The two *Xenopus Gbx2* genes exhibit similar, but not identical expression patterns and can affect head formation[1]

Ella Tour[a,b], Graciela Pillemer[a], Yosef Gruenbaum[b], Abraham Fainsod[a,*]

[a] *Department of Cellular Biochemistry and Human Genetics, Hebrew University-Hadassah Medical School, P.O. Box 12272, Jerusalem 91120, Israel*
[b] *Department of Genetics, Institute of Life Sciences, Hebrew University, Jerusalem 91904, Israel*

Received 13 July 2001; revised 19 September 2001; accepted 20 September 2001

First published online 11 October 2001

Edited by Takashi Gojobori

Abstract ***Gbx2*** **homeobox genes are important for formation and function of the midbrain/hindbrain boundary, namely the isthmic organizer. Two *Gbx2* genes were identified in *Xenopus laevis*, differing in 13 amino acids, including a change in the homeodomain. *Xgbx2a* is activated earlier during gastrulation and reaches higher levels of expression while *Xgbx2b* is expressed later, at lower levels and has an additional domain in the ventral blood islands. Their overexpression results in microcephalic embryos with shortened axes and defects in brain and notochord formation. Both genes encode functionally homologous proteins, which differ primarily in their temporal and spatial expression patterns.**

Key words: Embryonic development; Anterior–posterior axis; Midbrain–hindbrain boundary; Isthmic organizer; Homeobox

1. Introduction

Gbx homeobox genes are expressed in the gastrulating embryo and later in the brain, where they play important roles in anterior–posterior patterning [1,2]. *Gbx* homologues have been described from chicken, human and mouse [3–5]. Two subclasses, *Gbx1* and *Gbx2*, have been defined based on homeodomain amino sequences [6]. The murine *Gbx2* gene is first expressed during gastrulation and later in the midbrain/hindbrain boundary (MHB), the otic vesicles and the spinal cord [4,7]. Expression in the neural tube, and especially in the region of the MHB, suggests that this gene might play an important role in the patterning of the embryonic neural system. Mouse embryos homozygous for the *Gbx2* knockout are non-viable, dying shortly after birth [8]. These embryos exhibit severe brain malformations and show changes in the expression of molecular markers of the MHB region [8]. These data suggest that *Gbx2* is necessary for the normal development of the isthmic organizer.

In *Xenopus laevis*, due to its tetraploid genetic composition, two very similar copies of many genes can be identified. In this paper we report the cloning of the two pseudo-allelic copies of the *Gbx2* gene from *Xenopus*. The genes, *Xgbx2a* and *Xgbx2b*, exhibit differences in their regulation and patterns of expression. *Xgbx2a* exhibits higher expression levels, while only *Xgbx2b* is expressed in the anterior blood islands. The different patterns of expression, spatial, temporal or quantitative, suggest that the two genes might be under differential regulation. There are a number of amino acid changes between the two proteins, but despite these differences, overexpression of either *Xgbx2a* or *Xgbx2b* induces very similar head malformations.

2. Materials and methods

2.1. Cloning of Xgbx2a and Xgbx2b

A *Xenopus* stage 18 embryonic cDNA library in Lambda-ZAP II (Stratagene) was screened with a 90 bp PCR-derived probe from a putative *Xenopus Gbx2* sequence [9] under moderate stringency conditions, according to standard protocols. Sequencing of *Xgbx2b* and *Xgbx2a* was performed synthesizing new primers as required to sequence both strands.

2.2. Production of Xgbx2a- and Xgbx2b-based constructs

For the generation of gene-specific probes, parts of the 5′-UTR regions of *Xgbx2a* and *Xgbx2b* were subcloned into the pBluescript KS$^-$ vector. The *Xgbx2a* 5′-UTR region was PCR-amplified using the SK primer (Stratagene) and an internal primer from the *Xgbx2a* gene (5′-GAATCTAGAGGGTGATGGATGCCAAAG). The PCR product was cloned in the *Xba*I site to generate pBS-Xgbx2a-5′. The *Xgbx2b* 5′-UTR region was PCR-amplified using the KS primer (Stratagene) and a *Xgbx2b*-specific primer (5′-GAATCTAGAGGGTGATGGATGCCAAAG), digested with *Hin*dIII and *Xba*I and inserted into pBluescript KS$^-$ to generate pBS-Xgbx2b-5′. These two plasmids were utilized to prepare *Xgbx2a*- and *Xgbx2b*-specific probes. The full-length *Xgbx2a* probe cDNA was prepared from the pBS-Xgbx2a plasmid.

