Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus*

Joachim Wittbrodt, Dieter Adam, Barbara Malitschek, Winfried Mäueler, Friedrich Raulf, Agnes Telling, Scott M. Robertson & Manfred Schartl*

Genzentrum, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-8033 Martinsried, FRG

Malignant melanoma in *Xiphophorus* fish hybrids is caused by the activity of a dominant oncogene *Tu*. By combining genetic and molecular approaches, we have isolated the melanoma oncogene. We show that its level of expression correlates with the degree of malignancy of the tumour. The corresponding proto-oncogene is developmentally regulated. The *Tu* gene codes for a novel receptor tyrosine kinase which is closely related to the receptor for epidermal growth factor.

TUMORIGENESIS is generally considered to be controlled by oncogenes, which specify many of the malignant features of tumour cells. Oncogenes arise mainly by activation of their normal cellular counterparts, proto-oncogenes, which themselves seem to have essential roles in normal cellular physiology and are often involved in the regulation of normal cell proliferation and differentiation. By contrast, there are other genes, which are believed to be critically involved in the negative control of cell growth[1], that must be mutated at both alleles for the neoplastic phenotype of the cell to develop[2,3]. These observations have led to the distinction between recessive oncogenes (otherwise known as tumour suppressors or anti-oncogenes) and dominant oncogenes. Although the structure and molecular function of many oncogenes are well established, the evidence that their activation is indeed the primary cause and not a consequence of tumorigenesis in the intact organism is generally circumstantial rather than direct. In *Drosophila*, genetic and molecular techniques have defined the causative role of the recessive oncogene *lethal (2) giant larvae* in tumorigenesis[4,5]. Until now, however, comparable studies of dominant oncogenes have been limited to tumour viruses and tissue- and organ culture. Here we describe a genetic system, melanoma formation in the fish *Xiphophorus*, which allows the study of the induction and progression of malignant melanoma in vertebrates. We used this system to clone and characterize the dominant tumour-inducing gene, which we show encodes a novel putative receptor tyrosine kinase (RTK) of epidermal growth factor (EGF)-receptor type.

Genetics of melanoma formation

The platyfish (*Xiphophorus maculatus*) and the swordtail (*Xiphophorus helleri*) have a uniform grey body colouration due to small black pigment cells ('micromelanophores'). Some platyfish, however, exhibit spot patterns of melanophores that are much larger than normal ('macromelanophores')[6]. In hybrid offspring that result from the mating of spotted platyfish with non-spotted swordtails, these macromelanophore spots develop into malignant melanomas[6-8]. Genetic analyses showed that this

tumour formation depends on the abnormal regulation of a single genetic locus, the macromelanophore locus of the platyfish in hybrids[9-12]. Further genetic studies led to a model for this melanoma development[13], in which a dominant tumour gene *Tu* (equated with the macromelanophore locus) is under the control of a regulatory gene *R* which acts as a tumour suppressor. The gene *Tu* is located on sex chromosomes, whereas *R* is located on an autosome[14,15]. In the platyfish both *Tu* and *R* are present, but they are presumed to be absent in the swordtail. The concerted action of *Tu* and *R* results in non-proliferating macromelanophore spots. A variety of spot patterns have been identified in wild-type fish[16], and the corresponding gene loci are accordingly named *Tu-Sd* (spotted dorsal pattern), *Tu-N* (nigra pattern), *Tu-Sr* (striped pattern), and so on. These loci are located either on the platyfish X or Y chromosome. In this report we mainly look at the *Tu-Sd* locus, which maps to the X chromosome. Serially repeated backcrossing of *Tu-Sd*-bearing platyfish to swordtails results in the progressive replacement of *R*-bearing platyfish chromosomes by *R*-free swordtail chromosomes (Fig. 1a). The stepwise elimination of regulatory genes is thought to lead to a corresponding stepwise deregulation of *Tu*. This leads to either benign melanoma if only one functional allele of *R* remains, or to malignant melanoma if *R* is totally absent (for review see ref. 17).

Mapping of a marker to the *Tu–Sd* gene locus

The *Tu–Sd* locus has been mapped relative to several other known loci on the X chromosome of the platyfish as follows: cen-*P*-1%-*Ir*-(*Pt*–*Dr*)-0.03%-(*Tu*–*Sd*)-ter, where the physical distance between the loci is represented by the recombination fraction (%) between them, as indicated (see refs 16, 18-20, and our own unpublished data). We have previously identified a restriction fragment length polymorphism (RFLP)[21], with limited sequence similarity to a v-*erb-B* gene probe, as a diagnostic marker for sex chromosomes bearing *Tu*. We have cloned a fragment[22] of the corresponding RFLP sequence of the platyfish and used it for Southern blot analyses. DNA digested with *Eco*RI gave an invariant band of 7.0 kilobases (kb), which we have called INV, in all platyfish, swordtails and their hybrids tested (*n* = 135), regardless of the presence or absence of a *Tu* locus. For fish with the X-chromosomal *Tu–Sd* locus (*n* = 84) an additional 5.0 kb band was detected, whereas for fish with the Y-chromosomal *Tu–Sr* locus (*n* = 6), a 6.5 kb band was detected in addition to the 7.0-kb band (Fig. 1b).

To further define the linkage between the *Tu–Sd* locus and the 5.0 kb RFLP marker, we analysed backcross (BC) hybrids of platyfish and swordtails for segregation; swordtails were used as the recurrent parent (BC$_4$-BC$_7$, crossed as in Fig. 1a). No recombination between the 5.0 kb RFLP marker and the *Tu–Sd* locus was observed for all 117 fish tested (Table 1). Both markers, therefore, map on the X chromosome of the platyfish with a maximal separation of 0.85 centimorgans (a recombination fraction of 1% between loci indicates a physical map distance of 1 centimorgan). Our data therefore localize the 5.0 kb marker to the *Tu–Sd*-containing region of the X chromosome. Whereas

