

# Cloning of the Human Interferon-Related Developmental Regulator (IFRD1) Gene Coding for the PC4 Protein, a Member of a Novel Family of Developmentally Regulated Genes

Pasquale Buanne,<sup>\*,†</sup> Barbara Incerti,<sup>†</sup> Daniele Guardavaccaro,<sup>\*</sup> Virginia Avvantaggiato,<sup>‡</sup> Antonio Simeone,<sup>‡</sup> and Felice Tirone<sup>\*,1</sup>

<sup>\*</sup>Istituto di Neurobiologia CNR, Viale Carlo Marx 43, 00137 Rome; <sup>†</sup>Telethon Institute of Genetics and Medicine, Via Olgettina 58, 20132 Milan; and <sup>‡</sup>Istituto Internazionale di Genetica e Biofisica, Via G. Marconi 10, 80100 Naples, Italy

Received December 1, 1997; accepted February 10, 1998

The rat PC4 gene had been initially isolated as a nerve growth factor-inducible sequence in PC12 cells. Although its function remains unknown, recently it has been shown that PC4 is necessary to muscle differentiation and that it might have a role in signal transduction. We report the isolation of the human homolog of the rat PC4 gene, renamed here IFRD1 (interferon-related developmental regulator 1). Several human IFRD1 clones were identified by searching the EST database using the rat IFRD1 (PC4) cDNA as a query. An EST clone containing the entire ORF was chosen for sequencing. Human IFRD1 presented a predicted protein product of 453 amino acids, highly conserved (90.2% identity) compared to the rat IFRD1 (PC4) protein sequences. The mapping assignment of human IFRD1 to chromosome 7q22–q31 was retrieved from the UniGene database maintained at NCBI. A comparison of human IFRD1 (PC4) protein to databases revealed 47% identity to the protein encoded by the human gene SKMc15, originally isolated from a chromosome 3-specific library. Therefore, SKMc15 is a gene related to IFRD1, being the second member of a novel family. We analyzed their expression during murine development, and we found that mouse IFRD1 appears more expressed in specific differentiating structures at midgestation, while mouse SKMc15 is highly expressed soon after gastrulation and in the hepatic primordium, suggesting an involvement in early hematopoiesis. © 1998 Academic Press

## INTRODUCTION

With the aim of identifying genes involved in the process of neuronal differentiation, we have isolated in

recent years several immediate-early genes activated in the rat cell line PC12 by nerve growth factor (NGF; Tirone and Shooter, 1989; Bradbury *et al.*, 1991). This cell line, which is derived from a pheochromocytoma, differentiates into sympathetic neurons in the presence of NGF, thus representing a widely used *in vitro* model of the differentiative action of NGF on chromaffin cells *in vivo* (Greene and Tischler, 1976). We found that one of the genes isolated, PC4, presented a significant similarity with interferon- $\gamma$ , a molecule known for its role in cellular differentiation (Tirone and Shooter, 1989). Such similarity was between the carboxy-terminal half of the PC4 cDNA-deduced protein and the whole interferon- $\gamma$  protein, suggesting the existence of a functional domain within the PC4 protein. Furthermore, the mouse PC4 homolog (TIS7) was identified as a tetradecanoyl phorbol acetate-induced gene in mouse NIH3T3 cells (Varnum *et al.*, 1989). Although PC4 function remains unknown, its expression was found to be regulated not only during neuronal differentiation *in vitro* and *in vivo* (Guardavaccaro *et al.*, 1994; Iacopetti *et al.*, 1996), but also during myoblast differentiation. PC4 is in fact expressed in the myoblast and in the differentiated myotube, with a transient decrease after the onset of differentiation. This aspect appears to have an important functional significance in muscle differentiation, as directly demonstrated by studies performed in the myoblast cell line C2C12. In fact, stable inhibition of PC4 expression, using antisense PC4 cDNA transfection or microinjection of anti-PC4 polyclonal antibodies, leads to impairment of myogenin and myosin gene expression as well as of morphological differentiation (Guardavaccaro *et al.*, 1995). Microinjection data in myoblast pointed to the necessity of PC4 presence when the stimulus for differentiation is delivered, suggesting the involvement of PC4 in signal transduction (Guardavaccaro *et al.*, 1995). This latter possibility was also implied by the observation that a fraction of the PC4 protein, which appears to be cytosolic (Guardavaccaro *et al.*,

Sequence data from this article have been deposited with the EMBL Data Library under Accession No. Y10313.

<sup>1</sup> To whom correspondence should be addressed at the Institute of Neurobiology, National Research Council, Viale Marx 43, 00137 Rome, Italy. Telephone: 39-6-86090290. Fax: 39-6-86090370. E-mail: tirone@biocell.irmkant.rm.cnr.it.

1994, 1995), translocates transiently to the inner side of the plasma membrane at the beginning of neuronal differentiation by NGF in the PC12 cell line (Guardavaccaro *et al.*, 1994). As a whole, this evidence suggests that PC4 might play a modulatory role in development, during the transition of a determined cell lineage (e.g., muscle or neuron) to a committed phenotype.

Here we report the cloning, sequencing, and expression analysis in adult tissues of the human homolog of PC4. In compliance with the HGMW nomenclature rules, and with the properties observed so far for the protein product PC4, the gene has been renamed IFRD1 (interferon-related developmental regulator 1). In this study we analyzed in parallel, starting from early stages of development, the expression of IFRD1 and of the related gene SKMc15 (Latif *et al.*, 1997), obtaining hypotheses on the functional role of these genes of the same family.

## MATERIALS AND METHODS

**Sequence analysis and computer-assisted search of databases.** Automated fluorescent DNA sequencing was performed using Perkin-Elmer 377 Prism machines with both Dye Terminator and Dye Primer Cycle Sequencing chemistries on double-strand plasmid templates. Computer analysis of the sequences was performed with the Wisconsin Package version 8.1-UNIX (August 1995). Similarity searches were performed using BlastN, BlastP, and FastA algorithms against the GenBank (release 101, June 1997), EMBL (release 50.0, June 1997), Pir-Protein (release 52.0, March 1997), and Swiss-Prot (release 34.0, November 1996) databases.

**Northern analysis.** Similar amounts of mRNA from different human tissues, blotted on a nylon filter, were hybridized with human IFRD1 cDNA <sup>32</sup>P labeled with the hexamer primers procedure (Feinberg and Vogelstein, 1983). Hybridization was performed at 42°C for 18 h in 5× SSPE (3 M NaCl, 0.2 M NaH<sub>2</sub>PO<sub>4</sub>, 0.02 M EDTA), 2× Denhardt's solution (1% Ficoll 400, 1% polyvinylpyrrolidone, 1% BSA), 100 µg/ml sheared salmon sperm DNA, 2% SDS, 50% formamide. The mRNA amounts were controlled by hybridizing the filters to a β-actin probe.

