

The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/*forkhead* proteins

Siew-Lan Ang¹, Allison Wierda², David Wong², Kimberly A. Stevens², Stephanie Cascio², Janet Rossant^{1,3} and Kenneth S. Zaret²

¹Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Canada, M5G 1X5

²Section of Biochemistry, Division of Biology and Medicine, Brown University, Providence, RI 02912, USA

³Department of Molecular and Medical Genetics, University of Toronto, Ontario, Canada M5G 1X5

SUMMARY

Little is known about genes that govern the development of the definitive endoderm in mammals; this germ layer gives rise to the intestinal epithelium and various other cell types, such as hepatocytes, derived from the gut. The discovery that the rat hepatocyte transcription factor HNF3 is similar to the *Drosophila forkhead* gene, which plays a critical role in gut development in the fly, led us to isolate genes containing the HNF3/*forkhead* (HFH) domain that are expressed in mouse endoderm development. We recovered mouse HNF3 β from an embryo cDNA library and found that the gene is first expressed in the anterior portion of the primitive streak at the onset of gastrulation, in a region where definitive endoderm first arises. Its expression persists in axial structures derived from the mouse equivalent of Hensen's node, namely definitive endoderm and notochord, and in the ventral region of the developing neural tube. Expression of the highly related gene, HNF3 α , appears to initiate later than HNF3 β and is first seen in midline endoderm cells. Expression subsequently appears in notochord, ventral neural tube, and gut endoderm in patterns

similar to HNF3 β . Microscale DNA binding assays show that HNF3 proteins are detectable in the midgut at 9.5 days p.c. At later stages HNF3 mRNAs and protein are expressed strongly in endoderm-derived tissues such as the liver. HNF3 is also the only known hepatocyte-enriched transcription factor present in a highly de-differentiated liver cell line that retains the capacity to re-differentiate to the hepatic phenotype. Taken together, these studies suggest that HNF3 α and HNF3 β are involved in both the initiation and maintenance of the endodermal lineage. We also discovered a novel HFH-containing gene, HFH-E5.1, that is expressed transiently in posterior ectoderm and mesoderm at the primitive streak stage, and later predominantly in the neural tube. HFH-E5.1 is highly similar in structure and expression profile to the *Drosophila* HFH gene FD4, suggesting that HFH family members have different, evolutionarily conserved roles in development.

Key words: endoderm lineage, mammalian embryogenesis, transcription factor, HNF3, *forkhead*

INTRODUCTION

The molecular mechanisms underlying the development of the definitive endoderm in vertebrates have remained obscure; this important lineage gives rise to the lining of the gut and its many outpockets during organogenesis. Even in *Xenopus*, where there is considerable information on the signalling mechanisms involved in the formation of the mesoderm lineage, there is a paucity of studies on the specification and determination of the endoderm lineage.

Cell fate studies in chick and mouse show that the definitive endoderm does not arise from the primary hypoblast or primitive endoderm (Sanders et al., 1978; Gardner and Rossant, 1979), but arises de novo from ectoderm or epiblast at gastrulation (Rosenquist, 1971, 1972; Schoenwolf et al., 1992; Poelmann, 1981; Lawson and Pederson, 1987; Lawson et al., 1991). Single cell marking studies in the mouse have shown that, at the early primitive streak stage,

there is a subpopulation of cells in the axial endoderm overlying the anterior of the streak that contributes to the gut endoderm (Lawson and Pederson, 1987). The rest of the endoderm layer at this stage will be displaced from the embryonic region and become extraembryonic endoderm, whereas definitive endoderm cells are most likely derived from underlying epiblast cells (Lawson et al., 1991). By the mid-streak stage, the anterior end of the streak reaches the distal tip of the embryo and forms the node, by analogy to Hensen's node in the chick. The head process arises as a cranial extension of the node and appears to contribute to definitive endoderm by displacing primary endoderm cells in the midline (Poelmann, 1981). Cell fate studies in chick have shown that cells in early Hensen's node can contribute to gut endoderm, as well as to notochord and the floorplate of the neural tube (Selleck and Stern, 1991; Schoenwolf et al., 1992). The limited information available in mouse is consistent with this model (Lawson et al., 1991). Thus there

A. HFH-E5.1

```
TGTCCTACGGTCAAGGAGCGGAACCGAACACCCAGCAGCGTCTCTGGATCCCTGAG 60
GCGGGGTCTGCCACTGTCTCACGGGCACTCCGACGCGGAGTTCGGTCTCCAGCAGGGC 120
GCACGACCCACGAGCGCGCTGCGCAGCCCGGGAGAGACTCTACCGGGCTGAGTCCAA 180
GCCTAGCTGGCAGCGAGTCTCAGTCTCCCGACGCTGCGCTCCGGTCGCGCGCAGGGAG 240
GGACTCGGCCACCAGCCACGCCATCCCGGATGCGGACCGGCAACTTGAACAACCTTTAAG 300
ATTCGACCACTCTCCGGGCGCGGACCCCTGAGCAAGTGGCCAAGCAAGAAGATGCCT 360
M P
CGGCCGGGGCCGAACACGTACAGCGACCAGAAGCCGCCCTACTCATACTCTCGCTGACC 420
R P G R N T Y S D Q K P P Y S Y I S L T
GCGATGGCCATCCAGAGCTCTCCGGAGAAGATGCTGCGGCTCAGCGAGATCTACAAGTTC 480
A M A I Q S S P E K M L P L S E I Y K F
ATCATGGACCGCTTCCCCTACTACCGGGAGAACACGCGAGCGCTGGCAGAACAGCCTGCGC 540
I M D R F P Y R E N T Q R W Q N S L R
CCCTTGACCGTAAGACAT 559
P L T V R H
```

B.

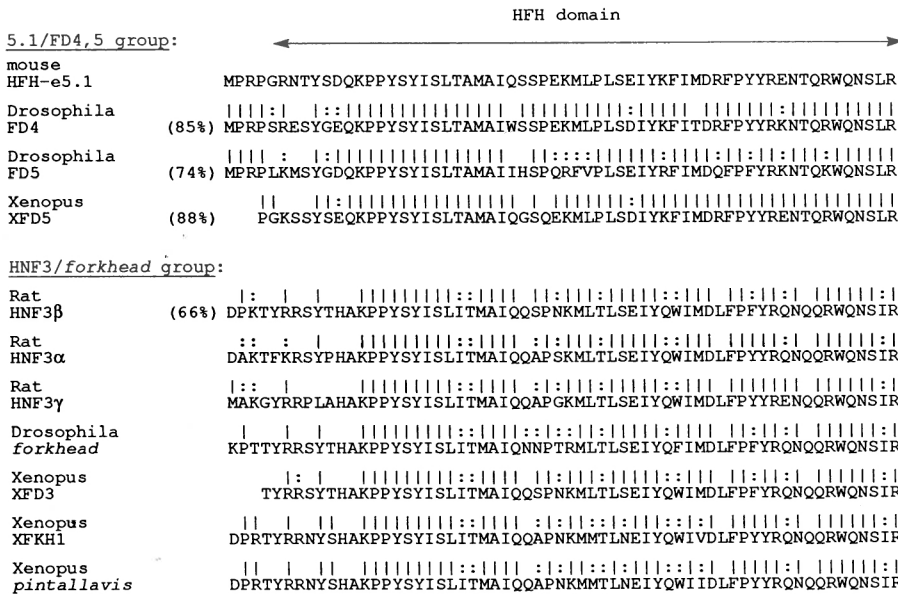


Fig. 1. Sequence analysis of isolated mouse HNF3β and HFH-E5.1 cDNAs. (A) cDNA sequence of HFH-E5.1 extends from the apparent 5' noncoding region through the middle of the HFH domain. (B) Amino acid sequence comparison of HFH proteins with the HFH-E5.1 sequence. HFH proteins and their sources are listed to the left of the sequence, with the percentage identity with HFH-E5.1 shown in parentheses. Vertical lines above amino acids indicate identity with HFH-E5.1; dots indicate conservative amino acid substitutions (Schwartz and Dayhoff, 1979). Sequences are grouped by parsimony analysis.

may be two early phases of definitive endoderm formation, one from anterior of the streak before formation of the node and one from the later head process.

With these cell fate studies in mind, one can predict where genes involved in the formation of the lineage should be expressed and thus focus attention on appropriate candidate genes. One source of candidate genes for endoderm formation are the tissue-specific regulatory factors involved in the later differentiation of endoderm-derived tissues. Most information is available on the liver, where a number of factors have been identified by their ability to bind to regulatory elements of liver-specific genes. These factors include HNF1, HNF3, HNF4, C/EBP, and DBP, all of which are highly enriched in expression in the adult liver (reviewed by Zaret, 1993). However, examination of their embryonic expression eliminates many of them from having a role in early definitive endoderm development. For example, the POU-homeodomain protein HNF1α, which is necessary for high-level expression of many liver-specific genes (Courtois et al., 1987), is first expressed in the early mouse liver at day 10.5 p.c., well after initiation of the endoderm lineage (Blumenfeld et al., 1991; De Simone et al., 1991). Similarly, mRNA for the liver-enriched leucine zipper protein, C/EBP,

is first detected in the mouse liver at day 12.5 (Kuo et al., 1990). HNF4 is not detected in definitive endoderm until about 8.5 to 9.5 days p.c. (Sladek et al., 1990; Cereghini et al., 1992), and DBP is expressed late in ontogeny, after birth (Mueller et al., 1990).

By contrast, HNF3 transcription factors display characteristics expected for proteins important for the endoderm lineage. HNF3α, HNF3β, and HNF3γ are products of three distinct genes, all of which are expressed exclusively in endoderm-derived tissues in the adult, such as liver, lung, and intestine (Lai et al., 1991), and play an important role in gene activation (Costa et al., 1989; Liu et al., 1991). The HNF3 family is characterized by a highly conserved DNA binding domain and two short carboxy-terminal domains (Lai et al., 1991). The DNA binding domain is 90% identical to a segment of the *Drosophila forkhead* gene (Weigel et al., 1989; Weigel and Jäckle, 1990); this conserved sequence does not conform to any previously defined DNA binding domain, and therefore defines a new class of transcription factors, designated HFH (Clevidence et al., 1993). Numerous HFH genes expressed in various cell types are being discovered in organisms as diverse as yeast, nematodes, zebrafish, and humans (see Clevidence et al.,

