The *yes*-Related Cellular Gene *lyn* Encodes a Possible Tyrosine Kinase Similar to *p56^{lck}*

YUJI YAMANASHI,¹ SHIN-ICHI FUKUSHIGE,² KENTARO SEMBA,¹ JUN SUKEGAWA,¹ NOBUYUKI MIYAJIMA,¹ KEN-ICHI MATSUBARA,² TADASHI YAMAMOTO,^{1*} AND KUMAO TOYOSHIMA¹

The Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108,¹ and Institute for Molecular and Cellular Biology, Osaka University, Suita, Osaka 565,² Japan

Received 21 July 1986/Accepted 24 September 1986

With *v-yes* DNA as the probe, a human cDNA library made from placental RNA was screened under relaxed conditions, and DNA clones derived from a novel genetic locus, termed *lyn*, were obtained. Nucleotide sequencing revealed that *lyn* could encode a novel tyrosine kinase that was very similar to mouse T-lymphocyte-specific tyrosine kinase *p56^{lck}* and the *v-yes* protein as well as to the gene products of *v-fgr* and *v-src*. Northern hybridization analysis revealed that a 3.2-kilobase *lyn* mRNA was expressed in a variety of tissues of the human fetus. The pattern of *lyn* mRNA expression was different from those of related genes, such as *yes* and *syn*. Hybridization analysis of DNA from sorted chromosomes showed that the *lyn* gene is located on human chromosome 8 q13-qter.

To date, protein products of eight retroviral transforming genes (*src*, *yes*, *fgr*, *fps/fes*, *abl*, *ros*, *fms*, and *erbB*) have been shown to exhibit tyrosine kinase activity. The catalytic domain of the protein tyrosine kinase was first identified in the *v-src* protein as a sequence of about 300 amino acids (kinase domain), which is conserved in the above gene products and has been shown to be responsible for transforming ability (19). Protein-tyrosine kinase activity is also associated with receptors of polypeptide growth factors, such as epidermal growth factor (8), colony-stimulating factor 1 (37), platelet-derived growth factor (33), insulin (22), and insulinlike growth factor 1 (20, 38). Most of these receptor proteins carry the kinase domain, and their protein-tyrosine kinase activities have been suggested to play a part in signal transduction (19). Recent studies showed an overlap between retroviral transforming proteins of the tyrosine kinase family and growth factor receptors. The genes encoding the epidermal growth factor receptor and colony-stimulating factor 1 receptor are the cellular counterparts of the *v-erbB* gene and *v-fms* gene, respectively (10, 43, 52). This suggests that deregulated expression of a receptor gene would lead to neoplastic transformation of cells.

Analysis of the *c-src* gene, the cellular homolog of the *v-src* gene, however, revealed that it apparently does not encode a typical growth factor receptor (46). The normal function of nonreceptor-type tyrosine kinase is unknown. For further understanding of the mechanism of cellular transformation by the nonreceptor-type oncogene, we examined the structure and expression of the cellular homolog of a *src*-related gene, *v-yes*. By screening a human genomic library with a *v-yes* probe, we recently obtained partial genomic clones derived from the *c-yes*, *c-src*, *c-fgr*, and *syn* genes, which have been mapped on human chromosomes 18 (42), 20 (25), 1 (34), and 6 (41), respectively. Nucleotide sequence analysis of these genes showed that their gene products are all highly homologous in the kinase domain (about 80% homology in amino acid sequence), but are less homologous to the receptor-type proto-oncogene. In addition, as far as examined, the intron-exon structures of the human *c-yes*-1, *c-fgr*, *c-src*, and *syn* genes are identical to

that of chicken *c-src* (41, 46) but different from those of the cellular homologs of other members of the tyrosine kinase family, mouse *c-abl* (51) and chicken *c-fps* (18). These data indicate that *c-yes*, *c-fgr*, *c-src*, and *syn* are closely related both in structure and in evolution. To obtain information on the structure and function of the *v-yes* gene, we screened a human cDNA library and obtained cDNA clones of the *c-yes* gene and also a novel *yes*-related gene.

MATERIALS AND METHODS

Isolation of RNA and Northern blot analysis. Total RNAs were extracted by the guanidine isothiocyanate-cesium chloride method (7) from tissues (brain, lung, liver, kidney) of a 12-week-old single human fetus, a human placenta at term, and human embryo fibroblast (cell line TIG-1) (35). The TIG-1 cells were cultured in Dulbecco modified Eagle medium with 10% fetal calf serum. Polyadenylated [poly(A)⁺] RNAs were selected and subjected to electrophoresis in an 0.8% agarose gel containing 2.2 M formaldehyde (26). The fractionated RNAs were then transferred directly to nitrocellulose filters and hybridized to the ³²P-labeled DNA fragment (specific activity, 2 × 10⁸ cpm/μg) as described previously (47). The 1.5-kilobase-pair (kbp) *Sal*I fragment of pYS-2 (42) was used as a *v-yes*-specific probe, the 0.6-kbp *Bgl*II-*Hinc*II fragment of pSN-2 (41) as a *syn*-specific probe, and the 2.2-kbp *Eco*RI-*Bam*HI fragment of pLYP10 as a *lyn*-specific probe (Fig. 1A). Hybridization was carried out at 42°C for 16 h in a mixture containing formamide (50%, stringent conditions; 35%, relaxed conditions), 5× SSC (1× SSC is 150 mM sodium chloride and 15 mM sodium citrate), 50 mM sodium phosphate buffer (pH 7.0), 10× Denhardt solution (1× Denhardt solution is 0.02% polyvinylpyrrolidone, 0.02% Ficoll, and 0.02% bovine serum albumin), yeast RNA (125 μg/μl), and denatured salmon sperm DNA (250 μg/ml). After hybridization, the filter was washed first with a solution of 2× SSC and 0.1% sodium dodecyl sulfate (SDS) at room temperature and then with either 0.2× SSC-0.1% SDS (stringent) or 0.5× SSC-0.1% SDS (relaxed) at 50°C.