FIG. 1 *a*, Crossing scheme of platyfish (*X. maculatus*) and swordtails (*X. helleri*) with the corresponding *Tu/R* constellations (see text). A female *X. maculatus* (A) homozygous for the X-chromosomal macromelanophore *Tu-Sd*-gene (spotted dorsal, small spots in the dorsal fin) is mated to *X. helleri* (B), which does not carry the corresponding locus and exhibits the uniform wild-type pigmentation. The heterozygous F_1 hybrid (C) shows enhancement of the spotted dorsal phenotype. Backcrossing of the F_1 hybrid to *X. helleri* results in offspring, 50% of which have not inherited the *Tu-Sd* locus and are phenotypically like the *X. helleri* parental strain (F), and 50% of which carry the macromelanophore locus and develop melanoma. The malignancy grade ranges from very benign in some individuals (D) (phenotype like the F_1 hybrids), to highly malignant in other individuals (E). Highly malignant melanomas occurring in such fish grow invasively and are fatal. *b*, Southern-blot analyses of *X. maculatus*, *X. helleri* and their hybrids for segregation analyses of invariant and *Tu*-associated RFLP loci. *Eco*RI-digested DNA from single fish was separated on 0.8% agarose gels, blotted and probed with random priming-labelled pXX21 (ref. 22). Hybridization was at 42 °C in 50% formamide, 5 ×SSC solution and the filters were washed at 68 °C in 0.1 ×SSC, 1% SDS. Lane 1; *X. maculatus*, male, *Tu-Sr* (Y-chromosome)/*Tu-Sd* (X-chromosome). Lane 2; *X. maculatus*, female, *Tu-Sd/Tu-Sd*. Lane 3; *X. maculatus/X. helleri* F_1 hybrid, *Tu-Sd/-*. Lane 4; *X. helleri, male, -/-*. Lane 5; BC_4-hybrid, melanoma free, *-/-*. Lane 6; BC_4-hybrid with melanoma, *Tu-Sd/-*.
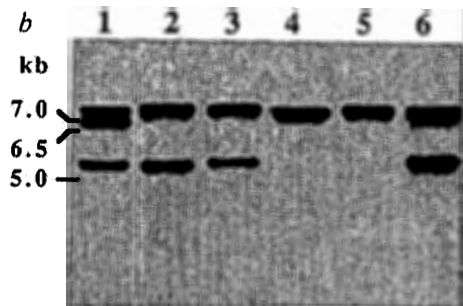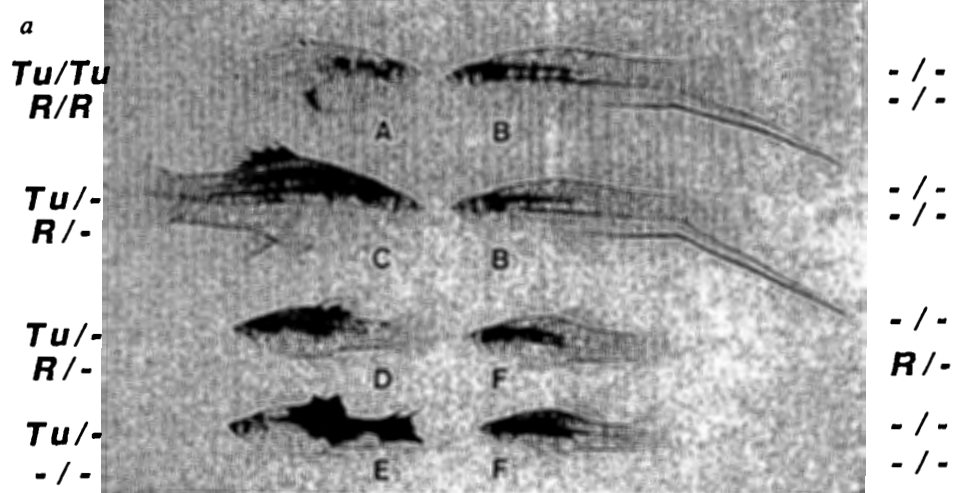
the invariant 7.0-kb sequence is present in all fish, the *Tu-Sd*-specific 5.0 kb sequence seems to be an additional genomic constituent of fish that have the *Tu-Sd* gene locus. By analogy we conclude that the 6.5 kb marker can be assigned to the *Tu-Sd*-containing region of the platyfish Y chromosome (Fig. 1*b*).

Cloning of the *Tu*-linked RFLP sequences

To clarify the relationship of the RFLP sequences to *Tu*, the invariant (INV), Y chromosomal (Y) and X chromosomal (X) *Eco*RI restriction fragments were cloned from a sub-genomic library of male platyfish (X^{Tu-Sd}/Y^{Tu-Sr}). The corresponding partial complementary DNAs, incomplete for the 5' end of the gene, were isolated from two separate cDNA libraries. The first library was generated from a melanoma that had arisen because of the presence of the X-chromosomal *Tu-Sd* locus, and the second library from a cell line (PSM; ref. 23) derived from a melanoma that had arisen because of the presence of a Y chromosomal *Tu* locus. To obtain the entire sequence of one of the cDNAs we used oligonucleotides for preparation of a primer extension library. A 2.4-kb clone, which overlapped the 1.7-kb PSM clone by 98 nucleotides, was isolated and found to contain the missing 5' sequence. The sequencing of the three

TABLE 1   Segregation of *Tu-Sd* and the 5.0 kb marker sequence

			n
Parental class 1	*Tu-Sd*	5.0-kb RFLP marker	76
Parental class 2	+	+*	41
Recombinant class 1	*Tu-Sd*	+	0
Recombinant class 2	+	5.0-kb RFLP marker	0
		Total = 117	

* The 5.0-kb *Eco*RI fragment is additional to the invariant 7.0-kb fragment of wild-type fish, therefore lack of the 5.0-kb RFLP marker is regarded as the wild-type situation (+).

genomic *Eco*RI fragments and their comparison with the cDNA sequences revealed that each genomic clone contained the 3' portion of the same gene (Fig. 2*b*, *c*). Despite a high degree of sequence conservation throughout the exons and the introns (99%) as well as their common arrangements and sizes in homologous regions, several structural differences confirmed the independent nature of the three clones and allowed their assignment to the individual cDNAs. The three *Eco*RI fragments are therefore regarded as deriving from three different genomic loci. Although all the three loci have almost identical restriction-site maps upstream of the *Sac*I site (at nucleotide 1,915) their physical maps downstream of this position differ. Most striking are the additional sequences in the loci of the INV fragment and the Y fragment of 1,349 and 1,395 bp, respectively (Fig. 2*b*). These additions contain intron sequences, exon y and the first five nucleotides of the large 3' exon (exon z) encoding the trailer region. These sequences are deleted from the X chromosomal locus. The cDNA clone from a melanoma that contained the X-chromosomal *Tu-Sd* locus, lacked exon y as expected, whereas the Y-chromosomal cDNA did not (Fig. 2*c*); this confirmed the genomic locus arrangement. The amino-acid sequence predicted from the X-chromosomal cDNA sequence indicated that the deletion does not change the reading frame, and thus represented a polypeptide with an internal deletion of 35 amino acids (see below). Neither cDNA extended to the poly(A) tail, although all three genomic loci contain a poly-adenylation-site signal at corresponding positions. This would add another 229 nucleotides downstream of the cDNA sequence (Fig. 3).

RFLPs are a structural component of the *Tu* locus

That the sex-chromosomal genes are actually encoded by the *Tu* locus is shown by the presence of 'loss of function' *Tu* alleles. We analysed two spontaneous *Tu* mutants that do not exhibit a single macromelanophore and have lost the capacity to develop malignant melanoma in backcross hybrids. In one mutant

derived from the Y-chromosomal *Tu–Sr* locus, the 6.5-kb fragment is totally-deleted (Fig. 2*d*). In another mutant, a rearranged *Eco*RI fragment of 12 kb instead of the expected 5.0 kb, in addition to the invariant 7-kb fragment, was detected with the RFLP probe (Fig. 2*d*). Because we have so far detected no intra-strain polymorphism for the X- or Y-chromosomal marker sequences (see Table 1, ref. 21, and unpublished data) this 12-kb fragment was cloned and analysed. Restriction mapping and sequence analysis precisely defined an insertion of ~7 kb between base-pairs 1,652 and 1,653 in exon w (Fig. 2*e*). This insertion, therefore, interrupts the coding sequence of the

X-chromosomal gene. The result of this alteration is the loss of the *Tu*-phenotype. We conclude therefore that the X-chromosomal 5.0-kb marker sequence is the 3' part of the *Tu–Sd* gene locus, the Y-chromosomal 6.5-kb marker by analogy is the homologous part of *Tu–Sr*, whereas the 7.0-kb marker represents a gene found in all fish independent of sex-chromosomal *Tu* loci.