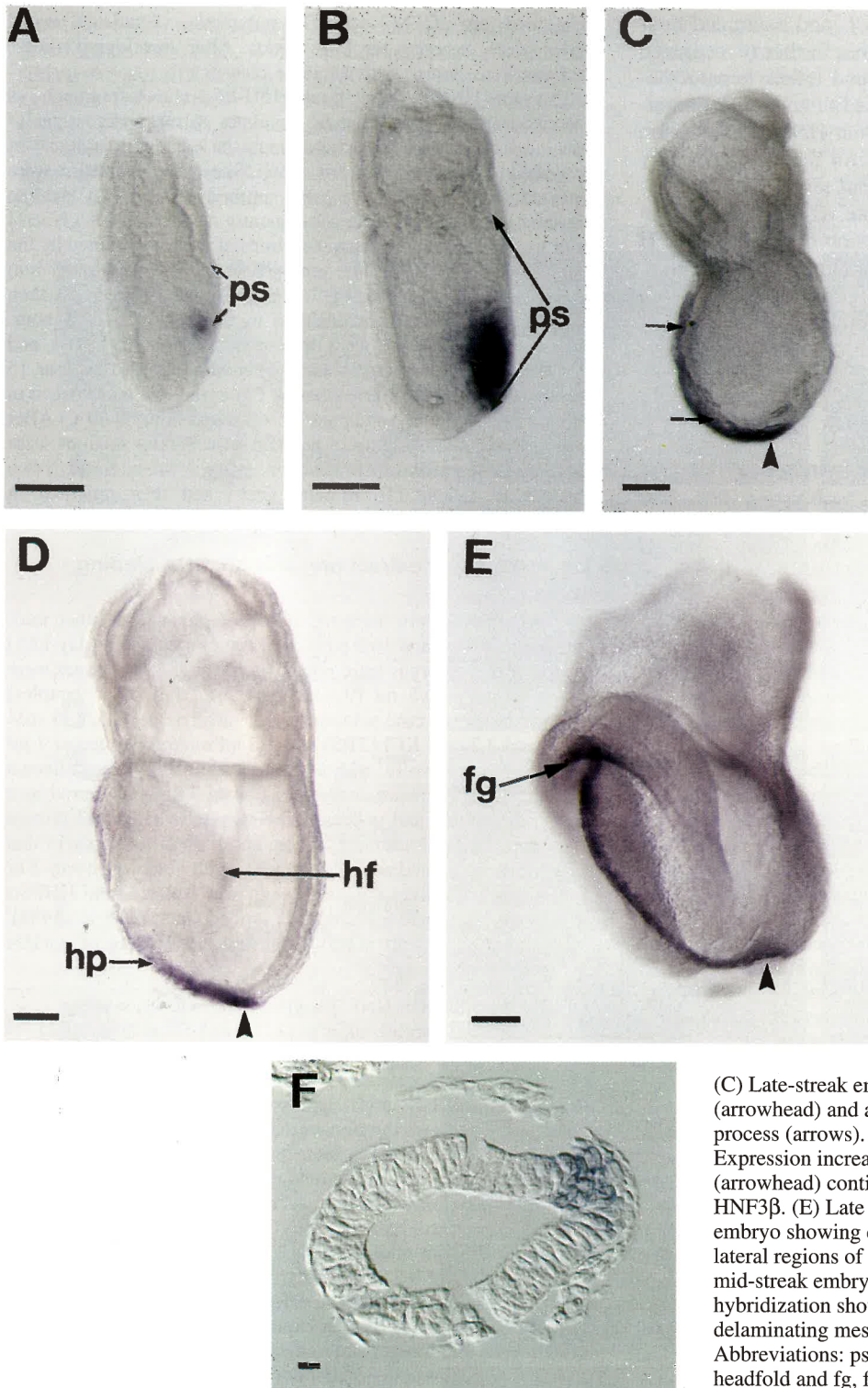


Fig. 2. Expression of HNF3 β transcripts during gastrulation and neurulation, as detected by whole-mount in situ hybridization. (A) Very early streak embryo showing small patch of cells expressing HNF3 β in the anterior part of primitive streak. (B) Mid-streak embryo. HNF3 β expression domain expands to include a larger group of cells in the anterior third of the primitive streak.

(C) Late-streak embryo showing expression in the node (arrowhead) and anteriorly migrating cells of the head process (arrows). (D) Headfold stage embryo. Expression increased in the anterior midline. The node (arrowhead) continued to show strong expression of HNF3 β . (E) Late headfold/early foregut pocket stage embryo showing expression of HNF3 β in medial and lateral regions of foregut. (F) Transverse section of a mid-streak embryo stained by whole-mount in situ hybridization showing that posterior epiblast and delaminating mesoderm cells express HNF3 β . Abbreviations: ps, primitive streak; hp, head process; hf, headfold and fg, foregut. Scale bar, 100 μ m.

1993; Miller et al., 1993; Strähle et al., 1993); the HFH family may be as extensive as the homeobox genes. Several different HFH genes have now been isolated from *Drosophila* and shown to have highly specific expression patterns, suggestive of roles in different developmental processes (Grossinklaus et al., 1992; Häcker et al., 1992). However, the original *forkhead* gene, which shows the strongest homology to the HNF3 genes, is required for the

formation of the anterior and posterior gut in *Drosophila* (Jürgens and Weigel, 1988). This suggested to us that HNF3 or related genes might be involved in the establishment of the endoderm lineage as well as its terminal differentiation.

Accordingly, we searched for HFH-related sequences in a cDNA library derived from late gastrulating mouse embryos, around the time of definitive endoderm formation. We found the mouse homolog of HNF3 β as well as a novel

HFH gene product, termed HFH-E5.1, and examined their expression in early embryogenesis. We further investigated the relative expression of HNF3 and other hepatocyte-enriched transcription factors in a conditionally differentiating hepatic cell line. We conclude that HNF3 β expression satisfies conditions expected for a factor that helps not only to establish the endoderm lineage, but also to maintain it. However, this may not be its only role, as it is expressed in other embryonic tissues. The expression of the novel HFH gene is also very restricted around the time of gastrulation, but this gene is not expressed in the endoderm, suggesting that HFH genes play multiple roles in development.

MATERIALS AND METHODS

Cloning of HFH cDNAs and DNA sequence analysis

Low-stringency hybridization conditions for isolating HFH genes were established by examining the Southern blot hybridization profiles of mouse genomic DNA probed with different segments of the rat HNF3 α cDNA (Lai et al., 1990). When the nitrocellulose blots (Sartorius) were hybridized in 5 \times SSC, 0.5% SDS, 1 mM EDTA at 32°C for 18 hours, and washed in 0.1 \times SSC, 0.1% SDS at 23°C four times for 5 minutes, then at 39°C twice for 1 hour, coding segment probes from both upstream and downstream of the DNA binding domain hybridized to single *Eco*RI and *Hind*III fragments, whereas a 473 bp *Ppu*M1 fragment encompassing the DNA binding domain hybridized to about 10 different DNA fragments in each digest (data not shown). The *Ppu*M1 fragment was used with the same hybridization conditions with BA85 filters (Schleicher & Schuell) to screen a total of 1.5 \times 10⁵ phage from a λ gt10 cDNA library prepared from mRNA of 8.5 day mouse embryos (kindly provided by Brigid Hogan; Farhner et al., 1987); *E. coli* C600 was the host. Restriction fragments encompassing the HFH homology were identified by hybridization and sequenced with the dideoxy procedure. cDNA sequences encoding the carboxy terminus of HNF3 β were isolated by PCR from a λ gt11 cDNA library prepared from mouse liver (kindly provided by R. Costa, E. Paulson, and J. Darnell). DNA sequences were analyzed with software from the Genetics Computer Group, Inc.

Staging of embryos

Mouse embryos at various stages of gestation were obtained by mating random-bred CD1 animals (Charles River Canada, Montreal). Noon on the day of vaginal plug appearance was considered to be embryonic day 0.5; gastrulation occurred between 6.5 and 8.5 days. Very early streak embryos showed asymmetry and were classified morphologically by the thickening of the epiblast in the posterior end, these lacked a distinguishable mesoderm layer. Mid-streak embryos contained mesoderm only in the posterior region. In late-streak embryos, mesoderm has reached its anterior-most extent, the head process has begun to extend, and the amnion is closed. Headfold stage embryos showed neural folds at the anterior end. Late headfold stage/foregut stage embryos contain a foregut pocket and more prominent neural folds.

Section and whole-mount in situ hybridization

Paraffin sections of different stages of mouse embryos were hybridized to ³⁵S-labelled anti-sense RNA probes, according to Yamaguchi et al. (1992). The HNF3 β probe was synthesized from a 566 bp cDNA clone consisting mostly of unique sequences outside the DNA binding domain, using T7 polymerase. RNA probe for HNF3 α was transcribed from a 480 bp cDNA clone, which does not include the HFH DNA binding domain. HFH-E5.1 probe was made from a 580 bp cDNA fragment (Fig. 1B)

subcloned into pGEM3 using T7 polymerase. Emulsion coated slides were exposed for 2-4 weeks. After developing slides, sections were stained with toluidine blue.

The same HNF3 α , HNF3 β and HFH-E5.1 cDNA fragments, as described above, were used as templates to make digoxigenin-labelled RNA probes for whole-mount in situ hybridizations, as described (Conlon and Rossant, 1992). Sites of hybridization were detected with an anti-digoxigenin antibody coupled to alkaline phosphatase. Embryos were subsequently cleared in 50% glycerol prior to photography. To obtain sections of embryos stained by the whole-mount procedure, the embryos were first rehydrated into phosphate-buffered saline (PBS) containing 0.1% Tween 20, then postfixed in 4% paraformaldehyde in PBS at 4°C for 1 hour. Following this, embryos were dehydrated through 70%, 80%, and 90% ethanol for 30 minutes each. This was followed by four 15 minute incubations in 100% ethanol. Embryos were then soaked in xylene (2 \times 15 minutes) and paraffin wax (3 \times 1 hour) at 60°C. After orientation and embedding in paraffin wax, 10 μ m sections were cut from the paraffin blocks. Sections were dewaxed through two changes of xylene (10 minutes each) and then mounted in Permount. The sections were photographed using Nomarski optics.

