Development 117, 1333-1343 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

Isolation and characterisation of murine homologues of the *Drosophila seven in absentia* gene *(sina)*

Neil G. Della[1,2], Paul V. Senior[1] and David D. L. Bowtell[1,*]

[1]Howard Florey Institute of Experimental Physiology and Medicine, and [2]Department of Ophthalmology, University of Melbourne, Parkville, Victoria, 3052, Australia

*Author for correspondence

SUMMARY

The *seven in absentia* gene *(sina)* is required for formation of the R7 photoreceptor cell in the developing eye of *Drosophila melanogaster*. The sina protein contains a putative zinc finger domain and localises to the cell nucleus in *Drosophila*. We report here the identification of a family of genes in the mouse (designated *Siah*) with extensive sequence homology to *Drosophila sina*. The *Siah* genes fall into two main groups: *Siah-1*, which consists of four closely related members, two of which appear to be functional, and *Siah-2*, which contains a single functional member. The predicted Siah proteins show an unusually high degree of conservation with sina over the majority of their lengths, diverging significantly only at their amino terminal ends. The *Siah-1* and *Siah-2* genes are widely expressed at a low level in the embryo and adult. Analysis of *Siah-2* by hybridisation histochemistry shows that it is expressed at a higher level in a restricted number of sites during development, including the olfactory epithelium, retina, forebrain and proliferating cartilage of developing bone. The striking degree of sequence homology observed between the *Drosophila* and murine genes implies strong conservation pressure on the *Siah* genes and suggests that they play a significant role in vertebrate development.

Key words: eye development, zinc finger protein, *Drosophila melanogaster*, mouse development, olfactory epithelium, tyrosine kinase signal transduction

INTRODUCTION

The investigation of mechanisms of development in *Drosophila melanogaster* has provided valuable insights into vertebrate development. Many of the genes discovered and characterised in *Drosophila* have been useful for the isolation of important developmental regulators in vertebrates. The homeobox genes provide a particularly striking example of conservation between *Drosophila* and vertebrates at the levels of gene sequence, chromosomal organisation and expression pattern (Duboule and Dollé, 1989; Graham et al., 1989). Similarly, the paired box motif found in several *Drosophila* genes is highly conserved in the murine *Pax* family (Deutsch et al., 1988; Jostes et al., 1991). Gene ablation by homologous recombination (Chisaka and Capecchi, 1991; Lufkin et al., 1991; Le Mouellic et al., 1992) and the molecular characterisation of certain murine and human developmental defects (Balling et al., 1988; Epstein et al., 1991; Tassabehji et al., 1992; Baldwin et al., 1992) has verified the importance of the *Hox* and *Pax* genes as developmental regulators in vertebrates.

Over the last few years, the *Drosophila* eye has emerged as a useful system in which the development of a precise neural array can be studied at the level of the single cell. Development of the *Drosophila* eye is the result of a cascade of short range inductive events, culminating in the formation of repeating units composed of photoreceptor and non-neuronal cells (Rubin, 1989). A number of the genes that regulate cellular communication and differentiation in the *Drosophila* retina have been cloned and characterised. We were interested to determine whether any homologues of these genes exist in vertebrates and, if so, whether they mediate similar cellular functions. We have concentrated on genes involved in photoreceptor specification in *Drosophila*, including the *seven in absentia* gene *(sina)*, which is necessary for the formation of the R7 photoreceptor cell (Carthew and Rubin, 1990). The sina protein localises to the nucleus and includes a potential metal-binding domain, suggesting that it may be a DNA-binding protein (Carthew and Rubin, 1990). Recent work has shown that the sina protein is required downstream of the tyrosine kinase receptor sevenless and *Drosophila* Ras1 in the specification of the R7 cell (Fortini et al., 1992; Gaul et al., 1992).

Here we describe the isolation and characterisation of a family of genes in the mouse that exhibit a remarkably high degree of sequence homology to *sina*. Like *sina*, these murine homologues display a widespread pattern of expression in both embryonic and adult tissues. In situ hybridisation to RNA shows elevated expression of one of the mouse homologues (*Siah-2*) in several sites, including the olfactory epithelium, forebrain, cartilage and retina. The

extensive conservation between *Drosophila sina* and its murine homologues suggests an important role for these genes in vertebrate development.

MATERIALS AND METHODS

Screening of mouse genomic and cDNA libraries

A genomic DNA library (random bred Swiss, Promega) was probed with a ^{32}P-labelled *Acc*I/*Nde*I *sina* cDNA fragment (Carthew and Rubin, 1990). Duplicate nitrocellulose filters were hybridised in 5× SSC, 5× Denhardt's, 5 mM EDTA, 100 μg/ml herring testis DNA, 0.1% SDS at 65°C for 18 hours then washed in 2× SSC, 0.1% SDS at 50°C and autoradiographed. A single genomic clone was isolated and subsequently used to screen a random bred Swiss E17.5 embryonic eye cDNA library in gt10 (D. Bowtell, unpublished data). Duplicate nitrocellulose filters were hybridised as above, then washed in 2× SSC, 0.1% SDS at 60°C and autoradiographed. Two populations of cDNA were identified based on the intensity of signals obtained and designated *Siah-1* and *Siah-2*. These were present at an approximate frequency of 1:25,000 and 1:100,000 recombinants, respectively. cDNA inserts were subcloned into Bluescript II (Stratagene) for further analysis. Genomic bacteriophage clones used for the characterisation of the number and structure of the *Siah* genes were obtained from an inbred 129/sj mouse genomic DNA library (D3 cell line, Stratagene). These were isolated using *Siah-1A* and *Siah-2* cDNA fragments as probes.

DNA sequence analysis

Double stranded dideoxy chain termination DNA sequencing was performed on cDNA by making nested deletions (Sambrook et al., 1989). The sequence of both strands was obtained. Compressions in G- and C-rich regions were resolved using 7-deaza-dGTP (Mizusawa et al., 1986). Sequence analysis was performed using the University of Wisconsin Genetics Computer Group (GCG) software packages (Devereux et al., 1984). Nucleic acid and protein sequences were compared using the FASTA program (Pearson and Lipman, 1988). An alignment of the predicted amino acid sequences was obtained using the CLUSTAL program (Higgins and Sharp, 1988) and displayed using the program BOXSHADE (obtained from the European Molecular Biology Library server; K. Hoffman, unpublished data). Sequences are lodged with the EMBL/GenBank/DDBJ databases and the nucleotide positions indicated below refer to sequences with the following accession numbers: *Siah-1A*, Z19579; *Siah-1B*, Z19580; *Siah-1C*, Z19582; *Siah-1D,* Z19583 and *Siah-2*, Z19581.

Southern blot analysis

Mouse (Balb/C), human (peripheral leucocyte) and *Drosophila melanogaster* (Oregon R) genomic DNA were prepared by disruption of cells in proteinase K and SDS followed by repeated phenol extractions, as described (Sambrook et al., 1989). Purified DNA was digested overnight with the appropriate restriction enzymes then separated on 0.8% agarose/Tris acetate gels and transferred overnight to Hybond N$^+$ membranes (Amersham) in 0.5 M NaOH, 1.5 M NaCl. Hybridisation with ^{32}P-labelled *Siah-1* or *Siah-2* cDNA probes was carried out as for the nitrocellulose filters above, except that the SDS concentration in the hybridisation solution was increased to 1.0%. A *Siah-1A* 1.19 kb *Sac*I/*Eco*RI fragment (nt 477-1667, comprising part of the coding and 3 untranslated regions) was used to probe mouse genomic DNA (Fig. 3A) and a 1.3 kb *Bam*HI/*Eco*RI fragment (nt 361-1667) from *Siah-1A* was used to probe mouse, *Drosophila* and human genomic DNA (Fig. 3B). A 2 kb *Siah-2* genomic *Bam*HI/*Eco*RI fragment (beginning at nt 830 in the *Siah-2* cDNA sequence and corresponding to the 3 half of the coding region and flanking 3 sequence) was used to probe mouse genomic DNA (Fig. 3A) and a 1.43 kb *Siah-2* cDNA fragment (nt 764-2197, comprising part of the coding region and the 3 untranslated sequence) was used to probe mouse, *Drosophila* and human genomic DNA (Fig. 3B). Washes were performed in 0.2× SSC, 0.3% SDS at 65°C (Fig. 3A) or 2× SSC, 0.3% SDS at 60°C (Fig. 3B) and the filters autoradiographed for 8 hours to 2 days at −70°C in the presence of an intensifying screen.