A novel putative receptor tyrosine kinase

The almost full-length cDNA clone isolated from the PSM cell line contains a single long open-reading frame (ORF) of 3498 nucleotides, which starts with ATG and terminates with an
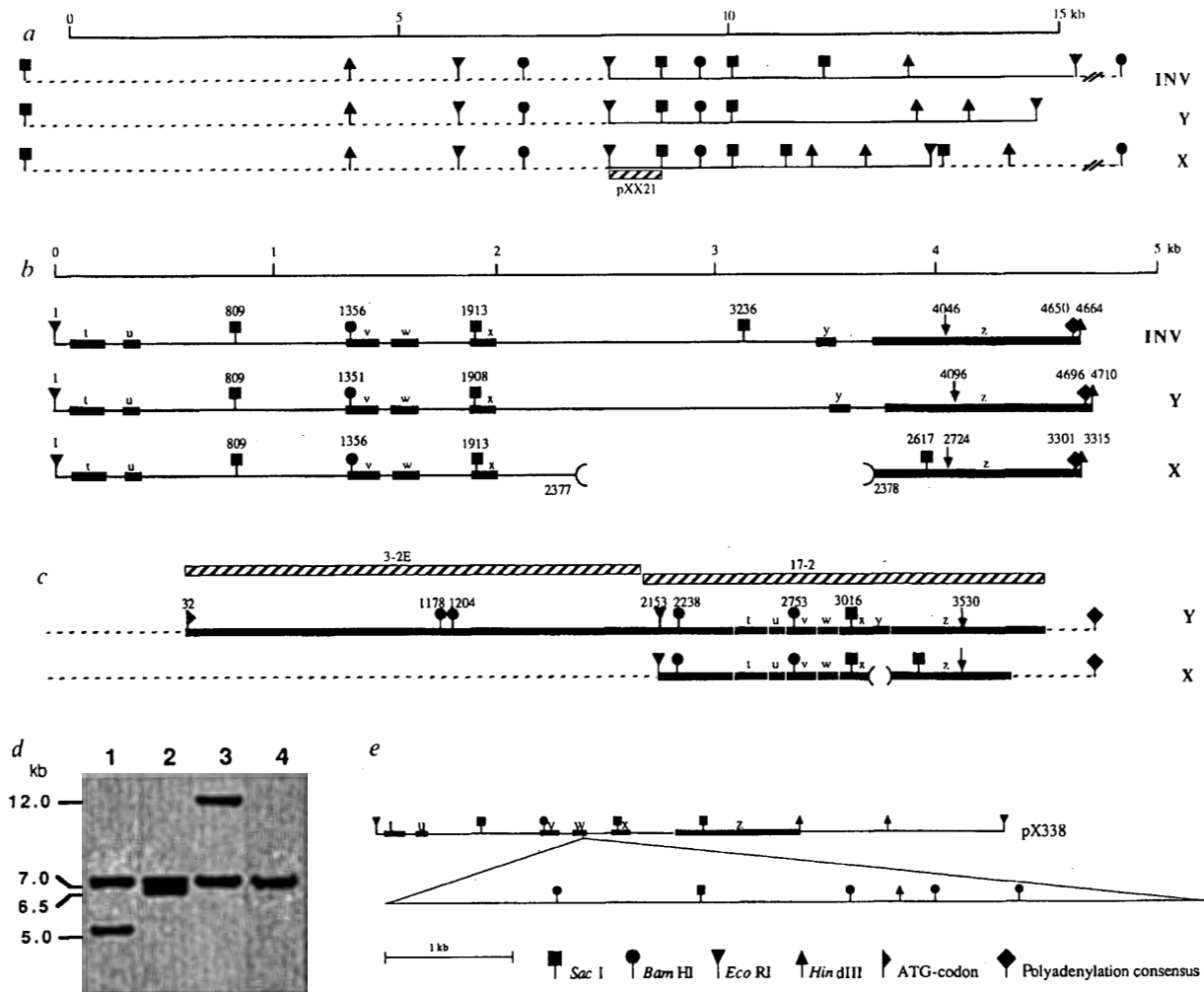

FIG. 2 Partial structure of the genomic region of INV-, Y- and X-linked RFLP loci and corresponding cDNAs. *a*, Restriction map of the genomic loci. *b*, Exon–intron arrangements in the 3' region of INV-, Y- and X-genomic loci. Differing positions of exons, stop codons, restriction sites and polyadenylation consensus sites downstream of nucleotide 1,277 between INV-, X- and Y loci result from sequence differences in the corresponding trailer and intron regions. *c*; Comparison of X- and Y-loci transcripts. Sequenced cDNAs are marked in black. *d*; RFLP analysis of the loss-of-function mutants. Southern-blot analysis of *Eco*RI-digested DNA of *X. maculatus*/*X. helleri* BC hybrids (recurrent parent *X. helleri*) carrying the parental wild-type *Tu–Sd*-bearing X chromosome (BC$_7$, lane 1), *Tu–Sr*-bearing Y chromosome (BC$_2$, lane 2) and the two independent mutant chromosomes (BC$_2$, lanes 3 and 4). For methods see Fig. 1. *e*; Restriction map of the rearranged 12-kb fragment from the loss-of-function mutant (see Fig. 2*d*, lane 3). Relevant stop codons are marked by arrows; deleted fragments are bounded by semi-circles; cloned fragments are indicated by solid lines; structural data obtained by Southern- or Northern-blot hybridizations are shown by broken lines; probes used for screening shown by hatched boxes.
METHODS. INV-, Y- and X-linked fragments were cloned as follows: High molecular weight genomic DNA from male *X. maculatus* (stock Rio Jamapa

X^{Tu-Sd}, Y^{Tu-Sr}) was prepared by standard methods[41]. The DNA was digested to completion with *Eco*RI. Fragments were separated on a LMP-agarose gel. Gel fractions with sizes of ~5 kb, 6.5 kb and 7 kb were cut out; purified DNA was ligated into λgt10 arms and packaged. The sub-libraries were screened with a v-*erb* B probe (600-bp *Bam*HI fragment D of pAE11 ref. 42) as well as with pXX21 (ref. 22). Hybridization conditions were 42 °C, 40% formamide, 5×SSC (wash at 60 °C, 2×SSC) for pAE11, and as described in Fig. 1 for pXX21. Genomic *Eco*RI-*Hin*dIII sub-fragments of INV, Y and X and the corresponding cDNAs were subcloned into pBluescript (Stratagene). For cloning of the 12-kb fragment from the loss-of-function mutant, DNA from a BC$_2$ hybrid containing the mutant chromosome was cloned as outlined above into λEMBL4. The subgenomic library was screened with the 17-2 the probe (Fig. 2*c*) under high stringency conditions. Sequencing: nested deletions were generated using either Exonuclease III (ref. 43), or T4 DNA polymerase[44] (using oligonucleotides adapted for pBluescript). Sequences were determined by dideoxy chain-termination sequencing using Sequenase (US Biochemical Corporation) or T7 DNA polymerase (Pharmacia). Sequence analysis was performed using the GCG sequence analysis software package[45].