Sense RNA for injection was prepared from the pSP64T-Xgbx2a and pSP64T-Xgbx2b plasmids by in vitro transcription using the RiboMax kit (Promega). Cap analog (Pharmacia) was added to the reaction mixture using a ratio of cap:GTP of 5:1.

2.3. Expression analysis by in situ hybridization

Whole mount in situ hybridization analysis of gene expression was performed as previously described by Epstein et al. [10]. Digoxigenin-labeled RNA probes were prepared from linearized plasmids transcribed in vitro using the RiboMax transcription kit (Promega). For the generation of specific probes of the 5′-UTR regions of *Xgbx2a* and *Xgbx2b*, pBS-Xgbx2a-5′ was restricted with *Sac*I and transcribed with T3 RNA polymerase, pBS-Xgbx2b-5′ was linearized with *Hin*cII and transcribed with T7.

3. Results

3.1. Cloning the Gbx2 homologues in X. laevis

Two *Xenopus Gbx2* cDNA clones were isolated and termed *Xgbx2a* and *Xgbx2b* (Fig. 1A, B). The *Xgbx2a* cDNA is 2310 bp in length with two putative polyadenylation signals close

*Corresponding author. Fax: (972)-2-6415848.
E-mail address: fainsod@cc.huji.ac.il (A. Fainsod).

[1] GenBank accession number *Xgbx2b*, AF395825.