RNA isolation and northern blot analysis

Adult tissues and embryos from E10.5 to E17.5 were collected from random bred Swiss mice. The gestational ages of embryos were estimated from the midnight preceding the appearance of copulation plugs. Total RNA was isolated from embryonic tissues by disruption in proteinase K and SDS (Gonda et al., 1985) and from adult tissues by disruption in guanidinium thiocyanate followed by extraction with acid phenol (Chomczynski and Sacchi, 1987). Polyadenylated RNA was isolated by oligo-(dT) affinity chromatography. 1-2 μg of polyadenylated RNA from each tissue was separated in formaldehyde agarose gels (Sambrook et al., 1989), transferred overnight to Hybond C-super membranes (Amersham) in 20× SSC and then baked for 2 hours. Filters were pre-hybridised for 2-5 hours in 50% formamide, 5× SSC, 5× Denhardt's, 5 mM EDTA, 100 μg/ml herring testis DNA and 0.5% SDS at 42°C and then hybridised for 12-16 hours. The filters were hybridised with random primed, ^{32}P-labelled 3 untranslated region probes which were specific for transcripts from the *Siah-1* or *Siah-2* genes. The *Siah-1A* probe used was a 676 base pair *Bal*I/*Eco*RI fragment (nt 991-1667), except for the filter shown in Fig. 4D where a probe 3 to the first polyadenylation signal of *Siah-1A* was used (an *Eco*RI fragment, nt 1668-1964). The *Siah-2* probe used was a 402 base pair *Sca*I/*Dra*I fragment (nt 1684-2086). Washes were performed in 0.2× SSC, 0.3% SDS at 65°C and the filters autoradiographed for 2-6 days at −70°C in the presence of an intensifying screen. RNA standards (BRL) were used to estimate the sizes of *Siah* transcripts.

In situ hybridisation to RNA

Random bred Swiss or C57/BJ mouse embryos from E10.5 to E17.5 and the heads of two week old C57/bl male mice were collected and fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) at room temperature for 18-24 hours, stored in 0.5 M sucrose in PBS, then processed and wax embedded by standard methods (Bancroft and Stevens, 1982). Sense and anti-sense RNA transcripts were produced in vitro from a linearised plasmid bearing a 628 base pair fragment of the *Siah-2* gene (nt 1158-1786) using RNA polymerase and either [^{32}P]UTP (3000 Ci/mmole, Bresatec)or [^{35}S]UTP (1500 Ci/mmole, Bresatec). The mean length of the in vitro transcribed RNA was reduced to 100-200 bases by alkaline hydrolysis (Cox et al., 1984). Wax sagittal and transverse sections were cut (5 μm), processed, pre-hybridised, hybridised with the probe and washed essentially as described by Senior et al. (1990). The slides were exposed to Kodak X-Omat film for 12-48 hours to assess the level and tissue distribution of labelling then dipped in liquid photographic emulsion (Ilford K5). The emulsion layer coating the slides was developed after 3 days to 3 weeks. Sections probed with anti-sense and control (sense) probes were developed at the same time. Finally, the slides were lightly stained with haematoxylin and eosin or haematoxylin alone, dehydrated, cleared with xylene and cover-slips applied. They were then examined by bright-field and dark-field illumination.