cDNA synthesis. Total poly(A)⁺ RNA derived from human placenta at term was separated into 24 fractions by sucrose density gradient centrifugation. An appropriate fraction for

* Corresponding author.

FIG. 1. Structure of pLYP10. (A) Schematic drawing of the pLYP10 cDNA clone. The *EcoRI*-*Bam*HI fragment of 2.2 kbp was used as a *lyn*-specific probe. The wavy lines indicate unspliced introns. The large boxes indicate coding sequences, and the shaded box indicates the putative tyrosine kinase domain (19). The solid line indicates the 3' untranslated sequence. The location of a stretch of 14 deoxyadenylate residues is indicated. The nucleotide sequences around the splicing sites indicated by arrows (a, b, and c) are shown below. (B) Nucleotide sequences around the splicing sites (a, b, and c) indicated in panel A are aligned with those of the chicken *c-src* gene. Arrows indicate splicing sites. The consensus sequences of the donor (GT) and acceptor (AG) sites (29) are underlined. The exon numbers of the chicken *c-src* gene are cited from Takeya and Hanafusa (46).

cDNA template was selected by Northern blot analysis. Complementary DNA libraries were made by a described method (12), with slight modifications, using pBR322 and pUC9 as vectors. The first strand of cDNA was synthesized in a reaction volume of 100 μ l containing 50 mM Tris hydrochloride (pH 7.5), 75 mM KCl, 10 mM dithiothreitol, 3 mM MgCl₂, 1 mM each dGTP, dATP, dTTP, and dCTP, 100 μ Ci of [α -³²P]dCTP (3×10^3 Ci/mmol) per ml, 50 μ g of oligo(dT)₁₂₋₁₈ (P-L Biochemicals) per ml, 800 U of RNasin (Seikagaku, Inc.) per ml, 2×10^4 U of reverse transcriptase from Moloney murine leukemia virus (Bethesda Research Laboratories), and 100 μ g of poly(A)⁺ RNA from an appropriate fraction per ml at 37°C for 60 min. The reaction was stopped by adding EDTA to a final concentration of 20 mM. The products were extracted with phenol-chloroform and precipitated twice with ethanol in 2 M ammonium acetate. The amounts of cDNAs synthesized were determined by measuring trichloroacetic acid-insoluble radioactivity, and their sizes were determined by alkaline agarose gel electrophoresis. For synthesis of the second strand, 500 ng of single-stranded cDNA was incubated in 100 μ l of 100 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid)-NaOH (pH 7.6)-4 mM MgCl₂-15 mM β -mercaptoethanol-67.5 mM KCl-125 μ M each dGTP, dATP, dTTP, and dCTP-150 mM NAD-100 μ Ci of [α -³²P]dCTP (3×10^3 Ci/mmol) per ml-25 U of *Escherichia coli* DNA ligase (Biolabs) per ml-300 U of *E. coli* DNA polymerase I (Biolabs) per ml-10 U of *E. coli* RNase H (P-L Biochemicals) per ml at 14°C for 60 min and then at 22°C for 60 min. The reaction was terminated with EDTA, and the products

were extracted with phenol-chloroform and precipitated twice with ethanol as described above.

cDNA cloning. Approximately 10 to 20 residues of dCMP were added to the 3' termini of the double-stranded cDNAs with terminal deoxynucleotidyl transferase (Bethesda Research Laboratories). These dC-tailed cDNAs were annealed with pBR322 (New England Nuclear Corp.) or pUC9 (P-L Biochemicals), both tailed with dG₈₋₁₂ at *Pst*I sites. The recombinant molecules were transfected into *E. coli* MC1061 (14). The transformants were screened by the high-density colony-screening method (15). Hybridization was performed in a solution containing formamide (50%, stringent; 35%, relaxed), 4 $\times$ SSC, 50 mM HEPES-NaOH (pH 7.0), 10 $\times$ Denhardt solution, yeast RNA (10 μ g/ml), and denatured *E. coli* DNA (70 μ g/ml) at 42°C for 16 h. Washing conditions were as described above.

Nucleotide sequencing. The nucleotide sequence was determined by the dideoxy chain termination method (39) and unidirectional deletion method (17) with M13mp19 and pUC19 as cloning vectors (53). Deoxy-7-deazaguanosine triphosphate was used in place of dGTP to determine GC-rich sequences (28).

