A *fork head* related multigene family is transcribed in *Xenopus laevis* embryos

JUTTA LEF, PETRA DEGE, MICHAELA SCHEUCHER, VERA FORSBACH-BIRK, JOACHIM H. CLEMENT and WALTER KNÖCHEL*

Abteilung Biochemie, Universität Ulm, Ulm, Germany

ABSTRACT We have isolated and sequenced ten different members of the fork head/HNF-3 multigene family from *Xenopus laevis* which have been termed *Xenopus* fork head domain related (XFD) genes 1 to 10. Another four isolated genes (XFD´ genes) represent pseudo-allelic variants which arose by an ancient tetraploidization within this species. Whereas all genes of this multigene family exhibit a high degree of sequence homology within the evolutionary conserved fork head domain, sequences outside this module are substantially different. Based upon sequence homologies over the entire coding sequences, XFD-7/7´ represent the *Xenopus* homologs to the rodent hepatocyte nuclear factor HNF-3α, while XFD-3/3´ encode the homologs to HNF-3ß. Here we present an analysis of the temporal transcription pattern of XFD genes 1 to 10 during embryogenesis and in some adult tissues. Eight of these XFD genes are activated during embryonic development, but show different and distinct transcription profiles. The localization of transcripts was determined by whole-mount in situ hybridization. Although transcription of individual XFD genes partially overlaps, each gene is characterized by means of a specific spatial pattern of transcriptional activity.

KEY WORDS: *Xenopus laevis, embryogenesis, fork head, whole-mount in situ hybridization*

Introduction

Control of gene expression during embryonic development in higher eucaryotic organisms depends on a multitude of different transcription factors which – in form of a complex network – regulate the temporal and spatial activation of their target genes. According to some evolutionary conserved sequence modules most of the hitherto known factors can be classified into a limited number of different multigene families. Examples for such modules are the homeobox (helix-turn-helix), the helix-loop-helix, the leucine zipper and the zinc finger motif. Another conserved module has been discovered by sequence comparison of the *Drosophila* gene *fork head* to rodent genes encoding hepatocyte nuclear factors 3 (HNF-3) (Weigel and Jäckle, 1990; Lai *et al.*, 1991). The corresponding proteins share a highly conserved DNA binding domain of about 110 amino acids, the fork head/HNF-3 domain. Meanwhile, this module has been detected in a variety of genes from different eucaryotic organisms ranging from yeast to primates and its tertiary structure has been determined (Clark *et al.*, 1993; for review see Lai *et al.*, 1993).

A previous search for fork head related genes in *Xenopus* revealed the existence of a multigene family (XFD genes: *Xenopus* fork head domain related genes), all members of which sharing this conserved module albeit at varying degrees of homology (Knöchel *et al.*, 1992). We have analysed so far the temporal and spatial transcription of XFD-1 and XFD-2 genes during embryogenesis (Knöchel *et al.*, 1992; Lef *et al.*, 1994). XFD-1 has independently been described as *pintallavis* (Ruiz i Altaba and Jessell, 1992) and the pseudo-allele XFD-1´ has independently been characterized as XFKH1 (Dirksen and Jamrich, 1992). Both genes are activated after midblastula transition (MBT; Newport and Kirschner, 1982) in the dorsal lip. Their transcripts are subsequently localized within notochord and neural floor plate. XFD-2/2´ is immediately activated at the onset of zygotic transcription in the animal hemisphere. At late blastula stage, transcripts are found in the marginal zone, i. e. within mesodermal cells which will invaginate into the blastoporus during gastrulation (Lef *et al.*, 1994). These results and recent reports on the embryonic expression of two *Xenopus* homologs of rodent HNF-3 factors (Bolce *et al.*, 1993; Ruiz i Altaba *et al.*, 1993) suggest, that members of the XFD family participate in the complex network of transcription factors which is required for pattern formation and tissue differentiation during early embryogenesis.

Here we report the sequences of some additional members of the XFD multigene family. Ten different genes have been analysed by RNase protection experiments for their temporal transcription pattern during embryogenesis. Except for two genes, where we failed to detect any transcripts, we show by

***Address for reprints:** Abteilung Biochemie, Universität Ulm, Albert Einstein Allee 11, D-89081 Ulm, Germany. FAX: 731.5023277.

