Conserved and acquired features of *neurogenin1* regulation

Patrick Blader*,†, Chen Sok Lam*, Sepand Rastegar*, Raffaella Scardigli‡, Jean-Christophe Nicod, Nicolas Simplicio§, Charles Plessy, Nadine Fischer, Carol Schuurmans¶, François Guillemot§ and Uwe Strähle,††**

Institut de Génétique et de Biologie Moléculaire et Cellulaire, 1 rue Laurent Fries, 67404 Illkirch Cedex, CU de Strasbourg, France

*These authors contributed equally to this work
†Present address: Centre de Biologie du Developpement, bat 4R3, Universite Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France
‡Present address: Institute of Cell Biology and Tissue Engineering, Via di Castel Romano 100/102, 00128 Rome, Italy
§Present address: Division of Molecular Neurobiology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK
¶Present address: 2277 HSC, University of Calgary, 3330 Hospital Drive NW, Calgary, Alberta T2N 1N8, Canada
**Present address: Institute for Toxicology and Genetics, Forschungszentrum Karlsruhe, Postfach 3640, 76021 Karlsruhe, Germany
††Author for correspondence (e-mail: uwe.straehle@itg.fzk.de)

Accepted 15 September 2004

Development 131, 5627-5637
Published by The Company of Biologists 2004
doi:10.1242/dev.01455

Summary

The telencephalon shows vast morphological variations among different vertebrate groups. The transcription factor *neurogenin1* (*ngn1*) controls neurogenesis in the mouse pallium and is also expressed in the dorsal telencephalon of the evolutionary distant zebrafish. The upstream regions of the zebrafish and mammalian *ngn1* loci harbour several stretches of conserved sequences. Here, we show that the upstream region of zebrafish *ngn1* is capable of faithfully recapitulating endogenous expression in the zebrafish and mouse telencephalon. A single conserved regulatory region is essential for dorsal telencephalic expression in the zebrafish, and for expression in the dorsal pallium of the mouse. However, a second conserved region that is inactive in the fish telencephalon is necessary for expression in the lateral pallium of mouse embryos. This regulatory region, which drives expression in the zebrafish diencephalon and hindbrain, is dependent on *Pax6* activity and binds recombinant Pax6 in vitro. Thus, the regulatory elements of *ngn1* appear to be conserved among vertebrates, with certain differences being incorporated in the utilisation of these enhancers, for the acquisition of more advanced features in amniotes. Our data provide evidence for the co-option of regulatory regions as a mechanism of evolutionary diversification of expression patterns, and suggest that an alteration in *Pax6* expression was crucial in neocortex evolution.

Key words: *neurogenin1*, Pax6, Telencephalon, Diencephalon, Evolution, Neurogenesis, Transcription, Regulatory elements, Co-option, Zebrafish, Mouse

Introduction

In vertebrates, neurogenesis is a complex process that controls the position, number, connectivity and neurophysiological properties of thousands of different neurones. It has become obvious in recent years that many features of the molecular mechanisms and principles controlling neurogenesis have been conserved throughout animal evolution. The neurogenic processes in the peripheral nervous system of the fruit fly *Drosophila melanogaster* served as important paradigms to unravel the regulatory principles of vertebrate neurogenesis (Bertrand et al., 2002; Chitnis, 1999; Hassan and Bellen, 2000). The *Drosophila* peripheral nervous system is specified by a hierarchical series of cell-fate decisions, in which pre-pattern genes first define regions, the so-called proneural clusters, that express proneural basic helix-loop-helix (bHLH) proteins and have the potential to develop into neurones (Chitnis, 1999). In the subsequent step of lateral inhibition that involves the Delta/Notch signalling system, committed neural precursors are selected from these proneural regions (Simpson, 1997). The regulatory events in the neural plate of lower vertebrates bear a strong resemblance to these patterning mechanisms in the imaginal discs of *Drosophila*. As in *Drosophila*, bHLH transcription factors related to the proneural genes demarcate regions of neuronal precursors, from which, by a Delta/Notch-mediated process, committed neurones are selected (Blader et al., 1997; Chitnis, 1999; Ma et al., 1996).

Central, but less well understood questions concern how the spatial aspects of neurogenesis are controlled in the vertebrate neural plate/tube and how these mechanisms have been modified during vertebrate evolution. Neurogenins (ngns), which belong to the bHLH family and which are most closely related to the *Drosophila* bHLH factors Biparous and Atonal, demarcate the regions of primary neurogenesis in the neural plate of zebrafish and *Xenopus* embryos (Blader et al., 1997; Ma et al., 1996). Accumulating evidence suggests that these *ngn1*-expressing regions in the neural plate of lower vertebrates are specified by pre-pattern genes that provide positional codes with their large and overlapping domains of expression (Bally-Cuif and Hammerschmidt, 2003). Genes of the *iroquois* family, members of which are involved in positioning the expression of the proneural genes *achaete* and *scute* in the *Drosophila* imaginal disc, have been suggested to play similar roles in the vertebrate neural plate (Bellefroid et al., 1998; Cheng et al.,

2001; Gomez-Skarmeta et al., 1998; Itoh et al., 2002; Wang et al., 2001). In *Xenopus* embryos, the zinc-finger transcription factor Zic2 acts as a repressor of *ngn1* expression in the longitudinal stripes, separating the *ngn1*-expressing precursors of motor neurones, interneurones and sensory neurones (Brewster et al., 1998). Another negatively acting factor, the bHLH factor Her5, prevents neurogenesis at the midbrain/hindbrain boundary (MHB). Inhibition of Her5 function leads to an expansion of *ngn1* expression and ectopic formation of a proneural field in the MHB area (Geling et al., 2003).

In the neural tube of mouse embryos, the paired-homeodomain transcription factor Pax6 has important regulatory functions (Ashery-Padan and Gruss, 2001; Simpson and Price, 2002). Pax6 mutant mice develop small eyes and have deficiency in neurogenesis in the brain and spinal cord. In both the telencephalon and the spinal cord, Pax6 is expressed in a graded fashion, suggesting that it provides concentration-dependent positional information for the region-specific differentiation of neural tissues (Stoykova et al., 2000; Scardigli et al., 2003). The zebrafish genome encodes two *pax6*-related genes, *pax6.1* and *pax6.2*, that are expressed in overlapping domains in the eye, dorsal diencephalon, hindbrain, spinal cord and pancreas, in a pattern reminiscent of the pattern of *pax6* expression in the mouse embryo (MacDonald et al., 1994; Nornes et al., 1998). Interestingly, prominent expression of *pax6.1* and *pax6.2* was not detected in the proliferating telencephalon of the zebrafish, in contrast to mouse embryos (MacDonald et al., 1994; Stoykova et al., 2000; Wullimann and Rink, 2002). The *pax6*-expressing cells in the zebrafish telencephalon constitute a small population of migrating, post-mitotic cells at the pallial/subpallial border. By contrast, *pax6* is abundantly expressed in the proliferating radial glia cells of the mouse telencephalon. This suggests that the pattern of *pax6* expression was modified during vertebrate brain evolution.

Expression of *ngn2* in the mouse depends on *pax6* activity in both the spinal cord and the telencephalon (Stoykova et al., 2000; Scardigli et al., 2003). It was recently shown that the *ngn2* upstream region contains a *pax6*-dependent regulatory region that drives expression in the spinal cord (Scardigli et al., 2003). Moreover, *ngn2* expression in the telencephalon of the mouse depends on *pax6* activity (Stoykova et al., 2000). Based on the expression of *pax6.1* and *pax6.2* in the zebrafish neural plate/tube, *pax6* may possess equivalent functions in the control of the related *ngn1* gene during primary neurogenesis in the zebrafish embryo.

We have previously mapped the regulatory regions of *ngn1* responsible for driving reporter expression in the neural plate. Two regions, the *lateral stripe element* (*LSE*) and the *anterior neural plate element* (*ANPE*) were identified in the *ngn1* upstream region (Blader et al., 2003). The *LSE* is required for expression in precursors of Rohon Beard sensory neurones and reticulospinal neurones in the anlage of the spinal cord and hindbrain, respectively; the *ANPE* is responsible for expression in the ventral caudal cluster in the midbrain anlage, the trigeminal ganglia and a few scattered nuclei in the anterior hindbrain (Blader et al., 2003). Further analysis of *ANPE* showed that it contained an E-box known to interact with bHLH factors. Indeed, Her5 was demonstrated to regulate the activity of *ANPE*, as in embryos that lacked Her5, expression of a transgene that contained the *ANPE* was expanded into the MHB area (Geling et al., 2004). Moreover, mutation of the E-box in the *ANPE* caused an expansion of reporter gene expression into the MHB area, suggesting that the E-box is required for the suppression of transgene activity in the MHB by Her5 (Geling et al., 2004).

These cis-regulatory regions show homology with sequences at the mouse and human *ngn1* loci, despite the fact that mammals do not express *ngn1* in the neural plate, but only later in the neural tube (Blader et al., 2003; Simmons et al., 2001). This suggests that these regions have shared functions in neurogenesis in mammals and teleosts. The preliminary analysis of transgenes lacking the *LSE* and *ANPE* in post-somitogenesis-stage zebrafish embryos suggested that more proximal regions of the zebrafish *ngn1* gene have regulatory activity at later stages when the neural tube has formed (Blader et al., 2003).