417

amber codon (Fig. 3). The deduced amino-acid sequence corresponding to this ORF is a polypeptide of 1,166 amino acids with a predicted relative molecular mass of 130,000 (Fig. 3). This hypothetical protein has two strongly hydrophobic regions (Fig. 3). One of these regions at the N terminus of the protein exhibits several features typical of signal peptides[24]. A possible cleavage site in the uncleaved peptide for signal peptidase, deduced from the −3 /−1 rule[25], is after residue 25. There is a second strongly hydrophobic region, of 23 amino acids, between residues 618 and 640 of the cleaved peptide. By the criteria of Klein et al.[26], this segment has a high probability of spanning a membrane. It is preceded by a positively charged histidine and is followed by four positively charged arginines that could

serve as a transfer-stop signal. These features of the primary structure indicate that the protein traverses a cellular membrane and is therefore comprised of an extracellular domain of 617 amino acids, a single transmembrane domain of 23 amino acids and a cytoplasmic domain of 501 amino acids. There are 12 putative glycosylation sites, all located in the hypothetical extracellular domain of the protein (Fig. 3). Comparison of the predicted polypeptide sequence with sequence data-banks indicates that this polypeptide has significant similarity with receptor tyrosine kinases (RTKs), most markedly with the human EGF-receptor[27]. All other RTKs such as the insulin receptor, the insulin-like growth-factor receptor, the platelet-derived growth factor (PDGF) receptor, and those encoded by HER2/neu and
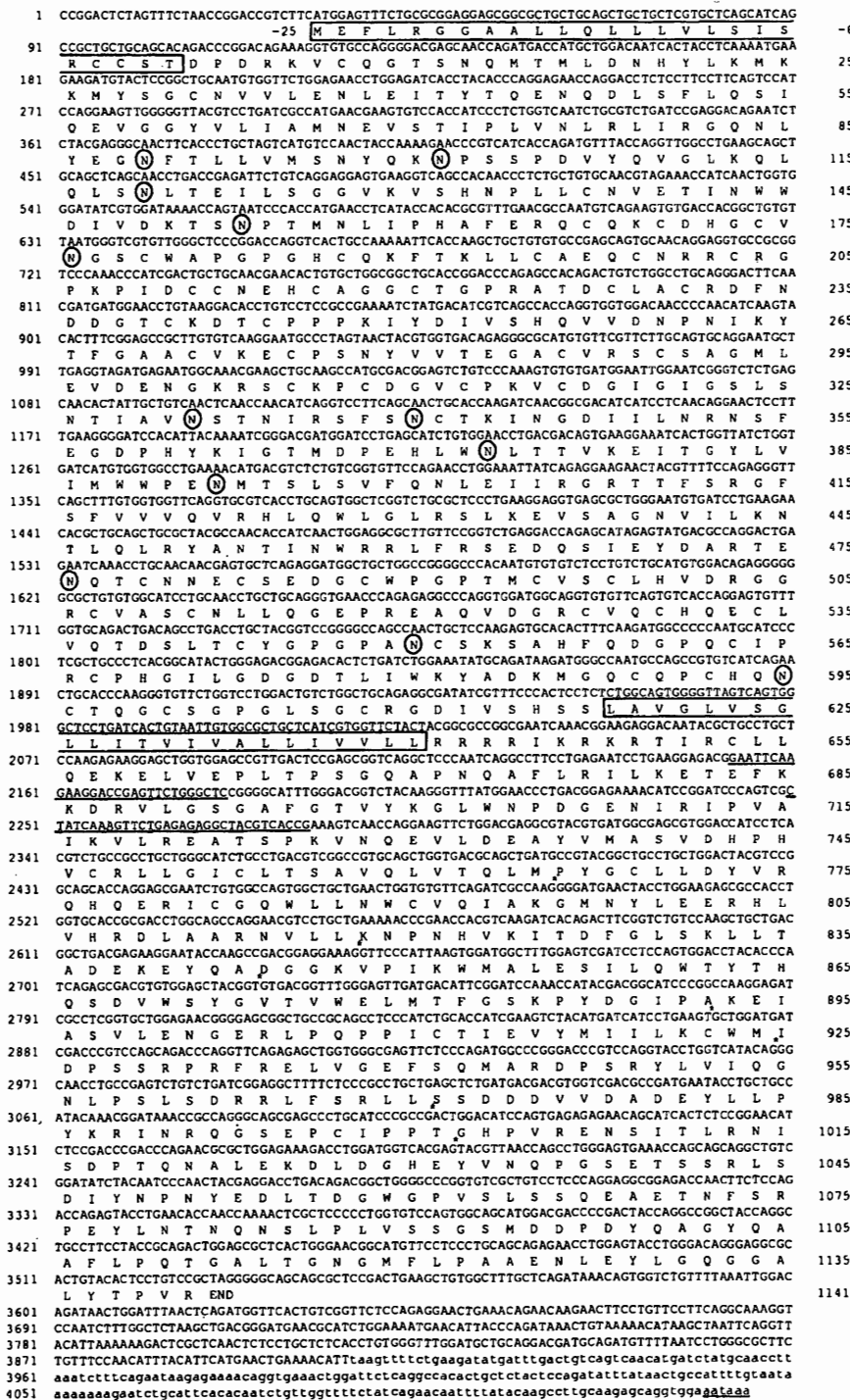
FIG. 3 Nucleotide sequence of the Tu gene and the predicted protein sequence. (Single letter amino acid code.) The nucleotide sequences (numbered left) of the overlapping cDNA clones 3-2 and 17-2 are in upper case. The sequence of the 3' end of exon z, which is not represented in the cDNA is derived from genomic DNA and shown in lower case. The predicted amino-acid sequence (numbered right) is shown below the nucleotide sequence. Potential glycosylation sites (N-X-(S/T)-X) are circled. The hydrophobic leader peptide (residues −25 to −1) and transmembrane domain (residues 618-640) are boxed. Oligonucleotides complementary to the sequences underlined were used for generation of a specifically-primed cDNA library (nt 2,281-2,250) and screening (nt 2,181-2,153), respectively. The putative polyadenylation signal at the 3' end is underlined. Exon boundaries derived from the genomic sequences are indicated by an asterisk (see also Fig. 2).

METHODS. An oligo(dT)-primed cDNA library was prepared by standard methods from the melanoma cell line (PSM) poly(A)+ RNA. The library was screened with random priming-labelled v-erb B probe (600-bp BamHI fragment D of pAE11; ref. 42) and pXX21 (ref. 22) under the conditions described in Fig. 2. The specifically-primed library was prepared from total RNA of PSM cells using an oligonucleotide complementary to nt 2,281-2,250 (see above). After denaturation of the RNA and pre-annealing of the oligonucleotide a cDNA-kit (Pharmacia) was used for preparation of double-stranded cDNA. Following ligation to λ, t10 arms, the DNA was packaged with Gigapack plus (Stratagene). The library was screened with a ^{32}P-labelled oligonucleotide complementary to nt 2,181-2,153 (see above). Generation of nested deletions and sequencing was performed as in Fig. 2.