0214-6282/96/$03.00
© UBC Press
Printed in Spain

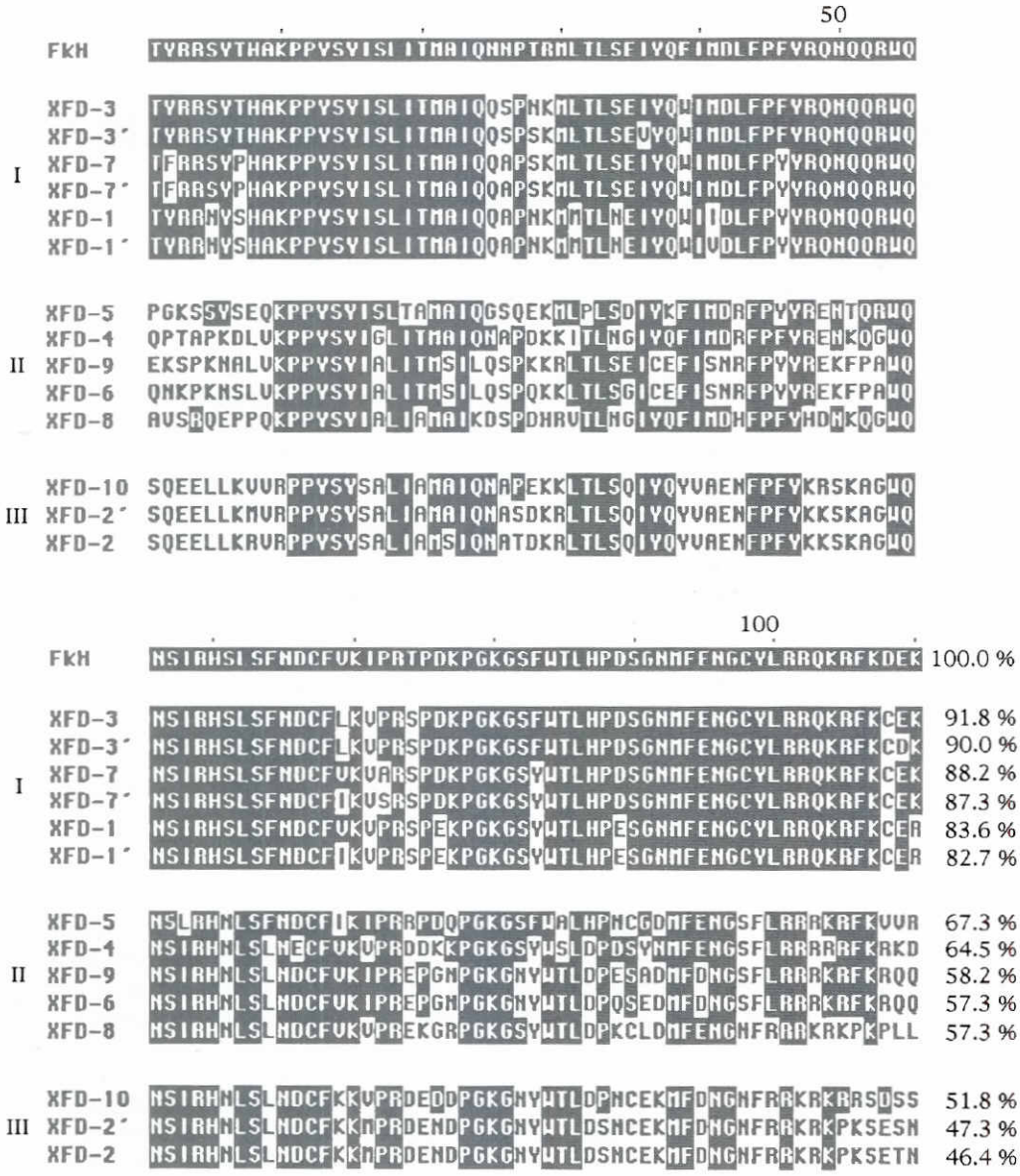


Fig. 1. Fork head domains (110 amino acids) of *Xenopus* XFD genes compared to the corresponding domain of the *Drosophila fkh* gene. *XFD genes are subdivided into three distinct subfamilies according to their degrees of homology. Rates of identity are shown for each gene with respect to the fkh gene. Amino acids which are identical to those of the fkh gene are highlighted. Note the sequence similarities (non-highlighted amino acids) within each type of subfamily.*

whole-mount in situ hybridization that transcription of each individual gene is restricted to distinct tissues. Even if we observe some partial tissue overlap, each gene displays a unique pattern of spatial transcriptional activity.

Results

A fork head related gene family in Xenopus laevis

Xenopus laevis gastrula stage cDNA and genomic libraries were hybridized under reduced stringency with a labeled probe encoding the fork head domain (*fkh*) of *Drosophila melanogaster* (Weigel *et al.*, 1989; Weigel and Jäckle, 1990). This approach led to the isolation of six different genes designated as *Xenopus* fork head domain related (XFD-1 to 6) genes (Knöchel *et al.*, 1992). Meanwhile, by performing further screenings, we succeeded in the isolation of additional four genes termed as XFD-7 to 10. Moreover, due to the known genome duplication in *Xenopus* (Bisbee *et al.*, 1977; Knöchel *et al.*, 1986), we have identified pseudo-allelic versions of XFD genes 1, 2, 3 and 7. These XFD´ sequences have also been isolated as cDNAs. By comparison to their individual counterparts it became obvious that a close sequence homology is not confined to their fork head domains but extends over their entire sequences. Since temporal and spatial expression patterns of these closely related genes are also indistinguishable, it is reasonable to assume that these sequences represent pseudo-alleles rather than different isoforms. Figure 1 shows a compilation of nucleotide derived amino acid sequences of 14 XFD fork head domains representing 10 different types of genes in comparison to that of the founder sequence *fkh*. According to their varying degrees of identity we have subdivided the XFD genes into three different subfamilies. Except for the second group which is rather heterogeneous all members of a given subfamily show striking sequence similarities. Whereas the DNA binding domains

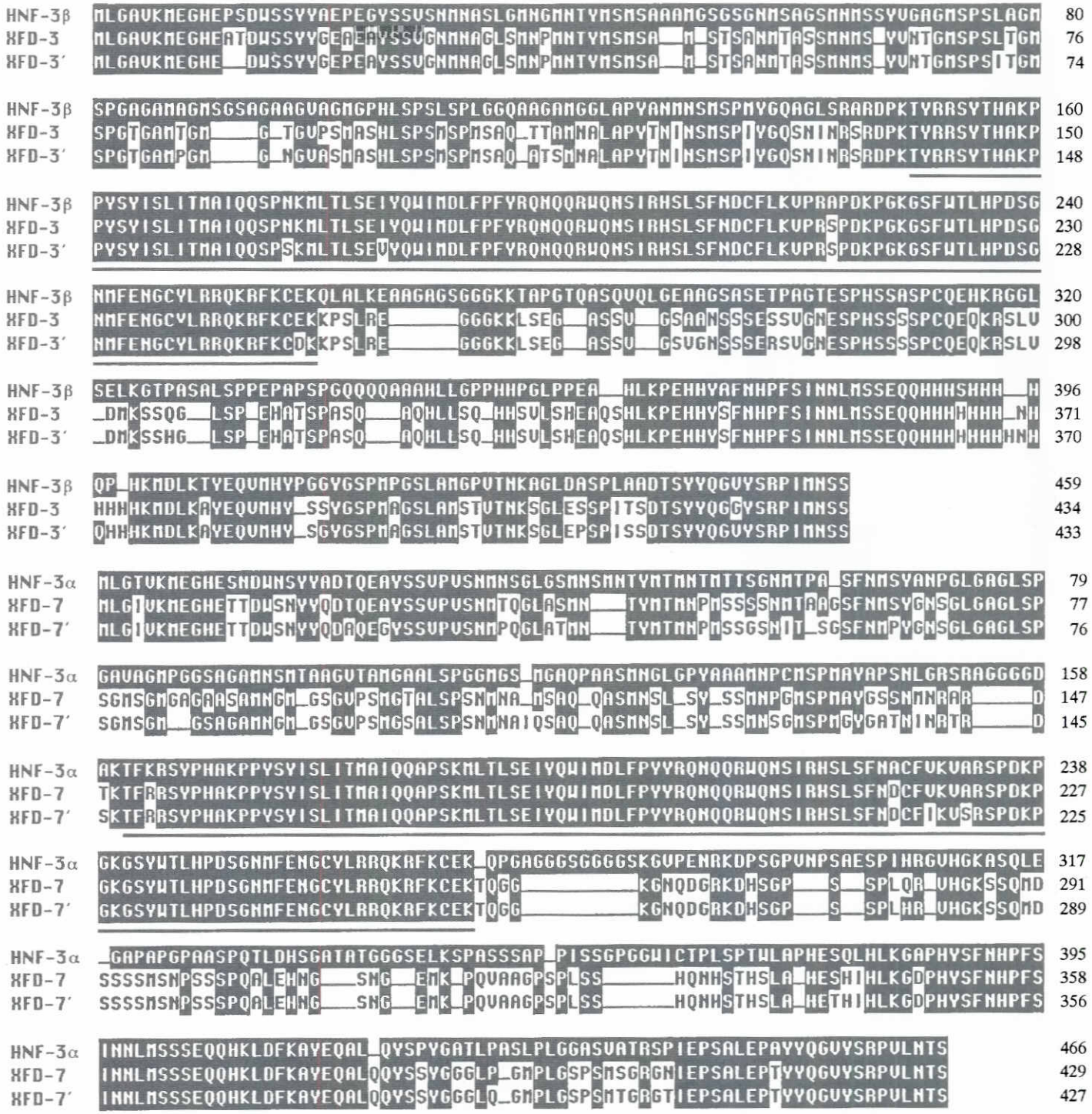


Fig. 2. Amino acid sequences of the pseudo-allelic variants XFD-3/3′ and XFD-7/7′, the *Xenopus* homologs of rodent HNF-3β and HNF-3α, respectively. *Amino acids are aligned to the rat HNF-3β and HNF-3α sequences at highest homology. Invariable amino acids are highlighted. XFD-7′ fully corresponds to XFKH2, XFD-3 deviates from the recently reported Xβ-1 sequence at three positions (50: S/R; 102: S/I and 311: E/K). Note the significant sequence homologies not only within the fork head domain (underlined) but also at the N- and C-terminal regions of the corresponding proteins.*

