Oncogene (1991), 6, 361–369

Conservation of structure and expression of the c-*yes* and *fyn* genes in lower vertebrates

Gerhard Hannig[1], Sabine Ottilie[1] & Manfred Schartl

Gene Center, Max-Planck-Institute for Biochemistry, D-8033 Martinsried, Germany

The *src*-gene family in mammals and birds consists of 9 closely related protein tyrosine kinases. We have cloned the c-*yes* and *fyn* homologues of the *src*-family from the teleost fish *Xiphophorus helleri*. Both genes show a high degree of sequence conservation and exhibit all structural motifs diagnostic for functional *src*-like protein tyrosine kinases. Sequence comparisons revealed three domains (exon 2, exons 3–6, exons 7–12) which evolve at different rates. Both genes exhibit an identical expression pattern, with preferential expression in neural tissues. No transcripts of c-*yes* were found in liver which is contrary to the situation in higher vertebrates. In malignant melanoma, elevated levels of c-*yes* and *fyn* were detected indicating a possible function during secondary steps of tumor progression for *src*-related tyrosine kinases.

Introduction

Protein kinases are a large group of enzymes, many of which have been implicated in mediating the response of eukaryotic cells to external stimuli. Besides this salient feature these enzymes have attracted additional attention, because several are considered to represent proto-oncogenes and/or oncogenes. Interestingly, approximately half of the proto-oncogenes/oncogenes known to date have been identified to code for proteins with kinase activity, mostly with specificity to phosphorylate tyrosine. Within these, the family of *src*-related tyrosine kinases is quite well studied (for review see Hunter & Cooper, 1985; Hunter, 1987). *src*-related tyrosine kinases have been found to appear first during phylogenesis in the most simple multi-cellular organisms, the sponges (Schartl & Barnekow, 1982). In higher vertebrates they constitute a closely related gene family of nine members identified to date: the prototypic c-*src*, c-*yes*, c-*fgr*, *fyn*, *lyn*, *lck*, *hck*, *tkl* and *bkl*. All encode proteins of similar size, which can be subdivided in a carboxyterminal catalytic domain of approximately 250–300 amino acids, being structurally extremely highly conserved among the different family members, and a 'regulative' amino terminal domain which is less conserved and thought to specify the supposed functional diversity within the different members. All avian and mammalian family members share a common genomic organization, i.e. common exon/intron arrangement and exon sizes. Besides exerting a normal physiological function, several genes have been found to act as oncogenes in tumors of viral and/or non-viral origins (see Hunter & Copper, 1985).

We have undertaken an evolutionary approach to contribute to an understanding of the function in normal and neoplastically transformed cells of *src*-related genes. Determination of amino acid residues or motifs — besides those generally diagnostic for tyrosine kinases (Hanks *et al.*, 1988) — which are conserved over large evolutionary distances might help to delineate functionally important structures. Comparative studies on changes or conservation in gene expression patterns should elucidate specific features of the respective gene's function. Molecular evolutionary analysis will unravel the structural development of this multi-gene family.

As an experimental system we use the teleost fish *Xiphophorus*, because of its uniqueness for studies on normal and neoplastic development in lower vertebrates. The *src*-gene has been cloned and its function has been studied during normal development and tumorigenesis (Raulf *et al.*, 1989a, b; Mäueler *et al.*, 1988a, b; Schartl *et al.*, 1985; Schartl & Barnekow, 1984) allowing comparative analysis with other family members. The *Xiphophorus* genome is thought to represent the basic vertebrate genome (Ohno *et al.*, 1967) making this organism also very suitable for evolutionary studies.

In this study we have concentrated on the *yes* and *fyn* genes. The *yes*-gene has been isolated from human (Sukegawa *et al.*, 1987), chicken (Sudol *et al.*, 1988b; Zheng *et al.*, 1989), frog (Steele *et al.*, 1989), and as a viral oncogene from Esh sarcoma virus (Wallbank *et al.*, 1966) and Yamaguchi 73 virus (Iothara *et al.*, 1987). The c-*yes* gene shows the highest structural similarity to c-*src*. Like $pp60^{c\text{-}src}$, $p61^{c\text{-}yes}$ is attached to the inner face of the cytoplasmic membrane via myristylation of the glycine-2 residue (Sudol & Hanafusa, 1986; Sudol *et al.*, 1988a). Like c-*src*, c-*yes* is preferentially expressed in neural tissues. However, high levels have also been reported from liver and kidney (Gessler & Barnekow, 1984; Semba *et al.*, 1986; Kawakami *et al.*, 1986). Unlike c-*src*, c-*yes* neuronal expression is considerably higher in adults than in embryos (Sudol *et al.*, 1988a). In general, c-*yes* transcripts are approximately 5 times more abundant than those from any other *src* gene family member (Gessler & Barnekow 1984; Shibuya *et al.*, 1982). Concerning activation of c-*yes* in tumors of non-viral origin, a single report demonstrates c-*yes* amplification in human gastric carcinoma (Seki *et al.*, 1985). High levels of c-*yes* expression have been reported in some tumor cell lines (Semba *et al.*, 1985; Kypta *et al.*, 1988).

The *fyn* gene has so far only been isolated from man (Kawakami *et al.*, 1986; Semba *et al.*, 1986) and frog (Steele *et al.*, 1990). It is most closely related to c-*src* and c-*yes* and $p59^{fyn}$ is analogously myristylated (Kypta *et*

[1] Present address: Harvard University, Department of Cell and Developmental Biology, Cambridge, Massachusetts 02138, USA
This manuscript contains part of the PhD thesis of G.H. and S.O.
Correspondence: M. Schartl, Max-Planck-Institut für Biochemie, Am Klopferspitz 18 A, D-8033 Martinsried, Germany
Received 24 July 1990; accepted in revised form 5 October 1990

al., 1988). In man, highest expression of *fyn* was found in brain, placenta and fibroblasts (Semba *et al.*, 1986). Although no naturally occurring *fyn*-containing transforming retrovirus is known, the oncogenic potential of the *fyn* protein tyrosine kinase has been demonstrated *in vitro* (Kawakami *et al.*, 1986; 1988). In addition, several tumor cell lines exhibit high levels of *fyn* expression (Kawakami *et al.*, 1986; Kypta *et al.*, 1988; Semba *et al.*, 1986).