**Mice.** C57/B16 mice were mated between 9 and 10 PM Day 0.5 postcoitum was assumed to begin at the middle of the day of vaginal plugging. Pregnant female mice were killed by cervical dislocation, and embryos were staged according to Theiler (1989), collected in ice-cold PBS under a dissection microscope (SV11; Carl Zeiss, Inc., Thornwood, NY), and fixed in 4% paraformaldehyde overnight.

**In situ hybridization.** *In situ* hybridization was carried out as described (Wilkinson and Green, 1990; Wilkinson, 1992), with minor modifications. Both paraffin-embedded and cryostat sections were analyzed. Dissected embryos were prefixed in 0.1 M sodium phosphate buffer (pH 7.3), 4% paraformaldehyde at 4°C overnight and embedded in Tissue-Tek (Miles Laboratories, Inc., Elkhart, IN). Cryosections (8 µm thick) were transferred onto gelatin/chromium (III) potassium sulfate-subbed and dried at room temperature. Before hybridization, slides were postfixed. To obtain probes, mouse SKMc15 and mouse IFRD1 (the latter previously isolated by Varnum *et al.*, 1989, named TIS7) were isolated respectively from mouse embryo (E13.5–14.5) and adult mouse cDNA libraries. The isolates, identified by searching the EST database using TIS7 and human SKMc15 as a query, corresponded to mouse EST W29669, i.e., TIS7, and to mouse EST W65790, i.e., mouse SKMc15. Their identity was checked and confirmed by complete sequencing and by Southern analysis (not shown). The probes used were antisense and sense strands spanning the SKMc15 gene from nt 1 to 699 (referred to the sequence obtained) and the IFRD1 (TIS7) gene from nt 748 to 1496 (referred to the published sequence of TIS7; Varnum *et al.*, 1989).

Transcription reactions with T7 or T3 polymerase (Riboprobe Kit; Promega Biotechnologies, Madison, WI) were carried out in the presence of [<sup>35</sup>S]CTP (Amersham Corp., Arlington Heights, IL). The template was then degraded with RNase-free DNase (Pharmacia), and the labeled RNA was purified through a Sephadex G-50 column and progressively degraded by random alkaline hydrolysis to improve access to RNA *in situ*. The probes were dissolved at a working concentration of 1×10<sup>5</sup> cpm/µl in the hybridization mix; 30 ml of the appropriate probe was added to each slide. Hybridization was carried out overnight at 55°C. The slides were then washed under stringent conditions (65°C, 2× SSC, 50% formamide) and treated with RNase. Autoradiography was performed with Kodak NT/B2 emulsion. Exposure times were between 15 and 25 days.

Negative controls using sense strands were tested on sections adjacent to those hybridized with the specific antisense strands to determine the basal background. These probes never gave a detectable signal.

## RESULTS

### Cloning and Sequencing of Human IFRD1

The rat IFRD1 (PC4) cDNA sequence was compared to EST databases (dbEST; Boguski *et al.*, 1993) using the BlastN algorithm (Altschul *et al.*, 1990). We detected 11 ESTs having significant homology to the query sequence, from Soares' infant brain, fetal liver-spleen, and placenta cDNA libraries as well as from other human muscle and testis cDNA libraries (October 1996). Sequence analysis was then carried out on cDNA clone 52803 (isolated from an infant brain cDNA library), corresponding to a human EST (H29134) homologous to rat IFRD1 with a *P* value in the BlastN output of 3.5×10<sup>-88</sup>. Clone 52803 has a cDNA insert of about 1800 nt, containing the entire coding region. The first in-frame ATG is located 219 nt from the 5' end of the clone and fulfills Kozak's criteria for an initiation codon (Kozak, 1984). Comparison of the complete sequence of clone 52803 with dbEST indicated the presence of several ESTs belonging to the same transcriptional unit and further extending the UTR 3' end of about 450 nt. These ESTs were assembled to obtain the complete sequence of human IFRD1 cDNA, which has a total length of 2252 nt, with a 5' UTR of 219 nt and a 3' UTR of 676 nt (Fig. 1). The predicted protein product encoded by the ORF is 453 residues long (Fig. 1) with a calculated molecular weight of 50,685. A search for active sites using the program Prosite (Bairoch, 1990) showed three potential phosphorylation sites by cAMP-dependent protein kinase (Thr-9, Thr-127, Thr-373), several potential phosphorylation sites by casein kinase II, four by protein kinase C (Ser-95, Thr-353, Thr-373, Thr 416), and one by tyrosine kinase (Tyr-364). Also, potential N-myristoylation sites were detected (residues 13–25 and 105–110), which, however, according to previous experiments performed with rat IFRD1 (PC4) cDNA, might be not functional (Guardavaccaro *et al.*, 1994).

### IFRD1 and SKMc15, Two Members of a Novel Gene Family

Human and rat IFRD1 (PC4) sequences are highly homologous, with a significant identity at both the cDNA and the amino acid levels (86.3 and 90.2% iden-