encoded by the first subfamily (XFD-3, 7 and 1) share more than 80% identity with the fkh protein, members of the third group (XFD-2 and XFD-10) show a rather weak homology (about 45%), thereby indicating only a distant relationship. The second group (XFD-4, 5, 6, 8 and 9) comprises rather different types of sequences; however, XFD-6 and XFD-9 fork head domains share 92% identity. Although they diverge from each other outside their fork head domains, they probably belong to a distinct

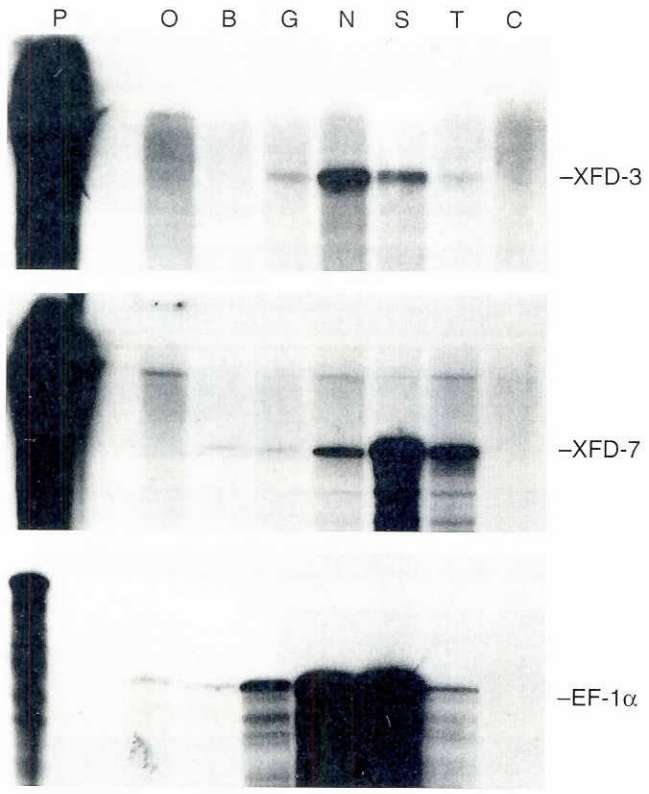


Fig. 3. RNase protection analysis of XFD-3 and XFD-7 transcripts during *Xenopus* embryogenesis. *Each 50 μg RNA from oocytes and embryos at different developmental stages (B, blastula; G, gastrula; N, neurula; S, somite segregation; T, hatched tadpole) were hybridized with corresponding antisense RNAs (see Materials and Methods). After RNase digestion, protected fragments were run on a polyacrylamide gel and visualized by autoradiography. (C) Control experiment with 50 mg tRNA; P, radiolabeled probes used. Integrity of RNA preparations were checked by RNase protection by using an antisense probe of EF-1a (Pöting et al., 1990).*

subfamily which is characterized by the recognition of similar target sequences.

Four of the presented sequences have independently been isolated in other laboratories: XFD-1′ corresponds to XFKH1 (Dirksen and Jamrich, 1992), XFD-1 to *pintallavis* (Ruiz i Altaba and Jessell, 1992), XFD-7′ to XFKH2 (Bolce *et al.*, 1993) and XFD-3 to frog HNF-3β (clone Xβ-1) (Ruiz i Altaba *et al.*, 1993). Furthermore, it is interesting to know whether any of the XFD sequences might correspond to those which have been described for other species. Table I shows the closest homologs that have been found by computer aided comparison of various fork head domains reported for other species. However, despite the striking homologies found for individual genes, it is not justified to draw any conclusion on species homologs; such a statement should also require sequence conservation outside the fork head domain. So far we have only observed such a conservation for XFD-4 which corresponds to the MFH-1 sequence of mouse (Miura *et al.*, 1992) and for XFD-3 and 7 which are the *Xenopus* homologs of rodent HNF-3β and α (Lai *et al.*, 1990; Lai *et al.*, 1991), respectively (see below).

Our data support the present view that the fork head domain defines an evolutionary conserved, DNA binding motif found to be present in many transcription factors from all eucaryotic organisms. Fork head related genes in *Xenopus* constitute a multigene family which, based upon sequence variation, may be subdivided into distinct subfamilies. The number of genes actually belonging to this *Xenopus* multigene family remains to be elucidated. Data presented here and from other laboratories suggest that at least 20 to 30 genes account for this motif.

XFD-3 and XFD-7 are Xenopus homologs to rodent HNF-3β and HNF-3α

We have isolated and sequenced four different cDNAs from a gastrula stage cDNA library which encode two pseudo-allelic pairs of *Xenopus* fork head related transcription factors (termed XFD-3/3′ and XFD-7/7′). Based upon sequence homology and transcription behaviour during embryogenesis and in adult tissues, XFD-7/7′ most likely represent the *Xenopus* homologs to the previously identified rodent hepatocyte nuclear factor HNF-3α (Lai *et al.*, 1990), while XFD-3/3′ encode the homologs to HNF-3β (Lai *et al.*, 1991). Interspecies comparison reveals that sequence homologies are not only apparent within the highly conserved fork head domains but also in the N- and C-terminal parts of corresponding proteins located outside this domain (see Fig. 2). It has to be noted that XFD-7′ exactly corresponds to XFKH2 (Bolce *et al.*, 1993) and the fork head domain of XFD-3 has recently been reported to be identical to that of the Xβ-1 sequence (Ruiz i Altaba *et al.*, 1993). Although for the latter we noticed three amino acid exchanges over the entire region, we suppose that both types of sequences are derived from allelic variants of the same gene. Here we show the complete amino acid sequences for the pseudo-alleles of both types of proteins. XFD-3/3′ exhibit 95% identity, XFD-7/7′ exhibit 94% identity. These values do well agree with our assumption that we have isolated the pseudo-allelic variants. Within the fork head domains we find three amino acid exchanges for XFD-3/3′ and two for XFD-7/7′. The overall homology between the rat and *Xenopus* proteins is about 70%. Interestingly, within the fork head domain there is only one exchange between rat HNF-3β and XFD-3 and only two exchanges between rat HNF-3α and XFD-7 which indicates a higher selective constraint on the DNA binding module than on the regions outside this motif. Although being derived from very different species, we predict that the resulting proteins will bind to identical target sequences.