Results

Isolation and characterization of the c-yes *and* fyn *proto-oncogenes of* Xiphophorus

The fish homologues of c-*yes* and *fyn* were isolated from a brain cDNA library of *Xiphophorus helleri* using a genomic fragment of the fish c-*yes* (X*yes*) and the tyrosine kinase domain of v-*src* for screening. The genomic clone of X*yes* (22-1) has been isolated from a *X. maculatus* genomic library due to its cross-hybridization to the v-*src* probe.

The X*yes* cDNA clone contains a single long open reading frame (ORF) of 1631 nt. There are two other methionine codons located upstream from the potential functional ATG (according to Kozak, 1984, 1986), but they are followed by a termination codon either six codons downstream or immediately. Conceptual translation of the long ORF predicts a protein of 544 aa with a relative molecular mass of 61288 dalton. Sequence comparisons of the X*yes* cDNA sequence with the genomic X*yes* clone revealed (a) that the trailer is encoded immediately adjacent to the translation stop signal in a large 3′ exon and (b) the consensus sequence TGTGTTT (McLauchlan *et al.*, 1985) is following 15 bp downstream the polyadenylation site which is generally located at this position and is assumed to have a regulatory function in 3′ mRNA processing. The exon/intron arrangement in the kinase domain (exons 7-12) of the X*yes* gene is identical to that of the *src* gene family members in higher vertebrates (Figure 1a, 2). The genomic X*yes* clone as probe under conditions of moderate stringency on chicken DNA in Southern blot analysis revealed a band which was also detected with the v-*yes* gene under conditions of high hybridization stringency (Figure 1b).

The X*fyn* clone (Figure 3) contains an ORF of 1614 nt starting with a methionine codon which shows a perfect match with Kozak's consensus sequence (Kozak, 1984; 1986). There are two further methionine codons located upstream, but none of these is flanked by nucleotides that favor initiation of translation, and the ORFs that follow are terminated eight and four codons downstream, respectively. Such short ORFs upstream of the translated ORF have been found to be characteristic of proto-oncogenes in higher vertebrates, and are obviously also conserved in fish (see also X*yes*). The predicted protein consists of 537 aa with a relative molecular mass of 60447 daltons.

The predicted proteins (Figure 4, 5) encoded by X*yes* and X*fyn* contain besides conserved SH2 and SH3-domains all structural motifs diagnostic for *src*-related kinases (see Hanks *et al.*, 1988) including tyrosine phosphorylation residues (X*yes*: tyr 538, tyr 427, X*fyn*: tyr 531, tyr: 421) and the lysine residue likely to be involved

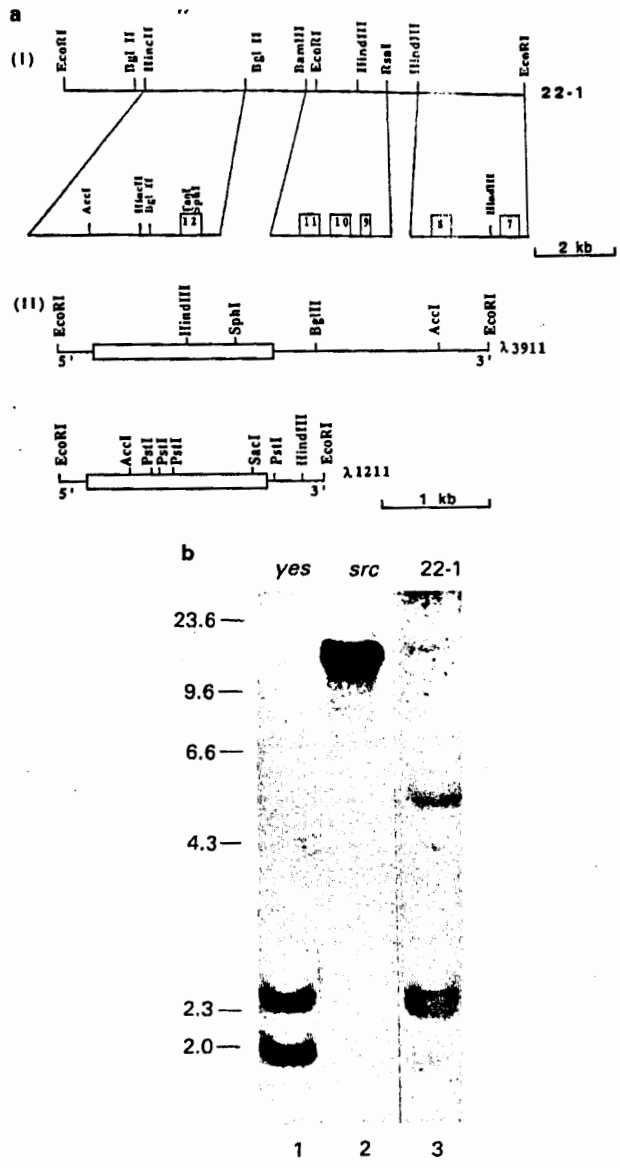

Figure 1 (a) (I) Restriction map of clone 22-1 containing the kinase domain encoding sequences of the X*yes* genomic locus. Numbered squares indicate the positions of exons 7–12. (II) Restriction maps of cDNA clones from the *Xiphophorus* c-*yes* (X*yes* lambda 3911) and *fyn* (X*fyn* lambda 1211) genes. Open boxes indicate the predicted translated regions. Kb, kilobase pairs. (b) Southern analysis of EcoRI digested chicken DNA probed with the v-*yes* (lane 1), v-*src* (lane 2) and the genomic clone 22-1 from *Xiphophorus maculatus* (lane 3). Only one filter was used for hybridization, which was stripped and then reused for the subsequent analysis. The faint signal of the 2 kb *yes* fragment in the 22-1 hybridization is due to the fact that the v-*yes* probe covers further amino terminal sequences which are not contained within clone 22-1. Cross-hybridization to the chicken *src* sequence and a so far unidentified sequence gives rises to the 5 kb and 15 kb bands

in nucleotide binding at analogous positions (X*yes*: lys 306, X*fyn*: lys 299). The deviation of the human *fyn* from a consensus motif in exon 10 (HRDLRSAN: *fyn*, HRDLRAAN: all other *src* related tyrosine kinases) is also conserved in X*fyn* (Figure 5). A similar situation is found for the myristilation recognition sequence where cys-6 (ser in all other *src* related tyrosine kinases) does not fit the consensus in human and fish *fyn*. However, human p59^{fyn} has been shown to be myristylated (Kypta *et al.*, 1988). Thus we anticipate that both the X*yes* and X*fyn* encoded proteins are myristylated and are bound to the plasma membrane.