1	CCTCGTGCCAGAGAAACATGTATCGTTCGATCACAGCTCTTCACGGGGAATTCTGCTGCCGCCACCGCCACTCTTACCCCGCGCTTCTCGACTCTGT	
103	TGTTAGCCGAAGACTCGCCTCTCAGCCGCCCGCGCACAGACGCACGAGTAAAAAGTGCAGCTCCATCGGCTGATCCTCGCTAAGCTCCGACTCTGGGCGGC	
1	M P K N K K R N T P H R G S S A G G G G S G A A A A T A A	29
205	ACCGGGCGTCCACGATGCCGAAGAACAAGAAGCGGAACACTCCCCACCGCGGTAGCAGTGCTGGCGGCGCGGGTCAGGAGCAGCCGACGACGCGCGCGC	
30	T A G G Q H R N V Q P F S D E D A S I E T M S H C S G Y S D P S S F	63
307	ACAGCAGGTGGCCAGCATCGAAATGTTACGCTTTTAGTGATGAAGATGCATCAATTGAAACAATGAGCCATTGCAGTGGTTATAGCGATCCTTCCAGTTTT	
64	A E D G P E V L D E E G T Q E D L E Y K R K G L I D L T L D K S A K	97
409	GCTGAAGATGGACCAGAAGTCTTGATGAGGAAGAACTCAAGAAGACCTAGAGTACAAGAGAAAGGATTAATTGAACCTAACCTGGATAAGAGTGCAGAG	
98	T R Q A A L E G I K N A L A S K M L Y E F I L E R R M T L T D S I E	131
511	ACAAGGCAGCAGCTCTTGAAGTATTAAAAATGCACCTGGCTTCAAAATGCTGTATGAATTTATTCTGGAAGGAGAATGACTTTAACTGATAGCATTGAA	
132	R C L K K G K S D E Q R A A A A L A S V L C I Q L G P G I E S E E I	165
613	CGCTGCCTGAAAAAGGTAAGAGTGATGAGCAACGTGCAGCTGCAGCGTTAGCATCTGTTCTTTGTATTAGCTGGGCCCTGGAATTGAAAGTGAAGAGATT	
166	L K T L G P I L K K I I C D G S A S M Q A R Q T C A T C F G V C C F	199
715	TTGAAACTCTTGGACCAATCCTAAAGAAAATCATTTGTGATGGGTACAGTAGTATGCAGGCTAGGCAAACCTGTGCAACTTGCTTTGGTGTGTTGCTGTTTT	
200	I A T D D I T E L Y S T L E C L E N I F T K S Y L K E K D T T V I C	233
817	ATTGCCACAGATGACATTACTGAACATACTCAACTCTGGAATGTTTGAAAAATATCTTCACTAAATCCTATCTCAAAGAGAAAGACACTACTGTTATTTCG	
234	S T P N T V L H I S S L L A W T L L L T I C P I N E V K K K L E M H	267
919	AGCACTCCTAATACAGTGCTTCATATCAGTCTCTTCTTGATGGACACTACTGCTGACCATATGCCCAATCAATGAAGTGAAGAAAAAGCTTGAGATGCAT	
268	F H K L P S L L S C D D V N M R I A A G E S L A L L F E L A R G I E	301
1021	TTCCATAAGCTTCCAAGCCTCCTCTCTTGTGATGATGTAACATGAGAATAGCTGCTGGTGAATCTTTGGCACTTCTCTTTGAATTGGCCAGAGGAATAGAG	
302	S D F F Y E D M E S L T Q M L R A L A T D G N K H R A K V D K R K Q	335
1123	AGTGACTTTTTTTATGAAGACATGGAGTCTTGACGCAGATGCTTAGGGCCTTGGCAACAGATGGAAATAAACACCGGGCCAAAGTGGACAAGAGAAAGCAG	
336	R S V F R D V L R A V E E R D F P T E T I K F G P E R M Y I D C W V	369
1225	CGGTGAGTTTTCAGAGATGCTCAGAGGCGAGTGGAGGAACGGGATTTTCCAACAGAAACCATTAATTTGGTCTGAAAGCATGTATATTGATGCTGGGTA	
370	K K H T Y D T F K E V L G S G M Q Y P L A V K M E F L E N V F E T W	403
1327	AAAAACACACCTATGACACCTTTAAGGAGGTTCTTGGATCAGGGATGCAGTACCACCTTGCACTCAAAATGGAATTCCTCGAAATGTATTGAAACTTGG	
404	T P S D A L M L Q R L K T M K I S R F E R H L Y N S A A F K A R T K	437
1429	ACCCCACTGATGCCTTGATGCTGCAACGCCTTAAACAGATGAAGATTTCCTGTTTCGAAAGGCGATTATATTAACCTGCAGCTTCAAAGCTCGAACCAAA	
438	A R S K C R D K R A D V G E F F	453
1531	GCTAGAAGCAATGTCGAGATAAGAGAGCAGATGTTGGAGAATTCTTCTAGATTTTTCAGAACTTGAAGACTATTTTCTAATTTCTATTTTTTTTCTATTTC	
1633	AATGTATTTAACTCTAGACACAGTTCCTTATCTTGGATTAACTTAGATAACTTTTGTAGCCAGTGGTTATATGCTTATAATTTAATGTACAATACTATTGA	
1735	AACTGGTGAGTTCTGATTATTAATATTCCTCTGTAATCAGTAAACATGTATAAAGTATTTGTAATGTTTGGTCATAATTTATTTATGAAGACAGCAAAAGA	
1837	CTGATTTCATGATGGGAAAACAATTAGCCAAAGTTTAATTTCTTACACTGTGGTTGTCAAGAATACTGATTTACTATAATGATATACATGCAAGATATT	
1939	TAACCTAATATCTTAGACAGAAGTCTTGGGTACAATTTGGGATCTAGTTCCTCGGAAAAGCTGCTGTATTTTAAATTTTAAATGGAATGTAGCTTTTAAA	
2041	ATCCTGTCACTGGCATCAACAAAGGAATTATACCATGAGACCTTATAGCTGTACTTAAAGCCATTCACTGATTCAGCTATGGGAGTTTCATGATGAATTAGCA	
2143	TATGCCAGAAAGGTTGCTAACCTTAACATCTGAGAGCAGTAACACTGATTTTATCTGCTGTATAAGACTTTGTGCAATTTACTTTGAAATAAGATTTTTT	
2245	CCCACTG	

**FIG. 1.** Nucleotide and predicted amino acid sequence of human IFRD1 (PC4). The ATG start codon and the polyadenylation signal are double underlined. The translation of the longest ORF is shown above the nucleotide sequence with the one-letter amino acid code. The nucleotide sequence has been submitted to the EMBL database under Accession No. Y10313.

tity, respectively; Fig. 2). The mouse IFRD1 sequence, TIS7 (Varnum *et al.*, 1989), is also shown to be highly conserved in relation to the human, with 87.1 and 90.6% identity between the nucleotide and the protein sequences, respectively (Fig. 2). The carboxy-terminal region of the human IFRD1 (PC4) protein shows significant similarity to the human interferon- $\gamma$  (47.5% similarity, alignment not shown), as previously seen for the rat IFRD1 (PC4) protein (49% similarity; Tirone and Shooter, 1989). Interestingly, a comparison of the full human IFRD1 sequence with the GenBank database also detected a human cDNA called SKMc15 (Accession No. U09585) whose sequence, although different from human IFRD1, presented significant homology. This cDNA had been isolated from a human

skeletal muscle cDNA library using as a probe an evolutionarily conserved sequence (LUCA14) from a chromosome 3-specific library (Latif *et al.*, 1997; Wei *et al.*, 1996). Percentage identities between nucleotide and protein sequences of human IFRD1 (PC4) and SKMc15 are respectively 48.1 and 47.6% (Fig. 2). Therefore, IFRD1 and SKMc15 appear to be different members of a family of genes. An interesting feature emerging by comparison between rat, mouse, and human IFRD1 and human SKMc15 sequences, which thus might be a distinguishing character of the family, is the existence of three domains with higher conservation (in SKMc15, aa 60–152, 285–341, 406–440). This is particularly evident in the carboxyl-terminal domain (Fig. 2), which is also the region of homology to the interferon- $\gamma$ .

