Springer

Dear Author:

Please find attached the final pdf file of your contribution, which can be viewed using the Acrobat Reader, version 3.0 or higher. We would kindly like to draw your attention to the fact that copyright law is also valid for electronic products. This means especially that:

- You may not alter the pdf file, as changes to the published contribution are prohibited by copyright law.

- You may print the file and distribute it amongst your colleagues in the scientific community for scientific and/or personal use.

- You may make an article published by Springer-Verlag available on your personal home page provided the source of the published article is cited and Springer-Verlag is mentioned as copyright holder. You are requested to create a link to the published article in LINK, Springer's internet service. The link must be accompanied by the following text: The original publication is available on LINK **http://link.springer.de.** Please use the appropriate URL and/or DOI for the article in LINK. Articles disseminated via LINK are indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks and consortia.

- You are not allowed to make the pdf file accessible to the general public, e.g. your institute/your company is not allowed to place this file on its homepage.

- Please address any queries to the production editor of the journal in question, giving your name, the journal title, volume and first page number.

Yours sincerely,

Springer-Verlag Berlin Heidelberg

Dev Genes Evol (2002) 212:293–297
DOI 10.1007/s00427-002-0244-1

SHORT COMMUNICATION

Tokiharu Takahashi · Peter W. H. Holland
Martin J. Cohn · Kiyoshi Shimizu · Mineo Kurokawa
Hisamaru Hirai

An orphan PRD class homeobox gene expressed in mouse brain and limb development

Received: 18 December 2001 / Accepted: 21 March 2002 / Published online: 24 May 2002

Abstract We report the cDNA sequence and expression of a mouse homeobox gene, *Dmbx1*, from the PRD class and comparison to its human orthologue. The gene defines a new homeobox gene family, Dmbx, phylogenetically distinct from the Ptx, Alx, Prx Otx, Gsc, Otp and Pax gene families. The *Dmbx1* gene is expressed in the developing mouse diencephalon, midbrain and hindbrain, and has dynamic expression during forelimb and hindlimb development. Unusually for homeobox genes, there is no orthologue in the *Drosophila* or *Caenorhabditis* genomes; we argue this reflects secondary loss.

Keywords Homeobox · Brain development · Limb development · K50 homeodomain · Paired

Introduction

Homeobox genes comprise a diverse and ancient gene superfamily. Molecular phylogenetic analyses of the homeodomain have identified two main "classes" (also called superclasses) of animal homeobox genes – ANTP and PRD – and several more divergent lineages (e.g. TALE, LIM, POU; Galliot et al. 1999). The ANTP and PRD classes can be subdivided, in turn, into numerous distinct gene families, again using homeodomain sequence comparisons. Examples include the Emx, En, Evx, Mnx or Msx families (in the ANTP class) and the Otx, Otp, Ptx, Pax3/7, Pax 1/9, Pax 2/5/8 or Prx families (in the PRD class). Each family includes one or more vertebrate genes and one or more genes from *Drosophila* and/or *Caenorhabditis elegans.* In practice, a gene family usually encompasses all the genes descendent from a single precursor gene in the most recent common ancestor of the Bilateria. The principal exception is the Hox genes, which are often considered a single family, even though the bilaterian ancestors certainly possessed multiple Hox genes (albeit in one cluster). Secondary loss of homeobox gene families has been noticed; for example, loss of Gsx genes in *C. elegans* or loss of Xlox in *Drosophila* and *C. elegans* (Ferrier and Holland 2001). Here we describe cloning and expression of a new member of the PRD class from mouse; this gene is not a member of any previously recognised homeobox gene family. The gene has a clear orthologue in the draft human genome sequence, but no orthologue in *Drosophila* or *C. elegans*.