0014-5793/01/$20.00
PII: S0014-5793(01)02963-5

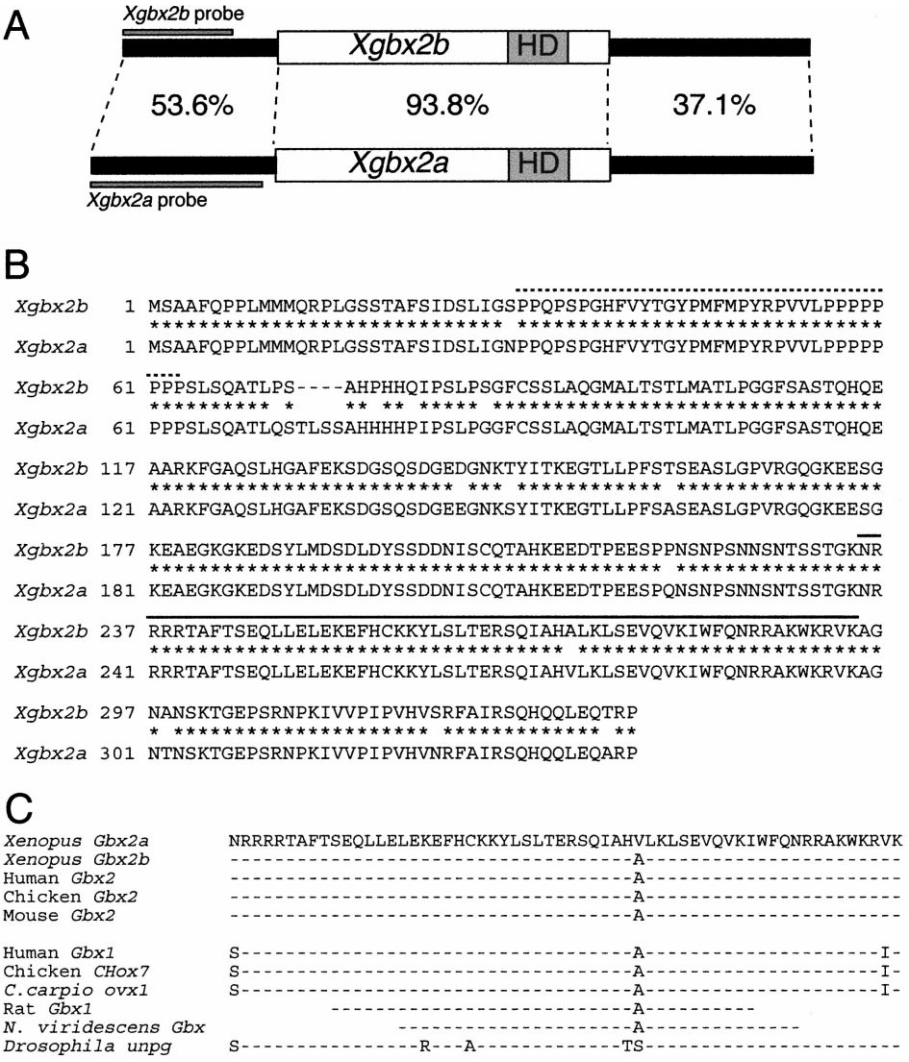

Fig. 1. *Xenopus Xgbx2a* and *Xgbx2b* genes. A: Schematic representation of the *Xgbx2a* and *Xgbx2b* cDNA clones. The percent identity in the 5′- and 3′-untranslated regions, the protein coding regions shown. The gene-specific probes are shown. B: Comparison of protein sequences. The homeodomain is overlined with a solid line and the proline-rich region with a dashed one. C: Comparison of the *Gbx*-class homeodomains.

to the 3′-end. The homeobox-containing open reading frame begins at position 667 of the cDNA, and the first in-frame stop codon downstream from the homeodomain is found at position 1689. The predicted protein is 340 amino acids long. It is identical to *Xlhox7a*/*Xgbx2* (Fig. 1C) [9,11], and was termed Xgbx2.

The *Xgbx2b* cDNA clone is 2097 bp in length. The open reading frame begins at position 474, and the first in-frame stop codon downstream from the homeodomain is at position 1482. A consensus signal for polyadenylation was not identified at the 3′-end of this cDNA suggesting that either the 3′-untranslated sequence is not full, or it utilizes one of the AT-rich sequences for polyadenylation. The predicted protein is 336 amino acids long with a *Gbx2*-class homeodomain (Fig. 1B, overlined) suggesting that it encodes a functional protein. Another interesting feature of the predicted Xgbx2b protein is a proline-rich region close to the amino-terminus of the protein spanning 33 amino acids, with 45% prolines (Fig. 1B, dashed line).

Alignment of the *Xgbx2a* and *Xgbx2b* DNA sequences shows 83% identity along the entire sequence, and 93.8% identity between the coding regions. In the non-coding regions, the similarity is much lower (Fig. 1A). At the amino acid level, the identity throughout the whole protein is 96%. There are 13 amino acids that differ between Xgbx2a and Xgbx2b, one of them a conservative change at position 37 of the homeodomain, from alanine in Xgbx2b to valine in Xgbx2a (Fig. 1B). Other changes localize to the amino-terminus and the short carboxy-terminal region downstream from the homeodomain, including a four amino acid deletion just downstream from the proline-rich domain in the Xgbx2b protein (Fig. 1B).

The sequence of *Xgbx2a* differs from that of other vertebrate *Gbx2* genes at position 37 where they all have an alanine and not a valine (Fig. 1C). In contrast, Xgbx2b has an alanine in this position making it identical to the other vertebrate *Gbx2* homeodomains. The closely related homeobox family, *Gbx1* also has an alanine at position 37 (Fig. 1C).

3.2. Comparison of the expression patterns of Xgbx2a and Xgbx2b

Gene-specific probes from the 5′-untranslated regions of the *Xgbx2a* and *Xgbx2b* genes were used to determine the extent

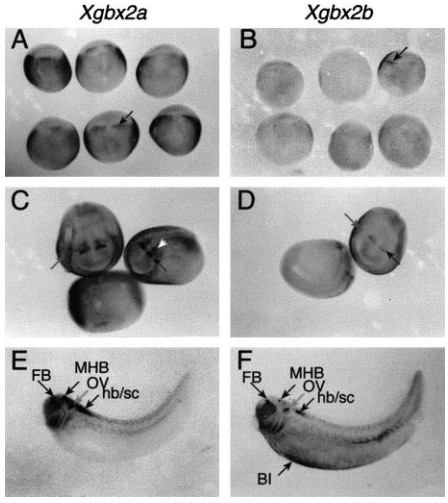

Fig. 2. Comparison of the expression patterns of *Xgbx2a* and *Xgbx2b*. A, C, E: *Xgbx2a*. B, D, F: *Xgbx2b*. A, B: Dorsal view of stage 12.5 embryos. Anterior to the top. The arrow marks the anterior border of expression in the neuroectoderm. C, D: Anterior view of stage 18 embryos. MHB, arrow on the right; future hindbrain and anterior spinal cord, white arrowhead; anterior lateral ectoderm, arrow on the left. E, F: Lateral view of stage 34 embryos. Anterior to the left. OV, otic vesicles; hb/sc, posterior hindbrain/anterior spinal cord; FB, forebrain; BI, ventral blood islands.