c-kit, were significantly less similar. Like the EGF receptor, the *Tu*-encoded polypeptide is preceded by a signal peptide (25 amino acids compared with 24 amino acids in EGF receptor). In both proteins the extracellular domain is composed of two cysteine-rich domains. Most cysteine residues (all 24 in domain I, 19 of 21 in domain II) in the domains of the EGF receptor are conserv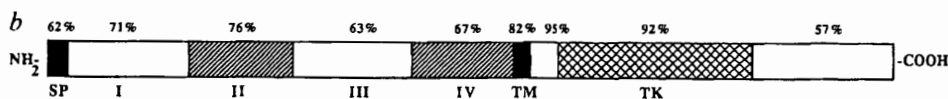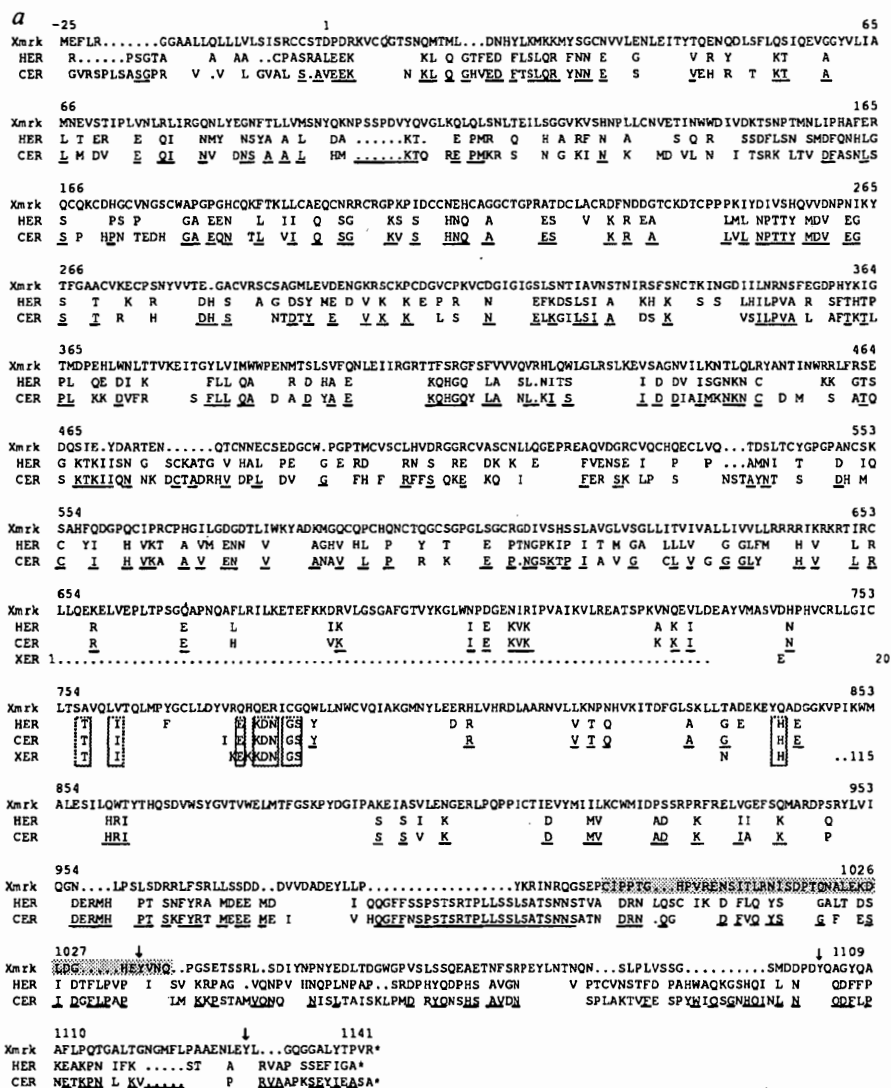ed between the two proteins (Fig. 4a). As with the EGF receptor, the *Tu*-encoded protein contains a single transmembrane domain, followed by a cytoplasmic juxtamembrane domain of 41 amino acids. The amino acids in this region are very highly conserved when compared with those of the corresponding region of the EGF receptor (95% similarity), including Thr 654 of the EGF receptor, which is a target for protein kinase C phosphorylation[28]. The remaining C-terminal portion of the predicted protein shows all the structural features of a functional tyrosine-kinase domain. All diagnostic structural motifs are conserved[29]. The C-terminal tail is less well-conserved on comparison with that of the EGF receptor and is heterogeneous in the predicted proteins encoded by the X- and Y chromosomes (see above). The Y-chromosomal C-terminus consists of 193 amino acids compared with 158 amino acids for the X-

chromosmal C-terminus and 235 for the EGF receptor. The Y-chromosomal *Tu*-encoded protein contains three tyrosine residues in the 3' domain which, by analogy for the EGF receptor, may serve as auto-phosphorylation sites (Fig. 4a), whereas the X-chromosomal *Tu*-encoded protein lacks the tyrosine residue corresponding to Tyr 1068 of the HER (human EGF-receptor) because of the in-frame deletion (see above).

We suggest, therefore, that the *Tu* oncogene encodes a transmembrane RTK with high similarity to the EGF receptor of higher vertebrates. We have called this gene *Xmrk* (*Xiphophorus* melanoma receptor tyrosine kinase).

Does *Xmrk* represent the fish EGF-receptor gene or is it a novel member of the subclass I RTK gene family, so far known to comprise the genes encoding the human, chicken and *Drosophila* EGF receptors (HER, CER and DER, respectively) and *HER2/neu*[30]? Sequences that hybridize most strongly with the v-*erb*-B probe have been isolated from a *Xiphophorus* genomic library and have been further characterized. In the central portion of the tyrosine-kinase domain, for which sequence information is available from the gene that encodes the *Xiphophorus* EGF-receptor (XER), there are motifs that the