Edited by B. Herrmann

T. Takahashi · P.W.H. Holland (✉) · M.J. Cohn
School of Animal and Microbial Science,
The University of Reading, Whiteknights, Reading, RG6 6AJ, UK
e-mail: p.w.h.holland@rdg.ac.uk
Tel.: +44-118-9318466, Fax: +44-118-9310180

T. Takahashi · M. Kurokawa
Department of Hematology and Oncology,
Graduate School of Medicine, University of Tokyo,
7–3-1 Hongo, Bunkyo-ku, Tokyo 113–8655, Japan

K. Shimizu
Pharmaceutical Division, Pharmaceutical Research Laboratories,
Kirin Brewery Co. Ltd., Miyahara 3 Takasaki, Gunma 370–1295, Japan

H. Hirai
Department of Cell Therapy and Transplantation Medicine,
Graduate School of Medicine, University of Tokyo,
7–3-1 Hongo, Bunkyo-ku, Tokyo 113–8655, Japan

Materials and methods

Degenerate primers were designed to two motifs conserved in a subset of PRD-like homeodomains (RRSRTTF and QVWF(K/S)NRR). These were used in RT-PCR reactions using a cDNA template from c-kit+ cells immunopurifed from the aorta-gonad-mesonephros region of 11.5 days post coitum (E11.5) C57B/6 mouse embryos. The expected 170-bp amplified product was cloned and nine recombinants sequenced. Eight derived from known genes (*Phox1, Pax3, S8*); one novel clone was used to screen an unamplified E9.5 mouse embryo cDNA library that we constructed in lambda ZAP II (Stratagene). Two 6-kb and four 4-kb independent cDNA clones were obtained; these all derived from the same gene. BLAST searches showed this to be a novel homeobox gene. The longer sequence is reported here. During preparation of this manuscript, another laboratory deposited a 1-kb sequence onto GenBank that is internal to our cDNAs, and named this *Dmbx1* (accession AF421858) (Ohtoshi et al. 2002, Miyamoto et al. 2002). To avoid confusion in the literature,

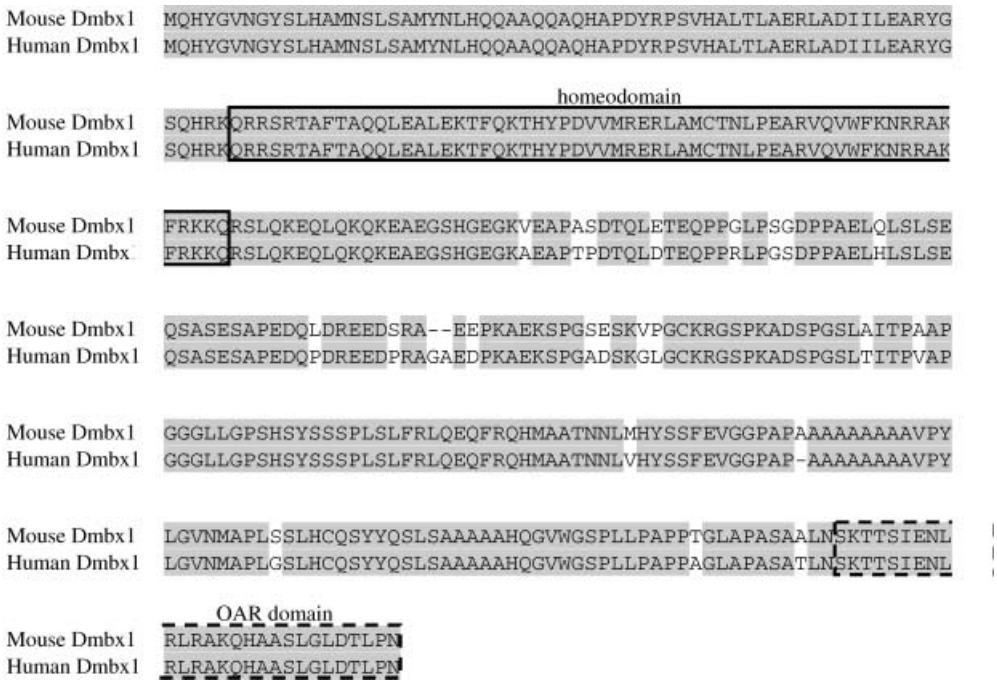

Fig. 1 Deduced protein sequence of mouse Dmbx1 aligned with its human orthologue predicted from genome sequence. *Dashes* indicate gaps inserted to maximise alignment. The *box* marks the homeodomain; the *dashed box* indicates the OAR domain

we follow this nomenclature in the work presented here. Our sequence is deposited with GenBank and given accession number AF499446. Phylogenetic analysis of amino acid sequence was performed using the Neighbour-Joining method implemented in ClustalX (Thompson et al. 1997), with outputs displayed using TreeView (Page 1996). We restricted analysis to the homeodomain to enable maximal representation of PRD class genes; analysis of a longer alignment including the homeodomain plus OAR domain gave similar results. The alignment and list of sequences used is available at http://www.rubic.rdg.ac.uk/amphioxus. Whole-mount in situ hybridisation to mouse embryos (strain CD-1) was performed as described by Nieto et al. (1996) with slight modifications using digoxygenin-labelled riboprobes from the complete 4-kb cDNA or a 990-bp subclone covering most of the open reading frame. The two probes gave identical results. After staining, hindbrains were dissected and flat-mounted under coverslips for photography.