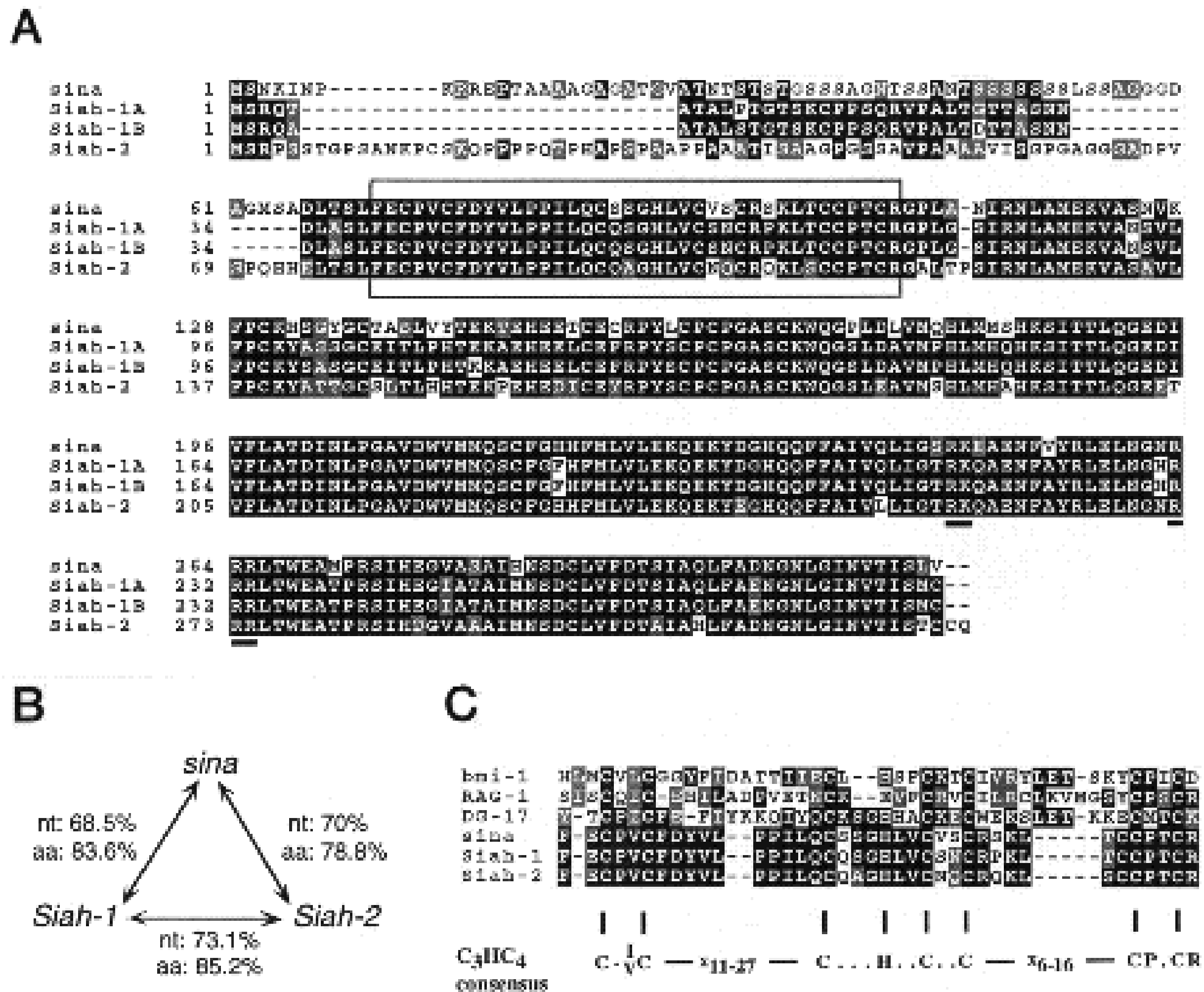

Fig. 1. Relationship between *Drosophila sina* and the murine *Siah-1* and *Siah-2* genes and proteins. (A) Alignment of the predicted amino acid sequence ofsina with that of the Siah-1A, Siah-1B and Siah-2 proteins. The individual proteins are highly homologous over the majority of their lengths, as indicated by the predominance of identical residues (black boxes) and conservative amino acid substitutions (grey boxes). *sina* had previously been shown to contain a 38 amino acid cysteine-rich region with homology to the *Dictyostelium discoideum* protein DG17 (Carthew and Rubin, 1990). This domain, outlined by the box and shown in C below, is also highly conserved in the murine Siah proteins. Two basic clusters (RK and RRR, indicated by solid bars) 14 amino acids apart and located in the C-terminal region of each protein correspond closely to bipartite nuclear localisation sequences identified in several genes, including nucleoplasmin and several steroid receptor genes (Robbins et al., 1991; Silver, 1991). (B) Summary of a three way nucleotide and amino acid alignment of the C-terminal halves of *sina*, *Siah-1A* and *Siah-2*, showing that the *Siah-1* and *Siah-2* genes have diverged from each other almost as much as they have from *sina*. These figures were calculated after excluding the N-terminal ends of the predicted proteins; i.e. the comparison includes residues 66-315 of sina and the corresponding residues for the murine proteins. (C) Comparison of the sina and Siah cysteine-rich domain with bmi-1 (Haupt et al., 1991), RAG-1 (Schatz et al., 1989) and DG17 (Driscoll and Williams, 1987). The cysteine-rich region in the predicted sina, DG17 and Siah proteins is most closely related to the consensus for a recently identified novel class of putative zinc finger protein, the C_3HC_4 class, which includes bmi-1 and RAG-1 (Freemont et al., 1991).

RESULTS

The *Drosophila sina* gene is highly conserved in the mouse

In an attempt to isolate a murine homologue of *sina*, a genomic DNA library was screened under conditions of reduced stringency with a fragment of the *Drosophila sina* gene corresponding to the complete coding region. A weakly hybridising clone was isolated and a small cross-hybridising fragment mapped by Southern analysis. Sequencing of this fragment identified an open reading frame which showed a high degree of homology to the C-terminal half of the predicted sina protein. Subsequent screening of an embryonic eye cDNA library with this genomic fragment identified three types of cDNA clones, designated *Siah-1A*, *Siah-1B* and *Siah-2* (***S****even* ***i****n* ***a****bsentia* ***h****omologue,* pronounced See·a). A comparison of *Siah-1A* and *Siah-1B* cDNA showed that they were very similar, differing in only 25 nucleotides (97% homology) and 6 amino acids (97.8% homology) within their open reading frame. The *Siah-2* cDNA sequence was more divergent, showing 73% nucleotide and 78% amino acid homology with *Siah-*

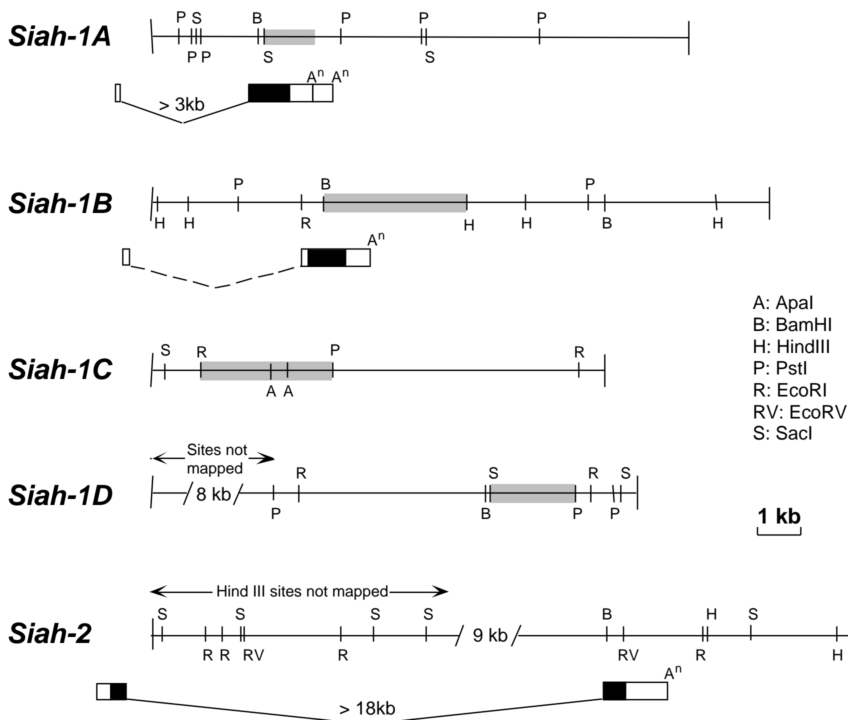

Fig. 2. Schematic representation of the *Drosophila sina* and murine *Siah* genes. Indivdiual genomic clones were isolated for each of the *Siah* genes and mapped with selected restriction enzymes. The region in each of the *Siah-1* phage which cross-hybridised with a *Siah-1A* cDNA probe (*Sac*I/*Eco*RI fragment, comprising the 3 half of the coding region and untranslated sequence) is shown as a stippled box. Comparison of the genomic sequence obtained from each of the stippled regions with *Siah-1A* and *Siah-1B* cDNA identified clones representing the *Siah-1A* and *Siah-1B* genes and the presence of introns within these genes. Solid boxes in the cDNA indicate coding region, open boxes indicate untranslated sequences and polyadenylation sites are indicated by A^n. Both *Siah-1A* and *Siah-1B* contain at least one intron within their 5 untranslated regions, although at different positions. This suggests that gene duplication, rather than retrotransposition of a processed *Siah-1* mRNA, has given rise to *Siah-1A* and *Siah-1B* (Wagner, 1986). This conclusion is consistent with the absence of any direct repeats or polyadenosine residues in either the *Siah-1A* or *Siah-1B* genomic sequences. Comparison of *Siah-1A* and *Siah-1B* cDNA shows a high degree of sequence similarity in their coding and untranslated sequences (92% and 66% nucleotide identity in the 3 and 5 untranslated sequences, respectively), indicating that they have diverged only recently. However, probes obtained from *Siah-1A* and *Siah-1B* genomic DNA flanking the stippled region do not cross-hybridise. This finding is at odds with a model of recent gene duplication of the ancestral *Siah-1* locus to give *Siah-1A* and *Siah-1B*. These conflicting results make the derivation of the *Siah-1A* and *Siah-1B* genes from their ancestral gene unclear. Sequencing of parts of the stippled area in *Siah-1C* and *Siah-1D* identified several frame-shifts and in-frame stop codons in the expected coding region, indicating that *Siah-1C* and *Siah-1D* are pseudogenes. Partial structure of the *Siah-2* gene is also shown. An intron was found within the coding region of *Siah-2* at a position corresponding to amino acid 140. The remaining *Siah-2* exon(s) lay outside the boundaries of the genomic phage obtained and therefore are located at least 18 kb 5 .

1A. (Owing to the high degree of homology between *Siah-1A* and *Siah-1B*, comparisons between *Siah-1A, Siah-2* and *sina* are equally valid for *Siah-1B*). There was a low degree of nucleotide homology between *Siah-1A* and *Siah-2* cDNA outside their coding regions (42% homology with multiple gaps). These findings indicated that the *Siah-1* and *Siah-2* genes had diverged a considerable time ago, whereas *Siah-1A* and *Siah-1B* appeared to have diverged relatively recently.