Embryo nuclear extract preparation, DNA binding assays

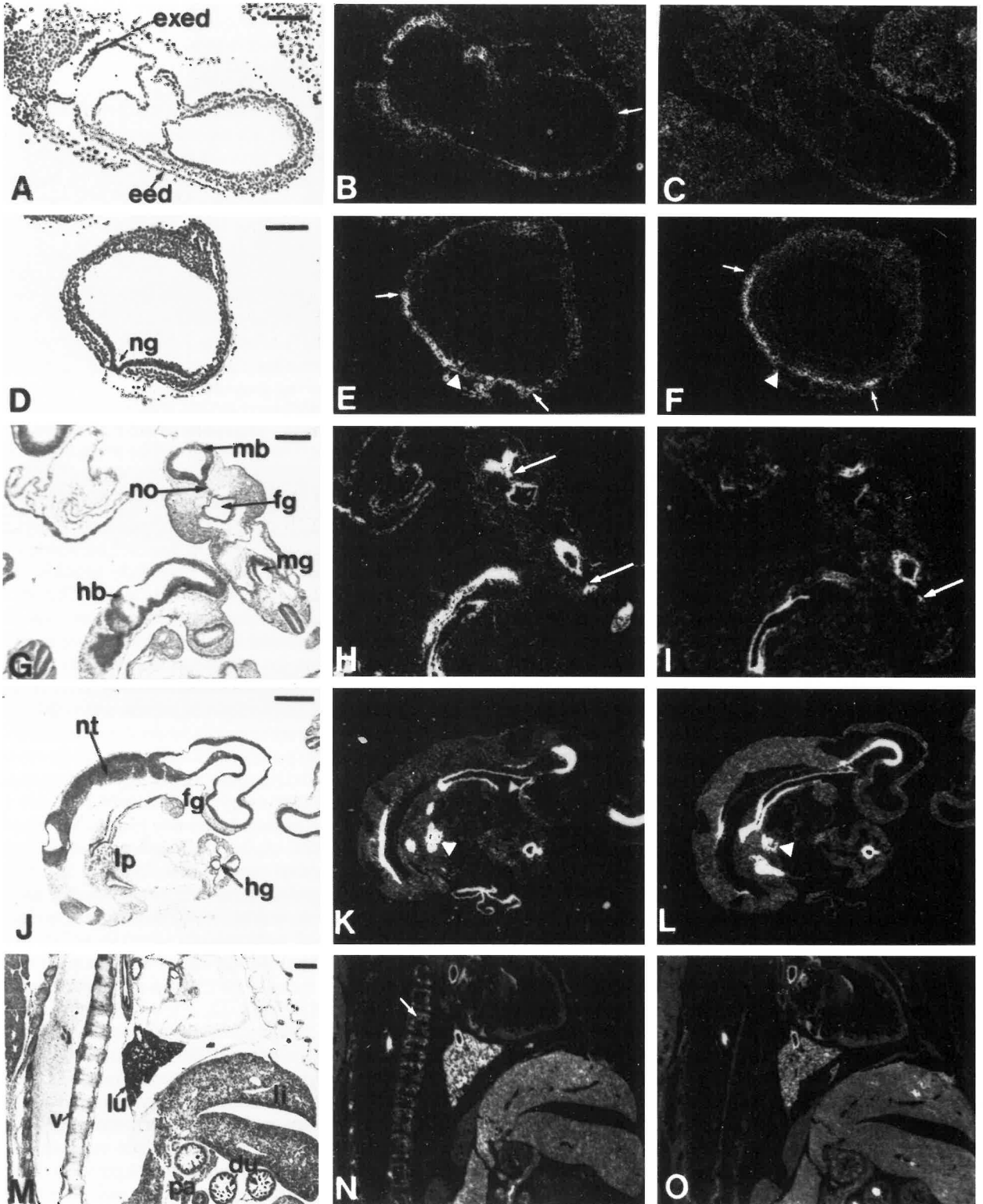
Mouse embryos were dissected free of decidua, and either mid-sections (days 9.5 and 10.5 p.c.) or newly formed livers (day 12.5) from multiple embryos were pooled and kept at 4°C. Tissues were homogenized in 0.5 ml (9.5 day) or 2 ml (all other samples) volumes of an ice cold solution of 25 mM Tris pH 7.5, 138 mM NaCl, and 2.7 mM KCl (TBS) in a 1.5 ml microfuge tube or 7 ml glass tube, respectively, with a teflon pestle. Cells were collected by centrifugation, resuspended into 0.5 ml TBS, transferred to a fresh 1.5 ml tube, and pelleted. Nuclei were isolated and protein was extracted as described by Schreiber et al. (1989), except that nuclei were incubated on ice in buffer C with vortexing every 2 or 3 minutes. Adult liver nuclear protein was isolated, and HNF3 α protein was translated in vitro, as described (Jackson et al., 1993).

Electromobility shift assays employed either 5 μ g of nuclear

Fig. 3. Comparison of HNF3 β and HNF3 α expression using sectioned in situ hybridization. A,D,G,J, and M are bright-field views. The remaining panels are dark-field views showing autoradiography signals. (B,C) At the late primitive streak stage HNF3 β (B) and HNF3 α (C) expression are both found in embryonic endoderm throughout the posterior of the embryo. HNF3 β is also expressed in anterior midline mesoderm (small arrow) and also in extraembryonic endoderm. (E,F) Transverse sections of headfold stage embryos showing expression of HNF3 β (E) in all three germ layers and HNF3 α (F) only in the endoderm layer in the anterior midline. (H,I) Transverse sections through the midbrain region of 8.5 days embryos showing HNF3 β expression (H) spreading more dorsally on either side of the floorplate than HNF3 α expression (I). Similar expression pattern for both genes were observed in foregut and midgut. Ventral spinal cord at posterior end shows HNF3 β but not HNF3 α expression. Both genes are also expressed in notochord (arrows). (K,L) The liver primordium (open arrowheads) in 9.5 day embryos expresses both HNF3 β (K) and HNF3 α (L) genes. (N,O) At 16.5 days, endoderm expression of HNF3 β (N) and HNF3 α (O) persists in lung, liver, pancreas and gut. In addition, HNF3 β expression was also seen in the cartilage of the vertebrae (N; arrow). Abbreviations: exed, extraembryonic endoderm; eed, visceral embryonic endoderm; ng, neural groove; mb, midbrain; no, notochord; fg, foregut; mg, midgut; hb, hindbrain; nt, neural tube; lp, liver primordium; lu, lung; li, liver; du, duodenum and pa, pancreas. Scale bar 100 μ m (A,D) and 5 mm (G,J,M).

protein or 1 μ l of in vitro translation product in a 20 μ l volume, as described (Jackson et al., 1993). Binding reactions included either 500 ng (9.5 and 10.5 day extracts), 750 ng (12.5 day extract), 1500 ng (adult liver extract), or 250 ng (HNF3 in vitro translation product) poly (dI-dC). Oligonucleotide probes, end-labelled with 32 P, consisted of the eG binding site for HNF3 from the mouse serum albumin enhancer (Liu et al., 1991), the APF-1 binding site for HNF4 from the apolipoprotein CIII promoter (Costa et al.,

1990), and the pC binding site for NF-Y from the mouse albumin promoter (Raymondjean et al., 1988). Nonlabelled competitor DNAs were used at a 100-fold molar excess, and included the eG and pC sites as well as a strong HNF3 binding site from the mouse transthyretin gene promoter (Costa et al., 1989). Western blot analysis was performed with nuclear extracts from liver and H2.35 cells, as described (Liu et al., 1991). The HNF1 α -specific antibody was kindly provided by Dr G. Crabtree.



Protein-DNA crosslinking

HNF3 binding site probes suitable for UV-crosslinking were prepared by annealing an 8 bp primer to an oligonucleotide of the bottom strand sequence of the eG site, and extending the primer with Klenow polymerase in the presence of dGTP, [³²P]dCTP, [³²P]dATP, and bromodeoxyuridine-substituted dUTP (Br-dUTP) (Chodosh et al., 1987). Binding reactions were performed in lids of 0.5 ml microfuge tubes and were subsequently exposed to ultraviolet light from a Spectroline EB-280C lamp with a 302 nm filter, yielding 860 mW/cm² at 15 cm, at a distance of 7 mm for 1 minute. Reaction products were resolved by the electromobility shift assay, identified by autoradiography, excised as gel fragments, and equilibrated in stacking gel loading buffer (Laemmli, 1970) for several hours. The fragments were then placed into the wells of 10% polyacrylamide gels containing SDS (Laemmli 1970); after electrophoresis, gels were dried and autoradiographed. In some experiments, the gel fragments were pre-equilibrated in 0.125 M Tris-HCl pH 6.8, 0.1% SDS, and 1 mM EDTA, then digested with 50-500 ng V8 protease in pre-equilibration buffer containing 10% glycerol (Gooderham, 1984), prior to electrophoresis in a 15% polyacrylamide-SDS gel.

RESULTS

Cloning HNF3/forkhead sequences expressed during mouse endoderm differentiation

To isolate HFH genes expressed during endoderm differentiation, we used low-stringency hybridization conditions to screen a λgt10 cDNA library prepared from mouse embryos at 8.5 days gestation (Fahrner et al., 1987), using a rat HNF3α DNA binding domain probe (Lai et al., 1990). DNA sequence analysis of the most strongly hybridizing cDNA showed that it contained 570 bp that were 96% identical to rat HNF3β over the carboxy-terminal sequences unique to HNF3β (sequence submitted to Genbank). By comparison, both rat HNF3β and the mouse clone only show extensive similarity to rat HNF3α or HNF3γ in the DNA binding domain (Lai et al., 1991). We conclude that this cDNA encodes mouse HNF3β. Additional segments of HNF3β cDNA were cloned by screening a mouse liver cDNA library; sequences of overlapping segments obtained from the liver and embryo libraries were identical. Northern blot analysis showed that mouse HNF3β is encoded by a 2.2 kb mRNA (data not shown), similar to HNF3β in the rat (Lai et al., 1991), indicating that the cDNA isolated is not full length. Comparison with the rat gene indicates that the cDNA sequence begins in the first third of the DNA binding domain and lacks amino-terminal sequences.

The second cDNA isolated from the mouse embryo library contained an open reading frame which encodes a novel HFH-related sequence, HFH-E5.1 (Fig. 1A). The open reading frame begins with an ATG at nucleotide 355, and has an A residue 3 bp upstream, as expected for an efficient translation start site (Kozak, 1986). Another ATG exists at position 270 in a different reading frame, but it contains a C residue at position -3, which is inhibitory to translation (Morle et al., 1985). There is a stop codon in frame 183 bp upstream of the ATG at 355, making it likely that this ATG is the initiation codon. The cDNA is incomplete, and ends in the middle of the putative HFH domain. The putative DNA binding domain of HFH-E5.1 is 66%

identical to HNF3, compared to the 90% identity between HNF3 proteins and *forkhead* proteins (Fig. 1B). The HFH-E5.1 cDNA is similar in structure to that for HTLF (Li et al., 1992) and *lin-31* (Miller et al., 1993), in that all appear to have relatively long 5' noncoding regions and only several amino acids prior to their HFH domains. However, of the HFH genes that have been cloned to date, HFH-E5.1 exhibits the most similarity, 85% over the HFH domain, with the FD4 cDNA segment recently isolated from *Drosophila* (Häcker et al., 1992), and with the XFD5 genomic sequence from *Xenopus* (Knöchel et al., 1992).

Alignment of published HFH domain amino acid sequences with the HFH-E5.1 clone shows that they can be arranged in groups that contain common amino acid substitutions; the distinctions are emphasized when amino acids upstream of the HFH domain are considered (Fig. 1B). HFH-E5.1, FD4 and FD5 (Häcker et al., 1992), and XFD5 (Knöchel et al., 1992) contain a highly similar sequence at the amino terminus of the HFH domain, whereas HNF3 proteins, *forkhead* protein, and XFD3 (Knöchel et al., 1992), XFKH1 (Dirksen and Jamrich, 1992) and *pintallavis* proteins (Ruiz i Altaba and Jessel, 1992) from *Xenopus* comprise a different, highly related group. We conclude that HFH-E5.1 is distantly related to HNF3/*forkhead* genes and that it represents another class of HFH genes expressed in mouse embryos during gastrulation.