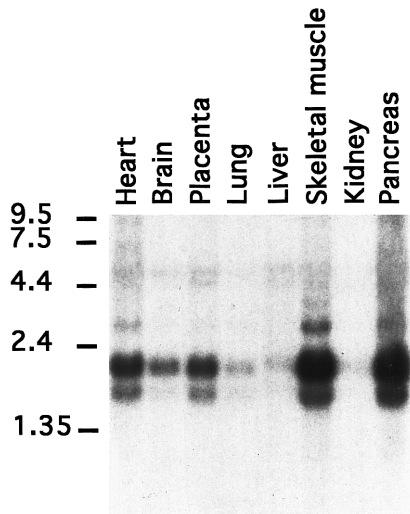
1	MPRARKSNTLRKGGQRRGGGARSSAQADSGSSDDEAASEARS-----T	43	SM15
1	MPKNKKRNTPHRGSSAGGGGSGAAAATAATAGGQHRNVQPFSDDEDASLET	50	hPC4
1	MPKNKKRNAPHRG..GGGGGSGAATSAATAGGPHRTVQPFSDDEDASLET	48	rPC4
1	MPKNKKRNAPHRG..GGGGGSGAATSAATAGGPHRTVQPFSDDEDASLET	48	TIS7
44	ASECPSLLSTTAEDSLGGDVDEQGGQEDLEKKEYVDCLTDKSAKTRQ	93	SM15
51	MSHCSGYSDPSSFAEDGPEVLDEEGTQEDLEYKRRKGLIDLTLDKSAKTRQ	100	hPC4
49	MSHCSGYSDPSSFAEDGPEVLDEEGTQEDLEYKRRKGLIDLTLDKSAKTRQ	98	rPC4
49	MSHCSGYSDPSSFAEDGPEVLDEEGTQEDLEYKRRKGLIDLTLDKSAKTRQ	98	TIS7
94	GALESRLRLALASRLLPDFILERRRLTLADALEKCLKKGKGEQALAAAVLG	143	SM15
101	AALEGKKNALASKMLYEFILERRMTLTDSIERCLKKGKSDEQRAAAALAS	150	hPC4
99	AALEGVKNALSSKVLVEFLERRMTLTDSIERCLKKGKSDEQRAAAALAS	148	rPC4
99	AALEGVKNALSSKVLVEFLERRMTLTDSIERCLKKGKSDEQRAAAALAS	148	TIS7
144	LLCVQLGPGPKGEELFHSLOPLLVSVLSDSTASPAARLHCAALGLGCVV	193	SM15
151	VLCIQLGPGIESEEILKTLGPILKKIICDGSASVQARQTCATCFGVCCFI	200	hPC4
149	VLCIQLGPGIESEEILKTLGPILKKIICDGSASVQARQTCATCFGVCCFI	198	rPC4
149	VLCIQLGPGIESEEILKTLGPILKKIICDGAASVQARQTCATCFGVCCFI	198	TIS7
194	AAADIQDLYSLACLESVFSRFYGLGGSSTSPVVPASLHGLLSAALQAWA	243	SM15
201	ATDDITELYSTLECLENIFTKSY-LKEKDTNVCSTPNTVLHISLLAWT	249	hPC4
199	ATDDITELYSTLECLEGIFTKSY-LKEKDTNVCSTPNTVLHISLLAWT	247	rPC4
199	ATDDITELYSTLECFENIFTKSY-LKEKDTNVCSTPNTVLHISLLAWT	247	TIS7
244	LLLTICPSTQISHILDRLPRLPOLLSSSEVMRIAAGETIALLFELARD	293	SM15
250	LLLTICPINEVKKKLEMFHFKLPSSLSCDDVMRIAAGESLALLFELARG	299	hPC4
248	LLLTICPISIEVKKKLEMFHFKLPSSLSCDDVMRIAAGESLALLFELARG	297	rPC4
248	LLLTICPINEVKKKLEMFHFKLPSSLSCDDVMRIAAGESLALLFELARG	297	TIS7
294	LEEEFYVEDMEALCSVLRILATDSNKYRAKADRRRQRSIFRAVLHSVEGG	343	SM15
300	IESDFFYEDMESLTQMLRALATDGNKHRAKVDKRKQRSVFRDVLRAVEER	349	hPC4
298	MESDFFYEDMDSLTQMLRALATDGNKHRAKVDKRKQRSVFRDVLRAVEER	347	rPC4
298	MESDFFYEDMDSLTQMLRALATDGNKHRAKVDKRKQRSVFRDVLRAVEER	347	TIS7
344	ECEEEIVRFGEFVLYMDSWARHRIYAAFKEVLGSGMHHLQNN-ELLRDI	392	SM15
350	DFPTETIKFGPERMYIDCWVKKHTYDTFKEVLGSGMQYPLAVKMEFLENV	399	hPC4
348	DFPTETVKFGPERMYIDSWVKKHTYDTFKEALGSGMQYHLQTN.EFLRNV	396	rPC4
348	DFPTETVKFGPERMYIDSWVKKHTYDTFKEVLGSGMQYHLQTN.EFLRNV	396	TIS7
393	FGL-GPVLILLDATAALKACKVERFEKHLYNAAAFKARTKARSVRDKRADI	441	SM15
400	FETWTPSDALMLQRLKTMKISRFERHLYNSAAAFKARTKARSKCRDKRADV	449	hPC4
397	FE.LGPPVMDAATLKTMKISRFERHLYNSAAAFKARTKARSKCRDKRADV	445	rPC4
397	FE.LGPPVMDAATLKTMKISRFERHLYNSAAAFKARTKARSKCRDKRADV	445	TIS7
442	L	442	SM15
450	GEFF	453	hPC4
446	GEFF	449	rPC4
446	GEFL	449	TIS7

**FIG. 2.** Comparison of human IFRD1 (hPC4) deduced amino acid sequence with rat IFRD1 (rPC4), mouse IFRD1 (TIS7), and human SKMc15 (SM15) proteins. The computer program Align was used to generate multiple alignments. Amino acids identical to human IFRD1 residues have been shaded in black.

### *Tissue Expression of Human IFRD1 in Adult Tissues*

The expression pattern of human IFRD1 was tested by hybridizing the <sup>32</sup>P-labeled cDNA clone 52803 to a Northern blot containing poly(A)<sup>+</sup> RNA samples from several human adult tissues (heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas). A

2.2-kb transcript, in agreement with the sequence length of IFRD1, was detected in several human tissues although at different levels of expression (Fig. 3). The human tissue expression of IFRD1 mRNA is very similar to that observed in rat (Tirone and Shooter, 1989), being very low in liver, lung, and kidney (see



**FIG. 3.** Expression of IFRD1 in adult human tissues. The blots contained 2  $\mu$ g of poly(A)<sup>+</sup> RNA/lane, obtained from eight different human tissues. The filters were hybridized with the human IFRD1 (PC4) clone. Control hybridizations with a  $\beta$ -actin cDNA probe indicated the presence of approximately equal amounts of RNA in all lanes (data not shown).

also Iacopetti *et al.*, 1996) and maximal in pancreas, skeletal muscle, and heart. Furthermore, the IFRD1 pattern of high expression in skeletal and cardiac muscle and in pancreas is common also to SKMc15 (Latif *et al.*, 1997). A lower and weaker band corresponding to a smaller transcript of about 1.8 kb was also detected (Fig. 3). Such signal did not correspond to SKMc15 mRNA, which has the same size of human IFRD1 mRNA (Latif *et al.*, 1997). Excluding the possibility of degradation, this might imply the presence of a transcript from another related gene.

