Mammalian Genome 8, 479–485 (1997).

Mammalian Genome

© Springer-Verlag New York Inc. 1997

Eyes absent: a gene family found in several metazoan phyla

Melinda K. Duncan,[1] Lidia Kos,[2] Nancy A. Jenkins,[3] Debra J. Gilbert,[3] Neal G. Copeland,[3] Stanislav I. Tomarev[1]

[1]National Eye Institute, Laboratory of Molecular and Developmental Biology, Building 6, Room 203, 6 Center Drive, MSC 2730, Bethesda, Maryland 20892, USA
[2]Laboratory of Genetic Disease Research, National Center for Human Genome Research, Bethesda, Maryland 20892, USA
[3]Mammalian Genetics Laboratory, ABL Basic Research Program NCI/FCRDC, Building 539, Room 229, Frederick, Maryland 21702, USA

Received: 23 November 1996 / Accepted: 25 February 1997

Abstract. Genes related to the *Drosophila* eyes absent gene were identified in vertebrates (mouse and human), mollusks (squid), and nematodes (*C. elegans*). Proteins encoded by these genes consist of conserved C-terminal and variable N-terminal domains. In the conserved 271-amino acid C-terminal region, *Drosophila* and vertebrate proteins are 65–67% identical. A vertebrate homolog of eyes absent, designated *Eya2,* was mapped to Chromosome (Chr) 2 in the mouse and to Chr 20q13.1 in human. *Eya2* shows a dynamic pattern of expression during development. In the mouse, expression of *Eya2* was first detected in 8.5-day embryos in the region of head ectoderm fated to become the forebrain. At later stages of development, *Eya2* is expressed in the olfactory placode and in a variety of neural crest derivatives. In the eye, expression of *Eya2* was first detected after formation of the lens vesicle. At day 17.5, the highest level of *Eya2* mRNA was observed in primary lens fibers. Low levels of *Eya2* expression was detected in retina, sclera, and cornea. By postnatal day 10, *Eya2* was expressed in secondary lens fibers, cornea, and retina. Although *Eya2* is expressed relatively late in eye development, it belongs to the growing list of factors that may be essential for eye development across metazoan phyla. Like members of the *Pax-6* gene family, eyes absent gene family members were probably first involved in functions not related to vision, with recruitment for visual system formation and function occurring later.

Introduction

Most metazoan phyla include species with the ability to detect the difference between light and dark and make behavioral responses to this information. In many cases, these light-detecting systems are composed of only a simple eye spot. A more elaborate optical system can be found in only six of the major animal phyla; however, they constitute about 96% of the known metazoan species (Land and Fernald 1992). While the eyes in these phyla often have drastically different structures (Land 1988), it has long been known that vision in divergent phyla was mediated by conformational changes in proteins related to opsin (see Yarfitz and Hurley 1994 for review). More recently, it was discovered that animals from a wide variety of phyla express orthologs of the *Pax-6/* eyeless gene during the development of their light-sensing organs, brain, and other sensory structures (Walther and Gruss 1991; Quiring et al. 1994; Loosli et al. 1996; Czerny and Busslinger 1995) and that mutations in the *Pax-6* gene disrupt the development of these organs (Quiring et al. 1994; Hill et al. 1991; Hanson and van Heyningen 1995; Glaser et al., 1994). Moreover, targeted expression of *Drosophila* eyeless (*ey*) and mouse and squid *Pax-6* in different imaginal disk primordia of *Drosophila* results in ectopic compound eyes (Halder et al., 1995b; Tomarev et al. 1997). These observations have led to the proposal that *Pax-6*-related genes are necessary for eye formation throughout the animal kingdom and that the molecular mechanisms responsible for early eye development may be similar in different systematic groups (Quiring et al. 1994; Halder et al. 1995a, 1995b).

It has been estimated that as many as 2500 genes may be essential for *Drosophila* eye morphogenesis (Halder et al., 1995a). Some genes acting downstream of *Pax-6* in the developing eye may also be homologous in different systematic groups, and several candidate genes were identified. Sine oculis/*Six3* (Cheyette et al. 1994; Oliver et al., 1995a), prospero/Prox 1 (Doe et al. 1991; Tomarev et al. 1996) and decapentaplegic/*BMP-7* (Heberlein et al. 1993; Dudley et al. 1995) are examples of such genes. The expression of none of these genes is eye-specific. The eyes absent (*eya*) mutation of *Drosophila* causes increased apoptosis of precursor cells anterior to the morphogenetic furrow of the eye portion of the eye-antennal imaginal disk. This results in a decrease in the overall number of cells available to differentiate into ommatidia, but does not specifically deplete any particular cell type (Bonini et al. 1993). The mutation causing the *eya* phenotype resides in a regulatory element necessary to direct the expression of a novel nuclear protein to cells anterior to the morphogenetic furrow (Bonini et al. 1993; Leiserson et al. 1994). A complete disruption of the *eya* (the clift phenotype) results in early embryonic lethality, possibly due to defects in head development (Bonini et al. 1993; Nusslein-Vollard et al. 1984). Since expression studies of *eya* indicated that it acts downstream of *ey* and upstream of sine oculis (*so*) in the eye-antennal imaginal disc (Cheyette et al. 1994; Bonini et al., 1993, 1995), we proposed that relatives of *eya* playing a role in eye and head development would be found in other metazoan phyla. We have subsequently identified cognates of *eya* in mice, humans, nematodes, and cephalopods and determined the expression pattern of the mouse *eya* homolog, *Eya2,* during embryonic development. The expression pattern of *Eya2* is consistent with its possible role in vertebrate sensory systems and head development.

Materials and methods

Cloning of the cDNAs for human and mouse Eya2. The peptide sequence for *Drosophila* eya (18) was used to search the expressed sequence tag (EST) database subset of Genbank. Clones exhibiting significant similarity to eya (GenBank accession numbers T80220, R72695, R20007, R76303) were purchased from Research Genetics (Huntsville, Ala.) and double-strand sequenced with Sequenase (Amersham, Cleveland, Ohio). The complete cDNA sequence was obtained by 5′ rapid amplification of cDNA ends (RACE) PCR of Marathon RACE Ready 19–23 fetal week human brain cDNA purchased from Clontech (Palo Alto, Calif.) following the manufacturer's instructions with gene-specific primers 3370

Correspondence to: S.I. Tomarev

Data deposition: The sequences reported in this paper have been deposited in the GenBank data base with the accession numbers U71207 and U71208, and the map portion for mouse *eya2* has been submitted to the Mouse Genome Database, accession number MGD CREX-695.