An alignment of each of the predicted Siah proteins with *Drosophila* sina (Fig. 1A) demonstrated a remarkably high degree of amino acid identity, with the individual proteins diverging significantly only at their N-terminal ends (77% overall amino acid identity between Siah-1A and sina and 66% between Siah-2 and sina). There was no significant homology between *sina* and any of the *Siah* genes outside of their coding regions. The only region of the sina protein that had previously been found to show amino acid similarity to other proteins was a 38 amino acid cysteine-rich domain also found in the product of the DG17 gene of *Dic - tyostelium discoideum* (Driscoll and Williams, 1987; Carthew and Rubin, 1990). This region is also highly conserved in the murine Siah-1 and Siah-2 proteins. The amino acid sequence of this cysteine-rich domain is similar to that of a recently recognised class of zinc finger, the C_3HC_4 motif, present in proteins such as RAG-1 and bmi-1 (Freemont et al., 1991; Haupt et al., 1991; van Lohuizen et al., 1991; see Fig. 1C). Although the cysteine-rich domain is highly conserved between the sina and Siah proteins, it is noteworthy that an equally high degree of conservation exists over the majority of the *Siah* coding regions outside this domain (85-88% identity over 172 amino acids).

The mouse genome contains three, apparently functional, *sina* homologues and two pseudogenes

In order to characterise the structure of the *Siah* genes, we isolated a number of genomic DNA fragments from a bacteriophage library prepared from an inbred mouse strain. The library was screened with separate *Siah-1A* and *Siah-2* coding region probes under conditions where these genes did not cross-hybridise significantly. The resulting clones, together with Southern blot analysis, allowed us to identify genes corresponding to *Siah-1A*, *Siah-1B* and *Siah-2* and two *Siah-1* pseudogenes.

Initial restriction mapping demonstrated that the *Siah-1* clones fell into four groups, rather than the expected two.

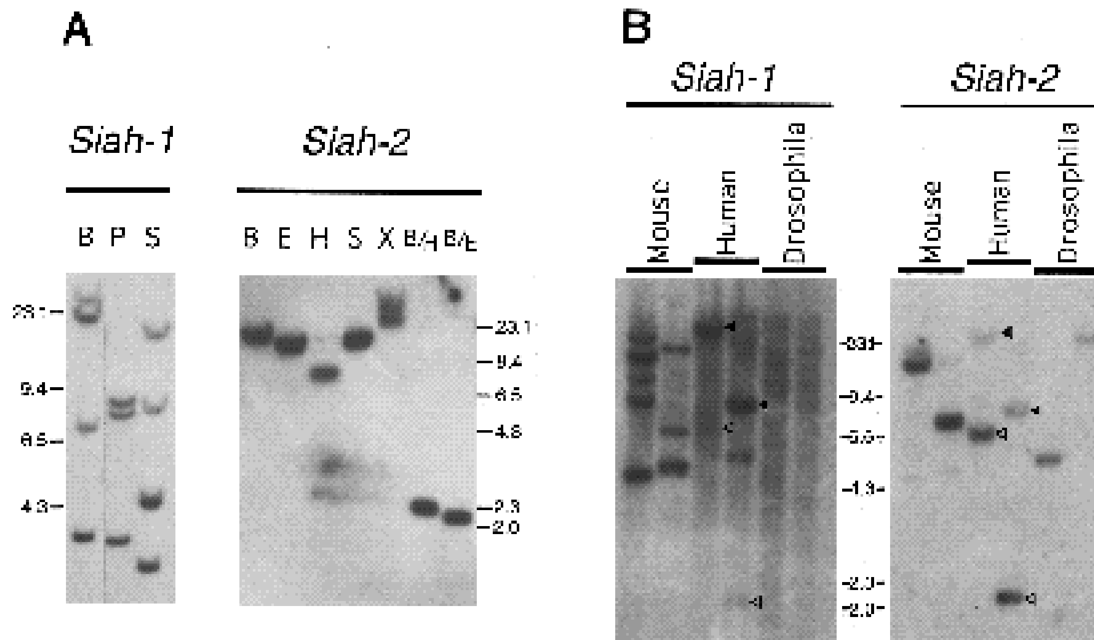

Fig. 3. Genomic Southern analysis of the *Siah* genes in the mouse and human. (A) The presence of four strongly hybridising bands in lanes containing mouse genomic DNA restricted with *Bam*HI or *Sac*I and hybridised with the *Siah-1A* probe was consistent with our analysis of genomic phage (see Results and Fig. 2). Only three bands were apparent in the lane containing *Pst*I restricted genomic DNA due to co-migration of two bands, each of 8.2 kb. Single bands of the predicted sizes were apparent in all lanes of genomic DNA probed with a *Siah-2* coding region fragment. *Siah-1* and *Siah-2* do not cross-hybridise detectably under the conditions used in A. Enzymes were: B, *Bam*HI; P, *Pst*I; S, *Sac*I; E, *Eco*RI; H, *Hin*dIII; X, *Xho*I. Sizes were estimated using lambda DNA restricted with *Hin*dIII. (B) Cross-species genomic Southern blot showing the presence of related genes in the human and *Drosophila*. For each species, the left lane contained genomic DNA digested with *Eco*RI, the right lane with *Hin*dIII. Strongly hybridising bands were detected in each lane containing human DNA when probed for the murine *Siah-1* and *Siah-2* genes, under conditions of reduced stringency. Under these conditions *Siah-1* and *Siah-2* do cross hybridise detectably. Filled arrowheads indicate human *Siah-1* homologues, open arrowheads indicate *Siah-2* homologues. The panels in B were derived from duplicate samples run on a single gel. See Materials and Methods for details of the probes used.

Representatives of each group were mapped more extensively and the region in each phage which hybridised with the *Siah-1A* coding region probe was identified. Probes which lay 5 and 3 of this region were then isolated and hybridised to the array of *Siah-1* clones. Probes from a given phage only hybridised strongly with the other members of its group, as assigned by our initial restriction analysis. This demonstrated that each of the *Siah-1A* hybridsing regions in the individual phage groups were embedded in unique genomic DNA. Subsequent sequencing of genomic fragments from representatives of two of these groups identified sequences that were identical to those of the *Siah-1A* and *Siah-1B* cDNA, respectively. Comparison of the *Siah-1A* cDNA and cloned genomic DNA revealed the presence of an intron immediately 5 of the probable initiation codon (Fig. 2). The *sina* gene contains a single intron in a similar, though not identical, position (Carthew and Rubin, 1990). An intron was also found in the 5 untranslated region of *Siah-1B*, although at a different position to *Siah-1A* (see Fig. 2 legend for further discussion). Sequencing of representatives of the remaining two types of *Siah-1* genomic phage groups identified two distinct pseudogenes. Both contained a number of frame shifts and in-frame stop codons within the expected coding region (data not shown).

Restriction mapping and sequencing of the *Siah-2* hybridising genomic clones indicated that *Siah-2* is a single copy gene in the mouse. Comparison of genomic and cDNA sequences revealed that the *Siah-2* gene contains an intron of at least 18 kb within the coding region, not present in *Siah-1A* or *Siah-1B*. (Fig. 2).

Our estimate of the total number of *Siah-1* and *Siah-2* genes within the mouse genome was confirmed by genomic Southern analysis. Four bands were obtained when restriction enzyme digests of mouse genomic DNA were probed with a single exon fragment which recognised all four *Siah-1* genomic clones (Fig. 3A). Flanking genomic DNA probes specific for each of the above phage types were also used to identify individual bands on Southern blots and their sizes corresponded with those predicted from the analysis of the genomic clones (data not shown). A similar analysis, using a *Siah-2* probe at high stringency, confirmed that *Siah-2* exists in single copy in the mouse genome (Fig. 3A). Genomic Southern analysis of human DNA revealed separate, strongly cross-hybridising bands when probed with mouse *Siah-1A* and *Siah-2* fragments (Fig. 3B). Fewer bands were seen in human than in mouse DNA with the *Siah-1A* probe, suggesting that a less complex *Siah-1* gene family exists in the human. Recently, the partial sequenc-

ing of a large number of randomly isolated, brain derived cDNA identified a fragment of a human *sina* homologue (Adams et al., 1992). Comparison of its available sequence with that of *Siah-1* and *Siah-2* demonstrated nucleotide homologies of 72% and 92%, respectively, over 150 nucleotides (data not shown). Therefore, the fragment of the human gene identified by Adams et al. (1992) appears to be homologous to mouse *Siah-2.*