Temporal transcription of XFD genes during embryonic development

We have analyzed next the temporal transcription patterns of XFD genes during embryonic development. Figure 3 shows RNase protection experiments performed with antisense probes derived from XFD-3 and XFD-7 sequences. To check for the integrity of our RNA preparations we routinely performed controls with an antisense probe of the translation factor EF-1α (Krieg *et al.*, 1989; Pöting *et al.*, 1990). This gene is activated at midblastula transition and transcripts are ubiquitously distributed in all adult tissues. The results demonstrate that transcription of XFD-7 precedes that of XFD-3. While XFD-7 transcripts are

already detected at blastula stage, XFD-3 is first detectable within gastrula stage RNA. However, XFD-3 transcripts accumulate more rapidly, so that they are prevailing at neurula stage but become less abundant at later stages. XFD-7 transcripts accumulate until hatching but there is a decrease at tadpole stages. Both genes are transcribed in the liver of adult frogs. We also detected XFD-3 and XFD-7 transcripts in lung (see Fig. 4) which corresponds to previous findings on rat HNF-3α and β, respectively (Lai *et al.*, 1991; Xanthopoulos *et al.*, 1991), as well as XFD-3 transcripts in brain.

The results obtained from RNase protection experiments performed with 10 different XFD genes are summarized in figure 4. It is interesting to note, that none of these genes is maternally expressed but most of them are zygotically transcribed during embryogenesis. Except for XFD-5 and XFD-8, where we failed to detect any transcripts during embryogenesis, the remaining eight genes are activated at specific developmental stages. Each XFD gene is characterized by a distinct time interval at which transcriptional activity is observed (since the pseudo-allelic variants behave similar or identical, they are not included in Fig. 4). The very early transcription of XFD-2 and the activation of XFD-1/1′ genes after midblastula transition have already been reported elsewhere (Knöchel *et al.*, 1992; Lef *et al.*, 1994). The temporal transcription patterns of XFD-6 and XFD-9 genes significantly differ, although both genes share very similar fork head domains. The same holds true for XFD-2 and XFD-10. In this case we also observe very different patterns of transcription despite a 90% homology between their fork head domains.

Our data reveal that various members of the fork head multigene family are transcribed during *Xenopus* embryogenesis. Transcription of each gene is characterized by a distinct temporal pattern and even in case of highly related fork head domains these patterns are significantly different. They differ by the time point of activation and by the stage at which the highest transcript level is accumulated. Some genes are most abundantly transcribed already at blastula and gastrula stages (XFD-1, XFD-2, XFD-4 and XFD-6), others at neurula stage (XFD-3) and some are activated later reaching highest levels of transcription during tailbud stages. The differential patterns observed give rise to the question whether transcription of these genes is ubiquitous or whether it is restricted to defined embryonic tissues.

Localized transcription of XFD genes

We have analyzed next the spatial distribution of XFD transcripts in embryos at different developmental stages by use of the whole-mount in situ hybridization technique (see Fig. 5). The most remarkable features determined for each of the XFD genes are outlined as follows.

XFD-1

This gene is activated after MBT in the dorsal lip and transcripts accumulate during gastrulation within the notochord. Neurula stage embryos do also show transcription within the neural floor plate. In accordance to its transcription in dorsal axial mesoderm this gene has been reported to be activated by activin A in isolated animal caps but not by basic fibroblast growth factor (Dirksen and Jamrich, 1992; Knöchel *et al.*, 1992; Ruiz i Altaba and Jessell, 1992).

	embryogenesis						adult tissues							
	oocyte	blastula	gastrula	neurula	tailbud	tadpole	liver	heart	intestine	spleen	kidney	ovary	brain	lung
XFD-1	-	+	+++	+	-	-	-	ND	ND	ND	ND	ND	ND	ND
XFD-2	-	+++	+++	+	+	-	-	ND	ND	ND	ND	ND	ND	ND
XFD-3	-	-	+	+++	++	+	+++	-	ND	-	ND	ND	++	+
XFD-4	-	++	+++	++	+	+	-	-	-	+	++	+	ND	ND
XFD-5	-	-	-	-	-	-	-	-	-	-	-	-	ND	ND
XFD-6	-	+	+++	++	+	++	±	±	±	±	±	±	ND	ND
XFD-7	-	+	+	++	+++	++	+++	-	ND	-	ND	ND	ND	++
XFD-8	-	-	-	-	-	-	ND	ND	ND	ND	ND	ND	ND	ND
XFD-9	-	-	+	+	+++	++	-	-	+	-	++	-	ND	ND
XFD-10	-	-	+	+	++	+	ND	ND	ND	ND	ND	ND	ND	ND

Fig. 4. Transcription patterns of XFD genes during embryogenesis and in adult tissues. *Transcription of XFD genes 1 to 10 was investigated by RNase protection analysis using RNAs isolated from different embryonic stages and adult tissues as indicated. Observed transcript levels are as following: (-) not detected; (±) barely detectable; (+) low level; (++) high level; (+++) very high level. ND: not determined.*

XFD-2

There is strikingly high transcriptional activity already at blastula stage. The burst of transcriptional activity is documented by intense colour within nuclei of the animal hemisphere. In late blastula embryos transcripts accumulate along the marginal zone. These cells enter at gastrula stage the blastoporus not only from dorsal but also from ventral and lateral sides.

XFD-3

At stage 23, the *Xenopus* homolog of HNF-3β is preferentially transcribed within the neural floor plate. There is no transcription in the notochord but transcripts are visible in the midbrain, the hindbrain and in cranial neural crest cells. The hatching larvae show transcription in the foregut which is indistinguishable from that presented for XFD-7.

XFD-4

Gastrula and neurula stage embryos show transcription in pre-somitogenic mesoderm but not in the notochord. While at later developmental stages transcripts in somites become less abundant, they continue to be present in the tip of the tail. We also observe transcripts in pronephros and pronephric duct. At hatching we detect transcripts within heart and foregut but also in the hindbrain and pharyngeal pouches.

XFD-6

Transcription of this gene is restricted to neural crest cells. There are two major sites of transcription at the lateral border of the anterior neural plate in neurula stage embryos. One represents a population of more superficially located cells which differentiate to neural crest cells originating from the rhombencephalon. These cells segregate as stripes during neurulation. The second locus is more anterior and within a deeper layer of cells representing neural crest cells which probably originate from the mesencephalon.

