). This interspecific backcross mapping panel has been typed for over 2200 loci that are well distributed among all the autosomes as well as the X chromosome. C57BL/6J and *M. spretus* DNAs were digested with several restriction enzymes and analyzed by Southern blot hybridization for informative restriction fragment length polymorphisms (RFLPs) using the exon probe YLAET2512/9. The 6.6-kb *Pvu*II RFLP was then used to follow the

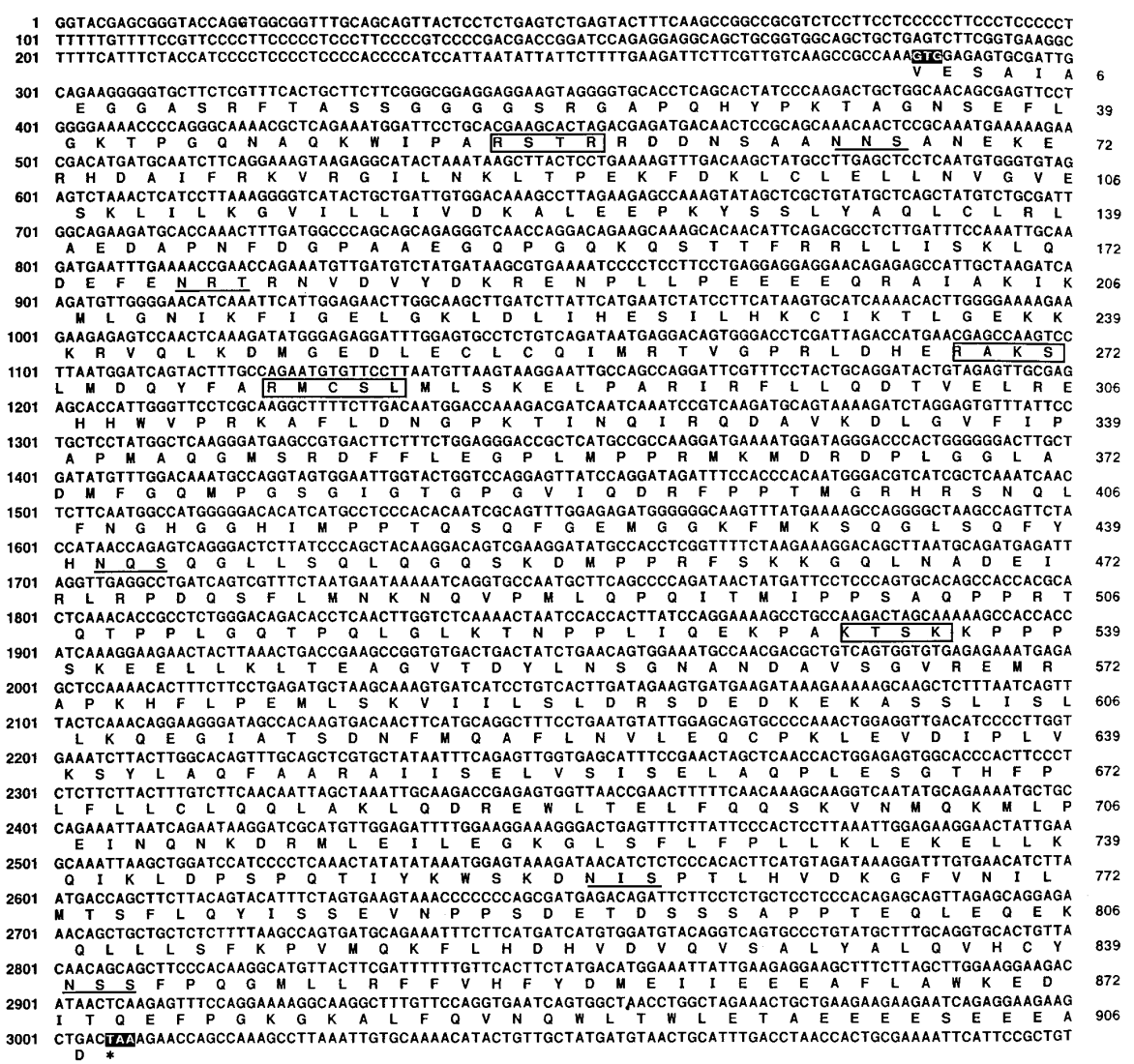

FIG. 3. Nucleotide and predicted amino acid sequence of *Eif4g2*. The GTG start codon and TAA stop codon are in boldface. The polyadenylation signal is underlined. Protein subsequence motifs are noted. Potential N-glycosylation sites are underlined. Potential protein kinase C and casein kinase II phosphorylation sites are boxed. Nucleotide number and amino acid number are indicated to the left and right, respectively. The sequence of *Eif4g2* was deposited with the EMBL/GenBank database under Accession No. U63323.