Chromosome mapping. Chromosomal location of the *lyn* gene was determined by Southern blot analysis of sorted chromosomes. GM0131 and GM0845 cells were provided by the Human Genetic Mutant Cell Repository (Institute for Medical Research, Camden, N.Y.). DNA probes specific to the factor IX and thyroglobulin genes were kindly provided by G. G. Brownlee and F. Baas, respectively.

RESULTS

Isolation of a cDNA clone derived from precursor mRNA of the *yes*-related gene. Poly(A)⁺ RNA prepared from human placenta was fractionated by sucrose density gradient centrifugation. A cDNA library was made from an RNA fraction that contained 4.5 to 5.5-kb poly(A)⁺ RNA and hybridized with the *v-yes* probe under conditions of reduced stringency. By screening 3×10^4 independent clones of this library with the *v-yes* probe, a unique cDNA clone (pLYP10) was obtained in addition to a *c-yes* cDNA clone. Nucleotide sequencing of the 2,741-base-pair (bp) insert of pLYP10 revealed that this clone contains sequences which encode a polypeptide related to the *v-yes* protein with interruption by what appeared to be unspliced introns (Fig. 1A). Sequences unique to the splicing donor and acceptor sites were found within the 1-kbp *EcoRI*-*Hind*III fragment of the insert. The predicted intron sequences were shown to be spliced from mature mRNA (see below, analysis of pLY30). The sequence of a predicted exon in this region corresponded to that of exon 4 of chicken *c-src*. The splicing sites were identical to the corresponding sites of chicken *c-src* (46) (Fig. 1B). These data suggested that the gene from which pLYP10 was derived and the *c-src* gene originated from the same ancestral gene. No sequence that corresponded to the exon 3 sequence of chicken *c-src* was identified in the pLYP10 insert. Part of the amino acid sequence predicted from the sequence of the pLYP10 insert was highly homologous to that of the kinase domains of *lck* (75%) (the name of the *lsk*¹/*tck* gene has been changed to *lck*) (27, 50) and *v-yes* (69%) (23), suggesting that it could specify tyrosine kinase. Thus, the gene from which the pLYP10 insert originated was termed *lyn* (*lck/yes*-related novel tyrosine kinase). The above data showed that the pLYP10 insert was generated from partially spliced *lyn* mRNA.

Isolation and characterization of cDNA clones derived from *lyn* mature mRNA. Northern hybridization analysis of hu-

FIG. 2. *lyn* cDNA nucleotide sequence. (A) Restriction maps of the human *lyn* cDNA inserts. The wavy lines for the pLYP10 clone represent unspliced introns (Fig. 1A). Abbreviations: E, *Eco*RI; P, *Pst*I; H, *Hind*III; B, *Bam*HI. (B) Complete nucleotide sequence of the 2,312-bp cDNA insert of pLY30 and the deduced amino acid sequence. Amino acids are numbered from the predicted initiation methionine. The termination codon 57 bp upstream of this methionine is boxed. The AATAAC sequence, an alternative polyadenylation signal (3), is underlined.

man placental RNA with the 2.2-kb *Eco*RI-*Bam*HI DNA probe prepared from pLYP10 showed that the *lyn* gene is transcribed into a 3.2-kb mature mRNA. We constructed a cDNA library with 3.0 to 3.5-kb RNA fractionated by sucrose density gradient centrifugation. Screening of 6×10^4 independent clones of this library with the *lyn*-specific probe resulted in the isolation of 16 positive clones. Of these, four clones (pLY30, pLY12'A, pLY25', and pLY14B) containing inserts longer than 1.5 kbp were analyzed further (Fig. 2A). A restriction map of the spliced region of the pLYP10 insert was identical to those of the four clones.

The 2,312-nucleotide sequence of the pLY30 insert was determined (Fig. 2B) and found to contain the sequence of predicted exons, but not introns, identified in the pLYP10 insert. An open reading frame of 1,536 bp starting with the first ATG codon at position 298 was identified. Although the flanking nucleotides of this ATG codon do not show a perfect match with the consensus sequence of Kozak (24), this codon seems to be the initiation codon, since a stop codon, TGA, was found 57 bp upstream in the same reading frame as the ATG. In addition, a sequence of 10 amino acids of the amino terminus of *lyn* was highly conserved in the amino terminus of *c-src* (Fig. 3). The predicted open reading frame could encode a protein of 512 amino acid residues, whose calculated molecular weight was 58,574.

The 14 deoxyadenylate residues at the extreme 3' termi-

nus of the pLY30 insert are not preceded by the typical consensus sequence, AATAAA, for the polyadenylation signal. Thus, this poly(dA) stretch may be an internal sequence located in the 3' untranslated region of 3.2-kb *lyn* mRNA. However, the AATAAC sequence located 55 nucleotides upstream of the poly(dA) stretch might be a signal for polyadenylation (3). This possibility cannot be excluded because we do not have cDNA clones that extend downstream of the poly(dA) stretch.