XFD-7

The *Xenopus* homolog of HNF-3α is transcribed at neurula stage in the notochord but not in the neural floor plate. During

tailbud stages it becomes also activated in the neural floor plate. At stage 35, transcripts are present in rhombencephalon, mesencephalon, pharyngeal pouches, foregut and pronephros. At stage 44 we detect transcripts in a restricted area of the gut located at the right side of the embryo.

XFD-9

This gene is first transcribed at gastrula stage within the neuroectoderm. While transcripts of the anterior part are later found in brain structures, cells along the trunk will differentiate into neural crest cells. At somite segregation stages we observe intense staining in the region of rotating somites. While somite rotation proceeds, staining is shifted towards the posterior end of the embryo until, at tailbud stages, it is located near the posterior pole. Transcripts are also visible in cells surrounding the pronephric duct and in neural crest cells migrating to the dorsal fin but not in the notochord.

XFD-10

Neurula stage embryos show transcripts in neuroectoderm but also in pre-somitogenic mesoderm. Strong signals are observed in the anterior but also in the posterior region of neuroectoderm and in the dorsal and ventral circumblastoporal collar. At somite segregation stages we observe transcripts in all parts of the differentiating brain and, especially, in neural crest cells which probably give rise to the hyoid and anterior branchial arches.

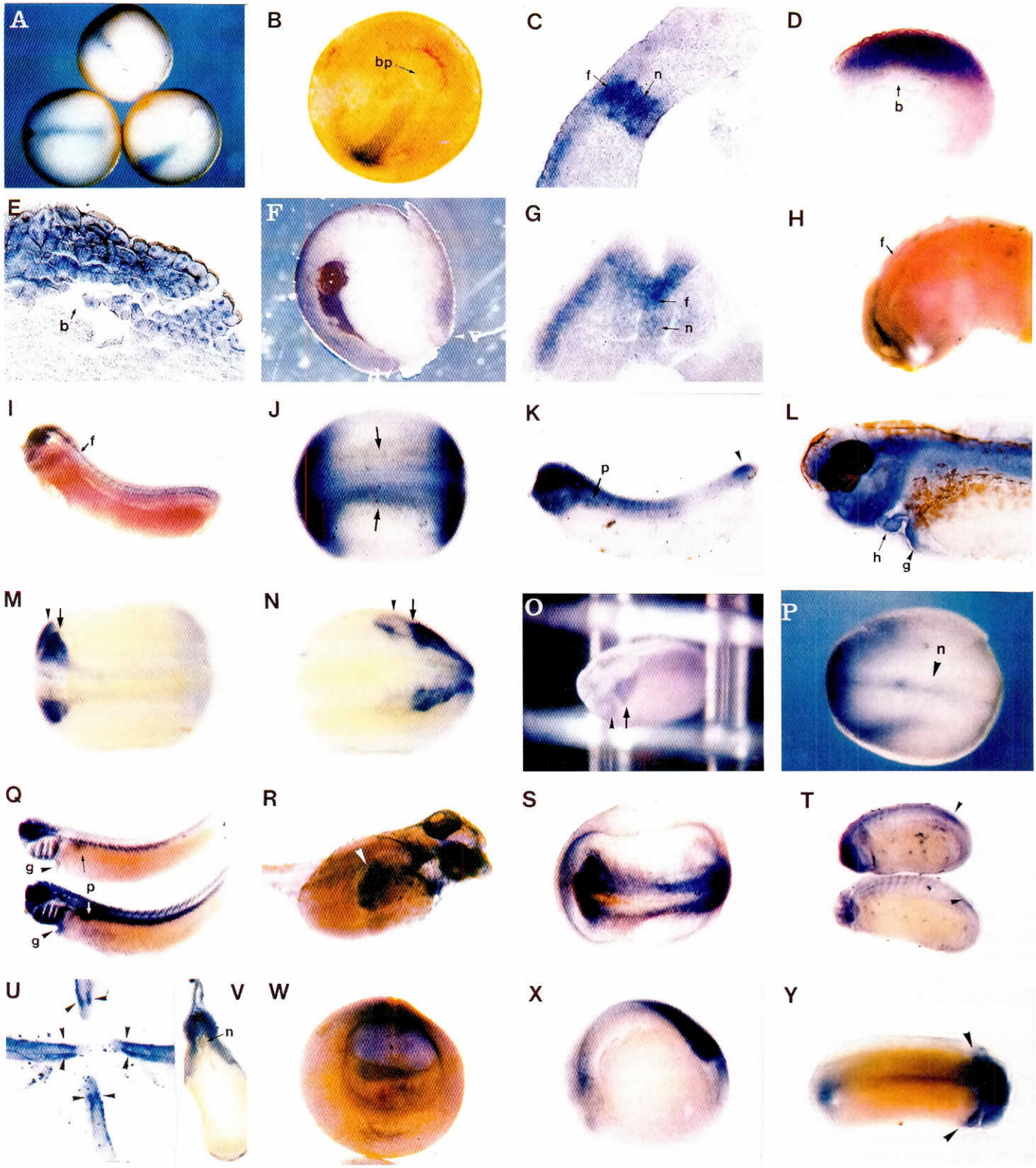
In summary, we have shown that each individual XFD gene is characterized by a specific pattern of tissues which display transcriptional activity. Although in case of some distinct tissues we observe simultaneous transcription of two or more XFD genes, none of these genes behaves identical to another one regarding its overall activity, i. e. the complete set of tissues where it is transcribed.

Discussion

Here we describe a *Xenopus* multigene family which is related to the *Drosophila* gene fork head by sharing a conserved 110 amino acid DNA-binding domain, the fork head domain. Ten different members of this XFD multigene family have been characterized by means of their coding sequence and by their temporal and spatial expression during embryogenesis. Obviously, the majority of these genes is transcribed during development which suggests that corresponding gene products contribute to the multitude of transcription factors required for normal gene expression during embryogenesis. Further, we have isolated some additional cDNAs displaying more than 90% homology over their entire sequence length to those already isolated. Thus, it seems reasonable to assume that they represent pseudo-allelic variants due to the known genome duplication event in this species some 50 million years ago (Bisbee *et al.*, 1977; Knöchel *et al.*, 1986). According to the varying extent of homology within the fork head domain the XFD genes have been subdivided into three different subfamilies. Such subdivision corresponds to findings from other organisms and to sequence alignments reported previously (Clevidence *et al.*, 1993; Sasaki and Hogan, 1993; Murphy *et al.*, 1994). However, except for XFD-3 and XFD-7 which clearly are the *Xenopus* homologs of rodent HNF-3α and 3β (Lai *et al.*, 1991) and probably for XFD-4 which might represent the counterpart to mouse MFH-1 (Miura *et al.*, 1993) it is rather uncertain which of the remaining *Xenopus* genes represents a homolog to one of those genes having been described in other species.