In situ hybridization analysis of HNF3β gene expression

Both whole-mount and sectioned in situ hybridization techniques were used to determine the temporal and spatial expression of HNF3β during gastrulation and neurulation. The earliest expression of HNF3β detected by whole-mount in situ hybridization was at the very early primitive streak stage, in a small patch of epiblast cells at the anterior of the streak (Fig. 2A). This initial staining pattern expands to include more cells so that by the mid-streak stage, a large group of cells within the anterior third of the streak express HNF3β (Fig. 2B). Sections of such embryos showed that expression at these early stages is confined to epiblast cells and delaminating mesoderm cells in the anterior streak (Fig. 2F). At the late streak stage, as the definitive node becomes apparent at the distal tip of the embryo, expression of HNF3β was found in both layers of the definitive node (Fig. 2C and data not shown). Anteriorly migrating mesoderm and endoderm cells of the head process, which begins around this stage, also express HNF3β (Fig. 2C). Slightly later, at the headfold stage, whole-mount stained embryos revealed that the neural plate, notochord and the underlying midline endoderm all expressed the gene (Figs. 2D). By late headfold/early foregut pocket stage, endoderm expression became broader at the anterior end and spread into the developing foregut pocket (Fig. 2E). ³⁵S-labelled in situ hybridization to tissue sections of late streak and headfold stage embryos basically confirmed these results, but showed broader HNF3β expression in the endoderm lineage (Fig. 3B,E). At the late streak stage, expression was seen in patches of cells in the endoderm layer overlying the entire extent of the streak, and not just at the anterior end. At headfold stages, anterior endoderm expression was not restricted to the midline but extended laterally on either side

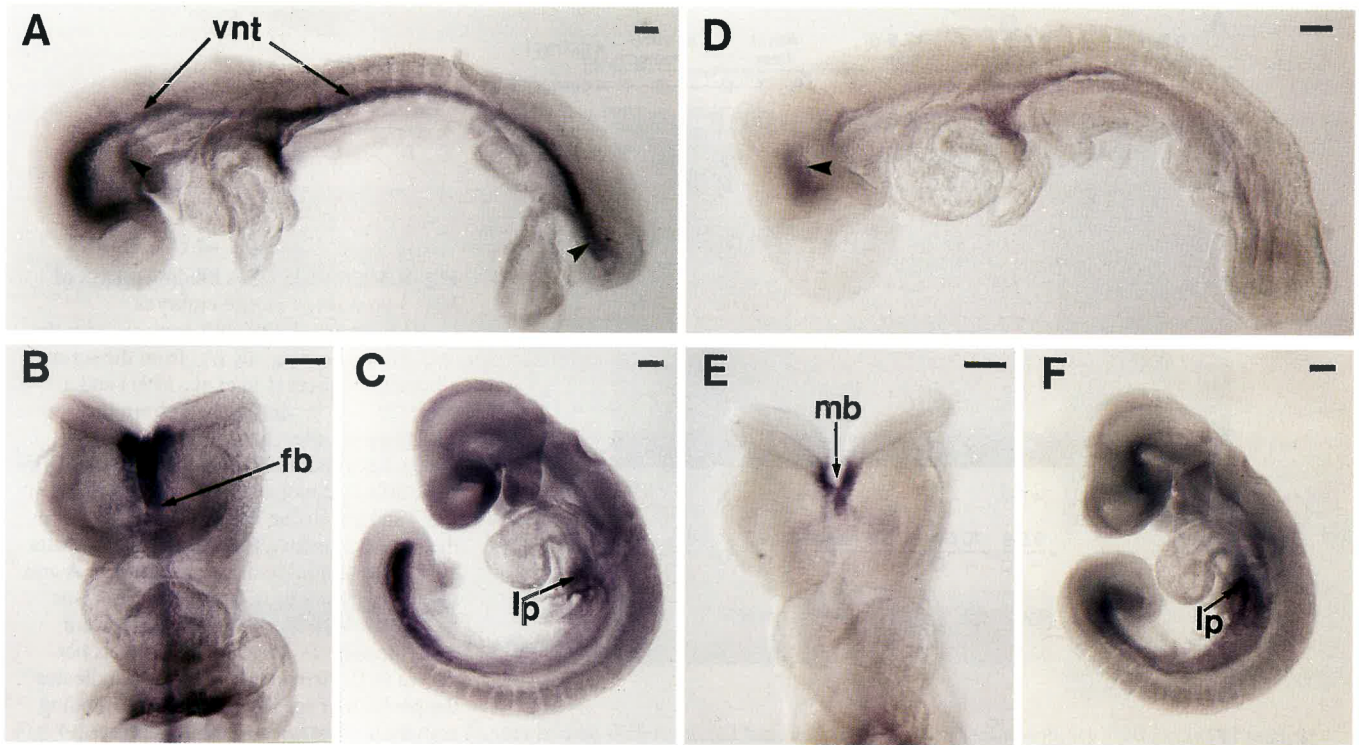


Fig. 4. Comparison of HNF3 β (A-C) and HNF3 α (D-F) expression in 8.5-9.0 day embryos using whole-mount in situ hybridization. HNF3 β and HNF3 α transcripts (A and D respectively) are found in ventral neural tube, notochord and entire gut in 8-9 somite embryo. (A,B) HNF3 β expression includes forebrain, and within the neural tube, expression extends more laterally than the floorplate. (D,E) HNF3 α is expressed in ventral midbrain but not in the rest of the neural tube. The midbrain staining encompasses floorplate and some lateral areas but is narrower than the HNF3 β domain. (C,F) Both HNF3 α (C) and HNF3 β (F) are more strongly expressed in liver primordium of 9.0 day (about 15 somites) embryos than in the surrounding endoderm. Abbreviations: vnt, ventral neural tube; fb, forebrain; mb, midbrain and lp, liver primordium. Scale bar, 100 μ m.

(Fig. 3E). The broader domains of endoderm expression in late-streak and headfold stages detected on sections, compared to whole embryos, presumably reflects a higher sensitivity of the former method.

At 8.5 days (6-10 somites), expression persisted in the ventral part of the developing neural tube, from the forebrain to the posterior end, in a region encompassing the floorplate and cells lateral to it (Fig. 4A). Within the midbrain region, sections revealed that the staining spread even more dorsally on either side of the floorplate (Fig. 3H). These sections also showed expression in notochord and gut. In the gut, stronger expression was observed in foregut and hindgut at early somite stage (data not shown), but by the 10 somite stage the entire gut showed similar levels of expression (Fig. 4A). Interestingly, HNF3 β expression was elevated in a portion of the gut beneath the cardiac region (Fig. 4A,C). The locally elevated levels of HNF3 β appear to coincide with or precede morphological differentiation of the hepatic endoderm.

At 9.0-9.5 days (15-25 somites) of development, the expression domains of HNF3 β established earlier basically persisted. Expression was seen throughout the notochord and the ventral part of the neural tube, as well as the entire gut (data not shown). Strong expression of HNF3 β was also observed by 9.0-9.5 days in the liver primordium, which begins to be morphologically distinguishable at this stage (Figs 3K, 4C).

Later in development, at 16.5 days, HNF3 β expression continued to mark the developing endoderm-derived tissues, such as lung, liver, pancreas and gut, with expression in the lung restricted to the developing respiratory epithelium; no expression was observed in mesoderm-derived organs such as the heart and kidney (Fig. 3N). Thus, early expression of HNF3 β in endoderm tissue persists throughout development and remains expressed in these tissues in the adult (Lai et al., 1991). However, even by 16.5 days, expression was not entirely endoderm-specific, because expression was still observed in the posterior hypothalamus (data not shown) and cartilage of the vertebrae (Fig. 3N).

Comparison of HNF3 β and HNF3 α

We examined the expression of the related gene, HNF3 α , to see if early expression in the endoderm lineage was a common feature of HNF3 genes. Localized expression of HNF3 α was not detectable by whole-mount in situ hybridization in the early primitive streak stage. Expression was first detected in the endoderm, by in situ hybridization to tissue sections at the late-streak stage of development (Fig. 3C). In contrast to HNF3 β expression, no staining was detected in axial mesoderm and neural plate at the late-streak and headfold stages respectively (Fig. 3C,F). From the 4-10 somite stage onwards, expression of HNF3 α in the ventral floorplate, notochord, and gut appeared very similar to HNF3 β , but weaker (Fig. 4B). Some differences in

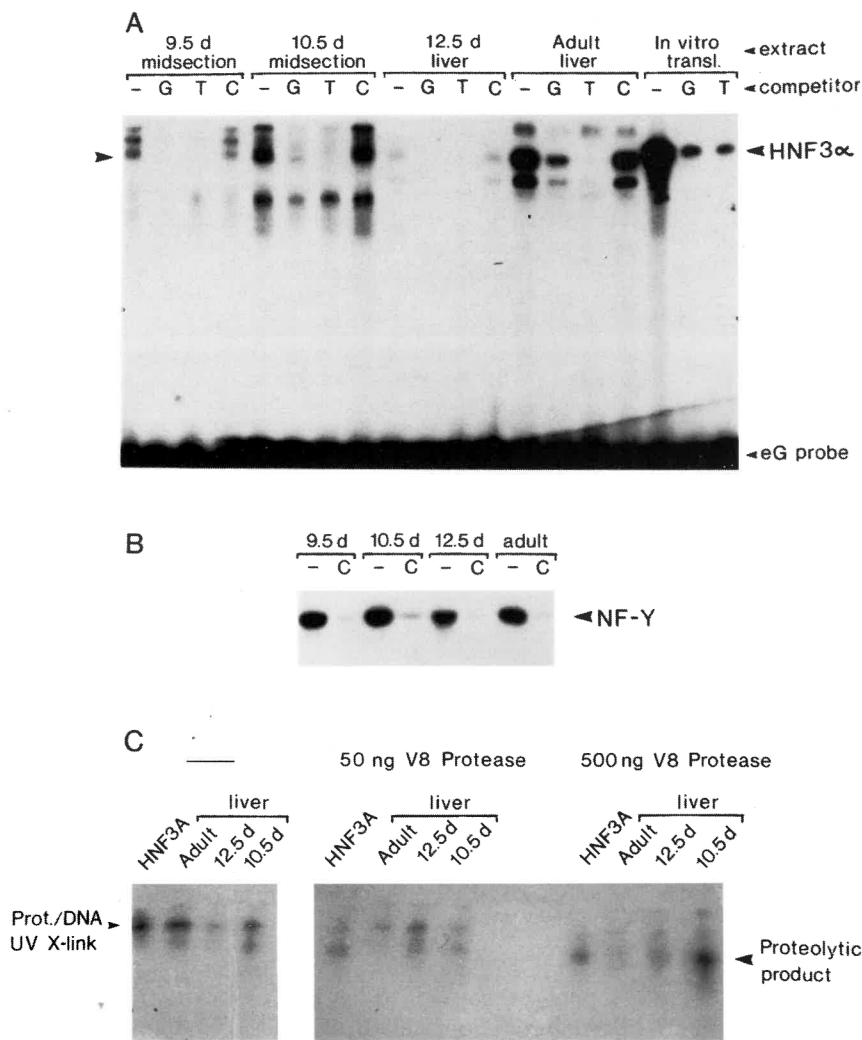


Fig. 5. Microscale DNA binding assays of HNF3 proteins in mouse embryos. (A,B) Electromobility shift analysis with the 'eG' HNF3 binding site (A) from the serum albumin enhancer (Liu et al., 1991) and a control NF-Y binding site (B) from the albumin promoter (Raymondjean et al., 1989). Specific binding competitors for HNF3 were 100-fold molar excesses of the eG site ('G') and the strong HNF3 binding site from the mouse transthyretin promoter ('T'; Costa et al., 1989); nonspecific competitor in A and specific competitor in B was the NF-Y site ('C'). Nuclear extracts for the assay were from the sources shown. Free probe is not shown in B. Arrowhead on the left indicates the position of specific HNF3 DNA binding activity present in early mouse embryo midsections. (C) SDS-polyacrylamide gel analysis of UV-crosslinked electromobility shift products from adult liver, 12.5 day embryo livers, and 10.5 day embryo midsections, that comigrate with HNF3. Certain samples were digested with V8 protease prior to electrophoresis, as indicated. Arrowheads indicate the positions of migration of the initial crosslinked products (left) and the terminal proteolytic digestion products (right).

expression patterns, however, were also observed. At the 10 somite stage, expression of HNF3 α was only observed in midbrain, but not in forebrain or hindbrain (Fig. 4D,E). Ventral floorplate expression of HNF3 α was also not detected at this stage in the spinal cord (Figs 3I, 4D). Expression in the gut at 8.5 days and 9.5 days extended further anteriorly for HNF3 β than HNF3 α (Fig. 3K,L).

