New 3′ elements control Pax6 expression in the developing pretectum, neural retina and olfactory region

Caroline Griffin, Dirk A. Kleinjan, Brendan Doe, Veronica van Heyningen*

MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

Received 10 October 2001; received in revised form 28 November 2001; accepted 28 November 2001

Abstract

Pax6 is a key transcriptional regulator in eye, olfactory system, forebrain, pituitary cerebellum, spinal cord and pancreas development. Alternative splicing, promoter usage and multiple enhancers regulate the complex Pax6 spatio-temporal expression pattern. Chromosomal rearrangements which abolish PAX6 gene expression have been characterised downstream of the coding region. Through evolutionary sequence comparison and transgenic reporter studies, we have identified a new Pax6 3′ cis-regulatory region. This region, C1170 Box 123, contains three distinct modules of human–mouse sequence conservation, while only Box 1 is conserved to Fugu. Both the human and the orthologous Fugu sequence direct similar reporter gene expression in the developing pretectum, neural retina and olfactory region, indicating evolutionary conservation of Pax6 regulatory mechanisms despite the low level of overall sequence conservation. © 2002 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Pax6; Evolutionary sequence conservation; Transgenic analysis; cis-Acting element; Pufferfish

1. Introduction

The transcription factor Pax6 is necessary for the correct development of the eyes, olfactory system, brain, spinal cord, pituitary and pancreas (Walther and Gruss, 1991; Callaerts et al., 1997; Mansouri et al., 1998; Kioussi et al., 1999). Heterozygous loss-of-function mutations in Pax6 have been shown to result in aniridia in humans and Small eye phenotype in rodents (Hogan et al., 1986, 1988; Hill et al., 1991; Glaser et al., 1992, 1994; Jordan et al., 1992; Fujiwara et al., 1994). A complex spatio-temporal expression pattern has been observed for this important developmental regulator in various animal phyla (reviewed in Callaerts et al., 1997; Kammermeier et al., 2001). From the earliest stages of vertebrate eye morphogenesis Pax6 is expressed in the optic vesicle, giving rise to the retina and pigmented epithelium, as well as in the overlying ectoderm that forms the lens and cornea (Hogan et al., 1986, 1988; Hill et al., 1991; Glaser et al., 1992, 1994; Jordan et al., 1992; Fujiwara et al., 1994). A complex spatio-temporal expression pattern has been observed for this important developmental regulator in various animal phyla (reviewed in Callaerts et al., 1997; Kammermeier et al., 2001). From the earliest stages of vertebrate eye morphogenesis Pax6 is expressed in the optic vesicle, giving rise to the retina and pigmented epithelium, as well as in the overlying ectoderm that forms the lens and cornea (Walther and Gruss, 1991; Martin et al., 1992; Grindley et al., 1995; Davis and Reed, 1996; Koroma et al., 1997; Gehring and Ikeo, 1999). Recent studies have shown a requirement for Pax6 in the vertebrate retina during the formation of the optic vesicle, in the maintenance of multipotent retinal progenitor cells and in neurogenesis (Marquardt et al., 2001).

In addition to its involvement in ocular morphogenesis, Pax6 is necessary for olfactory development. Beginning on E8.5 (8.5 days post fertilisation) Pax6 mRNA is found in the nasal placode and expression continues in placodal epithelium during the formation of the nasal pits (E10.5) (Walther and Gruss, 1991; Grindley et al., 1995). Strong Pax6 expression continues in the developing olfactory epithelium and olfactory bulbs, and into the adult-like olfactory structure visible at E18.5 (Walther and Gruss, 1991).

Induction of Pax6 during embryogenesis also contributes to the correct patterning of the murine forebrain (prosencephalon) which develops from the most anterior portion of the neural tube. Pax6 mRNA is first detected at the neural plate stage (E8.0) in the undifferentiated neuroepithelium of the forebrain as well as in the hindbrain and spinal cord (Walther and Gruss, 1991; Puelles and Rubenstein, 1993; Schwarz et al., 1999). At the neuromere stage (E10.5), the forebrain is divided into six transverse domains termed prosomeres (see Warren and Price, 1997; Fig. 5). During this stage a broad Pax6 expression pattern is observed (Stoykova and Gruss, 1994). Most of the Pax6-positive cells are observed in the ventricular zone, where undifferentiated neuronal progenitors of the central nervous system (CNS) exist. The primitive dense cell layer then begins to
where, in common with in the only documented compound heterozygous human (Stoykova et al., 1996; Grindley et al., 1997; Mastick et al., 1997; Warren and Price, 1997; Pratt et al., 2000; Mastick and Andrews, 2001). Brain malformations are also observed at E12.5 the primary prosoencephalon has subdivided into the secondary prosoencephalon (which includes prosomeres p4–p6) and the diencephalon (p1–p3). In the diencephalon Pax6 is expressed in many cells within p1, p2 and p3. The pretectum, dorsal thalamus and ventral thalamus develop from prosomeres p1, p2 and p3, respectively, where Pax6 expression becomes restricted to specific forebrain nuclei. Downregulation of Pax6 at E15 precedes the developmental stage when differentiation of functional neuronal nuclei occurs.