of similarity between their patterns of expression. The homology between the gene-specific probes is 44.2%, as they don't include the entire 5′-untranslated region (Fig. 1A). Hybridization with these probes revealed a remarkable similarity in the *Xgbx2b* and *Xgbx2a* expression patterns (Fig. 2). The main regions of expression of both *Xgbx2a* and *Xgbx2b*, i.e. MHB, otic vesicles, posterior hindbrain/anterior spinal chord, and the forebrain, are stained by the two probes (Fig. 2). However, *Xgbx2b* expression is detected only later in embryogenesis, at stage 12 as compared to stage 10.5 for *Xgbx2a* (Fig. 2A, B). The level of expression of *Xgbx2b* was lower at all stages examined. During later stages, *Xgbx2b* transcripts are also detected in the anterior blood islands (Fig. 2F). This domain of expression could not be identified with the *Xgbx2a*-specific probe.

3.3. Overexpression of the Xgbx2 genes affects head development

Overexpression of *Xgbx2a* and *Xgbx2b* was performed in order to determine the functional overlap in activity between both genes. Sense RNA of the *Xgbx2a* or *Xgbx2b* genes was injected in the marginal zone of 2–4 cell embryos (400 pg). The injected embryos were incubated until tailbud stages (stages 32–36), and phenotypic and histological analysis was performed. When overexpressed, both *Xgbx2a* and *Xgbx2b* genes induce the same range of phenotypes, resulting in shorter and microcephalic embryos (Fig. 3B, C). The anterior-most part of the head was reduced, with rudimentary or absent cement gland and eyes (Fig. 3B, C). Histological sections revealed that the forebrain of embryos overexpressing *Xgbx2a* or *Xgbx2b* was significantly reduced, and lacked the normal ventricular cavity (Fig. 3E, F). Eye structures were absent or rudimentary (Fig. 3E, F). Conversely, the hindbrain in these embryos was enlarged compared to control sibling embryos (Fig. 3G–I). The IV ventricle developed as an unusually wide cavity. Malformations in the notochord, which was significantly enlarged and developed side branches, were observed (Fig. 3H, I).

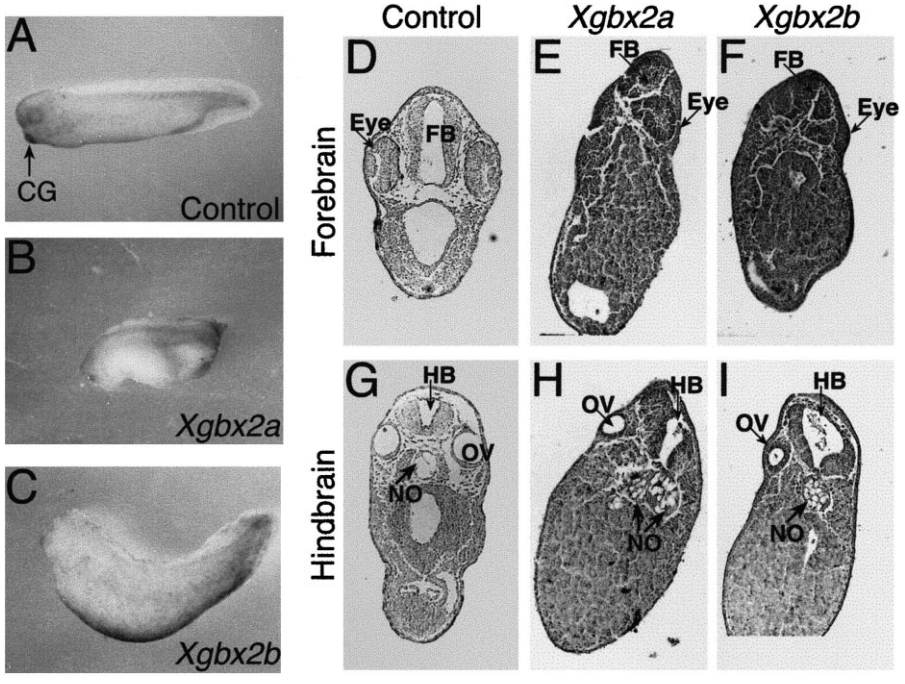

Fig. 3. Overexpression of *Xgbx2a* and *Xgbx2b* causes head malformations. Embryos were injected with *Xgbx2a* or *Xgbx2b* mRNA and subjected to phenotypic (A–C) and histological (D–I) analysis. A: Control embryo, stage 32. B: *Xgx2a*-injected embryo. C: *Xgbx2b*-injected embryo. D–F: Sections at the level of the forebrain. G–I: Sections at the level of the hindbrain. D: Normal forebrain of stage 30 embryo. E: Anterior head malformations in embryo injected with *Xgbx2a* mRNA. F: Embryo injected with *Xgbx2b* mRNA exhibits similar malformations. G: Control embryo, stage 30. H: *Xgbx2a*-injected embryo with hindbrain and notochord malformations. I: *Xgbx2b*-injected embryo with similar malformations. FB, forebrain; HB, hindbrain; NO, notochord; OV, otic vesicle; CG, cement gland.

