Int. J. Dev. Biol. 40: 545-555 (1996)

Original Article

Structure and developmental expression of mouse *Garp*, a gene encoding a new leucine-rich repeat-containing protein

RÉGINE ROUBIN*, SANDRINE PIZETTE, VINCENT OLLENDORFF[1], JACQUELINE PLANCHE, DANIEL BIRNBAUM* and ODILE DELAPEYRIERE[2]

Laboratoire d'Oncologie Moléculaire, U. 119 INSERM, Marseille, France

ABSTRACT Proteins with leucine-rich repeats (LRR) constitute a large family of molecules playing a role in protein-protein interactions and signal transduction. They are involved in various cellular processes in different species. We characterized the organization and pattern of expression of the mouse *Garp* gene. It is composed of two coding exons, expressed as a major 4.3 kb mRNA, and encodes a putative LRR transmembrane protein with an extracellular region almost entirely made of 20 repeats, and a short intracytoplasmic region. The mouse GARP deduced amino-acid sequence is highly similar to that of the human protein. The *Garp* gene is expressed in various areas in the mid-gestation developing embryo, including skin, lens fibre cells, nasal cavity, smooth and skeletal muscles, lung, and megakaryocytes of the fetal liver. In the adult it is expressed in the megakaryocytes of the spleen and in endothelial cells of the placenta. The data suggests that GARP might be involved in platelet-endothelium interactions.

KEY WORDS: *adhesion, embryogenesis, endothelial cells, leucine-rich repeats, mouse genome, platelet*

Introduction

Proteins with leucine-rich repeats (LRR) constitute an important superfamily involved in protein-protein interactions necessary for several processes (for reviews see Ollendorff *et al.*, 1993 and Kobe and Deisenhofer, 1994). It comprises proteins with as diverse functions and origins as yeast adenylate cyclase (Kataoka *et al.*, 1985), human platelet von Willebrand factor receptor complex (Lopez *et al.*, 1987), ribonuclease/angiogenin inhibitor (Schneider *et al.*, 1988), a murine antigen involved in B cell activation (Miyake *et al.*, 1995), *D. melanogaster* and *C. elegans* proteins Toll (Hashimoto *et al.*, 1988), Tartan (Chang *et al.*, 1993), Connectin (Nose *et al.*, 1992), Chaoptin (Reinke *et al.*, 1988), Slit (Rothberg *et al.*, 1990) and Flightless-1 (Campbell *et al.*, 1993), extracellular matrix components such as decorin (Krusius and Ruoslahti, 1986) and biglycan (Fischer *et al.*, 1989), and plant disease resistance genes (reviewed in Dangl, 1995). All these proteins share a common domain containing several copies of a 24 amino acids stretch with the consensus sequence LXLXXNXLXXLXXXXLXXLXXL (Ollendorff *et al.*, 1993; Kobe and Deisenhofer, 1994).

The human *GARP* gene has been described as coding for an 80 kDa transmembrane protein with an extracellular region made almost exclusively of 20 LRR (Ollendorff *et al.*, 1994). It is located in the chromosomal bands 11q13.3-q14 in humans, and 7F in the mouse, in a region of synteny conserved between the two species (Ollendorff *et al.*, 1992). The 11q13-q14 region is amplified in some types of tumors such as breast tumors, and the *GARP* gene is sometimes included in the amplification units (Gaudray *et al.*, 1992; Szepetowski *et al.*, 1992).

As a means to obtain more information on the structure and function of the *GARP* gene, we cloned and sequenced the murine gene and studied by *in situ* hybridization its spatio-temporal expression in the mouse. Because of its cellular distribution, GARP may play a role in various tissues including platelet-endothelium interaction.

Results

Isolation of a mouse Garp cosmid clone and sequencing of the Garp gene coding region

The screening of a genomic library with cDNA probes led to the obtention of one type of cosmid, referred to as 3.2. Mapping of cosmid 3.2 was performed and a tentative representation of the mouse *Garp* locus is shown in Figure 1. EcoRI cosmid digests were probed with various fragments of the human *GARP* cDNA. Two genomic fragments of 1.3 and 4 kb, respectively, were positive for *GARP* (Fig. 1) and were further subcloned in Bluescript vector. A restriction map of the two fragments was established and allowed further subcloning and sequencing. Sequence of the two genomic clones and the position of splice

Abbreviations used in this paper: LRR, leucine rich-repeats; dpc, days post conception; PMA, phorbol-myristate-acetate.

***Address for reprints:** Laboratoire d'Oncologie Moléculaire, U. 119 INSERM, 27 Bd. Leï Roure, 13009 Marseille, France. FAX: 33.91260364. e-mail: birnbaum@infobiogen.fr

Present addresses: [1]UCSD, San Diego, CA, USA; and [2]U.382 Inserm, IBDM, Marseille, France.

0214-6282/96/$03.00

junctions enabled us to define an open reading frame of 1992 nucleotides coding for a putative protein of 663 residues.

Like in humans, the mouse *Garp* gene comprises two coding exons, the majority of the coding region being included in the second exon. Analysis of the deduced amino acid sequence (Fig. 2) revealed the presence of a hydrophobic leader sequence extending from residues 1 to 19, which is likely to represent a signal peptide for the secretion of the protein. Two potential protease cleavage sites, located between residues 17 and 18 and between 19 and 20, are present downstream of the hydrophobic leader sequence. The signal peptide and nine additional residues constitute the first exon. The second exon encodes 20 repeats of the leucine-rich motif, each composed of approximately 24 residues (Fig. 2). The consensus sequence contains leucine or other aliphatic residues at position 1, 3, 8, 11, 19, and 22 and an asparagine at position 6. Eighty percent of the GARP structure is made of these units. A proline-rich region is found in between the tenth and the eleventh repeat. It may also correspond to a leucine-rich motif which has diverged from the consensus. In common with other LRR molecules (Kobe and Deisenhofer, 1994), *GARP* is also characterized by a proline residue and four similarly spaced cysteines in a stretch of 44 amino acids in the carboxy-LRR

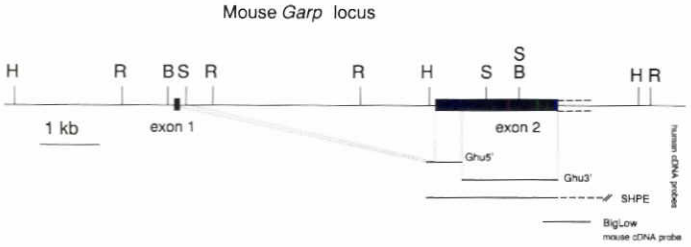

Fig. 1. Structure of the mouse *Garp* gene. *A tentative physical map of the Garp gene, as derived by restriction enzyme mapping of the cosmid 3.2 insert, and the location of the cDNA probes used, are shown. The coding (sequenced) portion of exons are represented by black boxes; the portion extending 3' (not sequenced) is shown in stipplings, as is the corresponding portion of the human SHPE clone. Selected restriction sites are as follows: B, BamHI; H, HincII; R, EcoRI; S, SmaI.*

flanking region. There are five potential N- linked glycosylation sites.

The deduced secondary structure indicates the presence of a succession of ß-sheets and α-helices in a configuration different from that described for the porcine ribonuclease inhibitor (Kobe and Deisenhofer, 1995).

When compared (Fig. 2), the human and mouse deduced amino acid sequences are very similar. The nucleotide sequence of the murine *Garp* presents 80% similarity with the human *GARP*. The murine *Garp* has three additional nucleotides coding for an extra isoleucine residue, inserted at position 18 in the signal peptide sequence. The major difference is observed in the vicinity of the first and second exon boundary with only a 25% difference over 60 nucleotides but a 60% difference at the protein level.