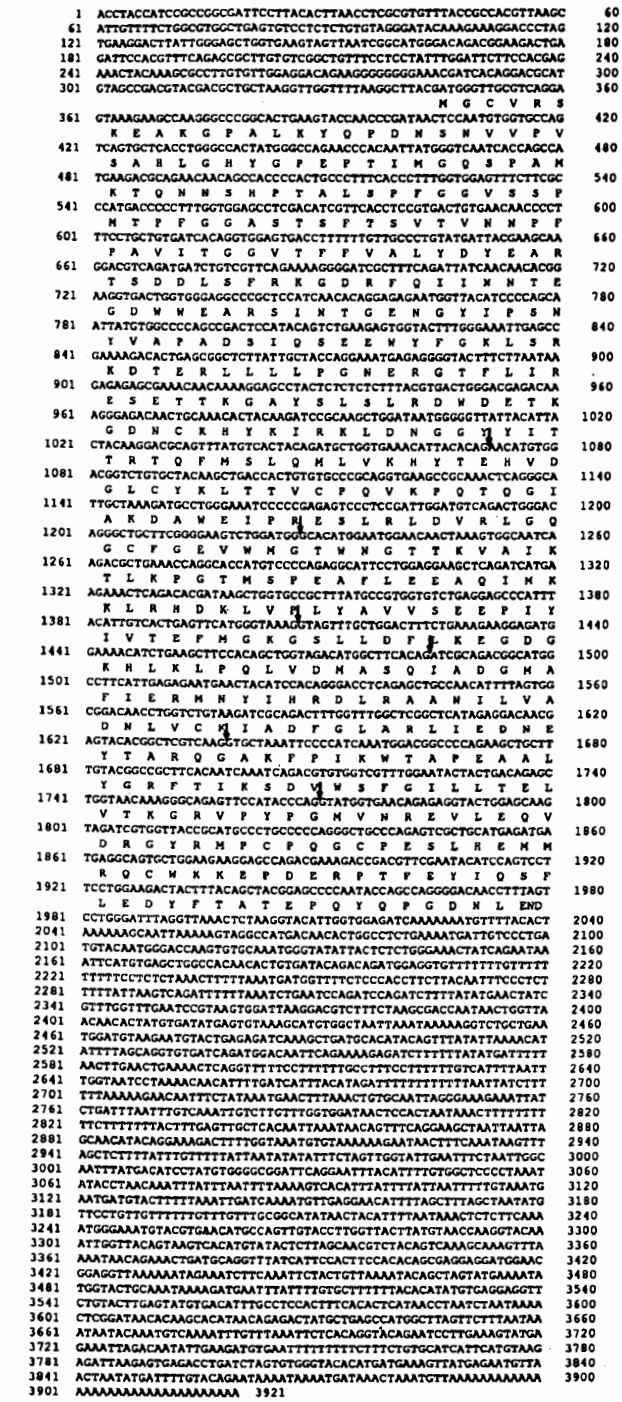

Figure 2 Sequence of the *Xiphophorus* c-*yes* (X*yes*) and predicted protein p61^{Xyes}. Exon boundaries of exons 7–12 as derived from the genomic sequence are marked by arrows above the nt sequence. In the 3'-untranslated region besides the predicted functional hexanucleotide sequence (AATAAA) for polyadenylation of mRNAs (Proudfoot & Brownlee, 1976) at position 3866, eight additional sequences following the polyadenylation consensus were identified

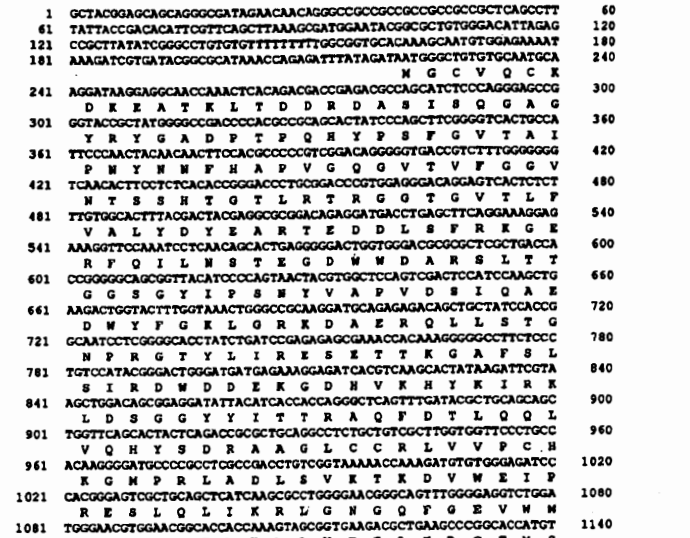

Figure 3 Sequence of the *Xiphophorus fyn* (X*fyn*) and predicted protein p60^{Xfyn}. In the 3'-untranslated region no polyadenylation signal consensus is found, but two hexanucleotide motifs (AAGAAA: positions 2406–2411; ATTAAA: positions 2369–2374) only differing in one position from the common consensus sequence were identified. Such motifs have already been demonstratd to serve as functional polyadenylation sites (Montell *et al.*, 1983; Birnstiel *et al.*, 1985)

Sequence comparison and estimations on evolutionary rates

Sequence comparison of the two fish genes with eachother and with the members of the *src*-gene family of higher vertebrates revealed three domains with respect to similarity values: one, a region which corresponds to the exon 2 encoded sequence of *src*-related genes, two, a domain, encoded by exons 3–6 and containing the SH-2 and SH-3 regions, three, a carboxy terminal domain, corresponding to exons 7–12, which encompasses the kinase domain. Both fish genes show a considerable similarity (Table 2) to each other comparable to the values obtained if the members of the *src*-gene family of birds and mammals are compared with eachother. In the exon 2 encoded domain as in the higher vertebrate genes, the similarity values drop to very low at the nt level, and are even insignificant with respect to the amino acid sequences. If the fish genes are compared with the individual members of the *src*-gene family of higher vertebrates (Tables 1 and 2) highest values are obtained for the kinase domain. In the amino terminal domain values are lower. Although several avian and mammalian genes show more or less equally high similarity with one of the fish genes within these two domains, unequivocal identification of the fish tyrosine kinases as the c-*yes* and *fyn* homologues was possible

Table 1 Similarity of the *Xiphophorus* c-*yes* gene to genes of the *src* tyrosine-kinase family of higher vertebrates