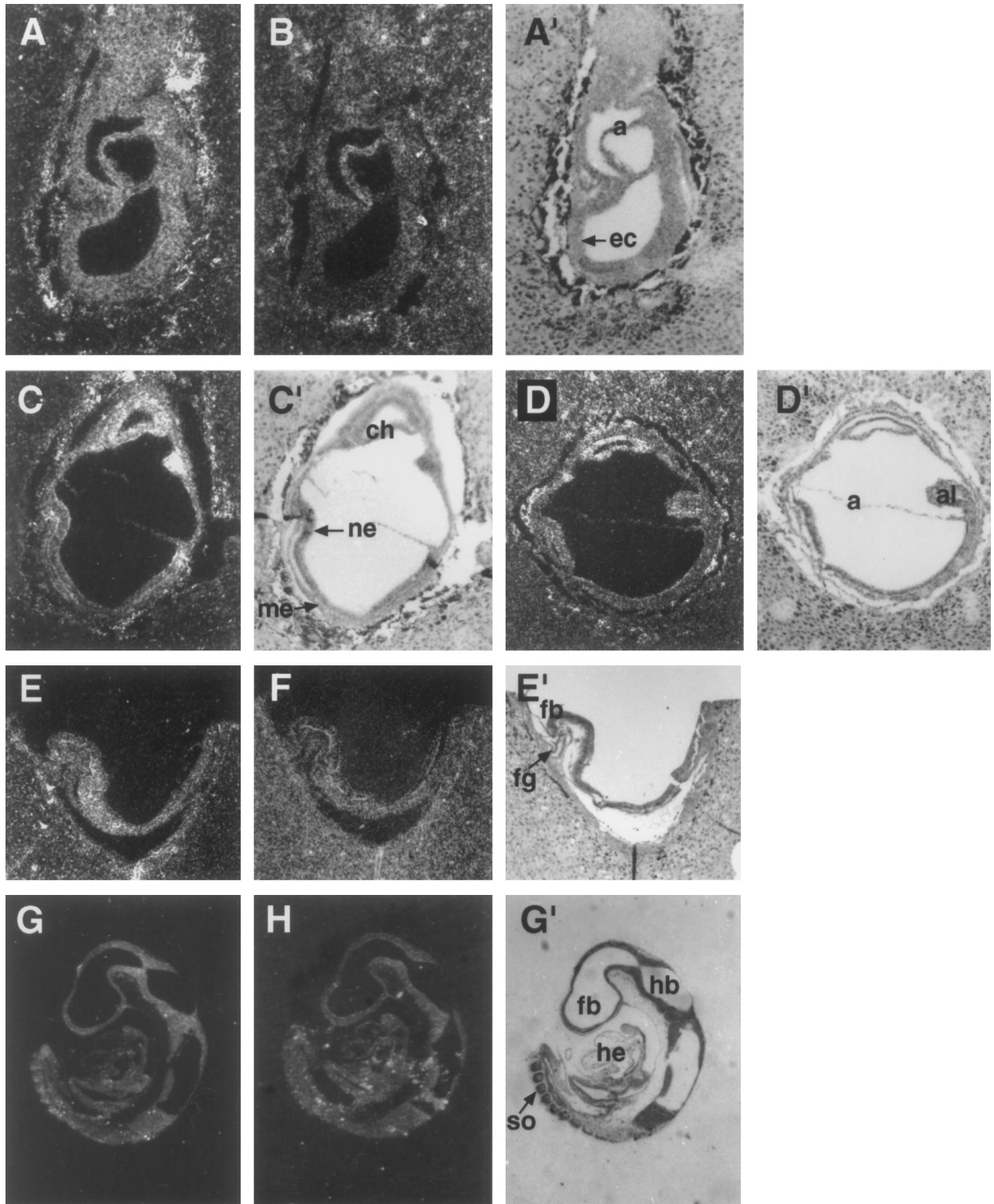
#### *Expression of IFRD1 (TIS7) and SKMc15 During Development*

To gain insight into a possible role for the SKMc15 and IFRD1 genes during embryonic development, we studied their expression in mouse from gestation day 6.7 (E 6.7) to E 17. Fragments 700 and 750 bp long were used as probes, respectively, for murine SKMc15 and murine IFRD1 (i.e., TIS7; see Materials and Methods). The sense strand of the same fragments did not reveal any detectable signal above the background level (data not shown). At E 6.7, SKMc15 was expressed in all the embryo, including extraembryonic and embryonic components (Fig. 4A), while at E 7.5 its embryonic expression was confined to the anterior and posterior third (Fig. 4C) of the embryos, with the highest signal appearing along the chorion close to the ectoplacental cavity (Fig. 4C). At E 8.2, SKMc15 was detected in the mesenchyme cells and in the surrounding neuroectoderm (Fig. 4E), while at E 9.5 a weak signal was detected along the midhindbrain region and in the somites (Fig. 4G).

At the same stages the distribution of IFRD1 (TIS7) transcripts was quite different from that of SKMc15. In

fact, at E 6.7 IFRD1 (TIS7) transcripts were restricted to the posterior amniotic fold (Fig. 4B), and at the headfold stage (E 7.7) they appeared in the chorion and at a lower level in the rostral neuroectoderm and in the allantois (Fig. 4D). From E 8.2 until E 9.5, IFRD1 (TIS7) transcripts were not seen above the background signal. Only at E 10.5 IFRD1 (TIS7) were transcripts again detectable in defined tissues and organs such as the differentiating spinal ganglia (arrows in Figs. 5D and 5E), the spinal cord, the basal neuroepithelium of the hindbrain and forebrain regions (Figs. 5D and 5E), and the hepatobiliary primordium (Figs. 5D–5E). At E 12.5 the expression pattern of IFRD1 (TIS7) was more complex, but restricted to a number of tissues and organs. These included the spinal ganglia, the developing kidney, the lung primordium, the olfactory and respiratory neuroepithelium, the whole CNS with a higher level in the developing telencephalon and diencephalon, and, finally, a restricted structure deriving from the mandibular arch (Fig. 5F).