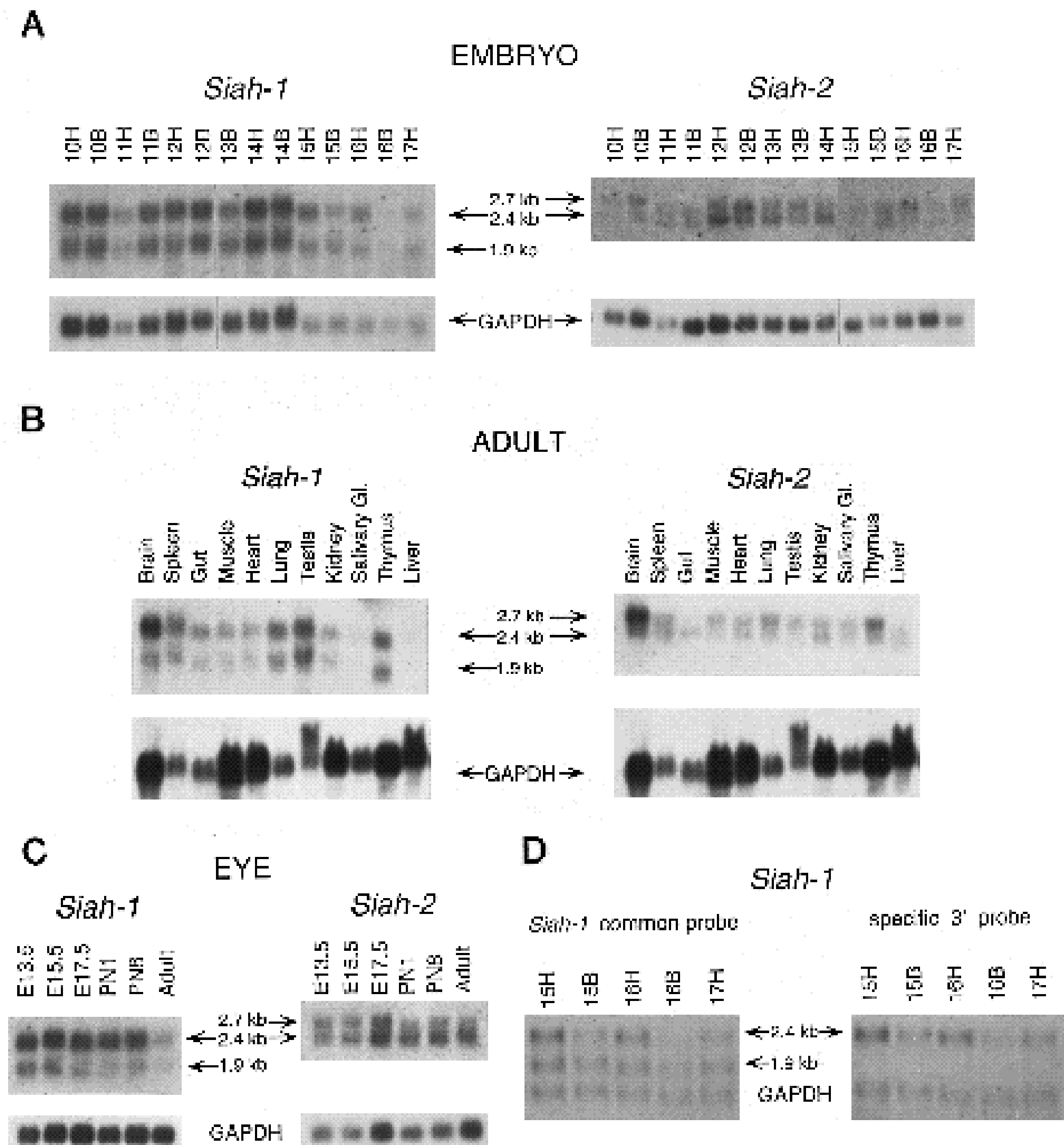

Fig. 4. Northern blot analysis of *Siah-1* and *Siah-2* transcripts in embryonic and adult tissues. (A) Embryonic head and body RNA from E10.5 to E17.5 showing uniform expression of both *Siah-1* and *Siah-2* throughout this period. Two transcripts were detected for each gene. (B) Adult mouse RNA from various tissues. *Siah-1* and *Siah-2* appear to be expressed at relatively uniform levels in adult tissues with the exception of the liver. *Siah-1* appears to be expressed at relatively higher levels than *Siah-2* in most tissues, with the exception of the brain. (C) Developmental pattern of expression of *Siah-1* and *Siah-2* in the mouse eye, beginning at embryonic day 13.5 and continuing through to adulthood. Uniform expression of both genes was apparent throughout this period of eye development, except that *Siah-1* expression appeared to be reduced in adult eyes compared with the other stages examined. (D) Use of a probe 3 to the first polyadenylation signal of the *Siah-1* gene demonstrated that the variation in size of the two *Siah-1* transcripts was associated with differential polyadenylation. See Materials and Methods for details of the probes used. 10H, E10.5 head;10B, E10.5 body, etc.; PN1, Postnatal day 1 etc.