To delineate the regulatory regions responsible for brain expression of *ngn1* in older zebrafish embryos, we analysed transgenic lines carrying wild-type and deletion variants of *ngn1* transgenes. We mapped two regulatory regions that are required for transgene expression in the brain of post-somitogenesis-stage embryos. The first region, residing at position –6702 to –6490 bp upstream of the ATG start site, which was also previously shown to harbour the *LSE*, drives expression in the dorsal telencephalon. A second regulatory region referred to as *LATE* was mapped to position –1775 to –1368. The *LATE* region, like the *LSE*, is highly conserved in mouse and human homologues of *ngn1*. We carried out comparative functional studies in mouse embryos to investigate the activity of these conserved regulatory elements. We focused on the dorsal telencephalon of the mouse, as this is undoubtedly the most derived brain region to have arisen during vertebrate evolution (Nieuwenhuys, 1994; Wullimann and Rink, 2002). The *LSE* drives expression in the dorsal telencephalon in both mouse and zebrafish embryos, indicating a conserved function with respect to telencephalic expression. Curiously, we found that the *LATE* region of the zebrafish *ngn1* gene drives expression in the lateral telencephalon of the mouse embryo but not in the zebrafish telencephalon. The area of activity of *LATE* overlaps with that of the paired-homeodomain transcription factor Pax6, suggesting a role of Pax6 in regulating the activity of *LATE*. We demonstrate that Pax6 binds to a conserved Pax6-binding site in the *LATE* region. Moreover, the lack of *pax6* activity in zebrafish by simultaneous knockdown of both *pax6.1* and *pax6.2* leads to a small eye phenotype and strongly reduces endogenous *ngn1* and transgene expression. These results are consistent with a direct regulatory role of Pax6 on the activity of *LATE*. Based on the highly modular structure of vertebrate regulatory regions, which are usually composed of multiple short and degenerate binding sites for transcription factors, it is commonly assumed that elaboration of novel patterns of gene expression is accomplished by changes in the regulatory sequence (Ludwig, 2002; Stone and Wray, 2001; Tautz, 2000). Our data suggest that a pre-existing enhancer was co-opted, and that the evolution of the *pax6* expression pattern led to the recruitment of *LATE* into the newly developed territories of the mouse telencephalon.

Materials and methods

Reporter constructs

The *–8.4, –8.4(del1-9), –6.3, –5.9* and *–3.1 ngn1:gfp* transgenes were published previously (Blader et al., 2003). Versions of all reporter constructs were also made using a nuclear-localised β-galactosidase reporter. The *delLATE* deletions were generated by a PCR strategy. Two PCR products 5′ and 3′ of the homology region were amplified using the oligonucleotide pairs: 5′-TAATACCCGGGGATTAATGC-3′/5′-GATCGTCGACCACCCCGCTTCTGAGACACG-3′; and 5′-CTGAGTCGACAATAAAACTTAAGCCACTGG-3′/5′-CTGTCCTG-CATGCAACAAGC-3′. They were ligated using *Sal*I sites, which replace the homology region between –1775 and –1368, and re-amplified with the 5′-TAATACCCGGGGATTAATGC-3′/5′-CTG-TCCTGCATGCAACAAGC-3′ oligonucleotide pair. The resulting PCR fragment was cloned into pBL3.1, by replacing the *Xma*I/*Eco*RI fragment between –3122 and –673. A 2.7-kb *Xma*I/*Nco*I fragment containing the deletion [–3.1del(LATE)] was then cloned into a vector with the GFP:SV40 poly-A cassette. The 8.4(delLATE):GFP was made by introducing a 5283 bp *Not*I/*Xma*I fragment containing the promoter-distal *ngn1* region (Blader et al., 2003) in pBL-3.1del(LATE), and inserting the resulting –8.4del(LATE) *Not*I/*Nco*I fragment into the vector with the GFP:SV40 poly-A cassette. In the zebrafish/mouse chimeric constructs, the regions of conservation (*LSE*, *LATE*) were replaced by the cognate mouse sequence using standard PCR-based cloning strategies. Details on the constructions are available upon request.

Transgenic animals and morpholino knockdown

Reporter fragments for generating transgenic fish were excised from plasmids and separated by agarose gel electrophoresis, followed by purification with the Qiaex II Kit (Qiagen), according to the manufacturer's instructions. Fragments were diluted to 50 ng/μl in TE and injected into freshly fertilised zebrafish embryos, as previously described (Blader et al., 2003). Mouse transgenics were generated as previously described (Scardigli et al., 2001). Morpholinos (GeneTools) were designed complementary to the 5′ region of *pax6.1* (5′-TTTGTATCCTCGCYGAAGTTCTTCG-3′) and *pax6.2* (5′-CTGAGCCCTTCCGAGCAAAACAGTG-3′) mRNA. They were resuspended in 1×Danieau buffer [58 mM NaCl, 0.7 mM KCl, 0.4 mM $MgSO_4$, 0.6 mM $Ca(NO_3)_2$, 5.0 mM HEPES, pH 7.6] and stored at 4°C. The morpholinos were injected into the yolk of zebrafish embryos at the one- to two-cell stage, and at concentrations of between 0.6 mM and 1.2 mM.

In situ analysis

In situ hybridisation and immunohistochemistry was performed using *ngn1* (Blader et al., 1997) and *gfp* (Blader et al., 2003) antisense riboprobes and antibodies directed against *E.coli* β-galactosidase or GFP (Scardigli et al., 2001), respectively.

Sequence comparisons

Human (AC005738), chicken (AY518341) and zebrafish (AF017301) sequences were extracted from GenBank. *Xenopus tropicalis* homologous sequences (scaffold_9761) were found by blasting the genome assembly one on the JGI website (http://genome.jgi-psf.org/cgi-bin/runBlast?db=xenopus0). Mouse sequences were kindly provided by J. Johnson (Nakada et al., 2004). Multiple alignment was performed using T-Coffee (Notredame et al., 2000). Pax6 putative-binding sites (Epstein et al., 1994) were searched using GCG findpattern (Accelrys).

Electrophoretic mobility shift assays

Pax6 protein was produced in vitro using the Sp6TNT kit (Promega) according to the manufacturer's protocol. Oligonucleotides containing the site C sequence (5′-GGCTTTGATATATCATACATGCCTGAA-GACTCCC-3′), or clusters of point mutations (shown in bold) (5′-GGCTTTGATA**GCGACGCACGTAAGTCC**GACTCCC-3′) were annealed by heating to 90°C with an equimolar mixture of the upper and the lower strands, and cooling slowly to room temperature. Annealed oligonucleotides were labelled with [γ^{32}ATP] using T4 polynucleotide kinase. Binding reactions were performed in a total volume of 25 μl containing 10 mM Hepes (pH 7.9), 100 mM KCl, 4% Ficoll, 1 mM EDTA, 1 mM DTT and 2.5 μg of poly(dI-dC). The reactions contained 4 μl of protein and 30,000 cpm of probe. The reactions were allowed to proceed for 30 minutes at 4°C and were analysed on a 6% polyacrylamide gel containing 0.25×Tris borate-EDTA (TBE).

Results

The zebrafish 8.4-kb *ngn1* upstream regulatory region recapitulates expression of endogenous transcripts

To identify the regulatory elements for forebrain expression, transgenic zebrafish that harbour the 8.4-kb sequence upstream of the coding region of the zebrafish *ngn1* locus fused to a *green fluorescent protein* (*gfp*) reporter were analysed (Blader et al., 1997; Blader et al., 2003). GFP expression was assayed at 28 hours post-fertilisation (hpf), a stage when the forebrain is predominantly proliferative and expresses endogenous *ngn1* (Korzh et al., 1998; Wullimann and Knipp, 2000). Four stable transgenic lines were analysed, whose embryos displayed identical patterns of GFP expression with strong activity in the telencephalon (Fig. 1A). In general, endogenous *ngn1* transcripts are abundantly expressed at this stage and present a complex pattern of expression (Fig. 1B). Besides the telencephalon, *ngn1* transcripts are also located in several clusters within the diencephalon. At least six populations occupying distinct dorsoventral regions can be recognised on whole-mounted embryos (Fig. 1B). In the dorsal diencephalon, *ngn1* mRNA is detected in the epiphysis and the pretectum. A larger, more prominent cluster in the dorsal thalamus,

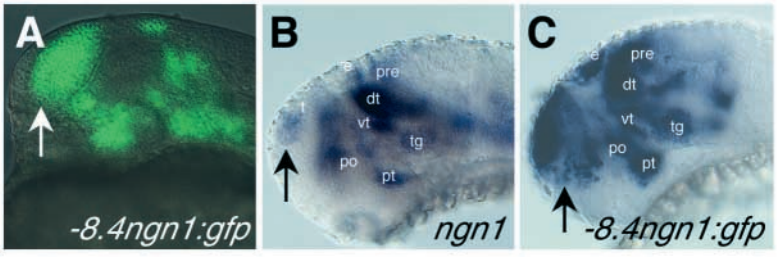

Fig. 1. The *–8.4ngn1:gfp* zebrafish transgene recapitulates the pattern of endogenous *ngn1* transcripts and drives telencephalic expression in the zebrafish embryo. (A) The 8.4 kb upstream zebrafish *ngn1* regulatory sequence drives GFP expression in the telencephalon (arrow) and in the diencephalon. (B,C) Comparison of the endogenous *ngn1* gene (B) and the *gfp* transgene (C) by in situ hybridisation indicates that the *gfp* transgene is capable of recapitulating the endogenous telencephalic (t, indicated by arrows) and diencephalic expression of *ngn1*. Expression is detected in at least six regions of the diencephalon, comprising the epiphysis (e), pretectum (pre), dorsal thalamus (dt), ventral thalamus (vt), preoptic area (po) and posterior tuberculum (pt). Transcripts are also localised in the midbrain tegmentum (tg). Zebrafish embryos are at stage 28 hpf. Panels show lateral views of whole-mounted embryos oriented anterior to the left and dorsal up.

presumably also comprising the future habenular nuclei, is located immediately adjacent to the epiphysis. Further ventrally, *ngn1* is found in the ventral thalamus, the preoptic area and the posterior tuberculum. In the mesencephalon, transcripts are localised predominantly in the tegmentum. With the exception of the epiphysis, these domains of expression broadly agree with the regions identified by Mueller and Wullimann (Mueller and Wullimann, 2003) based on the analysis of sections at 48 hpf. Comparison of the pattern of expression of the endogenous *ngn1* gene with that of the *–8.4ngn1:gfp* transgene indicates that the domains demarcated by GFP reproduced the pattern of endogenous *ngn1* mRNA (Fig. 1A,C). The expression of the transgene is more intense than that of the endogenous gene. However, this difference appears to be due to the higher copy number of the transgene, and to the higher stability of the *gfp* reporter mRNA in comparison with the endogenous *ngn1* mRNA, which is expressed in a highly dynamic pattern (Blader et al., 1997).