Possible tyrosine kinase encoded by the *lyn* gene. The predicted 512-amino-acid sequence of the *lyn* protein is shown in Fig. 3 in comparison with the sequences of chicken *c-src* (46) and mouse *lck* (27, 50). Except for 70 to 80 residues at their amino termini, these amino acid sequences are quite similar. In particular, the amino acid sequence between positions 252 and 490 of *lyn* is highly homologous to the sequence of the kinase domain of the *c-src* protein, *v-fgr* protein, *v-yes* protein, and *lck* protein and is distantly related to those of other tyrosine kinases (Table 1). Additionally, homology between this sequence and the sequence of the kinase domain of *raf*mil, which specifies serine kinase, is further reduced (31% homology) (4). Of 38 amino acid residues conserved in the kinase domain of previously described tyrosine kinases (19), 37 residues are conserved in the *lyn* protein. These include a lysine residue at position 275 and three glycine residues in the sequence Gly-X-Gly-X-X-

lyn	1	MGCIKSKGKD-----SL-SDDGVDLKTQPVNRT-ERT-IYVRDPT-SNKQQRVPESQLLPQ-QRFQT-KDP EEGQ---
lck	1	VC ---- NPE --WMENID ---C -NCH PIV LD - ISL IRNGSEVRDPLVTEGSL -PASPLQ
csrc	1	SS P PSQRRR EPP STHHGGF ASQ PNK AAPDTHR P RSFGTVAT PK FG FNTSD VTS QRA ALA
lyn	66	DIV---VALYPYDGIHPDDLSPFKKE-KMKVLEEHEGWKAKSLTKKEGFIPSNYVAKLNTLETEWFFKDITRKDAER
lck	63	--NLVI HS EPS DG G E QLRI QS Q T GQ F F A S P P NLS
csrc	81	GG TTF D ESRTET RLQI NNTD D L H T GQT Y PSDSIQA Y GK RES
lyn	142	QLLAPGNSAGAFIIRESETLKGFSLSVRDFDPVHGDVHKYKIRSLDNGGYIISPRITFCISDMIKHYQKQADGLCRR
lck	140	TH S STA QNQ E V N F GLH LVR TNAS TK
csrc	161	L N E PR T V T AYC S NAK LNV K S F TS TQ SSLQQLVAY S H H
lyn	222	LEKACISPKPQ-KPWDKDAWEIPRESIKLVKRLGAGQFGVEVWVGYYNNSTKVAVKTLKPGTMSVQAFLEENLMKTLQHD
lck	220	SRP QTQ WE E V TL E GH S Q S PVP A Q P
csrc	241	TNV PTS TQGLA LR EVK Q C TW GT R I N PE Q QV K R E
lyn	301	KLVRLYAVVTREPIYIITEYMAKGSLLDFLKSDEGGKVLLPKLIDFSAQIAEGMAYIERKNYIHRDLRAANVLVSES LM
lck	299	R -Q EN V TPS I LNVN L MA F EQ I DT S
csrc	321	Q -S V S GEM KYLR Q V MA S V M V I G N V
lyn	381	CKIADFGLARVIEDNEYTAREGAKFPIKWTAPEAINFGCFTIKSDVVSFGILLYEIVTYGKIPYPGRTNADVMTALSQGY
lck	378	L Y T T H R M PE IQN ER
csrc	400	V L Q ALY R T LT K RV MV RE LDQVER
lyn	461	RMPRVENCPDELYDIMKMCWKEAEERPTFDYLSQVLDDFYTATEGQYQQQP* 512
lck	458	V PD E HL ML RP D R F P * 509
csrc	480	CPPE ES H L CQ RRDP E AF E YF S P PGENL* 533

FIG. 3. Comparison of the amino acid sequences of human *lyn*, mouse *lck*, and chicken *c-src*. Only nonidentical residues are shown for *lck* and *c-src*. Gaps (—) are introduced to optimize alignment. Dots indicate residues that are conserved in all known tyrosine kinases (19). Arrows indicate the glycine of the *c-src* protein that is myristylated (5) and tyrosine phosphorylated in vivo (9). The open triangles indicate residues involved in ATP binding (21).

Gly at positions 254 to 259, both of which are suggested to be involved in ATP binding (21). The tyrosine at position 397 of *lyn* corresponds to Tyr-416 of the *src* protein, which is autophosphorylated in vitro (45). Moreover, the tyrosine at position 508 of *lyn* corresponds to Tyr-527 of the *src* protein, which is phosphorylated in vivo and plays a role in both the kinase activity and transforming ability of the *src* protein (9). These data suggest that the *lyn* protein should exhibit protein-tyrosine kinase activity. As the *lyn* protein does not have a stretch of hydrophobic amino acid residues characteristic of a transmembrane protein, it is not a typical receptor of polypeptide growth factors.

Expression of *lyn* mRNA in normal tissues. Northern hybridization of RNAs derived from human tissues was carried

out with the *lyn*-specific probe under stringent conditions. The 3.2-kb *lyn* mRNA was detected at various levels of expression in placenta at term and in human tissues (brain, lung, liver, and kidney) obtained from a single fetus, but was not detected in cultured human embryo fibroblasts (Fig. 4A). The level of *lyn* mRNA was higher in the liver than in other tissues. In the placenta, two additional mRNA species of 3.6 and 2.6 kb hybridized weakly with the *lyn* probe. The synthesis of these RNAs may be due to multiple poly(A) addition signals or alternative splicing events. It is also possible that these mRNAs are transcribed from a related gene(s) that has not yet been identified.