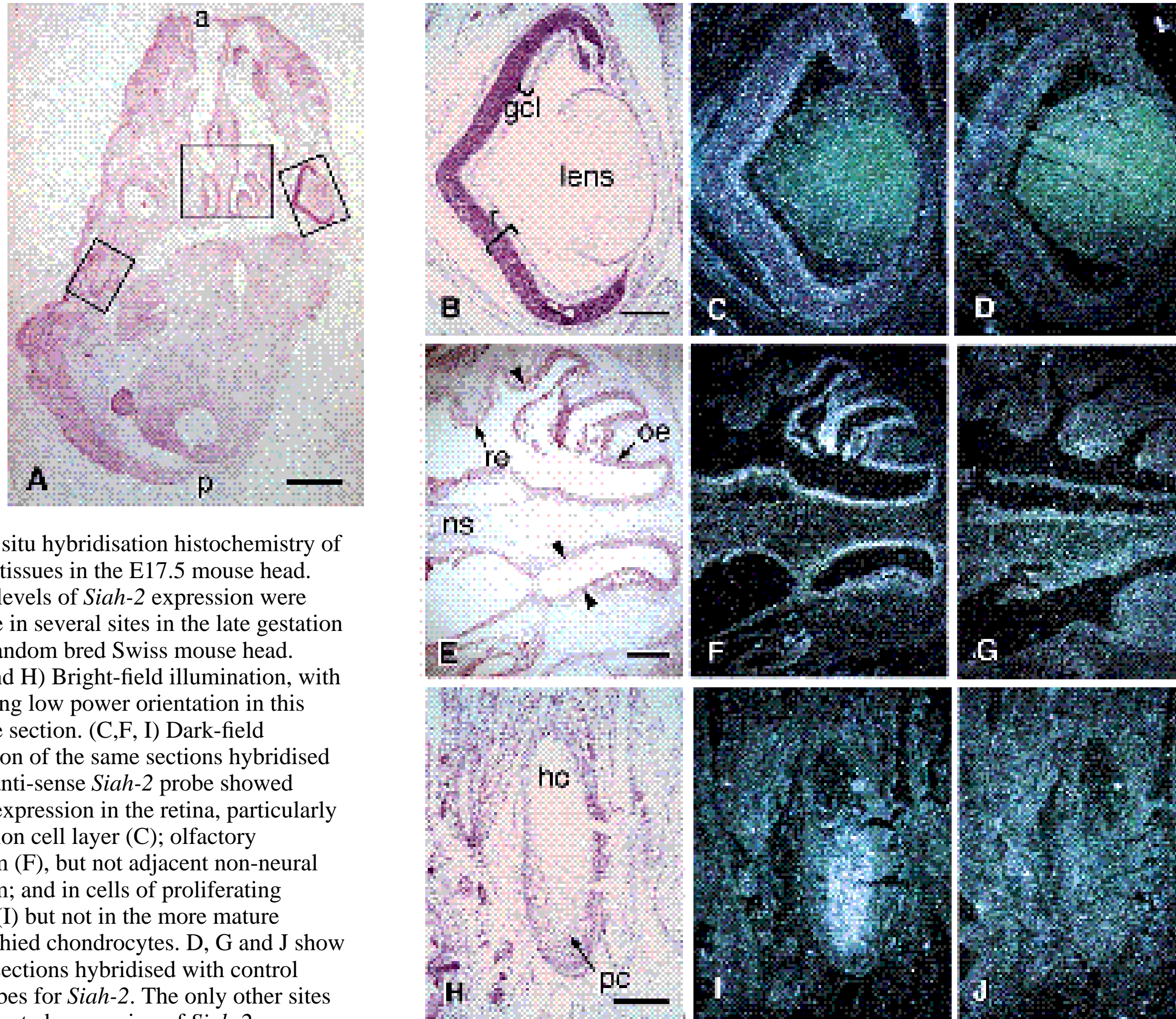

Fig. 5. In situ hybridisation histochemistry of *Siah-2* to tissues in the E17.5 mouse head. Elevated levels of *Siah-2* expression were detectable in several sites in the late gestation (E17.5) random bred Swiss mouse head. (A,B,E and H) Bright-field illumination, with A providing low power orientation in this transverse section. (C,F, I) Dark-field illumination of the same sections hybridised with the anti-sense *Siah-2* probe showed elevated expression in the retina, particularly the ganglion cell layer (C); olfactory epithelium (F), but not adjacent non-neural epithelium; and in cells of proliferating cartilage (I) but not in the more mature hypertrophied chondrocytes. D, G and J show adjacent sections hybridised with control sense probes for *Siah-2*. The only other sites where elevated expression of *Siah-2* was detected in late gestation embryos were the forebrain and gonad (data not shown). In E, arrowheads indicate the junction between olfactory and respiratory epithelium, and by comparison with F, demonstrate that the area of enhanced *Siah-2* expression was limited to the neuronal epithelium. Expression of *Siah-2* in the olfactory epithelium was located over the sensory neurons, however the level of resolution of the technique was insufficient to determine whether the support or basal cells also expressed *Siah-2* at an elevated level. a, anterior; gcl, ganglion cell layer; hc, hypertrophied chondrocytes; ns, nasal septum; oe, olfactory epithelium; p, posterior; pc, proliferating cartilage cells; r, retina; re, respiratory epithelium. Bars indicate 500 μm in A and 200 μm in the remaining panels.

The *Siah* genes are expressed widely at a low level in embryonic and adult mouse tissues but are expressed at a higher level in a restricted number of sites

The expression pattern of the murine *Siah* genes was initially examined by northern analysis of various embryonic stages and adult tissues using probes specific for all members of the *Siah-1* group or for *Siah-2*. The expression of both *Siah-1* and *Siah-2* was approximately equal in mouse embryos isolated from embryonic day 10.5 (E10.5) to E17.5 (Fig. 4A). Widespread and relatively uniform expression of *Siah-1* and *Siah-2* was also seen in a range of adult tissues (Fig. 4B). Both genes were expressed throughout the development of the eye as judged by northern analysis of a series of mouse eyes isolated from E13.5 to adulthood (Fig. 4C). The expression of *Siah-1* appeared lower in the adult mouse eye than at earlier time points, when compared to the GAPDH control probe. Northern analysis detected two transcripts for each gene. In the case of *Siah-1,* the difference in transcript size could be accounted for by alternate polyadenylation, as shown by the use of a probe 3 to the first polyadenylation site in *Siah-1A* (Figs 2, 4D).

Further analysis of *Siah* expression was performed by in situ hybridisation and revealed a level of complexity that was not apparent by northern analysis. Elevated expression of *Siah-2*, relative to surrounding tissue, was reproducibly detected in late gestation (E17.5) mouse embryos in the following sites: the developing retina, particularly the differentiating ganglion cell layer (Fig. 5C); the olfactory epithelium (Fig. 5F); areas of proliferating cartilage destined to later undergo ossification, including the base of the skull, the vertebral bodies and cartilage in the region of the temporo-mandibular joint (Fig. 5I); developing neurons in the forebrain; and germ cells in the developing gonad (data not shown). The olfactory epithelium, which undergoes neural differentiation, showed the highest level of *Siah-2* expres-