The regulatory region required for telencephalic expression maps to the *LSE*

To identify sequences responsible for the activity of the *–8.4ngn1:gfp* transgene in the zebrafish telencephalon, 5′ deletions retaining 6.9 kb (1 transgenic line), 5.3 kb (1 transgenic line) and 3.1 kb (2 transgenic lines) of the original 8.4 kb fragment upstream of the GFP reporter (Blader et al., 2003) were analysed. Telencephalic expression of GFP was lost in the transgenic line containing 5.3 kb of upstream sequence, indicating that the required regulatory elements lie between –6.9 and –5.3 kb upstream of the *ngn1* start codon (Fig. 2A,B). The *LSE*, which drives expression in the lateral neural plate, was previously mapped to this region (Blader et al., 2003). To assess whether the *LSE* and the region mediating telencephalic expression co-localise, a series of eight overlapping 400-bp deletions spanning this region [*–8.4(del1-9)ngn1:gfp*] was analysed (Fig. 2C). The stable lines carrying *del2* (–6886 to –6490 nucleotides) and *del3* (–6702 to –6294 nucleotides) were found to lack telencephalic GFP expression (Fig. 2C), identifying sequences between –6702 and –6490 nucleotides upstream of the *ngn1* ATG as necessary for transgene activity in the zebrafish telencephalon. The same region was shown to be required for expression in the lateral neural plate, and was found to be highly homologous (60%) to an upstream sequence of the murine and human *Ngn1* locus (Blader et al., 2003).

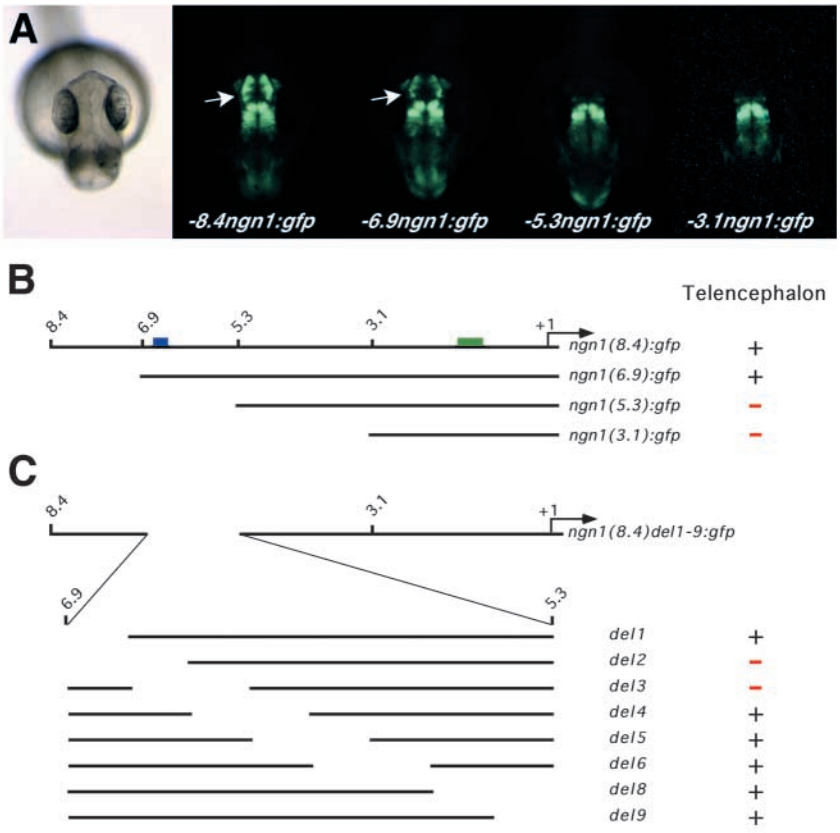

Fig. 2. Identification of the regulatory elements required for telencephalic expression of the *–8.4ngn1:gfp* transgene in the zebrafish. (A,B) Analysis of zebrafish embryos transgenic for truncated versions of the *–8.4ngn1:gfp* reporter fragment. Transgene expression in the telencephalon (arrows) is no longer detected in embryos carrying transgenes with 5.3 kb of *ngn1* sequence. (C) Overlapping deletions spanning the region between 6.9 and 5.3 kb reveal a 200-bp element, the *LSE*, that is required for telencephalic expression of the *–8.4 ngn1:gfp* transgene. The heads of embryos in A are seen from the dorsal aspect, with the telencephalon up, and are staged between 28 and 30 hpf. Boxes in B, corresponding to the *LSE* (blue) and *LATE* (green), indicate regions of homology between the zebrafish, mouse and human *ngn1* loci. Each construct was analysed in one to three transgenic lines. Note that expression patterns from the overlapping deletions provided an independent further confirmation of the results. Transgenic lines varied in the intensity of reporter expression, but no significant variations in the expression pattern due to integration site effects were observed.

The 8.4-kb zebrafish transgene is capable of recapitulating *ngn1* expression in the mouse telencephalon

A potential functional conservation of *ngn1* regulation between mammals and zebrafish was tested by generating transgenic mouse embryos with the zebrafish *–8.4ngn1:gfp* transgene. At 12.5 days post-coitum (dpc), at the onset of neurogenesis in the mouse telencephalon and at a stage comparable to the 28-hpf zebrafish forebrain, a comparison of GFP expression in transgenic embryos with embryos stained with a *Ngn1*-antisense probe indicated that the zebrafish transgene drives reporter expression comparable to endogenous *Ngn1* in the telencephalon (Fig. 3A-C). Furthermore, the transgene reproduced the sharp ventral border, just ventral to the corticostriatal angle, and the lateral-to-medial gradient of *Ngn1* expression seen in the mouse at this stage (Fig. 3B,C). In contrast to the exclusive localisation of *Ngn1* transcripts in the ventricular zone, GFP activity was also detected in the pre-plate layer of postmitotic neurones, probably because of the stability of the reporter protein (Fig. 3C).

The *LSE* is not sufficient to drive the entire expression in the mouse telencephalon

The conservation of the regulatory sequences in mammalian *Ngn1* orthologues, together with the ability of the zebrafish transgene to recapitulate the expression of the mouse *ngn1* gene in the telencephalon, suggests that zebrafish and mammals employ a similar regulatory strategy to drive expression of *Ngn1* in the telencephalon. To confirm that expression of *–8.4ngn1:gfp* in the mouse telencephalon does require the *LSE*, transgenic mouse embryos were generated carrying the *–8.4(del3)ngn1:gfp* fragment. Expression of GFP was lost in dorsal and medial regions of the telencephalon in *–8.4(del3)ngn1:gfp* transgenics (Fig. 3D), in comparison with the *–8.4ngn1:gfp* line (Fig. 3C). Surprisingly, the transgene carrying the deletion retained high levels of gene expression in the region of the pallial-subpallial border, including the lateral telencephalon, prospective basolateral amygdalar complex and claustrum-endopiriform nucleus. Hence, the *LSE* appears to recapitulate only the dorsal-most expression of endogenous *Ngn1* in the mouse telencephalon, and additional regulatory elements are responsible for the lateral domain of telencephalic expression.

Given that the complete pattern of endogenous *Ngn1* expression can be obtained with the *–8.4ngn1:gfp* zebrafish transgene in mouse, regulatory sequences controlling lateral expression must be included in this sequence that are apparently inactive in the zebrafish telencephalon. Sequence comparison between the zebrafish and mammalian *ngn1* loci revealed a region of homology located between –1775 to –1368 kb upstream of the zebrafish *ngn1* start codon (called *LATE*; 61% homologous to mouse *Ngn1*; Fig. 7A). Consistent with the hypothesis of *LATE* being necessary for reporter expression in the lateral pallium of the mouse, E12.5 embryos transgenic for *–3.1ngn1:nlacZ* exhibited transgene activity in the telencephalon in a pattern very similar to that of *–8.4(del3)ngn1:gfp* transgenics (Fig. 3D,E). Deletion of the *LATE* region [*–3.1(delLATE)ngn1:nlacZ*] rendered the transgene inactive in the telencephalon (Fig. 3F), indicating that *LATE* is indispensable for this activity in the mouse. Thus, expression of the *–8.4ngn1:gfp* zebrafish transgene in the telencephalon of mouse is a composite of the activities of two regulatory regions, *LSE* and *LATE* (Fig. 3G).