For examination of the specificity of expression of the related genes, the same filters were then hybridized to

TABLE 1. Amino acid sequence homology between the kinase domains^a of the *lyn* product and other tyrosine kinases^b

Tyrosine kinase	% Homology with <i>lyn</i>
<i>lck</i>	75
<i>c-src</i>	67
<i>v-fgr</i>	66
<i>v-yes</i>	69
<i>v-abl</i>	51
<i>v-fps</i>	44
<i>v-ros</i>	43
<i>v-fms</i>	38
Insulin receptor	36
Epidermal growth factor receptor	33

^a The kinase domain represents the amino acid residues corresponding to those between positions 252 and 490 of the *lyn* product (see Fig. 3).

^b The tyrosine kinases compared with the *lyn* product were mouse *lck* (27, 50), chicken *c-src* (46), *v-fgr* (31), *v-yes* (23), *v-abl* (36), *v-fps* (44), *v-ros* (32), and *v-fms* (13), and human insulin receptor (48) and epidermal growth factor receptor (49).

FIG. 4. Northern blot hybridization of RNAs from normal human cells. RNAs were isolated from the brain (lanes 1), lung (lanes 2), liver (lanes 3), and kidney (lanes 4) of a human fetus, from placenta at term (lanes 5 and 6), and from TIG-1 cells (lanes 7). Samples (5 µg) of poly(A)⁺ RNAs were subjected to blot hybridization with the ³²P-labeled *lyn* probe (panel A), *c-yes* probe (panel B, lanes 6 and 7), and *syn* probe (panel C) under stringent conditions and with the *v-yes* probe (panel B, lanes 1 to 5) under relaxed conditions. The length of each transcript was determined by comparison with transcripts (10.5, 5.6, and 2.6 kbp) of the human epidermal growth factor receptor gene (49).

FIG. 5. Assignment of the *lyn* gene to chromosome 8 q13-qter. (A) Sorting chromosomes from lymphoblast GM0131 cells and analysis of DNAs from the sorted fractions (A to H). Each DNA sample was cleaved with *Eco*RI and subjected to Southern blot analysis with the 32 P-labeled *lyn* probe. Total DNA cleaved with *Eco*RI was used as the standard (lanes L). (B) Sorting chromosomes from fibroblast GM0845 cells containing the translocation t(2;8) (q37; q13), and analysis of DNAs from the sorted fractions (a to h).

32 P-labeled probes specific to *c-yes* and *syn* (Fig. 4B and C). The *c-yes* mRNA was detected at higher levels in the brain than in other tissues and embryo fibroblasts. The *syn* mRNA was also detected at a high level in the brain. Thus, expression of the *lyn* gene is regulated differently in different cell types than are the related genes.

Allocation of the *lyn* gene to human chromosome 8. Meta-phase chromosomes were prepared from two cell lines, GM0131 and GM0845, and separated into eight fractions with a fluorescence-activated cell sorter (11). DNAs were prepared from each fraction of the sorted chromosomes and subjected to Southern blot analyses. By using the *lyn*-specific probe, karyotypically normal lymphoblast GM0131 cells produced two positives (Fig. 5A), of which the one with the stronger signal was recovered in fraction C, which contained chromosomes 6, 7, X, and 8, and the one with the weaker signal was recovered in fraction D, which contained chromosomes 9, 10, 11, and 12. This distribution suggests that the *lyn* gene is carried by chromosome X or 8, the major portion of which is fraction C and some in fraction D. By hybridizing the same Southern filter with DNA probes for the factor IX gene on chromosome X and the thyroglobulin gene on chromosome 8, the thyroglobulin probe produced a hybridization pattern similar to that obtained with the *lyn* probe (data not shown), suggesting that the *lyn* gene is located on chromosome 8. To confirm this, chromosomes were similarly sorted from GM0845 fibroblasts, which carry reciprocal translocation t(2q37;8q13). This cell line carries the derivative chromosome 2, der(2), which contains q13-qter of chromosome 8. The derivative chromosome can be sorted at the position overlapping with chromosomes 1 and

2, whereas the other derivative, der(8), which contains pter-q13 of chromosome 8, can be sorted at the position overlapping with chromosomes 19 and 20. Southern blot analysis of the sorted chromosomes from GM0845 cells revealed three positive signals, one in fraction a, in which der(2) was recovered, and in fractions c and d, in which normal chromosome 8 was recovered (Fig. 5B). These data clearly demonstrate that the *lyn* gene is located on chromosome 8 q13-qter. Since chromosomal translocations such as t(2p11-13; 8q24), t(8q24; 14q32), t(8q24; 22q11), t(3p25; 8q21), t(3p14; 8q24), and t(8q24; 22q11) have been found to be associated with Burkitt's lymphomas, mixed salivary gland tumors, familial renal cell carcinomas, and acute myelogenous leukemias (30), it seems of interest to examine the possible involvement of the *lyn* gene in these tumors.