The presumptive liver primordium between 8.5 and 9.5 days showed similarly strong expression of HNF3 α and HNF3 β (Figs 3K,L, 4C,F). Expression of HNF3 α in the ventral neural tube spread caudally to hindbrain and spinal cord regions by 9.0-9.5 days, but expression was restricted to the floorplate region (Fig. 3L). By 16.5 days, expression patterns of both genes were similar, except for the lack of expression of HNF3 α and HNF3 β in the vertebrae and the lining of the bladder, respectively (Fig. 3N and data not shown). Thus, although the patterns of expression were similar for the two genes, HNF3 β was generally expressed earlier and more broadly than HNF3 α . While the expression pattern in endoderm lineages appears to be very similar for both genes from the late-streak stage to the 16.5 day embryo, the earlier expression of HNF3 β in the primitive streak suggests that it could play an earlier role in specifying the

endoderm lineage. The difference in expression patterns detected by the HNF3 β and HNF3 α probes also confirms their hybridization specificity. In addition, a shorter HNF3 β probe, which does not include the HFH DNA binding domain, gave similar results in whole-mount in situ hybridization analysis of 7.5 and 8.5 day embryos (data not shown).

Presence of HNF3 proteins at the onset of endoderm differentiation

High levels of mRNA for liver-enriched transcription factors are sometimes found in various non-liver tissues where the corresponding proteins are undetectable (e.g., Birkenmeier et al., 1989); these precedents demonstrate that conclusions about active transcription factors cannot be drawn from mRNA studies alone. Furthermore, it is important to establish that HNF3 proteins are not only present in the developing endoderm lineage, but also competent to bind target DNA response elements.

To address this issue, nuclear proteins from midsections of mouse embryos at 9.5 and 10.5 days gestation, and from nascent livers at 12.5 days, were subjected to an electromobility shift assay with the eG binding site for HNF3 from the

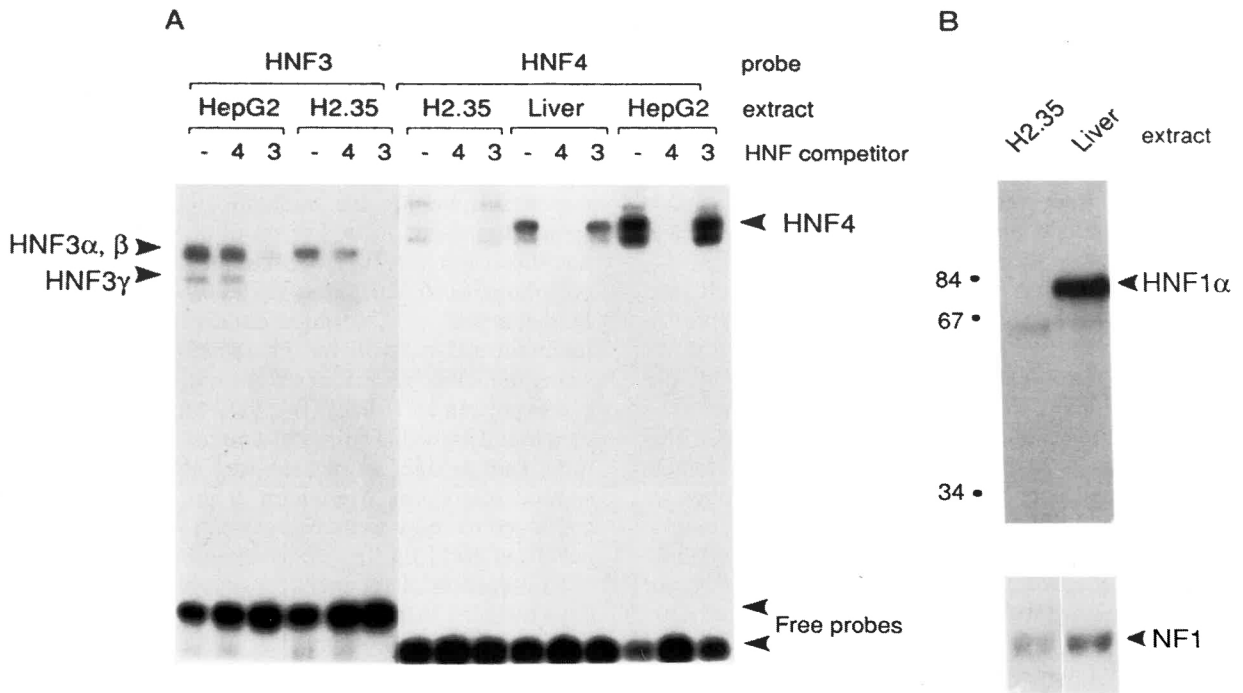


Fig. 6. Analysis of liver-enriched transcription factors in a de-differentiated hepatic cell line. (A) Electromobility shift assays of nuclear extracts from H2.35 cells, HepG2 cells, and mouse livers, using the albumin enhancer eG site as a probe ('HNF3'), or the HNF4 binding site from the apolipoprotein CIII gene. Specific protein-DNA complexes are indicated by the arrowheads at the side of the autoradiograph. Certain binding reactions contained a 100-fold molar excess of nonradioactive HNF4 ('4') or HNF3 ('3') sites as competitors, to demonstrate binding specificity. (B) Western blot analysis of H2.35 cell and liver nuclear extracts, probed with an antibody specific to HNF1 α (upper panel) or NF1 (lower panel). The molecular size ($M_r \times 10^{-3}$) of migration standards is shown on the left (upper panel).

mouse serum albumin enhancer (Liu et al., 1991). As controls for the mobility of eG-HNF3 complexes, we used an adult liver nuclear extract and *in vitro* translated HNF3 α . As expected, the liver extract produced an abundant protein-DNA complex that migrated slightly faster than the synthetic HNF3 α -DNA complex, and a less abundant, faster migrating complex (Fig. 5A); the upper complex contains both HNF3 α and HNF3 β , while the lower complex contains HNF3 γ (Lai et al., 1991). Extracts from the 9.5 and 10.5 day embryo midsections produced several protein-DNA complexes; one of these comigrated with the HNF3 α /HNF3 β complexes from liver (left arrow, Fig. 5A), with an additional faster-migrating complex in the 10.5 day midsections. All embryonic complexes were specifically competed by two different HNF3 sites, but not by a site for the ubiquitous factor NF-Y (Fig. 5A). The integrity of the embryonic extracts was demonstrated by the presence of NF-Y binding activity similar to that observed in adult liver (Fig. 5B). We conclude that the 9.5 day embryo midsection contain proteins with the DNA binding properties of HNF3 α and HNF3 β ; no evidence was obtained for the presence of HNF3 γ . The assay also shows that 9.5-10.5 day embryos contain other proteins with a similar DNA binding specificity.

By 12.5 days of development, the DNA binding complexes present in the newly formed liver contained HNF3 γ in addition to HNF3 α and β , and were identical in pattern and binding specificity to those seen with adult liver extracts, albeit lower in abundance (Fig. 5A). The lower abundance could be due to about half of the cells of the 12.5

day liver being hematopoietic (Paul et al., 1969), reducing the concentration of hepatocyte transcription factors. Control binding reactions with the NF-Y probe also showed several-fold less binding activity (Fig. 5B).

To demonstrate conclusively that HNF3 proteins were present in the embryonic extracts, we analyzed the migration and protease sensitivity of UV-crosslinked protein-DNA complexes. Specific binding complexes were cross-linked, excised from the gel, denatured *in situ*, and subjected to electrophoresis in a second dimension in an SDS-polyacrylamide gel. As seen in Fig. 5B, the most prominent band from each nuclear extract comigrates with *in vitro* translated HNF3 α . In addition, when the excised bands were incubated with increasing concentrations of V8 protease, the main terminal digestion products were identical in migration to that of *in vitro* translated HNF3 (Fig. 5B). These data show that HNF3 proteins are present and competent to bind their target sequences during endoderm differentiation, and that the adult pattern of HNF3 DNA binding activities is established by 12.5 days of development.

HNF3 is the only known liver-enriched transcription factor expressed in a cell line capable of hepatic differentiation

Detailed studies of differentiated and de-differentiated somatic cell lines have provided insight into transcription factors important for the endoderm lineage. For example, de-differentiated cell lines derived from liver tumors typically lack the expression of HNF1 α (Cereghini et al., 1988; Baumhueter et al., 1988), and a recent study has

shown this could be due to a lack of HNF4 (Kuo et al., 1992). By contrast, such studies have shown that members of the HNF3 family often continue to be expressed by the de-differentiated cells (Kuo et al., 1992; Nitsch et al., 1993). Possibly, HNF3 proteins confer the ability of rare, re-differentiated variants to be recovered from the cell population during selection or screening (Cereghini et al., 1988; Kuo et al., 1992). We sought to investigate this issue in greater detail.

Previously, we demonstrated that the H2.35 cell line, cloned from SV40-infected mouse hepatocytes, expresses extremely low levels of serum albumin mRNA under standard culture conditions, but when cultured on a collagen gel substratum and in a serum-free medium, albumin mRNA is induced over a hundred fold (Zaret et al., 1988). This system provides the opportunity to examine factors expressed in a cell population capable of homogenous re-differentiation within the endodermal lineage.

In earlier studies we showed that de-differentiated H2.35 cells lack C/EBP and related DNA binding activities, but that they contain low levels of HNF3 that are increased upon re-differentiation (Liu et al., 1991). Importantly, the low amounts of HNF3 in de-differentiated H2.35 cells are sufficient to activate transcription of a transfected reporter gene bearing HNF3 binding sites (DiPersio et al., 1991). We therefore wanted to address whether or not HNF4 and HNF1 α are present in the de-differentiated state. Electromobility shift assays readily detected HNF3 in de-differentiated H2.35 cell nuclear extracts, but no HNF4 DNA binding activity was found, using a high affinity HNF4 binding site from the apolipoprotein CIII gene (Fig. 6A). Northern blot analysis of H2.35 mRNA also failed to detect HNF4 mRNA (data not shown). As expected, abundant HNF4 DNA binding activity was observed in nuclear extracts from mouse liver and a partially differentiated hepatoma cell line, HepG2 (Fig. 6A). No HNF1 α protein was detected in de-differentiated H2.35 cells in a western blot assay, while abundant levels of the protein were readily detected in liver extracts (Fig. 6B). Subsequent detection of NF1, a ubiquitous transcription factor (Santoro et al., 1988), demonstrated the integrity of the H2.35 cell extracts. We conclude that under de-differentiating culture conditions, H2.35 cells lack HNF1, HNF4, and C/EBP, and that HNF3 is the only known liver transcription factor family expressed and active in the cells. These findings suggest that all liver transcription factors other than HNF3 can be lost from cells in culture without affecting the ability of the cells to re-differentiate along the hepatic endodermal lineage.