4. Discussion

We isolated two homologues of the *Gbx2* gene in *Xenopus*, *Xgbx2a* and *Xgbx2b* that apparently arose as part of the tetraploid condition in this species [12]. Sequence analysis and temporal and spatial expression patterns of these genes revealed that both share extensive similarities and represent active genes. Differences between the *Xgbx2a* and *Xgbx2b* genes suggest divergence in their regulation since their duplication during evolution.

4.1. Isolation and sequence characteristics of the Xenopus Gbx2 homologues

Two cognates of the *Gbx2* homeobox gene in *Xenopus* were isolated. One of them, *Xgbx2a*, was previously known [9,13], while the other, *Xgbx2b*, represents a novel gene. The coding regions of *Xgbx2a* and *Xgbx2b* are very similar, but the identity rapidly drops in the non-coding regions. The overall similarity suggests that both genes originated from the same ancestral gene. The Xgbx2a and Xgbx2b proteins differ from each other in 13 amino acids, including a four amino acid deletion in the amino-terminal region of Xgbx2b. A significant difference between the two is a change from alanine in Xgbx2b to valine in Xgbx2a, at position 37 in the homeodomain. All known vertebrate *Gbx2*- and *Gbx1*-type homeodomains, have alanine in this position, suggesting that the valine change is a novel event. Since overexpression of *Xgbx2a* and *Xgbx2b* results in similar phenotypes, it would indicate that both genes can recognize the same target genes.

4.2. Xgbx2a and Xgbx2b expression patterns

The *Xgbx2a* pattern of expression is similar to that described by von Bubnoff et al. [13] and it exhibits temporal, spatial and quantitative differences to the pattern of *Xgbx2b*. Overall, the expression domains of the *Xenopus Gbx2* cognates are highly conserved with their homologues in chick and mice ([3,4,14]; Rangini, Gruenbaum and Fainsod, unpublished results). Expression of the *Xenopus Gbx2* genes begins around early/midgastrula, and is excluded from anterior regions of the embryo. However, unlike their counterparts in chick and mice, the *Xenopus* genes are excluded from the posterior-most part of the embryo, and they exhibit a wide gap in their expression encompassing the dorsal midline. These differences are probably the result of different developmental strategies, like the formation of a blastopore as opposed to the formation of the primitive streak. Comparative analysis of the patterns of expression of the two *Xenopus Gbx2* genes suggests that although both genes represent a duplication of an ancestral gene, they have diverged enough to exhibit temporal, quantitative and spatial differences. Unlike *Xgbx2a*, *Xgbx2b* is detected in stage 32 embryos in the ventral aspect of more posterior somites and in the ventral region, where it overlaps with the expression of *Taglobin* [15]. This overlap raises the possibility that only the *Xgbx2b* gene is active in the ventral blood island in *Xenopus*, and it might play a role in the differentiation of the blood cells. The chicken *Gbx2* is expressed in myelomonocytic, erythroid and lymphoid cells of the hematopoietic lineage in 13-day embryos [16]. The differences between the two *Xenopus* genes suggest a divergence in their regulatory sequences to restrict or achieve additional functions during development.

4.3. Xgbx2a and Xgbx2b overexpression studies

The overall phenotype of *Xgbx2a* and *Xgbx2b* overexpression showed two major defects: lack of anterior head structures and overall shortening of the embryo. The most evident and reproducible manifestation of the head phenotype was the absence of cement gland and eyes. In the more severely affected embryos, lack of the entire anterior head was observed. Histological sections of mildly affected embryos showed that the forebrain tissue undergoes dramatic changes, becoming small and consisting of only a dense mass of cells, without any cavity at all. The eyes are rudimentary or missing. The phenotype obtained is similar but stronger than that described by King et al. [11]. These malformations are reminiscent of the *Otx2* knockout phenotype in mice [17–19]. Negative regulation of *Otx2* by *Gbx2* could account for the phenotypes observed in *Xgbx2a* and *Xgbx2b* overexpressing embryos ([20,21]; Tour et al., submitted). The striking phenotype of enlarged and/or branching notochord observed in *Xbgx2a*- and *Xgbx2b*-injected embryos might also result from *Otx2* inhibition [19].