FIG. 4 Comparison of the *Xmrk*-gene product with human EGF receptor (EGF-R)[27], chicken EGF-R[46] and the *Xiphophorus* EGF receptor homologue (partial), and structural organization of the predicted *Xmrk*-encoded protein. a, Amino-acid sequence comparison (single-letter code): Residues identical between all proteins are shown only in the *Xmrk* line. Residues identical in human EGF-R and chicken EGF-R but different from the *Xmrk*-gene product are underlined; residues identical in human EGF-R, chicken EGF-R and *Xiphophorus* EGF-R but different from the *Xmrk*-encoded protein are boxed. The residue predicted to be the first in the mature protein is numbered 1. The deletion in the predicted C-terminus of the X-chromosomal *Xmrk* gene product is shaded. Tyrosine residues homologous to Tyr 1,068, Tyr 1,148 and Tyr 1,173 of the HER are marked by arrows. (asterisk, translation stop site) b, Structural organization of the *Xmrk*-encoded protein. Different domains of the predicted protein are presented schematically: the hydrophobic signal peptide (SP) and transmembrane domain (TM) are in black; the two cysteine-rich modules of the extracellular domain (II, IV) are hatched; the tyrosine-kinase domain (TK, subdomains I–XI according to ref. 29) is cross hatched. Similarity values shown above the domains are based on the amino-acid sequence comparisons in Fig. 4a for *Xmrk*-gene product compared with human EGF-R.

```
a    -25                          1                                             65
Xmrk  MEFLR.......GGAALLQLLLVLSISRCCSTDPDRKVCÇGTSNQMTML...DNHYLKMKKMYSGCNVVLENLEITYTQENQDLSFLQSIQEVGGYVLIA
HER   R......PSGTA    A  AA ..CPASRALEEK    KL Q GTFED FLSLQR FNN E    G      VR Y    KT    A
CER   GVRSPLSASGPR   V .V  L GVAL S.AVEEK   N KL Q GHVED FTSLQR YNN E   S      VEH R  T  KT    A

      66                                                                          165
Xmrk  MNEVSTIPLVNLRLIRGQNLYEGNFTLLVMSNYQKNPSSPDVYQVGLKQLQLSNLTEILSGGVKVSHNPLLCNVETINWWDIVDKTSNPTMNLIPHAFER
HER   L T ER     E  QI    NMY  NSYA A L    DA ......KT.    E PMR  Q     H ARF N A     S Q R    SSDFLSN SMDFQNHLG
CER   L M DV     E  QI    NV  DNS A A L    HM ......KTQ   RE PMKR S    N G KI N  K    MD VL N    I TSRK LTV DFASNLS

      166                                                                         265
Xmrk  QCCQKCDHGCVNGSCWAPGPGHCQKFTKLLCAEQCNRRCRGPKPIDCCNEHCAGGCTGPRATDCLACRDFNDDGTCKDTCPPPKIYDIVSHQVVDNPNIKY
HER   S      PS P    GA EEN  L II  Q  SG    KS S    HNQ  A      ES  V KR EA      LML NPTTY MDV  EG
CER   S P   HPN TEDH  GA EQN TL VI Q  SG    KV S    HNQ  A      ES    K R A      LVL NPTTY MDV  EG

      266                                                                         364
Xmrk  TFGAACVKECPSNYVVTE.GACVRSCSAGMLEVDENGKRSCKPCDGVCPKVCDGIGIGSLSNTIAVNSTNIRSFSNCTKINGDIILNRNSFEGDPHYKIG
HER   S   T  K R       DH S   A G DSY ME D V K   K E P R    N      EFKDSLSI A   KH K    S S  LHILPVA R  SFTHTP
CER   S   I  T R H     DH S    NTDTY E   VK K   L S  N      ELKGILSI A   DS K        VSILPVA L  AFTKTL

      365                                                                         464
Xmrk  TMDPEHLWNLTTVKEITGYLVIMWWPENMTSLSVFQNLEIIRGRTTFSRGFSFVVVQVRHLQWLGLRSLKEVSAGNVILKNTLQLRYANTINWRRLFRSE
HER   PL  QE DI K      FLL QA    R D HA E      KQHGQ  LA  SL.NITS      I D DV ISGNKN C       KK  GTS
CER   PL  KK DVFR     S FLL QA   D A D YA E     KQHGQY LA  NL.KI S      I D DIAIMKNKN C   D M  S  ATQ

      465                                                                         553
Xmrk  DQSTE.YDARTEN......QTCNNECSEDGCW.PGPTMCVSCLHVDRGGRCVASCNLLQGEPREAQVDGRCVQCHQECLVQ...TDSLTCYGPGPANCSK
HER   G KTKIISN G SCKATG V HAL  PE    G E RD    RN S  RE DK K E    FVENSE I   P    P ...AMNI T   D  IQ
CER   S KTKIIQN NK DCTADRHV DPL  DV    G  FH F  RFFS QKE KQ I      FER SK LP S     NSTAYNT S     DH M

      554                                                                         653
Xmrk  SAHFQDGPGPQCIPRCPHGILGDGDTLIWKYADKMGQCQPCHQNCTQGCSGPGLSGCRGDIVSHSSLAVGLVSGLLITVIVALLIVVLLRRRRIKRKRTIRC
HER   C  YI   H VKT A VM ENN   V        AGHV HL P  Y T    E PTNGPKIP I T M GA  LLLV   G GLFM   H V   L R
CER   C   I   H VKA A V EN  V      ANAV L P  R K    E P.NGSKTP I A V G    CL V G G GLY   H V   L R

      654                                                                         753
Xmrk  LLQEKELVEPLTPSGÇAPNQAFLRILKETEFKKDRVLGSGAFGTVYKGLWNPDGENIRIPVAIKVLREATSPKVNQEVLDEAYVMASVDHPHVCRLLGIC
HER   R     E    L        IK      I E  KVK        AK I      N
CER   R     E    H        VK      I E  KVK        K K I      N
XER   1.................................................................................E       20

      754                                                                         853
Xmrk  LTSAVQLVTQLMPYGCLLDYVRQHQERICGQWLLNWCVQIAKGMNYLEERHLVHRDLAARNVLLKNPNHVKITDFGLSKLLTADEKEYQADGGKVPIKWM
HER   [T][T]   F     [K]KDN[GS] Y        D R         V T Q      A  G E  [H] E
CER   [T][I]        I [K]KDN GS Y        R           V T Q      A  G    [H] E
XER   [T][I]          KKKKDN GS              N                  A G      [H] H    ..115

      854                                                                         953
Xmrk  ALESILQWTYTHQSDVWSYGVTVWELMTFGSKPYDGIPAKEIASVLENGERLPQPPICTIEVYMIILKCWMIDPSSRPRFRELVGEFSQMARDPSRYLVI
HER   HRI               S SI K          D  MV     AD  K   II  K     Q
CER   HRI               S SV K          D  MV     AD  K   IA  K     P

      954                                                                         1026
Xmrk  QGN....LPSLSDRRLFSRLLSSDD..DVVDADEYLLP.................YKRINRQGSEPÇIPPTG....HPVRENSITLANISDPTQNALEKD
HER   DERMH  PT SNFYRA MDEE MD    I QQGFFSSPSTSRTPLLSSLSATSNNSTVA  DRN LQSC IK D FLQ YS   GALT DS
CER   DERMH  PT SKFYRT MEEE ME I   V HQGFFNSPSTSRTPLLSSLSATSNNSATN  DRN .QG    D FVQ YS   G F ES

      1027                                                                        1109
Xmrk  LDG....HEYVNQ..PGSETSSRL.SDIYNPNYEDLTDGWGPVSLSSQEAETNFSRPEYLNTNQN...SLPLVSSG.........SMDDPDYQAGYQA
HER   I DTFLPVP I   SV KRPAG .VQNPV INQPLNPAP..SRDPHYQDPHS AVGN    V PTCVNSTFD PAHWAQKGSHQI L N    QDFFP
CER   I DGFLPAP    LM KKPSTAMVQNQ  NISLTAISKLPMD RYQNSHS AVDN      SPLAKTVEE SPYHIQSGNHQINL N   QDFLP

      1110                                                                        1141
Xmrk  AFLPQTGALTGNGMFLPAAENLEYL...GQGGALYTPVR*
HER   KEAKPN IFK .....ST  A  RVAP SSEFIGA*
CER   NETKPN L KV.....  P  RVAAPKSEYIEASA*

b    62%   71%      76%       63%     67%  82% 95%      92%            57%
NH₂  ■□□□//////□□□□//////□□□□□□//////□□□■■⬛⬛▨▨▨▨▨▨▨▨▨▨▨□□□□□-COOH
     SP  I      II        III     IV  TM       TK
```

XER protein has in common with the HER- and the CER proteins that are not common to the *Xmrk*-encoded protein (Fig. 4a). Within this region, *Xmrk* has only 76% identity to the gene that encodes XER, and the two proteins that these genes encode have 83% identity. Sequence comparisons shows that the chicken and human EGF-receptor proteins have extended motifs in common throughout their extracellular domains and C-terminal tails, which are divergent from the corresponding sequences predicted for the *Xmrk*-encoded protein (Fig. 4a). The *Xmrk*-encoded protein is not therefore, the fish EGF-receptor homologue, but a novel putative receptor tyrosine kinase.

Elevated expression in melanoma

To investigate the physiological function of *Xmrk* we studied its expression pattern during embryogenesis, in normal adult tissue and in melanomas, induced by sex-chromosomal *Tu* loci, that differ in their degree of malignancy. Northern-blot hybridizations of RNA obtained from swordtail embryos containing only the INV *Xmrk* locus revealed a single *Xmrk* transcript of 5.8 kb (Fig. 5a). This transcript was highly abundant as a maternal messenger RNA in unfertilized eggs (Fig. 5a, lane 1). The level decreased during early development (cleavage to neurula stages), increased during organogenesis and from then on decreased. In 4-week-old neonate fish, only barely detectable hybridization signals were obtained (Fig. 5a, lane 10). In normal organs of adult fish the transcript was seen at low levels in gills, fins and skin but not in any other organ tested (data not shown).