Loss-of-function mutants have revealed crucial roles for Pax6 in the normal development of the morphological forebrain–midbrain boundary, regional patterning, the specification of forebrain neuron identity and axon guidance (Stoykova et al., 1996; Grindley et al., 1997; Mastick et al., 1997; Warren and Price, 1997; Pratt et al., 2000; Mastick and Andrews, 2001). Brain malformations are also observed in the only documented compound heterozygous human where, in common with Pax6 null Small eye mice and rats, there is agenesis of the olfactory bulb and eye (Hogan et al., 1986; Hill et al., 1991; Matsuo et al., 1993; Glaser et al., 1994; Grindley et al., 1995; Dellovade et al., 1998). Eye defects are also seen when Pax6 is reduced or overexpressed, suggesting that eye development is extremely sensitive to levels of Pax6 expression (Schedl et al., 1996).

A highly complex system of transcriptional regulatory control elements exists for the Pax6 gene. Differential splicing and promoter usage, acting in combination with multiple cis-acting elements, direct the complex Pax6 gene expression pattern observed (Epstein et al., 1994; Plaza et al., 1995a; Jaworski et al., 1997; Okladnova et al., 1998). All cis-acting sequences required for fully regulated Pax6 expression are assumed to be contained in a 420 kb human YAC clone extending ~200 kb both upstream and downstream of the Pax6 P1 promoter, since this construct rescues the lethality of the Pax6-null phenotype (Schedl et al., 1996). Evolutionarily conserved enhancers, exhibiting developmental stage-specific activity during eye, pancreas, spinal cord, nasal, telencephalon, diencephalon and hindbrain development have been identified as 5′ and 3′ of the Pax6 gene and within its introns (Plaza et al., 1995a; Williams et al., 1998; Xu and Saunders, 1998; Kammandel et al., 1999; Xu et al., 1999; Kleinjan et al., 2001). Functional evidence that additional 3′ regulatory elements may be required for Pax6 transcription comes from aniridia cases (normally caused by heterozygous null mutations within human Pax6) with no intragenic mutations in the Pax6 coding region, but with chromosomal rearrangements downstream of the gene (Fanties et al., 1995; Lauderdale et al., 2000; Kleinjan et al., 2001). In this study, we have identified through evolutionary sequence comparison and transgenic reporter studies, a new cis-regulatory region, called C1170 Box 123, consisting of three sequence modules (Box 1–3). C1170 Box 123 is located 77 and 17 kb 3′ of human and pufferfish (Fugu rubripes) Pax6 polyA-addition sites, respectively. Both the human and orthologous Fugu C1170 sequence direct a similar reporter gene expression pattern in the pretectum, neural retina and olfactory region of transgenic mice in a temporally and spatially restricted manner. An individual human sequence module, C1170 Box 3, located at the distal end of the full-length C1170 Box 123 sequence, which is conserved in humans and mice but not in Fugu, was shown to be necessary and sufficient to produce the neural retina expression observed using the larger fragment. However, the 5′ sequence block Box 1, which is very highly conserved in mammals and fish, fails to enhance expression in transgenic mice. Therefore, further experiments are needed to test the in vivo function of C1170 Box 1 sequence. These data demonstrate the potential value and limitations of using cross-species sequence comparisons to identify new long-range regulatory elements in the Pax6 locus.