	Exon 2		*Exons 3–6*		*Exons 7–12*	
	nt	*aa*†	*nt*	*aa*†	*nt*	*aa*†
c-*yes* (human)	59.0%	51.7%	73.1%	84.8%	77.5%	93.1%
c-*yes* (chicken)	54.5%	51.7%	72.6%	84.1%	77.9%	93.7%
c-*yes* (frog)	51.2%	50.0%	73.3%	85.4%	76.3%	92.7%
c-*src* (human)	*	*	69.3%	72.7%	77.6%	85.8%
c-*src* (chicken)	*	*	69.1%	72.9%	76.0%	84.4%
fyn (human)	44.7%	32.9%	73.1%	80.0%	74.2%	81.8%
c-*fgr* (human)	42.5%	*	68.3%	75.8%	72.4%	76.1%
hck (human)	*	*	60.9%	58.7%	67.7%	68.3%
tkl (chicken)	*	*	56.3%	50.0%	66.5%	67.7%
lck (mouse)	44.9%	*	55.8%	51.4%	64.9%	66.2%
lyn (human)	*	*	59.7%	52.4%	64.8%	65.3%

* Similarity values below 30%. Such values were not regarded as significant, especially because multiple gaps had to be introduced into the sequence alignment
† Similarity values include conservative changes according to the UWGCG programme

Table 2 Similarity of the *Xiphophorus fyn* gene to genes of the *src* tyrosine-kinase family of higher vertebrates

	Exon 2		*Exons 3–6*		*Exons 7–12*	
	nt	*aa*†	*nt*	*aa*†	*nt*	*aa*†
fyn (human)	78.9%	84.1%	77.6%	91.3%	80.6%	95.4%
c-*src* (human)	44.0%	*	69.2%	73.3%	81.9%	81.8%
c-*src* (chicken)	51.5%	*	69.1%	74.7%	80.7%	84.8%
c-*yes* (human)	45.9%	*	66.4%	80.5%	71.0%	83.4%
c-*yes* (chicken)	46.3%	*	71.1%	79.3%	74.2%	84.8%
c-*yes* (frog)	41.2%	46.7%	70.0%	80.1%	70.0%	84.4%
c-*yes* (fish)	45.3%	*	70.2%	77.3%	78.9%	83.1%
c-*fgr* (human)	52.7%	40.3%	76.0%	81.2%	75.2%	76.1%
tkl (chicken)	*	*	61.6%	41.2%	73.9%	68.6%
hck (human)	38.9%	*	62.7%	67.6%	71.9%	68.7%
lyn (human)	44.5%	*	61.6%	53.0%	66.4%	67.2%
lck (mouse)	38.5%	*	64.5%	55.9%	69.9%	66.0%

* Similarity values below 30%. Such values were not regarded as significant, especially because multiple gaps had to be introduced into the sequence alignment
† Similarity values include conservative changes according to the UWGCG programme

due to the diagnostically high similarity in the exon 2 encoded domains.

The availability of full sequence information of c-*yes* and *fyn* of fish and higher vertebrates allows the comparison of the orthologous gene pairs (fish c-*yes*/human c-*yes*, fish *fyn*/human *fyn* etc.) for estimation of genetic distances and gene divergence rates (Table 3). Detailed analysis revealed that the three domains that became already apparent due to their differing similarity values, have to be considered independently for such kind of molecular evolutionary analysis. For c-*yes*, the divergence rates on the nt-level are constant for all domains. Lowest values are obtained for the kinase domain and the exon 3–6 encoded domain. The exon 2 encoded sequences diverge, however, much faster. On the aa-level in the exon 2 region, even higher rates are apparent while the other two domains diverge at lower rates. The *fyn* gene in general has a decreased speed of evolution in all domains, most marked for the exon 2 region reaching on the nt-level the values of the other two domains.

*Expression of X*fyn *and X*yes

To exclude cross-hybridization, gene specific probes from the 3′ untranslated regions of X*yes* and X*fyn* were used. Both genes showed an identical expression pattern (Figure 6). Highest levels of transcripts were found in melanoma biopsies and in a melanoma cell line (PSM). High amounts were also seen in eyes, brain and late organogenesis stage embryos. Low expression was detected in gills. No or barely detectable amounts were present in muscle, liver and fins. For X*src* in confirmation of earlier results (Mäueler *et al.*, 1988a), also a similar expression pattern was found (data not shown),

Table 3 Genetic distance measurements and divergence rates for the c-*yes* and *fyn* genes

	Nucleotides						*Amino acids*					
	*Genetic distance**			*Divergence rate*†			*Genetic distance**			*Divergence rate*†		
Orthologous gene pairs	*exon 2*	*exons 3–6*	*exons 7–12*	*exon 2*	*exons 3–6*	*exons 7–12*	*exon 2*	*exons 3–6*	*exons 7–12*	*exon 2*	*exons 3–6*	*exons 7–12*
c-*yes*												
fish/human	41.0(59.0)	26.9(32.9)	22.5(26.6)	7.4	4.1	3.3	47.3(72.5)	15.3(17.0)	6.3(6.4)	9.1	2.1	0.8
fish/chicken	45.5(69.5)	27.4(33.8)	22.1(26.0)	8.7	4.2	3.3	48.4(75.0)	15.3(17.0)	5.6(5.8)	9.4	2.1	0.7
fish/frog	48.8(78.2)	26.7(32.7)	23.7(28.4)	9.8	4.1	3.6	53.0(88.0)	14.6(16.1)	7.3(7.5)	11.4	2.0	0.9
frog/human	29.8(37.6)	16.8(18.7)	17.2(19.4)	6.3	3.1	3.2	32.9(42.8)	5.3(5.4)	4.7(4.8)	7.1	0.9	0.8
chicken/human	22.5(26.6)	13.0(14.1)	14.8(16.3)	6.3	3.6	3.9	33.3(43.6)	4.7(4.7)	2.3(2.3)	10.4	1.1	0.5
fyn												
fish/human	2.15(26.2)	22.4(26.5)	19.4(22.2)	3.2	3.3	2.8	15.9(17.7)	8.7(9.2)	4.6(4.7)	2.2	1.2	0.6

* Values corrected according to Dayhoff (1978) in brackets
† Changes/100 residues/100 Myr