Because the *Garp* gene is well conserved in humans and the mouse, we probed a zoo blot with the human cDNA. DNAs from human, mouse, chicken, Zebra fish, *Xenopus laevis* and *Drosophila melanogaster* were used. They were isolated, fractionated on a 0.8% agarose gel and transferred onto Nytran (Scleicher and

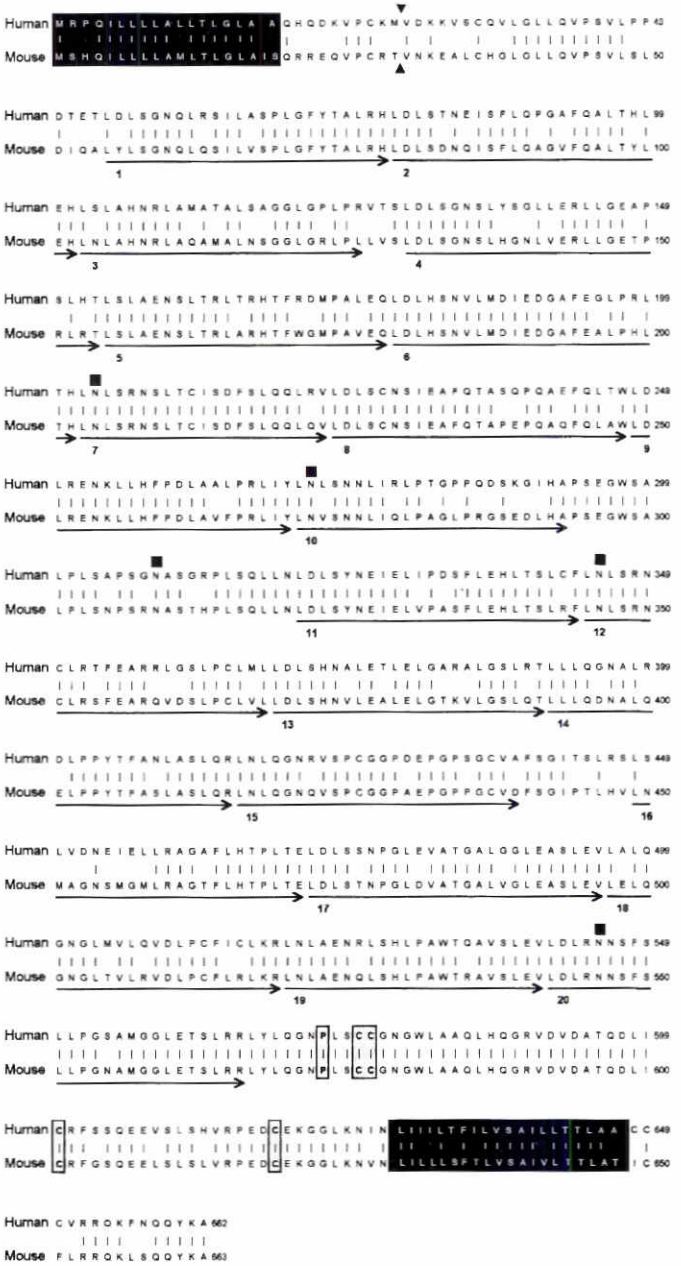

Fig. 2. Alignment of the deduced amino acid sequences of the mouse and human *GARP* genes. *The coding sequence is characterized by the presence of a putative signal peptide (black box, positions 1-19), two series of 10 blocks of leucine-rich motifs (numbered and underlined arbitrarily starting from the putative region of the ß-sheet, separated by a proline-rich sequence, a putative transmembrane domain (black box, positions 629-649) followed by a short, intracellular, carboxy-terminal tail of 15 residues. Also indicated are proline and cysteine residues (boxed) conserved in the C-terminal LRR flanking sequences of LRR-rich proteins. Arrowheads indicate the separation of the two exons. Black squares indicate putative N-glycosylation sites.*

Schuell, Dassell, Germany) membranes. Ten ug of the various DNAs digested by EcoRI were loaded. Blots were hybridyzed and washed at moderate stringency (55°C). The pattern observed on a Southern blot showed that *Garp* was present in all tested

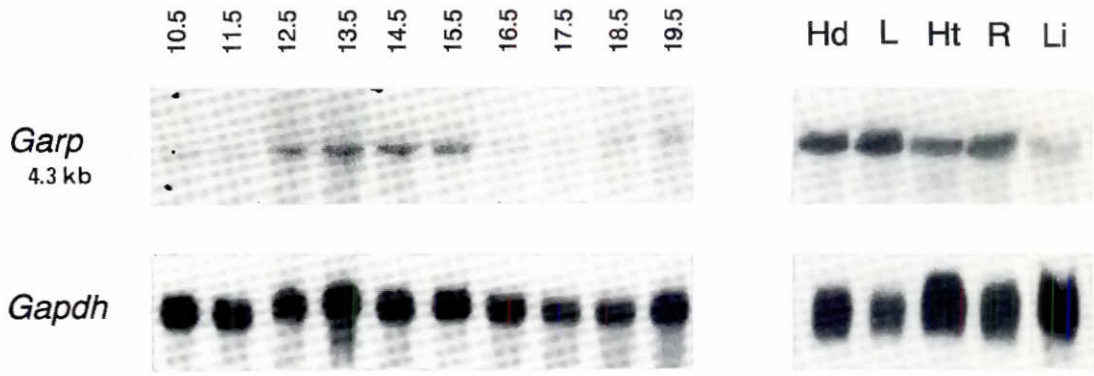

Fig. 3. Expression of *Garp* in mouse embryos. (Left panel) *Temporal expression of Garp in middle and late development was determined by Northern blot hybridizations of total RNAs from mouse embryos with a mouse cDNA Garp probe and a control Gapdh probe.* **(Right panel)** *Portions of a 15.5 dpc embryo were hybridized with the same probe; Hd, head; L, lung; Ht, heart; R, rest of the embryo; Li, liver.*

vertebrates but not in invertebrates, although the band was readily visible only in the human and murine DNAs (not shown).

Expression of Garp in the developing mouse embryo

Expression of *Garp* mRNA was examined in the mouse embryo. RNAs extracted at various stages of development were analyzed by Northern blot. Expression of a 4. 3 kb transcript was readily observed from 12.5 dpc (days post conception) to 15.5 dpc. Before and later on, *Garp* transcript was faintly detectable (Fig. 3, left panel). *Garp* mRNA were present in all tested dissected embryonic portions (Fig. 3, right panel).

To get further insight into the role of GARP during embryogenesis, paraffin sections of prefixed mouse embryos were hybridized *in situ* to either antisense or sense *Garp* probes. Since the maximum level of *Garp* expression was observed during the period of late organogenesis (see Fig. 3), the 14.5 dpc stage was chosen for a thorough investigation, though earlier and later stages were also analyzed for comparison. No signal above background was ever observed using the sense probe (data not shown).

Using the antisense probe, several major sites of expression were conspicuous. First, *Garp* RNA was detected in certain regions of the skin (Fig. 4), especially in regions where the skin is folding or invaginating, such as the neck region (Fig. 4A,B), the lip