In all melanomas that originated from six different *Tu* alleles from hybrid fish and in the melanoma cell line (PSM)[23], high levels of *Xmrk* transcripts were found (Fig. 5b). The RNA preparations included both early- and late-onset melanomas[23], as well as amelanotic tumours from fish homozygous for the albino (*a*) gene. Two large transcripts were seen, one that corresponds in size to the transcript from the INV locus (5.8 kb; see above) and a smaller transcript (4.7 kb) which is specific for the melanoma (Fig. 5b). Shorter transcripts (less than 3 kb) represented either specific degradation products, or transcripts of different length, but from the same gene, generated by differential splicing, alternative transcription start sites or alternative polyadenylation signals. The steady-state level of the 4.7 kb mRNA and the shorter transcripts was highly correlated to the the degree of malignancy of the melanoma. This was most obvious for tumours due to the *Tu–Sd* allele, for which backcross genotypes of defined malignancy were available. In extremely benign lesions, which occur in the presence of one intact allele of the melanoma suppressor gene *R*, only very low amounts of *Xmrk* mRNA were seen (Fig. 5c, lane 9) compared with the amounts seen in more malignant melanoma from fish that lacked both alleles of *R* (Fig. 5c, lane 10). In addition, fish that were heterozygous for *a*, which further enhances melanoma malignancy, had even higher transcript levels (Fig. 5b, lane 11). Expression of the INV locus was at a similarly low level in all tumours irrespective of their state of malignancy. The observation of a 5.8 kb INV *Xmrk* transcript in several normal tissues and the absence of the tumour-specific 4.7 kb transcripts indicates that the INV locus bears the *Xmrk* proto-oncogene.

Discussion

We have made use of the *Xiphophorus* melanoma system to study malignant tumours by both genetic and molecular techniques. We have cloned the oncogene *Tu* of *Xiphophorus*, known to induce malignant melanoma in the absence of the anti-oncogene *R*, and found that it encodes a novel transmembrane tyrosine kinase. We have called this gene *Xmrk*. In addition to the oncogenic *Xmrk* loci on the X- and Y-chromosomes of platyfish, we found that another copy of *Xmrk* at the INV locus is a normal constituent of the genome of all *Xiphophorus* fish, independent of the melanoma oncogene. This copy is probably the *Xmrk* proto-oncogene, which could have a normal physio-

logical function during development and in specific tissues.

The proposed structure of the protein encoded by *Xmrk* and its similarity to the subclass I growth-factor-receptor tyrosine kinases indicates that it is a transmembrane protein with an extracellular domain acting as a receptor, and a cytoplasmic tyrosine-kinase domain. RTKs are normally activated by the binding of a specific ligand which results in the phosphorylation of specific cellular proteins, thus producing a pleiotropic response, including proliferation and/or differentiation[28]. It is probable that *Xmrk* participates in such a signal-transduction mechanism. Although the predicted protein encoded by *Xmrk* shows structural similarity to the EGF-receptor of higher vertebrates, our data show that it does not represent the fish EGF-
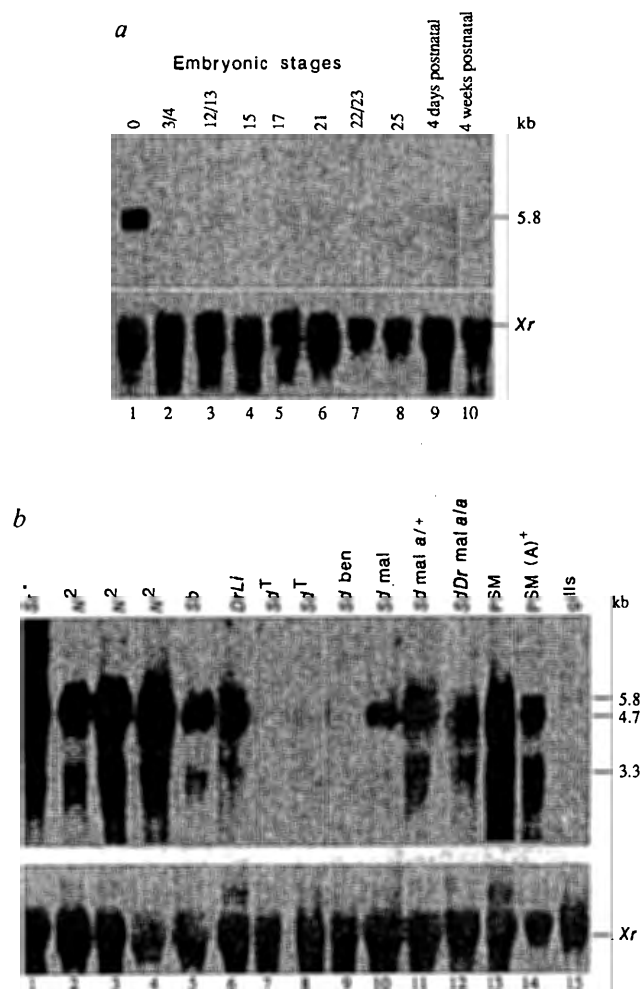

FIG. 5 Northern-blot analyses. *a*, Profile of *Xmrk* expression during embryonic development (for staging see ref. 47) (lanes 1–8), and at two different postnatal times (lanes 9 and 10) in fish containing only *Xmrk* at the INV locus. In each case, 20 μg of total RNA was analysed by Northern-blot hybridization. *b*, *Xmrk* expression in melanomas and in gills (lane 15). Total RNA (20 μg) from single tumours (lanes 1–4, 6, 11) or pooled melanoma biopsies (lanes 5, 7–10, 12) and a melanoma cell line (lanes 13 and 14) were analysed. PSM, melanoma cell line; *a*, albino gene; ben, benign; mal, malignant; all other abbreviations refer to melanomas due to different *Tu* alleles (lanes 1–5, Y-chromosomal; lanes 6, 9–12, X-chromosomal; lanes 7, 8, autosomal); *Sr"*, mutant *striped*; *N²*, nigra-2; *Sb*, spotted belly; *DrLi*, mutant lineatus; *Sd^T*, spotted dorsal translocation; *Sd*, spotted dorsal.
METHODS. RNA preparation, Northern blots and hybridizations for homologous probes were performed as described [38]. Probes used were random-priming-labelled 17-2 DNA (*b*) or a riboprobe of 17-2 (*a*). For size calibration a RNA ladder (BRL, Bethesda) was included. As a control for the amount of RNA blotted onto the membranes, the ubiquitously expressed *Xr* gene[48] was used as a hybridization probe.

receptor. The *Xmrk*-protein therefore, probably does not bind the fish EGF homologue. The *HER2/neu* gene product, which so far in mammals is the most closely related protein to the EGF-receptor, does not recognize EGF as its physiological ligand but binds to an as yet unidentified molecule[31-33]. Further experiments are needed to identify the possible ligand of the *Xmrk*-gene product RTK.

No satisfactory candidate for a RTK substrate has been identified[30]. The normal physiological function of the *Xmrk* encoded RTK is therefore beyond speculation, although the abundant *Xmrk* maternal message is intriguing with regard to recent reports that RTKs, like the *Drosophila* EGF-receptor and the *torso*-gene product, function as maternal-effect genes during early embryonic development[34-36].

The existence of the sex-chromosomal oncogenic *Xmrk* loci as additional genomic constituents in the platyfish poses the question of their evolutionary origin. It is conceivable that an allelic or duplicated copy of the INV-locus *Xmrk* gene was translocated to the sex-chromosomal pigmentary loci thereby acquiring an oncogenic potential. This is supported by the identification of macromelanophore pigment patterns in several poeciliid fish species and even in the genus *Xiphophorus* that never give rise to tumour formation after the hybridization (ref. 10, and our own unpublished data). In purebred platyfish the malignant phenotype is suppressed because of the presence of the *R* gene. After crossing-conditioned elimination of the *R* gene, melanoma develops. We have shown that the steady-state levels of sex-chromosomal *Xmrk* transcripts correlate with the genetically determined malignancy of the tumour, raising the possibility that transformation is affected by an increase in the level of *Xmrk* transcription. Such a situation is reminiscent of c-*myc* activation in Burkitt's lymphoma (reviewed in ref. 37) in which inappropriate regulation of *myc* is a consequence of a translocation.