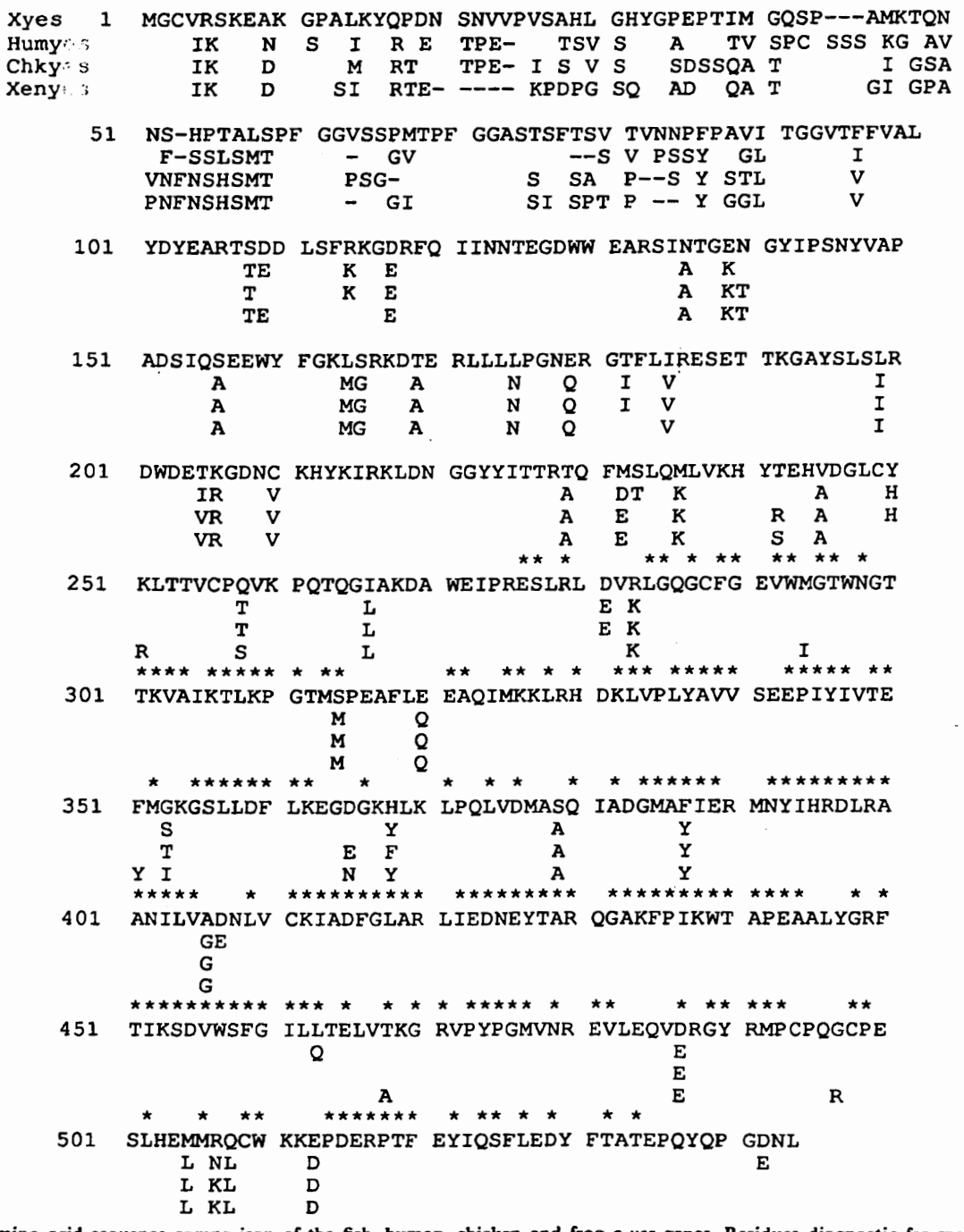

Figure 4 Amino acid sequence comparison of the fish, human, chicken and frog c-*yes* genes. Residues diagnostic for *src*-related tyrosine kinases of vertebrates (present in at least 7 of 8 gene family members, Hanks *et al.*, 1988) are marked by an asterisk above the fish sequence

with the exception of intermediate amounts of X*src* transcripts in gills and fins.

Discussion

We have cloned the c-*yes* and *fyn* homologues from the teleost fish *Xiphophorus*. Unambiguous identification of both genes was possible due to the high structural conservation in the exon 2 encoded region, which is shared only with the homologous gene family member of higher vertebrates. If conservation of structure means conservation of function this would point to the interpretation that some of the specific features which distinguish the individual *src* family members may be encoded in that region. Evidence for this comes from the c-*yes* gene comparisons (Figure 4) which delineate a variety of amino acids that are conserved in the different c-*yes* genes representing more than 800 million years of independent evolution. This conservation in exon 2 is even more apparent from the *fyn* gene comparisons (Figure 5). More than half of the amino acid exchanges are clustered in a small region of 11 amino acids adjacent to the myristilation motif, indicating again that the rest of the exon 2 encoded sequence displays some gene specific functions, which might be more specific than a simple 'spacer function' (Steele *et al.*, 1989) which would keep the rest of the protein from the membrane anchorage site.

The relatively constant divergence rates found for the

```
Xfyn    1  MGCVQCKDKE ATKLTDDRDA SISQGAGYRY GADPTPQHYP SFGVTA NY
Humfyn     S             EE  G  LN SS      T              S
Xenfyn                    E  N  LT SL      T            T  T

       51  NNFHAPVGQG VTVFGGVNTS SHTGTLRTRG GTGVTLFVAL YDYEAR DD
               AG     L        S
               TA     L        S

      101  LSFRKGERFQ ILNSTEGDWW DARSLTTGGS GYIPSNYVAP VDSIQA  Y
             H   K        S       E        ET
             Q   K        S       E         T

      151  FGKLGRKDAE RQLLSTGNPR GTYLIRESET TKGAFSLSIR DWDDEKGD V
                          F       F             Y          M
                          F                     Y          M

      201  KHYKIRKLDS GGYYITTRAQ FDTLQQLVQH YSDRAAGLCC RLVVPCHKGM
                    N             E           E
                    N             E           E
                                 **   *   ** *  ** ** **  *  **** **
      251  PRLADLSVKT KDVWEIPRES LQLIKRLGNG QFGEVWMGTW NGTTKVAVKT
             T                                          AN     I
             T                                           N     I
           **** **    ** **  **  * * *** ** ***  *****  ** *  *** *
      301  LKPGTMSPES FLEEAQIMKK LRHDKLVQLY AVVSEEPIYI VTEYMSKGSL
                                  K                          N
                                  K
           *****   *     *  * *   **  ***** **  ****** ********
      351  LDFLKDGEGR ALKLPNLVDM AAQVAAGMAY IERMNYIHRD LRSANILVGD
                                                                N
                                        R                       N
            * ******* *** ****** *** ****** *******    * ********
      401  NLVCKIADFG LARLIEDNEY TARQGAKFPI KWTAPEAALY GRFTIKSDVW
           G I
           G I
           ****** *   * ** *****  * **    *  *****     **  *   *
      451  SFGILLTELV TKGRVPYPGM NNREVLEQVE RGYRMPCPQD CPASLHELML
                                                        I       I
                                                        I

            **   **** *** * ** *  *  * *
      501  QCWKKDPEER PTFEYLQAFL EDYFTATEPQ YQPGDNL
           H                   S                E
           N                   G
```