Results and discussion

Using RT-PCR followed by cDNA library screening, we cloned a novel homeobox gene from c-kit+ cells isolated from the developing aorta-gonad-mesonephros (AGM) region of mouse embryos. These cells were chosen because we aimed to isolate genes for transcription factors involved in controlling haematopoiesis; the AGM region gives rise to definitive haematopoietic stem cells (Medvinsk and Dzierzak 1996). As described below, the gene we isolated is likely to have broader developmental regulatory roles. The longest cDNA has the potential to code for a 40.6-kDa protein of 376 amino acids, including a Paired-like homeodomain followed by an OAR domain. The gene, designated *Dmbx1*, has a clear human homologue in the draft human genome sequence, located on chromosome 1 (GenBank accession AL137797). Alignment of the mouse and human deduced proteins reveals 100% identity over the homeodomain and OAR domain, and 94% identity across the entire protein (Fig. 1). This is far higher than expected for paralogues within a gene family, and suggests these genes are direct orthologues. Using genome sequences accessible through FlyBase and WormBase, we did not detect an orthologue in either the *D. melanogaster* or the *C. elegans* genome.

The Dmbx1 homeodomain is equidistant from a range of genes in the PRD class, notably members of the Otx, Prx, Ptx and aristaless families. The presence of diagnostic P26, D27, E32, R44, O46 and A54 residues in the Dmbx1 homeodomain is diagnostic for the PRD class. The residue at homeodomain position 50 is important in determining DNA-binding specificity; some authors use its identity to subdivide the PRD class (Treisman et al. 1989), although these subdivisions are not monophyletic (Galliot et al. 1999). Dmbx1 has a lysine at this position, assigning it to the K50 Paired-like genes.

To investigate the evolutionary origin and relationships of *Dmbx1*, we conducted molecular phylogenetic analyses (Fig. 2). These confirmed that *Dmbx1* is not a member of any known homeobox gene family; indeed the gene is equidistant from several gene families. Considering the absence of a closely related gene in *Drosophila* or *C. elegans*, this could be interpreted as the gene having emerged during deuterostome, chordate, or vertebrate evolution. We suggest this is unlikely, because the phylogenetic node separating the *Dmbx1* lineage from other genes is deeper than the nodes separating *Drosophila*, nematode and vertebrate homologues *within* each known gene family. This implies that *Dmbx1* defines a new gene family within the PRD class, designated Dmbx, and that this gene family is as ancient as the Otx, Prx, Ptx, Gsc, al, Otp and Pax3/7 gene families, each of which have vertebrate and invertebrate members. We conclude that the origin of the Dmbx gene family