(GSP 1, 5′-TGAGGTGCTGAAGCTGGAGCCATA-3′, positions 357–380 in the final sequence) and 3371 (GSP2, 5′-TGAGAATCCACTCTG-GCCAG-3′, positions 334–356 in the final sequence). The resulting PCR product was cloned and sequenced. A contiguous DNA fragment containing the entire open reading frame (ORF) for human Eya2 was obtained by PCR from Marathon RACE Ready 19–23 fetal week human brain cDNA, cloned into a plasmid vector and resequenced.

The cDNA insert of human R20007 clone was radiolabeled with [^{32}P]dCTP by random priming and used to screen 50,000 plaques of a λEXlox 10.5 days post coitum (d.p.c) mouse embryo cDNA library (Novagen, Madison, Wis.). Two positive clones were converted to pEXlox (+) plasmids by the manufacturer's instructions and sequenced with the Fidelity DNA sequencing kit (Oncor, Gaithersburg, Md.). Additional sequences were obtained by PCR of 11-day mouse embryonic Marathon RACE ready cDNA (Clontech) (3′ primer 2864, 5′-ATTGACACAGTTGGGTCGAGA-3′, positions 1364–1384 of the mouse *Eya2* cDNA sequence and a 5′ primer 3073, 5′-CAGCTTGAACCATTCCCCTGGCCAGAGTG; nucleotides 411–440 of the human eyes *Eya2* cDNA sequence). This PCR product was cloned and sequenced. The remainder of the open reading frame was isolated by 5′RACE PCR of 11-day mouse embryonic Marathon RACE Ready cDNA (Clontech) with mouse *Eya2* gene-specific primer 3356 (5′-TCGGTCCGTAGCTCAGGAACCC-3′, positions 398–419 in the mouse sequence). This PCR product was cloned and sequenced with Fidelity. The complete ORF was then obtained by PCR from 11-day embryonic mouse Marathon Race Ready cDNA, cloned and resequenced.

Identification of an eya *family member in cephalopods and nematodes.* Embryonic squid (*Loligo opalescens*) were collected at the Hopkins Marine Station (Stanford University, Pacific Grove, Calif.), and total RNA was prepared with RNazol (Teltest, Friendswood, Tex.). Poly(A)$^{+}$RNA was prepared, annealed with random primers, and used for cDNA synthesis. This cDNA was used as a template for PCR with degenerate oligonucleotides derived from the sequence of human Eya2 and *Drosophila eya* with the addition of *Hin*dIII sites (underlined) (5′-primer, 5′-TGGTAAGCTTGAGCG(A/T)GTGTT(C/T)(G/C)T(G/C)TGGGA(C/T)(C/T)T(G/C)GA(C/T)GA-3′, positions 804–832 in the human Eya2 cDNA; 3′-primer, 5′-ACCTAAGCTTGGTA(A/G)CGGAA(G/T)GC(C/T)A(A/G)(C/T)TT(C/T)C(G/T)CATCCA-3′, positions 1128–1155) using the following PCR conditions: 94°C 2 min, followed by 30 cycles of 94°C 1 min, 55°C 1.5 min, 72°C 1 min. The PCR product was digested with *Hin*dIII, cloned into pBlueScript (Stratagene), and sequenced.

The peptide sequence of human Eya2 was used to search the *C. elegans* genomic database with the Sanger Center Blast server (Oxford University, Oxford, UK). Genomic cosmid C49A1, localized to the X Chr of *C. elegans,* was identified. The deduced amino acid sequence of the *C. elegans* eya-related protein was then used to search the *C. elegans* EST database by use of the Sanger Center Blast server. EST yk166c6.3 isolated from a N2 embryonic *C. elegans* cDNA library and EST yk74a9.5 were identified and their sequence was obtained from ACeDB (A *C. elegans* Data Base, USDA, Beltsville Md., USA). These nucleotide sequences were used to deduce a partial amino acid sequence for a *C. elegans* eya-related family member (see Fig. 1).

Determination of the chromosomal localization of human and mouse Eya2 *genes.* Specific PCR primers were designed for human Eya2 (5′-primer: 5′-CCCAACTGTGTCAATGTGCTGGT-3′, positions 1317–1339 and 3′-primer: 5′-CTGTCTTGGTTGCACTGTAGATGT-3′, positions 1414–1437) and were used to screen the Ceph B mega YAC library (service of Research Genetics).