Figure 5 Amino acid sequence comparison of the fish, frog and human *fyn* genes. Residues diagnostic for *src*-related tyrosine kinases of vertebrates (present in at least 7 of 8 gene family members, Hanks *et al.*, 1988) are marked by an asterisk above the fish sequence

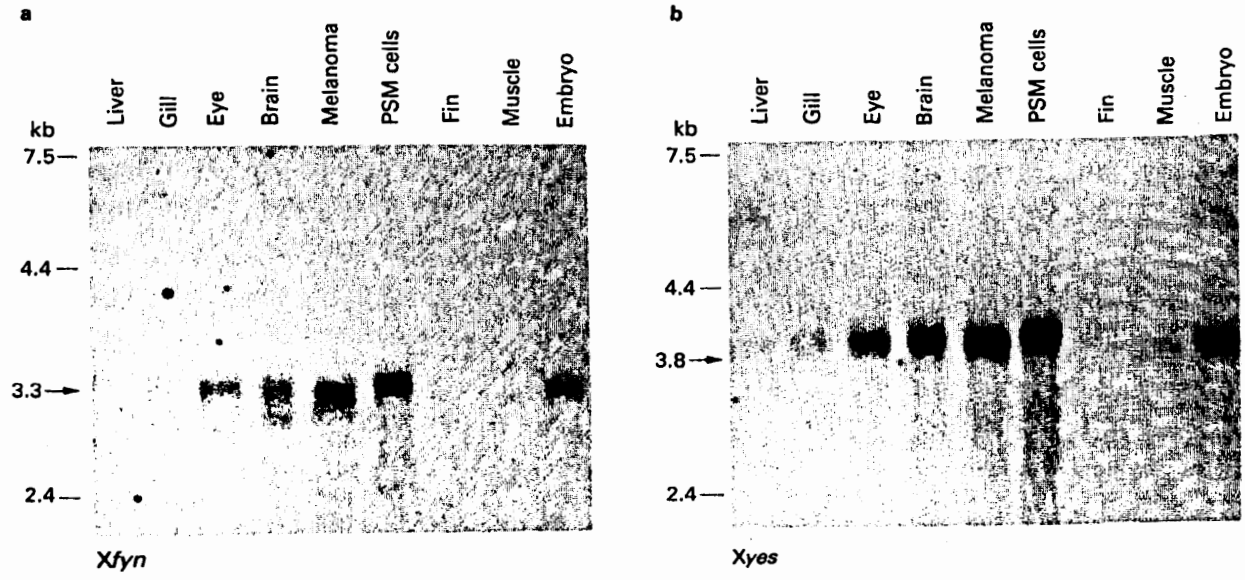

Figure 6 Northern blot analysis of *Xfyn* (**a**) and *Xyes* (**b**) expression in normal organs, melanoma cells and total embryos (stage 20, according to Tavolga, 1949). For hybridization in **a** the 0.75 kb EcoRI/SacI fragment and in **b** the 1.2 kb EcoRI/BglII fragment, both from the 3′ untranslated region were used. For size calibration an RNA ladder (BRL, Bethesda) was used. kb, kilobase pairs. PSM, platyfish-swordtail melanoma cell line

c-*yes* genes at the nt level are in accordance with the concept of the molecular clock (Wilson *et al.*, 1987). The different rates for the different domains, however, demonstrate that the clock has differing periods in specific regions of the gene, most likely evoked by domain-specific selective pressure. For the *fyn* gene the period of the molecular clock appears constant throughout the whole sequence, indicating, that since the divergence of the fish and mammal lineage only a similarily low variation for all three domains is tolerated. The slowing down of the divergence rates on the aa-level for the amino-terminal and kinase domains of c-*yes* during evolution indicates that the protein has reached an optimal structure to exert its function which apparently tolerates only very little variation. The comparable low values from the fish human *fyn* comparison may be interpreted accordingly as if *fyn* has reached such structural improvement on a much earlier phylogenetic level.

The *src*-family of protein tyrosine kinases has obviously arisen by repeated gene duplication events and subsequent sequence divergence (Hunter & Cooper, 1985; Hanks *et al.*, 1988). The deviation from a common ancestor for the vertebrate members of the *src*-family is not only evident from the high degree of sequence conservation, but also from common exon/intron arrangement and identical exon sizes. This genomic organization was also shown for the kinase domain of the fish c-*yes* gene. The presence of independent c-*src*, c-*yes* and *fyn* genes in *Xiphophorus* places at least two of the postulated gene duplication events prior to the divergence of the lineages leading to modern day teleosts and to the higher vertebrates (approx. 400 million years ago). Using the comparison dates from the paralogous and orthologous gene pairs from c-*yes* and *fyn* a divergence of both genes can be extrapolated back to -430 to -520 Myr. The very high conservation of the fish *fyn* gene tempts us to assume that this gene mirrors more the structure of the ancestral gene.

For c-*src* of higher vertebrates it has been found that neuronal cells, because of differential splicing, express an additional transcript which between the exon 3 and 4 encoded sequences has a 18 bp insertion coding for 6 hydrophobic amino acids (Martinez *et al.*, 1987; Levy *et al.*, 1987). Such a neuron-specific alternative splicing product of c-*src* has also been found in teleost fish (Raulf *et al.*, 1989a). The neuron-specific transcript is generally accepted to be instrumental in some neuronal function. As both c-*yes* and *fyn* are closely related to c-*src* and show a similar preferential expression in neuronal cells, the question arises whether they also encode a similar neuronal form. Both fish genes have been cloned from a brain cDNA library, but no evidence for an insertion between the putative exon 3/4 border was obtained. Although this does not exclude that other cDNAs might exist, it is consistent with the finding that fish brain does contain only one form of c-*yes* mRNA and with the view that the neuron-specific insert arose only after the c-*yes* gene has diverged from c-*src*, and is therefore unique to c-*src* (Raulf *et al.*, 1989a).