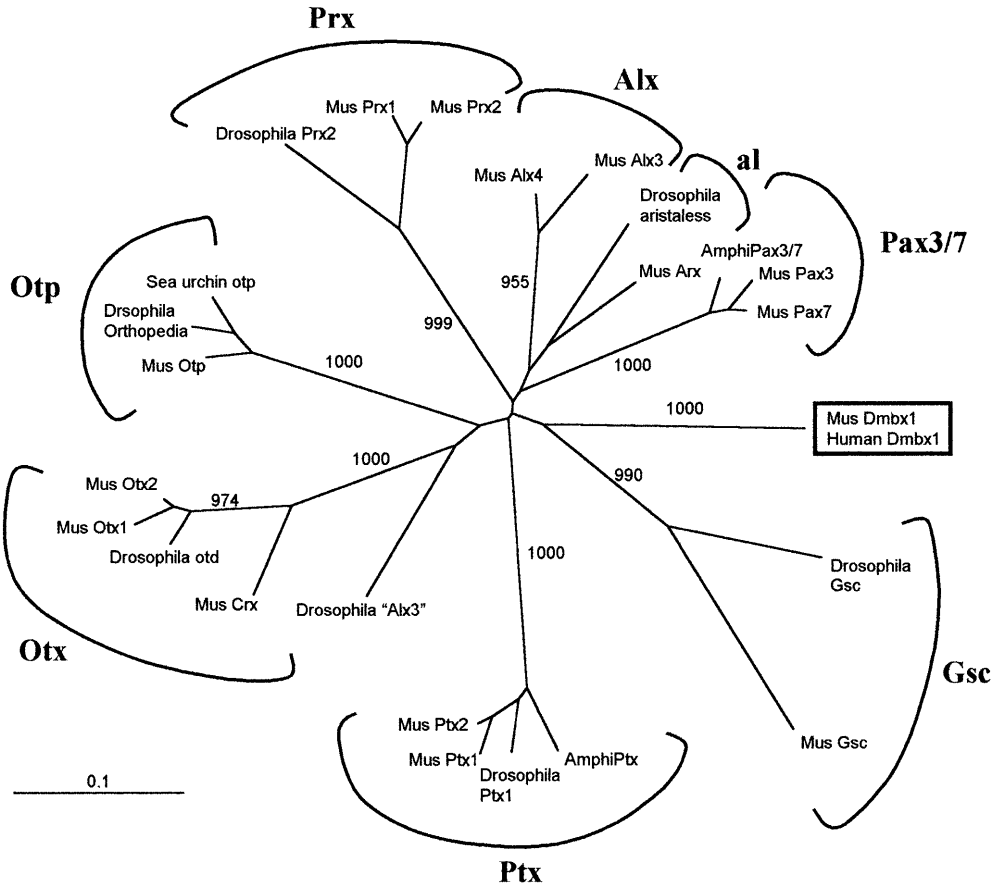

Fig. 2 Unrooted phylogenetic tree constructed from Paired class homeodomains. *Figures on nodes* indicate support values from 1,000 bootstrap resamplings of the data

pre-dates the divergence of arthropods, nematodes and vertebrates, and that Dmbx genes have been secondarily lost on the evolutionary lineages leading to *D. melanogaster* and *C. elegans*. We also note that the Dmbx gene family is (thus far) represented by only a single gene in the human and mouse genomes; most other homeobox families have two to four members (there are a few singletons, such as Xlox). If we accept the emerging view that the early vertebrate genome expanded by two rounds of whole genome duplication (Furlong and Holland 2002), then gene loss must also have occurred in the vertebrate Dmbx gene family.

Using whole-mount in situ hybridisation, we examined the spatiotemporal pattern of *Dmbx1* expression during mouse development. In mouse embryos at embryonic day (E) 7.5–8, expression of *Dmbx1* is detected around the prospective midbrain region (Fig. 3A). By E8–8.5, the expression becomes more definite and limited to the prospective midbrain region exclusively (Fig. 3B). In mouse embryos at E9.5, the domain of midbrain expression has expanded, extending partly into the optic eminence (Fig. 3C). In mouse embryos at E10, the clearest site of *Dmbx1* expression is still the midbrain, where mRNA is detected across all of the developing structure, and rostrally into the diencephalon (Fig. 3D). Clear expression also appears in the medial and lateral nasal pits, the dorsal half of the optic cup, and parts of the hindbrain. By E11, the hindbrain expression can be resolved into two anteroposteriorly oriented stripes along the lateral edges of rhombomeres, on each side of the midline (Fig. 3E). A small but distinct region of expression is also detected in the posterior, distal region of the forelimb buds (Fig. 3E, G). This expression is located near the distal edge of the *Sonic hedgehog*-expressing domain, but subjacent to the *Fgf4*-expressing domain in the posterior region of the apical ectodermal ridge. No expression is observed in hindlimb buds at this stage (Fig. 3H). Expression in the midbrain, diencephalon and hindbrain persists at E11.5, and the nasal expression in now detected on either side of the naso-lacrimal groove (Fig. 3F). Close examination reveals that the midbrain and hindbrain expression is not contiguous; *Dmbx1* transcripts are absent directly at the site of the midbrain-hindbrain boundary or MHB. At this stage, the expression in the optic cup has weakened. The expression in limbs also displays a dynamic pattern. Expression in forelimb buds observed half a day earlier has now disappeared (Fig. 3I); instead, *Dmbx1* transcripts are now detected along the distal edge of the hindlimb buds, subjacent to the apical ectodermal ridge (Fig. 3J). RT-PCR analysis detects *Dmbx1* expression in c-kit+ cells of the AGM region at this stage (data not shown). By flat-mounting the stained hindbrain at E11, the hindbrain expression can be seen to be in two parallel, bilateral stripes of cells running longitudinally (Fig. 3K, L). The more medial stripe runs from rhombomere 2 (r2) to the