For mapping of mouse *Eya2,* interspecific backcross progeny were generated by mating (C57BL/6J × *M. spretus*) F_1 females and C57BL/6J males as described (Copeland and Jenkins 1991). In total, 205 F_2 mice were used to map the *Eya2* locus (see text for details). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer, and hybridization were performed essentially as described (Jenkins et al. 1982). All blots were prepared with Hybond–N+ nylon membrane (Amersham). The probe, a 900-bp *Eco*RI/*Hin*dIII fragment of the mouse *Eya2* cDNA (positions 1273–2179 in the final cDNA sequence), was labeled with [^{32}P]dCTP and a nick translation labeling kit (Boehringer Mannheim); washing was done to a final stringency of 0.8 × SSCP, 0.1% SDS, 65°C. A major fragment of 2.4 kb was detected in *Pvu*II-digested C57BL/6J DNA, and a major fragment of 3.4 kb was detected in *Pvu*II-digested *M. spretus* DNA. The presence or absence of the 3.4-kb *Pvu*II *M. spretus*-specific fragment was followed in backcross mice.

```
            1                                                  50
Mouse       MLEVVTSPSL ATSSD..... .WSEHGAAVG TLSDREGIAK SAALSVPQLF
Human       *V*L*I**** TVN**CLDKL KFNRAD***W *****Q**T* **P*R*S***

            51                                                100
Mouse       VKSHPRVPPG QSSTAMAAYG QTQYSTGIQQ APPYTAYPTP AQAYGIPPYS
Human       SR*C***L*R *P******** *****A**** *T******P* *******S**

            101                                               150
Mouse       IKTEDSLNHS PSQSGFLSYG PSFSTAPAGQ SPYTYPVHST AGLYQGANGL
Human       ********** *G******** S****S*T** *****QM*G* T*F***G***

            151                                               200
Mouse       TNTAGFGSVH QDYPSYPSFS QNQYPQYFSP SYNPPYVPAS SLCSSPLSTS
Human       G*A******* *******G*P *S*****YGS ********** *I*P******

            201                                               250
Mouse       TYVLQEAPHN VPSQSSESLA GDYNTHNGPS TPAKEGDTER PHRASDGKLR
Human       *******S** **N******* *E******** ********D* **********
Drosophila  PLGNVS*AAA AAALN*SGGS SVGTAGS*GV ATS*TTP*GK TG**RGRRHQ

            251                                               300
Mouse       GRSKRNSDPS PAGDNEI... .ERVFVWDLD ETIIIFHSLL TGTFASRYGK
Human       *****S**** *******... .********* ********** **********
Squid                              ********* ********** ***Y*Q****
Drosophila  QP*PTR*TA* DT*NS*AVKP P********* **L****T** S*SY*N**T*

            301                                               350
Mouse       DTTTSVRIGL MMEEMIFNLA DTHLFFNDLE DCDQIHVDDV SSDDNGQDLS
Human       ********** ********** ********** ********** **********
Squid       *PPG**TL** R********* ********** E***V*I*** **********
Drosophila  *HSSLMT*AF R****V**M* ***F***EI* E***V*I*** **********

            351                                               400
Mouse       TYNFSTDGFH STAPGA...S LCLGTGVHGG VDWMRKLAFR YRRVKEMYNT
Human       *****A**** *S****...N ****S***** ********** **********
Squid       ****A***** AA*TNT...N **IA***C** ********** *.........
Drosophila  A********* TNT*PGAPPN ***P***R** ********** **KI*DI**S

            401                                               450
Mouse       YRNNVGGLIG APKRETWLQL RAELEALTDL WLTHSLKALN LINSRPNCVN
Human       *K******** T********* ********** ********** **********
Drosophila  **G***T*L* PG***A***I *S*I*VA**N *A*LA**C*S M*SQ*E****
C. elegans                                                  AEKYA

            451                                               500
Mouse       VLVTTTQLIP ALAKVLLYGL GSVFPIENIY SATKTGKESC FERIMQRFGI
Human       ********** ********** ********** ********** **********
Drosophila  ****S***A* *******F** *GI*N***** **H*I*H*T* Y***VT****
C. elegans  *VLSNDG*VL GA*QLMIS** N*SV*V**** *IS*Q****V **K*QS***I

            501                                          547
Mouse       KAVYIVIGDG VEEEQGAKKH NMPFWRISCH ADLEALRHAL ELEY*
Human       ****V***** ********** ********** *N******** *****
Drosophila  *ST*V***** N***TA**AM *F******A* S*IR**YT** DMGF*
C. elegans  *CSF*C*TS* .DTANS**RL *I*V*PLNSN T**DK*YS** DNFL*GG
```

Fig. 1. Sequence alignment of mouse and human Eya2 with the C-terminus of *Drosophila eya* and the partial amino acid sequences available for the squid and *C. elegans* family members. (* indicates sequence identity; . . . indicates gap inserted to generate an optimal alignment).

A description of the probes and RFLPs for the loci linked to *Eya2* including *Src* and *Nfat1* has been reported previously (Siracusa et al. 1989; Luo et al. 1996). Recombination distances were calculated as described (Green 1981) with the computer program SPRETUS MADNESS. Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns.

Northern and Southern blot hybridization analysis. Total RNA was prepared from 10.5- and 13.5-d.p.c. embryos (date of vaginal plug, 0.5 d.p.c.) and newborn mouse tissues using Razol (TelTest, Friendswood, Texas), and Northern blots prepared as described (Duncan et al. 1995). These blots, as well as an adult mouse multiple tissue Northern blot (Clontech), were hybridized to the same random-primed cDNA probe used for chromosomal localization (see above) as previously described (Duncan et al. 1995). The total mRNA blot was normalized by post-hybridization staining of the membrane for ribosomal RNA with methylene blue as described (Herrin and Schmidt 1988). The expression pattern of human *Eya2* in adult tissues was determined by hybridizing the insert from EST 34807 to a Clontech MTN blot as above.

In situ hybridization analysis of Eya2 *expression during mouse embryonic development.* Transcription templates were prepared from the mouse *Eya2* cDNA by PCR of cDNA clones with primers containing T3 RNA polymerase promoters. The antisense template was prepared with primer 5′-GGAGCAATTAACCCTCACTAAAGGCCGTCCG-CACTGCTATAGATACT-3′ (positions 1684–1662) and 5′-AAGAGAG-

ACCTGGCTGCAGCTGCGCGCCGA-3′ (positions 1273–1302), while the sense template was prepared with primers 5′-GGAGCAATTAAC-CCTCACTAAAGGAAGAGAGACCTGGCTGCAGCTGCG-CGCCGA-3′ (positions 123–1302) and 5′-CCGTCCGCACTGCTATA-GATACT-3′ (positions 1684–1662). The T3 RNA polymerase promoter is underlined. The PCR products were gel purified and used to generate digoxigenin-labeled riboprobes with the Ampliscribe T3 transcription kit (Epicentre Technologies, Madison, Wis., Kaplan et al. 1996).

Embryos (FVB/N or B6EiC3H) were obtained from superovulated females. Noon of the plug day was considered O.5 d.p.c. and the day of birth was considered 1 day post natal (d.p.n.).Whole-mount in situ hybridization was performed according to Wilkinson (1993). Cryostat sections of older embryos (17.5 d.p.c.) and 10 d.p.n. lenses were treated and hybridized with the same probes as described (Hodgkinson et al. 1993).