2. Results

2.1. Identification of a potential cis-regulatory region 3′ of the Pax6 gene

A powerful approach to finding functional segments in non-coding genomic regions is cross-species sequence comparison (Aparicio et al., 1995; Hardison, 2000). We used the program PipMaker to perform a detailed comparison between 140 kb of orthologous Pax6 human and mouse sequences to highlight conserved non-coding elements which may be candidate regulatory elements required for the correct control of Pax6 gene expression. Sequences are considered to be conserved if they align without a gap for >100 bp in the PipMaker alignment and have >70% nucleotide identity (Loots et al., 2000). The percentage identity plot (PIP) (Fig. 1) is a graphic display of the human–mouse pairwise alignment. The 13 Pax6 exons (shown in Fig. 1 as numbered black boxes above the 100% identity lines) and the most 3′ exon 10 of the neighbour gene Paxneb (C11ORF19) are all highly conserved. The analysis also revealed noteworthy matches in the proximal 5′ flanking region of the Pax6 gene, which contains the promoters P0 and P1 (Fig. 1). High level sequence identity is also observed at the phylogenetically conserved ectodermal and intron 4 retinal enhancers (dashed boxes marked EE and RE in Fig. 1) (Williams et al., 1998; Xu and Saunders, 1998; Xu et al., 1999). This validates the usefulness of our approach, using human–mouse genomic sequence comparison as the first step to highlighting conserved regions which may contribute to gene regulation. Examination of the PIP shows the presence of numerous other noteworthy horizontal lines corresponding to segments of strong sequence similarity within Pax6 3′ non-coding DNA. We focused on a region (positioned around 210 k in Fig. 1), located 77 kb downstream of the 3′ polyA-addition site of...
Fig. 1. PIP of the Pax6 locus. Comparison of 140 kb of human PAX6 DNA sequence (x-axis) with orthologous mouse DNA sequence using PipMaker. The plot shows the position of numerous dots and dashes representing the levels of conservation (y-axis indicates percent identity). The scale (in kb) is shown. Along the top line of the plot, the 13 PAX6 exons and the most 3' exon 10 of the neighbour gene PAXNEB are indicated as numbered black boxes. Pax6 is transcribed in the centromeric to telomeric direction whereas PAXNEB is transcribed in the opposite direction as shown by solid horizontal arrows. Small dashed boxes mark the phylogenetically conserved ectodermal (EE) and intron 4 retinal enhancer (RE). Also indicated is the position of promoters P0 and P1. Vertical dashed down-arrows mark the position of three known chromosomal rearrangements that abolish PAX6 gene expression. The large dashed box indicates the position of the highly conserved C1170 region, located between 205k and 215k, described in the text.
the PAX6 gene, and first reported as part of the sequence of human cosmid C1170 (EMBL: Z83306). This unique sequence has greater than 90% pairwise nucleotide identity over approximately 1 kb and does not appear to be an exon, a CpG island or part of an interspersed repeat (identified by Genescan, GRAIL, EST identity and RepeatMasker). It is located within the transcription unit of the unrelated neighbour gene PAXNEB. Interestingly, chromosomal rearrangements (indicated by vertical down-arrows in Fig. 1) which lead to aniridia, the PAX6 haploinsufficiency phenotype, have been described in the interval between the PAX6 polyA-addition site and downstream of the C1170 region (Fantes et al., 1995; Lauderdale et al., 2000), suggesting that cis-elements downstream of the breakpoints are required for normal PAX6 expression.

In order to investigate whether sequences at the human C1170 region are conserved over extended phylogenetic distance, 10 kb of orthologous human, mouse and Fugu genomic DNA sequence was compared using the VISTA program. Parameters were defined to select for ungapped alignment of at least 40 bp and greater than 50% identity. The VISTA plot reveals three positionally conserved islands of human–mouse sequence homology in the C1170 Box 123 cassette. Box 1 (Fig. 2A) shows the most extensive homology: 94% identity over nearly 800 bp, between human and mouse. Arrows in Fig. 2A indicate two short sequence blocks, 41 and 58 bp, with 92 and 87% identity, respectively, between human or mouse and Fugu. Box 2 (84% identical over 182 bp core) and Box 3 (82% identical over 242 bp core) show conservation among mammals only, but not over the extended phylogenetic distance between mammals and fish (Fig. 2A).

2.2. C1170 Box 123 cis-regulatory region directs Pax6 expression in the developing pretectal, neural retina and olfactory region

To study the biological significance of the C1170 Box 123 evolutionary sequence conservation, we generated a reporter construct to assess enhancer function of the whole conserved sequence fragment driving the hsp68 promoter, lacZ, and the SV40 poly(A) cassette. Box 1 (Fig. 2A) shows the most extensive homology: 94% identity over nearly 800 bp, between human and mouse. Arrows in Fig. 2A indicate two short sequence blocks, 41 and 58 bp, with 92 and 87% identity, respectively, between human or mouse and Fugu. Box 2 (84% identical over 182 bp core) and Box 3 (82% identical over 242 bp core) show conservation among mammals only, but not over the extended phylogenetic distance between mammals and fish (Fig. 2A).

or two lines. The developmental expression pattern consistently observed in all the expressing lines with the B123-Z reporter construct is illustrated in Fig. 3 for the single representative line A220. Restricted spatial and temporal lacZ expression is observed in the pretectum (p), neural retina (r) and olfactory region (o).