The *LATE* region is required for expression in the diencephalon and hindbrain of zebrafish embryos

Transgenic zebrafish carrying the *–3.1ngn1:gfp* construct express the reporter strongly in the diencephalon and in the hindbrain (Fig. 4A). To test whether the conserved *LATE* region controls these aspects of *ngn1* expression, we generated stable transgenic zebrafish lines in which *LATE* was removed by deletion of the sequence between –1775 and –1368 upstream of the ATG. When *LATE* was abolished [*–3.1(delLATE)ngn1:gfp*], expression was strongly reduced in the diencephalon and the hindbrain (Fig. 4B). Reporter gene expression in the posterior spinal cord was unaffected by the deletion of *LATE* (data not shown), indicating that the reduction of transgene expression is due to the lack of *LATE* and not to the integration site. Furthermore, when *LATE* was deleted from the *–8.4ngn1:gfp* construct, a similar reduction of reporter expression in the diencephalon was observed (Fig. 4D). However, expression in the telencephalon was not affected by the lack of *LATE* in the *–8.4ngn1* transgenes (compare Fig. 4C,D). This confirms the conclusion from the deletion series (Fig. 2C) that *LATE* does not mediate telencephalic expression in the zebrafish embryo, and demonstrates that the *LSE* can drive expression in the telencephalon in the absence of *LATE*.

The activity of *LATE* depends on Pax 6

The paired homeodomain transcription factor Pax6 is required for expression of the related *Ngn2* in the lateral pallium of the mouse embryo (Stoykova et al., 2000; Toresson et al., 2000). Moreover, the ventral boundary of *Pax6* expression (Stoykova et al., 2000) coincides with the ventral boundary of *LATE* activity (Fig. 5A,B). We tested whether *Pax6* is necessary for the activity of *LATE* in the murine telencephalon by crossing the transgene into a *Pax6* mutant background (*small eye*)

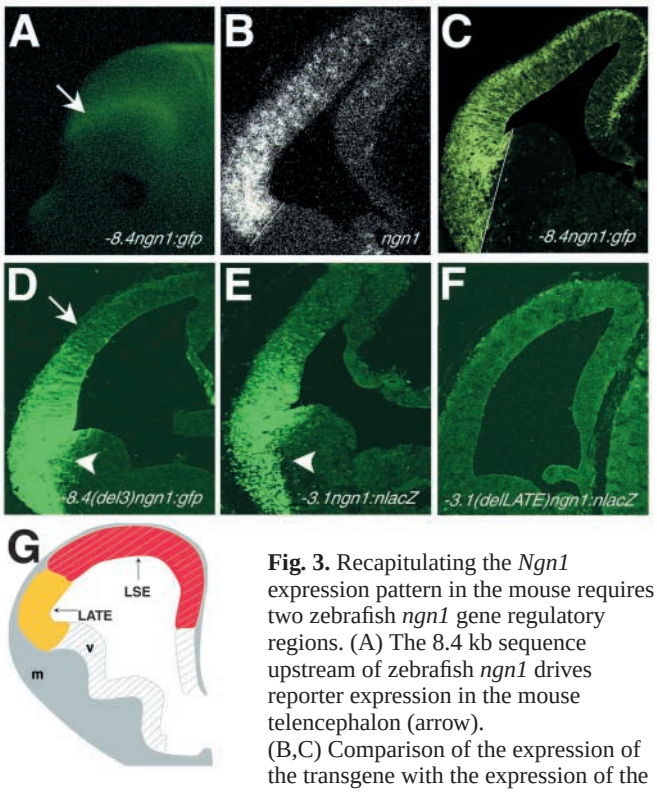

Fig. 3. Recapitulating the *Ngn1* expression pattern in the mouse requires two zebrafish *ngn1* gene regulatory regions. (A) The 8.4 kb sequence upstream of zebrafish *ngn1* drives reporter expression in the mouse telencephalon (arrow). (B,C) Comparison of the expression of the transgene with the expression of the endogenous *Ngn1* gene in coronal sections (dorsal up) shows that this fragment recapitulates fully the endogenous telencephalic expression of *Ngn1*. The sharp ventral boundary of Ngn1 and transgene expression is indicated by a line. (D) Whereas the complete *–8.4ngn1:gfp* transgene recapitulates the endogenous pattern of *Ngn1* (C), reporter activity is lost dorsally (arrow), but remains laterally (arrowhead), in transgenic lines with the *LSE* deleted [*–8.4(del3)ngn1:gfp*]. (E) Expression in the lateral telencephalon (arrowhead) is driven by an element within 3.1 kb of the *ngn1* regulatory region (*–3.1ngn1:nlacZ*). (F) Deletion of *LATE* [*–3.1(del LATE)ngn1:nlacZ*] abolishes the lateral activity of the *–3.1ngn1:nlacZ* transgene. Thus, two regulatory regions of the zebrafish transgene control the spatial expression of *ngn1* in the mouse pallium. (G) Summary of the activities of *LATE* and *LSE* in the dorsal telencephalon of mouse. m, mantle zone ; v, ventricular zone. Mouse embryos are at 12.5 dpc. Two transgenic lines were analysed per construct. With the expception of A, which is a lateral view of a whole-mounted mouse embryo, panels show coronal sections through the telencephalon, with dorsal up.

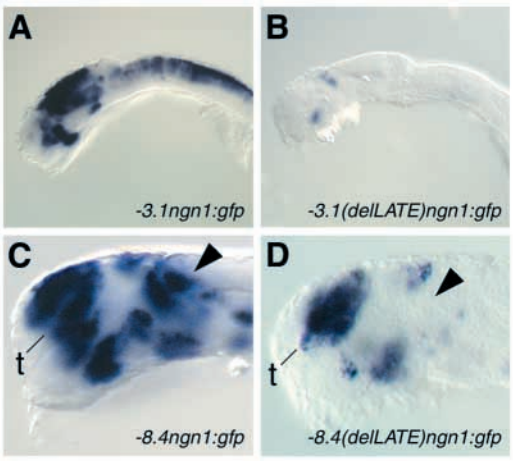

Fig. 4. The *LATE* region is required for expression in the diencephalon and hindbrain of zebrafish embryos. (A) The *–3.1ngn1:gfp* transgene, harbouring *LATE*, drives *gfp* expression extensively in the diencephalon and in the hindbrain. This pattern is partially reminiscent of the expression of the two zebrafish *pax6* genes (see Fig. 5D,E). (B) Upon deletion of *LATE* in *–3.1(delLATE)ngn1:gfp* transgenic embryos, expression of *gfp* mRNA in the diencephalon and in the hindbrain is strongly reduced. (C,D) Deletion of *LATE* in the *–8.4ngn1:gfp* transgenics reduced reporter expression in the diencephalon (arrowhead) but not in the telencephalon (t). Thus, the activity of *LATE* is required for transgene expression in the diencephalon. Embryos were oriented with anterior directed towards the left and dorsal up.

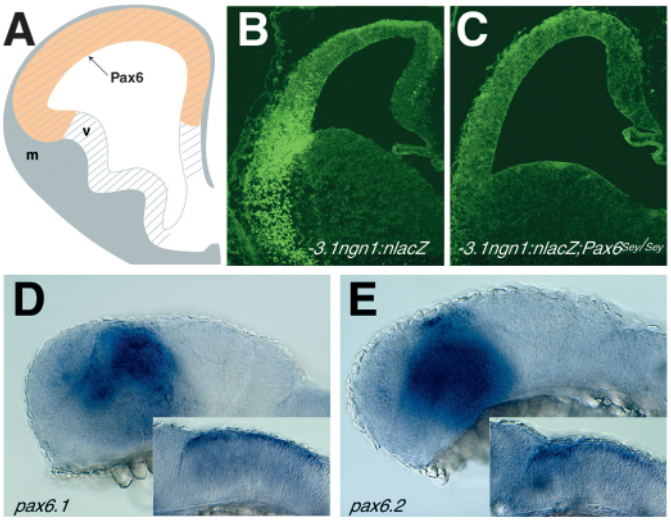

Fig. 5. *LATE* requires the activity of Pax6 in the mouse. (A) Schematic of the expression domain of *Pax6* in the telencephalon of the mouse [reproduced, with permission, from Stoykova et al. (Stoykova et al., 2000)]. (B) Control embryo carrying the *–3.1ngn1:nlacZ* transgene. (C) Pax6$^{sey/sey}$ embryo carrying the *–3.1ngn1:nlacZ* transgene, showing the loss of reporter expression, indicating that *LATE* activity depends on *Pax6*. Coronal sections through the telencephalon with dorsal up. (D,E) Lateral view of the two zebrafish pax *genes*, *pax6.1* (D) and *pax6.2* (E) shows overlapping expression in the fore- and hindbrain (insets). Staining is detected mainly in the dorsal diencephalon, reminiscent of the *–3.1ngn1:gfp* transgene (see Fig. 4A). m, mantle zone ; v, ventricular zone.

(Stoykova et al., 2000). Lack of *Pax6* activity abolishes expression of the *–3.1ngn1:nlacZ* transgene in the lateral area of the telencephalon (Fig. 5C) in the same manner as deletion of *LATE* does (Fig. 3F). Thus, Pax6 activity is necessary for expression of the zebrafish transgene in the mouse.