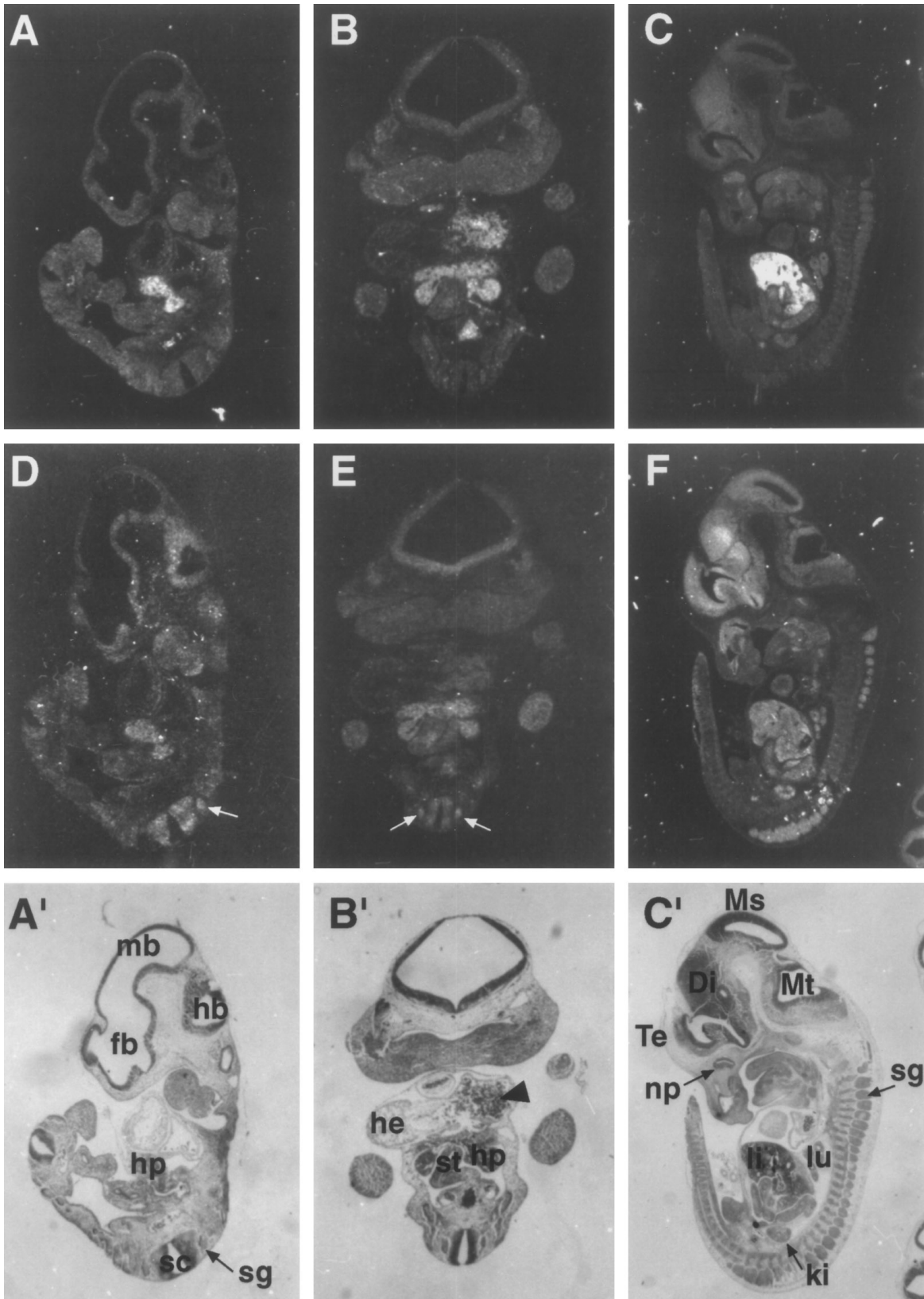
On the other hand, at E 10.5 SKMc15 was silent throughout the embryo, except for a strong signal restricted to the hepatic primordium (Figs. 5A and 5B) and to the early circulating hematopoietic cells actively generated in the differentiating liver (Fig. 5B and arrowhead in 5B'). From this stage onward, SKMc15 was detected at high levels in the hepatic primordium (Figs. 5C and 6A–6C). A more detailed observation at E 12.5 shows that, in addition to the liver, a lower signal of SKMc15 was present in the kidney and lung primordia as well as in the tongue and mandibular structures deriving from the first branchial arch (Fig. 5C). Therefore, during early and midgestation, the expression of SKMc15 and IFRD1 (TIS7) showed similar patterns and complementary levels. In fact, while SKMc15 was expressed at higher levels in early embryogenesis and then in the developing hepatic primordium, IFRD1 (TIS7) was higher in differentiating structures and body organs such as the developing kidney, lung, CNS, spinal ganglia, and nasal neuroepithelium, where SKMc15 was transcribed at lower level. Finally, SKMc15 and IFRD1 (TIS7) expression was studied at late gestation. At E 17, SKMc15 and IFRD1 (TIS7) were widely distributed throughout all the embryo, and their complementary expression previously observed at earlier stages was not evident. However, several organs still expressed high levels of both SKMc15 (Figs. 6A–6C) and IFRD1 (TIS7, Figs. 6E–6G), such as the liver, kidney, lung, spinal cord, telencephalon, and salivary glands. In addition, as shown at higher magnification, a relevant signal was detected in the back muscles for both genes (Figs. 6D and 6H). In sum, the expression patterns of SKMc15 and IFRD1 (TIS7) indicate that these genes are transcribed early, in gastrulating embryos mainly in extraembryonic tissues, then at midgestation in restricted structures and body organs, and at late gestation ubiquitously. This general feature, thus, suggests a complex regulation as well as different roles.



**FIG. 4.** SKMc15 and IFRD1 (TIS7) expression patterns in early murine embryogenesis. (A, C, E, and G) Sagittal section at E 6.7, 7.5, 8.2, and 9.5 hybridized with the SKMc15 probe. (B, D, F, and H) Sagittal sections at the same stages probed with the IFRD1 (TIS7) probe. The two genes were hybridized in adjacent sections of the same embryo except for the E 7.5 stage. (A' and C'–G') Bright fields of the corresponding sections. Abbreviations: a, amnios fold; ec, primitive ectoderm; ne, neuroectoderm; ch, chorion, me, mesendoderm; al, allantois; fg, foregut; so, somite; he, heart; fb, forebrain; hb, hindbrain.