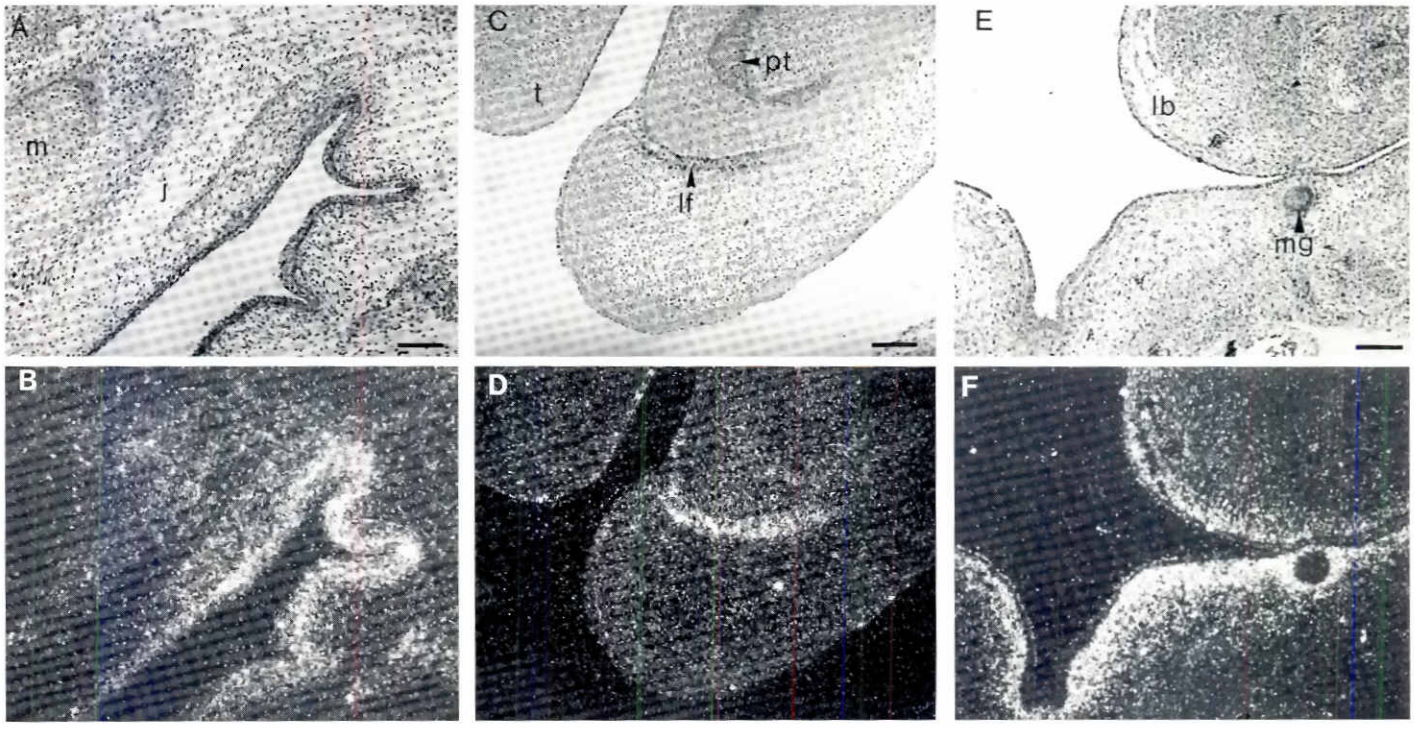

Fig. 4. Expression of *Garp* in the developing skin. (A,C,E) *Bright-field and* **(B,D,F)** *dark-field views of sagittal sections through the neck* **(A,B)** *and the mouth* **(C,D)** *of a 14.5 dpc embryo and of a transverse section through the trunk region* **(E,F)** *of a 12.5 dpc embryo. Hybridizations to the Garp probe show expression in the skin* **(B and F)** *and lip furrow band* **(D)**. *m, Meckel's cartilage; j, jaw; t, tongue; pt, primordium of tooth; lf, lip furrow band; lb, forelimb bud; mg, primordium of mammary gland. Bar, 100 μm.*

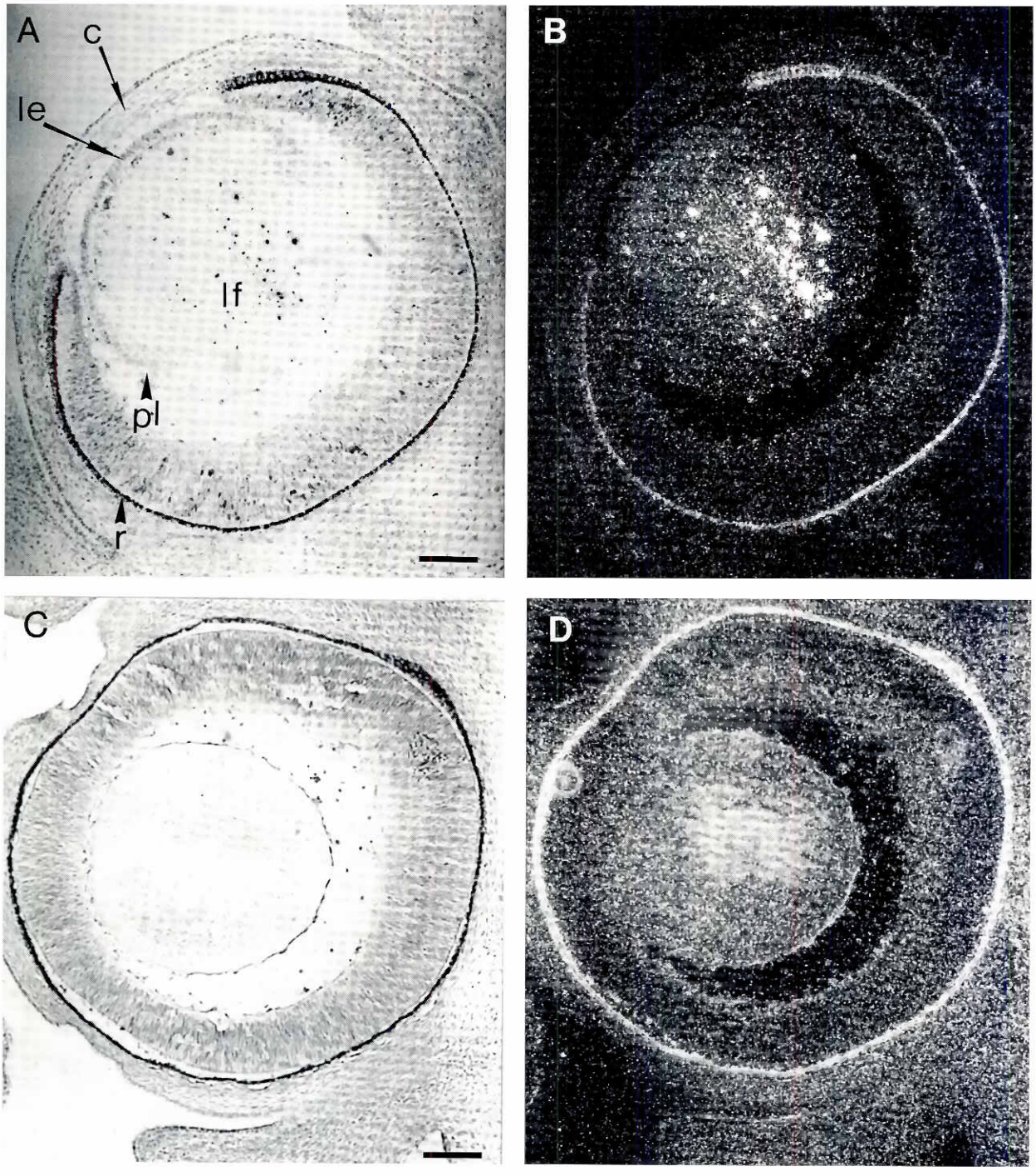

Fig. 5. Expression of *Garp* in the eye of a 14.5 dpc embryo. (A,C) *Bright-field and* **(B,D)** *dark-field views of sagittal sections through the eye. Hybridizations to the Garp probe show a strong expression in the lens fibres* **(A,B)**; *control hybridization with irrelevant probe is shown in* **(C,D)**; *a longer exposure shows a high homogenous background with no specific labeling. The pigmented layer of the retina (r) exhibits the usual false positive signal. pl, proliferating area of lens; lf, lens fibres; le, lens epithelium; c, cornea. Bar, 100 μm.*

furrow band (Fig. 4C,D), the interdigital web (not shown) or around the mammary gland primordium (Fig. 4D,E). The dermis seemed to express *Garp*. Control sense probe did not show positive labeling (not shown).

Second, *Garp* expression was observed in several cephalic structures. In the eye, lens fibre cells expressed *Garp* at a high level (Fig. 5). In contrast, the lens epithelial cells as well as the proliferative and early fibre differentiation zones were