The two zebrafish *pax6* genes, *pax6.1* and *pax6.2* (Fig. 5D,E) are expressed in an overlapping pattern in the diencephalon and the hindbrain (MacDonald et al., 1994; Nornes et al., 1998). Strikingly, these two domains are highly similar to the territory of *LATE* activity (Fig. 4A). Common to both anmiotes and anamiotes, *Pax6* mRNA is mainly detected in the eyes and in the alar plate of the forebrain (comprising the pretectum, and the dorsal and ventral thalamus), and in the spinal cord. However, in contrast to in mouse embryos, *pax6.1* and *pax6.2* are not significantly expressed in the telencephalon of zebrafish embryos (MacDonald et al., 1994; Nornes et al., 1998; Wullimann and Rink, 2001). As shown in Fig. 5D,E, only a small domain of expression in the telencephalon, at the so-called pallial-subpallial boundary, is detected using both pax6.1 and pax6.2 riboprobes.

To test whether *LATE* activity in the zebrafish embryo is also dependent on Pax6, as observed in the mouse, we knocked down Pax6 activity by injecting a cocktail of two antisense morpholinos complementary to *pax6.1* and *pax6.2* mRNA. Phenotypically, 60% (*n*=152) of the injected embryos showed a reduction in the size of their eyes (Fig. 6A,B), when compared with wild type. Expression of the endogenous *ngn1* gene (Fig. 6D) (22%, *n*=41), as well as the *–8.4ngn1:gfp* transgene (28%, *n*=71, Fig. 6F), was significantly reduced in the diencephalon and hindbrain of morpholino-injected embryos. Expression in the telencephalon was not affected, indicating that the *LSE* is not dependent on Pax6 activity. The injection of either the *pax6.1* or the *pax6.2* morpholino alone did not cause an effect, suggesting that the two *pax6* genes act redundantly. Moreover, the lack of an effect of each morpholino by itself, even at high concentrations (1.2 mM), is strongly suggestive of a specific interaction of the morpholinos with the *pax6.1* and *pax6.2* mRNA.

A conserved binding site in *LATE* interacts with Pax6 protein in vitro

Our results indicate that *LATE* activity is dependent on Pax6 in both the zebrafish and the mouse. Examination of the sequence of the *LATE* region (Fig. 7A) revealed three putative Pax6-binding sites (Epstein et al., 1994), site-A, -B and -C. Interestingly, Site-C is the most conserved between mouse, human, chicken, *Xenopus tropicalis* and zebrafish *ngn1*.

To test whether the three Pax6-binding site homologies in the *LATE* region can bind Pax6 protein, we performed electromobility-shift assays with recombinant mouse Pax6 protein that was synthesised by in vitro translation (Scardigli et al., 2003). Oligonucleotides comprising either site-A, site-B or site-C (Fig. 7B) were ^{32}P-labelled and incubated with recombinant Pax6 protein in the presence of unlabelled oligonucleotide containing the homologous Pax6-binding site, or an oligonucleotide in which the putative Pax6-binding site was mutated by a cluster of point mutations. As a positive control, we used a previously described consensus Pax6-binding site (Czerny et al., 1993). The site-C oligonucleotide (sC) gave a strongly shifted band that was not reduced by

unspecific competitor with a mutated binding site (msC), but was totally abolished by the presence of competitor oligonucleotide with an intact Pax6-binding site homology (Fig. 8). Site-A and site-B did not yield retarded protein-DNA complexes (data not shown), suggesting either that they do not interact with Pax6 or that they bind to the protein very inefficiently.

In summary, these results demonstrate that Pax6 can interact directly with the *LATE* region. The observed effects in the Pax6-deficient embryos are thus likely to be due to the failure of Pax6 to activate *ngn1* expression through interaction with the *LATE* region.

In zebrafish embryos, the mouse *LATE* and *LSE* regions have the same regulatory activities as their cognate zebrafish enhancers

The zebrafish *LATE* region drives expression in the lateral telencephalon of the mouse but not of the zebrafish. The high conservation of *LATE* suggests that it may be the target of similar regulatory principles that are used in different places in the forebrain of the zebrafish and the mouse. Its dependence on Pax6 activity in both mouse and zebrafish is in support of this notion. Hence, one prediction is that the mouse *LATE* region, like the zebrafish *LATE* region should be active in the zebrafish diencephalons, but should not drive expression in the telencephalon.

To test this hypothesis, zebrafish embryos were injected with constructs, in which the zebrafish *LATE* was replaced with the mouse element (Fig. 9). The embryos were analysed at 26 hours after injection. To overcome the moscaism of such transient expression patterns, we used the *Sce*I meganuclease protocol (Thermes et al., 2002), and collected, in addition, accumulative expression maps by overlaying the expression pattern from many independently injected embryos (Fig. 9A-F). The replacement of zebrafish *LATE* with the conserved mouse *LATE* sequences produced embryos in which the expression of GFP was restricted to the diencephalon (Fig. 9C,C′), in a pattern similar to that of the zebrafish *LATE* region (Fig. 9A,A′). This is reminiscent of the spatial activity of the *–3.1ngn:gfp* stable transgenic lines (Fig. 4A). In addition, injected embryos do not show expression in the telencephalon (Fig. 9C,C′). As seen in the stable transgenic lines, deletion of *LATE* abolished diencephalic expression almost completely (Fig. 9B,B′). Taken together, this suggests that mouse *LATE* is functionally similar to zebrafish *LATE* when introduced into the zebrafish embryo. Moreover, these findings suggest that the regulatory principles controlling *LATE* in the zebrafish diencephalon were co-opted in the evolution of the lateral telencephalon of the mouse. There is, however, some variation in the diencephalic pattern driven by the mouse and zebrafish *LATE* enhancers, indicating that changes have occurred during

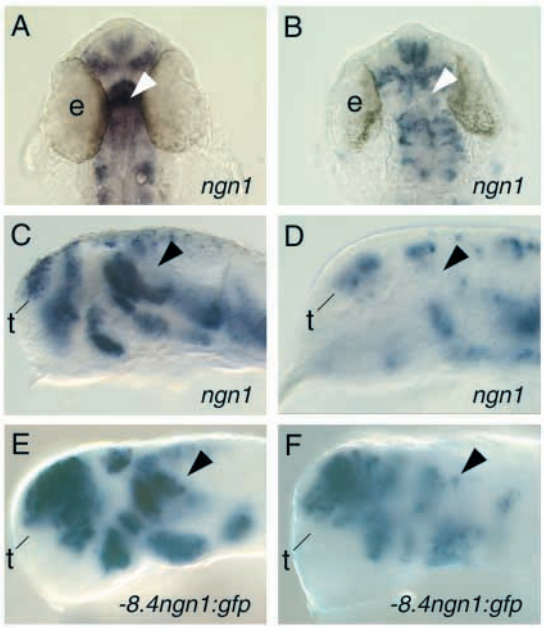

Fig. 6. Pax6 is required for *ngn1* and transgene expression in the diencephalon of the zebrafish embryo. (A-F) Control (A,C,E) and morpholino-injected (B,D,F) embryos hybridised to *ngn1* (A-D) and *gfp* (E,F) antisense probes. Expression of the endogenous *ngn1* mRNA and the transgene (*–8.4ngn1:gfp*) is reduced in the diencephalon (arrowheads) but not in the telencephalon (t). Embryos were injected with a cocktail of morpholino oligonucleotides directed against *pax6.1* and *pax6.2*. Embryos are 28 hpf. (A,B) Dorsal views of embryos with anterior oriented upwards; (C-F) lateral views. e, eye.

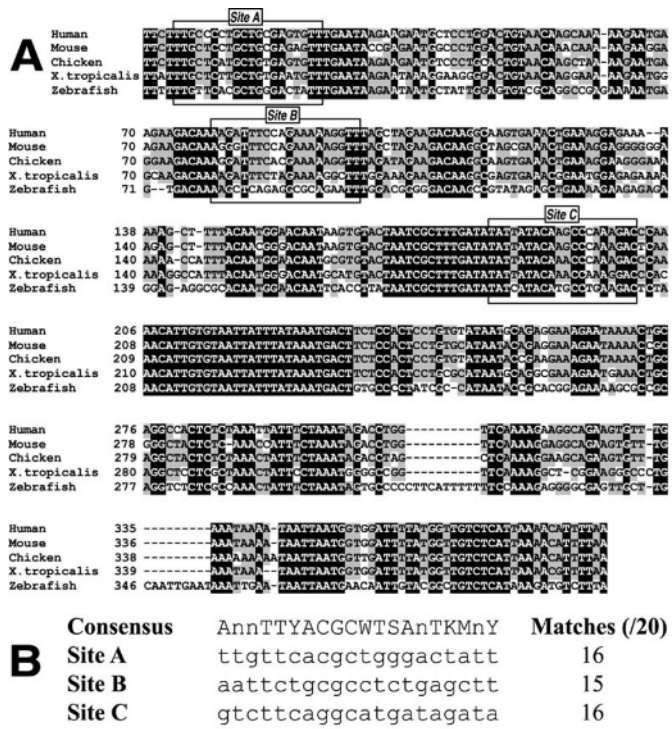

Fig. 7. The *LATE* region is conserved between tetrapods and zebrafish. (A) Sequence comparison of the *LATE* regulatory region of zebrafish *ngn1* with that of human, mouse, chicken and *Xenopus tropicalis*. The aligned sequences comprise nucleotides –1775 and –1368 in the zebrafish *ngn1* genomic regulatory region. In each column, the nucleotides are printed on a black background when they are all identical, or on grey background when one base is present in more than half of the sequences. Boxed regions indicate regions where the zebrafish sequence is similar to the Pax6 consensus-binding site (Epstein et al., 1994). Three sites (A, B and C) were scored. (B) Comparison of the three putative Pax6-binding sites with the consensus-binding sequence.

evolution at the level of fine-tuning of the expression pattern within the diencephalon.