DISCUSSION

In this work, we identified a novel *yes*-related gene, *lyn*. Characterization of cDNA clones suggested that *lyn* could encode a novel tyrosine kinase of about 58.5 kilodaltons. Comparison of the amino acid sequences of the kinase domains of *lyn* and other protein-tyrosine kinase genes revealed closer homology between *lyn* and *lck*. The homologies between the kinase domains of human *lyn* and *v-yes*, *v-fgr*, and chicken *c-src* were also high, but less than that between *lyn* and *lck* (Table 1). The *lyn* and *lck* genes encode proteins with 512 and 509 amino acid residues, respectively, which are apparently smaller than those encoded by chicken *c-src* (533 residues) (46), human *c-yes* (543 residues) (45a), and human *syn* (537 residues) (41). The sequences of

```

lyn   ---QYQQQPCOOH
lck   ---QYQPQPCOOH
c-src ---QYQPGENLCOOH
c-yes ---QYQPGENLCOOH
syn   ---QYQPGENLCOOH
c-fgr ---QYQPGDQTCOOH

```

FIG. 6. Alignment of carboxy-terminal amino acid sequences of *src*-related gene products. Comparison of carboxy-terminal amino acid sequences of human *lyn*, mouse *lck* (27, 50), human *c-src* (2), human *c-yes* (45a), human *c-fgr* (34), and human *syn* (41).

the *lyn* and *lck* proteins at their carboxy termini differ from those of the proteins of human *c-src*, *c-yes*, *c-fgr*, and *syn* (Fig. 6). These data suggest that *lyn* and *lck* are highly related in structure. Therefore it would be reasonable to speculate that the functions of their proteins may be similar. Southern blot hybridization of *Bam*HI-digested mouse DNA with a *lyn*-specific probe revealed four *lyn*-specific fragments (30, 14, 8, and 5 kbp) (data not shown), while two *Bam*HI fragments (22 and 1.6 kbp) have been shown to hybridize with an *lck* probe (27) that covers the entire sequence of the *lyn* probe. Thus, the *lyn* gene is not a human counterpart of mouse *lck*.

Our previous data (34, 41) showed that *src*, *fgr*, *yes*, and *syn* were formed in the most recent split within the proto-oncogenes of the tyrosine kinase family. As the intron-exon structure predicted from the cDNA clone derived from *lyn* precursor mRNA is identical to that of *src*, the *lyn* gene and *src* gene are suggested to be derived from the same ancestral gene. The close similarity of *lyn* to *lck* (discussed above) suggests that the *lck* gene also originated from the same archetype. However, we do not know how the expression of these similar but distinct genes is regulated. In a previous report, we showed that the modes of expression of *c-fgr*, *c-yes*, and *syn* differ in different types of cells (41). Here we showed that *syn* is highly expressed in the brain, while expression of *lyn* mRNA is higher in the liver than in other tissues of the human fetus. In addition, unlike *syn* and *yes* mRNA, *lyn* mRNA was not detectable in cultured human embryo fibroblasts. These data suggest that each of these closely related genes functions in distinct cell types. Since the protein products of these similar but distinct genes differ in their amino-terminal sequences, we assume that these sequences may be important in the interaction of these kinases with other factors that determine their specific roles.

The transforming protein pp60^{*c-src*} contains covalently linked myristate at its amino terminus, which allows its tight association with the cell membrane (40). In addition to pp60^{*c-src*}, its cellular homolog pp60^{*c-src*} (5), as well as the catalytic subunit of the cyclic AMP-dependent protein kinase (6), the B subunit of the phosphatase calcineurin (1), and the mammalian retroviral protein p15^{*gag*} (16), have been shown to be myristylated at the exposed amino-terminal glycine residue. Not only the *lyn* protein but also the *c-yes*, *syn*, and *lck* proteins carry glycine residues at position 2. Therefore, when the initiation methionine is removed, the amino-terminal glycine would be available for myristylation.

Although both pp60^{*c-src*} and pp60^{*v-src*} are autophosphorylated at Tyr-416 in vitro, the major site of tyrosine phosphorylation of pp60^{*c-src*} in vivo is Tyr-527, a residue not present in pp60^{*v-src*}. It is hypothesized that the lower protein-tyrosine kinase activity and transforming ability of pp60^{*c-src*} than of pp60^{*v-src*} are due to phosphorylation of Tyr-527 (9). Interestingly, Tyr-527 of pp60^{*c-src*} is conserved in the corresponding position of the *lyn* protein (Tyr-508), and also in the

c-yes, *syn*, and *lck* proteins. Examination of the in vivo phosphorylation of Tyr-508 of the *lyn* protein seems important.

ACKNOWLEDGMENTS

We thank S. Nishimura for kindly supplying dc⁷GTP, M. Nishizawa and S. Ikawa for valuable discussions, and M. Ide for typing this manuscript.

This work was supported in part by a Grant-in-Aid for Special Project Research Cancer-Bioscience from the Ministry of Education, Science, and Culture of Japan.