XFD-1 to 10 have been analysed for their temporal transcription pattern during embryogenesis. Except for XFD-5 and XFD-8 we find distinct temporal patterns for transcription of individual XFD genes during development. The earliest gene to be activated is XFD-2; at late blastula we also observe transcripts of XFD-1, XFD-4, XFD-6 and XFD-7. At gastrula stage we observe transcription of XFD-3, XFD-9 and XFD-10. Thus, there is a sequential activation of XFD genes and, for most cases, we show that transcripts accumulate at a specific developmental stage followed by a considerable decrease. Localized expression has already been reported for XFD-1, XFD-2, XFD-3 and XFD-7 (Knöchel *et al.*, 1992; Dirksen and Jamrich, 1992; Ruiz i Altaba and Jessell, 1992; Bolce *et al.*, 1993; Ruiz i Altaba *et al.*, 1993). Here we confirm these data and extend them by observation of additional features. For example we observe substantial amounts of XFD-2 transcripts within nuclei of animal caps and we demonstrate localization of XFD-7 transcripts in pronephros and later, in swimming tadpoles, at a defined part of the gut. XFD-3, the *Xenopus* homolog to rodent HNF-3β is transcribed in the neural floor plate but not in the notochord. HNF-3β expression in the floor plate has also been reported for other organisms (Ang *et al.*, 1993; Monaghan *et al.*, 1993; Sasaki and Hogan, 1993) and the corresponding protein has been postulated as a regulator of floor plate development (Sasaki and Hogan, 1994). However, in contrast to *Xenopus*, HNF-3β is also transcribed in the notochord. Despite considerable efforts, a mammalian homolog to the *Xenopus* XFD-1 gene which is transcribed in notochord has hitherto not been found. This failure led to the hypothesis that HNF-3β in higher organisms may compensate for the combined action of XFD-1 and XFD-3 in frogs (Ruiz i Altaba *et al.*, 1993). The distribution of XFD-6, XFD-9 and XFD-10 transcripts is mainly restricted to neuroectoderm and neural crest cells but transcription of each gene displays some unique features. Already at neurula stage, XFD-6 is only transcribed in two cell populations which later give rise to neural crest cells derived by the hindbrain and by the midbrain. XFD-9 is mainly observed in neural crest cells at the middle part of the embryo during somite segregation stages, especially at the bor-

Fig. 5. Localization of XFD gene transcripts in *Xenopus* embryos by whole-mount *in situ* hybridization. *XFD-1: embryos at gastrula (stage 11)* **(A,B)** *and transverse section through an embryo at early neurula (stage 13)* **(C)**. *XFD-2: blastula (stage 8) embryo* **(D)**, *section through the animal cap* **(E)** *and late gastrula (stage 12) embryo* **(F)**. *XFD-3: transverse section through an embryo (stage 23) at the anterior hindbrain region* **(G)**, *stage 23* **(H)** *and tailbud (stage 28)* **(I)** *embryos. XFD-4: embryos at stage 16* **(J)**, *stage 32* **(K)** *and stage 37* **(L)**. *Arrows* **(J)** *denote pre-somitogenic mesoderm, arrowhead* **(K)** *points towards intense staining at the tip of the tail. XFD-6: embryos at stage 17* **(M)**, *stage 19* **(N)** *and stage 23* **(O)**. *Arrows and arrowheads denote two types of cell populations differentiating to cranial neural crest cells of the hindbrain (arrows) and midbrain (arrowhead). XFD-7: neurula (stage 15)* **(P)**, *stage 35* **(Q)** *and stage 44* **(R)** *embryos. Arrowhead (R) points at localized transcription in a defined area of the gut at the right side of the embryo. XFD-9: Embryos at stage 16* **(S)**, *stage 25* **(T)**, *stage 35* **(U)** *and transverse section through a stage 36 embryo* **(V)** *at the middle part.*



Arrowheads in **(T)** *denote the border between segmented and unsegmented somitic mesoderm, arrowheads in* **(U)** *point at intense staining near the posterior pole. XFD-10: Neurula (stage 15) embryo* **(W)**, *lateral view of a neurula (stage 16) embryo* **(X)** *and dorsal view at an embryo at stage 24/25* **(Y)**. *Arrowheads in* **(Y)** *point at a cranial neural crest cells. b, blastocoel; bp, blastoporus; d and v, dorsal and ventral lip of blastoporus; f, neural floor plate; g, foregut; h, heart; n, notochord; p, pronephros.*

TABLE 1

INTERSPECIES COMPARISON OF XFD GENES WITH FORK HEAD RELATED GENES FROM OTHER SPECIES

X. *laevis* gene	highest homology (domain)	name of gene	from (organism)	reference
XFD-1	89%	HNF-3α	rat	Lai *et al.*, 1990
XFD-2	86%	fkh-10	mouse	Schütz, unpublished
XFD-3*	99%	HNF-3β*	rat	Lai *et al.*, 1991
XFD-4*	96%	MFH-1*	mouse	Miura *et al.*, 1993
XFD-5	94%	fkh-4	mouse	Kaestner *et al.*, 1993
XFD-6	99%	HFH-2	rat	Clevidence *et al.*, 1993
XFD-7*	98%	HNF-3α*	rat	Lai *et al.*, 1990
XFD-8	86%	fkh-6	mouse	Kaestner *et al.*, 1993
XFD-9	98%	HFH-B2	rat	Clevidence *et al.*, 1993
XFD-10	91%	fkh-9	mouse	Schütz, unpublished

Computer aided amino acid comparison of XFD genes was performed with the following sequences: *fork head* (Weigel *et al.*, 1989), HNF-3α/β/γ (Lai *et al.*, 1990; 1991), FD1 to 5, slp1, slp2 (Häcker *et al.*, 1992), BF-1 (Tao and Lai, 1992), HTLF (Li *et al.*, 1992b), ILF (Li *et al.*, 1992a), lin-31 (Miller *et al.*, 1993), fkh1 to 6 (Kaestner *et al.*, 1993), HFH-1 to 7, HFH-B2, HFH-B3 (Clevidence *et al.*, 1993), H3, H8, 5-3 (Hromas *et al.*, 1993), qin (Li and Vogt, 1993), HCM1 (Zhu *et al.*, 1993); HFH-E5.1 (Ang *et al.*, 1993), MFH-1 (Miura *et al.*, 1993), PES-1 (Hope, 1994), HFKH1, 2, 3 (Murphy *et al.*, 1994). Sequences yielding the highest rates of identity within their fork head domain to individual XFD genes are cited. For those genes which are indicated by an asterisk a striking sequence homology exists also for the regions which are located outside the fork head domain.

der of rotating somites, until it is finally present within a region near the posterior pole. XFD-10 is visualized at tailbud stages within distinct neural crest cells probably giving rise to formation of gill arches. Initial transcription of XFD-4 is similar to that reported for its putative mouse homolog MFH-1 (Miura *et al.*, 1993). At neurula stage, XFD-4 transcripts observed in non-notochordal structures of the mesoderm, mainly in pre-somitogenic mesoderm. At later development we observe transcripts in parts of the brain, in pronephros and, finally, in the tip of the tail. Since this gene is also transcribed in the heart and in the foregut, it is activated in derivatives of all three embryonic germ layers. The notion, that transcription of this gene during embryogenesis is not confined to derivatives of a specific germ layer, holds also true for many other XFD genes. Thus, it will be interesting to learn how these genes are activated and, on the other hand, how their protein products interfere with other factors in transcriptional control mechanisms during embryogenesis.