In all expressing lines analysed lacZ staining is first detected at E10.5 in the developing diencephalon in prosomere 1 (p1) which develops into the pretectum (p) (Figs. 2B and 3A, L). Prosomere 1 is located at the caudal part of the diencephalon which defines the forebrain–midbrain boundary. The staining pattern observed at this boundary is a very specific subset of the total endogenous Pax6 expression which extends over all three prosomeres (p1–p3) at this developmental stage (Warren and Price, 1997). The dense positively stained cell layer is located superficially where differentiating neurons reside (Fig. 3L–N). At E12.5 embryos exhibit an increase in pretectum lacZ activity (compare Fig. 3A, B). This is more clearly identified in sagittal sections (Fig. 3L, M). As development proceeds expression of the transgene in the pretectum is reduced (see E13.5–E15.5, Fig. 3C, D, F, G, N) and disappears completely by E18.5 (data not shown).

LacZ expression was also consistently seen in the developing eye (e). Staining is first detected at E11.5 (Fig. 2B). Sagittal sectioning of E12.5 whole-mount embryos indicates that staining is predominantly in a subpopulation of the outer layer of the neural retina (r) (Fig. 3I), confined to the central portion of the retina (Fig. 3J). No staining was seen in the ectodermal (lens and cornea) eye derivatives where endogenous Pax6 is strongly expressed. At E14.5 lacZ expression localised to the outermost portion of the retina corresponding to the prospective photoreceptor layer (Fig. 3J) and by E15.5 the number of lacZ staining positive cells within the outer layer were reduced (area marked by an asterisk in Fig. 3K). In keeping with the expected endogenous Pax6 expression pattern a clear boundary is observed between the expressing retinal subset of the optic cup and the non-expressing optic stalk (s in Fig. 3J).

Transgenic embryos also exhibited intense lacZ expression in the developing olfactory system (o) beginning at E12.5 (Figs. 2B and 3B). Staining was consistently observed in the developing olfactory region in five expressing lines between E12.5 (Fig. 3B–D, G) and E18.5 (Fig. 3H). Histological sections of a representative E14.5 embryo (Fig. 3E, F) show lacZ expression in a subset of cells of the olfactory bulb (ob) and olfactory epithelium (oe). This expression pattern is consistent with the endogenous Pax6 RNA expression pattern observed in the olfactory region (Stoykova and Gruss, 1994). It is known that cell types arise locally in the ventricular zone of the developing olfactory bulb and that the interneurons proliferate primarily along the subventricular zone of the lateral ventricles (Grindley et al., 1995). These progenitors migrate along a well-defined path known as the rostral migratory stream of the olfactory
Fig. 2. Identification of \textit{PAX6} C1170 conserved sequence elements responsible for the transgene expression in the developing pretectum, neural retina and olfactory region. (A) A VISTA plot showing a three-way comparison of human \textit{PAX6} C1170 region with orthologous mouse and \textit{Fugu} DNA sequences. The x-axis represents the base sequence. This corresponds to 10 kb of human DNA sequence (205–215k) in Fig. 1 for the top and middle alignments and mouse DNA sequence for the bottom alignment. The y-axis represents the percent identity. Conserved non-coding regions are highlighted under the curve. The parameters used select contiguous subsegments of at least 40 bp and minimum 50% identity. The approximate location and size of the conserved human elements (Box 123, Box 1 and Box 3) used to generate transgenic mice are indicated above the plot. Vertical down-arrows represent the two short non-coding sequences conserved across phyla. (B) A map of the reporter constructs used is shown to the left. B123-Z contains Box 123, 2.932 kb of human DNA; B1-Z contains Box 1, a 0.968 kb subset of Box 123; B3-Z contains Box 3, a 0.990 kb subset of this region. The expression patterns observed using these constructs are summarised in the columns to the right. Transgene expression was observed in developing pretectum, neural retina and olfactory region. Transient transgenic assays were performed for the construct B1-Z. Abbreviation: nd, not determined.
bulb, where they differentiate into neurons between E13 and E18. In keeping with this morphology Fig. 3G shows clear lacZ staining at the frontal end of the rostral migratory stream which terminates in the olfactory bulb (ob in Fig. 3G).