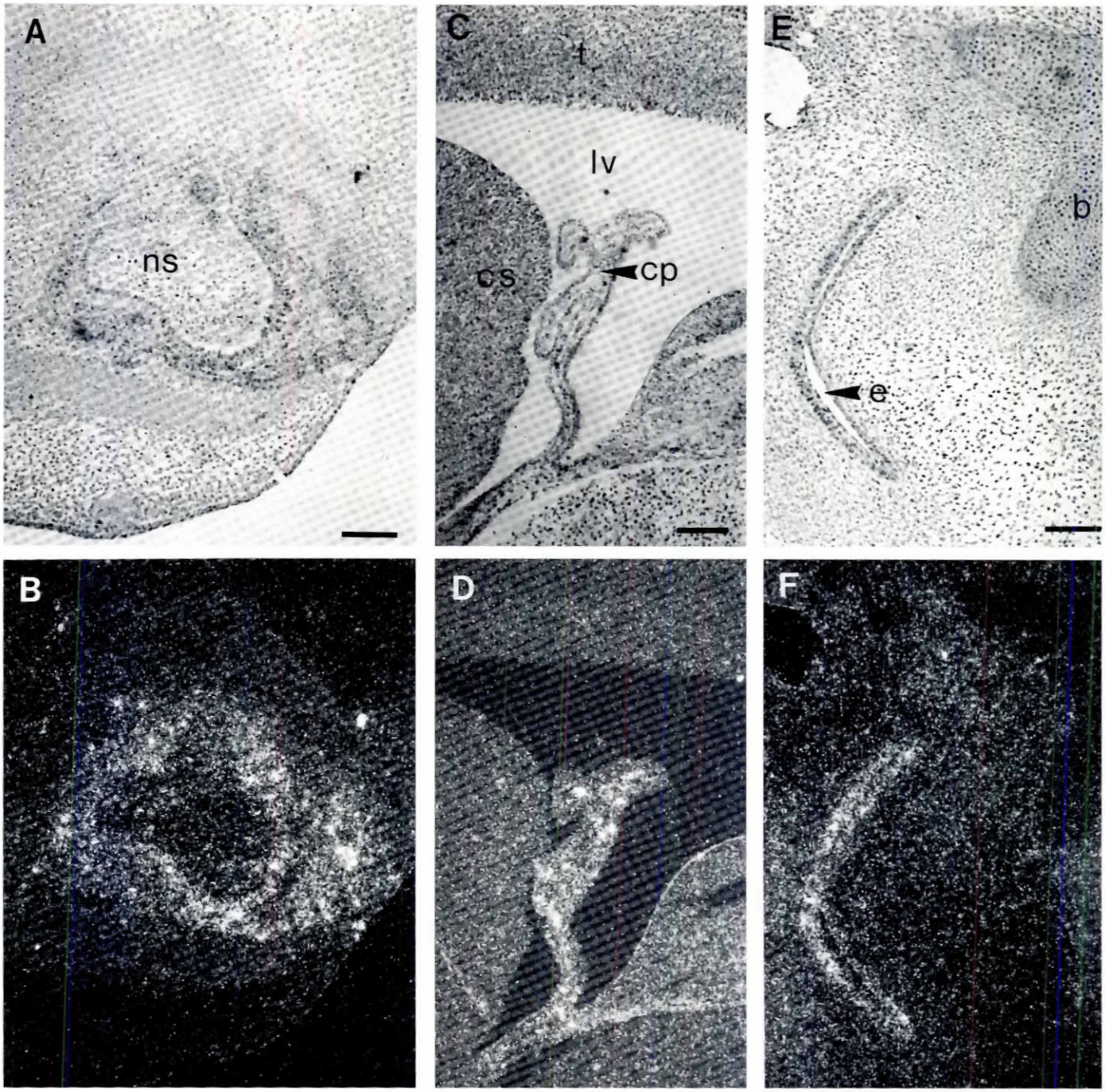

Fig. 6. Expression of *Garp* in regions of the head of a 14.5 dpc embryo. (A,C,E) *Bright-field and* **(B,D,F)** *dark-field views of sagittal sections through the head, and more specifically the snout* **(A,B)**, *the forebrain* **(C,D)**, *and the level of the ear* **(E,F)**. *Garp is expressed in the nasal cavity* **(B)**, *the lateral choroid plexus* **(D)**, *the pharyngo-tympanic (Eustachian) tube* **(F)**. *ns, nasal septum; t, telencephalon; lv, lateral ventricle; cp, lateral choroid plexus; cs, corpus striatum; b, basioccipital bone; e, pharyngo-tympanic (Eustachian) tube. Bar, 100 μm.*

negative. The region of the nasal cavity (Fig. 6A,B), the choroid plexus of the brain ventricles (Fig. 6C,D), the meninges (not shown) and the Eustachian tube of the ear (Fig. 6E,F) were also positive.

Muscle was recognized as another major site of expression (Fig. 7). Signals were detected in skeletal muscle as patches (Fig. 7A-D), and in the smooth muscle layers of the digestive tract (Fig. 7E,F). Whether the cells expressing *Garp* are muscle or nerve cells remains to be determined.

Finally, *Garp* expression was also found in lung (Fig. 8A,B), the urogenital sinus (Fig. 8C,D), and was very strong in the megakaryocytes of the fetal liver (see Fig. 11A,B).

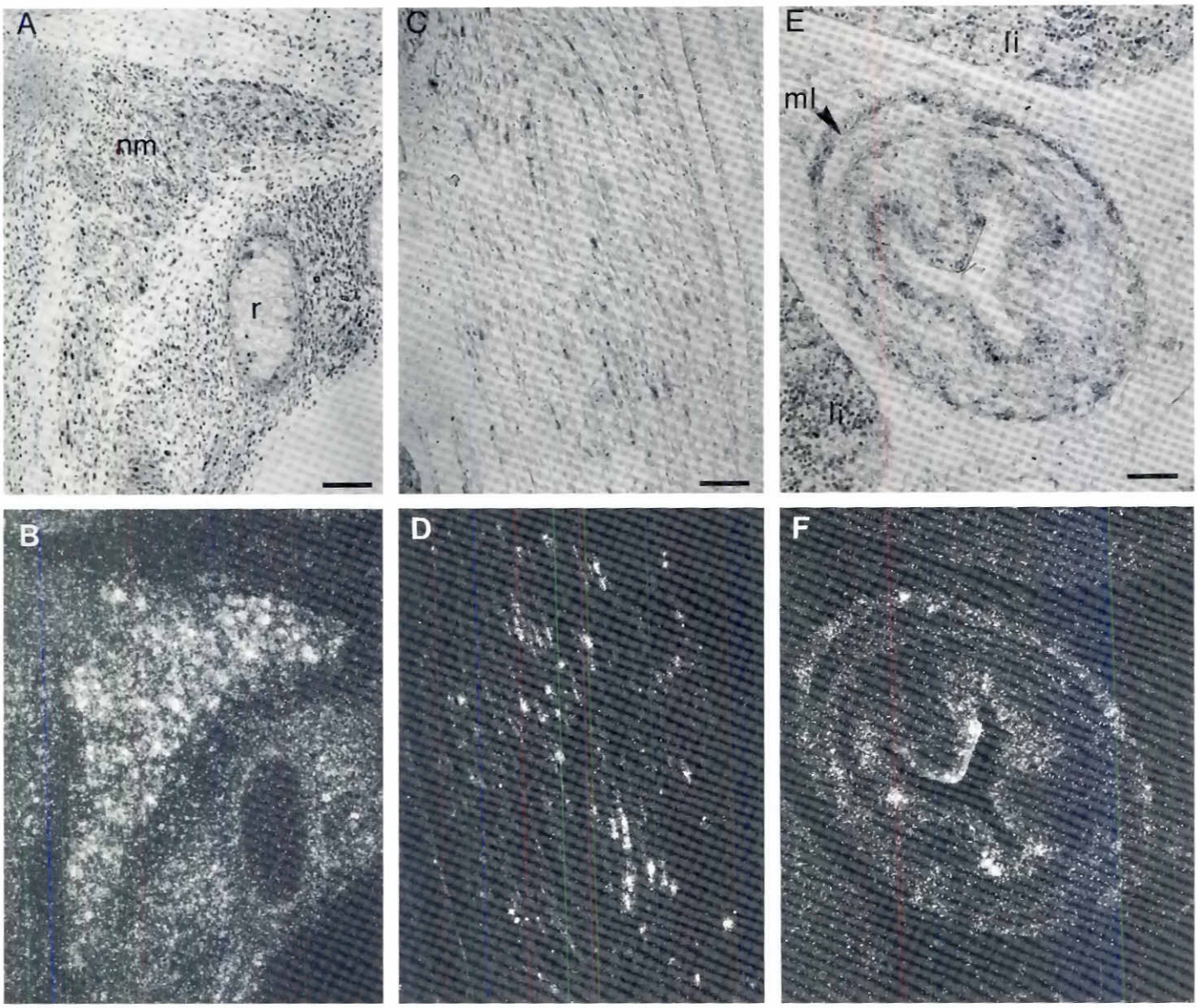

Fig. 7. Expression of *Garp* in developing skeletal and smooth muscles. (A,C,E) *Bright-field and* **(B,D,F)** *dark-field photographs of the same sagittal sections from a 14.5 dpc* **(B)** *and a 16.5 dpc* **(D,F)** *embryo hybridized to the Garp probe. Skeletal muscles from the neck (nm) or the trunk* **(C,D)** *show a strong positive signal in patches* **(B,D)**. *The muscular layer (ml) of the duodenum is also positive* **(F)**. *r, rib; li, liver. Bar, 100 μm.*

Expression of GARP in adult mouse tissues and in various cell lines undergoing in vitro differentiation

To determine the sites of expression of *Garp* in the adult mouse, selected tissues were dissected and total RNAs prepared for Northern analysis. Various tissues including placenta, kidney, heart, lung and ovary expressed a 4.3 kb transcript. When normalized to the *Gapdh* content, equivalent amounts of messengers were detected in these tissues (Fig. 9A). Little or no expression was detected in adrenal gland, testis, prostate, spleen, brain and cerebellum. Because in humans various transcripts have been described (Ollendorff *et al.*, 1994), we performed polyA$^+$ isolation in some selected samples. As shown in Figure 9B, although the 4.3 kb was always the major transcript, a 2.6 and a 2.2 kb transcripts, not visible with total RNA, were also detected. In embryo polyA$^+$ RNA the 2.6 kb transcript appeared strongly. In testis, no 4.3 kb transcript was seen but an additional 0.8 kb transcript was also observed.