LITERATURE CITED

- Aitken, A., P. Cohen, S. Santikarn, D. H. Williams, A. G. Calder, A. Smith, and C. B. Klee. 1982. Identification of the NH₂-terminal blocking group of calcineurin B as myristic acid. *FEBS Lett.* 150:314-318.
- Anderson, S. K., C. P. Gibbs, A. Tanaka, H.-J. Kung, and D. J. Fujita. 1985. Human cellular *src* gene: nucleotide sequence and derived amino acid sequence of the region coding for the carboxy-terminal two-thirds of pp60^{*c-src*}. *Mol. Cell. Biol.* 5:1122-1129.
- Birnsteil, M. L., and M. Busslinger. 1985. Transcription termination and 3' processing: the end is in site! *Cell* 41:349-359.
- Bonner, T. I., S. B. Kerby, P. Sutran, M. A. Gunnell, G. Mark, and U. R. Rapp. 1985. Structure and biological activity of human homologs of the *raf* oncogene. *Mol. Cell. Biol.* 5:1400-1407.
- Buss, J. E., and B. M. Sefton. 1985. Myristic acid, a rare fatty acid, is the lipid attached to the transforming protein of Rous sarcoma virus and its cellular homolog. *J. Virol.* 53:7-12.
- Carr, S. A., K. Biemann, S. Shoji, D. C. Parmelee, and K. Titani. 1982. *n*-Tetradecanoyl is the NH₂-terminal blocking group of the catalytic subunit of cyclic AMP-dependent protein kinase from bovine cardiac muscle. *Proc. Natl. Acad. Sci. USA* 79:6128-6131.
- Chirgwin, J. M., A. E. Pyzybyla, R. J. McDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-5299.
- Cohen, S., G. Carpenter, and L. King, Jr. 1980. Epidermal growth factor-receptor-protein kinase interactions: co-purification of receptor and epidermal growth factor-enhanced phosphorylation activity. *J. Biol. Chem.* 255:4834-4842.
- Cooper, J. A., K. L. Gould, C. A. Cartwright, and T. Hunter. 1986. Tyr⁵²⁷ is phosphorylated in pp60^{*c-src*}: implications for regulation. *Science* 231:1431-1434.
- Downward, J., Y. Yarden, E. Mayes, G. Scrase, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger, and M. D. Waterfield. 1984. Close similarity of epidermal growth factor receptor and *v-erb-B* oncogene protein sequences. *Nature (London)* 307:521-527.
- Fukushige, S., T. Murotsu, and K. Matsubara. 1986. Chromosomal assignment of human genes for gastrin, thyrotropin (TSH)- β subunit and *c-erb-2* by chromosome sorting combined with velocity sedimentation and Southern hybridization. *Biochem. Biophys. Res. Commun.* 134:477-483.
- Gubler, V., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* 25:263-269.
- Hampe, A., M. Gobet, C. J. Sherr, and F. Galibert. 1984. Nucleotide sequence of the feline retroviral oncogene *v-fms* shows unexpected homology with oncogenes encoding tyrosine-specific protein kinases. *Proc. Natl. Acad. Sci. USA* 81:85-89.
- Hanahan, D. 1983. Studies on transformation of *E. coli* with plasmids. *J. Mol. Biol.* 166:557-580.
- Hanahan, D., and M. Meselson. 1980. A protocol for high density screening of plasmids in 1776. *Gene* 10:63-67.
- Henderson, L. E., H. C. Krutzsch, and S. Oroszlan. 1983. Myristyl amino terminal acylation of murine retroviral proteins: an unusual post-translation protein modification. *Proc. Natl. Acad. Sci. USA* 80:339-343.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene*