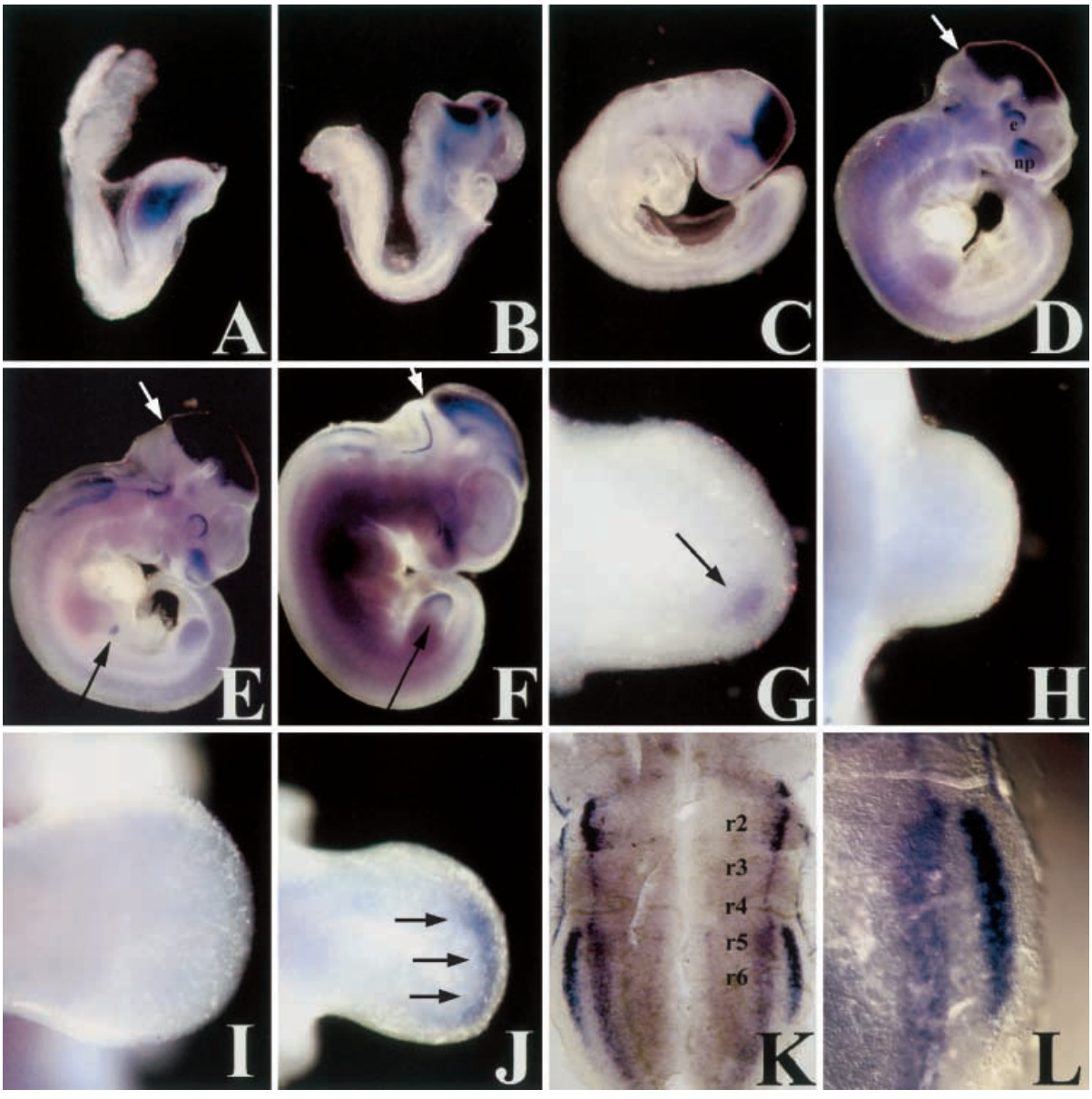

Fig. 3A–L Expression of *Dmbx1* in developing mouse embryos. **A**–**F** Whole-mouse in situ hybridisation showing *Dmbx1* expression at E7.5–8 (**A**), E8–8.5 (**B**), E9.5 (**C**), E10 (**D**), E11 (**E**) and E11.5 (**F**). **G**, **H** *Dmbx1* expression in limb buds at E11. **G** Forelimb and **H** hindlimb. **I**, **J** *Dmbx1* expression in limb buds at E11.5. **I** Forelimb and **J** hindlimb. **K**, **L** Dissected and flat-mounted hindbrain at E11 viewed from the dorsal surface, showing rhombomere identities. *White arrows* indicate the midbrain-hindbrain boundary and *black arrows* indicate limb bud expression (*e* eye, *np* nasal process)

spinal cord, changing its dorsoventral level and width. The more lateral stripe is more caudal, limited to r5 and r6. In adult mice, northern blot analysis detects a predominant 4.7-kb transcript in brain and stomach (data not shown).

In summary, we report the *Dmbx1* mouse homeobox gene and propose it as the founding member of the Dmbx family of PRD class genes. We infer that the gene family has an ancient origin in animal genomes, but has been secondarily lost from the *Drosophila* and *C. elegans* genomes. Paralogues of *Dmbx1* have also been lost from mammals. The broad expression in midbrain and diencephalon is compatible with a role in specification of regional territories. In contrast, the localised expression in hindbrain longitudinal stripes with distinct anterior and/or posterior boundaries suggests a more restricted developmental role in this structure, possibly in relation to specific neuronal populations as described for some other homeobox genes (Logan et al. 1998; Pattyn et al. 1997). The gene shows an intriguing temporal pattern of expression in limb development, being first restricted to forelimbs and subsequently to hindlimbs. The initial cloning of this gene from rare c-kit+ cells, and subsequent RT-PCR confirmation, is also consistent with a role in haematopoiesis.