As compared with the 14.5 dpc embryonic muscle (Fig. 7A-D), *Garp* was only slightly expressed in the adult skeletal muscle (Fig. 9). To determine whether *Garp* expression varies during myogenesis, we used the mouse C2 cell line which is capable of *in vitro* myogenic differentiation. *Garp* RNA was observed when C2 were grown as proliferating myoblasts (Fig. 10A, lane Mb). The transcript became undetectable when the culture conditions were switched to serum starved medium, i.e. in myotubes (lane Mt), a result in keeping with the detection of the low level of expression in adult muscle. Because the expression of several genes can be reinduced after denervation or during the process of regeneration, we examined *Garp* expression in animals in which a local injury had been inflicted by intramuscular injection of the drug notexin.

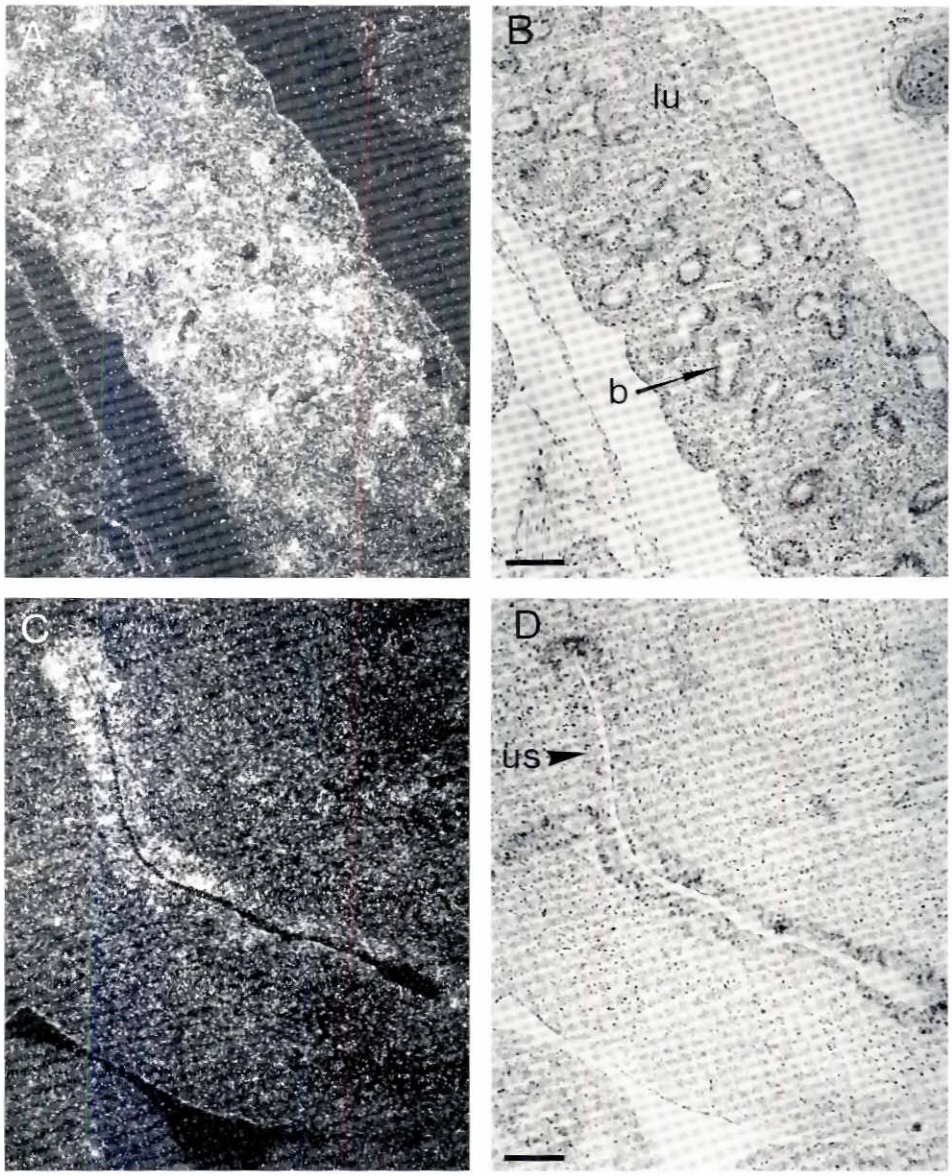

Fig. 8. Presence of *Garp* RNA in various developing organs. (A,C) *Dark-field and* **(B,D)** *bright-field views of sagittal sections through a 14.5 dpc embryo. Hybridizations to the Garp probe show expression in the bronchioles (b) of the lung* **(A,B)** *and the endodermal lining of the urogenital sinus (us)* **(C,D)**. *Scale bar, 100 μm.*

Mice were injected intramuscularly into the leg with 0.5 μg notexin (Sigma) as described (Mollier *et al.*, 1990), and were sacrificed 2 to 6 days later. RNA extracted from the leg muscles of mice treated for 3 to 6 days with notexin were compared to those extracted from untreated animals. No difference in the intensity of the *Garp* signal was noted, although histological analysis of sections of 4 day-treated legs indicated full muscle regeneration, with fibers exhibiting typical central nuclei (not shown).

Spleen was another adult murine tissue in which *Garp* expression was positive. The signal was barely detectable by Northern analysis. *In situ* hybridization revealed that *Garp* positive cells were megakaryocytes (Fig. 11C), a result in keeping with those on the megakaryocytes of the fetal liver (Fig. 11A,B). To strengthen this result, we used an *in vitro* model of cell differentiation where cells from HEL, a human erythroleukemia cell line, were allowed to differentiate into megakaryocytes after culture in the presence of phorbol esters (1 nM PMA). *GARP* expression was studied by Northern blot hybridization. As compared with untreated cells, PMA-treated HEL cells exhibited signals for both *GARP* and *GPIIb*, a control of megakaryocytic differentiation. In HEL cells, both *GARP* and *GPIIb* transcripts were absent (Fig. 10B).

Garp mRNA was also expressed in placenta (Fig. 9). To determine which cell type was responsible for this reactivity, we performed *in situ* hybridization of sections of the mouse decidua and placenta. *Garp* was found to be expressed in endothelial cells of decidua (Fig. 11D,E) and placenta (not shown). *Garp* mRNA appeared to be restricted to a subset of endothelial cells. Again, to further document this result, we studied freshly isolated cells from human umbilical cord vein (HUVEC). HUVEC were found to express a 4.4 Kb *GARP* transcript, as detected by Northern blot hybridization (Fig. 10C, lane 1). Similarly, tEnd1, an endothelial cell line isolated from the mouse thymus and immortalized by the polyoma middle T antigen, was positive (Fig. 10C, lane 3). Not every endothelial cell line tested was positive since ECV304, a spontaneously transformed endothelial cell line established from the vein of a human umbilical cord, did not express *GARP* transcript (Fig. 10C, lane 2), even when stimulated by IL1α or TNFα (not shown).

We never found a tumor expressing *GARP* mRNA, although 24 breast tumors selected for both their high content in blood vessels and/or the presence of an amplification of chromosome bands 11q13-q14, the location of the *GARP* gene, were examined using Northern blot hybridization under previously described conditions (Theillet *et al.*, 1993). Similarly, we failed to observe positive endothelial cells in muscle sections of notexin-treated mice although the neovascularization occurring during wound healing was morphologically observed.