Results

Identification of eya *family members in vertebrates, mollusks and nematodes.* A cDNA encoding the coding region of a human homolog of *Drosophila eya* was identified by a search of the EST database and subsequent analysis of 5′ RACE clones (see Materials and methods). This cDNA is 2327 bp long (excluding the poly(A) tail), has an ORF of 1614 nucleotides (Fig. 1), and is 64% identical to *Drosophila eya* in a conserved region of 813 nucleotides encoding 271 C-terminal amino acids. The deduced amino acid sequences in this region are 79% similar and 65% identical between the human and *Drosophila* proteins. While this paper was under review, the sequence of three eyes absent homologs was published (Xu et al. 1997), and the cDNA characterized here corresponds to *Eya2*.

The cDNA for the murine homolog of *Eya2* (see Materials and methods) is 2181 bp long (excluding the poly(A) tail), which is very close to the mRNA size obtained by Northern blot experiment (see below). The cDNA contains an ORF of 532 amino acids (Fig. 1). There is a stop codon in frame with the first methionine codon 48 nucleotides upstream, indicating that this cDNA contains the complete coding region. Comparison of the sequences for mouse and human *Eya2* indicates they are true orthologs. Mapping data (see below) confirm this conclusion. They are 78% identical at the nucleotide level and 93% identical at the amino acid level. The mouse protein is 67% identical and 80% similar to Drosophila eyes absent over the 271 C-terminal amino acids. During the cloning of mouse *Eya2,* an alternate cDNA with a deletion of nucleotides 421–497 was obtained, which would result in a protein with amino acids 114–142 deleted.

The region most highly conserved between vertebrate *Eya2* and *Drosophila eya* was used to design degenerate oligonucleotides (see Materials and methods), which were used for rt-PCR of embryonic squid (*Loligo opalescens*) RNA. The PCR product with the expected size of 351 bp possessed 68–72% sequence similarity to human, mouse, and *Drosophila eya* family members, while the deduced amino acid sequences demonstrated 74–91 % identities (Fig. 1). We concluded that the identified partial cDNA corresponds to the squid homolog of the *Drosophila eya.* The *C. elegans* eya-related protein identified in the *C. elegans* EST database is 38% identical and 60% similar over 102 known amino acids.

Sequence comparison between vertebrate and *Drosophila* eya family members demonstrated that eya proteins can be divided into two domains. The C-terminal 271 amino acids of these proteins are highly conserved between vertebrates and *Drosophila,* and the partial sequences available for the squid and *C. elegans* cognates overlap this region. The C-terminal region contains one cluster of basic and one cluster of acidic amino acids that were previously identified in the *Drosophila* eya protein (Bonini et al. 1993). Another cluster of basic amino acids is not conserved at the sequence level, but is located at the same position upstream of the conserved C-terminal domain in both *Drosophila* and vertebrate proteins. The N-terminal domains of the *Drosophila* and vertebrate proteins are of very different lengths (495 versus 261–267) and share little to no amino acid identity. However, the N-terminal domains of both the vertebrate and *Drosophila* proteins contain proline, serine, threonine (PST)-rich regions. Nearly 37% of the amino acids between positions 40 and 267 in vertebrate Eya2 are PST, and many of the amino acid changes seen between the mouse and human proteins are simply conversions among these three amino acids (See Fig. 1). Positions 349–440 of the *Drosophila* protein are 45% PST, while positions 266–440 are 33% PST.

Chromosomal localization of human and mouse Eya2. Screening of the human CEPH-B YAC library demonstrated that the human Eya2 is located in YAC 857H11. This 850-kb YAC has been localized to human Chr 20q13.1 by fluorescent in situ hybridization and lies within the 7 mb interval identified for the diabetes-susceptibility locus (Stoffel et al. 1996).

The mouse chromosomal location of *Eya2* was determined by interspecific backcross analysis with progeny derived from matings of [(C57BL/6J × *Mus spretus*) F_1 × C57BL/6J] mice. C57BL/6J and *M. spretus* DNAs were digested with several enzymes, and an informative 3.4-kb *PvuII M. spretus Eya2* RFLP (see Materials and methods) was used to follow the segregation of the *Eya2* locus in backcross mice. The mapping results indicated that *Eya2* is located in the distal region of mouse Chr 2 linked to *Src* and *Nfat.* Although 169 mice were analyzed for every marker and are shown in the segregation analysis (Fig. 2), up to 187 mice were typed for some pairs of markers. Each locus was analyzed in pairwise combinations for recombination frequencies using the additional data.

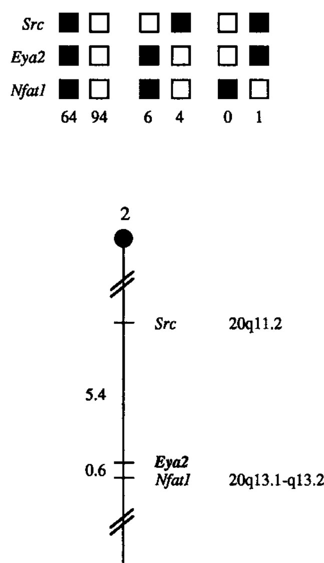

Fig. 2. Eya2 maps to the distal region of mouse Chr 2. *Eya2* was placed on mouse Chr 2 by interspecific backcross analysis. The segregation patterns of *Eya2* and flanking genes in 169 backcross animals that were typed for all loci are shown at the top of the figure. For individual pairs of loci, more than 169 animals were typed (see text). Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J × *M. spretus*) F_1 parent. The **shaded boxes** represent the presence of a C57BL/6J allele, and **white boxes** represent the presence of a *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. A partial Chr 2 linkage map showing the location of *Eya2* in relation to linked genes is shown at the bottom of the figure. Recombination distances between loci in centimorgans are shown to the left of the chromosome, and the positions of loci in human chromosomes, where known, are shown to the right. References for the human map positions of loci cited in this study can be obtained from GDB (Genome Data Base), a computerized database of human linkage information maintained by The William H. Welch Medical Library of The Johns Hopkins University (Baltimore, Md.).

The ratios of the total number of mice exhibiting recombinant chromosomes to the total number of mice analyzed for each pair of loci and the most likely gene order are: centromere–Src–10/187–*Eya2*–1/172–*Nfat1*. The recombination frequencies [expressed as genetic distances in centiMorgans (cM) + the standard error] are–*Src*–5.4 ± 1.7–*Eya2*–0.6 ± 0.6.