Distinct expression profile exhibited by a novel HFH family member

Expression of HFH-E5.1 was first detectable at the mid- to late-streak stage of development, in a specific domain in the posterior-distal region of the embryo (Fig. 7A). Optical sectioning through whole-mount stained embryos and in situ hybridization to tissue sections showed that expression at this stage was found in both ectoderm and mesoderm, but not in the endoderm (data not shown). By the early headfold stage, there was expansion of this expression to define a new anterior boundary. Immediately posterior to this boundary, a band of strong staining in mesoderm and overlying neural

plate was observed (Fig. 7B,C). From this stage onwards, expression diminished in the mesoderm, but increased in the developing neural tube so that by the early somite stage (1-2 somites), expression was strong in neuroectoderm in the anterior half of the embryo (Fig. 8B). The only staining in mesoderm anterior to the node at this stage was found in the region surrounding the midbrain, which probably corresponds to the initial band of strong staining anteriorly (data not shown). Posteriorly, weaker staining was observed in presomitic mesoderm and in the midline ectoderm (Fig. 8B). In the neural tube, stronger expression was observed in midbrain and parts of the hindbrain. By the 6-10 somite stage, the entire neural tube from midbrain to the posterior end expressed HFH-E5.1 (Fig. 7D). Mesoderm staining was then found in medial stripes of somites adjacent to the neural tube, and stronger expression was also found in the presomitic mesoderm. Mesoderm at the midbrain level also continued to show weak expression (data not shown). This pattern of HFH-E5.1 expression continued into 9.5 days.

Examination of sectioned in situ at 13.5 days and 15.5 days revealed that HFH-E5.1 continued to be expressed in the central nervous system (CNS). Specific groups of cells in the hypothalamus, midbrain, and medulla oblongata expressed HFH-E5.1 (Fig. 8F,H). In addition, staining was also observed in scattered cells of the ventral spinal cord (data not shown); no expression was detected outside the CNS at these later stages.

DISCUSSION

Knowledge of the forkhead family of regulatory proteins is rapidly expanding and reveals a group of genes that appear important in several developmental processes. The requirement of the *Drosophila forkhead* gene for proper gut development (Jürgens and Weigel, 1988) prompted us to ask whether the expression of *forkhead* homologs in the mouse could fulfill criteria expected for regulatory factors important for the formation of mammalian gut endoderm. While this work was in progress, Sasaki and Hogan (1993) independently reported the expression of HNF3 β and HNF3 α in mouse embryos between days 6.5 and 9.5 of gestation. Our focus on the expression of these genes in the endoderm lineage has revealed differences with respect to the early expression of HNF3 β reported previously (Sasaki and Hogan, 1993). We show that HNF3 β is expressed at the earliest stage of gastrulation in a group of cells that includes the progenitors of the definitive endoderm. Slightly later, at the late-streak stage, HNF3 α begins to be expressed in the endoderm, together with HNF3 β . Our studies of later stages showed that both HNF3 β and HNF3 α are expressed similarly throughout the subsequent differentiation of gut endoderm, during organogenesis, and that they appear to maintain the endoderm lineage after development.

To determine whether all progenitors of the definitive endoderm express HNF3 β , gene expression patterns must be compared with fate mapping studies. Fate mapping of the developing endoderm in mice has shown that there are already definitive endoderm cells on the surface of the anterior primitive streak by the mid-streak stage of development (Lawson and Pederson, 1987), although much of the definitive endoderm seems to arise later from the head

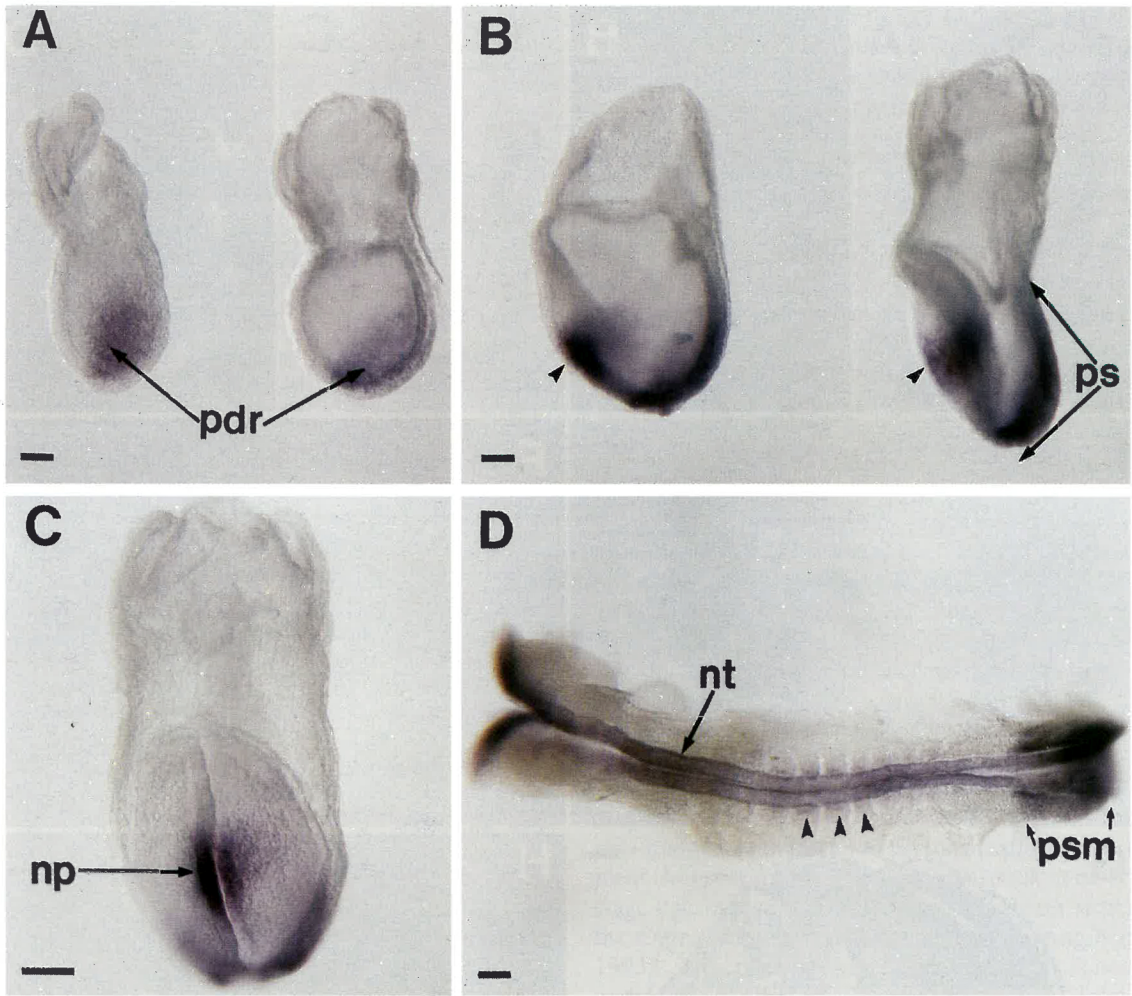


Fig. 7. Expression of HFH-E5.1 transcripts during gastrulation and neurulation as detected by whole-mount in situ hybridization. (A) Mid- to late streak embryos (7.0-day), showing expression in the distal portion of the embryo posteriorly. (B) Early and late headfold embryos. Expression of HFH-E5.1 extends anteriorly, resulting in a new band of strong staining in anterior mesoderm (arrowheads). Posterior to the node, strong expression was also observed in ectoderm and mesoderm. (C) Ventral view of late headfold embryo showing that the neural plate above the strong band of anterior mesoderm expression also expresses HFH-E5.1. (D) Somite stage embryo (9-10 somites). The entire neural tube and the presomitic mesoderm express HFH-E5.1. Weaker expression of HFH-E5.1 was also observed in medial stripes of the anterior somites (arrowheads). Abbreviations: pdr, posterior distal region; psm, presomitic mesoderm; nt, neural tube; np, neural plate and ps, primitive streak. Scale bar represents 100 nm.

process (Lawson et al., 1991). HNF3 β was clearly expressed in the group of cells at the anterior end of the developing primitive streak; this region of the epiblast is fated to give rise to the definitive endoderm (Lawson et al., 1991). Thus, HNF3 β expression marks at least a subset of the earliest endodermal precursors, suggesting that this gene plays a fundamental role in establishing the lineage. Given that HNF3 proteins represent the vertebrate HFH products with the greatest homology to the *Drosophila forkhead* gene (Weigel et al., 1989), there appears to be evolutionary conservation of mechanisms for specifying the developing gut.

As the head process forms, both HNF3 β and HNF3 α are strongly expressed in the midline endoderm cells, which are known to subsequently form part of the definitive gut (Poelmann, 1981), as well as in cells more anterior in the midline, which may represent the progeny of the early definitive endoderm precursors. Cells expressing lower levels of

HNF3 β were also seen lateral to the midline; the exact relationship between these cells and the later definitive endoderm is not yet clear. From this stage on, the expression of the two genes seems to be virtually identical in endoderm derivatives. Both genes are clearly expressed in the developing foregut and hindgut, although with different anterior boundaries, and expression is elevated in the early liver primordium. Later in development, expression persists in the liver and gut and then appears in the developing lung, a pattern consistent with the expression of these genes in the adult (Lai et al., 1991). These results suggest that both genes could play a role in the development and maintenance of the endoderm lineage.

The elevated level of expression of HNF3 α and HNF3 β mRNA in liver primordia is interesting in light of the known mechanisms of organogenesis. Tissue explant studies in the chick have shown that cardiac mesenchyme induces