Discussion

In this work, we have studied the organization and pattern of expression of the mouse *Garp* gene. It encodes a putative transmembrane molecule almost entirely made of leucine-rich motifs. The mouse and human *GARP* nucleotide and amino acid sequences are 80% and 82% identical, respectively. The similarity rises to 87% if one allows for conservative amino acid changes. The similarity is high throughout the extracellular domain: the

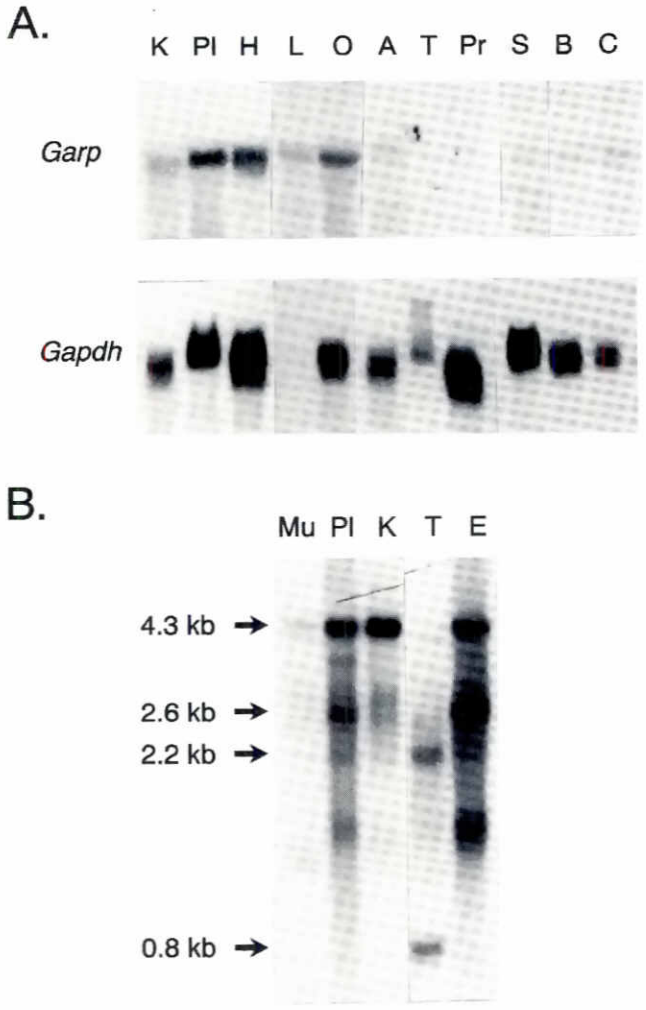

Fig. 9. Expression of the *Garp* gene in adult mouse tissues. (A) *Northern blot membranes were hybridized as described in Materials and Methods using the "BigLow" fragment probe. Each lane contained 20 μg of total RNA prepared from a variety of adult mouse tissues: K, kidney; Pl, placenta; H, heart; L, lung; O, ovary; A, adrenal gland; T, testis; Pr, prostate; S, spleen; B, brain; C, cerebellum. The same membranes were rehybridized with a Gapdh probe (note that this probe does not detect transcript in lung).* **(B)** *Same as in A except that each lane contained 10 μg of poly(A)⁺ RNA. Mu, skeletal muscle; Pl, placenta; K, kidney; T, testis; E, 14.5 dpc embryo. The sizes of the different transcripts, as determined from the migration of marker AvaI + NarI lambda DNA fragments (not shown), are indicated on the left.*

number of LRR is identical, and the cysteines in the C terminal flanking region are conserved; this is not the case in the intracytoplasmic part of the molecule, with 33% difference but a tyrosine residue, a putative phosphorylation site involved in signal transduction, is conserved. The sequence corresponding to the first-second exon boundary is not conserved, indicating that this part of the molecule might not be important for GARP conformation.

Proteins containing LRR domains are thought to have a role in specific protein-protein interaction. For example, the four LRR proteins from *D. melanogaster* all exhibit cell adhesion properties. Chaoptin is a photoreceptor cell specific glycoprotein known to be involved in cell adhesion. Connectin is a cell adhesion molecule expressed on muscles and their innervating motoneurons during embryogenesis. The Toll transmembrane protein is involved in the specification of dorsal-ventral polarity and the Slit protein is involved in the development of glial cells and commissural axons. Furthermore, the conservation of the LRR domain in such distantly separated organisms as yeasts, worms (Campbell *et al.*, 1993), flies and mammals indicate that proteins containing LRR may serve critical cellular functions in higher eukaryotes. It is likely that GARP is involved in protein-protein interaction.

Garp expression was weak before organogenesis but high in mid-gestation embryos. As a step towards understanding the role of GARP during organogenesis, *in situ* hybridization of embryos was carried out. Various sites of expression in structures of different origin were observed.

The putative expression of *Garp* at the neuromuscular junction or at the dermis/epidermis junction is suggestive of a role for GARP as an adhesion molecule necessary to bring together tissues of different origins during the developmental process.

A site of expression which retained our attention was endothelium. Interestingly, *Garp* was also found to be expressed in megakaryocytes, the platelet precursors. The presence of GARP both on platelets and endothelium suggests that GARP may play a role in platelet aggregation and clot formation. Further investigations are needed to confirm or refute this hypothesis. A family of LRR proteins are platelet adhesion glycoproteins GPIbα, GPIbß, GPV and GPIX. They are part of a complex playing a key

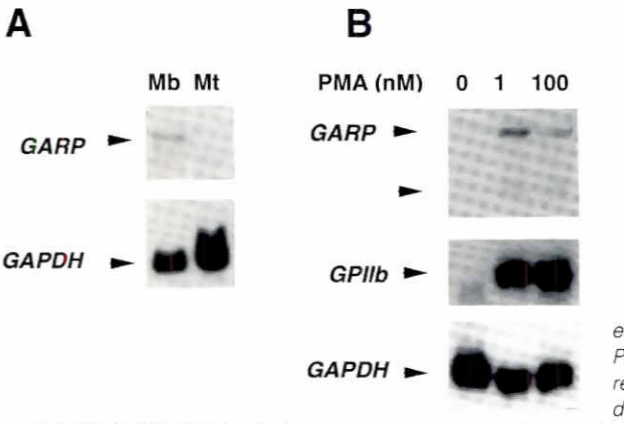

Fig. 10. Expression of *GARP* in various cell lines. *Northern blot hybridizations of mouse and human cell line RNAs show expression of GARP in distinct cellular models. 20 μg of total RNAs is loaded in each lane. For each analysis, blots were dehybridized and rehybridized with a mouse Gapdh cDNA probe, for control.* **(A)** *C2 cells were cultured in proliferation medium as myoblasts (Mb) then shifted to differentiation medium for 3 days where they fused as myotubes (Mt).* **(B)** *HEL erythroleukemia cells were cultivated in the absence or presence of phorbol ester PMA. With the latter they differentiate into megakaryocytic cells. Blots were rehybridized with the GPIIb probe, to control for the quality of differentiation.* **(C)** *Three different endothelial cell lines were tested for GARP expression. Lanes are as follows: 1: HUVEC, 2: ECV304, 3: tEnd1. As expected, two bands (arrowheads), one major of 4.4 kb and one minor (only visible on long exposures) of 2.8 kb are detected in RNAs of human origin.*