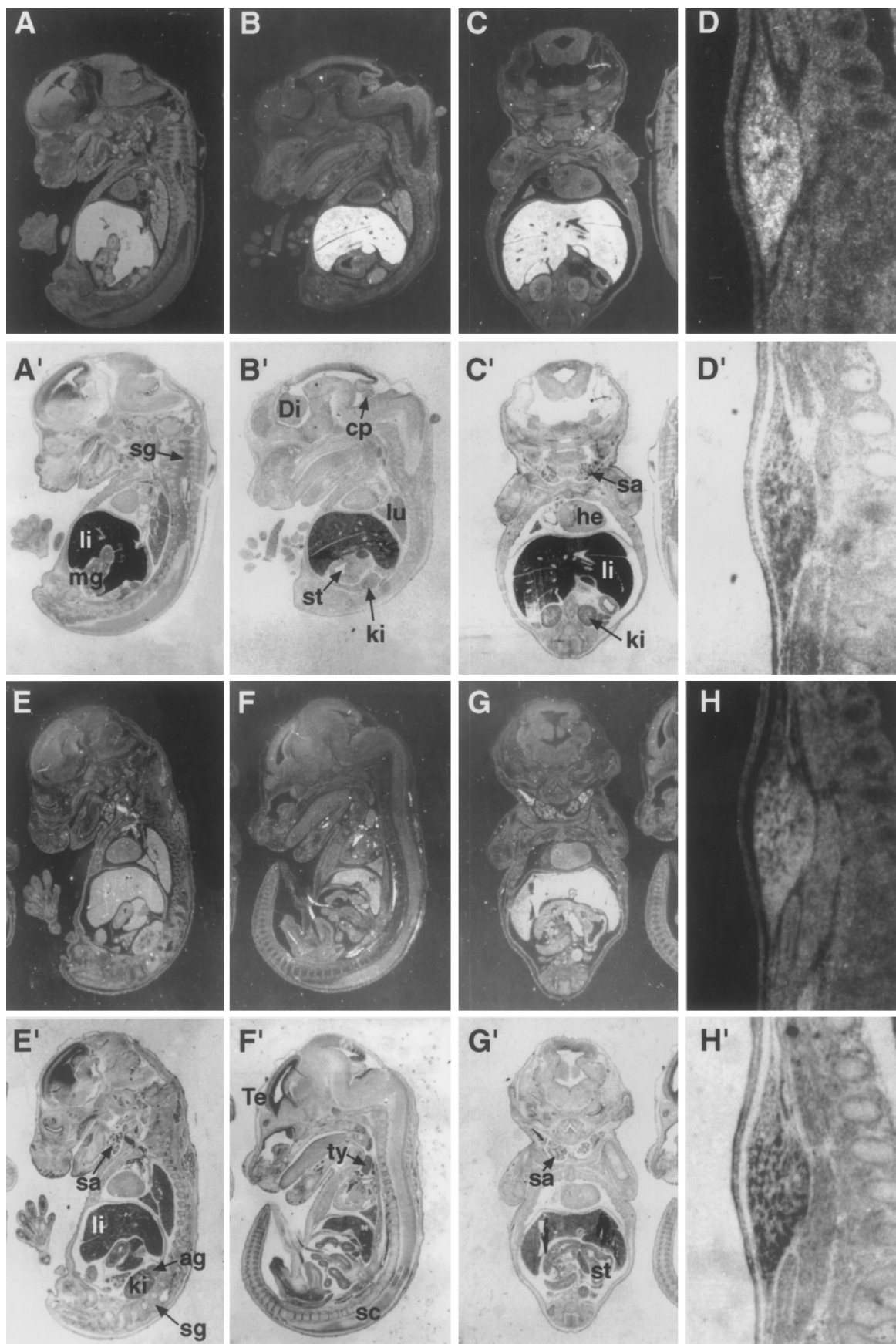
Among these, of particular interest is the highly restricted SKMc15 expression in the hematopoietic hepatic primordium as well as in early circulating meg-

aloblastic precursors at 10.5 dpc, thus providing evidence of a possible role for SKMc15 in early embryonic hematopoiesis.



**FIG. 5.** SKMc15 and IFRD1 (TIS7) expression patterns at early-midgestation. Adjacent sagittal (**A, D, C, and F**) and frontal (**B and E**) sections at E 10.5 (**A, B, D, and E**) and E 12.5 (**C and F**) probed with SKMc15 (**A–C**) and IFRD1 (TIS7) (**D–F**) genes. (**A'–C'**) Bright fields of the corresponding sections. Abbreviations as in Fig. 4 plus mb, midbrain; hp, hepatic primordium; sc, spinal cord; sg, spinal ganglia; st, stomach primordium; Te, telencephalon; Di, diencephalon; Mt, metencephalon; np, nasal pit; lu, lung; li, liver; ki, kidney. The arrows in **D** and **E** point to the IFRD1 (TIS7) signal in the spinal ganglia.





**FIG. 6.** SKMc15 and IFRD1 (TIS7) expression patterns at E 17. Sagittal (**A, B, E, and F**) and frontal (**C and G**) sections were hybridized with SKMc15 (**A–C**) and IFRD1 (TIS7) (**E–G**) probes, respectively. (**D–H**) Magnifications of back muscles probed with SKMc15 (**D**) and IFRD1 (TIS7) (**H**). (**A'–H'**) Bright fields of the corresponding sections. Abbreviations as in the previous figures, plus sa, salivary gland; mg, midgut; ag, adrenal gland; ty, thymus; cp, choroid plexus.



### *Chromosomal Localization of PC4*

Human EST clone 52803 is one of many EST sequences homologous to IFRD1 present in the Unigene database, at cluster 7879. STS 32586 (Généthon) and stSG401 (Sanger Centre) have been generated from 3' EST sequences belonging to the same cluster. Radiation hybrid screening results for these STSs, generated by both Généthon and the Sanger Centre, indicated that the human IFRD1 (PC4) gene maps between markers D7S523 and D7S486, corresponding to 7q22-q31 (cytogenetic position 1.25–1.39 M, as by Chumakov *et al.*, 1995).

### DISCUSSION

The human cDNA whose cloning and sequencing is reported here encodes a protein having 90% amino acid identity to the rat and mouse IFRD1 (PC4 and TIS7) proteins. It can thus be assumed that this cDNA encodes human IFRD1. Sequence and tissue distribution of IFRD1 appear to be highly conserved between species, suggesting that the function of IFRD1 is conserved during evolution. The SKMc15 gene, in virtue of its 47% identity to the IFRD1 sequences, appears to be the second member of a novel family of genes. Although the molecular function of IFRD1 and SKMc15 is still unknown, hypotheses on their functional role are suggested from our analysis of expression. In the first place, both genes have high expression in skeletal and cardiac muscle of the adult human, as seen by Northern analysis. Furthermore, both genes are expressed in the embryonic skeletal muscle, attaining an appreciable level at the late gestation period. This fact, in line with the requirement for IFRD1 (PC4) observed in the process of muscle differentiation (Guardavaccaro *et al.*, 1995), suggests that both proteins might share a common function(s), possibly encoded by the three regions of higher conservation. The same type of suggestion comes from the observation that during development both genes have qualitatively similar patterns, with early appearance in the embryo and extraembryonic tissues since gastrulation, followed by a common pattern of expression in restricted structures at midgestation (such as CNS, kidney, and lung primordia) or ubiquitous expression at late gestation. However, between the two genes there are strong and specific differences in the level of expression, since IFRD1 (TIS7) is expressed much more than SKMc15 in differentiating tissues, such as nervous tissues, kidney, and lung, while SKMc15 is highly expressed in the initial stages of embryogenesis and in the hepatic primordium. These findings agree with our previous analysis of IFRD1 (PC4) expression within the CNS, which showed that IFRD1 (PC4) was present not only in the ventricular zone of the neural tube at the moment of neuroblast proliferation, but also in the surrounding mantle zone where the differentiated neuron migrates (Iacopetti *et al.*, 1996). Considering all together, it is