The fish c-*yes* and *fyn* genes show an identical expression pattern. The same pattern has been detected for the fish c-*src* gene (Mäueler *et al.*, 1988a). The preferential expression in neuronal tissues which has been found in higher vertebrates is also evident in *Xiphophorus* and appears therefore as a evolutionary old feature of this subgroup of *src*-like tyrosine kinases, consistent with high *src*-like kinase activities in nerve cells of all metazoans tested (Schartl & Barnekow, 1984) even in the most primitive organism having a developed nerve system, the coelenterate Hydra (Schartl *et al.*, 1989). Similar to c-*src*, no expression of c-*yes* in liver of adult fish was seen. This is in contrary to findings in higher vertebrates (Gessler & Barnekow, 1984). Although we do not have the information as to which cell types in the liver are responsible for the high c-src and c-*yes* expression, it is reasonable to assume that this expression reflects an additional function for both genes, which has arisen during vertebrate evolution. The considerable transcript levels, which have been observed for *fyn* in human fibroblasts seem also not to be conserved. Fin tissue, which is a rich source of fibroblasts in fish did not show any detectable amount of *fyn* mRNA. The X*yes* and X*fyn* are like X*src* highly expressed in melanoma cells. This expression most likely does not reflect the embryonal origin of pigment cells from the neural crest as derivatives of the neuroectoderm. For X*src* it was shown that the melanoma cells express only the non-neuronal form of the transcript (Raulf *et al.*, 1989a) and that non-transformed, normal pigment cells do not express the gene at detectable levels (Raulf *et al.*, 1989b). By analogy, we propose that X*yes* and X*fyn* display a tumor specific expression. The melanoma inducing oncogene which is clearly defined in *Xiphophorus* encodes a receptor tyrosine kinase of the EGF-receptor family (Wittbrodt *et al.*, 1989) and is thus different from any *src*-like tyrosine kinase. It has therefore to be considered that if X*src*, X*yes* and X*fyn* have a function in tumorigenesis, that is more related to processes of tumor progression. The striking common expression pattern of X*src*, X*yes* and X*fyn* tempts us to speculate that all three genes might be subjected to a common regulatory machinery. More detailed expression studies and the identification and characterization of the corresponding regulatory sequences are needed in order to understand this phenomenon.

Materials and methods

Experimental animals

Fishes of the genus *Xiphophorus* (Teleostei: Poecliidae) were derived from natural populations that have been maintained as closed stocks under standard conditions. Fishes from the following populations were used: (1) *Xiphophorus helleri* (swordtail) from Rio Lancetilla, Belize and (2) *X. maculatus* (platyfish) from Rio Usumacinta, Mexico.

Northern blot analysis

Total cellular RNA was extracted following the LiCl/urea procedure (Auffray & Rougeon, 1980). 20 μg of total RNA were denatured with formamide/formaldehyde, subjected to electrophoresis in 1.2% agarose gels containing 2.2 M formaldehyde (Lehrach *et al.*, 1977), and electrotransferred to Gene Screen membranes (NEN) according to the supplier's instructions. For exact quantification of the RNA amount each filter was stained with methylene blue (Khandjian, 1986) prior to hybridization. Hybridization and washing was performed as described previously for homologous probes (Mäueler *et al.*, 1988a).

Screening of the cDNA library

A cDNA library (5.5×10^5 independent plaques) from polyA$^+$ RNA from *X. helleri* brain (supplied by W. Mäueler, Martinsried) was constructed as EcoRI fragments inserted into lambda gt10 (Clontech, Palo Alto, Ca). Approximately 2.5×10^5 recombinant phages on duplicate filters were screened under conditions of moderate stringency (hybridization conditions: 40% deionized formamide, 1 M NaCl, 50 mM Tris/HCl, pH 7.5, 5 × Denhardt's (1 × Denhardt's = 0.02% each of Ficoll, polyvinylpyrrolidine and BSA), 1% SDS, 100 μg ml^{-1} heat-denatured calf thymus DNA, 42°C; washing conditions: 60°C, 1 × SSC, 1% SDS) using an exon 10-containing fragment (123 bp HaeIII/RsaI fragment) from a genomic X*yes* clone (22-1) of *X. maculatus* and the viral *src* 612 bp PstI fragment F of RSV SRA-2 (DeLorbe *et al.*, 1980) encompassing most of the tyrosine kinase domain of the gene. The fragments were labelled using random oligonucleotides as primers (Feinberg & Vogelstein, 1983).

DNA sequence analysis

The cDNA inserts from lambda phages were subcloned into Bluescript KS+ (Stratagene GmbH, Heidelberg) for further subcloning or generation of deletion series using ExoIII/ExoVII nuclease (Yanisch-Perron *et al.*, 1985). The nucleotide sequence was determined by the dideoxy chain-termination procedure (Sanger *et al.*, 1977) using nodified T7 DNA polymerase (Sequenase; USB, Clevelan H) and specific oligonucleotide primers.

Evolutionary analysis

Sequence data were analyzed using the versity of Wisconsin sequence analysis program package 3 (Devereux *et al.*, 1984). Calculations of genetic distance ere corrected for potential unobservable changes using th rrection factors of Dayhoff (1978).

Acknowledgements

We thank F. Storch for technical assistan n the breeding of *Xiphophorus* fish, E. Buffler and C. Sitte t typing the manuscript, F. Raulf for PSM RNA and R. E son for critically reading the manuscript. This work was sup rted by grants to M.S. provided by the Stiftung Volksw enwerk through 'Wettbewerb Biowissenschaften' and the ndesministerium für Forschung und Technologie thro 'Schwerpunkt: Grundlagen und Anwendungen der Gente ologie' (No. 26).

Note added in proof

The sequences of the *Xiphophorus* c-*yes* *fyn* genes have been submitted to the EMBL Data Librar d are available under accession numbers X54970 (X*yes*) an 54971 (X*fyn*).