It has been shown that the differences between homeobox genes within a class focus mainly on the regulatory sequences of the genes and not the actual activity of the proteins they encode [22–25]. The *Gbx2* homeobox genes in *Xenopus* represent an example where, following duplication, the genes have diverged in their regulatory sequences to achieve differential spatial and temporal expression patterns, while their protein products remain functionally homologous.

Acknowledgements: We thank Dr. Dale Frank and Dr. Oded Khaner for discussions during the course of the work. This work was supported in part by grants from the March of Dimes Birth Defects Foundation, Volkswagen Stiftung and The Israel Academy of Sciences to A.F. and by the Sheinbrun Foundation to Y.G.

References

[1] Chapman, G. and Rathjen, P.D. (1995) FEBS Lett. 364, 289–292.
[2] Murtha, M.T., Leckman, J.F. and Ruddle, F.H. (1991) Proc. Natl. Acad. Sci. USA 88, 10711–10715.
[3] Fainsod, A. and Gruenbaum, Y. (1989) FEBS Lett. 250, 381–385.
[4] Bulfone, A., Puelles, L., Porteus, M.H., Frohman, M.A., Martin, G.R. and Rubenstein, J.L. (1993) J. Neurosci. 13, 3155–3172.
[5] Matsui, T., Hirai, M., Hirano, M. and Kurosawa, Y. (1993) FEBS Lett. 336, 107–110.
[6] Chapman, G. and Rathjen, P.D. (1995) FEBS Lett. 364, 289–292.
[7] Bouillet, P., Chazaud, C., Oulad-Abdelghani, M., Dolle, P. and Chambon, P. (1995) Dev. Dyn. 204, 372–382.
[8] Wassarman, K.M., Lewandoski, M., Campbell, K., Joyner, A.L., Rubenstein, J.L., Martinez, S. and Martin, G.R. (1997) Development 124, 2923–2934.
[9] King, M.W. and Moore, M.J. (1994) Nucleic Acids Res. 22, 3990–3996.
[10] Epstein, M., Pillemer, G., Yelin, R., Yisraeli, J.K. and Fainsod, A. (1997) Development 124, 3805–3814.
[11] King, M.W., Ndiema, M. and Neff, A.W. (1998) Dev. Dyn. 212, 563–579.
[12] Kobel, H.R., Loumont, C. and Tinsley, R.C. (1996) in: The Biology of *Xenopus* (Tinsley, R.C. and Kobel, H.R., Eds.), pp. 9–33, Clarendon Press, Oxford.
[13] von Bubnoff, A., Schmidt, J.E. and Kimelman, D. (1996) Mech. Dev. 54, 149–160.
[14] Niss, K. and Leutz, A. (1998) Mech. Dev. 76, 151–155.
[15] Xu, R.H., Ault, K.T., Kim, J., Park, M.J., Hwang, Y.S., Peng, Y., Sredni, D. and Kung, H. (1999) Dev. Biol. 208, 352–361.
[16] Kowenz-Leutz, E., Herr, P., Niss, K. and Leutz, A. (1997) Cell 91, 185–195.

[17] Ang, S.-L., Jin, O., Rhinn, M., Daigle, N., Steenson, L. and Rossant, J. (1996) Development 122, 243–252.
[18] Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A. and Brulet, P. (1995) Development 121, 3279–3290.
[19] Matsuo, I., Kuratani, S., Kimura, C., Takeda, N. and Aizawa, S. (1995) Genes Dev. 9, 2646–2658.
[20] Simeone, A. (2000) Trends Genet. 16, 237–240.
[21] Joyner, A.L., Liu, A. and Millet, S. (2000) Curr. Opin. Cell Biol. 12, 736–741.
[22] Hanks, M., Wurst, W., Anson-Cartwright, L., Auerbach, A.B. and Joyner, A.L. (1995) Science 269, 679–682.
[23] Suda, Y., Nakabayashi, J., Matsuo, I. and Aizawa, S. (1999) Development 126, 743–757.
[24] Greer, J.M., Puetz, J., Thomas, K.R. and Capecchi, M.R. (2000) Nature 403, 661–665.
[25] Acampora, D. (1999) Development 126, 1417–1426.