Acknowledgements We thank C. Boorman and R. Furlong for help with phylogenetic analysis, C. Perriton and R. Freitas for assistance with in situ hybridisation, A. Miyajima for technical support on AGM excision and A. Graham for useful advice on hindbrain anatomy. This research was funded by the Japan Society for the Promotion of Science and the BBSRC.

References

Ferrier DEK, Holland PWH (2001) Ancient origin of the Hox gene cluster. Nat Rev Genet 2:33–38

Furlong RF, Holland PWH (2002) Were vertebrates octoploid? Philos Trans R SocLond Biol Sci 357:531–544

Galliot B, Vargas CD, Miller D (1999) Evolution of homeobox genes: Q50 Paired-like genes founded the Paired class. Dev Genes Evol 209:187–197

Logan C, Wingate RJ, McKay IJ, Lumsden A (1998) Tlx-1 and Tlx-3 homeobox gene expression in cranial sensory ganglia and hindbrain of the chick embryo: markers of patterned connectivity. J Neurosci 18:5389–5402

Medvinsky A, Dzierzak E (1996) Definitive hematopoiesis is autonomously initiated by the AGM region. Cell 86:897–906

Miyamoto T, Kawahara A, Teufel A, Mukhopadhyay M, Zhao Y, Dawid IB, Westphal H (2002) Mbx, a novel mouse homeobox gene. Dev Genes Evol 212:104–106

Nieto MA, Patel K, Wilkinson DG (1996) In situ hybridization analysis of chick embryos in whole mount and tissue sections. Methods Cell Biol 51:219–235

Ohtoshi A, Nishijima I, Justice MJ, Behringer RR (2002) Dmbx1, a novel evolutionarily conserved paired-like homeobox gene expressed in the brain of mouse embryos. Mech Dev 110:241–244

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. CABIOS 12:357–358

Pattyn A, Morin X, Cremer H, Goridis C, Brunet JF (1997) Expression and interactions of the two closely related homeobox genes *Phox2a* and *Phox2b* during neurogenesis. Development 124:4065–4075

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Treisman J, Gonczy P, Vashishtha M, Harris E, Desplan C (1989) A single amino acid can determine the DNA binding specificity of homeodomain proteins. Cell 59:553–562