Comparison of the *Eya2* position with a composite mouse linkage map for Chr 2 indicates that this gene is located close to the cataract mutation *Lop-4,* which has been mapped to position 96.0 (West and Fisher 1986). Human Chr 20q13.1 and the distal portion of mouse Chr 2 have been previously shown to be syntenic chromosomal segments, confirming that human and mouse *Eya2* are orthologous genes (DeBry and Seldin 1996).

Expression pattern of Eya2 *in vertebrates.* The tissue-specific expression of *Eya2* was first determined in mice and humans by Northern blot hybridization. In both mouse and human, a single message with a length of about 2300–2400 nt was detected. In mouse embryos, expression was seen at 10.5 d.p.c. and decreased significantly by 13.5 d.p.c. In newborn mice, expression was largely confined to the eye, brain, and lung with little to no expression in the skin, liver, intestine, and kidney (Fig. 3A). In adult mice, the highest expression levels were detected in the eye lens with lower relative levels seen in the lung and brain. Relatively little eyes *Eya2* mRNA was detected in the adult mouse testis, kidney, muscle, liver, spleen, and heart (Fig. 3B). In adult humans, the highest expression was detected in muscle, with lesser amounts found in kidney, placenta, pancreas, brain, and heart Fig. 3C).

In order to localize *Eya2* expression at the cellular level, in situ hybridization analysis was performed with the mouse *Eya2* gene. At 11.5 d.p.c., the expression of *Eya2* mRNA is detected at appreciable levels in the nasal pit, nasal processes, cranial ganglia, and much of the head mesenchyme including the developing mandible and maxilla; however, no expression was detected in the developing eye (data not shown). While no *Eya2*-specific hybridization signals were observed during the early events of eye development, Northern blot analysis did detect appreciable levels of *Eya2* mRNA in the perinatal eye and adult lens of mice. Thus, we performed *in situ* hybridization analysis on tissue sections prepared from 17.5 d.p.c. and 10 d.p.n. mice to determine the cellular localization of *Eya2* expression in the eye. At. 17.5 d.p.c., the highest levels of *Eya2* mRNA in the eye were observed in the posterior portions of the primary lens fiber cells. Lower levels of expression were detected in the retina, sclera, and corneal epithelium as well as the cellular portion of the corneal stroma (Fig. 4 A–C). In the brain, hybridization was detected throughout the telencephalon including the hippocampus as well as the diencephalon (data not shown). The olfactory epithelium, nasal cartilage, and developing bone of 17.5 d.p.c. embryos did not express *Eya2* at detectable levels (data not shown). By 10 d.p.n., the expression of *Eya2* was maintained in the lens; however, most of the mRNA was in recently elongated secondary fibers with lower levels observed in the remainder of the fiber cell mass (Fig. 4D).

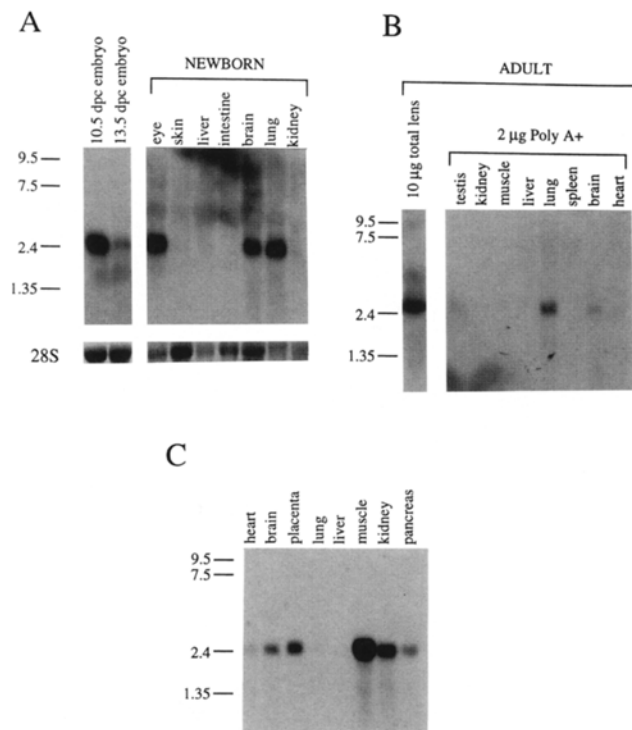

Fig. 3. Northern blot hybridization analysis of *Eya2*. **A.** Expression of *Eya2* mRNA in newborn and embryonic mice. Ten micrograms of total RNA was analyzed from the indicated tissues, and the 28S ribosomal RNA was visualized by post-hybridization staining of the filter to control for RNA loading and transfer. **B.** Expression of *Eya2* in adult mouse tissues. The blot was reprobed with actin, and approximately equal amounts of mRNA were found in each tissue (data not shown). **C.** Expression of *Eya2* in adult human tissues. The blot was reprobed with actin, and approximately equal amounts of mRNA were found in each tissue (data not shown).

Discussion

The present results indicate that homologs of the *Drosophila eya* gene are present across metazoan species. We have identified corresponding homologs in vertebrate, cephalopod mollusks, and nematodes. Eya/Eya2 are moderately conserved proteins showing 93% identity between human and mouse. The protein molecule may be divided into a conserved C-terminal domain and a variable N-terminal domain (see Fig. 1). A similar situation is observed when regulatory proteins from different species are compared. For example, *Pax-6* and eyeless gene products show a high percentage identity in the region of their paired- and homeodomains, and little similarity in the amino acid sequences in other regions (Quiring et al. 1994). It was proposed that similar secondary or tertiary folds may be adequate for conservation of function as opposed to direct identity in amino acid sequences (Xue and Noll 1996). The N-terminal domains of eya/Eya2 contain a PST-rich region that is present in the activator domain of many transcription factors (Walther and Gruss 1991; Mermod et al. 1989; Theill et al. 1989). *Drosophila* eya was demonstrated to be a nuclear protein (Bonini et al., 1993). We do not know whether Eya2 is a nuclear protein although we can assume it to be so if the functions of *Drosophila* and vertebrate proteins are conserved.