Apparently ectopic expression of the transgene in the midbrain (m) was observed in some lines (e.g. line A220, Fig. 3A–D, F–H). This extension of Pax6 expression beyond the usually tight forebrain–midbrain boundary may reveal a fundamental mechanism of expression control. The normal interaction of a broadly acting permissive element with a restrictive negative element may be disrupted in such transgenic reporter studies where normal cis-element interactions are absent.

2.3. A 990 bp fragment is sufficient for directing transgenic expression in the developing neural retina

Cross-species sequence comparisons highlighted three distinct subsets of the C1170 region displaying potentially significant sequence identity. To dissect the individual regulatory components of the complex pattern generated by the 2.932 kb B123-Z, we constructed B1-Z and B3-Z (Fig. 2B). The sub-fragments were selected on the basis of sequence conservation across phyla (Fig. 2A). Fifteen independent transgenic embryos, identified by polymerase chain reaction (PCR), were obtained using the B1-Z construct. Despite the high sequence conservation of the 968 bp Box 1, no consistent expression was observed at E12.5 and E13.5 (Fig. 2B).

B3-Z, containing the distal 990 bp Box 3 sequence, reveals an ocular expression which mirrors the spatiotemporal pattern observed with the full-length Box 123-Z (line D94, Fig. 4A–D; line A220, Fig. 3). Analysis of nine transgenic expressing lines first detects lacZ activity between E11.5 and E12.5, restricted to the central portion of the neural retina (Fig. 2B). Expression remains strong until E14.5 and then decreases in intensity by E15.5 as for the larger Box 123 construct (Figs. 2B and 4A–D, data not shown). No expression is detected in the peripheral neural retina or in the developing optic nerve, lens or cornea. Thus, human–mouse sequence homology has identified a new regulatory component, Box 3, within the human C1170 region. Box 3 is sufficient for the regulation of a distinct subset of neural retina expression while the phylogenetically more distantly conserved Box 1 (Fig. 2A) fails to confer enhancer function on its own (Fig. 2B).

2.4. Pufferfish and human C1170 genomic regions both direct similar expression patterns in transgenic mice

Pax6 regulatory mechanisms have been shown to be functionally conserved during evolution (Hauck et al., 1999; Xu et al., 1999). Therefore, we tested the ability of the topologically equivalent Box 123 Fugu region to direct reporter gene expression in transgenic mice (Fig. 5A). This Fugu region is located 17 kb 3′ of the Fugu Pax6 coding...
Fig. 5. LacZ expression patterns directed by the Fugu reporter construct FB123-Z, containing the presumptive FBox 123. (A) Schematic representation of Fugu Pax6 genomic organisation. Arrows indicate the opposite direction of transcription for Pax6 and Paxneb. The Fugu Pax6 exons and exons 10 and 9 of Paxneb are shown as numbered boxes. Fugu Pax6 exons 1–12 were found on a single cosmid 151J19 (EMBL: AL021531). An overlapping BAC, 293E01 (EMBL: AJ414048), contains downstream sequences including exons 10 and 9 of Paxneb. The Fugu genomic sequence that is topologically equivalent to the human C1170 Box 123 region is located 17 kb downstream of the Fugu Pax6 coding sequence. (B) The VISTA plot showing vertebrate C1170 Box 123 sequence comparisons was used as a guide to select a 1.404 kb Fugu region for incorporation into the reporter construct FB123-Z. (C–E) Arrows highlight specific staining observed driven by this construct in pretectum (p), olfactory region (o) and eye (e) at indicated stages. Neural retina (r) expression is shown at E13.5 (F). Arrowheads indicate the presence of additional midbrain (m) and lens (l) lacZ expression in (E) and (F). (D,F) Sagittal plane vibratome sections (150 μm).
region within the *Paxneb* final intron which covers 18 kb, and is unusually large for *Fugu* (Fig. 5A). The size of the fish sequence used was 50% smaller than its human counterpart: 1.404 and 2.932 kb, respectively (Fig. 5B), in keeping with the genome compaction reported for this region (Miles et al., 1998). At E11.5, four lacZ stained transient transgenic embryos consistently showed expression in the pretectum (p) and olfactory region (o) (Fig. 5C, D), which remains strong in two lines examined at E13.5 (Fig. 5E). At this developmental stage additional expression is seen in the neural retina (r) (Fig. 5F). The significance of staining observed in the lens at E13.5 is unclear and may represent stain trapping (Fig. 5F). As with the human B123 construct, ectopic midbrain expression is also seen at E13.5 (Fig. 5E). Despite the low level of overall sequence conservation there is a similar spatio-temporal expression pattern with homologous human and Fugu constructs (Figs. 3 and 5).