References

Auffray, C. & Rougeon, F. (1980). *Biochemistry*, **107**, 303–314.
Barnekow, A. & Schartl, M. (1984). *Mol. Cell. Biol.*, **4**, 1179–1181.
Birnstiel, M.L., Busslinger, M., & Strub, K. (1985). *Cell*, **41**, 349–359.
Dayhoff, M.O. (1978). *Atlas of Protein Sequence and Structure.* Vol. 5 (Suppl. 3), Natl. Biomed. Res. Foundation.
DeLorbe, W.J., Luciuw, P.A., Goodman, H.M., Varmus, H.E. & Bishop, J.M. (1980). *J. Virol.*, **36**, 50–61.
Devereux, J., Haeberli, P. & Smithies, O. (1984). *Nucleic Acids Res.*, **12**, 387–395.
Feinberg, A.P. & Vogelstein, B. (1983). *Anal. Biochem.*, **137**, 266–267.
Gessler, M. & Barnekow, A. (1984). *Bioscience Reports*, **4**, 757–770.
Hanks, S.K., Quinn, A.M. & Hunter, T. (1988). *Science*, **241**, 42–52.
Hunter, T. & Cooper, J.A. (1985). *Ann. Rev. Biochem.*, **54**, 897–930.
Hunter, T. (1987). *Cell*, **50**, 823–829.
Iothara, S., Hirata, K., Inone, M., Hatsuoka, M. & Sato, A. (1978). *Japanese J. Cancer R. (Gann)*, **69**, 825–830.
Kawakami, T., Pennington, C.Y. & Robbins, K.C. (1986). *Mol. Cell. Biol.*, **6**, 4195–4201.
Kawakami, T., Kawakami, Y., Aaronson, S.A. & Robbins, K.C. (1988). *Proc. Natl. Acad. Sci. USA*, **85**, 3870–3874.
Khandjian, E.W. (1986). *Mol. Biol. Rep.*, **11**, 105–115.
Kozak, M. (1984). *Nucleic. Acids Res.*, **12**, 857–872.
Kozak, M. (1986b). *Cell*, **44**, 283–292.
Kypta, R.M., Hemming, A. & Courtneidge, S.A. (1988). *EMBO J.*, **7**, 3837–3844.
Lehrach, H., Diamond, D., Wozney, J.M. & Boedtker, H. (1977). *Biochemistry*, **16**, 4743–4751.
Levy, J.B., Dorai, T., Wang, L.H. & Brugge, J.S. (1987). *Mol. Cell. Biol.*, **7**, 4142–4145.
Martinez, R., Mathey-Prevot, B., Bernards, A. & Baltimore, D. (1987). *Science*, **237**, 411–415.
Mäueler, W., Raulf, F. & Schartl, M. (1988a). *Oncogene*, **2**, 421–430.
Mäueler, W., Barnekow, A., Eigenbrodt, E., Raulf, F., Falk, H. F., Telling, A. & Schartl, M. (1988b). *Oncogene*, **3**, 113–122.
McLauchlan, J., Gaffney, D., Whitton, J.L. Clements, J.B. (1985). *Nuclei Acids Res.*, **13**, 1347–1368.
Montell, C., Fisher, E.F., Caruthers, M.H. & rk, A.J. (1983). *Nature*, **305**, 600–605.
Ohno, S., Wolf, U. & Atkin, N.B. (1967). *He ditas*, **59**, 169–187.
Proudfoot, N.J. & Brownlee, G.G. (1976). *Nature*, **263**, 211–214.
Raulf, F., Robertson, S.M. & Schartl, M. (1989). *J. Neurosc. Res.*, **24**, 81–88.
Raulf, F., Mäueler, W., Robertson, S.M. & Schartl, M. (1989). *Oncogene Res.*, **5**, 39–47.
Sanger, F., Nicklen, S. & Coulson, A. P. (1977). *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schartl, M. & Barnekow, A. (1982). *Differentiation*, **23**, 109–114.
Schartl, M., Schmidt, C.-R., Anders, A. & Barnekow, A. (1985). *Int. J. Cancer*, **36**, 199–207.
Schartl, M. & Barnekow, A. (1984). *Dev. Biol.*, **105**, 415–422.
Schartl, M., Holstein, T., Robertson, S.M. & Barnekow, A. (1989). *Oncogene*, **4**, 1185–1191.
Semba, K., Nishizawa, M., Miyajima, N., Yoshida, M.C., Sukegawa, J., Yamanashi, Y., Sasaki, M., Yamamoto, T. & Toyoshima, K. (1986). *Proc. Natl. Acad. Sci. USA*, **83**, 5459–5463.
Semba, K., Yamanashi, Y., Nishizawa, M., Sukegawa, J., Yoshida, M., Sasaki, M., Yamamoto, T. & Toyoshima, K. (1985). *Science*, **227**, 1038–1040.
Seki, T., Fujii, G., Mori, S., Tamaoki, N. & Shibuya, M. (1985). *Jpn. J. Cancer Res. (Gann)*, **76**, 907–910.
Shibuya, M., Hanafusa, H. & Baluzzi, P.C. (1982). *J. Virol.*, **42**, 143–152.
Steele, R.E., Irwin, M.Y., Knudson, C.L., Collett, J.W. & Fero, J.B. (1989). *Oncogene Res.*, **4**, 223–233.
Steele, R.E., Deny, I.C., Ghosn, C.R. & Fero, I.B. (1990). *Oncogene*, **5**, 369–376.
Sudol, M. & Hanafusa, H. (1986). *Mol. Cell. Biol.*, **6**, 2839–2846.
Sudol, M., Alvarez-Byulla, A. & Hanafusa, H. (1988). *Oncogene Res.*, **2**, 345–355.
Sudol, M., Kieswetter, C., Zhao, Y.-H., Dorai, T., Wang, L.-H.

& Hanafusa, H. (1988). *Nucleic Acids Res.*, **16**, 9876.
Sukegawa, J., Semba, K., Yamanashi, Y., Nishizawa, M., Miyajima, N., Yamamoto, T. & Toyoshima, K. (1987). *Mol. Cell. Biol.*, **7**, 41–47.
Tavolga, W.N. (1949). *Bull. Americ. Mus. Natl. Hist.*, **94**, 163–229.
Wallbank, A.M., Sperling, F.G., Hubben, K. & Stubbs, E.L. (1966). *Nature*, **209**, 1265.
Wilson, A.C., Ochman, H. & Prager, E.M. (1987). *TIG*, **3**, 241–247.
Wittbrodt, J., Adam, D., Malitschek, B., Mäueler, W., Raulf, F., Telling, A., Robertson, S.M. & Schartl, M. (1989). *Nature*, **341**, 415–421.
Yanisch-Perron, C., Vierra, J. & Messing, J. (1985). *Gene*, **33**, 103–119.
Zheng, X., Podell, S., Sefton, B.M. & Kaplan, P.L. (1989). *Oncogene*, **4**, 99–104.