Materials and Methods

cDNA and genomic library screening

Two gastrula stage cDNA libraries and three genomic libraries were hybridized with a 350 bp DNA fragment encoding the entire fork head domain of the *Drosophila* gene *fork head* (kindly provided by H. Jäckle, Göttingen, Germany). Hybridization was performed with the random primed ^{32}P labeled probe in 0,5 M sodium phosphate (pH 7.2), 7% SDS, 1% BSA and 1 mM EDTA at 58°C for 16 h. Final washing was done in 75 mM NaH_2PO_4, 1 mM EDTA and 0.1% SDS at 60°C for 30 min. Inserts of recombinant phages with positive autoradiographic signals were subcloned into pUC 18 (Boehringer, Mannheim). Nucleotide sequences were determined from both directions on an ABI 373 A sequencer using fluorescent labeled primers or terminators.

RNase protection assay

To analyze the temporal expression of XFD genes we have cloned the following restriction fragments into pSPT 18/19 vectors. XFD-1: a 484 bp Eco RI/Bgl II fragment starting 379 nucleotides 5′ to the fork head domain; XFD-2: a 241 bp Apa I/Bgl II fragment starting 136 nucleotides 5′ to the fork head domain; XFD-3: a 339 bp Afl III/Bgl II fragment starting 228 nucleotides 5′ to the fork head domain; XFD-4: a 207 bp Bal I/Rsa I fragment located inside the fork head domain; XFD-5: a 240 bp Sst I/Hae III fragment located inside the fork head domain; XFD-6: a 238 bp Sma I/Pst I fragment ending 116 nucleotides 3′ to the fork head domain; XFD-7: a 304 bp Eco RI/Bgl II fragment starting 102 nucleotides 5' to the fork head domain; XFD-8: a 200 bp Pst I/Bal I fragment starting 134 nucleotides 5′ to the fork head domain; XFD-9: a 340 bp Pst I/Pvu II fragment ending 78 nucleotides 3′ to the fork head domain; XFD-10: a 163 bp Bam HI/Hpa I fragment starting 72 nucleotides 5′ to the fork head domain. The *in vitro* transcription was performed with a commercially available kit (Boehringer, Mannheim) according to the manufactures protocol. ^{32}P-CTP labeled antisense RNA was hybridized with 50 µg RNA, each of oocyte, different developmental stages [blastula (stage 7-9), gastrula (stage 10 -12), neurula (stage 13-16), early somite segregation stage (20 - 26), tailbud (stage 25-30) (stage classification according to Nieuwkoop and Faber, 1967)], adult tissues or yeast tRNA. The same RNA preparations were used in separate control experiments with labeled antisense transcripts from a 311 bp Pvu II/Pst I fragment of the EF-1α sequence pXEF 7 (Pöting *et al.*, 1990). Hybridization and RNA digestion were performed as described (Melton *et al.*, 1984).

Whole-mount in situ hybridizations

The localization of XFD transcripts in *Xenopus* embryos was analysed by using the whole-mount *in situ* hybridization technique (Tautz and Pfeifle, 1989; Hemmati-Brivanlou *et al.*, 1990; Harland, 1991). Some technical modifications were introduced. After puncturing the blastocoel or gastrocoel with a fine needle the embryos were transferred into distilled water for 5 min, fixed in freshly prepared MEMPFA (0.1 M MOPS (pH 7.4), 2 mM EGTA, 1 mM $MgSO_4$, 4% paraformaldehyde) at room temperature for 90 min and stored at - 20°C in ethanol. Antibody incubation was done in maleic acid buffer (100 mM maleic acid (pH 7.5), 150 mM NaCl) with 2% blocking reagent (Boehringer, Mannheim) and 20% heat-treated lamb serum.

Acknowledgments

We thank G. Schütz, Heidelberg, for the communication of unpublished results. We are indebted to S. Hille and M. Köster for help in sequencing. This work was supported by Deutsche Forschungsgemeinschaft (Kn 200/4-1) and by Fonds der Chemischen Industrie.

References

ANG, S-L., WIERDA, A., WONG, D., STEVENS, K.A., CASCIO, S., ROSSANT, J. and ZARET, K.S. (1993). The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead protein. *Development 119:* 1301-1315

BISBEE, C.A., BAKER, M.A., WILSON, A.C., HADJI-AZIMI, I. and FISCHBERG, M. (1977). Albumin phylogeny for clawed frogs. *Science 195:* 785-787.

BOLCE, M.E., HEMMATI-BRIVANLOU, A. and HARLAND, R.M. (1993). XFKH2, a *Xenopus* HNF-3α homologue, exhibits both activin-inducible and autonomous phases of expression in early embryos. *Dev. Biol. 160:* 413-423.

CLARK, K.L., HALAY, E.E., LAI, E. and BURLEY, S.K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature 364:* 412-420.

CLEVIDENCE, D.E., OVERDIER, D.G., TAO, W., QIAN, X., PANI, L., LAI, E. and COSTA, R.H. (1993). Identification of nine tissue-specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA-binding-domain family. *Proc. Natl. Acad. Sci. USA 90:* 3948-3952.

DIRKSEN, M.L. and JAMRICH, M. (1992). A novel, activin-inducible, blastopore lip-specific gene of *Xenopus laevis* contains a fork head DNA-binding domain. *Genes Dev. 6:* 599-608.

HÄCKER, U., GROSSNIKLAUS, U., GEHRING, W. and JÄCKLE, H. (1992). A novel developmentally regulated *Drosophila* gene family encoding the fork head domain. *Proc. Natl. Acad. Sci. USA 89:* 8754-8758.

HARLAND, R.M. (1991). *In situ* hybridisation: an improved whole mount method for *Xenopus* embryos. In *Methods in Cell Biology*, Vol. 36 (Ed. B.K. Kay and H.B. Peng). Academic Press, San Diego, pp. 685-695.

HEMMATI-BRIVANLOU, A., FRANK, D., BOLCE, M.E., BROWN, B.D., SIVE, H.L. and HARLAND, R.M. (1990). Localisation of specific mRNAs in *Xenopus* embryos by whole mount *in situ* hybridization. *Development 110:* 325-330.