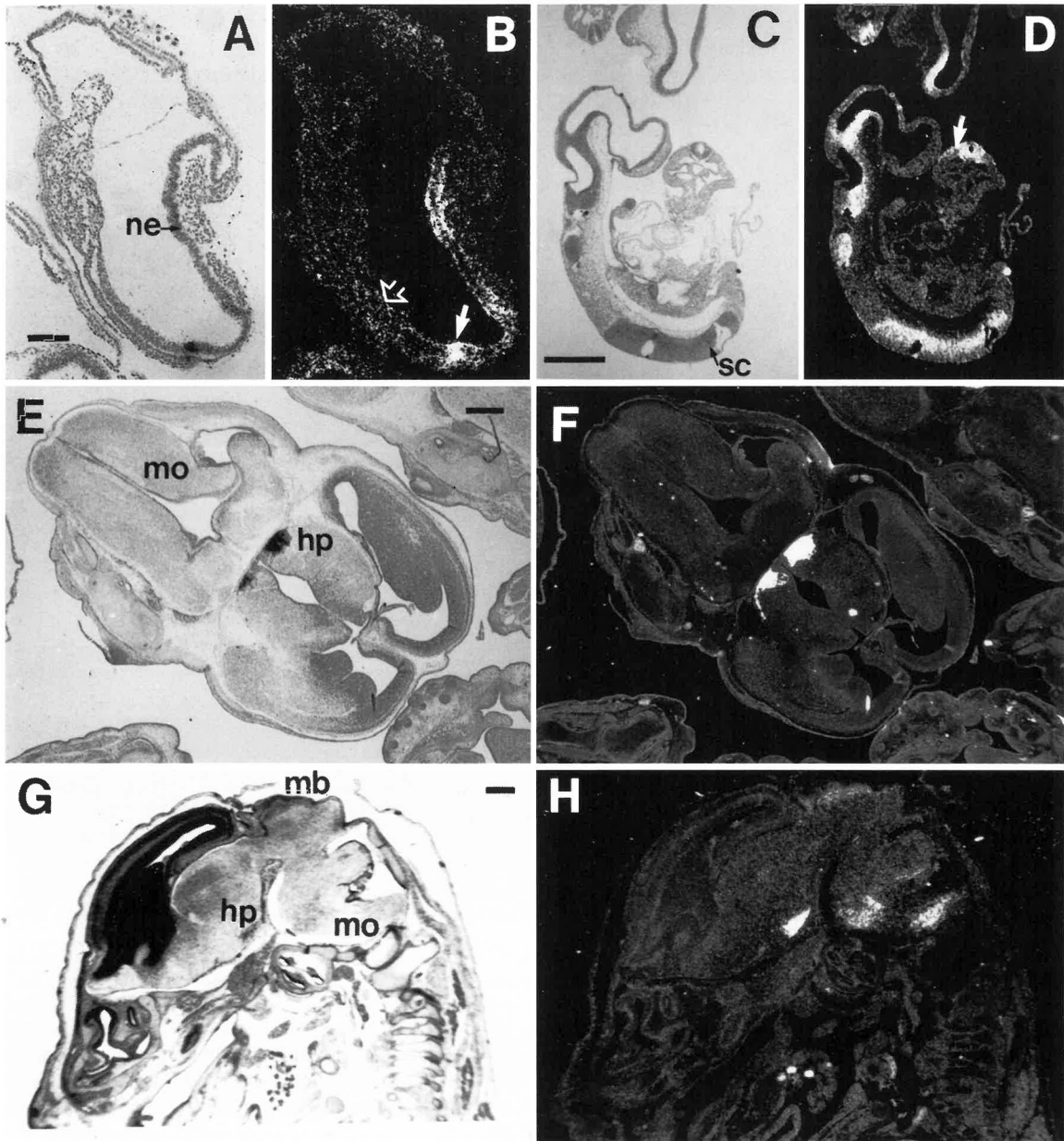


Fig. 8. HFHE5.1 expression during organogenesis analyzed by sectioned in situ hybridization. (A,B) 8.5 day (1-2 somites) embryo. Neural tube showing HFHE5.1 expression. Posterior to the node (solid arrow), staining was found in ectoderm and mesoderm (open arrow). (C,D) 9.5 day embryo showing continued expression in entire CNS and also in presomitic mesoderm. (E,F) Transverse section of 13.5 day embryo showing specific expression in a group of cells in the hypothalamus (hp). (G,H) Sagittal section of 16.5 day embryo showing neural expression in hypothalamus, midbrain (mb) and medulla oblongata (mo). Scale bar, 100 μ m (A); 5 mm (C,E,G).

adjacent endodermal cells to assume a hepatic fate (Le Douarin, 1975); limited studies with mouse embryo tissues are consistent with this model (Houssaint, 1980; Cascio and Zaret, 1991). We therefore suggest that one of the first responses of endoderm to hepatic induction is an increase in the level of expression of HNF3 genes. Indeed, this response seems to parallel or precede morphological differentiation of the endoderm. Our microscale assays showed that the region containing mouse liver primordium at 9.5 days gestation contains HNF3 proteins competent to bind DNA. This indicates that HNF3 is capable of functioning very early in the definitive endoderm lineage. We further showed that

HNF3 γ DNA binding activity was not detectable at 9.5 days, but that it was detectable in the early liver at day 12.5; thus, HNF3 γ may act later in development than HNF3 β or HNF3 α .

A role in early endoderm specification is supported by the finding that only HNF3 α continues to be expressed and active when a hepatocyte-derived cell line is maintained in the de-differentiated state; all other known liver-specific transcription factors, which are readily detectable in mature hepatocytes, are undetectable in the de-differentiated H2.35 cells. HNF3 may serve, therefore, to impart an endoderm identity to cells both in vivo and in vitro, perhaps by maintaining liver genes in a state that is permissive for induction

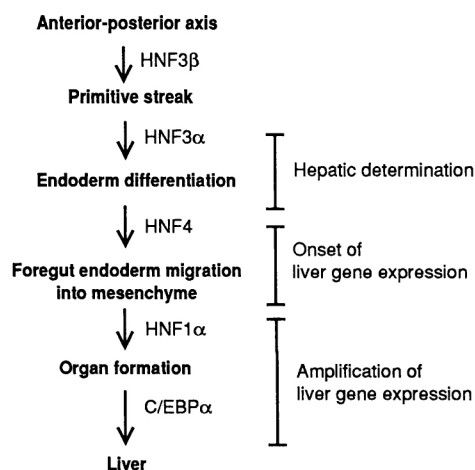


Fig. 9. Possible hierarchy of transcription factors acting in endoderm and liver differentiation. Developmental signals and consequences are indicated in bold face, with arrows indicating the progression in time. The time of appearance of liver-enriched transcription factors is shown alongside the arrows, and discernable stages of liver development are indicated at the side, with brackets denoting the appropriate events.

during differentiation. The ability of HNF3 β to bind its own promoter (Pani et al., 1992) fits nicely with this model; once the HNF3 β gene is activated, it would be sufficient to maintain its own synthesis as well as potentiate the expression of genes that are downstream in the regulatory cascade (Fig. 9).

HNF3 α and HNF3 β are both expressed in tissues other than the developing gut, making it possible that they play roles in other developmental processes. HNF3 β , in particular, shows strong expression throughout the developing node at the anterior end of the streak, and in the developing notochord and floorplate, which are structures known to derive in part from the node. The node in mammals is thought to be equivalent to Hensen's node in chick, or the Spemann organizer in amphibians, both of which can induce axial structures when grafted to ectopic sites. In *Xenopus*, there are two HFH genes whose expression domains include the organizer; *pintallavis* (Ruiz i Altaba and Jessell, 1992) and *Xfkd* (Dirksen and Jamrich, 1992), the latter of which is highly similar to HNF3 β . The *axial* gene in zebrafish is also highly similar to HNF3 β , and is required for the normal development of structures derived from the fish 'organizer' (Strähle et al., 1993). Thus, there may be a conserved role in the organizer for this subgroup of HFH genes. Other classes of transcription factors are also expressed in the organizer (Blum et al., 1992; Taira et al., 1992; von Dassow et al., 1993), suggesting that it possesses considerable information content.

HNF3 α and HNF3 β are also expressed in the developing ventral neural tube, from the midbrain to the posterior end. Expression in the midbrain clearly extends beyond the limits of the floorplate into neural tissue. Expression of mammalian HFH genes in neural tissue has been reported previously: rat BF-1 expression is restricted to the telencephalon in late stage embryos and remains specific to the brain in the adult (Tao and Lai, 1992). However, the

expression of BF-1 prior to day 11.5 has not been reported. The finding of early neural domains of expression of the HNF3 genes opens the possibility that there may be sequential activation, and possibly interaction, between HFH genes in the development of the nervous system.

During our search for HFH genes potentially involved in endoderm development, we isolated a novel partial cDNA, HFH-E5.1, whose sequence helps define a different subgroup of HFH genes. Expression of HFH-E5.1 is first detectable at the late primitive streak stage and is restricted to the ectoderm and mesoderm, with a sharp anterior boundary and a broad posterior region of expression. As development proceeds to the head fold stages, expression in the anterior mesoderm is matched by overlying expression in the neural plate. At the early somite stage, expression is found throughout the neural tube, being strongest in the future mid-brain. By days 13.5 and 16.5, the gene is expressed in specific groups of cells in the central nervous system. HFH-E5.1 shares a high degree of sequence similarity with the *Drosophila* HFH gene FD4 (Häcker et al., 1992), which is expressed in embryogenesis in neuroblasts along the longitudinal axis, and subsequently in sensory neurons in the head. Considering that both FD4 and HFH-E5.1 have such highly conserved structures and neural patterns of expression, this subgroup of HFH proteins, like the HNF3/forkhead subgroup, may have a developmental function that has been conserved through evolution.

The developmental pattern of expression of HFH-E5.1 also implicates the gene in regionalization along the neural plate. Anterior mesoderm at the late streak to early head fold stage can induce markers of mid-hindbrain structures, like the *Engrailed* genes, in naive ectoderm (Ang and Rossant, 1993), but posterior mesoderm cannot. The localized expression of HFH-E5.1 in the anterior mesoderm, followed by its neural expression, suggests that this gene may help restrict the inducing capacity of mesoderm. The posterior mesoderm expression of HFH-E5.1 persists and eventually becomes restricted to the presomitic mesoderm, suggesting that the gene may also play a role in somite formation, along with other genes expressed strongly in this area, such as FGFR1 (Yamaguchi et al., 1992) and mouse *Notch* (Reaume et al., 1992).

We have shown that HFH3 α and HNF3 β show a strong association with the development of the definitive endoderm, but are also expressed in restricted areas in other germ layer derivatives, and we have identified HFH-E5.1, which shows no expression in the endoderm, but has interesting domains of expression in the developing mesoderm and neurectoderm. Clearly the HFH gene family in mammals is large and probably not yet fully defined, and the varied and highly specific expression patterns suggests that these genes will play roles in a number of developmental processes, including early lineage specification as well as later cell type differentiation. Mutational analysis of these genes should elucidate their individual roles and possible interactions between them.

We thank Brigid Hogan for the mouse embryo cDNA library, Rob Costa, Eseng Lai, and James Darnell for the rat HNF3 α cDNA, Gerald Crabtree for the HNF1 α antibody, François Guillemot for help with in situ hybridization on sections, Kendra-

pasad Harpal for providing tissue sections, Kirstie Lawson for discussion on the endoderm lineage, David Jackson for advice on protein-DNA crosslinking, and John Burch for advice on cloning by PCR. The research was supported by grants from the MRC and NCI of Canada to J. R. and from the NIH (GM36477) to K. S. Z. J. R. is a Terry Fox Cancer Research Scientist of the NCIC and an International Research Scholar of the HHMI.