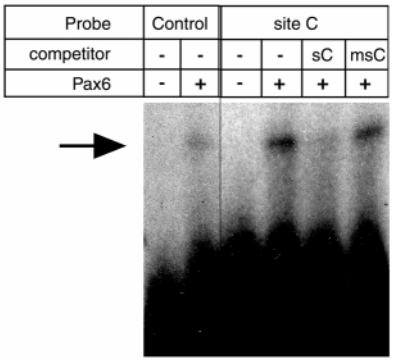

Fig. 8. Pax6 binds to site-C of the *LATE* region. Electromobility-shift assays with recombinant mouse Pax6 protein. Radioactively labelled oligonucleotide containing site-C of the *LATE* region was incubated with Pax6 protein without competitor oligonucleotide, or with a 50-fold molar excess of cold site-C oligonucleotide (sC), or with an oligonucleotide that contained a cluster of point mutations (msC) in the Pax6-binding site homology domain. The site-C competitor abolished the shifted band, whereas the complex formation was not affected by the presence of the mutated oligonucleotide, demonstrating that the interaction of Pax6 protein is dependent on an intact Pax6-binding site. As a positive control, an oligonucletide (Control) harbouring the Pax6-binding site described by Czerny et al. (Czerny et al., 1993) was used.

Next, we tested whether the mouse *LSE* would also be active in the zebrafish embryo (Fig. 9D-F′). In contrast to *LATE*, the zebrafish *LSE* drives expression in homologous structures, the dorsal telencephalon, in mouse and zebrafish. Replacement of the zebrafish *LSE* (*–8.4ngn1:gfp*) with the mouse *LSE* [*–8.4ngn1(msLSE):gfp*] produced embryos with transgene expression in the telencephalon (compare Fig. 9D,D′ with 9F,F′). Transient expression of constructs without the *LSE* only rarely drove expression in the dorsal telencephalon (Fig. 9E,E′). As expected from the presence of *LATE* in the parental transgene (*–8.4ngn1:gfp*), these derived constructs also showed prominent expression in the diencephalon. Thus, the structure and the function of the *LSE* from mouse and zebrafish *ngn1* are evolutionarily conserved.

Discussion

We previously mapped two enhancer regions that control *ngn1* expression in the zebrafish neural plate (Blader et al., 2003). Here, we have characterised the regulatory elements that drive expression in the embryonic brain at later post-somitogenesis stages. One of the previously mapped regulatory regions, the *LSE*, is required for expression in the zebrafish telencephalon. An additional, more proximally located region, *LATE*, mediates expression in the diencephalon and hindbrain. Both regions are conserved in mouse and human *Ngn1* genes, and both regions are active in the mouse and zebrafish brain, indicating that not only structural but also functional aspects of the two regulatory

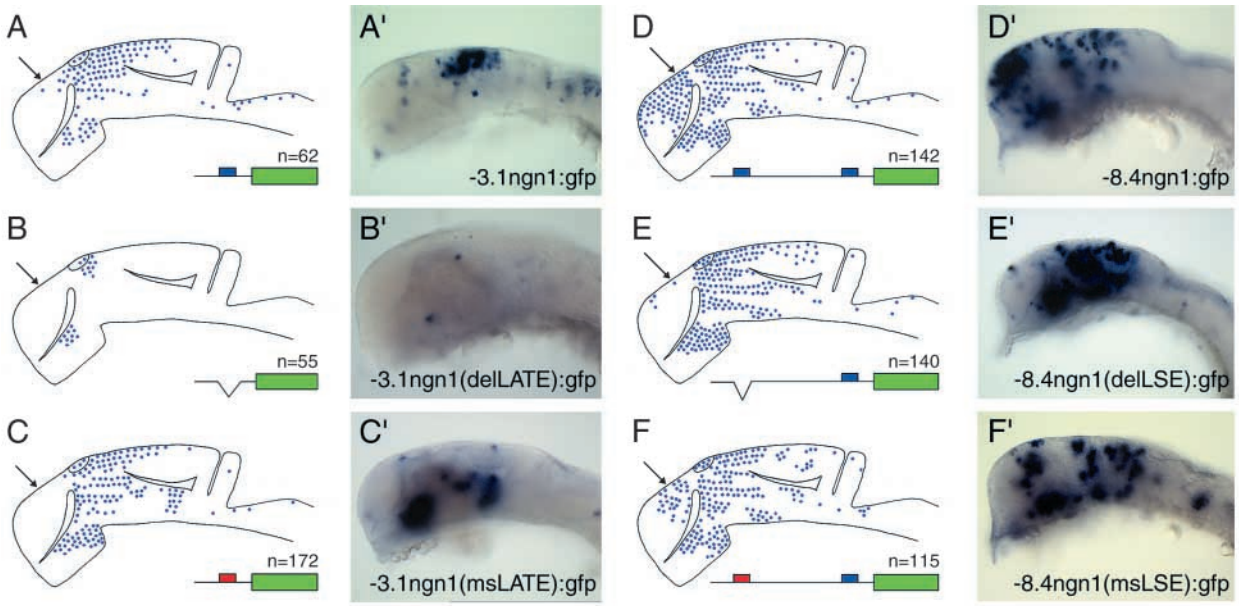

Fig. 9. The cis-regulatory elements of mouse *LATE* and *LSE* direct GFP expression in the zebrafish diencephalon and the telencephalon, respectively. The left panels show a cumulative map of the expression (blue dots; *n*, number of embryos analysed) of the reporter and the right panels are representative images of embryos injected with the construct indicated in the right bottom corner. Reporter gene expression was revealed by hybridising embryos to the antisense *gfp* probe. Blue and red boxes in the schematic drawings indicate zebrafish and mouse enhancers, respectively. (A-C) Embryos injected with –3.1*ngn1:gfp*, with mutant derivatives without the zebrafish LATE [*–3.1ngn1(delLATE):gfp*], or with a replacement with the mouse *LATE* [*–3.1ngn1(msLATE):gfp*]. Mouse *LATE* drives expression in the zebrafish diencephalon in a pattern very similar to zebrafish *LATE*. Like its zebrafish homologue, mouse *LATE* does not mediate expression in the zebrafish telencephalon (indicated by arrows). (D-F) Embryos were injected with the *–8.4ngn1:gfp* transgenes, with mutants without the *LSE* [*–8.4ngn1(delLSE):gfp*], or with the mouse *LSE* in place of the homologous zebrafish sequence [*–8.4ngn1(msLSE):gfp*]. As shown in stable expression experiments, the absence of the *LSE* significantly reduced the expression of the reporter in the telencephalon. The replacement of the zebrafish *LSE* with the homologous mouse *LSE* restored reporter expression in the telencephalon.

regions have been maintained during vertebrate evolution. Moreover, we provide evidence for a role of Pax6 proteins as regulators of *ngn1* expression in the zebrafish that probably involves a direct interaction of Pax6 with *LATE*.

Pax6 as a pre-pattern gene in the zebrafish brain

Several lines of evidence suggest that Pax6 is a regulator of *LATE*. First, the pattern of expression of *pax6* precedes and subsequently overlaps with that of *LATE* activity in the dorsal diencephalon and in the hindbrain. Moreover, *LATE* contains a *pax6*-binding site that is conserved in mammalian homologues of *ngn1* and that binds recombinant Pax6 in vitro. The simultaneous knockdown of *pax6.1* and *pax6.2* in the zebrafish embryo leads to the reduction of *ngn1* and transgene expression in the hindbrain and diencephalon, in a manner very similar to that observed following the deletion of *LATE* from transgenes. Finally, transgene expression in the mouse depends on a functional *pax6* gene.

Like the pre-pattern genes *her5*, *iroquois1* and *iroquois7* (Geling et al., 2003; Itoh et al., 2002), the zebrafish *pax6* genes (Krauss et al., 1991; Nornes et al., 1998) are expressed in broad domains in the neural plate from early stages onwards. However, not all cells within these broad domains express the transgene or the endogenous *ngn1* gene. Rather *ngn1* expression is restricted to distinct clusters of neurones that are located in at least six regions of the diencephalon, including non-*pax6*-expressing territories. In particular, the transcripts are detected in the epiphysis, pretectum, dorsal thalamus, ventral thalamus, preoptic area and the posterior tuberculum, in contrast to the expression of *pax6.1* and *pax6.2*, which is predominantly located in the alar plate of the forebrain. This suggests that *pax6* genes define a broad domain of competence in the diencephalon, in which other factors cooperate to specify the precise sub-region of neurogenesis. Hence, *pax6* could be regarded to act as a pre-pattern gene in the zebrafish neural plate/tube, in a similar fashion to *iroquois* or *her5* genes.

Humans and mice carrying a loss-of-function allele of *Pax6* show a severe haplo-insufficiency, causing aniridia and small eye phenotypes, respectively (Engelkamp and van Heyningen, 1996). This seems to be in contrast to the situation in zebrafish embryos, where reduction of *pax6* activity by knockdown of an individual *pax6* gene did not cause a visible effect in the embryo. This suggests that the two *pax6* genes in the zebrafish can act redundantly, at least during early embryonic stages. However, knockdown of gene function by a morpholino approach can impair gene function only transiently: defects obvious at later stages could thus not be scored.