segregation of *Eif4g2* in backcross mice. The mapping results indicated that *Eif4g2* was located on chromosome 7 linked to *Nup98*, *Lmo1*, and *Pth* (Fig. 5). Although 113 mice were analyzed for every marker and are shown in the segregation analysis (Fig. 5), up to 153 mice were typed for some pairs of markers. Each locus was analyzed in pairwise combinations for recombination frequencies using the additional data. The ratios of the total number of mice exhibiting recombinant chromosomes to the total number of mice analyzed for each pair of loci and the most likely gene order are centromere–*Nup98*–1/146–*Lmo1*–1/122–*Eif4g2*–1/153–*Pth*. The recombination frequencies (expressed as genetic distances in centimorgans ± the standard error) are *Nup98*–0.7 ± 0.7–*Lmo1*–0.8 ± 0.8–*Eif4g2*–0.7 ± 0.7–*Pth*. The mapping data also indicate that *Eif4g2* maps to human 11p15 based on the fact that it is flanked on either side by *NUP98* and *PTH*, two genes that map to 11p15.

DISCUSSION

Here we have described the cloning and analysis of a gene that has a high degree of homology to the p82 subunit of wheat germ eukaryotic initiation factor eIF-(iso)4F, an isozyme of the m^7GTP cap binding complex, and to mammalian eIF-4G. The wheat germ isozyme form of eIF-4F, eIF-(iso)4F, contains two polypeptides, p28 and p82, that are antigenically distinct from the

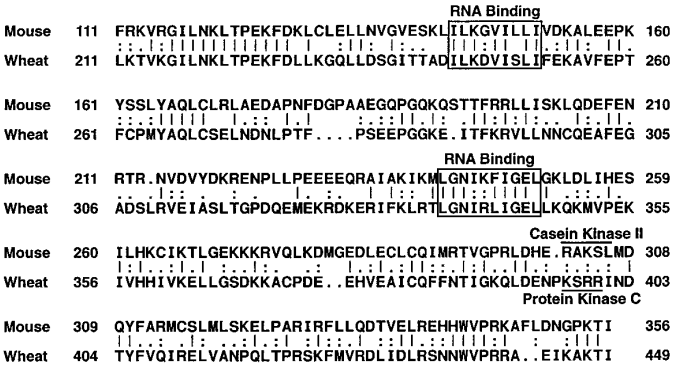

FIG. 4. Partial comparison of the amino acid sequence of the mouse *Eif4g2* (**Top**) and wheat eIF-(iso)4F p82 (**Bottom**). The 248-amino-acid region in the amino terminal of the two polypeptides, sharing 63% similarity, is compared. The positions of two potential RNA binding domains and a phosphorylation site are indicated, showing the conservation of both the primary sequence and the relative positions of the motifs. Amino acid identity is represented by a vertical line between the two sequences. Conservative amino acid changes are represented by two dots and semi-conservative changes by a single dot between the two sequences. Gaps created in the sequences to optimize alignments are represented by a dot in the sequence string. The amino acid positions within the full-length polypeptides are noted to the right and left of the sequence.

p26 and p220 polypeptides of eIF-4F (Browning *et al.*, 1987). Studies have shown that eIF-(iso)4F can (1) substitute for eIF-4F in *in vitro* translation systems deficient in eIF-4F; (2) substitute for eIF-4F in supporting the binding of the mRNA to the 40S ribosomal subunit; and (3) exhibit a RNA-dependent ATPase activity and ATP-dependent RNA helicase activity in the presence of eIF-4A and eIF-4B (Lax *et al.*, 1985, 1986; Browning *et al.*, 1987; Abramson *et al.*, 1988; Allen *et al.*, 1992). Additionally, equilibrium binding assays have shown that the two isoforms have different mechanisms for interacting with the cap structure and exhibit differences in their affinity for hypermethylated cap structures (Carberry *et al.*, 1991). It has therefore been suggested that there are two distinct isoforms of the eIF-4F complex in plants.

The discovery of what appears to be a mammalian counterpart to the plant eIF-4F isoform generates several questions. What is the reason for the evolutionary conservation of two isoforms of the cap binding complex? What distinct roles do the two isoforms play in the initiation of protein translation? One possible answer is that the two isoforms initiate the translation of two distinct types of mRNA. Regulation of the translation of mRNA appears to occur in two ways: a quantitative or global regulation that leads to alterations in the amount of total protein synthesized with relatively equal effects on the translation of all mRNAs and a qualitative mechanism that leads to an alteration in the relative translation of specific mRNAs (Merrick, 1990). Increasing the overall rate of initiation results in a greater rate of synthesis of all proteins, but those encoded by mRNAs containing a high degree of 5′-terminal secondary structure, the so-called "weak" mRNAs, are preferentially enhanced (Pelletier and Sonenberg, 1987). Several lines of evidence suggest that the factor responsible for the discrimination between weak and strong mRNAs is eIF-4F (Rhoads, 1991). It is possible that quantitative and/or qualitative regulation of protein synthesis during cell growth and differentiation is mediated by the actions of the two eIF-4F isoforms.