Many vertebrate regulatory proteins are encoded by gene families. They often possess one or more conserved domains, such as a DNA-binding domain, and show little similarity elsewhere. Following this pattern, there are at least two other *eya*-related genes in vertebrates (S.I. Tomarev, unpublished; Xu et al. 1997). The proteins encoded by the other genes also consist of two domains: a C-terminal domain, which shows high similarity to the C-terminal domain of Eya2, and an N-terminal domain, which is quite different from the N-terminal domain of Eya2. As is common for regulatory proteins, alternatively spliced forms of *Eya2* were identified: one which lacks the region coding for amino acids 114–142 in the variable N-terminal domain and a second one (GenBank accession number AA002292) that has a different 5′-untranslated region. The role of these alternatively spliced forms is not clear at present.

Since cells of the *Drosophila* eye imaginal discs deficient in *eya* activity undergo apoptosis instead of differentiating into ommatidia, and since *eya* is expressed very early in eye development, probably after *ey* but before *so* (Cheyette et al. 1994; Bonini et al. 1993), expression of *Eya2* in the eye was investigated in detail. We could not detect expression of *Eya2* in the eye at or before 11.5

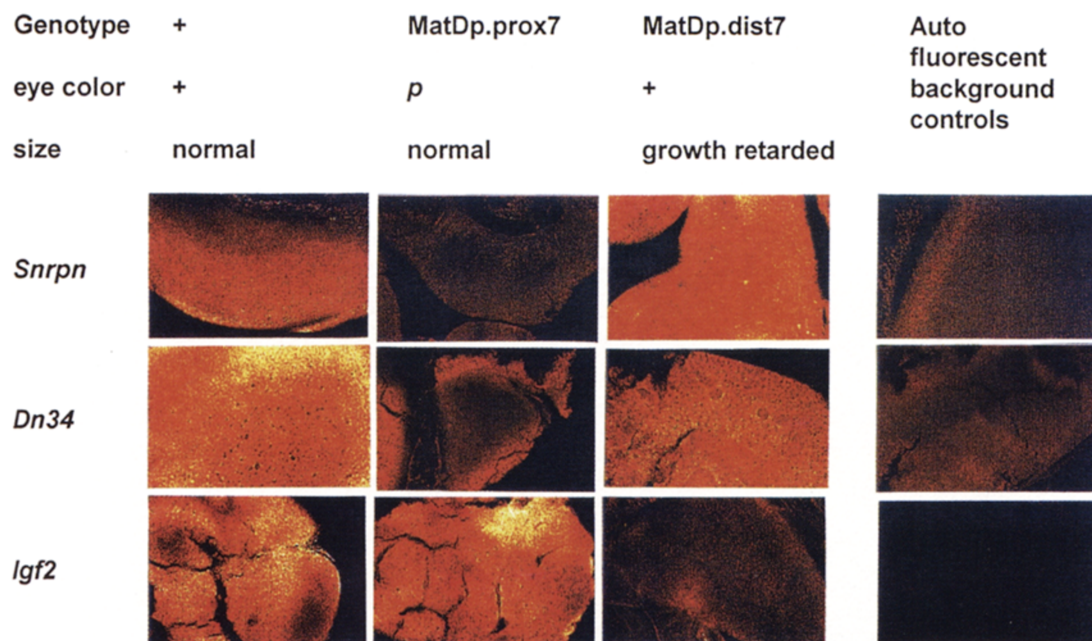

Fig. 4. In situ hybridization analysis of *Eya2* expression in the mouse eye. **A.** Brightfield photomicrograph of an eye from a 17.5-d.p.c. mouse embryo hybridized to an antisense probe for *Eya2*. Expression (blue/black) was detected in the cornea, retina, sclera, and the primary fiber cells of the lens. (100×). **B.** Brightfield photomicrograph of an eye from a 17.5-d.p.c. mouse embryo hybridized to a sense probe for *Eya2*. (100×). **C.** Differential interference contrast photomicrograph of the cornea from a 17.5-d.p.c. mouse embryo hybridized to the antisense strand of *Eya2*. Note the strong hybridization signal (blue/black) detected in the corneal epithelium and cellular aspects of the stroma. No hybridization signal was detected in the corneal endothelium. (400×). **D.** Differential interference contrast photomicrograph of the lens from a 10-d.p.n. mouse hybridized to the antisense strand of *Eya2*. Note the strong hybridization signal (blue/black) detected in the secondary fibers just underlying the lens epithelium. (1000×). **Abbreviations:** a, artifactual signal due to nonspecific probe binding to the acellular lens capsule; C, lens capsule; CE, corneal epithelium; Co, cornea; L, lens; LE, lens epithelium; N, corneal endothelium; p, pigmented epithelium; PF, primary lens fibers; R, retina; Sc, sclera; SF, secondary lens fibers; St, corneal stroma.

d.p.c. (data not shown; Xu et al. 1997). At this stage of mouse development, both eye and lens vesicles have already formed, and lens determination is complete (Grainger 1992; Grindley et al. 1995). On the basis of these results, we can conclude that *Eya2* is probably not involved in the early inductive events of eye development. In the mouse, expression of *Pax-6* and *Six 3* gene is detected in the region of the future optic vesicle starting from 8.2 d.p.c., much earlier than *Eya2* (Walther and Gruss 1991; Oliver et al. 1995b). Xu and associates (1997) have reported that *eya2* expression commences in the eye between 11.2 and 12.5 d.p.c., while no expression was detected in the lens as late at 14.5 d.p.c. Although at present we do not know how early the onset of *Eya2* expression occurs in the lens, high levels of mRNA were detected in the lenses of 17.5 d.p.c. mouse embryos. Furthermore, *Eya2* RNA is still abundant in the 10 d.p.n. and adult eye lens. The expression of *Eya2* in the primary lens fiber cells of 17.5 d.p.c. embryos is of particular note, since the nuclei and organelles of these cells are breaking down to facilitate lens transparency (Kuwabara and Imaizumi 1974, Vrensen et al. 1991). By 10 d.p.n., *Eya2* mRNA is detected in secondary lens fibers, which will also lose their organelles during their terminal differentiation. It is noteworthy that early cell death events occur in the developing retina (Potts et al. 1982) and dorsal root ganglia (Carr and Simpson 1981), and denucleation of lens fiber cells is reminiscent of apoptosis in many respects (Gao et al., 1995). Expression of *Eya2* in these tissues is consistent with a critical role proposed for its homolog, *eya,* in the selection between cell differentiation and cell death (Bonini et al. 1993).