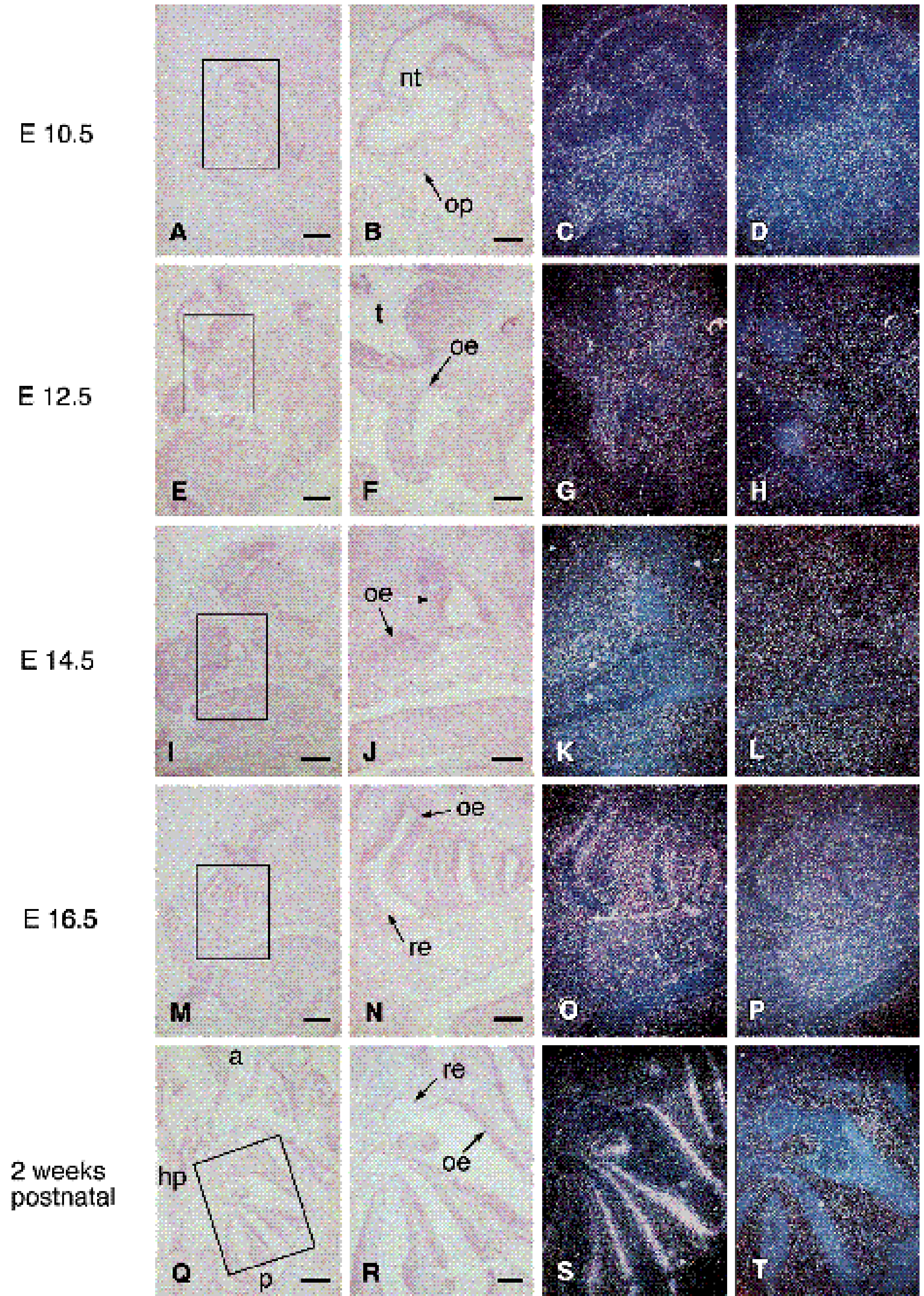

Fig. 6

Fig. 6. Time course of elevated *Siah-2* expression in the olfactory epithelium of the C57/bl mouse. Elevated expression of *Siah-2* in the olfactory epithelium, relative to surrounding tissue was first reproducibly detected at E14.5 (K), increased during the latter half of gestation (O, E16.5) and was particularly pronounced in a two week old animal (S). The E12.5 section appears to show a slight increase of *Siah-2* expression in the general area of the developing olfactory epithelium but this was not seen reproducibly. For each time point, two bright-field views of the same sagittal section are shown, one at low power for orientation and a second to display the olfactory epithelium (oe). This is followed in each case by two dark-field views, the first of the same section hybridised with the anti-sense *Siah-2* probe and the second of an adjacent section hybridised with the sense control probe. All sections shown here were hybridised in the same batch and so the various levels of expression at different times are directly comparable. Sections were counterstained with haematoxylin very lightly, to avoid quenching of the dark-field signal in areas with a high concentration of nuclei. a, anterior; hp, hard palate; nt, neural tube; oe, olfactory epithelium; op, olfactory placode; p, posterior; re, respiratory epithelium; t, telencephalon. Bars indicate 500µm in the low power bright-field views (A, E, I, M, Q) and 200µm in all other panels.

sion. Analysis of *Siah-1* expression was hampered by an inability to obtain probes that were specific for the individual *Siah-1* genes or pseudogenes and also by a relatively high level of uniform hybridisation obtained with the variety of probes tested (data not shown). We therefore concentrated further analysis on *Siah-2*.

The timing of *Siah-2* expression was analysed in embryos obtained from E10.5 to E17.5 and in the heads of two week old mice (Fig. 6). Saggital sections of whole embryos revealed a uniform low level of expression at E10.5 (Fig. 6C). Elevated *Siah-2* expression in the olfactory epithelium became apparent between E12.5 and E14.5 and increased in this site towards the end of gestation (Fig. 6C,G,K,O). A very high level of *Siah-2* expression was observed over the sensory neurons in the olfactory epithelium of two week old mice (Fig. 6S), compared with other tissues in the head and with that seen at E17.5. Elevated expression of *Siah-2* in the retina, forebrain and cartilage also became apparent during the latter stages of gestation. However, the overall level of expression in these tissues was lower than the olfactory epithelium, making it difficult to determine when elevated expression of *Siah-2* first began in these sites. The level of *Siah-2* expression in the retina and forebrain at two weeks of age was less than at E17.5 (data not shown).

DISCUSSION

The mouse genome contains a family of highly conserved *sina* homologues

The *sina* gene is one of a number of genes involved in *Drosophila* eye development for which we have sought mammalian counterparts. *sina* was first identified in *Drosophila* during a genetic screen for viable recessive mutations affecting eye development (Carthew and Rubin, 1990) and is required in a cell autonomous manner for the specification of the R7 photoreceptor. Mutations in *sina* result in the R7 precursor cell adopting a non-neuronal cell fate. Other phenotypic features include partially penetrant effects on the differentiation of other photoreceptor cells, disrupted adult sensory bristle formation and lethargic adult behaviour, sub-viability and infertility. Although the onset of the *sina* phenotype occurs late in *Drosophila* development, the gene is expressed widely in embryos as well as in larval, pupal and adult tissues (Carthew and Rubin, 1990).

We have shown that the mouse genome contains a family of three, apparently functional, genes that are highly related to *sina*. The degree of amino acid homology between the predicted proteins is unusually high compared with many of the other vertebrate homologues of *Drosophila* genes characterised to date (for comparison see review by Lobe and Gruss, 1989). The high degree of conservation between the individual Siah proteins, and between these proteins and sina, indicates that there are very strong constraints on the amino acid sequence required for sina/Siah function. In particular, the C-terminal halves of all four proteins show a striking degree of homology (85-88% identity over 172 amino acids, Fig. 1). It is noteworthy that all mutations in previously identified mutant alleles of *sina* map to the C-terminal half of the sina protein (Carthew and Rubin, 1990). Three of the alleles involve mis-sense mutations, altering the predicted amino acid sequence at residues that are completely conserved between the predicted sina and Siah proteins. The high degree of conservation of the C-terminal half of these proteins may reflect the presence of a separate functional domain, distinct from the more N-terminal cysteine-rich region noted previously (Carthew and Rubin, 1990).

The high degree of amino acid sequence identity between the *Drosophila* sina protein and its vertebrate counterparts suggests that the cellular components that interact with sina may be conserved in vertebrates and that the biochemical pathways involving sina and the Siah proteins may be similar. Recent data indicate that sina may be downstream of Ras in the sevenless tyrosine kinase signalling pathway in *Drosophila*. (Fortini et al., 1992; Gaul et al., 1992). Other studies have shown that there is a significant degree of conservation between the more proximal components of tyrosine kinase receptor signalling in *Drosophila* and vertebrates (reviewed by Williams, 1992). Given the high degree of homology between *sina* and the *Siah* genes, it seems possible that the Siah proteins may also play a role in signal transduction, downstream of tyrosine kinases and Ras in vertebrate cells.

Siah expression in the mouse

Our northern blot analyses demonstrated that the *Siah* genes are widely expressed throughout embryonic development and in adult tissues of the mouse. Although a widespread pattern of expression of the *Siah* genes may suggest that they perform general functions, it is of interest to compare their pattern of expression with *sina*. Both northern blot and immunocytochemical analyses show that *sina* is widely expressed throughout *Drosophila* development, yet the phenotype of the *sina* mutant is relatively restricted and occurs late in the *Drosophila* life cycle. (Carthew and Rubin, 1990). It is possible that an analogous situation occurs with

the *Siah* genes in the mouse. If so, then tissues which have substantially elevated levels of *Siah* expression may represent sites of critical gene function. We have identified a number of sites in late gestation mouse embryos that showed enhanced *Siah-2* gene expression, including the olfactory epithelium, retina, forebrain, gonad and areas of proliferating cartilage. Of these sites, *Siah-2* expression was particularly pronounced in the olfactory epithelium.

Olfactory neurons first appear at E10, establish synaptic contact with the olfactory bulb at E13 and begin to undergo final maturation at E17 (Cushieri and Bannister, 1975). They are unique among the neurons of vertebrates in that they are continually replaced throughout life from a pool of stem cells in the basal epithelial layer (Farbman, 1990). Elevated expression of *Siah-2* in the olfactory epithelium was first apparent between E12.5 and E14.5 and increased through gestation. A relatively high level of expression was seen in two week old mice over the sensory neurons of the olfactory epithelium. Interestingly, the pattern of *Siah-2* expression appeared to parallel the previously described expression pattern of the olfactory sensory neuron marker, OMP, (Monti Graziadei et al., 1980). This protein is first detectable in a few cells in the murine olfactory epithelium at E14, increases slowly through late gestation and early postnatal life and then dramatically increases between postnatal day 10 and 19 (Monti Graziadei et al., 1980). OMP has been cloned (Rogers et al., 1987) and is a 19×10^3 M_r cytoplasmic protein which is expressed predominantly in mature olfactory neurons and not in the neural precursor basal cells. A more detailed examination of the relationship between *Siah-2* and OMP expression is in progress.

The expression of *Siah-2* was also elevated in the retina, particularly in the emerging ganglion cell layer in late gestation mouse embryos. Unlike the olfactory epithelium, elevated *Siah-2* expression in the retina appeared to decrease with time. Whether the elevated expression of *Siah-2* in the retina reflects an important role in retinal development and the possibililty of shared evolutionary origins of the vertebrate and insect eye is not clear. Current experiments in mice aimed at disrupting the *Siah* genes by homologous recombination are at an advanced stage and should help to determine whether the *Siah* genes play a role in the retina, olfactory epithelium or any of the other sites where elevated expression was observed. These experiments should also test the degree of functional redundancy between the individual *Siah* genes in the mouse.