HOPE, I.A. (1994). PES-1 is expressed during early embryogenesis in Caenorhabditis elegans and has homology to the fork head family of transcription factors. *Development 120:* 505-514

HROMAS, R., MOORE, J., JOHNSTON, T., SOCHA, C. and KLEMSZ, M. (1993). *Drosophila* forkhead homologues are expressed in a lineage-restricted manner in human hematopoietic cells. *Blood 81:* 2854-2859

KAESTNER, K.H., LEE, K.H., SCHLÖNDORFF, J., HIEMISCH, H., MONAGHAN, A.P. and SCHÜTZ, G. (1993). Six novel members of the murine forkhead gene are developmentally regulated. *Proc. Natl. Acad. Sci. USA 90:* 7628-7631.

KNÖCHEL, S., LEF, J., CLEMENT, J.H., HILLE, S., KLOCKE, B., KÖSTER, M. and KNÖCHEL, W. (1992). Activin A induced expression of a fork head related gene in posterior chordamesoderm (notochord) of *Xenopus laevis* embryos. *Mech. Dev. 38:* 157-165.

KNÖCHEL, W., KORGE, E., BASNER, A. and MEYERHOF, W. (1986). Globin evolution in the genus *Xenopus*. Comparative analysis of cDNAs coding for adult globin polypeptides of *Xenopus borealis* and *Xenopus tropicalis*. *J. Mol. Evol. 23:* 211-223.

KRIEG, P.A., VARNUM, S.M., WORMINGTON, W.M. and MELTON, D.A. (1989). The mRNA encoding elongation factor 1-a (EF-1α) is a major transcript at the midblastula transition in *Xenopus*. *Dev. Biol. 133:* 93-100.

LAI, E., CLARK, K.L., BURLEY, S.K. and DARNELL Jr, J.E. (1993). Hepatocyte nuclear factor 3/fork head or "winged helix" proteins: A family of transcription factors of diverse biologic function. *Proc. Natl. Acad. Sci. USA 90:* 104421-10423.

LAI, E., PREZIOSO, V.R., SMITH, E., LITVIN, O., COSTA, R.H. and DARNELL, J.E. (1990). HNF-3A, a hepatocyte enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev. 4:* 1427-1436.

LAI, E., PREZIOSO, V.R., TAO, W., CHEN, W.S. and DARNELL, J.E. (1991). Hepatocyte nuclear factor 3α belongs to a gene family in mammals that is homologous to the *Drosophila* homeotic gene fork head. *Genes Dev. 5:* 416-427.

LEF, J., CLEMENT, J.H., OSCHWALD, R., KÖSTER, M. and KNÖCHEL, W. (1994). Spatial and temporal transcription patterns of the forkhead related XFD-2/XFD-2' genes in *Xenopus laevis* embryos. *Mech. Dev. 45:* 117-126.

LI, J. and VOGT, P.K. (1993). The retroviral oncogene qin belongs to the transcription factor family that includes the homeotic gene fork head. *Proc. Natl. Acad. Sci. USA 90:* 4490-4494.

LI, C., LUSIS, A.J., SPARKES, R., NIRULA, A. and GAYNOR, R. (1992a). Characterization and chromosomal mapping of the gene encoding the cellular DNA binding protein ILF. *Genomics 13:* 665-671.

LI, C., LUSIS, A.J., SPARKES, R., TRAN, S-M. and GAYNOR, R. (1992b). Characterization and chromosomal mapping of the gene encoding the cellular DNA binding protein HTLF. *Genomics 13:* 658-664.

MELTON, D.A., KRIEG, P.A., REBAGLIATI, M.R., MANIATIS, T., ZINN, K. and GREEN, M.R. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res. 12:* 7035-7056.

MILLER, L.M., GALLEGOS, M.E., MORISSEAU, B.A. and KIM, S.K. (1993). lin-31, a Caenorhabditis elegans HNF-3/fork head transcription factor homolog, specifies three alternative cell fates in vulval development. *Genes Dev. 7: 933-947.*

MIURA, N., WANAKA, A., TOHYAMA, M. and TANAKA, K. (1993). MFH-1, a new member of the fork head domain family, is expressed in developing mesenchyme. *FEBS Lett. 326:* 171-176.

MONAGHAN, A.P., KAESTNER, K.H., GRAU, E. and SCHÜTZ, G. (1993). Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 α, β and γ genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development 119:* 567-578.

MURPHY, D.B., WIESE, S., BURFEIN, P., SCHMUNDT, D., MATTEI, M-G., SCHULZ-SCHAEFFER, W. and THIES, U. (1994). Human brain factor 1, a new member of the fork head gene family. *Genomics 21:* 551-557.

NEWPORT, J. and KIRSCHNER, M. (1982). A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell 30:* 687-696.

NIEUWKOOP, P.D. and FABER, J. (1967). *Normal Table of Xenopus laevis (Daudin).* North Holland Publishing Co., Amsterdam.

PÖTING, A., DANKER, K., HARTMANN, L., KÖSTER, M., WEDLICH, D. and KNÖCHEL, W. (1990). Two different mRNAs coding for identical elongation factor 1a (EF-1α) polypeptides in *Xenopus laevis* embryos. *Differentiation 44:* 103-110.

RUIZ I ALTABA, A. and JESSELL, T.M. (1992). *Pintallavis*, a gene expressed in the organizer and midline of frog embryos: involvement in the development of neural axis. *Development 116:* 81-93.

RUIZ I ALTABA, A., PREZIOSO, V.R., DARNELL, J.E. and JESSELL, T.M. (1993). Sequential expression of HNF-3β and HNF-3α by embryonic organizing centers: the dorsal lip/node, notochord and floor plate. *Mech. Dev. 44:* 91 - 108.

SASAKI, H. and HOGAN, B.L.M. (1993). Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development 118:* 47-59.

SASAKI, H. and HOGAN, B.L.M. (1994). HNF-3β as a regulator of floor plate development. *Cell 76: 103-115.*

TAO, W. and LAI, E. (1992). Telencephalon restricted expression of BF-1, a new member of the HNF-3/fork head gene family, in the developing rat brain. *Neuron 8:* 957-966.

TAUTZ, D. and PFEIFLE, C. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma 98:* 81-85.

WEIGEL, D., JÜRGENS, G., KÜTTNER, F., SEIFERT, E. and JÄCKLE, H. (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. *Cell 57:* 645-658.

WEIGEL, D. and JÄCKLE, H. (1990). Fork head: A new eukaryotic DNA binding motif? *Cell 63:* 455-456.

XANTHOPOULOS, K.G., PREZIOSO, V.R., CHEN, W.S., SLADEK, F.M., CORTESE, R. and DARNELL Jr, J.E. (1991). The different tissue transcription patterns of genes for HNF-1, C/EBP, HNF-3, and HNF-4, protein factors that govern liver-specific transcription. *Proc. Natl. Acad. Sci. USA 88:* 3807-3811.

ZHU, G., MULLER, E.G.D., AMACHER, S.L., NORTHROP, J.L. and DAVIS, T.N. (1993). A dosage-dependent suppressor of a temperature-sensitive calmodulin mutant encodes a protein related to the fork head family of DNA-binding proteins. *Mol. Cell. Biol. 13:* 1779-1787.