plausible to imagine a role for IFRD1 in terminal differentiation of different cellular districts. This hypothesis is consistent with our previous observation of nuclear translocation of the IFRD1 (PC4) protein during neuronal differentiation (Guardavaccaro *et al.*, 1994), which suggests a role for IFRD1 (PC4) in this process. Similarly, our present data are very suggestive of an involvement of SKMc15 in the early differentiation of hematic precursors. It is interesting that SKMc15, isolated as a putative tumor suppressor gene, has been excluded from being a such candidate, based on mutational analysis (Latif *et al.*, 1997). Furthermore, we have observed that IFRD1 does not affect cellular proliferation (D. Guardavaccaro, unpublished results). It is therefore likely that SKMc15 and IFRD1 are not impinging on cell cycle progression. Certainly further analyses, as well as functional comparisons between IFRD1 and SKMc15, might help to assign a functional role to this gene family. Concerning the similarity between IFRD1 (PC4) and interferon- $\gamma$  that we have previously observed (Tirone and Shooter, 1989), a comparison of interferon- $\gamma$  proteins of different species with rat and human IFRD1 (PC4) still bespeaks common ancestry of these sequences (R. Doolittle, San Diego, pers. commun., December 1996). The interspecies conservation of the interferon protein, one of the fastest changing mammalian proteins, is, however, lower than observed for IFRD1, indicating different functional aspects and different positive selection for change. A comparison of the three-dimensional structures will thus be necessary for a better understanding of the evolutionary relation between IFRD1 and interferon- $\gamma$  proteins. Concerning the chromosomal localization of IFRD1, which evidently differs from that of SKMc15 (assigned to chromosome 3p21; Latif *et al.*, 1997), we cannot at present exclude a functional correlation of IFRD1 with any of the diseases mapping to the IFRD1 locus. Further studies will therefore be necessary to ascertain whether IFRD1 is involved in diseases related to development and/or cellular differentiation.

### ACKNOWLEDGMENTS

We are grateful to Russel Doolittle for the computer-assisted analysis of the human PC4 protein and to Giuseppe Borsani and Andrea Ballabio for critically reading the paper. We thank Sergio Cattadori for technical assistance. This work was supported by a Telethon grant to F.T. and to A.S. and by an AIRC grant to A.S.; P.B. was the recipient of a fellowship from Telethon. This work was also carried out under a research contract with NEFAC, Pomezia, Italy, within the Neurobiological Systems National Research Plan of the Ministero dell'Università e della Ricerca Scientifica e Tecnologica.

*Note added in proof.* After submission of the manuscript, a report appeared indicating that the locus linked to a severe genetic speech and language disorder colocalizes with the chromosomal region where IFRD1/PC4 gene maps, thus making IFRD1/PC4 a potential candidate responsible for such disease [Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., and Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nat. Genet.* **18**: 168–170.]

## REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
- Bairoch, A. (1990). "Dictionary of Protein Size and Patterns," Ph.D. thesis, Université de Genève.
- Boguski, M. S., Lowe, T. M., and Tolstoshev, C. M. (1993). dbEST—Database for "expressed sequence tags." *Nat. Genet.* **4**: 332–333.
- Bradbury, A., Possenti, R., Shooter, E. M., and Tirone, F. (1991). Molecular cloning of PC3, a putatively secreted protein whose mRNA is induced by nerve growth factor and depolarization. *Proc. Natl. Acad. Sci. USA* **88**: 3353–3357.
- Chumakov, I. M., Rigault, P., Le Gall, I., Bellanne-Chantelot, C., Billault, A., Guillou, S., Soularue, P., Guasconi, G., Poullier, E., Gros, I., *et al.* (1995). A YAC contig map of the human genome. *Nature* **377**: 175–297.
- Doolittle, R. F., and Feng, D. F. (1990). Nearest neighbor procedure for relating progressively aligned amino acid sequences. *Methods Enzymol.* **183**: 659–669.
- Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6–13.
- Greene, L. A., and Tischler, A.S. (1976). Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. *Proc. Natl. Acad. Sci. USA* **73**: 2424–2428.
- Guardavaccaro, D., Ciotti, M. T., Schafer, B. W., Montagnoli, A., and Tirone, F. (1995). Inhibition of differentiation in myoblasts deprived of the interferon-related protein PC4. *Cell Growth Differ.* **6**: 159–169.
- Guardavaccaro, D., Montagnoli, A., Ciotti, M. T., Lotti, L., Di Lazaro, C., Torrisi, M-R., Gatti, A., and Tirone, F. (1994). Nerve growth factor regulates the sub-cellular localization of the nerve growth factor-inducible protein PC4 in PC12 cells. *J. Neurosci. Res.* **37**: 660–674.
- Hollenberg, S. M., Cheng, P. F., and Weintraub, H. (1993). Use of a conditional MyoD transcription factor in studies of MyoD transactivation and muscle determination. *Proc. Natl. Acad. Sci. USA* **90**: 8028–8032.
- Iacopetti, P., Barsacchi, G., Tirone, F., and Cremisi, F. (1996). Expression of the PC4 gene in the developing rat nervous system. *Brain Res.* **707**: 293–297.
- Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**: 857–872.
- Latif, F., Duh, F.-M., Bader, S., Sekido, Y., Geil, H.-L. L., Zbar, B., Minna, J. D., and Lerman, M. (1997). The human homolog of the rodent immediate early response genes, PC4 and Tis7, resides in the lung cancer tumor suppressor gene region on chromosome 3p21. *Hum. Genet.* **99**: 334–341.
- Theiler, K. (1989). The house mouse. In "Atlas of Embryonic Development," Springer-Verlag, Zurich.
- Tirone, F., and Shooter, E. M. (1989). Early gene regulation by nerve growth factor: Induction of an interferon related gene. *Proc. Natl. Acad. Sci. USA* **86**: 2088–2092.
- Varnum, B. C., Lim, R. W., and Herschman, H. R. (1989). Characterization of TIS 7, a gene induced in Swiss 3T3 cells by the tumor promoter tetradecanoyl phorbol acetate. *Oncogene* **4**: 1263–1265.
- Wei, M. H., Latif, F., Bader, S., Kashuba, V., Chen, J. Y., Duh, F. M., Sekido, Y., Lee, C. C., Geil, L., Kuzmin, I., Zabarovsky, E., Klein, G., Zbar, B., Minna, J. D., and Lerman, M. I. (1996). Construction of a 600-kilobase cosmid clone contig and generation of a transcriptional map surrounding the lung cancer tumor suppressor gene (TSG) locus on human chromosome 3p21.3: Progress toward the isolation of a lung cancer TSG. *Cancer Res.* **56**: 1487–1492.
- Wilkinson, D. G. (1992). Whole mount in situ hybridization of vertebrate embryos. In "In situ Hybridization: A Practical Approach," IRL Press, Oxford.
- Wilkinson, D. G., and Green, J. (1990). In "Postimplantation Mouse Embryos: A Practical Approach," (D. Rickwood and D. L. Cockcroft, Eds.), pp. 155–171, IRL Press, Oxford.