Over-expression of RTKs of subclass I in tissue-culture leads potentially to oncogenesis, and increased levels are frequently, associated with certain human malignancies (for a review see

ref. 30). It should be noted, however, that in these tumours increased expression is mostly due to gene amplification, whereas in *Xiphophorus* melanomas no evidence for an amplification of *Xmrk* or related genes was obtained[21,38]. Recent studies indicate that in addition to over-expression of RTKs (in conjunction with autocrine ligand production), structural alterations may have a key role in the initiation as well as the progression of certain neoplasms (for a review see ref. 30). The *Xmrk* melanoma-specific transcripts are ~1 kb shorter than the transcript from the proto-oncogenic *Xmrk* at the INV locus. This points to such a structural alteration, possibly associated with the proposed translocation event. The full significance of the differing transcripts awaits further elucidation. The C-terminal deletion observed for the *Xmrk*-gene products encoded by the X-chromosomal locus seems not to affect its transforming capability in general, because in this region the potentially oncogenic *Xmrk* at the Y-chromosomal locus contains an identical exon as the proto-oncogenic *Xmrk* at the INV locus. This difference, however, may account for an observed variability in the malignancy of the resulting melanoma; the X-chromosomal *Tu-Sd* locus is responsible for much more severe tumourous lesions than the Y-chromosomal *Tu-Sr* locus. For the subclass I RTKs of higher vertebrates, a variety of minor or severe alterations of the C-terminal domains do not alter transforming potentials in general, but confer higher sensitivity to ligand-induced mitogenesis than that found in the wild-type situation[39,40] (for a discussion see ref. 30).

We propose that the melanoma oncogene of *Xiphophorus* encodes an activated version of a novel putative RTK, whose oncogenic potential is suppressed in wild-type platyfish. The suitability of fish for developmental biological studies and the availability of numerous mutants and wild-type alleles that affect the melanoma oncogene itself (or the phenotype of the resulting melanoma) makes *Xiphophorus* a unique *in vivo* system for studies of RTK-mediated oncogenesis as well as possible mechanisms for the suppression of this process. □

Received 29 June; accepted 5 September 1989.

1. Friend, S. H., Dryja, T. P. & Weinberg, R. *New Engl. J. Med.* **318**, 618–622 (1988).
2. Ponder, B. *Nature* **335**, 400–402 (1988).
3. Knudson, A. G. *Cancer Res.* **45**, 1437–1443 (1985).
4. Mechler, B. M., McGinnis, W. & Gehring, W. J. *EMBO J.* **4**, 1551–1557 (1985).
5. Jacob, L. *et al. Cell* **50**, 215–225 (1987).
6. Gordon, M. *Genetics* **12**, 253–283 (1927).
7. Häussler, G. *Klin. Wschr.* **7**, 1561–1562 (1928).
8. Kosswig, C. Z. *Indukt. Abstammungs-Vererbungsl.* **47**, 150–158 (1928).
9. Gordon, M. *Ann. N.Y. Acad. Sci.* **71**, 1213–1222 (1958).
10. Atz, J. W. *Zoologica* **47**, 153–181 (1962).
11. Kosswig, C. *Zool. Anz.* **175**, 21–50 (1965).
12. Zander, C. D. *Mitt. Hamburg. Zool. Inst.* **66**, 241–271 (1969).
13. Ahuja, M. R. & Anders, F. *Prog. Exp. Tumour Res.* **20**, 380–397 (1976).
14. Morizot, D. C. & Siciliano, M. J. *J. Natn. Cancer Inst.* **71**, 809–813 (1983).
15. Ahuja, M. R., Schwab, M. & Anders, F. *J. Hered.* **71**, 403–407 (1980).
16. Kallmann, K. D. in *Handbook of Genetics* (ed. King, I.) 81–132 (Plenum, New York, 1975).
17. Anders, F., Schartl, M., Barnekow, A. & Anders, A. *Adv. Cancer Res.* **42**, 191–275 (1984).
18. Kallmann, K. D. *Zoologica* **55**, 1–16 (1970).
19. Gordon, M. *Genetics* **22**, 376–392 (1937).
20. Kallmann, K. D. & Atz, J. W. *Zoologica* **51**, 107–135 (1966).
21. Schartl, M. *Genetics* **119**, 679–685 (1988).
22. Adam, D., Wittbrodt, J., Telling, A. & Schartl, M. *Nucleic Acids Res.* **16**, 7212 (1988).
23. Wakamatsu, Y. *Dev. Growth Differ.* **22**, 731–740 (1980).
24. Heijne, G. v. *J. Molec. Biol.* **184**, 99–105 (1985).
25. Heijne, G. v. *Nucleic Acids Res.* **14**, 4683–4690 (1986).
26. Klein, P., Kanehisa, M. & De Lisi, C. *Biochem. Biophys. Acta* **815**, 468–476 (1985).
27. Ullrich, A. *et al. Nature* **309**, 418–425 (1984).
28. Hunter, T. & Cooper, J. A. *A. Rev. Biochem.* **54**, 897–930 (1985).

29. Hanks, S. K., Quinn, A. M., & Hunter, T. *Science* **241**, 42–52 (1988).
30. Yarden, Y. & Ullrich, A. *A. Rev. Biochem.* **57**, 443–478 (1988).
31. Akiyama, T. *et al. Molec. Cell Biol.* **8**, 1019–1026 (1988).
32. King, C. R. *et al. EMBO J.* **7**, 1647–1651 (1988).
33. Stern, D. F. & Kamps, M. P. *EMBO J.* **7**, 995–1001 (1988).
34. Sprenger, F., Stevens, L. M. & Nüsslein-Vollhard, C. *Nature* **338**, 478–483 (1989).
35. Price, J. V., Clifford, R. J. & Schupbach, T. *Cell* **56**, 1085–1092 (1989).
36. Schejter, E. D. & Shilo, B.-Z. *Cell* **56**, 1093–1104 (1989).
37. Bornkamm, G. W., Polack, A. & Eick, D. in *Cellular Oncogene Activation* (ed. Klein, G.) 223–273 (Marcel Dekker, New York, 1988).
38. Mäueler, W., Raulf, F. & Schartl, M. *Oncogene* **2**, 421–430 (1988).
39. Honegger, A. *et al. EMBO J.* **7**, 3045–3052 (1988).
40. Khazaie, K. *et al. EMBO J.* **7**, 3061–3071 (1988).
41. Blin, N. & Stafford, W. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
42. Vennström, B., Fanshiw, L., Moscovici, C. & Bishop, J. M. *J. Virol.* **36**, 575–585 (1980).
43. Henikoff, S. *Gene* **28**, 351–359 (1984).
44. Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13**, 31–40 (1985).
45. Devereux, J. R. *et al. Nucleic Acids Res.* **12**, 387–395 (1984).
46. Lax, I. *et al. Molec. Cell Biol.* **8**, 1970–1978 (1988).
47. Tavolga, W. N. *Bull. Am. Mus. Nat. His.* **94**, 161–230 (1949).
48. Schwab, M. *Molec. Gen. Genet.* **188**, 410–417 (1982).