3. Discussion

To begin to understand the transcriptional mechanisms responsible for regulating the correct levels of *Pax6* in a tissue- and developmental-stage specific manner we have looked for novel cis-regulatory elements associated with the *Pax6* gene. In this study, we focus on genomic sequence approximately 77 kb downstream of the human *PAX6* polyA-addition site. Short sequences of this downstream region share high homology with orthologous mouse and *Fugu* sequence. The functional significance of this sequence conservation was explored using the mouse transgenic reporter approach. Our analysis revealed that this conserved human sequence, C1170 Box 123, consistently directs reporter expression to discrete parts of the brain, eye and olfactory region during development, in a restricted temporal manner, recapitulating part of the normal *Pax6* expression pattern. The orthologous *Fugu* sequence exhibited a similar expression pattern. Using the results of cross-species sequence comparisons as a guide we have identified a subregion, Box 3, that is responsible for restricted endogenous neural retina *Pax6* expression. Further analysis may localise the individual olfactory and pretectum elements within the C1170 sequence. The absence of detectable expression pattern using the Box 1-Z construct suggests that additional sequence may be required to confer function on this element. Molecular identification of new cis-regulatory elements located downstream of the *PAX6* coding region confirms recent somatic cell hybrid work which shows that 3’ sequences are essential for the full *PAX6* expression (Lauderdale et al., 2000). This raises the possibility that mutations affecting downstream regulatory elements may be associated with particular clinical phenotypes.

Consistent with the observation that *Pax6* is essential for the normal development and regionalisation of the diencephalon (Grindley et al., 1997), the present study describes a new long-range cis-acting brain element. All results suggest that this element enhances *Pax6* transcription in a restricted subset of the total endogenous *Pax6* expression in prosomere 1, at the boundary of the forebrain and the midbrain, at developmental stages E10.5–E15.5 when *Pax6* is normally widely expressed in all three prosomeres (p1–p3) of the diencephalon. Ectopic midbrain expression may be a normal function of this element that is normally curbed by superimposed negative regulation when the whole locus is intact. This new p1-specific cis-regulatory element will be a useful tool to further our understanding of how the diencephalon–mesencephalon boundary, which is lost in *Small eye* mice, is established (Warren and Price, 1997). Overlapping domains of expression have been demonstrated for a second forebrain enhancer element located 150 kb downstream of the *PAX6* polyA-addition site. This element is capable of driving diencephalic expression in transgenic mice in prosomere 1 from E9.5 throughout development (Kleijnen et al., 2001). A third mouse *Pax6* region located between exon 0 and exon 4 was also observed to drive pretectum expression (Kammandel et al., 1999). Therefore, multiple diencephalic enhancers are required to direct the correct spatio-temporal *Pax6* expression. Interestingly, the upstream telencephalon element was also reported to drive ectopic midbrain expression in a proportion of cases (Kammandel et al., 1999).

This study further reveals that expression of *Pax6* in distinct regions of the developing olfactory bulb and epithelium from E12.5 onwards is directed by the human C1170 Box 123 region. A regulatory element located 5’ of exon 0 has also been shown to direct restricted olfactory *Pax6* expression between E12.5 and E13.5 (Kammandel et al., 1999). Both nasal enhancers account for only a subset of the total *Pax6* olfactory transcripts demonstrated by in situ hybridisation experiments, which show expression commencing earlier in development at E8.5, coincident with the formation of the nasal placode (Walther and Gruss, 1991; Grindley et al., 1995). Therefore, we predict that additional control elements in the *Pax6* locus are necessary to direct the earlier endogenous *Pax6* nasal expression pattern (Kammandel et al., 1999).