We wish to thank M. Kafali for excellent technical assistance and also members of the Walter and Eliza Hall Institute and ANGIS computer centres and R. Flegg for help with sequence analysis. We are grateful to F. Beck and B. Key for help with the interpretation of hybridisation histochemistry results and to P. Bartlett, F. Beck, R. Carthew, J. Coghlan and S. Cory for helpful comments on the manuscript. In addition, we thank R. Carthew for initially providing the *sina* gene fragment and for many useful discussions involving unpublished results and observations. We also thank J. Haralambidis, S. Khoury and L. Lagniton for their synthesis of oligonucleotides. The Howard Florey Institute is supported by an Institute block grant from the National Health and Medical Research Council of Australia. N. G. D. is the recipient of a NH & MRC Medical Postgraduate Research Scholarship. P. V. S. is a NH & MRC R. Douglas Wright Fellow. D. D. L. B. is a Wellcome Trust Senior Research Fellow in Medical Research in Australia.

REFERENCES

Adams, M. D., Dubnick, M., Kerlavage, A. R., Moreno, R., Kelley, J. M., Utterback, T. R., Nagle, J. W., Fields, C. and Venter, J. C. (1992). Sequence identification of 2,375 human brain genes. *Nature* **355**, 632-634.

Baldwin, C. T., Hoth, C. F., Amos, J. E., da-Silva, E. O. and Milunsky, A. (1992). An exonic mutation in the *HuP2* paired domain gene causes Waardenburg's syndrome. *Nature* **355**, 637-638.

Balling, R., Deutsch, U. and Gruss, P. (1988). *undulated*, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of *Pax 1*. *Cell* **55**, 531-535.

Bancroft, J. D. and Stevens, A. (1982). *Theory and Practice of Histological Techniques*. Second edition. New York: Churchill-Livingstone.

Carthew, R. W. and Rubin, G. M. (1990). *seven in absentia*, a gene required for specification of R7 cell fate in the *Drosophila* eye. *Cell* **63**, 561-577.

Chisaka, O. and Capecchi, M. R. (1991). Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene *hox-1.5*. *Science* **350**, 473-479.

Chomczynski, P. and Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem.* **162**, 156-159.

Cox, K. H., Deleon, D. V., Angerer, L. M. and Angerer, R. C. (1984). Detection of mRNAs in sea urchin embryos by in situ hybridisation using asymmetric RNA probes. *Dev. Biol.* **101**, 485-502.

Cuschieri, A. and Bannister, L. H. (1975). The development of the olfactory mucosa in the mouse; light microscopy. *J. Anat.* **119**, 277-286.

Deutsch, U., Dressler, G. R. and Gruss, P. (1988). *Pax 1*, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. *Cell* **53**, 617-625.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**, 387-395.

Driscoll, D. M. and Williams, J. G. (1987). Two divergently transcribed genes of *Dictyostelium discoideum* are cyclic AMP-inducible and coregulated during development. *Mol. Cell. Biol.* **7**, 4482-4489.

Duboule, D. and Dollé, P. (1989). The structural and functional organization of the murine *HOX* family resembles that of *Drosophila* homeotic genes. *EMBO J.* **8**, 1497-1505.

Epstein, D. J., Vekemans, M. and Gros, P. (1991). *splotch* (*Sp*^{2H}), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of *Pax-3*. *Cell* **67**, 767-774.

Farbman, A. I. (1990). Olfactory neurogenesis: genetic or environmental controls? *Trends Neurosci.* **13**, 362-365.

Fortini, M. E., Simon, M. A. and Rubin, G. M. (1992). Signaling by the *sevenless* protein tyrosine kinase is mimicked by Ras1 activation. *Nature* **355**, 559-561.

Freemont, P. S., Hanson, I. M. and Trowsdale, J. (1991). A novel cysteine-rich sequence motif. *Cell* **64**, 483-484.

Gaul, U., Mardon, G. and Rubin, G. M. (1992). A putative Ras GTPase activating protein acts as a negative regulator of signaling by the sevenless receptor tyrosine kinase. *Cell* **68**, 1007-1019.

Gonda, T. J., Gough, N. M., Dunn, A. R. and de Blaquiere, J. (1985). Nucleotide sequence of cDNA clones of the murine *myb* proto-oncogene. *EMBO J.* **4**, 2003-2008.

Graham, A., Papalopulu, N. and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression. *Cell* **57**, 367-378.

Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P. and Adams, J. A. (1991). Novel zinc finger gene implicated as *myc* collaborator by retrovirally accelerated lymphomagenesis in Eμ-*myc* transgenic mice. *Cell* **65**, 753-763.

Higgins, D. G. and Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237-244.

Jostes, B., Walther, C. and Gruss, P. (1991). The murine paired box gene,

Pax 7, is expressed specifically during the development of the nervous and muscular system. *Mech. Dev*. **33**, 27-38.

Le Mouellic, H., Lallemand, Y. and Brûlet, P. (1992). Homeosis in the mouse induced by a null mutation in the *Hox-3.1* gene. *Cell* **69**, 251-264.

Lobe, C. G. and Gruss, P. (1989).Mouse versions of fly developmental control genes: Legitimate or illegitimate relatives ? *New Biol.* **1**, 9-18.

Lufkin, T., Dierich, A., LeMeur, M., Mark, M. and Chambon, P. (1991). Disruption of the *Hox-1.6* homeobox gene results in defects in a region corresponding to its rostral domain of expression. *Cell* **66**, 1105-1119.

Mizusawa, S., Nishimura, S. and Seela, F. (1986). Improvement of the dideoxy chain termination method of cDNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. *Nucl. Acids Res*. **14**, 1319-1324.

Monti Graziadei, G. A., Stanley, R. S. and Graziadei, P. P. C. (1980). The Olfactory Marker Protein in the olfactory system of the mouse during development. *Neurosci*. **5**,1239-1252.

Pearson, W. R. and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natn. Acad. Sci. USA* **85**, 2444-2448.

Robbins, J., Dilworth, S.M., Laskey, R.A. and Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615-623.

Rogers, K.E., Dasgupta, P., Gubler, U., Grillo, M., Khew-Goodall, Y.S. and Margolis, F.L. (1987). Molecular cloning and sequencing of a cDNA for olfactory marker protein. *Proc. natn. Acad. Sci. USA* **84**, 1704-1708.

Rubin, G. M. (1989). Development of the *Drosophila* retina: inductive events studied at single cell resolution. *Cell* **57**, 519-520.

Sambrook, J., Fritsch, F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* Second edition. New York: Cold Spring Harbor Laboratory Publishers.

Schatz, D. G., Oettinger, M. A. and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. *Cell* **59**, 1035-48.

Senior, P. V., Byrne, B., Brammar, W. J. and Beck, F. (1990). Expression of the IGFII/mannose-6-phosphate receptor mRNA and protein in the developing rat. *Development* **109**, 67-73.

Silver, P. A. (1991). How proteins enter the nucleus. *Cell* **64**, 489-497.

Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P. and Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the *Pax-3* paired box gene. *Nature* **355**, 635-636.

Wagner, M. (1986). A consideration of the orgins of processed pseudogenes. *Trends Genet.* **2**, 134-137.

Williams, L. T. (1992). Missing links between receptors and Ras. *Current Biol*. **2**, 601-603.

van Lohuizen, M., Verbeek, S., Scheijen, B., Wientjens, E., van der Gulden, H. and Berns, A. (1991). Identification of cooperating oncogenes in Eμ-*myc* transgenic mice by provirus tagging. *Cell* **65**, 737-752.

(Accepted 28 December 1992)