Expression of the mouse *Ngn2* gene is dependent on *Pax6* activity in the telencephalon and in the spinal cord (Stoykova et al., 2000; Scardigli et al., 2003). Mouse *Ngn1* and *Ngn2* are expressed in similar but not identical patterns in the mouse nervous system, suggesting that the two *Ngn* genes are regulated by related mechanisms in the mouse that are derived from a common ancestral gene. However, despite the strong conservation of other regulatory sequences, such as the ANPE (Blader et al., 2003), the zebrafish *LATE* enhancer sequence is not conserved in *Ngn2*.

The recruitment of *LATE* in the mouse appears to be linked to evolution of the telencephalon

The telencephalon is one region of the brain where molecular changes in regulatory activity are likely to have occurred most extensively, given the vast expansion and morphological variations of these forebrain structures among different vertebrate groups (Nieuwenhuys, 1994). These differences are particularly striking in regions flanking the pallial (dorsal telencephalon)-subpallial (ventral telencephalon) boundary, which plays a pivotal role in the establishment of neuronal diversity, and in the reception of diverse developmental signals from dorsal and ventral domains of the telencephalon (Molnar and Butler, 2002). Despite the anatomical differences in telencephalon structure, comparative gene expression studies suggest that development of the forebrain follows a similar 'Bauplan' in all vertebrates, raising the question of how morphological differences evolved (Fernandez et al., 1998; Reiner, 2000; Striedter, 1997; Zerucha et al., 2000). *ngn1* is expressed in the dorsal telencephalon in both zebrafish and mouse (Blader et al., 2003; Fode et al., 2000), and thus represents an interesting case to study the modification of cis-regulatory elements during evolution of the telencephalon.

We demonstrate that expression of *ngn1* transgenes in the dorsal telencephalon of zebrafish embryos is dependent on the activity of one regulatory region (*LSE*), whereas expression in the developing isocortex of mouse requires the activity of two distinct regulatory regions (*LSE* and *LATE*). The zebrafish regulatory sequences faithfully recapitulate the endogenous pattern of *Ngn1* expression in the mouse, including the position of the sharp ventral boundary and the high-lateral to low-medial gradient in the pallium. *LATE* is highly conserved in the murine *Ngn1* and *Ngn2* genes (Blader et al., 2003; Scardigli et al., 2003). The homologous region of *Ngn2* is a direct target of *Pax6* and also drives expression in the lateral telencephalon (Blader et al., 2003; Scardigli et al., 2003). As *LATE* is not required for lateral telencephalon expression in the zebrafish, what maintained *LATE* over 450 million years of evolution? Our data show that zebrafish *LATE* is employed to drive expression in the diencephalon of the zebrafish, and that this function has also been retained by the mouse *LATE* when placed into the context of the zebrafish embryo.

Evolutionary modification of expression patterns

Eukaryotic regulatory regions are usually composed of multiple protein-binding sites clustered within a few hundred base pairs or less. The internal organisation of these regions can be rather flexible, as individual protein-binding modules can vary in position and orientation, and the actual DNA sequences bound by regulatory proteins are usually rather short and degenerate. In addition, multiple regulatory regions that can be scattered over megabases in vertebrate genomes contribute frequently to the expression of a gene (Davidson et al., 2000). Given this flexibility in the cis-regulatory organisation of vertebrate genes, the strong conservation of the position and sequences of the individual regulatory regions (*LSE*, *ANPE*, *LATE*) of the *ngn1* gene is remarkable. Regulatory regions of different genes of the same species can change during evolution at varying speeds (Davidson, 2001). The conservation of regulatory regions over 450 million years of independent evolution is not only restricted to *ngn* genes but also includes other regulators of neurogenesis, such as *delta-d* (Dornseifer et al., 1997), *sonic hedgehog* (Müller et al., 2002; Müller et al., 1999) and *pax6* (Kammandel et al., 1999), which

is suggestive of a strong selection pressure to maintain the structure of the regulatory regions of some genes.

It is believed that the evolutionary diversification of the body plan was driven to a large extent by changes in gene expression, rather than by the emergence of novel regulatory proteins (Dermitzakis and Clark, 2002; Ludwig, 2002; Stone and Wray, 2001; Tautz, 2000). In principal, three ways of how the cis-regulatory region evolved can be envisaged. Cis-regulatory regions could have emerged de novo by the accidental clustering of protein-binding modules in close proximity in non-coding sequences. Alternatively, existing regulatory regions could have been modified by the deletion or addition of protein-binding sites, giving novel patterns of regulatory activity. As a third possibility, enhancer sequences could have been co-opted. In this case, the expression of the interacting transcription factors was altered, placing a pre-existing enhancer into a different spatial and/or temporal regulatory context. In addition, rearrangements of the genome may have played crucial roles in redistributing these novel regulatory activities among genes.

Our findings for the *ngn1* enhancers suggest a scenario in which a regulatory sequence has been co-opted in order to drive expression in a novel context. Pax6 is not widely expressed in the telencephalon of early post-somitogenesis-stage zebrafish embryos. Only a few migratory post-mitotic telencephalic cells express Pax6 in the zebrafish embryo (Wullimann and Rink, 2002). This is in striking contrast to the widespread expression of Pax6 in the proliferative radial glia in the telencephalon of the mouse embryo (Gotz et al., 1998; Stoykova et al., 2000). This, together with the dependence of *LATE* activity on Pax6, indicates that evolutionary modulation of Pax6 expression could have been involved in the recruitment of *LATE* for *Ngn1* expression in the lateral telencephalon of the mouse. Co-option of a regulatory sequence to drive expression at the pallial/subpallial border and in immediately adjacent structures in the mouse is particularly intriguing, as this region is believed to be absent from the zebrafish telencephalon, and to have appeared first in the amniote lineage as a major prerequisite for the emergence of the cortex in mammals (Molnar and Butler, 2002). However, it is unlikely that Pax6 is the only factor involved in this recruitment. Other regions of the embryo that express Pax6 do not show expression of endogenous *ngn1* or the *LATE*-containing transgenes. Moreover, even in the small group of cells that express *pax6* in the zebrafish telencephalon, Pax6 is not sufficient to activate *ngn1* expression, indicating that cooperating factors are also necessary. This is reflected in the extended regions of conserved sequence flanking the Pax6-binding site in *LATE* that presumably represent the phylogenetic footprints of other conserved transcription factors (see Fig. 7). Nevertheless, our data, together with the appearance of extensive *Pax6* expression in the radial glial cells of the mammalian neocortex, suggest an important role of Pax6 as one of the factors that have recruited *Ngn1* expression in the telencephalon of mammals.

Conclusion

Our results provide evidence that the co-option of pre-existing enhancers is a mechanism to diversify regulatory patterns during evolution. These results have further implications: the attempts to delineate fields of evolutionary homology on the basis of shared gene expression can be misleading, as expression territories may be composites of the activities of distinct regulatory regions that have evolved independently, as demonstrated here for the *LSE* and the *LATE* regions of *ngn1*. Moreover, in comparative genomic approaches to identify regulatory regions by sequence conservation, the regulatory function cannot be inferred from conserved sequences, as regulatory regions may have been co-opted in distinct processes during evolution.

We thank A. Simmons, J. Johnson, M. Mione, M. Wullimann. M. Wassef, T. Dickmeis, D. Hutcheson and S. Wilson for discussions; J. Johnson for mouse *Ngn1* sequences, N. Bäumer and P. Gruss for the Pax6 plasmid. We are grateful to O. Nkwunda, A. Karmin, D. Biellmann, L. Arnoux, C. Vialle, and the IGBMC mouse facility for technical assistance and animal care. C.S.L. is a scholar of ASTAR, Singapore. This work was supported by INSERM, CNRS, HUS, FZK, ARC, AFM, VW Stiftung, AICR, European commission IP ZF-Models and ACI.

References

Ashery-Padan, R. and Gruss, P. (2001). Pax6 lights-up the way for eye development. *Curr. Opin. Cell Biol.* **13**, 706-714.

Bally-Cuif, L. and Hammerschmidt, M. (2003). Induction and patterning of neuronal development, and its connection to cell cycle control. *Curr. Opin. Neurobiol.* **13**, 16-25.

Bellefroid, E. J., Kobbe, A., Gruss, P., Pieler, T., Gurdon, J. B. and Papalopulu, N. (1998). Xiro3 encodes a Xenopus homolog of the Drosophila Iroquois genes and functions in neural specification. *EMBO J.* **17**, 191-203.

Bertrand, N., Castro, D. S. and Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.* **3**, 517-530.

Blader, P., Fischer, N., Gradwohl, G., Guillemot, F. and Strähle, U. (1997). The activity of neurogenin1 is controlled by local cues in the zebrafish embryo. *Development* **124**, 4557-4569.

Blader, P., Plessy, C. and Strähle, U. (2003). Multiple regulatory elements with spatially and temporally distinct activities control neurogenin1 expression in primary neurons of the zebrafish embryo. *Mech. Dev.* **120**, 211-218.

Brewster, R., Lee, J. and Ruiz i Altaba, A. (1998). Gli/Zic factors pattern the neural plate by defining domains of cell differentiation. *Nature* **393**, 579-583.

Cheng, C. W., Hui, C., Strähle, U. and Cheng, S. H. (2001). Identification and expression of zebrafish Iroquois homeobox gene irx1. *Dev. Genes Evol.* **211**, 442-444.

Chitnis, A. B. (1999). Control of neurogenesis–lessons from frogs, fish and flies. *Curr. Opin. Neurobiol.* **9**, 18-25.

Czerny, T., Schaffner, G. and Busslinger, M. (1993). DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev.* **7**, 2048-2061.