Individual sequence modules located in intron 4 are required for *Pax6* expression in the neural retina of different species, including pufferfish (Plaza et al., 1995b; Kammandel et al., 1999; Xu et al., 1999). In mice carrying this neural retina regulatory element expression is predominantly confined to the peripheral retina, whereas the central retina appears negative (Kammandel et al., 1999). In contrast, using the C1170 Box 123 neural retina control element described in this study, reporter transgene expression is confined to the central retina and absent in the peripheral retina between E11.5 and E15.5 of development. The latter is a subset of that recently reported for the *PAX6* DRR which directs expression throughout the entire retina from E10.5 to adulthood (Kleijnen et al., 2001). To date numerous cis-regulatory elements have been shown to drive *Pax6*
expression in the developing neural retina. Further co-localisation experiments are required to study the possibility that elements acting as enhancers in the neural retina in an overlapping temporal pattern might regulate Pax6 expression in discrete cell types within this tissue.

The mechanism by which the newly identified Pax6 cis-regulatory elements act over a distance of >77 kb is not understood. Although strong mammalian sequence conservation was used to highlight the functional role of the C1170 element, sequence similarity is much reduced when the phylogenetic distance is extended to fish, although functional conservation is preserved. This implies that upstream transactivating factors and their target sequences have been conserved throughout evolution. Computer assisted analysis of Box 123 sequence identified multiple POU domain binding sites conserved between human and fish (Quandt et al., 1995). Therefore, POU domain transcription factors may be positive or negative regulators of Pax6 expression. Bm3b (Pou4f2) is required for the activity of the quail Pax6 EP enhancer (Plaza et al., 1999). Further studies, such as the analysis of transcription factor binding sites using in vivo footprinting and gel shift experiments will be required to confirm the validity of computer predictions. Such approaches have demonstrated that binding of Pax2 protein to the Pax6 intron 4 neural retina enhancer represses Pax6 expression in the optic stalk, while in the absence of Pax2 binding Pax6 expression is permitted in the optic cup (Schwarz et al., 2000). In this way, Pax2 and Pax6 contribute to the regionalisation and subsequent boundary formation between the optic cup and optic stalk observed in the developing visual system.

The potential value of using cross-species sequence comparisons to highlight long-range regulatory elements in the non-coding genome has been widely advocated (Aparicio et al., 1995; Dubchak et al., 2000; Loots et al., 2000). Many of the computationally identified non-coding sequences have been tested in vivo by a reductionist transgenic reporter approach and some have been shown to regulate a specific subset of the full spatio-temporal expression pattern observed for the respective gene (Gottgens et al., 1995; Dubchak et al., 2000; Loots et al., 2000). However, there are conserved sequence sites, which fail to direct reporter expression in this way including Box 1 sequence described here (Loots et al., 2000). Additional limitations of the transgenic approach include its failure to detect negative regulatory elements, origins of replication and signals for nuclear sub-compartmentalisation. It also ignores the importance of the spatial organisation of the gene within the native chromatin environment. A converse approach to studying the function of isolated elements is to assess the phenotypic effects of removing them by targeted deletion from the endogenous genomic locus. However, a recent targeting study showed that deletion of the mouse Pax6 upstream ectodermal enhancer generated only mild phenotypic effects (Dimanlig et al., 2001). These findings suggest that isolated enhancer elements cannot adequately define the expression and the function of a gene. Spatio-temporal expression is controlled through complex interactions between multiple cis-regulatory elements.

4. Materials and methods

4.1. Sequence analysis

Human DNA sequence for cosmids C1170 and H1281 was obtained from EMBL (Z83306 and Z83309, respectively). Working draft mouse DNA sequence was obtained from the Mouse Sequencing Group, HGMP-RC, Hinxton, (EMBL: AL512589). Fugu sequencing over a large conserved syntenic region around the Pax6 gene, was carried out by the group of Greg Elgar, MRC HGMP-RC, Hinxton (Miles et al., 1998; unpublished). Comparative genomic analysis was carried out using the PipMaker (http://bio.cse.psu.edu/pipmaker Schwartz et al., 2000) and VISTA (http://www-gsd.lbl.gov/vista/ Brudno et al., 2000; Mayor et al., 2000) programs. We masked the sequences for high complexity repeats using RepeatMasker (http://ftp.release.Fugu.EDU/CGI-BIN/RepeatMasker) before performing the analysis. Repeat-masked sequences were analysed for exon content using Genescan (http://CRR-081.MIT.EDU/Genescan.html), Grail (http://grail.isd.ornl.gov/Grail-1-3) and BLAST database searches using fragments of either the human, mouse or Fugu sequences.