Eif4g2 is physically linked to a common site of viral integration found in the myeloid leukemias of the BXH2 inbred strain of mice and may be affected by the viral integrations (J. Shaughnessy *et al.*, manuscript in preparation). Additionally, *Eif4g2* is located within a cluster of genes that have chromosomal synteny with

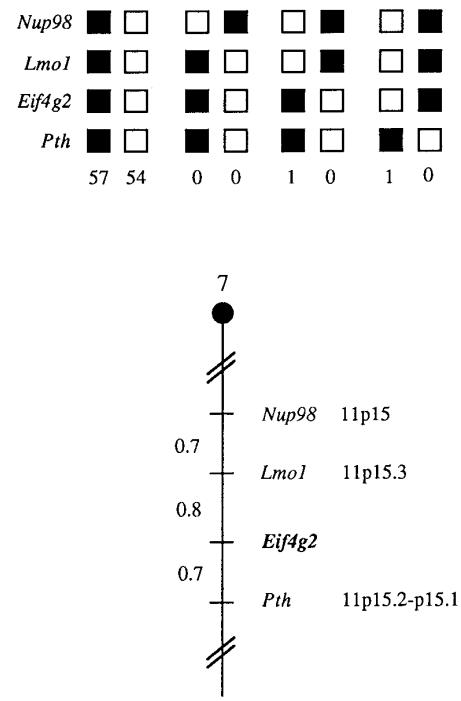

FIG. 5. A partial chromosome linkage map showing the chromosomal location of *Eif4g2*. (**Top**) The segregation patterns of *Eif4g2* and flanking genes in 113 backcross animals that were typed for all loci are shown. For individual pairs of loci more than 113 animals were typed (see text). Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J × *M. spretus*)F1 parent. The shaded boxes represent the presence of a C57BL/6J allele, and the white boxes represent the presence of a *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. (**Bottom**) A partial chromosome linkage map of chromosome 7 indicating the location of *Eif4g2*. Representative probes and RFLPs for loci used to position *Eif4g2* in this analysis have been previously described. These include *Lmo1* (Forino *et al.*, 1992) and *Pth* (Pathak *et al.*, 1996). One locus, *Nup98*, has not been reported previously for the Frederick IB. Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns. No double or multiple recombination events were observed. The number of recombinant N2 animals over the total number of N2 animals typed plus the recombination frequencies, expressed as genetic distances in centimorgans (± one standard error), is shown for each loci to the left of the chromosome map. The positions of loci in human chromosome 11 are shown to the right of the chromosome map. References to map positions for most human loci can be obtained from the GDB (Genome Data Base), a computerized database of human linkage information maintained by the William H. Welch Medical Library of the Johns Hopkins University (Baltimore, MD).

human chromosome 11p15, a region that harbors several unidentified tumor suppresser genes. Based on the chromosomal map position of *Eif4g2* and studies that show that alterations in the translation initiation complex can affect cell growth, it is possible that *Eif4g2* is itself a tumor suppressor gene. Brenner *et al.* (1988) have shown that the yeast cell cycle mutant CDC33 encodes the *Saccharomyces cerevisiae* 24-kDa p25 cap binding protein eIF-4E. In addition, Lazaris-Karatzas *et al.* (1990) have shown that overexpression of eIF-4E in NIH-3T3 and Rat-2 fibroblast causes their tumorigenic transformation in that the cells form transformed foci on cell monolayers, exhibit anchorage-independent growth, and form tumors when injected into nude mice. Future studies will be aimed at determining the exact role of *Eif4g2* in protein translation initiation and whether mutations in the gene might be associated with genetic disease.

ACKNOWLEDGMENTS

This work was sponsored by the National Cancer Institute, DHHS, under contract with ABL. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We thank Nahum Sonenberg for corresponding unpublished data. We thank David Largaespada for the *Nup98* probe. We thank Debbie Gilbert and Deborah Householder for excellent technical assistance.

REFERENCES

Abramson, R. D., Browning, K. S., Dever, T. E., Lawson, T. G., Thach, R. E., Ravel, J. M., and Merrick, W. C. (1988). Initiation factors that bind mRNA a comparison of mammalian fators with wheat germ factors. *J. Biol. Chem.* **263:** 5462–5467.