In the mouse, *Eya2* was mapped near to the *Lop4* mutation. This mutation leads to nuclear or perinuclear opacities that extended into the cortex of the lens, but in some cases anterior cortical or anterior polar cataracts have been observed (West and Fisher 1986). *Eya2* may be considered as a candidate gene for the *Lop4* mutation. In humans, we mapped *Eya2* to Chr 20q13.1 by screening of a human Mega Yac library. Recently, the human homolog of *eya* was independently mapped to the same region by FISH (Banfi et al. 1996). The corneal dystrophy locus (Heon et al. 1995) was proposed to be close to the position of the human *eya* homolog (Banfi et al. 1996), but our data indicate that *Eya2* is located outside of the corneal dystrophy region and is found in the recombination interval containing the diabetes-susceptibility locus (Stoffel et al. 1996). While *Eya2* is expressed in the adult human pancreas, a recent report has determined that mutations in hepatocyte nuclear factor-4α are responsible for MODY1 (Yamagata et al. 1996).

The present study demonstrates that *eya* family members are expressed in animals belonging to at least four highly divergent phyla. Since these genes have maintained a high degree of similarity with each other over vast distances of evolutionary time, it is possible that they play important developmental roles in vertebrates, mollusks, and nematodes as well as in arthropods. It should be noted that while vertebrates, mollusks, and arthropods all have complex eyes (Land 1988), the nematode *C. elegans* does not appear to possess photoreceptor cells and has only a weak response to visible light (Chisholm and Horvitz 1995; Zhang and Emmons 1995). It is possible that an *eya*-related gene, like *Pax-6,* was first involved in head development, and recruitment for visual system development occurred later.

Acknowledgments. We thank Drs. J. Piatigorsky for helpful discussion, suggestions, and support, W. Gilly of the Hopkins Marine Station for advice and help with the collection of *Loligo opalescens* embryos; Ernst Tamm and Csaba Saloman for helpful discussions, Mary Barnstead for excellent technical assistance; and the Frederick Cancer Research Center, The United States Department of Agriculture ACeB server, The Sanger Center, and the National Center for Biotechnology Information for com-

putational resources. This research was supported, in part, by the National Cancer Institute, Department of Health and Human Services, with The ABL Basic Research Program.

References

Banfi S, Borsani G, Rossi E, Bernard L, Guffanti A, Rubboli F, Marchitiello A, Giglio S, Coluccia E, Zollo M, Zuffardi O, Ballabio A (1996) Identification and mapping of human cDNAs homologous to Drosophila mutant genes through EST database searching. Nature Genet 13, 167–174

Bonini NM, Choi K-W (1995) Early decisions in *Drosophila* eye morphogenesis. Curr Opin Genet Dev 5, 507–515

Bonini NM, Leiserson WM, Benzer S (1993) The eyes absent gene: genetic control of cell survival and differentiation in the developing *Drosophila* eye. Cell 72, 379–395

Carr VM, Simpson SB (1981) Rapid appearance of labeled degenerating cells in the dorsal root ganglia after exposure of chick embryos to tritiated thymidine. Brain Res 254, 157–162

Cheyette BNR, Green PJ, Martin K, Garren H, Hartenstein V, Zipursky SL (1994) The Drosophila sine oculis locus encodes a homeodomain-containing protein required for the development of the entire visual system. Neuron 12, 977–996

Chisholm AD, Horvitz HR (1995) Patterning of the *Caenohabditis elegans* head region by the Pax-6 family member *vab-3*. Nature 377, 52–55

Copeland NG, Jenkins NA (1991) Development and applications of a molecular genetic linkage map of the mouse genome. Trends Genet 7, 113–118

Czerny T, Busslinger M (1995) DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). Mol Cell Biol 15, 2858–2871

DeBry RW, Seldin MF (1996) Human/mouse homology relationships. Genomics 33, 337–351

Doe CQ, Chu-Lagraff Q, Wright DM, Scott MP (1991) The prospero gene specifies cell fates in the *Drosophila* central nervous system. Cell 65, 451–464

Dudley AT, Lyons KM, Robertson EJ (1995) A requirement for bone morphogenetic protein-7 during development of the mammalian kidney and eye. Genes Dev 9, 2795–2807

Duncan MK, Roth HJ, Thompson M, Kantorow M, Piatigorsky J (1995) Chicken βB1-crystallin: gene sequence and evidence for functional conservation of promoter activity between chicken and mouse. Biochim Biophys Acta 1261, 68–76

Gao CY, Bassnett S, Zelenka PS (1995) CyclinB, p35^{cdc2}, and H1-kinase activity in terminally differentiating lens fiber cells. Dev Biol 169, 185–194

Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, Maas RL (1994) Pax-6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. Nature Genet 7, 463–471

Grainger RM (1992) Embryonic lens induction: shedding light on vertebrate tissue determination. Trends Genet 8, 349–355

Green EL (1981) Linkage, recombination and mapping. In *Genetics and Probability in Animal Breeding Experiments.* (New York: Oxford University Press) pp 77–113

Grindley JC, Davidson DR, Hill R (1995) The role of Pax-6 in eye and nasal development. Development 121, 1433–1442

Halder G, Callaerts P, Gehring W (1995a) Induction of ectopic eyes by targeted expression of the eyeless gene in *Drosophila.* Science 267, 1788–1792

Halder G, Callaerts P, Gehring WJ (1995b) New perspectives on eye evolution. Curr Opin Genet Dev 5, 602–609

Hanson I, van Heyningen V (1995) Pax-6: more than meets the eye. Trends Genet 11, 268–272

Heberlein U, Wolff T, Rubin GM (1993) The TGFβ homolog *dpp* and the segment polarity gene *hedgehog* are required for propagation of a morphogenic wave in the *Drosophila* retina. Cell 75, 913–926