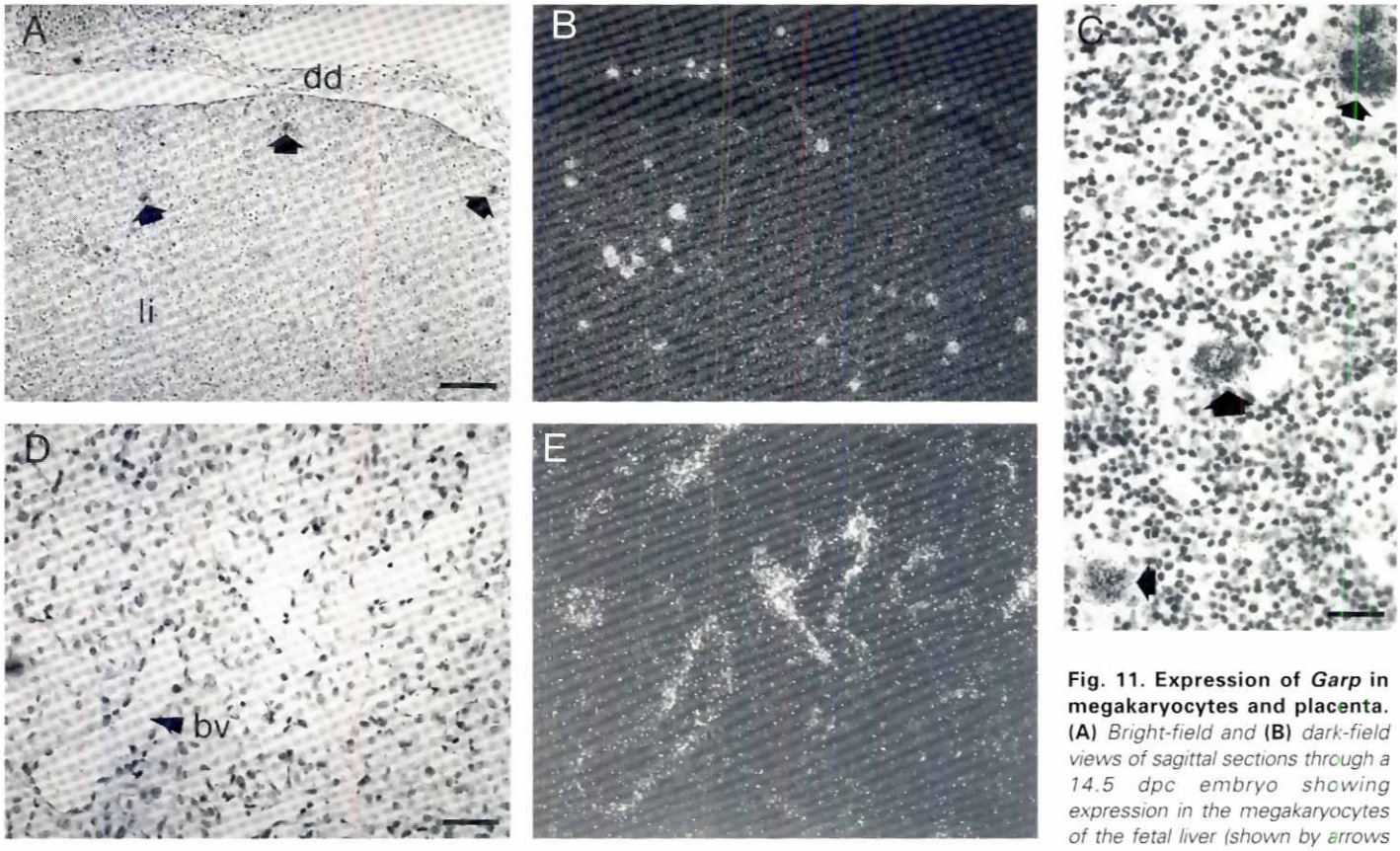

Fig. 11. Expression of *Garp* in megakaryocytes and placenta. **(A)** *Bright-field and* **(B)** *dark-field views of sagittal sections through a 14.5 dpc embryo showing expression in the megakaryocytes of the fetal liver (shown by arrows in* **A***); dd, dome of the diaphragm; li, liver.* **(C)** *Bright-field view through section of adult spleen. Arrows indicate the megakaryocytes present in this organ in the adult in this species.* **(D)** *Bright-field and* **(E)** *dark-field views of decidua (9.5 dpc) sections. bv, blood vessel. Bar, 50 μm (C), 100 μm (D), 200 μm (A).*

role in the shear-dependent initial events involved in attachment of platelets to injured arterial endothelium (reviewed in Roth, 1991, 1992). In particular, although no primary sequence identity is noticeable, GARP organization is similar to that of GPV, which is an 83 kDa protein with 15 LRR, a cluster of carboxy-flanking cysteines and a short intracytoplasmic tail (Hickey *et al.*, 1993).

Endothelial cells in tissues and organs are functionally and morphologically heterogeneous. However, *GARP* was never found to be expressed in the endothelial cell of blood vessels of normal tissues known to be highly vascularized. In contrast, *Garp* was found to be expressed on a subset of endothelial cells in some neovascularized tissues such as the placenta or the decidua. The growth of new capillaries is essential for normal embryogenesis but also occurs in the development of many diseases including tumor growth or wound healing. However, we were unable to demonstrate the presence of *GARP* positive endothelial cells in tumor biopsies or in muscle regeneration after notexin treatment. Whether GARP, as many other adhesion molecules such as integrins, selectins or extracellular matrix components, is required for the growth of new blood vessels of some types and in certain circumstances, remains to be determined.

This work provides the first evidence for *Garp* gene structure and involvement in development. The pattern of expression described here brings some clues to where and when GARP could play a role. This will be helpful for further studies including gene knock-out, which could only establish its specific function. Since LRR comprises almost the entire primary structure of GARP it is likely that the essential property of this protein is to bind other proteins. As for many other LRR proteins, whether it does so in a specific manner is unknown yet. Thus, a number of other questions need to be answered, e.g. is GARP part of a complex of proteins? Does it interact with itself or a specific ligand? Could it participate in mediating transduction of a signal? The fact that *Garp* may play a role in platelet-endothelial cell interactions while being expressed in both types of cells would be in favor of a homophilic GARP-GARP binding.

Materials and Methods

Genomic DNA screening

A cosmid library constructed from EcoRI partially digested DNA extracted from 129 mouse embryonic stem cells and inserted into SuperCos 1 cosmid vector (Stratagene) was a gift from Malek Djabali (CIML, Marseille, France). Clones were screened without prior amplification of the library by filter hybridization at 68°C in 6XSSC, 1X Denhardt with a ^{32}P oligolabeled human cDNA probe. The probes used were fragment derived from a BamHI digest of the complete cDNA clone SHPE (Ollendorff *et al.*, 1994) further digested by EcoRI. EcoRI digest of SHPE led to two fragments: Ghu5' and Ghu3' (see Fig. 1). Mouse genomic fragments from cosmid inserts that hybridized to cDNA probes were subcloned in Bluescript vector and mapped by standard procedures.

Nucleotide sequencing

Nucleotide sequences were determined by the dideoxynucleotide method, using Bluescript (Stratagene) single-stranded or double-stranded

templates and the modified T7 polymerase protocol (Sequenase; USBiochemicals). Sequence analysis was carried out using the PC-gene software (Intelligenetics).

Cells, tissues, tumors

Human erythroleukemia HEL cells were obtained from the American Type Culture Collection (ATCC). They were cultivated in RPMI-1640 medium containing 10% fetal calf serum (FCS). They were allowed to differentiate into megakaryocytes by adding $1x10^{-9}$ to $1x10^{-7}$ M phorbol myristyl acetate (PMA) to serum-containing medium. RNA was extracted 2 days later. Human umbilical vein endothelial cells (HUVECs) were isolated and subcultured as published (Bevilacqua *et al.*, 1985). They were a gift from J. Sampol (Faculté de Pharmacie, Marseille, France). ECV304, a spontaneously transformed immortal endothelial cell line established from the vein of a human umbilical cord, was purchased from the ATCC. It was cultivated in minimum Eagle's medium (MEM) containing 10% FCS. The mouse middle T antigen-expressing endothelioma cell line derived from the thymus, tEnd1 (Williams *et al.*, 1988), was cultivated in MEM containing 10% FCS. The mouse C2 muscle cell line was cultivated as proliferating myoblasts in MEM containing 20% FCS. The cells were allowed to differentiate into myotubes by culture in 2% FCS for 3 days and extraction of the RNA was performed 4 to 6 days later.

Human breast tumor biopsies were collected at the Institut Paoli-Calmette in Marseille (France). They were snap frozen upon surgical removal and stored at -80°C until RNA was extracted.

Northern blot hybridizations

RNAs from cells and adult and embryonic tissues were isolated, fractionated and transferred as described previously (Rosnet *et al.*, 1991). Total RNA was isolated from cells using RNasol (Bioprobe). The Biglow mouse cDNA fragment (Ollendorff *et al.*, 1992, 1994) and the BamHI fragment of the SHPE cDNA clone (Ollendorff *et al.*, 1994; Fig. 1) were used as mouse and human probes respectively after labeling (Feinberg and Vogelstein, 1984). Blots were hybridized as described (Stewart and Walker, 1989). Membranes were rehybridized with a mouse *Gapdh* probe (Galland *et al.*, 1990) to control for the non-degradation and for the loading of RNA samples. Human *GPIIb* was used to control for PMA-induced megakaryocytic differentiation of HEL cells. The probe, a kind gift of M.A. Mouthon (U. 362 Inserm, Villejuif, France), was a 0.6 kb cDNA fragment from plasmid pCRII-GPIIb.

In situ hybridization to mouse embryos

The Biglow cDNA fragment was cloned into Bluescript vector and radiolabeled probes were generated after linearization of the plasmid DNA with appropriate restriction enzymes and *in vitro* transcription, in the presence of ^{35}S-labeled UTP, with the Stratagene RNA transcription kit using either T3 or T7 RNA polymerase, depending on the orientation of the probe. Serial sections were deparaffinized and hybridized with 50-75,000 cpm/µl ^{35}S-labeled sense or antisense cRNA probes, according to deLapeyrière *et al.* (1993). After the washes, the slides were dipped into Kodak NTB-2 nuclear track emulsion diluted 1:1 and autoradiographed for 3 to 4 weeks. After photographic development, the slides were stained with toluidine blue and analyzed using both bright- and dark-field optics on a Zeiss Axiophot microscope.