4.2. Reporter construct preparation

The modified hsp68-lacZ vector p610+ (described in Kleinjan et al., 2001) was used to generate lacZ reporter constructs B123-Z, B1-Z and B3-Z shown in Fig. 4. Vector p610+ contains promoter sequences from the mouse hsp68 gene and includes the translational start codon fused in-frame to a lacZ open reading frame and an SV40 polyadenylation signal (Kothary et al., 1989). To create the lacZ reporter plasmids B123-Z, B1-Z and B3-Z, high fidelity Taq DNA polymerase (Roche) was used to PCR amplify, from cosmid DNA clone C1170 (van Heyningen and Little, 1995), regions from nucleotides 36,058–38,990, 36,530–37,498 and 38,000–38,990, respectively. For each construction, the two synthetic primers used for the amplification were designed to introduce 5′ Not I/Kpn I and 3′ Sal I restriction sites (underlined within the primer sequence) at the ends of the PCR products. Primer pairs for amplifying human Box 123 sequence were (FP1 5′-TATGGGCGCCGCTAGTTTCCCTGTACCTCCCGCTGG-3′) and (RP1 5′-A-CATGTCGACTATGAGCAGCGAGTCTAAATAGG-3′), for Box 1 ((FP2 5′-ACATGCGCCGCTATTTTCTGCGCAGGAAATGAT-3′) and (RP2 5′-ACATGTCGACGTGTCGCCCGCTTACATTTAAAAATGAT-3′)) and for Box 3 ((FP3 5′-CATGCGCCGCTTGGTACATTTCTGGCAGGAAATGAT-3′) and (RP3 5′-CATGCGCCGCTTGGTACATTTCTGGCAGGAAATGAT-3′)). To generate the orthologous Fugu construct FB123-Z, primer pairs used were FP4 5′-GATGGTACCAACACAGCTGGGATTTCAAGGGTA-3′ and
RP 4 5′-GCATGTCGACAGCTGATGCTGCTACGT-
TAGCA-3′. The template used was BAC 293E01 (EMBL: AJ414048). The amplified fragments were digested with restriction enzyme and directionally cloned into the poly-
linker of the p610 + vector upstream of the hsp68 promoter.

4.3. Production and genotyping of transgenic mice

Microinjection fragments, prepared from lacZ-fusion gene constructs by digestion with restriction enzymes (Nru I and Sma I (human) or Kpn I and Sal I (Fugu)), were gel purified using a Qiagen gel extraction kit. Transgenic animals were produced by microinjecting DNA fragments into mouse zygotes using standard procedures (Hogan et al., 1994). All procedures were performed under UK Home Office regulations. Integration of the transgene was confirmed by PCR using primers to lacZ (forward primer 5′-GTGTCGACAGCTGATGCTGCTACGT-
TAGCA-3′, reverse primer 5′-GCCGTCACTTCAACGCAGCA-3′) using genomic DNA prepared either from yolk sac or from tail biopsies.

4.4. Whole-mount beta galactosidase staining and histological analysis

Embryos allowed to develop to the appropriate developmental stage were dissected from the uterus, washed in phosphate buffered saline (PBS) and fixed for 1 h in a PBS solution containing 1% formaldehyde, 0.2% glutaraldehyde, 2 mM MgCl₂, 5 mM EGTA and 0.02% NP-40. Embryos were washed in PBS containing 0.02% NP-40, before staining overnight at 37°C in the dark in a solution containing 5 mM K₂Fe(CN)₆, 5 mM K₄Fe(CN)₆·3H₂O, 2 mM MgCl₂, 0.01% sodium deoxycholate, 0.02% NP-40 and 0.1% 5-bromo-4-chloro-3-indolyl-galactopyranoside. Stained samples were rinsed in PBS, dehydrated through ethanol, cleared in xylene and embedded in paraffin. Histological sections were cut at 8 μm and counterstained with eosin. Some embryos were prepared for vibratome sectioning (150 μm) by sequential overnight equilibration in 4% sucrose–PBS solution, 20% sucrose–PBS solution and a bovine serum albumin (BSA)–G mixture (90 ml PBS, 0.44 g gelatine, 14 g BSA and 18 g sucrose). Embryos were transferred to 25% glutaraldehyde for 20–40 min and were then returned to BSA–G mixture which was set for embedding by the addition of 100 μl of 25% glutaraldehyde per ml BSA–G.

Acknowledgements

We are grateful to Greg Elgar and Philippe Gautier for Fugu sequence data and interpretation. We wish to thank David Price for helpful discussions and Douglas Stewart for figure preparation. This work was supported by an EMBO long-term fellowship to Caroline Griffin.

References


Talbot, C., Grif