- 28:351-359.
18. 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.
 19. Hunter, T., and J. A. Cooper. 1985. Protein-tyrosine kinases. *Annu. Rev. Biochem.* 54:897-930.
 20. Jacobs, S., F. C. Kull, Jr., H. S. Earps, M. E. Svoboda, J. J. Van Wyk, and P. Cuatrecasas. 1983. Somatomedin-C stimulates the phosphorylation of the β -subunit of its own receptor. *J. Biol. Chem.* 258:9581-9584.
 21. Kamps, M. P., S. S. Taylor, and B. M. Sefton. 1984. Direct evidence that oncogenic tyrosine kinases and cyclic AMP-dependent protein kinase have homologous ATP-binding sites. *Nature (London)* 310:589-592.
 22. Kasuga, M., Y. Zick, D. L. Blithe, M. Crettaz, and C. R. Kahn. 1982. Insulin stimulates tyrosine phosphorylation of the insulin receptor in a cell-free system. *Nature (London)* 298:667-669.
 23. Kitamura, N., A. Kitamura, K. Toyoshima, Y. Hirayama, and M. Yoshida. 1982. Avian sarcoma virus Y73 genome sequence and structural similarity of its transforming gene product to that of Rous sarcoma virus. *Nature (London)* 312:205-208.
 24. Kozak, M. 1984. Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin *in vivo*. *Nature (London)* 308:241-246.
 25. Le Beau, M. M., C. A. Westbrook, M. O. Diaz, and J. D. Rowley. 1984. Evidence for two distinct *c-src* loci on human chromosome 1 and 20. *Nature (London)* 312:70-71.
 26. Lehrach, H., K. Diamond, J. M. Wozney, and H. Boedtker. 1977. RNA molecular weight determination by gel electrophoresis under denaturing conditions: a critical reexamination. *Biochemistry* 16:4743-4751.
 27. Marth, J. D., R. Peet, E. G. Krebs, and R. M. Perlmutter. 1985. A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed in the murine T cell lymphoma LSTRA. *Cell* 43:393-404.
 28. Mizusawa, S., S. Nishimura, and F. Seela. 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. *Nucleic Acids Res.* 14:1319-1324.
 29. Mount, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* 10:459-472.
 30. Mulvihill, J. J., and P. Madigan. 1984. Neoplasia of man (*Homo sapiens*), p. 447-448. In S. J. O'Brien (ed.), *Genetic maps*, vol. 3. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 31. Naharro, G., K. C. Robbins, and E. P. Reddy. 1984. Gene product of *v-fgr* onc: hybrid protein containing a portion of actin and a tyrosine-specific protein kinase. *Science* 223:63-66.
 32. Neckameyer, W. S., and L.-H. Wang. 1984. Molecular cloning and characterization of avian sarcoma virus UR2 and comparison of its transforming sequences with those of other avian sarcoma viruses. *J. Virol.* 50:914-921.
 33. Nishimura, J., J. S. Huang, and T. F. Devel. 1982. Platelet derived growth factor stimulates tyrosine-specific protein kinase activity in Swiss mouse 3T3 cell membranes. *Proc. Natl. Acad. Sci. USA* 79:4303-4307.
 34. Nishizawa, M., K. Semba, M. C. Yoshida, T. Yamamoto, M. Sasaki, and K. Toyoshima. 1986. Structure, expression, and chromosomal location of the human *c-fgr* gene. *Mol. Cell. Biol.* 6:511-517.
 35. Ohashi, M., S. Aizawa, H. Ooka, T. Ohsawa, K. Kaji, H. Kondo, T. Kobayashi, T. Noumura, M. Matsuo, Y. Mitsui, S. Murota, K. Yamamoto, H. Itoh, H. Shimada, and T. Utakoji. 1980. A new human diploid cell strain, TIG-1, for the research on cellular aging. *Exp. Gerontol.* 15:121-133.
 36. Reddy, E. P., M. J. Smith, and A. Srinivasan. 1983. Nucleotide sequence of Abelson murine leukemia virus genome: structural similarity of its transforming gene product to other *onc* gene products with tyrosine-specific kinase activity. *Proc. Natl. Acad. Sci. USA* 80:3623-3627.
 37. Rottenmire, C. W., J. H. Chen, M. F. Roussel, and C. J. Sherr. 1985. The product of the *c-fms* protooncogene: a glycoprotein with associated tyrosine kinase activity. *Science* 228:320-322.
 38. Rubin, J. B., M. A. Shia, and P. F. Pilch. 1983. Stimulation of tyrosine-specific phosphorylation *in vitro* of insulin-like growth factor-I. *Nature (London)* 325:438-440.
 39. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
 40. Schultz, A. M., L. E. Henderson, and S. Oroszlan. 1985. Amino terminal myristylation of the protein kinase $p60^{src}$, a retroviral transforming protein. *Science* 227:427-429.
 41. Semba, K., M. Nishizawa, N. Miyajima, M. C. Yoshida, J. Sukegawa, Y. Yamanashi, M. Sasaki, T. Yamamoto, and K. Toyoshima. 1986. *yes*-related protooncogene, *syn*, belongs to the protein-tyrosine kinase family. *Proc. Natl. Acad. Sci. USA* 83:5459-5463.
 42. Semba, K., Y. Yamanashi, M. Nishizawa, J. Sukegawa, M. Yoshida, M. Sasaki, T. Yamamoto, and K. Toyoshima. 1985. Location of the *c-yes* gene on the human chromosome and its expression in various tissues. *Science* 227:1038-1040.
 43. Sherr, C. J., C. W. Rettenmire, 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.
 44. 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.
 45. 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 ($pp60^{v-src}$) and its normal cellular homologue ($pp60^{c-src}$). *Proc. Natl. Acad. Sci. USA* 78:6013-6017.
 - 45a. Sukegawa, J., K. Semba, Y. Yamanashi, M. Nishizawa, N. Miyajima, T. Yamamoto, and K. Toyoshima. 1987. Characterization of cDNA clones for the human *c-yes* gene. *Mol. Cell. Biol.* 7:41-47.
 46. Takeya, T., and H. Hanafusa. 1983. Structure and sequence of the cellular gene homologous to the RSV *src* gene and the mechanism for generating the transforming virus. *Cell* 32:881-890.
 47. Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* 77:5201-5205.
 48. Ullrich, A., J. R. Bell, E. Y. Chen, R. Herrera, L. Liao, M. Tsubokawa, A. Mason, P. H. Seeburg, C. Grunfeld, O. M. Rosen, and J. Ramachandran. 1985. Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature (London)* 313:756-761.
 49. 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 epidermal carcinoma cells. *Nature (London)* 309:418-425.
 50. Voronova, A. F., and B. M. Sefton. 1986. Expression of a new tyrosine protein kinase is stimulated by retrovirus promoter insertion. *Nature (London)* 319:682-685.
 51. Wang, J. Y. J., F. Ledley, S. Gaff, R. Lee, Y. Groner, and D. Baltimore. 1984. The mouse *c-abl* locus: molecular cloning and characterization. *Cell* 36:349-356.
 52. Yamamoto, T., T. Nishida, N. Miyajima, S. Kawai, T. Ooi, and K. Toyoshima. 1983. The *erbB* gene of avian erythroblastosis virus is a member of the *src* gene family. *Cell* 35:71-78.
 53. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33:103-119.