Allen, M. L., Metz, A. M., Timmer, R. T., Rhoads, R. E., and Browning, K. S. (1992). Isolation and sequence of the cDNAs encoding the subunits of the isozyme form of wheat protein synthesis initiation factor 4F. *J. Biol. Chem.* **267:** 23232–23236.

Brenner, C., Nakayama, N., Goebl, M., Tanaka, K., Toh-E, A., and Matsumoto, K. (1988) CDC33 encodes mRNA cap-binding protein eIF-4E of *Saccharomyces cerevisiae. Mol. Cell. Biol.* **8:** 3556–3559.

Browning, K. S., Lax, S. R., and Ravel, J. M. (1987). Identification of two messenger RNA cap binding proteins in wheat germ. Evidence that the 28-kDa subunit of eIF-4B and the 26-kDa subunit of eIF-4F are antigenically distinct polypeptides. *J. Biol. Chem.* **262:** 11228–11232.

Browning, K. S., Humphreys, J., Hobbs, W., Smith, G. B., and Ravel, J. M. (1990). Determination of the amounts of the protein synthesis initiation and elongation factors in wheat germ. *J. Biol. Chem.* **265:** 17967–17973.

Browning, K. S., Webster, C., Roberts, J. K. M., and Ravel, J. M. (1992). Identification of an isozyme form of protein synthesis initiation factor 4F in plants. *J. Biol. Chem.* **267:** 10096–10100.

Carberry, S. E., Darzynkiewicz, E., and Goss, D. J. (1991). A comparison of the binding of methylated cap analogues to wheat germ protein synthesis initiation factors 4F and (iso) 4F. *Biochemistry* **30:** 1624–1627.

Copeland, N. G., and Jenkins, N. A. (1991). Development and applications of a molecular genetic linkage map of the mouse genome. *Trends Genet.* **7:** 44–50.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12:** 387–395.

Feng, D.-F., and Doolittle, R. F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **25:** 351–360.

Forino, L., Boehm, T., White, L., Forester, A., Sherrington, P., Liao, X. B., Brannan, C. I., Jenkins, N. A., Copeland, N. G., and Rabbits, T. H. (1992). The rhombotin gene family encode related LIM-domain proteins whose differing expression suggests multiple roles in mouse development. *J. Mol. Biol.* **226:** 747–761.

Green, E. L. (1981). Linkage, recombination, and mapping. *In* "Genetics and Probability in Animal Breeding Experiments," pp. 77–113, Oxford Univ. Press, New York.

Jenkins, N. A., Copeland, N. G., Taylor, B. A., and Lee, B. K. (1982). Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of Mus musculus. *J. Virol.* **43:** 26–36.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44:** 283–292.

Lax, S., Fritz, W., Browning, K. S., and Ravel, J. (1985). Isolation and characterization of factors from wheat germ that exhibit eukaryotic initiation factor 4B activity and overcome 7-methylguanosine 5′-triphosphate inhibition of polypeptide synthesis. *Proc. Natl. Acad. Sci. USA* **82:** 330–333.

Lax, S. R., Browning, K. S., Maia, D. M., and Ravel, J. M. (1986) ATPase activities of wheat germ initiation factors 4A, 4B, and 4F. *J. Biol. Chem.* **261:** 15632–15636.

Lazaris-Karatzas, A., Montine, K. S., and Sonenberg, N. (1990). Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5′ cap. *Nature* **345:** 544–547.

Merrick, W. C. (1990). Overview: Mechanism of translation initiation in eukaryotes. *Enzyme* **44:** 7–16.

Nissan, P. E., Ally, A., and Watkins, P. C. (1994). Protocols for trapping internal and 3′-terminal exons. *PCR Methods Appl.* **4:** S24–S39.

Pathak, B. G., Shaughnessy J. D., Jr., Meneton, P., Greeb, J., Shull, G. E., Jenkins, N. A., and Copeland, N. G. (1996). Mouse chromosomal location of three epithelial sodium channel subunit genes and and apical sodium chloride cotransporter gene. *Genomics* **33:** 124–127.

Pelletier, J., and Sonenberg, N. (1987). The involvement of mRNA secondary structure in protein synthesis. *Biochem. Cell Biol.* **65:** 576–581.

Rhoads, R. E. (1991). Protein synthesis, cell growth and oncogenesis. *Curr. Opin. Cell Biol.* **3:** 1019–1024.

Shatkin, A. J. (1976). Capping of eukaryotic mRNAs. *Cell* **9:** 645–653.

Thach, R. E. (1992). Cap recap: The involvement of eIF-4F in regulating gene expression. *Cell* **68:** 177–180.

Yan, R., Rychlik, W., Etchison, D., and Rhodes, R. E. (1992). Amino acid sequence of the human protein synthesis initiation factor eIF-4g. *J. Biol. Chem.* **267:** 23226–23231.