Acknowledgments

This work was supported by INSERM and ARC. We thank C. Mawas, F. Birg, F. Coulier, J. Jacquemier and A. Sebille for helpful discussions and their advice and encouragement. We are grateful to J. Adélaïde for analyzing tumor RNAs, T. Noguchi for occasional help, J. Sampol for providing the HUVEC cells and M. Djabali for the gift of the genomic mouse library. S.P. was supported by a Fellowship from ARC.

References

BEVILACQUA, M.P., POBER, J.S., WHEELER, M.E., COTRAN, R.S. and GIMBRONE, M.A. Jr. (1985). Interleukin 1 acts on cultured human vascular endothelium to increase the adhesion of polymorphonuclear leukocytes, monocytes, and related leukocyte cell lines. *J. Clin. Invest. 76:* 2003-2011.

CAMPBELL, H.D., SCHIMANSKY, T., CLAUDIANOS, C., OZSARAC, N., KASPRZAK, A.B., COTSELL, J.N., YOUNG, I.G., DE COUET, H.G. and GABOR MIKLOS, G.L. (1993). The *Drosophila melanogaster flightless-I* gene involved in gastrulation and muscle degeneration encodes gelsolin-like and leucine-rich repeat domains and is conserved in *Caenorhabditis elegans* and humans. *Proc. Natl. Acad. Sci. USA 90:* 11386-11390.

CHANG, Z., PRICE, B.D., BOCKHEIM, S., BOEDIGHERMER, M.J., SMITH, R. and LAUGHON, A. (1993). Molecular and genetic characterization of the *Drosophila tartan* gene. *Dev. Biol. 160:* 315-321.

DANGL, J.L. (1995). Pièce de résistance: novel classes of plant disease resistance genes. *Cell 80:* 363-366.

deLAPEYRIÈRE, O., OLLENDORFF, V., PLANCHE, J., OTT, M-O., PIZETTE, S., COULIER, F. and BIRNBAUM, D. (1993). Expression of the *Fgf6* gene is restricted to developing skeletal muscle in the mouse embryo. *Development 118:* 601-611.

FEINBERG, A.P. and VOGELSTEIN, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem. 137:* 266-267.

FISHER, L., TERMINE, J. and YOUNG, M. (1989). Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *J. Biol. Chem. 264:* 4571-4576.

GALLAND, F., STEFANOVA, M., PIRISI, V. and BIRNBAUM, D. (1990). Characterization of a murine glyceraldehyde-3-phosphate dehydrogenase pseudogene. *Biochimie 72:* 759-762.

GAUDRAY, P., SZEPETOWSKI, P., ESCOT, C., BIRNBAUM, D. and THEILLET, C. (1992). DNA amplification at 11q13 in human cancer: from complexity to perplexity. *Mutat. Res. 276:* 317-328.

HASHIMOTO, C., HUDSON, K. and ANDERSON, K. (1988). The *Toll* gene of *Drosophila*, required for dorso-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell 52:* 269-279.

HICKEY, M., HAGEN, F., YAGI, M. and ROTH, G. (1993). Human platelet glycoprotein V: characterization of the polypeptide and the related Ib-V-IX receptor system of adhesive, leucine-rich glycoproteins. *Proc. Natl. Acad. Sci. USA. 90:* 8327-8331.

KATAOKA, T., BROEK, D. and WIGLER, M. (1985). DNA sequence and characterization of the *S. cerevisiae* gene encoding adenylate cyclase. *Cell 43:* 493-505.

KOBE, B. and DEISENHOFER, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biol. Sci. 19:* 415-421.

KOBE, B. and DEISENHOFER, J. (1995). A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature 374:* 183-186.

KRUSIUS, T. and RUOSLAHTI, E. (1986). Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. *Proc. Natl. Acad. Sci. USA 83:* 7683-7687.

LOPEZ, J.A., CHUNG, D., FUJIKAWA, K., HAGEN, F., PAPAYANNOPOULOU, T. and ROTH, G. (1987). Cloning of the α chain of human platelet glycoprotein Ib: a transmembrane protein with homology to leucine-rich α2-glycoprotein. *Proc. Natl. Acad. Sci. USA. 84:* 5615-5619.

MIYAKE, K., YAMASHITA, Y., OGATA, M., SUDO, T. and KIMOTO, M. (1995). RP105, a novel B cell surface molecule implicated in B cell activation, is a member of the Leucine-rich repeat protein family. *J. Immunol. 154:* 3333-3340

MOLLIER, P., BROCHIER, G. and MOROT GAUDRY-TALARMAIN, Y. (1990). The action of notexin from tiger snake venom *(Notechis scutatus scutatus)* on acetylcholine release and compartmentation in synaptosomes from electric organ of *Torpedo marmorata. Toxicon 28:* 1039-1052.

NOSE, A., MAHAJAN, V. and GOODMAN, C. (1992). Connectin: a homophilic cell adhesion molecule expressed on a subset of muscles and the motoneurons that innervate them in *Drosophila. Cell 70:* 553-567.

OLLENDORFF, V., NOGUCHI, T. and BIRNBAUM, D. (1993). Des protéines à motifs riches en leucines définissent une cinquième famille de molécules d'adhérence. *Méd. Sci. 9:* 1102-11O9.

OLLENDORFF, V., NOGUCHI, T., deLAPEYRIERE, O. and BIRNBAUM, D. (1994). The *GARP* gene encodes a new member of the family of leucine-rich repeat-containing proteins. *Cell Growth Differ. 5:* 213-219.

OLLENDORFF, V., SZEPETWOSKI, P., MATTEI, M-G., GAUDRAY, P. and BIRNBAUM, D. (1992). New gene in the homologous human 11q13-q14 and mouse 7F chromosomal regions. *Mammal. Genome 2:* 195-200.

REINKE, R., KRANTZ, D., YEN, D. and ZIPURSKY, L. (1988). Chaoptin, a cell surface glycoprotein required for *Drosophila* photoreceptor cell morphogenesis, contains a repeat motif found in yeast and human. *Cell 52:* 291-301.

ROSNET, O., MARCHETTO, S., deLAPEYRIÈRE, O. and BIRNBAUM, D. (1991). Murine *Flt3*, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1R family. *Oncogene 6:* 1641-1650.

ROTH, G.J. (1991). Developing relationships: arterial platelet adhesion, glycoprotein Ib, and leucine-rich glycoproteins. *Blood 77:* 5-19.

ROTH, G.J. (1992). Platelets and blood vessels: the adhesion event. *Immunol. Today 13:* 100-105.

ROTHBERG, J., JACOBS, R., GOODMAN, C. and ARTAVANIS-TSAKONAS, S. (1990). *Slit:* an extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains. *Genes Dev. 4:* 2169-2187.

SCHNEIDER, R., SCHNEIDER-SCHERZER, E., THURNHER, M., AUER, B. and SCHWEIGER, M. (1988). The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses a novel highly diversified protein superfamily with a common repetitive module. *EMBO J. 7:* 4151-4156.

STEWART, W. and WALKER, C. (1989). Comparison of nylon membranes. *Meth. Mol. Cell. Biol. 2:* 73-76.

SZEPETOWSKI, P., OLLENDORFF, V., GROSGEORGE, J., COURSEAUX, A., BIRNBAUM, D., THEILLET, C. and GAUDRAY, P. (1992). DNA amplification at 11q13.5-q14 in human breast cancer. *Oncogene 7:* 2513-2517.

THEILLET, C., ADELAIDE, J., LOUASON, G., BONNET-DORION, F., JACQUEMIER, J., ADNANE, J., LONGY, M., KATSAROS, D., SISMONDI, P., GAUDRAY, P. and BIRNBAUM, D. (1993). *FGFR1* and *PLAT* genes and DNA amplification at 8p12 in breast and ovarian cancers. *Genes Chromosom. Cancer 7:* 219-226.

WILLIAMS, R.L., COURTNEIDGE, S.A. and WAGNER, E.F. (1988). Embryonic lethalities and endothelial tumors in chimeric mice expressing polyoma virus middle T oncogene. *Cell 52:* 121-131.

Received: March 1996
Accepted for publication: April 1996