Heon E, Mathers WD, Alward WL, Weisenthal RW, Sunden SL, Fishbaugh JA, Taylor CM, Krachmer JH, Sheffield VC, Stone EM (1995) Linkage of posterior polymorphous corneal dystrophy to 20q11. Hum Mol Genet 4, 485–488

Herrin DL, Schmidt GW (1988) Rapid, reversible staining of northern blots prior to hybridization. Biotechniques 6, 196–200

Hill RE, Favor J, Hogan BLM, Ton CCT, Saunders GF, Hanson IM, Prosser J, Jordan T, Hastie ND, van Heyningen V (1991) Mouse *Small eye* results from mutations in a paired-like homeobox-containing gene. Nature 354, 522–525

Hodgkinson CA, Moore KJ, Nakayama A, Steingrimsson E, Copeland NG, Jenkins NA, Arnheiter H (1993) Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. Cell 74, 395–404

Jenkins NA, Copeland NG, Taylor BA, Lee BK (1982) Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of Mus musculus. J Virol 43, 26–36

Kaplan ED, Lara S, Wight TN (1996) Non-radioactive in situ hybridization using digoxigenin-labeled RNA probes generated with the AmpliScribe™ T7, SP6 and T3 high yield transcription kits. Epicentre Forum 2, 1–4

Kuwabara T, Imaizumi M (1974) Denucleation process of the lens. Investig Opthalmol Vis Sci 13, 973–981

Land MF (1988) The optics of animal eyes. Contemp Phys 29, 435–455

Land MF, Fernald RD (1992) The evolution of eyes. Annu Rev Neurosci 15, 1–29

Leiserson WM, Bonini NM, Benzer S (1994) Transvection at the *eyes absent* gene of *Drosophila.* Genetics 138, 1171–1179

Loosli F, Kmita-Cunisse M, Gehring WJ (1996) Isolation of a Pax-6 homolog from the ribbonworm *Lineus sanguineus.* Proc Natl Acad Sci USA 93, 2658–2663

Luo C, Edelhoff S, Disteche C, Copeland NG, Jenkins NA, Hogan PG, Rao A (1996) Normal function of the transcription factor NFAT1 in *Wasted* mice: Chromosome localization of NFAT1 gene. Gene, in press.

Mermod N, O'Neill EA, Kelly TJ, Tijan R (1989) The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell 58, 741–753.

Nusslein-Vollard C, Wieschaus E, Kluding H (1984) Mutations affecting the pattern of the larval cuticle of *Drosophila melanogaster:* I. Zygotic loci on the second chromosome. Roux's Arch Dev Biol 193, 267–282

Oliver G, Mailhos A, Wehr R, Copeland NG, Jenkins NA, Gruss P (1995a) Six3, a murine homologue of the *sine oculis* gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development. Development 121, 4045–4055

Oliver G, Wehr R, Jenkins NA, Copeland NG, Cheyette BNR, Hartenstein V, Zipursky SL, Gruss P (1995b) Homeobox genes and connective tissue patterning. Development 121, 693–705

Potts RA, Dreher B, Bennett MR (1982) The loss of ganglion cells in the developing retina of the rat. Dev Brain Res 3, 481–486

Quiring R, Walldorf U, Kloter U, Gehring WJ (1994) Homology of the *eyeless* gene of *Drosophila* to the *Small eye* gene in mice and Aniridia in humans. Science 265, 785–789

Siracusa LD, Buchberg AM, Copeland NG, Jenkins NA (1989) Recombinant inbred strain and interspecific backcross analysis of molecular markers flanking the murine *agouti* coat color locus. Genetics 122, 669–679

Stoffel M, Le Beau MM, Espinosa III R, Bohlander SF, Le Paslier D, Cohen D, Xiang K-S, Cox NJ, Fajans SS, Bell GI (1996) A yeast artificial chromosome based map of the region of chromosome 20 containing the diabetes-susceptibility gene, MODY1, and a myeloid leukemia related gene. Proc Natl Acad Sci USA 93, 3937–3941

Theill LE, Castrillo J-L, Wu D, Karin M (1989) Dissection of functional domains of the pituitary-specific transcription factor GHF-1. Nature 342, 945–949

Tomarev SI, Sundin O, Banerjee-Basu S, Duncan MK, Yang J-M, Piatigorsky J (1996) Chicken homeobox gene Prox 1 related to Drosophila prospero is expressed in the developing lens and retina. Dev Dyn 206, 354–367

Tomarev SI, Callaerts P, Kos L, Zinovieva R, Halder G, Gehring W, Piatigorsky J (1997) Squid Pax-6 and eye development. Proc Natl Acad Sci USA 94, 2421–2426

Vrensen GFJM, Graw J, De Wolf A (1991) Nuclear breakdown during terminal differentiation of primary lens fibers in mice: a transmission electron microscopic study. Exp Eye Res 52, 647–659

Walther C, Gruss P (1991) Pax-6, a murine paired box gene, is expressed in the developing CNS. Development 113, 1435–1449

West JD, Fisher G (1986) Further experience of the mouse dominant cataract mutation test from an experiment with ethylnitrosourea. Mutat Res 164, 126–136

Wilkinson DG (1993) In *In situ Hybridization: a Practical Approach.* (Oxford: IRL Press), p74–83

Xu P-X, Woo I, Her H, Beier DR, Maas RL (1997) Mouse *Eya* homologues of the *Drosophila eyes absent* gene require *Pax-6* for expression in the lens and nasal placode. Development 124, 219–231

Xue L, Noll M (1996) The functional conservation of proteins in evolutionary alleles and the dominant role of enhancers in evolution. EMBO J 15, 3722–3731

Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI (1996) Mutations in the hepatocyte nuclear factor-4α gene in maturity-onset diabetes of the young (MODY1). Nature 384, 458–460

Yarfitz S, Hurley JB (1994) Transduction mechanisms of vertebrate and invertebrate photoreceptors. J Biol Chem 269, 14329–14332

Zhang Y, Emmons SW (1995) Specification of sense-organ identity by a *Caenorhabditis elegans* Pax-6 homologue. Nature 377, 55–59