Davidson, E. (2001). *Genomic regulatory systems: Development and evolution*. San Diego, CA: Academic Press.

Davidson, H., Taylor, M. S., Doherty, A., Boyd, A. C. and Porteous, D. J. (2000). Genomic sequence analysis of Fugu rubripes CFTR and flanking genes in a 60 kb region conserving synteny with 800 kb of human chromosome 7. *Genome Res.* **10**, 1194-1203.

Dermitzakis, E. T. and Clark, A. G. (2002). Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. *Mol. Biol. Evol.* **19**, 1114-1121.

Dornseifer, P., Takke, C. and Campos-Ortega, J. A. (1997). Overexpression of a zebrafish homologue of the Drosophila neurogenic gene Delta perturbs differentiation of primary neurons and somite development. *Mech. Dev.* **63**, 159-171.

Engelkamp, D. and van Heyningen, V. (1996). Transcription factors in disease. *Curr. Opin. Genet. Dev.* **6**, 334-342.

Epstein, J., Cai, J., Glaser, T., Jepeal, L. and Maas, R. (1994). Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. *J. Biol. Chem.* **269**, 8355-8361.

Fernandez, A. S., Pieau, C., Reperant, J., Boncinelli, E. and Wassef, M. (1998). Expression of the Emx-1 and Dlx-1 homeobox genes define three

molecularly distinct domains in the telencephalon of mouse, chick, turtle and frog embryos: implications for the evolution of telencephalic subdivisions in amniotes. *Development* **125**, 2099-2111.

Fode, C., Ma, Q., Casarosa, S., Ang, S. L., Anderson, D. J. and Guillemot, F. (2000). A role for neural determination genes in specifying the dorsoventral identity of telencephalic neurons. *Genes Dev.* **14**, 67-80.

Geling, A., Itoh, M., Tallafuss, A., Chapouton, P., Tannhauser, B., Kuwada, J. Y., Chitnis, A. B. and Bally-Cuif, L. (2003). bHLH transcription factor Her5 links patterning to regional inhibition of neurogenesis at the midbrain-hindbrain boundary. *Development* **130**, 1591-1604.

Geling, A., Plessy, C., Rastegar, S., Strähle, U. and Bally-Cuif, L. (2004). Her5 acts as a prepattern factor that blocks neurogenin1 and coe2 expression upstrem of notch to inhibit neurogenesis at the midbrain-hindbrain boundry. *Development* **131**, 1993-2006.

Gomez-Skarmeta, J. L., Glavic, A., de la Calle-Mustienes, E., Modolell, J. and Mayor, R. (1998). Xiro, a Xenopus homolog of the Drosophila Iroquois complex genes, controls development at the neural plate. *EMBO J.* **17**, 181-190.

Gotz, M., Stoykova, A. and Gruss, P. (1998). Pax6 controls radial glia differentiation in the cerebral cortex. *Neuron* **21**, 1031-1044.

Hassan, B. A. and Bellen, H. J. (2000). Doing the MATH: is the mouse a good model for fly development? *Genes Dev.* **14**, 1852-1865.

Itoh, M., Kudoh, T., Dedekian, M., Kim, C. H. and Chitnis, A. B. (2002). A role for iro1 and iro7 in the establishment of an anteroposterior compartment of the ectoderm adjacent to the midbrain-hindbrain boundary. *Development* **129**, 2317-2327.

Kammandel, B., Chowdhury, K., Stoykova, A., Aparicio, S., Brenner, S. and Gruss, P. (1999). Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. *Dev. Biol.* **205**, 79-97.

Korzh, V., Sleptsova, I., Liao, J., He, J. and Gong, Z. (1998). Expression of zebrafish bHLH genes ngn1 and nrd defines distinct stages of neural differentiation. *Dev. Dyn.* **213**, 92-104.

Krauss, S., Johansen, T., Korzh, V., Moens, U., Ericson, J. U. and Fjose, A. (1991). Zebrafish pax[zf-a]: a paired box-containing gene expressed in the neural tube. *EMBO J.* **10**, 3609-3619.

Ludwig, M. Z. (2002). Functional evolution of noncoding DNA. *Curr. Opin. Genet. Dev.* **12**, 634-639.

Ma, Q., Kintner, C. and Anderson, D. J. (1996). Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* **87**, 43-52.

MacDonald, R., Qiling, X., Barth, A. R., Mikkola, I., Holder, N., Fjose, A., Krauss, S. and Wilson, S. W. (1994). Regulatory gene expression boundaries demarcate sites of neuronal differentiation in the embryonic zebrafish forebrain. *Neuron* **13**, 1039-1053.

Molnar, Z. and Butler, A. B. (2002). The corticostriatal junction: a crucial region for forebrain development and evolution. *Bioessays* **24**, 530-541.

Mueller, T. and Wullimann, M. F. (2003). Anatomy of neurogenesis in the early zebrafish brain. *Brain Res. Dev. Brain Res.* **140**, 137-155.

Müller, F., Chang, B., Albert, S., Fischer, N., Tora, L. and Strähle, U. (1999). Intronic enhancers control expression of zebrafish sonic hedgehog in floor plate and notochord. *Development* **126**, 2103-2116.

Müller, F., Blader, P. and Strähle, U. (2002). Search for enhancers: teleost models in comparative genomic and transgenic analysis of cis regulatory elements. *Bioessays* **24**, 564-572.

Nakada, Y., Parab, P., Simmons, A., Omer-Abdalla, A. and Johnson, J. E. (2004). Separable enhancer sequences regulate the expression of the neural bHLH transcription factor neurogenin 1. *Dev. Biol.* **271**, 479-487.

Nieuwenhuys, R. (1994). The neocortex. An overview of its evolutionary development, structural organization and synaptology. *Anat. Embryol.* **190**, 307-337.

Nornes, S., Clarkson, M., Mikkola, I., Pederson, M., Bardsley, A., Martinez, J. P., Krauss, S. and Johansen, T. (1998). Zebrafish contains two Pax6 genes involved in eye development. *Mech. Dev.* **77**, 185-196.

Notredame, C., Higgins, D. G. and Heringa, J. (2000). T coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205-217.

Reiner, A. J. (2000). A hypothesis as to the organization of cerebral cortex in the common amniote ancestor of modern reptiles and mammals. *Novartis Found. Symp.* **228**, 83-102.

Scardigli, R., Schuurmans, C., Gradwohl, G. and Guillemot, F. (2001). Crossregulation between Neurogenin2 and pathways specifying neuronal identity in the spinal cord. *Neuron* **31**, 203-217.

Scardigli, R., Bäumer, N., Gruss, P., Guillemot, F. and le Roux, I. (2003). Direct and concentration-dpendent regulation of the proneural gene Neurogenin2 by Pax6. *Development* **130**, 3269-3281.

Simmons, A. D., Horton, S., Abney, A. L. and Johnson, J. E. (2001). Neurogenin2 expression in ventral and dorsal spinal neural tube progenitor cells is regulated by distinct enhancers. *Dev. Biol.* **229**, 327-339.

Simpson, P. (1997). Notch signalling in development: on equivalence groups and asymmetric developmental potential. *Curr. Opin. Genet. Dev.* **7**, 537-542.

Simpson, T. I. and Price, D. J. (2002). Pax6; a pleiotropic player in development. *Bioessays* **24**, 1041-1051.

Stone, J. R. and Wray, G. A. (2001). Rapid evolution of cis-regulatory sequences via local point mutations. *Mol. Biol. Evol.* **18**, 1764-1770.

Stoykova, A., Treichel, D., Hallonet, M. and Gruss, P. (2000). Pax6 modulates the dorsoventral patterning of the mammalian telencephalon. *J. Neurosci.* **20**, 8042-8050.

Striedter, G. F. (1997). The telencephalon of tetrapods in evolution. *Brain Behav. Evol.* **49**, 179-213.

Tautz, D. (2000). Evolution of transcriptional regulation. *Curr. Opin. Genet. Dev.* **10**, 575-579.

Thermes, V., Grabher, C., Ristoratore, F., Bourrat, F., Choulika, A., Wittbrodt, J. and Joly, J. S. (2002). I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech. Dev.* **118**, 91-98.

Toresson, H., Potter, S. S. and Campbell, K. (2000). Genetic control of dorsal-ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2. *Development* **127**, 4361-4371.

Wang, X., Emelyanov, A., Sleptsova-Friedrich, I., Korzh, V. and Gong, Z. (2001). Expression of two novel zebrafish iroquois homologues (ziro1 and ziro5) during early development of axial structures and central nervous system. *Mech. Dev.* **105**, 191-195.

Wullimann, M. F. and Knipp, S. (2000). Proliferation pattern changes in the zebrafish brain from embryonic through early postembryonic stages. *Anat. Embryol.* **202**, 385-400.

Wullimann, M. F. and Rink, E. (2001). Detailed immunohistology of Pax6 protein and tyrosine hydroxylase in the early zebrafish brain suggests role of Pax6 gene in development of dopaminergic diencephalic neurons. *Brain Res. Dev. Brain Res.* **131**, 173-191.

Wullimann, M. F. and Rink, E. (2002). The teleostean forebrain: a comparative and developmental view based on early proliferation, Pax6 activity and catecholaminergic organization. *Brain Res. Bull.* **57**, 363-370.

Zerucha, T., Stuhmer, T., Hatch, G., Park, B. K., Long, Q., Yu, G., Gambarotta, A., Schultz, J. R., Rubenstein, J. L. and Ekker, M. (2000). A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. *J. Neurosci.* **20**, 709-721.