REFERENCES

- Ang, S.-L. and Rossant, J. (1993). Anterior mesoderm induces mouse *Engrailed* genes in explant cultures. *Development* **118**, 139-149.
- Baumheuter, S., Courtois, G. and Crabtree, G. R. (1988). A variant nuclear protein in dedifferentiated hepatoma cells binds to the same functional sequences in the β fibrinogen gene promoter as HNF-1. *EMBO J.* **7**, 2485-2493.
- Birkenmeier, E. H., Gwynn, B., Howard, S., Jerry, J., Gordon, J. I., Landschulz, W. H., and McKnight, S. L. (1989). Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein. *Genes Dev.* **3**, 1146-1156.
- Blum, M., Gaunt, S. J., Cho, K. W. Y., Steinbeisser, H., Blumberg, B., Bittner, D., and De Robertis, E. M. (1992). Gastrulation in the mouse: The role of the homeobox gene *gooseoid*. *Cell* **69**, 1097-1106.
- Blumenfeld M., Maury M., Chouard T., Yaniv M. and Condamine H. (1991). Hepatic nuclear factor 1 (HNF1) shows a wider distribution than products of its known target genes in developing mouse. *Development* **113**, 589-599.
- Cascio S. and Zaret K. S. (1991). Hepatocyte differentiation initiates during endodermal-mesenchymal interactions prior to liver formation. *Development* **113**, 217-225.
- Cereghini, S., Blumenfeld, M. and Yaniv, M. (1988). A liver-specific factor essential for albumin transcription differs between differentiated and dedifferentiated rat hepatoma cells. *Genes Dev.* **2**, 957-974.
- Cereghini, S., Ott, M.-O., Power, S. and Maury, M. (1992). Expression patterns of vHNF1 and HNF1 homeoproteins in early postimplantation embryos suggest distinct and sequential developmental roles. *Development* **116**, 783-797.
- Chodosh, L. A., Carthew, R. W. and Sharp, P. A. (1987). A single polypeptide possesses the binding and transcription activities of the major late transcription factor of adenovirus. *Mol. Cell. Biol.* **6**, 4723-4733.
- Clevidence D. E., Overdier D. G., Tao W., Qian X., Pani L., Lai E. and Costa, R. H. (1993). Identification of nine novel tissue-specific transcription factors of the HNF-3/forkhead DNA binding domain family. *Proc. Natl. Acad. Sci. USA* **90**, 3948-3952.
- Conlon, R. A. and Rossant, J. (1992). Exogenous retinoic acid rapidly induces anterior ectopic expression of murine *Hox-2* genes in vivo. *Development* **116**, 357-358.
- Costa R. H., Grayson D. R. and Darnell, Jr., J. E. (1989). Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α 1-antitrypsin genes. *Mol. Cell Biol.* **9**, 1415-1425.
- Costa, R. H., Van Dyke, T. A., Yan, C., Kuo, R. and Darnell, Jr., J. E. (1990). Similarities in transthyretin gene expression and differences in transcription factors: Liver and yolk sac compared to choroid plexus. *Proc. Natl. Acad. Sci. USA* **87**, 6589-6593.
- Courtois G., Morgan J. G., Campbell L. A., Fourel G. and Crabtree G. R. (1987). Interaction of a liver-specific nuclear factor with the fibrinogen and α 1-antitrypsin promoters. *Science* **238**, 688-692.
- DeSimone V., DeMagistris L., Lazzaro D., Gerstner J., Monaci P., Nicosia A. and Cortese R. (1991). LFB3, a heterodimer-forming homeoprotein of the LFB1 family, is expressed in specialized epithelia. *EMBO J.* **10**, 1435-1443.
- DiPersio C. M., Jackson D. A. and Zaret K. S. (1991). The extracellular matrix coordinately modulates liver transcription factors and hepatocyte morphology. *Mol. Cell Biol.* **11**, 4405-4414.
- 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.
- Fahrner, K., Hogan, B. L. M. and Flavell, R. A. (1987). Transcription of *H-2* and *Qa* genes in embryonic and adult mice. *EMBO J.* **6**, 1265-1271.
- Gardner, R. L. and Rossant, J. (1979). Investigation of the fate of 4, 5 day post-coitum mouse inner cell mass cells by blastocyst injection. *J. Embryol. Exp. Morph.* **52**, 141-152.
- Gooderham, K. (1984). In situ peptide mapping of proteins following polyacrylamide gel electrophoresis. In *Molecular Biology* (ed. John M. Walker), pp. 193-202. New Jersey: Humana Press.
- Grossinklaus, U., Pearson, R. K. and Gehring, W. J. (1992). The *Drosophila sloppy paired* locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes Dev.* **6**, 1030-1051.
- Häcker, U., Grossinklaus, U., Gehring, W. J. and Jackle, H. (1992). Developmentally regulated *Drosophila* gene family encoding the forkhead domain. *Proc. Natl. Acad. Sci. USA* **89**, 8754-8758.
- Houssaint, E. (1980). Differentiation of the mouse hepatic primordium. I. An analysis of tissue interactions in hepatocyte differentiation. *Cell Differentiation* **9**, 269-279.
- Jackson, D. A., Rowader, K. E., Stevens, K., Jiang, C., Milos, P. and Zaret, K. S. (1993). Modulation of liver-specific transcription by interactions between hepatocyte nuclear factor 3 and nuclear factor 1 binding DNA in close apposition. *Mol. Cell Biol.* **13**, 2401-2410.
- Jürgens, G. and Weigel, D. (1988). Terminal versus segmental development in the *Drosophila* embryo: the role of the homeotic gene fork head. *Roux's Arch. Dev. Biol.* **197**, 345-354.
- Knöchel, S., Lef, J., Clement, J., Klocke, B., Fill, S., Köster, M., and Knöchel, W. (1992). Activin A induced expression of fork head related genes in posterior chordamesoderm (notochord) of *Xenopus laevis*. *Mech. Dev.* **38**, 157-165.
- Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292.
- Kuo, C. F., Xanthopoulos, K. G. and Darnell, Jr., J. E. (1990). Fetal and adult localization of C/EBP: evidence for combinatorial action of transcription factors in cell-specific gene expression. *Development* **109**, 473-481.
- Kuo, C. J., Conley, P. B., Chen, L., Sladek, F. M., Darnell, Jr., J. E. and Crabtree, G. R. (1992). A transcriptional hierarchy involved in mammalian cell-type specification. *Nature* **355**, 457-461.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lai E., Prezioso V. R., Smith E., Litvin O., Costa R. H. and Darnell, Jr., 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, Jr., J. E. (1991). Hepatocyte nuclear factor 3 α belongs to a gene family in mammals that is homologous to the *Drosophila* homeotic gene *fork head*. *Genes Dev.* **5**, 416-427.
- Lawson, K. A., Meneses, J. J. and Pedersen, R. A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113**, 891-911.
- Lawson, K. A. and Pedersen, R. A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development* **101**, 627-652.
- Le Douarin, N. M. (1975). An experimental analysis of liver development. *Medical Biology* **53**, 427-455.
- Li, C., Lusic, A. J., Sparkes, R. Tran, S.-M., and Gaynor, R. (1992). Characterization and chromosomal mapping of the gene encoding the cellular DNA binding protein HTLF. *Genomics* **13**, 659-664.
- Liu, J.-K., DiPersio C. M. and Zaret K. S. (1991). Extracellular signals that regulate liver transcription factors during hepatic differentiation in vitro. *Mol. Cell Biol.* **11**, 773-784.
- 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.
- Morle, F., Lopez, B., Henni, T., and Godet, J. (1985). α -Thalassaemia associated with the deletion of two nucleotides at position -2 and -3 preceding the AUG codon. *EMBO J.* **4**, 1245-1250.
- Mueller, C. R., Maire, P. and Schibler, U. (1990). DBP, a liver-enriched transcriptional activator, is expressed late in ontogeny and its tissue specificity is determined posttranscriptionally. *Cell* **61**, 279-291.
- Nitch, D., Boshart, M. and Schütz, G. (1993). Extinction of tyrosine aminotransferase gene activity in somatic cell hybrids involves modification and loss of several essential transcriptional activators. *Genes Dev.* **7**, 308-319.
- Pani L., Qian X., Clevidence D. and Costa R. H. (1992). The restricted

- promoter activity of the liver transcription factor hepatocyte nuclear factor 3b involves a cell-specific factor and positive autoactivation. *Mol. Cell Biol.* **12**, 552-562.
- Paul, J., Conkie, D. and Freshney, R. I.** (1969). Erythropoietic cell population changes during the hepatic phase of erythropoiesis in the fetal mouse. *Cell Tissue Kinet.* **2**, 283-294.
- Poelmann, R. E.** (1981). The head-process and the formation of the definitive endoderm in the mouse embryo. *Anat. Rec.* **162**, 41-49.
- Raymondjean, M., Cereghini, S. and Yaniv, M.** (1988). Several distinct 'CCAAT' box binding proteins coexist in eukaryotic cells. *Proc. Natl. Acad. Sci. USA* **85**, 757-761.
- Reaume, A. G., Conlon, R. A., Zirngibl, R., Yamaguchi, T. P. and Rossant, J.** (1992). Expression analysis of a *Notch* homologue in the mouse embryo. *Dev. Biol.* **154**, 377-387.
- Rosenquist, G. C.** (1971). The location of the pregut endoderm in the chick embryo at the primitive streak stage as determined by radioautographic mapping. *Dev. Biol.* **26**, 323-335.
- Rosenquist, G. C.** (1972). Endoderm movements in the chick embryo between the early short streak and head process stages. *J. Exp. Zool.* **180**, 95-104.
- Ruiz i Altaba, A. and Jessell, T. M.** (1992). *Pintallavis*, a gene expressed in the organizer and midline cells of frog embryos: involvement in the development of the neural axis. *Development* **116**, 81-93.
- Sanders, E. E., Bellairs, R. and Portch, P. A.** (1978). In vivo and in vitro studies on the hypoblast and definitive endoblast of avian embryos. *J. Embryol. Exp. Morph.* **46**, 187-205.
- Santoro, C., Mermod, N., Andrews, P. C. and Tjian, R.** (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. *Nature* **334**, 218-223.
- Sasaki, H., and Hógan, B. L. M.** (1993). Differential expression of multiple fork head related genes during gastrulation and pattern formation in the mouse embryo. *Development* **118**, 47-59.
- Schoenwolf, G. C., Garcia-Martinez, V. and Dias, M. S.** (1992). Mesoderm movement and fate during avian gastrulation and neurulation. *Dev. Dynam.* **193**, 235-248.
- Schreiber, E., Matthias, P., Müller, M. M. and Schaffner, W.** (1989). Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nuc. Acids Res.* **17**, 6419.
- Schwartz, R. M. and Dayhoff, M. O.** (1979). *Atlas of Protein Sequence and Structure*, (ed. M. O. Dayhoff). pp. 353-358. Washington, D. C.: National Biomedical Research Foundation.
- Selleck, M. A. J. and Stern, C. D.** (1991). Fate mapping and cell lineage analysis of Hensen's node in the chick embryo. *Development* **112**, 615-626.
- Sladek F. M., Zhong W., Lai E. and Darnell Jr., J. E.** (1990). Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev.* **4**, 2353-2364.
- Strähle U., Blader P., Henrique D. and Ingham P. W.** (1993). *Axial*, a zebrafish gene expressed along the developing body axis, shows altered expression in *cyclops* mutant embryos. *Genes Dev.* **7**, 1436-1446.
- Taira, M., Jamrich, M., Good, P. J. and David, I. B.** (1992). The LIM domain-containing homeobox gene *Xlim-1* is expressed specifically in the organizer region of *Xenopus* gastrula embryos. *Genes Dev.* **6**, 356-366.
- 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.
- von Dassow, G., Schmidt, J. E. and Kimehanan, D.** (1993). Induction of the *Xenopus* organizer: expression and regulation of *Xnot*, a novel FGF and activity-regulator homeobox gene. *Genes Dev.* **7**, 355-367.
- Weigel D. and Jäckle H.** (1990). The *fork head* domain: a novel DNA binding motif of eukaryotic transcription factors. *Cell* **63**, 455-456.
- 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.
- Yamaguchi, T. P., Conlon, R. A. and Rossant, J.** (1992). Expression of the fibroblast growth factor receptor FGFR-1/flg during gastrulation and segmentation in the mouse embryo. *Dev. Biol.* **152**, 75-88.
- Zaret, K. S.** (1993). Control of hepatocyte differentiation by liver-enriched transcription factors. In *Hepatic Transport and Bile Secretion: Physiology and Pathophysiology* (ed. N. Tavoloni and P. D. Berk), pp. 135-143, New York: Raven Press, Ltd.
- Zaret K. S., DiPersio C. M., Jackson D. A., Montigny W. J. and Weinstat D. L.** (1988). Conditional enhancement of liver-specific gene transcription. *Proc. Natl. Acad. Sci. USA* **85**, 9076-9080.

(Accepted 20 September 1993)